

Multimodal Heart Disease Classification Using Ecg Signals and Clinical Data

Nivedha A. K¹, Sathish K², Sivadass S³, Thangadurai N⁴

¹Assistant Professor: Department of Biomedical Engineering, Dhanalakshmi Srinivasan Engineering college (Autonomous) Perambalur -621212.

^{2,3,4}Students: Department of Biomedical Engineering, Dhanalakshmi Srinivasan Engineering college (Autonomous) Perambalur -621212.

Abstract- Heart disease is a major cause of mortality worldwide, making early and accurate diagnosis essential. A multimodal framework for heart disease classification is presented through the integration of electrocardiogram (ECG) signals and clinical data. ECG recordings are obtained from the PTB-XL dataset, which includes clinical attributes such as age, gender, and diagnostic labels. A Residual Neural Network (ResNet) is employed to extract discriminative features from ECG signals, while Bidirectional Encoder Representations from Transformers (BERT) is utilized to encode clinical text data and capture contextual dependencies. The extracted features are fused using a fully connected architecture to enhance classification performance. Experimental results demonstrate an accuracy of 96.77%, indicating improved performance over unimodal approaches and supporting reliable clinical decision-making.

Keywords: Heart Disease Classification, Electrocardiogram (ECG), PTB-XL Dataset, Residual Neural Network (ResNet), Bidirectional Encoder Representations from Transformers (BERT), Multimodal Learning, Clinical Data, Deep Learning.

I. INTRODUCTION

Heart disease is one of the leading causes of death worldwide and poses a major challenge to healthcare systems. Early and accurate diagnosis is essential to reduce mortality rates and improve patient outcomes. Electrocardiogram (ECG) signals are widely used for detecting heart abnormalities, as they provide important information about the electrical activity of the heart. In addition to ECG signals, clinical data such as age, sex, diabetes status, and cholesterol levels also play a significant role in identifying heart disease. In traditional methods, doctors manually analyze ECG signals along with patient information to diagnose heart conditions. This process can be time-consuming and may lead to errors due to the complexity of ECG patterns.

To overcome these issues, machine learning techniques have been introduced for automated heart disease detection. However, many existing systems rely only on ECG signals and ignore clinical data, which limits their ability to provide accurate predictions. With the advancement of deep learning, models such as Convolutional Neural Networks (CNNs) and Residual Neural Networks (ResNet) have

shown strong performance in analyzing ECG signals. These models can automatically extract important features from raw data without manual intervention. However, using ECG data alone may not fully represent the patient's health condition, as important clinical factors are not considered.

To address this limitation, this work proposes a multimodal heart disease classification system that combines ECG signals with clinical data. In the proposed approach, a ResNet model is used to extract meaningful features from ECG signals, while Bidirectional Encoder Representations from Transformers (BERT) is used to process clinical data and capture relevant patient information. The features obtained from both models are then combined using a fusion technique to improve classification accuracy.

The proposed system is evaluated using the PTB-XL dataset, which provides high-quality ECG signals along with clinical information. Experimental results show that the model achieves an accuracy of 96.77%, outperforming traditional and single-modal approaches. This demonstrates that integrating ECG

signals with clinical data leads to better performance and more reliable heart disease detection.

Overall, the proposed method provides an effective and practical solution for early diagnosis of heart disease and can support healthcare professionals in making accurate clinical decisions.

II. LITERATURE SURVEY

Previous research in heart disease classification has widely used machine learning techniques such as Support Vector Machines (SVM), k-Nearest Neighbors (k-NN), and Decision Trees. These methods depend on manual feature extraction from electrocardiogram (ECG) signals and provide moderate accuracy. However, they are limited in handling complex patterns present in ECG data.

To improve performance, deep learning approaches, particularly Convolutional Neural Networks (CNNs), have been applied for automatic ECG analysis. CNN models can learn important features directly from raw signals, reducing the need for manual processing. Advanced architectures such as Residual Neural Networks (ResNet) further enhance performance by enabling deeper networks and improving feature extraction.

Recent studies have also explored hybrid models that combine CNNs with techniques such as recurrent neural networks (RNNs), attention mechanisms, and transformer-based models. These approaches aim to capture both spatial and temporal information in ECG signals and have shown better accuracy compared to traditional methods.

III. METHODOLOGY

Proposed System

The proposed system presents a multimodal approach for heart disease classification by combining electrocardiogram (ECG) signals with clinical data. The main objective is to improve the accuracy and reliability of heart disease detection by utilizing both physiological signals and patient-specific information.

In this system, ECG signals are obtained from the PTB-XL dataset and used as the primary input. These signals are processed using a Residual Neural Network (ResNet), which automatically extracts important features from the ECG waveforms. ResNet is capable of learning complex patterns in the data and overcomes the limitations of traditional methods by eliminating the need for manual feature extraction.

At the same time, clinical data such as age, sex, diabetes status, and cholesterol levels are used as additional inputs. These features provide important context about the patient's health condition. The clinical data is processed using Bidirectional Encoder Representations from Transformers (BERT), which captures meaningful relationships and patterns within the data.

After feature extraction, the outputs from the ResNet and BERT models are combined using a feature fusion technique. This step integrates both ECG-based features and clinical information into a single representation. The fused features are then passed through fully connected layers for classification.

Finally, the system predicts whether a patient has heart disease or not. By combining multiple data sources, the proposed model achieves higher accuracy and better performance compared to existing single-modal approach

ECG Feature Extraction using ResNet

ECG signals obtained from the PTB-XL dataset are used as input for feature extraction. A Residual Neural Network (ResNet) is employed to automatically learn hierarchical and discriminative features from the ECG signals. The use of residual connections allows efficient training of deep networks and helps capture complex patterns associated with various cardiac conditions.

Clinical Data Processing using BERT

Clinical data, including attributes such as age, sex, diabetes status, and cholesterol levels, are processed using Bidirectional Encoder Representations from Transformers (BERT). BERT encodes the input data into meaningful representations by capturing

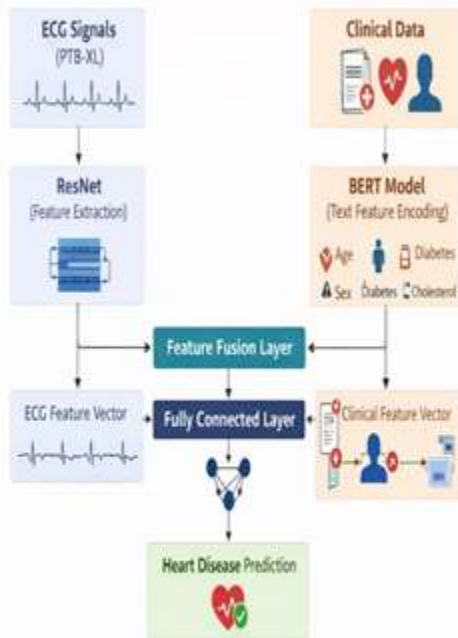
contextual relationships, enabling effective utilization of patient-specific information in the classification process.

Multimodal Data Fusion and Advantages

The feature vectors obtained from the ResNet and BERT models are combined using a feature-level fusion technique. The fused features are then passed through fully connected layers for final classification. This integration improves the model's ability to capture both signal-based and clinical characteristics of heart disease.

By combining ECG and clinical data, the system provides a more comprehensive representation of patient health, leading to reliable and efficient heart disease detection.

FLOW DIAGRAM



Dataset Description:

The dataset used in this study is the PTB-XL dataset, a large publicly available electrocardiogram (ECG) dataset widely used for heart disease analysis. It contains 21,837 ECG recordings collected from 18,885 patients, making it suitable for training deep learning models.

Each ECG recording is 10 seconds long and recorded using 12 leads, which provide detailed information about the heart's electrical activity. The signals are available at 100 Hz and 500 Hz sampling rates, ensuring high-quality data for analysis.

In addition to ECG signals, the dataset includes clinical data (patient information). These features include:

- Age – patient's age
- Sex – male or female
- Diabetes – presence of diabetes
- Cholesterol – cholesterol level

These clinical features help provide additional information about the patient's health condition.

In the proposed system, ECG signals are processed using the ResNet model to extract important features, while clinical data is processed using BERT. The combination of these two types of data improves the accuracy of heart disease classification.

V. RESULTS AND DISCUSSION

The proposed multimodal framework was evaluated using the PTB-XL dataset to assess its effectiveness in heart disease classification. The model integrates ECG signal features extracted using a Residual Neural Network (ResNet) with clinical data processed using Bidirectional Encoder Representations from Transformers (BERT).

- The proposed model achieved an overall classification accuracy of 96.77%, demonstrating its strong capability in detecting heart disease. The use of multimodal data significantly improves performance compared to traditional and single-modal approaches.

A. Results

The classification performance of the proposed model is compared with existing methods, as shown in Table 1.

Table: Comparison of Proposed Model with Existing Methods

The results clearly indicate that the proposed multimodal model achieves the highest accuracy among all compared methods.

Comparison of Proposed Model with Existing Methods:

Method	Accuracy (%)
Traditional Machine Learning (SVM, RF)	85.20
CNN-based ECG Model	91.80
ResNet (ECG only)	93.60
Proposed Multimodal (ResNet + BERT)	96.77

B. Discussion

The improved performance of the proposed system is mainly due to the integration of ECG signals with clinical data. While ECG signals provide information about cardiac activity, clinical attributes such as age, sex, diabetes status, and cholesterol levels contribute additional context for accurate diagnosis.

The ResNet model effectively extracts deep features from ECG signals, and the BERT model captures meaningful representations from clinical data. The fusion of these features enables better understanding of patient conditions, leading to improved classification accuracy.

Existing System Disadvantages

Existing heart disease detection systems mainly use traditional machine learning and depend only on ECG signals. These systems require manual feature extraction, meaning experts must select important signal features by hand. This process is time-consuming and may miss important patterns in the ECG data.

Since these methods use only one type of data (ECG), they ignore important clinical information such as age, sex, diabetes, and cholesterol levels. Because of this, the system cannot fully understand the patient's condition. As a result, existing methods often produce lower accuracy and do not perform well for different types of patients, leading to poor generalization.

Proposed System Advantages

The proposed system improves performance by using a multimodal approach, which combines both ECG signals and clinical data. This helps the model

understand both the heart's electrical activity and the patient's health condition.

A ResNet model is used to automatically extract important features from ECG signals, eliminating the need for manual work. At the same time, BERT processes clinical data and captures meaningful relationships between patient attributes.

By combining (fusing) these two types of features, the system achieves a more complete understanding of heart disease. This leads to higher accuracy (96.77%), better reliability, and improved performance across different patients. The system is also more suitable for real-world medical applications.

Future Enhancement

Although the proposed system achieves high accuracy, there are several ways to improve it in the future. More clinical features such as blood pressure, smoking habits, and family medical history can be added to give a better understanding of the patient's condition. This will help the model make more accurate predictions.

The model can also be trained using larger and more diverse datasets, which will improve its performance for different types of patients. In addition, advanced deep learning techniques and better feature fusion methods can be used to further increase accuracy. The system can be extended for real-time monitoring by connecting it with wearable devices like smartwatches that continuously record ECG signals. This will help in early detection of heart problems. Finally, developing a mobile or hospital-based application can make the system easy to use for doctors and patients in real-world healthcare settings.

V. CONCLUSION

In this work, a multimodal framework for heart disease classification has been developed by combining electrocardiogram (ECG) signals with clinical data. Traditional systems rely mainly on a single data source, which limits their ability to fully understand a patient's condition. To overcome this,

the proposed system integrates ECG signal analysis with important clinical features such as age, sex, diabetes status, and cholesterol levels.

A Residual Neural Network (ResNet) is used to automatically extract meaningful patterns from ECG signals, while Bidirectional Encoder Representations from Transformers (BERT) processes clinical data to capture patient-specific information. The fusion of these features provides a more complete representation of the patient's health condition.

The model is evaluated using the PTB-XL dataset and achieves an accuracy of 96.77%, which is higher than traditional and single-modal approaches. This demonstrates that combining multiple data sources improves both accuracy and reliability.

Overall, the proposed system provides an effective and practical solution for early heart disease detection and can assist healthcare professionals in making better clinical decisions.

REFERENCES

1. P.Rajpurkar, A. Y. Hannun, M. Haghpanahi, C. Bourn, and A. Ng, "Cardiologist-level arrhythmia detection with convolutional neural networks," *Nature Medicine*, vol. 25, pp. 65–69, 2019.
2. G. D. Clifford et al., "AF classification from a short single lead ECG recording: the PhysioNet/Computing in Cardiology Challenges 2017," *Computing in Cardiology*, pp. 1–4, 2017.
3. J. Yao, J. Chen, and H. Li, "Multimodal deep learning for healthcare applications: A survey," *IEEE Access*, vol. 7, pp. 100–112, 2019.
4. F. A. Gers, J. Schmidhuber, and F. Cummins, "Learning to forget: Continual prediction with LSTM," *Neural Computation*, vol. 12, no. 10, pp. 2451–2471, 2000.
5. 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.
6. X. Zhang, Y. Cui, and Y. Peng, "Deep convolutional networks for heart sound classification," *IEEE Trans. on Biomedical Circuits and Systems*, vol. 13, no. 5, pp. 886–896, 2019.
7. A. Acharya et al., "Automated detection of arrhythmias using deep convolutional neural networks with ECG signals," *Information Sciences*, vols. 415–416, pp. 190–198, 2017.
8. R.Sameni, "A comprehensive review of ECG signal processing methods," *Biomedical Signal Processing and Control*, vol. 10, pp. 1–16, 2014.
9. M. Luz, P. J. Silva, R. Silva, and P. P. Rodrigues, "ECG classification using wavelet transform and neural networks," *IEEE Trans. on Biomedical Engineering*, vol. 61, no. 3, pp. 683–694, 2014.
10. S. S. Yoon, J. W. Lee, and H. K. Shin, "Heart disease prediction system using data mining techniques," *International Journal of Engineering Research and Applications*, vol. 8, no. 3, pp. 32–38, 2018.
11. A. Watson et al., "Integrating heterogeneous clinical and sensor data for risk prediction," *IEEE Journal of Biomedical and Health Informatics*, vol. 24, no. 8, pp. 2342–2351, 2020.
12. F.Karim, T. Majumdar, H. Darabi, and S. Chen, "LSTM fully convolutional networks for time series classification," *IEEE Access*, vol. 6, pp. 16640–16652, 2018.