IRSTI 28.23.37 https://doi.org/10.26577/jpcsit2024020104

O.Zh. Mamyrbayev[1], D.O. Oralbekova[1], Sh.P. Zhumagulova[1*], E.Sh. Azanbekov[2]

[1]Institute of Information and Computational Technologies, Almaty, Kazakhstan
[2]Academy of the Border Guard Service of the National Security Committee of the Republic of Kazakhstan, Almaty, Kazakhstan
*e-mail: sh.zhumagulovakz@gmail.com

CLASSIFICATION OF DANGEROUS ARRHYTHMIAS USING ECG SCALOGRAMS WITH DEEP CONVOLUTIONAL NEURAL NETWORKS

Abstract. In modern medicine, the problem of detecting and classifying life-threatening arrhythmias based on ECG data remains relevant and critically important for continuous patient monitoring. This study is dedicated to developing a method for the automatic classification of six classes of dangerous arrhythmias using short ECG segments of 2 seconds duration. Existing methods for detecting dangerous arrhythmias require additional improvements to ensure high accuracy and efficiency. The goal of this research is to develop an effective method for the classification of dangerous arrhythmias to facilitate timely medical intervention. A unique method is proposed, based on transforming ECG signals into scalograms using continuous wavelet transformation. For arrhythmia classification, the AlexNet neural network is employed. The study utilizes data from the PhysioNet database and synthesized ECG data using the SMOTE method. Experimental investigations demonstrated a high accuracy of the proposed method, with an average accuracy of 98.7% for all arrhythmia classes, surpassing previously achieved maximum estimates by other researchers (93.18%). The study has been successfully completed, showcasing scientific novelty and practical significance of the results. The proposed method not only improved existing accuracy estimates but also emphasized the potential of using scalograms and neural networks for recognizing dangerous arrhythmias from ECG data. This opens new horizons for continuous monitoring and timely medical intervention, enhancing the quality of patient care.

Key words: heart disease, arrhythmia, classification, deep convolutional neural networks, scalograms.

1 Introduction

Cardiovascular diseases are the leading cause of death globally, claiming millions of lives each year. Within this broad category, cardiac arrhythmias—irregularities in the heart's rhythm—pose significant risks, ranging from benign palpitations to severe conditions like ventricular fibrillation and atrial fibrillation. These life-threatening arrhythmias require immediate recognition and intervention to prevent catastrophic outcomes such as cardiac arrest or stroke. Electrocardiography (ECG) is the gold standard for detecting and diagnosing arrhythmias. It provides a non-invasive and real-time snapshot of the heart's electrical activity, enabling clinicians to identify abnormal rhythms. Despite its widespread use, interpreting ECG results accurately can be complex and requires considerable expertise. In emergency situations, the need for rapid and precise interpretation becomes even more critical. Recent advancements in medical technology and computational methods have introduced new possibilities for enhancing ECG analysis. Among these innovations, ECG scalograms have emerged as a powerful tool. Scalograms are time-frequency representations of ECG signals, created through wavelet transforms. Unlike traditional time-domain ECG readings, scalograms offer a detailed and multi-dimensional view of the signal, capturing both frequency and temporal information. This enriched representation can reveal subtle patterns and anomalies indicative of arrhythmias that might be overlooked in standard ECG analysis. The integration of machine learning and artificial intelligence (AI) with ECG scalograms further amplifies their diagnostic potential. AI algorithms can be trained to recognize specific arrhythmic patterns within scalograms, enabling automated and highly accurate detection. This synergy between advanced signal processing and AI not only enhances diagnostic precision but also significantly reduces the time required for analysis, which is crucial in emergency care.In this article, we will explore the methodology behind ECG scalogram generation and their application in arrhythmia recognition. We will discuss the advantages of scalograms over traditional ECG interpretations and examine case studies where AI-driven scalogram analysis has been successfully implemented. By highlighting the cutting-edge developments in this field, we aim to underscore the transformative impact of ECG scalograms on improving the detection and management of life-threatening arrhythmias, ultimately contributing to better patient care and outcomes. Throughout this research, our focus is on the problem of recognizing critically important cardiac arrhythmias using electrocardiogram (ECG) data. Arrhythmia disease presents a significant threat to cardiovascular health, with potentially life-threatening conditions such as ventricular flutter (VF), ventricular fibrillation (VF), and ventricular tachycardia (VT) demanding precise and timely diagnosis. Central to this diagnostic endeavor is the analysis of the QRS complex, a fundamental component of the electrocardiogram (ECG) signal that reflects ventricular depolarization. [1 – 4]. Recent advances in signal processing, particularly the adoption of Continuous Wavelet Transform (CWT), have provided a deeper understanding of arrhythmic patterns embedded within ECG signals [5]. Arrhythmia analysis also involves studying the characteristics of P- and T-waves [6], RR intervals [7], intervals between ECG waves. In addition to the morphological analysis of the ECG wave complex, spectral components of the signal are applied [8], continuous wavelet transformations, and independent component analysis. Through the generation of scalograms, CWT unveils intricate temporal and frequency characteristics inherent in these irregular rhythms, offering clinicians a richer dataset for classification. In tandem, the emergence of deep learning methodologies, exemplified by architectures such as AlexNet, has revolutionized the landscape of arrhythmia classification. By harnessing the power of Convolutional Neural Networks (CNNs), AlexNet can automatically extract discriminative features from ECG scalograms, facilitating accurate and efficient classification of dangerous arrhythmic patterns[9, 10, 11]. In this article, we embark on an exploration of the intersection between signal processing and deep learning in the classification of deadly arrhythmias. We delve into the crucial role of the QRS complex in identifying VF, VF, and VT, elucidating the subtle nuances crucial for accurate diagnosis. Additionally, we examine how CWT and scalogram analysis enrich our understanding of ECG signals, providing clinicians with a comprehensive framework for classification. This article addresses the problem of classifying four categories of life-threatening arrhythmias based on short (2 seconds) segments of an electrocardiogram (ECG). The proposed classification reflects the level of threat to the patient's life, covering a scale from A1 to A4. Class A1 represents the highest level of danger, requiring immediate resuscitation. Next are classes of dangerous ventricular arrhythmias (A2), supraventricular arrhythmias (A3), and sinus rhythm (A4) [1]. Each of these classes has its subclasses, as presented in Table 1. Examples of ECG fragments for different classes of life-threatening arrhythmias are provided in Figures 1-4. Moreover, we investigate the transformative potential of deep learning architectures like AlexNet in reshaping arrhythmia diagnosis. By leveraging large datasets of ECG scalograms, CNNs can discern complex patterns with unprecedented accuracy, empowering clinicians with actionable insights for patient care and intervention. Through this comprehensive review, we aim to underscore the profound impact of ad bvanced signal processing and deep learning in the classification of dangerous arrhythmias. By integrating these innovative methodologies, clinicians can enhance diagnostic accuracy, improve patient outcomes, and ultimately mitigate the risks associated with arrhythmia disease.

In the past few years, machine learning neural network models, especially those employing deep learning techniques, have emerged as remarkably effective instruments for analyzing electrocardiograms (ECG). Within this realm, convolutional neural networks (CNN) [9, 10, 11] are extensively utilized. These models can analyze both one-dimensional fragments of the original temporal ECG signal (1D-CNN) [10, 12] and employ a two-dimensional representation of temporal segments (2D-CNN) [9, 11]. The transformation of the ECG signal into an image has become feasible due to remarkable successes achieved in image analysis using deep neural networks.

Various methods are employed to convert a one-dimensional signal into a two-dimensional image, such as Short-Time Fourier Transform (STFT) – spectrograms [13], Continuous Wavelet Transform (CWT) – scalograms [9, 11], as well as Markov Transition Fields (MTF). These techniques allow efficient representation of temporal segments of the ECG signal as images, thereby simplifying the analysis process with the involvement of convolutional neural networks.

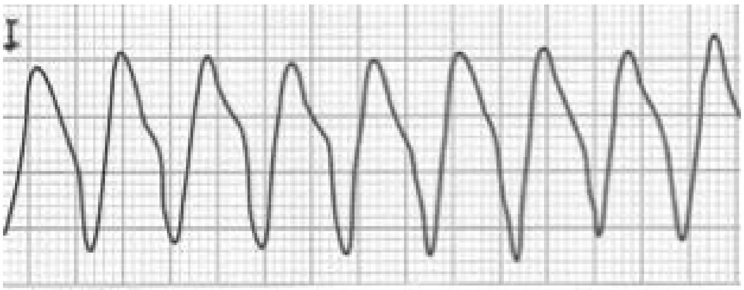


Figure 1 – Fragment of critical arrhythmias necessitating immediate resuscitation (A1)

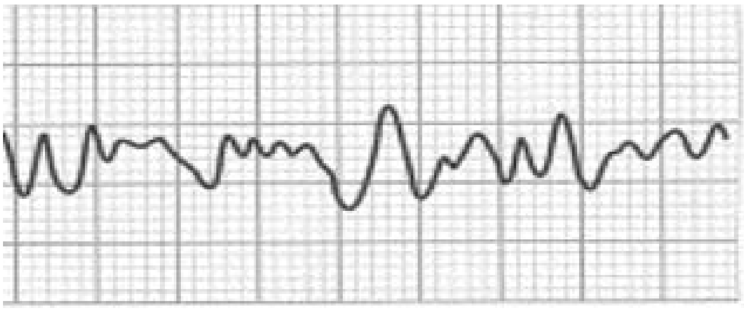


Figure 2 – Fragment of critical ventricular arrhythmias (A2)

Figure 3 – Fragment characterized by supraventricular arrhythmias (A3)

Figure 4 – Fragment of ventricular arrhythmias with potential life-threatening consequences (C4)

In recent years, there has been active research exploring the combination of Continuous Wavelet Transform (CWT) – scalograms with Convolutional Neural Networks (CNN). These studies are particularly relevant and closely related to the discussed topic.

In their publication [15], the authors introduced a technique aimed at classifying five categories of cardiovascular ailments by utilizing orthogonal leads from ECG. Employing the PTB (Physikalisch-Technische Bundesanstalt) database, they employed signal preprocessing, segmentation, and Continuous Wavelet Transform (CWT). Classification was performed using a pre-trained convolutional neural network, specifically AlexNet. Remarkably, the model achieved exceptional accuracy in classification across the specified categories when operating on 3-second scalograms. This methodology laid the groundwork for the current investigation into identifying life-threatening arrhythmias.

In study [11], researchers conducted a comparative examination of deep machine learning models applied to biometrics utilizing ECG scalograms. They introduced a biometric recognition system that employed ECG signal scalograms and deep learning models to attain remarkable accuracy in biometric identification. Findings from the research suggest that the proposed method surpasses traditional ECG-based biometric recognition techniques, highlighting its efficacy.

The study outlined in reference [12] introduces an innovative approach for automatically distinguishing between shockable and non-shockable ventricular arrhythmias, employing a one-dimensional Convolutional Neural Network (1D CNN). The authors processed 2-second ECG fragments using an eleven-layer CNN model for precise identification of life-threatening ventricular arrhythmias. The experiments demonstrated the effectiveness of the proposed approach, achieving a maximum accuracy of 93.18%.

In reference [14], an automated ECG classification method is introduced, which combines Continuous Wavelet Transform (CWT) with Convolutional Neural Networks (CNN). Alongside scalograms, the authors incorporated four RR interval features extracted during preprocessing. These features were integrated with CNN operations, feeding into a fully connected layer specifically designed for ECG classification. The study showcased the effectiveness of this approach in automating ECG classification.

In their study [9], researchers introduced a model that integrates Continuous Wavelet Transform (CWT) with the deep neural network CNN AlexNet. This model aims to classify various types of cardiac arrhythmias and congestive heart failure using ECG data. The research emphasized the effectiveness of this approach in accurately predicting prevalent heart conditions such as arrhythmias and congestive heart failure. Such investigations significantly progress ECG analysis methodologies by combining the benefits of continuous wavelet transform with the capabilities of convolutional neural networks.

In this work, we apply Continuous Wavelet Transform to 2-second ECG fragments and utilize transfer learning with a pre-trained convolutional neural network, AlexNet. The experimental results show an average classification accuracy of 96.2% across all classes. This performance significantly surpasses previous maximum accuracy estimates of 93.18% using a 1D-CNN neural network [12] and 94.12% using a Long Short-Term Memory (LSTM) neural network [8], both on similar ECG data classes.

2 Description of ECG Data

This research utilizes a database of ECG fragments, which is an essential resource for medical professionals and researchers studying cardiac arrhythmias. This database [1] contains a set of 2-second fragments with rhythm disturbances. The fragments are categorized into distinct classes based on the level of danger posed to the patient's life. This database is designed for practical application in the development and testing of algorithms capable of detecting dangerous arrhythmias in continuous monitoring systems. It contains 1016 electrocardiogram ECG fragments, each marked with one of four classes. These classes encompass: A1 - Life-threatening arrhythmias necessitating immediate resuscitation, A2 - Life-threatening ventricular arrhythmias, A3 - Supraventricular arrhythmias, and A4 - Sinus rhythm. The quantitative composition of various types of arrhythmias is presented in Table 1.

Table 1 – Content of the ECG database

Class	Types of arrhythmias	Number of segments	In total within the class of segments
A1	VF	240	337
	VFL	97	
A2	VTHR	169	169
A3	NOD	12	106
	BI	8	
	SBR	1	
	SVTA	39	
	AFIB	46	
A4	Ne	40	200
	N	107	
	BBB	53	

In table 1 the following abbreviations are used: VF - ventricular fibrillation, VFL - ventricular flutter, VTHR - high-frequency ventricular tachycardia, NOD – nodal (a-v) rhythm, BI – first degree heart block, SBR – sinus bradycardia, SVTA – supraventricular tachycardia, AFIB – atrial fibrillation , Ne – normal rhythm with single extrasystoles, N – normal sinus rhythm , BBB – sinus rhythm with bundle branch block. Apart from utilizing fragments extracted from the ECG database, we expanded our dataset by integrating synthetic data generated through the SMOTE method. This strategic augmentation was employed to rectify the inherent imbalance in our dataset, thereby fortifying the reliability and comprehensiveness of our analysis. SMOTE (Synthetic Minority Over-sampling Technique) is a popular algorithm used in the field of machine learning for handling imbalanced datasets. Imbalanced datasets are those where the number of examples in one class (the minority class) is much smaller than the number of examples in another class (the majority class). In the real world, oftentimes we end up trying to train a model on a dataset with very few examples of a given class (e.g. rare disease diagnosis, manufacturing defects, fradulent transactions) which results in poor performance. Due to the nature of the data (occurrences are so rare), it's not always realistic to go out and acquire more. One way of solving this issue is to under-sample the majority class. That is to say, we would exclude rows corresponding to the majority class such that there are roughly the same amount of rows for both the majority and minority classes. However, in doing so, we lose out on a lot of data that could be used to train our model thus improving its accuracy (e.g. higher bias). Another other option is to over-sample the minority class. In other words, we randomly duplicate observations of the minority class. The problem with this approach is that it leads to overfitting because the model learns from the same examples. This is where SMOTE comes in. At a high level, the SMOTE algorithm can be described as follows:

- Take difference between a sample and its nearest neighbour
- Multiply the difference by a random number between 0 and 1
- Add this difference to the sample to generate a new synthetic example in feature space
- Continue on with next nearest neighbour up to user-defined number

Advantages of SMOTE:

Improved Model Performance: By balancing the dataset, models can better learn the characteristics of both classes, improving metrics like recall and F1-score for the minority class.

Versatility: SMOTE can be applied to various types of data, including numerical, categorical (with modifications), and even text or image data with appropriate feature representations.

Simple and Effective: SMOTE is relatively easy to implement and often provides substantial performance improvements for imbalanced datasets.

SMOTE works by creating synthetic examples from the minority class rather than by over-sampling with replacement. It does this by selecting examples that are close in the feature space, drawing a line between the examples in the feature space and generating new examples along this line. This helps to balance the class distribution in the dataset, which can improve the performance of machine learning models, especially when the minority class is important. The Synthetic Minority Over-sampling Technique algorithm works:

• Identify Minority Class Instances: First, the algorithm identifies the instances belonging to the minority class in the dataset. These are the instances that are relatively rare compared to the majority class.

• Select Nearest Neighbors: For each minority class instance, the algorithm finds its k nearest neighbors. The number of neighbors to consider (k) is usually specified by the user.

• Randomly Generate Synthetic Instances: For each minority class instance, SMOTE randomly selects one of its nearest neighbors and creates a synthetic instance along the line joining the two instances in the feature space. The algorithm generates synthetic instances until the desired balance between the minority and majority class is achieved.

• Repeat if Necessary: Depending on the desired level of imbalance reduction, SMOTE may repeat the process multiple times, creating more synthetic instances for each minority class instance.

By generating synthetic instances, SMOTE effectively increases the representation of the minority class in the dataset, making it more balanced and improving the performance of machine learning algorithms trained on imbalanced datasets.

3 Continuous Wavelet transform

The continuous wavelet transform (CWT) of an ECG (Electrocardiogram) signal is a mathematical operation used to analyze the signal's time-frequency characteristics. It involves decomposing the signal into different scales and frequencies using wavelets, which are small, localized functions. This transformation helps in identifying important features of the ECG signal across both time and frequency domains, aiding in tasks such as heartbeat detection, arrhythmia diagnosis, and signal denoising. Converting an ECG signal from the time domain to the time-frequency domain using Continuous Wavelet Transform (CWT) offers several advantages:

• Time-Frequency Localization: CWT provides excellent localization in both time and frequency domains, allowing for precise identification of signal components and features at different scales.

• Multiresolution Analysis: CWT decomposes the signal into multiple scales, enabling the detection of both low and high-frequency components simultaneously. This is crucial for capturing important details in ECG signals, such as P-waves, QRS complexes, and T-waves, which occur at different frequencies.

• Adaptability to Signal Variability: ECG signals can vary significantly in terms of amplitude, frequency, and duration due to factors like patient condition and electrode placement. CWT adapts to these variations by adjusting the scale of the wavelet basis function, ensuring effective analysis across different signal characteristics.

• Feature Extraction: The time-frequency representation obtained through CWT facilitates efficient feature extraction from ECG signals. This is essential for automated ECG analysis tasks, including heartbeat classification, arrhythmia detection, and ischemia monitoring.

• Noise Robustness: CWT offers inherent noise robustness by focusing on localized signal characteristics. This helps in reducing the impact of noise and interference present in ECG recordings, leading to more accurate analysis results.

• Interpretability: The time-frequency representation obtained from CWT provides intuitive visual insights into the signal's characteristics, making it easier for clinicians and researchers to interpret and analyze ECG data.

The continuous wavelet transform at different time scales characterizes the signal in different frequency ranges, while the discrete wavelet transform (DWT) is limited to scales that are powers of two. Using CWT instead of DWT provides more options. Let s be a signal and ψ a wavelet. With continuous transformation, the wavelet coefficients of the signal s, corresponding to the scale factor a and position b, are determined by formula (1)[17]:

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

Here, s(t)- represents the given signal.

In this study, the results of Continuous Wavelet Transform (CWT) yield a collection of wavelet coefficients dependent on scale (a) and shift (b). These coefficients are utilized in the form of a scalogram, serving as the input [15] to the deep neural network AlexNet for disease classification. Figure 5a displays the original ECG signal, while Figure 5b illustrates the wavelet scalogram of the ECG signal. The wavelet scalogram serves as a three-dimensional representation of one-dimensional time signal data. Time is depicted on the X-axis, frequency on the Y-axis, and the Z-axis (color-coded from low values in blue to high values in red) represents the outcome of the wavelet transform of the signal at each time and frequency point.

4 AlexNet – the utilized machine learning model

Cardiac arrhythmia is a condition characterized by irregularities in the rhythm, rate, and sequence of heart contractions. It can arise from disorders within the heart or disruptions in the functionality of other bodily organs and systems. Heart irregularities pose significant risks, potentially leading to severe health complications or even patient fatality. Hence, initiating treatment for arrhythmia promptly upon the onset of initial symptoms is crucial for mitigating potential harm. In this section, the use of the AlexNet model for the classification of four classes of dangerous arrhythmias is discussed, employing a database of ECG signal fragments [1] and transforming 2-second signal fragments into scalograms using CWT [9,11,15,18].

AlexNet, a convolutional neural network, has significantly impacted the advancement of machine learning, particularly in computer vision algorithms. It achieved a remarkable victory in the ImageNet LSVRC-2012 image recognition competition, surpassing competitors by a considerable margin . Its architecture bears resemblance to Yann LeCun's LeNet network, albeit with more filters per layer and nested convolutional layers. AlexNet's structure encompasses convolutions, max pooling, dropout, data augmentation, ReLU activation functions, and stochastic gradient descent.

The architecture of AlexNet consists of eight layers, including five convolutional layers, two fully connected layers, and one softmax output layer. The network has a total of 60 million parameters and was trained on a dataset of 1.2 million images from 1,000 different classes. Below is a concise overview of each layer [19]:

• Convolutional Layer 1 (Conv1):This layer convolves the input image with 96 filters, each having a size of 11×11×3. The stride of 4 indicates that the filter moves 4 pixels at a time horizontally and vertically. ReLU activation is applied to introduce non-linearity, enhancing the network's ability to learn complex patterns.

• Max Pooling Layer 1: Following Conv1, max pooling is applied with a pool size of 3×3 and a stride of 2. This operation reduces the spatial dimensions of the feature maps by a factor of 2, aiding in computational efficiency and creating translation invariance.

• Convolutional Layer 2 (Conv2): Conv2 applies 256 filters of size 5×5×96 to the feature maps produced by the previous layer. The stride of 1 preserves spatial resolution, while a padding of 2 maintains the spatial dimensions of the feature maps. ReLU activation introduces non-linearity.maintains the spatial dimensions of the feature maps. ReLU activation introduces non-linearity.

• Max Pooling Layer 2: Similar to Max Pooling Layer 1, this layer applies max pooling with a 3×3 window and a stride of 2, further reducing the spatial dimensions of the feature maps.

• Convolutional Layer 3 (Conv3): Conv3 consists of 384 filters with a size of 3×3×256. The stride of 1 and padding of 1 ensure that the spatial dimensions of the feature maps remain the same. ReLU activation is applied for introducing non-linearity.

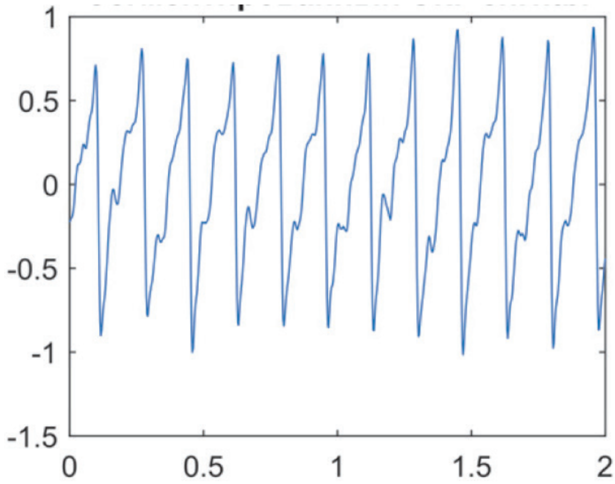
• Convolutional Layer 4 (Conv4): This layer has 384 filters of size 3×3×384. It operates similarly to Conv3, preserving spatial dimensions with a stride of 1 and padding of 1, while ReLU activation introduces non-linearity.

• Convolutional Layer 5 (Conv5): Conv5 applies 256 filters of size 3×3×384. With a stride of 1 and padding of 1, it maintains spatial dimensions. ReLU activation functions are used for non-linearity.

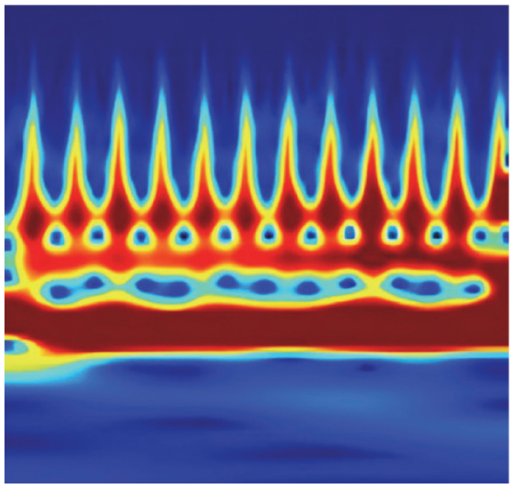
• Fully Connected Layers (FC1 and FC2): FC1 and FC2 are fully connected layers with 4096 neurons each. They take the flattened output of the last convolutional layer as input. ReLU activation functions introduce non-linearity, aiding in learning complex patterns in the data.

• Output Layer (Softmax): The output layer produces class probabilities using the softmax activation function, allowing the model to output probabilities for each of the 1000 ImageNet classes.

Throughout the network, local response normalization (LRN) and dropout are also applied to regularize the model and prevent overfitting during training. These techniques contribute to the overall robustness and generalization ability of the AlexNet architecture.



a)



b)

Figure 5 – a) Segmented ECG signal; b) its scalogram

A pivotal innovation in AlexNet is its adoption of rectified linear unit (ReLU) functions as activation functions. ReLU, a widely employed nonlinear activation function in deep learning, maps input values to a range between 0 and positive infinity. If the input value is zero or negative, ReLU outputs zero; otherwise, it outputs the original input value.

The impressive performance of AlexNet in the ImageNet competition underscored the promise of deep learning in image recognition endeavors. Comprising eight layers—five convolutional and three fully connected - AlexNet is celebrated for its depth and utilization of rectified linear units (ReLU) as activation functions, a key factor contributing to its success. These layers autonomously and adaptively acquire spatial feature hierarchies based on input images. For instance, the initial layer captures basic edge features, while subsequent layers progressively discern more intricate patterns using features from prior layers. Pooling layers interspersed amid convolutional layers decrease spatial dimensions, thus reducing computational demands and imparting a degree of translation invariance. The final three layers in AlexNet are fully connected, meaning each neuron in one layer is linked to every neuron in the subsequent layer. The ultimate layer comprises 1,000 blocks corresponding to the 1,000 classes in the ImageNet dataset, facilitating the derivation of the probability distribution for these classes.

5 Experimental results

In this study, we introduced a deep convolutional neural network architecture to classify a dataset containing four categories of life-threatening arrhythmias. There were a total of 4000 samples, each category included 1000 samples. The dataset was split into training, validation, and test sets following an 80:10:10 ratio. Specifically, 2,900 samples were designated for training, 550 for validation, and 550 for testing. The model was tested on an independent dataset comprising 550 samples, with equal representation of each class. We achieved an overall average classification accuracy of 98.2% on this test set. This performance significantly surpasses previous estimates, which reported a maximum accuracy of 93.18% using a one-dimensional convolutional neural network model on the same ECG classes [12], and 94.12% using a long short-term memory neural network on the same ECG dataset [8]."

The classifier's performance is evaluated using various metrics such as accuracy, sensitivity, specificity, precision, and F1 Score. These metrics can be calculated as follows:

$$\begin{gathered} Accuracy=\frac{TP+TN}{TP+FP+TN+FN} \\ Sensitivity=\frac{TP}{TP+FN} \\ Specificity=\frac{TN}{TN+FP} \\ Precision=\frac{TP}{FP+TP} \\ F1Score=2*\frac{Precision*Sensitivity}{Precision+Sensiticity} \end{gathered} \tag{2}$$

Our model demonstrated exceptional average precision, recall, and F1-score, resulting in an impressive average accuracy of 98.7% on the testing dataset, underscoring its capability to effectively classify the data. The remarkable accuracy of our model can be attributed to the utilization of sophisticated machine learning methodologies like neural networks and deep learning. Figure 6 displays the error matrix for the test set and table 2 shows the results of the proposed method.

Table 2 – Results of the proposed method

Accuracy	Sensitivity	Specificity	Precision	F1score
98.2%	98.2%	98.2%	98.2%	98.0%

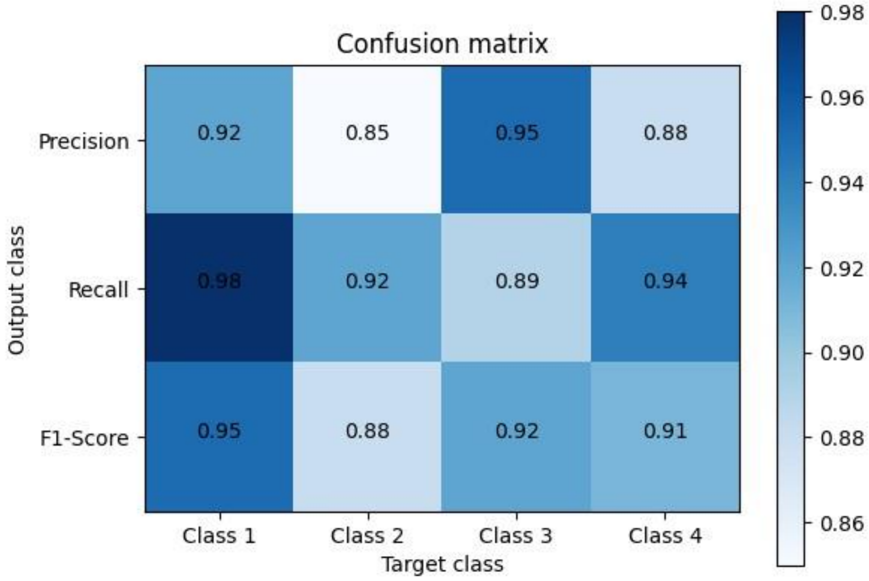


Figure 6 – Confusion matrix for the test set

6 Conclusion

This study has achieved a significant breakthrough in accuracy when classifying dangerous cardiac arrhythmias based on ECG data. Our unique method, which is centered around recognizing four classes of arrhythmias in 2-second ECG segments, combines continuous wavelet transformation with the application of the AlexNet neural network using transfer learning. The research results have demonstrated outstanding average accuracy of 98.2% on the test set, notably surpassing previous maximum values of 93.18% reported by other researchers for similar ECG classes. The innovative approach of representing one-dimensional signals as two-dimensional images introduces substantial novelty to the obtained results. This approach not only holds practical significance for enhancing cardiac monitoring systems but also provides scientific support for the effectiveness of continuous wavelet transformation in the automatic recognition of one-dimensional signals. The obtained results can be applied for the further development of neural network architectures in ECG classification systems, enhancing their efficiency, and introducing new standards in the field of medical technology.

Acknowledgements

This research was funded by the Science Committee of the Ministry of Science and Higher Education of the Republic of Kazakhstan, grant number BR19675574.

References

1. Nemirko A., Manilo L., Tatarinova A., Alekseev B., Evdakova E. ECG Fragment Database for the Exploration of Dangerous Arrhythmia (version 1.0.0). PhysioNet 2022.

2. Pandit D., Li Z., Chengyu L., Chattopadhyay S., Aslam N., Peng LC. A lightweight QRS detector for single lead ECG signals using a max-min difference algorithm. Comput Methods Programs Biomed 2017; 144: 61-75. DOI:10.1016/j.cmpb.2017.02.028.

3. Acharya UR., Fujita H., Oh SL., Hagiwara Y., Tan JH., Adam M., Tan RS. Deep convolutional neural network for the automated diagnosis of congestive heart failure using ECG signals. Appl Intell 2019; 49(1): 16-27. DOI:10.1007/s10489-018-1179-1.

4. Haberl R., Jilge G., Pulter R., Steinbeck G. Spectral mapping of the electrocardiogram with Fourier transform for identification of patients with sustained ventricular tachycardia and coronary artery disease. Eur Heart J 1989;10(4): 316-322.

5. Tuncer T., Dogan S., Pławiak P., Acharya UR. Automated arrhythmia detection using novel hexadecimal local pattern and multilevel wavelet transform with ECG signals. Knowledge-Based Systems 2019; 186: 104923. DOI: 10.1016/j.knosys.2019.104923.

6. Madeiro JPV., Santos EMB., Cortez PC., Felix JHS., Schlindwein FS. Evaluating Gaussian and Rayleigh-based mathematical models for T and P-waves in ECG. IEEE Lat Am Trans 2017; 15(5): 843-853. DOI:10.1109/TLA.2017.7910197.

7. Kennedy A., Finlay DD., Guldenring D., Bond RR., Moran K., McLaughlin J. Automated detection of atrial fibrillation using RR intervals and multivariate-based classification. J Electrocardiol 2016; 49(6): 871-876. DOI: 10.1016/j.jelectrocard.2016.07.033.

8. Nemirko AP., Popadina AO. Recognition of lifethreatening arrhythmias by the spectral characteristics of a short electrocardiogram recording [In Russian]. Biotechnosfera 2022; 67(1): 3-8. DOI: 10.25960/bts.2022.1.3.

9. Olanrewaju R., Ibrahim S., Asnawi AL., Altaf H. Classification of ECG signals for detection of arrhythmia and congestive heart failure based on continuous wavelet transform and deep neural networks. Indones J Electr Eng Comput Sci 2021; 22: 1520. SOI:10.11591/ijeecs.v22.i3., pp1520-1528.

10. Acharya UR., Fujita H., Oh SL., Hagiwara Y., Tan JH., Adam M., Tan RS. Deep convolutional neural network for the automated diagnosis of congestive heart failure using ECG signals. Appl Intell 2019; 49(1): 16-27. DOI:10.1007/s10489-018-1179-1.

11. Byeon YH., Pan SB., Kwak KC. Intelligent deep models based on scalograms of electrocardiogram signals for biometrics. Sensors 2019; 19: 935. DOI: 10.3390/s19040935.

12. Acharya UR., Fujita H., Oh SL., Raghavendra U., Tan JH., Adam M., Gertych A., Hagiwara Y. Automated identification of shockable and non-shockable lifethreatening ventricular arrhythmias using convolutional neural network. Future Generation Computer Systems 2018; 79(3): 952-959. DOI: 10.1016/j.future.2017.08.039.

13. Gupta V., Mittal M. QRS complex detection using STFT, chaos analysis, and PCA in standard and real-time ECG databases. Journal of the Institution of Engineers (India): Series B 2019; 100(5): 489-497. DOI: 10.1007/s40031-019-00398-9.

14. Wang T., Lu C., Sun Y., Yang M., Liu C., Ou C. Automatic ECG classification using continuous wavelet transform and convolutional neural network. Entropy (Basel) 2021; 23(1): 119. DOI: 10.3390/e23010119.

15. Ba Mahel AS., Harold N., Solieman H. Arrhythmia classification using Alexnet model based on orthogonal leads and different time segments. 2022 Conf of Russian Young Researchers in Electrical and Electronic Engineering (ElConRus) 2022:1312-1315. DOI: 10.1109/ElConRus54750.2022.9755708.

16. Chawla NV., Bowyer KW., Hall LO., Kegelmeyer WP. smote: synthetic minority over-sampling technique. J Artif Intell Res 2002; 16: 321-357.

17. Ozaltin O., Yeniay O. A novel proposed CNN–SVM architecture for ECG scalograms classification. Soft Comput 2023; 27: 4639-4658. DOI: 10.1007/s00500-022-07729-x.

18. Li C., Zheng C., Tai C. Detection of ECG characteristic points using wavelet transforms. IEEE Trans Biomed Eng 1995; 42(1): 21-28.

19. Krizhevsky A., Sutskever I., Hinton GE. ImageNet classification with deep convolutional neural networks. Commun ACM 2012; 60(6): 84-90. DOI: 10.1145/3065386.

Information about authors:

Orken Mamyrbayev – PhD, Associate professor, Institute of Information and Computational Technologies (Almaty, Kazakhstan, e-mail: morkenj@mail.ru).

Dina Oralbekova – PhD, Institute of Information and Computational Technologies (Almaty, Kazakhstan, e-mail: dinaoral@mail.ru).

Sholpan Zhumagulova (corresponding author) – Institute of Information and Computational Technologies (Almaty, Kazakhstan, e-mail: sh.zhumagulovakz@gmail.com).

Ernat Azanbekov – PhD, Academy of the Border Guard Service of the National Security Committee of the Republic of Kazakhstan (Almaty, Kazakhstan, e-mail: ernatazanbekov@mail.ru).

Submission received: 12 March, 2024.
Revised: 18 March, 2024.
Accepted: 19 March, 2024.