

Date of publication xxxx 00, 0000, date of current version xxxx 00, 0000.

Digital Object Identifier 10.1109/ACCESS.2017.DOI

Innovations in Stroke Identification: A Machine Learning-Based Diagnostic Model Using Neuroimages

MUHAMMAD ASIM SALEEM¹, ASHIR JAVEED², WASAN AKARATHANAWAT^{3,4}, AURAUMA CHUTINET^{3,4}, NIJASRI CHARNNARONG SUWANWELA^{3,4}, WIDHYAKORN ASDORNWISE¹, SURACHAI CHAITUSANEY¹, SUNCHAI DEELERTPAIBOON¹, WATTANASAK SRISIRI¹, WATIT BENJAPOLAKUL¹ and PASU KAEWPLUNG¹

¹Center of Excellence in Artificial Intelligence, Machine Learning, and Smart Grid Technology, Department of Electrical Engineering, Faculty of Engineering, Chulalongkorn University, Bangkok 10330, Thailand

²Aging Research Center, Karolinska Institutet, 171 65 Stockholm, Sweden

³Division of Neurology, Department of Medicine, Faculty of Medicine, Chulalongkorn University, Bangkok 10330, Thailand

⁴Chulalongkorn Stroke Center, Chula Neuroscience Center, King Chulalongkorn Memorial Hospital, Thai Red Cross Society, Bangkok 10330, Thailand

Corresponding author: Pasu Kaewplung (Pasu.K@chula.ac.th) and Aurauma Chutinet (Aurauma.C@chula.ac.th)

This research has received funding support from the National Science, Research and Innovation Fund (NSRF) via the Program Management Unit for Human Resources and Institutional Development, Research and Innovation No. B04G640068]. This research is also supported by Ratchadapisek Somphot Fund for Center of Excellence in Artificial Intelligence, Machine Learning, and Smart Grid Technology, and by the Second Century Fund (C2F), Chulalongkorn University, Bangkok, Thailand.

ABSTRACT Cerebrovascular diseases such as stroke are among the most common causes of death and disability worldwide and are preventable and treatable. Early detection of strokes and their rapid intervention play an important role in reducing the burden of disease and improving clinical outcomes. In recent years, machine learning methods have attracted a lot of attention as they can be used to detect strokes. The aim of this study is to identify reliable methods, algorithms, and features that help medical professionals make informed decisions about stroke treatment and prevention. To achieve this goal, we have developed an early stroke detection system based on CT images of the brain coupled with a genetic algorithm and a bidirectional long short-term Memory (BiLSTM) to detect strokes at a very early stage. For image classification, a genetic approach based on neural networks is used to select the most relevant features for classification. The BiLSTM model is then fed with these features. Cross-validation was used to evaluate the accuracy of the diagnostic system, precision, recall, F1 score, ROC (Receiver Operating Characteristic Curve), and AUC (Area Under The Curve). All of these metrics were used to determine the system's overall effectiveness. The proposed diagnostic system achieved an accuracy of 96.5%. We also compared the performance of the proposed model with Logistic Regression, Decision Trees, Random Forests, Naive Bayes, and Support Vector Machines. With the proposed diagnosis system, physicians can make an informed decision about stroke.

INDEX TERMS Stroke, Feature Selection, Genetic Algorithm, LSTM, BiLSTM, CT images

I. INTRODUCTION

A stroke occurs when the blood supply to the brain is blocked or a blood vessel in the brain bursts. It is the most common cause of death worldwide, causing over 6.2 million deaths annually [1]. Survivors often suffer from disabilities that severely impair their quality of life. Preventive measures and timely intervention can help to reduce the severity of stroke and its impact on the individual. Early identification of people at risk is crucial for stroke prevention [2]. There are

two types of stroke: ischemic stroke and hemorrhage-related stroke. About 87% of strokes are ischemic strokes, the most common type [3]. When a blood vessel supplying blood to the brain is blocked or narrowed, blood flow to a specific part of the brain is reduced or interrupted. The blockage may be caused by a blood clot, cholesterol deposits, or other debris entering the brain from other parts of the body [4]. Strokes caused by hemorrhages are less common, but they are more serious and often fatal if not treated properly. A

cerebral hemorrhage occurs when a blood vessel in the brain bursts or leaks, causing blood to leak into the surrounding brain tissue. This leads to swelling and pressure on the brain structures. This can damage brain cells and disrupt normal brain function [5], as shown in the figure 1. This study uses machine learning techniques to develop a systematic method for stroke detection. The study's main aim is to understand the causes of strokes better and create reliable detection models that help physicians make informed decisions about stroke prevention and treatment. The study will analyze electronic health records, genetic information, and lifestyle data to develop predictive models identifying individuals at risk of stroke. The research will use feature selection algorithms to extract valuable insights from the data and improve detection performance. A machine learning model will also support clinicians.

The aim of this study is to evaluate the effectiveness, clarity, adaptability, and scalability of the proposed algorithms. The proposed models were compared with established clinical risk detection approaches using numerous real-world datasets to determine their comparative performance. This study can provide valuable insights for clinical practice and tailored interventions to reduce the impact of stroke on patients and the healthcare system. The aim of this study is to use machine learning techniques to predict the frequency of strokes and ultimately solve the problem of stroke occurrence. This study uses advanced machine learning techniques to improve the accuracy and comprehensibility of stroke detection models. The results of this study have the potential to revolutionize stroke prevention by enabling precise interventions to prevent this debilitating disease in the future, thus improving treatment outcomes for patients suffering from this debilitating disease. The key contributions of this study are listed below:

- 1) This study presents a diagnostic system for stroke detection using an image-based dataset.
- 2) The proposed diagnostic system extracts useful features from the CT images via genetically optimized CNN.
- 3) LSTM, BiLSTM, and genetically optimized CNNs were compared for classifying strokes based on extracted features.
- 4) The performance of the proposed diagnostic system (GA_BiLSTM) is also compared with other ML and DL methods.
- 5) The proposed system aims to aid healthcare professionals in making informed decisions about stroke treatment and prevention by providing early detection.

II. RELATED WORKS

Machine learning algorithms have been used to develop highly accurate stroke risk detection models. These models use historical data from electronic medical records to predict stroke risk in hypertensive patients [6]. Various machine learning algorithms, such as Extreme Gradient Boosting (XGBoost), were used to achieve optimal performance in

predicting stroke risk [7], [8]. The use of machine learning models, including ensemble models that combine different approaches, has shown promising results in identifying individuals at high risk of stroke [9].

The authors of [10] thoroughly review skull stripping techniques, including more modern deep learning-based techniques. This review divides the existing approaches to skull stripping into two categories: Deep learning or convolutional neural networks and traditional or classical approaches. Several techniques are emphasized for their potential as they can be integrated into common clinical imaging techniques.

The authors in [11] present a time-based link prediction model called TDRL that uses deep reinforcement learning techniques to learn from a real-world evolving crime dataset. The experiments show that the TDRL model trained on a temporal dataset has better prediction accuracy than other machine learning models trained on a single point in time. Although the TDRL-CNA model accurately predicted most edges in the network topology during the year of the bombing, it failed to predict disappearing edges. Feature analysis revealed that the TDRL-CNA model was least affected by learning based on the removed edges.

Criminal network analysis (CNA) is difficult due to incomplete datasets, and most machine learning techniques rely on supervised learning. The authors investigated the use of deep reinforcement learning (DRL) to build a model for predicting hidden connections in criminal networks. The results show that this approach performs better than traditional supervised machine learning techniques [12].

A study in [13] examined clinical brain data CT and predicted multiple stroke scores after 24 hours or a Modified Rankin Scale score from 0 to 1 over 90 days ("mRS90") using a National Institutes of Health-derived CNN hybrid structure for artificial neural networks. With this structure, they achieved a detection accuracy of 74% for the mRS90.

An integrated wavelet entropy-based spider network graph was used with a probabilistic neural network to classify MRI images of the brain into normal brains, strokes, and degenerative diseases [14]. The authors systematically examined diseases, infectious diseases, and brain tumors as part of their study. The first step was to use a discrete wavelet transform to process brain images in two dimensions (2D). They extracted features using spider web diagrams and classified them using a probabilistic neural network. They reported that they achieved a classification accuracy of 100%.

The authors of [15], [16] want to improve the diagnosis and management of strokes by analyzing various factors in electronic health records. To better predict strokes, they use statistics and principal component analysis. Strokes are most commonly found in patients with advanced age, heart disease, average blood sugar levels, and high blood pressure. The paper also suggests using a perceptron neural network with these four attributes to achieve the highest accuracy and lowest error rate compared to other benchmarking algorithms. In addition, the authors address the problem of unbalanced datasets by presenting results on a balanced dataset

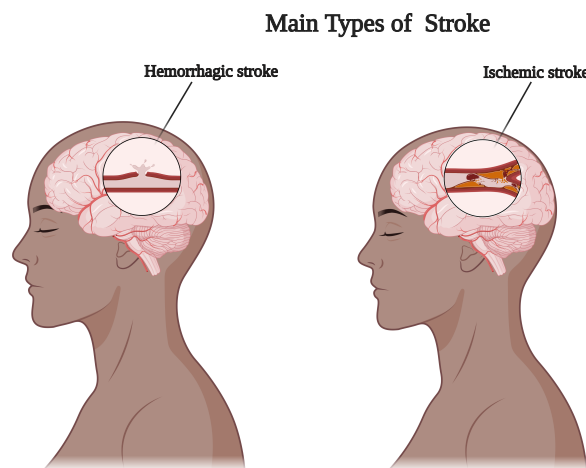


FIGURE 1: Main types of stroke.

created using sub-sampling techniques.

The deep learning framework proposed in [17] has been shown to be highly accurate in detecting different subtypes of intracranial hemorrhage (ICH) in CT images of the head. The system achieves an average accuracy of 96.21% for three types of hemorrhages - epidural, subdural, and intraparenchymal. Compared to existing work, the false positive rate is significantly reduced. In addition, the system includes a quantitative scoring algorithm that automatically measures the thickness and volume of hemorrhagic lesions. This enables clinically relevant quantification, which is important for the decision on emergency surgical treatment.

In [18], a hybrid feature selection approach uses transformed CT image features and grayscale co-occurrence matrix texture features. Features are extracted using discrete wavelet transforms, discrete cosine curves, and GLCMs. The machine learning algorithms Random Tree, Random Forest, and REPTree are used for classification, with Random Forest achieving the highest accuracy of 87.97% for the combination of discrete wavelet transforms and GLCM features.

In [19], an approach to classify CT images with intracranial hemorrhages using machine learning techniques is presented [19]. A set of common features consists of the Gray-Level Co-Occurrence Matrix (GLCM), Discrete Wavelet Techniques (DWT), and Discrete Cosine Techniques (DCT). SMOTE is used to solve the oversampling problem, and sequential forward selection of features is used to obtain subsets of features. Classification accuracy is evaluated using a confusion matrix, precision, and recall. Random Forest achieved the highest accuracy of 87.22 percent in combination with the proposed feature extraction mechanism.

In [20], the authors measure the performance of a deep learning model for detecting intracranial hemorrhage on CT head scans and compare the effects of different preprocessing and model design methods. Implementing preprocessing

techniques and a CNN-RNN framework significantly improved the model's performance and demonstrated its potential as a radiologist decision-support tool.

Based on the literature review conducted, we found that current research can classify strokes with less accuracy. Moreover, the proposed methods were very computationally intensive. However, we propose a diagnostic system that uses a genetic algorithm with bidirectional long short-term memory (BiLSTM) to predict strokes by analyzing brain images on a CT scan. The proposed method solves the problems of lower accuracy and higher computational complexity of the previously proposed models.

III. METHODOLOGY

This section will discuss the methodology we adopted in this study and the framework we have employed for this research. The proposed framework is shown in Figure 2. The framework describes the architecture of the proposed computer-aided stroke diagnosis system. This benchmark dataset is loaded into the system and divided into two sub-datasets: the training and validation datasets. To split the dataset, we use holdout cross-validation. We randomly select 70% of the images from five classes to train the model and split the remaining data. To test the performance of the proposed model, 30% of the images with the corresponding labels are used as validation sets. This study extracts robust and non-invariant features from RGB images using five advanced convolutional network architectures. These architectures include AlexNet, NASNet-Large, VGG-19, Inception V3 and ShuffleNet to extract features from training and validation images. The proposed feature extraction system takes an image as input to the CNN model and extracts a feature vector of 1000 features from the last fully connected layer of the CNN model. CNN models consist of two parts: Feature extraction and classification. For the proposed feature extraction system, an image

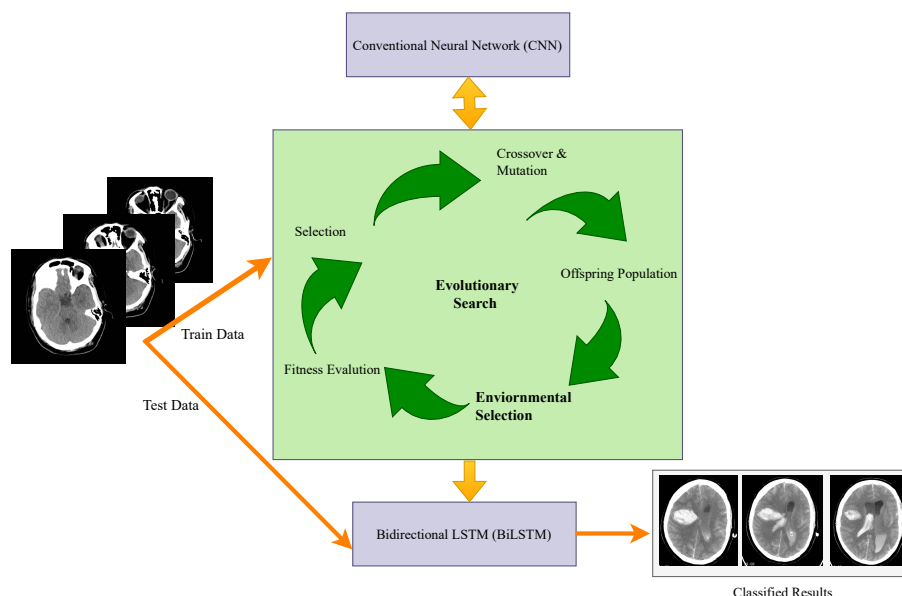


FIGURE 2: Framework of the proposed model.

is used as input. Genetic algorithms based on tournament selection are proposed to determine the ranking of features in terms of their optimal and non-optimal ranking. The vectors of training and validation features are analyzed to remove non-optimal attributes.

A. CNN ARCHITECTURE

CNN is a type of network design commonly used in deep learning algorithms and is mainly used in image recognition and pixel processing to analyze images [21], [22]. Various CNN models and transfer learning [23], [24] approaches have been used to categorize patients with diabetic retinopathy (AlexNet, Inception V3, NASNet-Large, ShuffleNet, and VGG 19). Figure 2 shows that the models and genetic algorithms have adapted to systematically select features [25].

1) AlexNet

The CNN model has about 650000 neurons and 60 million parameters. Many components are used in AlexNet, including fully connected layers, softmax activation, convolutional filters and ReLU transfer functions. AlexNet was introduced in 2012 by Alex Krizhevsky. By adjusting the number of filters and layers in AlexNet, different levels of efficiency can be achieved. Different effects can be achieved by connecting layers in a way that changes the input size. The connections between the layers can be set to different effect strengths. Neural networks consist of neurons that have weights and biases, which makes them convolutional neural networks. The architecture of AlexNet can be seen in Figure 3, which illustrates the architecture of AlexNet by giving an overview of the inputs, adding weighted amounts, and then activating, transmitting, and releasing the neurons.

2) Inception V3

In 2014, Google developed a CNN architecture that can estimate the local sparsity of a network by assembling several available modules. Using these complex modules, the network can estimate the optimal local sparsity. In December 2015, Szegedy released an improved version of the Inception V3 algorithm. Its performance is much better than that of the benchmark. It can identify objects, recognize human actions, segment videos, track objects and identify moving objects. By using convolutional factorizations and auxiliary classifiers, Inception V3 allows the grid size to be reduced to the absolute minimum. The network filtering activation scale is extended to prevent a display bottleneck before maximum or average pooling is used. Auxiliary classifiers can improve accuracy and provide insight into the network. The architecture of Inception V3 can be illustrated in Figure 4.

3) VGG 19

A number of networks are used in the analysis of the CNN-based ImageNet dataset, including the VGGNet network, which stands for Visual Geometry Group Networks. Figure 5 illustrates the VGG 19 network, a simple but versatile network for reducing the volume of datasets. It is a popular imaging method widely used in medical research and commercial imaging applications because it has convolutional layers at the top and max-pooling stages that reduce the volume.

4) NASNet-Large

Artificial neural networks (ANNs) have been influenced by human biases even though they offer valuable insights, even though their architecture is flawed [26]. It is evident from

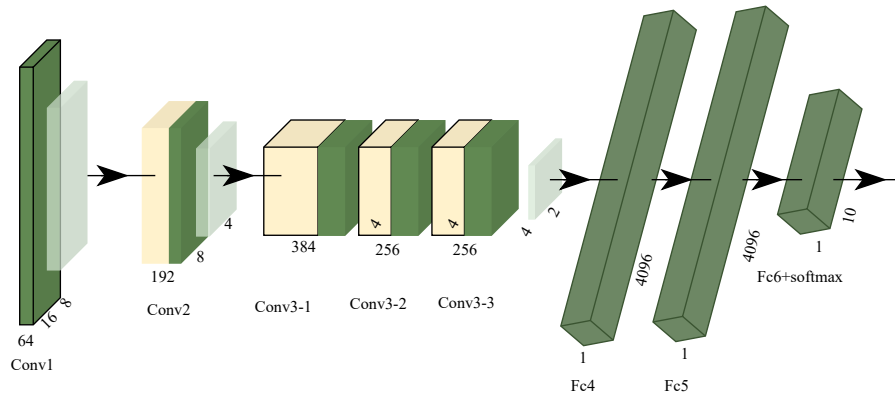


FIGURE 3: The pre-trained architecture of AlexNet.

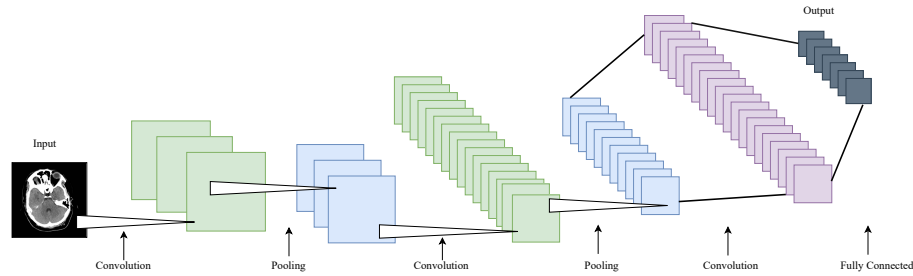


FIGURE 4: Pre-trained Inception V3 Architecture.

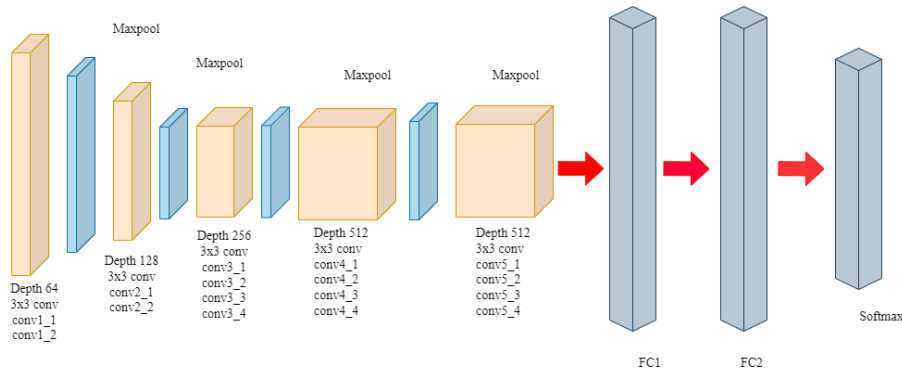


FIGURE 5: Pre-trained VGG 19 architecture.

Figure 6 that NASNet-Large (Google's algorithm for finding the optimal combination of features in a search space) produces extraordinary results when applied to large datasets, such as ImageNet.

5) ShuffleNet

This architecture was developed specifically for mobile devices in 2017. The ShuffleNet architecture can be seen in Figure 7, which is illustrated below. Megvii Inc. research group members worked on the ShuffleNet architecture in 2017. Convolutional methods are used within a network to speed up computations, and shuffled paths are used to reduce the size of the network, which means that the network will

perform more accurately.

B. GENETIC ALGORITHM-BASED FEATURE SELECTION

The genetic algorithm is a metaheuristic method often used for feature selection in deep learning research. This method aims to emulate the selection, mutation and crossover processes of nature in order to find the most reliable and distinct traits while reducing the complexity of the trait space. In the first step, the proposed system generates a random uniform population. The following equation defines the chromosomes as gene features, each of which is assigned a real number. The following equation illustrates how chromosomes are

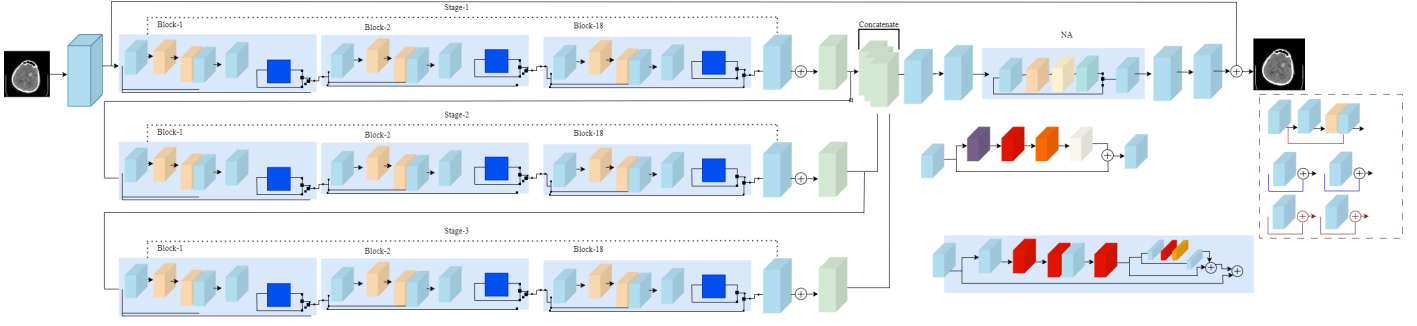
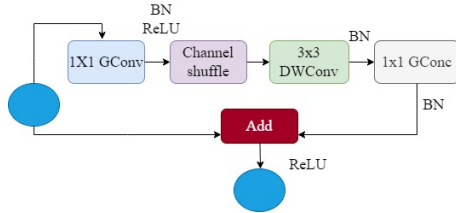
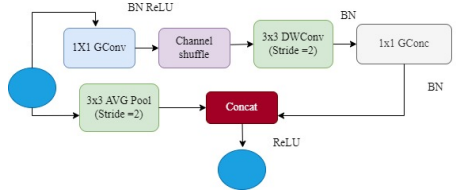


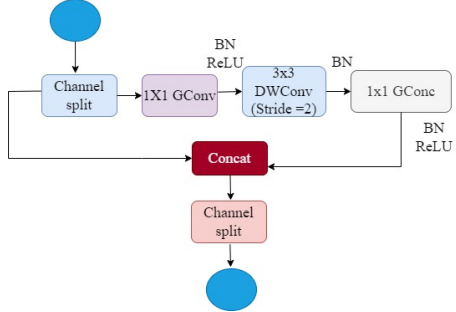
FIGURE 6: NASNet-Large Pre-trained Architecture NASNet-Large.



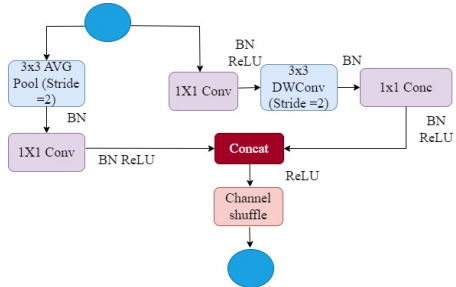
(a) Fundamental ShuffleNet unit



(b) ShuffleNet unit for spatial down sampling



(c) The proposed fundamental ShuffleNet unit



(d) The proposed ShuffleNet unit for spatial down sampling

FIGURE 7: The pre-trained ShuffleNet

represented [27]–[29].

C. CONVENTIONAL LSTM

A typical LSTM analyzes and records information as it propagates, similar to an RNN. An LSTM has a core composed of different gates and cell states, with different cell activities allowing it to ignore or store information depending on the activity level [30]. Different neural networks act as gates and control the information that can be applied to a cell's state during the processing procedure and the way in which it is applied. The state of the cell is a memory channel that is used to transmit relevant data during the processing of the cell. A gate learns which information to store and which to ignore during the training process by the process of training. LSTM cells use input gates output gates, and forget gates to communicate information. A gate determines which information is added to the current state, which is saved from the previous state, and which is added to the next state. An example of an LSTM architecture can be found in Figure 8. The following equations describe the inputs and outputs at time t and time $t-1$.

$$\zeta_t = \vartheta [(\zeta_{\varrho\zeta} * \varrho_t) + (\varsigma_{v\zeta} * v_{t-1}) + (\varsigma_{\nu\zeta} * \nu_{t-1}) + \kappa_{\zeta}] \quad (1)$$

$$\eta_t = \vartheta [(\zeta_{\varrho\eta} * \varrho_t) + (\varsigma_{v\eta} * v_{t-1}) + (\varsigma_{\nu\eta} * \nu_{t-1}) + \kappa_{\eta}] \quad (2)$$

$$\gamma_t = (\eta_t * \gamma_{t-1}) + \zeta_t \tanh [(\varsigma_{v\gamma} * v_{t-1}) + (\varsigma_{\nu\gamma} * \nu_{t-1}) + \kappa_{\gamma}] \quad (3)$$

$$\varphi_t = \vartheta [(\zeta_{\varrho\varphi} * \varrho_t) + (\varsigma_{v\varphi} * v_{t-1}) + (\varsigma_{\nu\varphi} * \nu_{t-1}) + \kappa_{\varphi}] \quad (4)$$

$$\sigma_t = \varphi_t \tanh(\sigma_t) \quad (5)$$

Where the input gate is represented by ζ_t , the input vector is denoted by ϱ , the output gate is φ , the output is ν_t , and the forgetting function is represented by η_t . Cell state is expressed as γ_t , where γ and κ define bias and weight, respectively.

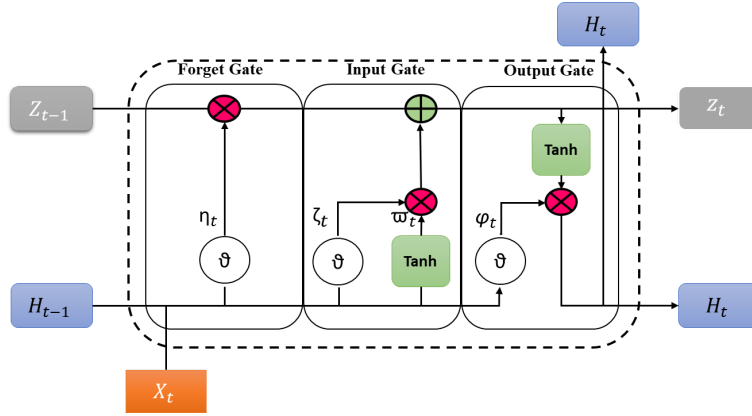


FIGURE 8: LSTM memory cell

1) Bidirectional LSTM (BiLSTM)

The use of bidirectional LSTMs as a supplement to conventional LSTMs improves classification performance. Two LSTMs are trained with the input data. Two sets of input data were trained, one with the original input data and one with the reversed input data. This makes the network more meaningful, and the results are obtained quickly. It is quite straightforward to understand how the BiLSTM works. Creating a duplicate recurrent layer in the network is necessary, forwarding the input data from the original layer to the duplicate layer in reverse order before returning to the original layer. In conventional RNNs, there are vanishing gradients [31] due to this concept.

The BiLSTM is trained over a period of time using all available past and present data for the training period. BiLSTMs are trained with both forward and backward shifts of the input data (from left to right and from right to left). Python's Keras library implements BiLSTMs as a bidirectional layer wrapper that accepts an LSTM layer as input. A combined forward and backward output can be sent to the next layer by specifying the fusion mode, as shown in Figure 9. It is derived from the forward hidden layer $\vec{\lambda}_\tau$ and the backward hidden layer $\overleftarrow{\lambda}_\tau$ as follows: [32]:

$$\vec{\Sigma}_l = L(\varsigma_{\vec{l}} \vec{l}_l + \varsigma_{\vec{l}} \vec{l}_{l-1} + \kappa_{\vec{l}}) \quad (6)$$

$$\overleftarrow{\Sigma}_l = L(\varsigma_{\overleftarrow{l}} \overleftarrow{l}_l + \varsigma_{\overleftarrow{l}} \overleftarrow{l}_{l-1} + \kappa_{\overleftarrow{l}}) \quad (7)$$

$$Z_l = \varsigma_{\vec{l}} \vec{l}_l + \varsigma_{\overleftarrow{l}} \overleftarrow{l}_l + \kappa_z \quad (8)$$

The hidden layer is represented by l , and the matrices of the input weights (forward, backward) and the vectors of the hidden weight bias for both directions are denoted by ς . These vectors are given by $\kappa(\kappa_{\vec{l}}$ and $\kappa_{\overleftarrow{l}}$).

IV. RESULTS AND DISCUSSION

The data sets used in this study, and the experimental evaluations that show the value of the proposed model are presented in this part. Researchers have used a number of validation techniques for statistical evaluation and validation of machine learning models. Some of these are holdout, k-fold cross-validation, and leave-one-out (LOO) cross-validation. However, the main drawback of holdout validation on small datasets is that the training dataset may not accurately reflect the distribution of the dataset; as a result, the results of this method would be biased. Therefore, in situations where only a small amount of data is available, cross-validation is considered the most practical validation scheme [33]–[35]. In this way, we can avoid overfitting and obtain an accurate estimate of the model's performance [36]–[38]. Following previous research, we used the ten-fold cross-validation scheme as well as accuracy, precision, recall, and F1-score as evaluation metrics to evaluate the classification results of the proposed model.

A. DATASET DETAILS

The data set used in this study was obtained from Kaggle [39]. The aim of the study was to identify risk factors for stroke, including age, gender, smoking, diabetes and high blood pressure. It included 100 patients aged 16 years and older, 68% of whom were men and 32% of whom were women. Each patient had CT images of the brain classified as normal or stroke; there were 1551 normal and 950 stroke images. The images were in grayscale and had a size of 650 x 650 pixels. To avoid overfitting during training, the data set was randomly equalised. The result was a new data set with 950 normal and 950 stroke images, each with a resolution of 227 x 227 pixels.

B. EVALUATION METRIC

In this study, along with accuracy, precision, recall, and F1-score have been used as an evaluation parameter [40]–[43]. The findings are included in every report. Following is a

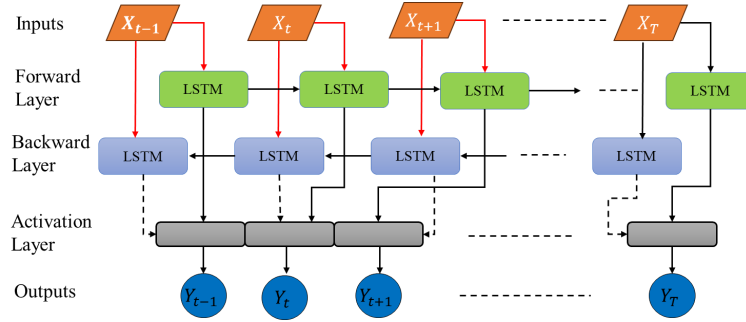


FIGURE 9: BiLSTM architecture

TABLE 1: Performance of ML classifiers.

ML models	Accuracy	Precision	Recall	F1-score
Logistic Regression	75.62	68.56	78.40	75.00
Decision Tree	85.00	80.65	88.34	85.00
Random Forest	89.53	91.50	86.73	90.00
Naive Bayes	66.21	70.00	62.15	61.00
Support Vector Machine	91.45	89.88	92.00	91.00

TABLE 2: Performance of DL classifiers.

DL models	Accuracy	Precision	Recall	F1-score
ANN	92.50	88.56	91.40	92.00
DNN	95.78	95.26	88.62	95.00
RNN	93.66	91.50	89.00	94.00

definition of used evaluation metrics are given as follows:

$$Accuracy = \frac{No. \text{ Correct Prediction}}{Total \text{ Samples}} \quad (9)$$

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

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

$$F1_score = \frac{2TP}{2TP + FP + FN} \quad (12)$$

where TP stands for the number of true positives, FP for the number of false positives, TN for the number of true negatives and FN for the number of false negatives.

C. EXPERIMENTAL RESULTS AND DISCUSSION

This section summarizes the results and identifies the baseline models that are compared with the proposed model. We compare the performance of the proposed model with different Machine Learning (ML) classifiers and Deep Learning (DL) methods to evaluate its performance. In this research, all algorithms were used with their default values. Tables 1 and 2 show the results of all ML and DL classifiers used in the study.

Table 1 shows that the Support Vector Machine outperforms all ML-based classifiers in terms of diagnosis with 91.45% accuracy on the dataset. In contrast, the worst classifier for the diagnosis dataset is Naive Bayes, with an accuracy of 66.21%. Among the DL-based classifiers, DNN performs well with an accuracy of 95.78% for the diagnosis dataset, compared to ANN with an accuracy of 92.50% and RNN with an accuracy of 93.66%, as shown in Table 2.

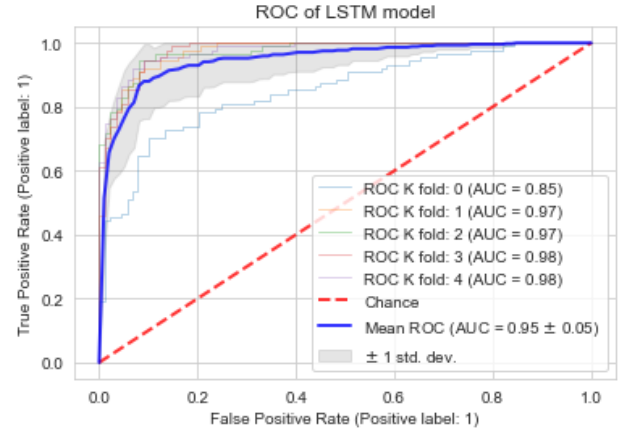


FIGURE 10: ROC of GA_LSTM

To solve the problem of uneven distribution of data and small sample size, we evaluated the experiments of the proposed model where the features extracted from the NN-based genetic algorithm are fed into the LSTM and BiLSTM for stroke detection.

We performed two further experiments using a neural network-based genetic algorithm to extract features from the brain images. In the first experiment, the LSTM was used to classify stroke patients. The best features are selected based on a threshold set after ranking the features. The best-extracted features from the neural network-based genetic algorithm were fed into the LSTM. To evaluate the performance of the GA_LSTM model, we used a cross-validation scheme (K=5) based on the Receiver Operating Characteristic Curve (ROC) [44]–[47]. Figure 10 shows the area under the curve (AUC) of 95% achieved by the GA_LSTM model.

In the next experiments, we used the BiLSTM model with the main features of the neural network-based genetic algorithm; therefore, we named the proposed

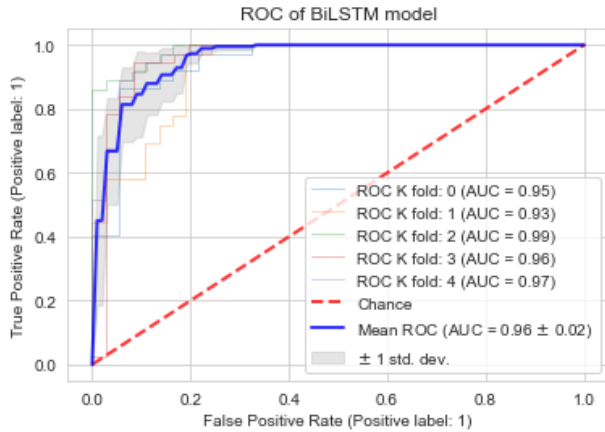


FIGURE 11: ROC of GA_BiLSTM

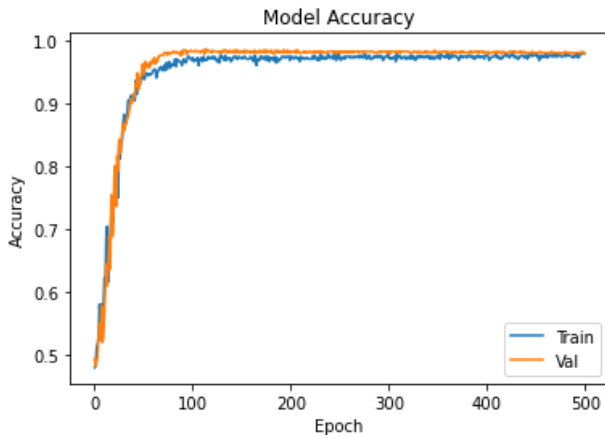


FIGURE 12: Demonstration of validation accuracy of GA_BiLSTM

model GA_BiLSTM. From Figure 11, it's evident that GA_BiLSTM achieved the highest AUC on the ROC curve. In addition, we also evaluated the accuracy performance of the newly developed model(GA_BiLSTM), which is shown in Figure 12.

Furthermore, we have also evaluated the performance of the GA_BiLSTM model along with the GA_LSTM based on accuracy, precision, recall, and F1-score, as given in Table 3. From Table 3, it can be seen that the proposed GA_BiLSTM model shows better performance in comparison to GA_LSTM model as well as other DL models 2 and other ML models from Table 1.

TABLE 3: Performance of proposed work

Models	Accuracy	Precision	Recall	F1-score
GA_LSTM	95.35	92.00	90.89	95.00
GA_BiLSTM	96.45	98.00	93.50	96.00

TABLE 4: Performance comparison between the proposed method and the methods presented in literature

Study (Year)	Method	Accuracy(%)
Sailasya et.al., (2021) [48]	LR	78.00
Sailasya et.al., (2021) [48]	DT	66.00
Sailasya et.al., (2021) [48]	RF	73.00
Aniwat et.al., (2021) [17]	DNN	96.21
Devet et.al., (2022) [15]	SVM	68.00
Devet et.al., (2022) [15]	RF	74.00
Devet et.al., (2022) [15]	CNN	74.00
Santwana et.al., (2022) [18]	RF	87.97
Akter et.al., (2022) [49]	RF	95.30
Tursynova et.al., (2023) [50]	CNN	81.00
Santwana et.al., (2023) [19]	RF	87.22
Yeo et.al., (2023) [20]	RNN + CNN	93.00
Gupta et.al., (2023) [51]	DenseNet-121	96.00
Luis et.al., (2023) [52]	DNN	92.70
Current Study	GA + LSTM	95.35
Current Study	GA + BiLSTM	96.45

D. PREVIOUSLY PROPOSED METHODS FOR STROKE PREDICTION

In this section, we discuss the performance of a newly proposed method for stroke detection and compare it with various machine learning-based methods that researchers have proposed in the recent past. We compared the classification accuracy of the proposed method with other methods that used the same Kaggle dataset of brain CT images for stroke prediction. The comparison of accuracy between the proposed method (GA_BiLSTM) and other methods can be found in Table 4. It is worth mentioning that the newly proposed method has shown a significant improvement in accuracy compared to the studies of researchers (10) who have used machine learning and deep learning techniques for stroke prediction.

V. CONCLUSION

This study proposes a method for stroke detection using machine learning techniques. An image-based dataset is used to validate the performance of the newly developed model. The proposed model is based on a genetic algorithm and BiLSTM. A genetic algorithm based on a neural network is applied to recognize the key features of CT brain images. These features are input into the LSTM and BiLSTM models for stroke prediction. The performance of different K-folds was evaluated to determine which is the most effective classification. We also tested different machine-learning algorithms for stroke prediction. The results of the experiment show that the proposed machine-learning model is more efficient than other models. To improve stroke detection in the future, we aim to use more complex models that can predict strokes automatically. This study used deep learning models on a modest dataset, with large datasets generally performing better. Therefore, to improve the performance of the model, we need to collect more samples in the future to achieve better results. In addition, data quality plays an important role in the performance of deep learning models. So, we need to develop new methods that will help improve

the data quality in the future. It is also necessary to build trust among healthcare professionals through explainable AI techniques to explain the model's decision-making processes. In the future, we will develop these explainable AI methods to help healthcare professionals make more informed decisions based on these AI results.

ACKNOWLEDGMENT

This research has received funding support from the National Science, Research and Innovation Fund (NSRF) via the Program Management Unit for Human Resources and Institutional Development, Research and Innovation [Grant No. B04G640068]. This research is also supported by Ratchadapisek Somphot Fund for Center of Excellence in Artificial Intelligence, Machine Learning, and Smart Grid Technology, and by the Second Century Fund (C2F), Chulalongkorn University, Bangkok, Thailand.

REFERENCES

- [1] W. S. Organization, "Vision impairment and blindness," 2023, accessed: 2023-02-03. [Online]. Available: <https://www.world-stroke.org/world-stroke-day-campaign/why-stroke-matters/learn-about-stroke>
- [2] A. Khosla, Y. Cao, C. C.-Y. Lin, H.-K. Chiu, J. Hu, and H. Lee, "An integrated machine learning approach to stroke prediction," in *Proceedings of the 16th ACM SIGKDD international conference on Knowledge discovery and data mining*, 2010, pp. 183–192.
- [3] C. f. D. C. Statistics of Stroke and Prevention, "Stroke facts," 2023, accessed: 2023-25-03. [Online]. Available: <https://www.cdc.gov/stroke/facts.htm>
- [4] Y. Xu, K. Li, Y. Zhao, L. Zhou, Y. Liu, and J. Zhao, "Role of ferroptosis in stroke," *Cellular and Molecular Neurobiology*, vol. 43, no. 1, pp. 205–222, 2023.
- [5] X. Li and G. Chen, "Cns-peripheral immune interactions in hemorrhagic stroke," *Journal of Cerebral Blood Flow & Metabolism*, vol. 43, no. 2, pp. 185–197, 2023.
- [6] Y. Yang, J. Zheng, Z. Du, Y. Li, Y. Cai et al., "Accurate prediction of stroke for hypertensive patients based on medical big data and machine learning algorithms: retrospective study," *JMIR Medical Informatics*, vol. 9, no. 11, p. e30277, 2021.
- [7] M. Chun, R. Clarke, B. J. Cairns, D. Clifton, D. Bennett, Y. Chen, Y. Guo, P. Pei, J. Lv, C. Yu et al., "Stroke risk prediction using machine learning: a prospective cohort study of 0.5 million chinese adults," *Journal of the American Medical Informatics Association*, vol. 28, no. 8, pp. 1719–1727, 2021.
- [8] G. Ntaios, D. Sagrais, A. Kallipolitis, E. Karagkiozi, E. Korompoki, E. Manios, V. Plagianakos, K. Vemmos, and I. Maglogiannis, "Machine-learning-derived model for the stratification of cardiovascular risk in patients with ischemic stroke," *Journal of Stroke and Cerebrovascular Diseases*, vol. 30, no. 10, p. 106018, 2021.
- [9] K. Uchida, J. Kouno, S. Yoshimura, N. Kinjo, F. Sakakibara, H. Araki, and T. Morimoto, "Development of machine learning models to predict probabilities and types of stroke at prehospital stage: The japan urgent stroke triage score using machine learning (just-ml)," *Translational Stroke Research*, pp. 1–12, 2022.
- [10] H. Z. U. Rehman, H. Hwang, and S. Lee, "Conventional and deep learning methods for skull stripping in brain mri," *Applied Sciences*, vol. 10, no. 5, p. 1773, 2020.
- [11] M. Lim, A. Abdullah, N. Jhanjhi, M. K. Khan, and M. Supramaniam, "Link prediction in time-evolving criminal network with deep reinforcement learning technique," *IEEE Access*, vol. 7, pp. 184 797–184 807, 2019.
- [12] M. Lim, A. Abdullah, N. Jhanjhi, and M. Supramaniam, "Hidden link prediction in criminal networks using the deep reinforcement learning technique," *Computers*, vol. 8, no. 1, p. 8, 2019.
- [13] S. Bacchi, T. Zerner, L. Oakden-Rayner, T. Kleinig, S. Patel, and J. Jannes, "Deep learning in the prediction of ischaemic stroke thrombolysis functional outcomes: a pilot study," *Academic radiology*, vol. 27, no. 2, pp. e19–e23, 2020.
- [14] M. Saritha, K. P. Joseph, and A. T. Mathew, "Classification of mri brain images using combined wavelet entropy based spider web plots and probabilistic neural network," *Pattern Recognition Letters*, vol. 34, no. 16, pp. 2151–2156, 2013.
- [15] S. Dev, H. Wang, C. S. Nwosu, N. Jain, B. Veeravalli, and D. John, "A predictive analytics approach for stroke prediction using machine learning and neural networks," *Healthcare Analytics*, vol. 2, p. 100032, 2022.
- [16] L. Ali, A. Rahman, A. Khan, M. Zhou, A. Javed, and J. A. Khan, "An automated diagnostic system for heart disease prediction based on chi square statistical model and optimally configured deep neural network," *Ieee Access*, vol. 7, pp. 34 938–34 945, 2019.
- [17] A. Phaphuangwittayakul, Y. Guo, F. Ying, A. Y. Dawod, S. Angkurawaranon, and C. Angkurawaranon, "An optimal deep learning framework for multi-type hemorrhagic lesions detection and quantification in head ct images for traumatic brain injury," *Applied Intelligence*, pp. 1–19, 2022.
- [18] S. Gudadhe and A. Thakare, "Classification of intracranial hemorrhage ct images for stroke analysis with transformed and image-based glm features," 2022.
- [19] S. Gudadhe, A. Thakare, and A. M. Anter, "A novel machine learning-based feature extraction method for classifying intracranial hemorrhage computed tomography images," *Healthcare Analytics*, vol. 3, p. 100196, 2023.
- [20] M. Yeo, B. Tahayori, H. K. Kok, J. Maingard, N. Kutaiba, J. Russell, V. Thijs, A. Jhamb, R. V. Chandra, M. Brooks et al., "Evaluation of techniques to improve a deep learning algorithm for the automatic detection of intracranial haemorrhage on ct head imaging," *European Radiology Experimental*, vol. 7, no. 1, p. 17, 2023.
- [21] K. Khosravi, F. Rezaie, J. R. Cooper, Z. Kalantari, S. Abolfathi, and J. Hatamiakouei, "Soil water erosion susceptibility assessment using deep learning algorithms," *Journal of Hydrology*, vol. 618, p. 129229, 2023.
- [22] J. Donnelly, S. Abolfathi, J. Pearson, O. Chatrabgoun, and A. Daneshkhah, "Gaussian process emulation of spatio-temporal outputs of a 2d inland flood model," *Water Research*, vol. 225, p. 119100, 2022.
- [23] E. Rasool, M. J. Anwar, B. Shaker, M. H. Hashmi, K. U. Rehman, and Y. Seed, "Breast microcalcification detection in digital mammograms using deep transfer learning approaches," in *Proceedings of the 2023 9th International Conference on Computing and Data Engineering*, 2023, pp. 58–65.
- [24] J. Donnelly, A. Daneshkhah, and S. Abolfathi, "Forecasting global climate drivers using gaussian processes and convolutional autoencoders," *Engineering Applications of Artificial Intelligence*, vol. 128, p. 107536, 2024.
- [25] M. A. Saleem, N. Thien Le, W. Asdornwised, S. Chaitusaney, A. Javed, and W. Benjapolakul, "Sooty tern optimization algorithm-based deep learning model for diagnosing nscic tumours," *Sensors*, vol. 23, no. 4, p. 2147, 2023.
- [26] B. Ghiasi, R. Noori, H. Sheikhan, A. Zeynolabedin, Y. Sun, C. Jun, M. Hamouda, S. M. Bateni, and S. Abolfathi, "Uncertainty quantification of granular computing-neural network model for prediction of pollutant longitudinal dispersion coefficient in aquatic streams," *Scientific reports*, vol. 12, no. 1, p. 4610, 2022.
- [27] S. Sindhiya and S. Gunasundari, "A survey on genetic algorithm based feature selection for disease diagnosis system," in *Proceedings of IEEE international conference on computer communication and systems IC-CCS14*. IEEE, 2014, pp. 164–169.
- [28] A. Javed, A. L. Dallora, J. S. Berglund, A. Ali, P. Anderberg, and L. Ali, "Predicting dementia risk factors based on feature selection and neural networks," *Computers, Materials and Continua*, vol. 75, no. 2, pp. 2491–2508, 2023.
- [29] P. Alirezazadeh, A. Fathi, and F. Abdali-Mohammadi, "A genetic algorithm-based feature selection for kinship verification," *IEEE Signal Processing Letters*, vol. 22, no. 12, pp. 2459–2463, 2015.
- [30] R. Noori, B. Ghiasi, S. Salehi, M. Esmaili Bidhendi, A. Raeisi, S. Partani, R. Meysami, M. Mahdian, M. Hosseinzadeh, and S. Abolfathi, "An efficient data driven-based model for prediction of the total sediment load in rivers," *Hydrology*, vol. 9, no. 2, p. 36, 2022.
- [31] A. Javed, J. S. Berglund, A. L. Dallora, M. A. Saleem, and P. Anderberg, "Predictive power of xgboost_bilstm model: A machine-learning approach for accurate sleep apnea detection using electronic health data," *International Journal of Computational Intelligence Systems*, vol. 16, no. 1, p. 188, 2023.
- [32] A. Mousa and B. Schuller, "Contextual bidirectional long short-term memory recurrent neural network language models: A generative approach to sentiment analysis," 2017.

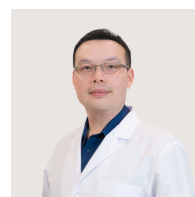
- [33] A. Javeed, A. L. Dallora, J. S. Berglund, A. Idrisoglu, L. Ali, H. T. Rauf, and P. Anderberg, "Early prediction of dementia using feature extraction battery (feb) and optimized support vector machine (svm) for classification," *Biomedicine*, vol. 11, no. 2, p. 439, 2023.
- [34] S. Yadav and S. Shukla, "Analysis of k-fold cross-validation over hold-out validation on colossal datasets for quality classification," in *2016 IEEE 6th International conference on advanced computing (IACC)*. IEEE, 2016, pp. 78–83.
- [35] A. Javeed, P. Anderberg, A. Nauman Ghazi, A. Noor, S. Elmst hl, and J. Sanmartin Berglund, "Breaking barriers: A statistical and machine learning-based hybrid system for predicting dementia," *Frontiers in Bioengineering and Biotechnology*, vol. 11, p. 1336255.
- [36] A. Javeed, A. L. Dallora, J. S. Berglund, and P. Anderberg, "An intelligent learning system for unbiased prediction of dementia based on autoencoder and adaboost ensemble learning," *Life*, vol. 12, no. 7, p. 1097, 2022.
- [37] K. Pal and B. V. Patel, "Data classification with k-fold cross validation and holdout accuracy estimation methods with 5 different machine learning techniques," in *2020 fourth international conference on computing methodologies and communication (ICCMC)*. IEEE, 2020, pp. 83–87.
- [38] A. Javeed, L. Ali, A. Mohammed Seid, A. Ali, D. Khan, and Y. Imrana, "A clinical decision support system (cdss) for unbiased prediction of caesarean section based on features extraction and optimized classification," *Computational Intelligence and Neuroscience*, vol. 2022, 2022.
- [39] B. S. C. I. Dataset, "Available online;," 2021, accessed: 2023-07-03. [Online]. Available: <https://www.kaggle.com/datasets/afdirahman/brain-stroke-ct-image-dataset>
- [40] J. O. Westgard and J. A. Lott, "Precision and accuracy: concepts and assessment by method evaluation testing," *CRC Critical Reviews in Clinical Laboratory Sciences*, vol. 13, no. 4, pp. 283–330, 1981.
- [41] H. Dalianis and H. Dalianis, "Evaluation metrics and evaluation," *Clinical Text Mining: secondary use of electronic patient records*, pp. 45–53, 2018.
- [42] L. Ali, A. Javeed, A. Noor, H. T. Rauf, S. Kadry, and A. H. Gandomi, "Parkinson's disease detection based on features refinement through 11 regularized svm and deep neural network," *Scientific Reports*, vol. 14, no. 1, p. 1333, 2024.
- [43] A. Javeed, M. A. Saleem, A. L. Dallora, L. Ali, J. S. Berglund, and P. Anderberg, "Decision support system for predicting mortality in cardiac patients based on machine learning," *Applied Sciences*, vol. 13, no. 8, p. 5188, 2023.
- [44] A. Javeed, P. Anderberg, A. N. Ghazi, A. Noor, S. Elmst hl, and J. S. Berglund, "Breaking barriers: a statistical and machine learning-based hybrid system for predicting dementia," *Frontiers in Bioengineering and Biotechnology*, vol. 11, 2023.
- [45] X. Zhang, X. Li, Y. Feng, and Z. Liu, "The use of roc and auc in the validation of objective image fusion evaluation metrics," *Signal processing*, vol. 115, pp. 38–48, 2015.
- [46] L. Ali, C. Zhu, N. A. Golilarz, A. Javeed, M. Zhou, and Y. Liu, "Reliable parkinson's disease detection by analyzing handwritten drawings: construction of an unbiased cascaded learning system based on feature selection and adaptive boosting model," *Ieee Access*, vol. 7, pp. 116 480–116 489, 2019.
- [47] T. Fawcett, "An introduction to roc analysis," *Pattern recognition letters*, vol. 27, no. 8, pp. 861–874, 2006.
- [48] G. Sailasya and G. L. A. Kumari, "Analyzing the performance of stroke prediction using ml classification algorithms," *International Journal of Advanced Computer Science and Applications*, vol. 12, no. 6, 2021.
- [49] B. Akter, A. Rajbongshi, S. Sazzad, R. Shakil, J. Biswas, and U. Sara, "A machine learning approach to detect the brain stroke disease," in *2022 4th International Conference on Smart Systems and Inventive Technology (ICSSIT)*. IEEE, 2022, pp. 897–901.
- [50] A. Tursynova, B. Omarov, N. Tukenova, I. Salgozha, O. Khaaval, R. Ramazanov, and B. Ospanov, "Deep learning-enabled brain stroke classification on computed tomography images."
- [51] M. Gupta, P. Meghana, K. H. Reddy, and P. Supraja, "Predicting brain stroke using iot-enabled deep learning and machine learning: Advancing sustainable healthcare," in *International Conference on Sustainable Development through Machine Learning, AI and IoT*. Springer, 2023, pp. 113–122.
- [52] L. Cort s-Ferre, M. A. Guti rrez-Naranjo, J. J. Egea-Guerrero, S. P rez-S nchez, and M. Balcerzyk, "Deep learning applied to intracranial hemorrhage detection," *Journal of Imaging*, vol. 9, no. 2, p. 37, 2023.



MUHAMMAD ASIM SALEEM received his Bachelor's degree from Government College University Faisalabad, Pakistan. The MS Information Technology Degree from Government College University of Faisalabad, Pakistan, 2017. He holds a PhD in Software Engineering from the School of Information and Software Engineering, University of Electronic Science and Technology, China. He is a postdoctoral researcher at the Faculty of Engineering, Chulalongkorn University, Thailand. He has served as a reviewer in many journals, including IEEE. His research interests include Code Optimization, Machine Learning, Deep Learning, Health Informatics and Biomedical.



ASHIR JAVEED received his BS in Software Engineering and MS in Computer Science. He has recently received his PhD in Software Engineering from the University of Electronic Science and Technology, Chengdu, China. Dr. Ashir has researched several areas, including artificial intelligence, data mining and applications in the medical field for designing diagnostic systems to identify heart failure and other diseases.



WASAN AKARATHANAWAT received the master's degree in science from Chulalongkorn University, Bangkok, Thailand, in 2018, and the M.D. degree (Hons.) from Mahidol University, Bangkok, in 2005. He held a Residency Training and a Fellowship Training with the King Chulalongkorn Memorial Hospital, Bangkok, in 2011 and 2018, respectively. He is currently with the Division of Neurology, Department of Medicine, Faculty of Medicine, Chulalongkorn University.



AURAUMA CHUTINET is a neurologist and stroke specialist. She is currently an Associate Professor and a director of Chulalongkorn Stroke Center, King Chulalongkorn Memorial Hospital. She is also the Former Head of the Division of Neurology, Department of Medicine, Faculty of Medicine, Chulalongkorn University, Bangkok, Thailand. She received her medical degree from Chulalongkorn University in 1995. She completed her internal medicine and neurology training at King Chulalongkorn Memorial Hospital in 2002 and 2006. Dr. Chutinet received a certificate in fellowship training in Stroke and Neurosonology (Thailand) in 2010. She was a Stroke Research Fellow at the Massachusetts General Hospital (MGH) in Boston for two years, from 2009 to 2011, under NINDS-funded Specialized Program of Translational Research in Acute Stroke (SPOTRIAS) projects. Her research interests include large vessel atherosclerosis, thrombolysis, stroke risk factors, genetics, biomarkers, neuroprotective agents, neuroimaging, telestroke and AI in Medicine.



DR. NIJASRI C. SUWANWELA is a Professor of Neurology at the Chulalongkorn University in Bangkok, Thailand. She received her medical degree and completed her residency in neurology from Chulalongkorn University. She was awarded by the Anandamahidol Foundation and had her fellowship training in cerebrovascular disease at the Massachusetts General Hospital in Boston, USA. She is the former Head of the Division of Neurology at Chulalongkorn University and Director of Chulalongkorn Comprehensive Stroke Center. She is currently the Director of Chula Neuroscience Center. She is also the past President of the Neurological Society of Thailand. For international societies, she is the treasurer of the Asia Pacific Stroke Organization (APSO) and Vice president of the ASEAN Neurological Association (ASNA). Prof. Suwanwela was recently appointed as the Vice Dean of Academic Service Affairs at the Faculty of Medicine, Chulalongkorn University and Program director of the Chulalongkorn University International Medical Program (CUMEDI). She published many articles in international peer-reviewed journals. She also delivered many local and international lectures.



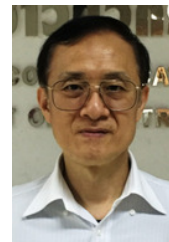
SUNCHAI DEELERTPAIBOON received M. Eng. degree in electrical engineering from Chulalongkorn University, Thailand, in 2006, and B. Eng. in electrical engineering from Prince of Songkhla University, Thailand, in 2002. At present, he is Head of the Technology and Information Department at Wattanapat Hospital Ao Nang. His research interests include communication networks, data analysis and the application of artificial intelligence.



WATTANASAK SRISIRI received the B. Eng. and M. Eng. degree in electrical engineering from the Khon Kaen University of Thailand. His current research interests include artificial intelligence and Internet of Things for medicine, smart grid, and renewable energy.



WIDHYAKORN ASDORNWISED received M.Sci. Electrical Engineering, Drexel University in 1999 and the D.Eng. Electrical Engineering, Chulalongkorn University in 2005. He is now Assistant Professor at the Department of Electrical Engineering at Chulalongkorn University. He does research in Artificial Intelligence, Data Mining and Artificial Neural Networks. The current projects are '5G Chula', Micro-operator, and 5G/6G Private Network.



WATIT BENJAPOLAKUL received the D. Eng. degree in electrical engineering from the University of Tokyo in 1989. He is now a Professor at the Department of Electrical Engineering, Chulalongkorn University, Thailand. He is the head of the Center of Excellence in Artificial Intelligence, Machine Learning, and Smart Grid Technology, Faculty of Engineering, Chulalongkorn University. His current research interests include mobile communication systems, broadband networks, applications of artificial intelligence in communication systems, wireless networks, and applications of communication networks in smart grids.



SURACHAI CHAITUSANEY obtained the Ph.D. degree from The University of Tokyo, Japan, with JICA scholarship in 2007. He is an Assistant Professor and the Head of Power System Research Laboratory (PSRL) at Chulalongkorn University. His research interests include renewable energy, solar PV, distributed generation, power system planning and reliability, and smart grid regulation. In addition, he has served as a member of subcommittee, committee and working group in several professional organizations, such as Engineering Institute of Thailand (EIT), Energy Policy and Planning Office (EPPO), Energy Regulatory Commission (ERC), and Electrical Engineering/ Electronics, Computer, Telecommunications and Information Technology (ECTI) Association.



PASU KAEWPLUNG received Ph.D. Eng. degree in electrical engineering from Chulalongkorn University, Thailand, in 2004, and M. Eng. in electrical engineering from Yokohama National University, Japan, in 1998. He is now Assistant Professor at the Department of Electrical Engineering, Chulalongkorn University. His research interests include mobile communications, signal processing, applications to optical devices, long-haul fiber transmission, fiber access network and application of artificial intelligence.

...