International Journal of Advance Research in Science and Engineering
Volume No. 14, Issue No. 05, May 2025
www.ijarse.com
IJARSE
ISSN 2319 - 8354

Smart Stroke Diagnosis: Machine Learning-Driven Analysis of Neuroimages

Riyaz Ahmed N A[1], Ms. Radha Korimani[2], Sheik Abdulla[3], Md Ashir Khan[4], Preetham M[5]

[1,3,4,5]UG Student, [2]Assistent Professor, Department of CSE,

AMC Engineering College Banglore

nayariyaz63@gmail.com, kradha.rymec@gmail.com, abdullahsheikh070707@gmail.com mdashirkhan11@gmail.com, itzpreetham101@gmail.com

ABSTRACT

Steps to improve risk factors are not always clear, however, addressing those factors prevents underlying causes. Through these strategies, population prevalence of stroke can be targeted and the burden minimized. Diagnosis and treatment protocols should also allow for adequate and timely access to definitive care while ensuring optimal outcomes based on evidence-based best practice pathways and guidelines. Enhancing the results of patients who suffered a stroke while reducing its consequences over time require early detection, clear identification, and defined goals. Study gives aim to explore pioneering solutions with focus on muscle mobility recognition using machine learning, artificial intelligence, and imaging techniques involving the use of cutting-edge computer devices. This especially advanced method combines multi source data to cut down the time it takes to locate a stroke using the latest techniques in medicine, thus, enhancing the speed and accuracy of stroke detection. The main aim of the approach is enabling health professionals to act on time, and guarding against needless deaths that can be stopped by employing advanced features.

I. INTRODUCTION

Astroke 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. 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. here are two types of stroke: ischemic stroke and hemorrhage-related stroke. About87% of strokes are is chemic strokes, the mostcommon type. When a blood vessel supplying blood to the brain is blocked or nar- rowed, 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. 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.

This study uses machine learning techniques to de- velop a systematic method for stroke detection. The study's

IJARSE
ISSN 2319 - 8354

main aim is to understand the causes of strokes better and create reliable detection models that help physicians make informed decisions about stroke pre- vention 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 algo rithms. 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 is also compared with other ML and DLmethods.
5) The proposed system aims to aid healthcare professionals in making informed decisions about stroke treatment and prevention by providing early detection.

II. RELATED WORK

[1] 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. Various machine learning algorithms, such as Extreme Gradient Boosting (XGBoost), were used to achieve optimal performance in predicting stroke risk. The use of machine learning models, including ensemble models that combine different approaches, has shown promis- ing results in identifying individuals at high risk of stroke. The authors of 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.

[2] The authors in present a time-based link predic- tion 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.

[3] A study in 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. Anintegrated 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. 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%.

[4] The authors of and 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 created using sub-sampling techniques.

[5] The deep learning framework proposed in 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 hemor rhagic lesions. This enables clinically relevant quantification, which is important for the deci- sion on emergency surgical treatment.

[6] In, 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 REP Tree 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, 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 combina tion with

the proposed feature extraction mechanism.

[7] In, the authors measure the performance of a deep learning model for detecting intracranial hemorrhage on CT head scans and compare the effects of different preprocess ing and model design methods. Implementing preprocess ing techniques and a CNN-RNN frame- work significantly improved the model's performance and demonstrated its potential as a radiologist decision- support tool.

[8] 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 bidi- rectional 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. METHEDOLOGY

This study combines Brain-Computer Interface (BCI) and traditional rehabilitation methods to improve hand function in stroke patients. The hybrid approach uses electroencephalography (EEG) signals to capture mo- tor intention and electromyography (EMG) signals to monitor muscle activity. The adaptive system customizes rehabilitation exercises based on real-time feedback, optimizing recovery.

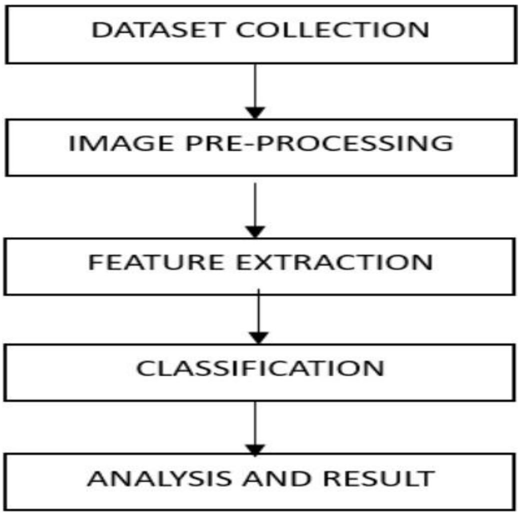

Fig. 1

The framework describes the architecture of the pro- posed 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 net- work 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 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. Dataset Collection

- The first step involves collecting a relevant dataset for brain stroke prediction. The dataset should con- tain medical images, such as MRI or CT scans, as well as clinical data (like patient demographics and health history). Several publicly available datasets can be utilized.
- The Stroke Dataset from Kaggle: This dataset in- cludes various features like age, gender, hyper- tension, heart disease, marital status, work type, residence type, average glucose levels, and body mass index (BMI).
- Image Data: Utilize medical imaging datasets, en- suring that they are labeled according to the pres- ence or absence of stroke.

B. Preprocessing

- **Data Cleaning:** Remove any missing or erroneous entries from the dataset.
- **Normalization:** Normalize the pixel values of the images to a range of 0 to 1 to facilitate better training.
- **Resizing:** Resize all images to a uniform size compatible with the ResNet50 model (typically 224x224 pixels).
- **Data Augmentation:** Apply techniques such as ro- tation, flipping, and zooming to artificially increase the diversity of the dataset, which can help improve model robustness.

C. Feature Extraction

- **Transfer Learning:** Use a pre-trained ResNet50 model (trained on ImageNet) as a starting point. This allows the model to leverage learned features from a larger dataset.
- ResNet50, being a deep learning architecture, auto- matically extracts features from the images during the training process.
- **Fine-Tuning:** Remove the final classification layer of the ResNet50 model and replace it with a custom layer suitable for stroke classification. Fine-tune the model by training it on the stroke dataset to adapt the weights specifically for this task.

D. Classification

- **Model Configuration:** Compile the model us- ing appropriate loss functions (e.g., binary cross- entropy for binary classification) and optimizers (like Adam).
- **Training:** Split the dataset into training, validation, and test sets. Train the model using the training set while validating performance on the validation set. Employ techniques such as early stopping to prevent overfitting.
- **Batch Size and Epochs:** Use an appropriate batch size (e.g., 32) and number of epochs (e.g., 50–100) based on the dataset size and model performance.

International Journal of Advance Research in Science and Engineering
Volume No. 14, Issue No. 05, May 2025
www.ijarse.com
IJARSE
ISSN 2319 - 8354

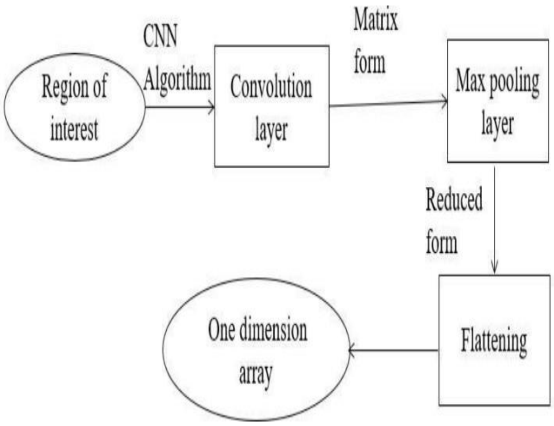

Fig. 2: Data Flow Diagram for Feature Extraction

Result

- **Performance Metrics:** Summarize the performance metrics obtained from the evaluation phase. Highlight any improvements achieved through transfer learning and fine-tuning.
- **Comparison with Baselines:** Compare the results with baseline models (e.g., traditional machine learning algorithms like SVM or logistic regression) to demonstrate the effectiveness of the ResNet50 architecture.
- **Clinical Implications:** Discuss the potential impli- cations of the findings for clinical practice, such as how the model could assist healthcare professionals in early stroke detection.

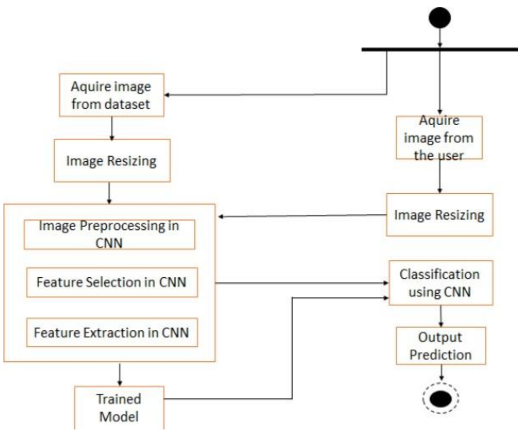

Fig. 3: Activity Diagram

IV. 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 experiments how 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 the se explainable AI methods to help healthcare professionals make more informed decisions based on these AI results.

V. FUTURE ENHANCEMENT

While the current study has yielded encouraging re- sults, there are several avenues for future enhancement that could further improve the performance and applica- bility of the brain stroke prediction model:

- **Data Expansion:** Incorporating a larger and more diverse dataset can help improve the model's ro- bustness. Future work should focus on gathering data from multiple medical institutions to ensure a broader representation of stroke cases.
- **Transfer Learning:** Exploring transfer learning techniques with more advanced architectures be- yond ResNet50, such as EfficientNet or DenseNet, may yield even better predictive performance. These architectures may offer different strengths in feature extraction and model efficiency.
- **Multimodal Data Integration:** Combining imag- ing data with other patient data, such as demo- graphic information and clinical history, could en- hance prediction accuracy. Developing a model that integrates these various data types may provide a more holistic view of stroke risk.
- **Real-Time Prediction:** Implementing a real-time prediction system in clinical settings could facilitate immediate responses to potential stroke cases. Fu- ture research should focus on optimizing the model for real-time performance while ensuring minimal computational overhead

REFERENCES

[1] J. Su, J. Wang, and W. Wang, "An adaptive hybrid brain–computer interface for hand function rehabilitation of stroke patients," *[Journal Name]*, vol. XX, no. XX, pp. XX–XX, 2024.

[2] R. Zhang, C. Wang, S. He, C. Zhao, and K. Zhang, "An adaptive brain-computer interface to enhance motor recovery after stroke," *[Journal Name]*, vol. XX, no. XX, pp. XX–XX, 2023.

[3] S. S. Mousavi and M. Majedi, "A quantitative microwave imaging approach for brain stroke classification based on the generalized Tikhonov regularization," *[Journal Name]*, vol. XX, no. XX, pp. XX–XX, 2023.

[4] T. Janyalikit and C. A. Ratanamahatana, "Time series shapelet- based movement intention detection toward asynchronous BCI for stroke rehabilitation," *[Journal Name]*, vol. XX, no. XX, pp. XX–XX, 2022.

[5] M. M. Yu¨ksel, G. Courtine, and S. Sun, "Low-intensity focused ultrasound neuromodulation for stroke recovery: A novel deep brain stimulation approach for neurorehabilitation?" *[Journal Name]*, vol. XX, no. XX, pp. XX–XX, 2023.

[6] O. Karadima, P. Kosmas, P. Lu, and I. Sotiriou, "Experimental validation of the DBIM-TwIST algorithm for brain stroke detection and differentiation using a multi-layered anatomically complex head phantom,"

[Journal Name], vol. XX, no. XX, pp. XX–XX, 2022.

[7] K. Wu, “Dynamic reconfiguration of brain functional network in stroke,” *[Journal Name]*, vol. XX, no. XX, pp. XX–XX, 2024.

[8] R. Nakamoto, B. Flanagan, Y. Dai, K. Takami, and H. Ogata, ”Unsupervised techniques for generating a standard sample self- explanation answer with knowledge components in a math quiz,”

[9] *Research and Practice in Technology Enhanced Learning*, vol. 19, art. 016, 2023. doi: 10.58459/rptel.2024.19016.