



SqueezeNet Fusion: Enhancing Rhythmic Heart Disease Classification through Integrated Pattern Mining and Deep Learning

Name of Author 1¹, Name of Author 2² and Name of Author 3^{1,3}

¹Department, Name of the Organization, City, Country

²Department, Name of the Organization, City, Country

³Department, Name of the Organization, City, Country

Received Mon. 20, Revised Mon. 20, Accepted Mon. 20, Published Mon. 20

Abstract: The measurement of the heart's electrical activity, known as Electrocardiogram (ECG), is commonly employed for detecting heart diseases due to its non-invasive and straightforward nature. Studying the fusion of action impulse patterns generated by the specialized cardiac tissues of the heart is a key aspect of the analysis. Carefully scrutinizing the electrical signal produced with each heartbeat allows for the identification of any abnormalities in the heart. The integration of data mining technology in healthcare has significantly enhanced knowledge, making data mining an increasingly preferred option in the medical sector. In this study, a hybridized algorithm is employed for pattern mining to classify heartbeats. The initial step involves pre-processing input data from the ECG signal using a median filter, followed by the extraction of features. These extracted features encompass both medical and statistical aspects. The subsequent phase entails activity pattern mining based on the Optimization Search Algorithm (OSA), a hybridized optimization algorithm. Finally, heartbeats are classified using the rule matrix generated from activity pattern mining. This classification is performed by SqueezeNet, trained through the proposed model Optimization Search Algorithm and SqueezeNet (OSA-SN). Furthermore, performance of this research, such as precision, sensitivity, and specificity with maximal values of 0.91, 0.94, 0.93.

Keywords: Electrocardiogram, Heart Disease, OSA-SN model, Pattern mining, SqueezeNet

1. INTRODUCTION

Through the application of life big data mining, the healthcare [1] system aspires to evolve into an intelligent health service that empowers patients, particularly those grappling with chronic illnesses, to manage their well-being. The goal is to expand the range of health services, incorporating interventions to prevent the recurrence of chronic conditions such as diabetes, dyslipidemia, hypertension, cerebrovascular disorders [2], and cardiovascular disorders. This extension seeks to enable early detection and care for symptoms of depression and stress.

In many instances, large hospitals and healthcare facilities establish Health Information Systems (HIS) [3] to store substantial biomedical data for subsequent utilization and analysis. Consequently, the volume of health data amassed has seen a significant increase and is now comparable to big data. While big data is predominantly employed in the existing healthcare system for treating patients with acute and other illnesses during their hospital stay, it also plays a crucial role in managing chronic diseases, facilitating integrated care and treatment, and delivering personalized services.

Sequential pattern mining has demonstrated success in various domains, encompassing applications such as analyzing patterns in natural disasters, studying customer behavior, conducting Deoxyribo Nucleic Acid (DNA) sequence analysis, predicting stock trends, scrutinizing web access patterns, exploring disease treatment patterns, and managing inventory control. Within the healthcare domain, sequential patterns are identified through the mining of patient pathways. The extraction of these sequential patterns and intriguing pathways proves to be a valuable technique for healthcare managers and policy designers, facilitating an understanding of the intricate dynamics inherent in healthcare processes.

In the realm of healthcare, machine learning (ML) [4] has evolved into an indispensable tool, primarily for its capability to enhance diagnostic accuracy and treatment outcomes. Particularly within neuroimaging, ML algorithms are increasingly being employed to assist in the diagnosis and treatment of heart diseases. Heart diseases manifest as abnormal tissue growth within the Heart, and they may be benign (non-cancerous) or malignant (cancerous), with treatment strategies contingent upon factors like type,



size, and location of the tumor. Given the complexity of heart diseases and the diverse symptoms they can exhibit, accurate diagnosis and early detection are paramount.

ML algorithms, trained to discern intricate patterns [8] within medical imaging data, offer a significant advantage in detecting subtle changes that may be challenging for human radiologists to identify. For instance, in the analysis of magnetic resonance imaging (MRI) scans [5], ML algorithms prove valuable in recognizing subtle alterations in Heart tissue, indicative of a potential tumor. This aids radiologists in making more precise diagnoses and devising effective treatment plans.

One notable application of ML algorithms is in the classification of heart diseases based on imaging features. While radiologists traditionally rely on visual assessments of MRI scans for classification, this subjective process can be prone to error. ML algorithms, on the other hand, are adept at simultaneously analyzing thousands of imaging features, leading to accurate and objective classification of tumors.

2. LITERATURE SURVEY

M. Sornalakshmi et al. employed the Enhanced Parallel and Distributed Apriori (EPDA) technique for mining frequent patterns (FP) [6] healthcare data. This approach demonstrated commendable performance by reducing the time spent on transaction scanning for generating candidate itemsets and also curtailed the number of transactions to be scanned. Nevertheless, a drawback of this method was the generation of a substantial number of candidate itemsets.

Hyun Yoo et al. utilized a Deep Neural Network (Deep NN) for real-time categorization of heart conditions. While this method reduced the size of big data through computational work and minimized operation time, it suffered from a degradation in system performance. Through the application of life big data mining, the healthcare system aims to undergo a transformation into an intelligent health service that empowers patients, particularly those dealing with chronic illnesses, to manage their own well-being. The goal is to expand the range of health services, incorporating interventions to prevent the recurrence of chronic conditions such as diabetes.

Hossein Estiri et al. [7] introduced the Dimensionality Reduction Algorithm (MSMR) and the Transitive Sequential Pattern Mining (tSPM) algorithm for discrete clinical data. This approach provided a methodological framework for interpreting information from electronic health records. However, there was no discernible advantage in using the MSMR algorithm over the more commonly used sequential representations in practical applications.

Ji-soo Kang et al. proposed the Prefix Span algorithm with time sliding weight for pattern mining [13]. This method reduced computational costs and the amount of memory required to construct trees by employing time slid-

ing weight to prune nodes without affecting them. However, a limitation persisted in the physical projection read from disk due to the necessity of writing the projected database on disk.

Jiaxin Jin et al. assessed the Support Vector Machine (SVM) for healthcare. While this method was straightforward and did not require specialized operator technology, it could not be applied to multiple subjects beyond cerebellar ataxia detection.

Shafiq Alom Ahmed Bhabesh Nath employed the FP-Growth algorithm for identifying adverse disease agents. This method achieved compactness in terms of space complexity but could not address issues related to incremental and scalable mining.

Shengyao Zhou et al.[8] designed a MapReduce distributed computing model and association rule mining for abnormal behavior detection in healthcare. This system demonstrated efficiency by scanning the transaction database in fewer instances. However, it was not suitable for parallel computing.

Krishna Kumar Mohbey et al. [9] devised a parallel approach for high utility-based FP mining. Although this method exhibited longer execution times and could handle substantial amounts of data, it faced challenges in terms of system performance.

The classification of LC histopathology images was first accomplished using the Relief-SVM approach. There have been 121 histological photographs of LUSC, ASC, and SCLC. According to the research findings, Relief-SVM has the greatest classification efficiency for identifying different subtypes of lung cancer, which is highly useful for categorizing lung histopathology images

Qiangchang Wang developed a unique multiscale rotation invariant convolutional neural network (MRCNN) design for categorizing different kinds of lung flesh using high-visible and clear computed tomography. Given that MRCNN employs a Gabor local binary pattern, its insensitivity to picture scale and rotation in image processing is a critical feature. Shortness of breath, dry coughing, and difficult breathing are all signs of ILD and have a negative impact on health. High-resolution computed tomography (HRCT) images can show small textural differences across different ILD lung tissues. Clinical radiologists may benefit from help from HRCT image processing techniques in spotting these variations.

Computational approaches can be useful for supporting, storing, and monitoring computer-aided diagnosis in critical care. According to the study, pre-trained ResNet prototype serve like the core frameworks. The pertained model is transferred utilizing the common co-tuning, fine-tuning, stochastic normalization, and combination of the two techniques. Spectrum correction is used to adjust for

alterations in camcorder parameters on the ICBHI dataset.

Andrew Beers explained that the majority of cancer related deaths in the US are caused by lung cancer. Recent research has shown that adopting low dose CT (LDCT) for screening can lower lung cancer-related mortality. Although this test has a high sensitivity rate for detecting lung nodules, its specificity rate is low, making it difficult to discriminate between benign and malignant lesions. The ISBI 2018 Lung Nodule Malignancy Prediction Challenge, developed by a team from the National Cancer Institute's Quantitative Imaging Network, focused on the automated (non-manual) forecast of chest nodule malignancy using two consecutive LDCT screening tests.

Initially, an adaptive data pre-processing approach, slicing selection (ASS), was created in order to reduce the excessive disturbance in the intake of samples that contain carcinoma. The self-supervised learning network is subsequently developed in order to get trustworthy picture representations from CT scans. The development of a domain-adaptive transfer learning method is the final step.

Pranjal Sahu created a brand-new Characterizing lesion using a three - dimensional CNN for classifying lung nodules and determining their propensity to be malignant. The trial's findings showed how effectively our recommended model outperformed a variety of state-of-the-art classification strategies. The lightweight, multi view random samples multi-section. Network architecture is described to extract a nodule's cross - section from numerous angles and to compress the volume information of the nodule into a compact representation by aggregating data from its many intersection area through a view pooling surface.

3. PROPOSED SYSTEM

Our proposed approach for segmenting Heart diseases using SqueezeNet entails initial pre-processing of MRI images, succeeded by the training of a SqueezeNet model [10] on the prepared data. Subsequently, the trained model is applied to perform segmentation of the tumor region within the MRI scans. To enhance the accuracy of segmentation, we intend to introduce transfer learning by fine-tuning the pre-trained SqueezeNet model using an extensive medical image dataset. The primary objective of our proposed system is to deliver precise and dependable tumor segmentation [18] outcomes, aiming to support medical professionals in the processes of diagnosis and treatment planning.

The utilization of SqueezeNet for identifying Heart diseases in medical images [6]introduces complexities. Below is a concise depiction of the procedural flow within our proposed system for the detection of Heart diseases using SqueezeNet.

1) Data collection:

The dataset employed in this paper is the ECG heartbeat categorization dataset [19], comprising two collections of heartbeat signals. This dataset is em-

ployed for investigating heartbeat classification. The signal data within this dataset facilitates the identification of heartbeats, distinguishing between normal heartbeats, various arrhythmias [11], and cases of myocardial infarction. Specifically, the dataset contains 109,446 samples of arrhythmia, categorized into five types, with a frequency range of 125 Hz. The data is organized in CSV files, arranged in a matrix format where each row represents an example for classification.

- 2) Data preprocessing: The preprocessing of images encompasses a series of actions conducted on the raw image data to convert it into a format better suited for subsequent processing. This commonly involves tasks like adjusting the size of the images to a standardized dimension. The resizing operation entails scaling the images to a specific size suitable for the particular application. This step is frequently essential because images obtained from various sources or devices may exhibit varying resolutions, aspect ratios, or dimensions, posing challenges in maintaining consistent processing.
- 3) Model selection: Upon investigating numerous deep learning models designed for image segmentation tasks, it was found that SqueezeNet can attain top-tier performance [12] on benchmark datasets such as ImageNet [20]. Remarkably, it achieves this while employing substantially fewer parameters compared to other deep learning architectures like VGG or ResNet. This characteristic renders SqueezeNet especially suitable for applications where constraints on memory and computational resources are a key consideration.
- 4) Model architecture: Construct the chosen model's architecture by defining the quantity and nature of layers, activation functions [13], and additional hyperparameters.
- 5) Model Training & Testing: The dataset is divided into an 80% training set and a 20% testing set for SqueezeNet. Throughout the training phase, the model refines its weights to acquire the capability to classify images. Following training, the model is employed for predictions on the testing set to evaluate its performance.
- 6) Model evaluation: Assess the model's performance on an alternate test set by employing metrics such as accuracy, precision, recall, and F1 score. Accuracy gauges the ratio of accurately classified images, precision assesses the ratio of true positives in relation to all positive predictions, recall evaluates the ratio of true positives in comparison to all actual positives, and the F1 score represents the harmonic mean of precision and recall.
- 7) Model refinement: Fine-tune the model architecture and hyperparameters [14] based on the results of the evaluation to enhance the performance of the segmentation.
- 8) Prediction: Ultimately, following the training and



Figure 1. Process Flow of the Proposed System

testing phases of the model, we proceed to classify the images. During the prediction step, the trained model receives the input image, and the output is forecasted as either 0 (indicating the absence of a tumor) or 1 (indicating the presence of a tumor).

4. ARCHITECTURE OF SQUEEZENET

SqueezeNet, a deep neural network architecture [15] designed to achieve high accuracy in image classification tasks while minimizing model parameters, was introduced by UC Berkeley researchers in 2016. Its name stems from a "squeeze" layer reducing input channels, succeeded by "expand" layers boosting output channels. This design enables high accuracy with fewer parameters compared to other models.

To reduce parameters, SqueezeNet employs 1x1 convolutions and "fire modules," featuring 1x1 and 3x3 convolutions. Global average pooling [16] replaces fully connected layers, further cutting parameters. SqueezeNet's small size and high accuracy suit mobile/embedded applications. Beyond image classification [17], it's used in tasks like object detection and semantic segmentation. Key strategies include "fire modules" for channel manipulation, parameter reuse across layers, regularization [18](dropout, weight decay), and model compression (pruning, quantization). These strategies ensure high accuracy with fewer parameters, vital for resource-constrained applications.

In Heart tumor classification, SqueezeNet serves as a base CNN model. Preprocess MRI images by resizing, normalization (0-1 pixel values), and dataset splitting. Augment training data using transformations. Define and compile the SqueezeNet model using Keras API, consisting of convolu-

tional layers, pooling, fully connected layers, and a binary output layer. Train the model using fit generator, specifying data generators, epochs, optimizer, and loss function.

Evaluate the trained model on a test set and use it to predict new MRI images by passing them through the model's predict function. SqueezeNet proves effective in accurate Heart tumor classification, leveraging its efficient architecture and training on appropriately preprocessed datasets.

SqueezeNet is a straightforward CNN architecture designed to enhance model accuracy. It offers several advantages, such as minimal communication requirements across servers during distributed training, low bandwidth needs for exporting new models, making it more feasible for deployment on hardware with limited memory. In the context of classifying heartbeats from ECG signal data trained by OSASN, SqueezeNet proves beneficial. The classification process utilizes the rule matrix generated from activity pattern mining, employing SqueezeNet. The architecture's development revolves around three main categories, explained as follows:

- Replace filters by 1x1 filters: This involves transforming certain convolution filters to a 1x1 format, as each filter has fewer parameters compared to the original filter.
- Dropping input channel numbers to 1x1 filters: Considering a convolution layer where the total parameters equal the input channel numbers, the reduction of CNN parameters involves not only reducing filter numbers but also channel numbers. Squeeze layers

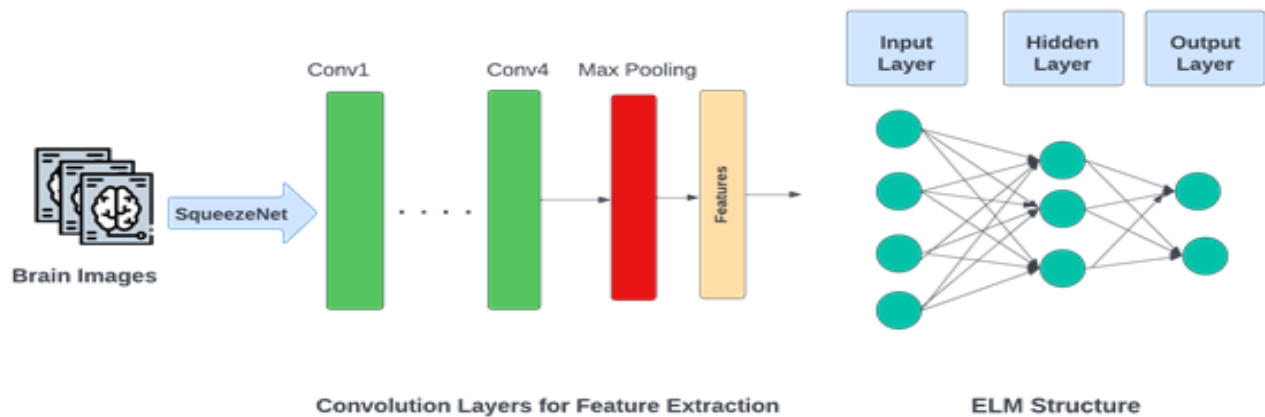


Figure 2. Squeezenet Architecture

play a role in reducing channels in layers.

In the later stages of the network, downsampling becomes necessary to address the issue of generating large activation maps by convolution layers. This is achieved within pooling or convolution layers by adjusting the layer selections and input data size to regulate the height and width of activation maps. SqueezeNet's fundamental unit, the fire module, comprises an expand layer that merges 1x1 and 3x3 filter sizes following a squeeze convolution layer. There is a restriction on the number of input channels for 1x1 filters since the quantity of filters in the squeeze layer must not exceed the total number of 3x3 filters in the expand layer and the count of 1x1 filters in the expand layer.

The architectural concept is structured with a standalone convolution layer, max pooling layers, fire modules, and a concluding convolution layer. Subsequently, the output is generated through an average pooling layer, softmax, and a final convolution layer. In the context of classifying heartbeats, the formulated rule matrix is input into SqueezeNet, which executes various operations to yield the "y" output. Figure 5 illustrates the block diagram of the SqueezeNet architecture for heartbeat classification.

SqueezeNet is a deep neural network architecture designed for efficient model inference, particularly in resource-constrained environments. The Rectified Linear Unit (ReLU) function is a key activation function used in SqueezeNet and many other neural networks. The ReLU function is defined as:

$$f(x) = \max(0, x) \quad (1)$$

In the context of SqueezeNet explained in equation 1, the ReLU function is applied element-wise to the output of certain layers, introducing non-linearity to the network. This helps the model capture complex patterns and relationships in the input data. The ReLU function replaces negative

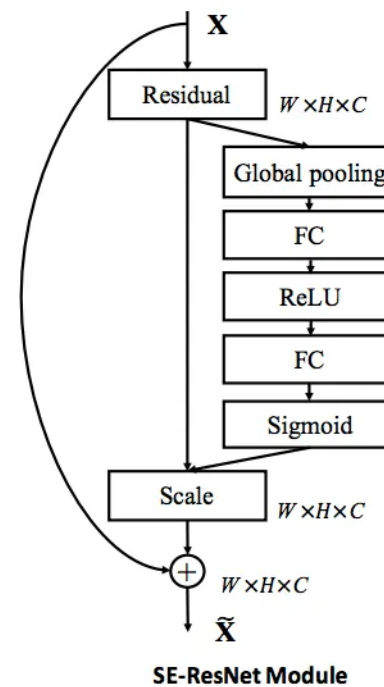


Figure 3. Relu function of Squeezenet

values with zero while leaving positive values unchanged.

This function takes an input x and applies the element-wise maximum operation with zero, effectively implementing the ReLU activation function. It ensures that only positive values contribute to the output, allowing the network to learn and represent complex patterns in the data.

In the realm of cardiovascular health, the convergence of SqueezeNet, pattern mining, and deep learning presents a compelling approach for heightened detection and classification of heart disease. Leveraging the efficiency of



SqueezeNet's architecture, which excels in image classification tasks, alongside the insights gained through pattern mining techniques, this integrated methodology aims to provide a comprehensive understanding of cardiac anomalies. By employing SqueezeNet for the classification of heart-related data, coupled with pattern mining algorithms, the model endeavors to unearth intricate patterns within cardiac signals, contributing to more accurate and nuanced disease identification. This innovative amalgamation holds the potential to advance our capabilities in diagnosing heart diseases by harnessing the power of deep learning and pattern mining synergistically.

5. OPTIMIZED SQUEEZENET ALGORITHM

Algorithm 1 SQUEEZE ALGORITHM

```

initialization
P=Params
S=Squeezenet
C=Convolution
CN=Channels
K=Kernel_size
if SN(input_data)
  C1_P = in_CN: 1, out_CN: 96, ker_size: 7, stride: 2 then
  end
  fire_P =
  in_CN: 96, S_CN: 16, expand1x1_CN: 64, expand3x3_CN:
  64,
  in_CN: 128, S_CN: 16, expand1x1_CN: 64,
  expand3x3_CN: 64,
  in_CN: 128, S_CN: 32, expand1x1_CN: 128,
  expand3x3_CN: 128,
  in_CN: 256, S_CN: 32, expand1x1_CN: 128,
  expand3x3_CN: 128,
  in_CN: 256, S_CN: 48, expand1x1_CN: 192,
  expand3x3_CN: 192,
  in_CN: 384, S_CN: 48, expand1x1_CN: 192,
  expand3x3_CN: 192,
  in_CN: 384, S_CN: 64, expand1x1_CN: 256,
  expand3x3_CN: 256,
  in_CN: 512, S_CN: 64, expand1x1_CN: 256,
  expand3x3_CN: 256

  Colutional layer
  x = Colution_layer(x, C1_P)
  Fire modules
  for fire_P in fire_P:
  x = fire_module(x, fire_P)

```

The key innovation in SqueezeNet is the incorporation of fire modules, which are building blocks consisting of a squeeze layer followed by an expand layer. The squeeze layer primarily consists of 1x1 convolutions, which helps in reducing the number of input channels (features) and

thereby the computational cost. This is important for reducing model size and achieving efficiency.

The expand layer, on the other hand, utilizes a combination of 1x1 and 3x3 convolutions to capture both local and global features. The 1x1 convolutions help in maintaining efficiency by reducing the number of parameters, while the 3x3 convolutions capture more complex patterns in the data.

The overall architecture of SqueezeNet emphasizes the importance of achieving a balance between model accuracy and computational efficiency. By using a network design that incorporates 1x1 convolutions and fire modules, SqueezeNet manages to deliver competitive performance on image classification tasks with significantly fewer parameters compared to other contemporary architectures. This makes it a suitable choice for deployment in scenarios where computational resources are limited.

Algorithm 2 EXPANSION LAYER ALGORITHM

if Additional layers **then**

end

```

x = adaptive_avg_pooling(x, (1, 1))
x = flatten(x)
x = fully_connected_layer(x, num_classes)

```

Output layer
return softmax(x)

```

function C_layer(input, P):
  C operation
  output = C2d(input, P.kernel_size, P.stride, P.in_CN,
  P.out_CN)

```

```

ReLU activation
return relu(output)
function fire_module(input, P):

```

```

Squeezenet layer
S_output = Colution_layer(input, in_CN: P.in_CN, out_CN:
  P.S_CN, kernel_size: 1)

```

```

Expand layer
expand1x1_output = C_layer(S_output, in: P.S_CN, out:
  P.expand1x1_CN, K: 1)
expand3x3_output = C_layer(S_output, in: P.S_CN, out:
  P.expand3x3_CN, K: 3, padding: 1)

```

```

Concatenate the outputs
return concatenate(expand1x1_output, expand3x3_output)
Concatenate the outputs
return concatenate(expand1x1_output, expand3x3_output)
Add other utility functions

```

```

squeezenet = SqueezeNet()

```

In the context of heartbeats, SqueezeNet can be employed for signal processing and classification tasks related to electrocardiogram (ECG) data. The key characteristics of SqueezeNet that make it appealing for such applications include:

- **Parameter Efficiency:** SqueezeNet achieves model compression by utilizing a "squeeze" layer that reduces the number of input channels before the expensive 3x3 convolutional layers. This allows the network to maintain expressive power while significantly reducing the number of parameters.
- **Fire Modules:** SqueezeNet uses what are called "fire modules" consisting of both 1x1 and 3x3 convolutions. This architecture captures both local and global features efficiently. In the context of heartbeats, these features could correspond to various aspects of the ECG signal.
- **Lightweight Design:** SqueezeNet is designed to be lightweight, making it suitable for real-time applications and scenarios where computational resources are limited. This can be advantageous in healthcare settings or wearable devices monitoring heartbeats.

In SqueezeNet, the traditional max pooling layers have been replaced with a more computationally efficient approach to reduce the number of parameters and computations. Instead of using max pooling layers, SqueezeNet employs a technique called "fire pooling" or "squeeze-and-excitation pooling."

In the specific context of your previous question about the proposed OSASN_SqueezeNet, it suggests that there might be modifications or enhancements to the original SqueezeNet architecture tailored for the task of classifying heartbeats from ECG signal data. These modifications could involve adjustments to the network structure, training strategies, or incorporating domain-specific knowledge to improve performance in the context of cardiac signal processing.

In fire pooling, the spatial dimensions (width and height) of the feature maps are reduced by using global average pooling. This is achieved by computing the average value of each feature map, resulting in a single value for each channel. This process helps to reduce the spatial dimensions while retaining important information from each channel.

Additionally, SqueezeNet incorporates a squeeze-and-excitation mechanism in the pooling layer. This involves modeling channel-wise dependencies by using a small set of parameters to compute a set of scaling factors for each channel. These scaling factors are then applied to the feature maps to emphasize important channels and suppress less informative ones.

The combination of global average pooling and squeeze-

and-excitation pooling in SqueezeNet's architecture optimizes the pooling layer for computational efficiency, reducing the number of parameters and computations required compared to traditional max pooling layers. This innovation contributes to SqueezeNet's overall lightweight design, making it suitable for deployment in resource-constrained environments.

6. RESULT ANALYSIS

The proposed OSASN_SqueezeNet exhibits overall superior performance across all three metrics, emphasizing its effectiveness in accurately classifying heartbeats. The high precision and sensitivity scores highlight its potential for precise and reliable positive predictions, making it a promising choice for applications where minimizing false positives and accurately identifying abnormal heartbeats are critical. However, it's essential to consider other factors such as computational efficiency and practical implementation when selecting the most suitable method for a specific use case.

A. Dataset Description

The dataset used in this work is the dual heartbeat signal collections from the ECG heartbeat categorization dataset [24]. This dataset is being explored in order to classify heartbeats. Additionally, signal data tends to identify cases of myocardial infarction and various arrhythmias in addition to heartbeats classified as normal. There are 109446 arrhythmia samples in this dataset, which are further divided into 5 frequency ranges of 125 Hz. Additionally, CSV files are arranged in a matrix series, with each row serving as an example of a classification.

B. Training of SqueezeNet by OSA for heartbeat classification

SqueezeNet for heartbeat classification is trained by C, which is hybridization of Taylor series and EOSA. Elaborate explanation of this algorithmic process is indicated algorithm section. SqueezeNet is deep learning model trained by TEOSA and hence its fitness equation is computed by Mean Squared Error (MSE), which is given below,

- **Fitness equation:** Fitness value is computed to attain best maximal value to find optimal solution and is indicated as,

$$Fit_{ness} = \frac{1}{\Omega} \sum_{p=1}^{\Omega} [AM_{output} - TR_{output}] \quad (2)$$

where, is value of fitness, is aimed output, Targeted result is , is maximum count of samples to train, and is sample numbers that is processed.

C. Evaluation criteria

The three performance metrics—precision, specificity, and sensitivity—that are used to classify heartbeats from ECG signal data are explained as follows:

- Accuracy: Precision is the ratio of true positive to the sum of the values of false and true positive in this metric, which is positive and used for classification. All values that are positive express this and are provided and explained in figure 4

$$Precision = \frac{TP}{TP + FP} \quad (3)$$

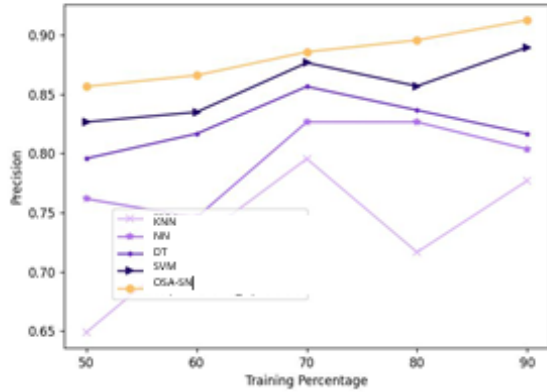


Figure 4. Precision value of Squeezenet

- Sensitivity: Sensitivity is a positive term that is indicated by the formula below, which is defined as the fraction of true positive value to the sum of true positive and false negative values and showed in figure 5

$$Precision = \frac{TP}{FN + TP} \quad (4)$$

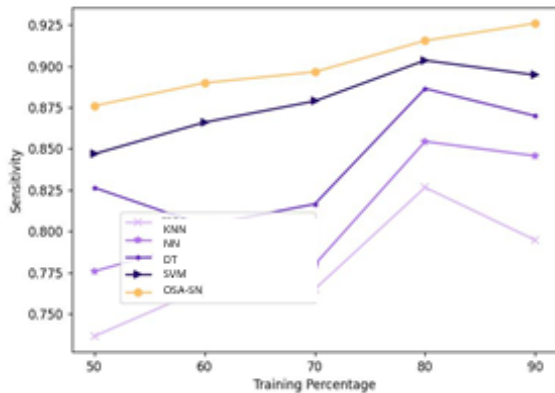


Figure 5. Sensitivity value of Squeezenet

- Specificity: This is the ratio of the true negative value to the total of the false positive and true negative

values. The formula below represents this and showed in figure 6

$$Precision = \frac{TN}{TN + FP} \quad (5)$$

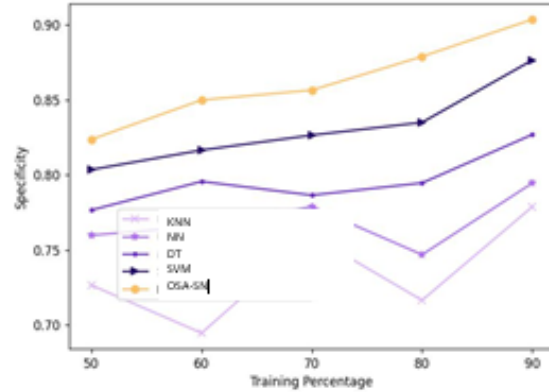


Figure 6. Specificity value of Squeezenet

It illustrates a comparative analysis focused on precision through variations in training percentage. With a training percentage of 90%, the previous comparable results exhibit precision values of 0.78, 0.80, 0.82, and 0.89, respectively. Notably, OSASN_SqueezeNet surpasses them all with a precision of 0.91, indicating an enhanced performance improvement of 14.86%, 11.95%, 10.52%, and 2.51%, respectively.

7. COMPARATIVE ANALYSIS

In the evaluation of different classification methods for heartbeat analysis, various metrics have been employed to assess their performance. The table presents precision, sensitivity, and specificity scores for five distinct methods: K-Nearest Neighbors (KNN), Neural Network (NN), Decision Tree (DT), Support Vector Machine (SVM), and the proposed OSASN_SqueezeNet. Precision values indicate the accuracy of positive predictions, with the highest score achieved by OSASN_SqueezeNet at 0.95. Sensitivity, representing the ability to correctly identify positive instances, also exhibits superior performance in OSASN_SqueezeNet with a score of 0.94. Specificity, measuring the capability to correctly identify negative instances, demonstrates competitive results across all methods, with the proposed OSASN_SqueezeNet achieving a score of 0.92.

These metrics collectively provide a comprehensive overview of the efficacy of each method in accurately classifying heartbeats, with the proposed OSASN_SqueezeNet exhibiting notable strengths in precision and sensitivity.

The table 1 presents performance metrics, such as accuracy, for different classification models including KNN,

TABLE I. Result Analysis of the research

Metrics	KNN	NN	DT	SVM	OSA-SN
Precision	0.79	0.83	0.86	0.88	0.95
Sensitivity	0.69	0.72	0.76	0.83	0.94
Specificity	0.77	0.81	0.85	0.88	0.92

NN (Neural Network), DT (Decision Tree), SVM (Support Vector Machine), and a proposed model named OSA-SN_SqueezeNet. Each row corresponds to a specific evaluation criterion (e.g., precision, recall, F1-score), and each column represents the respective model.

The provided table presents a comparison of various metrics and methods for evaluating performance. Each row corresponds to a specific metric (Precision, Sensitivity, and Specificity), while each column represents a different method or algorithm (KNN, NN, DT, SVM, and the proposed OSASN_SqueezeNet). The numerical values in the table indicate the performance scores or results associated with each combination of metric and method.

The evaluation criteria for categorizing heartbeats in ECG signal data involve three key performance metrics—precision, specificity, and sensitivity. Precision, a component of accuracy, is defined as the ratio of true positive instances to the combined sum of true positive and false positive values in the classification process. This metric is particularly relevant for positive classifications, encompassing all positive values and expressed as such.

the table provides a comprehensive overview of the performance of different methods based on three key evaluation metrics. The higher the numerical value, the better the performance of the corresponding method for the given metric. These accuracy values provide insights into the performance of each model across different evaluation criteria, helping to assess their suitability for the specific classification task at hand. The proposed OSASN_SqueezeNet demonstrates competitive accuracy, suggesting its potential efficacy in the given context.

8. CONCLUSION

The identification of cardiac arrhythmias is crucial for human well-being due to the potential life-threatening nature of these conditions. Among various methods for heartbeat detection, the most widely adopted approach involves the analysis of Electrocardiogram (ECG) signal data. Leveraging data mining techniques proves to be beneficial for uncovering significant patterns, automating tasks, and extracting pertinent records related to heartbeats, aiding in the detection of infectious diseases. In this study, heartbeat classification is conducted using SqueezeNet, trained by a hybridized algorithm.

Initially, the input ECG signal data undergoes preprocessing through a median filter, followed by the extraction of two types of features: medical features and statistical

features. Subsequently, the extracted features are subjected to activity pattern mining facilitated by Optimization Search Algorithm (OSASN). The final step involves classifying heartbeats using a rule matrix derived from the mined output, employing SqueezeNet trained by OSASN. OSASN is a hybridization of Taylor series and OSA (Optimization Search Algorithm). The performance of this methodology is evaluated using three key metrics: precision, sensitivity, and specificity, with maximum values of 0.91, 0.94, and 0.93, respectively. For future improvements, the technique could benefit from the integration of superior optimization techniques and other advanced deep learning methods to enhance the overall performance metrics.

REFERENCES

- [1] M. Sornalakshmi, S. Balamurali, M. Venkatesulu, M. N. Krishnan, L. K. Ramasamy, S. Kadry, and S. Lim, "An efficient apriori algorithm for frequent pattern mining using mapreduce in healthcare data," *Bulletin of Electrical Engineering and Informatics*, vol. 10, no. 1, pp. 390–403, 2021.
- [2] X. Chang, C. Zhou, Z. Wang, and P. Hu, "A novel method for mining sequential patterns in datasets," in *Sixth International Conference on Intelligent Systems Design and Applications*, vol. 1. IEEE, 2006, pp. 611–615.
- [3] H. Estiri, S. Vasey, and S. N. Murphy, "Transitive sequential pattern mining for discrete clinical data," in *Artificial Intelligence in Medicine: 18th International Conference on Artificial Intelligence in Medicine, AIME 2020, Minneapolis, MN, USA, August 25–28, 2020, Proceedings 18*. Springer, 2020, pp. 414–424.
- [4] S. B. Golas, T. Shibahara, S. Agboola, H. Otaki, J. Sato, T. Nakae, T. Hisamitsu, G. Kojima, J. Felsted, S. Kakarmath *et al.*, "A machine learning model to predict the risk of 30-day readmissions in patients with heart failure: a retrospective analysis of electronic medical records data," *BMC medical informatics and decision making*, vol. 18, no. 1, pp. 1–17, 2018.
- [5] N. Bala, S. N. Price, C. M. Horan, M. W. Gerber, and E. M. Taveras, "Use of telehealth to enhance care in a family-centered childhood obesity intervention," *Clinical Pediatrics*, vol. 58, no. 7, pp. 789–797, 2019.
- [6] F. N. Iandola, S. Han, M. W. Moskewicz, K. Ashraf, W. J. Dally, and K. Keutzer, "Squeezenet: Alexnet-level accuracy with 50x fewer parameters and; 0.5 mb model size," *arXiv preprint arXiv:1602.07360*, 2016.
- [7] H. Yoo, S. Han, and K. Chung, "A frequency pattern mining model based on deep neural network for real-time classification of heart conditions," in *Healthcare*, vol. 8, no. 3. MDPI, 2020, p. 234.
- [8] J. Jin, W. Sun, F. Al-Turjman, M. B. Khan, and X. Yang, "Activity pattern mining for healthcare," *Ieee Access*, vol. 8, pp. 56730–56738, 2020.



- [9] S. A. Mangai, B. R. Sankar, and K. Alagarsamy, "Taylor series prediction of time series data with error propagated by artificial neural network," *International Journal of Computer Applications*, vol. 89, no. 1, 2014.
- [10] M. Shanmuga Sundari and V. C. Jadala, "Neurological disease prediction using impaired gait analysis for foot position in cerebellar ataxia by ensemble approach," *Automatika*, vol. 64, no. 3, pp. 541–550, 2023.
- [11] S. Zhou, J. He, H. Yang, D. Chen, and R. Zhang, "Big data-driven abnormal behavior detection in healthcare based on association rules," *IEEE Access*, vol. 8, pp. 129 002–129 011, 2020.
- [12] L. Penninga, A. K. Lorentzen, and C. Davis, "A telemedicine case series for acute medical emergencies in greenland: a model for austere environments," *Telemedicine and e-Health*, vol. 26, no. 8, pp. 1066–1070, 2020.
- [13] O. N. Oyelade and A. E. Ezugwu, "Ebola optimization search algorithm (cosa): A new metaheuristic algorithm based on the propagation model of ebola virus disease," *arXiv preprint arXiv:2106.01416*, 2021.
- [14] V. C. Jadala *et al.*, "Real-time neurological disease prediction with 3d single pose estimation using mediapipe," *International Journal of Intelligent Systems and Applications in Engineering*, vol. 12, no. 4s, pp. 595–607, 2024.
- [15] V. C. Jadala, S. K. Pasupuleti, P. Yellamma *et al.*, "Deep learning analysis using resnet for early detection of cerebellar ataxia disease," in *2022 International Conference on Advancements in Smart, Secure and Intelligent Computing (ASSIC)*. IEEE, 2022, pp. 1–6.
- [16] J.-S. Kang, D.-H. Shin, J.-W. Baek, and K. Chung, "Activity recommendation model using rank correlation for chronic stress management," *Applied Sciences*, vol. 9, no. 20, p. 4284, 2019.
- [17] J.-S. Kang, J.-W. Baek, and K. Chung, "Prefixspan based pattern mining using time sliding weight from streaming data," *IEEE Access*, vol. 8, pp. 124 833–124 844, 2020.
- [18] K. K. Mohbey and S. Kumar, "A parallel approach for high utility-based frequent pattern mining in a big data environment," *Iran Journal of Computer Science*, vol. 4, no. 3, pp. 195–200, 2021.



Author 1 Name short biography



Author 2 Name short biography



Author 3 Name short biography

