JOURNAL OF ALGEBRAIC STATISTICS

Volume 13, No. 3, 2022, p.580-596

https://publishoa.com

ISSN: 1309-3452

Comparative Evaluation for Two and Five Classes ECG Signal Classification: Applied Deep Learning

Omar Mohammed Amin Ali
Department of Applied Computer
College of Medicals and Applied Sciences
Charmo University
Chamchamal, Sulaimani, Iraq
omer.muhamad@charmouniversity.org

Shahab Wahhab Kareem
Department of Technical Information Systems Engineering
Technical Engineering College
Erbil Polytechnic University
Erbil, Iraq
shahab.kareem@epu.edu.iq
Department of Information Technology
College of Engineering and Computer Science
Lebanese French University
Erbil, Iraq

Amin Salih Mohammed
Department of Computer Networking
College of Engineering and Computer science
Lebanese French University
Erbil,Iraq
amin.mohammed@su.edu.krd

ABSTRACT

ECG can be using to reliably monitor the health of a cardiovascular system. The proper classification of heartbeats has received a lot of attention recently. While there are numerous similarities among ECG situations, instead of learning and to use transferable knowledge across tasks, most research has concentrated on classify a group of situations using a dataset labeled for that task. This research suggested a technique for heartbeat classification based using deep CNN models that can accurately categorize distinct arrhythmias for two and five classes ECG signals in line to AAMI EC57 standards. The authors also propose a strategy for transferring this job's knowledge to the myocardial infarction (MI) categorization problem. Physician Net's MIT-BIH and PTB Diagnostic databases were used to evaluate the suggested technique. The ECG data was gathered from various Physio Bank databases, which provide clinical research data freely available. Images of ECG signals with time-frequency encoding were fed into architecture such as CNN, LSTM, Alex Net, VGG-16, Resnet50, and Inception. The categorization of ECGs was completed, as well as the performance of CNN, LSTM, Alex Net, VGG-16, Resnet50, and Inception architectures were evaluated using a transfer learning technique and modifications in particular output layers for five designs.

Keywords- ECG Signal, Machine Learning, Deep Learning, Transfer Learning, MIT-BIH and PTB Dataset, Deep Neural Networks (DNNs).

1. Introduction

Cardiovascular disorders are among top causes of death globally, accounting for one-fifth of all fatalities. The linked health-related concerns, notably heart-related diseases, have developed as a result of an increase in sedentary lifestyle among a big population. ECG is a most often used technique by doctors for assessing and interpreting a patient's heart condition. Different components of a healthy person's cardiac activity may be identified, like the P-wave, PR intervals, QRS complexes, QT intervals, and T-wave, which are all shown in Figure 1. The varied patterns and anomalies in these waves can be used to diagnose and identify any cardiac ailment. The identification of numerous heart diseases using computer-based methodology has become commonplace because to improvements in deep learning approaches over the last decade. Arrhythmia (ARR) is a cardiac ailment that affects the rate of the heartbeat. One of the primary causes is the improper production of electrical impulses that control heartbeats. The heart beats very slowly, very quickly, or completely randomly as a result of these faulty electrical impulses. As a result, if proper medical treatment is not received, it might result in a heart attack, heart failure, or sudden cardiac arrest. As a result, it's critical to recognize and characterize anomalies in the ECG signal [1]. As a consequence, physicians and health professionals will benefit from the detection of exact ECG anomalies in order to take appropriate medical action [2]. Neural networks with numerous layers in the neural are used in Deep Learning techniques. Deep learning achieves feature extraction by employing a large number of layers that function as processing units. Each layer extracts a certain characteristic using the output of the layer before it. [3]. In several pattern recognition challenges, deep learning approaches have surpassed previously available methods, motivating the research community to employ these cutting-edge technics of medical image processing.

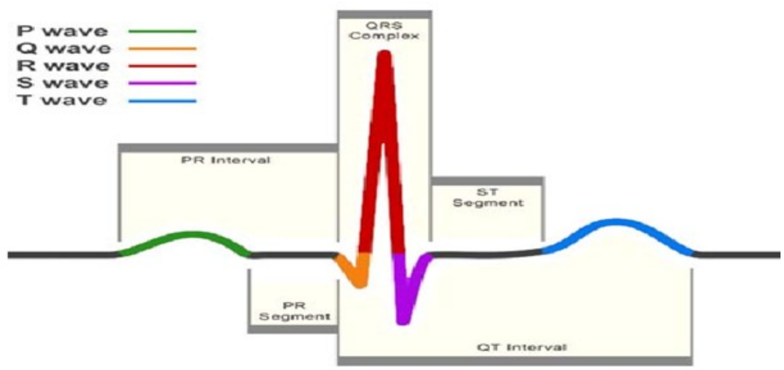

Fig. 1. Wave of ECG signal

For identifying ECG data, a LSTM model network with a novel wavelet-based layer was developed. A wavelet-based layer was used to split the ECG signal in to frequency sub-bands with different scales. The sub-bands are utilized as LSTM model input data. A MIT-BIH arrhythmias dataset has five different types of heartbeats that are collected for classification, totaling 7376 ECG signals [4].
The key benefit given by Deep Neural Networks (DNNs) is that they automatically extract and identify complicated and detailed properties of pictures, reducing the need for human feature extraction, which is a need of conventional Machine Learning methods. This benefit enables the design of an end-to-end pipeline with just an ECG signals as input to the classifying result of output. There are two types of ECG classification tasks: binary and multiple. Is from the other hand, greater accuracy and human-level output can be reached when a large quantity data are available for extracting precise ECG properties and learning from a range of sources. The lack of vast amounts of data necessary to train and execute ECG classifications is one of the most significant factors impeding the widespread use of deep learning algorithms. In comparison to traditional ML-based classification algorithms, DNNs and CNNs require a massive quantity of data for training. This issue arises due to a shortage of publicly accessible databases in the ECG analysis sector, resulting in a gap between appropriate difficult ECG properties and the quantity of the database [5].
To address the aforementioned problem, this study provides a multimodal classifier technique for ECGs based on transfer learning from different classes. which is unique to the ECG domain. Specifically, rather of starting from scratch with ECG data to train CNNs, architectures pre-trained using picture classification and object identification data (1000 classes such as a chair, mouse, table light, and so on) are used.
Because these fields have large datasets, it is possible to train and extract feature maps in a short amount of time that reflect intricate patterns and characteristics in images. Only by employing Continuous Wavelet Turn to It is possible to

JOURNAL OF ALGEBRAIC STATISTICS
Volume 13, No. 3, 2022, p.580-596
https://publishoa.com
ISSN: 1309-3452

learn to convert 1-Dimensional ECG signals samples into a 2-Dimensional representation of the ECG signal, known as a Scalogram, and to send accessible feature maps for ECG classify (CWT). Using 900 ECG signals scalograms from a small database, experiments reveal that, CNN, LSTM, Alex Net, VGG-16, Resnet50, and Inception trained beforehand using ImageNet dataset can extract characteristics well. The results of the trials using the pre trained CNN architectures are presented in this study.

2. Literature Review

In 2015 Serkan Kiranyaz et al [6] used the MIT-BIH arrhythmia database to create a classification model using one-dimensional CNNs. "The suggested technique only requires 1-D convolutions (multiplications and additions), which simplifies and reduces the cost of any hardware implementation," they write in their study. Furthermore, once a specialized CNN has been trained for a specific patient, it may be utilized purely to categorize the patient's lengthy ECG recordings." Their model (Figure 2) has an accuracy rate of more than 90%.

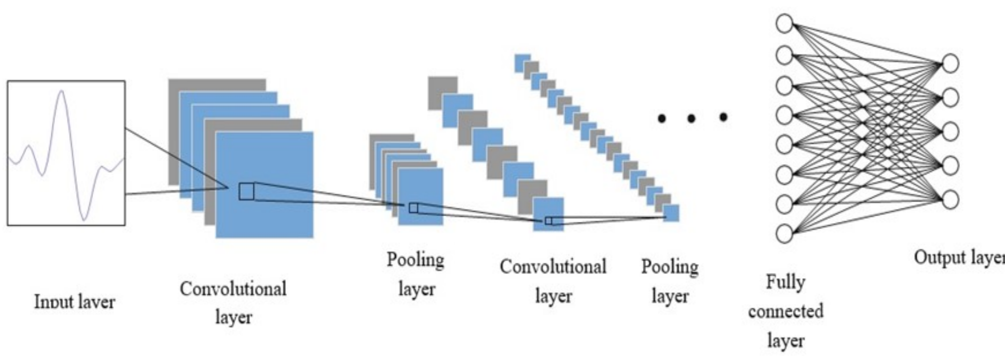

Fig 2. CNN model overall structure

In 2018, Mohammad Kachuee et al [7] made a model (Figure 3) with CNNs. They used the same dataset in the Serkan, Turker and Moncef's paper. This model Exceeded 90% accuracy. They divided the dataset are to two classifications: 80 % are used to training and 20 % are used to testing.

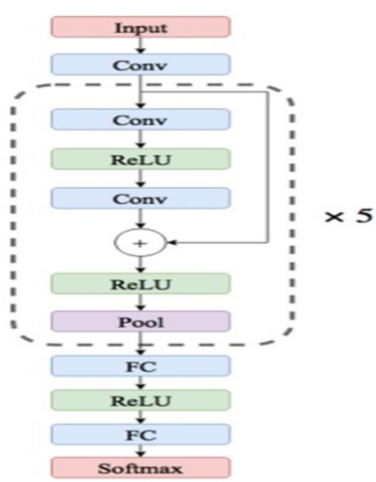

Fig. 3 suggested network's design Model of CNN

In 2020's Physionet Challenge, Bjørn-Jostein Singstad and Christian Tronstad [8]. Tested several models on a dataset that contains 43101 ECG recordings with corresponding information files describing the recording, patient attributes, and diagnosis. This dataset had 27 classes. They employed Fully Convolutional Neural Networks (FCN) and Encoder types of CNN models. They reported that FCN outperformed eight other CNN architectures compared. They also wanted to test the second-best architecture from their study which was an Encoder network. Finally, they assessed the integration of a rule-based algorithm within these models to test the performance of a CNN and rule-based hybrid classifier. They also used a rule-based model. The rule-based algorithm used the raw ECG signal as input, with no padding or truncation. Then algorithm performed classification independent of the deep learning models. If there was disagreement between the rule-based algorithm and the CNN model, the rule-based algorithm overwrote the classification from the CNN model. They divided the data set 90% for training, 10% for testing.

In 2016, Zubair et al. [9] proposed, the raw data was supplied to a model-based 1D-CNN to classification ECG signal into many classifications determined by model and the Association of Advancement and Medical Instrumentation (AAMI). They provide an ECG beat classifier based on CNN. A suggested model combines the ECG pattern recognition system's two primary sections, feature extraction and classification. From the raw ECG data, they algorithm automatically finds the suitable features representations, eliminating the requirement for hand-crafted features. The suggested classification method efficiently categorized ECG beats into five distinct groups by using minimal and patient-specific training data. To evaluate classifications performance, 44 records from MIT-BIH dataset's ECG signals were used, and the findings show the suggested methodology out performs most state the arts ECG signal classification approaches in term of classification accuracy to computing efficiency.
In 2020, Wangin [10] The model was built using a 1D CNN Added to an 11-layer Modified Elman Neural Network (MENN), which was a neural network, were used to create the model, as well as a non-transform ECG data. One of the most prevalent causes of life-threatening heart sickness is atrial fibrillation (AF), and for ECG, several researchers have suggested using the CNN model interpretation based of forms in 2- Dimensions. For identifying AF, professional cardiologists must visually analyze electrocardiogram (ECG) data, which is inefficient and time-consuming. In this paper, he proposes a breakthrough technique for the automatic while categorizing end-to-end
signals. Two important deep network models were constructed specifically to prove the superior of a suggested model by way of comparison.

In 2020, Zheng et al [11] They converted 1-Dimensional ECG signal data into 2-Dimensional images before entering the CNN model. The experimental data are separated into three sets: training, testing and validation, with each set accounting for 60%, 20%, and 20% of the total. They used a CNN and LSTM to construct an arrhythmia classification approach, which was then used to detect eight ECG signals.
In 2019, M.Talo et al. [12] A model for identifying a diabetic person was created by merging the 2D-CNN model that has already been trained, using frequency images made from heartbeat signal. 20% of the dataset in the research was utilized for testing, while the remaining 80% was used for training. They proposed developing of deep transfer learning approach for automated identification of diabetic utilizing heart rate (HR) signal collected with ECG data ((DM). Recently advancements in Deep learning have made a substantial contribution to the improvement of health care quality. Therefore, they suggested deep learning to make their model work well.
In 2020, Liu G. et al. [13] They in order to automatically analyze six different types of ECG data, the researchers developed a deep learning approach combining CNN and LSTM. ECG signals from two datasets were classified as Atrial Fibrillation (AFIB), Ventricular Bigeminy (B), Normal (N), pacing rhythm (P), Sinus Bradycardia (SBR) and Atrial Fibrillation (AFL).

In 2018, Oh SL. et al. [14] Using a CNN and LSTM, researchers proposed an automated technique for identifying healthy sinus rhythm, Left Bundle Branch Block (LBBB), and Right Bundle Branch Block (RBBB). ECG readings show early atrial premature beat (APB) and ventricular contraction (PVC). This study is unusual in that it makes use of ECG data of varied length.

In 2018, Swapnae G. et al. [15] The CNN-LSTM approach was suggested to discriminate between abnormal and normal ECGs. An accuracy parameters are 0.834. when using five-fold cross-validation. They use twisted CNN, recurrent structures like RNN, LSTM, and gated repeated unit (GRU), as well as a mix of CNN and repeated components, to auto identify the issue. Deep learning methods, unlike conventional analysis approaches, Never use analytical techniques based in feature extraction. The finest of these are chosen as the parameters for the deep learning approach after a series of tests.

In 2020, Sangaiah AK. et al. [16] suggested an artificial learning technique for arrhythmia analysis by extracting characteristics from ECG data. The Hidden Markov Model is used in the suggested strategy (HMM). ECG signals were used to extract and select features, and five separate arrhythmia types were identified.

In 2020, Wang H Shi.et al. [17] They provided details of a comprehensive binary neural network model for precise heart rate categorization. ECG data were divided into five categories: supraventricular ectopic beat (S), normal beat (N), mixed beat (F), ventricular ectopic beat (V), and unknown beat using a two-step method (Q). The signals were initially given a 105 feature analysis. A two-layer classifier was utilized in the subsequent stage.

In 2020, Sharma Garg N et al. [18] ECG signals were classified using the LSTM model. First, ECG data were used to compute RR-interval sequences. A property vectors is then extracted use Fourier-Bessel (FB) expansions from the RR interval sequences. To categorize the obtained vectors, an LSTM model is employed. Classification was done using the MIT-BIH Arrhythmia Database.

3. Methodology

The Alex Net, VGG-16, Resnet50, and Inception were deep CNNs that were created to categorize the ImageNet dataset into 1000 different categories. These CNN network topologies are reused to categorize ECG data based on pictures derived from the CWT of these signals. the goal in the research is to train a variety of deep model to categorize ECG data and assess their transferability.

Simple networks contain only dense units. Neurons are dense units that theoretically taken to biological neurons. The synthetic neuron receives inputs and generates one output that may be shared across several other neurons. the outputs of other neurons or feature values from a sample of outside information, like documents or images, might be used as inputs. The last a neural network's output neurons finish the job, like identifying a thing from an image. An author's start by taking the weighted total of all the inputs and adding a bias factor to get the neuron's output. The activation is the name given to this weighted sum.

The 19 layers of the 1D CNN model include four convolutions layers, two half percent dropout layer, and two fully-connected layers. Cross validation methods are used an optimize some model parameters for instance the number convolution layer, filters and epochs. Table 1 shows the CNN-layers and their explanations. A thirteenth layer of the CNN model provides a good recovery of the spatial and local feature map, this has ten filters on a convolutional layer. Edges, vertical and horizontal lines are recognized by the filters, and other properties in an ECG signal.

JOURNAL OF ALGEBRAIC STATISTICS
Volume 13, No. 3, 2022, p.580-596
https://publishoa.com
ISSN: 1309-3452

Table I. Explanation models 1D CNN

No. of Layers	Names	Explanation
1	Inputs	Depends on the input with 'zerocenter' normalization.
2	Layer1 Convolution	20 7*7 convolutions layers with (1 * 1) paddings same
3	Batch normalization 1	Batch normalization1
4	ReLU ,Clipped	ReLU is clipped ceilings5
5	Drop Out	Drop Out 50%
6	Layer2 Convolutions	20 convolutions of 9*9 using length (1 * 1) and padding same.
7	Batch normalization 2	Batch normalization 2
8	ReLU Leaky	ReLU with a leaky scale of 0.01
9	Layer3 Convolution	30 5x5 convolution layer with (1 * 1) stride and padding'same'
10	Batch normalization 3	Batch normalization 3
11	Soft Max	Soft Max
12	Drop out	Drop out 50%
13	Layer4 Convolution	10 3x3 convolution layer with 1 * 1 stride and padding'same'.
14	Batch normalization 4	Batch normalization 4
15	ReLU	ReLU
16	Completely Connected 1	60 layers are totally connected.
17	Completely Connected 2	Two layers that are completely interconnected
18	Soft Max	The last Fully Connected layer's activation function.
19	Classifications	Entropy of the output Entropy the output

In the Neural Network, an activation function explains how a node or nodes in a layers convert the total of inputs in to output. Activation functions are classified into two categories: output layers' activation as well as hidden layers' activation.

In hidden layers, there are primarily three activation function to choose from:

ReLU, Logistic(Sigmoid) and Hyperbolic Tangent (Tanh).

The following is how the ReLU function is calculated:

A) Max [0.0, x].

If an input value (x) be negative, then value is returned; otherwise, the value is returned.

The following formula is used to determine the sigmoid activation function:

B) [1.0/ (1.0 + e^-x)].

JOURNAL OF ALGEBRAIC STATISTICS
Volume 13, No. 3, 2022, p.580-596
https://publishoa.com
ISSN: 1309-3452

The base of the natural logarithm is e, which is a mathematical constant. Finally, we compute the activation function tanh as follows:

C) [(e^x – e^-x)/ (e^x+ e^-x)].

Linear, Logistic (Sigmoid), Soft max are the three activation functions you should think about for the output layer. The functions listed above are the most prevalent. More activation options are available.
A weighted total of a input is unchanged by the linear activation function, which simply returns a result.
The following is how the softmax function are calculated:

D) [e^x/sum(e^x)]. Where x is just a set with outputs. [19]

Table II shows the LSTM model structure, which includes two LSTM layers, out layers, and two levels that are entirely interconnected. The proposed model yields two classifications for the MIT-BIH Arrhythmia Dataset: healthy and aberrant Arrhythmia, for the PTB Diagnostics ECG Database, normal and Myocardial Infarction.

Table II. Explanation models LSTM

No. of Layers	Names	Explanation
1	Inputs	Input dimensions should be sequenced.
2	Layer1 LSTM	120 hidden units in the LSTM .
3	Soft Max	Soft Max
4	Dropout	Drop out 30%
5	Layer2 LSTM	120 hidden units in the LSTM .
6	Completely Connected 1	There are ten completely linked layers in all.
7	Completely Connected 2	There are two completely linked layers in all.
8	Soft Max	Soft Max
9	Classifications	Entropy of the output Entropy of the output

Before proceeding any further, one more layer type must be discussed, as this layer type is employed in the basic network, which is pooling layers. Pooling layers feature maps are down sampled by summing the existence of parts in sections to the feature map.

The VGG-16 Model for ECG rhythms classifications, as a fully connected 16-layers CNN, which had already been trained on the ImageNet dataset, was used. A network model is made up of convolution layers, fully connected layers and pooling layers. Last completely linked layers were deleted replaced and deleted with dense layers for categorize ECG beats in class for the current study's implementation of modified VGG-16. All levels preceding completely linked layers were frozen since the model was not fine-tuned. Weights of frozen layers just weren't updated and were used in their original condition. shown in figure 4.

JOURNAL OF ALGEBRAIC STATISTICS
Volume 13, No. 3, 2022, p.580-596
https://publishoa.com
ISSN: 1

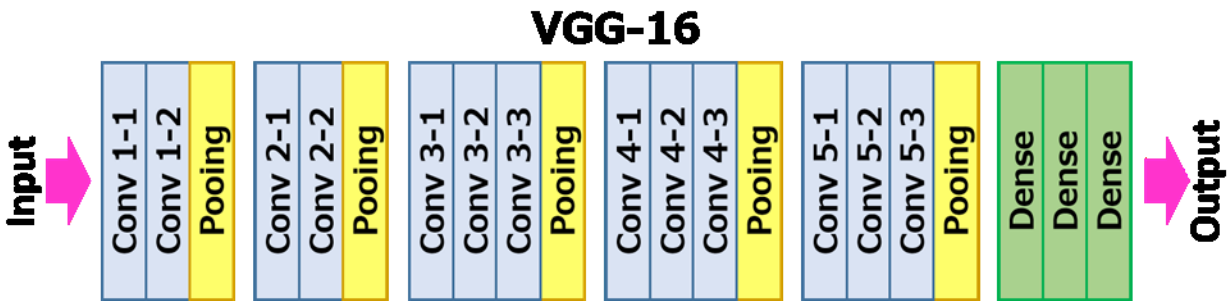

Fig 4. Define the VGG-16 network model

ResNet-50 model Batch normalization and max pooling were layered in the already trained to avoid overfitting convolution layers. Figure 5 shows a ResNet-50 network with 50 layers for first group classification. The ResNet model's last layer was configured to train, while all the layers before it was frozen, so the weights learned from prior data could be utilized to train new features and identify ECG beat types.

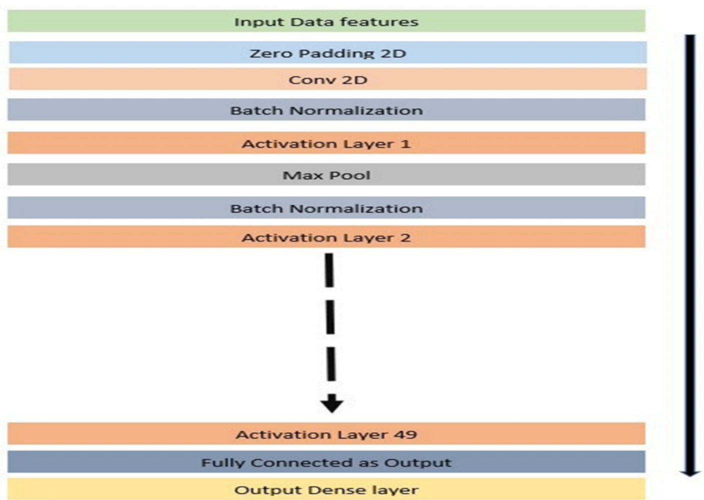

Fig 5. ResNet-50 architecture for binary classification

Alex Krzyzewski, Geoffrey Hinton and Elias Sutskiver created the 'Alex Net' in 2012 won the ILSVRC (Image Classification Challenge). AlexNet is made up of five convolutional layers, a maximum aggregation layer, three completely linked layers, and a 1000-way softmax classifier. [20] Like Lenet5, this model is updated by deleting pooling layers. As an extra, dropout layers are added to the end. Model is trained for 100 epochs for two class and 30 epochs for five class visualized the metrics with matplotlib (Figure 6).

Inception networks consists of multiple inception blocks. These blocks overlap each other to form the inception model. Model gets its name from of the popular internet joke "we need to dig deeper." The Inception module (Figure 7) operates by convolution ally filtering an input with three distinct size to filters (1*1, 3*3, 5*5). In addition, maximum pooling is using. The outputs are then mixed and delivered to next tier. [21]. In this model 5 Inception Modules is used (Figure 8). Then model is trained for 30 epochs and visualized the metrics with matplotlib.

JOURNAL OF ALGEBRAIC STATISTICS
Volume 13, No. 3, 2022, p.580-596
https://publishoa.com
ISSN: 1309-3452

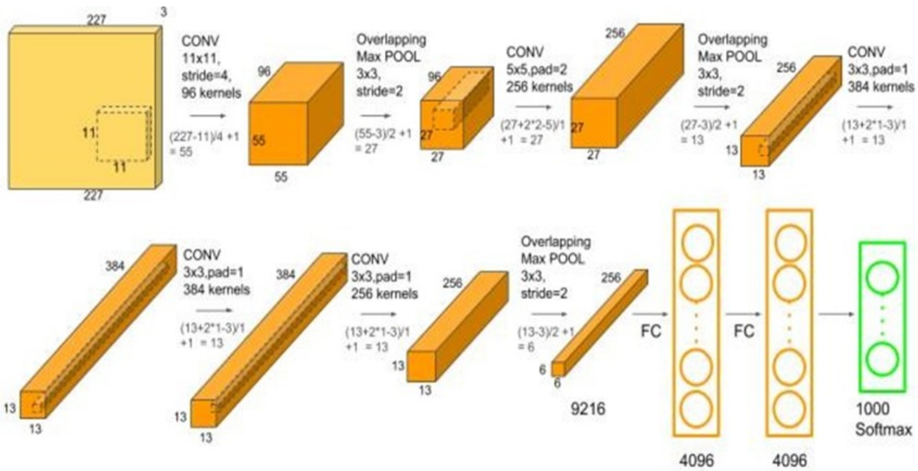

Fig 6. AlexNet Model

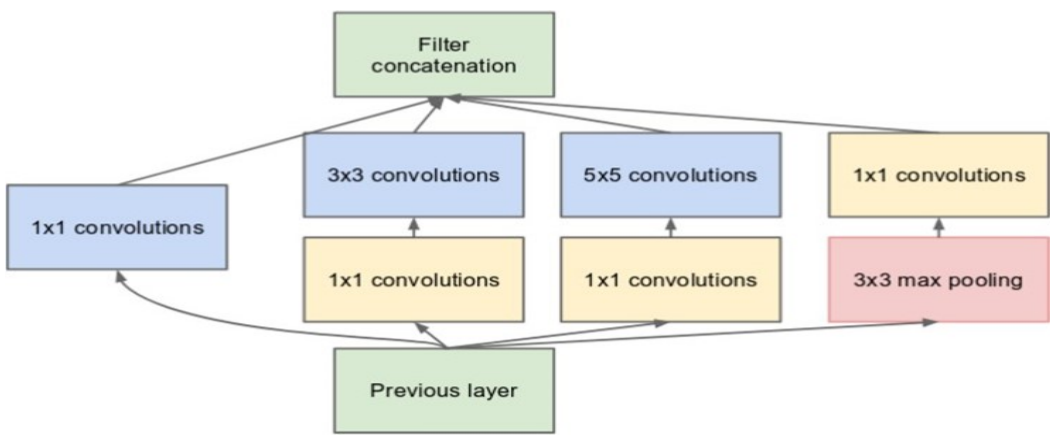

Fig 7. Inception module with dimension reductions

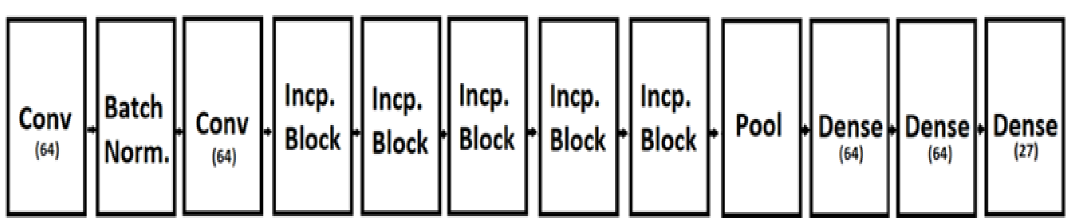

Fig 8. Inception Architecture

The two most popular pooling techniques. Average pooling, which takes the average of specified cell, and Max pooling, which gets the maximum value. [22]

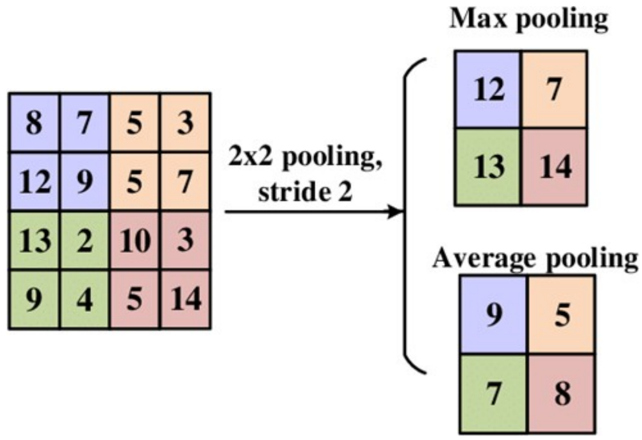

Fig 9. Pooling Layer operation approaches

After implementing model, it needs to be compiled. While compiling, there are several parameters to choose: a loss function, an optimizer and metrics.

The loss function is a way to assess well how a model works. The loss function will produce a bigger value if your forecasts are low. If they are satisfactory, the loss function will return a lower value.

The loss function would inform you if you're making progress as you adjust your model [23]. In this study, Equation 2 uses a binary cross-entropy loss function The sigmoid activation function and cross-entropy loss are combined in binaries cross-entropy loss, also referred to as sigmoid cross-entropy loss (Equation 1). In contrast to Softmax loss, the loss estimated to each neural component of the output vector is independent with the other component's values, indicating that the values of other components have no bearing on the loss calculated for each component. It's for this reason that it's employed for multi-label classifications [24].

$$CE = -\sum_{i}^{C} t_i \log(s_i) \tag{1}$$

$$CE = -\sum_{i=1}^{C\prime=2} t_i \log(s_i) = -t_1 \log\big(f(s_1)\big) - (1 - t_1)\log(1 - f(s_1) \tag{2}$$

Optimizers connect a loss function for models parameters by modifying the models in response to a loss function's output. Optimizers employ weights to in other words, shape the models as accurately as possible. Adam optimization function is utilized in this paper. In SGD, Adam is a replacement optimization method (stochastic gradient descent). Adam combines the greatest features of the Ada Grad and RMS Prop methods into a single algorithm. [25].

- Adaptive Gradient Algorithms (Ada Grad) increases performance for challenges with distributed gradients natural language processing (e.g., NLP) and by keeping a per-parameter learning rate, computer visions difficulties may be avoided.
- Root Mean Square Propagation (RMS Prop) is a learning algorithm than adjusts the parameter learning rates depend on a weight's average recent gradient magnitudes. This means that the technique works well for both stationary and concerns that aren't stationary (e.g., noisy).

JOURNAL OF ALGEBRAIC STATISTICS
Volume 13, No. 3, 2022, p.580-596
https://publishoa.com
ISSN: 1309-3452

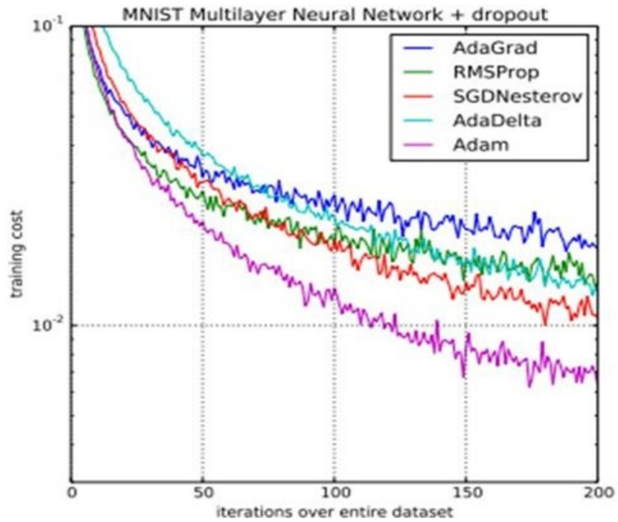

Fig 10. Training a Multilayer Perceptron with Adam vs. Others Optimization Algorithms.

The metrics used in this study "Accuracy", "Precision", "Recall", "AUC" and "Loss" metrics are monitored during training.

The metric by which categorization models are judged is accuracy. Informally, accuracy refers to the number the correct predictions made by a model. Divide the total numbers of estimates by a number of correct forecasts to arrive at an accuracy figure. [26].

A result when the model correctly predicts a positive goal is known as a true positive (TP). When the model accurately predicts the negative target, it is said to be a true negative (TN). A false positive (FP) happens when the model predicts the positive objective erroneously, while a false negative (FN) happens when the model predicts the negative objective incorrectly.

Precision aims the respond to the following query: How many right identifications were truly correct. [27].

Recall aims to provide a response the question following: How proportion of genuine positive were successfully tagged. [27].

$$\text{Accuracy} = \frac{\text{TP+TN}}{\text{TP+TN+FP+FN}} \tag{3}$$

$$\text{Precision} = \frac{\text{TP}}{\text{TP+FP}} \tag{4}$$

$$\text{Recall} = \frac{\text{TP}}{\text{TP+FN}} \tag{5}$$

JOURNAL OF ALGEBRAIC STATISTICS
Volume 13, No. 3, 2022, p.580-596
https://publishoa.com
ISSN: 1309-3452

Before discussing AUC, the ROC Curve must be described. A graph that displays how well a classifier model works at various levels of classification is called the receiver operating characteristic curve (ROC curve). This graph shows the True Positive Rate and False Positive Rate measurements. The AUC (Area Under the ROC Curve) is only a two-dimensional measurement of the region underneath the whole ROC curve between (0,0) and (1,1). (1,1). (1,1). [28].

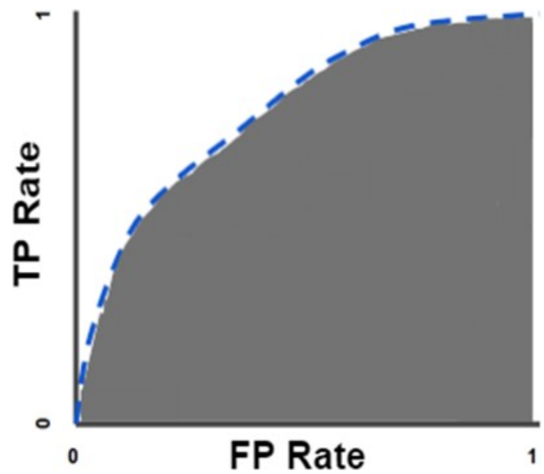

Fig7. AUCs (Area under to ROC Curves).

A loss is the cost of making a wrong forecast. To put it differently, loss is a metric that measures how accurate a model forecast in a certain circumstance. If a model's forecast is high, a loss is decreased; If not, there will be a bigger loss. Finding a set of weights and biases that causes the least amount of loss is the goal of model training. [29].

4. Dataset

In this study, ECG data of two group of people is used. The ECG records below were obtained from the Physio Net databases: MIT-BIH Arrhythmia Database and PTB Arrhythmia Database [30].

1) MIT-BIH Arrhythmia Dataset:

The MIT-BIH Arrhythmias database contains data from 47 arrhythmias patients who were investigated performed by the BIH Arrhythmias Laboratories. the first 23 recordings of acquired from a varied group of outpatients Among Boston's Beth Israel Hospital's outpatients (40%) and inpatients (60%); the final 25 are selected from a single set to capture uncommon though clinically important that would be arrhythmia underrepresented in a smaller random samples. [30].

Fig 11. MIT-BIH Arrhythmia

JOURNAL OF ALGEBRAIC STATISTICS
Volume 13, No. 3, 2022, p.580-596
https://publishoa.com
ISSN: 1309-3452

2) PTB Arrhythmia Dataset:

There are 549 records in the database from 290 different disciplines (aged 17 - 87, mean 57.2; 209 males, mean 55.5, and 81 women, mean 61.6; 1 female and 14 male respondents' ages also weren't reported). One to five recordings are assigned to each subject. There are no subjects with the numbers 124,132,134, or 161 assigned to them. Each record measures the standard There are 12 leads (I, II, III, AVR, AVF, V1, V2, V3, V4, V5, V6) and three Frank lead ECGs (vx, vy, vz). Each signal is digitalized across a range of 16.384 mV at a sampling rate of 1000 samples per second. Dataset contributors may get access to recordings with sample rate up to 10 KHz if they make a specific request. Most ECG recordings provide a full clinical explanation in the header file, which includes gender, age, and diagnosis, when appropriate, medical background, drugs and therapies, Echocardiography, ventriculography, hemodynamics, and coronary artery pathology [30].

5. Experiments and Results

A. System Configuration

The study is carried out with the help of the Computer Lenovo Intel Corei5 10400th generation processors and CPU @ 2.90 GHz, 12 GB RAM on a Windows operating system. The Python program language was used to create the model. In the Python 3.9.10 environment, deep learning tools were used in model creation and assessment.

B. Experimental Setting

A features were retrieved layer to layer, beginning with the input layer and progressing via hidden levels to the final classification output layer. The image's input resolution was changed to meet the model's requirements. With a faster learning rate, the learning process was accelerated. For classification issues, the feature map is elongated in the required shape by the fully connect layer, which is follow by that of the function for multiple-class activation. To lower the cost of pre-trained model, the continuous cross-entropy is utilized as a cost function. We employed a heuristic method to hyper parameter selections of our technique. Among the tested optimization methods, the RMS prop had the best response.

C. Result and Discussion

The suggested method employed to transfer learn idea a retrain of a shelf model that had previously been trained on natural medical image categorization images. An attempt is being made to decrease the model complexity that comes with training a model from the ground up. Medical image categorization is performed using to parameter gained during ImageNet training by the models. In addition, we conducted a thorough study using a variety Featuring cutting-edge pre-trained architectures with different depth. According to conventional wisdom, the performances to a deep networks improves as the architectural depth increases. Each layer extracts a collection of critical attributes that are passed on to the following level as input. The collected dataset is assessed on several depth pre-trained models in order to notice this behavior.

Volume 13, No. 3, 2022, p.580-596
https://publishoa.com
ISSN: 1309-3452

TABLE. III. ACCURACY AND TRAINING TIME for 2 Class			
Model	**Accuracy For 2 Class**	**No. of Epoch**	**Training Time for 1 Epoch**
CNN	**99.96**	**50 Epoch**	**13 sec.**
LSTM	**97.91**	**150 Epoch**	**12 sec.**
Alex Net	**99.55**	**100 Epoch**	**21 sec.**
VGG -16	**98.02**	**100 Epoch**	**1 min. 42 sec.**
Res Net-50	**98.53**	**100 Epoch**	**2 min. 48 sec.**
Inception	**97.98**	**30 Epoch**	**8 min. 35 sec.**

TABLE. IV. ACCURACY AND TRAINING TIME for 5 Class			
Model	**Accuracy For 5 Class**	**No. of Epoch**	**Training Time for 1 Epoch**
CNN	**98.44**	**200 Epoch**	**5 min. 1 sec.**
LSTM	**99.21**	**200 Epoch**	**5 min.**
Alex Net	**98.78**	**30 Epoch**	**2 min. 44 sec.**
VGG -16	**98.46**	**30 Epoch**	**12 min. 30 sec.**
Res Net -50	**97.90**	**30 Epoch**	**23 min. 28 sec.**
Inception	**98.10**	**30 Epoch**	**1h. 31 min. 12 sec.**

As a result of looking at Tables 3 and 4, it is clear that the results of the two classes are different in terms of time and results because the identification of two classes of data is clearer and easier to identify than five classes of ECG signal data.

6. Conclusion

In heart disease, the most crucial factor is early diagnosis. This has the potential to save lives. ECG stands for electrocardiogram. This information reveals whether or not a person is in good health. The computer may be trained to classification this data by looking at it. Neural networks can be used to do this. On these cardiac data, 6 distinct neural network models for two and five classes ECG signals are developed and assessed. Different numbers of epochs are used to train each model. Finally, the two-class CNN model has the greatest according the results, performance on this database. It is still unable to classify some illnesses, but it is improving. This is because the startup model is more complex than other models. It employs many convolutional procedures in a single layer, and when these layers are combined, the result is a model that is extremely deep, dense, and powerful. These findings can be improved by training on more data, training for longer periods of time, and tweaking a few more parameter (e.g., learning rate, number of epoch, hidden classes, units, and activation functions.).

7. References

[1] A. Diker, E. Avci, Z. Cömert, D. Avci, E. Kaçar and İ. Serhatlioğlu, "Classification of ECG signal by using machine learning methods", 2018 26th Signal Processing and Communications Applications Conference (SIU), pp. 1-4, 2018.

[2] W. Zhang, L H. Wang et al., "A Low-Power High-Data-Transmission multi-lead ECG Acquisition Sensor System", IEEE Sensors J., vol. 19, no. 22, pp. 1-3, Nov. 2019.

[3] L. Deng and D. Yu, "Deep Learning: Methods and Applications," Found. Trends® Signal Process., vol. 7, no. 3–4, pp. 197–387, 2014. Douangnoulack, Phonethep, and Veera Boonjing. "Building Minimal Classification Rules for Breast Cancer Diagnosis." 2018 10th International Conference on Knowledge and Smart Technology (KST). IEEE, 2018.

[4] Özal Yildirim, A novel wavelet sequence based on deep bidirectional LSTM network model for ECG signal classification, Computers in Biology and Medicine, Volume 96, 2018, Pages 189-202, ISSN 0010- 4825.

[5] Rajpurkar, Pranav & Hannun, Awni & Haghpanahi, Masoumeh & Bourn, Codie & Y. Ng, Andrew. (2017). Cardiologist-Level Arrhythmia Detection with Convolutional Neural Networks.

[6] Kiranyaz, Serkan & Ince, Turker & Gabbouj, Moncef. (2015). Real-Time PatientSpecific ECG Classification by 1D Convolutional Neural Networks. IEEE transactions on bio-medical engineering. 63. 10.1109/TBME.2015.2468589.

[7] M. Kachuee, S. Fazeli and M. Sarrafzadeh, "ECG Heartbeat Classification: A Deep Transferable Representation," 2018 IEEE International Conference on Healthcare Informatics (ICHI), 2018, pp. 443-444, doi: 10.1109/ICHI.2018.00092.

[8] Bjørn-Jostein Singstad and Christian Tronstad. "Convolutional Neural Network and Rule-Based Algorithms for Classifying 12-lead ECGs" ISSN: 2325-887X DOI: 10.22489/CinC.2020.227

[9] M. Zubair, J. Kim, and C. Yoon, "An automated ECG beat classification system using convolutional neural networks," in 2016 6th International Conference on IT Convergence and Security, ICITCS 2016. doi: 10.1109/ICITCS.2016.7740310.

[10] J. Wang, "A deep learning approach for atrial fibrillation signals classification based on convolutional and modified Elman neural network," Futur. Gener. Comput. Syst., 102, 670–679, 2020, doi: 10.1016/j.future.2019.09.012.

[11] Z. Zheng, Z. Chen, F. Hu, J. Zhu, Q. Tang, and Y. Liang, "An automatic diagnosis of arrhythmias using a combination of CNN and LSTM technology," Electron., 9(1), 1–15, 2020, doi: 10.3390/electronics9010121.

[12] O. Yildirim, M. Talo, B. Ay, U. B. Baloglu, G. Aydin, and U. R. Acharya, "Automated detection of diabetic subject using pre-trained 2D-CNN models with frequency spectrum images extracted from heart rate signals," Comput. Biol. Med113, 103387, Oct. 2019, doi: 10.1016/j.compbiomed.2019.103387.

[13] Chen C, Hua Z, Zhang R, Liu G, Wen W. 2020. Automated arrhythmia classification based on a combination network of CNN and LSTM. Biomed Signal Process Control. 57:101819.

[14] Oh SL, Eddie YKN, Tan RS, Acharya R. 2018. Automated diagnosis of arrhythmia using combination of CNN and LSTM techniques with variable length heart beats. Comput Biol Med. 102:278–287

[15] Swapna G, Soman KP, Vinayakumar R. 2018. Automated detection of cardiac arrhythmia using deep learning techniques. Procedia Comput Sci. 132:1192–1201.

[16] Sangaiah AK, Arumugam M, Bian G-B. 2020. An intelligent learning approach for improving ECG signal classification and arrhythmia analysis. Artif Intell Med. 103: 101788.

[17] Wang H, Shi H, Lin K, Qin C, Zhao L, Huang Y, Liu C. 2020. A high-precision arrhythmia classification method based on dual fully connected neural network. Biomed Signal Process Control. 58:101874

[18] Sharma A, Garg N, Patidar S, Tan RS, Acharya R. 2020. Automated pre-screening of arrhythmia using hybrid combination of Fourier–Bessel expansion and LSTM. Comput Biol Med. 120:103753.

[19] Jason Brownlee, "How to Choose an Activation Function for Deep Learning" machine learning mastery, 22 Jan. 2021, machinelearningmastery.com/choose-anactivation-function-for-deep-learning.

[20] Rizwan, M., 2018. AlexNet Implementation Using Keras. Medium. Available at: <https://medium.com/datadriveninvestor/alexnet-implementation-using-keras-7c10d1bb6715> 18 October 2018.

[21] Shaikh, J., 2018. Deep Learning in the Trenches: Understanding Inception Network from Scratch. Analytics Vidhya. Available at: https://www.analyticsvidhya.com/blog/2018/10/understanding-inception-network-from-scratch/> 29 May 2018.

[22] Jason Brownlee, "A Gentle Introduction to Pooling Layers for Convolutional Neural Networks" machine learning mastery, 5 July 2019, machinelearningmastery.com/pooling-layers-for-convolutional-neural-networks/ 36.

[23] Data Robot AI Cloud. 2018. Introduction to Loss Functions. Available at: <https://www.datarobot.com/blog/introduction-to-loss-functions/> 30 April 2018.

[24] Raul Gomez, "Understanding Categorical Cross-Entropy Loss, Binary Cross-Entropy Loss, Softmax Loss, Logistic Loss, Focal Loss and all those confusing names", 23 May 2018, gombru.github.io/2018/05/23/cross_entropy_loss/

[25] Jason Brownlee, "Gentle Introduction to the Adam Optimization Algorithm for Deep Learning", machinelearningmastery, 13 Jan. 2021, machinelearningmastery.com/adam-optimization-algorithm-for-deep-learning/

[26] Google Developers. 2021. Classification: Accuracy | Machine Learning Crash Course | Google Developers. Available at: <https://developers.google.com/machine-learning/crash-course/classification/accuracy> 5 February 2021.

[27] Google Developers. 2021. Classification: Precision and Recall | Machine Learning Crash Course | Google Developers. Available at: <https://developers.google.com/machine-learning/crash-course/classification/precision-and-recall> 5 February 2021.

Volume 13, No. 3, 2022, p.580-596
https://publishoa.com
ISSN: 1309-3452

[28] Google Developers. 2021. Classification: ROC Curve and AUC | Machine Learning Crash Course | Google Developers. Available at: <https://developers.google.com/machine-learning/crash-course/classification/roc-and-auc> 5 February 2021.

[29] Google Developers. 2021. Descending into ML: Training and Loss | Machine Learning Crash Course | Google Developers. Available at: <https://developers.google.com/machine-learning/crash-course/descending-into-ml/training-and-loss> 5 February 2021.

[30] A. L. Goldberger, L. A. Amaral, L. Glass, J. M. Hausdorff, P. C. Ivanov, R. G. Mark, J. E. Mietus, G. B. Moody, C.-K. Peng, and H. E. Stanley, "Physiobank, physiotoolkit, and physionet: components of a new research resource for complex physiologic signals," circulation, vol. 101, no. 23, pp. e215–e220, 2000.

[31] Guo Z, Chen Q, Wu G, Xu Y, Shibasaki R, Shao X., Village Building Identification Based on Ensemble Convolutional Neural Networks. Sensors.;17(11):2487, 2017.

[32] ImageNet. http://www.image-net.org /2021. Available at: <https://www.image-net.org/> 11 May 2021.

[33] C. Szegedy, W. Liu, Y. Jia, P. Sermanet, S. Reed, D. Anguelov, D. Erhan, V. Vanhoucke, and A. Rabinovich, "Going deeper with convolutions," in IEEE Conference on Computer Vision and Pattern Recognition, 2015, pp. 1–9.

[34] Huang, Gao, et al. "Densely connected convolutional networks." Proceedings of the IEEE conference on computer vision and pattern recognition. 2017.

[35] Iandola, F.N.; Han, S.; Moskewicz, M.W.; Ashraf, K.; Dally, W.J.; Keutzer, K. SqueezeNet: AlexNet-level accuracy with 50x fewer parameters and <0.5MB model size. arXiv 2016; arXiv: 1602.07360.

[36] Kareem, S., 2022. Bayesian Network Structure Discovery Using Antlion Optimization Algorithm. Ijosi.org. Available at: <https://www.ijosi.org/index.php/IJOSI/article/view/486> 31 March 2022.

[37] Kareem, S. and Okur, M., 2021. Falcon Optimization Algorithm for Bayesian Networks Structure Learning. Journals.agh.edu.pl. Available at: <https://journals.agh.edu.pl/csci/article/view/3773/2652> 11 June 2021.

[38] Kareem, S. and Okur, M., 2020. Structure Learning of Bayesian Networks Using Elephant Swarm Water Search Algorithm. International Journal of Swarm Intelligence Research, 11(2), pp.19-30.