Düzce University
Journal of Science & Technology

Research Article

Comparative Performance Analysis of Time-Frequency Domain Images and Raw Signal Data for Classification of ECG Signals

Turgut ÖZSEVEN[a,*]

[a] *Department of Computer Engineering, Tokat Gaziosmanpaşa University, Tokat, 60100 Turkey*
** Corresponding author e-mail: turgut.ozseven@gop.edu.tr*
DOI: 10.29130/dubited.1236072

ABSTRACT

ECG signals are one of the most common tools used to diagnose cardiovascular diseases. ECG signals are obtained by measuring electrical changes on the skin surface. Arrhythmias occurring in the heart are diagnosed because the expert evaluates ECG signals. This diagnosis depends on the experience of the specialist and is a subjective evaluation. With the widespread use of computer-aided diagnostic systems, evaluations dependent on the expert's experience are objectified, and support is provided to the physician for diagnosis. For computer-aided ECG classification, beats are detected from ECG signals, and arrhythmias are detected by analyzing the structure of these beats. In recent years, deep learning models have been successful in classifying ECG signals. The data to be used in the classification process is realized with the help of morphological features or images of the signal. The main objective of this study is to compare the classification performance of digital and visual heartbeat data for ECG signal classification. For this purpose, 1D-CNN and 2D-CNN architectures are used for the type of ECG signals. As inputs of the 1D-CNN model, numerical values of the heartbeat signal and hand-crafted features obtained from these numerical values were used. The inputs of the 2D-CNN model are the raw signal image, spectrogram, scalogram, Mel-spectrogram, GFCC, and CQT images, which are visual representations of the heartbeat signal. The results show that the numerical model of the ECG signal fails for classification, while the hand-crafted features provide 85.2% accuracy. The results obtained with the visual representation of the signal provided over 99% classification accuracy for all images. The highest success rate was 99.9% with the visualization of the raw signal. In line with these findings, the 2D-CNN architecture and the visual representation of the heartbeat signal were found to be the most suitable method for classifying ECG signals.

Keywords: *ECG, Deep learning, Heartbeat, Classification*

EKG Sinyallerinin Sınıflandırılmasında Zaman-Frekans Domenindeki Görüntülerin ve Ham Sinyal Verilerinin Karşılaştırmalı Performans Analizi

ÖZ

EKG sinyalleri kardiyovasküler hastalıkların klinik tanısı için kullanılan en yaygın araçlardan birisidir. Cilt yüzeyindeki elektriksel değişimlerin ölçülmesi ile EKG sinyalleri elde edilmektedir. EKG sinyallerinin uzmanın değerlendirmesi sonucu kalpte oluşan aritmiler teşhis edilmektedir. Bu teşhis uzmanın deneyimine bağlı olup subjektif bir değerlendirmedir. Bilgisayar destekli tanı sistemlerinin yaygınlaşması ile uzmanın deneyimine bağımlı değerlendirmeler objektifleşmekte ve hekime tanı için destek sağlanmaktadır. Bilgisayar destekli EKG sınıflandırma için EKG sinyallerinden atımlardan tespit edilmekte ve bu atımların yapısı incelenerek aritmiler tespit edilmektedir. Son yıllarda derin öğrenme modellerindeki yüksek başarı EKG sinyallerinin de

Geliş: 16/01/2023, Düzeltme: 10/05/2023, Kabul: 22/05/2023

sınıflandırılması için kullanılmaya başlanmıştır. Sınıflandırma sürecinde kullanılacak veri sinyalin morfolojik özellikleri veya görüntüsü yardımıyla gerçekleştirilmektedir. Bu çalışmanın temel amacı, EKG sinyallerinin sınıflandırılması için sayısal ve görsel kalp ritmi verilerinin sınıflandırma performanslarının karşılaştırılmasıdır. Bu amaçla, EKG sinyallerinin sınıflandırılması için 1D-CNN ve 2D-CNN mimarileri kullanılmıştır. 1D-CNN modelinin girdileri olarak kalp ritmi sinyalinin sayısal değerleri ve bu sayısal değerlerden elde edilen öznitelikler kullanılmıştır. 2D-CNN modelinin girdisi kalp ritmi sinyallinin görsel olarak temsilini içeren ham sinyal görüntüsü, spektrogram, skalogram, mel-spektrogram, GFCC ve CQT görüntüleridir. Elde edilen sonuçlar, EKG sinyallerinin sayısal temsilinin sınıflandırma için başarısız olduğunu, hand-crafted özniteliklerin %85.2 doğruluk sağladığını göstermiştir. Sinyalin görsel temsili ile elde edilen sonuçlar tüm görüntüler için %99 üzerinde sınıflandırma doğruluğu sağlamıştır. Bunlar içerisindeki en yüksek başarı ise sinyalin ham halinin görselleştirilmesi ile %99.9 olarak elde edilmiştir. Elde edilen bu bulgular doğrultusunda, EKG sinyallerinin sınıflandırılması için en uygun yöntemin 2D-CNN mimarisi ve kalp ritmi sinyalinin görsel temsili olduğunu göstermiştir.

Anahtar kelimeler: *EKG, derin öğrenme, kalp ritmi, sınıflandırma*

I. INTRODUCTION

Disorders of the heart and blood vessels are referred to as cardiovascular diseases (CVD). World Health Organization (WHO) reports and American Heart Association statistics show that CVD accounts for the majority of non-disease deaths [1]. It claims an estimated 17.9 million lives yearly, accounting for 44% of all non-communicable disease deaths worldwide [2]. Deaths from CVD are projected to reach 23.6 million in 2030 [3]. CVDs cause clotting and vascular occlusion, leading to cerebral or cardiac ischemic necrosis. As a result, the heart pumps blood poorly, and organs can be damaged [4]. As a result, early detection of cardiac arrhythmias is vital. Different ECG waveforms represent arrhythmias and contain information about heart function and condition. The Association for the Advancement of Medical Instrumentation (AAMI) has categorized arrhythmias into five main classes (N, S, V, F, Q) [5].

An electrocardiogram (ECG) is one of the most widely used methods for diagnosing CVD. Electrical activity in the heart causes electrical changes on the skin's surface. ECG provides visual monitoring of this change with the help of 12 electrodes attached to the patient's body. This facilitates diagnoses such as heart disease, high blood pressure, and heart failure using arrhythmias in the heartbeat. Furthermore, ECG is crucial in predicting short- and long-term outcomes [4]. Figure 1 shows the peaks of an ECG signal.

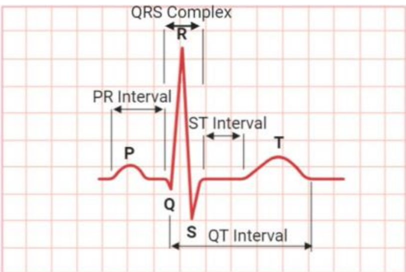

Figure 1. *The ECG constitution of a single heartbeat* [4].

Five peaks can characterize the heartbeat for the ECG waveform, as shown in Figure 1. The five points' values, distances, and various morphological characteristics are used for this purpose. P wave indicates atrial depolarization, the QRS complex wave indicates ventral depolarization, and the T wave indicates repolarization [5]. The Q, R, and S waves together indicate a single event. The length of the two intervals (PR and QT interval) means the time required for the respective electrical change to complete.

One of the main problems with ECG signal identification is that the signal varies according to the person and the disease. Another problem is that similar signals may be encountered for different diseases. Furthermore, the ECG signal has high noise and complexity characteristics, making it difficult to identify specific diseases [6]. Since the experience of the experts is an important factor in interpreting ECG signals, the result will be subjective. Therefore, computer-aided diagnosis (CAD) will provide objective evaluation for ECG, just as other medical fields do.

CAD systems have been under investigation in many medical fields for many years. For ECG, this process dates back to the 1960s [7]. Traditional CAD methods include the basic steps of pre-processing, feature extraction, and classification. The pre-processing step includes the removal of noise from the signal, framing, and windowing. Researchers have mostly used normalization and noise removal in the preprocessing step. These include z-score normalization, band pass, high pass filter, low pass filter, down-sampling, and DWT denoising [8]–[13]. The features are extracted from the time or frequency domain or the arrhythmia signal. Researchers using traditional methods have used raw data, RR intervals, discrete wavelet transforms, Fourier transforms, and morphological features [10], [11], [14]–[22]. Finally, classification is performed using the features obtained. Past studies have mainly used artificial neural networks, support vector machines (SVM), k-nearest neighbor (k-NN), and random forest (RF) classifiers [12], [21], [23]–[30]. In recent years, the high achievements obtained with deep learning models have started to be used for ECG interpretation. When deep learning studies are analyzed, studies using 1D-CNN, 2D-CNN, and transfer learning come to the forefront. 2D-CNN is generally used in image data, and the kernel is moved in 2 dimensions on the image. In 1D-CNN models, the kernel is moved in only one dimension and used in time series data. On the other hand, transfer learning involves applying CNN models available in the literature to the ECG signal. In studies using 1D-CNN, the raw data of the ECG signal and RR interval values are used [11], [15], [20], [31]–[33]. Spectrogram, log-scale spectrogram, Mel-spectrogram, bi-spectrum, ECG signal, and CQT images were used in studies using 2D-CNN and transfer learning [9], [10], [16], [18], [34]–[37]. Another deep learning model used in ECG classification studies is Long Short-Term Memory (LSTM). In studies using LSTM, raw data and RR interval values were used as input to the model [19], [21], [22], [32], [33], [38], [39].

If the studies involving the classification of ECG signals are evaluated in terms of success, the deep learning classification success is higher than the other methods, with success between 93%-99.7%. If the methods used in these studies are evaluated within themselves, the mel-spectrogram provides higher success than CQT [37]. 2D-CNN model provides higher success than 1D-CNN [31]. CNN models perform better than LSTM [33], [38]. Although recent studies have focused on deep learning models, the study using traditional classifiers (SVM and MLP) achieved 99.8% success using auto-regressive coefficients and discrete wavelet transform [12].

A general review of the research on ECG classification shows that most of the studies used CNN architectures, SVM, and k-NN classifiers proposed by the researchers. Raw data and RR intervals were mostly used as input for these methods. After the literature review, it was observed that the number of spectral image-based studies is limited, the current versions of traditional methods are not used, and there is no comparison between feature sets.

In this study, the performance of waveform, spectrogram (SPEC), Gammatone Frequency Cepstral Coefficients (GFCC), Mel-spectrogram (MEL), Constant Q-transform (CQT), and scalogram (SCL) images were compared for ECG signal classification. For this purpose, the MIT-BIH [40] dataset was used. CNN models were created for feature extraction and classification. Furthermore, the time series values of the ECG signal are used for classification using 1D-CNN, and the results are compared with 2D-CNN. Therefore, the contributions of this research paper are: 1) The classification performance of time-frequency image types of ECG signals is analyzed. 2) The classification performance of the 1D-CNN model with raw signal data and hand-crafted features is compared. 3) Guidance on using visual features is provided to researchers working in ECG classification. The architecture of this study is given in Figure 2.

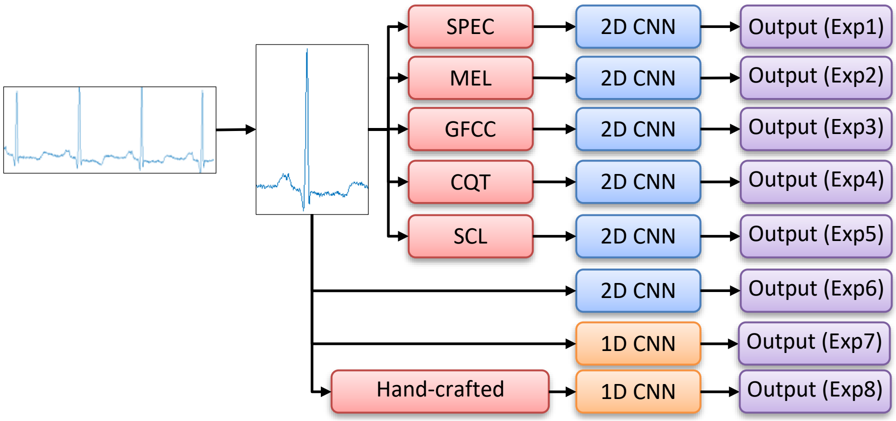

Figure 2. *Flow diagram we used for ECG classification.*

In this study, eight different experiments were conducted to classify ECG signals. The first six experiments involve the classification of time-frequency dome images with 2D-CNN. The last two experiments involve the time series of the ECG signal and the classification of the features obtained from this series with 1D-CNN. The rest of this paper is organized as follows: Section 2 presents the materials and methods used in the study. Section 3 presents the experimental results. Section 4 contains conclusions and discussion.

II. MATERIALS AND METHODS

A. DATA DESCRIPTIONS

This study used the MIT-BIH [40] database, widely used in the literature. The database contains heartbeat recordings collected by the Massachusetts Institute of Technology and Boston Hospital from 47 participants. The recordings are 30 minutes long and 360 Hz and contain 48 ECG recordings. Each signal was filtered with a 0.1-100 Hz band-pass filter. The database contains arrhythmic signals labeled by two or more cardiologists. Each ECG recording includes two channels (MLII and V1-V5). The QRS is usually more prominent in signals from the MLII lead. Therefore, data from the MLII channel were used in our study. The MIT-BIH database is unbalanced due to the unequal number of ECG beats for each arrhythmia [10]. MIT-BIH includes the classes N (normal), S (supraventricular ectopic), V (ventricular ectopic), F (Normal and V), and Q (undefined) [41]. Heartbeat types according to AAMI standards and tags in the MIT-BIH database are given in Table 1.

Table 1. *Heartbeat types according to AAMI and tags in the MIT-BIH.*

AAMI	MIT-BIH class[1]	Beat Count
N	N, L, R, e, j	90083
S	A, a, J, x	2972
V	V, !, E	7480
F	F	802
Q	Q	15

By AAMI recommendations, four records in this database (102, 104, 107, and 217) were not used in the study [14]. The heartbeat types and the selected recordings to be used in this study are given in Table 2. In the selection of the recordings, those commonly used in the literature were preferred.

Table 2. Types and numbers of heartbeats used.

Beat Types	Beat Count	Record Numbers
N	4868	103, 122, 220
LBBB	3856	109, 111, 214
RBBB	3613	118, 212, 231
PVC	2278	116, 119, 208, 213, 215, 221

The number of classes in the MIT-BIH database equals the number of tags. However, the database has an unbalanced distribution when tag-based beats are analyzed. Since data augmentation will not be used, the classes to be used were determined according to the balanced distribution. When selecting records, we tried to include records with a single beat type in the same class. In cases where this was not possible, other records were included. In this context, classification was made over four beat types.

B. DATA PRE-PROCESSING

Pre-processing includes normalization, filtering, and beat detection of the ECG signal before feature extraction and classification. Amplitude variations of ECG signals negatively affect the features. This variation creates significant variation in different patients for the same type of heartbeat [42]. With the normalization process, the amplitudes of the ECG signals are fixed to 1mV from peak to peak, and the offset of the signal is eliminated. Thus, the dependency of the features extracted from the ECG signal on demographic characteristics will be eliminated. The ECG signal will contain noise due to incorrect electrodes, patient movement, respiration, and various noises. Therefore, the DWT denoising method is used to remove the noise. Figure 3 shows the original, normalized, and denoised images of a sample ECG signal from the database.

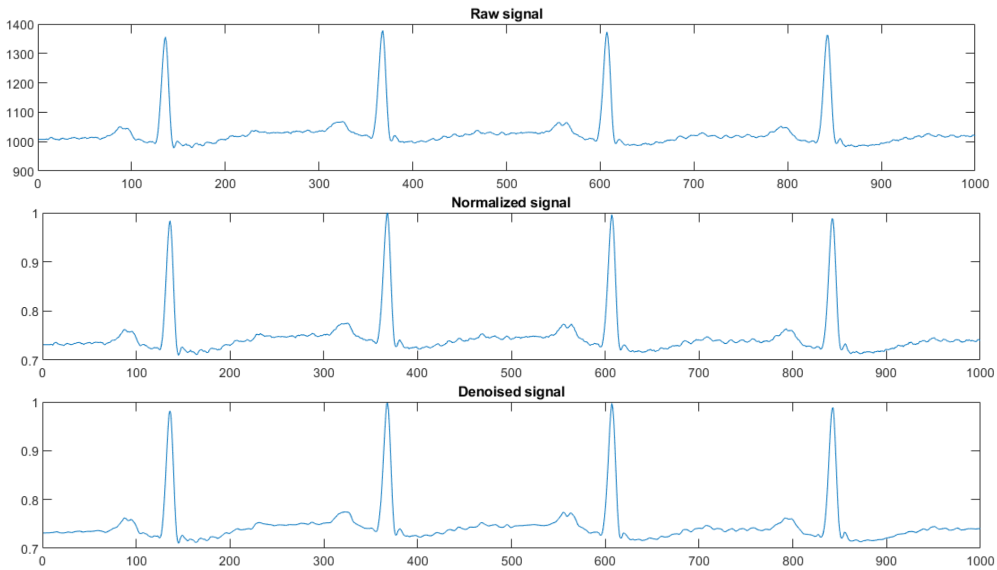

Figure 3. *Original, normalized, and noise-removed images, for example, ECG signal from the MIT-BIH database.*

The last stage of the preprocessing step is extracting heartbeat regions from the ECG signals. The R and RR interval values in the ECG signal are used for this process. Figure 4 shows an image of the extraction of heartbeat regions.

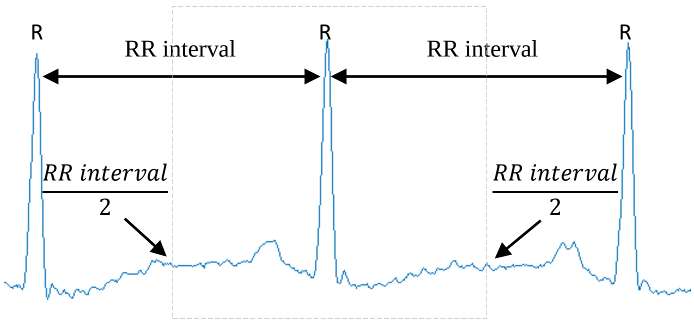

Figure 4. *Identification of heartbeat regions.*

RR intervals were calculated using the previous R-value and the next R-values for the heartbeat region to be determined. Then, the start and end points for the heartbeat to be received were determined by assessing the midpoint of these values.

C. VISUALIZATION OF HEARTBEATS

The main objective of this study is to classify ECG signals over images and to compare the performance of the methods used. For this purpose, SPEC, MEL, GFCC, CQT, and SCL images were obtained for each heartbeat obtained after the preprocessing step.

SPEC allows a time-varying signal to be moved into the frequency domain by applying a Fourier transform. Thus, the signal is transferred to the frequency domain, and the signal's energy is represented by colors in the spectrogram. To obtain MEL images, the frequencies in the SPEC are converted to the mel scale. If gamma tone filters are used instead of the mel scale, GFCC images are obtained. When the wavelet transform is used instead of the Fourier transform in the transition from the time domain to the frequency domain, the images to be obtained are also expressed as SCL. CQT transforms by creating a logarithmic gap between STFT and frequency transitions [43], [44] —sample images to be obtained after these transformations are given in Figure 5.

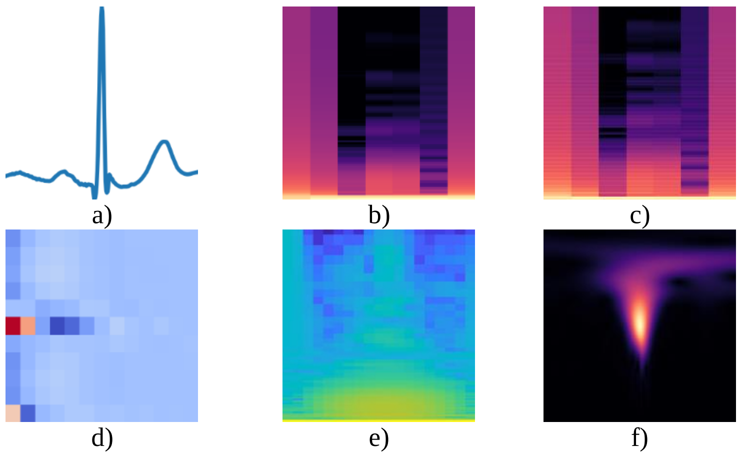

Figure 5. *Visualization for an example heartbeat from the MIT-BIH database;* ***a)*** *Raw signal,* ***b)*** *SPEC,* ***c)*** *MEL,* ***d)*** *GFCC,* ***e)*** *CQT and* ***f)*** *SCL*

D. HAND-CRAFTED FEATURES

Another experiment performed within the scope of the study is the hand-crafted features of the heartbeat signal. The features, linear spectrum, mel-spectrum, bark spectrum, MFCC, GTCC, spectral centroid, spectral entropy, pitch, and ZCR features were obtained from each heartbeat. These features were used as the input of the 1D-CNN model. In addition, the numerical values of the heartbeat signal were also used as input to the 1D-CNN model.

E. CONVOLUTIONAL NEURAL NETWORK (CNN)

Deep learning is an approach to machine learning in which the feature extraction process is performed in a network. The attributes to be used are realized in the network's learning process. Therefore, it provides higher success compared to classical methods [45]. Despite this high performance, deep learning models require more hardware and data. Thus, hardware with high processing power, such as GPUs, is needed.

CNN (Convolutional Neural Networks), often used in deep learning architecture, is a multilayer perceptron type. While CNN obtains more general features, such as edge information in the first layers, it obtains features representing the image in the advanced layers. CNN algorithms are used in many fields, such as image and audio processing, natural language processing, and biomedicine. The basic building blocks of any CNN consist of 4 main layers: convolution operator, ReLU, subsampling, and fully connected layer. These layers can be used more than once when building a CNN model. CNN models are modeled in 1D, 2D, and 3D. 1D-CNN is used for time series. 2D-CNN architecture is used in models using images as input. 3D-CNN, on the other hand, creates CNN models using 3D data as input.

In this study, a 1D-CNN model was created for the numerical values of the heart-beat signal and the hand-crafted features. The model is given in Figure 6.

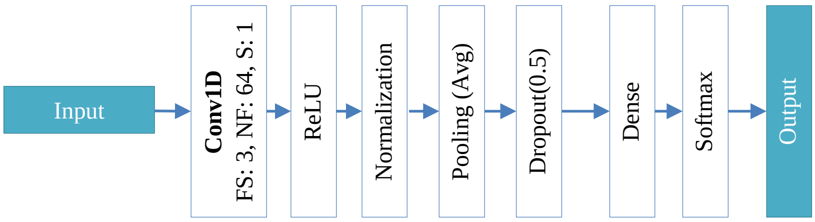

Figure 6. *Structure of the 1D-CNN model (FS: FilterSize, NF: NumFilters, S:Stride)*

For Exp7, the input to the model is the numerical data of the signal. For Exp8, the input is the hand-crafted features extracted from each signal.

A 2D-CNN model was created for the image of the heartbeat signal, and the SPEC, MEL, GFCC, CQT, and SCL images were obtained from this signal. The model is given in Figure 7.

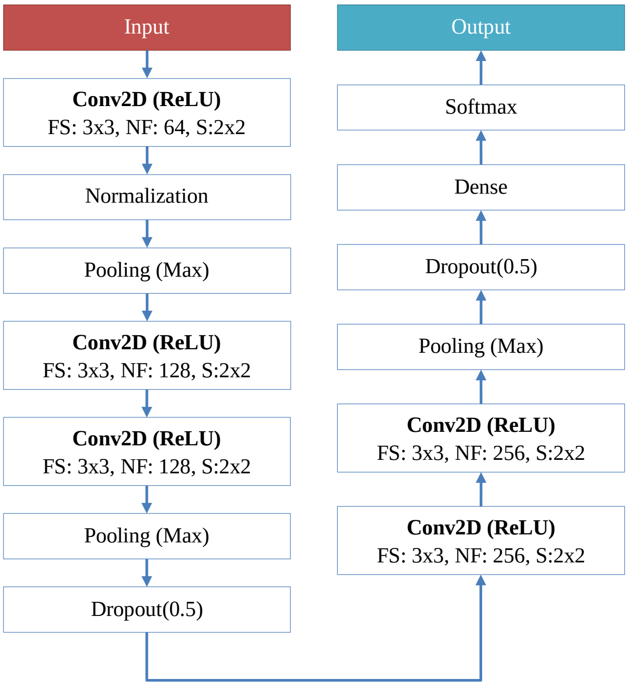

Figure 7. *Structure of the 2D-CNN model (FS: FilterSize, NF: NumFilters, S:Stride)*

The convolutional layer forms the basis of the CNN architecture and extracts features from the input. Since this process is performed by applying a filter to the data, the filter size, number, and stride value must be determined. Stride determines how many steps it takes to shift the filter on the input. The output of the convolutional layer is linear. However, the model also needs to learn non-linear problems. Therefore, activation functions are used, including various nonlinear activation functions such as sigmoid, tahn, and ReLU. However, ReLU is widely used because it gives faster results. Normalization is the normalization process between the layers of the CNN model and accelerates the training process. The pooling layer is used to reduce the number of features and computations. The dropout layer prevents overlearning by removing some neurons from the model during training. The dense layer is classified according to the information from convolution layers. The parameters used in the training phase of the CNN models are given in Table 3.

Table 3. *Parameters used for CNN models.*

Data selection for Training, Validation, and Testing	Random permutation
The portion of the data allocated for training	70%
The portion of the data allocated for validation	15%
The portion of the data allocated for the test	15%
Optimizer	Adam
Learning Rate	0.001
Epochs	30
Mini Batch Size	512

F. PERFORMANCE EVALUATION

There are various methods used to evaluate the success of classification problems. The confusion matrix gives the correctly and incorrectly classified examples for each class. With the help of the confusion matrix, measures such as precision, recall, f-score, and accuracy are used for performance evaluation. The recall is the ratio of the number of correctly classified positive samples to the sum of the number of correctly classified positive samples and the number of misclassified negative samples. Precision is the ratio of the number of correctly classified positive samples to the total number of positive samples. The F-score is the harmonic mean of the sensitivity and precision values. Accuracy is the ratio of the number of correctly classified samples to the total number of samples [46].

In the experiments, accuracy and loss graphs were obtained for training and validation during the training process of the models and given in the results section. A testing process was created and evaluated the success of the models with accuracy, precision, recall, and f-score criteria to evaluate the success of the models. For this purpose, 30% of the data was used for training, 15% for validation, and 15% for testing. The equations for the performance metrics used are given in Equations 1, 2, 3 and 4 [47].

$$Precision = \frac{TP}{TP+FP} \tag{1}$$

$$Recall = \frac{TP}{TP+FN} \tag{2}$$

$$f - score = \frac{2 \times Precision \times Recall}{Precision \times Recall} \tag{3}$$

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \tag{4}$$

II. EXPERIMENTAL RESULTS

In this study, six different images were obtained from heartbeat sound signals to classify ECG signals. ECG classification was performed with 2D-CNN using these six different images. In addition, to compare the performance of image and audio signal data, classification was made with 1D-CNN using the numerical values of the signal and the features obtained by signal processing (hand-crafted). Thus, eight experiments were conducted to classify ECG signals, and the results were compared. The test data classification metrics of the models trained with 1D-CNN and 2D-CNN are given in Table 4.

Table 4. *Performance values obtained with the test dataset.*

Exp	Model	Input	Precision	Recall	Accuracy	F-score
Exp1	2D-CNN	SPEC	0.9969	0.9971	0.9975	0.9970
Exp2	2D-CNN	MEL	0.9975	0.9970	0.9977	0.9972
Exp3	2D-CNN	GFCC	0.9935	0.9933	0.9934	0.9934
Exp4	2D-CNN	CQT	0.9946	0.9962	0.9963	0.9954
Exp5	2D-CNN	SCL	0.9958	0.9958	0.9963	0.9958
Exp6	2D-CNN	Raw Signal	0.9992	0.9985	0.9991	0.9989
Exp7	1D-CNN	Raw Signal	0.3461	0.4598	0.4753	0.3830
Exp8	1D-CNN	Hand-crafted features	0.8754	0.8330	0.8524	0.8377

When the test results in Table 4 are examined, the numerical signal data for ECG classification shows very low success. In addition, although the classification accuracy obtained with 1D-CNN in Exp8 is 85.24%, the accuracy rates obtained with 2D-CNN models are higher. ECG classification successes with the 2D-CNN model are higher than the results of the studies in the literature. If the results obtained with the image datasets are examined, a success of over 99% has been achieved in all datasets. Among these, the highest success was achieved with the visualization of the ECG sound signal, as 99.91%. Details of each experiment are given below as sub-items. All experiments were performed on MacBook Pro (i7 processor, 16GB RAM, and 512SSD).

A. CLASSIFICATION RESULTS WITH TIME-FREQUENCY DOMAIN IMAGES

In this section, the results obtained in Exp1 (SPEC), Exp2 (MEL), Exp3 (GFCC), Exp4 (CQT), Exp5 (SCL), and Exp6 are presented. The training, validation, and test data used in all six experiments were chosen to be the same for a more objective evaluation of the comparison results. The training and validation results obtained in the six experiments are given in Figure 8.

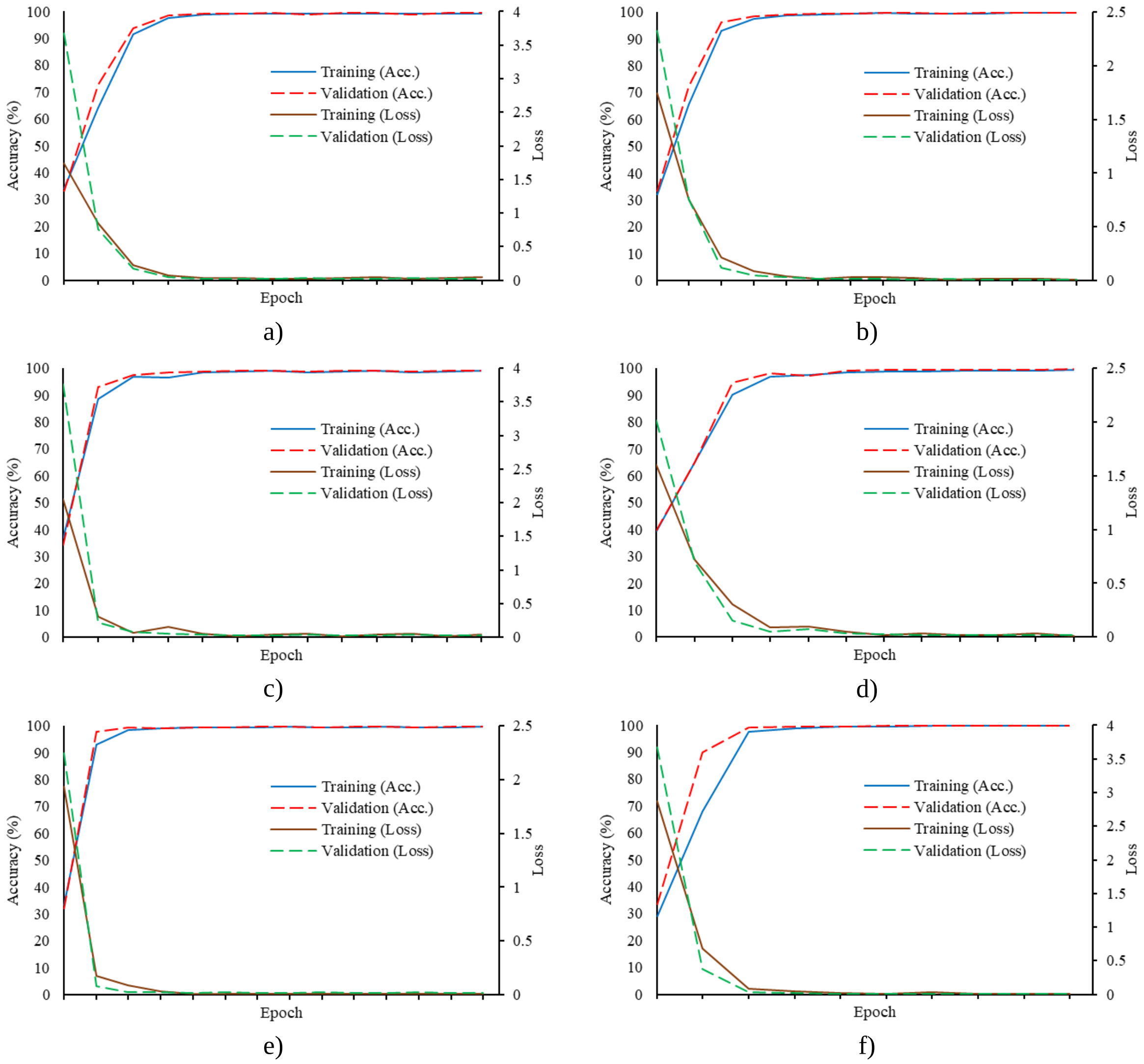

Figure 8. *Training and validation results for 2D-CNN model **a)** Exp1, **b)** Exp2, **c)** Exp3, **d)** Exp4, **e)** Exp5, **f)** Exp6*

As can be seen in the accuracy and loss graphs obtained on the 2D-CNN model and image datasets, the accuracy value is high, and the training value is low in all image datasets. Also, when the training and validation curves are analyzed, it is seen that there is no overfitting problem. The class-based classification achievements were examined using the test data (15% of the entire data set) for these six training experiments. The results obtained are given in Table 5.

Table 5. Class-based classification achievements were obtained with the test dataset for the first six experiments.

Experiments	Accuracy (%)				
	LBBB	RBBB	Normal	PVC	Overall
Exp1 (SPEC)	99.74	99.82	99.93	99.27	99.75
Exp2 (MEL)	99.66	100.00	99.79	99.56	99.77
Exp3 (GFCC)	99.48	99.54	99.10	99.27	99.34
Exp4 (CQT)	99.83	100.00	99.93	98.10	99.63
Exp5 (SCL)	99.22	99.91	100.00	99.19	99.63
Exp6 (Raw Signal)	99.74	100.00	99.93	100.0	99.91

The class-based accuracy rates in Table 5 are obtained from the confusion matrix obtained by classifying the test dataset. In the six experiments, the class-based success rates are evenly distributed. Another finding from the confusion matrix is that the LBBB and PVC classes are similar in 2D-CNN experiments. As in the training process, the highest success rate was obtained with the raw signal image (Exp6) in the testing process. These findings show that all six image types can be used successfully in heartbeat classification.

B. CLASSIFICATION RESULTS WITH 1D-CNN

The 1D-CNN model was used to classify the signal with raw digital data (Exp7) and hand-crafted features (Exp8). The training and validation graphs of the 1D-CNN model for Exp7 and Exp8 are given in Figure 9.

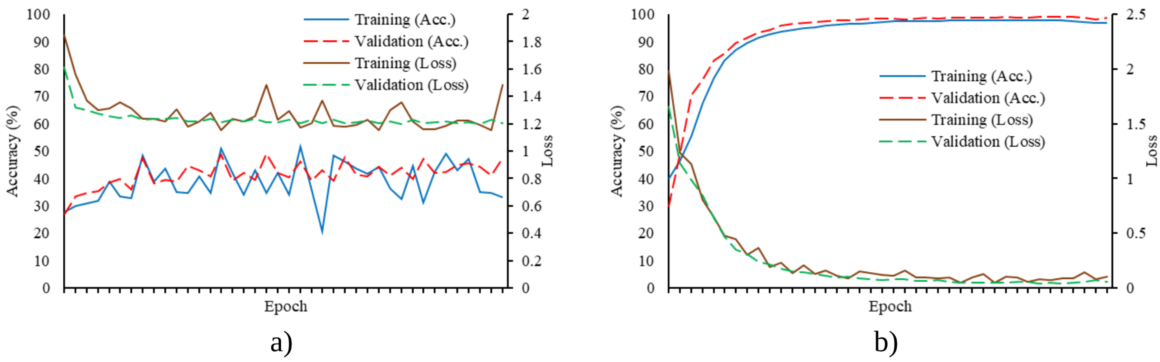

***Figure 9.** Training and validation results for 1D-CNN **a)** Exp7, **b)** Exp8*

The training and validation results of the experiment (Exp7) using the 1D-CNN model and the raw numerical data of the signal show that this dataset is unsuitable for heartbeat classification. Because both the loss value is high, and the training value is low. Therefore, the model fails. This situation is also seen more clearly in the class-based achievements obtained because of the classification of the test dataset for Exp7. Accuracy was 0% for LBBB, 25.46% for RBBB, 92.51% for Normal, and 65.95% for PVC. This model can be used for normal-abnormal heartbeat classification and is unsuitable for multi-class problems.

Exp8 used the 1D-CNN model and hand-crafted features extracted from the heartbeat signal. This experiment's accuracy and loss curves show that the model can be used successfully in heartbeat classification. After a particular iteration, the accuracy value increased and went horizontal, while the loss value decreased and went horizontal. When the model is tested using the dataset allocated for testing, it cannot be said that the class-based accuracy rates are fully balanced. Accuracy was 94.83% for LBBB, 96.19% for RBBB, 82.31% for Normal, and 59.87% for PVC. The overall accuracy of the model is 85.24%. The unbalanced class-based accuracy rates are due to the PVC class, which has a lower accuracy rate than the other classes. Also, the PVC class is most like the LBBB class.

IV. DISCUSSION AND CONCLUSION

This study uses raw signal data and time-frequency domain images for ECG signal classification. With these data, 1D-CNN and 2D-CNN models were created, and a comparative performance analysis was performed. For this purpose, ECG recordings from the MIT-BIH database were widely used in the literature. First, these recordings detected heartbeat regions, and digital and image datasets were created for these regions. The digital dataset contains raw digital signal data and hand-crafted features extracted from the signal. The heartbeat image, spectrogram, scalogram, CQT, GFCC, and mel-spectrogram images of the signal were used for the image dataset. In this context, eight experiments were performed, and each experiment's training, validation, testing, and loss results were analyzed.

The first six experiments with the image dataset achieved a classification accuracy of over 99%. Among these, the heartbeat signal's highest success rate was 99.91%. The class-based accuracy rates were evenly distributed in these six experiments, and no overfitting problem occurred. The other two experiments use the numerical values of the heartbeat signal. In the seventh experiment, numerical values for the heartbeat obtained from the MIT-BIH database were used, and classification was performed with 1D-CNN. The results obtained are unsuccessful according to the accuracy and loss values of the model. In addition, the class-based accuracy is high, especially for the normal class, but low for the other classes because of the classification performed on the test data. Especially in the LBBB class, the success rate is 0%. In the last experiment, feature extraction was performed on the numerical values obtained for a heartbeat, and training was performed with 1D-CNN. As a result of the training, both accuracy and loss curves show that the model can be used for the heartbeat. However, when the class-based accuracies are analyzed, the accuracy rate for the PVC class is low. The overall success of the model is 85.24%.

When the results of the experiments performed within the scope of the study are compared with the studies in the literature, the success rates of the experiments using 2D-CNN and images are higher than in the literature. The comparison of our results with the recent studies using the MIT-BIH database is given in Table 6.

Table 6. *Comparison of the results obtained with the literature.*

Method/Feature	Model/Classifier	Accuracy (%)
Our Exp. (Exp6-ECG signal)	**2D-CNN**	**99.9**
Our Exp. (Exp1-SPEC)	**2D-CNN**	**99.8**
Our Exp. (Exp2-MEL)	**2D-CNN**	**99.8**
TERMA and FrFT	SVM	99.8 [12]
Grayscale ECG Signal	2D-CNN	99.7 [10]
Our Exp. (Exp4-CQT)	**2D-CNN**	**99.6**
Our Exp. (Exp5-SCL)	**2D-CNN**	**99.6**
ECG Signal	CNN and GRU	99.6 [11]
ECG Signal, R-peaks, RR interval	CNN	99.6 [16]
SWT feature and RR interval	1D-CNN	99.4 [15]
ECG Signal	CNN	99.4 [35]
Our Exp. (Exp3-GFCC)	**2D-CNN**	**99.3**
ECG Signal	CNN+ELM	98.8 [13]
ECG Signal	2D-CNN	98.7 [48]
ECG Signal	1D-CNN	98.5 [20]
RR-intervals, higher-order-statistic features, DWT	Random Forest	95.7 [14]
Third Order Cumulant	SqueezeNet	94.6 [49]
Spectrogram	ResNet-18	91.0 [18]
Our Exp. (Exp8-handcrafted feat.)	**1D-CNN**	**85.2**
Our Exp. (Exp7-raw signal)	**1D-CNN**	**47.5**

When the results given in Table 6 are examined, the accuracy rate obtained with the image of the ECG signal is higher than the studies in the literature. The accuracy rate obtained with spectrogram and mel-spectrogram is the highest accuracy in the literature. The accuracy rates obtained with CQT, scalogram, and GFCC are between 0.2%-0.5% lower than the highest accuracy rates in the literature. Two experiments with 1D-CNN achieved lower accuracy than the studies in the literature.

The results of this study, which examined various inputs for deep learning, which has been widely used in the classification of ECG signals in recent years, showed that the numerical values of the signal for the heartbeat and the features to be obtained from these values showed low classification accuracy and could not entirely separate the similarity between the classes. The raw signal data and the 1D-CNN model showed low success for ECG classification. When 2D-CNN and six different image datasets were used for ECG classification, all images were highly successful for ECG classification. Among these, the highest success was achieved with the image of the heartbeat signal. An important issue encountered in the experiments is the high similarity rate in the classification of PVC and LBBB classes in image datasets compared to other classes.

The strengths of this study are the comparison of the 1D-CNN model with the 2D-CNN model and the performance analysis of ECG classification with different input data. In future studies, the imbalance in the datasets can be eliminated with data augmentation, and performance comparisons can be made. In addition, the success of texture analysis methods during the conversion of heartbeat signals into images can be examined.

CONFLICT OF INTEREST

The authors declare no conflict of interest.

V. REFERENCES

[1] O. M. A. Ali, S. W. Kareem, and A. S. Mohammed, “Evaluation of Electrocardiogram Signals Classification Using CNN, SVM, and LSTM Algorithm: A review,” presented at 8th International Engineering Conference on Sustainable Technology and Development (IEC), pp. 185–191, 2022.

[2] “World health statistics 2018: monitoring health for the SDGs, sustainable development goals - RELACSIS | OPS/OMS,” *Pan American Health Organization / World Health Organization*, 2018. https://www3.paho.org/relacsis/index.php/es/noticias-relacsis/906-report-world-health-statistics-2018-monitoring-health-for-the-sdgs-sustainable-development-goals (accessed Jan. 12, 2023).

[3] E. J. Benjamin *et al.*, “Heart disease and stroke statistics—2019 update: a report from the American Heart Association,” *Circulation*, vol. 139, no. 10, pp. e56–e528, 2019.

[4] X. Liu, H. Wang, Z. Li, and L. Qin, “Deep learning in ECG diagnosis: A review,” *Knowledge-Based Systems*, vol. 227, p. 107187, 2021.

[5] U. R. Acharya, S. L. Oh, Y. Hagiwara, J. H. Tan, M. Adam, A. Gertych, and R. San Tan, “A deep convolutional neural network model to classify heartbeats,” *Computers in Biology and Medicine*, vol. 89, pp. 389–396, 2017.

[6] G. De Lannoy, D. François, J. Delbeke, and M. Verleysen, “Weighted conditional random fields for supervised interpatient heartbeat classification,” *IEEE Transactions on Biomedical Engineering*, vol. 59, no. 1, pp. 241–247, 2011.

[7] H. V. Pipberger, R. J. Arms, and F. W. Stallmann, “Automatic screening of normal and abnormal electrocardiograms by means of a digital electronic computer,” *Proceedings of the Society for Experimental Biology and Medicine*, vol. 106, no. 1, pp. 130–132, 1961.

[8] H. Malik, U. Bashir, and A. Ahmad, “Multi-classification neural network model for detection of abnormal heartbeat audio signals,” *Biomedical Engineering Advances*, vol. 4, p. 100048, 2022.

[9] M. Cao, T. Zhao, Y. Li, W. Zhang, P. Benharash, and R. Ramezani, “ECG heartbeat classification using deep transfer learning with convolutional neural network and STFT technique,” *arXiv preprint arXiv:2206.14200*, 2022.

[10] M. Degirmenci, M. A. Ozdemir, E. Izci, and A. Akan, “Arrhythmic heartbeat classification using 2d convolutional neural networks,” *Irbm*, vol. 43, no. 5, pp. 422–433, 2022.

[11] G. Yao, X. Mao, N. Li, H. Xu, X. Xu, Y. Jiao, and J. Ni, “Interpretation of electrocardiogram heartbeat by CNN and GRU,” *Computational and Mathematical Methods in Medicine*, vol. 2021, 2021.

[12] S. Aziz, S. Ahmed, and M.-S. Alouini, “ECG-based machine-learning algorithms for heartbeat classification,” *Scientific reports*, vol. 11, no. 1, pp. 1–14, 2021.

[13] S. Zhou, and B. Tan, “Electrocardiogram soft computing using hybrid deep learning CNN-ELM,” *Applied soft computing*, vol. 86, p. 105778, 2020.

[14] T. Wang, C. Lu, W. Ju, and C. Liu, “Imbalanced heartbeat classification using EasyEnsemble technique and global heartbeat information,” *Biomedical Signal Processing and Control*, vol. 71, p. 103105, 2022.

[15] D. Zhang, H. Zhou, F. Li, L. Zhang, and J. Wang, “A reparameterization multifeature fusion CNN for arrhythmia heartbeats classification,” *Computational and Mathematical Methods in Medicine*, vol. 2022, 2022.

[16] A. Tyagi and R. Mehra, “Intellectual heartbeats classification model for diagnosis of heart disease from ECG signal using hybrid convolutional neural network with GOA,” *SN Applied Sciences*, vol. 3, no. 2, pp. 1–14, 2021.

[17] Y. Xu, S. Zhang, Z. Cao, Q. Chen, and W. Xiao, “Extreme learning machine for heartbeat classification with hybrid time-domain and wavelet time-frequency features,” *Journal of Healthcare Engineering*, vol. 2021, 2021.

[18] E. Jing, H. Zhang, Z. Li, Y. Liu, Z. Ji, and I. Ganchev, “ECG heartbeat classification based on an improved ResNet-18 model,” *Computational and Mathematical Methods in Medicine*, vol. 2021, 2021.

[19] E. Essa and X. Xie, "An ensemble of deep learning-based multi-model for ECG heartbeats arrhythmia classification," *IEEE access*, vol. 9, pp. 103452–103464, 2021.

[20] E. Maghawry, T. F. Gharib, R. Ismail, and M. J. Zaki, "An efficient heartbeats classifier based on optimizing convolutional neural network model," *IEEE access*, vol. 9, pp. 153266–153275, 2021.

[21] S. K. Pandey, R. R. Janghel, and V. Vani, "Patient specific machine learning models for ECG signal classification," *Procedia Computer Science*, vol. 167, pp. 2181–2190, 2020.

[22] S. L. Oh, E. Y. Ng, R. San Tan, and U. R. Acharya, "Automated diagnosis of arrhythmia using combination of CNN and LSTM techniques with variable length heart beats," *Computers in biology and medicine*, vol. 102, pp. 278–287, 2018.

[23] R. J. Martis, U. R. Acharya, and L. C. Min, "ECG beat classification using PCA, LDA, ICA and discrete wavelet transform," *Biomedical Signal Processing and Control*, vol. 8, no. 5, pp. 437–448, 2013.

[24] Y. H. Hu, W. J. Tompkins, J. L. Urrusti, and V. X. Afonso, "Applications of artificial neural networks for ECG signal detection and classification," *Journal of electrocardiology*, vol. 26, pp. 66–73, 1993.

[25] N. Izeboudjen and A. Farah, "A new neural network system for arrhythmia's classification," *NC*, vol. 98, pp. 23–25, 1998.

[26] J.-S. Wang, W.-C. Chiang, Y.-L. Hsu, and Y.-T. C. Yang, "ECG arrhythmia classification using a probabilistic neural network with a feature reduction method," *Neurocomputing*, vol. 116, pp. 38–45, 2013.

[27] C. V. Banupriya and S. Karpagavalli, "Electrocardiogram beat classification using probabilistic neural network," *Int. J. Comput. Appl.(IJCA)*, vol. 1, no. 7, pp. 31–37, 2014.

[28] N. Acır, "Classification of ECG beats by using a fast least square support vector machines with a dynamic programming feature selection algorithm," *Neural computing & applications*, vol. 14, no. 4, pp. 299–309, 2005.

[29] M. Moavenian and H. Khorrami, "A qualitative comparison of artificial neural networks and support vector machines in ECG arrhythmias classification," *Expert Systems with Applications*, vol. 37, no. 4, pp. 3088–3093, 2010.

[30] M. H. Song, J. Lee, S. P. Cho, K. J. Lee, and S. K. Yoo, "Support vector machine based arrhythmia classification using reduced features," *Artificial Intelligence in Medicine*, vol.44, no:1, pp. 51-64, 2005.

[31] A. B. A. Qayyum, T. Islam, and M. A. Haque, “ECG heartbeat classification: A comparative performance analysis between one and two dimensional convolutional neural network,” in *2019 IEEE International Conference on Biomedical Engineering, Computer and Information Technology for Health (BECITHCON)*, 2019, pp. 93–96.

[32] C. Chen, Z. Hua, R. Zhang, G. Liu, and W. Wen, “Automated arrhythmia classification based on a combination network of CNN and LSTM,” *Biomedical Signal Processing and Control*, vol. 57, p. 101819, 2020.

[33] H. Shi, C. Qin, D. Xiao, L. Zhao, and C. Liu, “Automated heartbeat classification based on deep neural network with multiple input layers,” *Knowledge-Based Systems*, vol. 188, p. 105036, 2020.

[34] Q. Xie, S. Tu, G. Wang, Y. Lian, and L. Xu, “Feature enrichment based convolutional neural network for heartbeat classification from electrocardiogram,” *IEEE Access*, vol. 7, pp. 153751–153760, 2019.

[35] X. Xu, and H. Liu, “ECG heartbeat classification using convolutional neural networks’, *IEEE Access*, vol. 8, pp. 8614–8619, 2020.

[36] T. F. Romdhane, H. Alhichri, R. Ouni, and M. Atri, “Electrocardiogram heartbeat classification based on a deep convolutional neural network and focal loss,” *Computers in Biology and Medicine*, vol. 123, p. 103866, 2020.

[37] K. Yadav, S. Tiwari, A. Jain, and A. K. Y. Dafhalla, “Deep learning based cardiovascular disease diagnosis system from heartbeat sound,” *International Journal of Speech Technology*, pp. 1-12, 2021.

[38] W. Ullah, I. Siddique, R. M. Zulqarnain, M. M. Alam, I. Ahmad, and U. A. Raza, “Classification of arrhythmia in heartbeat detection using deep learning,” *Computational Intelligence and Neuroscience*, pp. 1–13, 2021.

[39] Y. Liang, S. Yin, Q. Tang, Z. Zheng, M. Elgendi, and Z. Chen, “Deep learning algorithm classifies heartbeat events based on electrocardiogram signals,” *Frontiers in Physiology*, vol. 11, 2020.

[40] G. B. Moody, and R. G. Mark, “The impact of the MIT-BIH arrhythmia database,” *IEEE Engineering in Medicine and Biology Magazine*, vol. 20, no. 3, pp. 45–50, 2001.

[41] Ö. Yakut, S. Solak, and E. Bolat, “IIR based digital filter design for denoising the ECG signal,” *Journal Of Polytechnic*, vol. 21, no. 1, 2018.

[42] Z. Dokur, “Yapay sinir ağları ve genetik algoritmalar kullanılarak EKG vurularının sınıflandırılması,” Fen Bilimleri Enstitüsü, İstanbul Teknik Üniversitesi, İstanbul, Türkiye, 2023.

[43] J. K. Das, A. Ghosh, A. K. Pal, S. Dutta, and A. Chakrabarty, “Urban sound classification using convolutional neural network and long short term memory based on multiple features,” in *2020 Fourth International Conference On Intelligent Computing in Data Sciences (ICDS)*, 2020.

[44] Z. Huang, C. Liu, H. Fei, W. Li, J. Yu, and Y. Cao, “Urban sound classification based on 2-order dense convolutional network using dual features,” *Applied Acoustics*, vol. 164, p. 107243, 2020.

[45] M. A. Kızrak ve B. Bolat, “Derin öğrenme ile kalabalık analizi üzerine detaylı bir araştırma,” *Bilişim Teknolojileri Dergisi*, vol. 11, no. 3, pp. 263–286, 2018.

[46] A.-M. Šimundić, “Measures of Diagnostic Accuracy: Basic Definitions,” *EJIFCC*, vol. 19, no. 4, pp. 203–211, 2009.

[47] P. Eusebi, “Diagnostic Accuracy Measures,” *Cerebrovascular Diseases*, vol. 36, no. 4, pp. 267–272, 2013.

[48] H. Xie, H. Liu, S. Zhou, T. Gao, and M. Shu, “A lightweight 2-D CNN model with dual attention mechanism for heartbeat classification,” *Applied Intelligence*, 2022.

[49] A. M. Alqudah, S. Qazan, L. Al-Ebbini, H. Alquran, and I. A. Qasmieh, “ECG heartbeat arrhythmias classification: a comparison study between different types of spectrum representation and convolutional neural networks architectures,” *J Ambient Intell Human Comput*, vol. 13, no. 10, pp. 4877–4907, 2022.