A Novel Proposed CNN-SVM Architecture for ECG Scalograms Classification

Oznur Ozaltin (oznurozaltin@hacettepe.edu.tr)
Hacettepe Universitesi Beytepe Kampusu https://orcid.org/0000-0001-9841-1702

Yeniay Ozgur
Hacettepe Üniversitesi - Beytepe Kampüsü: Hacettepe Universitesi - Beytepe Kampusu

Research Article

Keywords: Convolutional Neural Networks (CNN), Continuous Wavelet Transform (CWT), Feature extraction, Scalograms, Support Vector Machine (SVM)

Posted Date: July 22nd, 2022

DOI: https://doi.org/10.21203/rs.3.rs-1364437/v1

A Novel Proposed CNN-SVM Architecture for ECG Scalograms Classification

Ozaltin Oznur, Yeniay Ozgur

(*Hacettepe University, Institute of Science, Department of Statistics*)

Abstract

Nowadays, the number of sudden deaths due to heart disease is increasing with the Coronavirus pandemic. Thus, Electrocardiogram (ECG) signals automatic classification is of vital importance for diagnosis and treatment. Thanks to deep learning algorithms, the classification can be performed without manual feature extraction. In this study, we propose a novel convolutional neural networks (CNN) architecture. Further, the proposed CNN can be extracted automatic features from images. Here, we classify a real ECC data set using our proposed CNN that includes 34 layers. While this dataset is one-dimensional signals, these are transformed into two-dimensional scalograms by continuous wavelet transform (CWT). In addition, we compare it with well-known architectures: AlexNet and SqueezeNet. When we classify ECG scalograms with proposed CNN, we find it more effectively than others. Although the results are very good, we benefit from support vector machines (SVM). Essentially, our main aim is to achieve the best classification results on account of health. Thus, we modify the proposed CNN with SVM. As a result, we achieve the highest success with an accuracy of 99.21% from our proposed CNN-SVM.

Keywords: Convolutional Neural Networks (CNN), Continuous Wavelet Transform (CWT), Feature extraction, Scalograms, Support Vector Machine (SVM).

1. Introduction

Qualitative processing and classification of biomedical signals are very important for diagnosis and treatment. Many methods are used to process biomedical signals.

Some important methods are Discrete Fourier Transform (DFT), Short-Term Fourier Transform (STFT), Continuous Wavelet Transform (CWT), and Discrete Wavelet Transform. For stationary signals, the Fourier transform provides a very good frequency domain. However, the time domain is almost non-existent. Especially, when it is wanted to infer time-dependent features, it can lead to serious problems. However, when signals are transformed with the wavelet transform, both frequency and time domains are distinguishable. In other words, the wavelet transform (WT) is a transformation method that divides signals into different frequency components and works each component with the time domain of the respective scale.

In this study, we focus on electrocardiogram (ECG) signals. The signals resulting from the electrical activity of the heart, the major vital organ of the human body, are called an electrocardiogram (ECG). Currently, sudden deaths due to heart disease increase with COVID -19. For this reason, the processing and analysis of signals received from the heart are very important for fast diagnosis and treatment.

In conventional methods, in the pre-processing phase of ECG signals, an appropriate sampling method is used and the signals are cleaned from the noise. Then the manual feature extraction phase is started where it is very important to seek expert opinions. This phase is very critical, since incorrect feature extraction may lead to misclassification of signals and result in serious errors in diagnosis and treatment. After all these phases are completed, classification is made by using traditional classification algorithms. However, the studies show that the situation is different for Deep Learning algorithms in recent years. Thanks to deep learning algorithms, successful classifications can be made automatically. Thus, the health status of patients can be monitored even without expert opinion using smartphones, watches, etc.

In this study, a novel Convolutional Neural Networks (CNN) architecture which is one of the deep learning algorithms, is proposed for automatically ECG signal classification. This novel CNN architecture is designed for two-dimensional images with having 34 layers which contain robust features. Actually, the new proposed CNN is not only considered as ECG signals classification but also considered as other biomedical signals, images, etc. classification. In this context, of course, the ECG signals are transformed from one-dimensional signals to two-dimensional scalograms (or images) by using continuous wavelet transform (CWT) in the pre-processing phase. This wavelet transform has three different mother wavelet functions: Amor, Bump, and Morse which are used mostly. These functions' effects on classification performance are examined as well. In general, the signal sample size is taken as 360 Hz by the researchers. In this study, as well as this generally used sampling size is also investigated with other sampling sizes: 500 Hz and 1000 Hz, whether the wave characteristics become more obvious. *Fig.1* shows the scalograms which are obtained with different sample sizes of ECG signals, 360 Hz, 500 Hz, and 1000 Hz respectively. In total, 9 different datasets are acquired in these conditions. In the classification phase, these datasets are classified separately with the same training options parameters using the proposed CNN architecture, AlexNet and SqueezeNet. Additionally, we work not only on CNN architectures but also support vector machines (SVM) algorithms, in this study.

CNN and SVM algorithms have high classification success separately, as known. Generally, the dropout technique, data augmentation, etc. are applied to overcome extreme learning for CNN architectures. However, it does not give a good result every time. These techniques may change depending on the data structure. In this study, the proposed CNN is tasked as an automatic feature extractor. Therefore, the proposed CNN is modified with the SVM algorithm and this structure can overcome extreme learning.

Nowadays, artificial intelligence is evolving day by day, many studies are also being carried out to classify ECG signals and other biomedical signals using CNN architectures. Khorrami and Moavenian[1] have applied the CWT, the discrete wavelet transforms (DWT), and the discrete cosine transform (DCT) to ECG signals. In addition, they have compared SVM with Multi-layer Perceptron (MLP) algorithms in the classification phase. Especially, they have found that combinations (CWT -MLP, DWT -MLP, DCT-MLP) created with MLP are superior to SVM. Rahhal et al. [2] have transformed signals from different datasets by using CWT to identify arrhythmias in ECG signals. Besides, they have used the CNN algorithm and obtained an accuracy of 99% in the classification phase. Huang et al. [3] have transformed ECG signals with STFT and obtained two-dimensional scalograms in their study. Moreover, they have benefited from the CNN architecture for classifying these scalograms and achieved an accuracy of 99%. In addition, they have also classified the one-dimensional ECG signals using CNN and found an accuracy of 90.93%. Krak et al. [4] have transformed ECG signals using CWT and DWT in their study. Furthermore, they have classified using the CNN architecture and obtained an accuracy of 96% in the classification phase. Baloglu et al. [5] have designed a 10-layer end-to-end CNN architecture for the classification of one-dimensional multiclass ECG data and achieved an accuracy of a 99.78%. Mahmud et al. [6] have created a CNN architecture for one-dimensional multiclass ECG data and obtained an accuracy rate of 99.28%. Salem et al. [7] have utilized DenseNet architecture to classify transformed two-dimensional ECG data and achieved an accuracy of 97.23%. Zhao et al. [8] have proposed a CNN which included 24 layers for classifying transformed ECG data and achieved an accuracy of 87.1%. Xu and Liu [9] have created a CNN architecture in order to analyze ECG data taken from a Holter device and achieved an accuracy of 99.4%. Rajkumar et al. [10] have suggested a CNN architecture for one-dimensional ECG data by using ELU activation layers and achieved an accuracy of 93.6%.

Hua et al. [11] have developed a CNN architecture for one-dimensional ECG signals and achieved an accuracy of 97.45 %. Kiranyaz et al. [12] have proposed a CNN architecture for real-time patient-specific one-dimensional ECG classification and achieved an accuracy of 96.4%. Chen et al. [13] have suggested CNN +Long-Short-Term Memory (LSTM) which can classify six kinds of ECG fragments. They have classified two ECG databases: MIT-BIH arrhythmia database and MIT-BIH arrhythmia database+ Challenge2017, and achieved an accuracy of 99.32% and 97.15%, respectively, by using CNN+LSTM. Sandeep et al. [14] have utilized the CNN architecture to classify ECG data and also achieved an accuracy of 90.63%. Furthermore, Machine Learning Algorithms such as Support Vectors Machine (SVM), K- Nearest Neighbors (KNN), Decision Tree (DT), Extreme Learning Machine (ELM), Ensemble Learning, Multi-layer Perceptron (MLP), etc. to classifying ECG signals were also used by many other researchers[15-19] .

Considering all the studies in the literature, a deeper CNN architecture is proposed which can be automatically done feature extraction without expert opinions. Besides, the proposed architecture is compared with two well-known pre-trained architectures: AlexNet and SqueezeNet on the created 9 different ECG datasets. Moreover, this study is highlighted the best sample size and the best mother wavelet function. As a result of the comparison, it is found that the proposed architecture is superior to the other architectures in terms of classification success and sensitivity. Essentially, the proposed CNN of classification success is well. However, the CNN architecture is modified with the SVM since it is aimed that minimum error to classification on ECG signals. Hereby, the classification success of the proposed CNN-SVM architecture is improved.

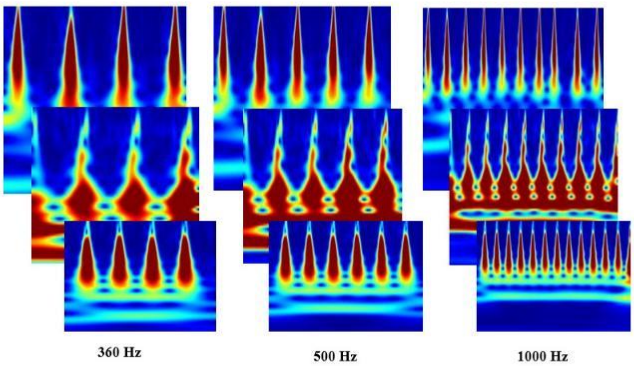

Fig. 1. 227x227x3 size scalograms of taken signal lengths with 360 Hz, 500 Hz, and 1000 Hz from data respectively.

2. Materials and Methods

In this section, firstly we present the details of ECG datasets. Then, we examine pre-processing method: CWT. Next, we introduce in general CNN, the proposed CNN, and pre-trained architectures: AlexNet[20] and SqueezeNet[21]. In the last, we present SVM and the proposed CNN-SVM architecture for the classification of ECG datasets. Fig.2 shows the framework for the classification of ECG datasets.

2.1. ECG Dataset

In this study, we benefit from three different ECG datasets off PhysioNet databases [22]. Each raw ECG dataset is taken as 1- hour signal length and sampled with 128 Hz. The first ECG dataset consists of 48 patients' ECG records which contain two leads. It is received from MIT -BIH Arrhythmia Database and named as "ARR". [23, 24]. The next ECG dataset consists of 15 patients' ECG records which contain two leads. It is taken from BIDMC Congestive Heart Failure Database and named "CHF"[23, 25]. The final ECG dataset consists of 18 patients' ECG records which include two leads. It is received from MIT -BIH Normal Sinus Rhythm and named as "NSR" [23]. Totally, 96 ARR, 30 CHF, and 36 NSR are in the ECG dataset.

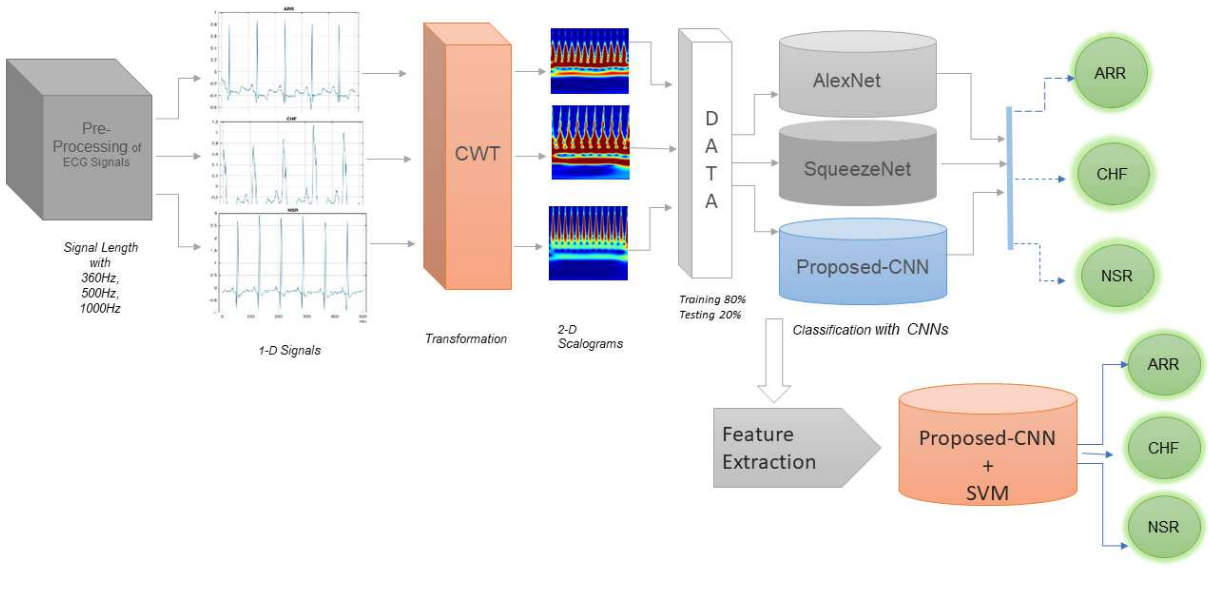

Fig. 2. Flowchart of ECG Signals Classification.

2.2 Continuous Wavelet Transform

The continuous wavelet transform (CWT) which is a transformation method that allows simple analysis of its frequency components, can transform a one-dimensional signal into a two-dimensional scalogram by providing a mapping of the signal also on the time axis. The mathematical formulation of the CWT and WT family is offered in Equation (1) and Equation (2), respectively:

$$CWT(a,b) = \left\langle f, \psi^*_{a,b} \right\rangle = \int_{-\infty}^{+\infty} f(t) \psi^*_{a,b}(t) dt \tag{1}$$

$$\psi_{a,b}(t) = \frac{1}{\sqrt{a}} \psi\left(\frac{t-b}{a}\right) \tag{2}$$

where $f(t)$ is a continuous signal function received in this study as an ECG signal function, $\psi_{a,b}(t)$ is the mother wavelet function, a indicates a scale parameter, and b indicates shift parameter or translation, the symbol of * indicates the complex conjugate function [28].

$\left\langle f, \psi_{a,b} \right\rangle$ is expressed as a function of the inner products of Equation (1). It $CWT(a,b)$ is regulated,

$$CWT(a,b) = \frac{1}{\sqrt{a}} \int_{-\infty}^{+\infty} f(t) \psi\left(\frac{t-b}{a}\right) dt \tag{3}$$

will be in the form like in Equation (3). The signal function $f(t)$ can be converted from the inverse of $CWT(a,b)$, as follows:

$$f(t) = \frac{1}{C} \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} CWT(a,b) \frac{\psi_{a,b}(t)}{|a|^{3/2}} da\, db \tag{4}$$

where C indicates the normalization constant depending on the choice of the mother wavelet function in Equation (4) [28].

Some mother wavelet functions as follows:

$$\psi_{Morl}(t) = e^{2\pi i t} e^{-\frac{t^2}{2\sigma^2}} = \left(\cos 2\pi t + i \sin 2\pi t\right) e^{-\frac{t^2}{2\sigma^2}} \tag{5}$$

$$\psi_{Mexh}(t) = \left(1 - \frac{t^2}{\sigma^2}\right) e^{-\frac{t^2}{2\sigma^2}} \tag{6}$$

$$\psi_{Bump}(ab) = e^{\left(1-\frac{1}{1-ab-\mu/\sigma^2}\right)} \chi[\mu-\sigma, \mu+\sigma] \tag{7}$$

will be in the form in Equation (5-7). Here $\psi_{Morl}(t)$, Morlet, $\psi_{Mexh}(t)$, Mexican hat and $\psi_{Bump}(ab)$, Bump shows the mother wavelet function [28].

2.3. Convolutional Neural Network (CNN)

Convolutional Neural Network (CNN) emerges as a specialized deep learning approach for analyzing two-dimensional data. Not only it is preferred algorithm in the analysis of multidimensional data but also one-dimensional data. Other classifications and clustering algorithms are difficult to apply to real-time data due to their computational complexity [26]. For this reason, deep learning technology that can overcome this complexity is evolving day by day. Moreover, CNN can perform feature extraction and classification automatically using raw data, so deep learning algorithms are very popular in the field of artificial intelligence. Further, it is found to give very good results of classification studies involving both big data and small data by researchers. Thanks to the CNN algorithm, ECG signals can be analyzed and observed on smartphones, watches, Holter monitoring devices, etc. [3].

The CNN processes an image in different layers and separates all its features. The most commonly used layers are:

1. Convolution layer
2. Nonlinear layer
3. Pooling layer
4. Flattening layer
5. Fully connected layer

expressed as [5, 27, 28].

1. Convolutional Layer: The convolution process is the layer where the features of the image are determined. To determine more than one feature, the number of convolutional layers increases in the same proportion. This layer is the main building block of CNN.
2. Non-Linear Layer: This layer is also known as the activation layer. It is used to realize the activation of the system with nonlinear functions. Rectified Linear Unit function (ReLU), which is widely used because it is faster than others, is preferred in recent years.
3. Pooling Layer: Smaller matrices are obtained while preserving the properties of the existing input. In this way, the computational complexity is reduced.
4. Flattening Layer: The matrix format data obtained from the previous step is prepared following the fully connected layer.
5. Fully-Connected Layer: It is the most important layer of convolutional neural network layers. The data is taken from the flattening layer and trained by the neural network and the learning process is performed.

2.4. Pre-trained architectures: AlexNet and SqueezeNet

AlexNet[20] was five convolution layers combined with max-pooling layers, 3 fully connected layers. It was also a dropout layer and a softmax. Moreover, each layer was activated with the ReLU activation function. Firstly, in 2012, it was used the ReLU activation function in place of the tanh function [29]. Thus, the architecture was accelerated. The total number of parameters is 62.3 million. The input image size is 227x227x3.

SqueezeNet[21] started with an independent convolutional layer (conv1), followed by eight firing modules, and ended with the last convolutional layer (conv10). In total, it consisted of ten convolutional layers, some max-pooling layers, and a SoftMax layer, in the recently presented version.

In this study, a novel CNN architecture that contains 34 layers is presented and it is trained much more than once. This proposed CNN classifies more effectively than pre-trained CNN architectures and overcomes overfitting thanks to the dropout layer. Fig.3 is a schematic of the proposed CNN.

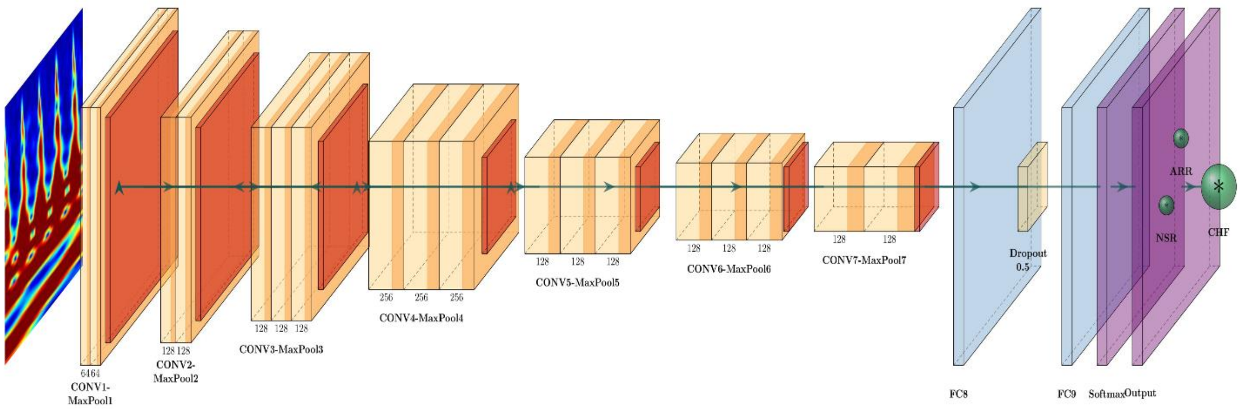

Fig. 3. Scheme of proposed CNN architecture

2.5 Novel Proposed CNN Architecture

A CNN architecture usually consists of an input layer, some convolutional layers, some pooling layers, and a fully connected layer [11]. In this study, we introduce a novel CNN architecture which are seven convolutional layers, seven batch normalization layers, seven activation layers (ReLU), seven maximum pooling layers, and two fully connected layers with one dropout layer. Additionally, a SoftMax layer and a classification layer with an entropy approach are used as well. The convolution layers are effectively utilized for feature extraction from ECG data. This is important since well feature extraction is also meaning very sensitive classification. Essentially, these layers are filtered to enhance the features of the primary signal while reducing the noise [11, 30]. The pooling layers reduce the dimension of the input images, and these are prepared for the next layer. Finally, extensive features in the fully connected layers are reduced with 0.5 probability by using the dropout layer and transferred to the SoftMax layer for the classification. Details of the parameters of the proposed CNN are given in Table 1.

The proposed CNN is a novel architecture that has different filter sizes, number of filters, strides, and padding. Fundamentally, we develop the architecture for biomedical image classification. However, it is tested on known classical datasets such as CIFAR-10, like pre-trained architectures. Additionally, it is also utilized on Physikalisch-Technische Bundesanstalt (PTB) Diagnostic ECG Database[23, 31]. And, this architecture is observed successfully in all these datasets. Moreover, this architecture is compared with the pre-trained AlexNet and SqueezeNet architectures with the same training options parameters. Here, the optimization method has been chosen as the stochastic gradient descent, and the momentum parameter is determined as 0.95, and the learning rate is also started with 0.0001. As a comparison result, our novel proposed CNN architecture is shown higher classification performance than others.

2.6. Support Vector Machine (SVM)

Support Vector Machine (SVM) is a machine learning algorithm that an effective separation with a kernel-based method to the datasets for classification or regression [32]. It is improved by Vapnik and Cortes [33] for two classes. Then the algorithm is advanced and generalized for multi-class and non-linear datasets. In general, the dataset can be separated in high-dimensional feature space with a kernel function. Also, SVM can be overcome confused datasets and extreme learning. The most common representation of the SVM function is $f(x) = w^T\phi(x) + b$ where $w \in R^n$ $b \in R$ and $\phi(x)$ is a feature map.

2.7. Modified Proposed CNN+ SVM architecture

In this present study, the proposed CNN has been designed with an SVM algorithm. In the modified model is assembled the best properties of the proposed CNN and SVM. Here, it is taken into consideration CNN whose extracts features from a dataset, automatically. Briefly, how CNN-SVM works are explained as follows: First, the proposed CNN is trained with all layers. Then, the fully connected layer (FC-8) is removed from the proposed CNN architecture. Here, 4096x (number of images) dimensional features are obtained from the dataset. In the following stage, the dataset is divided into a 30% training set and 70% testing set and these features are activated as training features and testing features. Highlighted in this study, a very large percentage is allocated to the test set and the robustness of the study is tried to be determined. After, these training features are trained with SVM which is a Gaussian kernel function and one versus all method. Finally, the classification of the testing features is obtained with an SVM classifier. Similarly, the same stages are performed for the maximum pooling layer (Max-Pooling 7) which is removed from the architecture. This study is essenced as a matter of fact that since the convolutional layers do not have a lot of parameters, extreme learning is an issue and so, the dropout layer would not have much influence [34, 35].

Both of them are observed to increase classification success. However, a higher success rate is achieved by removing the Max-Pooling 7 layer. Fig. 4 demonstrates the scheme modified with the SVM algorithm of the proposed CNN architecture.

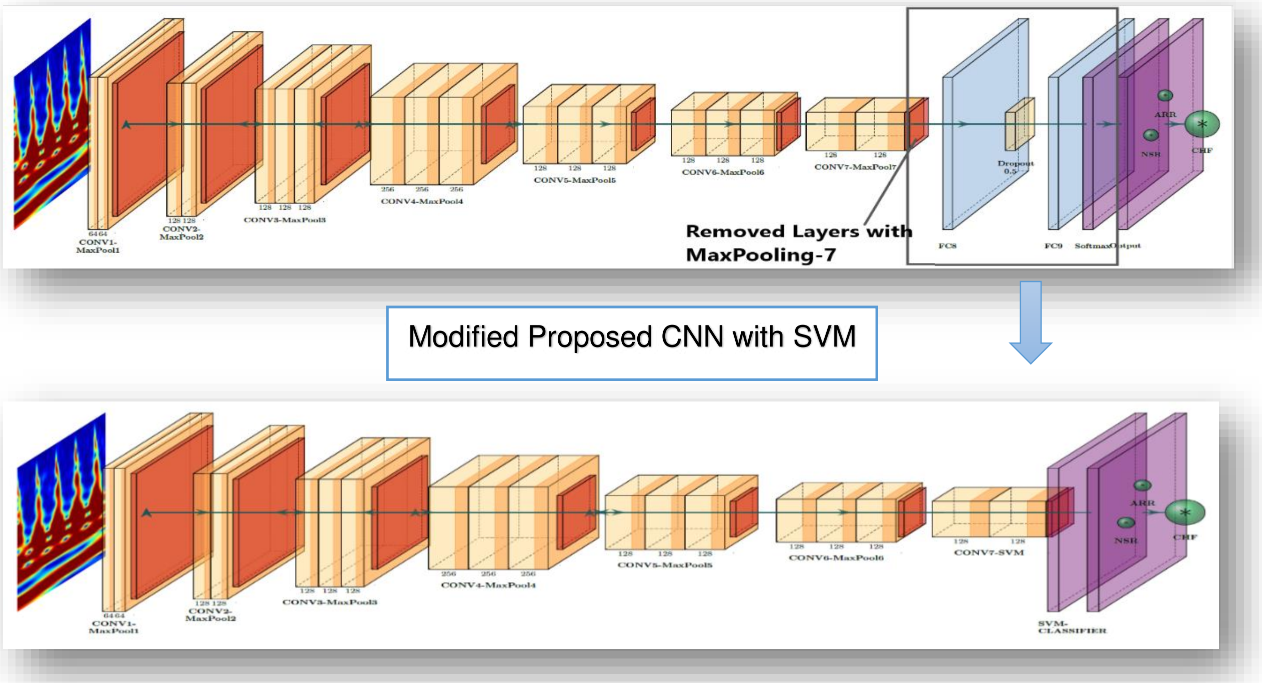

Fig. 4. Scheme of modified proposed CNN-SVM architecture

Table 1

Layers details of Proposed CNN architecture.

Layer Name	*Type*	*Layer Parameters*	*Output Shape*
Input	Image Input	227x227x3 images with 'zerocenter' normalization	227x227x3
Conv-1	Convolution 2D	Filter size=64, Number of filters=[5 5], Stride=[1 1], Padding=[1 1 1 1], BatchNormalization, ReLU	225x225x64
MaxPool-1	Max Pooling	Pool size=[3 3], Stride=[2 2], Padding=[0 0 0 0]	112x112x64
Conv-2	Convolution 2D	Filter size=128, Number of filters=[3 3],Stride=[1 1], Padding=[1 1 1 1],BatchNormalization, ReLU	112x112x128
MaxPool-2	Max Pooling	Pool size=[3 3], Stride=[2 2], Padding=[0 0 0 0]	55x55x128
Conv-3	Convolution 2D	Filter size=128,Number of filters=[13 13],Stride=[1 1], Padding=[0 0 0 0],BatchNormalization, ReLU	55x55x128
MaxPool-3	Max Pooling	Pool size=[3 3],Stride=[2 2], Padding=[0 0 0 0]	27x27x128
Conv-4	Convolution 2D	Filter size=256, Number of filters=[7 7],Stride=[1 1], Padding=[1 1 1 1],BatchNormalization, ReLU	27x27x256
MaxPool-4	Max Pooling	Pool size=[2 2], Stride=[2 2], Padding=[0 0 0 0]	13x13x256
Conv-5	Convolution 2D	Filter size=128, Number of filters=[3 3],Stride=[1 1], Padding=[1 1 1 1],BatchNormalization, ReLU	13x13x128
MaxPool-5	Max Pooling	Pool size=[3 3],Stride=[2 2], Padding=[0 0 0 0]	6x6x128
Conv-6	Convolution 2D	Filter size=128, Number of filters=[3 3],Stride=[1 1], Padding=[1 1 1 1],BatchNormalization, ReLU	6x6x128
MaxPool-6	Max Pooling	Pool size=[3 3],Stride=[2 2], Padding=[0 0 0 0]	3x3x128
Conv-7	Convolution 2D	Filter size=128, Number of filters=[3 3],Stride=[1 1], Padding=[1 1 1 1],BatchNormalization, ReLU	3x3x128
MaxPool-7	Max Pooling	Pool size=[2 2],Stride=[2 2], Padding=[0 0 0 0]	1x1x128
FC-8	Fully Connected	4096	1x1x4096
Drop-8	Dropout	50%	
FC-9	Fully Connected	3 (number of class)	1x1x3
Softmax	Softmax		1x1x3
Output	Classification	Cross entropy	

2.8. Performance Metrics

CNN architectures are evaluated in terms of performance metrics which are overall accuracy, sensitivity, specificity, precision, and F1-Score, as follows [9, 29]:

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \times 100\% \tag{8}$$

$$Sensitivity = \frac{TP}{TP + FN} \times 100\% \tag{9}$$

$$Specificity = \frac{TN}{TN + FP} \times 100\% \tag{10}$$

$$Precision = \frac{TP}{TP + FP} \times 100\% \tag{11}$$

$$F1 - Score = \frac{2 \times Precision \times Sensitivity}{Precision + Sensitivity} \times 100\% \tag{12}$$

where, TP: True Positive, FP: False Positive, TN: True Negative, and FN: False Negative is expressed in Equation (8-12).

3. Experimental Results

In recent studies, it is observed that the CWT-CNN hybrid structure is used in the classification of ECG signals. In this regard, this study is also benefited from this hybrid structure. In the present study, a novel CNN architecture is suggested using an ECG dataset. The highlighted of the study can be grouped under four main heading as follows:

- **i.** Pre-processing phase: ECG signals are normalized with the minimum-maximum normalization method. Following, one-dimensional signals are transformed to two-dimensional scalograms by using CWT.
- **ii.** Comparison phase: Chosen the best sampling size, the best mother wavelet function, and the best architecture via performance metrics.
- **iii.** Feature Extraction phase: ECG scalograms are trained with proposed CNN architecture.
- **iv.** Classification phase: ECG scalograms are classified by using SoftMax (cross-entropy approach), SVM, and their features are classified with proposed CNN-SVM.

3.1. Pre-Processing Phase

In first, raw one-dimensional ECG signals are normalized with the minimum-maximum normalization method.

$$X = \frac{signal - \min(signal)}{\max(signal) - \min(signal)} \tag{13}$$

where X is denoted normalized ECG signal. Here, $\min(.)$ is minimum function and $\max(.)$ is maximum function. Next, one-dimensional normalized ECG signals taken randomly with different sampling lengths of signals, 360hz, 500hz, and 1000hz are transformed into two-dimensional scalograms with CWT. Besides, three different mother wavelet functions namely Amor, Bump, and Morse are applied to each sampling length. Also, the size of the scalograms is set as 227x227x3 and .jpg format. For each class (ARR, CHF, and NSR) are created randomly 300 scalograms. In total, there are 900 scalograms for each ECG dataset. Therefore, 3 different ECG scalogram datasets are generated with 360 Hz sampling lengths of signals, 3 different ECG scalogram datasets are designed with 500 Hz sampling lengths of signals and 3 different ECG scalogram datasets are composed with 1000 Hz sampling lengths of signals. Finally, 9 different ECG scalograms datasets are prepared for classification. Further, these datasets randomly are divided 80% for training and 20% for testing.

3.2. Comparison phase

In this phase, transformed 9 ECG datasets are classified with AlexNet, SqueezeNet, and the proposed CNN. Here, the same options parameters such as epochs, learning rate, batch size, optimization method, etc. are utilized for each training phase in order to a fair comparison. When these architectures are trained separately for all datasets throughout 476 iterations, the overall test accuracy results for classification success are shown in Fig. 5 and Fig.6.

When Fig.5 is investigated in the matter of sampling size, 360 Hz is not contained sufficient information to classify scalograms because all architecture's performance metrics are lower than others. Besides, 1000 Hz is not increased its distinctiveness for classification performances. However, this situation is different for the 500 Hz sampling size. According to performance metrics of these architectures, scalograms wavelet with 500 Hz sampling size or signal length is observed that it became clear. Thus, the best sampling size or signal length is determined as 500 Hz.

When Fig.5 is also examined in terms of a mother wavelet function, Amor and Morse give nearly similar results to classify scalograms for AlexNet and our proposed CNN. However, these results do not give for SqueezeNet. When SqueezeNet is investigated in regards to mother wavelet function, Bump is found the best for its. Hence, if researchers would like to utilize SqueezeNet, they can choose the Bump wavelet function when they perform CWT. When Fig.6 is inspected for the proposed CNN in the matter of the mother wavelet function, choosing Amor to classify the scalograms is the best one.

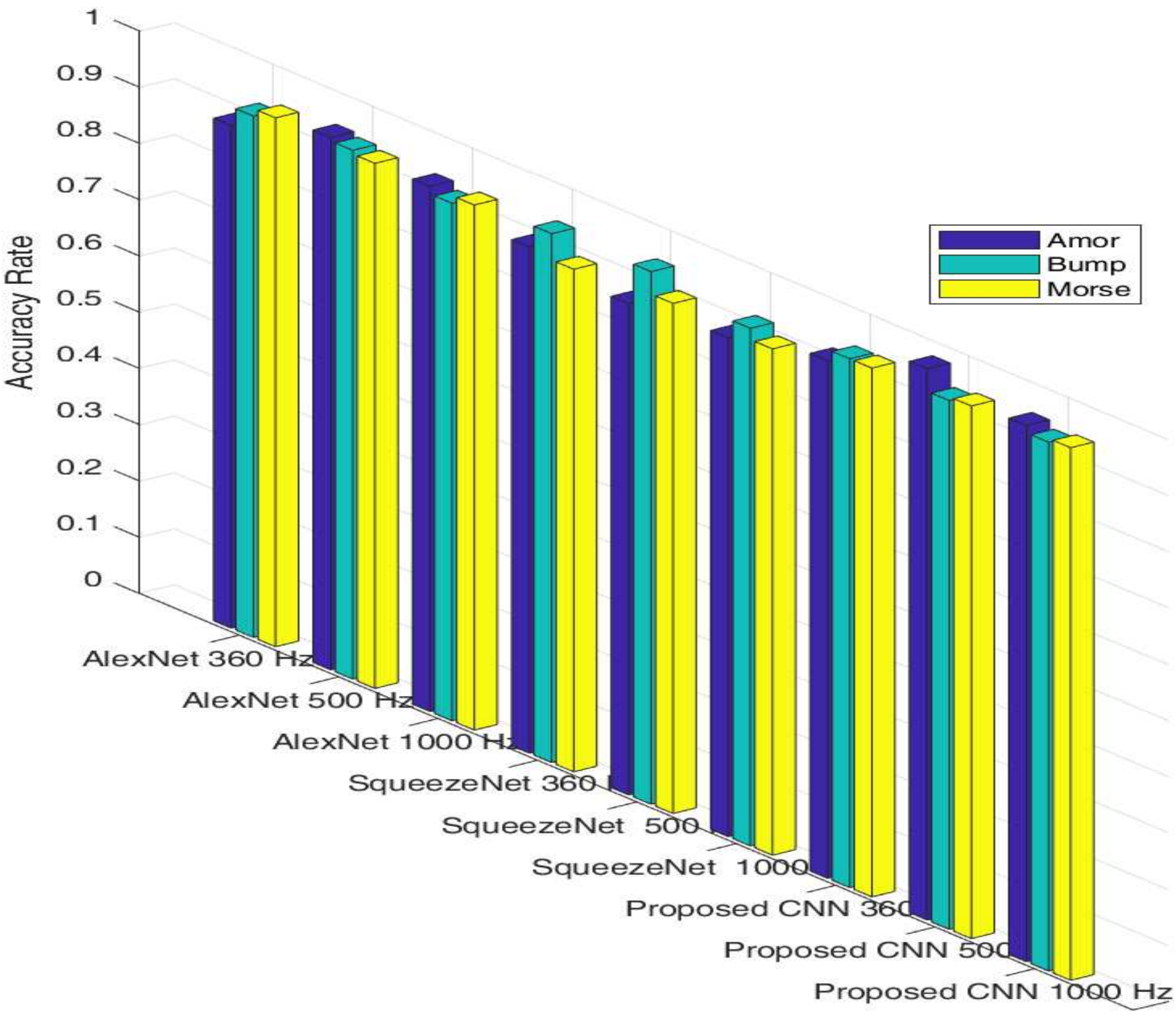

Fig. 5. The comparison of classification performance on different sampling lengths and mother wavelet function for CNNs architectures

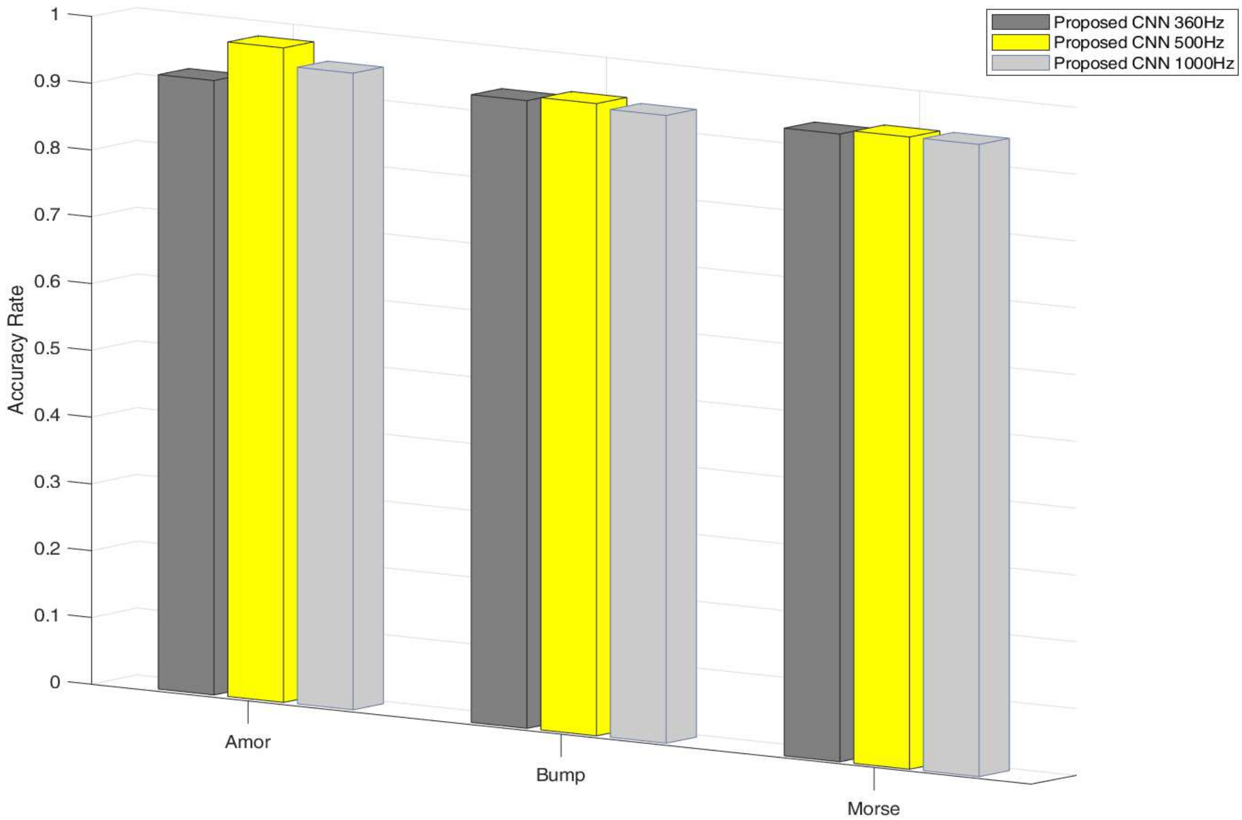

Fig. 6. The comparison of classification performance on different sampling lengths and different mother wavelet functions for Proposed CNN

In the comparisons is demonstrated that the proposed CNN is offered successful results in terms of overall accuracy. As well as the overall accuracy, the other performance metrics are also provided good results for the proposed CNN. It is shown the details for each comparison performed, in Table 2.

When all performance metrics are investigated, this proposed CNN's metrics are observed over 96%. Notably, NSR performances regarding specificity and precision value are viewed as %100. Additionally, its performances in terms of other metrics are also over than %98.

When it is examined to metrics what the classifiers performed well, it is prominent that the F1 Score of the proposed CNN is superior to the others, in Table2. Hence, the proposed CNN is determined as the best classifier concerning performance metrics.

As a result of this part, the best signal length, the best mother wavelet function, and the best architecture are determined as 500 Hz, Amor, and the proposed CNN, respectively. Thus, these foundations have shown that just an ECG dataset is classified. In addition, Fig.7 shows the accuracy rate graph and loss graph for Proposed CNN while the signal length is 500 Hz and the wavelet function is "Amor".

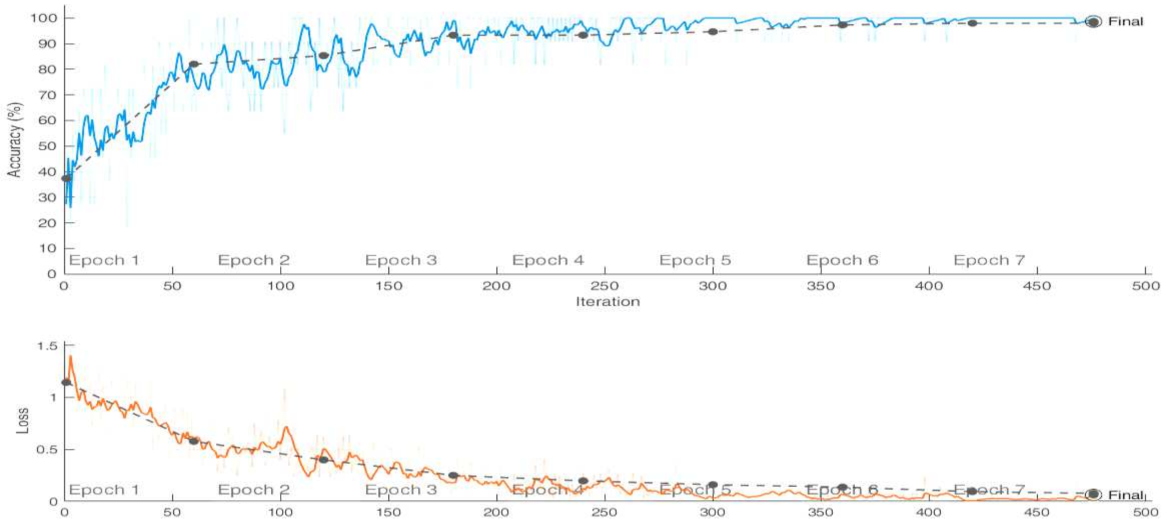

Fig. 7. Accuracy rate and loss graph of training progress for proposed CNN

Table 2

Performance metrics of proposed CNN, AlexNet, and SqueezeNet architectures.

CNN Architecture	Class Name	Sensitivity (%)	Specificity (%)	Precision (%)	F1-Score (%)	Test Accuracy Rate(%)
Proposed-CNN[a]	*ARR*	97.96	98.02	96	96.97	**98**
	CHF	98	99.02	98	98	**98**
	NSR	98.04	**100**	**100**	99.01	**98**
Proposed-CNN[b]	*ARR*	93.88	96.04	92	92.93	95.33
	CHF	96.15	**100**	**100**	98.04	95.33
	NSR	95.92	98.01	94	94.95	95.33
AlexNet[a]	*ARR*	92.31	97.96	96	94.12	94.67
	CHF	97.87	96.12	92	94.85	94.67
	NSR	94.12	98.6	96	95.05	94.67
SqueezeNet[a]	*ARR*	94.34	100	100	97.09	94.67
	CHF	90.57	97.94	96	93.20	94.67
	NSR	100	96.05	88	93.62	94.67

[a]ECG Signal Length 500Hz ,[b]ECG Signal Length 1000Hz

3.3. Feature Extraction

Here, our proposed CNN is trained for the ECG dataset. Essentially, the proposed CNN is considered both as a feature extractor and a classifier.

In this phase, the proposed CNN is trained 5 times in a loop to measure its performance, detailed in Table 3. As a result, its mean values and standard deviations are calculated for all performance metrics, also shown in Table 3. All mean performance metrics are observed over 96.53% and also maximum standards deviation (Std) has been 0.0173. Therefore, the proposed architecture is traditionally trained and saved to classify scalograms.

Table 3

Performance metrics of proposed CNN when it is trained 5 times.

Training Number	Sensitivity (%)	Specificity (%)	Precision (%)	F1-Score (%)	Test Accuracy Rate (%)
1	94,86	97,60	94,67	94,66	94,67
2	95,53	97,72	95,33	95,33	95,33
3	96,05	98,12	96,00	96,00	96
4	98,06	99,03	98,01	97,99	98
5	98,68	99,45	98,67	98,67	98,67
Mean + Std	**96,64± 1,65**	**98,38±0,82**	**96,54± 1,72**	**96,53± 1,73**	**96,53±1,73**

Then, the Fully connected (FC-8) layer is removed from the proposed CNN. Namely, this layer is included 4096x900 dimensional features. As known, every image has 227x227x3 dimensions and so 4096 features are provided by CNN via a fully connected layer for each image, automatically. Next, it is settled and activated the SVM method to the proposed CNN. In addition, Max-Pooling 7 layer is also removed from the proposed CNN owing to investigating the performance of classification, similarly.

3.4. Classification Phase

Consequently, all performance metric values are observed to increase for two different methods. However, the highest accuracy rate of the proposed CNN-SVM is achieved by removing the Max-Pooling 7 layer from the CNN, detailed in Table 4. For this reason, activation of SVM is preferred via removing the Max pooling-7 layer, in this study.

Table 4

Performance metrics of modified proposed CNN+SVM architecture

Removed Layer Name	Sensitivity (%)	Specificity (%)	Precision (%)	F1-Score (%)	Test Accuracy Rate(%)
MaxPooling-7	99,206	99,66	99,213	99,206	99,21
FC-8	98,72	99,5	98,75	98,732	98,73

This study is not only performed CNN but also performed SVM which has very high success in image classifying, is also studied. The combination of these two methods, which have separately high successes, performed very well. Table 5 shows a comparison of all methods in terms of performance inferences. Moreover, Fig. 8 indicates the confusion matrix of the modified proposed CNN-SVM which has the highest performing in terms of all performance inferences.

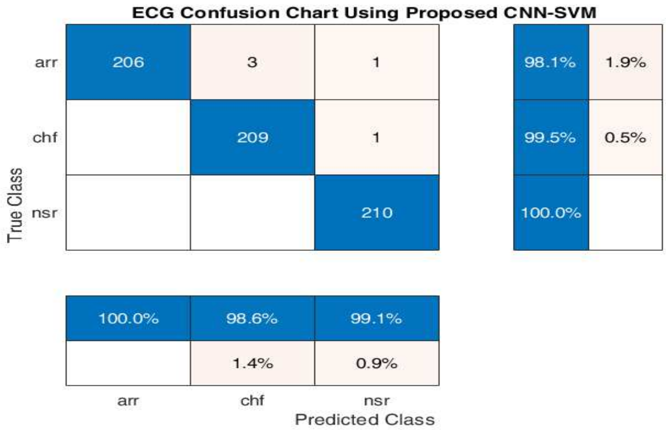

Fig. 8. Confusion matrix of modified proposed CNN+SVM

Table 5

Comparison of all methods in terms of performance inferences.

Classification Algorithm	Sensitivity (%)	Specificity (%)	Precision (%)	F1-Score (%)	Test Accuracy Rate(%)
Proposed CNN	96,64	98,38	96,54	96,53	96,53
SVM	85,56	93,68	85,56	85,51	85,56
Proposed CNN-SVM	99,206	99,66	99,213	99,206	99,21

4. Discussion

Many approaches are used for the classification of arrhythmia (ARR), congestive heart failure (CHF), and normal sinus rhythm (NSR) datasets. Essentially, successful classification is very important for diagnosis and treatment. For this reason, we suggest a novel deep learning algorithm that has 34 layers, mentioned as the proposed CNN, in this study. As well as this ECG dataset, other datasets have been also classified with our proposed CNN, such as the PTB ECG dataset and CIFAR-10 dataset. As known, the presented CNN (pre-trained) architectures are tested on the traditional dataset. Additionally, the proposed CNN architecture is also tested on the CIFAR-10 dataset and it is investigated whether it could make a successful classification, in this study.

CIFAR-10 dataset consists of 10 classes and 60,000 images. Similarly, this huge dataset is also split into 80% for training, 20% for testing as presented in the study. Thus, 50,000 images are trained and 10,000 images are tested as well. Besides, the same options parameters are applied for both datasets. In Table 6, it is demonstrated the proposed CNN of success on different datasets. Moreover, Fig.9 shows the confusion matrix for the CIFAR-10 dataset.

Table 6

The proposed CNN performance on different datasets.

Datasets	Number of Class	Sensitivity (%)	Specificity (%)	Precision (%)	F1-Score (%)	Test Accuracy Rate (%)
PTB ECG Dataset [31]	2	96,42	94,96	95	95,56	95,6
CIFAR-10	10	83,95	98,22	84,10	83,87	84
ECG Dataset in this Study	3	96,64	98,38	96,54	96,53	96,53

As seen, the performance of the proposed CNN is very well. However, as mentioned previously, this CNN must be wonderful for classification biomedical signals or images. Therefore, the proposed CNN is modified with SVM for perfect classification. In general, if any CNN architecture has a fully connected layer, this layer is removed from the CNN and settled to SVM. Of course, this method provides good advantages because of extracted features. However, Deep Learning Algorithm (also CNN) is a complex non-linear model and it is said a black box. [36]. Accordingly, it must be investigated what last layers contain good features within this probabilistic process. Under all these considerations, the features in the Max-pooling7 (just previous of the FC-8 layer) are also studied, in the present study. According to the findings obtained in this study, it is necessary to investigate the features in the last layers for more sensitive analysis, detailed in Table4.

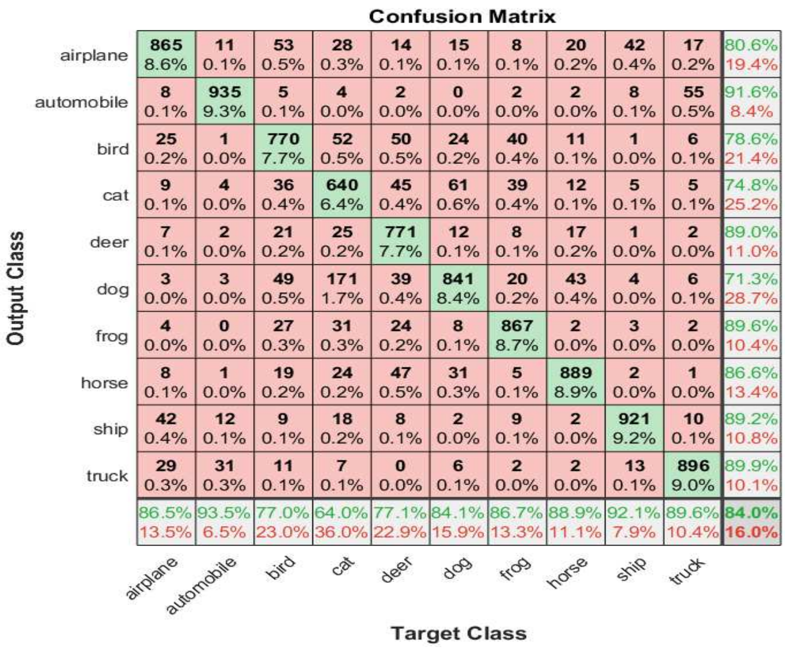

Fig. 9. Confusion matrix of the proposed CNN on the CIFAR-10 dataset

In this study, the aggregate classification is considered significant so the overall test accuracy is evaluated and detailed in Table 2-5. Apart from this, the inference of performance is appraised over the pre-trained architectures that are well-known in other studies. And, the proposed CNN is superior to them on the ECG scalograms dataset as well as sensitivity results are better than others. When the literature reviews are investigated in the classification on the same property ECG dataset which is shown in Table 7, the proposed CNN-SVM achieves the highest performance in terms of overall accuracy rate.

Table 7

The comparison of classification performances for different studies on ECG signals.

Study	Pre-processing Method	Algorithm	Accuracy Rate(%)
Çınar and Tuncer[37]	STFT	CNN (AlexNet-SVM)	96,77
Eltras et al. [38]	CQ-NSGT*	CNN(AlexNet)	98,82
Gaddam et al.[39]	CWT	CNN(AlexNet)	95,67
Golgowski and Osowski[40]	CWT	CNN	82,06
	DWT	Extra Random Forests	97,78
Krak et al. [4]	CWT	CNN	96
Krishnakumar et al.[41]	CWT	CNN (GoogleNet)	96,88
Kumari et al. [42]	DWT	SVM	95,92
Nahak and Saha [43]	RR	SVM	86,77
	Wavelet with AR	SVM	92,22
	Fusion of Features	SVM	93,33
Olanrewaju et al. [44]	CWT	CNN(AlexNet)	98,7
Rahuja and Valluru [45]	CWT	CNN(AlexNet)	97,3
Proposed CNN	**CWT**	**CNN**	**96,53**
Proposed CNN-SVM	**CWT**	**CNN-SVM**	**99,21**

*Constant-Q Non-Stationary Gabor Transform

Conclusion

Many of the sudden deaths due to heart disease continue to increase these days with Coronavirus (COVID-19). Based on this, the automatic classification of signals received from the heart is of great importance for diagnosis and treatment. In this study, we classify ECG data using our proposed CNN that overcame overfitting with the dropout layer. This CNN is also performed on other datasets, shown in Table 6. In addition, the proposed CNN is compared with AlexNet and SqueezeNet on 9 different ECG datasets whose is prepared via CWT using 3 different wavelet functions and 3 different signal lengths. All results show that the best signal length especially for two-dimensional scalograms is 500Hz and the best mother wavelet function is "Amor”. Besides, the proposed CNN, AlexNet, and SqueezeNet comparison of classification success in terms of overall accuracy rate are 98%, 94.67%, and 94.67% respectively. Hence, the proposed CNN architecture is performed classification on ECG dataset whose is generated with “Amor” wavelet function and 500 Hz signal lengths superior to others. Further, it is trained 5 times in a loop to measure the performance of the proposed CNN architecture, detailed in Table 3. Accordingly, it is observed all mean performance metrics are over 96.5%, and also maximum standard deviation (Std) is 0.0173 on testing the ECG dataset. The main purpose of the study is to find an excellent classification algorithm on the ECG dataset. Thus, the proposed CNN is modified with SVM. Here, it is pointed out that CNN is thought of as extracting features from the ECG dataset, automatically.

In general, if any CNN architecture has a fully connected layer, it is replaced with SVM, in the CNN-SVM studies. However, it is highlighted that it can be provided an advantage to examine features from the last layers of CNN, such as the max-pooling layer, in the present study. In order to improve the proposed CNN performance, it is removed max-pooling 7 and FC-8 layers from this CNN respectively, detailed in Table 4. As a result, the highest success with a 99.21% accuracy rate is achieved by removing the Max-Pooling 7 layer from its. When the comparison is examined with other studies on similar ECG datasets, the modified proposed CNN-SVM is observed as the highest performing for classification, detailed in Table 7.

In the next studies, the proposed CNN architecture will be tried on binary and multiclass biomedical datasets to see how it is well.

Acknowledgements

This study is based on Oznur Ozaltin's Ph.D. thesis and supervised by Ozgur Yeniay.

Authors' contributions

Ozaltin, Oznur idealized this study and analyzed the data. Yeniay, Ozgur supervised the research and approved the final draft.

Funding

The authors took on no certain funding for this study.

Declarations

Conflict of Interest

The authors announced that they had no conflicts of interest to report related to this study.

Data Availability

Data can be available from https://www.physionet.org/.

References

[1] H. Khorrami and M. Moavenian, "A comparative study of DWT, CWT and DCT transformations in ECG arrhythmias classification," *Expert systems with Applications,* vol. 37, no. 8, pp. 5751-5757, 2010.

[2] M. M. Al Rahhal, Y. Bazi, M. Al Zuair, E. Othman, and B. BenJdira, "Convolutional neural networks for electrocardiogram classification," *Journal of Medical and Biological Engineering,* vol. 38, no. 6, pp. 1014-1025, 2018.

[3] J. Huang, B. Chen, B. Yao, and W. He, "ECG arrhythmia classification using STFT-based spectrogram and convolutional neural network," *IEEE Access,* vol. 7, pp. 92871-92880, 2019.

[4] I. Krak, O. Stelia, A. Pashko, M. Efremov, and O. Khorozov, "Electrocardiogram Classification Using Wavelet Transformations," in *2020 IEEE 15th International Conference on Advanced Trends in Radioelectronics, Telecommunications and Computer Engineering (TCSET)*, 2020: IEEE, pp. 930-933.

[5] U. B. Baloglu, M. Talo, O. Yildirim, R. San Tan, and U. R. Acharya, "Classification of myocardial infarction with multi-lead ECG signals and deep CNN," *Pattern Recognition Letters,* vol. 122, pp. 23-30, 2019.

[6] T. Mahmud, S. A. Fattah, and M. Saquib, "Deeparrnet: An efficient deep cnn architecture for automatic arrhythmia detection and classification from denoised ecg beats," *IEEE Access,* vol. 8, pp. 104788-104800, 2020.

[7] M. Salem, S. Taheri, and J. S. Yuan, "ECG arrhythmia classification using transfer learning from 2-dimensional deep CNN features," in *2018 IEEE Biomedical Circuits and Systems Conference (BioCAS)*, 2018: IEEE, pp. 1-4.

[8] Y. Zhao, J. Cheng, P. Zhang, and X. Peng, "ECG classification using deep CNN improved by wavelet transform," *Computers, Materials & Continua,* vol. 64, no. 3, pp. 1615-1628, 2020.

[9] X. Xu and H. Liu, "ECG heartbeat classification using convolutional neural networks," *IEEE Access,* vol. 8, pp. 8614-8619, 2020.

[10] A. Rajkumar, M. Ganesan, and R. Lavanya, "Arrhythmia classification on ECG using Deep Learning," in *2019 5th International Conference on Advanced Computing & Communication Systems (ICACCS)*, 2019: IEEE, pp. 365-369.

[11] X. Hua *et al.*, "A novel method for ECG signal classification via one-dimensional convolutional neural network," *Multimedia Systems,* pp. 1-13, 2020.

[12] S. Kiranyaz, T. Ince, and M. Gabbouj, "Real-time patient-specific ECG classification by 1-D convolutional neural networks," *IEEE Transactions on Biomedical Engineering,* vol. 63, no. 3, pp. 664-675, 2015.

[13] 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.

[14] K. Sandeep, P. Kora, K. Swaraja, K. Meenakshi, and L. Pampana, "ECG classification using machine learning," *Int J Recent Technol Eng (IJRTE),* vol. 8, no. 4, pp. 2492-2494, 2019.

[15] E. Alickovic and A. Subasi, "Effect of multiscale PCA de-noising in ECG beat classification for diagnosis of cardiovascular diseases," *Circuits, Systems, and Signal Processing,* vol. 34, no. 2, pp. 513-533, 2015.

[16] S. M. Qaisar and A. Subasi, "Cloud-based ECG monitoring using event-driven ECG acquisition and machine learning techniques," *Physical and Engineering Sciences in Medicine,* vol. 43, no. 2, pp. 623-634, 2020.

[17] T. Tuncer, S. Dogan, P. Plawiak, and A. Subasi, "A novel Discrete Wavelet-Concatenated Mesh Tree and ternary chess pattern based ECG signal recognition method," *Biomedical Signal Processing and Control,* vol. 72, p. 103331, 2022.

[18] R. Ceylan and Y. Özbay, "Comparison of FCM, PCA and WT techniques for classification ECG arrhythmias using artificial neural network," *Expert Systems with Applications,* vol. 33, no. 2, pp. 286-295, 2007.

[19] P. Pławiak and U. R. Acharya, "Novel deep genetic ensemble of classifiers for arrhythmia detection using ECG signals," *Neural Computing and Applications,* vol. 32, no. 15, pp. 11137-11161, 2020.

[20] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "Imagenet classification with deep convolutional neural networks," *Advances in neural information processing systems,* vol. 25, pp. 1097-1105, 2012.

[21] F. N. Iandola, S. Han, M. W. Moskewicz, K. Ashraf, W. J. Dally, and K. Keutzer, "SqueezeNet: AlexNet-level accuracy with 50x fewer parameters and< 0.5 MB model size," *arXiv preprint arXiv:1602.07360,* 2016.

[22] "Physionet.". https://www.physionet.org/ (accessed (accessed 01.06.2020 14:00.).

[23] A. L. Goldberger *et al.*, "PhysioBank, PhysioToolkit, and PhysioNet: components of a new research resource for complex physiologic signals," *circulation,* vol. 101, no. 23, pp. e215-e220, 2000.
[24] 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.
[25] D. S. Baim *et al.*, "Survival of patients with severe congestive heart failure treated with oral milrinone," *Journal of the American College of Cardiology,* vol. 7, no. 3, pp. 661-670, 1986.
[26] A. Narin, "Detection of focal and non-focal epileptic seizure using continuous wavelet transform-based scalogram images and pre-trained deep neural networks," *IRBM,* 2020.
[27] U. R. Acharya, H. Fujita, S. L. Oh, Y. Hagiwara, J. H. Tan, and M. Adam, "Application of deep convolutional neural network for automated detection of myocardial infarction using ECG signals," *Information Sciences,* vol. 415, pp. 190-198, 2017.
[28] H. K. Lee and Y.-S. Choi, "Application of continuous wavelet transform and convolutional neural network in decoding motor imagery brain-computer interface," *Entropy,* vol. 21, no. 12, p. 1199, 2019.
[29] B. Abdelmalek, K. Ahmed, and T. M. Amine, "Lightweight CNNs-Based Object Detection for Embedded Systems implementation," in *1st Conference on Innovative Trends in Computer Science (CITCS'2019)*, 2019.
[30] J. Li, Y. Si, T. Xu, and S. Jiang, "Deep convolutional neural network based ECG classification system using information fusion and one-hot encoding techniques," *Mathematical Problems in Engineering,* vol. 2018, 2018.
[31] Ö. Özaltın and Ö. Yeniay, "Ecg classification performing feature extraction automatically using a hybrid cnn-svm algorithm," in *2021 3rd International Congress on Human-Computer Interaction, Optimization and Robotic Applications (HORA)*, 2021: IEEE, pp. 1-5.
[32] M. Koklu and I. A. Ozkan, "Multiclass classification of dry beans using computer vision and machine learning techniques," *Computers and Electronics in Agriculture,* vol. 174, p. 105507, 2020.
[33] C. Cortes and V. Vapnik, "Support-vector networks," *Machine learning,* vol. 20, no. 3, pp. 273-297, 1995.
[34] M. Elleuch, R. Maalej, and M. Kherallah, "A new design based-SVM of the CNN classifier architecture with dropout for offline Arabic handwritten recognition," *Procedia Computer Science,* vol. 80, pp. 1712-1723, 2016.
[35] N. Srivastava, G. Hinton, A. Krizhevsky, I. Sutskever, and R. Salakhutdinov, "Dropout: a simple way to prevent neural networks from overfitting," *The journal of machine learning research,* vol. 15, no. 1, pp. 1929-1958, 2014.
[36] R. Guidotti, A. Monreale, S. Ruggieri, F. Turini, F. Giannotti, and D. Pedreschi, "A survey of methods for explaining black box models," *ACM computing surveys (CSUR),* vol. 51, no. 5, pp. 1-42, 2018.
[37] A. Çınar and S. A. Tuncer, "Classification of normal sinus rhythm, abnormal arrhythmia and congestive heart failure ECG signals using LSTM and hybrid CNN-SVM deep neural networks," *Computer Methods in Biomechanics and Biomedical Engineering,* vol. 24, no. 2, pp. 203-214, 2021.
[38] A. S. Eltrass, M. B. Tayel, and A. I. Ammar, "A new automated CNN deep learning approach for identification of ECG congestive heart failure and arrhythmia using constant-Q non-stationary Gabor transform," *Biomedical Signal Processing and Control,* vol. 65, p. 102326, 2021.
[39] P. G. Gaddam and R. Sreehari, "Automatic Classification of Cardiac Arrhythmias based on ECG Signals Using Transferred Deep Learning Convolution Neural Network," in *Journal of Physics: Conference Series*, 2021, vol. 2089, no. 1: IOP Publishing, p. 012058.
[40] M. Golgowski and S. Osowski, "Anomaly detection in ECG using wavelet transformation," in *2020 IEEE 21st International Conference on Computational Problems of Electrical Engineering (CPEE)*, 2020: IEEE, pp. 1-4.
[41] S. Krishnakumar, M. Yasodha, J. V. Priyadharshini, J. B. Janney, S. Divakaran, and V. L. Christy, "Detection of Arrhythmia and Congestive Heart Failure Through Classification of ECG Signals Using Deep Learning Neural Network," in *2021 International Conference on Advancements in Electrical, Electronics, Communication, Computing and Automation (ICAECA)*, 2021: IEEE, pp. 1-7.

[42] C. U. Kumari *et al.*, "Heart Rhythm Abnormality Detection and Classification using Machine Learning Technique," in *2020 4th International Conference on Trends in Electronics and Informatics (ICOEI)(48184)*, 2020: IEEE, pp. 580-584.

[43] S. Nahak and G. Saha, "A fusion based classification of normal, arrhythmia and congestive heart failure in ECG," in *2020 National Conference on Communications (NCC)*, 2020: IEEE, pp. 1-6.

[44] R. F. Olanrewaju, S. N. Ibrahim, A. L. Asnawi, and H. Altaf, "Classification of ECG signals for detection of arrhythmia and congestive heart failure based on continuous wavelet transform and deep neural networks," *Indonesian Journal of Electrical Engineering and Computer Science,* vol. 22, no. 3, pp. 1520-1528, 2021.

[45] N. Rahuja and S. K. Valluru, "A Deep Neural Network Approach to Automatic Multi-Class Classification of Electrocardiogram Signals," in *2021 International Conference on Intelligent Technologies (CONIT)*, 2021: IEEE, pp. 1-4.