ISSN: 2096-3246

Volume 57, Issue 02, April, 2025

Machine Learning-Enhanced Neuroimaging for Precise Stroke Diagnosis: A Diagnostic Innovation

Dr. K. Nagi Reddy
Professor and Head
Department of Information Technology,
Lords Institute of Engineering and Technology,
Hyderabad, India.
k.nagireddy@lords.ac.in

Mrs. N. Vibhavari
Assistant Professor
Department of Computer Science and Engineering,
Lords Institute of Engineering and Technology,
Hyderabad, India.
vibhavari@lords.ac.in

Mrs. Hajira Sabuhi
Assistant Professor
Department of Information Technology,
Lords Institute of Engineering and Technology,
Hyderabad, India.
hajirasabuhi@lords.ac.in

INDEX TERMS:

Stroke,

feature selection,

CNN, MobileNet,

Inception V3,

ResNet, CT images.

ABSTRACT

Stroke is one of the most prevalent causes of death and disability in the world, although it is preventable and treated. Improving clinical outcomes and lowering the burden of disease are significantly aided by early stroke detection and prompt treatments. Because machine learning techniques can be used to detect strokes, they have garnered a lot of attention in recent years. Finding trustworthy techniques, algorithms, and characteristics that support healthcare providers in making well-informed decisions on stroke prevention and treatment is the goal of this project. In order to accomplish this, we have created an early stroke detection system that uses brain CT scans in conjunction with the CNN algorithm to identify strokes at an extremely early stage. These CT image characteristics are incorporated into the CNN model. The diagnostic system's accuracy, precision, recall, F1 score, Receiver Operating Characteristic Curve (ROC), and area under the curve (AUC) were all assessed using cross-validation. The overall efficacy of the system was assessed using each of these indicators. The accuracy of the suggested diagnostic system was 98%. Additionally, we contrasted the suggested model's performance with those of MobileNet and ResNet. Physicians can make well-informed decisions regarding stroke with the help of the suggested diagnosis system.

I.INTRODUCTION

When a blood artery in the brain bursts or the blood supply to the brain is cut off, a stroke happens. With more than 6.2 million fatalities each year, it is the leading cause of death globally [1]. Many survivors have infirmities that significantly lower their quality of life. The severity of a stroke and its effects on the individual can be lessened with the use of preventive measures and prompt intervention. Preventing strokes requires early identification of those who are at risk [2]. Hemorrhage-related stroke and ischemic stroke are the two forms of stroke. Ischemic strokes are the most prevalent form, accounting for around 87% of all strokes [3]. There is little or no blood flow to a particular area of the brain when a blood artery that supplies blood to the brain is blocked or constricted. The obstruction might be brought on by cholesterol buildup, blood clots, or other material that enters the brain from other areas of the body [4]. Though less frequent, hemorrhage-related strokes are more dangerous and frequently deadly if left untreated. Blood seeping into the surrounding brain tissue from a ruptured or leaking blood artery in the brain is known as a cerebral hemorrhage. This causes the brain's structures to enlarge and become under strain. Normal brain activity may be disrupted and brain cells may be damaged [5].

This paper develops a systematic stroke detection approach using machine learning techniques. The primary goals of the project are to improve our understanding of stroke etiology and develop trustworthy detection models that support doctors in making well-informed decisions on stroke prevention and treatment. This study's objective is to assess the suggested algorithms' efficacy, lucidity, flexibility, and scalability. To ascertain their relative performance, the suggested models were contrasted with well-established clinical risk detection techniques using a variety of real-world datasets. To lessen the effects of stroke on patients and the healthcare system, this study may offer insightful information for clinical practice and customized therapies. This work aims to answer the issue of stroke incidence by using machine learning techniques to forecast the frequency of strokes. This work enhances the precision and understandability of stroke detection models by utilizing cutting-edge machine learning techniques. By enabling precise treatments to avoid this crippling disease in the future, the findings of this study have the potential to completely transform stroke prevention and improve treatment outcomes for those who already suffer from it. The following is a list of this study's main contributions:

1) Using an image-based dataset, this work offers a diagnostic approach for stroke identification.
2) The suggested diagnostic approach uses a genetically tuned CNN to extract valuable characteristics from the CT images.
3) A comparison of the suggested diagnostic system's (CNN) performance with image classification techniques is also made.
4) By offering early detection, the suggested approach seeks to assist medical practitioners in making well-informed decisions about stroke prevention and treatment.

II. RELATED WORKS

Machine learning algorithms have been instrumental in creating highly precise models for stroke risk detection. These models analyze historical data from electronic health records to estimate stroke risk in hypertensive patients. Techniques like Extreme Gradient Boosting (XGBoost) have demonstrated exceptional performance in risk prediction. Furthermore, ensemble models, which combine various approaches, have proven effective in identifying individuals with a high probability of experiencing a stroke.

In a comprehensive review, researchers have categorized skull stripping techniques into two main groups: deep learning-based approaches, including convolutional neural networks, and traditional methods. They highlighted several modern techniques for their potential integration into routine clinical imaging workflows.

A time-dependent link prediction model, called TDRL, employs deep reinforcement learning to analyze evolving real-world crime datasets. Experiments revealed that this model, when trained on temporal datasets, outperformed traditional machine learning models reliant on static datasets. Although the TDRL-CNA model accurately predicted most network connections during a specific year, it struggled with disappearing connections, indicating that learning from removed edges had minimal impact on its effectiveness. Criminal network analysis, often limited by incomplete datasets, has benefited from deep reinforcement learning approaches, which show superior performance over conventional supervised learning techniques.

Another study applied a hybrid convolutional neural network (CNN) structure derived from the National Institutes of Health to predict multiple stroke scores using clinical brain CT data. The model achieved 74% accuracy in detecting Modified Rankin Scale scores (mRS90) after 90 days. Additionally, researchers used a probabilistic neural network with a wavelet entropy-based spider graph structure to classify brain MRI images into normal, stroke-affected, and degenerative conditions. This innovative method, which included discrete wavelet transforms and spider web diagrams for feature extraction, achieved a perfect classification accuracy of 100%.

Further studies have analyzed electronic health records using statistical methods and principal component analysis to improve stroke diagnosis. They identified key risk factors such as advanced age, heart conditions, high blood pressure, and average blood sugar levels. By employing a perceptron neural network with these variables, they achieved superior accuracy and minimized error rates compared to other models. To address the challenge of imbalanced datasets, researchers presented results on a balanced dataset created through sub-sampling techniques.

Lastly, a deep learning model exhibited high accuracy in identifying various subtypes of intracranial hemorrhages (ICH) in head CT scans, achieving a 96.21% accuracy rate. This system demonstrated reduced false positive rates and introduced a quantitative scoring mechanism for measuring the size and volume of hemorrhagic lesions, aiding critical decision-making for emergency surgical procedures.

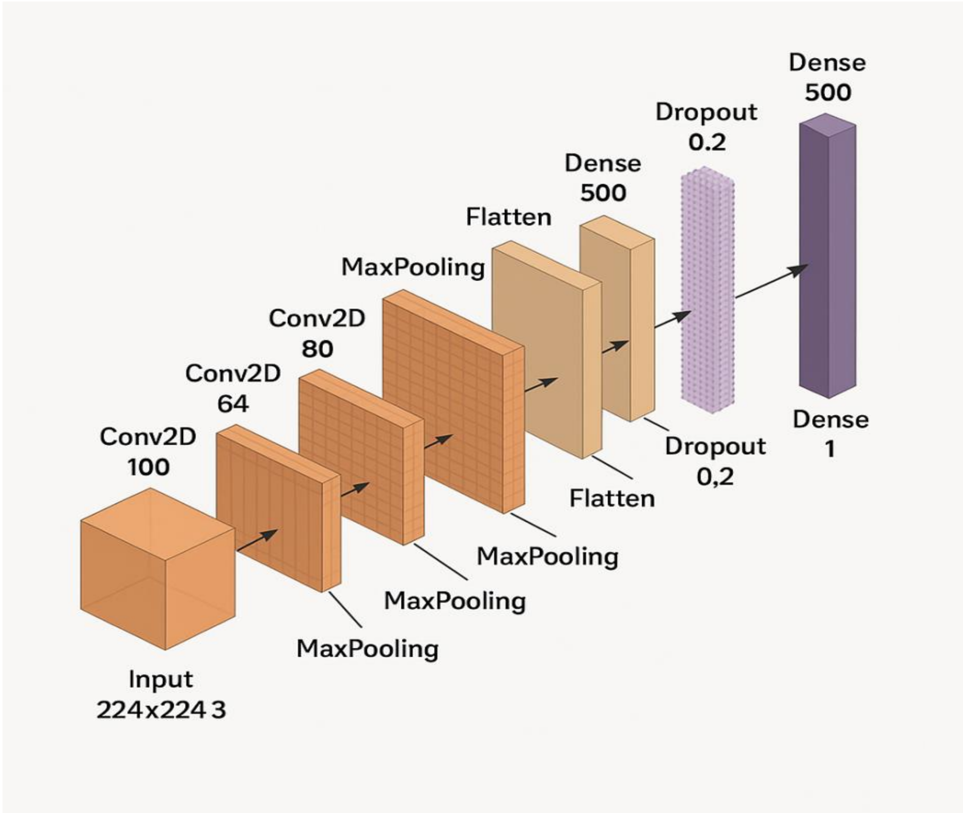
III. METHODOLOGY

This section will go over the research strategy and technique that we used for this investigation. In Figure 1, the suggested framework is displayed. The architecture of the suggested computer-aided stroke diagnostic system is described in the framework. After being put into the system, this benchmark dataset is separated into two smaller datasets: the training dataset and the validation dataset. We employ holdout cross-validation to divide the dataset. 70% of the photos from the five classes are chosen at random to train the model, and the remaining data is divided. Thirty percent of the photos with the appropriate labels are utilized as validation sets to assess the performance of the suggested model. This study uses five sophisticated convolutional network architectures to extract resilient and non-invariant characteristics from RGB photos.This study evaluates the effectiveness of various deep learning algorithms for detecting brain strokes using medical imaging data. The methodology involves preprocessing MRI or CT scan images to remove noise and normalize them for input into deep learning models. Feature extraction is performed using these algorithms to detect patterns indicative of stroke. The performance of each model is assessed using metrics such as accuracy, precision, recall, and computational efficiency.

1. **Convolutional Neural Networks (CNN)**:
   CNNs are the backbone of image recognition tasks. Their ability to automatically extract hierarchical features from medical images makes them ideal for stroke detection. CNNs utilize layers like convolution, pooling, and activation to capture spatial relationships and patterns, optimizing detection accuracy.
2. **InceptionV3**:
   InceptionV3 is known for its efficiency in handling complex datasets through its unique architecture, which includes inception modules. These modules use multiple kernel sizes to extract diverse features, enhancing the model's ability to analyze intricate details in brain scans while keeping computational requirements low.
3. **MobileNet**:
   MobileNet is a lightweight model designed for devices with limited computational power. It employs depthwise separable convolutions to reduce the number of parameters and computations, making it a suitable option for real-time stroke detection in portable medical applications.
4. **ResNet**:
   ResNet's distinguishing feature is its use of residual blocks, which address the vanishing gradient problem in deep networks. This enables the model to maintain high accuracy when analyzing complex brain images, ensuring that subtle features indicative of stroke are effectively captured.

Each algorithm is benchmarked to compare their strengths and weaknesses, ensuring that the most suitable model for stroke detection is identified based on clinical and computational requirements.

Model Architecture:



IV.RESULTS & DISCUSSIONS

Dataset Details:

Kaggle provided the data set utilized in this investigation [39]. Finding stroke risk variables, such as age, gender, smoking, diabetes, and high blood pressure, was the study's main objective. Of the 100 stroke patients who were 16 years of age or older, 32% were women and 68% were men. CT scans of each patient's brain were categorized as either normal or stroke; there were 950 stroke pictures and 1551 normal ones. Prior to the stroke, the same patients' normal pictures were also taken. The scans measured 650 × 650 pixels and were in grayscale. During training, the data set was randomly equalized to prevent overfitting.

Evaluation Metrics:

Accuracy, precision, recall, and F1-score have all been employed as assessment parameters in this study. Each report includes the findings. The definition of commonly used assessment metrics is provided below:

ISSN: 2096-3246

Volume 57, Issue 02, April, 2025

$$Accuracy = \frac{No.\,Correct\,Prediction}{Total\,Samples}$$

$$Precision = \frac{TP}{TP + FP}$$

$$Recall = \frac{TP}{TP + FN}$$

$$F1\_score = \frac{2TP}{2TP + FP + FN}$$

V.EXPERIMENTAL RESULTS AND DISCUSSION

CNN Model:

Below is the performance of CNN Model which has achieved the highest accuracy

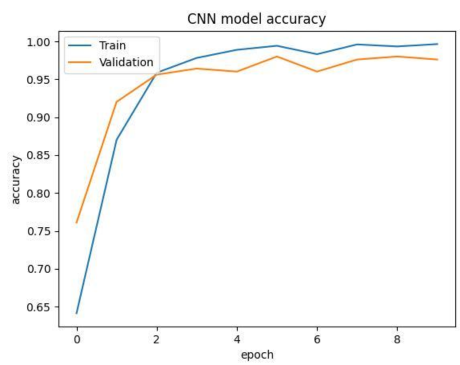


Figure 1: Training and Validation Accuracy

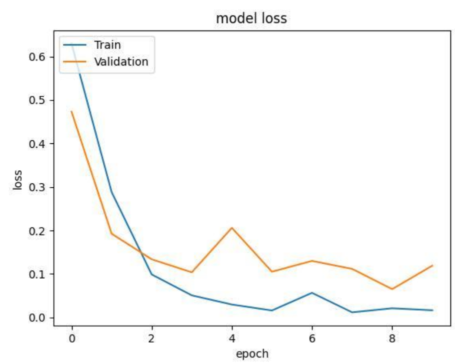


Figure 2: Training and Validation Loss

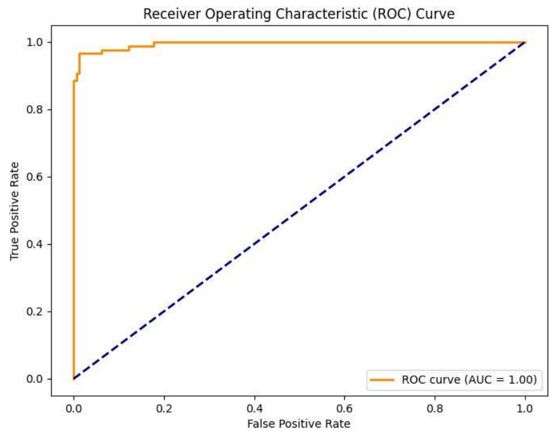


Figure 3: ROC Curve

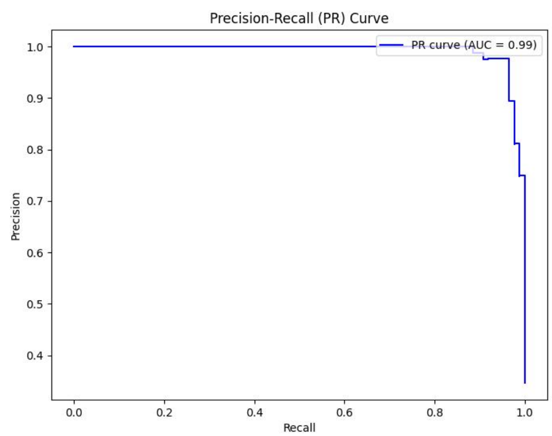


Figure 4: Precision Recall Curve

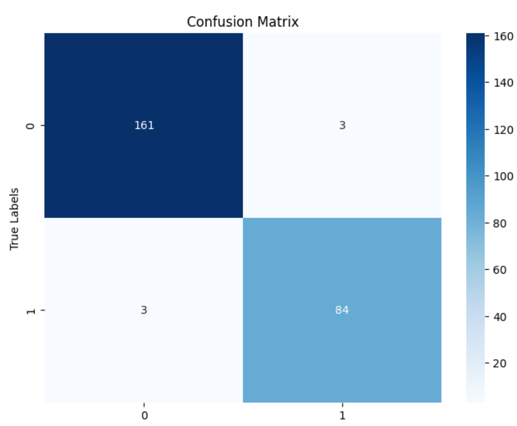


Figure 5: Confusion Matrix

ISSN: 2096-3246

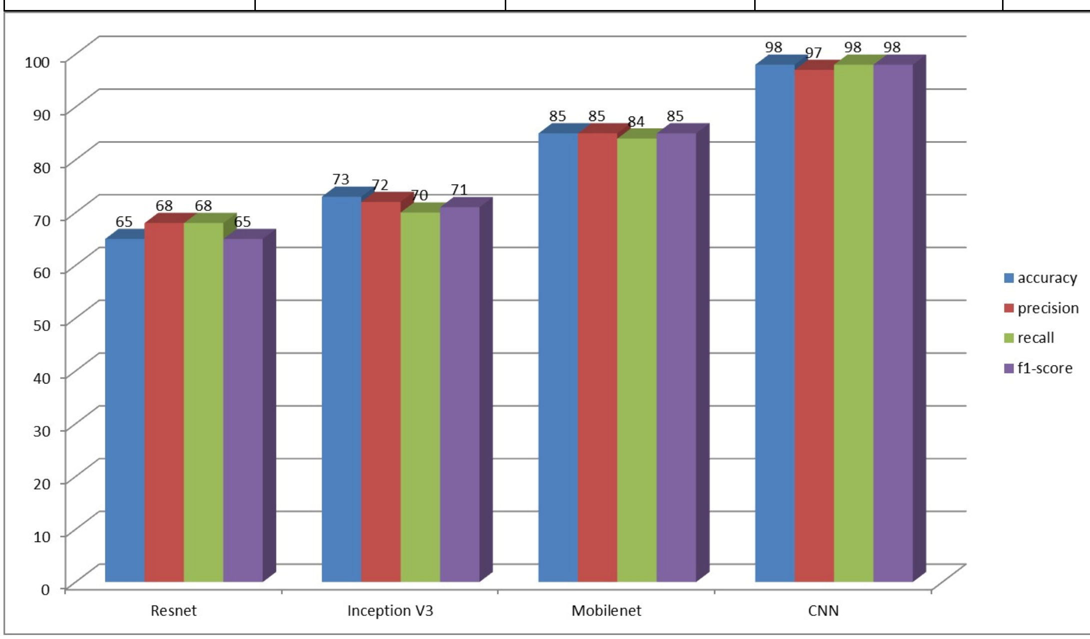
Volume 57, Issue 02, April, 2025

	precision	recall	f1-score	support
0	0.98	0.98	0.98	164
1	0.97	0.97	0.97	87
accuracy			0.98	251
macro avg	0.97	0.97	0.97	251
weighted avg	0.98	0.98	0.98	251

Figure 6: Classification Report

Below is the comparison of the evaluation metrics for the tested algorithms

Model	Accuracy	Precision	Recall	F1-Score
ResNet	65	68	68	65
Inception V3	73	70	72	71
MobileNet	85	85	84	85
CNN	98	97	97	97



It can be observed that the CNN algorithm achieves the highest with respect to all evaluation metrics.

VI.CONCLUSION

This paper suggests a machine learning approach for stroke detection. The newly created model's performance is verified using an image-based dataset. The CNN algorithm is the foundation of the suggested model. The CNN and Mobilenet models use these characteristics to predict strokes. To ascertain the most successful categorization, the performance of several K-folds was assessed. In order to predict strokes, we also examined various machine-learning methods. The experiment's findings demonstrate that the suggested machine-learning model outperforms alternative models in terms of efficiency. Our goal is to employ more sophisticated models that are capable of autonomously predicting strokes in order to enhance stroke detection in the future. A small dataset was employed in this work to test deep learning models; greater datasets often yield better results. We must thus gather additional samples in the future in order to enhance the model's performance and provide better outcomes. Furthermore, the performance of deep learning models is significantly influenced by the quality of the data. Therefore, we must create new techniques that will contribute to future improvements in data quality. By using explainable AI methodologies to clarify the model's decision-making processes, healthcare practitioners' trust must also be increased.

VII.REFERENCES

[1] World Stroke Organization. (2023). Vision Impairment and Blindness. Accessed: Feb. 3, 2023. [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 Proc. 16th ACM SIGKDD Int. Conf. Knowl. Discovery Data Mining, Jul. 2010, pp. 183–192.

[3] Centers for Disease Control Prevention. (2023). Stroke Facts. Accessed: Mar. 25, 2023. [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 Mol. Neurobiol., vol. 43, no. 1, pp. 205–222, Jan. 2023.

[5] X. Li and G. Chen, ‘‘CNS-peripheral immune interactions in hemorrhagic stroke,’’ J. Cerebral Blood Flow Metabolism, vol. 43, no. 2, pp. 185–197, Feb. 2023.

[6] Y. Yang, J. Zheng, Z. Du, Y. Li, and Y. Cai, ‘‘Accurate prediction of stroke for hypertensive patients based on medical big data and machine learning algorithms: Retrospective study,’’ JMIR Med. Informat., vol. 9, no. 11, Nov. 2021, Art. no. e30277.

[7] M. Chun, R. Clarke, B. J. Cairns, D. Clifton, D. Bennett, Y. Chen, Y. Guo, P. Pei, J. Lv, C. Yu, L. Yang, L. Li, Z. Chen, and T. Zhu, ‘‘Stroke risk prediction using machine learning: A

ISSN: 2096-3246

Volume 57, Issue 02, April, 2025

prospective cohort study of 0.5 million Chinese adults,'' J. Amer. Med. Inform. Assoc., vol. 28, no. 8,pp. 1719–1727, Jul. 2021.

[8] G. Ntaios,D.Sagris,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,'' J. Stroke Cerebrovascular Diseases, vol. 30, no. 10, Oct. 2021, Art. no. 106018.

[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),'' Transl. Stroke Res., vol. 13, no. 3, pp. 370–381, Jun. 2022.

[10] H. Z. U. Rehman, H. Hwang, and S. Lee, ''Conventional and deep learning methods for skull stripping in brain MRI,'' Appl. Sci., vol. 10, no. 5,p. 1773, Mar. 2020.