Balkan Journal of Electrical & Computer Engineering

An International Peer Reviewed, Referred, Indexed and Open Access Journal

Aim & Scope
The journal publishes original papers in the extensive field of Electrical-Electronics and Computer engineering. It accepts contributions which are fundamental for the development of electrical engineering, computer engineering and its applications, including overlaps to physics. Manuscripts on both theoretical and experimental work are welcome. Review articles and letters to the editors are also included.

Application areas include (but are not limited to): Electrical & Electronics Engineering, Computer Engineering, Software Engineering, Biomedical Engineering, Electrical Power Engineering, Control Engineering, Signal and Image Processing, Communications & Networking, Sensors, Actuators, Remote Sensing, Consumer Electronics, Fiber-Optics, Radar and Sonar Systems, Artificial Intelligence and its applications, Expert Systems, Medical Imaging, Biomedical Analysis and its applications, Computer Vision, Pattern Recognition, Robotics, Industrial Automation.

Balkan Journal of **Electrical & Computer Engineering**
BAJECE

An International Peer Reviewed, Indexed and Open Access Journal

CONTENTS

Brain Decoding over the MEG Signals Using Riemannian Approach and Machine Learning

Zeynep Ozer, Onursal Cetin, Kutlucan Gorur, Feyzullah Temurtas

Abstract—**Brain decoding is an emerging approach for understanding the face perception mechanism in the human brain. Face visual stimuli and perception mechanism are considered as a challenging ongoing research of the neuroscience field. In this study, face/scrambled face visual stimulations were implemented over the sixteen participants to be decoded the face or scrambled face classification using machine learning (ML) algorithms via magnetoencephalography (MEG) signals. This noninvasive and high spatial/temporal resolution signal is a neurophysiological technique which measures the magnetic fields generated by the neuronal activity of the brain. The Riemannian approach was used as a highly promising feature extraction technique. Then Long Short-Term Memory (LSTM), Gated Recurrent Unit (GRU), Convolutional Neural Network (CNN) were employed as deep learning algorithms, Linear Discriminant Analysis (LDA) and Quadratic Discriminant Analysis (QDA) were implemented as shallow algorithms. The improved classification performances are very encouraging, especially for deep learning algorithms. The LSTM and GRU have achieved 92.99% and 91.66% accuracy and 0.977 and 0.973 of the area under the curve (AUC) scores, respectively. Moreover, CNN has yielded 90.62% accuracy. As our best knowledge, the improved outcomes and the usage of the deep learning on the MEG dataset signals from 16 participants are critical to expand the literature of brain decoding after visual stimuli. And this study is the first attempt with these methods in systematic comparison. Moreover, MEG-based Brain-Computer Interface (BCI) approaches may also be implemented for Internet of Things (IoT) applications, including biometric authentication, thanks to the specific stimuli of individual's brainwaves.**

Index Terms—**Magnetoencephalography, Brain Decoding, Riemannian Approach, Deep Learning.**

ZEYNEP ÖZER, is with Department of Management Information Systems University Bandirma Onyedi Eylul University, Balıkesir, Turkey,(e-mail: zozer@bandirma.edu.tr)
https://orcid.org/0000-0001-8654-0902

ONURSAL ÇETİN, is with Department of Electrical and Electronics Engineering University of Bandirma Onyedi Eylul University, Balıkesir, Turkey,(e-mail: ocetin@bandirma.edu.tr)
https://orcid.org/0000-0001-5220-3959

KUTLUCAN GÖRÜR, is with Department of Electrical and Electronics Engineering University of Bandirma Onyedi Eylul University, Balıkesir, Turkey,(e-mail: kgorur@bandirma.edu.tr)
https://orcid.org/0000-0003-3578-0150

FEYZULLAH TEMURTAŞ, is with Department of Electrical and Electronics Engineering University of Bandirma Onyedi Eylul University, Balıkesir, Turkey,(e-mail: ftemurtas@bandirma.edu.tr)
https://orcid.org/0000-0002-3158-4032

I. INTRODUCTION

THE brain decoding has obtained great attention from scientific communities in medical applications and is observed as one of the primary goals of the brain analysis literature [1]. Face perception mechanism in individuals is generated by a sequence of cortical activities [2,3]. Neuropsychology and cognitive neuroscience research fields require magnetoencephalography (MEG) and electroencephalography (EEG) signals to decode these brain maps [4]. Different experiments can be implemented to arouse these brain patterns, such as visual stimuli. The brain reacts to different responses for different visual stimuli [4]. If the subject is stimulated by a visual cue, then the related brain activity is recorded from multiple noninvasive sensors. Then each recorded data is named as a trial [5]. In this study, face/scrambled face are used as a visual stimulus; during the same time, MEG signals are collected over the brain activities. MEG signal is a neurophysiological way of measuring magnetic fields generated by neural electrical activities [6]. These signals have some advantages compared to the EEG signals, functional magnetic resonance imaging (fMRI), and positron emission tomography (PET) methods because of the alleviated effects by cerebrospinal fluid, skull, and skin [3]. Moreover, MEG signals present great spatial and temporal resolution [7].

The low prediction performance in multivariate brain decoding is generally caused by the low signal-to-noise ratios (SNRs), high dimensionality recordings of the scalp, and cross-subject variations [8]. Recently machine learning (ML) algorithms offer these problems a very promising approach in signal processing techniques to recognize the activated brain patterns using noninvasive MEG signals [5,9-11]. In this research study, Convolutional Neural Network (CNN), Long Short-Term Memory (LSTM), and Gated Recurrent Unit (GRU) were implemented as deep learning algorithms when Linear Discriminant Analysis (LDA) and Quadratic Discriminant Analysis (QDA) were used as traditional (shallow) machine learning algorithms. CNN is considered as a state of the art machine learning algorithm especially used for computer vision, natural language processing, and pattern recognition. CNN has convolution layers to extract features of the signal-based images. These features consist of edges, lines, or corners [12]. LSTM is carried out in the prediction of sequential data. The key point of LSTM is that it can remember input for a long time while estimating outputs [13]. The GRU is another applied popular ML algorithm based on the Recurrent

Neural Network (RNN). The reason for the popularity is considered as computation cost and a simple model of topology. This technique is assembled into a single "update gate" with forget and input gates and mounted in the cell state and hidden state [14]. Furthermore, LDA is to project the dataset for classification in a supervised manner. This projected-based method aims to find the best projection direction during the classification task [15]. On the other hand, QDA assumes not equality of the covariance matrices that gain a quadratic decision boundary for two-class problems [16].

The Riemannian approach was implemented in the feature extraction step that enables direct manipulation of multichannel MEG signals to covariance matrices. Then these matrices were employed as features. The Riemannian approach is known as extremely competitive and superior to the other feature extraction techniques, including the Common Spatial Pattern [5,17]. The simultaneous MEG signal recordings were obtained from 16 participants who perform trials during the face and scrambled face visual stimuli to reveal the patterns over the brain dynamics [5]. After that, the Riemannian approach was employed for the feature extraction process [5]. In this study, deep learning algorithms, LSTM and GRU has achieved 92.99% accuracy and 0.977 of AUC (area under the curve) score, 91.66% accuracy, and 0.973 AUC score, respectively. Moreover, CNN has yielded 90.62% accuracy with 0.959 AUC score among the deep learning approaches. Then LDA has determined 78.23% accuracy and 0.861 AUC when QDA has obtained a classification accuracy of 72.24% and 0.796 of AUC score. These results show that LSTM, GRU, and CNN have noticable performances on the MEG signals compared to the previous research study (80.85% accuracy and 0.81 AUC with Deep Neural Network) [5]. Moreover, LDA has also offered improved performance than Support Vector Machine (78.01% accuracy) and the other traditional classifiers in the same study [5]. However, the satisfactory accuracy (79%) of the Generalized Regression Neural Network (GRNN) was not achieved by LDA and QDA results [11]. MEG-based BCI systems have high time resolution and high uniqueness for individuals [4,8-10]. These features can provide effectiveness for IoT applications, such as biometric authentication.

The rest of the article was organized as follows; the method chapter defines dataset description of the dataset details, preprocessing, feature extraction step, evaluation metrics, and machine learning algorithms. Furthermore, the results and discussion chapter explains the performance outcomes of the MLs and statements of the findings. Then the conclusion chapter is to point out the discriminations of the study among the literature of brain decoding in terms of the ML estimations.

II. MATERIAL AND METHOD

A. Definition of the dataset

Magnetoencephalography signals were recorded with an Elekta Neuromag VectorView system from 18 participants; thereby, the dataset was created by Henson et al. [18]. The triple sensor group, consisting of a magnetometer and two gradiometers, is located at 102 positions. In total, 306 sensors

record the magnetic field caused by brain currents. The z (radial) component of the magnetic field is measured by the magnetometer, while the x and y spatial derivative is measured by the gradiometer. Details of the measuring system are provided by Henson et al. [18].

The dataset, used in this study, was modified within the scope of the [19-20]. The modified dataset encompasses approximately 588 trials for each 16 subjects. Visual stimuli of faces and scrambled faces were randomly presented to the subjects for 1 second, and for each stimulation, subjects were rested for 0.5 seconds. Each trial consists of 1.5 seconds of MEG recording was sampled at 250 Hz from 306 channels. Thus, a total of 9414 trials were presented in a random order [5].

The flowchart showing the overall workflow from MEG recordings to the brain decoding process is given in Fig. 1. A series of operations carried out for this study can be listed as follows: Pre-processing of signals using a bandpass filter, source extraction by spatial filtering, extracting feature vector, classification of the feature vector.

The brain computer interfaces (BCIs) are used to perform various tasks (control, communication, biometric authentication) via processing of brain signals [21]. In the preprocessing phase, the signal was first high-pass filtered at 1Hz. The first 0.5 seconds of the signals were discarded when contemplating the rest time of the subjects. With the start of the feature extraction process, a bandpass filter having 1Hz-20Hz cut-off frequency points is used. Then, by applying a spatial filter to the signal, the dimensionality is reduced, and the signal-to-noise ratio is increased [5].

The Riemannian approach allows the direct manipulation of multichannel MEG signals to covariance matrices and subspaces with proper and special geometry, as defined by Yger et al. [22]. Through the Riemannian geometry, the 2176 features of the dataset were obtained using tangent space representation from the MEG covariance matrices at the end of the feature extraction process [23]. Tangent space is a vector field that allows all Euclidean statistical methods, and Riemannian metrics are more suitable for flat vector spaces rather than conventional Euclidean metrics. Thus, this approach can be applied to the MEG-based brain-computer interface, especially for the implementation of the classifier, and feature representation. Detailed computational issues about the Riemannian geometry can be found in [22].

B. Evaluation metrics for performances

Evaluation metrics are important factors to show the machine learning algorithm performances in the dataset. The suitable evaluation metrics are important keys to discriminate the optimal MLs [24]. In this study, accuracy (ACC), sensitivity (SENS), specificity (SPEC), and the area under the curve (AUC) score were employed as evaluation metrics with k-fold cross-validation technique. In this technique, the dataset is divided into k-subsets. Then the ML is trained with k-1 subsets, and the other one subset is tested. The process is repeated k times so that each subset sample will be trained and tested. After that, the average classification metrics are obtained for each ML algorithm [12]. The formulas for accuracy were presented in Eqs. (1-3) [25]:
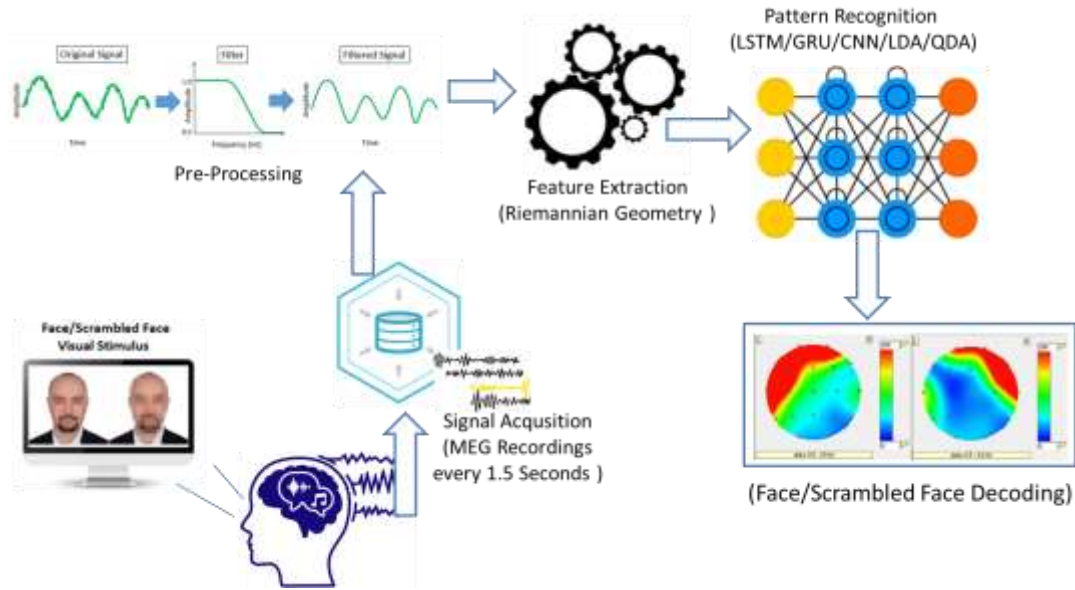
Fig.1. The experiment of MEG recordings and brain decoding process

$$Accuracy(N) = \frac{\sum_{i=1}^{|N|} estimate(n_i)}{|N|}, \quad n_i \in N \qquad (1)$$

$$Estimate(n) = \begin{cases} 1, & if\ estimate(n) = cn \\ 0, & otherwise \end{cases} \qquad (2)$$

$$Classification\ Accuracy(ML) = \frac{\sum_{i=1}^{|k|} Accuracy(N_i)}{|k|} \qquad (3)$$

where N refers the classified (test) dataset, cn defines the class of the value of n, $Estimate(n)$ describes the classification result of n, and then the k value is named for the k-fold cross-validation [12].

Different metrics reveal the different characteristics of the ML algorithms induced by the processing [26]. Therefore, it may help easier to make the comparison and analysis of ML algorithms in the robustness observation. In general, the sensitivity and specificity are well-known evaluation metrics for performance analysis of ML algorithms. Thus, they are described as the following equations [27]:

$$Sensitivity = \frac{TP}{TP + FN} \qquad (4)$$

$$Specificity = \frac{TN}{TN + FP} \qquad (5)$$

where: True Positive (TP): The number of face pattern decisions which are targeted as face pattern, True Negative (TN): The number of scrambled face pattern decisions which are targeted as scrambled face pattern, False Positive (FP): The number of scrambled face pattern decisions which are targeted as face pattern, False Negative (FN): The number of face pattern decisions which are targeted as scrambled face pattern.

Thanks to the AUC evaluation metrics, the classification results are presented across the interval of 0-1 scores under the curve of false positive rates and true positive rates. The higher AUC value means better classifier performance. This popular ranking type metric is to prove that the prediction and diagnostic ability of MLs are noteworthy (AUC>0.9) and good discrimination (0.8≥AUC>0.7) if the AUC scores are found in the range of the stated values [28].

C. Linear discriminant analysis and quadratic discriminant analysis Most

Linear Discriminant Analysis (LDA) and Quadratic Discriminant Analysis (QDA) machine learning methods are both based on the fundamentals of statistical and probabilistic learning. Basically, when LDA is used for linear classification, QDA is employed for quadratic decision problems of the classification process [16]. Kernel Fisher's Discriminant analysis has derived the LDA method that is a type of projection technique. LDA is implemented to classify the dataset in the manner of reducing the dimension. The aim of LDA is to maximize the between-class distance and to minimize within-class distance. If the class samples are defined as C1 and C2, LDA finds the projection direction (w) for maximum separability of the spatial pattern [15-29]. The related Equations are presented below:

$$z = w^T x \qquad (6)$$

where x (data samples) are employed to be projected onto w. The graphical presentation of LDA is shown in Fig. 2 [15].

where $m1$ to $\boldsymbol{m1}$ describes the means of samples in class $C1$ before and after the projection process, respectively. Therefore, $\boldsymbol{m1} \in \Re^d$ defines the multi-dimension, and $m1 \in \Re$ means the projected dimension. Then $m2$ and $\boldsymbol{m2}$ have a similar manner for class $C2$. The samples of scattered dataset around the means are presented as s_1^2 and s_2^2. Then the samples of training dataset are defined as $X\{x^t, r^t\}$:

$$X\{t\} = \begin{cases} r^t = 1, & x^t \in C1 \\ r^t = 0, & x^t \in C2 \end{cases} \qquad (7)$$
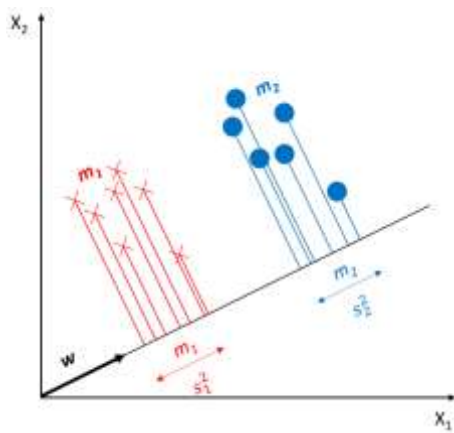
Fig.2. Classification process in LDA via the projection of data samples

$$J(w) = \frac{w^T S_B w}{w^T S_W w} = \frac{|w^T (m1 - m2)|^2}{w^T S_W w} \qquad (8)$$

where S_B and S_W are named as the between-class scatter matrix and within-class scatter matrix in $J(w)$, respectively [15].

QDA assumes that the equality of the covariance matrices is not necessary (as shown in Eq. 9) as LDA. This feature yields the advantage to the QDA to be used in the decision boundary of quadratic classification [16].

$$\Sigma_1 \neq \Sigma_2 \qquad (9)$$

D. Recurrent neural network

The use of artificial neural networks (ANN) in machine learning applications is very common. Over time, many different ANN models have been developed in line with needs. Accordingly, ANNs are specialized to process different types of data. Such as Convolutional Neural Networks are specialized for matrix type data like image. On the other hand, Recurrent Neural Networks (RNN) have been developed to process array data. Traditional feed-forward ANNs take into account existing samples to which they are exposed as input. RNNs, apart from this, apply the samples they perceive over time, as well as the existing ones.

An input sequence is given as $[x_1, x_2, \ldots, x_k]$ with $x_i \in \mathbb{R}^d$. Different examples can have different sequence lengths. Therefore, the k value may vary. In each step of the RNN model, a hidden state is generated as an array $[h_1, h_2, \ldots, h_k]$. Activation of hidden state at time t is calculated as a function of the current input x_t and previously hidden state h_t -1. This process can be expressed as follows:

$$h_t = f(x_t, h_{t-1}) \qquad (10)$$

Unlike traditional feed-forward ANN, RNNs have a repeat layer. By means of this layer, the state information generated by the feed-forward network is stored and re-applied to the network with the input information. That is, RNNs have a memory that holds what has been calculated so far [30]. Fig. 3 shows an exemplary RNN network unit and closed notation of this architecture.
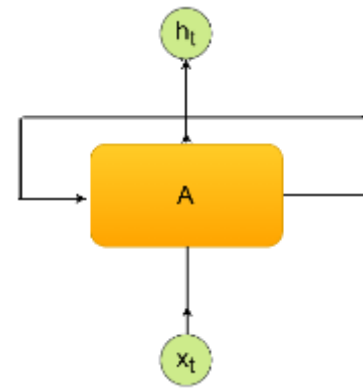
Fig.3. Closed notation of Recurrent Neural Network

1) Long short-term memory network

Long Short-Term Memory networks, often called LSTM, are a special type of RNN capable of learning long-term dependencies [21]. This model, which was first proposed in the mid-90s [31], is widely used today. While processing the sequences in RNNs, it is aimed to store and transfer ANN status information. However, it is unlikely to be transferred without disturbing long-term dependencies as a result of transferring the state information by continuously processing it. In other words, while short-term dependencies are transferred very successfully, there are problems in transferring long-term dependencies. LSTMs are designed to address long-term dependency problems.

All RNN networks consist of modules that are repeating like a chain. In standard RNNs, each of these modules usually consists of a *tanh* function or a similar function. The feature that distinguishes LSTMs from standard RNNs is that the internal structure of this module consists of 4 separate structures that interact with each other.
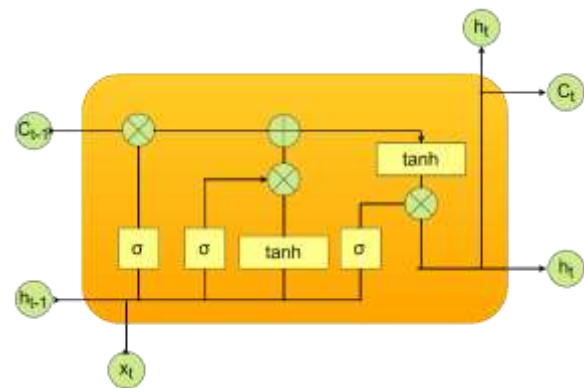
Fig.4. LSTM structure

LSTM module consists of 3 separate gates, as shown in Fig.4. These are the forget gate, the input gate, and the output gate, respectively. Forget gates decides how much of the information will be forgotten and how much of it should be transferred to the next stage. For this process, it uses the sigmoid function, which produces a value between 0 and 1. 0 means that the information will never be transmitted, while 1 means that all must be transmitted.

The next step is to decide what information should be stored. For this, the input layer firstly decides which values should be updated. Then the $tanh$ function forms a vector of the new candidate values of the memory cell defined as $\tilde{C}_t$. Then these two processes are combined. This process is expressed mathematically as follows:

$$i_t = \sigma(W_i \cdot [h_{t-1}, x_t] + b_i) \tag{11}$$

$$\tilde{C} = tanh(W_C \cdot [h_{t-1}, x_t] + b_C) \tag{12}$$

In general, W is named as the weight vector, b is called the bias term, σ describes the sigmoid activation function for non-linearity, x_t is used for the input sequence, h_{t-1} is implemented as the output of the neuron at time $t-1$ for feedback into the neuron. Furthermore, i_t, f_t and o_t are defined as the input, forget and output gate, respectively. After this process, the new status information of the memory cell is calculated. The new status information is calculated as follows:

$$C_t = f_t * C_{t-1} + i_t * \tilde{C}_t \tag{13}$$

Finally, the output of the system is calculated. This is done at the output gate. The output of the system h_t can be calculated as follows:

$$o_t = \sigma(W_o \cdot [h_{t-1}, x_t] + b_o) \tag{14}$$

$$h_t = o_t * tanh(C_t) \tag{15}$$

The LSTM architecture used in this study consists of two learnable layers. In this study, MEG data is adjusted to 128x17 dimensions and applied to the input layer. The LSTM layer contains 100 units. Dropout value is set to 0.1. At the last stage, there is a fully connected layer containing one neuron. The training process is carried out in 200 epochs. Also, Adam optimizer was used in the training phase.

2) Gated recurrent unit

The main difference of the Gated Recurrent Unit (GRU) network from the LSTM network is that each module consists of 2 gates instead of 3, as shown in Fig. 5. A GRU module consists of an update gate and a reset gate. The update gate decides how much of the past information should be transmitted, while the reset gate, on the contrary, decides how much of the past information should be discarded.
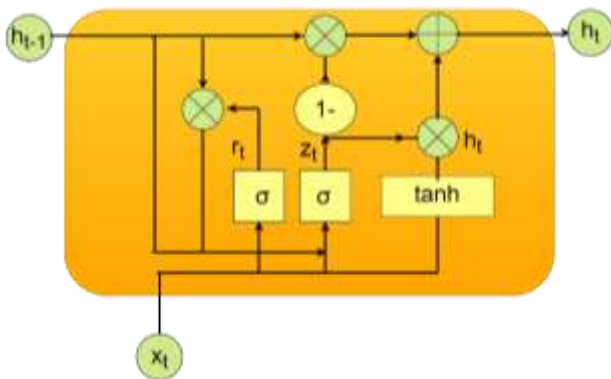
Fig.5. GRU structure

The sigmoid process representing the z_t update gate and the $\tilde{h}_t$ reset operation, GRU can be expressed mathematically as follows:

$$z_t = \sigma(W_z \cdot [h_{t-1}, x_t]) \tag{16}$$

$$r_t = \sigma(W_r \cdot [h_{t-1}, x_t]) \tag{17}$$

$$\tilde{h}_t = tanh(W \cdot [r_t * h_{t-1}, x_t]) \tag{18}$$

$$h_t = (1 - z_t) * h_{t-1} + z_t * \tilde{h}_t \tag{19}$$

GRU architecture also consists of two learnable layers, like LSTM. In the input layer, the dimensions are set to 128x17. The GRU layer contains 32 units. The dropout value is set to 0.1. Finally, there is a fully bonded layer connected to a neuron. The training process is carried out in 105 epochs. Also, Adam optimizer was used in the training phase.

E. Convolutional neural networks

CNN is a typical multi-layered neural network [32] that simulates the organization of the animal visual cortex [33]. It is widely used in image-related applications [31,34]. The operation of CNN models takes place in two stages as feature extraction and classification of these features in fully connected layers. CNN architectures are often created by combining convolution, pooling, and fully connected layers.

The convolution layer is the most important structure that makes up CNN. This layer is in principle based on the idea that an image of an object can be in any region above the image. Accordingly, neurons are attached to only a small part of the input and extend across the entire depth of the input. The filter size and number of maps produced are used to define this layer. Filters aim to extract different features related to lines, corners, and edges on the input images [12]. These filters containing pixel values are shifted on the image. During the sliding process, the filter values are multiplied by the values of the image. Then the obtained values are summed and a net result is produced. This process is applied to the whole image and feature maps are obtained. Calculation of feature map values can be expressed as follows:

$$y_l = \sum_{n=0}^{N-1} x_n h_{l-n} \tag{20}$$

where *y* is the feature map, *x* is signal, *h* is the filter, *N* is the number of elements in *x*, and the *nth* vector variable subscripts indicate the subscripts.

Another important structure that forms CNN is the pooling layer. There are different types of this layer commonly used in the literature, such as average pooling and max pooling. In this study, max-pooling was used. In the max-pooling process, the image is divided into blocks that do not overlap, and the biggest value of each block is taken. Therefore, calculation costs and overfitting possibilities are greatly reduced.

Another structure commonly used in CNN architectures is a fully connected layer. This layer is a typical artificial neural network layer. They have connections with all neurons before and after it.

Most of the problems that CNN is trying to solve are not linear. On the other hand, operations such as matrix multiplication and addition are linear. So, the non-saturating activation function is commonly used in CNN to provide non-linearity. This process can be expressed mathematically as follows:

$$f(x) = \begin{cases} x, & x \geq 0 \\ 0, & otherwise \end{cases} \quad (21)$$

The CNN architecture used in this study consists of four learnable layers, as seen in Fig. 6. In this study, 2D CNN architecture was used. MEG data is adjusted to 128x17 dimensions and applied to the input layer. The first convolutional layer takes place after the input layer. The filter number of this layer is 512, and the kernel size is 4. This layer is followed by the max-pooling layer, which is a size of 2. Then there is the second convolution layer. The values of this layer are the same as the first convolutional layer. After the second convolutional layer, there is a max-pooling layer with a size of 2. Then there is the fully connected layer of 50 neurons. Dropout is applied in this layer. The dropout value is 0.2. Finally, there is a fully connected layer with a single neuron.
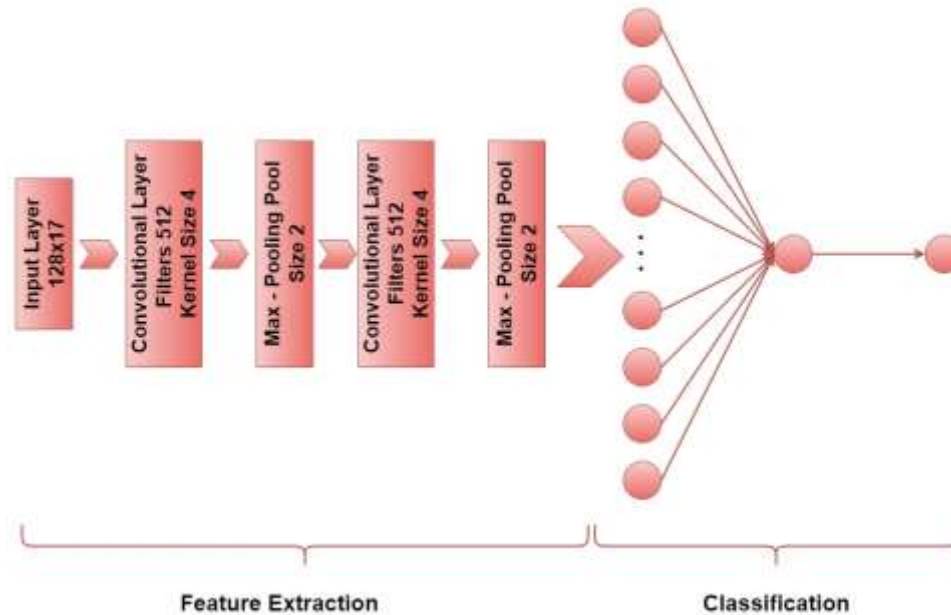
Fig.6. CNN architecture

III. RESULTS AND DISCUSSION

In this study, a binary classification process was carried out using MEG signals. Signals are grouped as a face or scrambled face. The dataset consists of 9414 stimuli belonging to 16 individuals. Approximately 588 stimuli were shown to each individual, and recording was performed through 306 channels. The noise-containing MEG signals are filtered as described in the sections above to extract 2176 features for each stimulus. For the classification process, 5 different supervised classifiers, LDA, QDA, LSTM, GRU, and CNN were used. The performance of each machine learning algorithm was compared by taking into account classification performance (ACC), Area under the curve (AUC) score, sensitivity (SENS), and specificity (SPEC) metrics. Results were given with 10k cross-validation for each algorithm. In addition, the results were compared with other studies carried out with the same dataset.

Values obtained as a result of 10k cross-verification are shown in Table 1. With an accuracy value of 92.99%, the best classification performance belongs to LSTM. Similarly, the best scores in specificity and AUC, respectively, with 92.33% and 0.977% belong to LSTM. LSTM and GRU showed quite similar performances in sensitivity value. Sensitivity was 93.64% for LSTM, while 93.73% was obtained in GRU. GRU's

performance and specificity values were 91.66% and 89.59%, respectively. CNN showed slightly lower performance in accuracy, sensitivity, and AUC values compared to RNN-based approaches. For all three metrics, 90.62%, 91.16%, and 0.959% values were obtained, respectively. On the other hand, specificity performed quite close to GRU. The specificity for CNN was 90.08%. On the other hand, QDA displayed the worst classification performance with a performance value of 72.24%. At the same time, the worst performance values in sensitivity, specificity and AUC were obtained with QDA. LDA achieved about 6% better classification performance compared to QDA. Fig. 7 shows graphs of the evaluation metrics for each machine learning algorithm.

TABLE I
AVERAGE 10K CROSS-VALIDATION RESULTS

	LSTM	GRU	CNN	LDA	QDA
ACC	**92.99**	91.66	90.62	78.23	72.24
SENS	93.64	93.73	91.16	79.56	78.56
SPEC	92.33	89.59	90.08	76.91	65.92
AUC	0.977	0.973	0.959	0.861	0.796

MEG signals may have adequate (sometimes poor) absolute locality. This means that a noticeable event can be seen at different times and slightly different frequency ranges. CNN emerged from the idea that the image of an object is independent of its location in the picture [32]. Each neuron binds only a small part of the entrance and extends across the entire depth of the entrance. In this case, it provides a distinct advantage compared to traditional machine learning algorithms such as LDA and QDA to detect patterns on poorly localized MEG signals. On the other hand, considering all comparison metrics, RNN-based machine learning algorithms perform significantly better than other methods. Over time, different neural network models have become specialized to process different data types. For example, Convolutional Neural Networks are specialized for processing matrix type information, such as image data, while Recurrent Neural Networks (RNN) have also been developed to process sequence data. The only input that traditional feed-forward neural networks take into account is existing examples to which it is exposed.
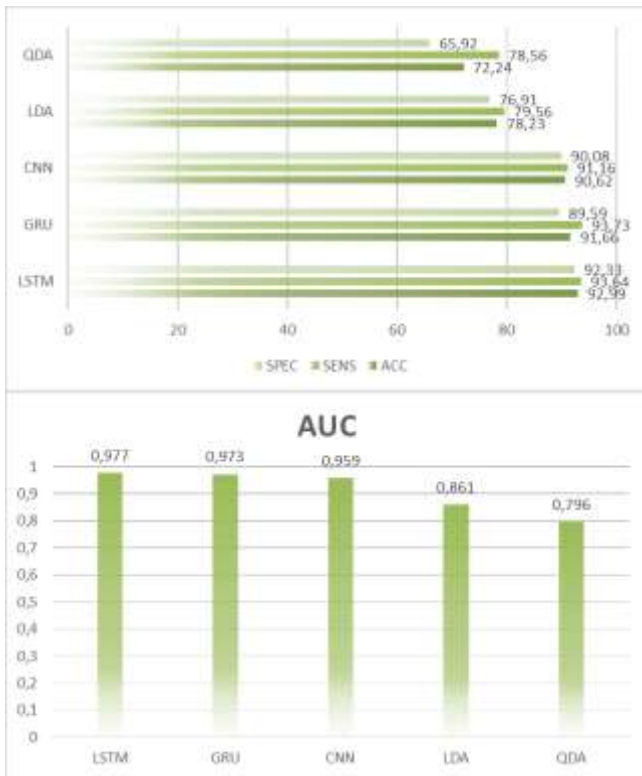

Fig.7. Performances of machine learning algorithms and AUC scores

On the other hand, RNNs also use the information for calculating the overtime, as well as the existing samples. Moreover this memory-based method (RNN) has a repetition layer, unlike traditional feed-forward neural networks. By means of this layer, the state information generated by the feed-forward network is stored and re-applied to the network with the input information. In other words, RNNs have a memory that holds what has been calculated so far. It is evaluated that these capabilities of RNNs cause them to perform better on MEG data compared to other machine learning algorithms.

Table 2 shows detailed results for LSTM 10k cross-validation results. Accordingly, the best classification performance for a fold was 94.37%. The lowest classification success was 91.61%. The highest and lowest scores for sensitivity were 95.12% and 91.30%, respectively. In specificity values, the lowest score was 90.66%, while the highest score was 94.04%. In general, the sensitivity value of the model is higher than the specificity value. In the results in Table 1, it is seen that the sensitivity values are higher than the specificity values. It can be said that all of the 5 different machine learning algorithms used generally differentiate true positive classes (meaningful face) more successfully.

TABLE 2
10K CROSS-VALIDATION RESULTS OF LSTM

	ACC	SENS	SPEC	AUC
1. Fold	91.61	91.30	91.93	0.972
2. Fold	92.46	94.27	90.66	0.978
3. Fold	91.40	90.87	91.93	0.966
4. Fold	94.16	**95.12**	93.21	0.983
5. Fold	93.62	94.06	93.19	0.982
6. Fold	92.88	94.69	91.06	0.975
7. Fold	93.41	92.78	**94.04**	0.978
8. Fold	**94.37**	95.11	93.63	**0.985**
9. Fold	92.56	94.04	91.08	0.977
10. Fold	93.41	94.26	92.57	0.975

In addition, based on Table 2, it is observed that the data distribution between folds occurs randomly in a mutually exclusive structure so that the accuracy between folds is close to each other.

TABLE 3
10K CROSS-VALIDATION RESULTS OF GRU

	ACC	SENS	SPEC	AUC
1. Fold	88.64	94.69	82.59	0.959
2. Fold	91.83	91.51	92.14	0.979
3. Fold	91.30	92.57	90.02	0.973
4. Fold	88.43	90.02	86.83	0.946
5. Fold	91.92	93.63	90.21	0.978
6. Fold	93.09	94.27	91.91	0.978
7. Fold	89.90	94.06	85.74	0.965
8. Fold	**94.58**	**95.96**	**93.21**	**0.985**
9. Fold	93.52	95.96	91.08	0.984
10. Fold	93.41	94.68	92.14	0.984

Table 3 also gives details about the 10k cross-validation results of the GRU algorithm. The highest performance value for GRU was 94.58%. On the other hand, the lowest performance value is 88.64%. The highest score for sensitivity was 95.96%. The lowest sensitivity score is 90.02%. The specificity score is generally lower than the sensitivity values, similar to other machine learning algorithms. The lowest and

http://dergipark.gov.tr/bajece

highest scores in the AUC score were 0.946 and 0.985, respectively.

The results for 10k cross-validation tests of the CNN algorithm are shown in Table 4. The highest accuracy value for CNN was 92.56%. The highest value obtained for a fold is lower compared to LSTM and GRU. The lowest accuracy value was obtained at 87.37%. Here, the lowest score was observed compared to LSTM and GRU. The lowest and highest score for sensitivity is 87.69% and 92.99%, respectively. The lowest score for specificity is 87.04%, and the highest score is 92.36%. The highest value for AUC was 0.973. The lowest score for AUC is 0.932.

Fig. 8 shows the boxplot representation of 10k cross-validation results obtained with LSTM, GRU, and CNN algorithms. Accordingly, the median value of the LSTM algorithm in accuracy is higher compared to GRU and CNN. On the other hand, the lowest median value for accuracy belongs to CNN. Similarly, the lowest median value in sensitivity is the CNN algorithm. Although the median value for Specificity is lower than LSTM, it is quite similar to GRU. The peak values of the CNN algorithm appear to be lower in all comparison metrics compared to RNN-based algorithms.

TABLE 4
10K CROSS-VALIDATION RESULTS OF CNN

	ACC	SENS	SPEC	AUC
1. Fold	91.08	90.45	91.72	0.963
2. Fold	90.76	92.99	88.54	0.962
3. Fold	88.43	87.89	88.95	0.939
4. Fold	87.37	87.69	87.04	0.932
5. Fold	91.07	91.72	90.42	0.962
6. Fold	90.75	90.44	91.06	0.970
7. Fold	91.39	92.14	90.64	0.968
8. Fold	92.56	**94.47**	90.66	0.970
9. Fold	90.44	91.49	89.38	0.957
10. Fold	**92.35**	92.34	**92.36**	**0.973**

In addition, if RNN based algorithms are analyzed, when Table 2, Table 3, and Fig. 13 are evaluated together, it is seen that there are much larger differences in the GRU algorithm between the highest and lowest values for each fold compared to LSTM. Results for GRU performance have been realized in a much wider range. The change interval of the sensitivity value occurred close to each other in both models. However, the peak value of GRU is higher. In addition, one outlier value was realized at the lower point in the GRU. Specificity value stands out as the main factor that reveals the difference between both models. The median of the specificity value of GRU is much lower than the LSTM, and the difference between the highest and lowest values is quite high. It was observed that the false-positive value was significantly higher in the GRU algorithm. Considering both the distribution of the sensitivity value and the specificity values, it can be considered that the positive class trend for the GRU algorithm is higher on the MEG dataset. GRU modules control the flow of information as in LSTM modules. But unlike LSTM, they don't have a memory unit. For this reason, LSTMs can remember longer sequences compared

to GRU [35]. Therefore, LSTMs are more successful in this task, as the evaluation of MEG signals also requires modeling long-term relationships.

In Fig. 9, the signal analysis of Subject-1 over the MEG signal recordings for face and scrambled face were shown to explore the cortical decoding at the first trial and the third trial, respectively. Notably, the face visual stimulation related power increment has been investigated between 8-18 Hz after the visual stimulation, as reported in the previous research study [18,36]. Moreover, this initial power increment of the evoked component has arisen around 170ms and/or 220ms. Then again, as expected, the negative deflection of N170 (the Event Related Potential-ERP) has occurred greater for face visual stimuli than scrambled visual stimuli around 170ms [37].
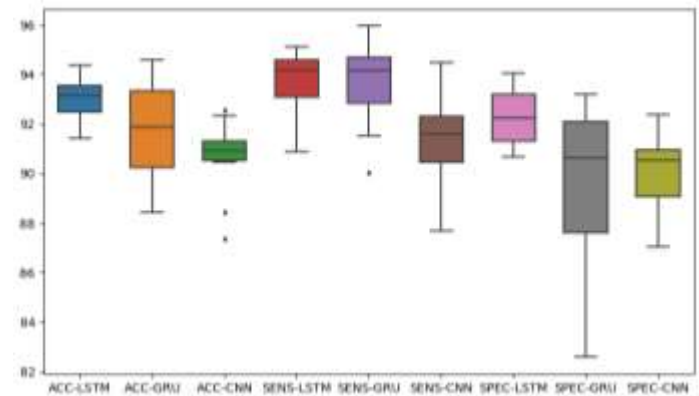
Fig.8. Boxplot graphics of LSTM, GRU and CNN algorithms

All these results are coherent with the previous outcomes connected to the N/M170 cortical activation [18,36-37]. Furthermore, the latency for scrambled face visual stimulation was observed that this might be due to the brain perception mechanism. The functioning of face perception has been observed as an automatic, rapid, and subconscious process which already has been seen in human newborns. Preferably, the simple schematic (such as scrambled face) may be seen for face-like patterns. Therefore, the face-like stimuli pattern can be perceived as faces for participants in the experiment. The tendency of the brain to see the face is called the phenomenon of pareidolia. Hence, the face-specificity of the N170 is a challenge for neurological studies [36]. Nonetheless, the development of more generative, complex and realistic comments from the neuroimaging data requires more multi-subject and multi-modal analysis.

Furthermore, according to the reported study, the stages of the stimulation performances in terms of the machine learning algorithms show that the perception stage can provide much higher accuracy than the pre-stimuli stage. In our study, deep learning algorithms may have yielded distinguished performances due to the better perception stage representation of the spatial and temporal features of the MEG signals [38]. t-SNE visualization was shown in Fig.10.

The confusion matrices were also determined for GRU and LSTM deep methods, as shown in Fig.11. Finally, some other studies performed with MEG dataset in Table 5 and the results of RNN-based models and CNN model that provide the best results in this study are presented together. Considering the

results in the table, it can be seen that RNN-based approaches perform significantly better. With the LSTM model, an improvement of 12.14% was achieved, which corresponds to the 63.39% relative error reduction rate compared to the DAE model, which provided the best results before. Besides, the CNN model appears to give significantly better results compared to other studies in the literature.

IV. CONCLUSION

The main purpose of this study is to investigate the classification performances of MLs over the MEG signals which were recorded during the human brain's response to visual stimuli to be decoded the brain functioning of face perception mechanism. There are two classes: face and scramble face in the classification process. MEG signals are very difficult to classify as they contain high amounts of noise. In this study, the classification performances were compared by using LSTM, GRU, CNN, LDA, and QDA algorithms. The CNN algorithm appears to provide a distinct advantage in capturing weakly localized MEG signals compared to LDA, QDA, and other studies with the same dataset. With the CNN algorithm, 90.62% and 0.959% values were obtained for accuracy and AUC, respectively.

On the other hand, the best results were obtained with RNN based algorithms. RNN algorithms cannot use only the existing information they are exposed to as input. In addition to this information, they use the information they calculate overtime. Therefore, they differ from traditional neural networks. In this study, the best results in all comparison metrics except sensitivity were obtained with the LSTM algorithm. Quite similar values were obtained with the GRU for the sensitivity metric. The LSTM model, 92.99%, 93.64%, 92.33%, and 0.977 values were obtained for accuracy, sensitivity, specificity, and AUC, respectively.
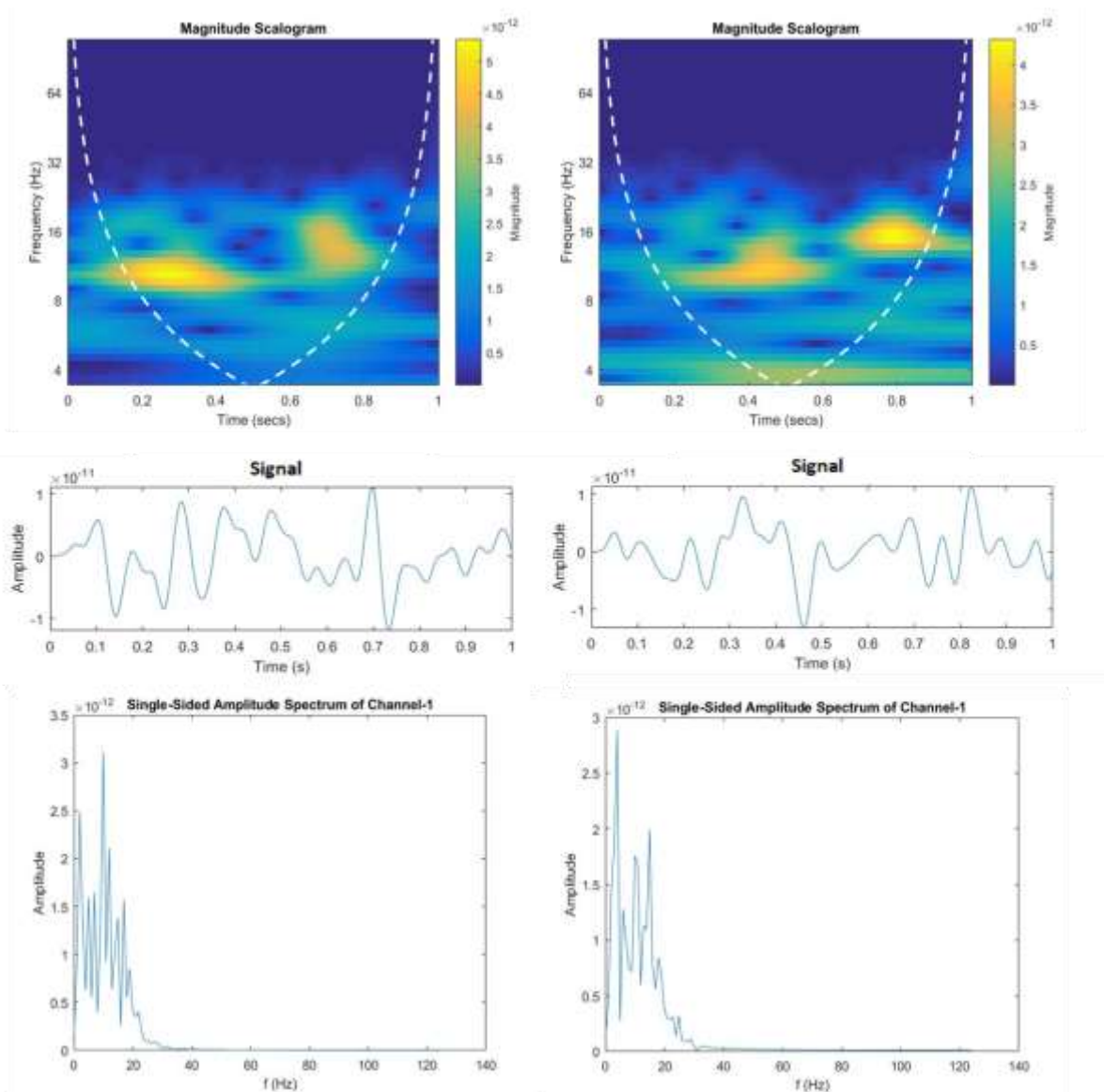

Fig.9. Signal, Scalogram and Single-Sided amplitude spectrum presentation for face visual stimuli (*left*) Signal, Scalogram and Single-Sided amplitude spectrum presentation for scrambled face visual stimuli (*right*) (Channel-1 signals are represented for all figures)
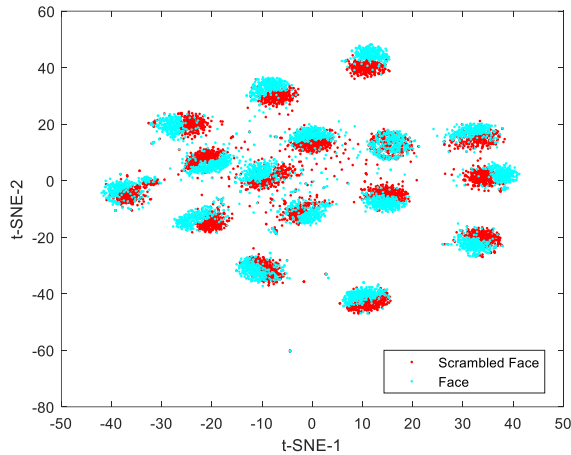
Fig.10. t-SNE visualization for feature extracted dataset via Riemannian approach

(a)

(b)
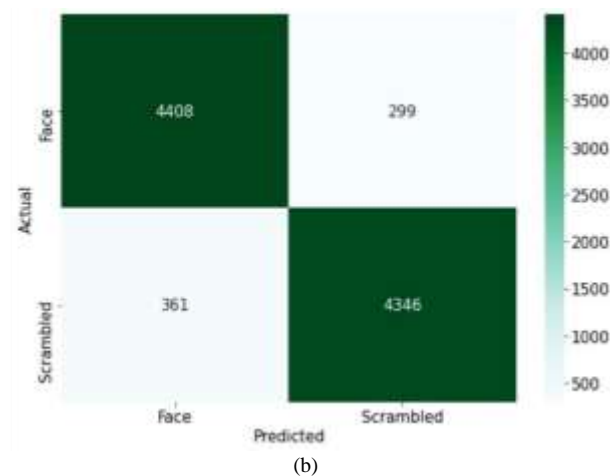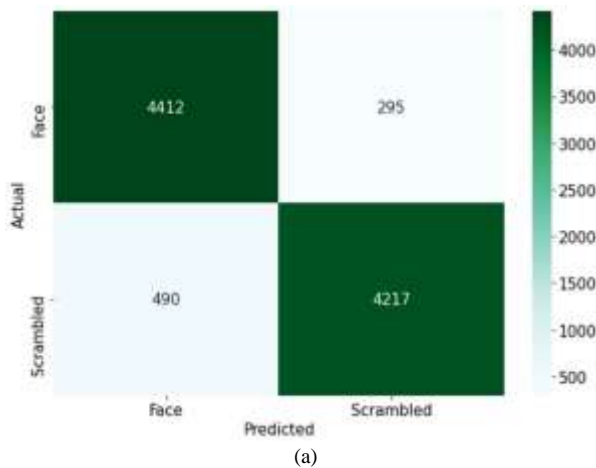
Fig. 11. The confusion matrices for GRU (*a*) and LSTM (*b*) related to the average results of 10-fold cross validation

TABLE 5
EXISTING METHODS USING THE MEG DATA AND CLASSIFICATION ACCURACIES FOR FACE/SCRAMBLED FACE RECOGNITION IN THE LITERATURE

STUDY	METHOD	NEURAL DECODING	ACC (%)
[5]	DNN	Face/Scramble Face Decoding	**80.85**
	SVM		78.01
	KNN		72.84
	NB		71.92
	DT		68.36
[10]	LVQ	Face/Scramble Face Decoding	**69.39**
[39]	SVM	Face/Scramble Face Decoding	**74.85**
[11]	GRNN	Face/Scramble Face Decoding	**79**
[40]	Hybrid GRU	Face/Scramble Face Decoding	**71.20**
[41]	SVM	Face, Tool, Animal, Scene Decoding	**84**
[9]	MLNN	Face/Scramble Face Decoding	77.78
	PNN		**82.36**
This Study	LSTM	Face/Scramble Face Decoding	**92.99**
	GRU		91.66
	CNN		90.62
	LDA		78.23
	QDA		72.24

The ability of LSTMs to learn long and short-term dependencies has provided a distinct advantage over other algorithms used in the MEG dataset. In future studies, the LSTM algorithm can be used to study and compare cortical activities of various regions of the brain. Moreover, applications of intelligent Internet of Things (IoT) need universal and trustworthy biometric authentication system [42-43]. To address these issues, paradigm of a visual presentation (face/scrambled face) can be proposed to use the MEG signals of subjects due to the great spatial and temporal resolution with specific stimulation of individual's brainwave pattern.

Human and animal rights

This article does not contain any studies with human participants or animals performed by any of the authors. The used data in this paper is taken from the Kaggle data science repository. It is a public dataset.

REFERENCES

[1] Zarief, C. N., & Hussein, W. (2019). Decoding the Human Brain Activity and Predicting the Visual Stimuli from Magnetoencephalography (MEG) Recordings. In Proceedings of the 2019 International Conference on Intelligent Medicine and Image Processing - IMIP '19 (pp. 35–42). New York, New York, USA: ACM Press. https://doi.org/10.1145/3332340.3332352

[2] Lin, J.-F. L., Silva-Pereyra, J., Chou, C.-C., & Lin, F.-H. (2018). The sequence of cortical activity inferred by response latency variability in the human ventral pathway of face processing. Scientific Reports, 8(1), 5836. https://doi.org/10.1038/s41598-018-23942-x

[3] Watanabe, S., Miki, K., & Kakigi, R. (2005). Mechanisms of face perception in humans: A magneto- and electro-encephalographic study. Neuropathology, 25(1), 8–20. https://doi.org/10.1111/j.1440-1789.2004.00603.x

[4] Tadel, F., Bock, E., Niso, G., Mosher, J. C., Cousineau, M., Pantazis, D., … Baillet, S. (2019). MEG/EEG Group Analysis With Brainstorm. Frontiers in Neuroscience, 13(FEB), 1–21. https://doi.org/10.3389/fnins.2019.00076

[5] Caliskan, A., Yuksel, M. E., Badem, H., & Basturk, A. (2017). A deep neural network classifier for decoding human brain activity based on magnetoencephalography. Elektronika Ir Elektrotechnika, 23(2), 63–67. https://doi.org/10.5755/j01.eie.23.2.18002

[6] Özkaya, Ş. N., & Yıldırım, T. (2018). Assessment of Components and Methods Used to Identify Responses and Regions of Brain Related with Face Recognition and Perception. Procedia Computer Science, 131, 38–44. https://doi.org/10.1016/j.procs.2018.04.183

[7] http://web.mit.edu/kitmitmeg/whatis.html. (n.d.). Retrieved from http://web.mit.edu/kitmitmeg/whatis.html

[8] Kia, S. M., Vega Pons, S., Weisz, N., & Passerini, A. (2017). Interpretability of Multivariate Brain Maps in Linear Brain Decoding: Definition, and Heuristic Quantification in Multivariate Analysis of MEG Time-Locked Effects. Frontiers in Neuroscience, 10. https://doi.org/10.3389/fnins.2016.00619.

[9] Cetin, O., & Temurtas, F. (2020). A comparative study on classification of magnetoencephalography signals using probabilistic neural network and multilayer neural network. Soft Computing. https://doi.org/10.1007/s00500-020-05296-7

[10] Çetin, O., & Temurtaş, F. (2018). A Study on Brain Computer Interface using Learning Vector Quantization. Sakarya University Journal of Computer and Information Sciences, 1(2), 1–7.

[11] Çetin, O., & Temurtaş, F. (2019). Classification of Magnetoencephalography Signals Regarding Visual Stimuli by Generalized Regression Neural Network. Dicle Tıp Dergisi, 45(3), 19–25. https://doi.org/10.5798/dicletip.534819.

[12] Gorur, K., Bozkurt, M., Bascil, M., & Temurtas, F. (2019). GKP Signal Processing Using Deep CNN and SVM for Tongue-Machine Interface. Traitement Du Signal, 36(4), 319–329. https://doi.org/10.18280/ts.360404

[13] Yeom, H. G., Kim, J. S., & Chung, C. K. (2020). LSTM Improves Accuracy of Reaching Trajectory Prediction From Magnetoencephalography Signals. IEEE Access, 8, 20146–20150. https://doi.org/10.1109/ACCESS.2020.2969720

[14] Alom, M. Z., Moody, A. T., Maruyama, N., Van Essen, B. C., & Taha, T. M. (2018). Effective Quantization Approaches for Recurrent Neural Networks. In Proceedings of the International Joint Conference on Neural Networks (Vol. 2018-July). https://doi.org/10.1109/IJCNN.2018.8489341

[15] Alpaydın, E. (2010). Introduction to Machine Learning. MIT Press, Cambridge, Massachusetts.

[16] Ghojogh, B., & Crowley, M. (2019). Linear and Quadratic Discriminant Analysis: Tutorial, (4), 1–16. Retrieved from http://arxiv.org/abs/1906.02590

[17] Barachant, A. (2014). MEG decoding using Riemannian Geometry and Unsupervised classification. Notes on the winner of the Kaggle "DecMeg2014 - Decoding the Human Brain" competition., 1–8.

[18] Henson, R. N., Wakeman, D. G., Litvak, V., & Friston, K. J. (2011). A Parametric Empirical Bayesian Framework for the EEG/MEG Inverse Problem: Generative Models for Multi-Subject and Multi-Modal Integration. Frontiers in Human Neuroscience, 5(August), 1–16. https://doi.org/10.3389/fnhum.2011.00076

[19] "DecMeg2014-Decoding the Human Brain". [Online]. Available: https://www.kaggle.com/c/decoding- the-human-brain. (n.d.), 2014.

[20] Olivetti, E., Kia, S. M., & Avesani, P. (2014). MEG decoding across subjects. Proceedings - 2014 International Workshop on Pattern Recognition in Neuroimaging, PRNI 2014. https://doi.org/10.1109/PRNI.2014.6858538

[21] Roy, P. P., Kumar, P., & Chang, V. (2020). A hybrid classifier combination for home automation using EEG signals. Neural Computing and Applications, 32(14), 1–19. https://doi.org/10.1007/s00521-020-04804-y

[22] Yger, F., Berar, M., & Lotte, F. (2017). Riemannian Approaches in Brain-Computer Interfaces: A Review. IEEE Transactions on Neural Systems and Rehabilitation Engineering, 25(10), 1753–1762. https://doi.org/10.1109/TNSRE.2016.2627016

[23] Barachant, A., Bonnet, S., Congedo, M., & Jutten, C. (2012). Multiclass brain-computer interface classification by Riemannian geometry. IEEE Transactions on Biomedical Engineering, 59(4), 920–928. https://doi.org/10.1109/TBME.2011.2172210

[24] Uçar, M. K. (2020). Classification Performance-Based Feature Selection Algorithm for Machine Learning: P-Score. IRBM, 41(4), 229–239. https://doi.org/10.1016/j.irbm.2020.01.006

[25] Bascil, M. S., Tesneli, A. Y., & Temurtas, F. (2015). Multi-channel EEG signal feature extraction and pattern recognition on horizontal mental imagination task of 1-D cursor movement for brain computer interface. Australasian Physical and Engineering Sciences in Medicine, 38(2), 229–239. https://doi.org/10.1007/s13246-015-0345-6

[26] Hossin, M., & Sulaiman, M. . (2015). A Review on Evaluation Metrics for Data Classification Evaluations. International Journal of Data Mining & Knowledge Management Process, 5(2), 01–11. https://doi.org/10.5121/ijdkp.2015.5201

[27] Toğaçar, M., Ergen, B., Cömert, Z., & Özyurt, F. (2020). A Deep Feature Learning Model for Pneumonia Detection Applying a Combination of mRMR Feature Selection and Machine Learning Models. IRBM, 41(4), 212–222. https://doi.org/10.1016/j.irbm.2019.10.006

[28] Yang, S., & Berdine, G. (2017). The receiver operating characteristic (ROC) curve. The Southwest Respiratory and Critical Care Chronicles, 5(19), 34. https://doi.org/10.12746/swrccc.v5i19.391

[29] Muller, K.-R., Mika, S., Ratsch, G., Tsuda, K., & Scholkopf, B. (2001). An introduction to kernel-based learning algorithms. IEEE Transactions on Neural Networks, 12(2), 181–201. https://doi.org/10.1109/72.914517

[30] Koudjonou, K. M., & Rout, M. (2020). A stateless deep learning framework to predict net asset value. Neural Computing and

Applications, 32(14), 1–19. https://doi.org/10.1007/s00521-019-04525-x

[31] Lecun, Y., Bottou, L., Bengio, Y., & Haffner, P. (1998). Gradient-based learning applied to document recognition. Proceedings of the IEEE, 86(11), 2278–2324. https://doi.org/10.1109/5.726791

[32] Ozer, I., Ozer, Z., & Findik, O. (2018). Noise robust sound event classification with convolutional neural network. Neurocomputing, 272, 505–512. https://doi.org/10.1016/j.neucom.2017.07.021

[33] Hubel, D. H., & Wiesel, T. N. (1968). Receptive fields and functional architecture of monkey striate cortex. The Journal of Physiology, 195(1), 215–243. https://doi.org/10.1113/jphysiol.1968.sp008455

[34] Ma, J., Wu, F., Zhu, J., Xu, D., & Kong, D. (2017). A pre-trained convolutional neural network based method for thyroid nodule diagnosis. Ultrasonics, 73, 221–230. https://doi.org/10.1016/j.ultras.2016.09.011

[35] Golmohammadi, M., Ziyabari, S., Shah, V., Von Weltin, E., Campbell, C., Obeid, I., & Picone, J. (2017). Gated recurrent networks for seizure detection. In 2017 IEEE Signal Processing in Medicine and Biology Symposium (SPMB) (Vol. 2018-Janua, pp. 1–5). IEEE. https://doi.org/10.1109/SPMB.2017.8257020

[36] Hadjikhani, N., Kveraga, K., Naik, P., & Ahlfors, S. P. (2009). Early (M170) activation of face-specific cortex by face-like objects. NeuroReport, 20(4), 403–407. https://doi.org/10.1097/WNR.0b013e328325a8e1

[37] Wakeman, D. G., & Henson, R. N. (2015). A multi-subject, multi-modal human neuroimaging dataset. Scientific Data, 2(1), 150001. https://doi.org/10.1038/sdata.2015.1

[38] Dash, D., Ferrari, P., & Wang, J. (2020). Decoding Imagined and Spoken Phrases From Non-invasive Neural (MEG) Signals. Frontiers in Neuroscience, 14(September 2004), 8–20. https://doi.org/10.3389/fnins.2020.00290

[39] Gupta, S., & Gandhi, T. (2020). Identification of Neural Correlates of Face Recognition Using Machine Learning Approach. In Computer Vision and Machine Intelligence in Medical Image Analysis (pp. 13–20). https://doi.org/10.1007/978-981-13-8798-2_2

[40] Li, J., Pan, J., Wang, F., & Yu, Z. (2021). Inter-Subject MEG Decoding for Visual Information with Hybrid Gated Recurrent Network. Applied Sciences, 11(3), 1215. https://doi.org/10.3390/app11031215

[41] Liu, C., Kang, Y., Zhang, L., & Zhang, J. (2021). Rapidly Decoding Image Categories From MEG Data Using a Multivariate Short-Time FC Pattern Analysis Approach. IEEE Journal of Biomedical and Health Informatics, 25(4), 1139–1150. https://doi.org/10.1109/JBHI.2020.3008731

[42] Huang, H., Hu, L., Xiao, F., Du, A., Ye, N., & He, F. (2019). An EEG-Based Identity Authentication System with Audiovisual Paradigm in IoT. Sensors, 19(7), 1664. https://doi.org/10.3390/s19071664

[43] Sooriyaarachchi, J., Seneviratne, S., Thilakarathna, K., & Zomaya, A. Y. (2021). MusicID: A Brainwave-Based User Authentication System for Internet of Things. IEEE Internet of Things Journal, 8(10), 8304–8313. https://doi.org/10.1109/JIOT.2020.3044726

BIOGRAPHIES

ZEYNEP ÖZER was born in Ankara, Turkey, in 1986.

In 2010, she graduated from Dumlupınar University with a bachelor's degree in mathematics. She received a M.Sc. degree from the Department of Computer Engineering at the University of Sakarya, Turkey, in 2012. Also, she received a Ph. D. degree from the Computer Engineering at the University of Karabük, Turkey, in 2019.

Since 2019, she has been working as an assistant professor in the Department of Management Information Systems at Bandırma Onyedi Eylül University.

Her current research interests include natural processing language, artificial intelligence, applied mathematics and advanced machine learning techniques.

ONURSAL ÇETİN was born in Kayseri, Turkey, in 1982. He received the B.S. and M.S. degrees from Erciyes University, Kayseri, Turkey, in 2006 and 2009, respectively, and the Ph.D. degree from Sakarya University, Sakarya, Turkey, in 2014, all in electrical and electronics engineering.

From 2007 to 2014, he was a Research Assistant with Bozok University, Yozgat, Turkey. From 2014 to 2018, he worked as an Assistant Professor with Bozok University. Since 2018, he has been an Assistant Professor with the Electrical and Electronics Engineering Department, Bandırma Onyedi Eylul University, Bandırma, Turkey. His research interests include artificial intelligence, signal processing, and embedded systems.

KUTLUCAN GÖRÜR was born on April 13, 1985 in Adana. He completed his primary, high school and university education in this city. He graduated from Adana Anatolian High School in 2003. After graduating from Çukurova University Electrical and Electronics Engineering in 2009, he worked as a Research Assistant at Bozok University Electrical and Electronics Engineering Department in 2010. started his life. After graduating from Bozok University Institute of Science and Mechatronics Engineering master's program in 2014, he enrolled in a doctoral program at Sakarya University Institute of Science, Department of Electrical and Electronics Engineering in 2015. He is currently working as an assistant professor at Bandırma Onyedi Eylül University, Department of Electrical and Electronics Engineering.

FEYZULLAH TEMURTAŞ received the B.Sc. degree in electrical and electronics engineering from Middle East Technical University, Ankara, Turkey, in 1991, the M.Sc. degree in electrical and electronics engineering from Kütahya Dumlupınar University, Kütahya, Turkey, in 1995, and the Ph.D. degree in electrical and electronics engineering from Sakarya University, Serdivan, Turkey, in 2000.

He is currently a Professor with Bandırma Onyedi Eylül University, Bandırma, Turkey. He has substantial experience of data processing, machine learning, electronic nose, biomedical, control and computer architecture.

A Detailed Comparison of Two Different Switched Reluctance Motor's Parameters and Dynamic Behaviors by applying PID Control in Matlab Simulink

Bekir Gecer, N.Fusun Serteller and Huseyin Calik

Abstract—The Switched Reluctance Motor (SRM) is one of the oldest types of electric motors. SRMs have high torque fluctuations as they have a salient pole structure. The popularity of Switched Reluctance Motors has increased in recent years due to their simple structure, ruggedness, reliability, as well as being inexpensive to manufacture and having a high torque to mass ratio. However, there are some disadvantages, such as high torque ripple, noise and the need for an advanced control system. So they are not extensively used in industry. This study focused on the torque ripple and the speed control aspects. Here, the two most popular and similar Switched Reluctance Motors with regard to motor size, called the 6/4 and 8/6 SRM, have been compared according to the dynamic behavior of their motor parameters. Also, the converters, loads and given power values are the same for the two motor systems. The motor parameters were controlled via MatLab Simulink software. Although the 6/4 and 8/6 models are identical and have the same power converters systems and position sensors, they show different motor behavior due to their dissimilar magnetic structures.

Index Terms—**Switched Reluctance Motor, Simulink, PID Control.**

BEKIR GECER, is with Machine and Metal Technology Department of Atasehir Adiguzel Myo, Istanbul, Turkey,(e-mail: bekirgecer@adiguzel.edu.tr).

https://orcid.org/ 0000-0002-7803-3844

N.FUSUN SERTELLER, is with Electrical and Electronic Engineering Department of Marmara University, Istanbul, Turkey,(e-mail: fserteller@marmara.edu.tr).

https://orcid.org/ 0000-0003-3147-2740

HUSEYIN CALIK, is with Electrical and Electronic Engineering Department of Giresun University, Istanbul, Turkey,(e-mail: huseyin.calik@giresun.edu.tr).

https://orcid.org/ 0000-0001-8298-8945

I. INTRODUCTION

THE Switched Reluctance Motors (SRMs) are early members of the electric machine family. SRMs have high torque fluctuations as they have a salient pole structure. The popularity of Switched Reluctance Motors has increased in recent years due to their simple structure, ruggedness, reliability, and being inexpensive to manufacture and having a high torque to mass ratio. The SRM is a singly or doubly salient motor, and torque is produced by reluctance. Therefore, the rotor of the SRM has a tendency to move to a position of alignment where the inductance of the excited windings is maximized [1]–[4].Because of the advantages of SRMs and improvements in their motor control and drive systems, they are used in industrial and engineering applications. With the development of driver systems, more effective and efficient control strategies can be designed in order to increase the motor performance and system power efficiencies by decreasing the size and cost of the driver components. In conventional terms, the SRM has some disadvantages: it needs a position sensor and an advanced control system. It has a double salient structure that causes high torque ripple and noise. The torque ripple is the primary problem of SRMs in applications that require stable dynamic performance and low torque ripple [2], [5], [6]. The Ansys/Maxwell software program was also used in the study at reference [19] to understand SRM by analyzing. As well as SRM design studies; novel driver methods applications were worked for SRM [16]. An induction machine and a SRM have been compared for E-Scooter based driving cycle design in [17].

This study compared the almost identical 6/4 and 8/6 doubly salient reluctance motors according to the effect on dynamic behavior of their motor parameters I (Amps), V (Volts), Te (induced electromagnetic torque), ω (angular velocity), Φ (flux) and Θ (angle) used as control parameters by the MatLab Simulink software, respectively. See Fig 1 (a) and (b) illustrating the physical structures.
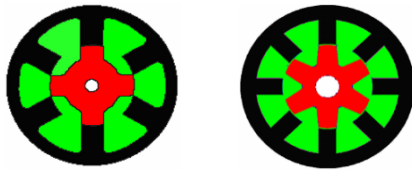
a) 6/4 Three Phase SRM   b) 8/6 Four Phase SRM
Fig 1. SRM Configurations [8].

The alignment position of both motors is given in Fig 2(a). In a magnetic circuit, the rotating part wants to reach the position of minimum reluctance at the instance of excitation. The unaligned positions of both motors are shown in Fig 2(b). Then, the set of stator poles is excited to bring the rotor poles into alignment.

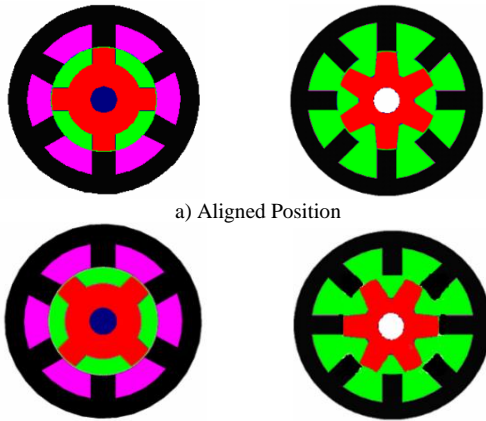a) Aligned Position

b) Unaligned Position
Fig 2. Operations of SRM [8].

II. DYNAMIC MODEL OF SRM

The dynamic model of SRM can be obtained as the following equations. These equations can obtain equivalent circuit seen in Fig 3.
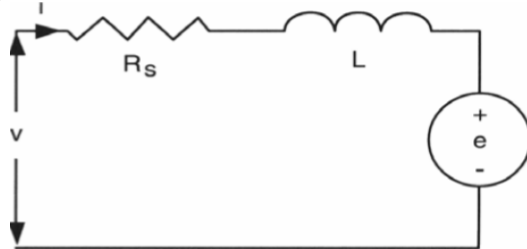

Fig 3. Single phase equivalent circuit of SRM.

The applied voltage to a phase is equal to the sum of the resistive voltage drop and the rate of the flux linkages and is given as Equation (1):

$$V = R_s i + \frac{d\lambda(\theta,i)}{dt}$$

(1)

where $R_s \, (\Omega)$ is the resistance per phase, i(A) is the phase current, θ(rad) is the rotor position and λ(Wb) the total flux linkage per phase given by:

$$\lambda = L(\theta,i)\,i$$

(2)

where L(H) is the inductance depended on the rotor position and phase current. Then, the phase voltage is derived:

$$V = R_s i + \frac{d(L(\theta,i)i)}{dt} = R_s i + L(\theta,i)\frac{di}{dt} + \frac{dL(\theta,i)}{d\theta} i\,\omega_m$$

(3)

Equation (3) consist of the three terms on te right-hand side: the resistive voltage drops, inductive voltage drop and induced emf, respectively. K_d is back-emf constant, $\omega_s \, (rad/s)$ is the angular speed of the rotor, the induced emf, e(V) can expressed briefly as:

$$e = \frac{dL(\theta,i)}{d\theta}\,\omega_m\,i = K_b\,\omega_m\,i$$

(4)

The electrical power is then:

$$V\,i = i^2\,R + L(\theta,i)i\frac{di}{dt} + i^2\,\omega_m\frac{dL(\theta,i)}{d\theta}$$

(5)

The mechanical power which is equivalent to where is the instantaneous electromagnetic torque [18].

$$T_e = \frac{1}{2}i^2\frac{dL(\theta,i)}{d\theta}$$

(6)

The mechanical equation is:

$$J\frac{d\omega_m}{dt} = T_e + T_L - B\omega_m$$

(7)

where $J\,(kg.m^2)$ is inertia moment, B is friction coefficient.

TABLE 1. SRM Motor Parameters

SRM	6/4	8/6
Stator poles number	6	8
Rotor poles number	4	6
Stator resistant(ohm)	0.05	0.05
Friction coefficient (B.N.m.s)	0.02	0.005
Inertia J (kg.m^2)	0.05	0.01
Unaligned Inductance(H)	0.67e-3	0.67e-3
Aligned Inductance(H)	23.6e-3	23.6e-3
Voltage(V)	240	240
Current(A)	0-50	0-30

III. SIMULINK MODEL OF SRM WITH SPEED CONTROL TECHNIQUES

The Switched Reluctance Motor Control Simulink model has four main blocks. These are the controller block, the converter block, the position sensor block and the SRM blocks. The position sensor block is linked to the rotor of the Switched Reluctance Motor to measure the turn-on and turn-off angles of the Switched Reluctance Motor phases. These angles are necessary for switching the converter system. At the same time, the measured current and reference current are compared to generate drive signals for the converter systems. Then, the hysteresis controller controls the currents independently [21-24].

To analyze the performance of the SRM, the model of the entire drive system was developed in the position sensor mode in the simulation software Matlab/Simulink. The parameters for the voltage supply, V= 240 volts (DC) and the stator resistance, RS = 0.05 ohms/phase, were used for both the 6/4 and 8/6 SRM, respectively. The motor was loaded at TL =10 Nm between 0.1s to 0.2s, and raised by 20 Nm after 0.2 s and kept there until the end of the simulation. A SRM

model is given in Fig 4 and the SRM Control Simulink Model of both motors is shown in Fig 5 and Fig 6. The motor parameters of both motors are seen Table 1. The PID controller values for 6/4 SRM are P=0.8, I=0.4 and D=0.03. Because of different poles number, 8/6 SRMs have different PID controller values are P=1.2, I=0.1 and D=0.01. If we change our reference speed, PID coefficient will be differ automatically. The PID gains of both SRMs are given in the Table 2.

TABLE 2. Controller PID Gains

PID GAIN	6/4	8/6
Proportional	0.8	1.2
Integral	0.4	0.1
Derivative	0.03	0.01
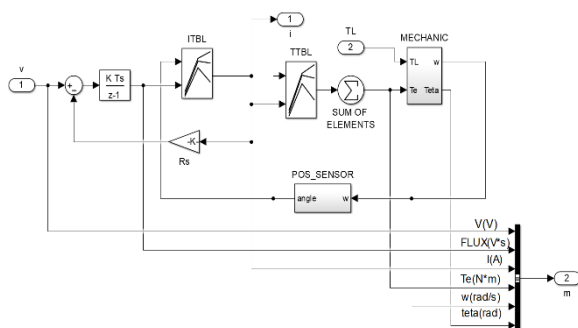
Fig 4. SRM Simulink Model.
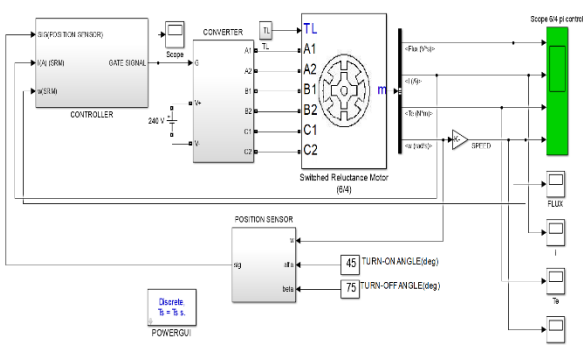
Fig 5. 6/4 SRM Control Simulink Model.
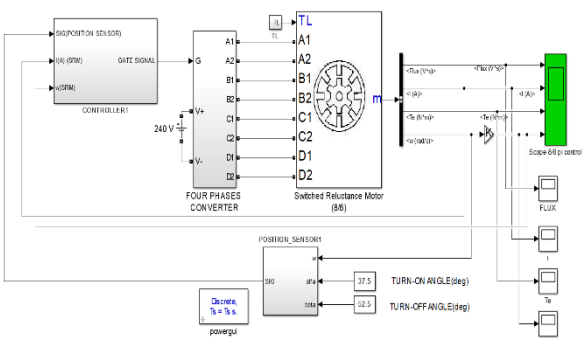
Fig 6. 8/6 SRM Control Simulink Model.

3. 1. Converter Topology

In this study, an asymmetric converter was used, which is the most usual and appropriate converter topology for the SRM. It is in the '2-switch per phase (2q switch)' converter group. The general structure of the asymmetric converter system is presented in Fig 7. Turning on the two power switches in each phase circulated a current in that phase of the SRM. If the current rose above the configd value, the switches were turned off. The energy stored in the motor phase winding keeps the current in the same direction until it is depleted [12, 19].
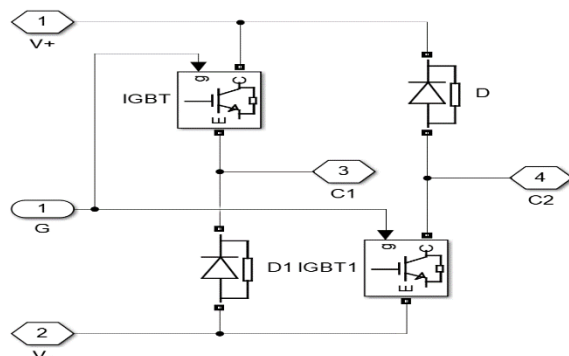
Fig 7. Converter Simulink Model for per phase.

3. 2. Controller Strategy

Fig 8 illustrates the Simulink diagram of the SRM controller that was used in simulations. It represents the Simulink diagrams for the SRM speed and position control. The position sensor value determines the command turn-on and turn-off, respectively, and generates the gate pulse for switching the SRM power converter. The gains of the PID controller, Kp, Ki and Kd, were adapted according to the rotor speed and they are adjustable. In the same way the SRM speed is controlled, the torque of the SRM is controlled by the feedback current. The 6/4 SRM shows more stable behavior than the 8/6 SRM in terms of speed control.
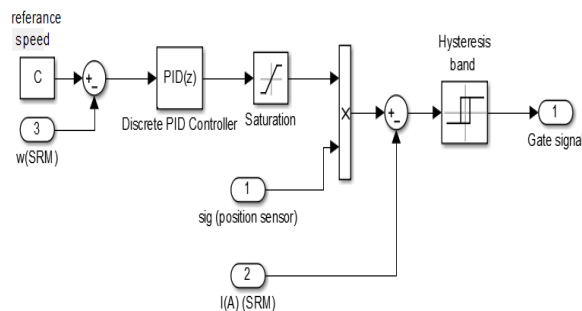
Fig 8. Controller Simulink Model.

3. 3. Position Sensor

In the SRM control systems, the rotor position is essential for the stator phase commutation and advanced angle control. The rotor position is acquired by the position sensors, which is shown in Fig 9. These blocks compare the rotor position with the reference turn-on and turn-off angle values to convert the output speed to the angle. For each motor, these

reference angles can be changed due to the different motor structures. The position of both the 6/4 and 8/6 SRM was determined by taking the integral of the angular velocity and power supplied according to the turn-on and turn-off angles.
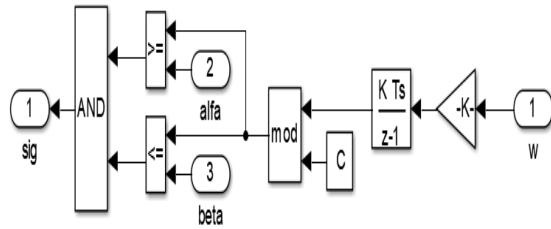
Fig 9. Position Sensor Simulink Model.

IV. SIMULINK RESULTS AND DISCUSSION OF SRM'S DYNAMIC BEHAVIOR

Although 6/4 and 8/6 SRM models are almost identical, they show different motor performance due to their individual dynamic behavior. The motor parameters compared were I (A), the phase current, Te (Nm), the electromagnetic torque produced, ω (rad/s), the angular velocity of the rotor and Φ (Wb), the magnetic flux. As seen in Fig10, both the 6/4 and 8/6 SRMs are subject to a 10 Nm load between 0.11s and 0.20s. A second load of 20 Nm is applied after 0.20s. The behavior of the flux is seen in Fig 10. Because the flux depends on the current, the flux of the 6/4 SRM is greater than the flux of the 8/6 SRM, by about 0.5 Wb. When the load is being increased, the 6/4 SRM has higher flux values than the 8/6 SRM.
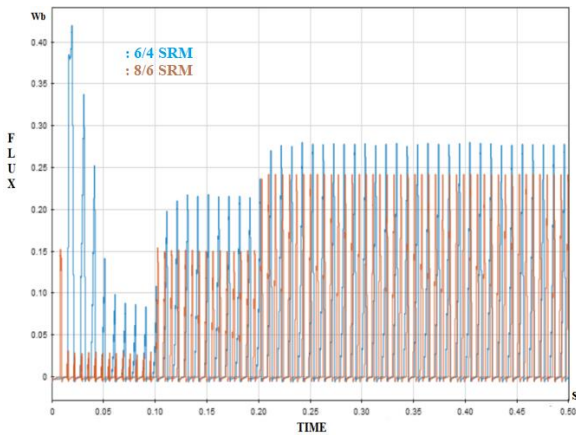
Fig 10. 6/4 and 8/6 SRM Flux Comparison.

When the loads are applied to the motors, changes in the motor currents are shown in Fig 11. It is seen that the 6/4 SRM draws 50% higher current than the 8/6 SRM. Because the 6/4 SRM's flux is higher than the 8/6 SRM.
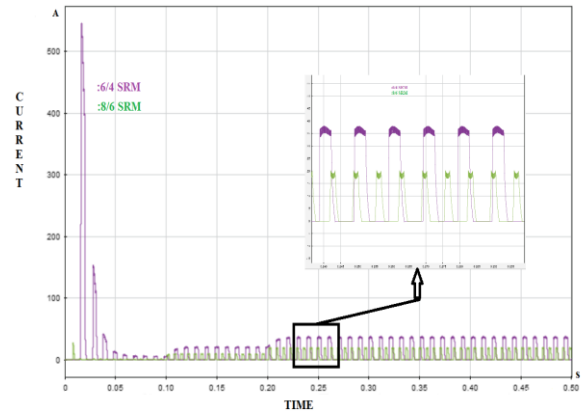
Fig 11. 6/4 and 8/6 SRM Current Comparison.

At a 10 Nm load the torque ripple 59.3% and at 20 Nm the torque ripple is 52%. At all load values, the 6/4 SRM current level was more than the 8/6 SRM current level due to the impact of its magnetic behavior.
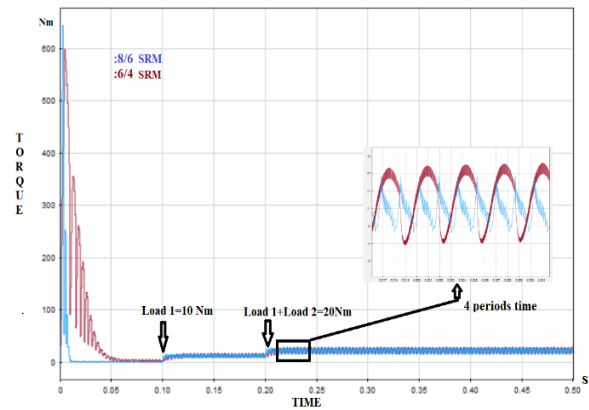
Fig 12. 6/4 and 8/6 SRM Torque Comparison.

Since the number of poles in the 8/6 SRM is more than the 6/4 SRM, the 8/6 torque ripple amplitude is less about 25% than the 6/4 torque ripple, as seen in Fig 12.

Under the loads conditions, the 8/6 SRM speed performance is higher than the 6/4 SRM and after 0.21 s the characteristics of the 8/6 SRM are more stable than the 6/4 SRM as seen in Fig 13.
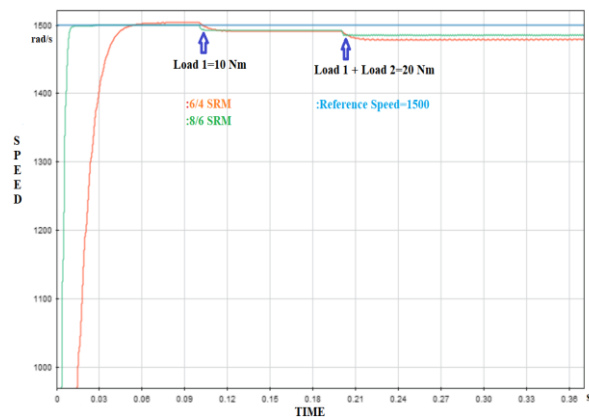
Fig 13. 6/4 and 8/6 SRM Speed Comparison.

Fig 14 shows the flux current variation of the 6/4 SRM. As can be seen from this fig, the continuous current value of the 6/4 SRM reaches 35 A. When the motor poles are in the centered position, the flux value is 0.025 Wb and when the motor poles are in the opposite position, it takes 0.27 Wb.
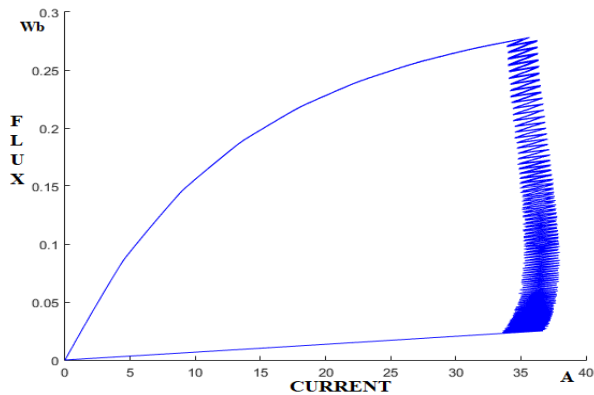
Fig 14. 6/4 SRM Flux-Current Behavior.

Fig 15 shows the flux current variation of the 8/6 SRM. As can be seen from the fig, the continuous current value of the 8/6 SRM motor reaches 18 A. The flux was 0.1 Wb when the motor poles were in the centered position and 0.24 Wb when the motor poles were in the opposite position.
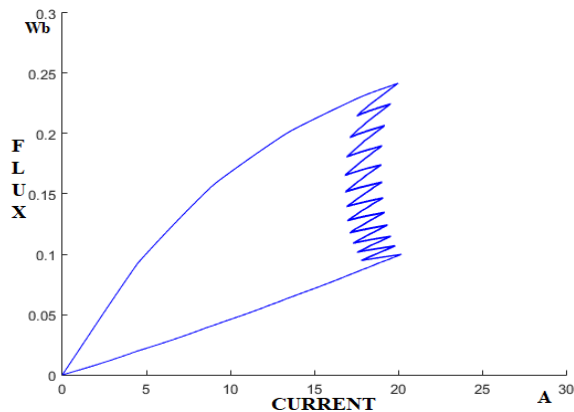
Fig 15. 8/6 SRM Flux-Current Behavior.

The 6/4 SRM characteristically draws more current than the 8/6 SRM, as shown in the current–flux graph. However, the 6/4 has higher flux Φ values, which causes more noise and vibration.

When the inductance graphs of the 6/4 SRM and the 8/6 SRM seen in Figs 16 and 17 were compared, we see that although they have same the minimum and maximum inductance values, their inductance frequencies due to the numbers of poles show different behaviors due to different current and flux values.
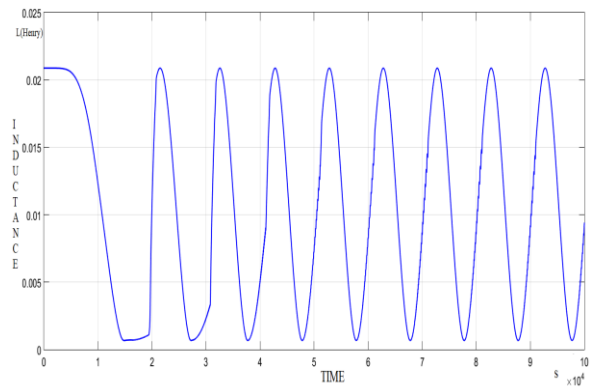
Fig 16. 6/4 SRM Inductance.
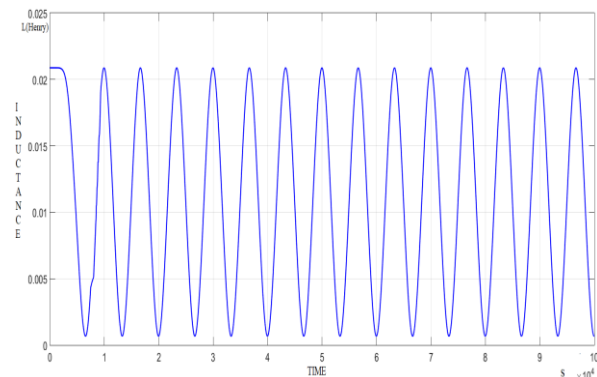
Fig 17. 8/6 SRM Inductance.

IV. CONCLUSION

This study illustrates the performance of two types of SRMs known as the 6/4 SRM and the 8/6 SRM. These motors are supplied by an asymmetric converter system and controlled by PID software. Although they are identical and have the same driver systems and position sensors, the motor performances are different due to their dissimilar magnetic structures. The simulation results show that 6/4 SRM carries more current than 8/6 SRM in the case of no load and loaded conditions, whereas 8/6 SRM has less 25 % torque ripple. Also, 6/4 SRM has 50% higher flux than 8/6 SRM due to the different poles number and unlike magnetic structure. Besides 6/4 SRM shows the better performance at low speeds.

REFERENCES

[1] Garlet, Tais B., Savian, Fernando S., Siluk, Julio C.M., (2019), "Research, development and innovation management in the energy sector", Advances in Energy Research, Volume 6, Issue 1, 17-33.
[2] Gujarathi, Pritam K., Shah, Varsha A., Lokhande, Makarand M., (2020), "Hybrid artificial bee colony-grey wolf algorithm for multi-objective engine optimization of converted plug-in hybrid electric vehicle", Advances in Energy Research, Volume 7, Issue 1, 35-52.
[3] Nematollahi, Amin Foroughi, Shiva, Navid, (2020), "An application of LAPO: Optimal design of a stand alone hybrid system consisting of WTG/PV/diesel generator/battery", Advances in Energy Research, Volume 7, Issue 1, 67-84.
[4] Kioskerides, I., Mademlis C., (2005), "Maximum efficiency in single-pulse controlled switched reluctance motor drives", IEEE Transactions on Energy Conversion, Volume 20, 809–817.
[5] Gorden R., Straughen A., (2000), "Electric machines", 1st Edition.
[6] Neuhaus, C., Fuengwarodsakul N. and Doncker R., (2006), "Predictive pwm-based direct instantaneous torque control of switched reluctance drives", Power Electronics Specialists Conference, 1–7.

[7]   Krishnan, R., (2001), "Switched reluctance motor drives modelling", CRC Press, London.
[8]   Aljaism, W., (2015), "Switched reluctance motor design simulation and control", University of Western Sydney School of Engineering DRG Power Conversion and Intelligent Motion Control.
[9]   Vikramarajan, J., (2016), "Matlab simulink implementation of switched reluctance motor with direct torque control technique", Volume 4, ISSN 2321–9939.
[10] Hashem, G., Hasanien H., (2010) "Speed control of switched reluctance motor based on fuzzy logic controller", Proceedings of the 14th International Middle East Power Systems Conference (MEPCON'10).
[11] Wichert T., Staszewski P., (2008), "Design and construction modifications of switched reluctance machines", Warsaw University of Technology Ph.D. Thesis.
[12] Farah N., Jurifa L., (2017), "Multilevel inverter fed switched reluctance motors (srms) 6/4, 8/6 and 10/8 srm geometric types", International Journal of Power Electronics and Drive System (IJPEDS), Volume 8, ISSN 2088–8694.
[13] Toker K., Serteller F.N., (2017), "A dynamic analysis of bldc motor by using matlab/simulink and mathematica", Conference on Electrical Electronics Engineering (CHILECON).
[14] Kentli F, Çalik H., (2011), "Matlab-Simulink modelling of 6/4 srm with static data produced using finite element method", Acta Polytechnica Hungarica, 23–46.
[15] Nutan S., Panda S., (2017), "Speed control with torque ripple reduction of switched reluctance motor by many optimizing liaison technique", Journel of Electrical Systems and Information Technology (JESIT).
[16] Grace L., Nguyen X., (2020), "Design of a low torque ripple three-phase srm for automotive shift-by-wire actuator", Energies.
[17] Honzhong M., Chaozhi H., (2020), "The effect of a single-sided pole shoe and slot on reducing torque ripple in a switched reluctance motor", Concurrency and Computation: Practice and Experience.
[18] Gecer B., Serteller F., (2019), "Investigation of effect of pole number change on operation parameters in switched reluctance motors", Marmara University Electrical and Electronic Engineering.
[19] Sun H., Moghaddam A., (2020), "A novel driving method for switched reluctance motor will standard full bridge inverter, IEEE Transactions on Energy Conversion.
[20] Tuan V., Kreuawan S., (2020), "Switched reluctance motor and induction machine for e-scooter based on driving cycles design comparasion", Transactions on Electrical and Electronic Engineering.
[21] Howey B., Liang J., (2020), "Making the case for switched reluctance motors for propulsion applications", IEEE Transactions on Vehicular Technology.
[22] Gecer B., Serteller F., (2020), "Understanding switched reluctance motor analysis using ansys/maxwell". IEEE 29th International Symposium on Industrial Electronics (ISIE).
[23] Gecer B., Tosun O., Apaydin H., Serteller F., (2021), "Comparative Analysis of SRM, BLDC and Induction Motor using ANSYS/Maxwell ", Electrical, Computer, Communications and Mechatronics Engineering (ICECCME), International Conference.
[24] Gecer B., Serteller F., (2020), "Understanding a switched reluctance motor control and analysis methods using Matlab/simulink". IEEE World Engineering Education Conference (EDUNINE).

BIOGRAPHIES

Bekir GECER Maltepe, Istanbul, in 2022. He received the B.S. degree in electrical-electronic engineering from the Fatih University, Istanbul, in 2014 and M.S. degree in electrical-electronic engineering from the Marmara University, Istanbul, in 2019. He is a Ph.D. student at Marmara University.   He works as a Lecturer at Atasehir Adiguzel MYO for the Machine and Metal Technology Department, Istanbul. His research areas are electric motors, power electronic and control systems.

N. Fusun SERTELLER She received her degree and M.Sc. in electrical engineering, from Istanbul Technical University (ITU), in Istanbul. She received her Ph.D. degree from Marmara University (MU), Technical Education Faculty Electrical Education Department in Istanbul, Turkey in 2000.

She is currently lecturing on magnetic field theory, electric machines, and numerical methods in Electric-Electronic Engineering Department at MU. Her research interests include electric machine modeling, electrical machine design, electric vehicle motors, mathematical modeling of electric machines and systems, and electromagnetic field theory applications.

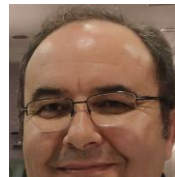**Hüseyin CALIK** Merkez, Giresun, in 2022. He received the B.S. degree in electrical-electronic engineering from the Marmara University, Istanbul, in 2001 and Ph.D. degree in electrical-electronic engineering from the Marmara University, Istanbul, in 2004. He is currently lecturing on magnetic electric machines, and power plant in Electric-Electronic Engineering Department at Giresun University.

Prediction of Monkeypox on the Skin Lesion with the Siamese Deep Learning Model

Talha Burak Alakus

Abstract— One of the viral diseases that has started to cause concern in various parts of the world is the monkeypox virus, which has emerged recently. Monkeypox is transmitted to humans from an animal infected with the virus or from another human being infected with monkeypox. Among the most basic symptoms are high fever, back and muscle aches, chills, and blisters on the skin. These blisters are sometimes confused with chickenpox and measles, resulting in incorrect diagnosis and treatment. Therefore, the need for computer-aided systems has increased and the need for more robust and reliable approaches has arisen. In this study, a deep learning model was used, and a distinction was made between monkeypox and other diseases. The study consisted of three stages. In the first stage, data were obtained and images of both chickenpox and other diseases were used. In the second stage, the Siamese deep learning model was employed, and data were classified. In the last stage, the performance of the classifier was evaluated and accuracy, precision, recall, and F1-score were calculated. At the end of the study, an accuracy score of 91.09% was obtained. This result showed that the developed deep learning-based model can be used in this field.

Index Terms— Artificial intelligence, Monkeypox, Siamese deep learning model

I. INTRODUCTION

MONKEYPOX, which is seen as a possibility of a new virus outbreak has begun to be followed in detail. Infection of monkeypox virus in humans was first observed in 1970 [1]. From these years to the present, monkeypox has been regularly occurring in the African continent as sporadic reported cases. Monkeypox, which has been seen for many years in West and Central Africa and it rarely reaches different parts of the world with the contamination caused by animals exported from the region [2]. However, recently, the disease has become more widespread than in the past, and it has been diagnosed in different people from different regions [3]. Monkeypox can be transmitted to humans from an infected animal or from another person infected with monkeypox. Although the symptoms are similar to smallpox, it progresses somewhat differently [4]. The most basic symptoms of monkeypox are high fever, weakness, headache, back pain, chills, and blisters on the skin [5].

TALHA BURAK ALAKUS, is with Department of Software Engineering University of Kirklareli, Turkey, (e-mail: talhaburakalakus@klu.edu.tr).

https://orcid.org/0000-0003-3136-3341

In monkeypox, rashes are first observed in the face area. It then spreads to other parts of the body. A low mortality rate is observed in monkeypox, and this value varies between 3-6% [6].

Infection from the monkeypox virus is contagious. Many situations such as being bitten by a virus-infected animal, touching the body fluids of these animals, consuming the undercooked meat of an infected animal, using its fur are among the ways of transmission of the disease. The monkeypox virus, which is transmitted to humans in this way, spreads among humans and creates a risk of turning into an epidemic. The rashes on the body of the sick individual and the fluids in these bubbles contain the disease factor. Therefore, situations such as touching the rashes of a sick person and sharing the clothes, sheets, towels, and similar items contaminated with these rashes cause the disease to be transmitted easily [7]. Paying attention to the hygiene rules that help protect against all diseases is the most important stage of protection. Since animal meat is an important route of transmission, only well-cooked meat should be consumed. Apart from this, people who spend time in nature should also be careful. Stray and wild animals, including dead animals, should not be approached. Animals that are seen in nature and do not look healthy should never be touched [8].

The disease can be confused with chickenpox and measles in some cases because of the similarity of the rashes [9, 10]. An expert interpretation is required to make a clear distinction between them. However, in the early stages of the disease, this difference cannot be fully evident and causes the diagnosis to be incorrect. Accordingly, the treatment is wrong, and this causes the patient to lose time and to be treated incorrectly. To avoid such problems, computer-aided systems are needed [11]. In this study, a computer-aided diagnostic system was created, and a decision-making mechanism was established to distinct monkeypox or other diseases (chickenpox and measles). In this direction, deep learning algorithm was used in the study and skin images of diseases were classified. The study consisted of three different stages. In the first stage, skin images of new monkeypox and other diseases were obtained. In the second stage, deep learning model was designed, and Siamese deep learning model was applied at this stage. In the final stage, the performance of the classifier was determined by accuracy, precision, recall, and F1-score evaluation metrics. The highlights of the study can be expressed as follows:

- Siamese deep learning model, which is one of the lesser-known deep learning methods, was used and monkeypox disease was predicted from skin images.

- With this study, the differences between monkeypox and other diseases were determined by computer- aided system.
- In this study, raw skin images were evaluated without any preprocessing and classified without complicating the feature vector. In this way, both the processing load and the data load did not increase.
- It has been shown that a computer-aided system can be effective in situations that require expert interpretation.

The remainder of the study is organized as follows. In the Section II, studies in this field in the literature were examined. In the Section III, the data set and method used in the study were emphasized. Section IV showed the results of the application and the discussions. Furthermore, the advantages and disadvantages of the study were explained. In the last section, the contribution of the study to the literature was given and its impact on future studies was mentioned.

II. RELATED STUDIES

In this part of the study, studies that have predicted monkeypox disease with artificial intelligence algorithms using skin images were included. In the study [11], researchers used deep learning models to predict monkeypox. Images of three different skin diseases were used in the study. These diseases were monkeypox, chickenpox, and measles. In the study three different classifiers were employed including VGG-16, ResNet50, InceptionV3. The performances of the models were determined by the 3-fold cross-validation method and for this purpose accuracy, precision, F1-score, and recall evaluation criteria were used. At the end of the study, the highest accuracy score was obtained with the ResNet50 deep learning model, and the result was 82.96%. In study [12], the researchers predicted monkeypox disease using pre-trained deep learning models. In the study, a total of 1,754 images belonging to four different classes, including chickenpox, monkeypox, measles and normal skin images, were used. The images were then classified using thirteen different deep learning models. These models were VGG-16, VGG-19, ResNet50, ResNet101, IncepResnetV2, MobileNetV2, InceptionV3, Xception, EfficientNetB0, EfficientNetB1, EfficientNetB2, DenseNet121 and DenseNet169. The performance of the classifiers was determined by the evaluation criteria of precision, F1-score, recall, and accuracy. In the validation process, 5-fold cross-validation was applied. The lowest accuracy score was obtained with VGG-16 and the result was 82.22%. The highest accuracy score was calculated with the Xception deep learning model, and 86.51% accuracy was obtained with this model. In study [13], researchers predicted monkeypox from images using the modified VGG-16 model. In the study, two different images, monkeypox and chickenpox, were used and two different scenarios were emphasized. While a total of 90 images were used in the first scenario, the images were augmented in the second scenario and classification was performed on 1,754 images. The performance of the classifier was determined by accuracy, precision, recall and F1-score evaluation criteria. While 83% accuracy rate was obtained for the first scenario, 78% accuracy score was observed for the second scenario. In study [14], researchers predicted monkeypox disease using CNN (Convolutional Neural Networks) based features. In the study, AlexNet, GoogleNet and VGG-16 architectures were used as CNN models and separate features were obtained. Then, these features were classified with five different machine learning algorithms. In the study, NB (Naive Bayes), SVM (Support Vector Machine), KNN (K Nearest Neighbor), DT (Decision Tree) and RF (Random Forest) algorithms were used as machine learning algorithms. The performance of the classifiers was determined by the evaluation criteria of accuracy, F1-score, precision, and recall. At the end of the study, an effective classification process was carried out with transfer learning. In study [15], researchers predicted monkeypox using CNN architectures and skin images. Five classifiers were used in the study, namely VGG-16, VGG-19, ResNet50, MobileNetV2 and EfficientNetB0. The performance of the classifiers was determined by four different evaluation criteria: accuracy, F1-score, precision, and recall. In addition to these, they also used Ensemble classifiers. The highest accuracy score was obtained with Ensemble, which includes ResNet, EfficientNet and MobileNet models, and an accuracy of 98.33% was observed.

III. MATERIAL AND METHODS

A. Dataset

In this study, the publicly available data set was used, and the data were obtained from study [11]. While there are 102 images of monkeypox in the original dataset, there are 126 images of chickenpox and measles. Furthermore, there are also augmented images in the dataset. In the augmented images, there are images of 1,428 monkeypox diseases, and 1,764 images of chickenpox and measles. Augmented images were used in the study and 80% of the data was used for training and the remaining 20% for testing. Some of these images are given in Fig. 1.

Fig.1. Images of the data set. The images at the top are of monkeypox, while the images at the bottom are photos of measles and chickenpox.

B. Siamese Deep Learning Model

A Siamese Deep Learning model (SDLM) was first proposed to solve signature verification problems [17]. It is a   type of neural network architecture that learns to differentiate between distinct classes by comparing new data points with existing data points. The term "Siamese" refers to the structure of the network: it contains two or more identical subnetworks with the same architecture and shared weights [16]. The SDLM decides whether the two images given as input are similar or different by training the network. SDLM determines whether the two image data are the same by matching two images given as input,

according to the weight parameters. If the approximate values of the weight parameters obtained from the two images produce similar results, the images are identical. Therefore, a threshold value is used in the SDLM [18]. In SDLM, classification is performed by considering the difference between the results obtained from the two neural networks. If this difference is below the threshold value, the images are the same or similar. If it is above the threshold value, the model shows that the results are different. Since each pixel on the image has numerical values, the distance between pixels in the SDLM is usually calculated with Euclidean. The working steps of the Siamese deep learning model can be summarized as follows:

Forward Pass: First, the model processes two inputs simultaneously. Each input x_1 and x_2 is handled by two separate networks (twins of a Siamese network) that share the same weights. This is done with a set of transformation and activation functions represented by weights W. Thus, the outputs are often expressed as $f(x_1; W)$ and $f(x_2; W)$ where the function f is the feature extraction that the model has learned.

Similarity Metric: A similarity metric is then used to measure the similarity between the outputs of the two inputs. One of the most used metrics is the L1 or L2 (Euclidean) distance between two vectors. Euclidean distance is determined by the formula given in Equation 1.

$$D = |f(x_1; W) - f(x_2; W)| \qquad (1)$$

Loss Function: Finally, a loss function is calculated based on the similarity metric. Generally, it is desirable that two inputs belonging to the same class ($y = 1$) have a low distance, and two inputs belonging to different classes ($y = 0$) should have a high distance. An example of this is the Contrastive Loss function. The calculation of the Contrastive Loss function is shown in Equation 2.

$$CL(d, Y) = \frac{1}{2} * Y * d^2 + (1 - Y) * \frac{1}{2} * \max(0, m - d)^2 \qquad (2)$$

In the Equation 2, d represents the distance of the outputs. While Y shows the label of the model inputs, m specifies the margin parameter. This loss function encourages to reduce the distance between instances of the same class and to increase the distance between instances of different classes.

The main reason for using SDLM in the study is that the model is effective even with a small number of images [19]. Since monkeypox is a new disease, there are not enough data yet. Existing data are either few or artificially augmented images. Therefore, the low amount of data in general lead to the use of the SDLM in this study. Furthermore, there is a semantic similarity situation in the SDLM. In this model, similar classes or concepts produce weight values close to each other [20]. Contrastive loss is generally used in Siamese networks. In this loss function, all positive images are forced to be close to each other, while all negative images are scattered a certain distance. Moreover, using the same adjustment distance for all images can be restrictive. This causes triplet loss, which requires negative images to be further away from any positive images [21]. The triplet loss function does not use a threshold value to

distinguish between similar and different images. Instead, it generates various variance values for different classes. Withal, in triplet loss, positive images and negative images are close to each other. However, the contrastive loss focuses only on positive images. For these reasons, the triplet loss function was used in the study.

Triplet loss is a loss function used when training a SDLM. Triplet loss uses three instances of "triplets": an "anchor" instance, a "positive" instance (belonging to the same class), and a "negative" instance (belonging to a different class). The loss function aims to increase the distance between the anchor and negative samples, while reducing the distance between the anchor and positive samples. The formula for the triplet loss function is given in Equation 3.

$$TL(W) = \max(d(A, P) - d(A, N) + margin, 0) \qquad (3)$$

In Equation 3, $d(A, P)$ is the distance between the anchor and positive samples. The expression $d(A, N)$ represents the distance between the anchor and negative samples. The margin expression is the size of the margin, which is a hyperparameter. This controls the distance between positive and negative samples. If $(d(A, P) - d(A, N)) + margin < 0$, the loss function is set to zero. SDLM was employed in this study and the images were evaluated according to the flowchart given in Fig. 2.
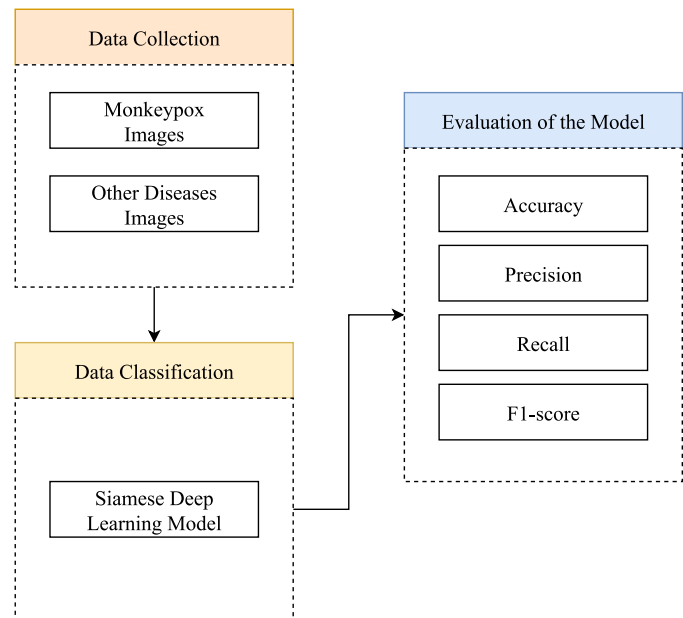
Fig. 2. Flowchart of the study

As can be seen in Fig. 2, the study consists of three stages. In the first stage (**Data Collection**), skin images of monkeypox and other diseases were obtained. Data augmentation was not performed in this study. The images in the dataset are already augmented images. Furthermore, since there are two different class labels, binary classification was made in the study. Then the images were normalized, and the pixel values were ensured to be in the same range. In the second stage (**Classification**),

       http://dergipark.gov.tr/bajece

SDLM was designed, and images were classified. In the last stage (**Evaluation of the Model**), the performance of the classifier was determined by accuracy, precision, F1-score, and recall values.

The classification process in the study was carried out in Anaconda, Spyder environment using the Python programming language. The version of the Spyder in the study is 5.4.2. Anaconda is a popular distribution of the Python programming language. It is a platform with a rich ecosystem that includes the Python programming language as well as various data science, machine learning and deep learning tools. Libraries like Pandas, NumPy, and SciPy are used for data manipulation, mathematical calculations, and statistical analysis. Visualization libraries such as Matplotlib and Seaborn allow to visualize data. Deep learning and machine learning libraries such as Scikit-learn, TensorFlow, and PyTorch can be used for a variety of modelling and prediction tasks. The Tensorflow library was used for the classification process and the *Model, Input, Flatten, Dense* of this library were used. Additionally, the Keras library was also employed and the optimization algorithm, loss function and evaluation metrics were also imported from this library. The study was carried out on a Mac computer. The operating system of the computer is Ventura 13.2.1 and consists of 8 GB of RAM. Finally, the Mac computer has the M1 chip. The inclusion of the Apple Neural Engine in the M1 chip significantly enhances the speed and performance of machine learning and deep learning tasks on Mac devices. With its cutting-edge 16-core architecture, the Neural Engine embedded in the M1 chip from Apple empowers Mac devices to execute an impressive 11 trillion operations per second [22]. This remarkable feature results in a staggering 15-fold increase in machine learning and deep learning performance compared to previous generations. The M1 chip is specifically engineered to deliver exceptional performance in machine learning tasks. It incorporates machine learning and deep learning accelerators within the CPU and a high-performance GPU, enabling unparalleled capabilities in areas such as video analysis, voice recognition, and image processing [23]

IV. APPLICATION RESULTS AND DISCUSSION

In this study, SDLM was used and prediction of monkeypox and other diseases was carried out. The parameters of the proposed method were determined as follows:

- The first layer contains the 8 x 8 convolution layer. At the end of these layers, the ReLU (Rectified Linear Unit) activation function is used.
- LRN (Local Response Normalization) has been applied in the second layer and the pixel values in the feature map have been square normalized.
- Pooling was done in the third layer and 4 x 4 maximum pooling was employed.
- In the fourth layer, the convolution layer is used again. The size of this layer was evaluated as 6 x 6. Then the ReLU activation function was used.
- LRN was applied in the fifth layer.
- In the sixth layer, 4 x 4 maximum pooling was utilized.
- In the seventh layer, dropout was made and 15% of the data was forgotten.

- In the eighth and ninth layers, convolutional layers were formed, and the dimensions of these layers were determined as 3 x 3). ReLU was used as the activation function.
- In the tenth layer, 2 x 2 maximum pooling was employed.
- 15% of the data was forgotten by dropout in the 11th layer.
- The fully connected layer is defined in the 12th and 13th layers, and 256 and 128 neurons are defined, respectively.
- Then, the distance between the outputs of the two models was determined by Euclidean and triplet loss was used for the loss of the model.
- The training was carried out with 500 iterations. However, early stop was applied, and the training was terminated when the epoch number reached 100 sine the training accuracy reached %95 accuracy.

The form representation of the parameters of the developed model is given in Fig. 3.
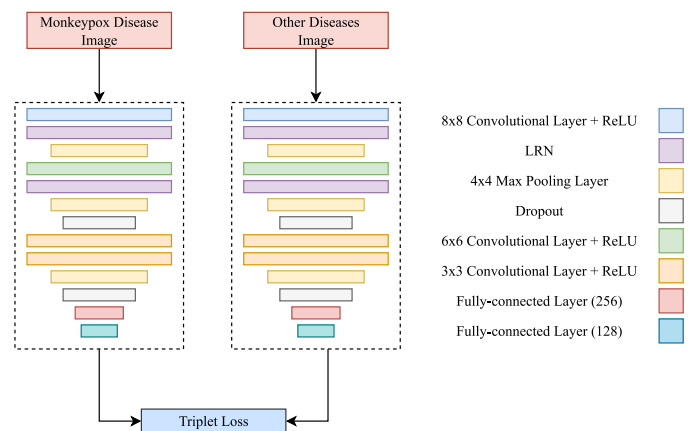
Fig.3. Parameters of the developed model

The performance of the classifier was determined by the accuracy, F1-score, precision and recall values. Classification results were given in Table 1.

TABLE I
CLASSIFICATION RESULTS

Classifier	Accuracy	Precision	Recall	F1-Score
SDLM	91.09%	89.03%	93.36%	91.14%

In line with the results obtained from Table I, an accuracy rate of 91.09% was obtained with the SDLM. Further, 89.03% precision and 93.36% recall values were observed with this deep learning model. The F1-score value was 91.14%. The data set used in the study is not a balanced data set. Accuracy score alone is not an adequate evaluation criterion in unbalanced data sets [24]. Therefore, in this study, F1-score evaluation criterion was also applied in addition to the accuracy score. The F1-score is determined by precision and recall values. The higher these values, the higher the F1-score value. The F1-score takes values between 0-1. If the value is close to 1, it indicates that the model has made a good classification. Therefore, looking at the F1-score, SDLM was successful.

The validation loss and the validation accuracy of the model were given in Fig. 4 and Fig. 5, respectively.
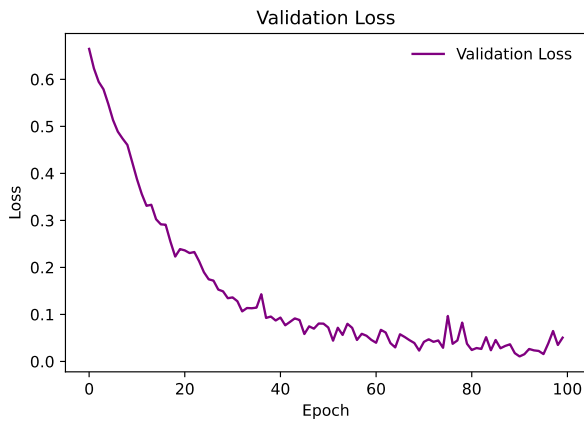
Fig.4. Validation loss of the developed SDLM
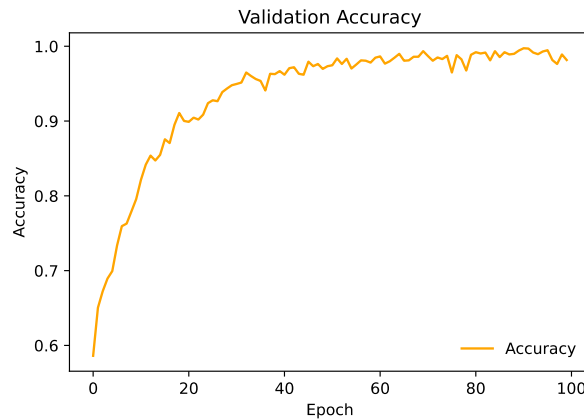
Fig.5. Validation accuracy of the developed SDLM

As seen in Fig. 4 and Fig. 5, the training was terminated when the epoch number reached 100 and the most successful validation accuracy and validation loss were obtained at this epoch value. The confusion matrix was used to show how much of the data were classified correctly and incorrectly in Figure 6.
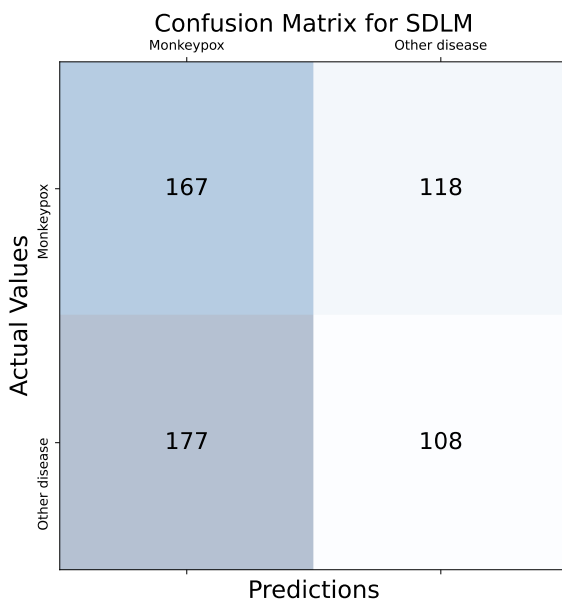
Fig. 6. Confusion matrix of the developed SDLM

The accuracy score of this study was also compared with other studies in the literature using the same dataset. The studies and their results were given in Table 2.

TABLE II
COMPARISON OF STUDIES IN THE LITERATURE (THE HIGHEST ACCURACY SCORES WERE GIVEN AND THE MOST SUCCESSFUL DEEP LEARNING MODEL WAS INCLUDED.

Ref.	Method	Accuracy	Precision	Recall	F1-Score
[11]	Ensemble	82.96%	87.00%	83.00%	84.00%
[12]	Ensemble	87.13%	85.44%	85.47%	85.40%
[13]	VGG16	78.00%	75.00%	75.00%	75.00%
[14]	Transfer learning	91.11%	93.10%	93.35%	93.15%
[15]	Ensemble	98.33%	98.33%	98.33%	98.33%
This study	**SDLM**	**91.09%**	**89.03%**	**93.36%**	**91.14%**

When the results in Table II were examined, it was observed that the SDLM was effective against some models and ineffective against others. The accuracy scores obtained in [14] and [15] were higher than this study. In the study [14], the researchers used transfer learning and obtained features with deep learning models and classified them with machine learning algorithms. This whole process takes time and increases the analysis time. Then again, obtaining features increases the size of the feature vector and causes the data to be complex. Although an accuracy score of 91.11% was obtained despite all this process, feature extraction was not used in this study and the raw data were classified only with SDLM and a result as effective as [14]. In the study [15], the most successful classification process was obtained with the Ensemble method. In that method, three deep learning models were combined (ResNet50+EfficientNetB0+MobileNetV2) and classification was made. Since three different deep learning models were used in that study, defining these models, analyzing each model with images and interpreting the results are time-consuming. However, in this study, only one model was designed, and these processes were carried out and obtained an accuracy of 91.90%. This shows how effective the method is. However, since the same data set was used in the comparison studies, there is an unbalanced data set problem in those studies. Therefore, F1-score was used in those studies and the classification of the model was evaluated. When compared according to the F1-score, the proposed method lagged behind the studies [14] and [15]. When the comparison was made according to the accuracy scores, the results of the proposed method and the study [14] were close to each other, yet when the model was evaluated according to the F1-score, it was observed that the study [14] was more effective. In study [14], transfer learning was applied and the data were not classified with raw images. In the first stage, features were obtained with deep learning models and then classification was made with machine learning algorithms. F1-score may have been more effective because of this procedure. Obtaining its properties may have caused the classification process to be more effective. Although the studies [14] and [15] are more effective than the proposed method, it has been shown that the SDLM method can be more effective

than the CNN models. One of the main reasons for this is the low number of images and labels, and the Siamese deep learning model's ability to make an effective classification even with low amount of data [19]. Moreover, triplet loss was used in this study and traditional loss was abandoned. This enabled the model to be better trained despite the low data number. The advantages and disadvantages of the study can be expressed as follows:

- Siamese deep learning model was used in this study instead of traditional feature extraction methods.
- In this study, images were not subjected to any preprocessing and raw images were used directly.
- The proposed method produced high accuracy results from some of the studies performed in the literature.
- The number of data affects the performance of deep learning models. The large number of data will positively affect the performance of this model. However, despite the small number of data, a successful classification process has taken place.
- No feature extraction was performed in this study. Performing various feature extraction or transfer learning can increase the performance of this model.
- More successful methods than the Siamese deep learning model are usually ensemble methods. Similarly, by using an ensemble structure, the classification performance obtained in this study can be increased.
- While the model was being trained, an average of 17 seconds has passed for each iteration. Since the training process was completed in 100 iterations, it took $17 * 100 = 1700$ seconds, approximately 28.3 minutes in total. Although it seems long in terms of total time, the time obtained for 1 iteration is fast enough. Of course, this time varies according to the hardware and software environment used. This issue needs to be tested in more detail and on different hardware.

When the studies in the literature were examined, researchers generally used transfer learning and ensemble. In the transfer learning method, the data load increases, and the data becomes complex. Because in this approach, features are obtained by deep learning models and used for classification. This extends the analysis time. Similarly, multiple models are combined in the ensemble approach. This is a process that prolongs the classification process. In addition, it takes time to find the most suitable combination. However, the SDLM used in this study only and classified raw data. It has performed an effective classification process from most of the studies in the literature.

V. CONCLUSION

The main purpose of this study is to lay the groundwork for the development of an auxiliary model that can be used to differentiate between monkeypox and other skin diseases. For this reason, deep learning, which is used effectively in medical imaging applications, has been applied and classification of monkeypox and other skin diseases has been made. The study consisted of three stages. In the first stage, skin images of monkeypox and other diseases were obtained. In the second stage, the Siamese deep learning model was designed, and the data were classified. In the last stage, the performance of the deep learning model was evaluated. Accordingly, accuracy, precision, recall, and F1-score were used. With the proposed method, 91.09% accuracy was obtained. When compared with other studies in the literature, it was observed that the proposed method was more effective than most studies. In line with the findings obtained in the study, it has been shown that the developed model can help experts. This will facilitate the work of specialists and will provide convenience to physicians when the distinction of blisters on the skin is not made clearly.

REFERENCES

[1] J.G. Breman, R. Kalisa, M.V. Steniowski, E. Zanotto, A.I. Gromyko, et al. "Human monkeypox, 1970-79." Bulletin of the World Health Organization, vol. 58. 2, 1980, pp. 165-182.

[2] D. Mileto, A. Riva, M. Cutrera, D. Moschese, D. Mancon, et al. "New challenges in human monkeypox outside Africa: A review and case report from Italy." Travel Medicine and Infectious Disease, vol. 49, 2022, pp. 1-10.

[3] J.P. Thornhill, S. Barkati, S. Walmsley, J. Rockstroh, A. Antinori, et al. "Monkeypox virus infection in human across 16 countries-April-June-2022." The New England Journal of Medicine, vol. 387, 2022, pp. 687-691.

[4] J.A. Cann, P.B. Jahrling, L.E. Hensley, V. Wahl-Jensen. "Comparative pathology of smallpox and monkeypox in man and macaques." Journal of Comparative Pathology, vol. 148. 1, 2013, pp. 6-21.

[5] J.G. Rizk, G. Lippi, B.M. Henry, D.N. Forthal, Y. Rizk. "Prevention and treatment of monkeypox." Drugs, vol. 82, 2022, pp. 957-963.

[6] A. Zumla, S.R. Valdoleiros, N. Haider, D. Asogun, F. Ntoumi, et al. "Monkeypox outbreaks outside endemic regions: Scientific and social priorities." The Lancet: Infectious Diseases, vol. 22. 7, 2022, pp. 929-931.

[7] K.C. Zachary, E.S. Shenoy. "Monkeypox transmission following exposure in healthcare facilities in nonendemic settings: Low risk but limited literature." Infection Control and Hospital Epidemiology, vol. 43. 7, 2022, pp. 920-924.

[8] D. Daskalakis, R.P. McClung, L. Mena, J. Mermin. "Monkeypox: Avoiding the mistakes of past infectious disease epidemics." Annals of Internal Medicine, vol. 175. 8, 2022, pp. 1177-1178.

[9] Z. Jezek, M. Szczeniowski, K.M. Palukulu, M. Mutombo, B. Grab. "Human monkeypox: Confusion with chickenpox." Acta Tropica, vol. 45. 4, 1988, pp. 297-307.

[10] S.N. Shchelkunov, A.V. Totmenin, I.V. Babkin, P.F. Safronov, O.I. Ryazankina, et al. "Human monkeypox and smallpox viruses: Genomic comparison." FEBS Letters, vol. 509. 1, 2011, pp. 66-70.

[11] S.N. Ali, T. Ahmed, J. Paul, T. Jahan, M.S. Sani, et al. "Monkeypox skin lesion detection using deep learning models: A feasibility study." arXiv, 2022.

[12] C. Sitaula, T.B. Shani. "Monkeypox virus detection using pre-trained deep learning-based approaches." arXiv, 2022.

[13] M. Ahsan, M. Farjana, K.A. Momin, M.R. Uddin, A.N. Sakib, et al. "Image data collection and implementation of deep learning-based model in detecting monkeypox disease using modified VGG16." arXiv, 2022.

[14] V. Kumar. "Analysis of CNN features with multiple machine learning classifiers in diagnosis of monkeypox from digital skin images". medRxiv, 2022.

[15] S.L. Munoz, L.E. Escobar, M.J. Civit, P.F. Luna, A. Civit, et al. "Monkeypox diagnostic-aid system with skin images using convolutional neural networks." SSRN, 2022.

[16] R. Vinayakumar, K.P. Soman. "Siamese neural networks architecture for homoglyph attacks detection." ICT Express, vol. 6. 1, 2020, pp. 16-19.

[17] J. Bromley, J.W. Bentz, L. Bottou, I. Guyon, Y. Lecun, et al. "Signature verification using a "siamese" time delay neural network." International Journal of Pattern Recognition and Artificial Intelligence, vol. 7. 4, 1993, pp. 669-686.

[18] M. Toğaçar, Z. Cömert, B. Ergen. "Recognition of the digits in Turkish sign language using siamese neural networks." Dokuz Eylül University Faculty of Engineering Journal of Science and Engineering, vol. 23. 68, 2021, pp. 349-356.

[19] I. Melekhov, J. Kannala, E. Rahtu. "Siamese network features for image matching." International Conference on Pattern Recognition. Cancun, Mexico, 2016.

[20] B. Wang, D. Wang. "Plant leaves classification: A few-shot learning method based on siamese network." IEEE Access, vol. 7, 2019, pp. 151754-151763.

[21] X. Dong, J. Shen. "Triplet loss in siamese network for object tracking." European Conference on Computer Vision. Munich, Germany, 2018.

[22] Apple   Unleashes   M1, https://www.apple.com/newsroom/2020/11/apple-unleashes-m1/, Erişim tarihi (29 Mayıs 2023)

[23] D. Kasperek, M. Podpora, A. Kawala-Sterniuk. "Comparison of the usability of apple M1 processors for various machine learning tasks." Sensors, vol. 22, 2022.

[24] L.A. Jeni, J.F. Cohn, F. De La Torre. "Facing imbalanced data-- recommendations for the use of performance metrics." International Conference on Affective Computing and Intelligent Interaction and Workshops. Geneva, Switzerland, 2013.

BIOGRAPHY

TALHA BURAK ALAKUS received his B.S. degree in software engineering from the Izmir University of Economics in 2010. Furthermore, he received his M.S. and Ph.D. degrees in software engineering from Fırat University in 2018 and 2022, respectively. From 2016 to 2022, he was a Research Assistant in Kırklareli University and Fırat University. Since 2022, he has been an Assistant Professor with the Software Engineering Department in Kırklareli University. His research interests include bioinformatics, software engineering, biomedical and genomic signal processing, pattern recognition and artificial intelligence.

Remote Maintenance and Software Update Methods for Connected Vehicles

Erkan Demir, Habib Kaymaz, Ilker Yildiz*

Abstract— While building smart living areas, the main purpose is to increase the quality and comfort of life without compromising health and safety. With the increase of new technologies used in growing and updated modern living areas, we witness that the needs are shaped and changed with this change. The primary purpose of each new technology is to offer more safety, health, comfort and convenience to the person, organization or structure that it addresses. The basic means of the connected vehicles has been to ensure the safety of the user in the best way, and people expect more than vehicle. With the widespread use of autonomous vehicles, that is without driver assistant, in the near future, vehicles are expected to connect and communicate with the ecosystem structured for the future and among themselves. Thanks to this communication that the vehicles have, the traffic flow on the route, the road condition, the location of the pedestrians around, possible malfunctions and failures that may occur and all such information will be shared by the vehicle with each other and the future ecosystem. On the other hand, while the vehicles are becoming more connected with any of useful technologies, thanks to the remote connection provided by their technology, manufacturers can easily develop their vehicles without recalling. In this study, information was given about the connected vehicles, the remote connection infrastructure and methods. The advantages provided by these systems are explained.

Index Terms— Connected Vehicles, Over The Air (OTA), Firmware Over the Air (FOTA), Software Over the Air (SOTA), Remote maintenance

I. INTRODUCTION

THE WORLD'S leading automakers and communications infrastructure providers, Mercedes-Benz, BMW, Audi, Volvo, Toyota, TIM, and Telstra, describe the importance of connectivity in cars and the ecosystem required to support connected vehicles [1]. According to Mercedes-Benz, "connecting vehicles is laying the foundation to offer new kinds of services for comfort, safety, and entertainment in a new dimension". For BMW, "the highly-developed smart city is characterized by all-encompassing networking, and cars are becoming a part of inner-city networking in this smart city" [2]. Audi thinks that 5G to make software downloads more reliable and faster for an increasing demand from a huge number of vehicles. For Volvo, connectivity was once a feature only, but now it is an essential part of all cars they offer. According to Toyota, "the challenge for connected cars receives and sends large amounts of data from the cloud". TIM forecasts, there will be an estimated 100 million connected vehicles on the road in 2025. Telstra considers that the possibilities are endless when all cars can communicate with each other, with the surrounding infrastructure and pedestrians.

Mercedes-Benz has unveiled its concept vehicle, Vision AVTR as shown in Fig. 1, which is directly correlated with Mercedes-Benz's plans for the future. The company designed the car to connect seamlessly with its passengers, and it embodies the vision of near future for mobility in the. A new era of the car, "The Vision AVTR shows a completely new interaction between human, machine, and nature," the automaker claims. We can see the future by looking at this car; it will become such a natural part of our lives in the near future.

ERKAN DEMIR, is with Department of Electrical and Electronics Engineering, Faculty of Technology, Marmara University, Istanbul, Turkey, (e-mail: erkan.demir@marun.edu.tr).

https://orcid.org/0000-0002-4832-5889

HABIB KAYMAZ is with Mercedes Benz Türk AŞ, Otobus Gelistirme, Esenyurt, Turkey, (e-mail: habib.kaymaz@gmail.com).

https://orcid.org/0000-0002-8338-004X

İLKER YILDIZ is with Ministry of National Education, İstanbul, Turkey, (e-mail: ilkeryildiz2000@hotmail.com).

https://orcid.org/0009-0004-9021-5909

Fig.1. Mercedes-Benz's VISION AVTR [3]

II. CONNECTED VEHICLE

Connected vehicle is a vehicle that can communicate bidirectional with other external objects, structures, and systems inside or outside. In this way, the vehicle can share

the information, which it has with other structures inside and outside of the vehicle or use the information from other structures [4]. It is expected to improve people safety and traffic management by adding communication to a vehicle [5].

General Motors as an automaker offered its connected vehicle features to the end-user market in 1996. The company developed its technology called OnStart in 1996 and introduced it to the public usage in order to ensure the safety of drivers and passengers [6]. General Motors firstly adopted OnStar technology on its Cadillac DeVille, Sevilla, and Eldorado models and improved it later. The primary purpose of using this technology in vehicles was safety and emergency assistance to the vehicle in the event of an accident. If the medical aid and support arrive at the scene as earlier as fastest in the event of an accident, the higher the likelihood that drivers and passengers will survive.

Thanks to OnStar technology, the connected vehicle transmits the help call to the emergency center, which will provide assistance to the vehicle via the phone integrated into the vehicle during the accident and enables communication between passengers and emergency centers.

After the success of General Motors in OnStar, many car manufacturers have included similar safety practices in new cars. And nowadays, after March 31, 2019, new EU-approved car models must have an emergency call assistant as a default function as figured out in Fig. 2.
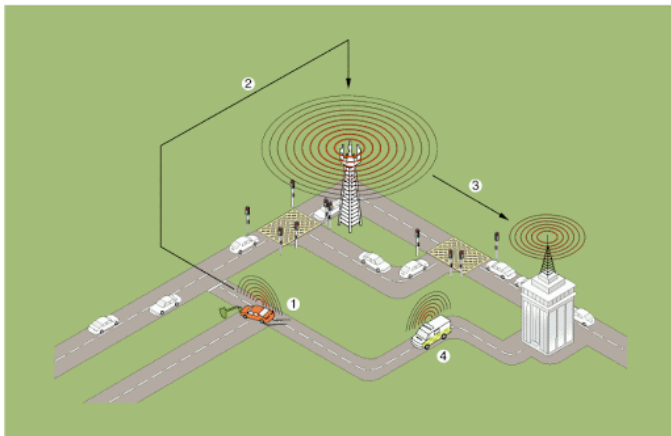
Fig.2. Emergency call system on vehicle [6]

Initially, this technology only provides voice calls via the phone, with the addition of data communication to the phone communication in 1999, GPS location signals have also been transmitted to the emergency center. However, since 2000, by adding the functionality of the GPS system to vehicles, GPS not only took the navigation technology to a new upper level, but also contributed to safety by monitoring and tracking the stolen vehicles [7,8].

III. CONNECTION TYPES

All these new technologies used in vehicles, while offer driver, passenger and vehicle safety, also support building a reliable relation between the end-user and manufacturer or service. For example, the connected vehicle diagnoses a non-critical but needs to be checked and informs the vehicle user as well as the service company. And a specialist person or

system would give more detailed information about the vehicle status to the user, so that the user can drive safely. In this way, the user will be provided with more reliable details of the vehicle and driving, and a long-term reliable relationship will be established between the manufacturer and the user.

As simulated in Fig. 3 there are 5 ways for a vehicle to connect around of it and communicate with them:

- V2I "Vehicle to Infrastructure": communication between vehicle and its environment
- V2V "Vehicle to Vehicle": a vehicle communicates with other vehicles, such as information about the speed and location of the surrounding vehicles
- V2N "Vehicle to Network": data communication between vehicle and cloud-connected structures
- V2P "Vehicle to Pedestrian": communication between vehicle pedestrians for environmental safety
- V2X "Vehicle to Everything": this is a summary of other connectivity's; communication between vehicle and all kinds of vehicles and structures (vehicles, trains, ships, aircraft etc.)
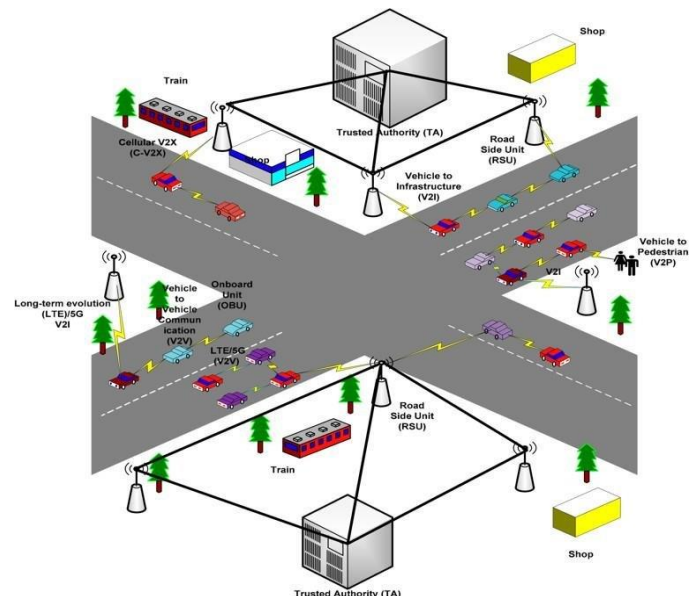
Fig.3. Vehicle to around [9]

The communication standard V2X as a protocol, that allows the connected vehicles to connect with each other and includes all other vehicle communication systems as a framework protocol where data can be transferred in real time at high speed and low loss [10].

Inter-vehicle connection network is one of the main technologies that can be used in the smart transportation system to provide wireless connection between vehicles, roadside drivers, passengers and pedestrians. Despite considerable potential market opportunities and gains, V2X communication technology is still in the field trial phase in general and new advanced algorithms should be developed for this type of communication [11].

Applications in vehicle networks; due to the different performance requirements, it can be classified according to passenger and road safety, traffic flow efficiency and

infotainment types. However, as an important point here is that road safety applications in vehicle-to-vehicle (V2V) communication require low latency and high reliability [12]. And due to safety reasons, priority must be given to minimize especially security risks of communication on the connected vehicle [13].

IV.   COMMUNICATION INFRASTRUCTURE

Various applications in the vehicle network pose specific requirements and challenges in the wireless communication

technology and it is needed multiple security control techniques to secure privacy sensitive data [14] and privacy of the ecosystem [15]. To provide the communication infrastructure required by vehicle communication as shown in Fig. 4, various wireless communication technologies are available, such as Wi-Fi, cellular network systems and IEEE 802.11p.
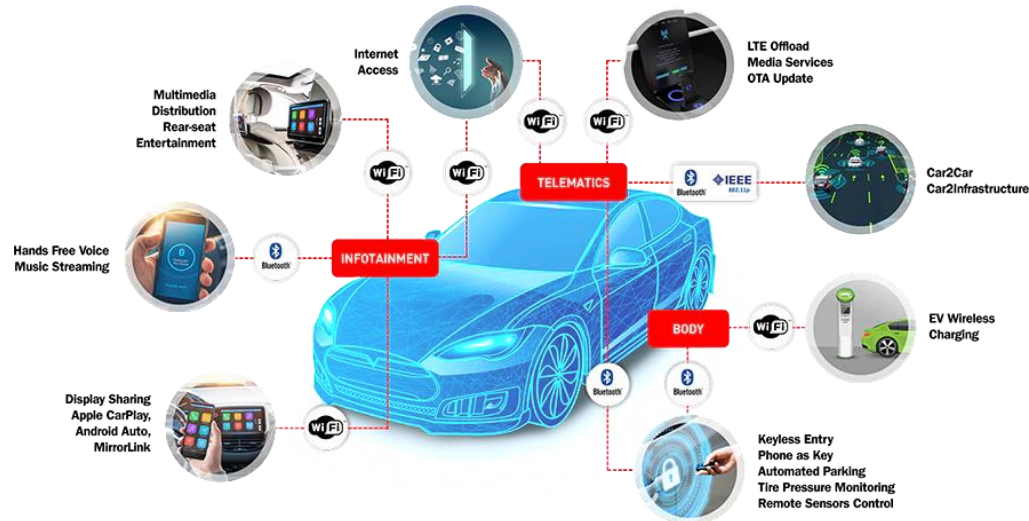
Fig.4. Communication technologies on the connection vehicle [16]

Wi-Fi communication is not suitable to support high mobility due to its short range. Although IEEE 802.11p is considered the first wireless technology standard specifically designed for road communication, it has obvious disadvantages such as low reliability, hidden node problem and unlimited latency. The widespread use of IEEE 802.11p requires major investments in network infrastructure. Due to these disadvantages of IEEE 802.11p, new efforts have been made to use 5G in the long-term evolution (LTE) as an optimal potential wireless access technology to support vehicle applications.

It is vital that the communication must take place instantly and without delay. At this point, it is very rational to look for the solution at 5G since the 4/4.5G technology we are currently using is insufficient. We face the importance of 5G for connected vehicles and IoT at this point.

5G technology, one of the indispensable technologies for connected vehicles, is extremely critical with its ultra-fast data transfer and latency below 1ms. Intel predicts that connected vehicles will spend unbelievable data.

The figures are hard to believe. Connected vehicles are expected to consume 4 TB of data per vehicle in a day when it enters our lives completely! Each vehicle will exchange about 4 TB of data, and this data traffic is only possible with 5G [16].

The challenge is how data can be continuously sent and received from the cloud. The volume of data is very important.

To realize sending and receiving huge amount of data, new technologies and new infrastructures are needed.

While the connected vehicles have many important features and benefits brings to next life, we will focus on easy function upgrade and maintenance of these connected vehicles via such processes like remote monitor and control.

According to BMW, "first remote-controlled function is used 2001. Vehicle diagnostics, also known as remote diagnosis, have produced a connection between users and their cars across vast distances since 2001. With vehicle diagnostics, vehicle manufacturers can examine the system's functionality and, if there are any problems, potentially recognize the cause more quickly".

A.   Remote maintenance methods

OTA, which is a process of the data transmission and reception for application-related information in a wireless communication system. OTA was initially associated with consumer electronics such as television and later it started entering every industry from telecommunication to household appliances and the automotive industry.

In the automotive industry, OTA process has some delivery solution ways. Software provider and manufacturer can coordinate in a system architecture [18]. In this system architecture data can be sent to the vehicle from a cloud-based server via a wireless network as shown in Fig. 5a, or mobile network directly to the vehicle as shown in Fig. 5b, or to the vehicle owner's mobile device and then installed directly to

vehicle from mobile device via a wireless connection such as Bluetooth as shown in Fig. 5c.

Along with the prevalence of connected vehicle technology, a usage of over the air software update is also progressing for easy function upgrade and maintenance [19].
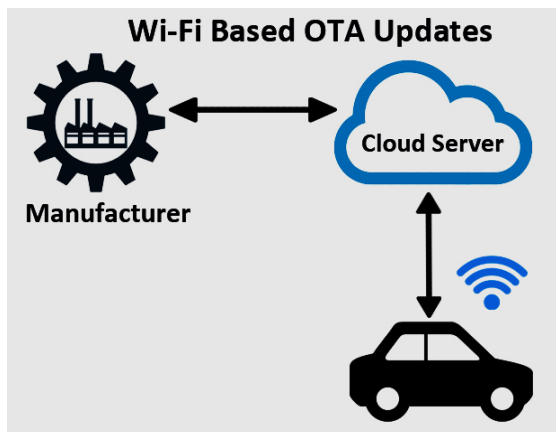
Fig.5a. Wi-Fi based solution delivery model for OTA updates [19]
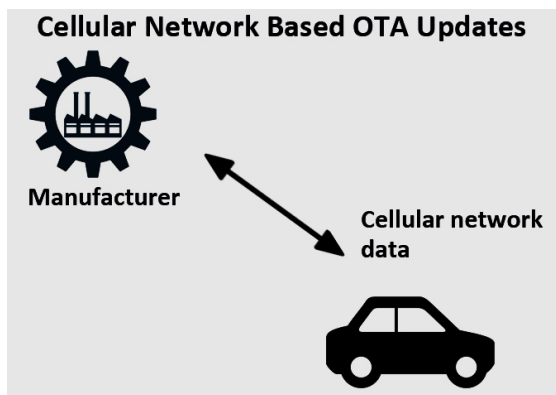
Fig.5b. Cellular network-based solution delivery model for OTA updates [19]
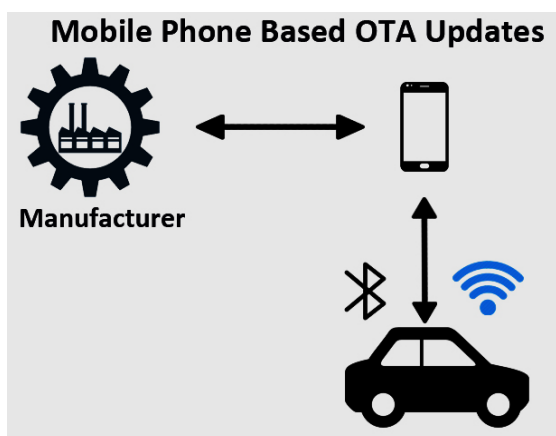
Fig.5c. Mobile phone-based solution delivery model for OTA updates [19]

OTA connection and updating process is critical and should be robust and reliable, both in terms of security and user safety. Thus, it needs detailed security architecture while the updating process [20]. When a vehicle's firmware or default software have a software-related security bug or vulnerability,

it must be corrected immediately. This kind of correction by OTA updates is also convenient and comfortable for vehicle owners because by this way no longer need to go a service center to receive the updated software [21]. All of the vehicle's software related its system must be up to date for security reasons [22]. So that, OTA updates are main issue for the safety and keeping vehicle's features and functions up to date.

Developers try to fix known software bugs or integrate several numbers of new additional features to the vehicle by software updates. The complexity of automotive software has made it more difficult to verify software updates. A strong and reliable verification methodology is required during this process. Firstly, software update source must be verified [20], second, it is ensured that any bug fixes are implemented correctly, then it is ensured that the new additional features are running correctly, and finally it is ensured that the entire software update does not harm any other functions [21].
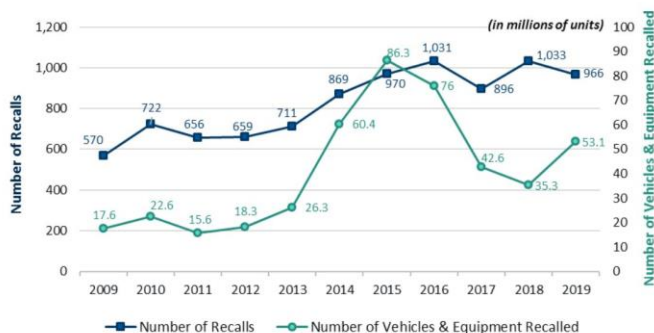
Fig.6. Comparison of recall issues with recalled vehicles [24]

Software dependency of vehicles increasing more and more regarding electronic components used on the vehicles. Because of the complexity of these hardware and software infrastructures, the amount of recalling has risen dramatically. As a comparison of the number of recalls with the number of recalled vehicles regarding various faults, bugs in the software, etc. by the period of 2009 to 2019 are shown in Fig. 6. It is prominently displayed on the figure, the number of issues and the number of vehicles related recalling has increased since 2013. While there is approximately 60 million increase in Overall recalled vehicle number from 2013 to 2016, there is approximately 51 million decrease from 2016 to 2018 [24]. Many manufacturers recalled such number of vehicles due to software issues. Remote OTA updates provide many benefits to keep the vehicle and its functions up to date. All these recalls could have been eliminated by the implementation of remote OTA software updates. As a valuable result of avoiding recalls sustainable consumer satisfaction could be built-in self.

OTA process offers two types of methods, one is SOTA and the other is the FOTA.

SOTA is the meaning of sending a software file from cloud to vehicle. The communication between cloud and vehicle can be implemented in some ways, wireless communication or mobile communication can be used to download the software file to the vehicle.

Instead of sending a full image of the software installation file, SOTA updates can be downloaded as a part of the whole file. Thus, the developer can forward only a part of a file, which is needed to be fixed on the vehicle. While this can reduce download time and make it safer with the reduction of traffic on the communication data line.

The installation file may contain several files to update different hardware and software on the vehicle. Especially critical hardware of vehicle must be considered while doing SOTA updates [25]. SOTA must be cyber-safe, making sure that car manufacturers only apply the appropriate updates to their vehicles. Updates, which can be sent by the SOTA method may include changes to software that controls the physical parts of the vehicle and to user interface software such as infotainment screens or instrument panels. The SOTA update method enables automotive manufacturers to fix, maintain, and check for disruptions in vehicles, also to improve them for future needs through remote software updates sent to the vehicle from a cloud-based server. These kinds of improvement-based technologies can offer both cost savings and revenue earnings to the manufacturers, while also improving customer satisfaction and brand value with remedies, modifications and enhancements that are applied remotely, timely and comfortably.

Vehicle that need fixing, upgrading, and maintenance are likely to pay attention to their software as if it's a mechanical thing; in this case SOTA eliminates the need to go to service center. SOTA will also provide an opportunity to be quickly updated and corrected when a problem arises that can pose major security concerns.

Remote enabled SOTA and FOTA include such kind of OTA enabled segments, such as navigation map, infotainment applications, telematics control unit (TCU) and electronics control unit (ECU) updates. Every functions of vehicle must be implemented individually for security reasons [26].

TCU updates are carried over FOTA in real time, subject to the condition that the TCU is connected to the server. FOTA checks for the correct update package and then finishes the process of updating the vehicle remotely, thereby reducing the correction cycle time. SOTA offers such updates as below:

- Updates based on applications
  - ➤ Navigation Updates
  - ➤ Head Unit Display Updates
  - ➤ Telematics System Updates
- Updates based on software platforms (infotainment operating system and operating platform)
  - ➤ Embedded Applications
  - ➤ Featured Software
  - ➤ Infotainment Software

FOTA offers updates such as TCU and ECU.

By 2022, analysts expect vehicle manufacturers to realize approximately 350 million software updates regarding application based, infotainment software platforms and telematic services for their customers' vehicles by using software SOTA process. And another point of view, ABI Research, leading in market intelligence, forecasts nearly 203 million OTA enabled vehicles will be produced with approximately number of 180 million vehicles featured with SOTA and approximately number of 22 million vehicles featured with FOTA by 2022. Components of OTA can be classified like that:

- Application providers
- Solution providers
- OTA Platform Providers
- Infrastructure providers;
  - ➤ Cloud infrastructure; such as Microsoft, Sierra Wireless
  - ➤ Communication infrastructure; such as AT&T, Verizon
  - ➤ Cyber security solution providers; such as Visual Threat, Airbiquity, Gemalto, Escrypt, Infineon
  - ➤ Component manufacturers

In the vehicle industry, many of the manufacturers are already involved with these structures, and telematics and electronic control units are going to be the primary focus areas in the future. For example, Hyundai has proof of many concept systems in terms of OTA map updates, on the other side BMW, Audi and Tesla have already started rolling out OTA procedures for updating navigation maps, Ford's Sync 3, which is interactive touch screen system, will be accomplished through an OTA using Wi-Fi. Autonomous car manufacturer, Tesla use OTA, their vehicles regularly receive over the air software updates which add new features and functionality.
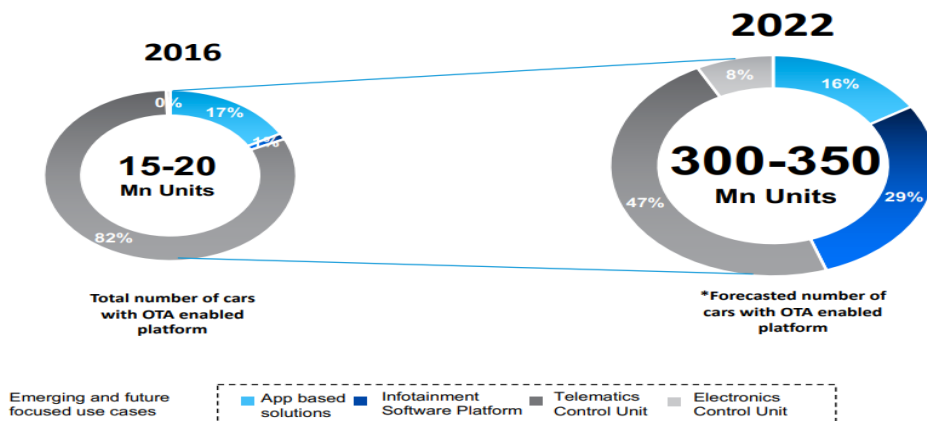

Fig.7. Total market size in terms of the number of vehicles [27]

Total market size in terms of the number of vehicles is curated from the IHS market report and GEIP platform data (source: DRAUP).

Fig. 7 shows the number of vehicle units that include OTA updates related to application-based solutions, infotainment software platforms, TCU and ECU. These OTA enabling updates will see the highest interest in the number of vehicles. Application-based solution programs take up less space in total core memory when compare to others and have low-level security issues; therefore, it can be implemented easily as OTA enabled platform.

Infotainment software OTA updates are complex because the programs can be quite large. These updates can take place over Wi-Fi communication instead of mobile communication because of the mobile network limitations.

Types of updates include binary files for ECUs or other firmware updates, mapping data for onboard navigation, user-generated contents such as music, videos and photos, operating systems, applications and software package updates, and such kind of other updates.

According to the benefits of OTA processes, total worldwide OEM cost savings from OTA software update events is forecasted by growing from 2.7 billion USD in 2015 to more than 35 billion USD in 2022 [28]. When comparing with others, telematics control unit and infotainment software platform updates are contributing most to these figures to become alive. It shows that value-added cost savings by OTA updates will be the most valuable part of this kind of new future solution. There are some success stories, which will be useful to mention:

- MinVolkswagen

Denmark based Connected Cars company, which is founded in 2016, aims to provide smart and effective tracking of vehicles for workshops and fleet managers, as well as a better driving and service experience for vehicle owner.

The software they developed as simulated in Fig. 8, helped Volkswagen Denmark to improve customer satisfaction and add a new life to the customers' old cars. Free connection service was offered to the customers with a Volkswagen model year 2008 and later vehicles. This allows vehicle owners to see the status and usage of their vehicles at first a glance and enables direct communication between the service and customers via chat and exchange of information. In addition, the services know the status of their customers' vehicles almost in real time and can proactively help them.

Fig.8. Mobile application for vehicle status [29]

Developed software called MinVolkswagen is a digital service which is offered to Volkswagen customers [29]. It aims to ensure that users are connected and under control of their vehicles. This technology also serves the service workshops, enabling the services to follow customer vehicles technically, to be more productive in production and service quality and to offer the highest level of customer satisfaction.

- Mercedes PRO Connect

Mercedes-Benz plans to offer a wide range of digital services to the fleet customers by a platform, called Mercedes PRO connect [30], which one makes the van transparent and allows it to be used even more efficiently by collecting vehicle data via a digital vehicle log. The communication equipment required for the internet connection is integrated into the vehicle hardware as a default component when the vehicle is manufactured. So, all fleet customers will be able to make the best use of Mercedes PRO connect in the future. Thanks to the web-based service, it provides fleet managers with an overview of their vehicles, for example, they can manage orders in real time, check which vehicle is where, whether it needs refueling or not, or go into the workshop for maintenance.

V. CONCLUSION

While remote connection technologies offer huge opportunities for the automotive industry, these will require manufacturers to review their plans, redesign their processes and release new software updates to bring these new technologies to life.

There will also be issues to be concerned and overcome. For example, manufacturers may be concerned about remote connection safety, because there may be security issues when new technologies become alive. While the vehicles are becoming more connected with any of useful technology, which has the ability to connect each other and everywhere on its ecosystem, a reassuring security infrastructure is needed to ensure the privacy and security of the vehicle and its occupants.

Nevertheless, as the advantages of OTA technology and the potential opportunities it brings to the manufacturer are noticed, the investments in this technology will increase and the obstacles we mentioned will be overcome easily. As connected vehicles' dependency to the software will increase, OTA will become an indispensable necessity.

On the other hand, OTA can help manufacturers to improve their vehicles without calling back it. For example, manufacturers can release OTA updates where a vehicle has the opportunity to improve the quality of reducing fuel consumption. Manufacturers can also use SOTA or FOTA for customer satisfaction, for example, to keep vehicle functions up to date with new software and firmware releases and vehicle infotainment services.

OTA technology has the ability of bidirectional communication from vehicle to manufacturer or vice versa. Manufacturers could collect data from vehicles for their production quality analyses for future improvements or to help vehicle owners by applying of any immediate correction of bugs to prevent failures before occurring.

REFERENCES

[1] Ericsson < https://www.ericsson.com/en/internet-of-things/automotive>, Accessed: 24 March 2023.

[2] BMW <https://www.bmw.com/en/innovation/connected-car.html>, Accessed: 24 March 2023.

[3] https://www.mercedes-benz.com/en/vehicles/passenger-cars/mercedes-benz-concept-cars/vision-avtr/ Accessed: 24 March 2023.

[4] E. Uhlemann. "Introducing Connected Vehicles." [Connected Vehicles] in IEEE Vehicular Technology Magazine, vol. 10, no. 1, pp. 23-31, March 2015, doi: 10.1109/MVT.2015.2390920.

[5] B. Masini, G. Ferrari, C. Silva, I. Thibault. "Connected Vehicles: Applications and Communication Challenges." Mobile Information Systems. 2017. 1-2. 10.1155/2017/1082183.

[6] G. S. Vasilash. "OnStar: 10 Years After." Automotive Design and Production, 2006.

[7] K. R. Chen, L. Chun-Chung, H. Cheng-Hung H. "Vehicle burglar alarm system with GPS recognition." U.S. Patent No. 7,151,441. 19.

[8] S. Lee, G. Tewolde, J. Kwon. "Design and implementation of vehicle tracking system using GPS/GSM/GPRS technology and smartphone application." IEEE World Forum on Internet of Things (WF-IoT), Seoul, 2014, pp. 353-358, doi: 10.1109/WF-IoT.2014.6803187.

[9] Kiran, Jonnalagadda Surya, and Pranay Reddy Jakkala. "Cyber Security and Risk Analysis on Connected Autonomous Vehicles." Solid State Technology (2020): 10161-10176.

[10] S. K. Datta, R. Costa, J. Härri, C. Bonnet. "Integrating connected vehicles in Internet of Things ecosystems: Challenges and solutions." 1-6. 10.1109/WoWMoM.2016.7523574.

[11] J. He, A. Radford, Z. Xiong, Z. Tang, X. Fu, S. Leng, F. Wu, K. Huang, J. Huang, J. Zhang, Y. Zhang. "Cooperative Connected Autonomous Vehicles (CAV): Research, Applications and Challenges." 1-6. 10.1109/ICNP.2019.8888126, 2019.

[12] M. Bennis, M. Debbah, H. V. Poor. "Ultrareliable and low-latency wireless communication: Tail, risk, and scale." Proceedings of the IEEE, 106(10), 1834-1853, 2018.

[13] S. Parkinson, P. Ward, K. Wilson, J. Miller. "Cyber Threats Facing Autonomous and Connected Vehicles: Future Challenges." IEEE Transactions on Intelligent Transportation Systems, vol. 18, no. 11, pp. 2898-2915, Nov. 2017, doi: 10.1109/TITS.2017.2665968.

[14] T. Nawrath, D. Fischer, B. Markscheffel. "Privacy-sensitive data in connected cars." 11th International Conference for Internet Technology and Secured Transactions (ICITST), Barcelona, 2016, pp. 392-393, doi: 10.1109/ICITST.2016.7856736.

[15] S. Chen, J. Hu, Y. Shi. "LTE-V: A TD-LTE based V2X Solution for Future Vehicular Network." IEEE Internet of Things Journal. PP. 1-1. 10.1109/JIOT.2016.2611605.

[16] https://www.cypress.com/blog/corporate/driving-connected-car-revolution Accessed: 24 March 2023.

[17] Karel <https://www.karel.com.tr/blog/5g-ve-iot-otonom-araclar-icin-neden-onemli>, Accessed: 01 May 2020.

[18] Y. Zhou, X. Wu, P. Wang. "Secure Software Updates for Intelligent Connected Vehicles." Electrical Engineering and Computer Science (EECS), 3, 109-112.

[19] A. Kanda, T. Kurafuji, K. Takeda, T. Ogawa, Y. Taito, K. Yoshihara, M. Nakano, T. Ito, H. Kondo, T. Kono. "A 24MB Embedded Flash System Based on 28nm SG-MONOS Featuring 240MHz Read Operations and Robust Over-The-Air Software Update for Automotive Applications." IEEE Solid-State Circuits Letters. PP. 1-1. 10.1109/LSSC.2019.2948813.

[20] M. Steger, A. Dorri, S. Kanhere, K. Römer, R. Jurdak, M. Karner. "Secure Wireless Automotive Software Up-dates using Blockchains – A Proof of Concept."

[21] D. Coe, J. Kulick, A. Milenkovic, L. Etzkorn. "Virtualized In Situ Software Update Verification: Verification of Over-the-Air Automotive Software Updates." IEEE Vehicular Technology Magazine. PP. 10.1109/MVT.2019.2954302.

[22] Ş. Okul, M. A. Aydin, F. Keleş. "Security Problems and Attacks on Smart Cars." In: Boyaci A., Ekti A., Aydin M., Yarkan S. (eds) International Telecommunications Conference. Lecture Notes in Electrical Engineering, vol 504. Springer, Singapore.

[23] K. Mayilsamy, N. Ramachandran, V. Raj. "An integrated approach for data security in vehicle diagnostics over internet protocol and software update over the air." Computers & Electrical Engineering. 71. 578-593. 10.1016/j.compeleceng.2018.08.002.

[24] Motor Vehicle Safety: Issues for Congress <https://sgp.fas.org/crs/misc/R46398.pdf> 24 March 2023.

[25] A. Freiwald, G. Hwang. "Safe and Secure Software Updates Over The Air for Electronic Brake Control Systems." SAE International Journal of Passenger Cars - Electronic and Electrical Systems. 10. 10.4271/2016-01-1948.

[26] A. Camek, C. Buckl, A. Knoll, A. "Future cars: necessity for an adaptive and distributed multiple independent levels of security architecture." 17-24. 10.1145/2461446.2461450.

[27] https://www.abiresearch.com/press/abi-research-anticipates-accelerated-adoption-auto/ Accessed: 24 March 2023.

[28] M. Khurram, H. Kumar, A. Chandak, V. Sarwade, N. Arora, T. Quach. "Enhancing connected car adoption: Security and over the air update framework." IEEE 3rd World Forum on Internet of Things (WF-IoT), Reston, VA, 2016, pp. 194-198, doi: 10.1109/WF-IoT.2016.7845430.

[29] Connectedcars <https://connectedcars.dk/>, Accessed: 24 March 2023.

[30] Daimler <https://www.daimler.com/products/vans/advance.html >, Accessed: 24 March 2023.

BIOGRAPHIES

ERKAN DEMIR received his B.S. in Electronics Engineering from the University of Erciyes, Kayseri, Turkey, in 2004. His research interests include automotive electronics and solutions.

HABİB KAYMAZ Kasımpaşa, İstanbul, in 1970. He received the B.S. and M.S. degrees in Electric Education from the University of Marmara. He has been working for Mercedes Benz Türk AŞ, Bus Development as a lead test engineer since 1997. He earned his doctorate degree with the study of "Development of Driving Cycle for Hybrid and Electrical BRT vehicle" in Marmara University Institute for Graduate Studies in Pure and Applied Sciences in 2018.

His research interests include Automotive Electric-Electronic Systems, Automotive Testing, Life Test, Reliability, Connectivity, Hybrid/Electrical Powertrains, Measurement Systems and Sensors, Data Acquisition, Data Mining and Analyzing.

İLKER YILDIZ received his M.S. degree in Institute of Turkic Studies from the University of Marmara, İstanbul, Turkey, in 2002. His research interests include History of Technology.

Recent Topology-based Routing Approaches in VANETs: A Review

Ali Fuat Gunes, Ipek Abasikeles-Turgut

Abstract—**Due to increasing both safety and efficiency of the traffic, Vehicular Ad Hoc Network (VANET) is a promising technology of Intelligent Transport Systems (ITS). Unique characteristics of VANETs including high mobility and strict delay constraints, require new routing solutions specific to these networks to be proposed. As one of those solutions, topology-based routing approaches aim to find the shortest path by managing routing tables. In this paper, recent topology-based routing approaches for VANETs are investigated in detail. Proactive, reactive and hybrid solutions are compared with respect to their advantages, disadvantages, updating procedures and network sizes. This paper will shed light on future studies since it provides detailed information about the current status of the literature in topology-based routing approaches in VANETs.**

Index Terms—**Proactive, reactive, survey, topology, vehicular ad hoc networks routing.**

I. INTRODUCTION

VEHICULAR AD HOC NETWORKS (VANETs), an increasingly important component of Intelligent Transport Systems (ITS), consist of vehicles in motion communicating with each other and road side units (RSU) as infrastructures. VANET, as seen in Fig 1, is a popular topic in both academia and industry due to its usage in critical areas such as enhancing traffic safety and efficiency, providing real-time and secure data to drivers and passengers, and managing emergencies such as accident prevention, ambulance guidance and fire brigade assistance [1].

VANET services can be summarized as follows [2]:
• Security service: Providing automatic driving functions and mitigating traffic risks,
• Data sharing services: Enhancing communication comfort for drivers and passengers by exchanging information about the surrounding road conditions.

ALİ FUAT GÜNEŞ, is with Department of Computer Engineering University of Iskenderun Technical University, Hatay, Turkey,(e-mail: aligunes.lee21@iste.edu.tr).

https://orcid.org/0009-0004-8656-3951

İPEK ABASIKELEŞ-TURGUT, is with Department of Computer Engineering University of Iskenderun Technical University, Hatay, Turkey (e-mail: ipek.abasikeles@iste.edu.tr).

https://orcid.org/ 0000-0002-5068-969X

In order to provide these services, VANETs need to process real-time data on the network. These networks exhibit a dynamic topology due to the high mobility of vehicles. As vehicles frequently change their positions within the network, interruptions in established routes occur frequently [3]. Therefore, setting up the route from source to destination is a difficult task in VANETs [4]. Routing protocols play a crucial role in reducing network congestion by enabling dynamic routes between vehicles. The reduction in network traffic can yield economic benefits such as reduced battery and fuel consumption in vehicles. Numerous studies in the literature are dedicated to develop a routing protocols specifically designed for VANETs [1],[2],[4].

Routing protocols in VANETs are categorized into different classes, including topology-based, position-based, broadcast-based, cluster-based, and geo-cast-based approaches [5]. Among these, topology-based routing focuses on determining the optimal route between the vehicles by utilizing and analyzing information such as the current road and traffic conditions. This approach enhances the security of data flow while reducing both packet losses and delays [6]. Consequently, topology-based routing architectures have garnered significant attention from researchers due to their pivotal role in improving the efficiency, security and performance of VANETs.

While there are existing review articles for position-based [7]-[12][7] below, broadcast-based [13],[14], cluster-based [15],[16] and geo-cast-based [17],[18] routing architectures in VANETs, the number of surveys analyzing recent topology-based protocols [19],[20] is relatively limited. Therefore the objective of this study is to classify topology-based VANET routing protocols and conduct a detailed examination of each class. The studies are compared using tables on various parameters, and open research areas are highlighted. It is anticipated that this paper will provide valuable insights into future developments of topology-based routing algorithms in VANETs.

Section II of the paper provides a review of the existing studies in the literature and highlights the contribution of this study. Section III presents a summary of the research methodology employed in this study. In Section IV, a comprehensive analysis of the examined topology-based studies is provided, clarifying the details of each study. Section V involves the general comments and remarks obtained from the reviewed studies. Finally, Section VI concludes the paper.
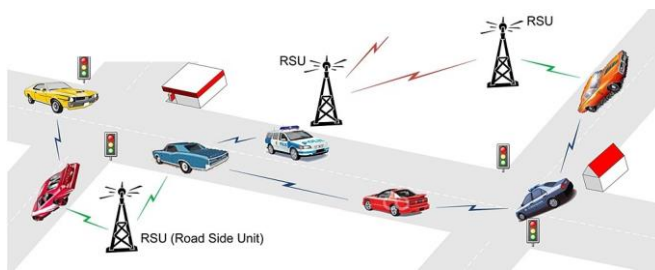
Fig.1. A VANET infrastructure consisting of vehicles and RSUs

II. LITERATURE REVIEW

In the literature, various classification types of routing algorithms and protocols in VANETs can be found. For instance, Srivastava et. al [[7]], classify the routing protocols into two classes based on transmission strategy and route formation. The transmission strategy class includes unicast, multicast (such as geo-cast and cluster-based approaches), and broadcast approaches. The route information class can be further divided into location-based and topology-based approaches. Other studies [[5], [13], [20]] categorize routing protocols in VANETs into five classes including position/geographic/location-based, cluster-based, broadcast-based, geo-cast based and topology-based protocols. In this study, five different class of routing architectures are considered as illustrated in Fig. 2.
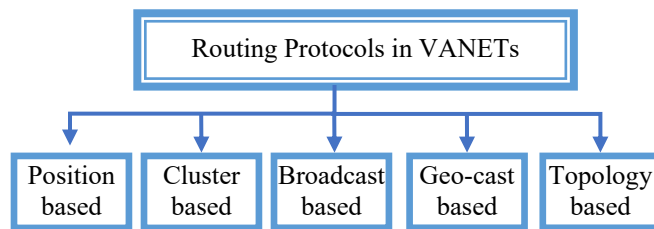

Fig.2. Classification of routing protocols in VANETs

In position based routing, the focus is on utilizing the geographic location information of the nodes rather than establishing end-to-end connectivity [[7]]. Cluster-based routing approaches divide the network into smaller regions known as clusters [[16]]. A designated node, known as the cluster head, is responsible for collecting and transmitting data from its member nodes within the cluster. In broadcast-based routing protocols [[13]], data packets are propagated throughout the network using broadcast transmission techniques. Geo-cast routing [[17]], a type of position-based routing, selects a specific area based on the location of the destination vehicle and transmits data through broadcasting within that relevant area. Topology-based routing protocols [[20]] operate by utilizing the network's connectivity information to create routing tables. These protocols are further classified into three classes: proactive, reactive and hybrid. Topology-based routing protocols offer advantages such as being suitable for smaller networks and endeavoring to find the optimal (i.e., shortest) route from the source to the destination node [[21]].

Most of the recent survey papers (published between 2017 and 2023) focusing on routing protocols in VANETs primarily concentrate on position-based approaches [[7]-[12]]. Some studies delve into broadcast-based [[13],[14]], cluster-based

[15],[16], and geo-cast-based [17],[18] methods. Additionally, there are reviews [14],[22],[23] that examine protocols from the perspective of quality of service (QoS). However, survey papers specifically addressing topology-based routing are relatively scarce. One study is tailored for FANETs [19], while another [20] examines only four different protocols, comparing their performances through simulation under various metrics. Thus, it is evident that there is a need in the literature for a review article that sheds light on recent papers focusing on topology-based routing approaches.

The objective of this study is to provide an overview of the state-of-the-art in recent topology-based routing architectures in the literature. The contributions of this study can be summarized as follows:

• Examination of 38 different topology-based routing approaches: The study conducts a comprehensive analysis of 38 distinct approaches, providing detailed insights into each approach.

• Classification and comparison: The papers are categorized based on their routing architectures, and a comparative analysis is performed within each class. The approaches are evaluated and compared with respect to various routing and network parameters. The advantages and disadvantages of each approach are emphasized.

• Identification of open research areas: The study identifies and highlights open research areas that assure further investigation and exploration in the field of topology-based routing in VANETs.

III. RESEARCH METHODOLOGY

By synergistically combining cutting-edge research, real-world experience, and deeply ingrained human values, Evidence-based Software Engineering (EBSE) endeavors to elevate decision-making practices in software development and maintenance to new heights [24]. The foundation of EBSE relies heavily on the utilization of systematic reviews. A systematic literature review (SLR) is a rigorous and exhaustive evaluation of all available research pertaining to a particular research question, topic area, or phenomenon of interest [24], [25]. The primary objective of conducting an SLR is to ensure a methodical, replicable, and thorough review process. By employing systematic reviews, EBSE ensures a robust and unbiased approach to synthesizing evidence for informed decision-making in software development and maintenance [26]. Therefore, we have followed one of the widely recognized systematic review guidelines [26] and we have applied the same strategy as in the existing literature, e.g., [27][28], [29].

The papers included in this study were collected from four databases: IEEE Xplore, ScienceDirect, Springer, and Google Scholar. Both conference and journal publications were considered in the search.

Based on the information provided in [30], which states that protocols like AODV, DSR, DSDV, and OLSR proposed for MANETs form the foundation of topology-based routing protocols in VANETs, the searches were divided into two groups:

1. The keywords "PROACTIVE OR REACTIVE OR TOPOLOGY AND ROUTING AND VANET" were used to find both hybrid protocols and topology-based approaches.

2. The keywords "AODV or DSR or DSDV or OLSR and VANET" were used to identify different versions or variants of these foundational protocols.

After obtaining 724 articles, those that solely compared the performance of existing protocols, that do not have full text, that are not written in English, that do not verify the proposed methodology and review/survey papers were excluded. Duplicate studies were also eliminated. Thus, the number of articles was reduced to 28. The original versions of the relevant studies were identified manually and included by thoroughly examining the articles. As a result, a total of 38 articles were obtained.

By conducting this search strategy, the study aims to gather a comprehensive collection of relevant articles on topology-based routing protocols in VANETs, ensuring the inclusion of original research and minimizing redundancy.

IV. OVERVIEW OF TOPOLOGY-BASED ROUTING

Topology-based routing protocols in VANETs employ routing tables on nodes to handle data transmission and update these tables as the network topology changes. Message transmission can be performed using three methods: unicast, multicast and broadcast. These protocols are classified into three classes: proactive, reactive and hybrid as depicted in Fig 3 [30]

Proactive routing protocols, also known as table-driven routing, manage and store routing information about all nodes in the network in tables located at every node. Nodes ensure that the routing information is kept up-to-date by exchanging tables with other nodes to accommodate changes in the network topology [19], [24].

In reactive routing protocols, also known as on-demand routing, the route to the destination node is discovered when a node wants to forward a data packet to another node. However, unlike proactive routing protocols, the routing table is not continuously updated. Instead, only active routes are logged and maintained [19], [24].

In this paper, proactive, reactive, and hybrid approaches are individually investigated and analyzed. Each approach is briefly summarized, highlighting its key characteristics. To provide a comprehensive understanding and facilitate comparison, tables are created to compare the updating strategies employed by these approaches. These tables allow for a systematic evaluation of the pros and cons of each approach, shedding light on their respective strengths and weaknesses. By presenting this comparative analysis, the paper aims to provide valuable insights into the different updating strategies and their implications in VANET routing protocols.

A. Proactive Solutions

Proactive routing protocols in VANETs can be classified into two main classes: Distance Vector Algorithm [31] and Link State Algorithm [32], based on their routing approaches. Distance Vector-based protocols utilize the Bellman-Ford Algorithm [33] for routing decision-making. On the other hand, Link State-based protocols employ the Dijkstra Algorithm [34] for route computation.

Distance vector based proactive routing protocols are summarized as below and the comparison of these protocols with their advantages and disadvantages are shown in Table I.

In Distance Vector Routing Protocol (DVRP), each node maintains a routing table that stores information about its distance to other nodes in the network and the optimal route to reach those nodes. The routing table is periodically exchanged with neighboring nodes or in the event of a network disconnection. This process continues iteratively until each node in the network has updated its routing table with the most optimal paths to all destinations [35].

In Destination-Sequenced Distance Vector Routing (DSDV), each node in the network determines the optimal path to destinations based on the distance vectors provided by its neighboring nodes. However, DSDV introduces the use of sequence numbers to prevent routing loops and ensure consistent routing table formation. Another key feature of DSDV is the partial update of routing tables. Instead of updating the entire routing table, DSDV only updates entries for the destination networks that have experienced changes. This approach reduces the overhead on network traffic, as nodes transmit and process only the necessary updates, rather than exchanging complete routing tables [36].
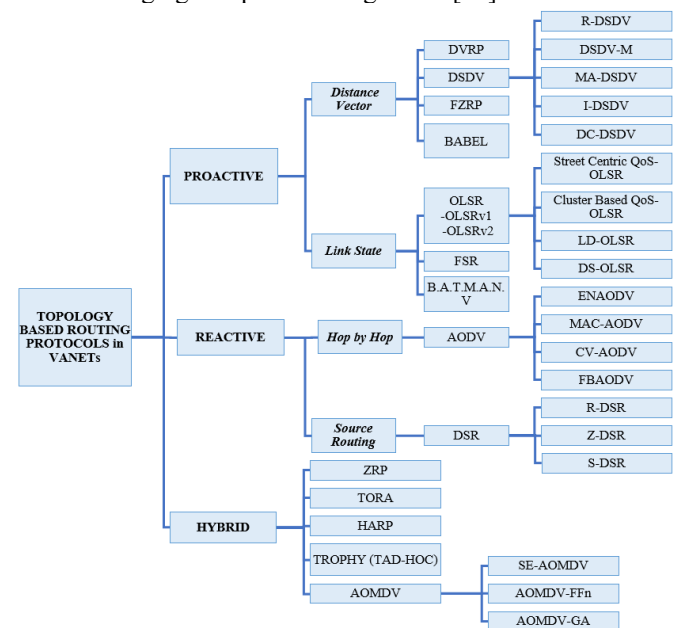
Fig.3. Topology-based routing protocols in VANETs

In Randomized-DSDV (R-DSDV), routing tables are updated at random time intervals to distribute network traffic and facilitate the rapid dissemination of updated routing information [37].

In Multi-Agent-DSDV (MA-DSDV), individual agents maintain their own routing tables and exchange routing information with neighboring agents to ensure up-to-date table information. This approach enables effective and efficient routing in complex network environments [38].

Dual-Interface Multiple Channels (DSDV-M) is a routing protocol designed for networks with devices having multiple wireless interfaces. It allows nodes to maintain updated

routing tables, enabling effective communication across multiple wireless channels simultaneously. This enhances network capacity and flexibility [39].

Improved-DSDV aims to enhance the stability of routing information in the network by responding more quickly to topology changes compared to the classical DSDV protocol. It achieves this by minimizing routing information changes between nodes, thereby reducing energy consumption. Timestamps are used in shared routing tables to mitigate routing errors and ensure reliable sharing of routing information. The protocol focuses on improving the overall reliability and efficiency of routing in the network [40].

region have more frequent updates in their routing tables, while nodes farther away have less frequent updates [42].

BABEL is a versatile routing protocol that enables the use of multiple paths with different metrics. It dynamically updates routing tables to optimize path selection based on factors such as link quality, available bandwidth, and delay. This enhances network reliability and adaptability [43]-[45].

Link state-based proactive routing protocols are summarized as below and the comparison of these protocols with their advantages and disadvantages are shown in Table II.

OLSR is used for determining the shortest path and routing network traffic by establishing the connection states between the nodes. This protocol follows the neighborhood relations and link states of the nodes; shares them among the nodes and creates the routing table accordingly [46]-[48].

Street Centric QoS-OLSR is an extension to the OLSR protocol, which provides QoS support. This protocol optimizes communication in the network by considering physical routes such as streets and avenues and performs routing accordingly. By leveraging street-centric information, Street Centric QoS-OLSR selects routes that take into account the specific characteristics and constraints of the physical environment [49].

Cluster Based QoS-OLSR is also an extended version of the OLSR protocol that incorporates QoS support. This protocol introduces a logical division of the network into clusters, where each cluster is assigned a cluster head. By centralizing the routing decisions within each cluster, the protocol enables better management of QoS requirements in the network [50].

An QoS Aware Link Defined-OLSR (LD-OLSR) is an extension of the OLSR protocol that incorporates QoS support. This protocol is designed to route traffic with different levels of service qualities in the network, ensuring that priority traffic, such as voice or video data, is given lower latency and reduced packet loss rates. By reducing the disparity between different levels of QoS, LD-OLSR aims to provide a more consistent and reliable service in the network [51].

Disaster Scenario Optimized-OLSR (DS-OLSR) is also an extended version of OLSR protocol specifically designed to address the challenges of communication network configuration in disaster scenarios. Since this protocol can replace disconnected nodes with other nodes after a natural disaster, it ensures the essential communication functionality required for emergency activities [52].

Divide Cluster-DSDV divides the network into smaller subnets, managed by dedicated routing nodes. Subnets share local routing information, reducing network traffic and enabling faster routing decisions. Bridge nodes facilitate communication between subnets when needed, improving overall network efficiency and scalability [41].

Fisheye Zone Routing Protocol (FZRP) is a routing protocol designed to control network traffic by limiting updating of routing tables between nodes. Each node in the network updates the routing information of its neighbors and maintains up-to-date routing tables based on this information. However, the frequency of updating the routing tables of neighboring nodes is dependent on their distance. Nodes in a specific

FSR performs the routing process by focusing on a specific zone within the network nodes. The protocol maintains two distinct routing tables, including global and local, to route the communication between the nodes. The global routing table contains general information about all nodes in the network including their connections. On the other hand, the local routing table includes detailed information about the connections between the nodes and their neighbors [53].

Better Approach to Mobile Ad-Hoc V (B.A.T.M.A.N V) constantly monitors the link quality between nodes in the network and makes the routing decisions by selecting the most reliable links. This approach aims to improve network efficiency and reliability [54]-[55].

B. Reactive Solutions

Reactive protocols allow the nodes in the network to communicate directly with each other. Basically, these protocols are divided into two classes according to routing logic, including node-to-node hop (Hop by Hop[56]) and source routing (Source Routing[57])[62]. In hop by hop routing, the data transmitted from one node to another is received and directed by each node. This way, data is routed by each node until it reaches the destination node. This technique is used to choose the shortest path [56]. On the other hand, data is directly routed to the destination node by the sending node in source routing. In this method, the sending node determines relay nodes through which the data must pass to reach the destination node, and relay information is sent along with the data. This routing technique can provide the best route selection when there are obstacles or constraints in the network [57].

Hop by Hop reactive routing protocols are summarized as below and the comparison of these protocols with their advantages and disadvantages are shown in Table III.

Ad-hoc On-demand Distance Vector (AODV) calls for only the necessary nodes to establish a direct route when a data packet needs to be routed from a source node to a destination node. Therefore, it effectively utilizes the resources of the nodes in the network and increases network efficiency [56].

Enhanced AODV (ENAODV) is an enhanced version of the AODV protocol, designed to achieve high performance, low latency and high efficiency. To achieve these objectives, ENAODV considers the distance between nodes and constructs the shortest path [58].

Medium Access Control-AODV (MAC-AODV) is a combination of protocols designed for both the MAC layer

and the routing layer. It is used to control data transmission and routing between the nodes in the network. MAC-AODV aims for low latency and high efficiency, making it particularly popular in low power consumption devices such as smart devices and smart sensors [59].

Compatibility Based Vehicular Ad-Hoc Networks-AODV (CV-AODV) is particularly used in autonomous vehicles, intelligent transportation systems and traffic routing applications. It is known for providing high efficiency, rapid data transmission and secure data communication [60].

Fitness Based AODV (FBAODV) is a protocol that constructs routes based on the physical states of the nodes in the network. It is designed to select the most suitable routing paths by considering the location, speed, energy level and physical state of the nodes [61].

Source routing reactive routing protocols are summarized as below and the comparison of these protocols with their advantages and disadvantages are shown in Table IV.

TABLE I
THE COMPARISON OF DISTANCE VECTOR BASED PROACTIVE ROUTING PROTOCOLS

Protocol	Network Size	Routing Table Updating Strategy and/or Frequency	Advantages	Disadvantages
DVRP[35]	Medium-Large	• updating when an event occurs (when a new node is discovered or lost) • updating with periodic time intervals	• high reliability • low latency	• high update traffic • high memory usage
DSDV[36]	Small-Medium	• updating when an event occurs (when a new node is discovered or lost) • updating with periodic time intervals	• low complexity • simple structure	• high update traffic • low mobility support
Randomized DSDV[37]	Small-Medium	• updating with random time intervals	• low update traffic • low complexity	• random update interval • low mobility support
Multi-Agent DSDV[38]	Medium-Large	• using multiple agents • updating when there is a change in the routing tables of the nodes • updating with periodic time intervals	• high scalability • local update	• agent coordination requirement • complex structure
DSDV- Dual Interface Multiple Channels[39]	Small-Medium	• using dual interface and multi-channel • updating when there is a change in the routing tables of the nodes • using active and passive routing tables	• channel variety • low latency	• high hardware requirement • complex structure
Improved-DSDV[40]	Medium-Large	• updating a limiting number of nodes • updating when there is a change in the routing tables of the nodes • updating with periodic time intervals	• low update traffic • low complexity	• additional communication cost
Divide Cluster DSDV[41]	Medium-Large	• dividing network into clusters • updating with periodic time intervals for the leader nodes • updating when there is a change in the routing tables of the other nodes	• scalability • local update	• management problems and difficulties in subnets
Fisheye Zone Routing Protocol[42]	Medium-Large	• updating remote zones using the "Fisheye" routing table • updating when there is a change in the routing tables of the nodes in the relevant domain • updating with periodic time intervals for all nodes	• low update traffic • rapid update of topology information	• high memory usage • routing complexity • high latency
BABEL[43][44][45]	Medium-Large	• updating based on the link states of the nodes • updating when there is a change in the routing tables of the neighbouring nodes • updating with periodic time intervals for all nodes • updating all affected nodes when a node in the network is disconnected from its neighbour	• high scalability • low memory usage	• complex routing policies • high update traffic

TABLE II
THE COMPARISON OF LINK STATE BASED PROACTIVE ROUTING PROTOCOLS

Protocol	Network Size	Routing Table Updating Strategy and/or Frequency	Advantages	Disadvantages
OLSR[46][47][48]	Small-Medium	• periodic topology control messages • allowing nodes to recognize each other and notify topology changes	• high efficiency • low latency	• complex structure
Street Centric QoS-OLSR[49]	Small-Medium	• periodic topology control messages • considering QoS requirements	• street centric QoS Support	• complex structure
Cluster Based QoS-OLSR[50]	Medium-Large	• periodic topology control messages • considering QoS requirements • primarily updating by choosing the nodes within cluster	• high scalability • low network traffic	• complex structure
LD-OLSR[51]	Small-Medium	• periodic topology control messages • considering QoS requirements • using high-speed connection if available	• routing with QoS information	• complex structure
DS-OLSR [52]	Small-Medium-Large	• periodic topology control messages • using emergency signal in case of disaster • considering the integrity of the network, energy efficiency and communication quality	• high durability • high scalability	• low efficiency • high latency

| FSR[53] | Small-Medium | • dynamic update when the link states change
• Frequent updating for closer nodes, infrequent updating for distant nodes | • low routing table size
• low network traffic | • high latency
• low routing table update frequency |
| B.A.T.M.A.N V [54][55] | Small-Medium | • dynamic update when needing to find a new route
• monitoring the link state of neighbours | • low configuration requirement
• low latency | • high routing table size |

TABLE III
THE COMPARISON OF HOP BY HOP REACTIVE ROUTING PROTOCOLS

Protocol	Network Size	Routing Table Updating Strategy and/or Frequency	Advantages	Disadvantages
AODV[56]	Large	• updating depends on the distance between the nodes and the energy level of the nodes • on-demand approach	• fast forwarding • efficient updating	• performance can be poor at high node density
ENAODV[58]	Large	• updating depends on the distance between the nodes • on-demand approach	• efficient updating • high network performance	• complex structure • requires much computation
MAC-AODV[59]	Small	• minimizing routing table • on-demand approach. • time-basis updating	• low energy consumption • low delay	• low scalability • low flexibility
CV-AODV[60]	Large	• updating depends on the availability between the nodes • on-demand approach	• effective routing • high network performance	• low scalability • low flexibility
FBAODV[61]	Small	• using genetic algorithm and simulated annealing for updating • on-demand approach	• low energy consumption • low delay	• low scalability • low flexibility

Reliable DSR (R-DSR) is a derivative of DSR protocol that aims to mitigate packet loss in wireless ad hoc networks. R-DSR operates similarly to DSR by utilizing routing packets containing route information. However, unlike DSR, the RDSR protocol incorporates a reliability mechanism that allows for retransmission the lost packet in event of packet losses [63].

Zone Based DSR (Z-DSR) updates the routing table by utilizing the concept of "zones" that are created based on the geographical locations of nodes in the network. The ZDSR protocol considers distances and geographic locations between nodes to gain a better understanding of the network topology. This approach reduces the number of required routing packets between nodes, resulting in reduced network traffic. Consequently, the performance of the network increases as congestion decreases [64].

Segment Based DSR (S-DSR) proposes a similar approach to the DSR protocol. However, S-DSR enhances data integrity by dividing data into segments rather instead of packaging it as a whole. This approach reduces the risk of data loss within the network and ensures higher data reliability [64].

C. Hybrid Solutions

Hybrid protocols are designed to minimize control overhead in proactive solutions and reduce the delay in searching for an initial path in reactive approaches [65]. The objective of hybrid routing protocols is to efficiently manage data traffic by combining several proactive and reactive routing algorithms and leveraging the strengths of both approaches [66]. Table V provides a comparison of these protocols, including their advantages and disadvantages.

Zone Routing Protocol (ZRP) divides the network into zones, where each node within a zone is informed about its neighbors using proactive routing. However, when it comes to routing data to remote nodes, zones employ reactive routing. ZRP utilizes the redundant path feature, allowing for the utilization of multiple paths in the network. However, these redundant paths can lead to increased network traffic, resulting in high latency [67].

Temporally-Ordered Routing Algorithm (TORA) creates a hierarchical routing tree from the nodes in the network using proactive routing. Reactive routing is employed for delivering the data to destination node. TORA is known for its effectiveness in reducing latency and improving network performance. However, as the number of nodes increases, updating the tables in TORA requires high bandwidth usage [68].

Hybrid Ad Hoc Routing Protocol (HARP) aims to enhance the performance of the network by combining proactive and reactive routing approaches. HARP maintains a proactive routing table where information about the overall topology of the network is stored. Reactive routing is employed for data transmission to destination node. While HARP boasts low energy consumption, the routing tables of all nodes need to be updated as the network topology changes [69].

TAD-Hoc/TROPHY is a routing protocol that leverages the location information of the nodes to determine the optimal routing path. By considering the battery states of the nodes, TROPHY aims to achieve energy efficiency and low power consumption The protocol utilizes reactive routing features to update the routing tables, with updates occurring only when the destination node is first accessed [70].

An Optimized Ad-Hoc On-Demand Multipath Distance Vector (AOMDV) is a hybrid protocol that combines the features of AODV and DSDV protocols. AOMDV uses an on-demand routing strategy along with the capability to establish multiple paths. This enables faster and more reliable transmission of data packets to destination nodes. AOMDV has been used as a foundation in various studies within the literature [71].

Secure and Efficient-AOMDV (SE-AOMDV) is a protocol specifically designed to enhance both the security and efficiency of the AOMDV protocol. It retains the hybrid structure of AOMDV while introducing additional features to

ensure secure data transmission. These features include secure route selection, data packet encryption, authentication, and message integrity mechanisms. While these enhancements improve the security of the protocol, they also come with the drawback of increased processing and storage costs due to the additional computational and storage requirements associated with the security mechanisms [72].

AOMDV- Fitness Function (FFn) introduces a fitness function to AOMDV protocol. The fitness function is designed to ensure communication security, to reduce energy consumption and to prevent packet losses by using node specific characteristics, including battery level and hop count. AOMDV-FFn employs a selective interfacing method to improve communication quality. However, AOMDV-FFn requires high processor power and memory usage [73].

AOMDV- Genetic Algorithm (GA) incorporates a genetic algorithm to elect the best route. The goal of this protocol is to minimize energy consumption, prevent packet losses and increase data transmission speed. However, utilization of a genetic algorithm comes at the cost of high processing power and memory usage [73].

TABLE IV
THE COMPARISON OF SOURCE REACTIVE ROUTING PROTOCOLS

Protocol	Network Size	Routing Table Updating Strategy and/or Frequency	Advantages	Disadvantages
DSR[57]	Small-Medium	• dynamic update for every packet	• low packet delay • high routing performance	• high network traffic • high network delay
R-DSR[63]	Medium	• dynamic update for every packet • using both regular and quick updating	• high reliability • low packet loss	• high network traffic • low flexibility
Z-DSR[64]	Large	• periodic update • using geographic regions and zone tables • using active and passive routing	• high forwarding performance • low network traffic	• complex structure • difficulty in determining node boundaries
S-DSR[64]	Large	• dynamic update for every packet • using segment table • using active and passive routing	• high routing performance • low network traffic	• complex structure • high computing power

TABLE V
THE COMPARISON OF HYBRID ROUTING PROTOCOLS

Protocol	Network Size	Routing Table Updating Strategy and/or Frequency	Advantages	Disadvantages
ZRP[67]	Medium-Large	• regular updating • using both proactive and reactive routing approaches	• several path usage	• high network traffic • high network delay
TORA[68]	Small-Medium	• regular updating • using proactive approach	• high network performance • low delay	• high bandwidth usage • high computation cost
HARP[69]	Medium-Large	• regular updating • using both proactive and reactive routing approaches	• low energy consumption	• reforming when topology changes
TROPHY(TAD-Hoc)[70]	Medium-Large	• updating when first access • using reactive approach	• Low bandwidth usage • low delay	• low scalability • low flexibility
AOMDV[71]	Medium-Large	• regular updating • using both proactive and reactive routing approaches	• low energy consumption • low delay • high reliability	• reforming when topology changes
SE-AOMDV[72]	Medium-Large	• updating when first access • using reactive approach	• high communication quality	• high processing power • high memory usage
AOMDV-FFn[73]	Medium-Large	• regular updating • using both proactive and reactive routing approaches	• high network performance	• high processing power • high memory usage
AOMDV-GA[73]	Medium-Large	• updating when first access • using reactive approach	• low energy consumption • low delay • low packet loss	• high processing power • high memory usage

V. DISCUSSION

In the previous section, we conducted a detailed evaluation of individual protocols within each group. In this section, we will provide general evaluations for each group of protocols. Then we will emphasize open research areas.

When evaluating distance vector based proactive routing protocols, it is evident that these protocols can effectively scaled to network of various sizes, ranging from small to large. The update strategy of the routing tables in these protocols is typically based on event-driven updates (such as the discovery or loss of a node) or periodic updates that occur independently of any specific event. While these protocols ensure that the routing tables remain up-to-date, they do have certain drawbacks, including high memory usage and increased traffic for updating the tables.

The link state approach is commonly used in smaller and medium-sized networks, although the DS-OLSR protocol can be scaled to networks of any size. Protocols such as OLSR and modified OSLR update their routing tables periodically, while protocols like FSR and B.A.T.M.A.N-V dynamically update their tables whenever there are changes in link status or route discovery. One advantage of link state proactive routing protocols is that scalability and QoS support can be added as additional features. However, these protocols also have some drawbacks, including complex configuration difficulties, high

latency, and large routing tables. As a result, they are generally not preferred for large-scale networks.

Hop-by-hop based reactive protocols are suitable for networks of any size. In this group, routing tables are updated as needed, resulting in efficient routing table updates and low energy consumption. As a result, these protocols can be preferred in large-scale networks. However, configuring the network and addressing scalability issues, especially in modified versions like increasing reliability in the ADOV protocol, can pose challenges that need to be addressed.

Source-based reactive protocols can be utilized in networks of any size, although they particularly suitable for large networks. Update tables are dynamically performed each time a packet is transmitted. The limitations of these protocols are that they can cause high processing power and high network traffic.

Hybrid protocols offer the flexibility to create application-specific solutions for networks of any size since they utilize both proactive and reactive approaches. Some protocols update their routing tables periodically, while others update them only upon the first transmission of a data packet. Additionally, there are approaches that employ both regular and on-demand updating. While the combination of proactive and reactive routing approaches brings about several advantages, it is crucial to address the issues of high processing power and memory usage associated with these protocols.

The evaluations have indicated that the following areas present opportunities for further research.

1. QoS support: Irrespective of the protocol type, it is essential to provide QoS support that caters to the specific requirements of the application, as well as ensuring independent QoS provision.

2. Scalability: The scalability of protocols poses challenges in large networks, as those designed for efficient operation in small networks often encounter issues such as memory shortage and increased network load. To address this, it is crucial to focus on resolving reliability and efficiency issues to prevent system degradation and ensure optimal performance in larger network environments.

3. Data losses: Efforts should be made to tackle the issue of routing table loss in nodes caused by frequent disconnections due to high mobility, especially in scenarios where long-term data storage is not available. Finding solutions to mitigate this problem is crucial to maintain efficient and reliable routing in such dynamic environments.

In order to mitigate the potential threat to the validity of our study, we have implemented proactive measures in relation to the selection of digital libraries and search terms. To overcome this concern, we utilized four esteemed digital libraries in the field of computer science. These libraries offered a wealth of resources and diverse search query structures, enabling us to establish precise search terms. By incorporating these robust digital libraries, our intention was to enhance the accuracy and comprehensiveness of our research while minimizing biases associated with the selection of libraries or formulation of search terms.

VI. Conclusion

VANET, a network designed for wireless communication between vehicles, is utilized for various purposes such as enhancing traffic safety, managing emergencies and increasing passenger comfort through communication between vehicles and vehicle-roadside units.

In this study, we have conducted a comparative analysis of topology-based routing protocols in VANETs. These protocols utilize the vehicle locations to automatically determine the network topology and ensure efficient data transmission through the shortest path. As a result, topology-based protocols play a crucial role in achieving reliable and efficient routing in VANETs.

We classified the topology-based routing protocols in VANETs and arranged them in chronological order from basic to the most up-to-date, considering their acceptance by academic communities. Comparative feature tables were created to showcase their usage characteristics. Detailed analyses and comments were conducted for both individual protocols and the overall class.

The following results can be obtained from a general perspective for topology-based routing approaches:

• Distance vector based proactive routing approaches provides up-to-date routing tables but suffer from high memory usage and high data traffic for updating the tables.

• Link-state based proactive approaches can support QoS. However, they have drawbacks of complex network configuration difficulties, high latency, and large size of routing tables.

• Although hop by hop reactive protocols, offer efficient routing table updating and low energy consumption, they have difficulty in configuring the network and scalability issues.

• Source routing based reactive approaches have the ability to address any network size. However, they can cause high processing power and high network traffic.

• Although hybrid approaches combine the beneficial aspects of proactive and reactive protocols, they may require additional processing power and memory usage to handle the integration of both approaches effectively.

References

[1] Giridhar, K., Anbuananth, C., & Krishnaraj, N. (2023). Energy efficient clustering with Heuristic optimization based Routing protocol for VANETs. Measurement: Sensors, 100745.

[2] Diaa, M. K., Mohamed, I. S., & Hassan, M. A. (2023). OPBRP-obstacle prediction based routing protocol in VANETs. Ain Shams Engineering Journal, 14(7), 101989.

[3] Al-Sultan, S., Al-Doori, M. M., Al-Bayatti, A. H., & Zedan, H. (2014). A comprehensive survey on vehicular ad hoc network. Journal of network and computer applications, 37, 380-392.

[4] Wahid, I., Ikram, A. A., Ahmad, M., Ali, S., & Ali, A. (2018). State of the art routing protocols in VANETs: A review. Procedia computer science, 130, 689-694.

[5] S. Yogarayan, S. F. A. Razak, A. Azman, M. F. A. Abdullah, S. Z. Ibrahim and K. J. Raman, "A Review of Routing Protocols for Vehicular Ad-Hoc Networks (VANETs)," 2020 8th International Conference on Information and Communication Technology (ICoICT), Yogyakarta, Indonesia, 2020, pp. 1-7, doi: 10.1109/ICoICT49345.2020.9166174.

[6] Ali, T. E., al Dulaimi, L. A. K., & Majeed, Y. E. (2016, May). Review and performance comparison of vanet protocols: Aodv, dsr, olsr, dymo, dsdv & zrp. In 2016 Al-Sadeq International Conference on

Multidisciplinary in IT and Communication Science and Applications (AIC-MITCSA) (pp. 1-6). IEEE

[7] Srivastava, A., Prakash, A., & Tripathi, R. (2020). Location based routing protocols in VANET: Issues and existing solutions. Vehicular Communications, 23, 100231.

[8] Boussoufa-Lahlah, S., Semchedine, F., & Bouallouche-Medjkoune, L. (2018). Geographic routing protocols for Vehicular Ad hoc NETworks (VANETs): A survey. Vehicular Communications, 11, 20-31.

[9] Ullah, A., Yao, X., Shaheen, S., & Ning, H. (2019). Advances in position based routing towards ITS enabled FoG-oriented VANET–A survey. IEEE Transactions on Intelligent Transportation Systems, 21(2), 828-840.

[10] Bilal, S. M., Khan, A. U. R., & Ali, S. (2017). Review and performance analysis of position based routing in VANETs. Wireless Personal Communications, 94, 559-578.

[11] Abbasi, I. A., & Shahid Khan, A. (2018). A review of vehicle to vehicle communication protocols for VANETs in the urban environment. future internet, 10(2), 14.

[12] Srivastava, A., Bagga, N., & Rakhra, M. (2021, September). Analysis of Cluster-Based and Position-based Routing Protocol in VANET. In 2021 9th International Conference on Reliability, Infocom Technologies and Optimization (Trends and Future Directions)(ICRITO) (pp. 1-5). IEEE.

[13] Shah, P., & Kasbe, T. (2021). A review on specification evaluation of broadcasting routing protocols in VANET. Computer Science Review, 41, 100418.

[14] Mchergui, A., Moulahi, T., Alaya, B., & Nasri, S. (2017). A survey and comparative study of QoS aware broadcasting techniques in VANET. Telecommunication Systems, 66, 253-281.

[15] Abuashour, A., & Kadoch, M. (2017). Performance improvement of cluster-based routing protocol in VANET. Ieee access, 5, 15354-15371.

[16] Katiyar, A., Singh, D., & Yadav, R. S. (2020). State-of-the-art approach to clustering protocols in VANET: a survey. Wireless Networks, 26, 5307-5336.

[17] Mokhayeri, S., & Kheirabadi, M. T. (2021). Zone selection strategy in Geocast routing algorithms in VANET: a review. The Journal of Supercomputing, 77(11), 12953-12986.

[18] Garrosi, M. T., Kalac, M., & Lorenzen, T. (2017, January). Geo-routing in urban Vehicular Ad-hoc Networks: A literature review. In 2017 International Conference on Computing, Networking and Communications (ICNC) (pp. 865-871). IEEE.

[19] Wheeb, A. H., Nordin, R., Samah, A. A., Alsharif, M. H., & Khan, M. A. (2021). Topology-based routing protocols and mobility models for flying ad hoc networks: A contemporary review and future research directions. Drones, 6(1), 9.

[20] Sehrawat, P., & Chawla, M. (2022). Interpretation and Investigations of Topology Based Routing Protocols Applied in Dynamic System of VANET. Wireless Personal Communications, 1-27.

[21] Shrivastava, P. K., & Vishwamitra, L. K. (2021). Comparative analysis of proactive and reactive routing protocols in VANET environment. Measurement: Sensors, 16, 100051.

[22] Belamri, F., Boulfekhar, S., & Aissani, D. (2021). A survey on QoS routing protocols in Vehicular Ad Hoc Network (VANET). Telecommunication Systems, 78(1), 117-153.

[23] Gawas, M. A., & Govekar, S. (2021). State-of-art and open issues of cross-layer design and QOS routing in internet of vehicles. Wireless Personal Communications, 116(3), 2261-2297.

[24] Dybå, T., Kitchenham, B.A., and Jørgensen, M. (2005). "Evidence-based Software Engineering for Practitioners," IEEE Software, 22(1): 58–65.

[25] Kitchenham, B., & Charters, S. (2007). Guidelines for performing systematic literature reviews in software engineering.

[26] Zhang, H., & Babar, M. A. (2013). Systematic reviews in software engineering: An empirical investigation. Information and Software Technology, 55(7), 1341-1354.

[27] Akdur, D., & Demirörs, O. (2020). Systematic Reviews in Model-Driven Engineering: A Tertiary Study. Journal of Aeronautics and Space Technologies, 13(1), 57-68.

[28] Akdur, D., Garousi, V., & Demirörs, O. (2018). A survey on modeling and model-driven engineering practices in the embedded software industry. Journal of Systems Architecture, 91, 62-82.

[29] Gurbuz, H. G., & Tekinerdogan, B. (2018). Model-based testing for software safety: a systematic mapping study. Software Quality Journal, 26, 1327-1372.

[30] Sharma, S., & Sharma, P. (2019, May). Comprehensive study of various routing protocols in VANET. In 2019 International Conference on Intelligent Computing and Control Systems (ICCS) (pp. 1272-1275). IEEE.

[31] Waitzman, D., Partridge, C., & Deering, S. E. (1988). RFC1075: Distance Vector Multicast Routing Protocol.

[32] Clausen, T., & Jacquet, P. (2003). Optimized link state routing protocol (OLSR) (No. rfc3626).

[33] Hedrick, C. (1988). Routing Information Protocol (RFC 1058). Network Working Group, IETF. http://www. ietf. org/rfc/rfc1058. txt.

[34] Bao, L., & Garcia-Luna-Aceves, J. J. (1999, October). Link-state routing in networks with unidirectional links. In Proceedings Eight International Conference on Computer Communications and Networks (Cat. No. 99EX370) (pp. 358-363). IEEE.

[35] Chen, G., Lau, F. C., Du, P., & Xie, L. (2001, September). Distance-vector routing protocols for networks with unidirectional link. In Proceedings International Conference on Parallel Processing Workshops (pp. 473-478). IEEE.

[36] Perkins, C. E., & Bhagwat, P. (1994). Highly dynamic destination-sequenced distance-vector routing (DSDV) for mobile computers. ACM SIGCOMM computer communication review, 24(4), 234-244.

[37] Boukerche, A., & Das, S. K. (2003). Congestion control performance of R-DSDV protocol in multihop wireless ad hoc networks. Wireless Networks, 9, 261-270.

[38] Harrabi, S., Chainbi, W., & Ghedira, K. (2014, June). A multi-agent proactive routing protocol for Vehicular Ad-Hoc Networks. In The 2014 International Symposium on Networks, Computers and Communications (pp. 1-6). IEEE.

[39] Luong, T. T., Lee, B. S., & Yeo, C. K. (2009, October). Dual-Interface Multiple Channels DSDV Protocol. In 2009 IEEE International Conference on Wireless and Mobile Computing, Networking and Communications (pp. 104-109). IEEE.

[40] Ahmed, G. F., Barskar, R., & Barskar, N. (2012). An improved DSDV routing protocol for wireless ad hoc networks. Procedia Technology, 6, 822-831.

[41] Fengjie, Y., Hui, Y., & Ying, Z. (2018, June). Research on DSDV routing protocol based on wireless Mesh network. In 2018 Chinese Control And Decision Conference (CCDC) (pp. 4292-4297). IEEE.

[42] Yang, C. C., & Tseng, L. P. (2005, January). Fisheye zone routing protocol for mobile ad hoc networks. In Second IEEE Consumer Communications and Networking Conference, 2005. CCNC. 2005 (pp. 1-6). IEEE.

[43] Chroboczek, J. (2011). The babel routing protocol (No. rfc6126).

[44] Chroboczek, J. (2015). Extension Mechanism for the Babel Routing Protocol (No. rfc7557).

[45] Chroboczek, J. & Schinazi D. (2021). The babel routing protocol (No. Rfc8966).

[46] Jacquet, P., Muhlethaler, P., Clausen, T., Laouiti, A., Qayyum, A., & Viennot, L. (2001, December). Optimized link state routing protocol for ad hoc networks. In Proceedings. IEEE International Multi Topic Conference, 2001. IEEE INMIC 2001. Technology for the 21st Century. (pp. 62-68). IEEE.

[47] Clausen, T., & Jacquet, P. (2003). Optimized link state routing protocol (OLSR) (No. rfc3626).

[48] Clausen, T., Dearlove, C., Jacquet, P., & Herberg, U. (2014). The optimized link state routing protocol version 2 (No. rfc7181).

[49] Kadadha, M., Otrok, H., Barada, H., Al-Qutayri, M., & Al-Hammadi, Y. (2017, June). A street-centric QoS-OLSR protocol for urban vehicular ad hoc networks. In 2017 13th International Wireless Communications and Mobile Computing Conference (IWCMC) (pp. 1477-1482). IEEE.

[50] Kadadha, M., Otrok, H., Barada, H., Al-Qutayri, M., & Al-Hammadi, Y. (2018, June). A cluster-based QoS-OLSR protocol for urban vehicular ad hoc networks. In 2018 14th International Wireless Communications & Mobile Computing Conference (IWCMC) (pp. 554-559). IEEE.

[51] Jain, R., & Kashyap, I. (2019). An QoS aware link defined OLSR (LD-OLSR) routing protocol for MANETs. Wireless Personal Communications, 108(3), 1745-1758.

[52] Aliyu, U., Takruri, H., Hope, M., & Halilu, A. G. (2020, July). DS-OLSR–Disaster Scenario Optimized Link State Routing Protocol. In 2020 12th International Symposium on Communication Systems, Networks and Digital Signal Processing (CSNDSP) (pp. 1-6). IEEE.

[53] Pei, G., Gerla, M., & Chen, T. W. (2000, June). Fisheye state routing: A routing scheme for ad hoc wireless networks. In 2000 IEEE International Conference on Communications. ICC 2000. Global Convergence

Through Communications. Conference Record (Vol. 1, pp. 70-74). IEEE.

[54] Neumann, A., Aichele, C., Lindner, M., & Wunderlich, S. (2008). Better approach to mobile ad-hoc networking (BATMAN). IETF draft, 1-24.

[55] Sliwa, B., Falten, S., & Wietfeld, C. (2019, April). Performance evaluation and optimization of batman v routing for aerial and ground-based mobile ad-hoc networks. In 2019 IEEE 89th Vehicular Technology Conference (VTC2019-Spring) (pp. 1-7). IEEE.

[56] Perkins, C., Belding-Royer, E., & Das, S. (2003). RFC3561: Ad hoc on-demand distance vector (AODV) routing.

[57] Johnson, D. B., & Maltz, D. A. (1996). Dynamic source routing in ad hoc wireless networks. Mobile computing, 153-181.

[58] Singhal, P. K., & Chaubey, V. K. (2019, February). An Enhanced Ad Hoc on Demand Distance Vector Routing Protocol for Vehicular Ad Hoc Networks (VANET's). In Proceedings of International Conference on Sustainable Computing in Science, Technology and Management (SUSCOM), Amity University Rajasthan, Jaipur-India.

[59] Adil, M., Khan, R., Almaiah, M. A., Al-Zahrani, M., Zakarya, M., Amjad, M. S., & Ahmed, R. (2020). MAC-AODV based mutual authentication scheme for constraint oriented networks. IEEE Access, 8, 44459-44469.

[60] Kumbhar, F. H., & Shin, S. Y. (2020, October). CV-AODV: Compatibility Based Vehicular Ad-hoc On Demand Distance Vector Routing Protocol. In 2020 International Conference on Information and Communication Technology Convergence (ICTC) (pp. 1004-1008). IEEE.

[61] Suganthi, B., & Ramamoorthy, P. (2020). An advanced fitness based routing protocol for improving QoS in VANET. Wireless Personal Communications, 114, 241-263.

[62] Khudayer, B. H., Anbar, M., Hanshi, S. M., & Wan, T. C. (2020). Efficient route discovery and link failure detection mechanisms for source routing protocol in mobile ad-hoc networks. IEEE Access, 8, 24019-24032.

[63] Al-Shora, M. A., Nouh, S. A., & Khalifa, A. R. (2018). Reliable dynamic source routing (RDSR) protocol with link failure prediction for mobile ad hoc networks (MANET). Journal of Network Communications and Emerging Technologies (JNCET) www. jncet. org, 8(3).

[64] Malwe, S. R., Taneja, N., & Biswas, G. P. (2017). Enhancement of DSR and AODV protocols using link availability prediction. Wireless Personal Communications, 97, 4451-4466.

[65] Hayat, S., Liu, X., Li, Y., & Zhou, Y. (2019, May). Comparative analysis of vanet's routing protocol classes: an overview of existing routing protocol classes and futuristic challenges. In 2019 IEEE 2nd International Conference on Electronics Technology (ICET) (pp. 1-7). IEEE.

[66] Kuppusamy, P., Thirunavukkarasu, K., & Kalaavathi, B. (2011, April). A study and comparison of OLSR, AODV and TORA routing protocols in ad hoc networks. In 2011 3rd International Conference on Electronics Computer Technology (Vol. 5, pp. 143-147). IEEE.

[67] Haas, Z. (1998). ZRP: The Zone Routing Protocol (ZRP) for Ad Hoc Networks. Internet Draft draft-zone-routing-protocol-01. txt.

[68] Park, V. D., & Corson, M. S. (1997, April). A highly adaptive distributed routing algorithm for mobile wireless networks. In Proceedings of INFOCOM'97 (Vol. 3, pp. 1405-1413). IEEE.

[69] Nikaein, N., Bonnet, C., & Nikaein, N. (2001, September). Harp-hybrid ad hoc routing protocol. In Proceedings of international symposium on telecommunications (IST) (pp. 56-67).

[70] Ragavi, B., Saranya, S., Pavithra, L., Mohanapriya, G. K., Nandhini, V. D., & Dhaarani, S. (2021, February). TROPHY: Efficient greedy transmission of data packet using trophy (TAD-hoc) protocol along with greedy algorithm in VANET. In 2021 Third International Conference on Intelligent Communication Technologies and Virtual Mobile Networks (ICICV) (pp. 23-28). IEEE.

[71] Yuan, Y., Chen, H., & Jia, M. (2005, October). An optimized ad-hoc on-demand multipath distance vector (AOMDV) routing protocol. In 2005 Asia-Pacific Conference on Communications (pp. 569-573). IEEE.

[72] Meddeb Makhlouf, A., & Guizani, M. (2019). SE-AOMDV: Secure and efficient AOMDV routing protocol for vehicular communications. International Journal of Information Security, 18(5), 665-676.

[73] Bhardwaj, A., & El-Ocla, H. (2020). Multipath routing protocol using genetic algorithm in mobile ad hoc networks. IEEE Access, 8, 177534-177548.

BIOGRAPHIES

ALİ FUAT GÜNEŞ Konya, Turkey, in 1987. He received his first B.S. in computer systems teaching department from Selçuk University, Konya, in 2009 and his second B.S. in computer engineering department from Iskenderun Technical University, Hatay, in 2022. He is currently a graduate student in computer engineering department of Iskenderun Technical University.

His research interests are routing approaches and QoS support in vehicular ad hoc networks.

İPEK ABASIKELEŞ-TURGUT Adana, Turkey, in 1985. She received the B.S. degree in computer engineering from the University of Istanbul Technical University (ITU), Istanbul, in 2007 and the M.S. and Ph.D. degrees in computer, electrical and electronics engineering from Çukurova University, Adana, in 2009 and 2013, respectively.

From 2013 to 2022, she was an Assistant Professor in Iskenderun Technical University with the Computer Engineering Department, Hatay. Since 2023, she has been an Associate Professor in the same department. She is the author of 17 Journal and 27 Conference papers. Her research interests include wireless ad hoc and sensor networks, routing algorithms, security issues in wireless sensor networks, distributed multiprocessor architectures, simulation and modelling.

Wearable Thimble-like Device for the Objective Follow-up and Therapy of Multiple Sclerosis

Elif Hocaoglu

Abstract—**Multiple sclerosis (MS) is an autoimmune disease that affects more than 1 million people worldwide. Since there is no definitive treatment for the disease, the treatment plan for each patient should be updated regularly according to the current level of the disease. There are standard clinical tests, each with its own scoring scale, used to monitor the deterioration of upper and lower extremity functions of MS patients under the supervision of a neurologist and physiotherapist. However, non-objective scoring based on the opinion of the physiotherapist is open to erroneous assessments and may vary from person to person. In addition, clinical tests do not provide detailed information about the functional impairment of the patient. Unfortunately, an objective evaluation system has not yet been implemented all over the world, and the treatment plan is still determined according to the disease in neurological-based disabilities, such as MS, which is of vital importance. personal assessment. To address the aforementioned problem, the design and experimental evaluation of a wearable thimble-like device that can be substituted for the standard clinical tests to assess the follow-up of MS are presented. The device provides the measurement of high sensitivity and opportunity for objective assessment and allows patients of all ages to use it in any desired place during their treatment phase.**

Index Terms—**Multiple sclerosis, wearable device, sensor fusion, objective assessment, rehabilitation.**

I. INTRODUCTION

MULTIPLE Sclerosis (MS) is a complex, progressive, immune-mediated disease and is the third most common disease-causing neurological disability, occurring in approximately 400,000 people in the United States and approximately 2.3 million worldwide [1]. MS disease, which is mostly diagnosed between the ages of 20-50, brings with it emotional and economic problems as well as physical limitations [2]. It is among the basic needs of MS patients to develop effective solutions for this disease, which is frequently encountered in the world and includes financial and moral difficulties, and to be able to follow the patient with high sensitivity.

As well as the treatment process of MS disease, the follow-up of the disease is of great importance in terms of observing the effect of the methods applied for treatment and the instantaneous condition of the physical functions of the patient. With the help of various clinical tests, (the 9-hole peg test [3], the expanded disability status scale [4], The Jebsen-Taylor hand function test (JTHFT) [5]) performed under the supervision of physiotherapists, the regression or

Elif Hocaoglu is with the Department of Electrical and Electronics Engineering, School of Engineering and Natural Sciences, Istanbul Medipol University, and Research Institute for Health Sciences and Technologies (SABITA), Istanbul Medipol University, Istanbul, 34810 TURKIYE e-mail: ehocaoglu@medipol.edu.tr

improvement in the physical functions (hand functions) of MS patients can be observed. The general approach in the evaluation of sensory-motor control is that physiotherapists, using the measurement values of the specified standard clinical tests, evaluate and score the patient's movements according to the defined tasks. Besides, the expanded disability status scale (EDSS) is still actively used for various reasons such as determining the medical treatment method and deciding on the rehabilitation process and the tools to be used in this process. The disadvantage of this measurement method is that it only evaluates the patient's gait disturbance and ignores the problems in the upper extremity [6]. Because these methods also depend on the personal perspective of the physiotherapist, assessments are inaccurate and cannot provide sufficient detail about the patient's disability level. However, objective, accurate and precise assessments provide valuable information in the evaluation of the sensory-motor abilities of the patient and allow for accurate treatment planning.

Considering the need for objective and instantaneous assessment of MS patients, accurate diagnostic techniques based on sensor data have been recommended by some research groups instead of individual evaluations [7]. An isometric force/torque measurement, in which subjects interact with a sensing system that measures the force/torque was proposed as an objective method applied in the static condition of the patients. In such studies, The expectation from the individual during the experiment is applying force to the experimental setup fixed to the planar surface by using their hands, and the normal and tangential force coordination of the individual was investigated based on the force sensor placed on the device [8]–[10]. In particular, the isometric strength of the patients was also evaluated utilizing a simple isometric setup considering the six different directions of exerted force/torque [11]. However, studies so far have been limited in terms of the analysis of the force exerted by the patient only, the grabbed object endowed with load cells fixed on a plane. In addition, the weight of the employed object during the experiment is always constant and the task assigned to the individual throughout the experiment is the same [8]–[10], [12], [13].

Recent studies show that various functional attributes of hand function can be evaluated through kinetic analysis. In particular, the manipulation of stationary and moving objects requires a proper contribution of two particular force components, grip force and load force which are perpendicular and parallel to the surface of the grabbed object, respectively. The findings in the literature validate the importance of the coordination between grip and load force applied to the object without any time delay between them. Moreover, the high coordination between these particular forces takes

place based on the anticipatory neural control mechanism of humans [14], [15]. In other words, disharmony between such forces results in the slippage of the object. However, although the detection of incipient slippage time of an object is an important indication of the moment when the exerted improper force, these studies did not consider the force level that causes the slippage. In the study of Krishnan's group, static and dynamic tasks were defined for the individual by using two plates fixed on the table where the force sensor was placed, and the ratio of normal and tangential forces applied to the plate by the individual compared to a healthy individual was compared [16], [17]. They argued that the tasks in dynamic conditions contribute to alleviating the problems of patients with MS by reconstructing the nerve connection corrupted by the degenerated nerve. Different from the study above, a cylindrical assessment device endowed with an accelerometer and load cell was proposed to compare the amount of force applied by the patients to move the object from one point to another [12]. The device provides an advantage in terms of testing the coordination of force components applied by the patient under dynamic conditions. On the other hand, there are a number of restrictions on performing tests of different difficulties. For example, holding an object with different weights, and needing more manipulation skills to successfully grasp different thicknesses of objects were not involved in the study. In other words, the study was limited by the single configuration of the fingers/ hand to hold a single object with the same load and geometry. Given the sensorimotor ability of the patients was evaluated based on the isometric strength in literature, the assessment of the patients was not properly performed as the assigned tasks and the used devices do not allow to reveal the important features of their movement. Thus, the clinical evaluation of the aforementioned studies does not allow the individual to fully express his/her daily life functions usually carried out under dynamic conditions. Furthermore, the determined metrics are insufficient to evaluate the success of the individual in performing such activities.

In this study, the proposed wearable robotic device addresses all the limitations mentioned earlier and facilitates patients to perform assigned tasks similar to daily activities. In such a way, these tasks require MS patients to use their upper extremities under both dynamic and static conditions. Furthermore, this device allows for objective assessments of patients through various performance metrics, such as comparing the force applied to an object against the required force (such as the force component ratios of a healthy individual) based on the ability to hold or release an object without slipping. This device endeavours not only to monitor and track the progress of patients but also to augment their performance through the utilization of tactile stimulations facilitated by the device. The proposed system for rehabilitation and clinical evaluation presents an electromechanical design that enables objective assessments, detached from subjective viewpoints. The efficacy of the proposed device was assessed through the involvement of two distinct cohorts of healthy participants. The control group, characterized by its normative status, actively engaged in the experimental procedures, while the other group experienced a significant symptom of MS by means of a specialized glove. Accordingly, the aforementioned performance metrics indicate two separate performance characteristics of each group. Thus, the aforementioned performance metrics delineate distinct performance attributes exhibited by each group.

mds
July 25, 2023

A. Materials and Methods

This section presents the design objectives for the thimble-like sensor fusion system, data gathering from the device and implementation of the proposed objectives.

1) Design Objectives: The proposed methodology in this study enables the wearable device to provide individuals with the autonomy to manoeuvre their hands in 3-D coordinate space and actively participate in dynamic activities that closely resemble those encountered in their daily routines. The performance of the device is evaluated based on the introduced terminology in the literature [18]. Accordingly, the imperative design requirement of the wearable sensor fusion system is the system's capability to perform realistic daily activities by hand.

Manipulating different thicknesses and weights of objects is an optimal design requirement in terms of the activities requiring various difficulty levels of manipulation skills. The findings in the literature also support the importance of objective assessment therapy evaluating the progression or regression of patients' hand response after completing a task of placing a cylindrical object from one place to another [12]. Such a therapy provides advantages in terms of testing the force coordination applied under dynamic load. However, the aforementioned therapy does not allow to change in the difficulty level of the tasks as the employed object has a single geometric shape and a constant load. Enhancing such tasks applied in dynamic conditions with various difficulty levels provides more efficacy to patients during their therapy. Accordingly, the optimal design requirement of the thimble-like wearable device allows the grabbing of objects with different thicknesses requiring more manipulation difficulties based on the real-time performance assessment of the patients.

The primary design requirement is the ability to detect the applied force, the incipient slippage time, the duration to hold the object stable, and also to observe the force coordination applied by the fingers of the patient to the object at such critical moments. That being said, although an object can be held without slippage, MS patients may apply more force than required, or sometimes cause it to slip due to the lack of exerted force. Such possibilities are frequently encountered in daily activities and can only be observed in experiments where realistic conditions are mimicked with a special device containing the above-mentioned features.

The secondary requirement of the device is that the wearable device is lightweight and adaptable to any size of the user. Accordingly, the lightweight design allows the patients to concentrate their focus only on manipulation. Besides, the wearable structure can be adjustable considering the ergonomics of the user. By following these criteria, the biomechatronic design is
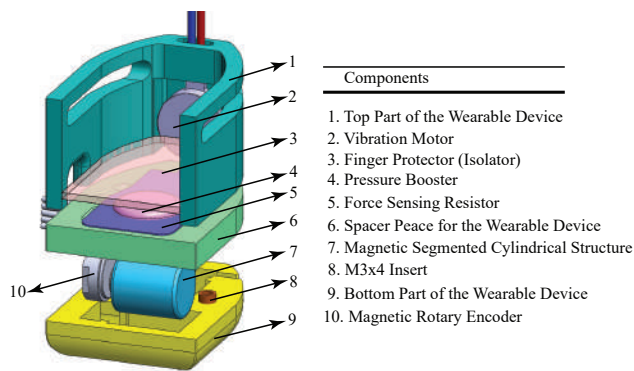
Components

1. Top Part of the Wearable Device
2. Vibration Motor
3. Finger Protector (Isolator)
4. Pressure Booster
5. Force Sensing Resistor
6. Spacer Peace for the Wearable Device
7. Magnetic Segmented Cylindrical Structure
8. M3x4 Insert
9. Bottom Part of the Wearable Device
10. Magnetic Rotary Encoder

Fig. 1. Exploded view of the solid model of the thimble-like device

Design Details

1. Hemispherical textured upper surfaces that provides balanced rotation
2. Vertical sliding detectors
3. Horizontal sliding detectors
4. Hemispherical textured lower surfaces that provides balanced rotation

Fig. 2. Housing of the sensors detecting incipient slippage

determined to be a solution to all the aforementioned needs. In light of the required design criteria, a wearable fingertip tactile sensor design capable of detecting the incipient slippage time and force variation while manipulating the objects is proposed.

The tertiary design requirement for the device is the capability of giving biofeedback to the user when the tasks are carried out with low performance. With the integration of such property into the device, the thimble-like device warns the patients via tactile feedback and forces them to regulate their motor control and accordingly improve the success rate of the task.

2) Electromechanical Design: The sensor fusion design is aimed to be wearable by fingertip and allows freedom to perform some of the activities of daily living as a means of pinch grasp. The device is basically composed of three segments, and the exploded view of the device's solid model is presented in Figure 1. The distal phalanx of the finger is placed into the allotted nest, enumerated by 1 and fixed by an adjustable tie. Thus, the finger is only responsible for carrying the rather light design it wears. Moreover, the wireless communication of the thimble-like device enables the patient to manipulate the fingers. Since the device is not fixed to a stationary place via wires, the use of the device does not compel the patient to stay nearby the device during the therapy. In both thin and lightweight thimble design and wireless communication, patients can perform realistic activities realized by pinch grasping anywhere they live, as targeted by the imperative design criterion.

With the objective of meeting the optimal design criterion, the dimensions of the sensor housings are determined considering the maximum size of a person's distal phalanx, as illustrated in Figure 2. Accordingly, minimizing the base of the device provides the patient with the ability to grasp objects of varying thicknesses and allows for spatial manipulation tailored to the specific requirements of different tasks.

Designing a highly responsive device to detect incipient slippage in both vertical and horizontal directions is one of the primary design criteria of the device. This crucial design aspect is accomplished by incorporating a high-resolution magnetic encoder, which empowers the device with exceptional accuracy and sensitivity in detecting even the slightest signs of slippage. In order to bring this feature to the system, two magnetic rotary encoders integrated into the system detect
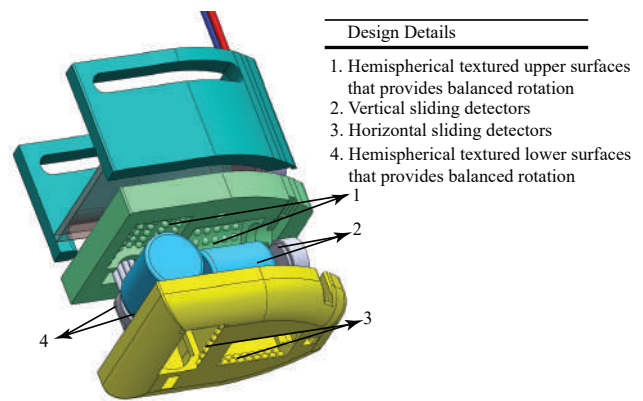
the angular movement occurring in the mechanical system and provide the necessary data about the slip moment. As detailed in Figure 3(a), the magnetic encoder enumerated by (1) detects the vertical sliding by sensing the rotation of the vertical cylinder embedded with a tiny magnet (2). Similarly, the horizontal sliding is sensed by the other pair of the magnetic encoder (4). The employed 12-bit non-contact magnetic encoders (RLS®Miniature Rotary Magnetic Encoder) are sensitive enough to detect nearly stationary motion (1 count equals 0.09° revolutions)). The nests of the cylindrical elements and the sensors are placed on part (3). The other half of the nests are also represented in Figure 3(b). The small hemispheres covered the nests to decrease the contact friction between the cylinders and their nests as much as possible. Accordingly, once the shear force is exerted by the object on the cylindrical elements, thanks to the negligibly small contact friction, the required slip moment as an indication of the slippage of the object is easily generated.

As for the measurement of the force exerted by the patient on the thimble-like device, a force-sensing resistor is placed on the protruding surface to be able to collect the force acting on this area efficiently. The system performs the necessary measurement during the follow-up of the patients based on the cooperation of the two sensors. The force sensing resistor (FSR), which is used to measure the force applied by the finger, has been selected in such a way that it can accurately detect the maximum force value [19] that men and women can apply in the pinch grasp category. The FSR is placed in its slot in the mechanical design of the wearable device, enumerated by 5 in Figure 1, to sense the force in the stressed direction. The FSR sensors (Tekscan®), which can detect a maximum force of 44N, can also measure the upper limit of the force of 12N exerted by a fingertip. The sensing diameter of the selected FSR is 3.8mm and thus does not exceed the dimensions of the device worn by the fingertip. That being said, the evaluation of data gathered from both types of sensors allows us to detect the amount of force exerted by the patient to grip the object while it tends to slip. Accordingly, the primary design criterion is satisfied by the collaborative work of slippage and FSR sensors by determining the moment/force to initiate the sliding of the object.

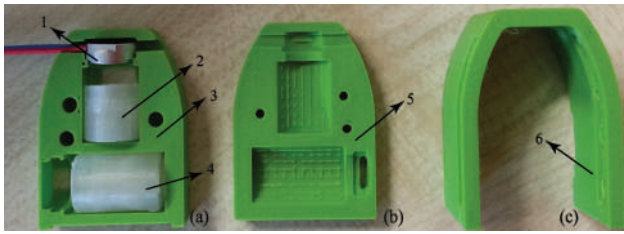The lightweight and ergonomic design criteria are met

Fig. 3. Fabrication of the lower part (a), spacer (b) and the finger nest (c) of the thimble-like device (a) The lower part of the wearable device also consists of two parts. The lower part houses the rotary encoder and the cylindrical parts in which the magnetic part is embedded. (b) The spacer of the wearable device incorporates the necessary spaces for mounting rotary encoders and force sensors. (c) The lateral sides of the top part are designed to accommodate the finger, which is designed according to the finger anthropomorphic model.

by avoiding more than the required size for the fingertips (18x27x11mm), choosing a flexible resin material with a density of 1.1 g/cm3, and designing the overall platform as durable yet thin enough. The criterion is satisfied as a means of UV LCD 3d printer manufacturing technique, which allows adjusting the size of the wearable device for any finger. From the ergonomic point of view, keeping the volume of the device as small as possible enables a wider workspace for the manipulation of the fingertips.

Lastly, the tertiary design criterion is addressed by adding a shaftless mini-vibration motor at the tip of the device (in Figure 1, enumerated by 2) whose vibration is sensed by the finger directly. For example, when the patient slides the object that is wanted to be held, the patient is expected to successfully perform the holding task by giving a vibration at the first moment of the incipient slippage. It is hypothesized that such simulations based on the performance of the patient help alleviate the patient's complaints as a means of recovery of the damaged nerve area in the long term.

II. RESULTS

The aim of the experimental procedure is to reveal the efficacy of the thimble-like wearable device for MS symptomatic patients when performing dynamic tasks that simulate activities of daily living (ADL). While our future endeavours include conducting experiments specifically with MS patients, it is of great importance that the experiments with healthy individuals can provide a preliminary assessment of the hand functions of MS patients in terms of shedding light on our future studies. Consequently, having a comprehensive understanding of the symptoms associated with MS is instrumental in shaping the experimental procedures. Additionally, to simulate the diminished sensory perception encountered by MS patients, the research group employs a strategy wherein a powder-free nitrile glove is worn by volunteers prior to using the robotic device. This approach intentionally restricts the sensory feedback experienced by the volunteers, aiming to replicate the limitations reported by individuals with MS.

A. Experimental Procedure

Somatosensory complaints are encountered in %70 of MS patients. These include symptoms such as decreased sensation,

numbness, and tingling. In this study, two groups of healthy volunteers contributed to the study. The first group who wears the latex glove represents the conditions in which MS patients live; whereas the second group is the control group responsible for unveiling the performance of the device considering the different grasping capacities of people. As indicated in Figure 6, the first group of volunteers represents healthy people and they wear the device directly on their bare fingers. Thus, the volunteer's perception of feeling the environment becomes clear to a large extent. This experimental group was named the control group.

The performance evaluation of the volunteers during the experiments is carried out based on their upper extremity (hand-arm) coordination and accordingly grasping ability. The performance evaluation is mainly concentrated on task-based force control and reaction time [20] against slippage. The experiments are performed based on conducting two different tasks. In the first task, as presented in Figure 7, it is expected for the subjects to grasp the object whose weight is gradually increased without slippage. In the second task as shown in Figure 8, the subjects need to achieve the grasping of the object while orienting their arm/ hand in the 3D space.

The salient aspects pertaining to the performance exhibited by the participants encompass the following key attributes:

- The time until successful grasp of the object whose weight is predicted.
- The dynamic nature of the applied force exerted in order to accomplish successful object grasping, considering the predicted weight of the object.
- The duration until the prevention of object slippage is achieved in response to the applied force exerted on the object.
- The duration necessary to avert slipping when administering vibrational stimulation to the subject in instances where their grip is deemed insufficient (also referred to as the reaction time)

B. Experimental Evaluations

Four healthy volunteers aged between 26 and 39 participated in the experiments performed in the study. The volunteers read and approved the informed consent form, which was approved by the Istanbul Medipol University Ethics Committee.

In the first experiment, volunteers are required to carry out the task of lifting objects with increasing weight until the last object does not touch the surface. The main concerns of the experiment are both the maximum displacement to shift the object and the dynamic behaviour of the force exerted on the grasped object. When the maximum sliding distance of the object was assessed based on the first performance metric in the initial experiment, it was observed that healthy volunteers demonstrated a more proficient performance by sliding the object at a lower rate compared to MS symptomatic volunteers, as depicted in Figure 9.

Upon conducting a t-test statistical analysis to evaluate the accuracy of performance assessment based on the sliding distance in the first experiment, we obtained p-values of 0.014, 0.0027, 0.0107, and 0.0137, respectively, all of which were
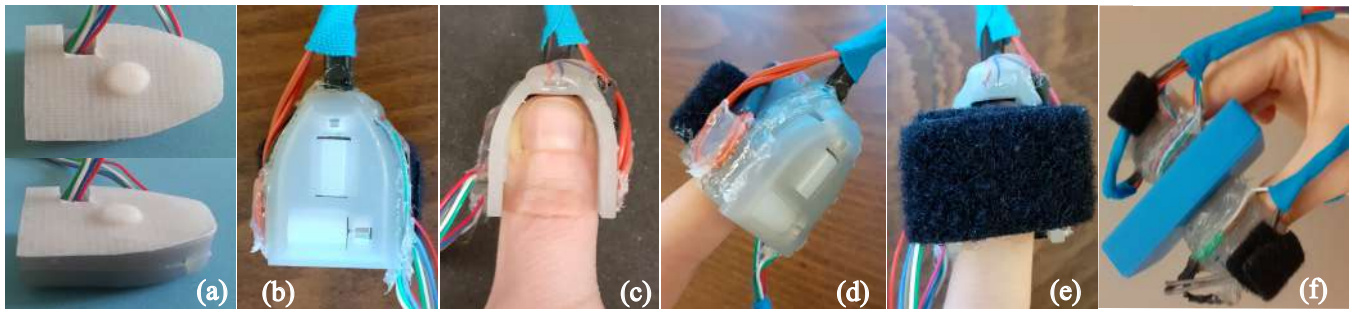
Fig. 4. Prototype of the thimble-like device (a) The protruding surface of the force-sensing sensor, which makes it possible to detect the exerted force more clearly (b) The magnetic encoder and rotary elements placed perpendicular to each other, viewed from the bottom of the device (c) The suitability of the wearable device for finger anthropomorphism (d) The ready-to-use device worn by the finger (e) A velcro tie fixes the device to the finger (e) A daily activity in the dynamic condition is performed using the device
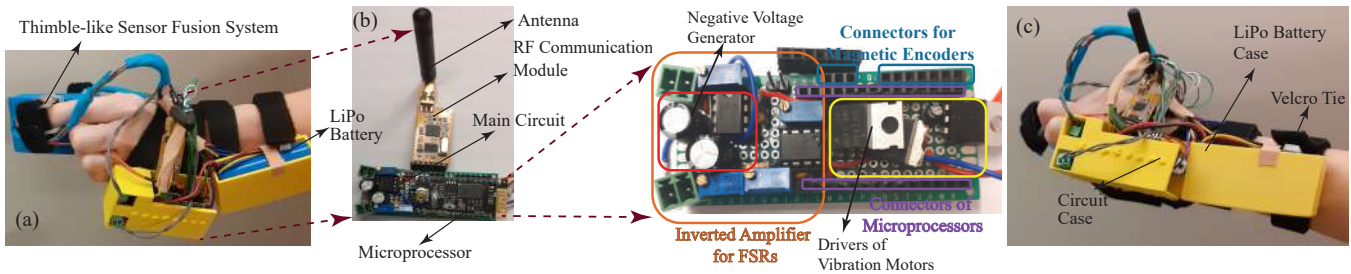
Fig. 5. Electronic unit of the thimble-like sensor fusion system

Fig. 6. Classification of the subjects based on their hand sensation capability

less than 0.05. This confirmed that the mean performances of each volunteer differed in the condition where they reduced hand sensitivity by using either bare hands or wearing nitrile gloves.

In the second experiment, volunteers are asked to place the object in slots opened in various orientations. The task of taking it from one slot and placing it in the other slot was repeated ten times. Two groups of volunteers carry out the task with bare hands and with wearing nitrile gloves. The performance metrics of this experiment are the sliding time and reaction time. The sliding time is the time spent inhibiting an object from sliding in response to the force exerted on the object; while the reaction time is the time elapsed to retain the object from slipping from the moment the subject is given vibratory stimuli.

The second experiment carried out while performing Task-

Fig. 7. The first task is evaluating the performance of the subjects while they are grasping and lifting the gradually increasing weight of the objects (a) equally separated three connected objects via wires (b-c) simultaneously grasping and lifting up the objects

2 involves measuring the duration for which the object is displaced during the course of executing the second task by the participants, using the performance metric known as "sliding
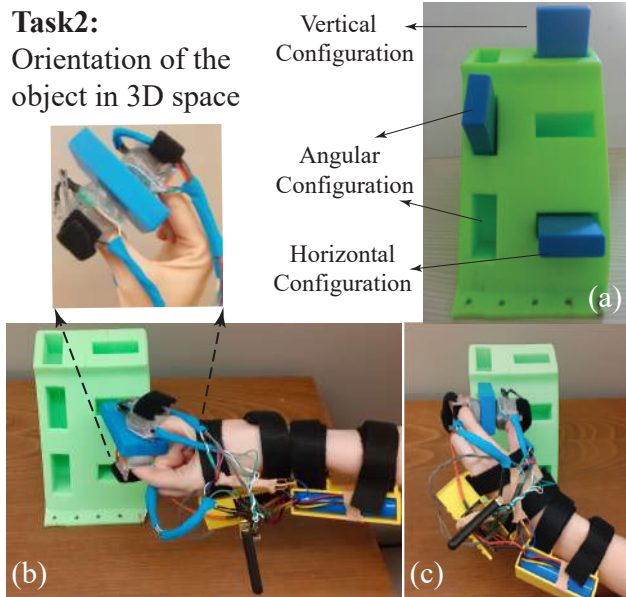
Fig. 8. The second task is evaluating the performance of the subjects while they are orienting their upper limbs in the 3D space (a) various options for different configurations of objects (b) orienting the object horizontally (c) orienting the object vertically

time". The performance of each volunteer is presented in Figure 10.

Based on the findings depicted in Figure 10 (a) and (b), the experiment involving the placement of an object from one nest to another nest with a different orientation resulted in significantly lower total displacement time for healthy individuals, as opposed to MS symptomatic volunteers. This difference was confirmed by a t-test, with the p-values obtained for each volunteer being 0.0186, 0.0113, 0.0333, and 0.0356, respectively, which are all lower than 5%.

The second experiment also sheds light on the effectiveness of vibrotactile stimulation, a form of biological feedback, in improving the performance of the volunteers during the execution of Task-2. During the object manipulation task in the final stage of the experiment, which involved picking up and placing the object from nests with varying configurations, both groups of volunteers received vibrational feedback in the event of slippage while holding the object. Accordingly, the volunteer reacts by perceiving the slipping of the object as a means of tactile stimulation and taking action to hold onto it. The time elapsed between the warning and the object stabilizing is defined as the reaction time. The reaction times of both the healthy and MS symptomatic groups are shown in Figures 11 (a) and (b), respectively. To verify this result, a t-test was conducted, which yielded p-values that are significantly greater than 0.05 for both groups (0.9374, 0.8489, 0.6653, 0.2152). This finding confirms that the data from the two groups are very similar to each other.

III. CONCLUSION

This study encompasses the design of a wearable robotic device intended to serve as a substitute for conventional clinical assessments in monitoring the progression of Multiple
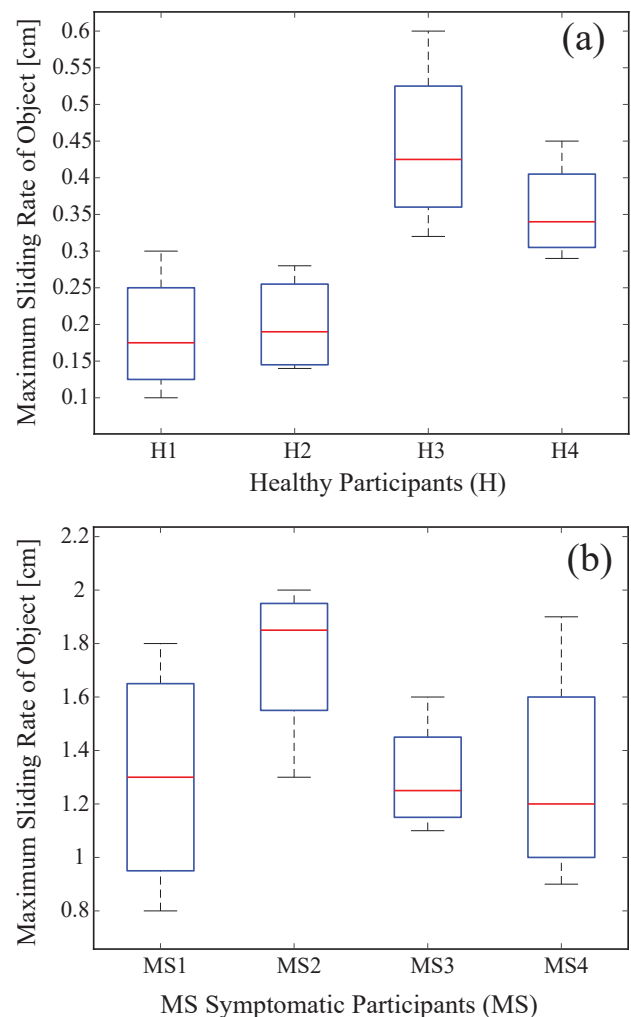
Fig. 9. The initial task involves assessing the participants' performance as they grasp and elevate the object with a gradual increase in weight a) The sliding rate of the healthy participants b) The sliding rate of the MS symptomatic participants

Sclerosis. The device offers high-sensitivity measurements, enabling objective evaluations, and allows patients of various age groups to utilize it conveniently in any desired place. Furthermore, the inclusion of an integrated vibrotactile stimulator enables patients to receive therapy concurrently in parallel with their performance evaluations. Within the scope of this study, the device undergoes experimental evaluation involving two distinct groups: healthy volunteers and individuals exhibiting artificially MS-related symptoms (MS symptomatic volunteers). While the primary focus of this study lies in substantiating the design criteria of the proposed wearable device, the current number of participants is deemed adequate.

Preliminary findings from the initial set of experiments corroborate the observation that individuals with multiple sclerosis tend to exert excessive force during object grasping, owing to diminished sensory perception. In contrast, healthy individuals demonstrate superior control modulation during activities of daily living (ADL) compared to MS-symptomatic people. Consequently, while the healthy cohort promptly rectifies any slipping occurrences by firmly gripping
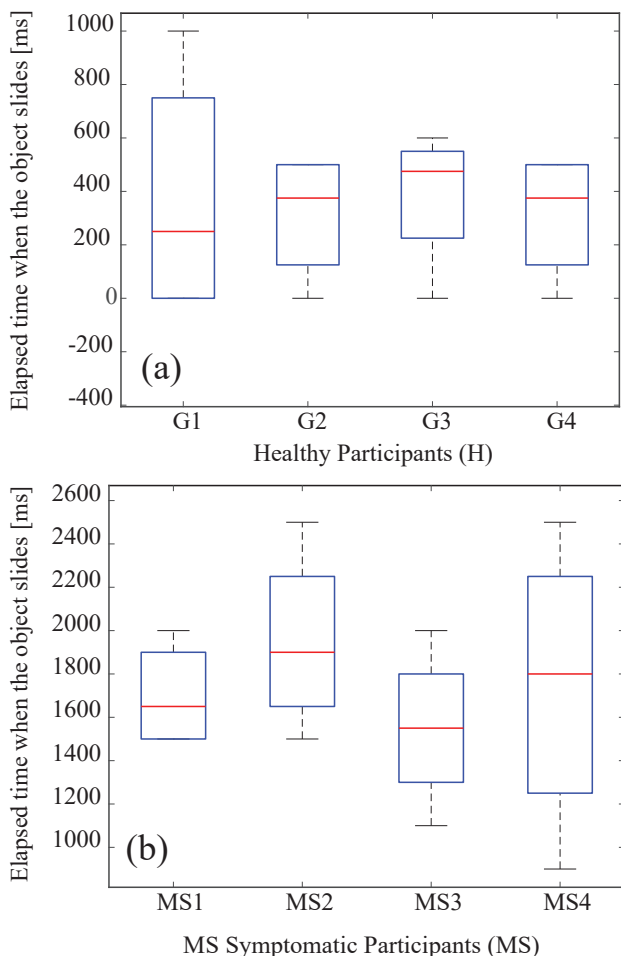
Fig. 10. In the assessment of the performance of each group of participants in the second task, the sliding rate, which denotes the time duration of the object's sliding, is the performance metric that is commonly employed. a) The sliding time of the healthy participants b) The sliding time of the MS symptomatic participants

Fig. 11. In the assessment of the performance of each group of participants in the second task, the reaction time, which measures the time it takes for a participant to respond to the tactile stimulation, is used as the third performance metric. a) The reaction time of the healthy participants b) The reaction time of the MS symptomatic participants

the object, MS-symptomatic individuals experience prolonged and frequent instances of slippage.

The subsequent experiment expands the assessment scope beyond a unidirectional upward movement, as observed in the first experiment, to enclose activities involving orientation and rotational motions. Notably, the MS symptomatic group exhibited a protracted duration of shifting during the task compared to the control group. However, the final experiment yielded a surprising outcome. With the influence of the stimulant, the MS symptomatic group demonstrated enhanced reaction speed and lessened reaction time. Intriguingly, the reaction times of both groups converged, reinforcing the hypothesis that a tactile stimulus system could potentially ameliorate neural dysfunctions.

Among the future aims of this study, there also lies the inclusion of conducting the aforementioned experiments with a larger cohort, thus substantiating the veracity of the experimental outcomes. Moreover, the future objective of this study also entails the integration of the proposed wearable robotic device with an immersive virtual reality environment to pro-

vide patients with diverse tasks. The aim is to investigate the therapeutic potential of the device in facilitating the recovery process when used consistently by patients. By combining the device with a motivating virtual reality setting, the study aims to assess the impact of this combined approach on the patient's rehabilitation progress. The evaluation will focus on the extent to which the device enhances patient outcomes after regular use and contributes to their overall recovery journey. The envisioned system for rehabilitation and clinical evaluation embodies an intricate electromechanical design that facilitates unbiased and objective assessments, transcending subjective perspectives.

REFERENCES

[1] A. Guo, M. Grabner, P. S. R., J. Elder, M. Sidovar, P. Aupperle, and S. Krieger, "Treatment patterns and health care resource utilization

associated with dalfampridine extended release in multiple sclerosis: a retrospective claims database analysis," *ClinicoEconomics and Outcomes Research: CEOR*, vol. 8, pp. 177–186, 2016.

[2] C. Asche, M. Singer, and M. e. a. Jhaveri, "All-cause health care utilization and costs associated with newly diagnosed multiple sclerosis in the united states," *J Manag Care Pharm.*, vol. 16, no. 9, p. 703–712, 2010.

[3] K. Grice, K. Vogel, V. Le, A. Mitchell, S. Muniz, and M. Vollmer, "Adults norms for a commercially available nine-hole peg test for finger dexterity," *American Journal of Occupational Therapy*, vol. 57, no. 3, p. 570–573, 2003.

[4] J. Kurtzke, "Rating neurologic impairment in multiple sclerosis: An expanded disability status scale (edss)," *Neurology*, vol. 33, no. 1, p. 1444–1452, 1983.

[5] R. Jebsen, N. Taylor, R. Trieschmann, M. Trotter, and L. Howard, "An objective and standardized test of hand function," *Arch Phys Med Rehabil.*, vol. 50, no. 6, p. 311–319, 1969.

[6] R. Kalb, *Multiple Sclerosis: The Questions You Have-The Answers You Need, Demos Health*. Demos Medical Publishing, 2011, ch. 2. How is multiple sclerosis treated, pp. 45–195.

[7] O. Lambercy, L. Dovat, H. Yun, and et al., "Robotic assessment of hand function with the hapticknob," in *International Convention on Rehabilitation Engineering and Assistive Technology*, 2010, p. 33:1–33:4.

[8] S. Gorniak, M. Plow, C. McDaniel, and J. Alberts, "Impaired object handling during bimanual task performance in multiple sclerosis," *Mult Scler Int.*, vol. 4, no. 450420, 2014.

[9] S. Jaric, C. Knight, J. Collins, and R. Marwaha, "Evaluation of a method for bimanual testing coordination of hand grip and load forces under isometric conditions," *J Electromyogr Kinesiol.*, vol. 15, no. 6, pp. 556–63, 2005.

[10] S. Jaric, E. Russell, J. Collins, and R. Marwaha, "Coordination of hand grip and load forces in uni- and bidirectional static force production tasks," *Neuroscience Letters*, vol. 381, no. 1-2, pp. 51–56, 2005.

[11] A. Hussain, S. Balasubramanian, I. Lamers, S. Guy, P. Feys, and E. Burdet, "Investigation of isometric strength and control of the upper extremities in multiple sclerosis," *J Rehabil Assist Technol Eng.*, vol. 3, no. 4, 2016.

[12] V. Iyengar, M. Santos, M. Ko, and A. Aruin, "Grip force control in individuals with multiple sclerosis," *Neurorehabil Neural Repair.*, vol. 23, no. 8, pp. 855–61, 2009.

[13] T. Platz, C. Pinkowski, and F. e. a. van Wijck, "Reliability and validity of arm function assessment with standardized guidelines for the Fugl-Meyer Test, Action Research Arm Test and Box and Block Test: A multicentre study," *Clin Rehabil*, vol. 19, no. 4, pp. 404–411, 2005.

[14] R. Johansson and G. Westling, "Roles of glabrous skin receptors and sensorimotor memory in automatic control of precision grip when lifting rougher or more slippery objects," *Exp Brain Res.*, vol. 56, no. 3, pp. 550–564, 1984.

[15] J. Flanagan and A. Wing, "The stability of precision grip forces during cyclic arm movements with a hand-held load," *Exp Brain Res.*, vol. 105, no. 3, p. 455–464, 1995.

[16] V. Krishnan and S. Jaric, "Hand function in multiple sclerosis: force coordination in manipulation tasks," *Clin Neurophysiol.*, vol. 119, no. 10, p. 2274–2281, 2008.

[17] V. Krishnan, P. Barbosa de Freitas, and S. Jaric, "Impaired object manipulation in mildly involved individuals with multiple sclerosis," *Clin Neurophysiol.*, vol. 12, no. 1, pp. 3–20, 2008.

[18] J. Merlet, *Parallel Robots*. 2nd Edn. Springer, 2006.

[19] T. Nilsen, M. Hermann, C. Eriksen, H. Dagfinrud, M. P., and I. Kjeken, "Grip force and pinch grip in an adult population: Reference values and factors associated with grip force," *Scandinavian Journal of Occupational Therapy*, vol. 19, no. 3, pp. 288–296, 2012.

[20] T. Ersoy and E. Hocaoglu, "A 3-DoF robotic platform for the rehabilitation and assessment of reaction time and balance skills of ms patients," *PLoS ONE*, vol. 18, no. 2, p. e0280505, 2023.

Elif Hocaoglu received her Ph.D. degree in Mechatronics Engineering from Sabanci University, Istanbul in 2014. She worked as a post-doctoral researcher at Advanced Robotics Laboratory at the Italian Institute of Technology, and then Human Robotics Laboratory at Imperial College London. Currently, she is an assistant professor at Istanbul Medipol University. Her research is in the area of physical human-machine interaction, in particular sEMG-based control interfaces with applications to upper extremity prostheses, rehabilitation and assistive devices, and wearable robotics. Her research extends to soft robotics.

Classification and Regression Using Automatic Machine Learning (AutoML) – Open Source Code for Quick Adaptation and Comparison

Oguzhan Topsakal and Tahir Cetin Akinci

Abstract— **This paper presents a comprehensive exploration of automatic machine learning (AutoML) tools in the context of classification and regression tasks. The focus lies on understanding and illustrating the potential of these tools to accelerate and optimize the process of machine learning, thereby making it more accessible to non-experts. Specifically, we delve into multiple popular open-source AutoML tools and provide illustrative examples of their application. We first discuss the fundamental principles of AutoML, including its key features such as automated data preprocessing, feature engineering, model selection, hyperparameter tuning, and model validation. We subsequently venture into the hands-on application of these tools, demonstrating the implementation of classification and regression tasks using multiple open-source AutoML tools. We provide open-source code samples for two data scenarios for classification and regression, designed to assist readers in quickly adapting AutoML tools for their own projects and in comparing the performance of different tools. We believe that this contribution will aid both practitioners and researchers in harnessing the power of AutoML for efficient and effective machine learning model development.**

Index Terms— **AutoML, Machine Learning, Artificial Intelligence, Code, Adaptation, Sample, Classification, Regression**

I. INTRODUCTION

MACHINE LEARNING (ML) has experienced substantial advancements in recent years, becoming an indispensable tool in an expanding array of academic and industrial disciplines [1-3]. The efficacy and accuracy of machine learning models are contingent on a multitude of factors, each of which necessitates significant expertise and labor.

OGUZHAN TOPSAKAL, Department of Computer Science, Florida Polytechnic University, 33805, FL, USA. (e-mail: otopsakal@floridapoly.edu).
https://orcid.org/0000-0002-9731-6946

TAHIR CETIN AKINCI, Department of Electrical Eng. Istanbul Technical University (ITU), Turkey. and University of California Riverside (UCR), CA, USA (e-mail: tahircetin.akinci@ucr.edu).
https://orcid.org/0000-0002-4657-6617

Such factors include the preliminary task of data preprocessing and cleansing, the selection and creation of pertinent features, the determination of an optimal model family, the tuning of model hyperparameters, and, in the case of deep learning applications, the design of neural networks. Furthermore, a critical evaluation of the results derived from these models forms an integral component of the process [4, 5].

With the primary objective of democratizing the application of machine learning models, research has been undertaken to automate these tasks, thus eliminating the need for extensive knowledge and skill in machine learning principles. The culmination of these endeavors has resulted in the development of Automated Machine Learning (AutoML) tools [1,6]. These tools are designed to enable individuals without specialized machine learning expertise to successfully implement and gain valuable insights from machine learning models.

Automated Machine Learning (AutoML) refers to the process of automating the end-to-end process of applying machine learning to real-world problems. Traditionally, building a machine learning model involves a lot of manual work such as feature selection, model selection, hyperparameter tuning, etc., which require considerable expertise and time [7]. AutoML aims to automate these manual, time-consuming aspects of machine learning, making the technology more accessible and efficient. It's designed to make machine learning more user-friendly by enabling individuals without specific knowledge in the field to build models, improve model efficiency, and speed up the process.

Some of the steps AutoML can automate include [8-11]:

- *Data preprocessing*: It helps in cleaning the data and selecting the right features to feed into the model. This might include imputing missing values, handling categorical variables, scaling and normalizing data, and more.

- *Feature engineering*: It is the process of creating new features from existing ones to improve the performance of the model. AutoML can create and choose the most effective features to use.

- *Model selection*: There are many types of machine learning models (e.g., decision trees, linear models, neural networks, ensemble models), and each has its strengths and weaknesses depending on the problem at hand. AutoML

can automatically test different models on your data to find the one that works best.

- *Hyperparameter tuning*: This involves finding the optimal configuration for a given model to maximize its performance. AutoML can systematically try many combinations of hyperparameters to find the best ones.
- *Model validation and selection*: After training multiple models, AutoML can evaluate their performance and select the best one.
- *Prediction and interpretation*: Finally, AutoML can generate predictions using the best model and provide the functionality to understand the insights into the model's decisions.

AutoML systems can be especially useful for businesses and researchers with large datasets but without the time or resources to manually build and tune machine learning models. However, it's worth noting that while AutoML can help automate many tasks, understanding the fundamentals of machine learning is still important to interpret results and ensure ethical and responsible use of the technology.

Many AutoML tools have been developed to achieve various tasks at different levels of performance. With all AutoML tools, users should consider their specific needs, the nature of their data, and the problem they're trying to solve when deciding which AutoML tool to use. To compare the AutoML tools for the task at hand, researchers or practitioners should learn each tool's capabilities and develop the code to test the task with the AutoML tool.

II. SELECTED AUTOML TOOLS

We have six most frequently used, well-known AutoML tools to study in this paper.

A. AutoGluon

AutoGluon is an open-source Automated Machine Learning (AutoML) library developed by Amazon. It aims to enable easy-to-use and easy-to-extend AutoML with robust machine learning techniques and advanced features.

AutoGluon automates various steps of the machine learning pipeline such as data preprocessing, feature engineering, model selection, model training, and hyperparameter tuning. It also supports automatic ensembling and stacking of models, a powerful technique for achieving higher predictive performance.

AutoGluon is particularly known for its efficiency and flexibility. It provides users the option to have full control over the machine learning process while also allowing them to leverage automation capabilities when needed.

AutoGluon is designed to work with different types of data including tabular data, image data, and text data. AutoGluon's functionality has demonstrated competitive performance in several machine learning competitions and has been utilized in real-world applications, making it a strong contender in the AutoML field [12].

B. AutoKeras

AutoKeras is an open-source Automated Machine Learning (AutoML) library developed by DATA Lab at Texas A&M University. It's built on top of the popular deep learning library Keras, and it aims to make machine learning accessible to non-experts and improve the efficiency of experts in model development [13].

The primary focus of AutoKeras is to automate the process of model selection and hyperparameter tuning, thereby reducing the manual, often time-consuming, trial-and-error involved in designing optimal neural network architectures.

The main features of AutoKeras include automated model architecture search, preprocessing (ex. handle missing data, categorical data), and hyperparameter tuning.

AutoKeras supports multiple types of data and tasks, including image classification, text classification, and regression problems, and continues to evolve and expand its capabilities to encompass a broader range of applications [14-16].

C. Auto-Sklearn

Auto-Sklearn is an open-source library in Python for performing Automated Machine Learning (AutoML). It is an extension of the popular Scikit-learn machine learning library and is built to automate the process of selecting the right machine learning model and tuning its hyperparameters [17].

The main features of Auto-sklearn include automated model selection, automated hyperparameter tuning, automated pre-processing. Auto-Sklearn uses ensemble methods to combine the predictions of multiple models, which can often lead to better predictive performance. Auto-sklearn is designed to be used with a minimal amount of code and has been engineered to fit into the Scikit-learn ecosystem, making it easy for those who are already familiar with Scikit-learn to start using Auto-sklearn [18, 19, 20].

D. H2O

H2O's AutoML is an automated machine learning tool developed by H2O.ai, which is well-known for its scalable and fast machine learning platform. The H2O AutoML library provides automated model selection, hyperparameter tuning, and ensemble learning capabilities, thus simplifying the process of building machine learning models [21].

The key features of H2O's AutoML include automated model selection, automated hyperparameter tuning, and automated ensemble learning: H2O's AutoML automatically trains two kinds of ensemble models at the end of its run - one is a simple ensemble using uniform weights, and the other uses stacking, a technique that uses a meta-learning algorithm [22-24] to learn how to best combine the predictions from multiple models. H2O's platform is designed for scalability and can handle large datasets and complex computations efficiently.

E. PyCaret

PyCaret is an open-source, low-code machine learning library in Python [24-25] that allows you to go from preparing your data to deploying your model within seconds. It is designed to expedite the process of building machine learning

pipelines for both binary and multiclass classification problems, regression problems, and clustering, among others [24-28].

The key features of PyCaret's AutoML include preprocessing, model selection, hyperparameter tuning, model analysis, and ensembling. It also provides a method for stacking models where a meta-model is trained on the predictions of base models.

F. TPOT

TPOT (Tree-based Pipeline Optimization Tool) is an open-source Automated Machine Learning (AutoML) tool developed in Python. The main goal of TPOT is to automate the process of building optimal machine learning pipelines, making it easier for researchers and data scientists to build efficient machine learning models [29].

One of the key features of TPOT includes automated pipeline optimization using genetic programming. TPOT also has the functionality of automated feature preprocessing, feature selection, model selection, and hyperparameter tuning.

TPOT is built on top of Scikit-learn, a popular machine learning library in Python. This makes it compatible with Scikit-learn's extensive range of functions and features.

Although TPOT automates the machine learning process, it also allows users to customize the search space and other parameters for the genetic programming algorithm, providing a balance between automation and control. After TPOT finds the optimal pipeline, it can export the corresponding Python code. This enables users to understand the pipeline's structure and make further customizations if necessary.

TPOT is computationally intensive and may require a lot of time and computational resources to find the optimal pipeline, particularly for large and complex datasets [29-32].

III. CODE SAMPLES

In this study, we also provide open-source code samples for regression and classification tasks for several well-known AutoML tools. The availability of code samples for AutoML tool will help compare and adapt the tools. The code samples provide practical examples of how to use the tool. This is especially useful for beginners or those transitioning from other tools, as it helps them understand how to apply the tool to real-world scenarios. Moreover, the code will help the researcher save time by letting them utilize the code rather than writing code from scratch. The samples also showcase different features of the tool, illustrating how to use and combine these features to achieve the desired results. A working example of the code is again beneficial when users encounter problems or errors while using the tool, they can refer to the code samples to check how certain functions or features should be used correctly. We hope the sample code will encourage adaptation by reducing the barrier to entry for using the tool.

Jupyter Notebook pages including sample code for AutoML tools AutoGluon, AutoKeras, Auto-Sklearn, H2O, PyCaret and TPOT can be downloaded from the GitHub page [33]. The code has been tested on Google Colab and includes samples for both regression and classification. The dataset used for regression includes features to predict used car values [34]. The dataset used for classification includes the data from the well-known Kaggle's Titanic competition [35].

IV. DISCUSSION

Adopting Automated Machine Learning (AutoML) tools into the machine learning process comes with several challenges:

- *Interpretability and transparency*: While AutoML tools can make machine learning more accessible, they can also act as "black boxes" where it's hard to understand why a particular decision or prediction has been made. This can make it difficult for users to trust the system and can be a barrier to adoption, especially in domains where explainability is critical, like healthcare or finance.

- *Lack of customization*: AutoML tools are designed to automate and simplify many steps in the machine learning process, but this can also limit their flexibility. Advanced users or those with specific needs may find that they don't have as much control or ability to customize the model as they would with traditional machine learning approaches.

- *Resource consumption*: AutoML methods often involve exploring a large number of different models and hyperparameters, which can be computationally intensive and time-consuming, especially for larger datasets or more complex models.

- *Data privacy and security*: Like all machine learning approaches, AutoML needs access to potentially sensitive data to train models. Ensuring that this data is used and stored securely is a critical challenge.

- *Quality of input data*: The quality of the results produced by AutoML is heavily dependent on the quality of the input data. AutoML can automate many parts of the machine learning process, but it may still struggle with poorly-formatted, inconsistent, or biased data. Users of AutoML tools may not always have the expertise to recognize and address these issues.

- *Evaluation of results*: While AutoML tools can automate the process of evaluating and comparing different models, interpreting these results can still be challenging, especially for non-experts. Users may need a solid understanding of machine learning concepts to make sense of the results and choose the best model for their needs.

In spite of these challenges, the field of AutoML is rapidly evolving, and many ongoing research efforts are aimed at addressing these issues. The ultimate goal is to create systems that can automate as much of the machine learning process as possible, while still being transparent, customizable, and easy to use.

V. CONCLUSION

In this research paper, we have embarked on a deep investigation into the realm of Automated Machine Learning (AutoML) tools, specifically focusing on their application in

classification and regression tasks. Our exploration aimed to reveal the immense potential of AutoML tools in streamlining and optimizing the process of machine learning, essentially democratizing the field by making it more approachable to non-experts.

Our discourse covered the fundamental principles of AutoML, elucidating key features such as automated data preprocessing, feature engineering, model selection, hyperparameter tuning, and model validation. We extended this theoretical understanding into practical application, demonstrating the implementation of AutoML tools in real-world classification and regression tasks.

In doing so, we illustrated the significant capabilities of these tools, highlighting their ability to quickly adapt to varied data scenarios and problem statements. Our findings confirmed the proficiency of AutoML tools in efficiently and effectively developing machine learning models.

To further aid in the comprehension and utilization of AutoML, we have made available a broad set of open-source code. This repository is designed to provide immediate assistance to readers, enabling them to swiftly adapt AutoML tools for their projects and perform comparative analysis between different tools.

In conclusion, our study underpins the value and impact of AutoML tools in the modern data-driven era. We posit that these tools will play a crucial role in the future of machine learning, enabling professionals and researchers alike to harness their power for efficient and effective model development. We hope that our contributions in this study will empower more individuals to embrace and exploit AutoML, leading to novel insights and breakthroughs in various fields.

REFERENCES

[1] Patil, P. S., Kappuram, K., Rumao, R., & Bari, P. (2022, May). Development Of AMES: Automated ML Expert System. In 2022 International Conference on Machine Learning, Big Data, Cloud and Parallel Computing (COM-IT-CON) (Vol. 1, pp. 208-213). IEEE.

[2] Glasby, L. T., Whaites, E. H., & Moghadam, P. Z. (2023). Machine Learning and Digital Manufacturing Approaches for Solid-State Materials Development. AI-Guided Design and Property Prediction for Zeolites and Nanoporous Materials, 377-409.

[3] Pugliese, R., Regondi, S., & Marini, R. (2021). Machine learning-based approach: Global trends, research directions, and regulatory standpoints. Data Science and Management, 4, 19-29.

[4] 4-Lu, S. C., Swisher, C. L., Chung, C., Jaffray, D., & Sidey-Gibbons, C. (2023). On the importance of interpretable machine learning predictions to inform clinical decision making in oncology. Frontiers in Oncology, 13, 780.

[5] Manduchi, E., & Moore, J. H. (2021). Leveraging automated machine learning for the analysis of global public health data: a case study in malaria. International Journal of Public Health, 31.

[6] Hutter, F., Kotthoff, L., & Vanschoren, J. (2019). Automated machine learning: methods, systems, challenges (p. 219). Springer Nature.

[7] Chauhan, K., Jani, S., Thakkar, D., Dave, R., Bhatia, J., Tanwar, S., & Obaidat, M. S. (2020, March). Automated machine learning: The new wave of machine learning. In 2020 2nd International Conference on Innovative Mechanisms for Industry Applications (ICIMIA) (pp. 205-212). IEEE.

[8] Singh, V. K., & Joshi, K. (2022). Automated Machine Learning (AutoML): An overview of opportunities for application and research. Journal of Information Technology Case and Application Research, 24(2), 75-85.

[9] Heizmann, M., Braun, A., Glitzner, M., Günther, M., Hasna, G., Klüver, C., ... & Ulrich, M. (2022). Implementing machine learning: chances and challenges. at-Automatisierungstechnik, 70(1), 90-101.

[10] Majidi, F., Openja, M., Khomh, F., & Li, H. (2022, October). An Empirical Study on the Usage of Automated Machine Learning Tools. In 2022 IEEE International Conference on Software Maintenance and Evolution (ICSME) (pp. 59-70). IEEE.

[11] Mengi, G., Singh, S. K., Kumar, S., Mahto, D., & Sharma, A. (2023, February). Automated Machine Learning (AutoML): The Future of Computational Intelligence. In International Conference on Cyber Security, Privacy and Networking (ICSPN 2022) (pp. 309-317). Cham: Springer International Publishing.

[12] Erickson, Nick & Mueller, Jonas & Shirkov, Alexander & Zhang, Hang & Larroy, Pedro & Li, Mu & Smola, Alexander. (2020). AutoGluon-Tabular: Robust and Accurate AutoML for Structured Data. ArXiv, abs/2003.06505.

[13] M. Feurer, A. Klein, K. Eggensperger, J. T. Springenberg, M. Blum, and F. Hutter, "Auto-sklearn: Efficient and Robust Automated Machine Learning," Automated Machine Learning, pp. 113–134, 2019, doi: https://doi.org/10.1007/978-3-030-05318-5_6.

[14] Alaiad, A., Migdady, A., Al-Khatib, R. E. M., Alzoubi, O., Zitar, R. A., & Abualigah, L. (2023). Autokeras Approach: A Robust Automated Deep Learning Network for Diagnosis Disease Cases in Medical Images. Journal of Imaging, 9(3), 64.

[15] Filippou, K., Aifantis, G., Papakostas, G. A., & Tsekouras, G. E. (2023). Structure Learning and Hyperparameter Optimization Using an Automated Machine Learning (AutoML) Pipeline. Information, 14(4), 232.

[16] Vincent, A. M., & Jidesh, P. (2023). An improved hyperparameter optimization framework for AutoML systems using evolutionary algorithms. Scientific Reports, 13(1), 4737.

[17] Jin, H., Chollet, F., Song, Q., & Hu, X. (2023). AutoKeras: An AutoML Library for Deep Learning. Journal of Machine Learning Research, 24(6), 1-6.

[18] Lee, S., Kim, J., Bae, J. H., Lee, G., Yang, D., Hong, J., & Lim, K. J. (2023). Development of Multi-Inflow Prediction Ensemble Model Based on Auto-Sklearn Using Combined Approach: Case Study of Soyang River Dam. Hydrology, 10(4), 90.

[19] Feurer, M., Eggensperger, K., Falkner, S., Lindauer, M., & Hutter, F. (2022). Auto-sklearn 2.0: Hands-free automl via meta-learning. The Journal of Machine Learning Research, 23(1), 11936-11996.

[20] Shi, M., & Shen, W. (2022). Automatic Modeling for Concrete Compressive Strength Prediction Using Auto-Sklearn. Buildings, 12(9), 1406.

[21] LeDell, E., & Poirier, S. (2020, July). H2o automl: Scalable automatic machine learning. In Proceedings of the AutoML Workshop at ICML (Vol. 2020).

[22] Singh, K., & Malhotra, D. (2023). Meta-Health: Learning-to-Learn (Meta-learning) as a Next Generation of Deep Learning Exploring Healthcare Challenges and Solutions for Rare Disorders: A Systematic Analysis. Archives of Computational Methods in Engineering, 1-32.

[23] Mohr, F., Wever, M., & Hüllermeier, E. (2018). ML-Plan: Automated machine learning via hierarchical planning. Machine Learning, 107, 1495-1515.

[24] Whig, P., Gupta, K., Jiwani, N., Jupalle, H., Kouser, S., & Alam, N. (2023). A novel method for diabetes classification and prediction with Pycaret. Microsystem Technologies, 1-9.

[25] Huynh, T., Mazumdar, H., Gohel, H., Emerson, H., & Kaplan, D. Evaluating the Predictive Power of Multiple Regression Models for Groundwater Contamination using PyCaret-23489.

[26] Liu, X., Wu, J., & Chen, S. (2023). Efficient hyperparameters optimization through model-based reinforcement learning with experience exploiting and meta-learning. Soft Computing, 1-18.

[27] Pol. U.R. and Sawant, T.U. "Automl: building an classification model with PyCaret", YMER, vol. 20, pp. 547-552, Dec. 2021, doi: 10.37896/YMER20.11/50

[28] N. Sarangpure, V. Dhamde, A. Roge, J. Doye, S. Patle and S. Tamboli, "Automating the Machine Learning Process using PyCaret and Streamlit," 2023 2nd International Conference for Innovation in Technology (INOCON), Bangalore, India, 2023, pp. 1-5, doi: 10.1109/INOCON57975.2023.10101357.

[29] R. S. Olson et al., "TPOT", Accessed on March 3, 2023, Available: http://epistasislab.github.io/tpot/

[30] Chen, X., Xu, J., Zhou, H., Zhao, Y., Wu, Y., Zhang, J., & Zhang, S. (2023). Tree-based machine learning models assisted fluorescent sensor array for detection of metal ions based on silver nanocluster probe. Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy, 297, 122738.

[31] Xiang, C. Y., Gao, F., Jakovlić, I., Lei, H. P., Hu, Y., Zhang, H., ... & Zhang, D. (2023). Using PhyloSuite for molecular phylogeny and tree-based analyses. iMeta, 2(1), e87.

[32] Grjazniha, M. (2023). Performance and Competitiveness of Tree-Based Pipeline Optimization Tool (Doctoral dissertation).

[33] Github: https://github.com/research-outcome/automl-sample

[34] Regression Dataset Based on a Used Card Dataset at Kaggle: https://www.kaggle.com/datasets/lepchenkov/usedcarscatalog

[35] Classification Dataset Adapted from Kaggle Titanic Competition: https://www.kaggle.com/competitions/titanic

BIOGRAPHIES

OGUZHAN TOPSAKAL (Senior Member IEEE) received his B.S. in Computer Engineering in 1996 from Istanbul Technical University, Turkey. He received his M.S. and Ph.D. in Computer Science from the University of Florida in 2003 and 2007. After gaining extensive experience in the software industry, Dr. Topsakal is currently an assistant professor in the Computer Science department at the Florida Polytechnic University since 2018. He teaches courses related to Machine Learning, Algorithm Design, Databases, and Mobile Development. Dr. Topsakal's research interests include applications of machine learning and deep learning in the medical field.

TAHIR CETIN AKINCI (Senior Member IEEE) pursued his Bachelor's degree in Electrical Engineering in 2000, followed by his Master's and Ph.D. degrees in 2005 and 2010, respectively. From 2003 to 2010, he worked as a Research Assistant at Marmara University in Istanbul, Turkey. Dr. Akinci is currently a full professor in the Electrical Engineering Department at ITU in 2021. Dr. Akinci assumed the role of a visiting scholar at the University of California Riverside (UCR). His research interests including artificial neural networks, deep learning, machine learning, cognitive systems, signal processing, and data analysis. In 2022, Dr. Akinci was honored with the International Young Scientist Excellence Award as well as the best researcher award for his exceptional research achievements.

Deep Belief Network Based Wireless Sensor Network Connectivity Analysis

Ayhan Akbas and Selim Buyrukoglu

Abstract—**Wireless sensor networks (WSNs) are widely used in various fields, and their deployment is critical to ensure area coverage and full network connectivity to achieve the maximum network lifetime. In this study, we present a mixed-integer programming (MIP) model that deeply investigates deployment parameters to optimize lifetime and analyze network connectivity. We further analyze the obtained results using Deep Belief Network (DBN) and Deep Neural Network (DNN) algorithms to achieve higher accuracy rates. Our evaluation shows that the DBN outperforms the DNN with an accuracy rate of 81.2%, precision of 81.2%, recall of 99.1%, and an F1-Score of 0.78. We also utilize two different datasets to justify the efficiency of the DBN in this research. The findings of this study emphasize the validity of our DBN algorithm and encourage further research into lifetime optimization and connectivity analysis in WSNs.**

Index Terms—**Deep Belief Network, Wireless Sensor Network, Connectivity**

I. INTRODUCTION

WIRELESS sensor networks (WSNs) are networks of dedicated sensors distributed in space that monitor physical conditions in a given environment, and collect and transmit the data they gather to a central base station. With the increasing use of WSNs, research into the deployment of sensor nodes has become more intense in recent years. The main goals of sensor node deployment are to achieve the best possible network lifetime and the widest possible coverage area at the lowest cost possible. For that reason. in the planning phase of a WSN deployment, the aim is to maximize network lifetime while achieving planned coverage within the designated budget.

Wireless sensor networks (WSNs) rely on a sparse placement of sensor nodes on a large scale to monitor physical conditions and transmit data to a central base station. Ensuring connectivity among these nodes is crucial for accurate data collection. To address this issue, several studies have been conducted in recent years. Sheikh-Hosseini and Hashemi,

AYHAN AKBAS is with the Department of Computer Engineering, Engineering Faculty, Abdullah Gul University, Kayseri, Turkey e-mail: (ayhan.akbas@agu.edu.tr)

[ID] https://orcid.org/0000-0002-6425-104X

SELİM BUYRUKOGLU is with the Department of Computer Engineering, Engineering Faculty, Cankiri Karatekin University, Cankiri, Turkey e-mail: (sbuyrukoglu@karatekin.edu.tr)

[ID] https://orcid.org/0000-0001-7844-3168

for instance, studied node connectivity and different coverage types in WSN deployments [1]. They proposed a new model that optimizes node placement to cover all targets with the required number of nodes while maximizing lifetime and coverage using genetic algorithms. Another study by Senouci and Mellouk [2] investigated optimal plan strategies for WSN topologies to achieve optimal network connectivity, sensing coverage quality, reliability, and lifetime with minimal cost. They developed a probabilistic-based communication cost model and demonstrated that their deployment approach can meet the needs of real-world fusion-based WSNs with predictable performance. Aitsaadi et al. [3] also addressed the deployment problem of WSNs with the objective of achieving the best network topology with minimal deployment cost while ensuring network connectivity and optimal network lifetime. They proposed a Tabu search metaheuristic and multi-objective deployment algorithm for their solution. Sengupta et al. [4] investigated the optimal trade-offs among coverage, lifetime, energy consumption, and connectivity using a multi-objective evolutionary algorithm-based model that outperforms single-objective node deployment schemes. Sevgi and Kocyigit [5] proposed a novel framework for the optimal deployment of WSNs from the connectivity perspective, particularly for random deployments. Overall, these studies contribute to the understanding and optimization of WSN deployments for connectivity, coverage, reliability, and lifetime, opening avenues for further research in the field [6], [7].

Although several studies have investigated connectivity issues in wireless sensor networks (WSNs), these studies have mostly utilized multilayer perceptron (MLP) and backpropagation (BP) [8] to address the problem. Notably, there is a lack of research that has employed Deep Belief Networks to address WSN connectivity issues. Given the need for enhanced performance and robust classification in WSNs, exploring the potential of DBN becomes imperative. DBN's superiority lies in its efficient utilization of hidden layers, yielding substantial performance gains when compared to Multilayer perceptron models. Furthermore, DBN exhibits specific robustness in classification tasks, effortlessly handling variations in topology, internode distances, transmission power levels and other crucial channel parameters. Therefore, incorporating DBN into the realm of WSN research can significantly contribute to overcoming connectivity challenges and advancing the capabilities of wireless sensor networks.
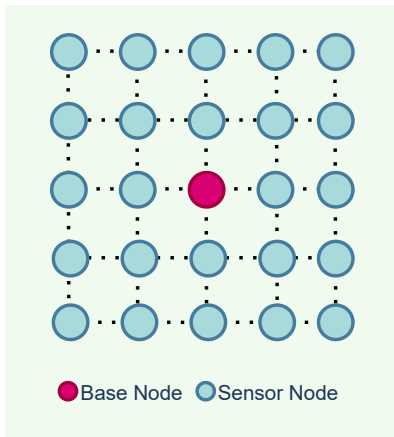
Fig. 1: Distributions of nodes in WSN grid topology

Fig. 2: Proposed Model Diagram

II. WSN SYSTEM MODEL

In this model, WSN network topology is constructed as a data flow graph and mathematically modelled using Mixed Integer Programming (MIP) in GAMS [9].

A. Wsn Model Overview

The WSN is modelled as a square topology (Fig.1) network consisting of a sink node in the center and sensor nodes around it. Sensor nodes in the network either transfer the data they collect directly to the sink node (single-hop) or relay it over other nodes (multi-hop). The 60-second round time is shared by all nodes in equal time slots, and each node transmits data in its own time slot. In each round, each node transmits its generated packet data in its own time slot. Internode transmission is performed by handshake, and every packet sent is confirmed with an ACK message, confirming that the transmission was successful. In the optimization, the maximum network lifetime is set as the objective in the linear programming. In the design of the model, it is assumed that the nodes are stationary, and all have clock synchronization with the sink node. Furthermore, we assume that the sink node's energy resources are limitless. All topology and route information is provided at the sink node. Furthermore, we assume that the data packets cannot be fragmented or aggregated.

B. Link Layer Model

In this study, the log-normal shadowing model [10] is assumed and a two-way handshaking link layer data transmission model is used, which is the extended version of the study by Akbas et. al.[11]. The MIP Model has been coded in GAMS [9] and MATLAB [12], and verified with real-world data.

C. Linear Programming Model

Each data transfer between nodes can be considered as an energy expenditure. Each bit of data sent is an energy cost that reduces the lifetime of the node. Therefore, calculating the maximum network time for the WSN can be considered
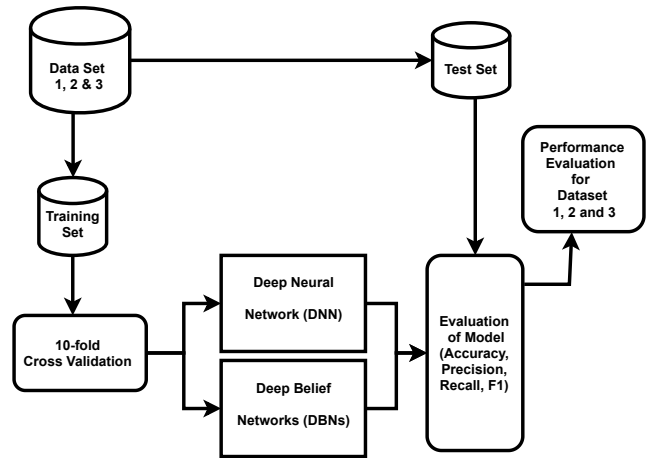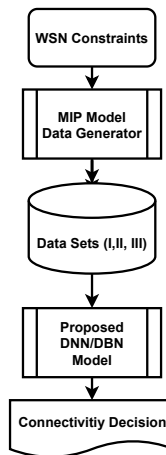
Fig. 3: WSN dataset creation

as a graph flow problem to be solved. In complex integer programming, we calculate the maximum network lifetime by defining the flow constraints on the network and modelling the flow on the WSN. The constraints we used in the model are:

- Balanced flow constraint states that for each node, the sum of the data produced in a node and the data coming from outside the node is equal to the amount of data sent from the node.
- The channel bandwidth constraint limits the channel bandwidth required to perform communication operations at each node.
- The interference matrix guarantees that the total duration of inbound data streams, outbound data streams, and interference streams is limited by the total network lifetime time.
- The non-negative constraint ensures that no data stream can be negative, i.e., all flows have to reach the sink node.

III. PROPOSED APPROACH

This section presents the proposed WSN classification approach, including four phases, which are data creation (Fig.3),

data cleaning and preprocessing, deep learning models, and evaluation metrics. The proposed model diagram is depicted in Fig.2.

A. Data Creation

Three different datasets were created and used in this study to reveal the efficiency of the deep neural network (DNN) and deep belief networks (DBNs) in the use of variety of datasets. The rest of this section explains the details of three different datasets.

We have created and used 3 sets of data as given in Fig. 3, which are:

- Dataset I: The dataset has been created with the shadowing factor taken into account. The dataset has the following features: Node-count (49,81 and 121 nodes), Packetsizes (1,2,3,4,5,6 and 8), Internode distances (50 to 109 meters with 1 meter increments), Full-Connectivity (1 or 0). It has 1260 records.
- Dataset II: Exactly the same as Dataset I except for the fact that the shadowing factor has not been considered in the calculations.
- Dataset III: Dataset-III obtained through the outcome of the model, contains the following features: Shadowing (1.5 to 5.5), Path Loss exponent (2 to 4), Packet Size (1 to 8), Transmission Powerlevel (6 to 26), Node Count (9 to 121), Internode Distance (20 to 100) and the Network lifetime (0 to 5M). The size of the dataset is 8200 records.

B. Deep Learning Models

In this study, Deep Belief Networks (DBNs) and Deep Neural Network (DNN) algorithms were used in this study. The dataset was split into training (80% - 941 for train) and testset (20% - 236 for test) based on the hold out method for datasets I and II. Also, dataset III has 39375 samples, which were split into training (80% - 31500 for training) and testset (20%-7875 for testing). Ten-fold cross-validation was used in the training process of the employed DBNs and DNN while the test set (unseen data) was used to obtain the evaluation performance of the employed models. In other words, nested cross-validation was used in this study. The rest of this section covers the models' background and creation process.

1) Deep Neural Network (DNN): The structure of artificial neural networks (ANN) consists of input, hidden, and output layers. If the structure of an ANN consists of more than one hidden layer, it is considered a deep neural network (DNN) [13]. The structure of the created DNN is presented in Fig. 4. As it can be seen from Fig. 4, one input, two hidden layers, and one output layer were used in the judgement of WSNs' connectivity. Three nodes for input layers were used, while one node was used in the created DNN. Also, six and three neurons in the hidden layers were used as a result of attempting to use the various numbers of neurons in the hidden layers. Sigmoid was used as an activation function in the deep neural network model.
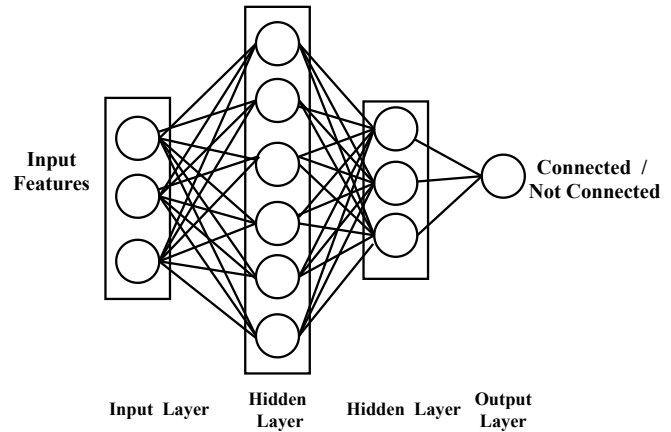
Fig. 4: Structure of the employed Deep Neural Network
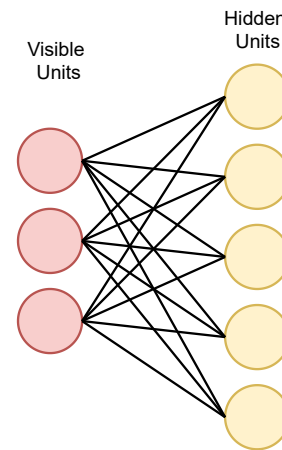
Fig. 5: RBM architecture

2) Deep Belief Networks (DBNS): The structure of a deep belief network (DBN) includes a sequence of restricted Boltzmann machines (RBMs), and they are connected sequentially [14]. Understanding RBMs is crucial for a thorough knowledge of DBNs. Thus, the rest of this section explains the RBMs and DBNs, respectively. The structure of an RBM consists of two units, which are the visible and hidden units [15]. Also, these two units are fully connected with both forward and backward connections. In an RBM structure, observable data is represented in one layer of visible units, while the hidden unit is used in terms of capturing dependencies of the visible units. Fig. 5 illustrates the structure of the employed RBMs. In an RBM, the weights related to each neuron are randomly initialized, and then the weights in a layer are updated based on the conditions of the visible and hidden units in the other layer. This is repeated until the system is sampling from its equilibrium distribution.

Deep Belief Networks (DBNs) are made up of multiple RBMs and a classifier [16]. A DBN can be considered a stacked model because it uses more than one RBM in its creation. Fig. 6 depicts the DBN architecture that was used. It

Fig. 6: The structure of the employed DBNs

TABLE I: Definition of TP, FP, TN, and FN

		Actual Class	
		Positive	Negative
Predicted Class	Positive	True Positive	False Positive
	Negative	False Negative	True Negative

TABLE II: Evaluation results of the models

Model	Accuracy	Precision	Recall	F1-Score
DNN	80.1	81.1	98.4	0.725
DBNs	81.2	81.2	99.1	0.728

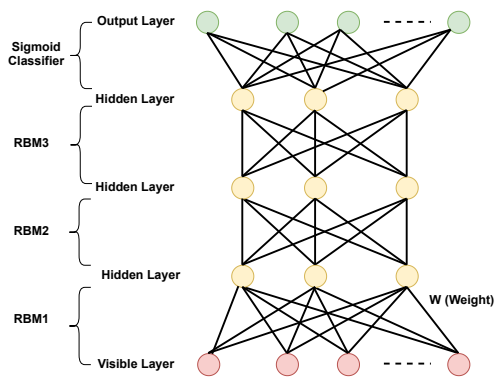contains three RBMs and a Sigmoid classifier. In the creation of the employed DBN structure, two and four RBMs were used, except for the three RBMs to obtain the optimum number of RBMs. In the end, the best performance was obtained through the use of three RBMs.

C. Scoring Metrics

The following scoring parameters are used in the evaluation of the employed deep belief networks (DBNs) and deep neural networks (DNNs). Formulas for the used scoring parameters are presented below [17]. The definition of TP, FP, TN, and FN is presented in Table I.

- Accuracy = (TN + TP)/(TN+TP+FN+FP),
- Precision = TN/(TN + FP),
- Sensitivity (Recall) = TP/(TP + FN)
- F1-Score = 2*((precision*recall)/(precision+recall))

IV. RESULTS AND DISCUSSION

A. Evaluation Of The Models

Table II presents the evaluation results of the employed deep neural networks (DNNs) and deep belief networks (DBNs) using Dataset I. Detailed information about the dataset was given in Sec. III-A. As it can be seen from Table II, DBNs provided better performance in terms of accuracy (81.2%), precision (81.2%), recall (100%), F1-score (72.8%), and Jaccard (81.2%). Even if the used DBNs provided better performance in the classification of fully-connectedness state analysis, the differences between the DBNs and DNN in terms of scoring metrics are not large.

TABLE III: Evaluation results of the models for different datasets

Dataset	Model	Accuracy	Precision	Recall	F1-Score
Dataset 2	DNN	81.6	84.4	92.7	89.4
	DBNs	84.4	86.4	100	91.5
Dataset 3	DNN	77.4	77.4	90.4	87.3
	DBNs	80.2	85	100	87.6

B. Evaluation Of The Models With Different Datasets

To reveal the efficiency of the employed DBNs in the prediction of whether a given WSN is fully connected, two different datasets were also created. Detailed information about these two datasets (Dataset II and III) was given in Section III-A.

The efficiency of DBNs is compared with that of DNNs in the prediction of the fully-connectedness of WSN based on datasets II and III. Table III shows the prediction performance scores of DBNs and DNNs. The results indicate that DBNs outperformed DNN if WSN is fully connected for database II and III.

In other words, DBNs have the ability to provide more accurate performance than DNNs. On the other hand, the overall prediction performance of DBNs for dataset III is considerably worse than the prediction performance of DBNs for dataset II. Moreover, the overall performance of DNN for Dataset III is also worse than the prediction performance of DNN for Dataset III.

It should be highlighted that, the number of features in datasets II and III is different. Initially, DNN was created based on Dataset II. Then, the created DNN was also used in the prediction of the connection state of the WSN network. using Dataset III. Due to the different number of features in the dataset, the number of neurons in the hidden layers was arranged to create an optimal DNN structure for Dataset III. As a result, the arranged DNN structure provides better performance than the previous DNN structure created for Database II. Table III shows the scoring performance of the arranged DNN.

C. General Discussion

The computational complexity of NNs rises exponentially with node count. Therefore, this turns into a challenging problem to resolve for constrained WSN devices and new computational methods are required to be employed and implemented. Though there are specialized libraries for use with limited capability and computing platforms [18]. Although there are specialized libraries for use with limited capabilities and computing platforms [18], the training phase of machine learning algorithms constitutes the dominant part of the high computational cost. Once the training is complete, the testing stage is a quick process that generally produces the desired results quickly and at a lower computational cost.

We've used NN models up to 121 nodes, beyond which a reasonable amount of time is not possible. For this reason, relatively small-sized WSNs were chosen. Nevertheless, increasing the number of nodes within the WSN does not endanger the forecast performance of the NN model, as the outcome. Consequently, it's possible to assume that the parameters of

the more extensive networks can also be forecasted likewise. New optimization methods might be suggested to be able to analyze larger topologies in future studies. As highlighted in Section I, to the best of the authors' knowledge, there is no research that has investigated the WSN connectivity. Thus, this study revealed that DBNs can be used in the connectivity prediction for a WSN with given parameters.

V. CONCLUSION

A WSN needs to be fully connected so as to collect and transmit the data to the sink node. As a node fails or dies, WSN connectivity may be questioned as to whether it operates properly. In this regard, it is critical to justify a given topology in relation to WSN connectivity.In this sense, the prediction of WSN connectivity plays a key role in determining if WSN is still connected or disconnected. With the motivation that connectivity estimation can be achieved with higher accuracy and precision at a lower computational complexity, we suggested a hybrid solution based on Deep Belief Networks. Algorithms with deep architectures mostly have advantages compared to single-based algorithms in terms of robust results and higher accuracy. Therefore, Deep Belief Networks (DBNs) and Deep Neural Network (DNN) classification algorithms were employed in this study.Three different datasets were used to compare the scoring performance of these algorithms. DBNs provided better scoring values in the classification of WSN connectivity decisions than DNNs (see Table II and III) in the use of these three datasets. The findings reflect the importance of DBN compared to DNN in the classification of WSN lifetimes. As a further study, a hybrid deep neural network can be developed to predict WSN connectivity.

REFERENCES

[1] M. Sheikh-Hosseini and S. R. S. Hashemi, "Connectivity and coverage constrained wireless sensor nodes deployment using steepest descent and genetic algorithms," *Expert Systems with Applications*, p. 116164, 2021.

[2] M. R. Senouci and A. Mellouk, "A robust uncertainty-aware cluster-based deployment approach for wsns: Coverage, connectivity, and lifespan," *Journal of Network and Computer Applications*, vol. 146, p. 102414, 2019.

[3] N. Aitsaadi, N. Achir, K. Boussetta, and G. Pujolle, "Artificial potential field approach in wsn deployment: Cost, qom, connectivity, and lifetime constraints," *Computer Networks*, vol. 55, no. 1, pp. 84–105, 2011.

[4] S. Sengupta, S. Das, M. Nasir, and B. K. Panigrahi, "Multi-objective node deployment in wsns: In search of an optimal trade-off among coverage, lifetime, energy consumption, and connectivity," *Engineering Applications of Artificial Intelligence*, vol. 26, no. 1, pp. 405–416, 2013.

[5] C. Sevgi and A. Koçyiğit, "Optimal deployment in randomly deployed heterogeneous wsns: A connected coverage approach," *Journal of Network and Computer Applications*, vol. 46, pp. 182–197, 2014.

[6] A. Akbas, H. U. Yildiz, and B. Tavli, "Data packet length optimization for wireless sensor network lifetime maximization," in *2014 10th International Conference on Communications (COMM)*. IEEE, 2014, pp. 1–6.

[7] O. G. Uyan, A. Akbas, and V. C. Gungor, "Machine learning approaches for underwater sensor network parameter prediction," *Ad Hoc Networks*, vol. 144, p. 103139, 2023.

[8] A. Akbas, H. U. Yildiz, A. M. Ozbayoglu, and B. Tavli, "Neural network based instant parameter prediction for wireless sensor network optimization models," *Wireless Networks*, vol. 25, no. 6, pp. 3405–3418, 2019.

[9] "Gams," https://www.gams.com/products/gams/gams-language/, 12 2021, (Accessed on 12/12/2021).

[10] "Log distance path loss or log normal shadowing model - gaussianwaves," https://www.gaussianwaves.com/2013/09/log-distance-path-loss-or-log-normal-shadowing-model/, 12 2021, (Accessed on 12/12/2021).

[11] A. Akbas, H. U. Yildiz, B. Tavli, and S. Uludag, "Joint optimization of transmission power level and packet size for wsn lifetime maximization," *IEEE Sensors Journal*, vol. 16, no. 12, pp. 5084–5094, 2016.

[12] "Matlab - mathworks - matlab & simulink," https://www.mathworks.com/products/matlab.html, 12 2021, (Accessed on 12/12/2021).

[13] M. Gentil, A. Galeazzi, F. Chiariotti, M. Polese, A. Zanella, and M. Zorzi, "A deep neural network approach for customized prediction of mobile devices discharging time," in *GLOBECOM 2017-2017 IEEE Global Communications Conference*. IEEE, 2017, pp. 1–6.

[14] C.-H. Zhu and J. Zhang, "Developing soft sensors for polymer melt index in an industrial polymerization process using deep belief networks," *International Journal of Automation and Computing*, vol. 17, no. 1, pp. 44–54, 2020.

[15] H. Lee, R. Grosse, R. Ranganath, and A. Y. Ng, "Unsupervised learning of hierarchical representations with convolutional deep belief networks," *Communications of the ACM*, vol. 54, no. 10, pp. 95–103, 2011.

[16] Y. Qin, X. Wang, and J. Zou, "The optimized deep belief networks with improved logistic sigmoid units and their application in fault diagnosis for planetary gearboxes of wind turbines," *IEEE Transactions on Industrial Electronics*, vol. 66, no. 5, pp. 3814–3824, 2018.

[17] M. Hossin and M. N. Sulaiman, "A review on evaluation metrics for data classification evaluations," *International journal of data mining & knowledge management process*, vol. 5, no. 2, p. 1, 2015.

[18] D. Justus, J. Brennan, S. Bonner, and A. S. McGough, "Predicting the computational cost of deep learning models," in *2018 IEEE international conference on big data (Big Data)*. IEEE, 2018, pp. 3873–3882.

BIOGRAPHIES

Ayhan Akbas holds a Bachelor of Science and a Master of Science in Electrical and Electronics Engineering, both obtained from the Middle East Technical University in 1991 and 1995, respectively. He earned his doctoral degree from the Department of Computer Engineering at TOBB University of Economics and Technology. With over 20 years of experience in the field of Information Technology, Dr. Akbaş has occupied both technical and managerial roles at multinational corporations, including SIEMENS, Sun Microsystems, and NEC. Currently, he is pursuing his career as Asst. Professor at Abdullah Gül University. His areas of interest are Wireless Sensor Networks, IoT, 5G, Wireless Communication, and Machine Learning.

Selim Buyrukoglu is an Asst. Prof. in Computer Engineering Department at Cankiri Karatekin University. He completed his PhD at Loughborough University, UK in 2019, He received an MSc degree in Advance Computer Science in 2014 from Leicester University, UK and also a BSc degree in Computer Engineering in 2010 from European University of Lefke, TRNC. He is particularly focusing on the application of Machine Learning and Deep Learning methods on interdisciplinary subjects.

LSTM-based approach for Classification of Myopathy and Normal Electromyogram (EMG) Data

Erdem Tuncer and Emine Dogru Polat

Abstract— **Electromyograms (EMG) are recorded movements of nerves and muscles that help diagnose muscles and nerve-related disorders. It is frequently used in the diagnosis of neuromuscular diseases such as myopathy, which causes many changes in EMG signal properties. The most useful auxiliary test in the diagnosis of myopathy is EMG. Therefore, it has become imperative to identify computer-assisted anomalies with full accuracy and to develop an efficient classifier. In this study, a new machine learning method with a deep learning architecture that can score normal and myopathy EMG from the EMGLAB database is proposed. Using the discrete wavelet transform Coiflets 5 (Coif 5) wavelet, the EMG signals are decomposed into subbands and various statistical features are obtained from the wavelet coefficients. The success rates of the decision tree C4.5, SVM and KNN-3, which are traditional learning architectures, and the Long Short-term Memory (LSTM) algorithm, which is one of the deep learning architectures, were compared. Unlike the studies in the literature, with the LSTM algorithm, a 100% success rate was achieved with the proposed model. In addition, a real-time approach is presented by analyzing the test data classification time of the model.**

Index Terms— **Deep learning, EMG, Myopathy, Neuromuscular disorder, Wavelet.**

I. INTRODUCTION

MYOPATHY IS a critical type of neuromuscular disorder called neurological disorders in skeletal muscle [1]. It is a disease caused by the improper functioning of muscle fibers [1, 2]. In the later stages, it even affects the respiratory muscles and makes life difficult. Both the disturbances in the muscle cells and the disturbances in the nerve cells that stimulate these cells cause changes in certain symptoms. However, it also makes changes in Electromyogram (EMG) signals.

ERDEM TUNCER, Bahcecik Vocational and Technical Anatolian High School, Kocaeli, Turkey,(e-mail: erdemtuncerr@gmail.com)

https://orcid.org/0000-0003-1234-7055

EMINE DOGRU BOLAT, Faculty of Technology, Biomedical Eng. Department of Kocaeli University, Turkey, (e-mail: ebolat@gmail.com).

https://orcid.org/0000-0002-8290-6812

EMG is a method that records the electrical activity occurring in muscle cells, that is, a large number of action potentials occurring in muscle tissue. Factors such as the variation of muscle structures from person to person and the degree of disease complicate the diagnosis of myopathy, which acts as a healthy EMG signal. Artificial intelligence algorithms are used to overcome these difficulties [2, 3]. Advanced artificial intelligence techniques have recently provided convenience in applications that will facilitate diagnosis. Because these techniques quickly summarize detailed and accurate information, they have allowed the development of tools to aid in diagnosis [5].

In the literature, there are various studies that include the classification of signals such as normal, myopathy and neuropathy from EMG data related to the diagnosis of neuromuscular diseases. Myopathy and Amyotrophic Lateral Sclerosis (ALS) signals in the EEGLAB database were classified by Bakiya A. et al. [6]. They applied the selected features to the deep neural network and artificial neural network using the bat algorithm. They compared the results of the classifier algorithms. In the study with the traditional single-layer artificial neural network, they reached an accuracy rate of 83,3%. Using the deep neural network modeled with layers 2 and 3 (neurons = 2 and 4), they achieved a 100% success rate in classifying abnormalities in EMG signals. Belkhou A. et al. [7] presented a new method for classifying normal and myopathy EMG signals in the EMGLAB database. They analyzed EMG signals with Symlet 6 (Sym6) wavelet using Continuous Wavelet Transform (CWT). Five features were obtained with CWT. Support Vector Machine (SVM), K-Nearest Neighbors (KNN), Decision Tree (DT), Discriminatory Analysis (DA) and Naive Bayes (NB) were used as classifier algorithms. It was revealed in the study that the KNN classifier reached the best performance with an accuracy of 93,68%. Patidar M. et al. [8] classified normal and myopathic EMG data with data obtained from Beth Israel Deaconess Medical Center. The data is sampled at 50 KHz. 60% of the data was used for training and 40% for testing. They achieved 96,75% success rate with the Neural Network Classifier. Jose S. et al. [9] classified healthy, myopathy and neuropathy EMG data using the EMGLAB database. They produced subsampled signals from the EMG signal using Lifting Wavelet Transform (LWT). They calculated Higuchi's fractal dimensions of the LWT

coefficients in the lower bands. They used 10-fold cross validation. They achieved success rate of 99,87% using a combination of Multilayer Perceptron Neural Network (MLPNN) and Boyer-Moore majority vote (BMMV). Normal, myopathic and ALS data were classified by Belkhou A. et al. [10]. The features were obtained using the Mel Frequency Cepstral Coefficients (MFCC) technique. The size of MFCC vectors has been reduced using statistical values. SVM and KNN algorithms were used as classifiers. According to the F-measure evaluation metrics, the performances of the classifiers were evaluated. Normal-ALS and Normal-Myopathy binary classifications had 99,34% and 99,07% accuracy rates, respectively.

When the literature studies are examined; It is seen that traditional learning algorithms are mostly used in the classification of EMG data. For this reason, deep learning architecture is included in this study and high success rates are aimed. At the same time, traditional classification algorithm and deep learning method are compared. The continuation of the article is organized as follows: In the second part, information about the EEG data sets to be used and the methods and techniques used are given. In the third part, the results of the analysis of the article and the studies in the literature are discussed. In the fourth part, the studies that are planned to be done in the future according to the results of the analysis are included.

II. MATERIALS AND METHODS

In this section, the methods and techniques used for feature extraction and classification from EMG data will be explained.

A. Subjects and data acquisition

The data were taken from the database that can be accessed free of charge on the EMGLAB.net website. There are 10 healthy and 6 myopathy patients' data in the database. Healthy people are between the ages of 21-37. There are no signs of neuromuscular disorders in the normal group of EMG data. People with myopathy are between the ages of 19-63. Myopathic patients have all the electrophysiological and clinical manifestations of myopathy. Although there are measurements from various muscles in the database, only the measurements taken from the Brachial biceps muscle were used to ensure homogeneity in this study. During the measurements, a concentric needle electrode was used and the low and high pass filter cut-off frequencies of the EMG recording device were adjusted to the range of 2Hz-10Khz. A total of 1196,5 seconds of EMG data were obtained from healthy subjects, all taken from the long head of the brachial biceps muscle. The number of measurements taken from this muscle group in the database is 3355,4 seconds for the myopathy group. The sampling frequency of the data is 23437,5 Hz [11].

B. Preprocessing

Various filter structures are used in order to remove unwanted electrical activities seen in the channel during recording, such as DC component, biological artifact and network noises that directly affect the signals [12, 13]. In this article, EMG signals were passed through a notch filter (50 Hz.) and a third order

band stop Butterworth filter. The filter degree was chosen as three by trial and error process.

C. Feature extraction

In today's conditions, the concept of data is important. The amount of data is increasing day by day, so the concept of data has an important place. However, the raw form of the data is not a meaningful sum of information. In order to become meaningful, it must go through several processes. One of these processes is feature extraction. The raw data space may contain redundant elements, and the data can be difficult to manipulate due to its size. Therefore, feature extraction is provided in most machine learning applications [14,15]. In this study, the feature vector was calculated for each epoch with the discrete wavelet method and all feature vectors were combined to form the feature matrix. This obtained feature matrix is given as input to the classifier algorithms for classification.

D. Discrete wavelet transform (DWT)

Wavelet transform is a tool whose functions separate the data into its components at different frequencies and allow to work on each component separately [16]. Wavelet transform has been used frequently in the fields of mathematics and biomedical engineering recently. Figure 1 shows the 3-levels wavelet tree. H(n), g(n) generate wavelet coefficients at each stage. Approximation (A) is called approximation coefficients, Detail (D) is called detail coefficients. For the spectral analysis of non-stationary EMG signals, feature extraction was performed with wavelet transform, which aims to provide the best time-frequency resolution by using small size window at high frequencies and large size window at low frequencies [17,18].
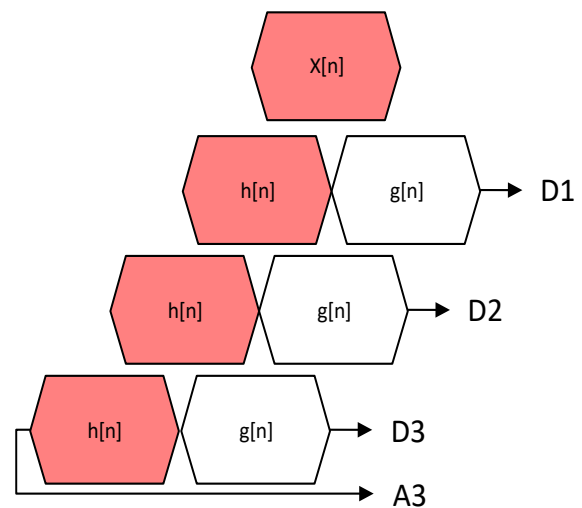

Fig.1. Wavelet decomposition tree with 3-levels

The most important parameter of the wavelet transform is the wavelet. The term wavelet is expressed as a wavelet defined as a window function of a certain length. There are many main wavelets with different properties and uses [19]. The drawing of the Coiflets (Coif5) wavelet form used in this study is given in Figure 2 [20].
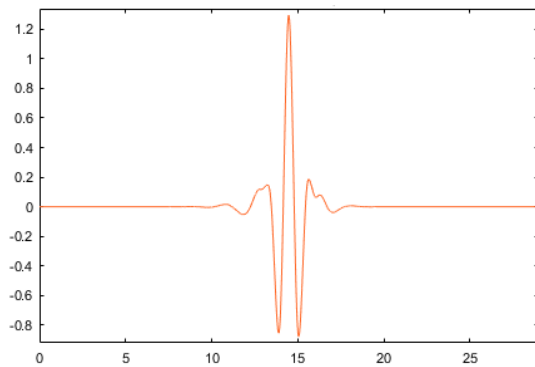
Fig.2. Plots of Coif5 wavelet function

E. Classification

In addition to extracting useful information from the data, it is also important to classify this information correctly. Classifier algorithms can make rational decisions in a very short time by evaluating all situations and examining millions of data very quickly. Millions of data bring great problems due to diversity, speed and volume [21]. Classifier algorithms prevent these problems by properly classifying big data [22]. In this study, traditional classification algorithms C4.5 decision tree, SVM, KNN-3 and deep learning algorithms LSTM were used.

1) C4.5 decision tree

Decision trees are one of the machine learning algorithms used for classification and value estimation. Many approaches have been proposed to construct decision trees. C4.5 decision trees method, which is one of these approaches, is frequently used in many areas. A decision tree structure consists of root, node, branch and leaf. The lowest part of the tree structure is called the leaf and the upper part is called the root. Each feature in the data set represents the nodes. The part that provides the connection between the nodes is called the branch. In this algorithm, the entropy of all columns in the data set is calculated according to Equation 1. Next, the entropy of each column in the dataset is divided by the entropy of the class to calculate the gain for each column. Finally, the information gain is calculated for each predictive variable/class. The root node with the highest gain is assigned as [23-26].

$$Entropi = -\sum_{i=1}^{n}(p_i log_2 p_i) \tag{1}$$

Where, n is the number of values that the target variable can take.

After the root node, the tree begins to branch. Thus, the data will be evenly distributed under each branch. After the first predictor variable is determined, the same process is calculated by repeating. This process continues until all predictive variables are placed in the tree [23-26].

2) Long short-term memory

Long short-term memory (LSTM), a sub-branch of recurrent neural networks (RNN), was developed to eliminate the problems in RNN. There are three gates in an LSTM structure: input, output and forget. The tasks of these ports are write, read and reset respectively. Changes in cell states are controlled by the three gates described. The task of the entrance gate is to control the information to be added to the memory, the task of the forget gate is to control how much of the old information will be transferred to the new data, and the output gate is to control how much of the information in the memory will be used at the output stage [27,28]. Figure 3 shows the structure and gates of the LSTM cell. Here, the input data at time t is xt and the output data from the previous cell is ht. The input at time t, xt, and the output of the previous step, ht-1, arrive at the forget gate. According to these values, forget gate makes a decision according to xt input and ht-1 output. The amount of information from the previous cell state information is checked in the cell at the current time t. At the input gate, it is decided how much of the newly incoming information as ht-1 and xt from the previous cell will be used in the memory cell. It can be understood from 0 and 1 outputs whether this information will be used or not. The output port, on the other hand, decides whether there will be an output. The output port also takes the inputs ht-1 and xt and decides whether to output 0 or 1 as in the input port. The state of the cell at time t is the sum of the information at t-1 and t times when it passes to the next cell (t+1 moment) [28,29].

3) Support Vector Machine

SVM is a pattern recognition method proposed by Vladimir Vapnik in the 1960s and whose algorithm was developed by Cortes and Vapnik in 1999 [30, 31]. The main principle of this method is to find the hyperplanes in the feature space that will allow to distinguish the two classes in the most appropriate way. It is aimed to complete the classification process by finding the most optimal hyperplane. In general, signals of non-physiological origin can be separated by a linear hyperplane, but it may be difficult to separate signals of physiological origin with a linear hyperplane. In order to overcome this difficulty, the data is transformed into a high-dimensional feature space [32].

4) KNN

KNN is one of the supervised data mining algorithms that classifies according to the distance between objects. This method makes the classification according to the proximity calculation. Its basic principle is based on the idea that objects that are close together in the sample space probably belong to the same category. The application steps of this method are given below.
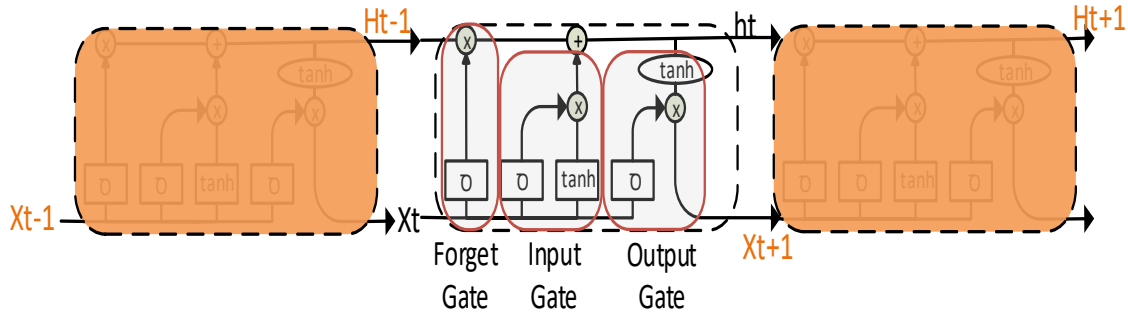
Fig.3. LSTM cell structure and gates

➢ The distance of the new observation to all the observations in the data set is calculated.
➢ The distance values are ordered.
➢ The k observations with the smallest distance are selected.
➢ In k observations, the majority voting category creates the class value [33].

III. RESULTS

This article presents a new machine learning algorithm for EMG signal analysis and classification. The flow diagram of the proposed algorithm is given in Figure 4. After the data was transferred to the Matlab environment, it was passed through a 3rd order band stop Butterworth filter (50 Hz) as a preprocessing step. After the preprocessing step, the EMG signal was subjected to a one-second windowing process to avoid overlapping and various features (19 features) were extracted in different frequency bands using DWT for each window. The obtained feature matrix was given as input to the classifier algorithms (C4.5, LSTM, KNN-3, SVM) and the results were discussed. Typical examples of EMG signals for the two groups (normal and myopathic) used in the article study are given in Figure 5.

After the preprocessing step, windowing was done and various features were obtained for each window. The obtained feature matrix was given as an input to the classifier algorithms and the results were discussed. Typical EMG signal examples for the two groups (normal and myopathic) used in the article study are given in Figure 5.

In the study, measurements taken from the Brachial biceps muscle in the EMGLAB database were used. The data used are given in Table 1. Biceps Brachii muscle data was not included in the M01 dataset, so it was not included in the study. In the study, the last 1,8 seconds of M07 myopathy data and the last four seconds of C10 Normal EMG data were excluded from the study for adjusting window sizes.

Features are obtained to classify each window data. The discrete wavelet method (Coif5 wavelet) was used to obtain the features. The lower frequency bands of the DWT used in this study are shown in Figure 7. With 10-levels wavelet transform, EMG signals are reduced to lower frequency bands and approximation and detail coefficients are calculated for each 1 second EMG data. Figure 7 shows the 10-levels DWT tree for obtaining the features. Various statistical features are obtained from all coefficients D1-A10.

The EMG data used in this study are divided into training dataset and test dataset. The training data represents 80% of the randomly selected dataset from the total data and is used to construct the classification model. Test data representing 20% of the total data are used to evaluate classifier performance. The distribution of records in the datasets for each class is presented in Table 2.
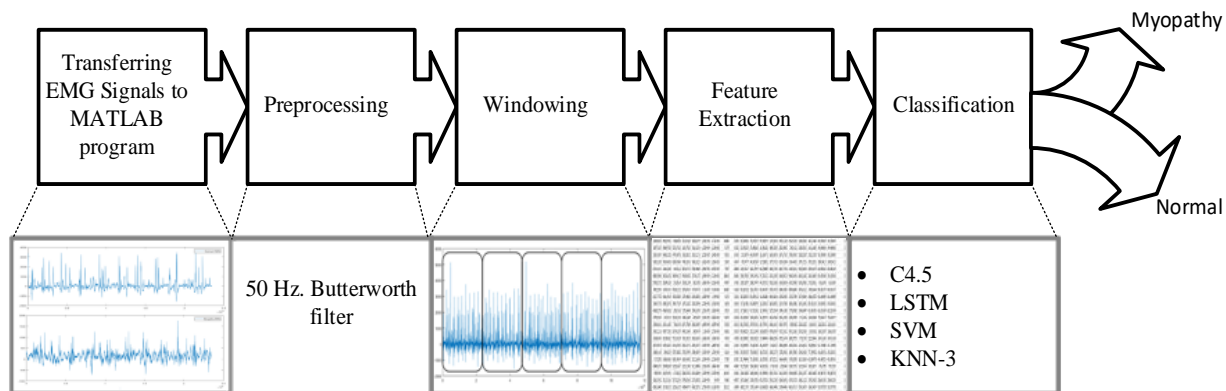
Fig.4. Schematic diagram of the signal processing

(a)

(b)

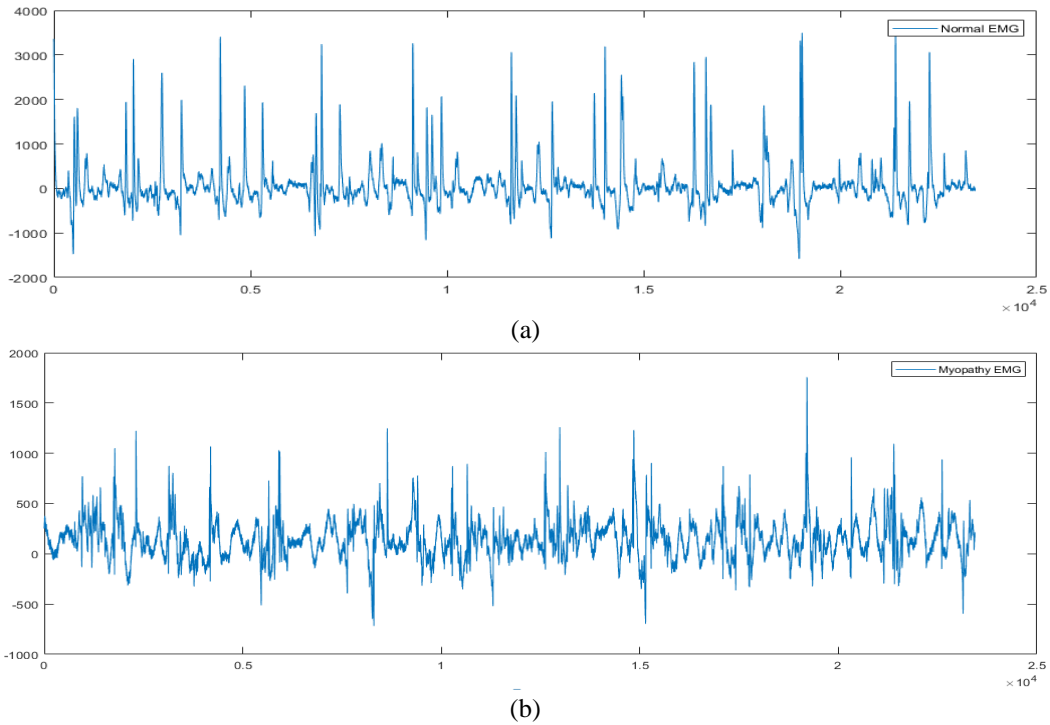Fig.5. (a) Normal EMG (b) Myopathy EMG

TABLE I
EMG DATASET USED IN THE STUDY

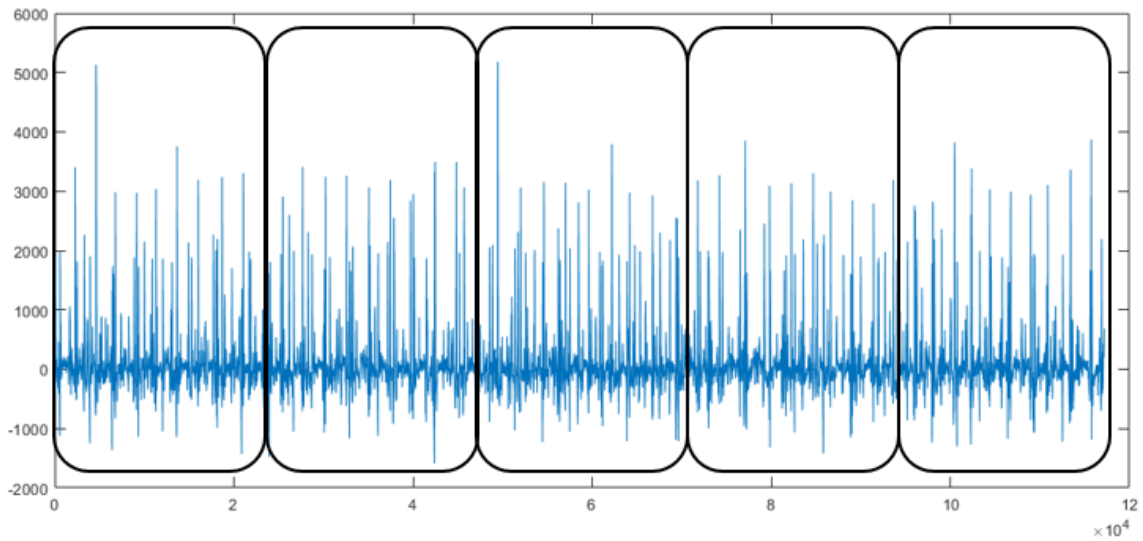| Types	Patient	Number of patient	Type of Muscle	Time (sn)
Myopathy	M02, M03, M04, M05, M06, M07.	6	Biceps Brachii (long head)	1196,8
Normal	C01, C02, C03, C04, C05, C06, C07, C08, C09, C10.	10	Biceps Brachii (long head)	3355,4

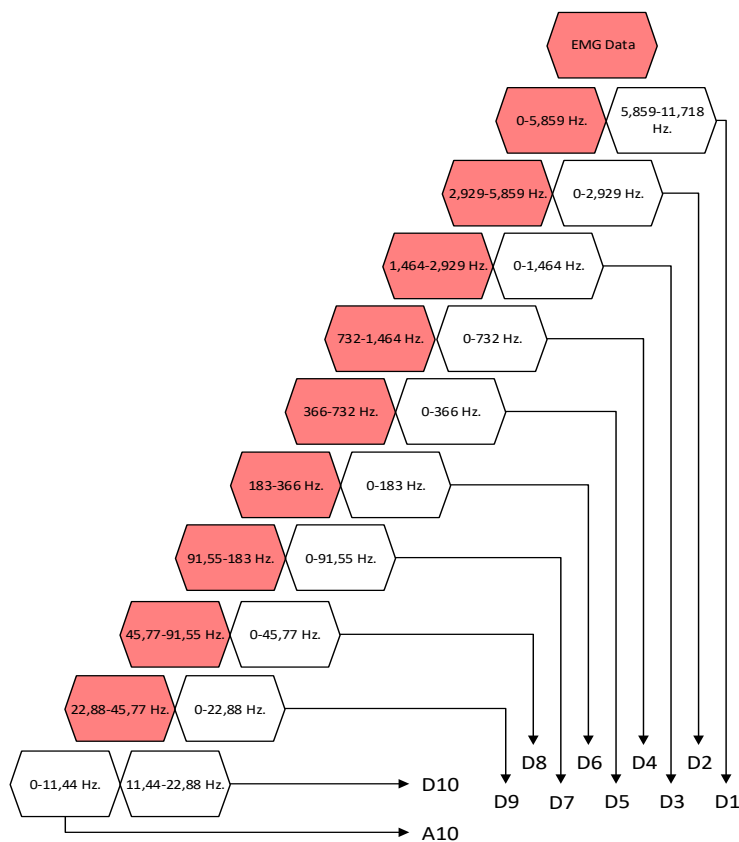Fig.6. Dividing EMG data into 1-second epochs

Fig.7. Wavelet decomposition tree with 10-levels

A10, D1, D2, D3, D4, wavelet coefficient variance value, energy value of all wavelet coefficients, maximum and minimum values of A10 wavelet coefficient, standard deviation values of D1, D2, D3, D4, D5, D6, D7, D8, D9, D10, A10 wavelet coefficients were calculated from the wavelet coefficients obtained from the EMG data in each epoch. A total of 19 features were extracted. The numbers 0 for normal EMG signals and 1 for myopathic EMG signals are set for the classifier response. The details of the obtained features are given in Table 3.

TABLE II
DISTRIBUTION OF RECORDS IN THE TWO DATASETS

Class	Training set (80%)	Test set (20%)	Total
Myopathy	956	239	1195
Normal	2684	671	3355
Total	**3640**	**910**	**4550**

As a result of the training carried out for the classification of healthy and myopathic signals, high success rates were achieved with the C4.5 decision tree and LSTM algorithms. In this article, the Confidence Factor (CF) coefficient used in the C4.5 decision tree algorithm was determined as 0,2, number of leaves: 7, size of the tree: 13, batch size: 50 in order to more effectively use the post-pruning process that prevents over-learning. Figure 8 shows the decision tree result screen for myopathy disease prediction of the C4.5 algorithm.
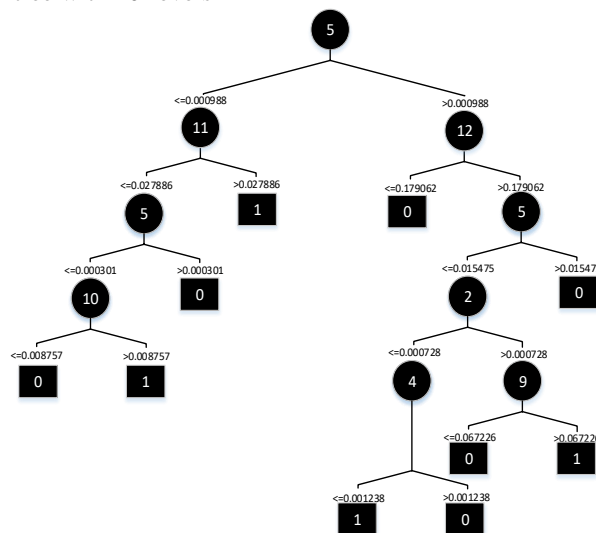
Fig.8. Decision tree created as a result of the study

According to Figure 8, if feature number 5 is greater than 0,000988 and feature number 12 is less than 0,179062, class 0 is assigned. If feature 5 is less than 0,000988 and feature number 11 is greater than 0,027886, class label 1 is assigned. Similarly, other situations occurring in the decision tree are given in Figure 8.

4-layer model (LSTM, dropout, activation and output layer) has been created for the LSTM algorithm. Algorithm parameters; optimizer: Adaptive Moment Estimation (ADAM), dropout rate: 0,2, batch size: 50, epochs: 50, 8 memory unit in LSTM

layer. Learning rate of 0,001 was used with the ADAM optimizer.

The kernel function used in this study for the SVM algorithm is PUK. The mathematical representation of the PUK kernel function is given in Equation (2).

$$f(x) = \frac{H}{[1+\left(\frac{2(X-X_O)\sqrt{2\left(\frac{1}{W}\right)^{-1}}}{\sigma}\right)^2]^w} \tag{2}$$

The w ve σ parameters indicate the width of the function, and the H indicates the maximum peak level between the peak values. The Euclidean function was used as the distance function for the KNN algorithm used in this study, which determines which class the data belong to. The mathematical representation of the Euclidean function is given in Equation (3). In the study, k=3 neighbourhood was used.

$$d_{(i,j)} = \sqrt{\sum_{k=1}^{p}(X_{ik} - X_{jk})^2}$$
(3)

The confusion matrix resulting from the training of both classifiers is given in Table 4 and Table 5.

TABLE III
FEATURE LIST OBTAINED FROM EMG DATA

No.	Wavelet coefficient	Feature Name	No.	Wavelet coefficient	Feature Name
1	A10	Variance	11	D6	Standard deviation
2	D4	Variance	12	D7	Standard deviation
3	D3	Variance	13	D8	Standard deviation
4	D2	Variance	14	D9	Standard deviation
5	D1	Variance	15	D10	Standard deviation
6	D1	Standard deviation	16	A10	Standard deviation
7	D2	Standard deviation	17	All coefficients	Energy
8	D3	Standard deviation	18	A10	Maximum value
9	D4	Standard deviation	19	A10	Minimum value
10	D5	Standard deviation		Classifier Response	[0,1]

The calculation results from this confusion matrix are given in Table 4 and Table 5. Different metrics such as accuracy, precision, recall, Receiver Operating Characteristic (ROC) area were calculated to measure the success criteria of the models. While the time taken to classify data in deep learning architecture was 1.25 seconds, it was measured as 0,01 seconds (SVM) in traditional learning algorithms. Calculations were made on a computer with "Intel(R) Core(TM) i3-6006U CPU @ 2.00GHz and 4GB ram". Accuracy value is calculated by the ratio of correctly predicted data in the model to the total data set. Precision is the metric that shows how many of the data predicted as positive are actually positive. Recall is called as the metric that shows how much of the transactions we need to predict positively [34]. The ROC area is defined as the True Positive Rate / False Positive Rate ratio. It is basically a metric

to understand whether the models established to solve the classification problems are working well. As the ROC value approaches 1, it means that the success of the model established to distinguish the classes increases [35]. According to Table 4, with the decision tree model, the accuracy rate was 98%, precision 98,20%, recall 98%, ROC value 0,99. With the LSTM model, the accuracy rate, precision and recall values were 100%, and the ROC value was 1.

According to Table 6, LSTM algorithm achieves more successful results, other algorithms It has been found that the performance is close to the LSTM model. As a result of the applications within the scope of the study, the deep learning architecture model, which has a 100% success rate, has been suggested for the classification of myopathy diagnosis because it has higher accuracy. At the same time, by examining the data in one-second windows, more data was classified separately, and the classification resolution was kept higher. When the data of Table 6 models are analyzed in terms of classification time; Traditional learning architectures achieved faster results than deep learning architecture. However, considering the classification time together with the success rate, we consider that 1,25 seconds is a feasible time for real-time applications. This study will shed more light on the investigation of deep learning architectures with high accuracy rates in future studies with EMG data. In addition, minimizing the need for physical consultation of the patients, blood tests etc. It will allow the design of a system or algorithm with high accuracy using only EMG signals without the need for operations.

Table IV
Confusion matrix (C4.5 and LSTM)

Classifier					
C4.5 (80% train, 20% test)			LSTM (80% train,20% test)		
Class Label	0	1	Class Label	0	1
0	688	18	0	706	0
1	0	204	1	0	204

Table V
Confusion matrix (SVM and KNN-3)

Classifier					
SVM (80% train, 20% test)			KNN-3 (80% train,20% test)		
Class Label	0	1	Class Label	0	1
0	703	3	0	699	7
1	1	203	1	8	196

Table VI
Success criteria and classification time of models

Classifier	Accuracy (%)	Precision (%)	Recall (%)	ROC Area	Classification time (sn.)
C4.5	98	98,20	98,00	0,99	0,03
LSTM	100	100	100	1	1,25
SVM	99,56	99,60	99,60	1	0,01
KNN-3	98,35	98,30	98,40	0,99	0,02

IV.   DISCUSSION

From studies in the literature, Bue DB et al. [36] presented a methodology to predict the presence of myopathy from EMG signals in their study. They classified myopathic disease with an average success rate of 90%. Data were collected in the EMG lab. at the Baylor College of Medicine Department of Neurology in Houston, TX. Collected by James Killian. 15

EMG recordings were taken from 8 different subjects measured in one or more different muscles. EMG data were classified with SVM using Fourier transform. The performance of the model was measured at intervals of 0,05-2 seconds of the window duration. It was stated that the success rate increased with the increase of the window duration.

From studies using EMGLAB data; Belkhou A. et al. [7] studied myopathy disease classification. They obtained 5 features (the mean scale, the median scale, the mean coefficient, the minimum coefficient, the maximum coefficient) using the CWT together with the Sym6 wavelet. As a result of the classification made with SVM, KNN, DT, DA and NB algorithms, they reached a success rate of 93,68% with the KNN classifier algorithm. Belkhou A. et al. [37] have classified myopathy disease. 4 features (mean coefficient, minimum coefficient, mean scale, median scale) were extracted using CWT. As a result of the classification using KNN and SVM classifier algorithms, they achieved a success rate of 91,11% with KNN-9. ALS and Myopathy were classified by Bakiya A. et al. [6]. They obtained features in the time domain (seventeen time features) and Wigner-Ville transformed time-frequency domain (nineteen time-frequency features) from the data. Feature selection was made using the bat algorithm. The performance of the deep neural network is compared with the traditional neural network. It has been demonstrated that the deep neural network modeled with layers 2 and 3 (neurons = 2 and 4) using time domain features classifies abnormalities of EMG signals with higher accuracy. Dubey R. et al. [2] classified myopathy, ALS and normal EMG data. EMG signals were decomposed by the Empirical Mode Decomposition (EMD) method. The proposed methodology was tested on a dataset of more than 900 EMG signals from three classes. Empirical mode decomposition method is applied to decompose EMG signals. Appropriate intrinsic mode functions for feature selection are selected using the t-test-based approach and a complex plane graph is created. The proposed algorithm is trained and validated using Feed Forward Neural Network (FFNN), SVM and DT. When the algorithm was tested with FFNN, a maximum classification accuracy of 99,53% was achieved. Torres-Castillo JR. et al. [3] classified myopathy, ALS, and normal EMG data. Using the Hilbert Transform, 234 features are extracted in the time-frequency domain. Non-parametric statistical analysis and unrelated linear discrimination analysis were used for feature selection. After feature selection, 103 features were given to the classifier. The data were classified with the KNN classifier and a success rate of 99,4% was achieved.

Table 7 compares the results of this study with the results of other studies conducted with various data sets and classifiers in the literature. Bue DB. [36] achieved 90% success rate with the SVM classifier using the Fourier transform. Belkhou A. [7] achieved a success rate of 96,68% by using KNN, one of the traditional learning architectures, in their 2019 study.

Table VII
The proposed methodology and other exiting methodologies are compared

Work	Dataset	Method	Classifier	Acc.	Validation method	Classes
Bue DB. et al. [36]	Department of Neurology of the Baylor College of Medicine	Fourier transform	SVM	90%	10 fold cross validation	Normal and myopathy
Belkhou A. et al. [7]	EMGLAB	CWT	SVM, KNN, DT, DA, NB	93,68%	75%, 25%	Normal and myopathy
Belkhou A. et al. [37]	EMGLAB	CWT	SVM, KNN	91,11%	10 fold cross validation	Normal and myopathy
Bakiya A. et al. [6]	EMGLAB	Statistical features in the time and frequency domain	Developed deep neural network, conventional artificial neural network	100%	80%, 20%	ALS and myopathy
Dubey R. et al. [2]	PhysioBank, EMGLAB	EMD	FFND, SVM, DT	99,53%	85%, 15%	Myopathy, ALS and normal
Torres-Castillo JR. et al. [3]	EMGLAB	Ensemble-EMD	Latent Dirichlet allocation, DT, KNN	99,50%	3 fold cross validation	Myopathy, ALS and normal
Jose S. et al. [9]	EMGLAB	LWT	MLPNN-BMMV	99,87%	10 fold cross validation	Myopathy, ALS and normal
This work	**EMGLAB**	**DWT**	**LSTM,** DT (C4.5), SVM, KNN	**100%**	**80%, 20%**	**Normal and myopathy**

Belkhou A. et al. [37] achieved a success rate of 91,11% in the classification study. Jose S. et al. [9] classified the data with a success rate of 99,87%. Bakiya A. et al. [6] performed ALS and Myopathy classification in their study and achieved a 100% success rate. Dubey R. et al. [2] achieved a 99,53% success rate in myopathy, ALS and normal EMG classification. Similarly, Torres-Castillo JR. et al. [3] achieved a high success rate of 99,50% in the study, but too many (103) features were used in the study. This will naturally increase the processing load in a real-time application. When the studies in the literature are examined in terms of methods and techniques, in the study by Bue DB et al. [36], fast fourier transform was used and the signal was examined in the frequency domain. The performance of the model was tested with the classifier algorithm. In the study conducted by Belkhou A. et al. [7], it was concluded that the continuous wavelet method can be used on EMG data. Belkhou A. et al. [37] calculated the average absolute coefficient to reduce the size of the CWT coefficients. Thus, they tried to solve the size problem in CWT. Bakiya A. et al. [6] tried to improve the classification performance by extracting features in the time-frequency domain. A bat algorithm is used to select the best features from the time and time-frequency feature sets extracted by the Wigner-Ville transform. They concluded that the Wigner-Ville transform and the bat algorithm are suitable for classification of myopathy and ALS signals. Dubey R. et al. [2] using EMD method arranged the signals from high frequency to low frequency components and decomposed Intrinsic Mode Functions (IMF)'s. The features were obtained using the hilbert transform of the IMFs. The obtained features were classified by various classifier algorithms. They measured the statistical significance of the IMFs using the rule-based learning proposition (statistical values such as t, h, and p were calculated) in the feature

extraction stage. Torres-Castillo JR. et al. [3] separated all signals into amplitude or frequency modulated subbands and extracted time-frequency features using Hilbert transform. Due to the large number of features, feature selection was made using linear discriminant analysis. Jose S. et al. [9] decomposed the signals using LWT. They calculated Higuchi's fractal dimensions (FD) of the LWT coefficients in the separated signals. The FDs of the LWT subband coefficients are combined in one dimension and given as input to the classifier algorithms. As stated above, studies have focused on high success rates. At the same time, the performances of different algorithms were compared and optimum models were proposed. In our study, lower frequency analysis of EMG signals was performed similar to the studies in the literature. Unlike the studies in the literature, wavelet method was preferred to examine the signal in lower frequency bands and features were obtained from the wavelet coefficients representing the signal form.

With this proposed study, the deep learning architecture, which is not available in most other studies, is tested on EMG data. At the same time, high success rates were achieved with 19 features. In addition, the results were compared with the traditional learning architecture. A better classification success rate (100%) was obtained in the diagnosis of myopathy with the deep learning architecture than the studies in the literature. According to the results obtained, the deep learning architecture LSTM model (4 layers: LSTM, dropout, activation and output layer) is proposed as the optimum model. At the same time, the parameters of the proposed model were determined as Optimizer: ADAM, learning rate: 0,001, dropout rate: 0,2, batch size: 50, epochs: 50, 8 memory units in a single layer.

V. CONCLUSION

The EMG signal is a non-linear, noisy signal. Therefore, it is difficult to distinguish various diseases from the EMG signal. In this study, a classifier algorithm from both traditional and deep learning architectures was used to diagnose neuropathy from the EMG signal. Numerous studies have classified neuropathy using traditional learning architectures. The proposed method has been effective in diagnosing neuropathy. Before extracting features from the EMG signals, the data were decomposed using the Coif5 wavelet. 19 features are obtained from wavelet coefficients and given as input to classifier algorithms. C4.5 decision tree, SVM and KNN-3 from traditional learning methods and LSTM model from deep learning method were used as classifiers. In this study, the focus is on achieving high accuracy rates. A higher success rate was obtained with the LSTM model compared to traditional learning methods. At the same time, the performance of traditional classifier algorithms is compared with deep learning architectures and the optimum model is proposed. A new contribution has been made to the literature with the application of deep learning architecture to the EMG dataset. It is considered important by us to be able to diagnose neuropathy using only EMG signals. It is possible to test this research article by increasing the number of data and classes. In future studies, the classification results of deep learning architectures can be evaluated by expanding the number of classes and data

sets. Thus, it will be possible to provide more detailed and accurate information that can help medical professionals with machine learning methods.

REFERENCES

[1] Frederikse A., "The role of different EMG methods in evaluating myopathy." Clinical Neurophysiology 2006, 117(6), 1173-1189.

[2] Dubey R., Kumar M, Upadhyay A, Pachori RB. "Automated diagnosis of muscle diseases from EMG signals using empirical mode decomposition based method." Biomedical Signal Processing and Control 2022, 71, 103098.

[3] Torres-Castillo J.R., López-López C.O., Padilla-Castañeda M.A., "Neuromuscular disorders detection through time-frequency analysis and classification of multi-muscular EMG signals using Hilbert-Huang transform." Biomedical Signal Processing and Control 2022, 71, 103037.

[4] Bentick G., Fairley J., Nadesapillai S., Wicks I., Day J., "Defining the clinical utility of PET or PET-CT in idiopathic inflammatory myopathies, A systematic literature review." Seminars in Arthritis and Rheumatism 2022, 57, 152107.

[5] Kukker A., Sharma R., Malik H. "Forearm movements classification of EMG signals using Hilbert Huang transform and artificial neural networks." IEEE 7th Power India International Conference (PIICON) 2016, 1-6.

[6] Bakiya A., Anitha A., Sridevi T., Kamalanand K., Classification of myopathy and amyotrophic lateral sclerosis electromyograms using bat algorithm and deep neural networks. Behavioural Neurology 2022, 3517872.

[7] Belkhou A., Achmamad A., Jbari A., "Myopathy detection and classification based on the continuous wavelet transform." Journal of Communıcatıons Software and Systems 2019, 15(4), 336-342.

[8] Patidar M., Jain N., Parikh A., "Classification of normal and myopathy EMG signals using BP neural network." International Journal of Computer Applications 2013, 69(8), 12-16.

[9] Jose S., George S.T., Subathra MSP, Handiru VS. "Robust classification of intramuscular EMG signals to aid the diagnosis of neuromuscular disorders." Engineering in Medicine and Biology 2020, 1, 235-242.

[10] Belkhou A., Jbari A., Badlaoui O.E., "A computer-aided-diagnosis system for neuromuscular diseases using mel frequency cepstral coefficients." Scientific African 2021, 13, e00904.

[11] Nikolic M., "Detailed analysis of clinical electromyography signals EMG decomposition, findings and firing pattern analysis in controls and patients with myopathy and amyotrophic lateral sclerosis," PhD Thesis 2021. Faculty of Health Science, University of Copenhagen.

[12] Hurtik P., Molek V., Hula H., "Data preprocessing technique for neural networks based on ımage represented by a fuzzy function." in IEEE Transactions on Fuzzy Systems 2020, 28(7), 1195-1204.

[13] Tuncer E., Bolat E.D. "EEG signal based sleep stage classification using discrete wavelet transform." International Conference on Chemistry, Biomedical and Environment Engineering (ICCBEE'14) 2014, 57-61.

[14] Tuncer E., Bolat E.D. "Classification of epileptic seizures from electroencephalogram (EEG) data using bidirectional short-term memory (Bi-LSTM) network architecture." Biomedical Signal Processing and Control 2022, 73, 103462.

[15] Phinyomark A., Quaine F., Charbonnier S., Serviere C., Tarpin-Bernard F., Laurillau Y., "Feature extraction of the first difference of EMG time series for EMG pattern recognition." Computer Methods and Programs in Biomedicine 2014, 117(2), 247-256.

[16] Tuncer E., "Ictal-interictal epileptic state classification with traditional and deep learning architectures." International Journal of Research Publication and Reviews 2022, 3(9), 1972-1977.

[17] Yilmaz M. "Wavelet Based and Statistical EEG Analysis in Patients with Schizophrenia." Traitement du Signal, 2021, 38(5), 1477-1483.

[18] Constable R., Thornhill R.J., Pandv M.G., "Using the continuous discrete wavelet transform for time-frequency analysis of the surface EMG signal." Journal of Biomechanics 1994, 27(6), 723.

[19] Oner IV, Yesilyurt M.K., Yilmaz E.C., "Wavelet analysis techniques and application areas." Ordu University Journal of Science and Tecnology 2017, 7(1), 42-56.

[20] URL , https,//www.mathworks.com/help/wavelet/gs/introduction-to-the-wavelet-families.html#f3-998398.

[21] Inik O., Ulker E., "Deep learning and deep learning models used in ımage analysis." Gaziosmanpasa Journal of Scientific Research 2017, 6(3), 85-104.

[22] Tosunoglu E., Yılmaz R., Ozeren E., Saglam Z., "Machine learning in education, a study on current trends in researchs." Journal of Ahmet Kelesoglu Education Faculty 2021, 3(2), 178-199.

[23] Gumuscu A., Tasaltin R., Aydilek I.B. "C4.5 decision tree pruning using genetic algorithm." Dicle University Journal of the Institute of Natural and Applied Science 2016, 5(2), 77-80.

[24] Kokver Y, Barıscı N., Ciftci A., Ekmekci Y., "Determining affecting factors of hypertension with data mining techniques." NWSA-Engineering Sciences 2014, 9(2), 15-25.

[25] Meng X., Zhang P., Xu Y., Xie H. "Construction of decision tree based on C4.5 algorithm for online voltage stability assessment." International Journal of Electrical Power & Energy Systems 2020, 118, 105793.

[26] Gokgoz E, Subasi A. "Comparison of decision tree algorithms for EMG signal classification using DWT." Biomedical Signal Processing and Control 2015, 18, 138-144.

[27] Tuncer E., Bolat E.D., "Channel based epilepsy seizure type detection from electroencephalography (EEG) signals with machine learning techniques." Biocybernetics and Biomedical Engineering, 2022, 42, 575–595.

[28] Arslankaya S., Toprak S., "Using machine learning and deep learning algorithms for stock price prediction." International Journal of Engineering Research and Development 2021, 13(1), 178-192.

[29] Supakar R., Satvaya P., Chakrabarti P., "A deep learning based model using RNN-LSTM for the Detection of Schizophrenia from EEG data." Computers in Biology and Medicine 2022, 151,106225.

[30] Vapnik V.N, "Methods of Pattern Recognition: In The Nature of Statistical Learning Theory." Springer, 2000, New York-USA, 123–180.

[31] Kecman V., "Learning and Soft Computing: Support Vector Machines, Neural Networks, and Fuzzy Logic Models." MIT Press, 2001, Cambridge, MA-USA, 121-189.

[32] Wu X, Kumar V, Ross Q.J., Ghosh J., Yang Q., Motoda H., Steinberg D., "Top 10 Algorithms in Data Mining." In Knowledge and Information Systems, 2008, 14 (1), 1-37.

[33] Cover T., Hart P., "Nearest Neighbor Pattern Classification." IEEE Transactions on Information Theory, 1967, 21-27.

[34] Demir F., "Deep autoencoder-based automated brain tumor detection from MRI data. Artificial Intelligence-Based Brain-Computer Interface" Academic Press, 2022, 317-351.

[35] Fawcett T. "An introduction to ROC analysis. Pattern Recognition Letters" 2006, 27(8), 861-874.

[36] Bue B.D., Merényi E., Killian J., "Classification and diagnosis of myopathy from EMG signals." 2nd Workshop on Data Mining for Medicine and Healthcare, in Conjunction with the 13th SIAM International Conference on Data Mining, 2013, Austin, TX.

[37] Belkhou A., Achmamad A., Jbari A., "Classification and diagnosis of myopathy EMG signals using the continuous wavelet transform." Scientific Meeting on Electrical-Electronics & Biomedical Engineering and Computer Science (EBBT) 2019, 1-4. DOI, 10.1109/EBBT.2019.8742051.

BIOGRAPHIES

Erdem TUNCER received M.Sc. degrees in Electronics & Computer Education from Kocaeli University, Turkey, in 2015. He received Ph.D. degree in Biomedical Engineering from Kocaeli University, Turkey in 2022. He is currently working as an electronics teacher at the Ministry of National Education. His active research interests are signal processing, artificial intelligence, machine learning.

EMİNE DOĞRU BOLAT received her M.Sc. degree in Electronics & Computer Education from Kocaeli University, Turkey in 1998 and Ph.D. degree in Electrical Education from Kocaeli University, Turkey in 2005. She is currently working as a professor in Biomedical Engineering in Kocaeli University, Turkey. Her research interests are biomedical engineering, automation systems, industrial control, embedded systems and engineering education.

Investigation of the Effects of Atmospheric Attenuation and Frequency on MIMO Channel Capacity

Ahmet Furkan Kola and Cetin Kurnaz

Abstract— The efficiencies of 5G channels, which are highly affected by atmospheric attenuation, are still being investigated. The effects of frequency and atmospheric attenuation parameters such as humidity, temperature, rain, and pressure were investigated in this study using the NYUSIM program. The spatial consistency mode of the NYUSIM channel simulator was turned off, and unnormalized channel capacities were calculated at 28, 45, 60, and 73 GHz frequencies. According to the research results, the rain rate was the atmospheric attenuation parameter that significantly affected MIMO channel capacity. In contrast, the humidity percentage had the slightest impact. The frequency where the channel capacity is most affected by the four determined atmospheric attenuation parameters is 60 GHz, while the frequency where it is least affected is 28 GHz. The study found that using frequencies with high atmospheric attenuation reduces communication efficiency significantly. Furthermore, rain rate has a significant impact on 5G channel performance.

I. INTRODUCTION

THE INTERNATIONAL Telecommunication Union (ITU) attempted to meet the growing demand for data rate with the fifth-generation (5G) technology. New technologies such as IoT, cloud, and intelligent systems add new bits to mobile communication has limited the use of the existing frequency region. Using the millimeter wave (mmWave) region with 5G has been motivated. The extremely high frequency (EHF) region, planned for 5G and beyond systems, aims to reach frequencies of up to 300 GHz and achieve extremely high data rates.

AHMET FURKAN KOLA, is with Department of Electrical and Electronics Engineering, Samsun University, Samsun, Turkey,(e-mail: ahmet.kola@samsun.edu.tr.
https://orcid.org/0000-0003-0510-3255

CETIN KURNAZ, is with Department of Electrical and Electronics Engineering, Ondokuz Mayıs University, Samsun, Turkey, (e-mail: ckurnaz@omu.edu.tr).
https://orcid.org/0000-0003-3436-899X

The communication system in mmWave technologies has been determined as Multiple-Input Multiple-Output (MIMO) Orthogonal Frequency-Division Multiplexing (OFDM). Existing MIMO systems have been converted to massive MIMO (m-MIMO) techniques to increase channel capacity and data rate by increasing the number of antennas. This technology can be developed due to the antenna sizes, which decrease inversely correspondingly to frequency increase. The OFDM technique, also used in existing communication systems, has been maintained by increasing the bandwidth to 800 MHz. Thus, the disruptive effect of the channel is eliminated by dividing the fast-fading channels into fixed-fading sub-channels [1].

Signal powers decreased as the frequency of mmWave communication increased, as did antenna lengths and communication distances between antennas. Furthermore, atmospheric attenuation begins to have a more significant impact on the signals. Meteorological factors such as temperature, humidity, air dryness, pressure, vapor, rain, and haze all affect the quality of mmWave communication. These innovations and disruptive effects necessitated the development of new mobile communication channel models. Many simulation programs are available to demonstrate how these models can be generated and how disruptions affect the models. WinProb [2] and QuaDRiGa [3], which create deterministic models, as well as MiliCar [4] and NYUSIM [5], which generate stochastic models, are among them. In this study, the NYUSIM channel simulator designed by Sun, MacCartney, and Rappaport in 2017 was used. Many studies have explored the power delay profile (PDP) and path loss models generated by the NYUSIM channel simulator [6-8]. There are also studies using this simulator to investigate the effects of meteorological variables on the channel [9-13] and the results of some of these parameters on channel capacity [14-16]. Aside from this simulator, the effects of meteorological variables on 5G systems have been studied [17,18]. Related researchs have demonstrated that mmwaves are significantly impacted by atmospheric processes, particularly as rain increases and absorption and scattering become more pronounced. However, the impact of all these variables on channel capacity has not been studied in detail.

This study investigated the effects of temperature, humidity, barometric pressure, and rain rate parameters on MIMO channel capacity at different frequencies using the NYUSIM

channel simulator. In this way, it has been determined which meteorological conditions affect the channels the most and what precautions should be taken. The responses of various high frequencies to these parameters were also investigated and evaluated.

II. MILLIMETER WAVE COMMUNICATION

The extreme congestion of the sub-6 GHz bands has encouraged different frequency bands in mobile communication. With its unlicensed use and wide bandwidths, mmWave transmission, the ITU's standards will be appropriate for the developing technology. MDue to the smaller antenna sizes in mmWave communication, more antennas can be placed at base stations. This way, the number of MIMO antennas for each communication can increase. Increasing the number of antennas to 64, 128, and 256 is called massive MIMO. The number of antennas per user decreases as the number of antennas increases. Furthermore, to communicate, these antennas, which increase in number with increasing frequency and declining signal strength, must be placed close to each other. Although costly, this system will reduce the intersymbol interference (ISI) and minimize the attenuation effects such as propagation and fading. The distortion effect of the channel increases with both the increase in carrier frequency and the increase in bandwidth in mmWave communication. To resolve this, the OFDM technique, which began with 4G, expanded to 5G. The deep fading points of the channel increase in agreement with the increase in bandwidth. The OFDM technique attempts to fix this problem by dividing the channel into sub-channels. It makes the fast fading channel locally flat faded and transmits with these sub-channels by dividing it into N sub-channels. In this way, it simplifies communication and avoids the channel's disruptive effects [1].

III. CHANNEL MODELLING

A communication model consists of a receiver, transmitter, and channel. The signal transmitted from the transmitter x(t) a signal at the receiver is y(t). The impulse response of the channel is h(t) and noise n(t) and is expressed by Eq. (1). These variables all depend on time represented with t. Also, τ is a time variable.

$$y(t)=x(t)*h(t)+n(t) \qquad (1)$$

$$= \int_{-\infty}^{\infty} h(\tau)x(t-\tau)d\tau + n(t)$$

If the communication has two inputs and two outputs, it is mathematically expressed as Eq. (2).

$$\begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} h_{11} & h_{12} \\ h_{21} & h_{22} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \begin{bmatrix} n_1 \\ n_2 \end{bmatrix} \qquad (2)$$

For m inputs and n outputs, the channel is denoted by Eq. (3).

$$H_{mxn} = \begin{bmatrix} h_{11} & h_{12} & ... \\ h_{21} & h_{22} & ... \\ \vdots & \vdots & h_{mn} \end{bmatrix} \qquad (3)$$

The channel between the first antennas at the receiver and the transmitter is denoted by h_{11}. This value is obtained when the signal is reflected, refracted, or directly reaches the receiver during transmission. If the signal goes in a direct path, it is referred to as the line of sight (LOS); if it goes as a result of reflections, it is referred to as a non-line of sight (NLOS). h_{11} and all other MIMO channels can be represented by only one of these transmitted signals (flat fading) or by the shifted sum of these signals arriving one after the other at different times (fast fading) [19]. Additionally, the signal can reach the receiver by being attenuated by the channel and interfering with other communication channels. As a result of these disruptive effects, the signal fades before going to the receiver. The use of high frequencies with 5G revealed the disruptive effects of atmospheric attenuation. Fig. 1 depicts atmospheric sea-level attenuation in the 1-100 GHz dB frequency range. The regions marked in red in the figure show fading at the 28, 45, 60, and 73 GHz frequencies used in this study. Because 28 and 73 GHz have the lowest fading locally, 60 GHz has the highest fading in the determined range, and 45 GHz is just before the fading slope increases.
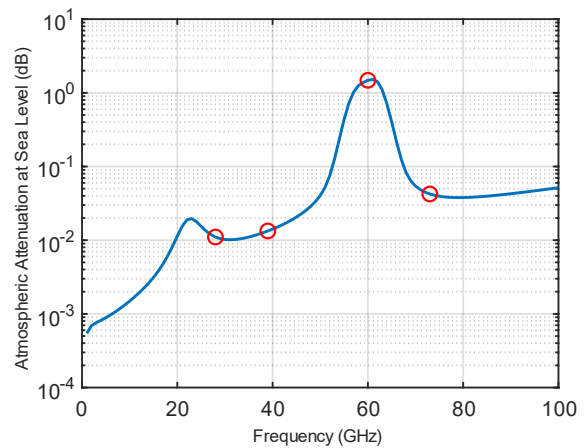

Fig.1. Atmospheric attenuation at sea level

A. NYUSIM Channel Model

In this study, the NYUSIM channel simulation V3.1 spatial consistency mode is turned off, and evaluations are performed for stationary users. Channel model measurements were taken for 28 and 73 GHz carrier frequencies. The channel model creates channel impulse responses (CIRs) using the concepts of time cluster (TC) and spatial lobe (SL). The simulator generates different stochastic channels at each run based on the parameters. Therefore, each scenario was created to repeat at least 2000 times and averaged during the analysis.

NYUSIM is a MATLAB program that calculates the frequency response of the channel. In the frequency (f) region, m is the number of transmitters, k is the receivers, and p is the components above -40 dB, which is considered stronger than noise, and all MIMO-OFDM subcarriers are represented in Eq. (4).

$$h_{mk}(f) = \sum_p \alpha_{m,k,p} e^{j\Phi} e^{-j2\pi f\tau} e^{-j2\pi d_T m\sin(\emptyset)} e^{-j2\pi d_R k\sin(\varphi)} \qquad (4)$$

Here, α is the amplitude of the antenna gain, Φ is the phase of each multipath component (MPC), τ is the delay of the MPC, d_T is the distance between MIMO transmitters, and d_R is the distance between each antenna at the receiver. $\emptyset$ and φ represent azimuth angles.

In NYUSIM, a close-in free space reference distance (CI) path loss model with a reference point of 1 m is used depending on various atmospheric attenuation factors. It is expressed by PL^{CI} and calculated by Eq. (5).

$$PL^{CI}(f, d) = FSPL(f, 1m) + 10nlog_{10}(d) + AT + X_\sigma^{CL} \qquad (5)$$

Here, d is the distance between the transceiver, n is the path loss exponent (PLE) value, and X_σ^{CL} is a zero-mean Gaussian random variable with a standard deviation σ in dB named shadow fading. The free space path loss model (FSPL) in dB can be calculated with Eq. (6).

$$FSPL(f, 1m)[dB] = 20\log_{10}\left(\frac{4\pi f \times 10^9}{c}\right) \qquad (6)$$

Where f represents the carrier frequency in GHz, and c represents the speed of light. Atmospheric attenuation in dB can be calculated in terms of α attenuation coefficient and distance with Eq. (7).

$$AT[dB] = \alpha[dB/m] + d[m] \qquad (7)$$

The coefficient α is expressed here as a function of dry air, haze, fog, rain, and vapor effect between 1 GHz and 100 GHz. Fig. 2 depicts the impact of these parameters individually, as shown by NYUSIM, on various frequencies [20].
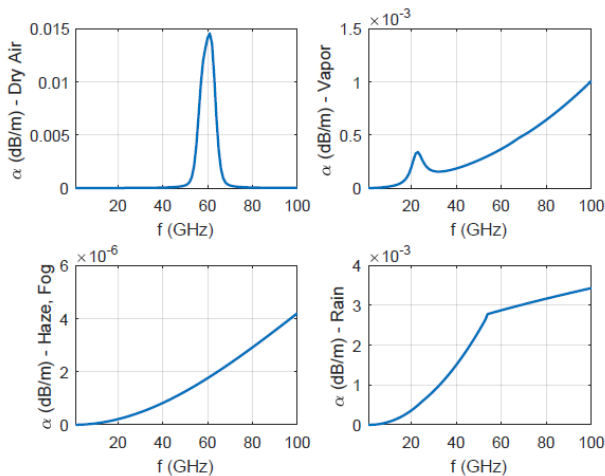

Fig.2. Propagation attenuation due to dry air, vapor, haze/fog, and rain at mmWave frequencies

B. Channel Capacity

This study made channel capacity calculations for the unnormalized channel condition. The measure of channel capacity can be expressed with Eq. (8) for MIMO channels.

$$C = B\log_2[\det(I_m + SNR.H.H')] \qquad (8)$$

Where H stands for the channel matrix, and H' stands for the conjugated transpose of the channel matrix. I_m represents the identity matrix with dimensions $m \times m$ and SNR means Signal-to-Noise ratio.

This study performed the capacity analysis in the frequency domain as given in Eq. (9).

$$C = B\sum_1^r \log_2(1 + SNR.\lambda_i^2) \qquad (9)$$

Here, λ_i^2 represents the eigenvalues of the $H.H'$ matrix.

IV. ANALYSIS RESULTS

The layers of Using the NYUSIM channel simulator, the effects of rain rate, barometric pressure, humidity, temperature, and frequency on MIMO channel capacity were investigated. The parameters used in the simulations are summarized in Table I. The Power Delay Profile (PDP) of the channel produced by the NYUSIM program, whose properties are determined by Table I (1013.25 mbar barometric pressure, 20 ℃ temperature, 80% humidity, 5 mm/hr rain rate), is shown in Fig. 3. PDP is made up of delayed reflections that arrive at the receiver at different strengths. The frequency response of this channel is shown in Fig. 4. The signal has an average power of -44.84 dBm and a bandwidth of 800 MHz around the carrier frequency of 28 GHz.

TABLE I
CHANNEL PARAMETERS

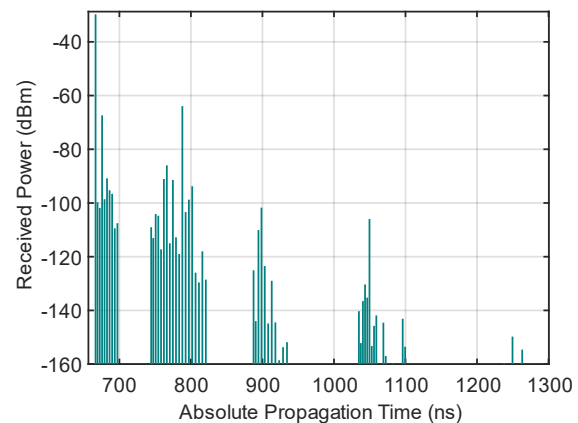| Channel Parameters	Scenarios			
	1	*2*	*3*	*4*
Frequencies	28, 45, 60, 73 GHz			
Environment	LOS			
Bandwidth	800 MHz			
MIMO	2 x 2			
Scenario	UMa			
Distance	100 m			
Tx Power	30 dBm			
Rain Rate	**Variable**	5 mm/hr	5 mm/hr	5 mm/hr
Barometric Pressure	1013.25 mbar	**Variable**	1013.25 mbar	1013.25 mbar
Humidity	80%	80%	**Variable**	80%
Temperature	20 ℃	20 ℃	20 ℃	**Variable**

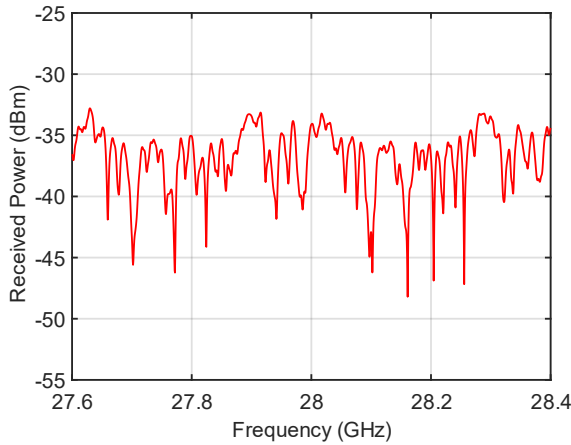Fig.3. A sample of channel impulse response value in time-domain

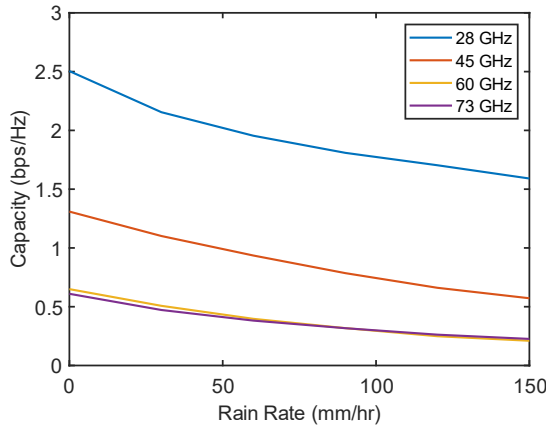Fig.4. An example of OFDM subcarrier presentation in the frequency domain

Fig.5. The effect of rain rate on channel capacity

A. Rain Rate Effect

All other parameters were held constant except for the rain rate, and the effect of this value on channel capacity was examined. Fig. 5 depicts the extent of the impact of the four different frequencies determined by the precipitation rate by channel capacity as a result of the analysis. The results were analyzed between conditions with no rain and a 150 mm rain rate per hour. When the simulation results are examined, the channel capacity at 28 GHz frequency is 2.5043 in the absence of rain and 1.5907 in the case of 150 mm/hr precipitation, representing a 36.48% decrease in channel capacity. When the 45 GHz frequency is examined, while the capacity is 1.3092 for 0 mm/hr, it is 0.5720 for 150 mm/hr, representing a 56.31% decrease in channel capacity. For 60 GHz, the capacity is 0.6497 in the absence of rain and 0.2104 in the case of the highest rain rate. The capacity was reduced by 67.62%. At 73 GHz, the capacity is 0.6099 for 0 mm/hr and 0.2260 for 150 mm/hr, representing a 62.94% reduction in channel capacity. The results showed that the frequency most affected by rain was 60 GHz. For 60 GHz frequency, 90 mm/hr, and above rain rate, the channel capacity is below the capacity of the 73 GHz channel. The frequency least affected by the rain rate as a percentage is 28 GHz.

B. Barometric Pressure Effect

The effect of Barometric Pressure on channel capacity was examined using the channel parameters shown in Table I. Fig. 6 depicts the decrease in the capacities of four different frequencies due to the analyses with increasing pressure. The results were examined between 100 mbar and 1013.25 mbar pressure values. As shown in Fig. 6, while the 60 GHz frequency is highly affected by pressure increases, such a decrease in capacity was not observed at other frequencies. The 28 GHz frequency capacity is 2.4837 for 100 mbar pressure, while it is 2.4696 for 1013.25 mbar, and there is a 0.56% decrease in channel capacity. When the frequency of 45 GHz is examined, while the capacity for 100 mbar is 1.2906, it is 1.2795 for 1013.25 mbar pressure. There is a 0.86% reduction in channel capacity. At 100 mbar pressure for 60 GHz, the capacity was 0.8011, while at 1013.25 mbar, it was 0.6243, representing a capacity decrease of 22.07%. At the 73 GHz frequency, while the capacity is 0.5963 for 100 mbar, it is 0.5752 for 1013.25 mbar, and there is a 3.54% reduction in channel capacity. The 60 GHz frequency experienced the most significant decrease in channel capacity as pressure increased. The least affected frequency is 28 GHz.
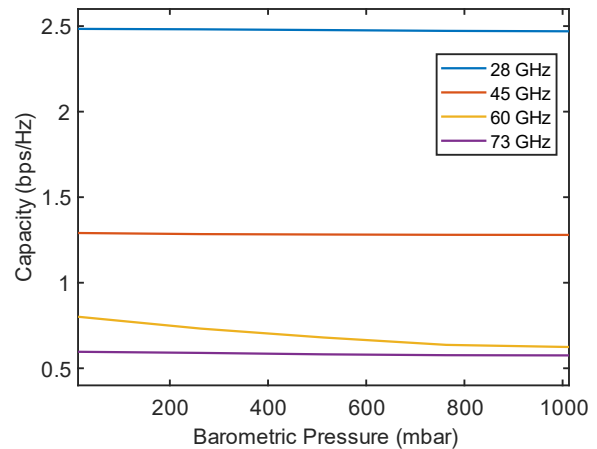

Fig.6. The effect of barometric pressure on channel capacity

C. Humidity Effect

The effect of humidity percentage on channel capacity was investigated for 28, 45, 60, and 73 GHz frequencies. Humidity was examined between 0% and 100%, and the results are shown in Fig. 7. When the simulation results are analyzed, the 28 GHz frequency has the highest channel capacity of 2.5006 in a 60% humidity environment and the lowest value of 2.4695 in a 100% humidity environment. Channel capacity was reduced by 1.24%. When the 45 GHz frequency is examined, while the capacity is 1.2912 for 0% humidity, it is 1.2542 for 100%, and there is a 2.86% decrease. For 60 GHz, the capacity is 0.5737 at 0 humidity, 0.6109 at 60% humidity, and a 6.09% change in capacity is seen. At the 73 GHz frequency, while the capacity is 0.5786 for 60% humidity, it is 0.5582 for 100% humidity, and there is a 3.52% reduction in channel capacity. As a result, the frequency most affected by humidity change is 60 GHz, while the least affected frequency is 28 GHz.
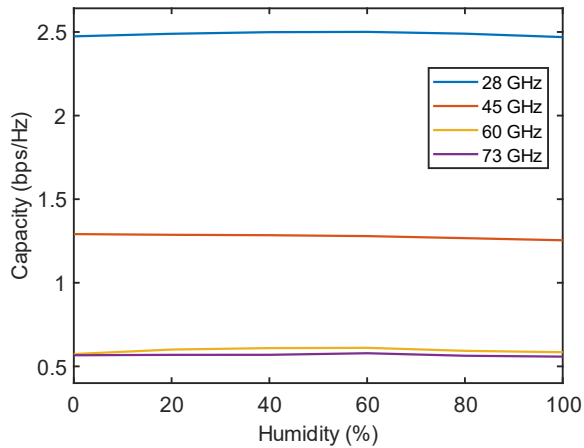
Fig.7. The effect of humidity percentage on channel capacity

D. Temperature Effect

The effect of temperature on the channel capacity was investigated between -10℃ and 40℃, and the results are shown in Fig. 7. The 28 GHz frequency capacity has a minimum value of 2.4692 at 20℃, and the capacity has increased for increasing and decreasing values from this temperature. The highest capacity was obtained as 2.4934 for -10℃, and the variation in capacity was 0.97%. The lowest capacity value for the 45 GHz frequency is 20℃, while the highest capacity value is -10℃. Capacity values are 1.2804 and 1.2758, respectively, and there is a 0.35% reduction in channel capacity. For 60 GHz, the capacity at -10℃ is calculated to be 0.5563, and the channel capacity increases with increasing temperature at this frequency. For 40℃ temperature, the capacity is 0.6382, and an increase of 12.83% was seen. At the 73 GHz frequency, the capacity is highest at 10℃, while the channel capacity decreases with increasing and decreasing temperatures. For 10℃, the capacity is 0.5784, while for 40℃, it is 0.5555, and there is a 3.96% reduction in channel capacity. The temperature effect has the most significant effect on 60 GHz, while 45 GHz has a minor impact.
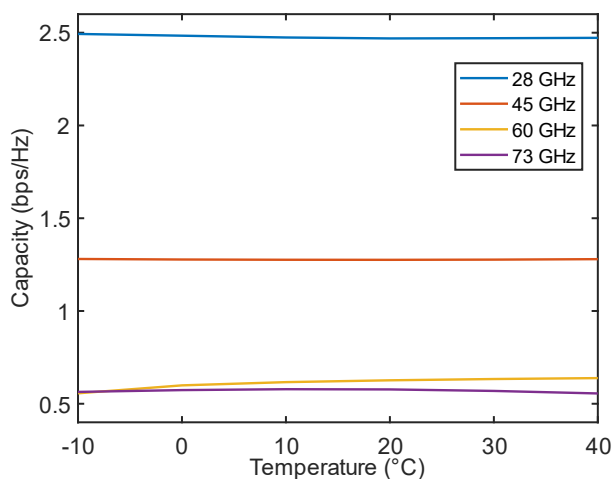

Fig.8. The effect of temperature on channel capacity

V. CONCLUSION

In this study, the effects of atmospheric attenuation parameters such as temperature, humidity percentage, rain rate, and barometric pressure on MIMO channel capacity were investigated using the NYUSIM channel simulator. In channel simulations, frequencies of 28, 45, 60, and 73 GHz were used. From the results of the analysis, it was seen that the 60 GHz frequency was the most affected by the four different parameters. 73 GHz is the second most affected frequency. While the frequency of 28 GHz was the least affected by the other three parameters except for temperature, it affected the frequency of 45 GHz at least by the temperature change. For the 60 GHz frequency, the rain rate had the most significant effect on channel capacity, with a variation of 67.62%, while humidity had the most negligible impact, with 6.09%. While a change in barometric pressure reduces capacity by 22.07%, a temperature change reduces capacity by 12.83%.

These results indicate that frequency selection is critical in next-generation mobile communication. It does not appear possible to achieve efficiency at frequencies with high atmospheric attenuation. Local minimum atmospheric attenuation (24-30 GHz/37-40 GHz/67-73 GHz) should be preferred when selecting a higher frequency. It is also understood that it will be more challenging to use at high frequencies with high attenuation, especially in heavy rain environments. In future studies, different channel parameters such as transceiver distance, different channel scenarios, different bandwidths will be evaluated, and analyses will be carried out with other comparisons such as path loss and received power. Apart from NYUSIM, it is also planned to compare different channel models with the simulator.

REFERENCES

[1] Wang, X., Kong, L., Kong, F., Qiu, F., Xia, M., Arnon, S., & Chen, G. (2018). Millimeter-wave communication: A comprehensive survey. IEEE Communications Surveys & Tutorials, 20(3), 1616-1653.
[2] WinProp, Wave Propagation and Radio Network Planning Software (part of Altair HyperWorks), www.altairhyperworks.com/WinProp.
[3] Jaeckel, S., Raschkowski, L., Börner, K., & Thiele, L. (2014). QuaDRiGa: A 3-D multi-cell channel model with time evolution for enabling virtual field trials. IEEE transactions on antennas and propagation, 62(6), 3242-3256.
[4] Drago, M., Zugno, T., Polese, M., Giordani, M., & Zorzi, M. (2020, June). MilliCar: An ns-3 module for mmWave NR V2X networks. In Proceedings of the 2020 Workshop on ns-3 (pp. 9-16).
[5] Samimi, M. K., & Rappaport, T. S. (2016). 3-D millimeter-wave statistical channel model for 5G wireless system design. IEEE Transactions on Microwave Theory and Techniques, 64(7), 2207-2225.
[6] Lodro, M. M., Majeed, N., Khuwaja, A. A., Sodhro, A. H., & Greedy, S. (2018, March). Statistical channel modeling of 5G mmWave MIMO wireless communication. In 2018 International Conference on Computing, Mathematics and Engineering Technologies (iCoMET) (pp. 1-5). IEEE.
[7] Momo, S. H. A., & Mowla, M. M. (2019, July). Statistical analysis of an outdoor mmWave channel model at 73 GHz for 5G networks. In 2019 International Conference on Computer, Communication, Chemical, Materials and Electronic Engineering (IC4ME2) (pp. 1-4). IEEE.
[8] Zekri, A. B., Ajgou, R., Chemsa, A., & Ghendir, S. (2020, May). Analysis of Outdoor to Indoor Penetration Loss for mmWave Channels. In 2020 1st International Conference on Communications, Control Systems and Signal Processing (CCSSP) (pp. 74-79). IEEE.
[9] Alfaresi, B., Nawawi, Z., Malik, R. F., Anwar, K., & Nur, L. O. (2020). Humidity Effect to 5G Performances Under Palembang Channel Model At 28 GHZ. Sinergi, 24(1), 49-56.

[10] Budalal, A. A., Rafiqul, I. M., Habaebi, M. H., & Rahman, T. A. (2019). The effects of rain fade on millimeter-wave channel in tropical climate. Bulletin of Electrical Engineering and Informatics, 8(2), 653-664.

[11] Rahman, M. N., Anwar, K., & Nur, L. O. (2019, November). Indonesia 5G channel model considering temperature effects at 28 GHz. In 2019 Symposium on Future Telecommunication Technologies (SOFTT) (Vol. 1, pp. 1-6). IEEE.

[12] Dimce, S., Amjad, M. S., & Dressler, F. (2021, March). mmwave on the road: Investigating the weather impact on 60 GHz v2x communication channels. In 2021 16th Annual Conference on Wireless On-demand Network Systems and Services Conference (WONS) (pp. 1-8). IEEE.

[13] Larasati, S., Yuliani, S. R., & Danisya, A. R. (2020). Outage Performances of 5G Channel Model Influenced by Barometric Pressure Effects in Yogyakarta. Jurnal Infotel, 12(1), 25-31.

[14] Prasetyo, A. H., Suryanegara, M., & Asvial, M. (2019, December). Evaluation of 5G Performance at 26 GHz and 41 GHz frequencies: The Case of Tropical Suburban Areas in Indonesia. In 2019 IEEE 14th Malaysia International Conference on Communication (MICC) (pp. 101-105). IEEE.

[15] Hikmaturokhman, A., Suryanegara, M., & Ramli, K. (2019, June). A comparative analysis of 5G channel model with varied frequency: a case study in Jakarta. In 2019 7th International Conference on Smart Computing & Communications (ICSCC) (pp. 1-5). IEEE.

[16] Kurniawan, A., Danisya, A. R., & Isnawati, A. F. (2020, December). Performance of mmWave Channel Model on 28 GHz Frequency Based on Temperature Effect in Wonosobo City. In 2020 IEEE International Conference on Communication, Networks, and Satellite (Comnetsat) (pp. 37-41). IEEE.

[17] Kourogiorgas, C., Sagkriotis, S., & Panagopoulos, A. D. (2015, April). Coverage and outage capacity evaluation in 5G millimeter wave cellular systems: impact of rain attenuation. In 2015 9th European Conference on Antennas and Propagation (EuCAP) (pp. 1-5). IEEE.

[18] Huang, J., Cao, Y., Raimundo, X., Cheema, A., & Salous, S. (2019). Rain statistics investigation and rain attenuation modeling for millimeter-wave short-range fixed links. IEEE Access, 7, 156110-156120.

[19] Popa, S., Draghiciu, N., & Reiz, R. (2008). Fading types in wireless communications systems. Journal of Electrical and Electronics Engineering, 1(1), 233-237.

[20] Sun, S., MacCartney, G. R., & Rappaport, T. S. (2017, May). A novel millimeter-wave channel simulator and applications for 5G wireless communications. In 2017 IEEE International Conference on Communications (ICC) (pp. 1-7). IEEE

BIOGRAPHIES

AHMET FURKAN KOLA received his B.Sc. degree in Electronic and Communication Engineering from Istanbul Technical University in 2019. He is an Ph.D. candidate in the Department of Electrical and Electronics Engineering at Ondokuz Mayıs University. He is currently working as a Research Assistant in Electrical and Electronics Engineering the Samsun University. His current research interests include Telecommunications, Wireless Communication channels, and Digital Signal Processing.

ÇETİN KURNAZ received his B.Sc. degree in Electrical and Electronic Engineering from the Ondokuz Mayıs University in 1999. He received an M.Sc. degree in Electrical and Electronics Engineering from the Ondokuz Mayıs University in 2002. He received a Ph.D. in Electrical and Electronics Engineering from the Ondokuz Mayıs University in 2009. He is currently working as an Associate Professor of Electrical and Electronics Engineering at Ondokuz Mayıs University. His research interests include Electromagnetic Fields, Microwave Techniques, Telecommunications, and Digital Signal Processing.

Design of the Integrated Cognitive Perception Model for Developing Situation-Awareness of an Autonomous Smart Agent

Evren Daglarli

Abstract— **This study explores the potential for autonomous agents to develop environmental awareness through perceptual attention. The main objective is to design a perception system architecture that mimics human-like perception, enabling smart agents to establish effective communication with humans and their surroundings. Overcoming the challenges of modeling the agent's environment and addressing the coordination issues of multi-modal perceptual stimuli is crucial for achieving this goal. Existing research falls short in meeting these requirements, prompting the introduction of a novel solution: a cognitive multi-modal integrated perception system. This computational framework incorporates fundamental feature extraction, recognition tasks, and spatial-temporal inference while facilitating the modeling of perceptual attention and awareness. To evaluate its performance, experimental tests and verification are conducted using a software framework integrated into a sandbox game platform. The model's effectiveness is assessed through a simple interaction scenario. The study's results demonstrate the successful validation of the proposed research questions.**

Index Terms— **Autonomous smart agents, Cognitive perception, Attention modelling, World model.**

I. INTRODUCTION

THE ABILITY to engage with their environment through perception is vital not just for humans but also for autonomous systems equipped with intelligent agents. These abilities rely on representations of the world model in terms of spatial and temporal, as well as perceptual cognition, situation awareness and attentional capabilities [1]. In biological systems, the cortical and cerebral lobes of the human brain play a significant role in providing these functions and characteristics. The cerebral cortex's anatomical structure consists of two primary cortical structures known as the frontal (anterior) and posterior (posterior) lobes [2]. Cognitive abilities related to perception functions are primarily located

EVREN DAGLARLI. is with Computer and Informatics Engineering Faculty, Istanbul Technical University, Istanbul, Turkey, (e-mail: evren.daglarli@itu.edu.tr, evrendaglarli@ieee.org).

https://orcid.org/0000-0002-8754-9527

in the posterior section of the cerebral cortex [2, 3]. This region is further divided into three subregions, namely the occipital, parietal, and temporal lobes [4]. The occipital lobe, housing the primary visual cortex regions, performs various functions on visual stimuli after extracting their features [5, 6]. The temporal lobe is responsible for pattern recognition in visual and auditory stimuli [3, 5]. Spatial perception, on the other hand, is handled by the parietal lobe, which receives visual and somatosensory stimuli [4, 7]. However, autonomous systems comprising intelligent agents, digital assistants, or social robots have encountered significant challenges in implementing these capabilities during human-machine interaction experiments [8, 9]. Therefore, cognitive perception systems are currently critical issues in the fields of human-computer interaction (HCI), aiming to enhance interaction between autonomous agents and humans [10, 11].
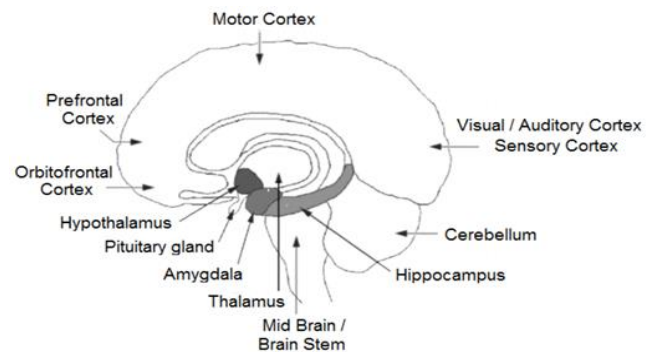

Fig.1. Regions of cerebral cortex in the human brain

In our daily lives, autonomous systems that possess cognitive perceptual abilities require various forms of interaction with their surroundings [12, 13]. Hence, it is necessary for these systems to be endowed with a perceptual system resembling that of humans, enabling them to interpret and sustain representations of the world model. Furthermore, they should be able to assess human-machine interaction through shared attention within a collaborative workspace [1, 14, 15].

Spatial cognition involves being aware of one's environment, encompassing the perception of objects in terms of their spatial aspects such as positions, orientations, distances, and movements [1, 16-18]. On the other hand, temporal cognition focuses on the intermediate processes that involve encoding non-spatial or temporal characteristics of objects, such as color and shape, as well as recognizing

patterns like objects, faces, and spoken words. Accomplishing higher-level cognitive abilities to effectively respond to multimodal perceptual stimuli, engage in pattern recognition, model attention, and exhibit environmental awareness presents a challenging task [1, 19, 20]. Developing perceptual models that represent the environment of an autonomous system, including spatial world models and the interaction of physical behavior models, is among the significant challenges faced [17-20]. Additionally, temporal perception plays a crucial role by incorporating event-based or situation-based representations of the world [21-23]. Errors in representing the world model or coordinating multimodal perceptual stimuli can lead to mistakes in interaction.

Fig.2. Sand-box game platform

The objective of this research is to design a comprehensive multimodal perception system for an intelligent agent that can effectively navigate and explore its surroundings within a virtual gaming platform. By leveraging the computational principles of the posterior neocortex, a software framework can be developed to guide the construction of this cognitive perception system. The proposed solution aims to demonstrate its efficacy in representing dynamic environments with uncertainties. This approach offers several notable contributions. For instance, it incorporates cognitive perceptual functions capable of processing various multimodal stimuli such as visual, auditory, and somatosensory inputs, enabling tasks such as feature extraction, pattern recognition, and spatial perception. Moreover, the system incorporates the ability to coordinate perceptual information, including perceptual association (or sensory fusion) and the management of competition between stimuli, which are vital for modeling perceptual attention. Achieving these objectives involves implementing supervised and unsupervised learning techniques across different modules within the cognitive perception system.

The subsequent section of the paper delves into the second part, where it provides an overview of the relevant research conducted in the field. Section 3 focuses on outlining the design principles governing the computational framework of the comprehensive multimodal perception system, which aims to achieve world model representation and spatial-temporal situational awareness within dynamic and uncertain environments. The article concludes with section 4, which encompasses a discussion of the findings, conclusions drawn from the study, and potential avenues for future research.

II. RELATED WORKS

Computational cognitive architectures have emerged as solutions for addressing perceptual and environmental modeling challenges in the context of autonomous aget systems. The number of projects in this domain has been rapidly increasing and is projected to continue growing in the future, signifying its significance. Notably, several commendable examples of computational architectures incorporating cognitive perception principles and facilitating world model representation and attention modeling have been introduced.

In one study, Inceoglu et al. [24] proposed a visual scene representation framework specifically designed for service robots. Their objective was to generate and maintain comprehensive workplace models to facilitate object manipulation. The framework employed various algorithms and vision data flow sources to cater to both humanoid and manipulator systems. It incorporated different detection algorithms that processed visual data, continually improving and updating the world model representation.

Another notable work by Kim et al. [25] explored a curiosity-driven framework called Dynamic World Model Learning (AWML). The study involved the development of a curious agent that constructed models of the world through visual exploration of a rich 3D physical environment [26]. The researchers focused on refining representative real-world agents to drive the AWML framework. They specifically emphasized efficient and adaptive learning progress-based curiosity indicators to guide the exploration process. The study demonstrated that the AWML framework, propelled by such progress-driven controllers, outperformed alternative approaches, including random network distillation and model mismatch, in terms of achieving higher AWML performance. These examples highlight the advancements made in computational architectures for cognitive perception, particularly in relation to world model representation and adaptive learning mechanisms.

Riedelbauch and Henrich proposed an adaptable method tailored for human-robot collaboration, where a robot dynamically selects actions that contribute to a shared objective based on a given behavioral pattern [27]. To gather information about task progress, they constructed a world model using camera images captured from an eye-to-eye perspective. Recognizing that data generated by fractional workspace perceptions can become obsolete over time due to human interaction with resources, they introduced a human-aware world model. This model maintains observations of ongoing human presence and stored item confidence in relation to past assignment progress. Their notable contribution was an action selection mechanism that utilized this confidence measure, combining mission operations with active vision to update the world model. The extensive testing of their system involved simulating various human interests by recreating modernized human models and evaluating the system's performance across different benchmark assignments, resulting in scores associated with various functions.

In a separate study, Rosinol et al. introduced an integrated model termed 3D Dynamic Environmental Networks for dynamic spatial perception [28]. This model represented the scene using environmental networks composed of nodes representing entities such as objects, walls, and rooms, along with the relationships between these nodes. To accommodate moving agents and incorporate dynamic data aiding planning and decision making, they extended this concept with Dynamic Scene Graphs (DSGs). Additionally, they developed an automated Spatial Perception Engine (SPIN) that leveraged visual inertial data to construct a DSG. The researchers focused on state-of-the-art strategies for human and object recognition, posture computation, and perception of objects, robot nodes, and human nodes in crowded environments. Their work incorporated visual-inertial SLAM and dense human network tracking. They also devised algorithms for generating hierarchical models of indoor environments, including places, structures, and rooms, along with their interrelationships. A noteworthy achievement was the demonstration of the spatial perception engine within a photo-realistic Unity-based simulator. The application of the 3D Dynamic Scene Graphics technique had significant implications for planning and decision making, human-robot interaction, long-term autonomy, and scene prediction.

Venkataraman et al. tackled the challenge of generating generic 3D models for original items using a robot capable of decluttering items to enhance organization [29]. Their approach involved creating models of grasped objects through simultaneous manipulation and tracking. These models were processed using a kinematic representation of the robot, which allowed for combining observations from multiple scenes and eliminating background noise. To evaluate their model, they employed a robot equipped with a mobile platform, a manipulator, and an RGBD camera. This setup facilitated the assembly of voxelized representations of unidentified items, which were then classified into new categories.

Persson et al. focused on semantic world representation by combining probabilistic thinking and item binding [30]. Their paradigm adopted a top-down item binding approach based on continuous attribute values obtained from perceptual sensor data. They trained a binding matching model to maintain item entities and validated its performance using a large ground truth dataset of manually labeled real-world items. To handle more complex scenarios, they integrated a high-probability item tracker into the binding architecture, enabling reasoning about the state of unobserved items. The effectiveness of their system was demonstrated through various scenarios, including a shell game scenario that showcased how binding items were preserved through probabilistic reasoning.

Martires et al. aimed to establish a semantic scene representation paradigm based on top-down item connectivity, utilizing an item-induced model of the world [31]. Their approach involved processing continuous perceptual sensor data to maintain perceptual connectivity, which correlated with a symbolic model. They extended the symbol binding model to incorporate binding annotations, enabling the

execution of multimodal probability distributions and probabilistic logic reasoning for making inferences. Additionally, they employed statistical associative learning to enable the binding system to acquire symbolic knowledge in the form of probabilistic logic rules from noisy and sub-symbolic sensor input. By leveraging logical rules to reason about the state of indirectly detected items, their system, incorporating perceptual connectivity and statistical associative learning, could maintain a semantic world model of all perceived items over time. They validated the performance of their system by evaluating the framework's probabilistic reasoning on multimodal likelihood and learning probabilistic logical rules from connected items obtained through perceptual observations.

III. BACKGROUND AND PRELIMINARY MATERIALS

Currently, in conventional deep learning techniques, training data comprising input data and corresponding target (class) information can be effectively trained and subsequently evaluated with new data inputs. These deep learning algorithms demonstrate remarkable efficiency in terms of data set size, data set quality, feature extraction methods, hyperparameter selection for deep learning models, activation functions, and optimization algorithms.
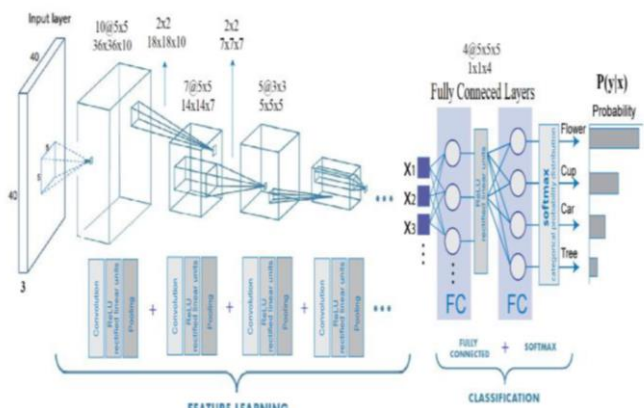

Fig.3. Convolutional neural network

A deep neural network comprising multiple layers enables the system to recognize objects at various levels of abstraction. Given the significant information processing demands associated with performing artificial cognitive functions and executing cognitive tasks, diverse deep neural network architectures such as multilayer perceptron, auto-encoder, convolutional neural network (CNN), and long-short-term memory (LSTM) recurrent neural networks are essential for learning models. These models can be further integrated within a hybrid AI framework, depending on the specific circumstances.

Regarding the traditional convolutional neural network model, which exhibits robust feature learning capabilities similar to high-level abstraction processes in the cortical regions of the human brain, the layer arrangement typically follows the sequence of [input (x) – convolution layer - ReLU

– max pool layer] illustrated in figure 3. The convolution process involves applying convolution filters (weights) arranged as a cubic tensor. The rectified linear unit (ReLU) serves as the activation function. Typically, the neural network is trained using momentum stochastic gradient descent (SGD), facilitating the hierarchical extraction of features to be encoded by the large-scale neural activations within the model [32-34].
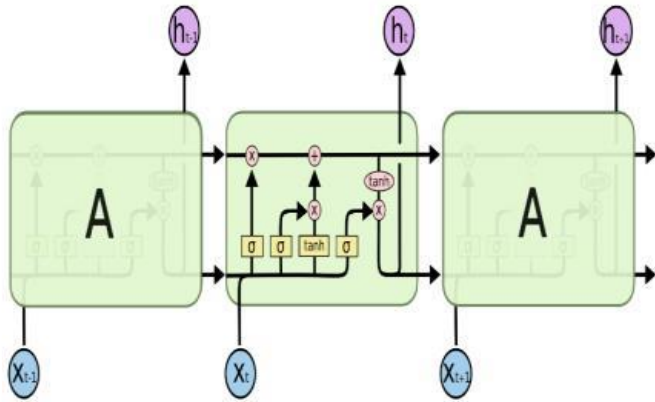

Fig.4. Long-Short Term Memory (LSTM) type neural network [35]

The LSTM (Long Short-Term Memory), a recurrent neural network (RNN) structure depicted in fig. 4, can be engineered using convolutional layers to mimic the memory mechanisms observed in the human brain. This architecture enables the LSTM to maintain short-term memory states over an extended duration, resembling episodic memory formations [32, 33, 35]. The LSTM comprises a fundamental component called the memory cell, which incorporates input, output, and forget gates. Training the neural model may involve employing backpropagation through time calculations [35, 36].

IV. COMPUTATIONAL MODEL OF PERCEPTUAL COGNITION

In this section, we provide an explanation of the cognitive architecture employed in an open-world game/simulation environment, which encompasses the perceptual mechanism of an autonomous intelligent agent engaging with various elements in its surroundings. The key aspect of the developed cognitive integrated perception system is its ability to construct a world model representation for dynamic and uncertain environments, while also supporting the agent's attention model. Throughout this study, the cognitive integrated perception system refers to three core components: the proposed framework, the environmental elements (objects, non-player characters, etc.), and the agent's internal states. The newly proposed structure aims to capture the spatio-temporal relationships and features arising from the dynamic interaction between the autonomous intelligent agent and the world model. To achieve a comprehensive model of the world for the autonomous agent, the attention model is integrated into higher-level perceptual processing, serving as a crucial element for assessing and detecting the level of spatio-temporal state awareness during the agent's interaction with its environment.
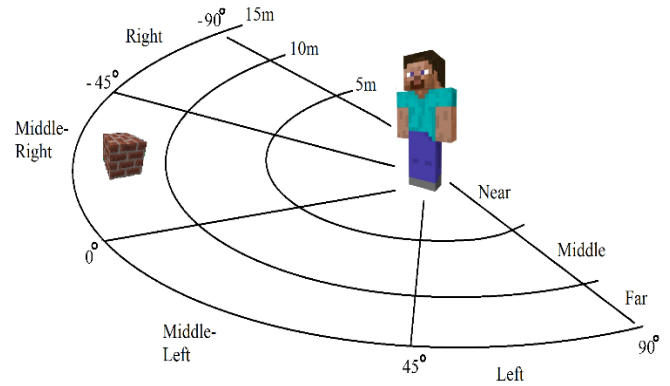

Fig.5. Spatial perception

In order to develop a comprehensive understanding of the environment and navigate through it, the cognitive architecture must create a world model that encompasses all perceptual relationships and incorporates semantic concepts [37-39]. The integration of perceptual data into this model requires the cognitive system to perform complex fusion tasks. Additionally, the attention model, which is an integral part of the cognitive process, enhances the agent's interaction with the environment by promoting situational awareness. However, the presence of non-structural dynamic uncertainties during the creation of the world model can introduce perceptual distortions, leading to limitations in the agent's ability to recognize and attend to different elements in its surroundings.
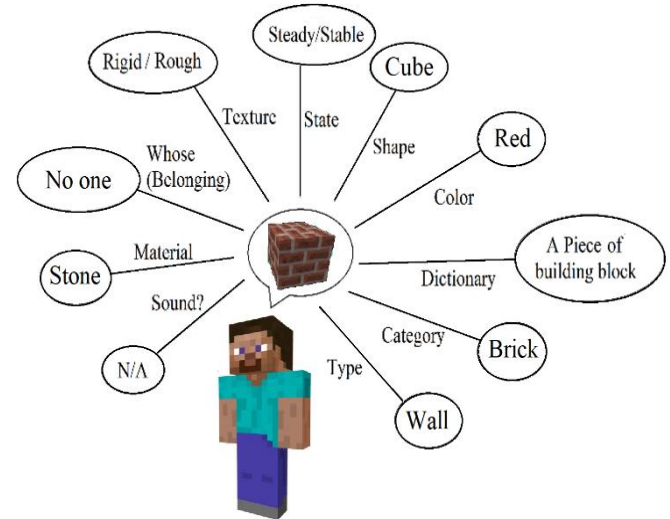

Fig.6. Temporal or non-spatial cognition

The cognitive model presented utilizes real-time image flow to enable visual perception during the autonomous agent's interaction with the environment. The general algorithm follows a sequential data flow, consisting of preprocessing, feature extraction, and basic perceptual operations such as spatial-temporal pattern recognition. The final stage of the perceptual cognition mechanism focuses on achieving situational awareness for the autonomous agent, which involves constructing the world model and implementing the attention model.

Our proposed cognitive perception system offers a comprehensive framework for achieving situational awareness in autonomous robots and intelligent agents. It encompasses various components such as spatio-temporal pattern recognition, world model, and attention model, with a strong emphasis on the relationship between semantic concepts and perceptual connections. The diagram of the cognitive architecture can be seen in Figure 7.
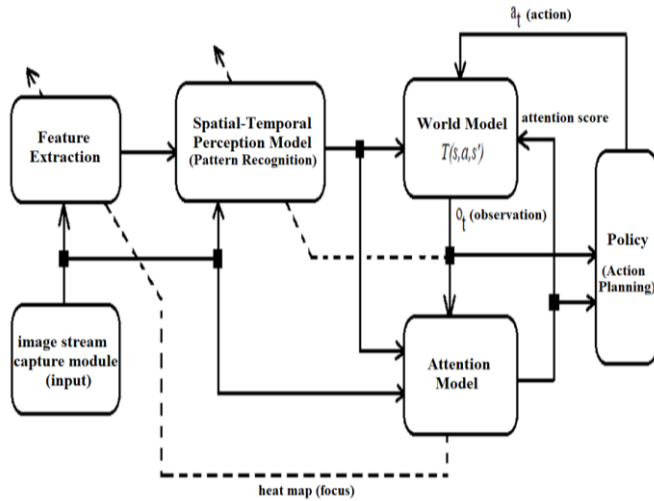

Fig.7. Cognitive architecture of the autonomous agent

Before engaging in cognitive perception processes involving the world model representation and attention model, it is necessary to establish the system parameters for the learning models of the cognitive architecture. This includes performing data attribute processing activities such as segmentation, edge/corner detection, and filtering for feature extraction. The generalization and clustering tasks in this context are guided by unsupervised learning techniques. As a result, two distinct attribute data streams, namely spatial attribute information and temporal attribute information, are obtained for utilization in the spatio-temporal pattern recognition model in figure 8.
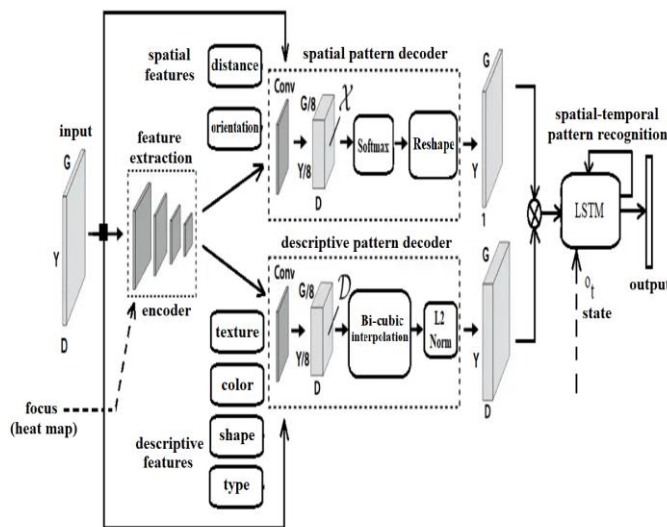

Fig.8. Feature extraction and spatial-temporal perception model

To facilitate the pattern identification process, a convolution filter (weight tensor) size is selected, consisting of a stack of four-time frames with dimensions of 83x158x3, representing the height, width, and depth of the image pattern, respectively. The initial alpha coefficient, which serves as a learning rate parameter, is set to 0.00025. The pattern recognition tasks incorporated in this model primarily employ supervised learning methodologies.

The neural network's weights are continuously updated using the backpropagation algorithm and the stochastic gradient descent optimization method. To extract features from the images and reduce their dimensionality, a VGG-type mesh is employed as the encoder. The encoder comprises convolutional layers, spatial subsampling achieved through maximum pooling, and nonlinear activation functions. The spatial and temporal features obtained from encoding are then separately forwarded to decoders that consist of convolutional layers. In the spatial decoder module, the information from the convolutional layer undergoes the softmax function and subsequently a resize operation. On the other hand, the descriptive decoder module applies bidirectional cubic interpolation, contrasting the spatial decoder, and concludes with the L2 norm on the output information of the convolutional layer. The outputs from the spatial and descriptive decoders are combined with the state information (observation) from the world model and transmitted to the long-short-term memory (LSTM) network. Finally, by adding a classifier layer to the output of this model, the spatio-temporal pattern recognition mechanism produces object recognition results.
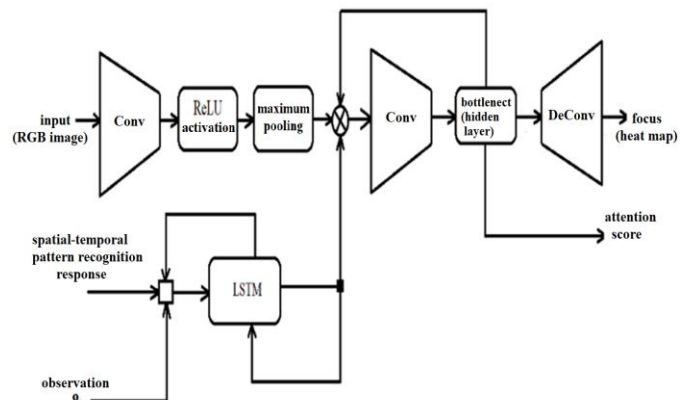

Fig.9. Attention model of the cognitive architecture

The attention model and the world model, crucial components of the situation awareness mechanism, engage in reciprocal interaction. The world model utilizes the attention value computed by the attention model and generates observation information to be relayed back to the attention model. The classification information obtained from the spatio-temporal pattern recognition module serves as a fundamental input for the attention model. Additionally, the captured RGB image stream acts as another input, which undergoes processing through convolutional neural networks, including convolutional, relu activation, and max pooling

layers. Subsequently, in conjunction with the spatio-temporal pattern output, the observation response from the world model is conveyed to the long-short-phase memory (LSTM) network. The outputs from this network and the convolutional neural network are merged and transmitted to an auto-encoder network structure. The output of the convolutional layer in this auto-encoder structure is transferred to a hidden layer, also known as the bottleneck, whose output is utilized as attention points in other modules. Furthermore, the output information of the said layer is linked back to the convolutional layer input of the auto-encoder network through internal feedback. In the final stage of the autoencoder within the attention model, the output of the hidden layer is directed to a deconvolutional layer, resulting in the generation of a heat map, denoting focus information to be shared with other modules. The calculation of attention values, crucial for providing situational awareness to the autonomous agent, employs supervised learning methodologies, akin to the spatio-temporal pattern identification model within the attention model. Unsupervised learning methods are employed to obtain the heat map (focus information) as feedback data for feature extraction.
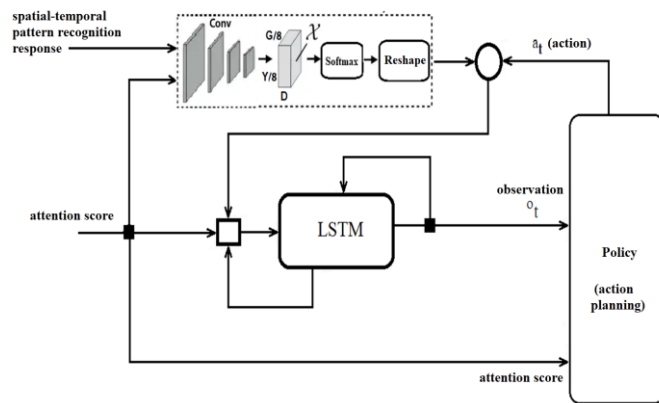

Fig.10. World model of the smart agent

The world model of the autonomous agent, which assumes the role of action planning, relies on both the action information provided by the policy module and the attention score generated by the attention module. Furthermore, the classification information derived from the spatio-temporal pattern recognition module serves as an additional input for the world model. The formation of the world model involves a convolutional layer that combines the attention score and the output of the spatio-temporal pattern recognition module. Subsequently, it undergoes softmax and size reduction operations. The disparity between the resulting output and the action information produced by the policy module is fed into an LSTM-based artificial neural network as input, which includes the attention score. The output of this neural network is utilized as observation information in the policy module, responsible for action planning. Another input to the policy module is the attention score itself. Reinforcement learning methodologies are employed in the policy module to facilitate action planning. As for the learning algorithm of the model in question, the initial discovery probability is set at 1.0, with a decay rate of 0.00001 and a minimum threshold of 0.01. The

reward reduction ratio (gamma) is determined as 0.9. Through the fusion of perceptual data using hybrid machine learning tools, this architecture creates a network of semantic relationships, granting it sensing capabilities similar to the human perception system. This enables intelligent agents designed for autonomous systems to engage in continuous learning by establishing a world model and achieving situational awareness.

V. IMPLEMENTATION AND RESULTS

This research explores various experimental and simulation environments to enable autonomous navigation for virtual characters. While conducting experiments with physical robots in real-world settings offers realistic evaluation benefits, it also presents challenges in terms of practical implementation. Conversely, utilizing simulators provides significant advantages, including customizable degrees of freedom, noise-free environments, lower costs, and reduced risk compared to deploying mobile robots.

A. Experimental Setup and Application Scenario

In order to implement the application, it is essential to establish the experimental setup and define the initial conditions. The application area serves as the setting for the game installation, where the fundamental mechanics of the game are introduced. Prior to proceeding with the installation, it is necessary to provide an overview of the general setup.

1) General Settings

The computational workload of the system architecture was supported by a workstation PC. The PC specifications include a quad-core Intel i7 CPU running at 3.9 GHz with 8MB cache, 32GB DDR4 RAM operating at 1600MHz, an NVIDIA GeForce GTX1080Ti graphics card with 11GB video memory, and 1TB SSD and 1TB 7200rpm HDD storage. The main framework processes were executed on the Ubuntu 18.04 LTS operating system. TensorFlow, a machine learning framework, was utilized for neural network processing and deep learning applications. OpenCV libraries were chosen for image processing and computer vision tasks.

2) Sand-box game platform

For this research, Minecraft was selected as the simulation environment. It is an open-world, first-person game that revolves around resource collection (such as wood from trees or stone walls) and the construction of structures and items. Players can engage in various actions, including movement, exploration, and building within the three-dimensional voxel space of the Minecraft map [40]. This game offers an infinitely dynamic environment that can be easily modified using a simplified physics engine. It can be played as a single-player or multiplayer open-world game, without any specific objectives. Instead, each player can create their own narrative with diverse sub-objectives, resulting in complex hierarchical structures.
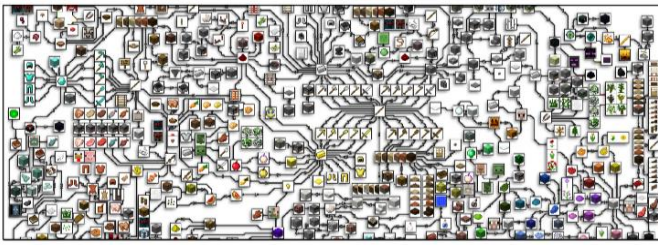
Fig.11. Tool crafting hierarchy of Minecraft game.

The virtual character model of the agent possesses extensive degrees of freedom, granting it a wide range of capabilities. Within the Minecraft environment, the agent can engage in various actions, including movement in any direction, turning, item manipulation (picking up/dropping), chopping, tool selection, and utilization. These actions, arranged in different sequences, form behaviors that represent complex tasks organized hierarchically [41]. Each behavior consists of a combination of interdependent actions that fulfill different needs and prioritize multiple objectives within the Minecraft map or world. Consequently, the difficulty level in Minecraft is influenced by the size and complexity of these hierarchical structures [42]. For instance, a navigation task involves actions such as moving forward/backward, turning right/left, which enable the virtual character model to reach specific locations or avoid obstacles and threats. Another example task, building a structure like a shelter, necessitates several actions such as item manipulation, chopping/destroying, and other equipment-related tasks involving item selection, modification, and usage [41, 42].

Fig.12. Inventory of the smart agent

In addition to engaging in missions that involve item collection and tool crafting, the experimental scenario presents more intricate and abstract hierarchies through various game features, which shape the agent's trajectory. For instance, interactive scenarios like combatting enemies, constructing shelters, and crafting tools from diverse resources for survival require extended durations or open-ended lifetimes to exhibit flexible hierarchies that allow for resource exploration. This facilitates the assembly of numerous resources and situational experiences [43, 44].

Data collection and feature extraction are crucial tasks in this context, necessitating extensive gameplay sessions with a large number of agents/humans [45, 46]. The dataset content encompasses substantial repetitions of memory utilizing observations, rewards, and actions. However, in nature, reward information is implicit and cannot be directly observed. This research paper utilizes the MineRL dataset, which stands as an extensive collection of imitation learning data, containing a staggering number of 60 million frames of human player recordings. This dataset consists of several sub-datasets and is employed to conduct experiments aimed at achieving a model that can adapt to diverse environments. Serving as a meta-dataset, it encompasses a wide range of tasks that showcase challenging problems, including exploration (such as navigation and item collection) and survival (such as tool crafting and combat).

The experimental setting involves creating a 600x600 Minecraft map that incorporates both high and low perceptual overlap. A typical naturalistic Minecraft map comprises elements such as mountains/hills, trenches, caves, valleys, rivers, lakes, trees, vegetation, rocks, and soil. Additionally, buildings like houses, shelters, and warehouses with walls, windows, doors, and furniture can be constructed using materials collected from the generated environment.

B. Implementation Scenario and Simulation Outputs

Moving on to the implementation scenario and simulation outputs, the proposed scenarios are realized using application platforms to assess the system's performance and validate its effectiveness in addressing the research questions at hand. To evaluate the system's results, it is necessary to capture a data stream comprising stacked image frames during the application scenarios in the experiments. In the initial stage, a snapshot is taken from the game's video stream, followed by preprocessing operations such as size reduction, grayscale image conversion, and optimization of the sampling rate. Once the system receives sensor data streams containing visual and auditory information, feature extraction tasks are performed to enhance cognitive perception skills. Subsequently, operations related to perceptual cognition, including spatial perceptions and object/event recognition, are conducted on the extracted feature data pertaining to salient attributes like color, texture, size, shape, 3D position, and audio features.

Fig.13. Snapshots from experiments using the Minecraft game platform. Entities and threat levels encountered by the autonomous agent in the experiment.
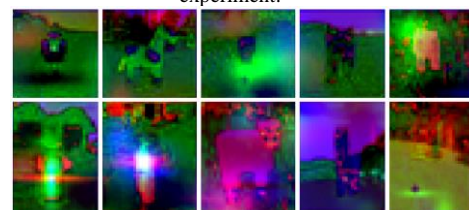

Fig.14. Focus (heat map matrix) information generated in the attention model of the cognitive perception architecture

Screenshots from the experiments implementing the interaction scenarios are presented in Figure 13. Two interaction scenarios such as "exploration" and "survival" were considered for the experiments. The "exploration" scenario involves spatio-temporal detection of hidden objects placed at various points on the map with the help of an attention model that takes into account different environmental stimuli. In doing so, it makes diagnoses based on feature information such as distance, direction, color, shape, size, and current state. The information produced by this process and attention score information contributes to the formation of the autonomous agent's world model. In performance evaluation, the number of objects discovered by the agent in the map is used as the main performance parameter. The "survival" scenario involves spatio-temporal risk detection tasks that evaluate the dangers (enemies, etc.) that the autonomous agent may encounter at various points of the island. Meanwhile, the autonomous agent builds a threat analysis-based world model in its memory using the risk detection data and feedback data from the attention model. The number of enemies defeated and the number of dangerous situations avoided by the agent in the map is used as the main performance criterion for efficiency measurement. During the experiments, the entities (non-player characters) that the autonomous agent encounters while navigating the map and their threat states are shown in Figure-13. Accordingly, the attention levels of the autonomous agent when it encounters zombie, wolf, killer and octopus characters are higher than the other entities and are 0.817, 0.683, 0.726 and 0.613, respectively. The threat levels of other entities were lower and defined as "harmless" for the autonomous agent. The attention levels of the autonomous agent when faced with chicken, cow, horse, sheep, pig, sheep, pig, and llama entities are 0.289, 0.352, 0.324, 0.331, 0.316, and 0.367, respectively. Figure-14 presents the focus (heat map matrix) information produced in the attention model of the cognitive perception architecture. Accordingly, when the autonomous agent encounters species with high threat level and species that can be considered as "harmful", more concentration (brightness in the heat map) is observed in the focus information compared to other species.

Fig.15. Minecraft in-game environmental terrain scene

Depending on the scenarios, at the end of these experiments, interaction data such as various experimental statistics (number of objects discovered on the map, scores and times for how many enemies defeated and how many dangerous situations escaped) as well as information about learning performances (accuracies, costs, scores, etc.) are

obtained and presented in tables to illustrate the performance of the autonomous agent's cognitive perception system, which includes spatio-temporal pattern recognition, attention model and situational awareness models.

TABLE I
SPATIAL-TEMPORAL PATTERN RECOGNITION MODEL

Model	Accuracy	Precision	Recall	F score
Proposed model	0,73	0,69	0,71	0,72
Regional Convolutional Neural Network (RCNN)	0,67	0,64	0,63	0,66
VGG16	0,58	0,65	0,63	0,61
AlexNet	0,62	0,66	0,62	0,64

Table-1 shows the performance values for spatio-temporal image recognition. In addition to the model proposed in this study, the performance values of the regional convolutional neural network (RCNN), VGG16, and AlexNet structures were also used for comparison. Accordingly, the model proposed in the paper was found to be advantageous with an accuracy of 73%. In terms of the performance evaluation, the runner-up model was observed to be the RCNN model.

TABLE II
THE ATTENTION MODEL'S EFFICIENCY

Model	Accuracy	Precision	Recall	F score
Proposed model	0,79	0,73	0,67	0,74
Regional Convolutional Neural Network (RCNN)	0,66	0,71	0,64	0,69
LSTM	0,75	0,71	0,62	0,72
ResNet	0,70	0,69	0,63	0,67

Table-2 shows the performance of the attention model, which is a part of the state awareness function. In addition to the model proposed in this study, the performance values of regional convolutional neural network (RCNN), LSTM, and ResNet structures were also used for comparison. Accordingly, the model proposed in the paper was found to be advantageous with an accuracy of 79%. The second-best model was found to be the LSTM model.

TABLE III
PERFORMANCE SCORES OF THE WORLD MODEL

Model	Accuracy	Precision	Recall	F score
Proposed model	0,71	0,66	0,63	0,67
Regional Convolutional Neural Network (RCNN)	0,68	0,65	0,62	0,66
LSTM	0,72	0,61	0,62	0,68
Resnet	0,69	0,57	0,59	0,62

Similarly, as in the evaluation and analysis for the attention model, convolutional neural network (RCNN), LSTM, and ResNet structures were used in addition to the proposed model in order to compare the performance values of the world model, which is another part of the situation awareness function. Although LSTM is slightly ahead of the proposed model in terms of accuracy in the world model, it is worse than the proposed model in terms of sharpness. Apart from these, the sharpness and sensitivity results of ResNet lagged

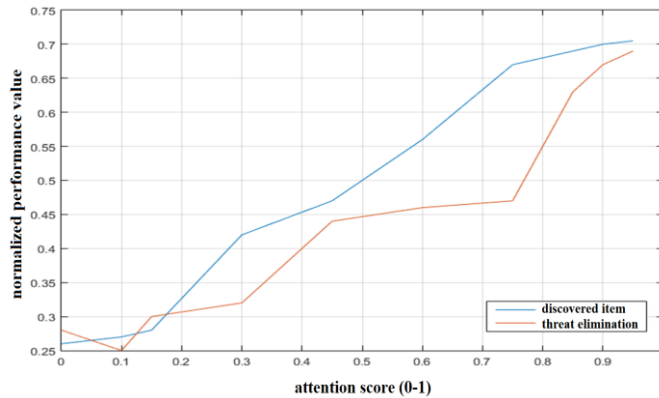behind the results of the other models.

Fig.16. Normalized task performances

The normalized action performances for the scenarios in the experiments are presented in figure-16. One of the parameters related to these actions is the number of hidden objects found in the map for the "exploration" scenario. Another parameter is the number of times the autonomous agent survives the threats encountered in the map for the "survival" scenario. For the convenience of the evaluations, the total number of hidden objects and the number of threatening characters were limited to 50. On the horizontal axis of the figure, the attention score is expressed in percentages between 0-1. The scaled values expressed on the vertical axis are the rate of change over time of the normalized values found by dividing the parameters (actions) in both scenarios by the total number of those parameters. Considering this figure, it is observed that at low attention levels, the parameters in both scenarios are close to each other and as the attention level increases, the number of object discoveries, i.e., the "exploration" task performance, dominates over the "survival" task performance. However, when the attention level reaches the highest levels, it is observed that the "survival" task performance catches up with the "exploration" task performance.

VI. CONCLUSION

In this study, we investigate the design principles of a novel cognitive perception system for autonomous agents. A cognitive perception system is a framework that includes spatio-temporal pattern recognition, an attention model and a world model.

The spatio-temporal pattern recognition model, which evaluates the 3D spatial environment representation and the dynamics based on the actual event, effectively served as one of the main components of the cognitive perception architecture. The information produced by this architecture was used in the attention model and the world model. The attention model successfully calculated both the attention score and the focus information (heat map matrix) using the information from the world model. The world model, which has a feedback data exchange structure with the attention model, produced both the observation information required for the action planning model.

The experiments focused on two different scenarios such as

"survival" and "exploration", and the main performance parameters were the number of dangers avoided in the "survival" task and the number of hidden objects found in the "exploration" task. In addition to the models in the proposed architecture, VGG16, AlexNet, ResNet, LSTM and RCNN models were also used for performance comparison. As a result of the experiments, the superiority and efficiency of the model proposed in the paper compared to other models are presented with the results obtained.

The designed framework represents the perceptual cognition system of autonomous agents in a sand-box game environment. Therefore, it can be used in other intelligent/autonomous systems or by social robots. The presented framework can be further improved in the future by integrating approximate models of other cortical regions of the human brain related to cognitive perception.

REFERENCES

[1] Yan, Z., Schreiberhuber, S., Halmetschlager, G., Duckett, T., Vincze, M., & Bellotto, N. (2020). Robot Perception of Static and Dynamic Objects with an Autonomous Floor Scrubber. arXiv preprint arXiv:2002.10158.

[2] Freud, E., Behrmann, M., & Snow, J. C. (2020). What Does Dorsal Cortex Contribute to Perception?. Open Mind, 1-18.

[3] Bear, M., Connors, B., & Paradiso, M. A. (2020). Neuroscience: Exploring the brain. Jones & Bartlett Learning, LLC.

[4] Chin, R., Chang, S. W., & Holmes, A. J. (2022). Beyond cortex: The evolution of the human brain. Psychological Review.

[5] Thiebaut de Schotten, M., & Forkel, S. J. (2022). The emergent properties of the connected brain. Science, 378(6619), 505-510.

[6] Li, B., Solanas, M. P., Marrazzo, G., Raman, R., Taubert, N., Giese, M., ... & de Gelder, B. (2023). A large-scale brain network of species-specific dynamic human body perception. Progress in Neurobiology, 221, 102398.

[7] Devia, C., Concha-Miranda, M., & Rodríguez, E. (2022). Bi-Stable Perception: Self-Coordinating Brain Regions to Make-Up the Mind. Frontiers in Neuroscience, 15, 805690.

[8] Taylor, A., Chan, D. M., & Riek, L. D. (2020). Robot-centric perception of human groups. ACM Transactions on Human-Robot Interaction (THRI), 9(3), 1-21.

[9] Ronchi, M. R. (2020). Vision for Social Robots: Human Perception and Pose Estimation (Doctoral dissertation, California Institute of Technology).

[10] Suzuki, R., Karim, A., Xia, T., Hedayati, H., & Marquardt, N. (2022, April). Augmented reality and robotics: A survey and taxonomy for ar-enhanced human-robot interaction and robotic interfaces. In Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems (pp. 1-33).

[11] Farouk, M. (2022). Studying Human Robot Interaction and Its Characteristics. International Journal of Computations, Information and Manufacturing (IJCIM), 2(1).

[12] Müller, S., Wengefeld, T., Trinh, T. Q., Aganian, D., Eisenbach, M., & Gross, H. M. (2020). A Multi-Modal Person Perception Framework for Socially Interactive Mobile Service Robots. Sensors, 20(3), 722.

[13] Russo, C., Madani, K., & Rinaldi, A. M. (2020). Knowledge Acquisition and Design Using Semantics and Perception: A Case Study for Autonomous Robots. Neural Processing Letters, 1-16.

[14] Cangelosi, A., & Asada, M. (Eds.). (2022). Cognitive robotics. MIT Press.

[15] Iosifidis, A., & Tefas, A. (Eds.). (2022). Deep Learning for Robot Perception and Cognition. Academic Press.

[16] Lee, C. Y., Lee, H., Hwang, I., & Zhang, B. T. (2020, June). Visual Perception Framework for an Intelligent Mobile Robot. In 2020 17th International Conference on Ubiquitous Robots (UR) (pp. 612-616). IEEE.

[17] Mazzola, C., Aroyo, A. M., Rea, F., & Sciutti, A. (2020, March). Interacting with a Social Robot Affects Visual Perception of Space. In Proceedings of the 2020 ACM/IEEE International Conference on Human-Robot Interaction (pp. 549-557).

[18] Mariacarla, B. Special Issue on Behavior Adaptation, Interaction, and Artificial Perception for Assistive Robotics.

[19] Sanneman, L., & Shah, J. A. (2020, May). A Situation Awareness-Based Framework for Design and Evaluation of Explainable AI. In International Workshop on Explainable, Transparent Autonomous Agents and Multi-Agent Systems (pp. 94-110). Springer, Cham.

[20] Kridalukmana, R., Lu, H. Y., & Naderpour, M. (2020). A supportive situation awareness model for human-autonomy teaming in collaborative driving. Theoretical Issues in Ergonomics Science, 1-26.

[21] Tropmann-Frick, M., & Clemen, T. (2020). Towards Enhancing of Situational Awareness for Cognitive Software Agents. In Modellierung (Companion) (pp. 178-184).

[22] Gu, R., Jensen, P. G., Poulsen, D. B., Seceleanu, C., Enoiu, E., & Lundqvist, K. (2022). Verifiable strategy synthesis for multiple autonomous agents: a scalable approach. International Journal on Software Tools for Technology Transfer, 24(3), 395-414.

[23] Sakai, T., & Nagai, T. (2022). Explainable autonomous robots: A survey and perspective. Advanced Robotics, 36(5-6), 219-238.

[24] Inceoglu, A., Koc, C., Kanat, B. O., Ersen, M., & Sariel, S. (2018). Continuous visual world modeling for autonomous robot manipulation. IEEE Transactions on Systems, Man, and Cybernetics: Systems, 49(1), 192-205.

[25] Kim, K., Sano, M., De Freitas, J., Haber, N., & Yamins, D. (2020). Active World Model Learning in Agent-rich Environments with Progress Curiosity. In Proceedings of the International Conference on Machine Learning (Vol. 8).

[26] Kim, K., Sano, M., De Freitas, J., Haber, N., & Yamins, D. (2020). Active World Model Learning with Progress Curiosity. arXiv preprint arXiv:2007.07853.

[27] Riedelbauch, D., & Henrich, D. (2019, May). Exploiting a Human-Aware World Model for Dynamic Task Allocation in Flexible Human-Robot Teams. In 2019 International Conference on Robotics and Automation (ICRA) (pp. 6511-6517). IEEE.

[28] Rosinol, A., Gupta, A., Abate, M., Shi, J., & Carlone, L. (2020). 3D Dynamic Scene Graphs: Actionable Spatial Perception with Places, Objects, and Humans. arXiv preprint arXiv:2002.06289.

[29] Venkataraman, A., Griffin, B., & Corso, J. J. (2019). Kinematically-Informed Interactive Perception: Robot-Generated 3D Models for Classification. arXiv preprint arXiv:1901.05580.

[30] Persson, A., Dos Martires, P. Z., De Raedt, L., & Loutfi, A. (2019). Semantic relational object tracking. IEEE Transactions on Cognitive and Developmental Systems, 12(1), 84-97.

[31] Zuidberg Dos Martires, P., Kumar, N., Persson, A., Loutfi, A., & De Raedt, L. (2020). Symbolic Learning and Reasoning with Noisy Data for Probabilistic Anchoring. arXiv, arXiv-2002.

[32] LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. nature, 521(7553), 436.

[33] Goodfellow, I., Bengio, Y., & Courville, A. (2016). Deep learning. MIT press.

[34] LeCun, Y., & Bengio, Y. (1995). Convolutional networks for images, speech, and time series. The handbook of brain theory and neural networks, 3361(10), 1995.

[35] Hochreiter, S., & Schmidhuber, J. (1997). Long short-term memory. Neural computation, 9(8), 1735-1780.

[36] Sainath, T. N., Vinyals, O., Senior, A., & Sak, H. (2015, April). Convolutional, long short-term memory, fully connected deep neural networks. In 2015 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP) (pp. 4580-4584). IEEE.

[37] Chiu, H. P., Samarasekera, S., Kumar, R., Matei, B. C., & Ramamurthy, B. (2020). U.S. Patent Application No. 16/523,313.

[38] Wang, S., Wu, T., & Vorobeychik, Y. (2020). Towards Robust Sensor Fusion in Visual Perception. arXiv preprint arXiv:2006.13192.

[39] Xue, T., Wang, W., Ma, J., Liu, W., Pan, Z., & Han, M. (2020). Progress and prospects of multi-modal fusion methods in physical human-robot interaction: A Review. IEEE Sensors Journal.

[40] Guss, W. H., Codel, C., Hofmann, K., Houghton, B., Kuno, N., Milani, S., ... & Wang, P. (2019). Neurips 2019 competition: The minerl competition on sample efficient reinforcement learning using human priors. arXiv preprint arXiv:1904.10079.

[41] MineRL: A Large-Scale Dataset of Minecraft Demonstrations

[42] Frazier, S., & Riedl, M. (2019, October). Improving deep reinforcement learning in Minecraft with action advice. In Proceedings of the AAAI Conference on Artificial Intelligence and Interactive Digital Entertainment (Vol. 15, No. 1, pp. 146-152).

[43] Aluru, K. C., Tellex, S., Oberlin, J., & MacGlashan, J. (2015, September). Minecraft as an experimental world for AI in robotics. In the 2015 AAAI fall symposium series.

[44] Angulo, E., Lahuerta, X., & Roca, O. (2020). Reinforcement Learning in Minecraft.

[45] Eraldemir, S. G., Arslan, M. T., & Yildirim, E. (2018). Investigation of feature selection algorithms on A cognitive task classification: a comparison study. Balkan Journal of Electrical and Computer Engineering, 6(2), 99-104.

[46] Akinci, T. Ç., & Martinez-Morales, A. A. (2022). Cognitive Based Electric Power Management System. Balkan Journal of Electrical and Computer Engineering, 10(1), 85-90.

BIOGRAPHIES

EVREN DAGLARLI earned his undergraduate degree in electrical and electronics engineering from Marmara University. He pursued a master's degree in mechatronics engineering at Istanbul Technical University (ITU), specializing in intelligent systems and robotics. His Ph.D. degree focused on control and automation engineering, specifically computational cognitive neuroscience and human-robot interaction, also from ITU. He served as a research assistant at Atilim University in the Department of Electrical and Electronics Engineering. Dr. Daglarli has contributed to numerous publications in international journals, conferences, and symposiums within the fields of mechatronics, intelligent control systems, and robotics. He has actively participated in various national and international projects, shouldering responsibilities as a researcher. He has also held positions as a project engineer and department manager in a private technology and engineering company. Dr. Daglarli is a senior member of IEEE. Presently, he serves as a faculty member and instructor in the Computer Engineering Department at Istanbul Technical University's Faculty of Computer and Informatics Engineering. He continues his research work in the Cognitive Systems Laboratory (CSL) and the Artificial Intelligence and Data Science Research Center (ITUAI).

Application of Sensor Fusion Techniques for Vehicle Condition and Position Analysis

Yasin Alyaprak and Gokhan Gokmen

Abstract—Sensor fusion is a method of processing data from raw data to meaningful outputs and getting quality output. Architectures used in sensor fusion are chosen depending on the application. The sensor fusion architecture that is frequently used today was found by the directors of the United States Joint Laboratory (JDL). Sensor fusion has been realized with this architecture. Using the axial data of a car, inertial movements such as acceleration, deceleration and stationary are classified as controlled. At the classification point, low level and high-level methods are used in the sensor fusion application. By pre-processing the received data, joint high-quality data was obtained with complementary sensor modeling, and high-level sensor fusion methods were used after recording the obtained data. Artificial intelligence algorithms are preferred for high-level sensor fusion. Various algorithms such as "Decision Tree", "Gradient Boosting", "Multi-Layer Perceptron", "Regression Algorithm" have been used. Real-time acquired data is stored after preprocessing and raw data fusion. The stored data has created a high-level sensor fusion at the point of decision making with supervised learning artificial intelligence algorithms.

Index Terms— Sensor Fusion, IMU, Gyroscope, Accelerometer, Artificial Intelligence.

I. INTRODUCTION

SENSOR FUSION is the combination of sensed data or data from sensed data to be better than would be possible when these sources were used individually. Data sources for a fusion operation are not expected to be exactly the same type. Sensor fusion can be created with different types of sensors [1].

It is not necessary to use different sensors for sensor fusion. It is sufficient to have different data belonging to the same sensor. Direct fusion is the merging of data from one or more

YASIN ALYAPRAK, Department of Mechatronics Engineering, Technology Faculty, University of Marmara, Istanbul, Turkey, (e-mail: alyaprakyasin@gmail.com).

http://orcid.org/0000-0002-6428-3973

GOKHAN GOKMEN, Department of Mechatronics Engineering, Technology Faculty, University of Marmara, Istanbul, Turkey, (e-mail: gokhang@marmara.edu.tr).

http://orcid.org/0000-0001-6054-5844

sensors, while Indirect fusion is the use of environmental information sources [2]. Consequently, sensor fusion defines direct fusion systems.

Sensor fusion also involves combining data received at different times or combining multiple received data
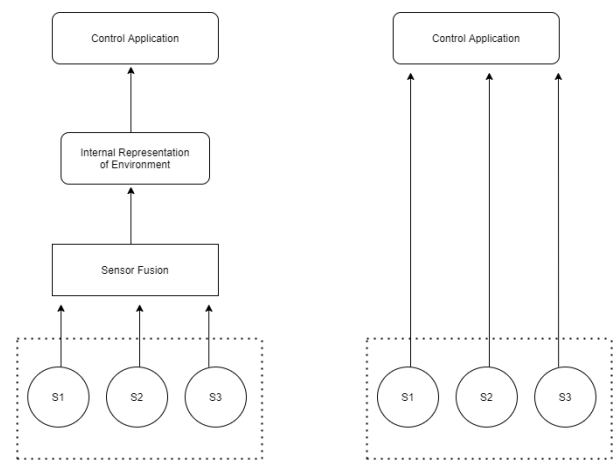

Fig.1. Sensor Fusion and Multiple Information Merging

A. Sensor fusion types

Fusion of sensors for C3I-oriented (Command, Control, Communication, Intelligence) applications and sensor fusion used in real-time microprocessor systems (embedded system) works on another level. C3I-based work usually deals with medium to high-level sensor fusion issues. Applications in embedded systems focus on low-level fusion [3].

Low-level aggregation in sensor fusion with three-level categorization combines raw data sources to extract more qualified data from raw data inputs. Intermediate aggregation or feature-level aggregation combines more meaningful data than raw in an entity map that may be used for significance. High level fusion (fusion of decision), unite the final decision and results. Among the decision fusion methods, there are artificial intelligence and statistical approaches [4].

II. RELATED WORKS

R. Olfati-Saber [5] provides consensus algorithms for networked dynamic systems, scalable algorithms for sensor fusion in sensor networks. This article introduces a distributed filter that checks the average of n sensor measurements of the nodes of a sensor network, called the consensus filter.

M. Kam [6] worked on the robot's self-positioning with sensor fusion techniques in robot navigation. The obstacles and routes around the robot were tried to be determined by

means of the sensors used. Among the sensors used are rangefinder, GPS, obstacle recognition sensors.

J.K. Hackett [7], a multisensor fusion research are presented for the latest technology. He classified the fusion definition into 6 categories: scene segmentation, representation, 3D shape, sensor modeling, autonomous robots, and object recognition.

He tried different strategies to derive more meaningful data from the data received from the sensors. These strategies range from simple set intersection, logical operations, heuristic generation rules, nonlinear least squares method to more complex methods that include maximum probability estimates. Sensor uncertainty is modeled by making sense with Bayesian probabilities and Dempster-Shafer algorithms.

O. Kreibich [8] presents an industrial wireless sensor network (IWSN) based machine condition monitoring (MCM) system that is capable of handling a false indication caused by temporary data loss, signal interference or invalid data. Multiple sensor fusion was used, driven by a quality parameter generated by each sensor node based on data history outliers and the actual state of the node. The fusion node also provides a quality assessment at its output. In this study, the effects of low-level sensor fusion on high-level sensor fusion were evaluated. The use of raw data in the decision-making algorithm is compared with the output of low-level sensor-fused data.

III.   METHOD

In this section, sensors and artificial intelligence algorithms to which sensor fusion algorithms will be examined. The general name of the sensors to be used is the inertial measurement unit (IMU). They are units consisting of at least 2 types of sensors to measure the 3-axis acceleration and 3-axis rotational force on the moving or stationary object. Today, these devices, which have many uses, from advanced aircraft to mobile phones to simple devices, perform the analysis of the movement of the platform on which they are located. Sensor types are protractor, gyroscope, and magnetometer. These sensors alone cannot give the axial state and the inertial state [9].

A. Accelerometer

Accelerometers are microelectronic devices that measure the applied force from the reference mass. Accelerometers measure the force of gravity. When the accelerometers are under constant speed or fixed, the 3 axis angles of the object can be measured with the force created by the gravitational force on the axes. The accelerometer measures two types of forces fixed and dynamic. The constant force is the gravitational force. Dynamic forces are such as acceleration, deceleration, and friction. In many fields such as aviation, automation, product electronics, the outputs given by serial communication are evaluated. Accelerometers output in terms of gravitational acceleration. If the measuring range is low, more precise measurements can be made. (Like vibration). If the measuring range is high, the sensitivity will decrease, but it can measure high acceleration values.

While the accelerometers are stationary, they are subjected to the acceleration of gravity on the Z-axis of 1 g on the earth. In this case, the resultant force is 1g.
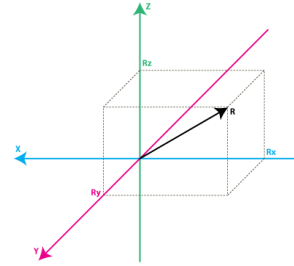

Fig.2. Acting resultant acceleration vector [9]

Thus, the tilt and roll angle can be found from the accelerometer standing still. But the roll angle cannot be found. Because rotations around the Z axis do not change the acceleration applied on the sensor [10]. Euler Angles are created for low-level sensor fusion as follows:

$$\tan \phi_{xyz} = \frac{-G_{px}}{\sqrt{G_{py}\sin\phi + G_{pz}\cos\phi}} \quad \tan\phi_{xyz} = \frac{G_{py}}{G_{pz}} \quad (1)$$

$$\tan \theta_{xyz} = \frac{-G_{px}}{\sqrt{G_{py}\sin\phi + G_{pz}\cos\phi}} \quad (2)$$

Here : G_{px} -Accelerometer Raw X Data
G_{py} - Accelerometer Raw Y Data
G_{pz} -Accelerometer Raw Z Data
ϕ -Pitch Raw Angle
θ -Roll Raw Angle
$\tan\phi_{xyz}$ -Euler Angle -Pitch
$\tan\theta_{xyz}$ -Euler Angle -Roll

B. Gyroscope

Gyroscopes are sensors used to measure the rotational speed at a specified reference orbit. The gyroscopes outputs are in degrees per second. Angular velocity is a measure of the speed of rotation about the axis. Gyroscopes use it to determine the axial state. For example, it can be used to balance an aircraft, to measure the angle at the equilibrium position and send it to the automatic control mechanism [11].

C. Magnetometer

Studies on the time-varying magnetic field of the earth have been carried out for years. Devices that measure this variable magnetic field are called magnetometers. Magnetometers are used extensively in industrial applications. agriculture, defense, biology, aerospace, space exploration, etc. It is also widely used in other fields such as, and currently, no electronic system is independent of magnetic field. Magnetometer is a commonly used sensor, especially in aviation.

D. Artificial intelligence algorithms

In this study, classification and interpretation studies were carried out on the stored data. After the low-level sensor fusion, artificial intelligence algorithms were used at the

decision-making point, that is, in the high-level sensor fusion part. Various supervised learning algorithms have been used. Decision Tree [12, 13] is a type of algorithm that aims to reduce the disorder in the system [14]. The Gradient Boosting algorithm, on the other hand, aims to create stronger algorithms by modeling the error between the results of any algorithm and the actual values. The Multi-Layer Perceptron algorithm is a method of generating output by summing the weighted percentage of the inputs and passing them through the threshold function in neuron logic [15].

IV.   EXPERIMENTS

In this section, data collection, filtering, application of sensor fusion and preparation as input to artificial intelligence algorithms will be explained.
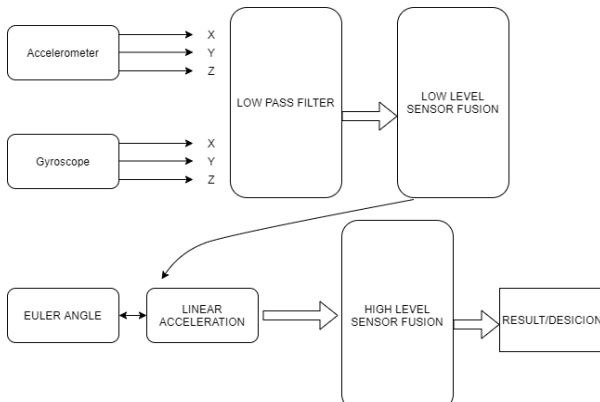
Fig.3. Flow Diagram

The sensors will be subjected to both low- and high-level fusion algorithms. Accelerometer data alone gives more meaningful data than Gyroscope data. The gyroscope sensor gives the angular velocity. It does not give information about direct axial changes. By taking its integral, the sum of the changes is reached. The sum of these changes is used to reach the desired axial angles. But there will be shifts in the gyroscope sensors. And it never outputs 0. Since the value of 0 cannot be taken, a cumulative angle change is observed from the integral calculation over time. An accelerometer is needed to compensate for this angle. Raw data were collected using the sensor kit shown in Figure 4. This development kit includes a 3-axis gyroscope, magnetometer and accelerometer sensor.
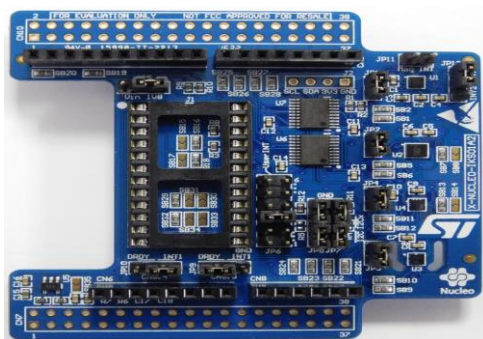
Fig.2. X-NUCLEO-IKS01A2 [16]

The collected data was first subjected to a sliding mean filter. Complementary filter is used for low level sensor fusion. Euler angles were reached after the complementary filter.

In order to use artificial intelligence algorithms, the data collection program shown in Figure 5 was written. With the written program, the data was transferred to the PC with the stm32L073 series microprocessor connected to the sensor kit. Saved in .csv format via the program.
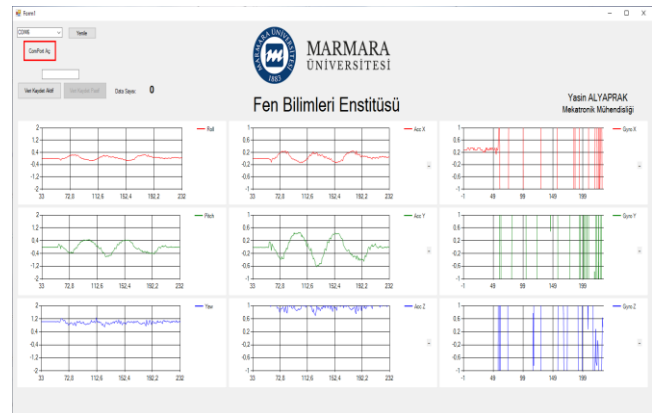

Fig.3. Data Collection Program Runtime Display

Fig.4. Real-Time Data Collection with Car

The high number of repetitions in algorithms such as artificial intelligence and machine learning will increase the percentage of classification and prediction accuracy. Within the scope of the study, this system test was carried out with a real car. Take-off acceleration in 3 repetitions, deceleration acceleration and fixed position data while traveling with a certain speed was taken. The slope factor for these 3 cases was added to the test.

Thus, a total of 6 different status data were recorded from 3 tests on the slope exit direction and on the non-slope road. While the data was included in the artificial intelligence algorithm, it was processed in certain groups. A 2-stage learning table was used. In the first group, the data was generated with raw data without low-level sensor fusion.

That is tilt and roll values are not included. In the second group, learning data was created without including the gyroscope data.

After the learning data was created, learning was carried out with 60% of the data. The test was also carried out with 40%

of the data. The writing and evaluation of artificial intelligence algorithms were carried out in Python language and Jupyter environment.

V.   EXPERIMENTAL RESULTS

While applying the sensor fusion experiment, the low-level sensor fusion process was applied first. Accelerometer and gyroscope data were filtered and then Euler angles were created and 2 data, inclination and roll angle, were obtained from 6 sensor data. In this way, high quality easily understandable data was obtained. Accelerometer data collected from the vehicle traveling at constant speed are shown in Figure 7, gyroscope data in Figure 8 and Euler angles obtained after low-level sensor fusion output are shown in Figure 9. Data were recorded for approximately 30 seconds.
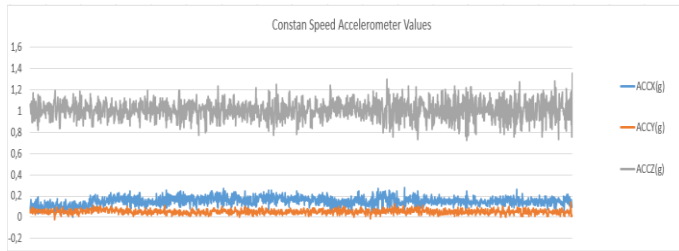
Fig.5. Accelerometer Raw Data
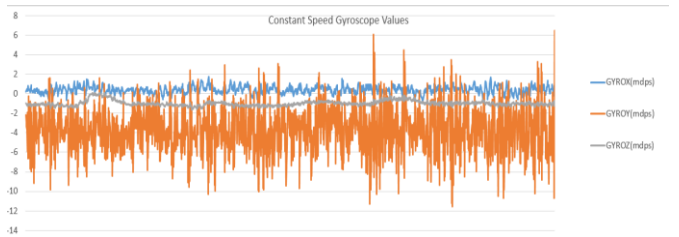
Fig.8. Gyroscope Raw Data
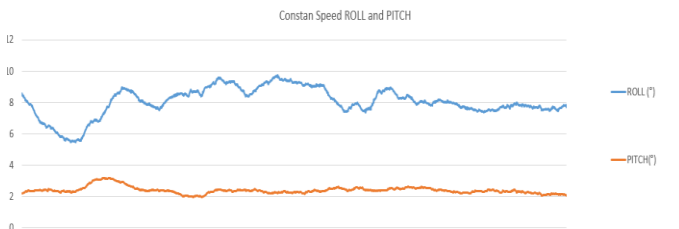
Fig.6. Low Level Fusion Output: Euler Angles

Low-level sensor fusion is planned on acceleration, deceleration and stop. Flat ground and sloping ground were used for these 3 movements. Each movement is numbered and classified.

TABLE 1
RESULTS INCLUDING LOW LEVEL FUSION

Algorithm Type	Success Rate
Decision Tree	%92.9479
Gradient Boost Regressor	%93.1992
Multi-Layer Perceptron	%87.3880
Classifier Accuracy	%83.440

After all, data were taken, the artificial intelligence algorithm was run with 2 different table data. The data collected for the artificial intelligence algorithm is divided into 60% training and 40% testing. As a result of comparing each numbered movement with any test data, the success percentages were compared. Table 1 and Table 2 were obtained by using the table with low level sensor fusion and the table without sensor fusion, respectively.

TABLE 2
RESULTS WITHOUT LOW LEVEL FUSION

Algorithm Type	Success Rate
Decision Tree	%65.483
Gradient Boost Regressor	%72.4950
Multi-Layer Perceptron	%73.0385
Classifier Accuracy	%63.24

VI.   CONCLUSION

For more than 30 years, sensor fusion techniques continue to be developed. With the renewed and newly produced sensors, the detection can be converted from very low to high levels. Self-driving vehicles, one of the most prominent examples today, use sensor fusion techniques at the point of detection and decision making.

In this study, pre-processing, processing and improvement were made by using the sensor fusion algorithm according to the JDL architecture. As a result of the experiments, high quality data that is easier to understand were obtained from the data obtained when sensor fusion algorithms are used. (Euler Angle and Linear Velocity).

High-level sensor fusion with supervised learning algorithms has been applied. At this point, it was observed that the error increased as a result of storing the data received from the sensors without processing and without being subjected to low-level sensor fusion.

All these processes are based on recording the data and then analysing the data. It has been observed that the accuracy percentage increases as the recorded data and the number of algorithm iterations increase.

REFERENCES

[1]   S.Birogul, Y. Sonmez, U.guvenc," Veri Füzyonuna Genel Bir Bakış", Politeknik Dergisi Vol.10, No.3, pp. 235-240, 2007.
[2]   W. Elmenreich," An Introduction to Sensor Fusion", Research Report 47/2001.
[3]   Weisstein W., E., "Eric Weissten's world of science", www.scienceworld. wolfram.com/physic (2003).
[4]   Llinas, J.,Waltz, E., "Multisensor data fusion", Boston, MA: Artech House, (1999).
[5]   O.Saber, R., J.S Shamma. "Consensus Filters for Sensor Networks and Distributed Sensor Fusion." Seville, Spain: IEEE, Dec. 2005.
[6]   M. Kam, X. Zhu, P. Kalata, "Sensor fusion for mobile robot navigation," 1997, Proceedings of the IEEE(Volume: 85, Issue: 1, Jan. 1997).
[7]   J.K. Hackett, M. Shah, "Multi-sensor fusion: a perspective," 1990 IEEE International Conference on Robotics and Automation, Cincinnati, OH, USA.
[8]   O. Kreibich, J. Neuzil, "Quality-Based Multiple-Sensor Fusion in an Industrial Wireless Sensor Network for MCM," IEEE Transactions on Industrial Electronics,. 09.2014.

[9]   Samancı.B,   "Accelerometer,   Gyroscope,   IMU   nedir?",
        "http://www.barissamanci.net/Makale/26/accelerometer-gyroscope-
        imu-nedir/",(05.2020).
[10]   https://learn.sparkfun.com/tutorials/accelerometer-basics/how-to-select-
        an-accelerometer,(06.05.2021).
[11]   https://learn.sparkfun.com/tutorials/gyroscope/all , (06.05.2021).
[12]   T.C. Akinci. "Applications of big data and AI in electric power systems
        engineering." *AI and Big Data's Potential for Disruptive Innovation*. IGI
        Global, 2020. 240-260.
[13]   H.S. Nogay, T.C. Akinci, and M. Yilmaz. "Comparative experimental
        investigation and application of five classic pre-trained deep
        convolutional neural networks via transfer learning for diagnosis of
        breast cancer." *Advances in Science and Technology. Research Journal,
        vol.* 15, no. 3, 2021, pp.1-8. https://doi.org/10.12913/22998624/137964
[14]   O. Akgun, A. Akan, H. Demir, T.C. Akinci, "Analysis of Gait Dynamics
        of ALS Disease and Classification of Artificial Neural
        Networks." *Tehnički vjesnik* 25.Supplement 1 (2018), pp.183-187.
        https://doi.org/10.17559/TV-20160914144554
[15]   O. Turk, D. Ozhan, E. Acar, T.C. Akinci, and M. Yilmaz. "Automatic
        detection of brain tumors with the aid of ensemble deep learning
        architectures and class activation map indicators by employing magnetic
        resonance images." *Zeitschrift für Medizinische Physik* (2022).
        https://doi.org/10.1016/j.zemedi.2022.11.010
[16]   https://www.st.com/en/ecosystems/x-nucleo-iks01a2.html,
        (04.04.2021).

BIOGRAPHIES

Yasin ALYRAPRAK was born in 1994, in Sivas, Turkey. He received the B.S. degree in Mechatronic Engineering from the University of Marmara, İstanbul, Turkey in 2016. From 2016 and he still works as Embedded System Developer.

Gokhan GOKMEN was born in 1974. He received B.S, M.S and PhD degrees from Marmara University, Istanbul, Turkey. He has been working as a full professor at Marmara University. His current interests are measurement method, signal processing and artificial intelligence techniques

ISSN: 2147-284X   http://dergipark.gov.tr/bajece

Publication Ethics

The journal publishes original papers in the extensive field of Electrical-electronics and Computer engineering. To that end, it is essential that all who participate in producing the journal conduct themselves as authors, reviewers, editors, and publishers in accord with the highest level of professional ethics and standards. Plagiarism or self-plagiarism constitutes unethical scientific behavior and is never acceptable.

By submitting a manuscript to this journal, each author explicitly confirms that the manuscript meets the highest ethical standards for authors and coauthors

The undersigned hereby assign(s) to *Balkan Journal of Electrical & Computer Engineering* (BAJECE) copyright ownership in the above Paper, effective if and when the Paper is accepted for publication by BAJECE and to the extent transferable under applicable national law. This assignment gives BAJECE the right to register copyright to the Paper in its name as claimant and to publish the Paper in any print or electronic medium.

Authors, or their employers in the case of works made for hire, retain the following rights:

1. All proprietary rights other than copyright, including patent rights.
2. The right to make and distribute copies of the Paper for internal purposes.
3. The right to use the material for lecture or classroom purposes.
4. The right to prepare derivative publications based on the Paper, including books or book chapters, journal papers, and magazine articles, provided that publication of a derivative work occurs subsequent to the official date of publication by BAJECE.
5. The right to post an author-prepared version or an of ficial version (preferred version) of the published paper on an i nternal or external server controlled exclusively by t he author/employer, pr ovided that (a) such posting is noncommercial in nature and the paper is made available to users without charge; (b) a copyright notice and full citation appear with the paper, and (c) a link to BAJECE's official online version of the abstract is provided using the DOI (Document Object Identifier) link.

CONTENTS

BALKAN JOURNAL OF
ELECTRICAL & COMPUTER ENGINEERING

(An International Peer Reviewed, Indexed and Open Access Journal)