Studies in Medical and Health Sciences
https://smhs.sabapub.com
ISSN: 3007-3707

2024 Volume 5, Issue 2: 58– 71

DOI: https://doi.org/10.48185/smhs.v1i2.1320

Convolutional Block Attention BiLSTM for Arrhythmia Detection

Wesley Chorney[1] *

[1]School of Medicine, University College Cork, Ireland

Received: 24.07.2024 • Accepted: 05.09.2024 • Published: 05.09.2024 • Final Version: 28.11.2024

Abstract: Cardiovascular diseases represent a significant cause of mortality, with millions of electrocardiograms being recorded each year. Therefore, methods of automated diagnosis for electrocardiograms are of particular interest. Since electrocardiograms have recognizable features and are time-dependent, we propose a model using convolutional layers, convolutional attention, and long short-term memory units. The model is trained and validated on the MIT-BIH Arrhythmia database, and achieves an accuracy of 99.10%, a precision of 99.09%, a specificity of 99.64%, a sensitivity of 95.90%, and an F1 score of 97.47%.

Keywords: LSTM, Attention, Convolutions, Electrocardiograms, Arrhythmia

1. Introduction

According to data from the World Health Organization, cardiovascular dis- eases are the leading cause of death [1]. Because many cardiovascular diseases arise due to or present with arrhythmias (irregular electrical activity of the heart), they can be detected via an electrocardiogram (ECG). Electrocardiograms are valuable medical tests, due to the fact that they can be used as di- agnostic tools for a wide range of cardiovascular diseases [2], and because they are a non-invasive test [3]. Non-invasive tests do not break the skin or enter the body, and thus can be administered to patients in a relatively straightforward manner.

As a result of the ease with which an ECG can be obtained, as well as the diagnostic value of the signal, data suggests that over 300 million ECGs are recorded annually [4]. The quantity of data recorded creates an extremely large burden for cardiologists and other doctors, who must analyze each signal to evaluate the presence of arrhythmias. Because of the sheer volume of ECGs recorded and the care required to accurately assess each signal, manual analysis of ECGs is a challenge. However, as ECGs record the electrical activity of the heart, which in a healthy heart follows a normal cyclical pattern (see Figure 1), and in a diseased heart can vary according to predictable patterns [5], ECGs are promising candidates for automated diagnosis. Automated diagnosis methods with high sensitivity and specificity could be used clinically, saving time for cardiologists and other doctors, who would then be able to spend time on more clinically relevant tasks. In a sense, the suitability of ECGs to automated diagnosis has already been demonstrated. Early approaches to automated ECG analysis relied on handcrafted feature detectors [6], where features could relate to morphological features of an ECG, time relationships between different detected features, wavelets, or more. With the advent of deep learning, automated ECG analysis began to focus less on handcrafted feature detectors and more on neural networks, where different studies have used multilayer perceptrons [7, 8], convolutional neural networks [9, 10], or recurrent neural networks [11, 12].

* Corresponding Author: wes.chorney@gmail.com

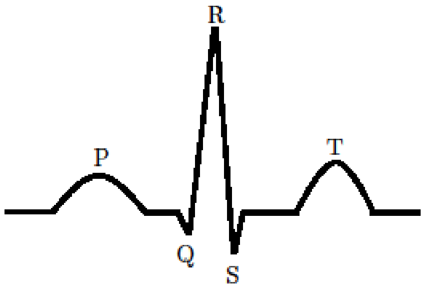

Figure 1: A simplified representation of the cardiac cycle, as captured by an ECG. The P wave, QRS complex, and T wave are labeled.

In this study, we present a model for classifying arrhythmias in ECGs that makes use of convolutional layers, convolutional block attention, and long-short term memory units. In particular, we

- propose a novel model architecture for classification of arrhythmias,
- validate the model performance on an extremely common dataset of ECG recordings, and
- demonstrate the superior performance of the proposed model.

The remainder of this study is structured as follows. In Section 2, we provide an overview of related studies that have explored automated diagnosis of ECGs, with a focus on those that use convolutional layers, recurrent layers, or both. In Section 3, we provide a description of the model, including its mathematical underpinnings. In Section 4, we introduce the dataset we use to validate the model, and discuss the results, with comparisons to other models in the literature. Finally, in Section 5, we conclude, providing a summary of what was achieved, as well as discussing future work.

2. Related Work

Automated diagnosis of arrhythmias is a well-studied field, with regular innovation and progress. Perhaps the most well-studied dataset is the Massachusetts Institute of Technology Beth Israel Hospital (MIT-BIH) arrhythmia database [13]. Owing to both its ubiquity in the field, as well as the number of high quality studies of which it plays a central part, we train and evaluate our proposed model on this dataset. We focus on those studies that make use of convolutional layers, recurrent layers, and attention in the following.

Methods from computer vision have been applied to arrhythmia detection by transforming ECGs (which are time-series) into images, following which convolutional neural networks may be directly applied to the image [14, 15]. One dimensional convolutional neural networks have also been applied to ECG analysis [16], as ECGs are inherently one-dimensional (albeit they can have different channels). Studies seem to indicate that deep networks perform better than shallow with respect to diagnosis [17], which is unsurprising as successive convolutional layers can capture more detailed features. One dimensional convolutions have the benefit of not requiring any transformation before they can be used to process an ECG.

Convolutional neural networks have also been combined with other methods from cardiology, machine learning and mathematics. For instance, convolutional neural networks have been used in conjunction with recurrence plots [18], which is a transformation of a time-series into phase space. Similarly, ECGs have been transformed into spectrograms to give frequency information before being processed by a convolutional network [19, 20]. Additionally, multiple convolutional neural networks have been combined in order to obtain a more accurate diagnosis of ECGs [21]. Convolutional neural networks have also been combined with extreme learning machines (ELMs) for ECG diagnosis [22]. Methods of isolating parametric features from ECGs (such as QRS complexes or P waves) have been used in conjunction with convolutional neural networks [23]. The efficacy of transfer learning has also been studied; for instance, with residual neural networks as in [24].

While convolutional neural networks are natural approaches in that they are feature detectors, they are unable to model sequence-dependent features. Therefore, some studies have used recurrent neural networks and their derivatives in order to incorporate modelling of the sequential nature of ECGs. Studies have used both algorithmic ECG preprocessing and ensembling [25], including CNNs, long short-term memory units (LSTMs) or other recurrent neural net- works (RNNs), and algorithm methods [26]. Similarly, CNNs and LSTMs have been combined with algorithmic methods of extracting features from ECGs [27]. Other studies have used wavelets and particle swarm optimization (PSO) in order to arrive at a diagnosis [28].

Another significant advancement in image processing was the use of attention [29]. Convolutional attention has been extremely successful in many different domains, including being used for improvements in ResNet [30], in facial recognition [31], and electroencephalography analysis [32]. It has also been used successfully in ECG analysis, denoising, and diagnosis [33, 34, 35]. Attention has also been studied with respect to the MIT-BIH Arrhythmia database. For instance, it has been used with both two-dimensional CNNs [36] and one- dimensional CNNs [37]. Finally, attention has been used with feature fusion and convolutions [38].

Table 1 gives an overview of these works, including the accuracy achieved by each study.

Table 1: Recent studies using the MIT-BIH Arrhythmia database.

Year	Techniques	Accuracy	Feature Detection	Sequence Aware	Attention	Disadvantages	Reference
2018	2D CNN	99.05\%	✓	✗	✗	Insensitive to sequential features	[15]
2019	2D CNN	97.42\%	✓	✗	✗	Insensitive to sequential features	[14]
2019	1D CNN	93.60\%	✓	✗	✗	Insensitive to sequential features	[16]
2019	2D CNN Spectrogram	97.10\%	✓	✗	✗	Insensitive to sequential features Additional transformation required	[19]
2019	2D CNN Spectrogram	99.00\%	✓	✗	✗	Insensitive to sequential features Additional transformation required	[20]
2020	CNN Ensemble	99.03\%	✓	✗	✗	Insensitive to sequential features	[21]
2020	Neural Network KNN Feature Detection	97.70\%	✓	✓	✗	Additional transformation required	[23]
2020	CNN LSTM	95.90\%	✓	✓	✗	Relatively low accuracy	[39]
2021	2D CNN Recurrence Plot	98.41\%	✓	✓	✗	Additional transformation required	[18]
2021	1D CNN	97.41\%	✓	✗	✗	Insensitive to sequential features Deep network	[17]
2021	CNN LSTM Feature Detection Ensembling	95.81\%	✓	✓	✗	Additional transformation required Multiple models	[26]
2022	CNN RNN Ensembling Data Resampling Feature Detection	94.43\%	✓	✓	✗	Complicated pipeline Relatively low accuracy	[25]
2022	2D CNN LSTM Scalograms	99.00\%	✓	✓	✗	Additional transformation required	[40]
2023	Particle Swarm Optimization Wavelet Transform KNN	98.50\%	✓	✗	✗	Heuristic optimization Additional transformation required	[28]
2023	1D CNN Extreme Learning Machine	98.82\%	✓	✗	✗	Insensitive to sequential features	[22]
2023	ResNet Transfer Learning	90.80\%	✓	✗	✗	Insensitive to sequential features Relatively low accuracy	[24]
2023	1D CNN LSTM Feature Detection	97.20\%	✓	✓	✗	Additional transformation required	[27]
2023	2D CNN Attention	98.68\%	✓	✗	✓	Insensitive to sequential features	[36]
2023	1D CNN Attention	96.19\%	✓	✗	✓	Insensitive to sequential features	[37]
2023	2D CNN Attention Feature Fusion	97.72\%	✓	✗	✓	Insensitive to sequential features Additional transformation required	[38]

Automated diagnosis of ECGs is a rapidly-developing field with a wealth of active research. This is likely due to its clinical relevancy, and that there is more ECG data than other healthcare data. Many studies use some com- bination of data preprocessing, CNNs, RNNs, and attention. Since ECGs have notable features which can serve as markers for pathology, it is natural to use convolutional layers. Furthermore, since ECGs are time series, it makes sense to use RNNs or LSTMs, so that they can be processed in a sequence-dependent manner. This is relevant since certain cardiovascular pathologies present in only in a sequence dependent manner. For instance, to differentiate between Mobitz type 1 and 2 second degree atrioventricular block, the relationship between the P wave and QRS complex needs to be investigated. Finally, while CNNs are capable of detecting features, attention allows emphasis of features based on their importance. Therefore, we propose using a model complete with convolutional layers augmented with convolutional block attention, followed by bidirectional LSTMs for sequential processing. To the best of our knowledge, this is the first study investigating such a model architecture.

3. Materials and Methods

Let $x \in R^{L \times n}$ be an ECG, where l is the number of sampled points, and n is the number of leads. For simplicity, in much of the following, we will assume that n = 1; that is, the ECG is recorded using a single lead, but it is relatively straightforward to extend the majority of the following to an arbitrary number of leads. The following discussion holds for a general signal $x \in R^{L \times n}$, but given the subject of the study, we will often refer to x as an ECG.

3.1. Convolutions and Attention

A one dimensional convolution makes use of a filter $k \in R^{L}{}_{k}$, where we refer to L_k as the size of the filter. The filter is the learned feature detector component of the convolutional layer. Letting x_i represent the *i*th entry of x, and $x_{i:j}$ represent entries i through j of x, then a single element of the

convolution operation is given by

$(x \star k)_j = \langle x_{i:i+Lk-1}, k \rangle$, (1)

where $x \star k$ is used to denote the convolution of x with the filter k, and $\langle \cdot, \cdot \rangle$ is the inner product in R. Typically, x is padded with zeros so that the convolution operation produces a product of equal dimension to the input. That is, if $x \in R^L$, then x is padded so that $(x \star k) \in R^L$ as well. One convolutional layer typically uses many filters $k_1, \ldots, k_m$, so that a number of different features can be extracted from the input at once. Stacking convolutional layers with multiple filters means that multiple higher-level features are captured.

While convolutional layers allow for multiple different features to be captured, these will typically vary in quality. For example, from the perspective of an ECG, a signal that captures delta waves [41] may be useful for diagnosing Wolff-Parkinson-White syndrome, while a signal that captures the peak of a QRS complex may not be as valuable diagnostically. While the feature detectors learned by a convolutional layer may be significantly more abstract, they still vary in diagnostic utility. Convolutional block attention addresses this issue, allowing for weighting of features both spatially (along the time axis), which is referred to as spatial attention, and relative to one another, which is referred to as channel attention.

Given $x \in R^{L \times n}$, spatial attention makes use of global maximum pooling and global average pooling across the channel dimension to produce two elements m_s, $a_s \in R^{L \times 1}$. These two elements are concatenated to create an element $m_s \oplus a_s \in R^{L \times 2}$, before being fed through convolutional layer with a single filter to obtain $s_a \in R^{L \times 1}$, which is then broadcast to $R^{L \times n}$. This creates a spatial attention map, which is able to emphasize or deemphasize elements of each learned feature. Finally, the Hadamard product (which we denote $\odot$) is taken between s_a and x to obtain the spatial attention A_s. That is,

$A_s(x) = C(\mathrm{AvgPool}(x) \oplus \mathrm{MaxPool}(x)) \odot x$, (2)

where C is a convolutional layer with a single filter, and AvgPool and MaxPool denote global average pooling and global maximum pooling, respectively.

Channel attention is very similar to spatial attention, except for the pooling is used across the time dimension, yielding elements m_c, $a_c \in R^{1 \times n}$. These are then concatenated and fed through a two-layer dense network, yielding a channel attention map $c_a \in R^{1 \times n}$, which is combined with x using the Hadamard product to obtain the channel attention A_c. That is,

$A_c = D(\mathrm{AvgPool}(x) \oplus \mathrm{MaxPool}(x)) \odot x$, (3)

where D is a two-layer dense network. Channel and spatial attention are used sequentially in the convolutional block attention module, which we write as

$\mathrm{CBAM}(x) = A_s(A_c(x))$. (4).

3.2. Long Short-Term Memory

While feature detectors have proven extremely useful in ECG diagnosis, they are time-insensitive. Recurrent neural networks are able to produce an output based on how a sequence evolves in time. Therefore, they could be used to capture time-sensitive aspects of an ECG — for example, the characteristic PR interval elongation of Mobitz type 1 AV block. Figure 2 visualizes the LSTM layer.

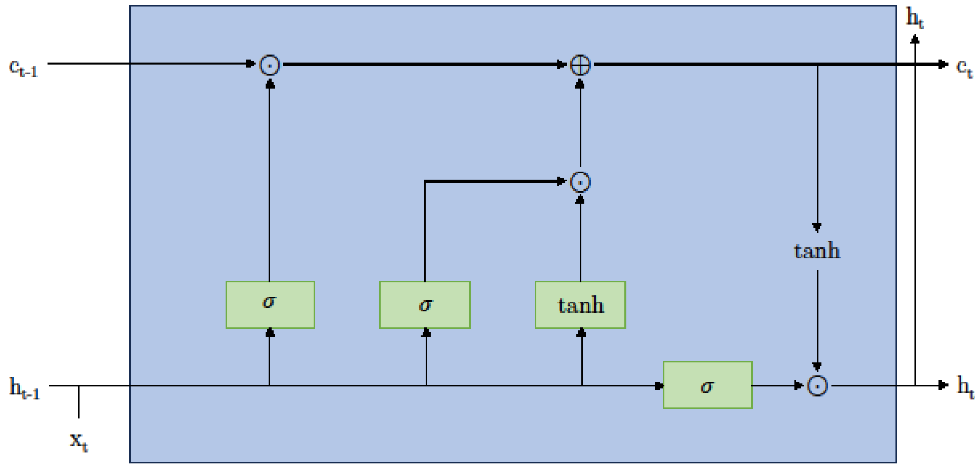

Figure 2: A long-short term memory (LSTM) unit. By ⊕ we denote component-wise addition, and by ⊙ we denote component-wise multiplication. Elements in green are layers with either sigmoid or tanh activations, and the tanh without a background is component-wise.

At time $t - 1$, the LSTM holds a state c_{t-1}, which is updated based on the hidden state h_{t-1} of the LSTM, as well as the input x_t at time t (which are concatenated). The LSTM can be more succinctly described via the following equations:

$f_t = \sigma(W_f x_t + U_f h_{t-1} + b_f)$ (5)

$i_t = \sigma(W_i x_t + U_i h_{t-1} + b_i)$ (6)

$o_t = \sigma(W_o x_t + U_o h_{t-1} + b_o)$ (7)

$c'_t = \tanh(W_c x_t + U_c h_{t-1} + b_c)$ (8)

$c_t = f_t \odot c_{t-1} + i_t \odot c'_t$ (9)

$h_t = o_t \odot \tanh(c_t),$ (10)

where W_f, W_i, W_o, W_c, U_f, U_i, U_o, U_c are learnable matrices with accompanying bias vectors.

Essentially, the LSTM is able to take as input a sequence and propagate a state c_t that captures information on the sequence that should be memorized. This state is influenced by the hidden state h_{t-1} at the previous timestep, as well as the new input x_t. Conversely, the state c_t will also affect the new hidden state h_t that the LSTM outputs. When the signal terminates at t_f, h_{tf} can be used (or the sequence of $h_1, \ldots, h_{tf}$) to perform a task of interest.

In the context of the current study, LSTMs are employed owing to the fact that convolutional layers output a sequence of features, which reflect features captured at different steps in the sampling process of an ECG. Therefore, they can be used to process the features in a time-sensitive manner, and the output can be used for classification that is both sensitive to detected features, as well as the sequence in which those features arose.

3.3. Proposed Architecture

Given the importance of using both convolutional layers as well as LSTM layers, and how attention could improve the diagnostic capability of such a model, we propose using a combination of all three layers. Table 2 gives an overview of the proposed model.

The model uses a sequence of three convolutional layers before using convolutional block attention to emphasize or deemphasize the features learned at this point. After the convolutional block attention, the sequence of features is input into two bidirectional LSTMs so that higher order time-dependent features can be learned based on the features detected from the convolutional layers. Finally, the output is fed into a two-layer dense neural network for classification.

Table 2: Summary of the proposed model architecture.

Layer (Type)	Output Shape	Parameters
1D Convolution	(None, 120, 64)	960
Batch Normalization	(None, 120, 64)	256
MaxPooling1D	(None, 40, 64)	0
1D Convolution	(None, 40, 64)	20,544
Convolutional Attention Layer	(None, 40, 64)	3,152
Batch Normalization	(None, 40, 64)	256
MaxPooling1D	(None, 13, 64)	0
LSTM	(None, 13, 128)	66,048
LSTM	(None, 64)	41,216
Dense	(None, 16)	1,040
Dense	(None, 5)	85
Total Parameters:	133,557	
Trainable Parameters:	133,301	
Non-Trainable Parameters:	256	

The model is trained on 60% of the data, with 20% reserved for validation, and the remaining 20% used as a testing set. Training proceeded for 15 epochs, using the Adam optimizer with a learning rate of 10^4 and a batch size of 32. Upon training end, the best weights (based on validation accuracy) were restored to the model. Cross entropy loss was used to train the model, which, given predicted class probabilities $\hat{y}$ and a true output y, can be expressed as

$$L(y, \hat{y}) = -\sum \delta(1 - y_j) \log(\hat{y}_j), \quad (11)$$

where δ is the Dirac delta function, and y_j is the *j*th element of $y \in R^L$.

Finally, models are evaluated on a number of different evaluation metrics in order to give a comprehensive view of performance. In particular, the MIT-BIH dataset is imbalanced, where the majority of beats are normal, and pathological beats are relatively rare. Furthermore, the number of different pathological beats differs as well. Therefore, a single accuracy metric can be misleading. We present results with respect to accuracy, precision, recall, specificity, and F1 score. Letting TP, FP, TN, and FN denote true positives, false positives, true negatives, and false negatives, respectively, the above metrics are defined as

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

$$\text{Precision} = \frac{\mathrm{TP}}{\mathrm{TP} + \mathrm{FP}}$$

$$\text{Recall} = \frac{\mathrm{TP}}{\mathrm{TP}+\mathrm{FN}}$$

$$\text{F1 Score} = 2 \cdot \frac{\text{Precision} \cdot \text{Recall}}{\text{Precision} + \text{Recall}}$$

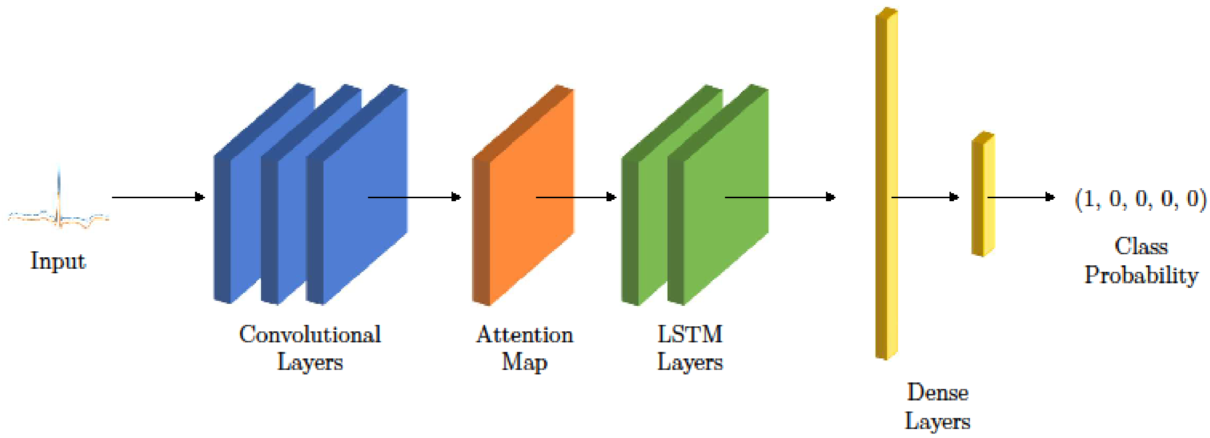

Figure 3: The proposed model architecture. A sequence of three convolutional layers before convolutional block attention is used. The output is fed into two bidirectional LSTMs before being processed by dense layers.

4. Results and Discussion

4.1. Dataset

Due to its ubiquity in the literature regarding automated ECG diagnosis, we make use of the MIT-BIH Arrhythmia database, which consists of a number of two-lead ECG recordings from several patients. The dataset itself is separated beat-by-beat, and the beats themselves are evaluated independently by two different cardiologists to ensure high quality labeling.

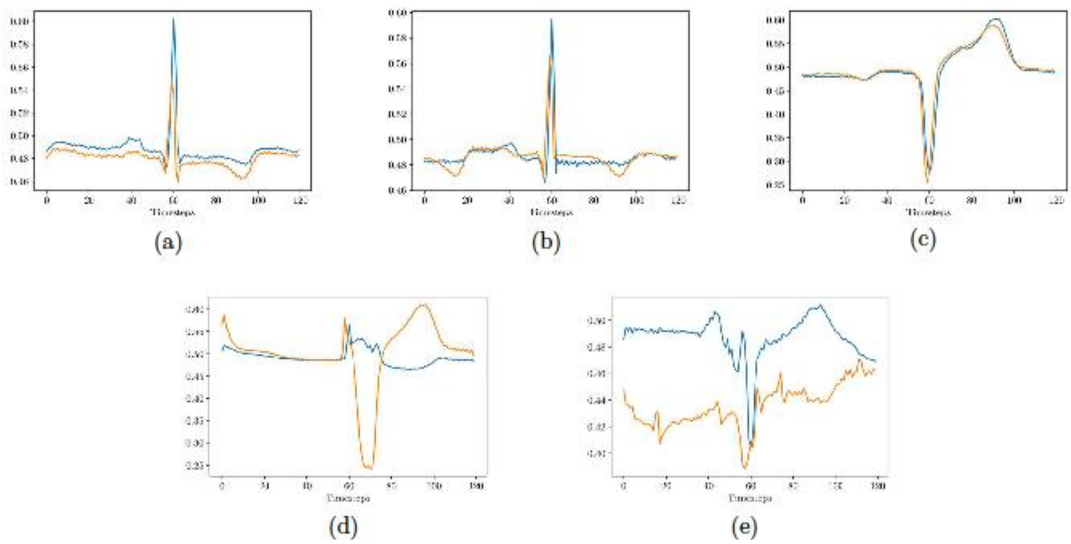

Figure 4: A sample of the beats from the dataset, including normal (a), ventricular ectopic (b), supraventricular ectopic (c), fusion (d) and unknown (e) beats.

Figure 4 gives a sample of each beat in the dataset. We note in particular that there are some clear differences between certain beats. For instance, normal and fusion beats could be easily distinguished. However, for different classes of beats, the differences can be more difficult to distinguish. In particular, the pictured normal beat and ventricular ectopic beat are quite similar, except for a noticeable difference in the P waves between each lead.

Table 3 gives an overview of the dataset. We note in particular that it is very imbalanced, with the vast majority of the dataset consisting of normal beats. For the minority of pathological beats, these differ rather significantly in number between the different classes. For this reason, we use an oversampling technique to increase the size of the minority classes.

Table 3: An overview of the types and numbers of beats contained in the dataset.

Beat Type	Size
Normal	92,596
Ventricular Ectopic	7,631
Supraventricular Ectopic	2,779
Fusion	802
Unknown	982
Total:	104,790

Synthetic minority oversampling (SMOTE) is frequently used to oversample minority classes, which relies on generating new points based on k nearest neighbors [42]. However, we make use of localized random affine shadowsampling [43] (LoRAS), which samples k points, calculates their covariance, and draws new points from the multivariate normal distribution described by the covariance. Importantly, LoRAS has been shown to outperform SMOTE [43]; therefore, we make use of it to oversample each class to a size equal to the majority class size.

4.2. Results

Table 4 gives an overview of the results. In particular, we note that the proposed model outperforms recent models in accuracy, precision, specificity, and F1 score. Besides this, it also performs very strongly with respect to recall. Figure 5 shows the loss as the model was trained over 15 epochs. After training concluded, the weights corresponding to the best validation accuracy were restored to the model, before the model was tested using the test data.

Table 4: Summary of model results, compared to recent results using the MIT-BIH Arrhythmia database. If a result is not reported, it is denoted via the ~ symbol.

Study	Accuracy	Precision	Specificity	Recall (Sensitivity)	F1 Score
[28]	97.85%	98.08%	~	96.23%	97.15%
[22]	98.82%	~	94.73%	93.14%	93.52%
[24]	90.80%	~	~	~	~
[27]	97.20%	97.20%	99.30%	**97.20%**	~
[36]	98.68%	~	~	92.16%	92.02%
[37]	96.19%	~	93.39%	96.41%	~
[38]	97.22%	81.87%	98.72%	83.29%	~
Current Study	**99.10%**	**99.09%**	**99.64%**	95.90%	**97.47%**

Figure 6 also displays the classification results as a heatmap. We observe that the model performs well in distinguishing between each class, particularly those with more samples. Nevertheless, the model is able to classify unknown beats with very high accuracy, despite the lower sample size. Future work could focus specifically on improving the accuracy for classes with less representation, since these are where the majority of beats are missed.

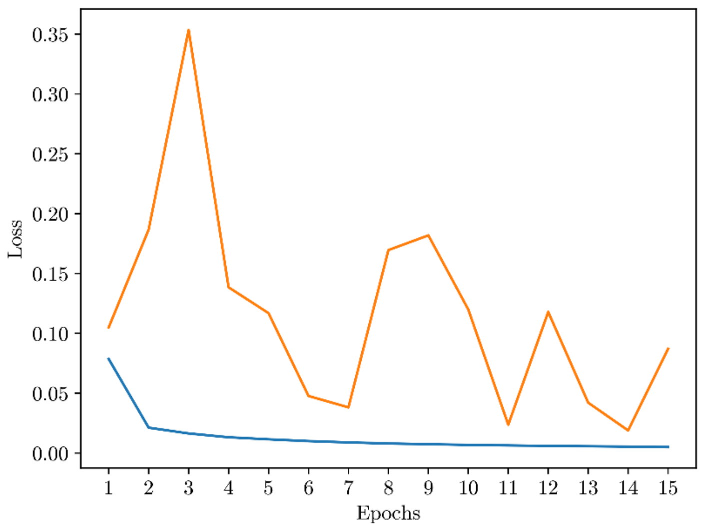

Figure 5: The model loss, with training loss shown in blue, and validation loss in orange.

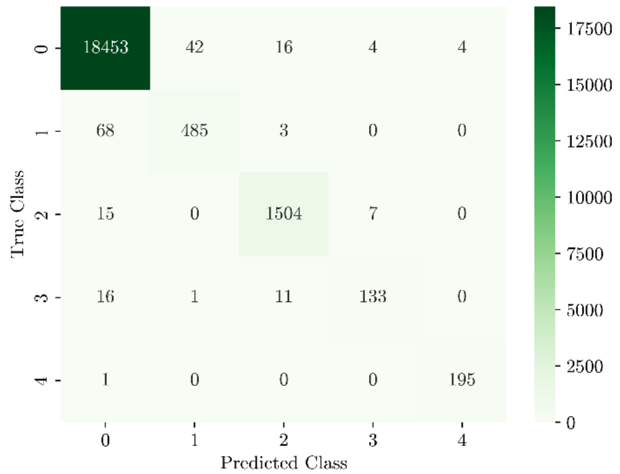

Figure 6: Model classification results in the form of a confusion matrix.

Figure 7 also gives the area under the receiver operating characteristic curve (AUC) for the proposed model, for the diagnosis of normal versus pathological beats. We note that the model achieves an excellent AUC of 0.997 in this case. Furthermore, when diagnosing normal versus pathological beats, the accuracy is 99.20%.

Overall, the proposed model demonstrates excellent results, and outperforms several recent models using the same dataset. We believe the combination of convolutional layers, attention, and LSTM units is able to capture important features of an ECG very well. The model can successfully differentiate between a number of different types of beats, and could be used clinically to diagnose arrhythmias. The beat-by-beat diagnosis could also lend additional explainability to the diagnosis.

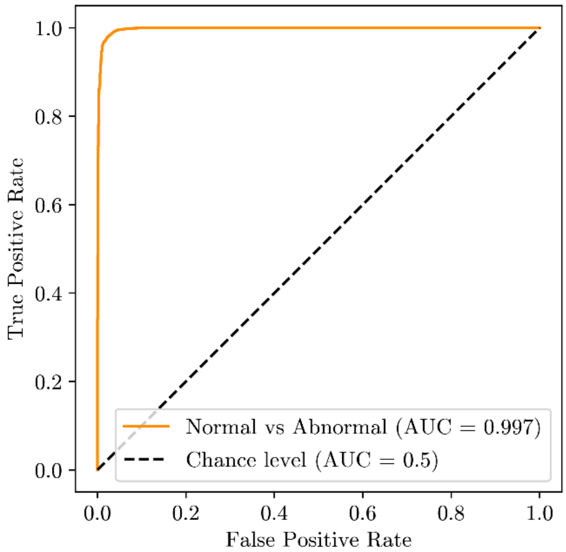

Figure 7: Area under the receiver operating characteristic curve for the proposed model.

4.3. Discussion

Electrocardiograms represent the electrical activity of the human heart. As the cardiac cycle has recognizable features even in a pathological heart, it is reasonable to use convolutional layers to capture these features. Attention can be used to further emphasize or deemphasize the learned features, so that the model can focus on the more important features. However, ECGs are also time-dependent, and different pathologies can present in a time-dependent manner. Therefore, while convolutional layers (and attention) are a natural element to include in a model, they could be accentuated by recurrent layers. We used bidirectional LSTMs after extracting features with convolutional layers so that the extracted features could also be modeled in a time-dependent manner.

The proposed model performs very well, outperforming recent models using the same dataset. Furthermore, the model is relatively lightweight, consisting of only three convolutions, one attention module, and two LSTM units. In addition, the model was trained for only 15 epochs; therefore, in addition to being a lightweight model, the training time is also relatively minimal. The model achieves competitive metrics and could likely be applied in a clinical setting (given segmented beats, which can be achieved algorithmically).

While the results we achieved are promising, they are limited in a sense by the dataset on which they were tested. The MIT-BIH Arrhythmia dataset is unique in that it is labeled beat-by-beat, making it an invaluable resource for the development of automated methods for ECG diagnosis. Being able to label an ECG beat-by-beat in an automated fashion lends a level of explainability to an overall diagnosis from an electrocardiogram, which is of major importance in medical artificial intelligence [44]. Nevertheless, the dataset we used consists of two-lead ECGs recorded from 47 subjects. Therefore, while we expect a model extended to the standard 12-lead ECG would perform quite well in a beat-by-beat analysis, no data exists to substantiate this claim. Furthermore, while the dataset consists of over 100,000 individual beats, they arise from 47 subjects; thus, the intra-class variation between beats may be rather low and not accurately represent the population distribution. The author hopes that this study and similar studies will stimulate the release of larger-scale datasets with beat-by-beat labeling, specifically for 12-lead ECGs drawn from a larger patient sample.

Nevertheless, because the heart functions based on the same principles irrespective of the patient, we hypothesize that the model would generalize fairly well to a new population.

5. Conclusion

In this study, we demonstrated the value of using convolutions combined with convolutional block attention, and bidirectional LSTMs. In particular, extrapolating from the fact that ECGs have important recurrent features that arise in a time-dependent manner, we proposed a model architecture that would be able to capture both of these elements of an ECG. We validated the model using a well-studied dataset, and demonstrated that the proposed model outperforms other state-of-the-art models.

Automated ECG diagnosis could save the time of many doctors, allowing them to focus on more pressing matters. Therefore, the clinical utility of these methods are clear. Models that can diagnose specific beats of an electrocardiogram are particularly valuable, since they provide additional clarity on a diagnosis — instead of outputting a single diagnosis for an entire ECG, they can provide a beat-by-beat analysis of an input.

The present study used a very well-studied database consisting of two-lead ECGs recorded from 47 patients. A limitation of the study is that two-lead ECGs are not as common as the standard 12-lead ECG. To the best of the author's knowledge, this is due to a gap in the available datasets, which we hope will be addressed in further dataset releases. While we used a rather shallow convolutional neural network consisting of three layers, future work could investigate how increasing the depth of the convolutional neural network (in conjunction with attention) affects the results. Furthermore, more advanced sequential methods could be explored, such as using transformers.

Funding Sources

None declared.

Declaration of Competing Interest

None declared.

References

1. WHO. Health topics. cardiovascular diseases. Cardiovascular Diseases, 2019.

2. Vennela Avula, Katherine C Wu, and Richard T Carrick. Clinical applications, methodology, and scientific reporting of electrocardiogram deep- learning models: a systematic review. JACC: Advances, 2(10):100686, 2023.

3. Gabor Kovacs, Alexander Avian, Vasile Foris, Maria Tscherner, Xhylsime Kqiku, Philipp Douschan, Gerhard Bachmaier, Andrea Olschewski, Marco Matucci-Cerinic, and Horst Olschewski. Use of ecg and other simple non- invasive tools to assess pulmonary hypertension. PloS one, 11(12):e0168706, 2016.

4. Hongling Zhu, Cheng Cheng, Hang Yin, Xingyi Li, Ping Zuo, Jia Ding, Fan Lin, Jingyi Wang, Beitong Zhou, Yonge Li, et al. Automatic multilabel electrocardiogram diagnosis of heart rhythm or conduction abnormalities with deep learning: a cohort study. The Lancet Digital Health, 2(7):e348– e357, 2020.

5. Renate B Schnabel, Ladislav Pecen, Daniel Engler, Markus Lucerna, Jean Marc Sellal, Francisco M Ojeda, Raffaele De Caterina, and Paulus Kirchhof. Atrial fibrillation patterns are associated with arrhythmia pro- gression and clinical outcomes. Heart, 104(19):1608–1614, 2018.

6. Larissa Montenegro, Mariana Abreu, Ana Fred, and Jose M Machado. Human-assisted vs. deep learning feature extraction: an evaluation of ecg features extraction methods for arrhythmia classification using machine learning. Applied Sciences, 12(15):7404, 2022.

7. VSR Kumari and P Rajesh Kumar. Cardiac arrhythmia prediction using improved multilayer perceptron neural network. International Journal of Electronics, Communication & Instrumentation Engineering Research and Development (IJECIERD), 3(4):73–80, 2013.

8. Gaurav Kumar, Urja Pawar, and Ruairi O'Reilly. Arrhythmia detection in ecg signals using a multilayer perceptron network. AICS, 2563:353–364, 2019.

9. U Rajendra Acharya, Hamido Fujita, Shu Lih Oh, Yuki Hagiwara, Jen Hong Tan, and Muhammad Adam. Application of deep convolutional neural network for automated detection of myocardial infarction using ecg signals. Information Sciences, 415:190–198, 2017.

10. Roberta Avanzato and Francesco Beritelli. Automatic ecg diagnosis using convolutional neural network. Electronics, 9(6):951, 2020.

11. Elif Derya U¨ beyli. Recurrent neural networks employing lyapunov exponents for analysis of ecg signals. Expert systems with applications, 37(2):1192–1199, 2010.

12. Shraddha Singh, Saroj Kumar Pandey, Urja Pawar, and Rekh Ram Janghel. Classification of ecg arrhythmia using recurrent neural networks. Procedia computer science, 132:1290–1297, 2018.

13. George B Moody and Roger G Mark. The impact of the mit-bih arrhythmia database. IEEE engineering in medicine and biology magazine, 20(3):45– 50, 2001.

14. Elif Izci, Mehmet Akif Ozdemir, Murside Degirmenci, and Aydin Akan. Cardiac arrhythmia detection from 2d ecg images by using deep learning technique. In 2019 Medical Technologies Congress (TIPTEKNO), pages 1–4, 2019.

15. Tae Joon Jun, Hoang Minh Nguyen, Daeyoun Kang, Dohyeun Kim, Daeyoung Kim, and Young-Hak Kim. Ecg arrhythmia classification using a 2-d convolutional neural network. arXiv preprint arXiv:1804.06812, 2018.

16. Rajkumar, M. Ganesan, and R. Lavanya. Arrhythmia classification on ecg using deep learning. In 2019 5th International Conference on Advanced Computing and Communication Systems (ICACCS), pages 365–369, 2019.

17. Mengze Wu, Yongdi Lu, Wenli Yang, and Shen Yuong Wong. A study on arrhythmia via ecg signal classification using the convolutional neural network. Frontiers in computational neuroscience, 14:564015, 2021.

18. Bhekumuzi M Mathunjwa, Yin-Tsong Lin, Chien-Hung Lin, Maysam F Abbod, and Jiann-Shing Shieh. Ecg arrhythmia classification by using a recurrence plot and convolutional neural network. Biomedical Signal Processing and Control, 64:102262, 2021.

19. Sena Ya˘gmur S¸EN and Nalan O¨ ZKURT. Ecg arrhythmia classification by using convolutional neural network and spectrogram. In 2019 Innovations in Intelligent Systems and Applications Conference (ASYU), pages 1–6. IEEE, 2019.

20. Jingshan Huang, Binqiang Chen, Bin Yao, and Wangpeng He. Ecg arrhythmia classification using stft-based spectrogram and convolutional neural network. IEEE access, 7:92871–92880, 2019.

21. Eko Ihsanto, Kalamullah Ramli, Dodi Sudiana, and Teddy Surya Gunawan. An efficient algorithm for cardiac arrhythmia classification using ensemble of depthwise separable convolutional neural networks. Applied Sciences, 10(2):483, 2020.

22. Sumanta Kuila, Namrata Dhanda, and Subhankar Joardar. Ecg signal classification to detect heart arrhythmia using elm and cnn. Multimedia Tools and Applications, 82(19):29857–29881, 2023.

23. Hui Yang and Zhiqiang Wei. Arrhythmia recognition and classification using combined parametric and visual pattern features of ecg morphology. IEEE Access, 8:47103–47117, 2020.

24. Minh Cao, Tianqi Zhao, Yanxun Li, Wenhao Zhang, Peyman Benharash, and Ramin Ramezani. Ecg heartbeat classification using deep transfer learning with convolutional neural network and stft technique. In Journal of Physics: Conference Series, volume 2547, page 012031. IOP Publishing, 2023.

25. Liang-Hung Wang, Yan-Ting Yu, Wei Liu, Lu Xu, Chao-Xin Xie, Tao Yang, I-Chun Kuo, Xin-Kang Wang, Jie Gao, Pao-Cheng Huang, et al. Three- heartbeat multilead ecg recognition method for arrhythmia classification. IEEE Access, 10:44046–44061, 2022.

26. Ehab Essa and Xianghua Xie. An ensemble of deep learning-based multi- model for ecg heartbeats arrhythmia classification. ieee access, 9:103452– 103464, 2021.

27. Wei Zeng, Bo Su, Yang Chen, and Chengzhi Yuan. Arrhythmia detection using tqwt, ceemd and deep cnn-lstm neural networks with ecg signals. Multimedia Tools and Applications, 82(19):29913–29941, 2023.

28. Ugur Demiroglu, Bilal Senol, and Radek Matusu. A fused electrocardiography arrhythmia detection method. Multimedia Tools and Applications, pages 1–33, 2023.

29. Sanghyun Woo, Jongchan Park, Joon-Young Lee, and In So Kweon. Cbam: Convolutional block attention module. In Proceedings of the European conference on computer vision (ECCV), pages 3–19, 2018.

30. Yana Luo and Zhongsheng Wang. An improved resnet algorithm based on cbam. In 2021 International Conference on Computer Network, Electronic and Automation (ICCNEA), pages 121–125. IEEE, 2021.

31. Wenhao Cao, Zhuoyu Feng, Dongyao Zhang, and Yisiyuan Huang. Facial expression recognition via a cbam embedded network. Procedia Computer Science, 174:463–477, 2020.

32. Zhongye Chen, Yijun Wang, and Zhongyan Song. Classification of motor imagery electroencephalography signals based on image processing method. Sensors, 21(14):4646, 2021.

33. Mohamed Hammad, Pawe-l P-lawiak, Kuanquan Wang, and Udyavara Ra- jendra Acharya. Resnet-attention model for human authentication using ecg signals. Expert Systems, 38(6):e12547, 2021.

34. Wesley Chorney, Haifeng Wang, Lu He, Seunghan Lee, and Lir-Wan Fan. Convolutional block attention autoencoder for denoising electrocardiograms. Biomedical Signal Processing and Control, 86:105242, 2023.

35. Wesley Chorney, Haifeng Wang, and Lir-Wan Fan. Attentioncovidnet: Efficient ecg-based diagnosis of covid-19. Computers in Biology and Medicine, page 107743, 2023.

36. Hongfu Xie, Hui Liu, Shuwang Zhou, Tianlei Gao, and Minglei Shu. A lightweight 2-d cnn model with dual attention mechanism for heartbeat classification. Applied Intelligence, 53(13):17178–17193, 2023.

37. Muhammad Zubair, Sungpil Woo, Sunhwan Lim, and Daeyoung Kim. Deep representation learning with sample generation and augmented attention module for imbalanced ecg classification. IEEE Journal of Biomedical and Health Informatics, 2023.

38. Zhimin Xu, Mujun Zang, Tong Liu, Shusen Zhou, Chanjuan Liu, and Qingjun Wang. Multi-modality multi-attention network for ventricular arrhythmia classification. In Proceedings of the 2023 3rd International Conference on Bioinformatics and Intelligent Computing, pages 331–336, 2023.

39. Xue Xu, Sohyun Jeong, and Jianqiang Li. Interpretation of electrocardiogram (ecg) rhythm by combined cnn and bilstm. Ieee Access, 8:125380– 125388, 2020.

40. Parul Madan, Vijay Singh, Devesh Pratap Singh, Manoj Diwakar, Bhaskar Pant, and Avadh Kishor. A hybrid deep learning approach for ecg-based arrhythmia classification. Bioengineering, 9(4):152, 2022.

41. Brian T Fengler, William J Brady, and Claire U Plautz. Atrial fibrillation in the wolff-parkinson-white syndrome: Ecg recognition and treatment in the ed. The American journal of emergency medicine, 25(5):576–583, 2007.

42. Nitesh V Chawla, Kevin W Bowyer, Lawrence O Hall, and W Philip Kegelmeyer. Smote: synthetic minority over-sampling technique. Jour- nal of artificial intelligence research, 16:321–357, 2002.

43. Saptarshi Bej, Narek Davtyan, Markus Wolfien, Mariam Nassar, and Olaf Wolkenhauer. Loras: An oversampling approach for imbalanced datasets. Machine Learning, 110:279–301, 2021.

44. Andreas Holzinger, Georg Langs, Helmut Denk, Kurt Zatloukal, and Heimo Mu¨ller. Causability and explainability of artificial intelligence in medicine. Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery, 9(4):e1312, 2019.