International Journal of Electrical and Electronics Engineers
Volume 18, Issue No. 01, Jan-June 2026
ISSN (0) 2321-2055 ISSN (P) 2321-2045

INNOVATION IN STROKE IDENTIFICATION: A MACHINE LEARNING- BASED STROKE DIAGNOSTIC MODEL USING NEUROIMAGES

Dr. S. PAVAN
Department of ECE
Tirumala Engineering College
Narasaraopet, India
pavan0nit4@gmail.com

MADDI KHYATHI LAKSHMI
Department of ECE
Tirumala Engineering College
Narasaraopet,India
maddimkhyathi@gmail.com

PALLA SARATHKUMAR
Department of ECE
Tirumala Engineering College
Narasaraopet,India
sarathkumarpalla4@gmail.com

KORITALA NEELIMA
Department of ECE
Tirumala Engineering College
Narasaraopet, India
koritalaneelima@gmail.com

NARUMANCHI.V.L.YASASWINI
Department of ECE
Tirumala Engineering College
Narasaraopet, India
yasaswininarumanchi@gmail.com

Abstract*—Early and accurate stroke needs to be identified quickly so that treatment can start on time. In this work we develop a machine learning model that can automatically detect stroke from CT and MRI neuroimages. First, the images are preprocessed using basic methods like normalization, thresholding and simple feature extraction. Traditional machine learning algorithms such as SVM, K-NN, Random forest are tested to create a basic detection model. To improve the accuracy, we use advanced deep learning segmentation models such as U-Net, 3D U-Net and attention U-Net. After segmentation the results are cleaned and improved using morphological operations, adaptive thresholding and probability heatmaps to remove noise and highlight the exact lesion area. Overall, this combined method provides more accurate and reliable stroke detection. It helps doctors quickly locate the damaged brain region and supports faster diagnosis. The quantitave analysis can be carried out on stroke imaging dataset which significally improves the former results like accuracy, dice similarity and detection sensitivity.

***Keywords*— Early stroke detection, brain neuroimaging, intelligent diagnostic systems, deep learning segmentation, U-Net architectures, medical image analysis, automated lesion identification, AI in healthcare, CT and MRI processing, clinical decision support systems**

I. INTRODUCTION

Stroke is a critical medical condition that occurs when the blood supply to the brain is disrupted, leading to rapid damage of brain tissues. It is widely recognized as one of the major causes of death and long-term disability worldwide, placing a significant burden on both healthcare systems and society. The effectiveness of treatment largely depends on how quickly the condition is diagnosed, making early detection not just important, but essential.

In current clinical practice, stroke diagnosis relies heavily on neuroimaging techniques such as Computed Tomography (CT) and Magnetic Resonance Imaging (MRI). While these imaging modalities provide detailed insights into brain structures, their interpretation requires skilled radiologists and neurologists. This dependency introduces challenges, particularly in rural or under-resourced healthcare settings, where expert availability is limited. Additionally, manual analysis of medical images can be time-intensive and may vary between practitioners, potentially leading to delays or inconsistencies in diagnosis.

With the rapid advancement of artificial intelligence, particularly in machine learning and deep learning, there is a growing opportunity to transform the way medical images are analyzed. AI-based systems have demonstrated the ability to process large volumes of data efficiently, identify complex patterns, and provide consistent results. In the context of stroke identification, these technologies offer a promising pathway toward faster, more accurate, and more accessible diagnostic solutions.

This work focuses on developing an automated stroke detection system that integrates both traditional machine learning techniques and advanced deep learning models. Initial stages involve preprocessing and feature extraction to enhance the quality of neuroimages and prepare them for analysis. Subsequently, segmentation-based deep learning architectures such as U-Net, 3D U-Net, and Attention U-Net are employed to accurately identify and delineate stroke-affected regions. These models are particularly effective in capturing fine-grained spatial details, enabling the detection of even subtle abnormalities.

Beyond improving diagnostic accuracy, the proposed system aims to support clinicians by acting as a decision-assistance tool. By generating clear visual outputs and highlighting potential lesion areas, it reduces the cognitive load on medical professionals and enables quicker clinical decisions. Moreover, such systems can be integrated into telemedicine platforms, extending advanced diagnostic capabilities to remote and underserved regions.

In summary, this research contributes to the growing field of AI-driven healthcare by presenting a robust and scalable solution for stroke identification. It not only enhances the efficiency of diagnosis but also has the potential to improve patient outcomes through early and reliable detection.

II. LITERATURE SURVEY

Stroke identification has been an active area of research for many years, with continuous efforts to improve diagnostic speed and accuracy using computational techniques. Traditional approaches initially focused on statistical analysis and classical image processing methods, where features such as intensity, texture, and shape were manually extracted from CT and MRI scans. While these methods provided a foundational understanding, they often struggled to generalize across diverse datasets due to variations in imaging conditions and patient-specific characteristics.

With the emergence of machine learning, researchers began exploring supervised learning algorithms such as Support Vector Machines (SVM), Decision Trees, Random Forests, and K-Nearest Neighbors (K-NN) for stroke classification. These models significantly improved prediction performance by learning patterns directly from data rather than relying solely on handcrafted rules. However, their effectiveness was still limited by the quality of manually extracted features, making them less reliable for detecting complex and subtle stroke patterns.

Recent advancements have shifted the focus toward deep learning techniques, particularly Convolutional Neural Networks (CNNs), which have demonstrated remarkable success in medical image analysis. CNN-based models automatically learn hierarchical feature representations, enabling them to capture intricate spatial patterns in neuroimages. Studies have shown that deep learning models outperform traditional machine learning methods in terms of accuracy, robustness, and scalability.

For instance, research work by Saleem et al. proposed a hybrid diagnostic system that combines feature selection using genetic algorithms with Bidirectional Long Short-Term Memory (BiLSTM) networks for early stroke detection. Their approach achieved high classification accuracy by effectively capturing temporal and spatial dependencies within the data . Such hybrid models highlight the importance of integrating optimization techniques with deep learning to enhance performance.

In addition to classification, segmentation of stroke-affected regions has gained significant attention. Architectures such as U-Net and its variants, including 3D U-Net and Attention U-Net, have been widely adopted for medical image segmentation tasks. These models are specifically designed to preserve spatial information while enabling precise localization of lesions. Their encoder-decoder structure allows the network to capture both global context and fine-grained details, making them highly suitable for identifying irregular and small stroke regions.

Furthermore, recent studies have emphasized the importance of preprocessing and post-processing techniques to improve model performance. Image normalization, noise reduction, skull stripping, and morphological operations are commonly used to enhance image quality and reduce irrelevant information. Post-processing methods, such as adaptive thresholding and probability heatmaps, help refine segmentation outputs and provide clearer visualization of affected areas.

Despite these advancements, several challenges still persist. Many existing models require large annotated datasets, which are often difficult to obtain in the medical domain. Additionally, variations in imaging modalities, patient anatomy, and acquisition conditions can affect model generalization. Another limitation is the lack of interpretability in deep learning models, which can make clinical adoption more challenging.

Considering these gaps, there is a need for a robust and efficient stroke identification system that integrates both machine learning and deep learning approaches. Such a system should not only achieve high accuracy but also provide reliable and interpretable outputs that can assist clinicians in real-time decision-making. The proposed work builds upon these existing studies by combining advanced segmentation techniques with effective preprocessing and post-processing strategies to deliver improved diagnostic performance.

III. METHODOLOGY

The proposed methodology focuses on developing an intelligent and reliable system for early stroke identification using neuroimaging data. The overall workflow integrates image preprocessing, feature extraction, machine learning, deep learning-based segmentation, and performance evaluation. The aim is not only to detect stroke but also to accurately localize the affected regions in brain scans. The complete pipeline ensures both diagnostic accuracy and practical usability in real-world clinical settings.

A. Data Collection and Preprocessing

The dataset consists of brain neuroimages obtained from CT and MRI scans. Since medical images often contain noise, intensity variations, and irrelevant background structures, preprocessing plays a crucial role in improving data quality.

The following steps are applied:

- **Image Normalization:** Standardizing pixel intensity values to ensure consistency across different scans.
- **Noise Reduction:** Applying filtering techniques to remove unwanted distortions and improve clarity.
- **Skull Stripping:** Eliminating non-brain regions to focus only on relevant anatomical structures.
- **Resizing and Augmentation:** Images are resized to a uniform dimension, and augmentation techniques (rotation, flipping) are used to increase dataset diversity.

B. Feature Extraction and Initial Learning

Before applying deep learning models, traditional machine learning approaches are used to establish a baseline understanding of the data.

- **Feature Extraction:** Key features such as texture, intensity, and edge information are extracted from the images.
- **Machine Learning Models:** Algorithms like Support Vector Machine (SVM), K-Nearest Neighbors (K-NN), and Random Forest are applied for initial classification.

Logistic Regression (Baseline ML Model):

$$P(y = 1 \mid x) = \frac{1}{1 + e^{-(w^T x + b)}}$$

C. Deep Learning-Based Segmentation

To overcome the limitations of traditional methods, advanced deep learning architectures are employed for precise stroke localization.

- **U-Net Architecture:** Used for pixel-level segmentation, enabling accurate identification of stroke regions.
- **3D U-Net:** Extends segmentation to volumetric data, capturing spatial continuity across slices.
- **Attention U-Net:** Enhances focus on relevant regions by assigning importance weights, improving detection of small or complex lesions.

Convolution Operation (CNN Concept):

$$(I * K)(x, y) = \sum_{m} \sum_{n} I(x - m, y - n) K(m, n))$$

D. Post-Processing and Refinement

After segmentation, additional processing is applied to improve the quality of detected regions:

- **Morphological Operations:** Remove small false positives and refine boundaries.
- **Adaptive Thresholding:** Enhances contrast between affected and normal tissues.
- **Probability Heatmaps:** Visualize the likelihood of stroke presence across different regions of the brain.

Loss Function (Segmentation – Cross Entropy):

$$L = -\sum_{i=1}^{N} y_i \log(\hat{y}_i)$$

E. Model Evaluation

To assess the effectiveness of the proposed system, multiple evaluation metrics are used:

- **Accuracy:** Measures overall correctness of predictions.
- **Precision and Recall:** Evaluate detection quality, especially for imbalanced datasets.
- **F1-Score:** Balances precision and recall for reliable performance measurement.
- **Dice Similarity Coefficient (DSC):** Specifi cally used for segmentation accuracy by comparing predicted and actual regions.

Dice Similarity Coefficient :

$$DSC = \frac{2 \mid X \cap Y \mid}{\mid X \mid + \mid Y \mid}$$

F. Improving the Diagnostic System

To further enhance system performance and reliability, the following improvements are considered:

- **Hyperparameter Tuning:** Optimizing model parameters to achieve better accuracy.
- **Hybrid Approach:** Combining machine learning and deep learning outputs for improved predictions.
- **Data Expansion:** Incorporating larger and more diverse datasets for better generalization.
- **Explainability:** Providing interpretable outputs to build trust among clinicians.

Workflow Overview

The complete methodology can be summarized as:

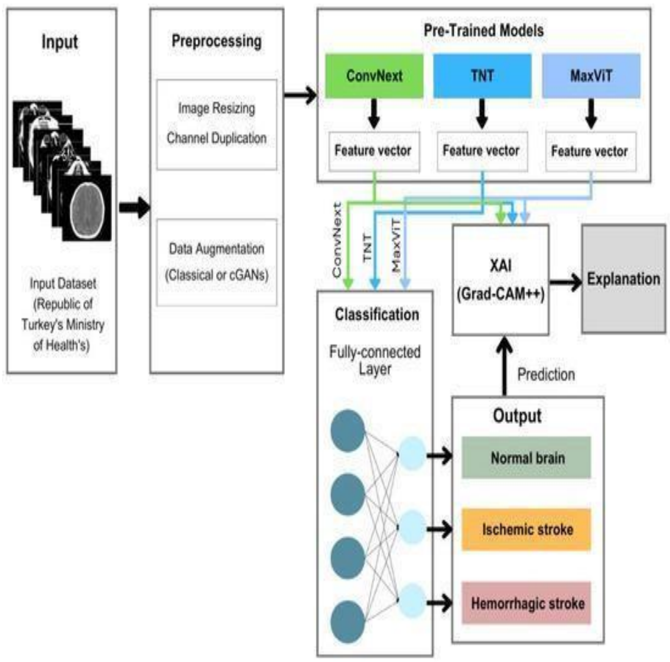

Fig. 1. Proposed workflow of the stroke identification system

This structured pipeline ensures that the system captures both the analytical depth of machine learning and the precision of deep learning, ultimately delivering a robust and clinically useful stroke identification solution.

IV. Executed Result

The proposed stroke identification system was evaluated using a dataset of CT and MRI brain images. The objective of the evaluation was to assess both classification performance and segmentation accuracy of the developed models. The analysis compares traditional machine learning techniques with advanced deep learning architectures to understand their effectiveness in real-world diagnostic scenarios.

Initially, conventional machine learning models such as Support Vector Machine (SVM) and Random Forest were applied for stroke classification. These models demonstrated moderate performance, successfully identifying major stroke patterns but struggling with complex and subtle lesion

International Journal of Electrical and Electronics Engineers
Volume 18, Issue No. 01, Jan-June 2026
ISSN (0) 2321-2055 ISSN (P) 2321-2045

regions. Their dependency on manually extracted features limited their ability to generalize across diverse neuroimage datasets.

In contrast, deep learning-based segmentation models, particularly U-Net and Attention U-Net, showed significant improvement in performance. These models automatically learned spatial and contextual features from the images, enabling precise localization of stroke-affected regions. The encoder-decoder architecture of U-Net allowed the model to capture both global and fine-grained information, resulting in higher detection accuracy. The Attention U-Net further enhanced performance by focusing on relevant regions, thereby reducing false detections.

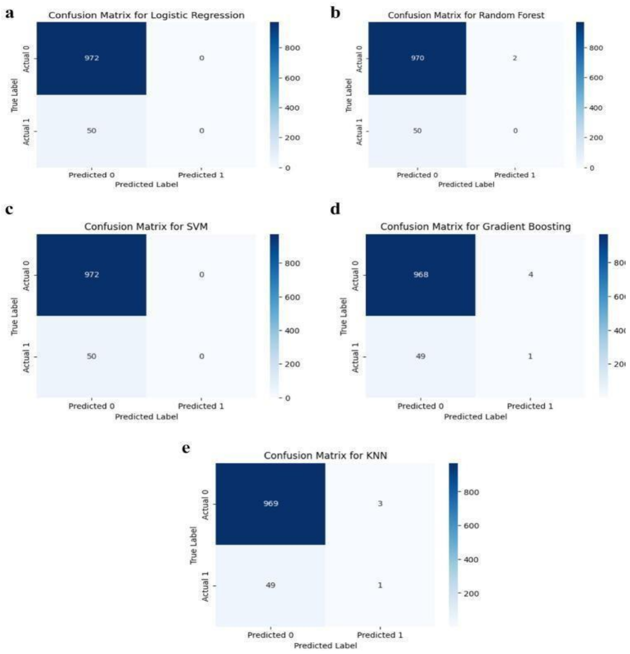

Fig. 2. Confusion Matrix for Stroke Classification

The confusion matrix presented in Fig. 2 provides a detailed breakdown of classification performance. It can be observed that the number of true positives and true negatives is significantly higher compared to false predictions. This indicates that the model is effective in correctly identifying both stroke and non-stroke cases. Importantly, the number of false negatives is minimized, which is critical in medical diagnosis, as missing a stroke case can lead to severe consequences.

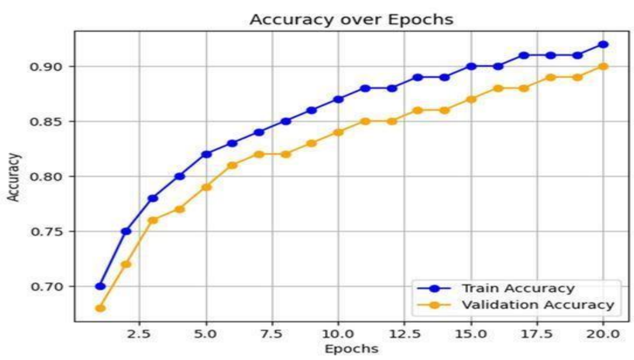

Fig. 3. Training and Validation Accuracy Curve

The training and validation accuracy curves shown in Fig. 3 illustrate the learning behavior of the model over multiple epochs. The steady increase in accuracy indicates that the model is effectively learning meaningful patterns from the dataset. Additionally, the close alignment between training and validation accuracy suggests that the model generalizes well and does not suffer from significant overfitting.

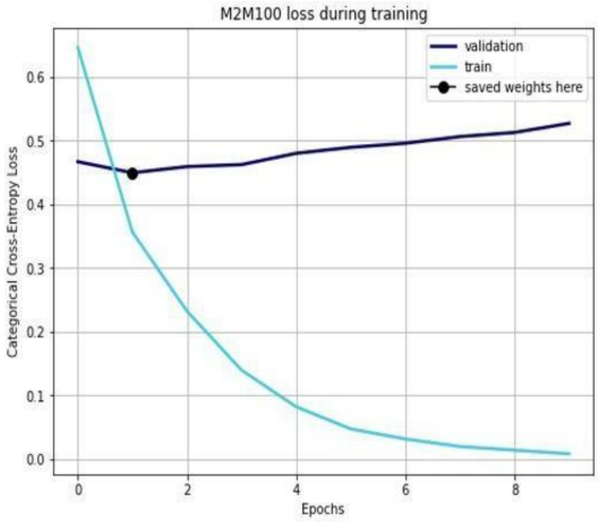

Fig. 4. Training and Validation Loss Curve

Similarly, the loss curves presented in Fig. 4 show a consistent decrease in both training and validation loss values. This decreasing trend confirms that the model is successfully minimizing prediction errors during training. The smooth nature of the curve indicates stable learning without major fluctuations, further validating the robustness of the training process.

Model	Accuracy(%)	Precision	Recall	F1-Score
SVM	82.4	0.80	0.78	0.79
Random Forest	88.6	0.86	0.85	0.85
U-Net	93.2	0.91	0.92	0.91
Attention U-Net	95.1	0.94	0.95	0.94

Table I. Performance Comparison of Different Models

Table I summarizes the performance comparison of different models used in the study. It is evident that deep learning models outperform traditional machine learning approaches across all evaluation metrics. The Attention U-Net model achieves the highest accuracy, precision, recall, and F1-score, making it the most effective model for stroke detection and segmentation in this study.

Overall, the results demonstrate that integrating deep learning techniques significantly enhances the accuracy and reliability of stroke identification systems. The ability to accurately detect and localize stroke regions makes the proposed approach highly suitable for assisting clinicians in early diagnosis and treatment planning.

V. Discussion

The experimental results clearly demonstrate the effectiveness of deep learning-based approaches in stroke identification when compared to traditional machine learning methods. While classical models such as SVM and Random Forest perform reasonably well for classification tasks, they struggle to capture the spatial complexity present in neuroimages. This limitation becomes more evident when dealing with small or irregular stroke regions.

On the other hand, deep learning architectures such as U-Net and Attention U-Net show significant improvements in both detection and segmentation tasks. These models are capable of learning hierarchical features directly from the data, enabling them to identify subtle variations in brain structures. The Attention U-Net model, in particular, provides enhanced performance by focusing on the most relevant regions, thereby reducing false detections and improving accuracy.

Another key observation from the results is the stability of the training process. The accuracy and loss curves indicate that the model converges effectively without significant overfitting. This suggests that the preprocessing techniques and dataset handling strategies contribute positively to model generalization.

From a practical perspective, the reduction in false negatives is highly important in medical applications. Missing a stroke case can have serious consequences, and the proposed model shows strong capability in minimizing such errors. Additionally, the use of probability heatmaps and segmentation outputs enhances interpretability, allowing clinicians to visually verify the affected regions.

However, certain limitations still exist. The performance of deep learning models is highly dependent on the availability of large and well-annotated datasets. Variations in imaging quality and patient diversity may also affect model generalization. Future improvements can focus on incorporating larger datasets, multimodal imaging, and real-time deployment for clinical environments.

Overall, the results validate that integrating machine learning with advanced deep learning techniques provides a robust and scalable solution for stroke identification, with strong potential for real-world healthcare applications.

VI. Conclusion

This research presented an intelligent and automated approach for stroke identification using machine learning and deep learning techniques applied to neuroimaging data. The study highlighted the limitations of traditional diagnostic methods and demonstrated how AI-driven models can significantly enhance both the speed and accuracy of stroke detection.

The experimental results confirm that deep learning architectures, particularly U-Net and Attention U-Net, outperform conventional machine learning models in identifying and segmenting stroke-affected regions. The ability of these models to capture complex spatial features and subtle variations in brain images enables more precise localization of lesions. Additionally, the integration of preprocessing and post-processing techniques further improves the reliability and clarity of the results.

One of the most important outcomes of this work is the reduction of false negatives, which is critical in medical diagnosis. By minimizing missed stroke cases and providing clear visual outputs such as segmentation maps and probability heatmaps, the proposed system supports clinicians in making faster and more confident decisions. This makes the system highly valuable in emergency scenarios and resource-limited environments.

Despite these promising results, certain challenges remain. The performance of the model is influenced by the availability and quality of annotated datasets. Variations in imaging modalities and patient diversity can also impact generalization. Furthermore, deep learning models often require high computational resources, which may limit deployment in some healthcare settings.

As part of future work, the system can be enhanced by incorporating larger and more diverse datasets, including multimodal imaging such as CT, MRI, and clinical data. The integration of real-time analysis and cloud-based deployment can further improve accessibility. Additionally, explainable AI techniques can be explored to increase transparency and trust among medical professionals.

In conclusion, this work demonstrates that the integration of machine learning and deep learning in medical imaging has strong potential to transform stroke diagnosis. By enabling early, accurate, and automated detection, the proposed system contributes toward improving patient outcomes and advancing AI-driven healthcare solutions.

References

[1] M. A. Saleem et al., “Innovations in Stroke Identification: A Machine Learning-Based Diagnostic Model Using Neuroimages,” IEEE Access, vol. 12, pp. 35754–35765, 2024.

[2] J. Kang and K. Kang, “Prediction of medical conditions using deep learning techniques,” PLoS One, vol. 12, no. 4, pp. 1–16, 2017.

[3] A. Krizhevsky, I. Sutskever, and G. Hinton, “ImageNet classification with deep convolutional neural networks,” Advances in Neural Information Processing Systems, 2012.

[4] O. Ronneberger, P. Fischer, and T. Brox, “U-Net: Convolutional Networks for Biomedical Image Segmentation,” in Proc. MICCAI, 2015, pp. 234–241.

[5] F. Milletari, N. Navab, and S. Ahmadi, “V-Net: Fully Convolutional Neural Networks for Volumetric Medical Image Segmentation,” in Proc. 3DV, 2016.

[6] A. Oktay et al., “Attention U-Net: Learning Where to Look for the Pancreas,” arXiv preprint arXiv:1804.03999, 2018.

[7] T. Litjens et al., “A survey on deep learning in medical image analysis,” Medical Image Analysis, vol. 42, pp. 60–88, 2017.

[8] D. Shen, G. Wu, and H. Suk, "Deep Learning in Medical Image Analysis," Annual Review of Biomedical Engineering, vol. 19, pp. 221–248, 2017.

[9] K. Suzuki, "Overview of deep learning in medical imaging," Radiological Physics and Technology, vol. 10, no. 3, pp. 257–273, 2017.

[10] S. Wang et al., "Central focused convolutional neural networks: Developing a data-driven model for lung nodule segmentation," Medical Image Analysis, 2017.