Research Article
Academic Platform Journal of Engineering and Smart Systems (APJESS) 13(2), 31–41, 2025
https://dergipark.org.tr/en/pub/apjess
Received: 06-06-2024, Accepted: 04-03-2025
https://doi.org/10.21541/apjess.1495405

Real-Time Detection of Turkish Sign Language Letters and Numbers with Deep Learning

*1 Abdil KARAKAN, 2 Yüksel OĞUZ

*1 Corresponding Author, Department of Electricity and Energy, Afyon Kocatepe University, Türkiye, akarakan@aku.edu.tr
2 Department of Electrical-Electronics Engineering, Afyon Kocatepe University, Türkiye, yukseloguz@aku.edu.tr

Abstract

The visual language that hearing or speech-impaired individuals communicate with through facial expressions and hand movements is called sign language. The rate of reading and writing sign language is very low. For this reason, hearing or speech-impaired individuals have great difficulty in communicating with other people, especially when benefiting from services such as hospitals and education. In this study, real-time sign language detection and display on the computer screen were performed with deep learning. The movements of hearing or speech-impaired individuals shown with their hands and fingers are detected in front of the camera. As a result of detection, the letter corresponding to the movement is recognized and displayed on the computer screen. YOLOv8 architecture was used in this method. First, a data set was created for the study. The data set consists of 29 letters and 10 numbers. Photographs of sign language movements from 100 different people were taken in the data set. Different changes were made to the photographs in the data set. With these additions, the error that may occur due to any distortion that may occur from the camera was minimized. With the changes made to the photographs, the number of photographs forming the data set increased to 11079. As a result of the study, average stability was 90.7%, mAP was 85.8%, and recall was 81.4%.

Keywords: Sign Language; Sign Language Recognition; Letter and Number Detection; Deep Learning; YOLOv8

1. INTRODUCTION

Communication is the medium in which the information to be transmitted is understood by both the sender and the receiver [1]. There is a process of transferring information from a sender to a receiver. Communication is a necessity for people to understand each other, exchange information, and meet their needs. Individuals with hearing or speech disabilities have difficulties in communication [2]. Sign language was developed to eliminate these problems. Sign language is a visual language in which hearing or speech-impaired individuals communicate with facial expressions and hand movements [3]. With this language, hearing or speech-impaired individuals can communicate with each other very easily. Hearing or speech-impaired individuals have difficulty understanding other people and expressing themselves. In order to overcome this difficulty, it is extremely important for every person, whether disabled or not, to learn sign language [4]. However, the number of people who know sign language is very low. For this reason, the need for sign language recognition systems is increasing day by day [5]. Sign language recognition systems detect and translate the letter corresponding to hand and finger movements. The reason for using these systems is that they make communication between hearings or speech-impaired individuals with other people more understandable and healthy through sign language [6].

In recent years, there has been an increase in studies on sign language recognition systems with the use of new hardware devices. Sign language recognition systems are generally seen to be hardware device-based, using electronic gloves, or image-based, using colored gloves [7-10]. Since such systems are not useful enough and hardware device-based systems are costly; especially many image-based studies have been done and continue to be done. Image-based sign language recognition systems are basically divided into two: systems that use auxiliary hardware and those that do not. In most of the studies using auxiliary hardware devices, the Microsoft Kinect sensor was used. Apart from this, hand-tracking devices such as Leap Motion and Intel RealSense have also been used.

2. LITERATURE REVIEW

In order to overcome the shortcomings of existing datasets, a new public dataset named LIBRAS-UFOP is introduced. The dataset has data that is categorized according to the concept of minimal pairs, correctly labeled, and verified by a sign language expert. The dataset contains full RGB-D (color and depth) and skeleton data. There are 56 different signs in total and these signs are divided into four categories with high similarity levels. In addition, a baseline method is presented that generates dynamic images for each multimodal data and uses these images as input to two-stream CNN architectures. An experimental protocol is proposed to evaluate the proposed dataset. According to the experimental results, the recognition rate is reported as 74.25% due to the high similarity of the signs in the dataset. This result reveals the difficulty of the dataset and provides opportunities for future research to improve the recognition rate [11]. It presents a hand gesture recognition framework for Pakistan Sign Language (PSL) that processes gesture images obtained using Kinect motion sensor and achieves 98.74% accuracy with an improved Convolutional Neural Network (CNN) model. The proposed model has demonstrated impressive performance with low error rate using features extracted from PSL images using SIFT algorithm [12,13]. In the study conducted by Quesada et al., Leap Motion and Intel RealSense hand-tracking devices were used. When classification was performed with the SVM method using more than 50 individuals, 93% success was achieved [14]. The accuracy rates of studies conducted with auxiliary equipment have not increased to very high levels [15-18]. In addition, the use of extra hardware in these systems is both costly and difficult to use.

In recent years, a lot of work has been done on systems that do not use auxiliary hardware devices in image-based sign language recognition [19-22]. Sign language recognition (SLR) was developed to facilitate the understanding of sign language used by deaf individuals by other people. This review describes SLR techniques, especially those based on algorithms in recent years, comparing datasets, traditional and deep learning methods. The study aims to guide future studies in sign language recognition research by presenting the basic principles of the methods [23]. The proposed method achieves 99.13% accuracy with the BMCNN architecture, which is created by combining one-dimensional CNN and the improved multi-head attention mechanism containing BiLSTM. The results show that the proposed approach is superior to traditional machine learning methods and other current techniques [24]. Aims to use computer vision algorithms and deep learning methods to recognize American and British Sign Languages (ASL and BSL) for the hearing impaired. The model offers the ability to convert sentences into text by recognizing gestures in sign language with high accuracy and speed with Mediapipe keypoint detection and CNN architecture. This system, which can be used in educational environments and sign language learning, helps users improve their sign language skills with its real-time feedback feature [25]. Sumaya and colleagues used YOLOv7 architecture for sign language detection in Bangladesh. YOLO architecture is among the latest deep-learning networks. It gives very good results in real-time detection. In their study, they prepared a data set consisting of 3760 images. The photographs in the dataset were first converted to 416 × 416 dimensions. As a result of the work done with YOLOv7 and its architecture, the accuracy rate varies between 85% and 97%. The biggest reason for this is that the detection rates of each sign in sign language are different. When some signs are detected very easily, the accuracy rate is very high. Some signs are very difficult to detect. This reduces the accuracy rate significantly. In the study, Detectron2 and EfficientDet-D0 architectures were also used. The best result was achieved with an accuracy rate of 94.915% in the Detectron2 architecture [26].

Sign language varies depending on the language used in each country. Sometimes even countries that use the same language have different sign languages. Although America and England use the same language, their sign languages are different. Turkish sign language is different. Many studies have been carried out on Turkish sign language. Çelik and Odabaş used Convolutional Artificial Networks (CNN) and Long Short Term Memory (LSTM) deep learning techniques in sign language detection. First, they created a data set with images taken from the camera. Care was taken to ensure that the head area and hands were visible in the photographs that constitute this data set. As a result of the implementation of the study, a success rate of 97% was achieved. In the detection process, the head is detected first. Then hand detection takes place. In this case, the detection process takes a lot of time. An accuracy rate of 97% was achieved on a single image. In real-time detection, the accuracy rate drops to 57%. The biggest reason for this is that the frame per-second-rate (FPS) is very low [27]. Toğaçar and his colleagues worked on detecting Turkish sign language from a ready-made image set. They used Siamese neural networks in this study. In their study, they identified only numbers from Turkish sign language. As a result of the study, an accuracy rate of 98.16% was obtained. Real-time detection was not performed in their study [28]. Since Özcan and Baştürk's studies were focused on classification, single images in RGB format were converted to grayscale in the last part of the pre-processing stage. In the next stage, parameter adjustment was carried out using the development set derived from the training set and GoogLeNet-based CNN. For parameter tuning, GS and RS were used as global search methods, and the genetic algorithm (GA) was used as the heuristic search algorithm. Although GS and RS methods have the risk of getting stuck in local minima, they are frequently used to reduce the temporal cost in parameter optimization of the deep learning model. In the experimental studies, an accuracy rate of 93.93% was achieved with the GA-supported GoogLeNet-based CNN model. An accuracy rate of 93.63% was achieved with the GS-supported GoogLeNet-based CNN model. On the other hand, success rates of 89.81% were achieved with GoogLeNet-based CNN with RS support and 88.62% with GoogLeNet-supported CNN using only default parameters [29]. In their work, Pacal and Alafttekin focus on the use of deep learning methods to classify numbers in Turkish sign language. The performance of the most popular and current deep learning methods for classifying the signals in the data set was examined. Transfer learning and data augmentation techniques have also been used to improve the performance of deep learning models. MobileNet, VGG, EfficientNet, DenseNet, and ResNet architectures were used in their studies [30].

Many studies on sign language have been conducted in the literature. Different methods were used in these studies. Table 1 shows literature studies.

Table 1. Literature

Writer	Architecture	mAP@0.5
Talukder et al., [37]	YOLOv5x	98.56
Hoque et al., [38]	FasterRCNN	98.20
Lipi et al., [39]	Custo CNN	95.20
Angona et al., [40]	MobileNet	95.74
Shi et al., [41]	AlexNet	42
Wadhawan et al., [42]	AlexNet	99.72
Miah et al., [43]	BenSignNet	97.60
Alam et al., [44]	Custom CNN	97.57
Angona et al., [45]	MobileNet	95.71
Shamrat et al., [46]	Custom CNN	99.80
Çelik et al., [27]	CNN	53.8
Almeida et all [18]	SWM	80.2
Ameen et all [23]	CNN	91.3
Rao et all [25]	CNN	87.8
Siddique et al., [26]	YOLOv7	94.91
Halbuni et al., [33]	ResNet50	98.50

It is seen that very different accuracy rates have been achieved in literature studies. The accuracy rate does not depend only on the architecture used in the study. The data set is the biggest factor affecting the accuracy rate. No ready data set was used in the study. Help was taken from 100 people to create the data set. Thus, a high accuracy rate was achieved.

Another aspect of the study is that the detection speed is very high. The frame rate per second (FPS) in the study is 60. Thus, the detection of movements in front of the camera was achieved at a very high speed.

With the study, 29 letters and 10 numbers in Turkish sign language were detected in real-time. The detection process result is shown on the computer screen. In this way, communication of hearing or speech-impaired individuals is ensured. To achieve this, images of 29 letters and 10 numbers in Turkish sign language were taken from 100 different people. Additionally, grayscale, tilt addition, blurring, variability addition, noise addition, image brightness change, color vibrancy change, perspective change, resizing, and position change have been added to the photos. With these additions, the error that may occur due to any distortion that may occur from the camera is minimized. With the changes made to the photographs, the number of photographs forming the data set increased to 11079. The work has been implemented in the YOLOv8 architecture. As a result of the application, average stability was 98.3%, mAP was 95.8% and recall was 91.4%.

In the study, a new dataset was created in Turkish sign language. The changes made to this dataset increased its suitability for real life. The latest deep learning architecture YOLOv8 was used in the study. In addition, the study was used in real-time. As a result of the study, a high level of accuracy was achieved.

3. MATERIALS AND METHODS

Although hearing and speech-impaired individuals can establish healthy communication among themselves, they have serious difficulties in communicating with other people. In order to minimize the miscommunication experienced in these situations, in our study, sign language was detected and the letters and numbers corresponding to the movement were determined. Thus, the communication problems experienced by hearing or speech-impaired individuals have been eliminated. With the study, 29 letters and 10 numbers in Turkish sign language were detected in real-time. Figure 1 shows the general operating scheme of the system.

3.1. Data Set

Letter-based sign language is also called finger alphabet. Each movement made with the finger alphabet corresponds to a letter in the alphabet. In communicating with the finger alphabet, each letter is expressed using the hands. Words are formed from letters. The finger alphabet is used in all communication. It is especially used in the expression of proper names, in expressing words that are not used very often, and in words borrowed from foreign languages. In addition, the finger alphabet is used in abbreviations, suffixes, scientific terms, and synonyms when the meaning cannot be fully expressed [31].

In the study, the data set consists of 100 different people. Each person was photographed with all the digits and numbers present in Turkish sign language. Figure 3 shows some of the photographs used in the data set.

A new and original dataset was prepared in the study. Different filters were made on the dataset. In this way, its suitability for real error was increased. Figure 4 shows the different filters applied to the dataset.

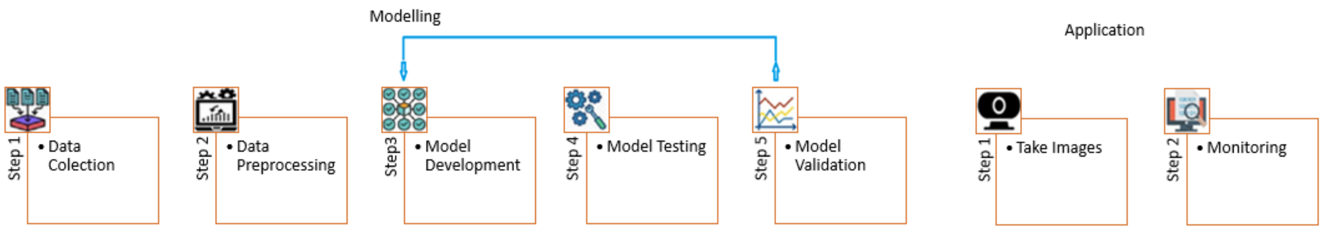

Figure 1. General operating scheme of the system.

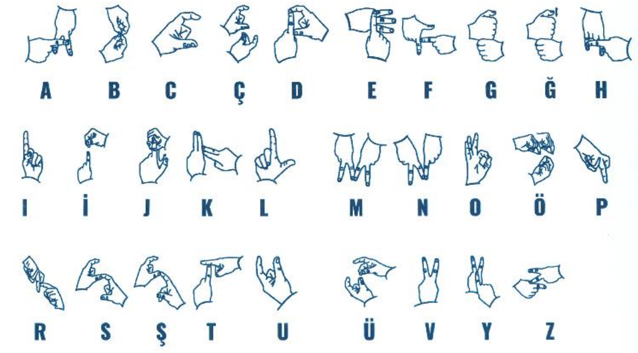

Figure 2. Turkish language finger alphabet.

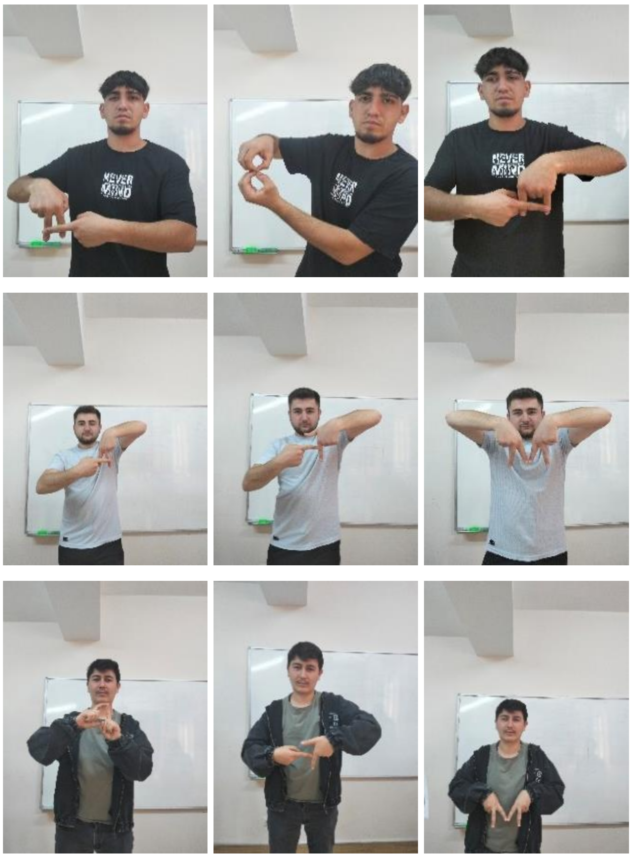

Figure 3. Some of the photographs are used in the data set.

In the study, real-time classification is performed. In this study, the snapshot taken from the camera is processed. Errors that may occur in the camera affect the classification result. For this reason, errors that may occur in the camera are added to the data set. The original image in the data set is seen in Figure 4a. Grayscale has been added to the data set in Figure 4b. A 15º rotation has been added to the images in Figure 4c. In Figure 4d, the image has been cropped by 10º in four different directions. Color fading has been added to the images in the data set in Figure 4e. Color flashing has been added to the images in Figure 4f. In Figure 4g, 2.5 pixels of blur has been added to the images in the data set. 0.1 pixels of noise has been added to Figure 4h. A 10% crop has been added to Figure 4k. The image has been inverted in Figure 4l. The characteristics of the dataset are shown in Table 2

.

Table 2. Dataset properties

Class	Train (%70)	Validation (%20)	Test (%10)
Letter	5684	1624	812
Number	2070	590	300
Total	7754	2214	1112

The dataset consists of 29 letters and 10 numbers. The letters and numbers in the dataset are divided into 70% train, 20% validation, and 10% test. There are 280 samples from each number. There are 300 samples from each number. The number of samples in each number is 20 more than the number of samples in the letters. The data used for training in the data set was not used in the testing phase.

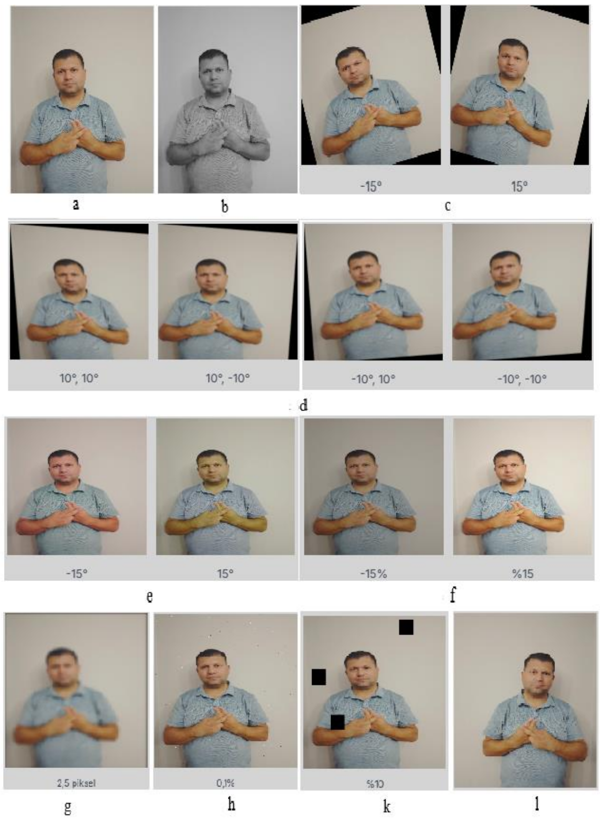

Figure 4. a) Original image **b)** Grayscale **c)** Rotation **d)** Cropping **e)** Color fade **f)** Color flash **g)** Blur **h)** Noise **k)** Cut **l)** Flip

3.2. Deep Learning

With the rapid development of computer technologies, the rapid spread of information and the abundance of data have emerged. It is aimed to obtain the desired information quickly and accurately in the abundance of data. For this purpose, learning techniques such as data mining, data analytics, and machine learning have been proposed. In recent years, deep learning, one of the subfields of artificial intelligence and machine learning, has become the most common computational approach, being used in many scientific disciplines. Very good results have been achieved in this field. [32]. Traditional machine-learning techniques are limited to extracting features from raw data [33]. Deep learning, on the other hand, gradually extracts features from unstructured or unlabeled raw data. It is an approach developed from machine learning that performs an accurate classification from this extracted data.

By processing pixels in an image, deep learning makes the learning task simpler by making the intensity vector value per pixel, a series of edges, or regions of a certain shape more abstract [34]. It facilitates the training of deep learning on huge data sets. With the advancement of this technology, the processing speed in deep learning has increased as the graphics processing units developed provide tremendous parallel computing power. This has made deep learning popular in real-world applications.

Deep learning utilizes machine learning methodologies. However learning mechanisms are different from each other. In the traditional machine learning process, training is performed using supervised learning algorithms on the processed data set. In machine learning, the feature extraction step is done externally, manually. In deep learning, the feature extraction step is automatically performed by the multilayer neural network during training using unlabeled raw data [35]. The ability to use raw data during training is one of the unique features of deep learning algorithms [36]. Providing the necessary information directly from the raw data by using hierarchical structures in multilayer neural networks leads to a more accurate result.

3.3. YOLOv5

The YOLOv5 algorithm was developed by Joseph Redmon. The YOLOv5 architecture was developed by an artificial intelligence research company called Ultralytics. This model was released in 2020 as a single-stage target recognition algorithm. It is an open-source software. The YOLOv5 algorithm has significantly higher detection accuracy and speed compared to previous versions [23]. The YOLOv5 algorithm is divided into four different network model architectures, increasing in size and amount of model parameters. These are; YOLOv5s, YOLOv5m, YOLOv5l and YOLOv5x architectures. Among these architectures, the YOLOv5s network has the fastest computational speed and the lowest average precision. The YOLOv5x network has the exact opposite feature of the YOLOv5s network [24]. The YOLOv5 model architecture is shown in Figure 5.

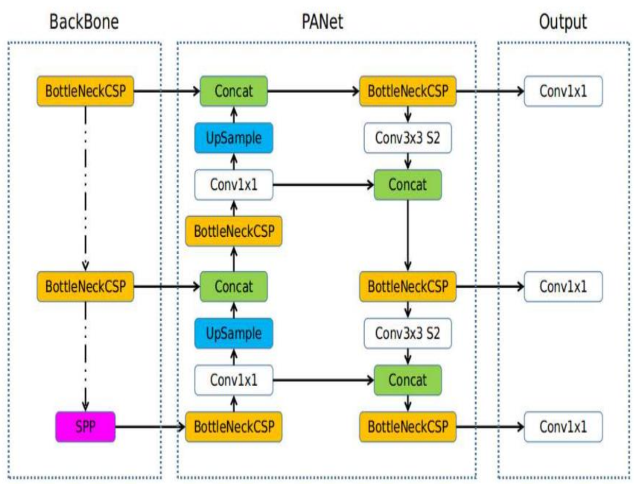

Figure 5. YOLOv5 model architecture

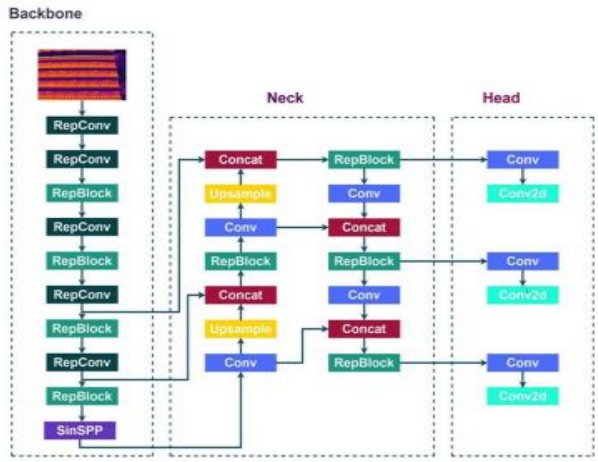

Figure 6. YOLOv6 model architecture.

3.4. YOLOv6

YOLOv6 is an advanced object detection model designed to provide both high accuracy and efficiency. It builds upon the foundations of previous YOLO (You Only Look Once) versions while introducing several optimizations to enhance speed and precision. Developed with real-time applications in mind, YOLOv6 is particularly suitable for scenarios where fast and accurate object detection is crucial, such as autonomous vehicles, video surveillance, and robotics.

One of the key features of YOLOv6 is its anchor-free detection mechanism, which eliminates the need for predefined anchor boxes, reducing computational complexity and improving inference speed. The model also benefits from techniques such as knowledge distillation, re-parameterization, and improved loss functions, which contribute to better feature representation and overall performance. These enhancements allow YOLOv6 to outperform previous versions while maintaining a lightweight structure, making it ideal for deployment on edge devices and embedded systems.

Furthermore, YOLOv6 supports multiple model sizes, enabling users to select an appropriate balance between accuracy and speed based on their specific application needs. It also incorporates advancements in convolutional network architectures, enabling better feature extraction and improved robustness across various datasets. YOLOv6 is a state-of-the-art object detection model that combines efficiency, flexibility, and accuracy. Its innovations make it a powerful tool in fields that require fast and precise detection, from industrial automation to smart surveillance and beyond [27]. Figure 6. The YOLOv6 model achieves higher accuracy by using fewer parameters than previous YOLO architectures.

3.5. YOLOv7

YOLOv7 is an advanced deep learning model that offers high speed and accuracy in object detection. Compared to previous YOLO versions, it features a more optimized architecture, making it highly effective for real-time object recognition tasks. Its fast processing capabilities make it suitable for various applications, including security cameras and autonomous vehicles. Tests conducted on the COCO dataset have shown that YOLOv7 outperforms its predecessors in both accuracy and inference speed. The model incorporates improved data processing and training strategies to enhance detection precision. Built on PyTorch, it provides flexibility and can be easily trained with different datasets. With applications in object tracking, facial recognition, healthcare, retail analytics, and autonomous systems, YOLOv7 stands out as one of the most efficient object detection models available today [19]. Figure 7 shows the YOLOv7 model architecture.

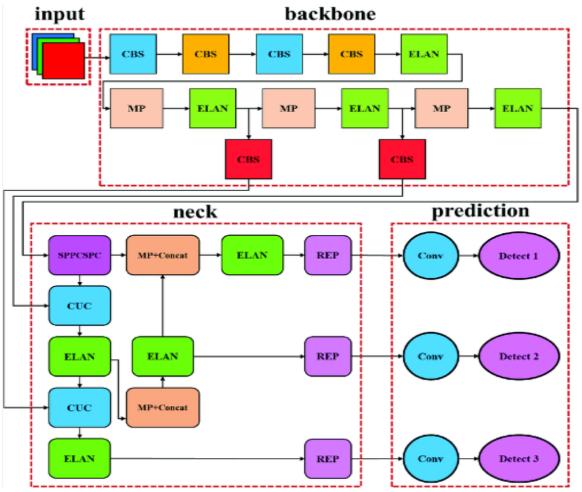

Figure 7. YOLOv7 model architecture.

3.6. YOLOv8

YOLOv8 is the latest evolution in the YOLO object detection series, designed for improved accuracy, efficiency, and ease of use. This version introduces a more refined neural network architecture, enabling better performance across a wide range of detection and segmentation tasks. One of its key advantages is its lightweight yet powerful design, making it suitable for both edge devices and high-performance computing environments. YOLOv8 supports tasks such as object detection, instance segmentation, and key point detection, expanding its usability beyond traditional applications. With enhanced training techniques and optimized inference speed, it delivers superior results in real-time scenarios. Built on Ultralytics' framework, it offers seamless integration with modern machine learning workflows, allowing developers to fine-tune models for specific applications with minimal effort. Whether used in autonomous systems, medical imaging, or industrial automation, YOLOv8 sets a new standard for high-performance computer vision solutions [32-34]. Figure 8 shows the YOLOv8 architecture.

YOLOv8 algorithm has a new architecture, improved convolutional layers backbone and a more advanced detection head [32]. YOLOv8 algorithm works by dividing an image into a grid of smaller regions and estimating a bounding box and class probabilities for each object [28]. YOLOv8 algorithm uses the Darknet-53 backbone network, which is faster and more accurate than the YOLOv7 network to improve the feature extraction process. The Darknet-53 is a 53-layer convolutional neural network and is more efficient than previous versions due to its convolutional network [23-35]. YOLOv8 algorithm combines DarkNet-53, PS and YOLOv4 architecture to improve the accuracy and robustness of object detection.

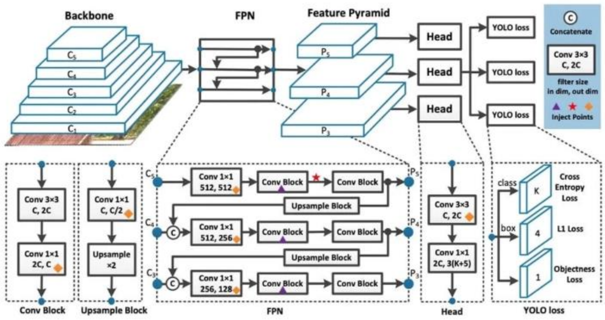

Figure 8. YOLOv8 model architecture.

3.7. Evaluation Metrics

Accuracy, recall, precision, and F1 score are widely used metrics for evaluating and comparing the performance of classification models. In this study, we employed these metrics to assess the classification effectiveness and overall performance of the proposed model. It is important to note that higher values for accuracy, precision, recall, and F1 score indicate better performance.

Accuracy provides an overall measure of a model's performance across all classes. It is calculated using the following formula:

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN}$$

Here, TP (true positives) and TN (true negatives) represent the number of correctly identified positive and negative instances, respectively, while FP (false positives) and FN (false negatives) denote the misclassified instances.

Recall, on the other hand, measures the proportion of correctly identified positive cases out of the total actual positive cases, including those that were missed. It is determined using the following formula:

$$Recall = \frac{TP}{TP + FN}$$

Precision is a metric that quantifies the proportion of correctly predicted positive instances out of all instances predicted as positive (including incorrect predictions). It is computed using the following formula:

$$Precision = \frac{TP}{TP + FP}$$

The F1 score is a balanced evaluation metric that combines recall and precision by calculating their harmonic mean. It is determined using the following formula:

$$F1\ score = 2 \times \frac{Precision \times Recall}{Precision + Recall}$$

4. RESULT AND DISCUSSION

In the literature, there are sign language recognition studies of different spoken languages. One of the most important factors showing the performance of the presented models is the data sets they use. When we examine the studies, each study uses its own special data set. This affects the objectivity of the models on different sign languages.

In our study, YOLOv8 architecture, one of the current and popular deep learning models, was used. In the study, an experimental study is presented to examine the performance of deep learning in Turkish sign language. As a result of the study, accuracy was 0.983, precision was 0.958, sensitivity was 0.981 and F1-score was 0.914. Figure 9 shows the ROC graph resulting from the study.

In the study, 29 letters and 10 numbers from Turkish sign language were identified. The representation of each letter and number is different. For this reason, there are differences in the accuracy rates of each letter and number. Table 3 shows the accuracy rates of 29 letters, and 10 numbers identified in the study. As a result of the YOLOV8 architecture, the complexity matrix consists of two parts: the predicted value and the actual value. Figure 10 shows the complexity matrix implemented in the system.

In the study, although there are very few incorrect classifications. The first reason for this is that the data set consists of movements made by different people. These people show differences when making the same sign. There may be a slight difference in the signs according to the age, gender, finger thickness and body of the people. Another reason is the images taken. Although the images are tried to be taken in the same way, there is a slight difference. The area covered in the same photo changes according to the structures of the people. Some of the signs are very similar to each other. This causes incorrect classifications. There is an error rate in the implemented architecture. This causes incorrect classification.

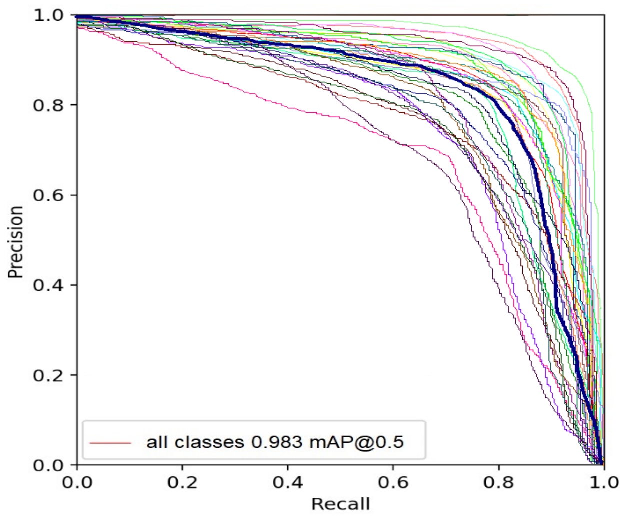

Figure 9. ROC chart resulting from the study

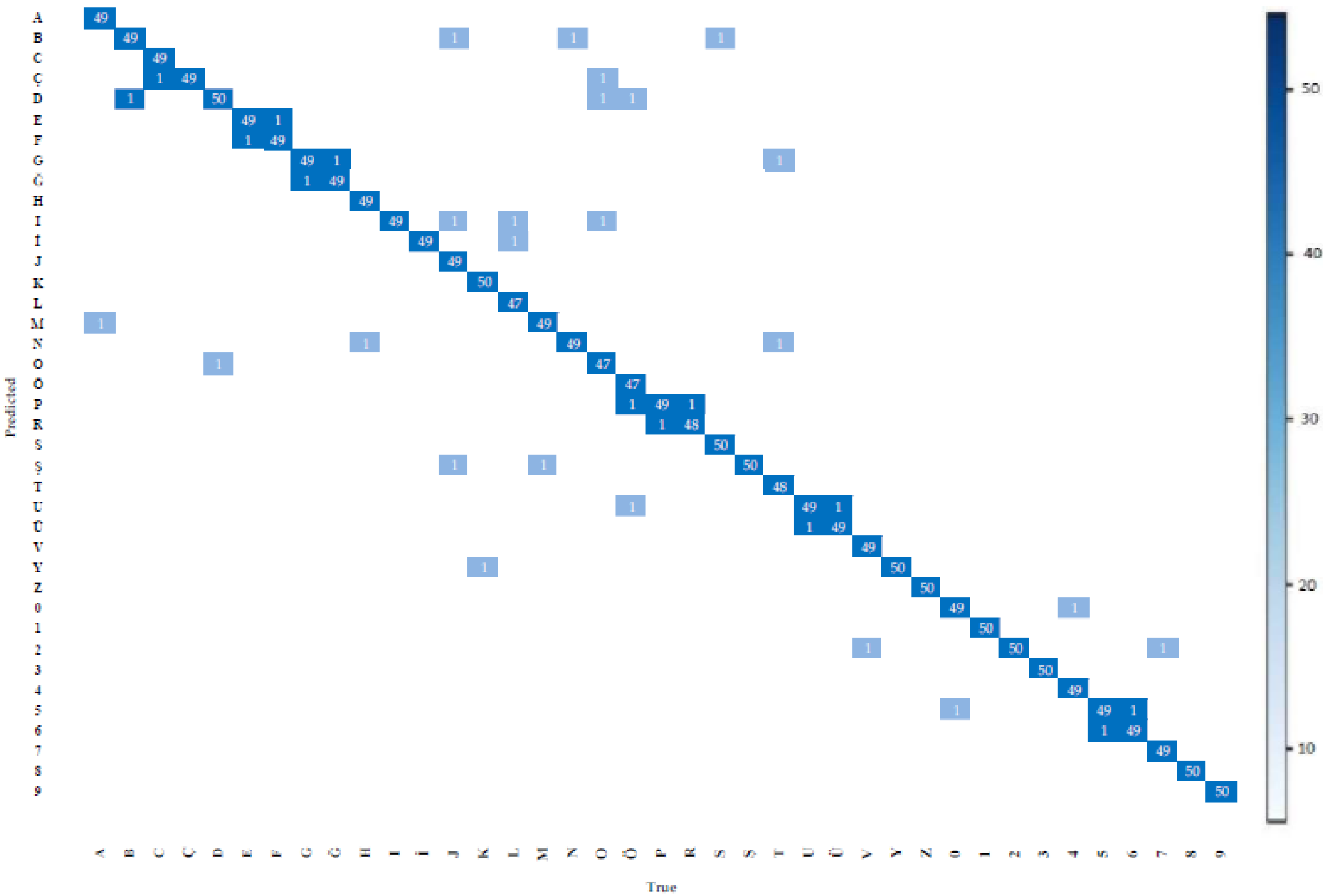

Figure 10. Complexity matrix implemented in the system.

Table 3. Accuracy rates of the 29 letters and 10 numbers detected in the study

Class	mAP@0.5(%)	Class	mAP@0.5(%)
A	99.5	R	97.5
B	99.0	S	97.7
C	97.3	Ş	98.3
Ç	98.8	T	98.8
D	99.1	U	99.1
E	99.7	Ü	98.3
F	98.2	V	97.5
G	98.7	Y	98.7
Ğ	99.2	Z	99.0
H	98.0	0	97.3
I	98.9	1	97.9
İ	99.0	2	98.5
J	99.5	3	98.1
K	99.9	4	98.6
L	98.7	5	98.8
M	99.2	6	97.9
N	97.7	7	98.5
O	97.7	8	99.1
Ö	98.3	9	98.7
P	98.5	**All Class**	**0.983**

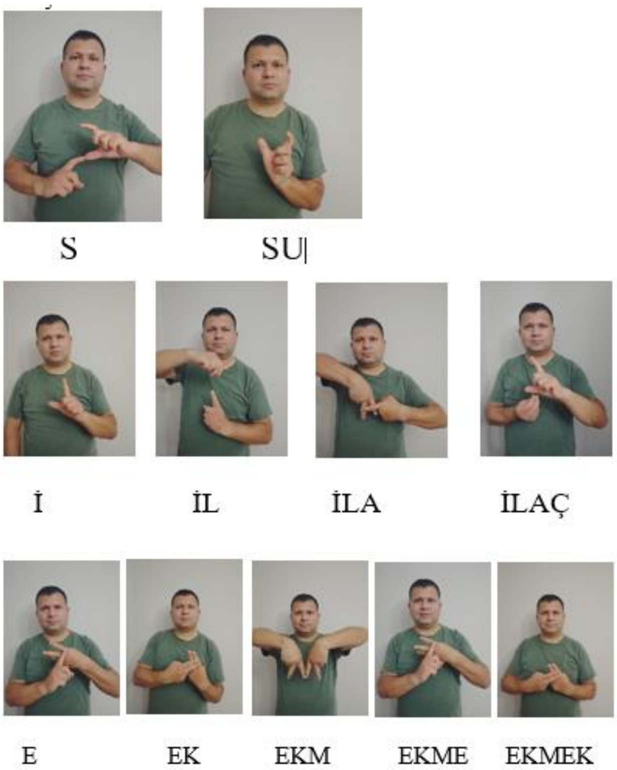

Figure 11. The study result is shown in real-time.

In the study, 29 letters and 10 numbers were classified. Different accuracy rates were obtained in this classification process. The highest classification of the letter K was achieved with 99.9% among the 29 letters. The lowest accuracy rate was achieved with 97.3% in the letter C. Among the numbers, the highest accuracy rate was 99.1% in the number 8. The lowest accuracy rate among the numbers was 97.3% in the number 0. As a result of the study, an average accuracy rate of 98.3% was achievedThe work has been tested in real-time. Images of the person were taken with a webcam. Then, the detection process was carried out in deep learning. The detection process result is shown on the computer screen. Figures 11 and 12 show the real-time study result.

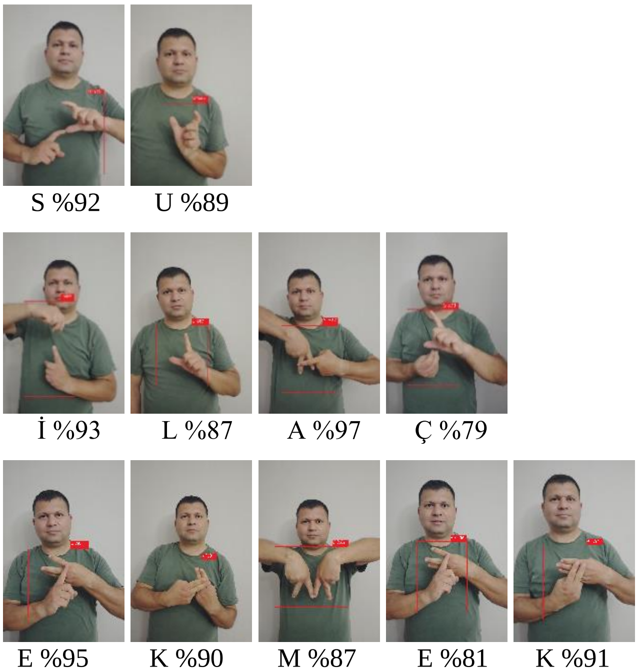

Figure 12. The study result is shown in real-time.

5. CONCLUSION

Sign language and systems have an important place in enabling deaf and mute people to communicate healthily and accurately with their environment without the need for an interpreter. However, another important factor is that sign language users adapt to daily life by using technological devices such as smartphones and tablets without the need for any external hardware. This has made sign language recognition systems popular in the field of computer vision. This study was carried out using the original data set we created for letters and numbers in Turkish sign language, using YOLOv8, one of the current deep learning architectures. In the study, an original data set was prepared instead of a ready-made data set. For this purpose, help was received from 100 different people. The data set consists of 29 letters and 10 numbers in Turkish sign language. The data set consists of a total of 11079 photographs. As a result of the study, an average accuracy rate of 98.3% was obtained. These experimental studies have revealed that existing deep-learning techniques are quite successful in classifying the images of numbers and letters in Turkish sign language. Sign language has a complex structure because it is a language that communicates using visual signs. Correctly classifying sign language, hand and gesture movements and extracting meaningful expressions from them is an important process. For this reason, classification has an important place in sign language recognition systems. Future studies will focus on word detection. The word will be identified by examining the continuous movements of deaf and mute people. This will ensure faster communication.

Author contributions: A. Karakan, data collection, processing, and writing; Y. Oğuz, analysis, interpretation and literature review. All authors have read and agreed to the published version of the manuscript.
Conflict of Interest: No conflict of interest was declared by the authors.
Financial Disclosure: The authors declared that this study has received no financial support.

REFERENCES

[1] T. Tao, Y. Zhao, J. Zhu, T. Liu, J. Kuang, "A survey on sign language recognition from perspectives of traditional and deep-learning methods", Journal of Visual Communication and Image Representation, 6, pp. 104363, 2024, https://doi.org/10.1016/j.jvcir.2024.104363.

[2] F. Nehal, T. Mohamed, F. S. Ahmed, A. M. Mahmoud "Efficient deep learning models based on tension techniques for sign language recognition", Intelligent Systems with Applications, 20, pp. 200284, 2023. https://doi.org/10.1016/j.iswa.2023.200284.

[3] J. Bora, S. Dehingia, A. Boruah, A. Chetia, D. Gogoi, "Real-time Assamese Sign Language Recognition using Media Pipe and Deep Learning", Procedia Computer Science, 218, pp. 1384-1393, 2023. https://doi.org/10.1016/j.procs.2023.01.117.

[4] Y. Zhang, X. Jiang, "Recent Advances on Deep Learning for Sign Language Recognition", Computer Modeling in Engineering and Sciences, pp. 2399-2450, 2024. https://doi.org/10.32604/cmes.2023.045731.

[5] Z. Katılmış, C. Karakuzu, "Double handed dynamic Turkish Sign Language recognition using Leap Motion with meta learning approach", Expert Systems with Applications, 228, pp. 120453, 2023. https://doi.org/10.1016/j.eswa.2023.120453.

[6] Y. Liu, X. Jiang, X. Yu, H. Ye, C. Ma, W. Wang, Y. Hu,"A wearable system for sign language recognition enabled by a convolutional neural network," Nano Energy Volume 116, 108767, 2023. https://doi.org/10.1016/j.nanoen.2023.108767.

[7] J. Hasanov, N. Alishzade, A. Nazimza, S. Dadashzad, T. Tahirov, "Development of a hybrid word recognition system and dataset for the Azerbaijani Sign Language dactyl alphabet," Speech Communication, Volume 153, 102960, 2023, https://doi.org/10.1016/j.specom.2023.102960.

[8] R. Sreemathy, MP. Turuk, S. Chaudhary, K. Lavate, A . Ushire, S. Khurana, "Continuous word level sign language recognition using an expert system based on machine learning," International Journal of Cognitive Computing in Engineering, Volume 4, pp. 170-178, 2023. https://doi.org/10.1016/j.ijcce.2023.04.002.

[9] Y. Qin, S. Pan, W. Zhou, D. Pan, Z. Li, "WiASL: American Sign Language writing recognition system using commercial WiFi devices," Measurement, Volume 218, 113125, 2023. https://doi.org/10.1016/j.measurement.2023.113125.

[10] Y. Obi, K.S. Claudio, V.M. Budiman, S. Achmad, A. Kurniawan, "Sign language recognition system for communicating to people with disabilities," Procedia Computer Science, Volume 216, pp. 13-20, 2023. https://doi.org/10.1016/j.procs.2022.12.106.

[11] L. R. Cerna, E. E. Cardenas, D. G. Miranda, D. Menotti, G. Camara-Chavez, "A multimodal LIBRAS-UFOP Brazilian sign language dataset of minimal pairs using a microsoft Kinect sensor", Expert Systems with Applications, Volume 167, 1l4179, 2021. https://doi.org/10.1016/j.eswa.2020.114179.

[12] S. Arooj, S. Altaf, Ş. Ahmed, H. Mahmud, A. Ş N. Muhammed, "Enhancing sign language recognition using CNN and SIFT: A case study on Pakistan sign language" Journal of King Saud University - Computer and Information Sciences, Volume 36, Issue 2, 101934 2024. https://doi.org/10.1016/j.jksuci.2024.101934.

[13] B. Alsharif, E. Alalwany, M. Ilyas, "Transfer learning with YOLOV8 for real-time recognition system of American Sign Language Alphabet" Franklin Open, Volume 8, 100165, 2024. https://doi.org/10.1016/j.fraope.2024.100165.

[14] S. Qahtan, H.A. Alsattar, A.A. Zaidan, M. Deveci, D. Pamucar, L. Martinez, "A comparative study of evaluating and benchmarking sign language recognition system-based wearable sensory devices using a single fuzzy set," Knowledge-Based Systems, Volume 269, 110519, 2023. https://doi.org/10.1016/j.knosys.2023.110519.

[15] N. Musthafa, C.G. Raji, "Real time Indian sign language recognition system," Materials Today: Proceedings, Volume 58, pp.504–508, 2022. https://doi.org/10.1016/j.matpr.2022.03.011.

[16] S. Katoch, V. Singh, U. S. Tiwary, "Indian Sign Language recognition system using SURF with SVM and CNN," Array, Volume 14, 100141, 2022. https://doi.org/10.1016/j.array.2022.100141

[17] K. Wangchuk, P. Riyamongkol, R. Waranusast, "Real-time Bhutanese Sign Language digits recognition system using Convolutional Neural Network," ICT Express, Volume 7, Issue 2, pp. 215-220, June 2021. https://doi.org/10.1016/j.icte.2020.08.002.

[18] M.A. Ahmed, B.B. Zaidan, A.A. Zaidan, M. Salih, Z. T. Al-qaysi, A.H. Alamoodi, “Based on wearable sensory device in 3D-printed humanoid: A new real-time sign language recognition system,” Measurement Volume 168, 108431, 2021. https://doi.org/10.1016/j.measurement.2020.108431.

[19] R. Rastgoo, K. Kiania, and S. Escalerab, “Sign Language Recognition: A Deep Survey,” Expert Systems with Applications, 164, pp. 113794, 2021. htpss://doi.org/10.1016/j.eswa.2020.113794.

[20] A. A. Mujeeb, A. H. Khan, S. Khalid, M. S. Mirza, S. J. Khan, “A neural-network based web application on real-time recognition of Pakistani sign language,” Engineering Applications of Artificial Intelligence, Volume 135, 108761, 2024. https://doi.org/10.1016/j.engappai.2024.108761.

[21] A. O. Salau, N. K. Tamiru, B. T. Abeje, “Derived Amharic alphabet sign language recognition using machine learning methods,” Helijon, Volume 10, Issue 19, 15 2024, https://doi.org/10.1016/j.heliyon.2024.e38265.

[22] Y. Nam, and C. Lee, “Cascaded convolutional neural network architecture for speech emotion recognition in noisy conditions,” Sensors, 21(13), pp. 4399, 2021. https://doi.org/10.3390/s21134399.

[23] T. Tao, Y. Zhao, J. Zhu, T. Liu, J. Kuang, “A survey on sign language recognition from perspectives of traditional and deep-learning methods” Journal of Visual Communication and Image Representation, 104363, 2024. https://doi.org/10.1016/j.jvcir.2024.104363.

[24] W. Hao, C. Hou, Z. Zhang, X, Zhai, L. Wang, G. Lv, “A sensing data and deep learning-based sign language recognition approach”, Computers and Electrical Engineering, Volume 118, 109339, Part A, 2024. https://doi.org/10.1016/j.compeleceng.2024.109339.

[25] S. N. Koyineni, G. K. Sai, K. Anvesh, T. Anjali, “Silent Expressions Unveiled: Deep Learning for British and American Sign Language Detection” Procedia Computer Science, Volume 233, 2024, Pages 269-278, https://doi.org/10.1016/j.procs.2024.03.216.

[26] S. Siddique, S. Islam, E. E. Neon, T. Sabbir, I. T. Naheen, and R. Khan, “Deep Learning-based Bangla Sign Language Detection with an Edge Device,” Intelligent Systems with Applications, 18, pp. 200224. 2023. https://doi.org/10.1016/j.iswa.2023.200224 30 March 2023.

[27] Ö. Çelik, and A. Odabaş, “Sign2Text: Turkish sign Language Recognition Using Convolutional Neural Networks” European Journal of Science Technology. 19, pp. 923-934, Aug 2020.

[28] M. Toğaçar, Z. Cömet and 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. 23(68), pp. 349-356, 2021. https://doi.org/10.35860/iarej.700564.

[29] T. Özcan and A. Baştürk, “ERUSLR a new Turkish sign language dataset and its recognition using hyper parameter optimization aided convolution neural network,” Journal of Faculty Engineering Architecture of Gazi Universty, 36:1, pp. 527-542, 2021. htpss://doi.org/ 10.17341/gazimmfd.746793.

[30] I. Pacal and M. Alaftekin, “CNN-Based approaches for automatic recognition of Turkish sign language,” Journal of the Instute of Science and Technology, 13(2), pp. 760-777, 2023. https://doi.org/10.55525/tjst.1073116.

[31] J. Li, L. Cheng, J. Lei, W. Xiang, “Uyghur Sign Language Recognition Based on Improved YOLOv7,” Procedia Computer Science, Volume 242, pp.512-519, 2024. https://doi.org/10.1016/j.procs.2024.08.095.

[32] L. Alzubaidi, J. Zhang, A. J. Humaidi, Y. D. Ayad Al-Dujaili, O. Al-Shamma, J. Santamaría, and L. Farhan, “Review of deep learning: concepts, CNN architectures, challenges, applications, future directions,” Journal of big Data, 8(1), pp. 1-74, 2021. htpss://doir.org/10.1186/s40537-021-00444-8.

[33] A. Halbouni, T. S. Gunawan, M. H. Habaebi, M. Halbouni, M. Kartiwi, and R. Ahmad, “Machine Learning and Deep Learning Approaches for CyberSecurity: A Review,” IEEE Access (10), pp. 19572 – 19585, 2022. htpss://doi.org/10.1109/ACCESS.2022.3151248.

[34] S. Renjith, M. Rashmi, S. Suresh, “Sign Language Recognition by using Spatio-Temporal Features,” Procedia Computer Science, Volume 233, pp. 353-362, 2024. https://doi.org/10.1016/j.procs.2024.03.225.

[35] A. Karaman, I. Pacal, A. Basturk, B. Akay, U. Nalbantoglu, S. Coskun, O. Sahin, and D. Karaboga, “Robust real-time polyp detection system design based on YOLO algorithms by optimizing activation functions and hyper-parameters with artificial bee colony (ABC),” Expert Systems with Applications, 221, pp. 119741, 2023. https://doi.org/10.1016/j.eswa.2023.119741.

[36] S. Chaudhuri, U. Dayal, and V. Narasayya, “An overview of business intelligence technology,” Communications of the ACM, 54(8), pp. 88-98, 2021. htpss:/doi.org/ 10.1145/1978542.1978562.

[37] D. Talukder, F. Jahara, S. Barua, and M. M. Haque, “OkkhorNama: BdSL image dataset for real time object detection algorithms,” IEEE Region 10 Symposium, pp. 1–6. Jeju, Korea, 23-25 August 2021. https://doi.org/10.1109/TENSYMP52854.2021.9550907.

[38] O. B. Hoque, M. I. Jubair, A. F. Akash, and M. S. Islam, “BDSL36: A dataset for Bangladeshi sign letters recognition,” Asian Conference on Computer Vision, pp. 71–86, Kyoto, Japan, 30 November- 4 December 2020. https://doi.org/10.1007/978-3-030-69756-3_6.

[39] K. A. Lipi, “Static-gesture word recognition in Bangla sign language using Convolutional Neural Network,” TELKOMNIKA (Telecommunication Computing Electronics and Control), 20, pp.1109–1116, 2022. https://doi.org/10.12928/telkomnika. v20i5.24096.

[40] T. M. Angona, “Automated Bangla sign language translation system for alphabets by means of MobileNet,” TELKOMNIKA (Telecommunication Computing Electronics and Control), 18, pp. 1292–1301, 2020. https://doi.org/10.12928/telkomnika. v18i3.15311.

[41] B. Shi, A. M. Del Rio, J. Keane, D. Brentari, G. Shakhnarovich, and K. Livescu, “Fingerspelling Recognition in the Wild with Iterative Visual Attention, 2019 IEEE/CVF International Conference on Computer Vision, Seoul, Korea, 27 October 2019 - 02 November 2019. htpss://doi.org/10.1109/ICCV.2019.00550.

[42] A. Wadhawan, P. Kumar, “Deep Learning-Based Sign Language Recognition System for Static Signs,” Neural Comput & Applic, 32(2), pp. 1-12, 2020, https://doi.org/10.1007/s00521- 019-04691-y.

[43] A. S. M. Miah, J. Shin, M. A. H. Hasan, and M. A. Rahim, “BenSignNet: Bengali sign language alphabet recognition using concatenated segmentation and convolutional neural network,” Applied Sciences, 12(8), pp. 3933, 2022. https://doi.org/10.3390/ app12083933.

[44] M. S. Alam, M. Tanvir, D. K. Saha, and S. K. Das, “Two-Dimensional convolutional neural network approach for real-time bangla sign language characters recognition and translation,” SN Computer Science, 2(5), 2022. https://doi.org/10.1007/s42979-021-00783-6.

[45] T. M. Angona, “Automated Bangla sign language translation system for alphabets by means of MobileNet,” TELKOMNIKA (Telecommunication Computing Electronics and Control), 18, pp. 1292–1301, 2020. https://doi.org/10.12928/telkomnika. v18i3.15311.

[46] F. M. Shamrat, “Bangla numerical sign language recognition using convolutional neural network CNNs,” Indonesian Journal of Electrical Engineering and Computer Science, 23, pp. 405–413, 2021. https://doi.org/10.11591/ijeecs.v23.i1.pp405-413.