

ECG Beat-By-Beat Classification Using Hybrid Transformer Neural Network Model in Smart Health

I Hua Tsai
Department of Computer Science
Texas Tech University
Lubbock, TX 79409, USA
i-hua.tsai@ttu.edu

Bashir I. Morshed
Department of Computer Science
Texas Tech University
Lubbock, TX 79409, USA
bmorshed@ttu.edu

Abstract—Wearable cardiac monitors can be used to detect potential heart attack by syncing with smartphone apps for instant data analysis and alerts. Our goal is to build an efficient smart health application to help patients prevent and early diagnose the risk of heart disease. Our novel hybrid transformer neural network model can effectively predict the occurrence of heart disease and prevent it early. We used Artificial Neural Network and Convolutional Neural Network and combined them with the transformer model to form a hybrid transformer neural network model, Transformer Artificial Neural Network and Transformer Convolutional Neural Network, respectively. Used three different feature modules, (1) top 10 ranked time series features, (2) noise-removed MIT BIH ECG data, (3) combination module (top 10 ranked time series features and noise-removed MIT BIH ECG data) as input data and sent into the hybrid transformer neural network model to analysis their accuracy and power consumption, respectively. Converted the best-performing hybrid transformer neural network model into a pre-trained model and applied it to our Smart-Health application to verify the correctness and functionality of the algorithm. TCNN model achieved the highest accuracy of 98% and an F1 score of 99%. The hybrid transformer neural network model can be used in our application and detects cardiac diseases more accurately than the neural network model.

Keywords—Transformer model, Deep learning, Smart healthcare, Feature selection, ECG classification

I. INTRODUCTION

Many healthcare wearable devices have been developed to quickly detect physical conditions. Wearable devices often sync with smartphone apps, allowing for instant data analysis and alerts. In today's wearable devices, heartbeat measurement functions are often seen. An electrocardiogram (ECG) is a diagnostic tool that measures and diagnoses the heart and records its electrical activity. ECG can effectively identify various heart diseases such as arrhythmia (AR) and myocardial infarction (MI). Using smart wearable health care devices to detect heart disease can enable early detection and prevention.

Analyzing heart disease through electrocardiogram can help understand the health of the heart and help medical staff provide immediate treatment. The transformer model in natural language processing (NLP) is to extract meaningful features that help predict future architectures [1]. The transformer model analyzes continuous heartbeat signals through electrocardiogram to detect abnormal heartbeats and provide immediate treatment [2-8]. The broad and deep Transformer Neural Network classifies

12-lead ECG sequences into 27 cardiac abnormalities by combining hand-crafted features with automatically learned discriminative features [2]. The transformer network algorithm combines the convolutional neural network to automatically extract high-dimensional ECG features through pre-processing [3]. Transformer networks capture temporal information in CNNs and enhance the classification capabilities of embedding vectors [4].

The multi-module Recurrent Convolutional Neural Network (RCNN) with transformer encoder architecture is an end-to-end deep framework with time series module, spectrogram module, metadata module and semantic fusion module [5]. The embedded layer of the unsupervised transformer and the standard transformer encoder evaluate and detect anomalies in the electrocardiogram (ECG) signal [6]. Use ECG signals to fuse hand-crafted temporal features with learned features through a sequence-to-sequence model and capture the rhythm pattern signal of the ECG through Transformer [7]. The component-aware transformer (CAT) uses ECG waveform length and type information as input to the transformer [8].

This work developed a hybrid model (Transformer Artificial Neural Network (TANN) and Transformer Convolutional Neural Network (TCNN)) to classify ECG signals and apply it to a real-time monitoring system. Figure 1 illustrated step by step the process of using hybrid transformer model of to classify ECG signals and apply them to the real-time monitoring system. Users provided real-time ECG signal data. Pretrained neural network models can be trained on large datasets of labeled ECG recordings to identify different types of heartbeats and then analyzed the real-time ECG signals. The ECG signal is analyzed to detect individual heartbeats, and then divided into five categories. When built the pre-trained model, we used combination module (top 10 ranked time series features and noise-removed MIT BIH ECG signal) as model input data. After the transformer encoder, there was a neural network processing that converted the multidimensional output of the transformer into a one-dimensional array that can be fed into the classifier. The classifier assigned each heartbeat to one of the predefined categories based on features learned from the material and finally outputs a classification. We used the MITBIH testing dataset for our test input to check the validation of the pretrained model. We deployed the pretrained model into our application for heart disease detection.

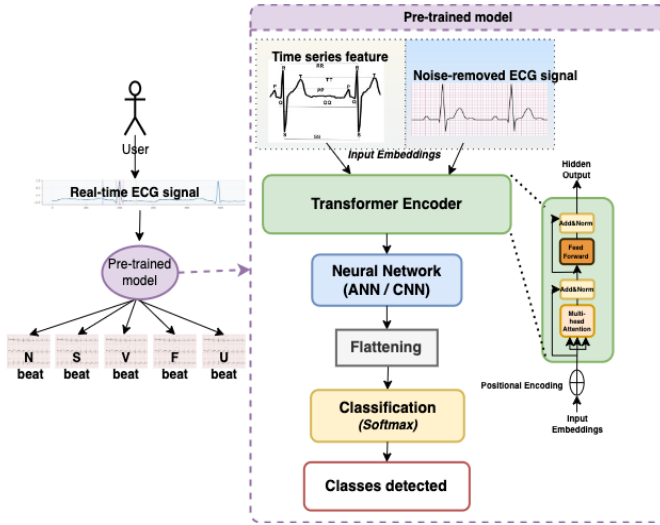


Fig 1: Analysis ECG signal using transformer model with deep learning algorithm.

II. METHODOLOGY

A. ECG data preprocessing

In this study, we utilized the MIT BIH arrhythmia repository from PhysioNet, an open source database [9]. Our focus was on the classification of beat types according to the AAMI criteria [10], which divided beats into five groups: normal and bundle branch block beats (N), supraventricular ectopic beats (S), ventricular ectopic beats (V), fusion beat (F), and unknown beat (Q). They have been defined by experts at MIT BIH to add annotations to the database. We use low pass filter to reduces high-frequency noise, and wavelet transform to decomposes the ECG signal into wavelet components to remove noisy ECG signal. The Pan Tompkins algorithm can accurately detect R peaks and is a pioneer in the accuracy of ECG heartbeat identification [11]. The open source "Neurophysiological Signal Processing Python Toolbox" in Neurokit 2 facilitates the implementation of this algorithm [12]. After successfully detecting the R peak, we calculate the RR interval and the average RR interval, and detect the P, Q, S, T, and T' peaks. Figure 2 represented each peak be detected using Pan Tompkins algorithm from interval 200 to 800 sampling rate. The R peak was the reference point used as a benchmark for identifying and labeling other peaks. P peak (blue), Q peak (green), R peak (red), S peak (pink), T peak (yellow).

B. Feature extraction

We leveraged the Time Series Feature Extraction Library (TSFEL) in Python [13]. A comprehensive set of 175 features was extracted from ECG beats processed using a sliding window technique, which sequentially processed two beats at a time before advancing. The selection of features was refined by applying the analysis of variation (ANOVA) algorithm, and the top 10 features that were most useful for classification purposes were identified. ANOVA is a suite of statistical models and estimation procedures that evaluates variability between group means to inform feature selection [14]. Table I introduced the technical description (TD) and clinical significance (CS) of the top 10 features selected by TSFEL and ANOVA.

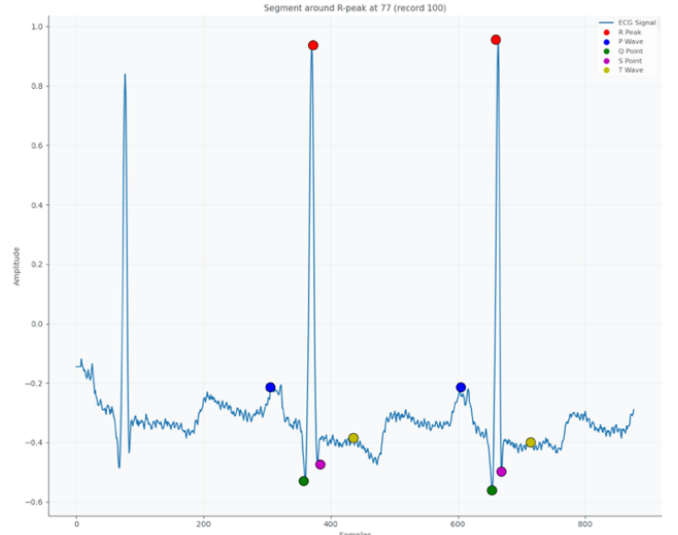


Fig 2: P, Q, R, S, and T peak detection.

Table I: Summary the technical description (TD) and clinical significance (CS) of ANOVA selection features.

ANOVA Selected features (top 10 ranked time series features)	
<i>Wavelet variance</i>	TD: The wavelet coefficient variance of the time frequency of the ECG signal at different scales. CS: Identify heart rhythm changes that may not be visible in the original ECG.
<i>Signal distance</i>	TD: Measures the similarity or variability between consecutive waveforms. CS: Detect heart rhythm changes and identify arrhythmia episodes.
<i>Root mean square</i>	TD: Calculates the square root of the squared mean of the signal's median value. CS: Overall energy of the ECG.
<i>Peak to peak distance</i>	TD: Measures the change in signal amplitude between the highest and lowest points in the ECG cycle. CS: Shows the electrical instability of the heart.
<i>Slope</i>	TD: Calculate the slope of each segment within the ECG signal. CS: Changes in slope can reveal heart conditions.
<i>QRS pos</i>	TD: The position of the QRS complex within the electrocardiogram cycle. CS: Changes in QRS positioning indicate conduction abnormalities or ectopic beats.
<i>RR interval</i>	TD: Measure the time interval between consecutive R peaks. CS: Variations in the RR interval can diagnose conditions such as sinus arrhythmia.
<i>Absolute energy</i>	TD: Calculate the total amount of energy in the ECG signal. CS: Higher energy levels may be related to thicker heart muscle or increased cardiac workload.
<i>Neighborhood peaks</i>	TD: Calculate the maximum number of peaks in a specific neighborhood around each point in the signal. CS: Identify arrhythmias in which normal rhythm patterns are disrupted.
<i>Spectral distance</i>	TD: Use the Fourier transform to measure the spectral differences between ECG segments. CS: Helps identify changes in heart rate.
Combination features	Noise-remove MIT BIH ECG signal + top 10 ranked time series features

C. Classification

We classified normal and bundle branch block beats (N), supraventricular ectopic beats (S), ventricular ectopic beats (V), fusion beats (F), and unknown beats (Q). To classify the data, we used all Tsfel features as input and train machine learning and deep learning models such as Decision Tree (DT), Support Vector Machine (SVM), Bagging Tree, Naive Bayes, K Nearest Neighbors (KNN), artificial neural network (ANN), convolutional neural network (CNN), recurrent neural network (RNN) and long short-term memory (LSTM). The performance of the classifier was evaluated using 10-fold cross-validation. We selected the best-performing ANN and CNN models and combined them with the Transformer model to form a hybrid model of Transformer Artificial Neural Network (TANN) and Transformer Convolutional Neural Network (TCNN). We selected three data combination modules to evaluate performance and power. (1) Top 10 features, (2) MIT BIH noise-removed ECG signal and (3) a combination module of (1) and (2) to evaluate performance and power consumption and adjusted parameters for comparison.

We used key parameters of the transformer model to combine with the neural network to form a hybrid model. Head size: 32. Number of head: 4. Feed forward dimension: 32.

1) *Transformer Artificial Neural Network (TANN)*: The output of this transformer, $H^{(L)}$ was full of features after traversing the L layers. They were then fed into an artificial neural network through one or more dense layers. Each layer in the ANN applied weights and biases, W_{ANN} and b_{ANN} , to the input.

2) *Transformer Convolutional Neural Network (TCNN)*: The output of the Transformer encoder $H^{(L)}$ was the feature input to the CNN. The feature applied filter W_c , and the feature F was generated through the convolution operation $F = ReLU(H * W_c + b_c)$. Features were max-pooled to capture the most important features and passed through additional convolutional layers or flattened into vectors for processing.

D. Performance

Accuracy measures the overall correctness of the model. F1 score is used to evaluate the accuracy of the model when the class distribution is uneven. Our deep learning model was trained on a High-Performance Computing Cluster (HPCC) equipped with an AMD EPYC™ 7702 processor, 64 cores, and 512 GB of memory, running on CentOS 8.1 within the Nocona partition. To weigh the model's efficiency, we measured power consumption by assessing computer usage, and running time of various neural network architectures.

E. Pre-trained model

We built Smart-Health applications using Android Studio Java and integrated pre-trained machine learning models for smart health diagnosis [15]. Use model.save() to save the model trained on the MIT-BIH dataset in Keras, .h5 file. Then converted the model to TensorFlow Lite format and add this .tflite model to the asset folder. Use ByteBuffer to store input data. We used the offline MIT-BIH test data set to verify the accuracy of the model. Passed the ECG signal data through

the interpreter to obtain predictions. This validation was critical to whether the application can make correct predictions in the field before conducting clinical trials.

III. RESULTS

Figure 3 provided a comparative analysis based on the accuracy of various machine learning models in classification tasks using total of 175 time series features. The ANN model achieved an accuracy of 84%, CNN performed slightly better than ANN, with an accuracy of 85%.

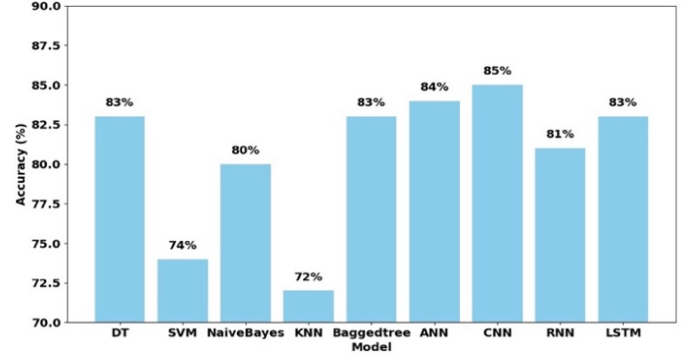


Fig 3: Analysis the accuracy of various machine learning models using total of 175 time series features.

Tables II summarized the performance of CNN, TCNN, ANN and TANN on the noise-removed ECG signal and the top 10 time series features, respectively. The TCNN model was overall better than the CNN, TANN and ANN models. The performance of these models all improved when trained on the first 10 features extracted by TSFEL. Transformer models outperform traditional ML models in the task of identifying time and space between features. CNN and ANN models provide capabilities in efficient feature selection.

Table II : Summary the performance of CNN, TCNN, ANN and TANN on the noise-removed ECG signal and the top 10 time series features, respectively.

Model Layers	Accuracy %	F1 score %	Accuracy %	F1 score %
Features	Noise-removed ECG signal		TSFEL top 10	
CNN / TCNN				
16	78 / 82	82 / 89	88/ 97	92/ 98
32	81 / 85	90 / 90	90/ 96	92/ 97
64	83 / 87	93 / 90	91/ 97	94/ 97
128	86 / 87	90/ 92	93/ 97	93/ 97
256	89 / 89	93/ 92	94/ 98	94/ 98
512	90 / 92	94/ 93	95/ 97	95/ 98
1024	91/ 94	96/ 93	96/ 98	97/ 98
ANN / TANN				
16	72 / 80	88 / 90	83 / 95	90 / 96
32	73 / 81	88 / 91	85 / 97	90/ 97
64	77 / 81	89 / 90	89 / 98	89 / 98
128	80 / 82	92 / 92	92 / 98	92 / 98
256	81 / 85	91 / 93	92 / 98	91 / 97
512	83 / 88	93 / 93	94 / 97	91 / 98
1024	89 / 90	93 / 94	95 / 97	92 / 97

Fig 4 compared the accuracy and CPU usage of four different neural network models using the combination model. As the number of layers increases, the accuracy of all models

showed an overall increase. The accuracy of CNN and TCNN models started from 83% and 88% respectively, and both reached a peak of 98%. CPU usage resource consumption also increased for all models. Higher CPU usage indicated that TCNN has higher computational complexity but also improved performance. TCNN started at around 10% and was just over 22.5% of CPU usage. TCNN handles local pattern recognition space and global context understanding to achieve better performance, but also loses more power consumption.

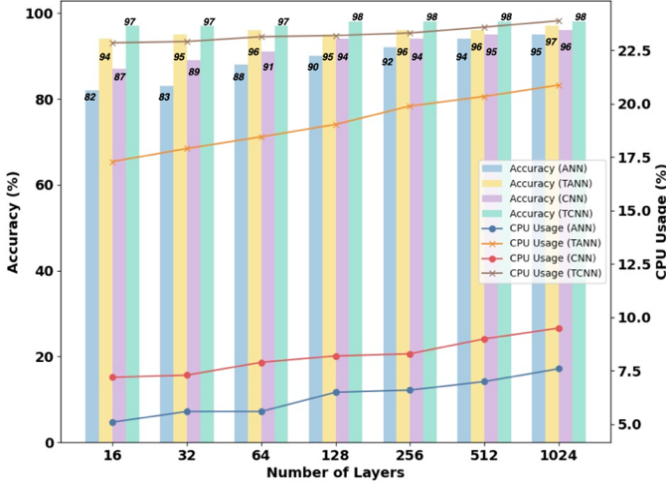


Fig 4: Comparison the accuracy and CPU usage of four different neural network models using combination module.

Fig 5 compared the runtime and CPU time used of four different neural network models using the combination module. The running time and CPU time of all models increase with the number of layers. TCNN and CNN can provide better performance, but will increase computing resources. ANN and TANN may not provide the same accuracy in processing complex patterns but consume less energy. Deep learning is good at processing spatial and temporal data and is suitable for complex applications that do not limit computing resources.

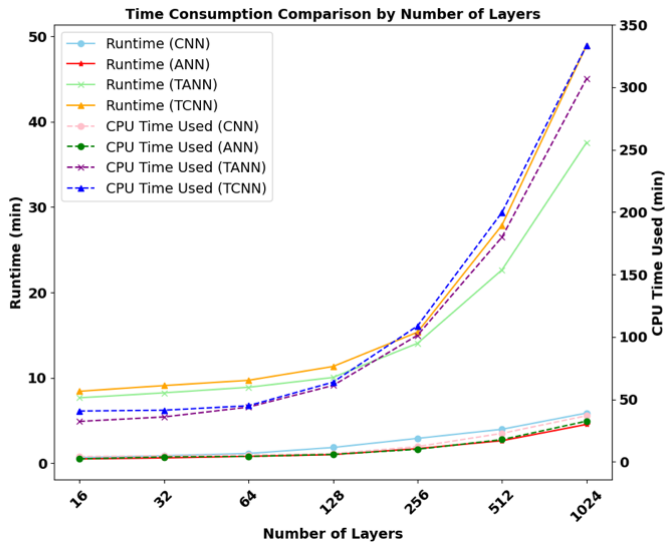


Fig 5: Comparison the runtime and CPU time used of four different neural network models using combination module.

Table III summarized the performance of four pre-trained models evaluated on the Smart-Health Applications using MIT BIH test dataset. TCNN performed best, achieving a top accuracy of 98% and an F1 score of 99%.

Table III : Summarizes the performance of four pre-trained models evaluation using MIT BIH test dataset.

MIT BIH test dataset		
Model	Accuracy %	F1 score %
ANN	92	96
CNN	95	96
TANN	96	98
TCNN	98	99

IV. DISCUSSION.

Our study of a hybrid transformer neural network model for a wearable cardiac monitor achieved significant accuracy and F1 scores, demonstrating its efficacy in detecting cardiac events. Our future work will focus on enhancing the robustness and generalization of the model across different patient populations. This includes employing data augmentation strategies, such as synthetic ECG generation and transformation techniques, to expand the model's exposure to a variety of cardiac diseases and patient demographics. Optimizing models to instantly process and integrate multimodal health data on mobile devices can significantly expand their applicability and effectiveness in real-world scenarios. Further research to improve model interpretability will also enhance trust and usability between clinicians and patients.

V. CONCLUSION.

Heart disease affects the body by impairing the functioning of the heart, leading to a range of health problems. Prompt prevention and detection are very important for people. Our study built a hybrid transformer model combined with neural networks to evaluate the efficiency and power consumption of algorithm in ECG heartbeat classification. We selected the top 10 ranked time series features and combined with the noise-removed MIT BIH ECG signal to feed into a hybrid transformer neural network model for performance and power usage comparison. TCNN model performed better than CNN, TANN and ANN models on noise-removed ECG signals and top 10 ranked features respectively. The accuracy of all models generally improved when testing the combination module, but CPU usage, runtime and CPU time resource consumption also increased across all models. TCNN has the best performance but consume the most power. We converted CNN, ANN, TCNN and TANN models into pre-trained models for Smart-Health application used and detected the validation of each model, respectively. TCNN model achieved the highest accuracy of 98% and an F1 score of 99%. Hybrid transformer neural network model can be use in our application and detect heart disease accurate than the trinational neural network model.

ACKNOWLEDGMENT

This material is based upon work supported by the National Science Foundation under Grant No. 2105766.

REFERENCES

- [1] Vaswani, Ashish, Noam M. Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N. Gomez, Lukasz Kaiser and Illia Polosukhin. "Attention is All you Need." *Neural Information Processing Systems* (2017).
- [2] A. Natarajan et al., "A Wide and Deep Transformer Neural Network for 12-Lead ECG Classification," *2020 Computing in Cardiology, Rimini, Italy, 2020*, pp. 1-4, doi: 10.22489/CinC.2020.107.
- [3] Naik, Manthan & Pancholi, Tirth & Achary, Rathnakar. (2021). Prediction of Congestive Heart Failure (CHF) ECG Data Using Machine Learning. 10.1007/978-981-15-9509-7_28.
- [4] Che, Chao & Zhang, Peiliang & Zhu, Min & Qu, Yue & Jin, Bo. (2021). Constrained transformer network for ECG signal processing and arrhythmia classification. *BMC Medical Informatics and Decision Making*. 21. 10.1186/s12911-021-01546-2.
- [5] M. D. Le, V. Singh Rathour, Q. S. Truong, Q. Mai, P. Brijesh and N. Le, "Multi-module Recurrent Convolutional Neural Network with Transformer Encoder for ECG Arrhythmia Classification," *2021 IEEE EMBS International Conference on Biomedical and Health Informatics (BHI)*, Athens, Greece, 2021, pp. 1-5, doi: 10.1109/BHI50953.2021.9508527.
- [6] Alamr, Abrar and Abdel Monim Artoli. "Unsupervised Transformer-Based Anomaly Detection in ECG Signals." *Algorithms* 16 (2023): 152.
- [7] G. Yan, S. Liang, Y. Zhang and F. Liu, "Fusing Transformer Model with Temporal Features for ECG Heartbeat Classification," *2019 IEEE International Conference on Bioinformatics and Biomedicine (BIBM)*, San Diego, CA, USA, 2019, pp. 898-905, doi: 10.1109/BIBM47256.2019.8983326.
- [8] Yang, M. U., Lee, D. I., & Park, S. (2022). Automated diagnosis of atrial fibrillation using ECG component-aware transformer. *Computers in biology and medicine*, 150, 106115. <https://doi.org/10.1016/j.combiomed.2022.106115>.
- [9] Goldberger AL, Amaral LAN, Glass L, Hausdorff JM, Ivanov PCh, Mark RG, Mietus JE, Moody GB, Peng C-K, Stanley HE. PhysioBank, PhysioToolkit, and PhysioNet: Components of a New Research Resource for Complex Physiologic Signals. *Circulation* 101(23):e215-e220 *Circulation Electronic Pages*; <http://circ.ahajournals.org/content/101/23/e215.full>; 2000 (June 13).
- [10] ANSI/AAMI. Testing and Reporting Performance Results of Cardiac Rhythm and ST Segment Measurement Algorithms, 1998. Association for the Advancement of Medical Instrumentation.
- [11] Pan and W.J. Tompkins, "A Real-time QRS Detection Algorithm," *Ieee Transactions on Biomedical Engineering*, vol. 32, no. 3, 1985, pp. 230-236.
- [12] Makowski, D., Pham, T., Lau, Z.J. et al. NeuroKit2: A Python toolbox for neurophysiological signal processing. *Behav Res* 53, 1689–1696 (2021). <https://doi.org/10.3758/s13428-020-01516-y>.
- [13] Barandas M., Folgado D., Fernandes L., Santos S., Abreu M., Bota P., Liu H., Schultz T., Gamboa H. Tsfel: Time series feature extraction library *SoftwareX*, 11 (2020), Article 100456, 10.1016/j.softx.2020.100456
- [14] I. Rojas, J. Gonzalez, H. Pomares, J. J. Merelo, P. A. Castillo and G. Romero, "Statistical analysis of the main parameters involved in the design of a genetic algorithm," in *IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews)*, vol. 32, no. 1, pp. 31-37, Feb. 2002, doi: 10.1109/TSMCC.2002.1009128.
- [15] Utsha, U.T., Hua Tsai, I., Morshed, B.I. (2024). A Smart Health Application for Real-Time Cardiac Disease Detection and Diagnosis Using Machine Learning on ECG Data. In: Puthal, D., Mohanty, S., Choi, B.Y. (eds) *Internet of Things. Advances in Information and Communication Technology. IFIP IoT 2023, IFIP Advances in Information and Communication Technology*, vol 683. Springer, Cham. https://doi.org/10.1007/978-3-031-45878-1_10.