AN ECG-BASED MODEL FOR HEART ARRHYTHMIA CLASSIFICATION USING DEEP LEARNING

Supported by tetfund

Okpor James [1], Olajide Blessing Olajide [2], Tsokwa Idibore Danasabe [3], Ehizuenlen Ejordenore Prudence [4]

[1, 2] Department of Computer Engineering, Federal University Wukari, Nigeria.

[3] Department of Computer Science, Federal University Wukari, Nigeria.

[4] Department of Computer Engineering, University of Benin, Benin City, Nigeria.

Corresponding email: okpor2004@gmail.com

Received: May 09, 2024 Accepted: June 25, 2024

Abstract: Heart Arrhythmia is a cardiovascular disease that Electrocardiogram (ECG) signal records can be used to identify and classify into categories including; normal heartbeats, supra-ventricular premature, ventricular escape, fusion of ventricular and normal, and unclassified heartbeats. However, interpreting these signals accurately remains intricate for medical professionals that examine patient ECG report to make diagnosis and to ascertain the risk level. To aid the medical professionals in making heart arrhythmia diagnosis and classification, this study leverages the prowess of machine and deep learning techniques, which have showcased their efficacy in medical domains by automating processes and enhancing result accuracy. On this pedestal, this research developed two diagnostic model using Convolutional Neural Network (CNN) and Long Short-Term Memory (LSTM). The two models were thereafter hybridized (LSTM-CNN) and evaluated. The three models were compared and it was revealed that the CNN and the hybrid LSTM-CNN emerge as notable performers, exhibiting remarkable accuracy results of up to 97% respectively compared to LSTM that showed 90%. This research signifies a significant stride toward leveraging advanced technologies for more effective and efficient cardiovascular disorder diagnosis and management.

Keywords: Heart, Arrhythmia, ECG, CNN, LSTM, Diagnosis.

Introduction

Heart arrhythmia is a cardiovascular disease (CVD) with the pathological features of irregular breathing which depicts abnormal rhythms of the electrical initiation and conduction of excitation waves in the heart (Amini *et al.,* 2021). Heart arrhythmia is of different variant such as Normal sinus rhythm (NSR), Premature (extra) beats, heart block, sinus node dysfunction, supraventricular arrhythmias, and ventricular arrhythmias, among others, some of these variants are harmless and some impacts the patient's health badly (Alday *et al.,* 2021). Therefore, timely identification of the risk type of the heart arrhythmia disease is imperative to averting and managing the attendant consequences. This disease is among the recognized leading causes of stroke and death if discovered late (Ko *et al.,* 2012). An electrocardiogram (ECG) is used to diagnose heart arrhythmia symptoms, it works by representing the expression of the myocardium's electrical activity perceived from the body on a monitor or paper. The ECG make use of electrodes which are attached to the chest and limbs of the patient to records the electrical impulse that makes the heart to beat **(**Xinwen *et al.,* 2021). Medical expert uses the information presented by the ECG test result of a patient to classify the abnormality symptoms and identify the risk level. Arrhythmia diagnosis and classification are commonly achieved through non-invasive manual interpretation of the electrocardiogram (ECG) by a medical specialist. But conventional ECG-based arrhythmia identification is prone to human error, hence in recent times, computer-aided ECG diagnosis based on deep learning and machine learning has emerged as a popular topic of research (Siontis *et al.,* 2021). Rapidly analyzing vast volumes of data, machine learning can spot patterns that are invisible to the human eye. Comparing this to conventional approaches can result in a more informed and accurate diagnosis of diseases. Also, deep learning algorithms are equipped with advanced training techniques such as stochastic gradient descent, optimization algorithms leading to better classification performance. They can also infer from large data to understand robust and generalizable patterns (Siontis *et al.,* 2021).

Over the past decade, several researchers have introduced a distinct number of automatic analysis algorithms to help decision making and diagnosis, especially electro-cardiogram disease identification (Olajide and Andrew, 2023). For instance, (Liu and Guo, 2019) employed the Support Vector Machine (SVM), (Bin-Heyat *et al.,* 2022) further employed the Decision Tree Method. Even though these techniques rationally combine feature extraction and classifier to increase the accuracy of ECG signal classification of heart arrhythmia, they still have some common flaws because they heavily rely on experts to design and extract the characteristics of ECG signals, leaving a gap where other important aspects of the original signal are overlooked. Moreover, the artificial definition of many illness features may vary somewhat, which could lead to limitations or generalizations in the model's capacity to identify heart diseases (Yao *et al.,* 2020); Olajide *et al*., 2020). An interesting problem is the continuous increase of the dataset features dimension making it more difficult to identify the relevant attributes (Bin-Heyat *et al*., 2022). Hence, the choice of model parameters increasingly broadens the complexity of selecting the best features. At the same time, as the feature dimension increases, the choice of model parameters has also become more difficult (Madugu *et al.,* 2023). These factors led to the investigation of the classification potentials of two deep learning algorithms in order to create a diagnostic model that can differentiate between the many heart arrhythmia disease categories, assisting professionals in making an accurate and timely diagnosis and enhancing previous research in this area. The two methods are called Long Short-Term Memory (LSTM) neural network and Convolutional Neural Network (CNN). The MIT-BIH Arrhythmia Electro Cardiogram dataset, which was downloaded from the Kaggle data source and comprises five classes of ECG signals, was utilized to create the model. An effective deep learning neural network for image and video analysis is called a convolutional neural network (CNN). CNNs extract information from pictures and videos using a sequence of convolution and pooling layers, and then utilize these features to identify or categorize objects or scenes (Madugu *et al.,* 2023). The pseudocode for the CNN algorithm is shown in Figure 1.

CNN Pseudocode

Input: $x = [x_1, x_2, \ldots\ x_n]$
***Function** CNN_MODEL(x)*
weights ← DEFINE WEIGHT
biases ← DEFINE BIASES
FOR** x to n (no of layers) **DO
*x ← **RESHAPE** (x, shape ←(x-dim, x-length, ...)*
*conv2 ← **RELU_ACTIVATION_FUNC** (conv3D (x, weight (0) + biases (0))*
*conv2 ← **DROPOUT** (0.5)*
*conv2 ← **MAX_POOL3D** (conv2)*
END FOR
*fully-connect ← **RESHAPE** (conv2 **INVERSE** (weights))*
*fully-connect ← **RELU_ACTIVATION_FUNC** (fully-connect + weight + biases)*
*fully-connect ← **DROPOUT** (0.5)*
*output ← fully-connect * weights + biases*
***RETURN** output*
END FUNCTION

Figure 1: CNN pseudocode

One kind of Recurrent Neural Network (RNN) that can learn and retain information over time is called Long Short-Term Memory (LSTM). An LSTM is a sort of RNN that has a memory cell that allows it to store and retrieve information over time. Traditional RNNs, on the other hand, have limited memory and can only hold data for a limited amount of time. The three fundamental parts of an LSTM are the input gate, forget gate, and output gate. The data flow into and out of the memory cell is controlled by these gates. The output gate uses the data stored in the memory cell for the present task, the forget gate removes unnecessary information, and the input gate adds new information to the memory cell. This makes it an effective algorithm for time series forecasting, speech recognition, and natural language processing, among other applications. Recently, merging of LSTMs with other deep learning techniques such as CNNs for image and video processing has become open research to study (Okpor et al., 2023). The LSTM algorithm simulates the following equations:

$$i_t = \sigma_s(w_{xi}\, x_{t-1} + w_{hi} h_{t-1} + b_i) \quad (1)$$
$$f_t = \sigma_s(w_{xf}\, X_t + w_{hf} h_{t-1} + b_f) \quad (2)$$
$$o_t = \sigma_s(w_{xo}\, x_t + w_{ho} h_{t-1} + b_o) \quad (3)$$
$$c_t^* = \tanh(w_{hc}\, h_{t-1} + w_{xc} x_t + b_c) \quad (4)$$
$$c_t = f_t c_{t-1} + i_t c_t^* \quad (5)$$
$$h_t = o_t \tanh(c_t) \quad (6)$$

Where the activations of the input, forget, and output gates at time-step t are denoted by the symbols $a\,t$, $i\,t$, and $o\,t$. $c\,t$ represents the protected cell state at time step t, while h_t represents the activation that will be transferred to the subsequent layer. Subscripts of its transform are used to explicitly indicate the notations of the weight matrices w (…). The bias vector of every gate is indicated by b (.). The matrices that convert the inputs x_t and h_{t-1}, respectively, to the forget gate dimension are w_{xf} and w_{hf}. Similarly, w_{xi} and w_{hi} convert the inputs and hidden state to the input gate dimension, and so forth. There is a single matrix for each of the gates w_z, w_i w_c, and $w\,o$ due to the simplified notation that combines the input and hidden state into a single matrix.

In data science, the feature extraction techniques are key to identifying the features that represent arrhythmia signals on an ECG and impact the resultant prediction from developed deep learning models. The ECG parameters used for this experiment are Wavelet-Based Functions (WBF). This makes the use Daubechies feature selection techniques applicable. Daubechies are wavelets with a number of vanishing moments (N), and the minimum filter size is 2N. The simplest and the oldest function in Daubechies is the Haar wavelet (DB_1). It has a value of either 1 for $0 \leq x \leq 0.5$ or −1 for $0.5 \leq x \leq 1$; otherwise, it will be 0 (Mandala *et al.,* 2023).

Individual studies have showcased the strengths of CNNs and LSTMs in ECG analysis, an opportunity exists to combine these techniques in a hybrid model. While LSTMs are experts at capturing temporal connections within time-series data, such as ECG signals, CNNs are excellent at identifying spatial patterns from sequential data. Combining these algorithms in a hybrid model could lead to synergistic advantages, addressing the complexities of ECG data. Hence, this study is aimed to achieve a more reliable heart arrhythmia classification using the hybrid of the aforementioned algorithms.

Related Works

Nazrul *et al.,* (2020) employed machine learning techniques such as linear discriminant analysis (LDA), linear and quadratic support vector machines (SVM), decision trees (DN), k-nearest neighbors (KN), and artificial neural networks (ANN) to predict cardiovascular disease (CVD) based on electrocardiograms. ANN performs well compared to other algorithms with 90% specificity, 90% sensitivity, and 90% accuracy. In order to automatically classify four different types of ECG data, Liu and Kim (2018) created a deep learning (DL) model-based system. The ECG signals in four different classes correspond to labeled classes in the Physio Net open-source ECG dataset. There are four categories: arrhythmias, sudden death, supraventricular arrhythmias, and normal. Among the four classes, the results showed that their ECG waveform has the highest predictive of sudden death with 100% of the samples correctly classified (100 accuracies). The lowest forecast, on the other hand, was for typical sinus ECG waveforms. Marcel et al. (2021) used single-lead electrocardiogram (ECG) monitoring to construct and test machine learning models for the detection of sleep apnea. The two-machine learning algorithm used are two-way Gated recurrent units (GRU) and LSTM. Their findings demonstrated that in predicting sleep apnea in patients, the two-way GRU model and the LSTM model both reached an accuracy of 97.1 with excellent sensitivity and specificity.

Zhang *et al*., (2021) used an Inception-ResNet-v2 network in conjunction with the recurrence plot (RP) to develop a method for classifying cardiac arrhythmias (CAs). The 2018 China Physiological Signal Challenge (CPSC) dataset was also utilized in their work. The best leads (lead II and lead aVR) were chosen for implementation, and the RP technique was then used to convert 1D ECG segments into 2D texture images. The time waveform of the ECG recordings and time-frequency pictures based on continuous wavelet transform (CWT) were two conventional ECG transform into 2D image approaches that were compared with their methodology. According to experimental results, their suggested solution used only two leads from the original 12-lead ECG data, yet it still scored the greatest average F1-score of 0.844. Rahman *et al.*, (2022) used machine learning to perform feature extraction of human electrocardiogram (ECG) signals. The authors take into account different approaches to performing human classification of ECG signals. Undefined signals generated by ECG were corrected by transforming them into smaller feature sets using a classification process. The Physio-Net records of 162 distinct patients with congestive heart failure (CHF), arrhythmia (ARR), and normal sinus rhythm (NSR) provided the data needed to perform feature extraction. The goal of this feature extraction model implementation was to use machine learning to train and test a classifier that would distinguish between patients with arrhythmia (ARR), patients with congestive heart failure (CHF), and those with normal sinus rhythms (NSR). The ECG signal categorization process uses wavelets and Support Vector Machines (SVM) for feature extraction. The results obtained from this demonstration showed that feature extraction significantly reduces the size of large data sets, as it also captures differences between CHF, ARR, and NSR classes. From the results obtained, it was concluded from cross-validation and the performance his SVM classifier on the test set that the SVM classifier exhibits an accuracy of 96.51%. Belal *et al.,* (2022) use machine learning classifiers to create an automatic system for researchers to detect mental stress based on electrocardiogram (ECG) signals from twenty selected male researchers who are free of mental and cardiac illnesses. Five machine learning classifiers were used to categorize the intra-subject (mental stress and normal) and inter-subject (mental stress and normal), namely Decision Tree (DT), Naive Bayes (NB), Random Forest (RF), and Logistic Regression (LR). Following its implementation and evaluation, Decision Trees outperformed all other models in the intra-subject classification with respect to recall (93.30%), specificity (96.70%), precision (94.40%), accuracy (93.30%), and F1 (93.50%). Furthermore, the system's classification accuracy while utilizing a DT classifier for inter-subject classification is 94.10%. The findings indicate that the DT classifier attained an average accuracy of 94.10% for inter-subject classification and the greatest accuracy of 93.30% for intra-subject classification.
Ghasem et al. (2022) classified electrocardiogram signals on students ranging in age from 6 to 18 using manual characteristics and machine learning algorithms. The Z-score was used to normalize the gathered features, and then the student's t-test and chi-square test were used to determine their relevance. In order to further implement the model, eight different classifiers were implemented: eXtreme Gradient Boosting (XGB), Support Vector Machine (SVM), Quadratic Discriminant Analysis (QDA); Random Forest (RF), Decision Tree (DT), K-nearest neighbors (KNN), Logistic Regression (LR), and Stack Learning (SL). The Boruta algorithm was applied to the data collected for feature selection. The research findings indicate that manual ECG feature measurement yielded superior results compared to automated measurement. However, in certain models, automated measurement demonstrated potential in differentiating between normal and abnormal instances. A machine-learning framework and a bag-of-features derived from electrocardiogram spectrograms were used by Cheng-Yu *et al.,* (2022) to construct a sleep apnea classification algorithm. The National Cheng Kung University Hospital Sleep Center Apnea Database (NCKUHSCAD) and the Physionet Apnea-ECG Database (PAED) provided the two datasets that were used in the development process. While PAED was utilized to validate the suggested technique, NCKUHSCAD was utilized to track variations in the ECG spectrogram's apnea and normal periods.

The aforementioned research trends have recorded remarkable accuracy, however, there are problems common to their studies such as; the model still relying on experts to extract the characteristics of ECG signals creating an avenue where other vital attributes from the original signal are neglected, and continuous increase of the dataset features dimension making it more difficult to identify the relevant attributes. To remedy this drawback, this work presents two deep neural models namely the convolutional neural network and the long short-term memory neural network for the effective detection of heart arrhythmia diseases.

Methodology

To develop the ECG prediction model for heart arrhythmia disease, this study performed feature selection using Daubechies feature selection techniques, apply CNN and LSTM network to the MIT-BIH Arrhythmia Electro Cardiogram dataset, and performed a comparative analysis of the used algorithms. The proposed electro-cardiogram prediction models for heart arrhythmia were implemented using the Anaconda programming environment via the Python programming language on a Windows operating system with a dual-core Intel Core 15 processor and 4 GB RAM. The integrated model comprises the CNN and LSTM. Given that the focus of the subject is machine learning with a deep neural component, the TensorFlow API is one of the Python packages used. Aside from NumPy modules for numerical dimensional manipulation, other Python packages used are pandas for reading ECG datasets and Matplotlib for data visualization.

MIT-BIH Arrhythmia Electro Cardiogram dataset Data Preprocessing

The MIT-BIH Arrhythmia ECG dataset includes signals that match the electrocardiogram (ECG) heartbeat morphologies for both normal and arrhythmia and myocardial infarction-affected cases. It contains 875,544 electrocardiogram records categorized into 5 classes with a sampling frequency of 125Hz (Mohammad *et al.*, 2019). The classes are represented by the letters "N" for normal heartbeats, "S" for supraventricular premature heartbeats, "V" for ventricular escape, "F" for ventricular and normal heartbeats combined, and "Q" for heartbeats that are not classified. The noise in the acquired dataset was removed using the Daubechies Wavelet (DWT) filter. Using wavelet type and thresholding, the DWT was able to separate the patient's ECG signals into approximation sections and details. At certain scales, the details include largely inconsequential noise, which can be eliminated or blank out without compromising the signal. Subsequently, the reconstruction filters are computed by the wavelet class by applying its symmetry properties. The basic steps are captured in the algorithm in Figure 2.

Daubechies Wavelet (DWT) filter.
Step 1: Use the DWT to first decompose the signal. Step 2: Apply thresholding to filter the signal in the wavelet space. Step 3: Use the inverse DWT to invert the filtered signal in order to recreate the original, now filtered signal.

Figure 2: Daubechies Wavelet (DWT) filter

After completing the previously described stage, the dataset is split into two parts: 70% training and 30% test. The training data is used to train the two deep learning network models, while the test data is used to verify the model's accuracy.

Implementing CNN Algorithm for Heart Arrythmia Classification

The CNN model for cardiac arrhythmia consists of the following layers: normalization, SoftMax, fully connected, pooling, and convolutional. The output of the trained network is provided by the activation function SoftMax. The approach employs the rectified linear unit (ReLU) as the default activation function for all activation tools in all convolutional layers, with the exception of the output layers. The max pooling layer spatially resizes its forms while operating independently for every row and column of the input layer. The preprocessed dataset is read, the labels are created, the dependent variables are encoded, and the dataset is divided into training and test portions. The loss function and category cross entropy are then calculated, and a gradient descent optimizer is used to lower the cost function. The subsequent phase includes implementing the model via four max convolutional and pooling layers that integrates a vertical filter to convolve the electro-cardiogram signals using a max pooling layer for each convolutional layer with a stride size of 3 x 3. These filters enable the exploitation of the spatial locality of the signals by enforcing a local connectivity pattern between neurons. Therefore, the suggested convolutional neural network design consists of a layer, completely linked layer, and a pooling two-dimensional (2D) layer. The filter slides over the 1D sequence data with the convolutional 2D layer, extracting the best features. Each filter's retrieved features are combined to create a new feature set known as a feature map for the concealed layers. Therefore, using a non-linear activation function, such as SoftMax for the output layer's activation function and Rectified Linear Unit (RELU) for the input and hidden layer, a hyperparameter tuning method is used to determine the length and number of filters. Adjusted linear unit as a non-linear activation as aforementioned reduces the vanishing gradient and error state. It has been found that the rectified linear unit by default is more effective and can speed up the learning process in its entirety. The rectified linear unit is denoted mathematically as follows:

$$softmax(z_i) = \frac{\exp(z_i)}{\sum_j \exp(z_i)} \quad (7)$$

In this case, the values from the output layer's neurons are represented by the z. As a non-linear function, the exponential is used. SoftMax was proposed for the activation function as it is meant for case of multiclass classification problem, as in the case of this study with multi-lead electro-cardiogram labels. The following equation is used to calculate the convolution operation:

$$x^n = \sum_{k=0}^{N-1} y_k f_{n-k} \quad (8)$$

where the signal, filter, and number of elements in y are represented by the variables y, f, and N, respectively. x is a representation of the output vector.

Implementing LSTM Algorithm for Heart Arrythmia Classification

The gates of the LSTM are implemented using the sigmoid and rectified linear unit activation function. As far as algorithms are concerned, the gates are just like weighting the input data. The values of the variables that are involved in triggering the sigmoid nonlinearity determine the weights, which are situated within the [0,1] range. In other words, contextual factors are taken into account when weighing information. As seen in Figure 3, the vector s, also referred to as the cell or unit state, and the vector h, sometimes referred to as the hidden variables vector, are the two sets of variables around which the suggested LSTM is constructed. In other words, this fundamental unit will be successively concatenated to create the network. After receiving s_(n-1) and h_(n-1) from the previous stage, the unit corresponding to time n passes s_n and h_n to the subsequent one. This is in addition to the input vector.

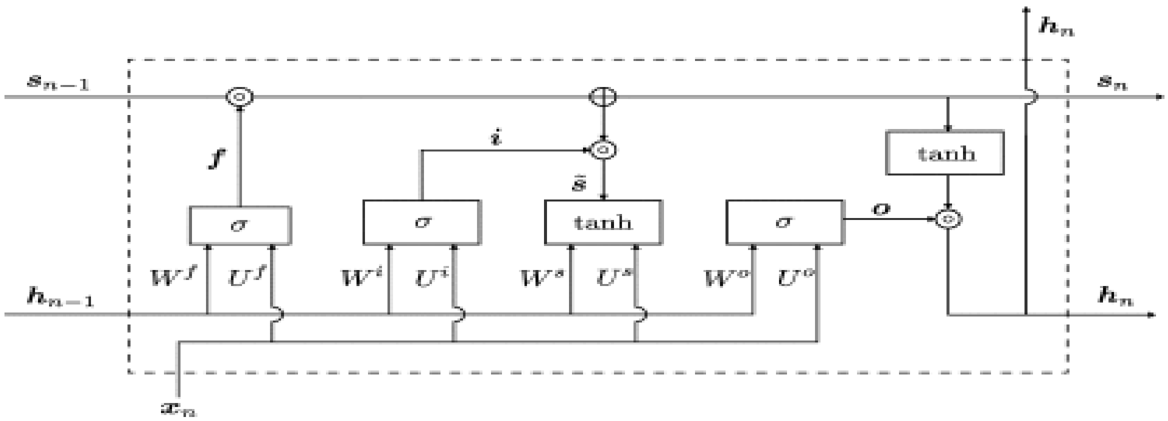

Figure 3: Long Short-Term Memory (LSTM)

The network's hidden activation layers can be safeguarded by putting gates between each of its transaction points and the remainder of its layer, which solves the vanishing gradient issue in specifics. The protected concealed activation is referred to as the cell state. The forget, input, and output gates of the LSTM are responsible for guarding the state of the cell. The forget gate thus operates upon the cell first during information forwarding, deciding which of the cell's activations are forgotten and by how much. This can be done by multiplying each element of the cell by the vector $f_t \in (0,1)^{mh}$. If the forget gate produces a value that is almost equal to zero, the corresponding element in the cell state is cleared and set to zero; on the other hand, if the value is nearly equal to one, the cell will fully retain the value of that element. The input gate is the next gate to operate on the cell; it chooses how much new data will be added to the protected state. This occurs simultaneously with the computation of a novel candidate cell state, c_t^*. Similar to the forget gate, the input gate $i_t \in (0,1)^{mh}$ is multiplied by the candidate state c_t^* and added to the cell state. This stops the cell state from being added unnecessarily.

Finally, the output gate, which is the last and most crucial part of the long short memory for the backwards pass (backpropagation) is present. The output gate chooses which portions of the cell state should be included in the network's output or which parts need to be pushed forward.

Evaluation Metric

To evaluate the performance of the selected classifiers. This study used standard evaluation metrics such as accuracy, precision, recall, and F1-score. These are based on True Positive (TP), True Negative (TN), False Positive (FP), and False Negative (FN). Accuracy, precision and recall are represented by equation 9, 10 and 11.

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \quad (9)$$

$$precision = \frac{TP}{TP+FP} \quad (10)$$

$$Recall = \frac{TP}{TP+FN} \quad (11)$$

Result and Discussions

The evaluation metrics report is shown in Figures 4 to 7 and the metrics encompass the precision, recall, and Fl-score evaluation metrics for the respective five classes of the electrocardiogram signal which are:

i. 0 - “N” for normal heartbeats
ii. 1 - “S” for supra-ventricular premature
iii. 2 - “V” for ventricular escape
iv. 3 - “F” for the fusion of ventricular and normal
v. 4 - “Q” for unclassified heartbeats

Taking into consideration the Convolutional Neural Network performance in relation to the five classes of the electrocardiogram. The CNN precision result for the label N, S, V, F, and Q are 0.84, 0.97, 0.98, 0.93, and 0.94. The CNN recall results for the N, S, V, F, and Q classes are 0.64, 0.99, 0.94, 0.59, and 0.83 respectively, whereas the f1-score for the classes are 0.73, 0.98, 0.96, 0.72, and 0.88. In correspondence with the CNN accuracy score of 0.97, the classification report metric confirms an accuracy of 0.97 as obtained by the CNN algorithm.

```
# cnn_model evaluation
cnn_predit = cnn_model.predict(X_val)
cnn_predit = np.argmax(cnn_predit, axis=1)
print("CNN Classification Report")
print(classification_report(y_val, cnn_predit))
```

```
274/274 [==============================] - 1s 4ms/step
CNN Classification Report
              precision    recall  f1-score   support

           0       0.84      0.64      0.73        64
           1       0.97      0.99      0.98      7248
           2       0.98      0.94      0.96       643
           3       0.93      0.59      0.72       222
           4       0.94      0.83      0.88       579

    accuracy                           0.97      8756
   macro avg       0.93      0.80      0.85      8756
weighted avg       0.97      0.97      0.96      8756
```

Figure 4: Classification Metric for the developed CNN Model for Heart Arrhythmia Diagnosis

Considering the Long Short-Term Memory (LSTM) in correspondence to the five classes of the electrocardiogram. The LSTM model precision result for the signal label N, S, V, F, and Q are 0.00, 0.90, 0.95, 0.87, and 0.73. The LSTM recall for the N, S, V, F, and Q classes are 0.00, 1.00, 0.67, 0.07, and 0.36 respectively, whereas the f1-score for the classes is 0.00, 0.95, 0.78, 0.16, and 0.48. In conformity with the LSTM accuracy, the classification report metric confirms an accuracy of 0.90 as obtained by the LSTM algorithm. The result of the LSTM classification result is seen in Figure 5

```
# lstm_model evaluation
lstm_pred = lstm_model.predict(X_val)
lstm_pred = np.argmax(lstm_pred, axis=1)
print("LSTM Classification Report")
print(classification_report(y_val, lstm_pred))
```

```
274/274 [==============================] - 4s 11ms/step
LSTM Classification Report
              precision    recall  f1-score   support

           0       0.00      0.00      0.00        64
           1       0.90      1.00      0.95      7248
           2       0.95      0.67      0.78       643
           3       0.87      0.09      0.16       222
           4       0.73      0.36      0.48       579

    accuracy                           0.90      8756
   macro avg       0.69      0.42      0.47      8756
weighted avg       0.89      0.90      0.88      8756
```

Figure 5: Classification Metric for the developed LSTM Model for Heart Arrhythmia Diagnosis

Considering that the LSTM and CNN algorithm were hybridized. The hybrid model (figure 6) precision results for the signal class labels N, S, V, F, and Q are 0.93, 0.98, 0.98, 0.90, and 0.88. The hybrid recalls result for the N, S, V, F and Q electrocardiogram signal classes are 0.59, 0.99, 0.97, 0.65, and 0.91 respectively, whereas the f1-score for the classes is 0.72, 0.98, 0.98, 0.76, and 0.89. In conformity with the result obtained by the hybrid CNN-LSTM, the classification report metric confirms an accuracy of 0.97 as obtained by the hybrid algorithm.

```
# Hybrid  evaluation
hyb_pred = model.predict(X_val)
hyb_pred = np.argmax(hyb_pred, axis=1)
print("CNN-LSTM Classification Report")
print(classification_report(y_val, hyb_pred))
```

```
274/274 [==============================] - 2s 7ms/step
CNN-LSTM Classification Report
              precision    recall  f1-score   support

           0       0.93      0.59      0.72        64
           1       0.98      0.99      0.98      7248
           2       0.98      0.97      0.98       643
           3       0.90      0.65      0.76       222
           4       0.88      0.91      0.89       579

    accuracy                           0.97      8756
   macro avg       0.93      0.82      0.87      8756
weighted avg       0.97      0.97      0.97      8756
```

Figure 6: Classification Metric for the developed Hybrid (CNN-LSTM) Model for Heart Arrythmia Diagnosis

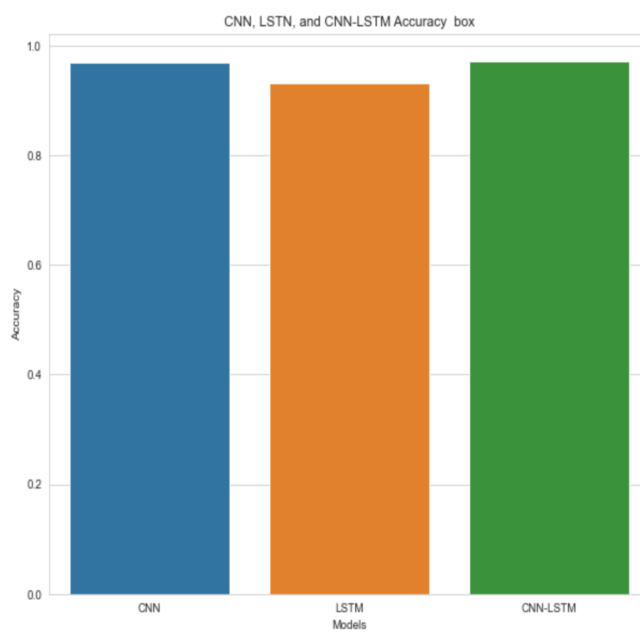

Figure 7: Accuracy of the three developed Heart Arrhythmia's Classification Models

The bar plot in Figure 7 depicts the respective accuracy score of each algorithm. The x-axis of the plot shows the algorithm experimented on. As aforementioned, the algorithm includes LSTM, CNN, and the hybrid of LSTM and CNN. The y-axis displays the accuracy score measured in percentile ranging from 0.0 to 1.0. The CNN algorithm is shown in blue with an accuracy score of 0.97 in percentile, the LSTM model is shown in orange with an accuracy score of 0.90 also in percentile, and also the hybrid of both the LSTM and CNN model in green with an accuracy score of 0.97 in percentile. The graphical plot of the model visualizes both the CNN and LSTM-CNN (hybrid) to outperform the LSTM algorithm with both having an accuracy of 0.97%.

Conclusion and Recommendations

This research has developed three deep learning models; the CNN model, LSTM model and a hybrid of the CNN and LSTM network named (LSTM-CNN) to predict the heart arrhythmia disease. Amongst the three models, the hybrid (LSTM-CNN) in particular have proven to outperform the LSTM model with an outstanding result of 97%. After an extensive experimental study and review of several kinds of literature relating to cardiovascular disease prediction, this study recommends future research as follows:

1. The design of models that predicts and also administer necessary treatment or diagnosis for cardiovascular disease patients.
2. The study used the Daubechies feature selection techniques to determine which features of the dataset were crucial for the experiment that was run. Hence, the performance of the models can be improved by conducting feature selection using several meta-heuristic algorithms, such as the artificial bee colony algorithm, bat algorithm, and so on.

References

Amini, M., Zayeri, F., and Salehi, M. (2021). Trend analysis of cardiovascular disease mortality, incidence, and mortality-to-incidence ratio: Results from the global burden of disease study 2017. *BMC Public Health,* 21(1), 401.

Alday, E. A. P., Gu, A., Shah, A. J., Robichaux C., Ian, Wong, A.K., and Liu, C., (2021). Classification of 12-lead ECGs: The PhysioNet/Computing in Cardiology challenge 2020. *Physiol. Meas.,* 41 (12), 124003.

Belal B. H., Faijan A., Syed J. A., and Abdulrahman A. (2022). Wearable Flexible Electronics Based Cardiac Electrode for Researcher Mental Stress Detection System using Machine Learning Models on Single Lead Electrocardiogram Signal. Biosensor. 12(6): 427. DOI:10.3390/bios12060427.

Bin-Heyat, M. B., Akhtar, F., Abbas, S. J., Al-Sarem, M., Alqarafi, A., Stalin, A., and Wu, K. (2022). Wearable flexible electronics-based cardiac electrode for researcher mental stress detection system using machine learning models on single lead electrocardiogram signal. *Biosensors*, *12*(6), 427.

Cheng-yu L., Yi-Wen W., Febryan S., and Nguyen T.T. (2022). Sleep Apnea Classification Algorithm Development Using a Machine Learning Framework and Bag of Features Derived from Electrocardiogram Spectrogram. Journal of Clinical Medicine. 11(1) :192. DOI:10.3390/jcm11010192.

Ghasem H., Mohammad R. K., Yosef R., and Mehdi A. (2022). Electrocardiogram Signal Classification Using Manual or Automated Features and Machine Learning Algorithms: A large Cohort Study on Students Aging from 6 to 18 Years Old. DOI:10.21203/rs.3.rs-1551010/v1.

Ko, S. Y., Wang, K. M., Lian, W. C., and Kao, C. H. (2012). A portable ECG recorder. *IEEE," in 2012 2nd International Conference on Consumer Electronics, Communications and Networks (CECNet),* 3063–3067.

Liu, G., and Guo, J. (2019). Bidirectional LSTM with attention mechanism and convolutional layer for text classification. *Neurocomputing*, 337, 325–338.

Liu, M., & Kim, Y. (2018). Classification of heart diseases based on ECG signals using long short-term memory. In 2018 40th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC) (pp. 2707-2710). IEEE.

Madugu J. O., Olajide B. O., Okpor J., and Andrew I. W. (2023). Smart User Consumption Analysis Using Support Vector Machine and Multilayer Perceptron. Asian Research Journal of Current Science. 5(1) 163-170.

Mandala S, Pratiwi Wibowo AR, Adiwijaya, Suyanto, Zahid MSM, Rizal A. (2023). The Effects of Daubechies Wavelet Basis Function (DWBF) and Decomposition Level on the Performance of Artificial Intelligence-Based Atrial Fibrillation (AF) Detection Based on Electrocardiogram (ECG) Signals. *Applied Sciences*., 13(5):3036. https://doi.org/10.3390/app13053036

Marcel H., Alex R., Anmol A., and David O. (2021). Developing and comparing machine learning models to detect sleep apnoea using single-lead electrocardiogram (ECG) monitoring. Master's Thesis submitted to 3School of Clinical Medicine, University of Cambridge, United Kingdom. 10-34.

Nazrul A. N., Ab H. H., Mohd S. Z., and Abdullah N. (2020). Cardiovascular Diseases Prediction from Electrocardiogram by Using Machine Learning. International Journal of Online and Biomedical Engineering (IJOE) 16(07): 34, DOI:10.3991/ijoe.v16i07.13569

Olajide B. O., Olawale O., Ngozi C. E. and Peter A. I. (2020). Land Suitability Prognostic Model for Crop Planting Using Data Mining Technique. International Journal of Theoretical and Applied Mathematics. 6(3), 31-38.

Olajide B. O., and Andrew I. S. (2023). A Predictive Model for Car-Loan Repayment Credibility of Customers. Asian Basic and Applied Research Journal. 5(1), 81-89.

Okpor J., Olajide B. O., and Madugu J. O. (2023). Over-the-top Applications Traffic Analysis Model for Networks Using Multilayer Perceptron (MLP) and Long Short-term Memory (LSTM). Asian Journal of Pure and Applied Mathematics. 5(1) 242-250.

Rahman, M., Mehedi Shamrat, F. M., Kashem, M. A., Akter, M., Chakraborty, S., Ahmed, M., & Mustary, S. (2022). Internet of things-based electrocardiogram monitoring system using a machine learning algorithm. *International Journal of Electrical & Computer Engineering, 12*(4).

Siontis, K. C., Noseworthy, P. A., Attia, Z. I., and Friedman, P. A. (2021). Artificial intelligence-enhanced electrocardiography in cardiovascular disease management. *Nat. Rev. Cardiol.,* 18 (7), 465–478.

Xinwen, L., Huan, W., Zongjin, L., and Lang, Q. (2021). Deep learning in ECG diagnosis. *Knowledge-Based Systems*, 227,107-187

Yao, Q., Wang, R., Fan, X., Liu, J., and Li, Y. (2020). Multi-class Arrhythmia detection from 12-lead varied-length ECG using attention-based time-incremental convolutional neural network. *Information Fusion,* 53, 174–182.

Zhang, H., Liu, C., Zhang, Z., Xing, Y., Liu, X., Dong, R., ... & Liu, F. (2021). Recurrence plot-based approach for cardiac arrhythmia classification using inception-ResNet-v2. *Frontiers in physiology*, *12*, 648950.