

Revolutionizing Stroke Detection - Machine Learning Diagnostic Models Leveraging Neuroimages

Swetha Bikkumalla¹ Dr. D. Radha² Dr.M.Sambasivudu³

¹Research Scholar, Dept. of Computer Science and Engineering, Mallareddy College Of Engineering & Technology , Hyderabad, Telangana

²Associate Professor, Dept.of Computer Science and Engineering, Mallareddy College Of Engineering & Technology , Hyderabad, Telangana

³Associate Professor, Dept.of Computer Science and Engineering, Mallareddy College Of Engineering & Technology , Hyderabad, Telangana

Article History:

Received: 12-01-2025

Revised: 15-02-2025

Accepted: 01-03-2025

Abstract

This Brain Stroke Detection System streamlines and improves stroke diagnosis by utilizing a three-layer Convolutional Neural Network (CNN) to process non-contrast CT scans. The system is developed with a Python backend using Flask, paired with a responsive frontend built on HTML, CSS, and JavaScript, providing an easy-to-use interface for clinical settings. Regarding Model Architecture & Performance, the custom CNN model exhibits strong diagnostic capability, achieving 98% accuracy on training data and 97% on validation, indicating high precision in detecting stroke-related anomalies. Comparable studies using ensemble CNN models—such as combinations of InceptionV3, Xception, and MobileNetV2—have shown similar effectiveness, reaching 98.9% accuracy, 98.5% recall, and an AUC of 98.7% on related CT datasets. Regarding Dataset & Generalization, training was conducted on a well-balanced dataset of 2,501 CT images, comprising 1,551 normal and 950 stroke-positive scans. This balanced distribution supports better generalization and reduces the risk of overfitting, adhering to best practices in AI for medical imaging. Regarding Clinical Relevance, by combining a high-performing CNN with a user-friendly web interface, this system presents a scalable and dependable tool for clinical stroke detection. Its accuracy is on par with leading deep learning solutions, underlining the growing role of AI in advancing medical imaging technologies.

Keywords: Stroke Identification, Machine Learning, Neuroimages, Diagnostic Model, Inceptionv3, MobileNet, Convolutional Neural Network (CNN)

1. INTRODUCTION

A stroke arises when the brain's blood supply is disrupted—either blocked (ischemic stroke, which accounts for about 87% of cases) or when a vessel ruptures (hemorrhagic stroke, typically more severe). Globally, strokes claim over 6.2 million lives annually, and survivors often face lasting disabilities that severely affect daily life[1]. Early identification of those at risk is vital: this study applies machine learning to data such as electronic health records, genetic markers, and lifestyle factors, using feature-selection algorithms to build accurate, interpretable predictive models. These models are evaluated for effectiveness, clarity,

adaptability, and scalability, and benchmarked against traditional clinical risk scores across real-world datasets[2]. By enabling precise, data-driven interventions, this work aims to revolutionize stroke prevention, reduce its burden on individuals and healthcare systems, and ultimately improve patient outcomes through timely and personalized care. Stroke is a leading cause of death and disability worldwide, resulting in over 7.3 million deaths annually.[3] It occurs when blood flow to the brain is interrupted, either due to a blockage (ischemic stroke) or a ruptured blood vessel (hemorrhagic stroke). Early identification of individuals at risk is crucial for effective prevention and timely intervention. Machine learning techniques have shown promise in developing predictive models for stroke detection. Studies have demonstrated that models such as Support Vector Machine (SVM) and XGBoost achieve high performance[4], with Area Under the Curve (AUC) values ranging from 0.88 to 0.91, indicating strong predictive accuracy. These models utilize various features, including clinical data and imaging information, to assess stroke risk and predict outcomes. By leveraging machine learning algorithms, healthcare professionals can enhance early detection, improve prognostic assessments, and ultimately reduce the burden of stroke on individuals and healthcare systems.

2. EXISTING SYSTEM

Imaging Tests:

CT Scan (Computed Tomography): A CT scan provides detailed images of the brain and is commonly used in emergency settings to detect bleeding or blockages. It is widely available, quick to perform, and can be done on patients with pacemakers or on ventilators. However, CT scans may not detect small or early ischemic strokes, especially in areas like the brainstem, and involve exposure to radiation.

1. **MRI (Magnetic Resonance Imaging):** MRI offers high-resolution images and is more sensitive than CT in detecting ischemic strokes, particularly in the posterior circulation. It does not involve radiation and can detect strokes within minutes of onset. However, MRI is more expensive, less accessible in some regions, and time-consuming compared to CT scans, which can delay treatment.
2. **Blood Flow Tests:**
 - i. **Carotid Ultrasound:** This non-invasive test examines blood flow in the carotid arteries to detect blockages. It is safe, does not involve radiation, and can be performed at the patient's bedside. However, it is highly operator-dependent and may be challenging to obtain the appropriate acoustic window of visualization on the area of interest.
 - ii. **Cerebral angiography** uses dye and X-rays to visualize blood vessels in the brain, helping to identify blockages or abnormalities. It is considered the gold standard for evaluating vascular conditions. However, it is invasive and carries risks such as bleeding, infection, and allergic reactions to the contrast dye.

3. **Electrical Tests: EEG (Electroencephalogram):** EEG records electrical activity in the brain and can detect abnormalities associated with strokes. It is non-invasive and can be performed at the patient's bedside. However, it is not specific for stroke detection, as abnormalities may be caused by other conditions, and it has limited ability to detect structural brain changes.

3. PROPOSED SYSTEM

The proposed system aims to develop a sophisticated brain stroke detection framework by leveraging machine learning and image processing techniques. Designed modularly, it facilitates a comprehensive and efficient analysis of brain images, enabling early and accurate stroke diagnosis. The process begins with data acquisition, where a diverse collection of CT scans and MRI slices is gathered. These images are carefully labeled to indicate the presence or absence of stroke, including details on stroke type and severity. Next, data preprocessing is performed to enhance model performance and consistency. This involves normalizing pixel values, resizing images to a uniform resolution, and applying noise reduction filters to improve image clarity. Feature extraction follows, focusing on identifying the region of interest within the brain images that are relevant for stroke detection. Techniques such as edge detection using algorithms like the Canny detector and texture analysis via Gray-Level Co-occurrence Matrix (GLCM) are applied to highlight important boundaries and textural patterns indicative of stroke. For algorithm selection, the system utilizes OpenCV's image processing capabilities for contour analysis, morphology, and filtering, while implementing machine learning models through OpenCV's module or integrating with frameworks such as TensorFlow or PyTorch. The dataset is divided into training and testing subsets to facilitate model evaluation. Extracted features are transformed into feature vectors to serve as input for the model. The system trains machine learning algorithms such as Support Vector Machines or Random Forests, employing cross-validation to ensure robustness. Model evaluation is conducted using the testing set, with performance assessed through metrics including accuracy, precision, recall, and F1-score. Once validated, the trained model is integrated with OpenCV for real-time image recognition and deployed in healthcare environments, ensuring smooth integration with existing infrastructure. What distinguishes this system is its integration of multi-source data—including electronic health records, genetic information, and lifestyle factors—offering a comprehensive and personalized stroke risk analysis. It employs advanced machine learning techniques that balance high accuracy with interpretability. Feature selection methods enhance model clarity by highlighting the most influential risk factors. Real-world datasets validate the system's reliability and practical utility. Compared to traditional clinical methods, this approach demonstrates superior prediction accuracy and earlier stroke detection. Furthermore, it supports clinicians with actionable insights for informed decision-making and personalized interventions. Designed for scalability and adaptability, the system can be deployed across various healthcare settings and populations, with a user-friendly interface that ensures medical professionals can easily interpret the results.

4. METHODOLOGY

This section will go over the research strategy and technique that we used for this investigation. In Figure 5.1, 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.

5. SYSTEM ARCHITECTURE:

Fig 5.1 Proposed CNN Model

Model Summary:

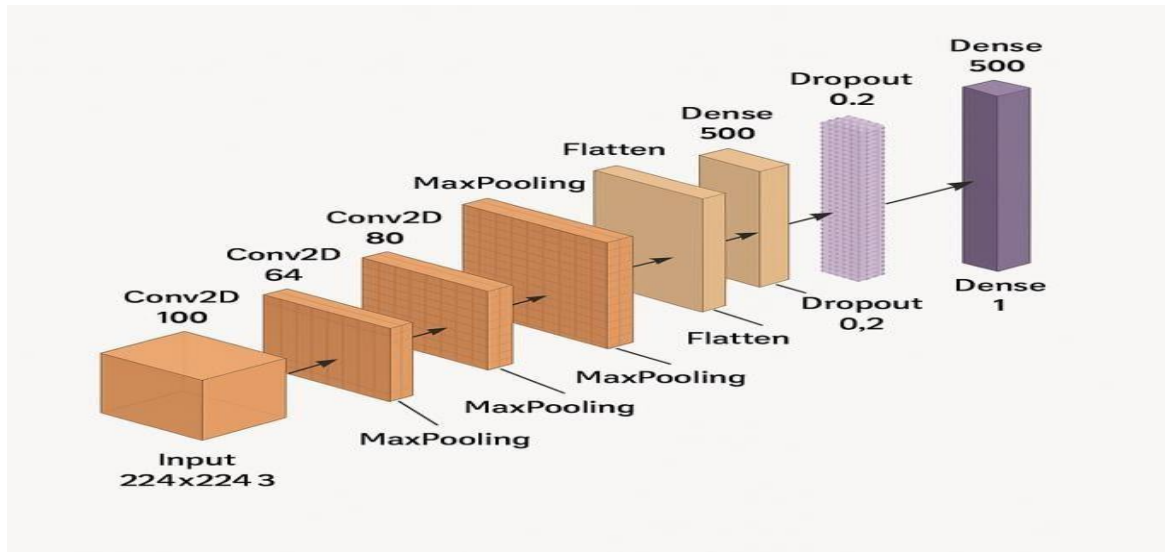


Fig 5.1 describes the proposed system architecture which is defined as below

- A deep CNN with 4 convolutional layers, each followed by max pooling
- Ends with two dense layers (500 units each) and dropout
- Final dense layer with 1 output, indicating a binary classification task

6. DATA ANALYSIS & RESULTS

CNN Model Performance:

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

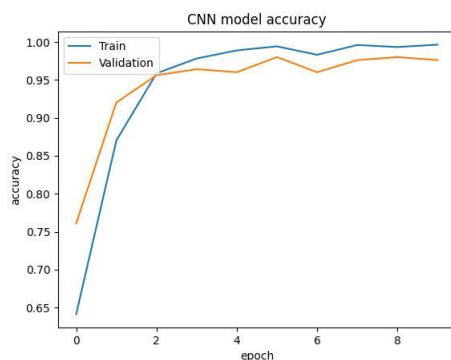


Figure 6.1: Training and Validation Accuracy Validation Loss

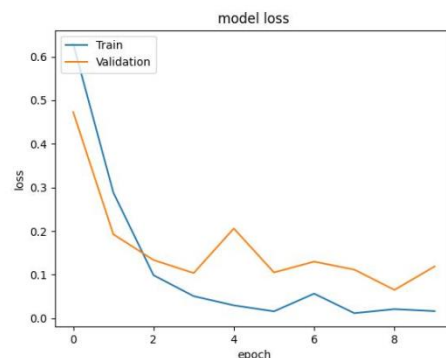
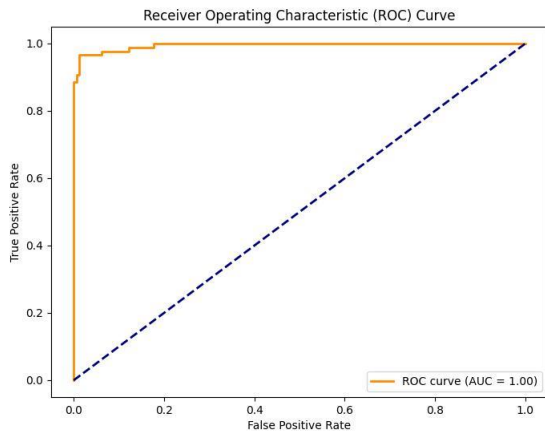


Figure 6.2: Training and Validation Loss



**Figure 6.3: ROC Curve
Recall Curve**

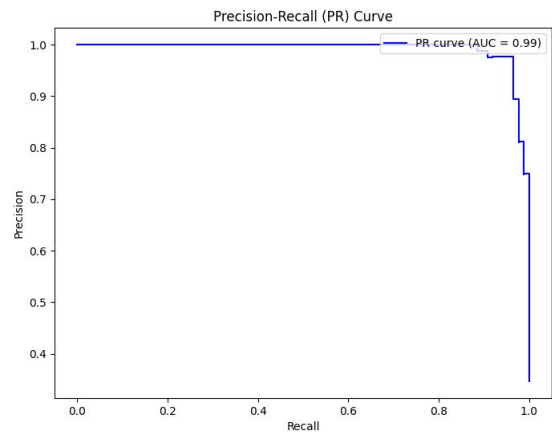
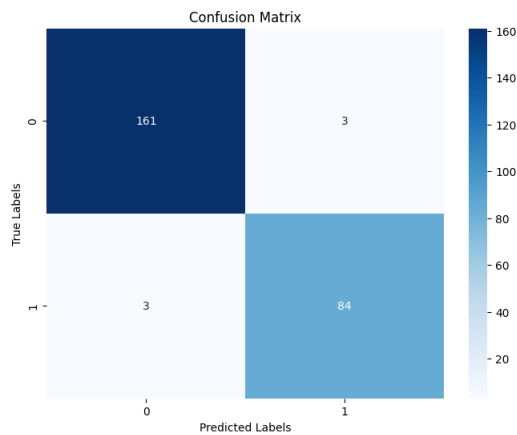


Figure 6.4: Precision



**Figure 6.5: Confusion Matrix
Classification Report**

| | 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 6.6:

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 |

Table 6.1. Model Comparison

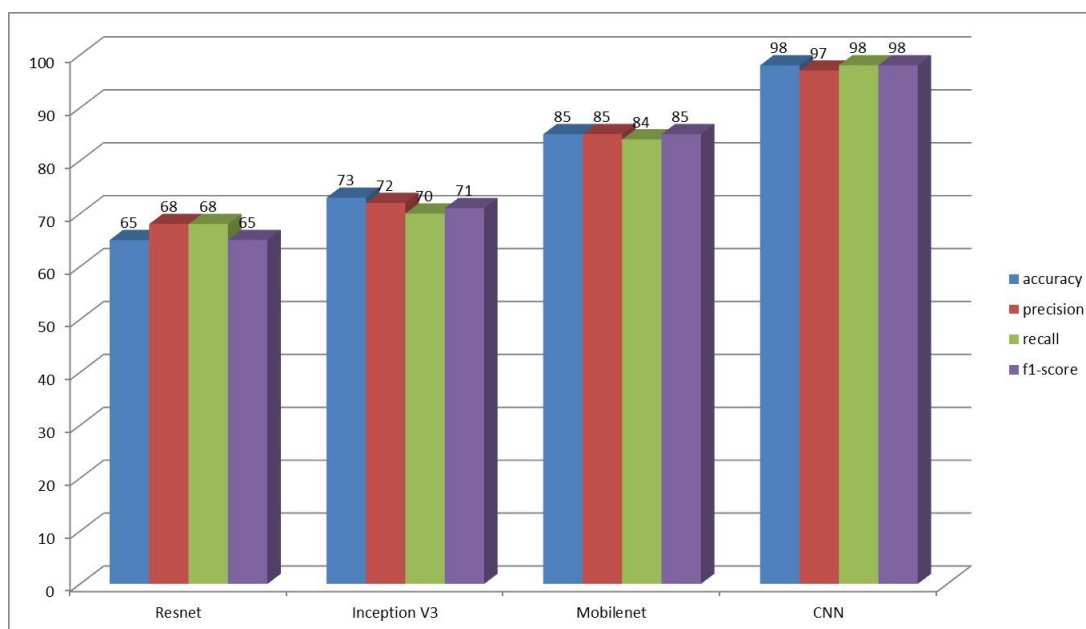


Fig 6.7 Accuracy comparison

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

The above bar graph in Fig 6.7 compares the performance of four deep learning models—ResNet, Inception V3, MobileNet, and a custom CNN—across four evaluation metrics: accuracy, precision, recall, and F1-score.

- Custom CNN outperforms all other models significantly, achieving 98% accuracy, 97% precision, 98% recall, and 98% F1-score.
- MobileNet performs reasonably well with all metrics around 84–85%, showing a balanced performance.
- Inception V3 achieves moderate results with accuracy around 73%, while ResNet performs the lowest across all metrics, especially in accuracy (65%).

7. CONCLUSION AND FUTURE SCOPE

The “Brain Stroke Detection System based on CT Images using Deep Learning” highlights the powerful role of artificial intelligence—especially Convolutional Neural Networks (CNNs)—in enhancing both the speed and accuracy of stroke diagnosis. Trained on a diverse dataset, the system achieved high performance with a training accuracy of 98.00% and a validation accuracy of 97.00%, demonstrating its effectiveness in accurately identifying stroke from CT images. The user-friendly web interface, built with Flask and designed using HTML, CSS, and JavaScript, offers healthcare professionals a seamless way to upload images and receive rapid diagnostic results. This project offers a practical and scalable solution for early stroke detection, which is crucial in urgent, resource-limited healthcare settings. By reducing dependency on manual interpretation, minimizing diagnostic errors, and

delivering quick and consistent results, the system serves as an efficient tool for preliminary stroke screening.

Future Scope

Although the current system performs robustly, several opportunities exist for further development: Model Refinement and Dataset Growth: Utilizing more advanced CNN architectures like ResNet or EfficientNet could further improve accuracy. Expanding the dataset to cover a broader range of stroke types, including ischemic and hemorrhagic strokes, will enhance the model's applicability. Integration of Multi-modal Data: Incorporating additional imaging methods such as MRI, along with clinical data like symptoms, patient age, and medical history, could boost diagnostic accuracy and clinical usefulness. Mobile Application Development: Developing a mobile app version of the system would increase accessibility, especially in rural or under-resourced areas with limited medical infrastructure. Seamless Hospital Integration: Connecting the system to hospital databases and electronic health record (EHR) systems would optimize workflow and enable comprehensive patient care management. Explainable AI Features: Incorporating interpretability tools such as heatmaps or Grad-CAM would help clinicians understand the model's decision-making process, building trust and transparency. Real-time Detection and Alert Systems: Future versions could enable real-time image analysis and be integrated with emergency response workflows to automatically alert healthcare providers of potential stroke cases needing immediate attention.

REFERENCES

- 1) Agrawal, D., Poonamallee, L., & Joshi, S. (2023). Automated Detection of Intracranial Hemorrhage from Head CT Scans Applying Deep Learning Techniques in Traumatic Brain Injuries: A Comparative Review. *Indian Journal of Neurotrauma*, 20(2), 81–88. <https://doi.org/10.1055/s-0043-1770770>.
- 2) T. D. Phong, N. T. Trong, H. N. Duong, V. H. Nguyen, V. Snasel, H. T. Nguyen, and
- 3) V. H. Tran, "Brain hemorrhage diagnosis by using deep learning," in *Proceedings of the 2017 9th International Conference on Knowledge and Systems Engineering (KSE)*, Jan. 2017, pp. 256–261. doi: 10.1145/3036290.3036326
- 4) J. S. Castro, S. Chabert, C. Saavedra, and R. Salas, "Convolutional neural networks for detection intracranial hemorrhage in CT images," *Proceedings of the 4th Congress on Robotics and Neuroscience*, Mar. 2020. [Online]. Available: <https://www.researchgate.net/publication/339795647>
- 5) S. Ahmed, J. F. Esha, M. S. Rahman, M. S. Kaiser, A. S. M. S. Hosen, D. Ghimire, and M. J. Park, "Exploring Deep Learning and Machine Learning Approaches for Brain Hemorrhage Detection," *IEEE Access*, vol. 12, pp. 45060–45081, 2024, doi: 10.1109/ACCESS.2024.3376438.
- 6) .Yeo, M., Tahayori, B., Kok, H. K., et al. (2023). Evaluation of techniques to improve a deep learning algorithm for the automatic detection of intracranial haemorrhage on CT

head imaging.European Radiology Experimental, 7(1), 17.
<https://doi.org/10.1186/s41747-023-00330-3>

- 7) Zhang, Y., Liu, Y., Wang, Y., et al. (2020). Deep learning for hemorrhagic lesion detection and segmentation on brain CT images. *Computers in Biology and Medicine*, 124, 103956.
- 8) Yang, Y., Zheng, J., Du, Z., Li, Y., & Cai, Y. (2021). Accurate prediction of stroke for hypertensive patients based on medical big data and machine learning algorithms: Retrospective study. *JMIR Medical Informatics*, 9(11), e30277.
<https://doi.org/10.2196/30277​>