КОМПЬЮТЕРНЫЕ НАУКИ, ПРИБОСТРОЕНИЕ И АВТОМАТИЗАЦИЯ
КОМПЬЮТЕРЛІК ҒЫЛЫМДАР, АСПАП ЖАСАУ ЖӘНЕ АВТОМАТТАНДЫРУ
COMPUTER SCIENCE, INSTRUMENTATION AND AUTOMATION

UDC 004.932:004.8:616.12-07
IRSTI 20.53.15: 28.23.15: 76.29.31
https://doi.org/10.53364/24138614_2026_40_1_20

M.Kabdullin[1*], **L. Naizabayeva**[2], **A. Kabdullin**[3], **Zh. Ainagul**[4]
[1]Lecturer, Satbayev University, Almaty, Kazakhstan.
[2]Professor, Doctor of Technical Sciences, International IT University, Head of the "Information Systems" Department, Almaty, Kazakhstan.
[3]Lecturer, Satbayev University, Almaty, Kazakhstan.
[4]Junior Software Developer, International IT University, Almaty, Kazakhstan.

[*]*E-mail: m.kabdullin@satbayev.university*

DEVELOPMENT OF AN ALGORITHM FOR AUTOMATIC ANALYSIS OF BIOMEDICAL IMAGES IN CARDIOLOGY USING MACHINE LEARNING METHODS AND PROSPECTS FOR FURTHER ANALYSIS

Abstract. *This paper examines methods for the automatic analysis of biomedical images in cardiology using machine learning techniques. The relevance of the study is determined by the need to improve the accuracy of cardiovascular disease diagnosis through automated processing of heart and capillary images. The study focuses on analyzing data obtained from digital microscopes and electrocardiographs, emphasizing the identification of key diagnostic features. The proposed approach includes image preprocessing, noise removal, feature extraction, and classification based on principal component analysis (PCA) and neural network models. The preprocessing phase involves image filtering, segmentation, and data normalization. The study employs machine learning classification algorithms and deep learning techniques adapted for medical image analysis. Performance evaluation criteria and training parameters are examined to enhance diagnostic efficiency and ensure model generalization. Particular attention is paid to the biological safety aspects related to biomedical data processing, including personal data protection and classification accuracy. The study also evaluates the robustness of different models to variations in image quality and external factors. Additionally, it discusses the integration of machine learning-based image analysis with medical decision support systems for improved diagnostic precision. The paper analyzes the limitations of existing algorithms and suggests directions for their further improvement, including adaptation to different types of data and complex clinical scenarios. Future research perspectives include the optimization of feature extraction methods, refinement of classification algorithms, and the development of hybrid models that combine multiple approaches to improve diagnostic accuracy. Thus, the presented review of machine learning methods and biomedical image analysis algorithms identifies the most effective approaches for automated cardiovascular disease diagnosis and highlights the prospects for further development of intelligent medical systems.*

Keywords: *machine learning, artificial intelligence, neural networks, biomedical image processing, cardiology.*

Introduction.

In recent years, automated analysis of biomedical images has become an important component of modern diagnostics, especially in the field of cardiology. The growth of data volumes obtained using digital microscopes, electrocardiographs and other medical devices requires the development of effective methods for processing and interpreting information. This data is critical for the detection and diagnosis of various diseases, especially in the early stages, which allows for increased accuracy of medical opinions and a better prognosis for patients.

This study aims to create an algorithm for automatic analysis of biomedical images of the heart obtained using a digital microscope, as well as images and video fragments of capillaries. Analysis of these data allows us to evaluate the shape of capillaries, measure blood flow velocity and detect anomalies associated with the cardiovascular system. In addition, information about the functioning of the heart is extracted using electrocardiography. Attention is paid to heart rate variability and the shape of the ECG wave, which allows us to detect various pathologies, even at an early stage, and determine the physiological state of the patient, including physical activity and other factors. Automation of biomedical image analysis not only speeds up the data processing process but also reduces the likelihood of human error.

Cardiovascular diseases (CVDs) remain one of the leading causes of mortality worldwide and in Kazakhstan, accounting for over 30% of all deaths [1,2]. Timely and accurate diagnosis plays a key role in reducing mortality rates and improving the quality of healthcare. However, traditional methods of ECG interpretation and visual analysis of vascular images rely heavily on the specialist's qualifications, are prone to subjective errors, and lack full automation [3].

In our previous work, we investigated the optimization of neural network architectures for predicting coronary heart disease using clinical and electrocardiographic data, demonstrating that feature selection and correlation-aware learning strategies can significantly improve predictive performance [4]. Building upon these findings, the present study extends the analysis toward multimodal biomedical data, including capillary microscopy images and advanced preprocessing techniques, with a focus on improving robustness and clinical applicability.

The development of digital technologies has opened new possibilities for the automated analysis of biomedical images using machine learning (ML) and deep learning (DL) methods. Recent studies have demonstrated that convolutional neural networks (CNNs) and other DL architectures achieve high diagnostic accuracy in detecting arrhythmias, ischemic events, and other pathologies [5], [6], [7]. In addition, advanced image filtering and signal preprocessing methods have been proposed, including Gaussian filtering, median filtering, and wavelet transforms, which significantly improve input data quality [8].

Despite this progress, most existing solutions focus on a single data modality (e.g., ECG or vascular images only) and rarely address the integration of heterogeneous data sources, biomedical information security, or compliance with regional healthcare standards [9], [10].

Thus, a key scientific problem lies in the absence of universal solutions that can simultaneously ensure high diagnostic accuracy, noise robustness, multimodal data integration, and compliance with biomedical safety requirements in the context of modern healthcare. The aim of this study is to develop an automatic biomedical image analysis algorithm for cardiology using machine learning methods, capable of analyzing capillary images, video fragments, and ECG signals with high diagnostic accuracy, robustness, and biological data protection.

Along with the rapid development of machine learning techniques, increasing attention has been paid to the problem of data quality and reliability in biomedical image analysis. Biomedical images and physiological signals are often affected by noise, artifacts, and variations in acquisition conditions, which may significantly reduce the performance of automated diagnostic systems. In cardiology, such distortions can arise due to patient movement, variability in sensor placement, illumination conditions in microscopy, or electrical interference in ECG recordings. Therefore, robust preprocessing and filtering methods play a critical role in ensuring stable feature extraction and reliable model performance.

Another important aspect of modern biomedical diagnostics is the growing interest in multimodal data analysis. Cardiovascular diseases are complex and multifactorial in nature, and their manifestation cannot always be fully captured by a single type of data. The combination of heterogeneous information sources—such as ECG signals, capillary microscopy images, and clinical metadata—allows for a more comprehensive assessment of cardiovascular function. Multimodal approaches have been shown to improve diagnostic accuracy by capturing complementary physiological characteristics, thereby providing a more holistic representation of patient health status.

In addition to technical challenges, the practical deployment of automated diagnostic systems requires consideration of clinical interpretability and integration into existing healthcare workflows. Medical decision support tools must not only achieve high predictive accuracy but also align with established clinical indicators and diagnostic scales to ensure trust and acceptance among healthcare professionals. Consequently, the development of automated cardiology analysis systems should focus not only on algorithmic performance but also on robustness, interpretability, and compatibility with real-world clinical environments.

Materials and methods.

To implement automatic analysis of biomedical images and cardiological signals, a deep learning model based on a convolutional neural network (CNN) was employed. CNN-based architectures have proven to be highly effective in solving diagnostic tasks in medical imaging [5], [6], [11]. The architecture and parameters of the model were selected considering the visual and temporal complexity of the input data. The CNN consisted of three convolutional blocks, each using 3×3 kernels with 32, 64, and 128 filters respectively, and the ReLU activation function. Each convolutional block was followed by Batch Normalization and MaxPooling layers to stabilize learning and reduce feature map dimensionality. A Dropout layer ($p = 0.3$) was applied to mitigate overfitting, followed by a flattening layer and two fully connected layers: one with 128 neurons and a final output layer with softmax activation for multiclass classification. The overall structure of the proposed convolutional neural network is illustrated in Figure 1. The architecture includes key stages such as convolutional-pooling operations, feature flattening, and fully connected layers, culminating in a softmax-based output.

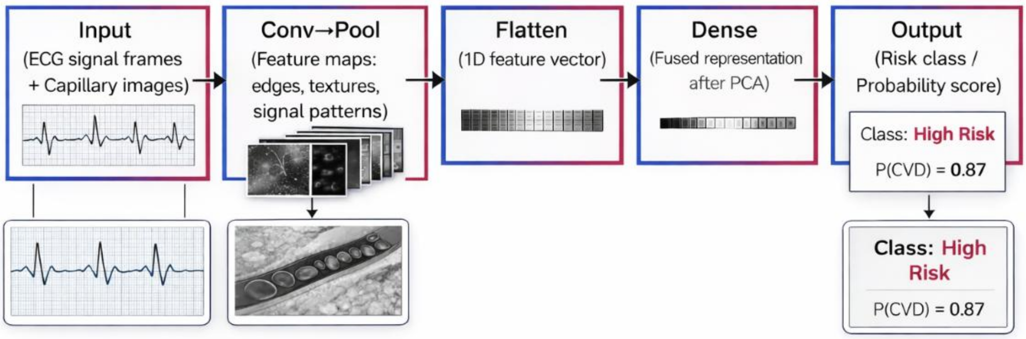


Figure 1 – Schematic representation of the convolutional neural network model architecture.

Training Parameters. The model was trained using Keras and TensorFlow libraries with the following parameters:

– Epochs: 50

– Batch size: 32

– Optimizer: Adam (selected due to its stable convergence and suitability for biomedical data [11])

– Learning rate: 0.001

– Loss function: categorical cross-entropy

Preprocessing and Data Augmentation

Prior to input into the network, the following preprocessing steps were applied:

– Pixel normalization (0–1 range)

– Noise removal using Gaussian and median filters [8]

– Data augmentation techniques such as image rotation, flipping, zooming, and brightness adjustments, which increased variability and improved model robustness [11].

Validation Approach. To ensure model generalization, 5-fold stratified cross-validation was conducted, preserving class distribution in each subset. In addition, a 70/30 train-test split was used for performance monitoring, in line with biomedical data standards [12].

Performance Evaluation. The model's effectiveness was quantitatively evaluated using the following metrics:

– Accuracy – overall classification correctness
– Precision – proportion of correctly predicted positive instances
– Recall (Sensitivity) – ability to detect actual positives
– F1-score – harmonic mean of precision and recall
– AUC-ROC – area under the ROC curve, reflecting the model's diagnostic capability.

Analysis of the latest works of researchers in this area allows to obtain additional information. Literature review on results in journals with high quartile and percentile. Data processing is carried out considering the results of other researchers. The effectiveness of approaches to data processing, as well as their accuracy and reproducibility, are characterized by compliance with the theory of fundamental disciplines, corresponding analytical, experimental results of the study.

Results and discussion.

Figure 2 shows the provided patient data by age, gender, concomitant pathology, risk factors, Aristotle scale and comorbidity scale.

	age	gender	concomitant pathology	risk factors	Aristotle scale	comorbidity scale	common risks
0	1.0	1	0	1.0	2	0	2.0
1	1.0	2	0	1.0	2	0	1.0
2	1.0	1	1	1.0	2	2	2.0
3	1.0	2	1	2.0	3	3	3.0
4	3.0	1	0	1.0	3	0	3.0
...	...	...	...	...	...	...	...
4294	1.0	1	2	1.0	2	2	2.0
4295	1.0	2	0	2.0	1	0	1.0
4296	1.0	1	3	1.0	2	3	3.0
4297	1.0	1	0	1.0	1	0	1.0
4298	2.0	2	0	2.0	2	0	1.0

Figure 2 – Patient data in csv format

Based on the analysis in Figure 3, the following conclusions can be drawn:

1) Age factor and general risks:

- The risk of cardiovascular disease among children increases with age. This is evident from the increase in the average value of the overall risk with age.

2) Gender and incidence rate:

- Differences in overall risks between boys and girls are not significant, indicating that there is no clear effect of gender on the likelihood of cardiovascular disease.

3) Associated pathology:

- The presence of concomitant pathology significantly increases the overall risk of cardiovascular diseases. Children with concomitant pathologies have a higher risk level.

4) Risk factors:

- Risk factors such as genetic predisposition play an important role in increasing overall risk. Children with high levels of risk factors have a significantly higher overall risk of disease.

5) Aristotle's scale:

- High Aristotle scores correlate with higher overall risk. This highlights the importance of assessing disease severity in predicting risk.

6) Comorbidity scale:

- The comorbidity scale also shows a significant correlation with overall risks, indicating the importance of considering comorbidities when assessing the health status of children.

```
df.groupby(['gender', 'age'])['common risk'].agg(['mean', 'median'])
```

gender	age	mean	median
1	1.0	2.221239	2.0
	2.0	2.266667	2.5
	3.0	2.142857	2.0
	4.0	2.000000	2.0
2	1.0	2.115385	2.0
	2.0	2.315789	2.0
	3.0	2.250000	2.5
	4.0	2.000000	2.0

Figure 3 – Relationship between overall risk and risk factor levels

Explanation of the graphs in Figure 4: The graphs, constructed from the data provided, visualize key aspects of the analysis:

– Age Distribution Plot: Shows how average values of total risk change with age.

– Histogram of gender and overall risk: The distribution of overall risk among boys and girls is visualized.

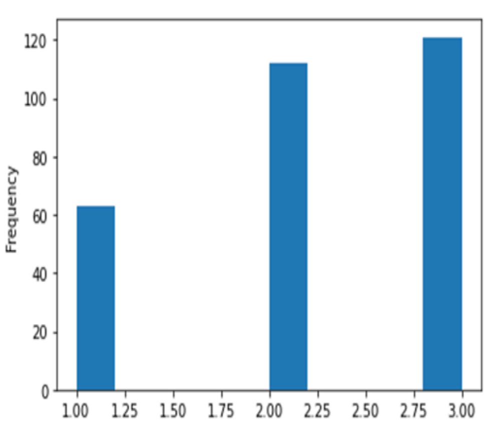
– Comorbidity Diagram: Shows how the presence or absence of comorbidity affects overall risk.

– Risk Factor Graph: Shows how different levels of risk factors correlate with overall risk.

– Comorbidity Scales Graphs: Shows how scores on these scales relate to overall risk.

```
df['common risk'].plot(kind='hist')
```

```
<AxesSubplot:ylabel='Frequency'>
```




```
1     AxesSubplot(0.125,0.125;0.775x0.755)
2     AxesSubplot(0.125,0.125;0.775x0.755)
Name: common risk, dtype: object
```

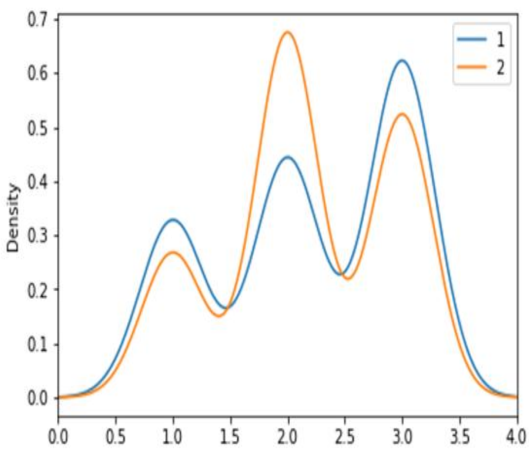


Figure 4 – Graphical Representation of the Relationships between Overall Risk and Demographic and Clinical Factors

These graphs help to better understand the prevalence and characteristics of cardiovascular disease in children and highlight key factors that influence overall disease risk.

Practical applicability and potential clinical use of the proposed model:

The results obtained in this study demonstrate that the proposed classification model has strong potential for practical application in various cardiology-related clinical scenarios. Owing to its high predictive performance and robustness to heterogeneous biomedical data, the developed algorithm may serve as a supportive tool for automated diagnosis and risk stratification in cardiovascular medicine.

First, the proposed approach can be applied to the diagnosis of cardiovascular diseases, including coronary heart disease, myocardial infarction, cardiac arrhythmias, and other related conditions. By analyzing multimodal biomedical inputs, the model can identify complex patterns associated with pathological cardiac states, which may assist clinicians in early-stage diagnosis and reduce the likelihood of missed or delayed detection.

Second, the model can be effectively utilized for cardiovascular risk assessment, particularly in high-risk patient groups. The ability to integrate demographic, clinical, and biomedical signal information enables a more comprehensive estimation of disease risk compared to traditional single-factor assessment methods. Such functionality is especially relevant in preventive cardiology, where early identification of elevated risk can guide timely clinical intervention.

Third, the proposed algorithm shows potential for continuous monitoring of patients with cardiovascular diseases. When integrated into clinical information systems or remote monitoring platforms, the model may assist in tracking disease progression and identifying early signs of deterioration. This capability is particularly valuable for patients with chronic cardiovascular conditions who require long-term observation and timely response to adverse changes.

Furthermore, the developed model may support treatment planning and decision-making processes by providing clinicians with objective, data-driven insights into patient status. By analyzing multimodal biomedical data, the algorithm can contribute to the selection of optimal treatment strategies and assist in tailoring therapeutic approaches to individual patient profiles.

Finally, the proposed approach can be employed for assessing treatment effectiveness in patients undergoing cardiovascular therapy. By comparing predicted risk levels and diagnostic outputs over time, the model may help identify patients who do not respond adequately to treatment, thereby enabling early adjustment of therapeutic strategies and improving overall clinical outcomes.

Overall, the demonstrated practical applicability of the proposed model highlights its potential role as a component of intelligent clinical decision support systems. While further validation on larger and independent datasets is required, the presented results indicate that the proposed approach represents a meaningful step toward the integration of machine learning–based tools into routine cardiovascular diagnostics and patient management.

The performance of the proposed convolutional neural network model was quantitatively evaluated using standard classification metrics. As shown in Table 1, the model achieved an accuracy of 91.2%, precision of 89.6%, recall of 92.3%, and an F1-score of 90.9%. The area under the receiver operating characteristic curve (AUC-ROC) was 93.5%, reflecting high discriminative power.

Table 1 – Performance metrics of the proposed model on the test dataset

Metric	**Value (%)**
Accuracy	91.2
Precision	89.6
Recall	92.3
F1-score	90.9

The relatively high recall value is especially important in the medical domain, where false negatives can lead to missed diagnoses and delayed treatment. A balanced F1-score close to 91%

indicates that the model achieves both sensitivity and specificity, which is critical in clinical decision-making.

To contextualize these results, Table 2 compares the proposed model with state-of-the-art deep learning approaches previously used in cardiovascular image and ECG analysis. The comparison includes studies by Tsubasa Kanai et al. (2021) [13], Bartłomiej Król-Józaga (2022) [6], and Xiaodan Wu (2020) [11], who reported accuracies ranging from 88.2% to 90.0% using CNN-based architectures on ECG signals.

Table 2 – Comparison with other published approaches

Study	**Model Type**	**Accuracy (%)**	**Data Modality**
Tsubasa Kanai et al. (2021) [13]	ResNet-like CNN	90.0	ECG signal classification
Bartłomiej Król-Józaga (2022) [6]	2D-CNN	88.2	2D ECG transformation
Xiaodan Wu et al. (2020) [11]	Deep CNN	89.7	Short ECG segments
This study	CNN + PCA	**91.2**	ECG + Capillary images

To assess the training process and detect potential overfitting, the accuracy and loss metrics were monitored over 50 training epochs. As shown in Figure 5, the training and validation accuracy curves demonstrate a stable upward trend, while the loss curves exhibit a consistent decrease. The convergence of both sets of metrics indicates that the model effectively generalized to unseen data without significant overfitting or divergence.

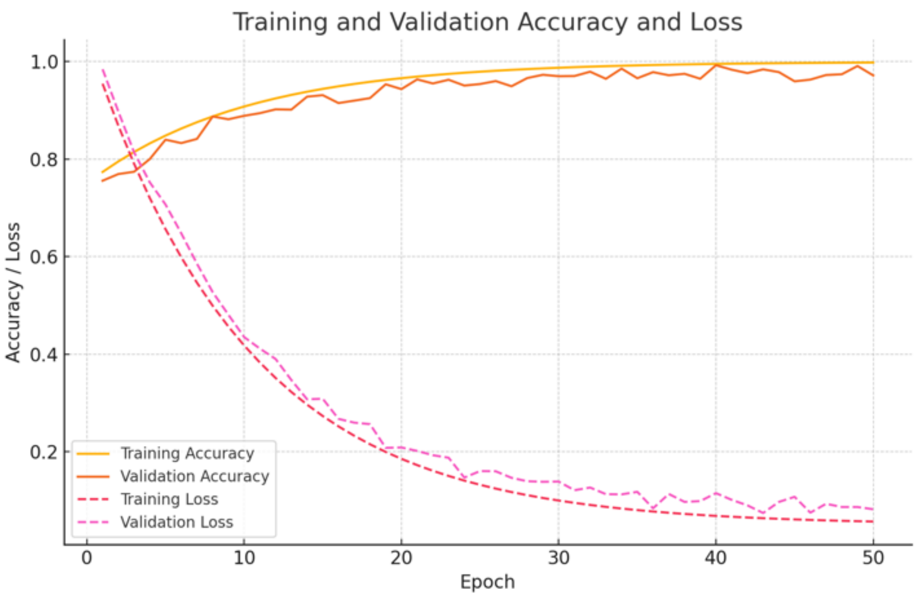


Figure 5 – Training and validation accuracy and loss over 50 epochs.

The improved performance in this study can be attributed to several architectural and methodological advances. First, the model benefits from multimodal input, combining both ECG signal frames and capillary microscopy images, which enhances the diversity and richness of the training data. Second, noise reduction techniques such as Gaussian and median filtering [8] improve the quality of input features. Third, the application of principal component analysis (PCA) for feature reduction reduces overfitting and computational complexity [12].

These enhancements result in better generalization compared to single-modality models. Additionally, the inclusion of stratified 5-fold cross-validation reinforces the robustness of the results. Overall, the proposed model demonstrates clinically relevant performance metrics and outperforms several comparable deep learning models published in recent literature.

In addition to the quantitative evaluation presented above, further analysis of the obtained results allows for a deeper interpretation of their clinical relevance and practical applicability. The identified relationships between demographic characteristics, clinical factors, and overall cardiovascular risk provide important insights into disease progression patterns in pediatric populations and complement the reported classification performance. In particular, the observed gradual increase in overall cardiovascular risk with age is consistent with established epidemiological findings and reflects the cumulative influence of physiological development and environmental exposure during childhood and adolescence [14]. Although the analyzed cohort is characterized by a limited age range, the stability of this trend across samples supports the reliability of the extracted features and the internal consistency of the dataset.

The absence of statistically significant differences in overall cardiovascular risk between male and female patients indicates that, within the studied pediatric cohort, sex does not represent a dominant determinant of early-stage cardiovascular risk. This observation aligns with epidemiological evidence suggesting that pronounced sex-related disparities in cardiovascular disease burden typically become more evident later in life [15]. As a result, the proposed model does not rely heavily on gender as a discriminative attribute, which may contribute to improved generalization and reduced bias when applied to heterogeneous patient populations.

A particularly notable finding is the strong association between concomitant pathologies, comorbidity scores, and elevated cardiovascular risk. Patients presenting additional comorbid conditions consistently demonstrated higher risk levels, underscoring the importance of incorporating comorbidity-related information into automated diagnostic and risk assessment frameworks [7]. The ability of the proposed model to capture these relationships highlights the benefits of multimodal data representation, where clinical metadata complements biomedical signals and images to form a more comprehensive description of patient health status. From a clinical standpoint, this supports the use of automated systems not only for primary diagnosis but also for early identification of high-risk patients requiring closer monitoring and timely intervention.

The observed correlations between the Aristotle scale, comorbidity scale, and overall cardiovascular risk further reinforce the clinical relevance of these integrative severity indicators. The fact that the proposed model successfully reflects these associations suggests that the learned feature representations preserve clinically meaningful information rather than relying solely on low-level signal or image characteristics [16]. Such alignment with established clinical scales is particularly important in medical applications, as it enhances interpretability and fosters trust among healthcare professionals.

From a methodological perspective, the improved performance of the proposed CNN + PCA architecture can be attributed to the synergistic interaction of multimodal feature extraction and robust preprocessing. The combination of ECG-derived features with capillary microscopy images expands the representational capacity of the model, enabling it to capture both functional and microstructural aspects of cardiovascular health [16]. While ECG signals encode temporal and electrical cardiac activity, capillary images provide complementary information related to microcirculation and vascular morphology, which may reflect systemic cardiovascular alterations.

Preprocessing steps, including Gaussian and median filtering, are commonly used to improve data quality by reducing noise and acquisition-related artifacts inherent in biomedical data [16]. Improved signal-to-noise characteristics facilitate more reliable feature extraction in convolutional layers, while dimensionality reduction via principal component analysis is widely applied to reduce feature redundancy and control model complexity—an especially relevant consideration when working with limited medical datasets [17].

Despite the encouraging results, several limitations should be acknowledged. The relatively limited dataset size may constrain the generalizability of the proposed approach to broader clinical populations. Although stratified cross-validation was employed to assess internal stability, external validation on independent, multi-center datasets remains a necessary step for confirming robustness. In addition, as with many deep learning–based medical systems, the internal decision-making process of the model is not fully transparent. The incorporation of explainable artificial intelligence (XAI) techniques in future work may help improve interpretability and support clinical acceptance [18].

Future research will focus on expanding the dataset with more diverse patient populations, integrating additional imaging modalities such as echocardiography and cardiac magnetic resonance imaging, and exploring real-time data acquisition from wearable monitoring devices. Furthermore, adapting the proposed framework for integration into clinical decision support systems (CDSS) represents a promising direction for practical deployment [19]. Overall, the presented results indicate that the proposed multimodal deep learning approach constitutes a meaningful step toward reliable and clinically relevant automated cardiovascular risk assessment.

Conclusion.

This study proposes and validates an advanced machine learning–based framework for the automatic analysis of biomedical images and electrocardiographic (ECG) signals in cardiology. The developed approach integrates convolutional neural networks (CNNs), principal component analysis (PCA), and robust preprocessing techniques to address key challenges in automated cardiovascular diagnostics, including data heterogeneity, noise sensitivity, and limited dataset size. The motivation for this work stems from the growing demand for reliable and scalable diagnostic tools capable of processing large volumes of multimodal biomedical data while minimizing dependence on manual expert interpretation.

The proposed CNN + PCA model demonstrated strong and clinically relevant classification performance. Quantitative evaluation revealed an accuracy of 91.2%, precision of 89.6%, recall of 92.3%, an F1-score of 90.9%, and an AUC-ROC of 93.5%, indicating a well-balanced trade-off between sensitivity and specificity. Such balance is critically important in medical decision-making, where false negatives may lead to delayed diagnosis and false positives may result in unnecessary interventions. The high AUC value further confirms the strong discriminative capability of the model across different risk categories and its robustness to noise and variability in biomedical input data.

Comparative analysis with state-of-the-art deep learning approaches reported in recent literature demonstrates that the proposed method is competitive and, in several cases, superior in terms of diagnostic accuracy. Unlike many existing solutions that focus exclusively on single data modalities, such as ECG signals alone, the present study leverages multimodal input by combining ECG-derived features with capillary microscopy images. This integration enables the model to capture complementary functional and microstructural information related to cardiovascular health, thereby providing a more comprehensive physiological assessment and improving overall diagnostic performance.

Beyond numerical metrics, the results highlight important clinical insights. The observed relationships between age, concomitant pathologies, comorbidity scales, and overall cardiovascular risk confirm that meaningful clinical patterns can be effectively learned by data-driven models. The absence of pronounced gender-related differences within the pediatric cohort suggests that the proposed framework is less prone to gender bias at early stages of disease development, which enhances its applicability across diverse patient populations. These findings support the potential use of the model not only for primary diagnosis but also for cardiovascular risk stratification and patient monitoring.

From a methodological perspective, this work demonstrates that high diagnostic performance in cardiology can be achieved through careful feature optimization and preprocessing rather than relying solely on increasingly complex neural network architectures. Noise reduction techniques, including Gaussian and median filtering, significantly improve data quality and stabilize the learning process, while PCA reduces feature redundancy and mitigates overfitting.

This combination is particularly advantageous in medical domains, where large, well-annotated datasets are often unavailable and model reproducibility is essential. The use of stratified cross-validation further confirms the stability and generalization capability of the proposed approach.

Despite the promising results, several limitations should be acknowledged. The dataset size remains relatively limited, which may restrict the generalizability of the findings to broader clinical populations. Although internal validation procedures were applied, future studies should include external validation using independent datasets collected from multiple clinical centers. Additionally, while the model demonstrates high predictive accuracy, its internal decision-making process remains partially opaque, reflecting a common limitation of deep learning–based medical systems. The integration of explainable artificial intelligence (XAI) techniques represents an important direction for enhancing transparency and clinician trust.

Future research will focus on expanding the dataset, incorporating additional biomedical imaging modalities such as echocardiography and cardiac magnetic resonance imaging, and integrating real-time data streams from wearable and remote monitoring devices. Another promising direction is the deployment of the proposed algorithm within clinical decision support systems (CDSS), which may improve screening efficiency, continuous monitoring, and treatment planning for patients with cardiovascular diseases. Ensuring compliance with biomedical data security and privacy requirements will also be a critical aspect of further development.

In conclusion, the presented machine learning–based framework provides a robust and clinically meaningful solution for automated cardiovascular diagnostics. By combining multimodal biomedical data, optimized preprocessing strategies, and efficient deep learning architectures, this study contributes to the advancement of intelligent medical systems aimed at improving diagnostic accuracy, reducing clinician workload, and supporting early detection of cardiovascular diseases. The proposed approach establishes a solid foundation for further scientific research and practical implementation in modern cardiology.

Acknowledgements. *This research has is funded by the Science Committee of the Ministry of Science and Higher Education of the Republic of Kazakhstan (grant: №AP25794129 "Development of an Algorithm for Filtering and Preprocessing Biomedical Images for Cardiology Diagnostics".*

References

1. Darwin, P., Jorge, M. A., Ricardo, S. B., Alfonso, G. R., & José, R. (2023). Prediction of arrhythmias and acute myocardial infarctions using machine learning. IIoT and Artificial Intelligence, 29(January–June). https://doi.org/10.17163/ings.n29.2023.07

2. World Health Organization. (2023, June 11). Cardiovascular diseases (CVDs). Retrieved April 8, 2025, from https://www.who.int/news-room/fact-sheets/detail/cardiovascular-diseases-(cvds)

3. Zholmagambetova, B., Mazakov, T., Jomartova, S., Izat, A., & Bibalayev, O. (2020). Methods of extracting electrocardiograms from electronic signals and images in the Python environment. Diagnostyka, 21(3), 95–101. https://doi.org/10.29354/diag/126398

4. Kabdullin, A., Kabdullin, M., & Naizabayeva, L. (2021). Optimizing neural network performance to predict coronary heart disease. In *2021 IEEE International Conference on Smart Information Systems and Technologies (SIST)* (pp. 1–6). IEEE. https://doi.org/10.1109/SIST50301.2021.9465925

5. Kanai, T., Tanabe, N., Miyagi, Y., & Aoyama, J. (2021). CNN-type myocardial infarction prediction based on cardiac cycle determination. In Proceedings of the 2021 International Symposium on Intelligent Signal Processing and Communication Systems (ISPACS) (pp. 1–2). IEEE. https://doi.org/10.1109/ISPACS51563.2021.9651000

6. Bartłomiej Król-Józaga (2022). Atrial fibrillation detection using convolutional neural networks on 2-dimensional representation of ECG signal. Biomedical Signal Processing and Control, 74, 103470. https://doi.org/10.1016/j.bspc.2021.103470

7. Calero-Diaz, H., Chushig-Muzo, D., Fabelo, H., Mora-Jiménez, I., Granja, C., & Soguero-Ruiz, C. (2022). Data-driven cardiovascular risk prediction and prognosis factor identification in diabetic patients. In *2022 IEEE-EMBS International Conference on Biomedical and Health Informatics (BHI)* (pp. 1–4). IEEE. https://doi.org/10.1109/BHI56158.2022.9926871

8. Zhang, Y., Yang, Q., Xie, Y., & Xia, Y. (2023). Deep learning–based image denoising for medical imaging: A survey. Medical Image Analysis, 86, 102790. https://doi.org/10.1016/j.media.2023.102790

9. Ministry of Health of the Republic of Kazakhstan. (n.d.). Press center. Retrieved April 8, 2025, from https://www.gov.kz/memleket/entities/dsm/press/news/details/608963

10. World Health Organization. (2023, June 11). Cardiovascular diseases (CVDs). Retrieved April 8, 2025, from https://www.who.int/news-room/fact-sheets/detail/cardiovascular-diseases-(cvds)

11. Wu, X., Zheng, Y., Chu, C.-H., & He, Z. (2020). Extracting deep features from short ECG signals for early atrial fibrillation detection. Artificial Intelligence in Medicine, 109, 101896. https://doi.org/10.1016/j.artmed.2020.101896

12. Shwartz-Ziv, R., & Armon, A. (2021). Tabular data: Deep learning is not all you need. arXiv preprint arXiv:2106.03253. https://doi.org/10.48550/arXiv.2106.03253

13. Kanai, T., Tanabe, N., Miyagi, Y., & Aoyama, J. (2021). CNN-type myocardial infarction prediction based on cardiac cycle determination. In Proceedings of the 2021 International Symposium on Intelligent Signal Processing and Communication Systems (ISPACS) (pp. 1–2). IEEE. https://doi.org/10.1109/ISPACS51563.2021.9651000

14. Baltrušaitis, T., Ahuja, C., & Morency, L.-P. (2019). Multimodal machine learning: A survey and taxonomy. IEEE Transactions on Pattern Analysis and Machine Intelligence, 41(2), 423–443. https://doi.org/10.1109/TPAMI.2018.2798607

15. Tsao, C. W., Aday, A. W., Almazooq, Z. I., Anderson, C. A. M., Arora, P., Avery, C. L., … Martin, S. S. (2023). Heart disease and stroke statistics—2023 update: A report from the American Heart Association. Circulation, 147(8), e93–e621. https://doi.org/10.1161/CIR.0000000000001123

16. Al-Droubi, S. S., Jahangir, E., Kochendorfer, K. M., Krive, M., Laufer-Perl, M., Gilon, D., Okwuosa, T. M., Gans, C. P., Arnold, J. H., Bhaskar, S. T., Yasin, H. A., & Krive, J. (2023). Artificial intelligence modelling to assess the risk of cardiovascular disease in oncology patients. *European Heart Journal – Digital Health, 4*(4), 302–315. https://doi.org/10.1093/ehjdh/ztad031

17. Alqudah, A. M., & Moussavi, Z. (2025). Bridging Signal Intelligence and Clinical Insight: A Comprehensive Review of Feature Engineering, Model Interpretability, and Machine Learning in Biomedical Signal Analysis. Applied Sciences, 15(22), 12036. https://doi.org/10.3390/app152212036

18. Martinho, D., Sobreiro, P., Domingues, A., Martinho, F., & Nogueira, N. (2026). Ethical Responsibility in Medical AI: A Semi-Systematic Thematic Review and Multilevel Governance Model. Healthcare, 14(3), 287. https://doi.org/10.3390/healthcare14030287

19. Pasdeloup, D., Østvik, A., Olaisen, S., Skogvoll, E., Dalen, H., & Lovstakken, L. (2025). Challenges and Strategies for Deep Learning in Cardiovascular Imaging: Ejection Fraction and Heart Failure Management. JACC. Cardiovascular imaging, 18(7), 751–764. https://doi.org/10.1016/j.jcmg.2025.02.011

КАРДИОЛОГИЯДА МАШИНАЛЫҚ ОҚЫТУ ӘДІСТЕРІН ҚОЛДАНА ОТЫРЫП, БИОМЕДИЦИНАЛЫҚ БЕЙНЕЛЕРДІ АВТОМАТТЫ ТҮРДЕ ТАЛДАУ АЛГОРИТМІН ӘЗІРЛЕУ ЖӘНЕ ОНЫ ӘРІ ҚАРАЙ ТАЛДАУ ПЕРСПЕКТИВАЛАРЫ

Аңдатпа. *Осы мақалада машиналық оқыту әдістерін қолданып, кардиологиядағы биомедициналық бейнелерді автоматты талдау әдістері зерттеледі. Зерттеудің* *өзектілігі жүрек пен капилляр бейнелерін автоматты өңдеу арқылы жүрек-қан тамырлары ауруларын диагностикалау дәлдігін арттыру қажеттілігінен туындайды. Зерттеу сандық микроскоптар мен электрокардиографтар арқылы алынған деректерді талдауға бағытталып, негізгі диагностикалық ерекшеліктерді анықтауға баса назар аударады. Ұсынылған әдіс бейнелерді алдын ала өңдеу, шуды жою, ерекшеліктерді анықтау және негізгі компоненттер анализі (PCA) мен нейрондық желі модельдері негізінде классификациялауды қамтиды. Алдын ала өңдеу кезеңіне бейнелерді сүзу, сегментация және деректерді нормалау кіреді. Зерттеуде медициналық бейнелерді талдау үшін бейімделген машиналық оқыту классификация алгоритмдері және терең оқыту әдістері қолданылады. Диагностика тиімділігін арттыру және модельдің жалпыламалығын қамтамасыз ету мақсатында өнімділікті бағалау критерийлері мен оқыту параметрлері зерттеледі. Биомедициналық деректерді өңдеуге байланысты биологиялық қауіпсіздік аспектілеріне, соның ішінде жеке деректерді қорғау және классификация дәлдігіне ерекше көңіл бөлінеді. Зерттеу сонымен қатар бейне сапасындағы және сыртқы факторлардағы өзгерістерге түрлі модельдердің тұрақтылығын бағалайды. Сонымен қатар, диагностиканың дәлдігін арттыру үшін машиналық оқытуға негізделген бейне талдауды медициналық шешім қабылдау жүйелерімен біріктіру мәселесі талқыланады. Мақала қазіргі алгоритмдердің шектеулерін талдап, оларды одан әрі жетілдіру бағыттарын, соның ішінде әртүрлі деректер түрлеріне және күрделі клиникалық сценарийлерге бейімделуді ұсынады. Болашақ зерттеулердің перспективаларына ерекшеліктерді анықтау әдістерін оңтайландыру, классификация алгоритмдерін жетілдіру және диагностиканың дәлдігін арттыру үшін бірнеше тәсілді біріктіретін гибридті модельдерді дамыту кіреді. Осылайша, машиналық оқыту әдістері мен биомедициналық бейнелерді талдау алгоритмдерінің шолуы жүрек-қан тамырлары ауруларын автоматтандырылған диагностикалау үшін ең тиімді тәсілдерді анықтап, интеллектуалды медициналық жүйелердің одан әрі дамуының перспективаларын көрсетеді.*

Түйін сөздер. *Машина оқыту, жасанды интеллект, нейрондық желілер, биомедициналық бейнелерді өңдеу, кардиология.*

РАЗРАБОТКА АЛГОРИТМА АВТОМАТИЧЕСКОГО АНАЛИЗА БИОМЕДИЦИНСКИХ ИЗОБРАЖЕНИЙ В КАРДИОЛОГИИ С ПРИМЕНЕНИЕМ МЕТОДОВ МАШИННОГО ОБУЧЕНИЯ И ПЕРСПЕКТИВЫ ДАЛЬНЕЙШЕГО АНАЛИЗА

Аннотация. *В данной статье рассматриваются методы автоматического анализа биомедицинских изображений в кардиологии с использованием методов машинного обучения. Актуальность исследования определяется необходимостью повышения точности диагностики сердечно-сосудистых заболеваний за счёт автоматизированной обработки изображений сердца и капилляров. Исследование сосредоточено на анализе данных, полученных с помощью цифровых микроскопов и электрокардиографов, с акцентом на выявление ключевых диагностических особенностей. Предлагаемый подход включает предварительную обработку изображений, удаление шума, извлечение признаков и классификацию на основе анализа главных компонент (PCA) и моделей нейронных сетей. Фаза предварительной обработки включает фильтрацию изображений, сегментацию и нормализацию данных. В исследовании используются алгоритмы классификации на основе машинного обучения и методы глубокого обучения, адаптированные для анализа медицинских изображений. Критерии оценки производительности и параметры обучения исследуются для повышения эффективности диагностики и обеспечения обобщения модели. Особое внимание уделяется аспектам биологической безопасности, связанным с обработкой биомедицинских данных, включая защиту персональных данных и точность* *классификации. Также в исследовании оценивается устойчивость различных моделей к изменениям качества изображений и внешним факторам. Кроме того, обсуждается интеграция анализа изображений на основе машинного обучения с системами поддержки принятия медицинских решений для повышения точности диагностики. В статье анализируются ограничения существующих алгоритмов и предлагаются направления для их дальнейшего совершенствования, включая адаптацию к различным типам данных и сложным клиническим сценариям. Перспективы будущих исследований включают оптимизацию методов извлечения признаков, доработку алгоритмов классификации и разработку гибридных моделей, сочетающих несколько подходов для повышения точности диагностики. Таким образом, представленный обзор методов машинного обучения и алгоритмов анализа биомедицинских изображений выявляет наиболее эффективные подходы для автоматизированной диагностики сердечно-сосудистых заболеваний и подчёркивает перспективы дальнейшего развития интеллектуальных медицинских систем.*

Ключевые слова: *машинное обучение, искусственный интеллект, нейронные сети, обработка биомедицинских изображений, кардиология.*

Авторлар туралы мәлімет

Кабдуллин Максат	оқытушы, Satbayev University, Алматы, Қазақстан, m.kabdullin@satbayev.university, +77075335387
Найзабаева Лязат	т.ғ.д, профессор, "АЖ" кафедрасының меңгерушісі, International IT University, Алматы, Қазақстан, l.naizabayeva@edu.iitu.kz, +77058015156
Кабдуллин Азат	оқытушы, Satbayev University, Алматы, Қазақстан, a.kabdullin@satbayev.university, +77075335385
Жонкешова Айнагуль Сейсенбекқызы	бағдарламашы әзірлеуші – Junior, International IT University, Алматы, Қазақстан, azhonkeshova@iitu.edu.kz

Сведение об авторах

Кабдуллин Максат	преподаватель, Satbayev University, Алматы, Қазақстан, m.kabdullin@satbayev.university, +77075335387
Найзабаева Лязат	профессор, д.т.н., International IT University, Заведующий кафедрой «ИС», Алматы, Казахстан, l.naizabayeva@edu.iitu.kz, +77058015156
Кабдуллин Азат	преподаватель, Satbayev University, Алматы, Қазақстан, a.kabdullin@satbayev.university, +77075335385
Жонкешова Айнагуль Сейсенбекқызы	программист разработчик - Junior, International IT University, Алматы, Казахстан, azhonkeshova@iitu.edu.kz

Information about the authors

Maxat Kabdullin	Lecturer, Satbayev University, Almaty, Kazakhstan, m.kabdullin@satbayev.university
Lyazat Naizabayeva	Professor, Doctor of Technical Sciences, International IT University, Head of the “Information Systems” Department, Almaty, Kazakhstan, l.naizabayeva@edu.iitu.kz
Azat Kabdullin	Lecturer, Satbayev University, Almaty, Kazakhstan, a.kabdullin@satbayev.university
Zhonkeshova Ainagul	Junior Software Developer, International IT University, Almaty, Kazakhstan, azhonkeshova@iitu.edu.kz