Medical & Biological Engineering & Computing (2023) 61:1993–2000
https://doi.org/10.1007/s11517-023-02858-3

ORIGINAL ARTICLE

Transformer-based temporal sequence learners for arrhythmia classification

Ann Varghese[1] · **Suraj Kamal**[1] · **James Kurian**[1]

Received: 22 July 2022 / Accepted: 25 May 2023 / Published online: 6 June 2023

Abstract
An electrocardiogram (ECG) plays a crucial role in identifying and classifying cardiac arrhythmia. Traditional methods employ handcrafted features, and more recently, deep learning methods use convolution and recursive structures to classify heart signals. Considering the time sequence nature of the ECG signal, a transformer-based model with its high parallelism is proposed to classify ECG arrhythmia. The DistilBERT transformer model, pre-trained for natural language processing tasks, is used in the proposed work. The signals are denoised and then segmented around the R peak and oversampled to get a balanced dataset. The input embedding step is skipped, and only positional encoding is done. The final probabilities are obtained by adding a classification head to the transformer encoder output. The experiments on the MIT-BIH dataset show that the suggested model is excellent in classifying various arrhythmias. The model achieved 99.92% accuracy, 0.99 precision, sensitivity, and F1 score on the augmented dataset with a ROC-AUC score of 0.999.

Keywords Time-series · Attention · ECG · Transformer

1 Introduction

In its 2020 report, the World Health Organization ranked cardiac diseases first and second among the leading causes of mortality [1]. Arrhythmia is a predominant class of cardiovascular diseases that manifest as slight variations in the normal heart rhythm and beat. The ECG is a non-invasive recording of electrophysiological events in the heart, often of extremely low amplitude. The heart signal depicts the nuances caused by various arrhythmia's, with heart rate, amplitude, and other attributes. Cardiologists must monitor patients' recordings for prolonged periods to discover and distinguish these variations. Even such an inspection is error-prone and involves human bias. Hence, long-term measurements, computer-driven methods, and subtle pattern detection are required to determine the heart's functional variability. Introduction of devices like Zio patch [2] or Apple Watch for continuous monitoring accents the need for an intelligent and robust arrhythmia classification system.

Classification of arrhythmia is possible only by analyzing the sinus rhythm produced by the cardiac conduction system. Figure 1 depicts the cardiac conduction system. The sino-atrial (SA) node, which acts as the heart's "natural pacemaker," initiates cardiac electrical conduction. It causes atrial depolarization by releasing electrical stimuli (denoted by point P). These pulses are delayed at the atrioventricular (AV) node to allow the atria to drain the blood. The stimuli then pass through the septum, separating the ventricles (Q). It then spreads through the ventricular muscles, depolarizing them (R) and the Purkinje fibers (S). The ventricular repolarization (T) is followed by the Purkinje fiber repolarization (U). The correct synchronization of the SA node and AV node maintains the smooth working of the heart.

The subtle morphological variations in ECG signal components, such as the P wave, QRS complex, and T wave, are used by experts in detecting cardiac abnormalities. Accordingly, earlier attempts were to detect and extract features based on these components to aid diagnosis. Machine learning methods like support vector machines were used for classification, whereas techniques like discrete wavelet transforms were used for feature extraction [3-5]. With the

✉ Ann Varghese
ann.doe@cusat.ac.in

Suraj Kamal
surajkamal@cusat.ac.in

James Kurian
james.cusat@gmail.com

[1] Department of Electronics, Cochin University of Science and Technology, Cochin 682022, Kerala, India

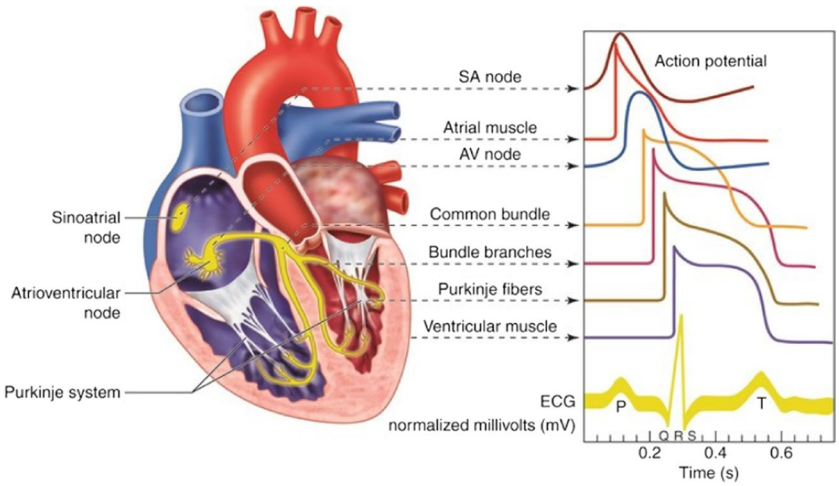

Fig. 1 Cardiac conduction system. Image courtesy: "OLI—Drawing Cardiac conduction system and action potentials—English labels" by Open Learning Initiative, license: CC BYNC-SA

advent of deep learning, the focus shifted from these multi-step processes to single-step classification techniques. The convolutional neural network [6] model using the morphological features or sequence module structures like recurrent neural networks (RNN) [7] using the ECG signal's temporal information is commonly employed.

The attention mechanism that imitates the cognitive attention has proven to improve the performance of many classifiers by attributing higher weights to the most relevant parts of the input sequence [8, 9]. Adding attention will help the classifier to utilize the characteristics of the long-term time series of the signal, thereby assisting in a more accurate analysis [10]. An encoder-decoder architecture called transformer [11] using attention can match the performance of RNNs with attention without the recurrent structures. The attention provides context for input sequence positions and can thus process the entire input concurrently, resulting in higher parallelism and reduced training times than RNNs. Transformers are mainly used for natural language processing (NLP) tasks such as text summarization or translation and computer vision applications. A classification approach using language modelling technique on symbolically represented time series data has been utilized by Li et al. [12]. This provided the impetus to classify time series data directly using an NLP model. The premise behind this is that time-series patterns in a specific class of a given domain are analogous to how texts from various languages or dialects can be distinguished. The input ECG segments are handled like text sentences, where certain words and their combinations will define certain classes.

ECG signal's time series nature and bounded values make it similar to the output of word tokenization and embedding, allowing to skip the input embedding step, reducing the computational overhead. Another aspect that influences the computational cost and training time is the parameter count, but with performance enhancement. The training time is abated using pretrained transformers, which account for the signal's local and global similarities. DistilBERT [13] provides a filtered BERT version [14], which is fine-tunable for purposes where there are constraints on computational training or inference budgets.

2 Methodology

The proposed model architecture is given in Fig. 2. DistilBERT has a 40% reduction in size from BERT while retaining 97% of its language abilities and being 60% faster. The approach is to perform transfer learning from a DistilBERT model pretrained on a language task, and once it converges, finetune the model. That significantly reduces the time it takes to build a model from the ground up. Only the encoder section of the transformer is used as there is no translation for the ECG signal.

Since input data is time-series in nature, no specific tokenization is used. A start and stop tokens are added to the input segments to make it to the configured input size of 256 for the particular model. Since transformers use multi-head attention, it is important to encode the input position information to preserve the time series nature of the signal. The input embeddings are superimposed with sinusoidal positional encodings (PE) to help the model discern the input order. The PE's should be unique, deterministic, generalizable for longer sequences, and equally distant from time steps for different input lengths. This can be expressed mathematically as:

$$PE_{(pos,2i)} = sin\left(\omega_i \cdot pos\right) \tag{1}$$

$$PE_{(pos,2i+1)} = cos\left(\omega_i \cdot pos\right) \tag{2}$$

$$where\ \omega_i = \frac{1}{10000^{2i/d_{\text{model}}}} \tag{3}$$

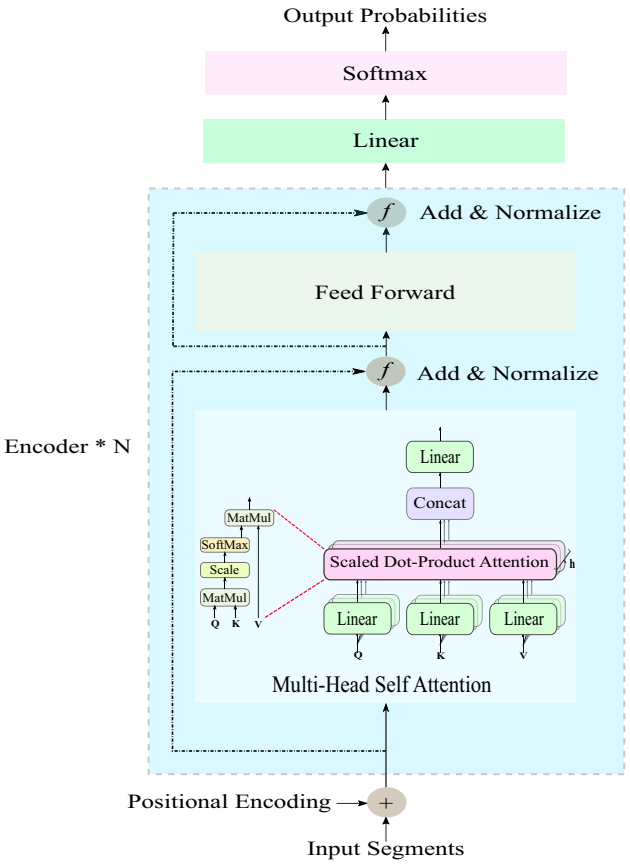

Fig. 2 The proposed transformer model architecture. Number of encoders $N=6$

and d_{model} the encoding dimension. Thus, the frequencies decrease along the encoding vector dimension, whereas the wavelengths form a geometric progression from 2π to $10{,}000 \cdot 2\pi$. For every sine–cosine pair corresponding to frequency ω_i, there is a linear transformation of PE_{pos}.

The encoder section consists of a self-attention layer followed by a feedforward layer. In the attention part, the position encoded input matrix is multiplied by weight matrices $\mathbf{W}^Q$, $\mathbf{W}^K$ and $\mathbf{W}^V$ to obtain the query $\mathbf{Q}$, key $\mathbf{K}$ and value $\mathbf{V}$ matrices respectively. The self-attention score at each position is obtained by multiplying together the query and key matrix. The score is divided by the square root of the dimension of the key vector ($\sqrt{d_k}$) to get more stable gradients. The result is normalized by passing through a softmax layer summing up to 1. The value matrix is multiplied with the softmax output to get the weighted value vectors. The self-attention output at each position is obtained by summing up the weighted value vectors and is expressed as,

$$\text{Attention}(\mathbf{Q},\mathbf{K},\mathbf{V}) = \text{softmax}\left(\frac{\mathbf{Q}\mathbf{K}^T}{\sqrt{d_k}}\right)\mathbf{V} \tag{4}$$

Multiple attention heads are employed to increase the ability of the model to focus on different positions. For this purpose, multiple query, key, and value vectors are randomly initialized and trained. The outputs from these attention heads are concatenated together and multiplied with another weight matrix $\mathbf{W}^O$ to capture all the information from the different attention heads.

$$\text{Multihead}(\mathbf{Q},\mathbf{K},\mathbf{V}) = \text{concat}\left(\text{head}_1, \text{head}_2, \ldots, \text{head}_h\right)\mathbf{W}^O \tag{5}$$

The feed-forward network consists of two 1D convolution layers with an activation layer (ReLU) in between. The self-attention and feed-forward sublayers are followed by residual connections and layer normalization.

$$\text{Sublayeroutput} = \text{LayerNorm}(x + \text{Sublayer}(x)) \tag{6}$$

where x is the sub layer input. A classification head is added to the output of the transformer model consisting of a linear (convolution) layer followed by a softmax layer.

The raw input is denoised and then segmented around the R peak. The beat segments are inputted into the proposed model, and the output probabilities of each selected class are computed. The following sections describe each process in detail.

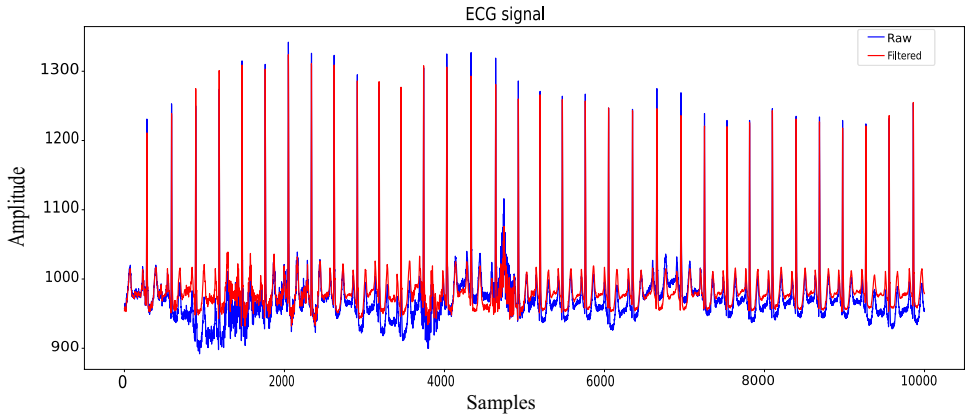

Fig. 3 The MIT-BIH signal before and after filtering with the help of low pass and high pass Butterworth filters

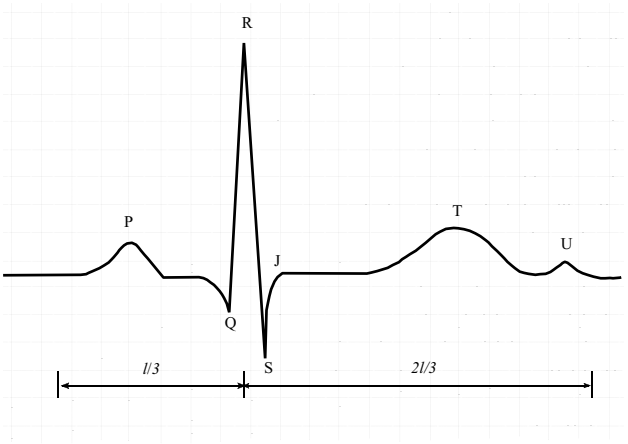

Fig. 4 Cardiac beat. The sampling window is taken such that one third of signal is sampled before the R peak and two third after it

2.1 Data

In the proposed work, the MIT-BIH arrhythmia database [15] consisting of 48 two-channel recordings from 47 patients digitized at 360 Hz is used. Four records belonging to patients with pacemakers are discarded. The modified limb lead II (ML II) is used in this work. The raw input is first denoised, for which two butterworth filters are employed. The butterworth filter calculation formula is given as:

$$H_{(j\omega)} = \frac{1}{\sqrt{1+\varepsilon^2\left(\frac{\omega}{\omega_p}\right)^{2n}}} \tag{7}$$

where n represents the filter order, ω is equal to 2πf and ε is the maximum pass band gain, $A_{\max}$. The order of the filters is five with cutoff frequencies 0.5 Hz and 45 Hz. The combination of the filters gets rid of the baseline wander effect and powerline interference [16]. The raw and filtered ECG signal waveforms are given in Fig. 3.

The twenty beat types in the data set can be grouped into the five AAMI (Association for the Advancement of Medical Instrumentation) super classes: Normal (N), Supraventricular ectopic (S), Ventricular ectopic (V), Fusion (F) & Unknown (Q) beats. Since the number of Q beats is negligible compared to the other classes, only four groups are taken. The data is extracted with the sliding window approach with a length of 254. The MIT-BIH data is annotated at the R peak, and it is used as a marker for the sliding window. The window size is divided into three and positioned such that one part is to the left and two to the right of the R peak. Thus, the data is automatically aligned in all the windows as shown in Fig. 4.

Table 1 SMOTE augmentation details for N, S, V, F beat classes

Class	Number of beats	Balanced set
N	20,000	20,000
S	7006	20,000
V	3025	20,000
F	802	20,000
Total	30,833	80,000

Table 2 Proposed model parameters

Hyper parameter	Value
Batch size	64
d_{model}	256
Number of encoding layers	6
Number of heads	12
Dimension of feed forward layer	768
Dropout	0.2
Loss function	Cross-Entropy
Optimizer	Adam
Learning rate	1e-6

2.2 Implementation details

With imbalanced data, the model prediction priority for the majority class will be high compared to the minority class [17], impacting the ECG-based disease prediction reliability. To overcome underfitting due to data imbalance, SMOTE (Synthetic Minority Oversampling Technique) augmentation [18] is employed. Details of the beat classes are given in Table 1. The data is divided into an 8:1:1 ratio for training, test, and validation.

The model parameters are listed in Table 2. The final convolution layer consists of 64 filters with a kernel size of 3 and ReLu activation. Adam optimizer with learning rate $1e-6$, $\beta_1 = 0.9$, $\beta_2 = 0.98$ and $\epsilon = 1e-9$ is used. The model training is done with 8 CPU Intel Xeon Gold 5315Y system with 45 GB RAM, base processor frequency of 3.20 GHz, and a max turbo frequency of 3.60 GHz. An NVIDIA RTX A4000 Graphics Card – 16 GB DDR6 RAM, 6144 cuda cores, single precision, and compute performance of 19.2 TFLOPS is used.

3 Results and discussions

Any disruption in the smooth conduction of electrical stimuli generated by the SA node manifests as an arrhythmia. Of the four arrhythmia beats considered in this work, the S beats originating from the upper chambers or atria are distinguished by the narrow QRS complex and PR interval. The P wave might overlap with the preceding T wave making

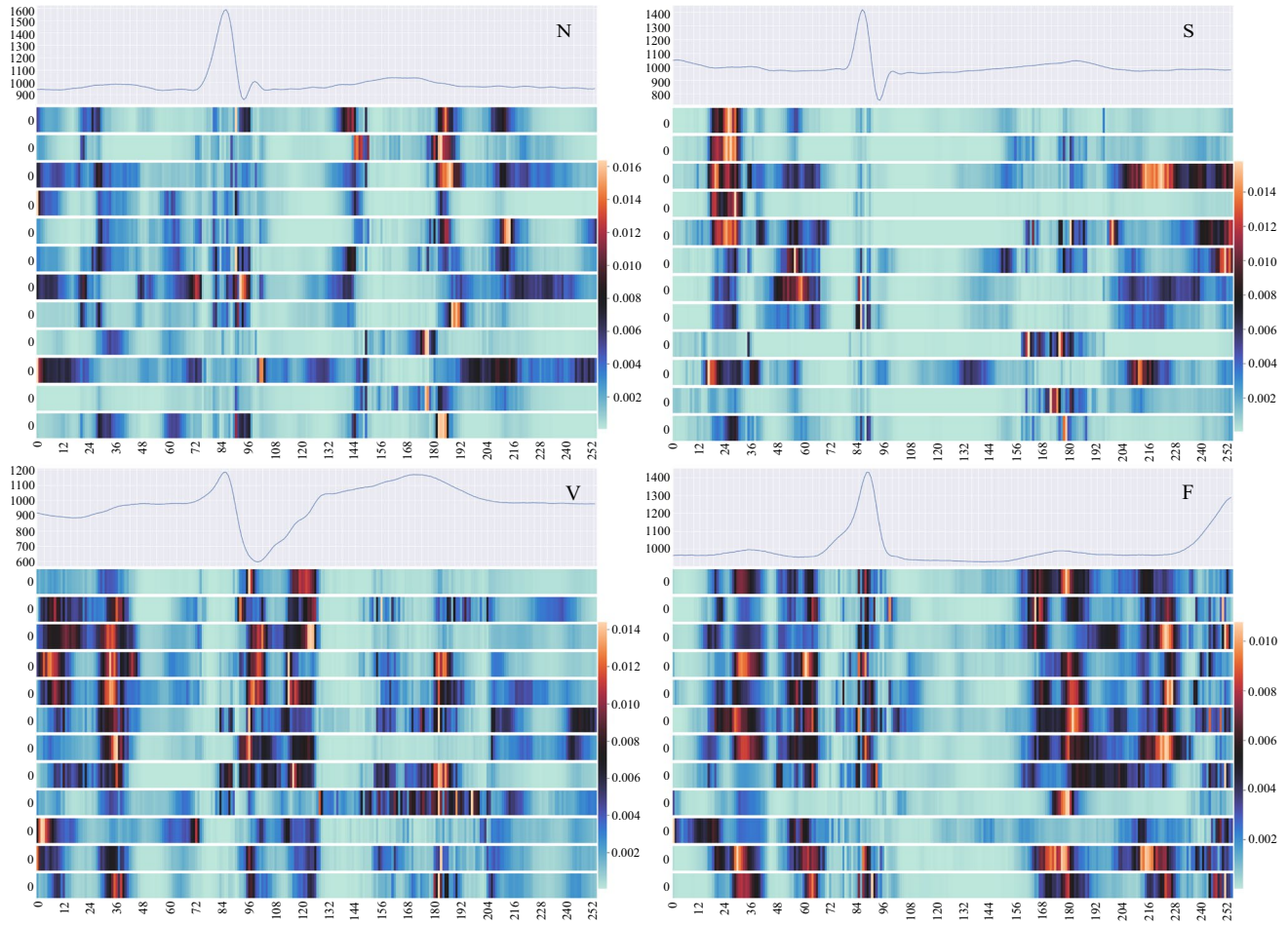

Fig. 5 Attention maps for the four classes

it indistinguishable and manifesting as a deformed T wave. The V beats are extra beats prematurely originating from the ventricles. These are set apart by the broad QRS deflection, discordant QRS and T wave axis, and an absent P wave. The F beats are a fusion of all the other beat types. The main features considered by specialists in differentiating the various beat types are pre-RR interval, the energy within the QRS complex, T wave duration, and R peak location.

From analyzing the attention map of the four beat types, attention values (Fig. 5) are high around the QRS complex, T and P wave locations, indicating that the proposed model is accurately discerning the features. The attention maps for the

Table 3 Performance comparison in terms of accuracy, F1-score, precision, and recall

Model	Method	No. of classes	Pre	Rec	F1-score	Accuracy (%)
Proposed	DistilBERT without input embedding	4	0.99	0.99	0.99	99.92
[21]	Transformer with LightConv attention + CNN embedding	3	0.9506	0.9244	0.9363	99.32
[19]	Fusing transformer	4	0.9974	0.9974	0.9974	99.86
[20]	Low-dimensional denoising embedding transformer (LDTF)	5	0.9839	0.9841	-	-
[22]	Transformer network embedded in CNN	8	-	-	0.7860	-
[23]	BiRNN + CNN embedding	5	0.9720	0.9618	-	99.53
[24]	CNN Bi-LSTM	2	0.4935	0.9182	0.6404	-

*The MIT-BIH data set is used by all except [22] and [24]

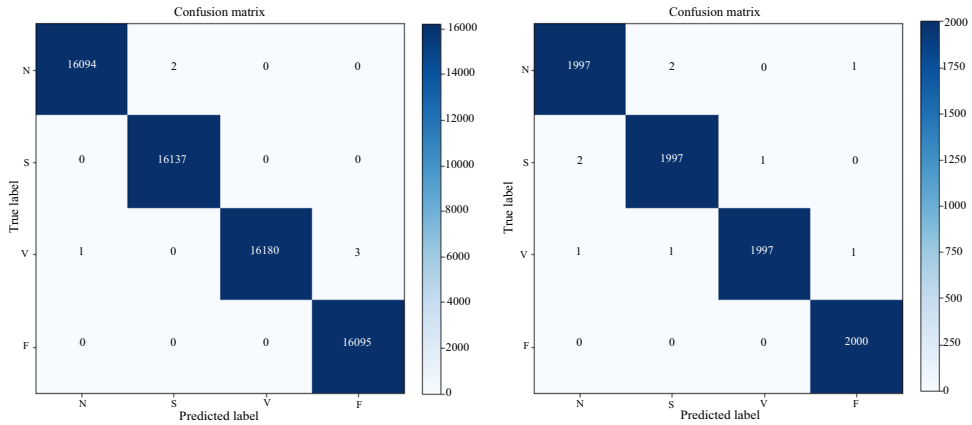

Fig. 6 Confusion matrix for the training data DS1 (with 3% of DS2 data) and testing data DS2

12 heads in the final encoder layer are plotted along with the corresponding inputs for the four arrhythmia types. The attention values obtained from each head are of shape (256×256). The mean value along the columns is taken to get the one-dimensional attention values plotted in the figure.

Accuracy, precision, recall, and F1-score are used to evaluate the classification algorithm performance. The formulae for these metrics is given in Eqs. (8, 9, 10, 11).

$$\text{Accuracy} = \frac{TP + TN}{TP + FP + FN + TN} \tag{8}$$

$$\text{Recall} = \frac{TP}{TP + FN} \tag{9}$$

$$\text{Precision} = \frac{TP}{TP + FP} \tag{10}$$

$$F1\ \text{Score} = 2 \times \frac{\text{Recall} \times \text{Precision}}{\text{Recall} + \text{Precision}} \tag{11}$$

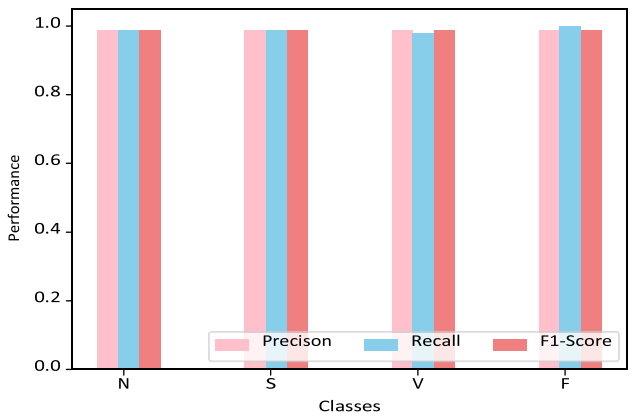

Fig. 7 Performance comparison for the 4 classes in terms of precision, recall, and F1-score

A comparison of the proposed work with those in literature is given in Table 3. Yan et al. [19] introduced a fusing transformer encoder model with an accuracy of 99.86%, where the ECG signal is embedded into 64 dimensions via a CNN followed by transformers and a convolution and softmax layer to perform the classification. A low-dimensional denoising embedding followed by a transformer encoder that is deep and narrow for signal classification is put forward by Guan et al. [20] with a precision-recall value of 0.98.

The proposed model converges in 15 epochs with an accuracy of 99.92% and is superior to other reported works. To ensure the obtained accuracy is not a result of overfitting, another trial is performed by splitting the records patient-wise. The MIT-BIH dataset includes randomly chosen records and those with less frequent but clinically significant arrhythmias that would not be well-represented in a random sample. The 44 usable records are split into two categories based on the characteristics and symptoms of the subjects

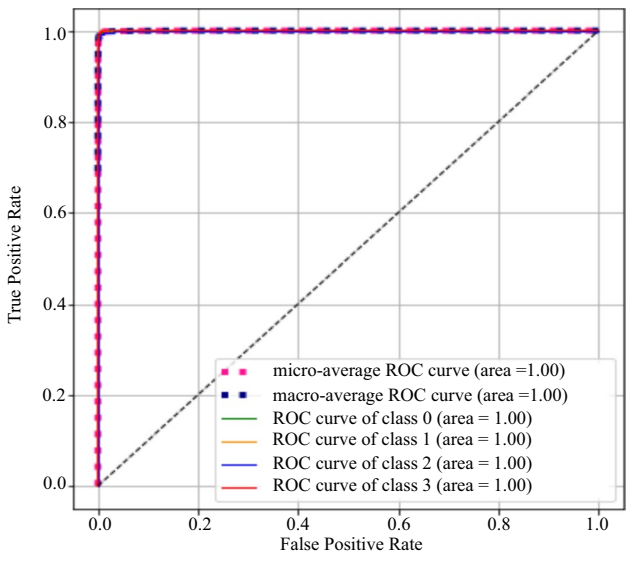

Fig. 8 ROC-AUC plot for the four output classes where 0, 1, 2, and 3 represents N, S, V, and F, respectively

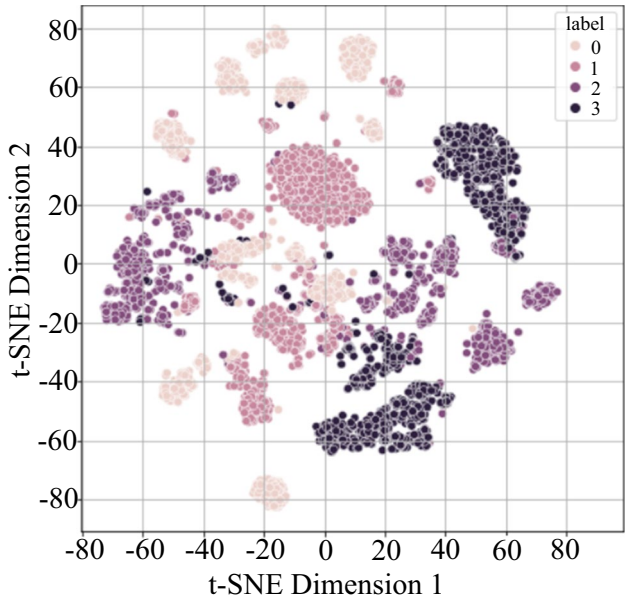

Fig. 9 t-SNE plot for visualizing the four output classes where 0, 1, 2, and 3 represents N, S, V, and F, respectively

[15]. The first set, DS1 (consisting of 20 records with labels that begin with #1), provides a representative sample of the many waveforms and artifacts that an arrhythmia detector may encounter in typical clinical use. The following 24 records (with labels beginning with #2) make up DS2 and comprise complex junctional, supraventricular, and ventricular arrhythmias [25–27]. DS1 and DS2 serve as training and testing sets, respectively.

Since some of the subclasses among AAMI superclasses are primarily or only present in DS2 [15, 19], the classifier, which depends on the morphological information, may not be able to learn the characteristic features of those classes. The AAMI suggested practice is to use no more than 5 min of recordings from a subject during classifier training [28]. For this reason, 3% of the test data was added to the training data to resolve the problem [19]. Testing with this patient-split data gave an accuracy of 99.89% compared with the original accuracy of 99.92% with no split. So with sufficient confidence, we can say this eliminates the concern of overfitting.

The train and test data confusion matrix is given in Fig. 6. Comparing the confusion matrices shows consistent performance for the different beat types. N and S beats are very similar and are difficult to differentiate, which explains the slight increase in the wrong classification for these beats. V and F beats have similar performance in both matrices.

Figure 7 lists the precision, recall, and F1- score for the four classes and the ROC-AUC curve is given in Fig. 8. It achieved an overall precision, recall, and F1 score of 0.99 and an AUC score of 0.999. The high recall obtained indicates that the classes are correctly classified. Similarly, the high precision value denotes the correctness of the labels assigned. The F1 score quantifies how many times a model predicted correctly across the entire dataset. The efficiency of the model in distinguishing between the classes is denoted by the high AUC value. To further visualize the model output classes, a t-Distributed Stochastic Neighbor Embedding (t-SNE) plot is given in Fig. 9. The t-SNE helps visualize high dimensional data in 2D and 3D maps using the dimensionality reduction technique with the help of gradient descent. t-SNE has been trained for 1000 iterations with a learning rate of 200 and a perplexity parameter of 50. The results strongly suggest that the proposed model is a strong candidate for arrhythmia classification with its superior performance in terms of precision, recall, F1 score, and accuracy.

4 Conclusion

This work proposes a pretrained transformer-based model with higher parallelism for arrhythmia classification. Moreover, with inherent self-attention, this model can efficiently detect even the subtle intricacies of the ECG signal. The temporal nature of ECG helps in discarding the embedding step. Along with that using a pretrained filtered version helps in further reduction of parameter count, resource requirement, and training time. The results demonstrate the superiority and effectiveness of the proposed method compared to the state-of-the-art methods.

Declarations

Competing interests The authors declare no competing interests.

References

1. Organization WH (2020) The top 10 causes of death https://www.who.int/news-room/fact-sheets/detail/the-top-10-causes-of-death. Accessed 10 Jan 2021
2. Barrett PM et al (2014) Comparison of 24-hour holter monitoring with 14day novel adhesive patch electrocardiographic monitoring. Am J Med 127(1):95-e11
3. da S Luz EJ, Schwartz WR, C'amara-Ch'avez G, Menotti D (2016) Ecg-based heartbeat classification for arrhythmia detection: a survey. Comput Methods Programs Biomed 127:144–164
4. Afkhami RG, Azarnia G, Tinati MA (2016) Cardiac arrhythmia classification using statistical and mixture modeling features of ecg signals. Pattern Recogn Lett 70:45–51
5. Xie J et al (2021) A signal quality assessment–based ecg waveform delineation method used for wearable monitoring systems. Med Biol Eng Compu 59(10):2073–2084
6. Varghese A, Sylaja MM, Kurian J (2022) Conception and realization of an iot-enabled deep cnn decision support system for automated arrhythmia classification. J Intell Syst 31(1):407–419

7. Ebrahimi Z, Loni M, Daneshtalab M, Gharehbaghi A (2020) A review on deep learning methods for ecg arrhythmia classification. Expert Syst Appl: X 7:100033
8. Majid S et al (2022) Attention based cnn model for fire detection and localization in real-world images. Expert Syst Appl 189:116114
9. Tomar NK, Jha D, Riegler MA, Johansen HD, Johansen D, Rittscher J, Halvorsen P, Ali S (2022) Fanet: A feedback attention network for improved biomedical image segmentation. IEEE Transactions on Neural Networks and Learning Systems. IEEE
10. Zhang J et al (2020) Ecg-based multi-class arrhythmia detection using spatio-temporal attention-based convolutional recurrent neural network. Artif Intell Med 106:101856
11. Vaswani A, Shazeer N, Parmar N, Uszkoreit J, Jones L, Gomez AN, Kaiser L, Polosukhin I (2017) Attention is all you need. Adv Neural Inform Process Syst 30
12. Li D, Li L, Bissyand´e TF, Klein J, and Le Traon Y (2016) Dsco: A language modeling approach for time series classification. International Conference on Machine Learning and Data Mining in Pattern Recognition. Springer pp 294–310
13. Sanh V, Debut L, Chaumond J, and Wolf T (2019) Distilbert, a distilled version of bert: smaller, faster, cheaper and lighter. arXiv preprint arXiv:1910.01108. Accessed 10 Oct 2020
14. Devlin J, Chang M-W, Lee K, and Toutanova K (2018) Bert: Pre-training of deep bidirectional transformers for language understanding. arXiv preprint arXiv:1810.04805. Accessed 9 Sept 2020
15. Moody GB, Mark RG (2001) The impact of the mit-bih arrhythmia database. IEEE Eng Med Biol Mag 20(3):45–50
16. Ahmad I, Ansari F, Dey U (2013) Power line noise reduction in ecg by butterworth notch filters: A comparative study. Development (IJECIERD) 3(3):65–74
17. He H, Garcia EA (2009) Learning from imbalanced data. IEEE Trans Knowl Data Eng 21(9):1263–1284
18. Chawla NV, Bowyer KW, Hall LO, Kegelmeyer WP (2002) Smote: synthetic minority over-sampling technique. J Artif Intell Res 16:321–357
19. Yan G, Liang S, Zhang Y, and Liu F (2019) Fusing transformer model with temporal features for ecg heartbeat classification. 2019 IEEE International Conference on Bioinformatics and Biomedicine (BIBM). IEEE, pp 898–905
20. Guan J, Wang W, Feng P, Wang X, and Wang W (2021) Low-dimensional denoising embedding transformer for ecg classification. ICASSP 2021-2021 IEEE International Conference on Acoustics, Speech and Signal Processing(ICASSP). IEEE, pp 1285–1289
21. Meng L et al (2022) Enhancing dynamic ecg heartbeat classification with lightweight transformer model. Artif Intell Med 124:102236
22. Che C, Zhang P, Zhu M, Qu Y, Jin B (2021) Constrained transformer network for ecg signal processing and arrhythmia classification. BMC Med Inform Decis Mak 21(1):1–13
23. Mousavi S, Afghah F (2019) Inter-and intra-patient ecg heartbeat classification for arrhythmia detection: a sequence to sequence deep learning approach. ICASSP 2019-2019 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP). IEEE, pp 1308–1312
24. Yin Y et al (2021) An Algorithm for Locating PVC and SPB in Wearable ECGs. 2021 13th International Conference on Communication Software and Networks (ICCSN). IEEE, pp 89–93
25. Li J, Si Y, Xu T, Jiang S (2018) Deep convolutional neural network based ecg classification system using information fusion and one-hot encoding techniques. Math Probl Eng 2018:1
26. Zhai X, Tin C (2018) Automated ecg classification using dual heartbeat coupling based on convolutional neural network. IEEE Access 6:27465–27472
27. Kiranyaz S, Ince T, Gabbouj M (2017) Personalized monitoring and advance warning system for cardiac arrhythmias. Sci Rep 7(1):1–8
28. ECAR, AAMI (1987) Recommended practice for testing and reporting performance results of ventricular arrhythmia detection algorithms. Assoc Adv Med Instrumen 69

Ann Varghese completed her Master of Technology in VLSI & Embedded Systems in 2014 from Mahatma Gandhi University, Kerala, India. She is currently pursuing PhD at the Department of Electronics at Cochin University of Science & Technology, Kerala, India. Her research interest includes Deep Learning, Biomedical Pattern Recognition, and Signal Processing.

Suraj Kamal completed his Master of Science Programme in Electronics from College of Applied Science, Thodupuzha, affiliated to MG University, Kerala, India, in 2006. He is presently working as Project Scientist in the Department of Electronics, Cochin University of Science and Technology, Kerala. His main areas of research are Applied Underwater Acoustics, Artificial Intelligence and Machine Learning.

Dr James Kurian received a Master's degree in physics specializing in solid state electronics from Kerala University in 1983 and a Master'ws degree in technology (Digital Electronics) and a PhD in Robotics from Cochin University of Science and Technology, Kerala, India, in 1990 and 2009 respectively. He joined the Department of Electronics as Lecturer in 1992 and was elevated to Professor in Electronics in 2009. He has performed research on Non-destructive Testing Systems, Mobile Robot Localisation and Control Systems, Embedded Control and Automation Systems, Digital Instrumentation, and Machine Learning. He has more than 40 research publications on these areas to his credit.