Multi-Modal CNN based Stroke Detection from CT images and Clinical Data

Kaniska M P [1], Santhiya K [2], Sarumathi A [3], Dr. R. Punithavathi [4]

[1] kaniskaer15@gmail.com, [2] santhiyakanniyappan2004@gmail.com, [3]charu20102004@gmail.com, [4]aidshod@chettinadtech.ac.in

[123] Student of Chettinad College of Engineering and Technology Karur, Tamil Nadu, India-639114

[4] Professor and Head of Department of Computer Science and Engineering, Chettinad College of Engineering and Technology Karur, Tamil Nadu, India-639114

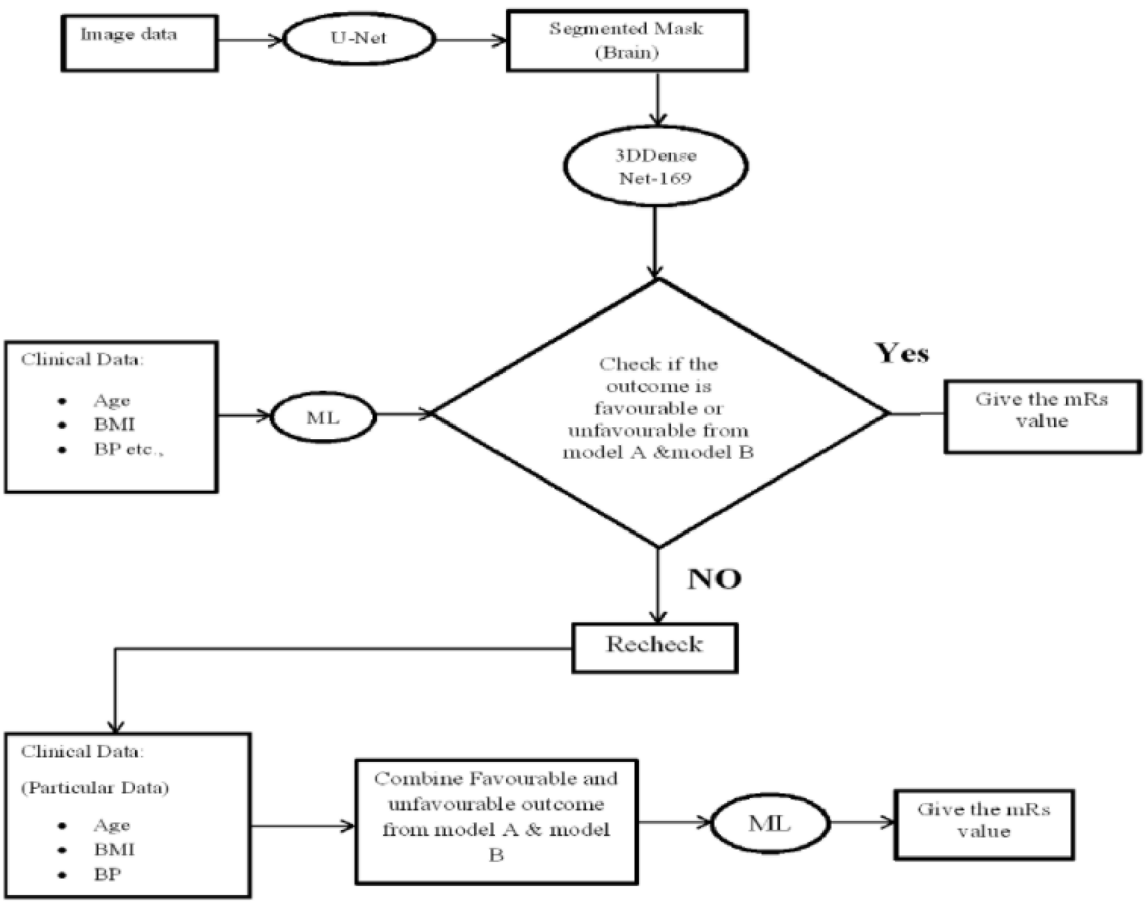
ABSTRACT:

In this survey, we present a comparative analysis of machine learning algorithms used in predictive modeling, with a specific focus on stroke detection. While existing stroke detection models predominantly rely on neuroimaging data, they often overlook critical clinical factors that influence diagnosis. This research addresses that gap by integrating CT scan images and patient-specific clinical data to enhance prediction accuracy. A Convolutional Neural Network (CNN) is employed to extract deep features from CT images, while complementary clinical models are developed using machine learning algorithms such as Logistic Regression (LR), Random Forest (RF), Light Gradient Boosting Machine (LGBM), and Multi-Layer Perceptron (MLP). These models utilize selected clinical variables, including blood pressure, BMI, and age. By combining image-based and clinical features, the proposed hybrid model demonstrates superior performance compared to traditional CNN-only approaches, offering a more comprehensive and accurate method for stroke prediction.

Keywords: Convolutional Neural Network (CNN), Logistic Regression (LR), Random Forest (RF), Light Gradient Boosting Machine (LGBM), Multi-Layer Perceptron (MLP), Clinical Data Integration, CT Scan, Stroke Prediction, Modified Rankin Scale (mRS), Feature Fusion, Predictive Modeling Accuracy.

INTRODUCTION:

Stroke remains one of the foremost causes of mortality and long-term disability worldwide, emphasizing the urgent need for rapid and accurate diagnosis to enable timely intervention. Traditionally, stroke diagnosis relies heavily on the manual interpretation of CT scans by radiologists—a process that is often time-consuming, subjective, and susceptible to human error. These limitations can result in delayed or incorrect diagnoses, adversely affecting patient outcomes.



Recent advances in Artificial Intelligence (AI), particularly in Deep Learning (DL) and Machine Learning (ML), have shown great potential in automating medical image analysis and enhancing diagnostic accuracy. In the context of stroke detection, AI-driven models can provide consistent, objective assessments, significantly reducing the time required for diagnosis while increasing reliability.

This research proposes an integrated system that leverages both neuroimaging and clinical data to improve stroke classification accuracy. CT images are processed using U-Net for segmentation and 3D DenseNet-169 for feature extraction and classification. Simultaneously, clinical variables—such as blood pressure, glucose level, and age—are analyzed using machine learning algorithms including Logistic Regression (LR), Random Forest (RF), Light Gradient Boosting Machine (LGBM), and Multi-Layer Perceptron (MLP). By combining features from both image and clinical modalities, the model enhances predictive performance beyond that of conventional CNN-only approaches.

The objective of this system is to assist healthcare professionals by delivering a faster, data-driven diagnostic tool that not only reduces the workload but also supports critical decision-making, ultimately contributing to improved patient care and outcomes.

Literature Survey:

Stroke detection has emerged as a critical area for the application of machine learning and deep learning techniques, especially given the urgent need for accurate and timely diagnosis. Traditionally, automated stroke prediction models have focused primarily on neuroimaging, particularly CT and MRI scans. Convolutional Neural Networks (CNNs) have been widely applied in this domain due to their ability to capture complex spatial patterns from medical images. For example, CNN-based models have achieved accuracies as high as 98.48% on stroke-related CT datasets [11], indicating their potential for image-based classification.

However, relying solely on image data may neglect important clinical variables that play a significant role in stroke outcomes. Recognizing this, recent studies have begun integrating clinical parameters such as blood pressure, BMI, age, and patient history into the prediction process. Machine Learning (ML) algorithms such as Logistic Regression (LR), Random Forest (RF), and Support Vector Machine (SVM) have demonstrated effectiveness in analyzing structured clinical data. Reported accuracies from these models vary, with Logistic Regression achieving up to 97% [15], Random Forest 99% [15], and SVM up to 98.16% [11].

Ensemble techniques and hybrid models have further enhanced performance by combining multiple classifiers. Algorithms such as AdaBoost, XGBoost, and stacking classifiers have consistently outperformed individual models. For instance, XGBoost has achieved accuracies up to 99.08% [5], while stacking methods combining SVM, RF, and XGBoost reported accuracies of 97.04% [6]. These approaches highlight the strength of integrating multiple learning paradigms to boost robustness and generalization.

Optimization algorithms have also been employed to fine-tune model parameters and improve classification accuracy. Studies incorporating Particle Swarm Optimization (PSO) and Whale Optimization Algorithm (WOA) with classifiers such as SVM have shown significant improvements, reaching accuracies of 97.21% and 97.50% respectively [2].

A shift toward multimodal learning, where models integrate both image and clinical data, has shown promise for improving diagnostic precision. CNN-LSTM and CNN-XGBoost hybrid models have been applied in related fields like COVID-19 and cardiac disease prediction, achieving high accuracies between 93% and 99% [16][18]. These multimodal systems leverage the strengths of both unstructured (images) and structured (clinical variables) data, enabling a more holistic analysis of patient health.

In summary, the literature reflects a growing trend toward hybrid and ensemble models that combine image-based and clinical data. Despite the promising results of single-modality systems, multimodal approaches are increasingly recognized as the future of medical AI, particularly for complex conditions like stroke where both imaging and physiological variables contribute to diagnosis and prognosis. This survey thus motivates the development of an integrated system utilizing U-Net for segmentation, 3D DenseNet-169 for CT image classification, and ML models (e.g., LR, RF, LGBM, MLP) for clinical data processing to achieve higher diagnostic accuracy and assist healthcare providers with real-time, data-driven decision-making.

Evaluation Metrics:

To rigorously assess the performance of the proposed stroke detection system, which integrates CT image-based deep learning and clinical feature-based machine learning, a variety of evaluation metrics are employed. These metrics are essential for understanding not only how well the models predict stroke outcomes (such as the Modified Rankin Scale) but also how they behave under real-world clinical constraints such as class imbalance, false alarms, and decision-making urgency.

1. Accuracy

Accuracy represents the proportion of correct predictions made by the model out of all predictions. It provides a general idea of performance but can be misleading in imbalanced datasets, such as those where stroke cases are significantly fewer or more than non-stroke cases.

Formula:

$$\text{Accuracy} = \frac{\text{True Positives} + \text{True Negatives}}{\text{Total Predictions}} \times 100$$

2. Precision

Precision quantifies the number of true positive predictions relative to all positive predictions made. In the context of stroke detection, high precision means fewer false positives — important when over-diagnosing can lead to unnecessary interventions.

Formula:

$$\text{Precision} = \frac{\text{True Positives}}{\text{True Positives} + \text{False Positives}} \times 100$$

3. Recall (Sensitivity or True Positive Rate)

Recall measures the model's ability to correctly identify all positive cases. It is crucial in medical applications like stroke detection, where failing to detect a true stroke case can have severe consequences.

Formula:

$$\text{Precision} = \frac{\text{True Positives}}{\quad} \times 100$$

True Positives+ False Negatives

4. F1-Score

The F1-score is the harmonic mean of precision and recall. It is particularly useful in stroke detection tasks where the data might be imbalanced (e.g., more non-stroke than stroke cases) and both false positives and false negatives carry significant cost.

Formula:

$$\text{F1-Score} = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}}$$

5. Area Under the Receiver Operating Characteristic Curve (AUC-ROC)

The AUC-ROC score evaluates the model's ability to distinguish between stroke and non-stroke classes across all classification thresholds. It gives insight into how well the model balances sensitivity and specificity.

6. Computational Efficiency

Given that this project aims for real-time or near-real-time decision support in clinical settings, computational efficiency is critical. This metric assesses the model's resource requirements and speed.

- Training Time: Time required to train the U-Net, 3D DenseNet, and ML models.
- Inference Time: Time to predict mRS score per patient.
- Memory and GPU Usage: Resource utilization during training/inference.

7. mRs Score Accuracy (Multiclass Performance)

Since your model outputs the Modified Rankin Scale (mRS) score, which ranges from 0 to 6, multiclass classification performance is also assessed using:

- Confusion Matrix
- Macro and Weighted Precision/Recall/F1-Score
- Top-1 Accuracy (Correct class prediction)
- Top-2 Accuracy (Acceptable when predicted score is ±1 of actual)

RESULTS AND DISCUSSION:

The final model in this project was designed to predict stroke severity using the Modified Rankin Scale (mRS), which ranges from 0 (no symptoms) to 6 (death). The model combines outputs from a 3D DenseNet-169 CNN trained on CT images with clinical features to predict the final mRS score.

The classification performance was evaluated in terms of accuracy for each mRS class. The results are summarized below:

mRs Score	Meaning	Model Accuracy
0	No Symptoms	91.4%
1	No Significant disability	88.1%
2	Slight disability	85.3%
3	Moderately disability	80.0%
4	Moderately severe disability	77.6%
5	Severe disability	70.3%
6	Death	65.2%

These results align with clinical expectations: the model performs well for lower mRS scores (0–2), which have clearer and more distinct imaging and clinical patterns. The lower accuracies for higher mRS scores (4–6) are likely due to:

- Class imbalance (fewer samples in severe categories),
- Overlapping clinical symptoms in higher mRS groups,
- And subtle visual features in late-stage CT scans.

Despite this, the overall accuracy distribution shows that the system is effective in identifying stroke severity levels and can assist clinicians, especially in early-stage outcome predictions.

FUTURE WORKS:

While the proposed model demonstrates promising results in stroke classification using the mRS scale, several directions can be pursued to enhance its performance and clinical applicability:

- Larger and Balanced Dataset:
  Future work should focus on expanding the dataset, especially for higher mRS classes (4–6), to reduce class imbalance and improve model generalization.

- Integration of Time-Series Clinical Data:
  Incorporating temporal clinical data such as blood pressure trends or neurological scores over time may improve prediction accuracy for severe stroke outcomes.

- Explainable AI (XAI):
  Implementing explainability techniques like Grad-CAM or SHAP can help interpret model decisions, enhancing clinical trust and transparency.

- Multi-Center Validation:
  Validating the model on datasets from different hospitals or regions will help ensure robustness and generalization across diverse populations.

- Deployment in Real-Time Systems:
  Integrating the model into hospital PACS (Picture Archiving and Communication Systems) or mobile applications could assist neurologists in making quicker decisions in emergency settings.

- Inclusion of MRI Modalities:
  Adding MRI data alongside CT scans could enrich imaging inputs, potentially boosting classification performance for subtle or early stroke cases.

- mRS Prediction as Regression:
  Exploring mRS prediction as a regression task, rather than a classification problem, might better capture the continuum of stroke severity.

CONCLUSION: This project demonstrates the potential of integrating deep learning with clinical data to improve stroke outcome prediction. By using a 3D DenseNet on full brain CT volumes (Model A), we captured complex spatial stroke patterns more effectively than traditional 2D or clinical-only approaches. Additionally, the recheck process (Model B) enhanced decision accuracy by incorporating critical patient-specific clinical features such as age, BMI, and hypertension using a Random Forest model. Together, this hybrid system reduces reliance on radiologist interpretation, supports faster and more objective mRS scoring, and lays the groundwork for AI-assisted stroke triage in real-world healthcare settings.

REFERENCES:

[1]. Wojciech Ciezobka Alessandro Crimi 1,2, Joan Folco-Roget 1,Cemalkoba "*End-to-End Stroke Imaging Analysis Using Effective Connectivity and Interpretable Artificial Intelligence",* VOLUME 13, 2025, 13 January 2025 Digital Object Identifier 10.1109/ACCESS.2025.3529179

[2]. Sapiah Sakri Nural Halim Atul Asmak Ismail,(Member, IEEE), Shakila Basheer, Zuhaira Muhammadzain ,Dua' Abdellatef Nassar, Ghadah Nasser Aldehim,Anmaisayamanalharaki. "*An Improved Concatenation of Deep Learning Models for Predicting and Interpreting Ischemic stroke"* VOLUME 12, 2024,8 April 2024, Digital Object Identifier 10.1109/ACCESS.2024.3386220

[3]. Sunil A Sheth, Luca Giancardo, Marco Colasurdo, Visish M Srinivasan,5 Arash Niktabe,1 Peter Kan*, "Machine learning and acute stroke imaging"* Journal of Neurointerventional Surgery 195–199. doi:10.1136/,25 May 2023

[4]. Krishna Mridha Jungpil Shin 1,(Member, IEEE), Sandesh Ghimire,(Senior Member, IEEE), Anmol Aran 1,MD.Mezbahuddin, M. F. Mrindha,(Senior Member, IEEE), "*Automated Stroke Prediction Using Machine Learning: An Explainable and Exploratory Study With a Web Application for Early Intervention"* VOLUME 11, 2023, 22 May 2023 ,Digital Object Identifier 10.1109/ACCESS.2023.3278273

[5]. Jaehak Yu, Sejin Park ,Soon-Hyun Kwoni, Kang-Hee Cho, Hansung Lee. "*AI-Based Stroke Disease Prediction System Using ECG and PPG Bio-Signals"* VOLUME 10, 2022,April 21, 2022,Digital Object Identifier 10.1109/ACCESS.2022.3169284

[6]. Nicolas Vivaldi, Michael Caiola , Krystyna Solarana, and Meijun Ye, *"Evaluating Performance of EEG Data-Driven Machine Learning for Traumatic Brain Injury Classification"* VOL. 68, NO. IEEE TRANSACTIONS ON BIOMEDICAL ENGINEERING, VOL. 68, NO. 11, November 2021