International Journal of Computer Information Systems and Industrial Management Applications
ISSN 2150-7988 Volume 18(17s) 2026 pp. 583-595
https://doi.org/10.70917/ijcisim-2026-4783
Article

Development of an Efficient Stacking-Based Machine Learning Framework for Brain Stroke Diagnosis Using CT Images

Ajay Kumar Mehta[1], Shipra Khandelwal [2], Nagendra Singh[3], Dimpy Singh[4], Rupayali Swaroop[5], Sunita Bhati[6]

[1,2,4,5]Department of Computer Science and Engineering, JECRC University, Jaipur, Rajasthan , India
[3]Department of Computer Science and Engineering, MNIT, Jaipur, Rajasthan , India
[6]Department of Computer Application, JECRC University, Jaipur, Rajasthan, India
Email: ajaymehta7684@gmail.com1 , shiprakhandelwal.khandelwal@gmail.com 2, 2025rcp9032@mnit.ac.in 3, dimpy0892@gmail.com4, srupayali@gmail.com5, bhati.sunita01@gmail.com6

Abstract: Brain stroke is a severe pathological condition caused by interrupted blood flow to the brain, resulting in possible neurological damage or death. Accurate and timely diagnosis through Computed Tomography (CT) scan images is a prerequisite for immediate medical intervention. The current research is proposing an effective stacking-based machine learning framework for brain stroke diagnosis automatically using feature extraction, feature selection, dimensionality reduction, and ensemble learning methods. The method starts with CT image preprocessing to convert images into grayscale, downsize them into a uniform 224×224 resolution, and normalize pixel intensity. Feature extraction is done for texture-based statistical features using Gray Level Co-occurrence Matrix (GLCM) and Local Binary Patterns (LBP) for texture pattern description. For improved model performance, SelectKBest (ANOVA F-test) is utilized for the most discriminative feature selection, and Principal Component Analysis (PCA) for dimensionality reduction. Synthetic Minority Over-sampling Technique (SMOTE) is utilized to address class imbalance and enhance generalization.A stacking ensemble learning model is built for classification purposes, with Logistic Regression (LR), Support Vector Machine (SVM), and Random Forest (RF) as base learners and XGBoost as the meta-learner. Hyperparameter tuning is carried out using GridSearchCV for model performance optimization. The results of the experiment, carried out on a labeled CT image dataset, show high accuracy at 96.21%, with precision at 0.95 (Normal) and 0.98 (Stroke), recall at 0.99 (Normal) and 0.92 (Stroke), and F1-score at 0.97 (Normal) and 0.95 (Stroke). These findings indicate that the proposed framework significantly improves stroke classification accuracy compared to conventional machine learning models.This research illustrates the potential of stacking-based ensemble learning in medical imaging and provides an efficient method for automatic, non-invasive, and accurate diagnosis of stroke. The proposed system has the ability to assist radiologists and healthcare professionals in achieving quick and accurate decisions, subsequently leading to improved patient outcomes.
Keywords: Brain Stroke, CT Imaging, Stacking Ensemble, Machine Learning, Feature Extraction, GLCM, LBP, PCA, SMOTE, Random Forest, SVM, XGBoost, Feature Selection

1. Introduction

Stroke is a significant cause of death and chronic disability globally, accounting for millions of people every year. It happens when the flow of blood to the brain is disrupted because of a blockage in one of the arteries (ischemic stroke) or the rupturing of a blood vessel (hemorrhagic stroke). The consequences of a stroke are typically devastating, leading to permanent neurological damage, loss of motor function, and, in the worst cases, death. Early and proper diagnosis is critical to the advent of early medical treatments, such as thrombolytic therapy or surgery, to prevent subsequent brain damage and improve patient outcomes.However, stroke diagnosis by traditional means is still a major issue, especially in the emergency department where timely and correct decision-making is critical [1].

Computed Tomography (CT) [2] scans are the most common imaging modalities for stroke detection because they can offer immediate and high-resolution imaging of the brain anatomy. CT scans enable radiologists to differentiate between normal and abnormal brain tissue, identify blood clots, and detect hemorrhagic strokes. Yet, the time-consuming and subjective manual reading of CT scans by neurologists and radiologists is prone to inter-observer variability, thus resulting in inconsistency in diagnosis. Moreover, the occurrence of complex patterns and minute abnormalities in CT scans of stroke patients renders it challenging for clinicians to maintain high diagnostic accuracy across all cases. The escalating demand for automated and accurate diagnostic solutions has promoted considerable interest in the use of machine learning (ML) and deep learning (DL) methods for supporting stroke detection [3].

Medical image analysis has, over the past few years, witnessed machine learning-based methodologies as potent agents for extracting features automatically, recognizing patterns, and classifying. Conventional machine learning algorithms tend to use hand-designed feature extraction techniques like texture analysis, histogram-based features, and morphological descriptors, which might not be comprehensive in accounting for the complex variations observable in stroke-related CT images.Moreover, the capability of generalizing stand-alone classifiers is weak and thus may be poor on new datasets. These limitations are improved by using ensemble learning methods, particularly stacking-based approaches, that have been proposed to increase accuracy and stability of classification [4].

This work introduces an efficient stacking-based machine learning system for computer-aided brain stroke diagnosis from CT images. The system uses several steps such as image preprocessing, feature extraction, feature selection, class balancing, and ensemble classification to create a high-performance diagnostic model. Preprocessing consists of converting the CT images into grayscale, resizing to a default size of 224×224 pixels, and normalizing the pixel intensity values for consistency. For feature extraction, texture descriptors like Gray Level Co-occurrence Matrix (GLCM) and Local Binary Patterns (LBP) are utilized to encode statistical patterns and neighborhood grey scale variation. Due to the high dimensionality of extracted features, the most discriminatory features are preserved using SelectKBest (ANOVA F-test) and dimensionality is reduced using Principal Component Analysis (PCA). As medical datasets are generally imbalanced, causing biased model predictions, Synthetic Minority Over-sampling Technique (SMOTE) is used to create synthetic samples for minority classes to have balanced training data [5].

The heart of the presented framework is the ensemble learning framework based on stacking, where the several base classifiers such as Logistic Regression (LR), Support Vector Machine (SVM), and Random Forest (RF) are stacked together to enhance predictive accuracy. The ultimate predictions are done by a meta-learner, in this research implemented through XGBoost, a gradient-boosting algorithm that is renowned for accuracy and efficiency. Hyperparameter tuning is done to optimize settings of classifiers and achieve optimal performance using GridSearchCV. The performance is measured with an extensive range of metrics, which are accuracy, precision, recall, F1-score, confusion matrix, Receiver Operating Characteristic (ROC) curve, and Precision-Recall curve for a complete measurement of the performance of the model.Experimental results reveal that the proposed stacking-based method in the paper achieves an accuracy of 96.21%, which is higher than simple machine learning algorithms with a large gap. Precision, recall, and F1-score values indicate very high reliability in stroke and non-stroke classification. They emphasize the capability of ensemble learning methods to revolutionize medical image-based stroke diagnosis, with a scalable, accurate, and automatic decision-making support tool for medical practitioners. The incorporation of machine learning into stroke diagnosis not only improves the accuracy of diagnoses but also relieves radiologists of some burden, enabling more rapid and objective medical decision-making [7].

Finally, this study introduces a new stacking-based machine learning model that accurately diagnoses stroke-related abnormalities in CT images. Through the utilization of ensemble learning, feature optimization, and class balancing algorithms, the developed model provides a very accurate and stable diagnostic tool for early stroke detection. This work contributes to the expanding frontier of AI-based medical imaging and can be potentially adopted in clinical workflows for the betterment of patient outcomes through early and accurate stroke diagnosis [7].

1.1 Objectives of this Research

The objectives of this research study can be summarised as

1. to foster a stacking-type machine learning model effectively for the diagnosis of strokes in the brain using CT scans.
2. to leverage machines and ensemble learning to maximize classification accuracy.
3. to extract meaningful features based on texture from CT scans, LBP (Local Binary Patterns) and GLCM (Gray Level Co-occurrence Matrix)
4. to apply SMOTE (Synthetic Minority Over-sampling Technique) to overcome the imbalanced data problem and enhancing model performance.
5. Assess the performance of the proposed model with respect to measures such as ROC curve analysis, accuracy, precision, recall, and F1-score.

The objectives are to bridge the gap between traditional image processing techniques and modern deep learning methods, offering a comprehensive solution for brain stroke detection.

2. Literature Survey

Freitas, C., et al. (2025) [8] explored a point-of-care diagnostic platform for stroke diagnosis based on silver nanostars (AgNS) and Surface-Enhanced Raman Spectroscopy (SERS) combined with machine learning (ML). The research distinguished hemorrhagic from ischemic strokes by classifying plasma samples spiked with glial fibrillary acidic protein (GFAP), a hemorrhagic stroke biomarker. The low-cost aluminum foil substrate permitted rapid classification within 15 minutes. The ML model with ensemble methods and feature engineering significantly improved classification performance and showed promise for point-of-care diagnostics.

Buchlak, Q. D., et al. (2024) [9] assessed the impact of a deep learning-aided system on brain non-contrast computed tomography interpretation. The research involved 32 radiologists interpreting 2,848 scans, comparing independent and supported performance. The deep learning model, which was trained on 212,484 scans, produced an AUC of 0.93. Assisted radiologists improved their AUC to 0.79 from 0.73 for parent interpretation and to 0.72 from 0.68 for child interpretation. The system greatly improved radiologist accuracy and decreased reading time.

Kanna, R. K., et al. (2024) [10] established a machine learning stroke prediction model from patient medical data. The research compared Random Forest, SVM, and Decision Tree classifiers with the highest accuracy of 94.30% for Random Forest. The integrated model in the Flask application rendered easy-to-understand stroke risk evaluations and lifestyle advice. The system exhibited efficient and cost-saving predictive power and supported early prevention and intervention.

Abujaber, A. A., et al. (2024) [11] aimed to enhance Posterior Circulation Syndrome (PCS) diagnosis with machine learning models trained on a national Stroke Registry dataset of 15,859 patients.Five ML models were compared, and the best was XGBoost (AUC: 0.81, accuracy: 0.79). SHAP analysis revealed important diagnostic factors, such as BMI, blood glucose, ataxia, dysarthria, and blood pressure. The study noted the application of ML in aiding early PCS diagnosis and enhancing clinical outcomes.

Nagarajan, A., et al. (2024) [12]explored the viability of using transfer learning for Motor Imagery (MI) Brain-Computer Interface (BCI) stroke rehabilitation. By pre-training models on healthy individuals and fine-tuning them for stroke patients, the research recorded a 71.15% MI detection accuracy, with a 18.15% enhancement after transfer learning. The strategy surpassed subject-specific models and performed similarly to stroke-to-stroke transfer learning. Explainable AI analysis revealed pertinent cortical areas involved in MI detection, highlighting the clinical value of this approach to stroke rehabilitation.

Conclusively, **Gudadhe S. S., et al. (2023)** [13] discuss the classification of intracranial hemorrhages (ICH), a life-threatening condition caused by ruptured blood vessels in brain tissues. The authors utilize texture-based methods, such as Local Binary Pattern (LBP), Local Ternary Pattern (LTP), and Weber Local Descriptor (WLD), to extricate features from CT images in one study. The feature vectors are generated by concatenating histogram bins of the respective texture codes. As found through testing with ten different classifiers, WLD performed better than all other descriptors in ensemble-based classification tasks. Among all classifiers, Random Forest delivered the best performance, providing accuracy of 86.55%, recall of 86.31%, precision of 87.23%, sensitivity of 86.31%, specificity of 86.81%, and F1-score of 86.77%.

Gudadhe, S., et al. (2023) [14] also propose a joint feature selection approach mixing texture and transform features (GLCM, DWT, and DCT) to classify ICH in CT images. They apply SMOTE to tackle class imbalance in the dataset, while applying Sequential Forward Selection to fine-tune the feature set of interest in the classification of ICH contents. The classifiers are evaluated via ensemble machine learning through accuracy, precision, and recall. Random Forest fits the model best, securing an accuracy of 87.22% while using a feature subset of six key features.

Together, these studies highlight the contributions of machine learning and AI to the improvement of stroke detection, diagnosis, and rehabilitation, augmenting clinical decision-making and patient outcomes.

3. Research Methodology

The research methodology of this study is a methodical procedure to design an efficient stacking-based machine learning model for diagnosing brain stroke from CT scans [15]. The study begins with the collection of CT scan images from public domain medical imaging repositories and hospital records. The images are split into two categories: stroke-impacted and healthy brain scans. Since medical image datasets are often class-imbalanced, data augmentation techniques such as Synthetic Over-sampling Technique (SMOTE) are employed to generate synthetic samples of the minority class so that they are equally represented. Moreover, all images are preprocessed by converting them to grayscale, resizing them to a common size of 224×224 pixels, and normalizing them to improve computational efficiency while training the model [15].

Feature extraction becomes an important role in identifying regions affected by a stroke from standard brain structures. Two main techniques of feature extraction, Gray Level Co-occurrence Matrix (GLCM) [14] and Local Binary Patterns (LBP), are employed. The GLCM processes textural features like contrast, correlation, energy, and homogeneity, from which useful information regarding the co-occurrences of pixel intensity values in space is obtained through the CT scans. In parallel, LBP converts pixel intensity changes to define fine texture detail, which is extremely suitable for identifying patterns in medical images. The feature sets from both approaches are merged to yield an entire description of all CT images. In order to further improve the performance of the model, Principal Component Analysis (PCA) is implemented to compress the feature space and preserve the essential information [15].

The work utilizes the stacking-based ensemble learning for enhancing the stroke diagnosis accuracy. The underlying models in the ensemble are Logistic Regression (LR), Support Vector Machine (SVM), and Random Forest (RF). The underlying models are trained separately on the found feature sets so that predictions are made by them based on their learned parameters.The output of all these base classifiers is then provided as input to a meta-learner, an XGBoost classifier. The meta-learner combines the outputs of all the base models to produce the ultimate classification decision, taking advantage of the strength of each model but not their disadvantage. Hyperparameter tuning is performed with GridSearchCV to ensure best-of-breed settings for each model, thereby enhancing the overall prediction accuracy [16].

The trained model is measured using several performance metrics to determine its efficiency in stroke diagnosis. Accuracy is applied to quantify the general correctness of the model's predictions, while precision and recall show the model's capabilities to accurately diagnose stroke cases while ensuring minimum false positives and false negatives. The F1-score as the harmonic mean of recall and precision ensures there is an even assessment of model performance. Confusion matrix is also used to visually represent classification output, detailing the true positives, false positives, true negatives, and false negatives. The Receiver Operating Characteristic (ROC) curve and the Area Under the Curve (AUC) are studied to graphically illustrate the sensitivity versus specificity trade-off, verifying the capability of the model in discriminating between normal and stroke, shown in Figure 1 [17].

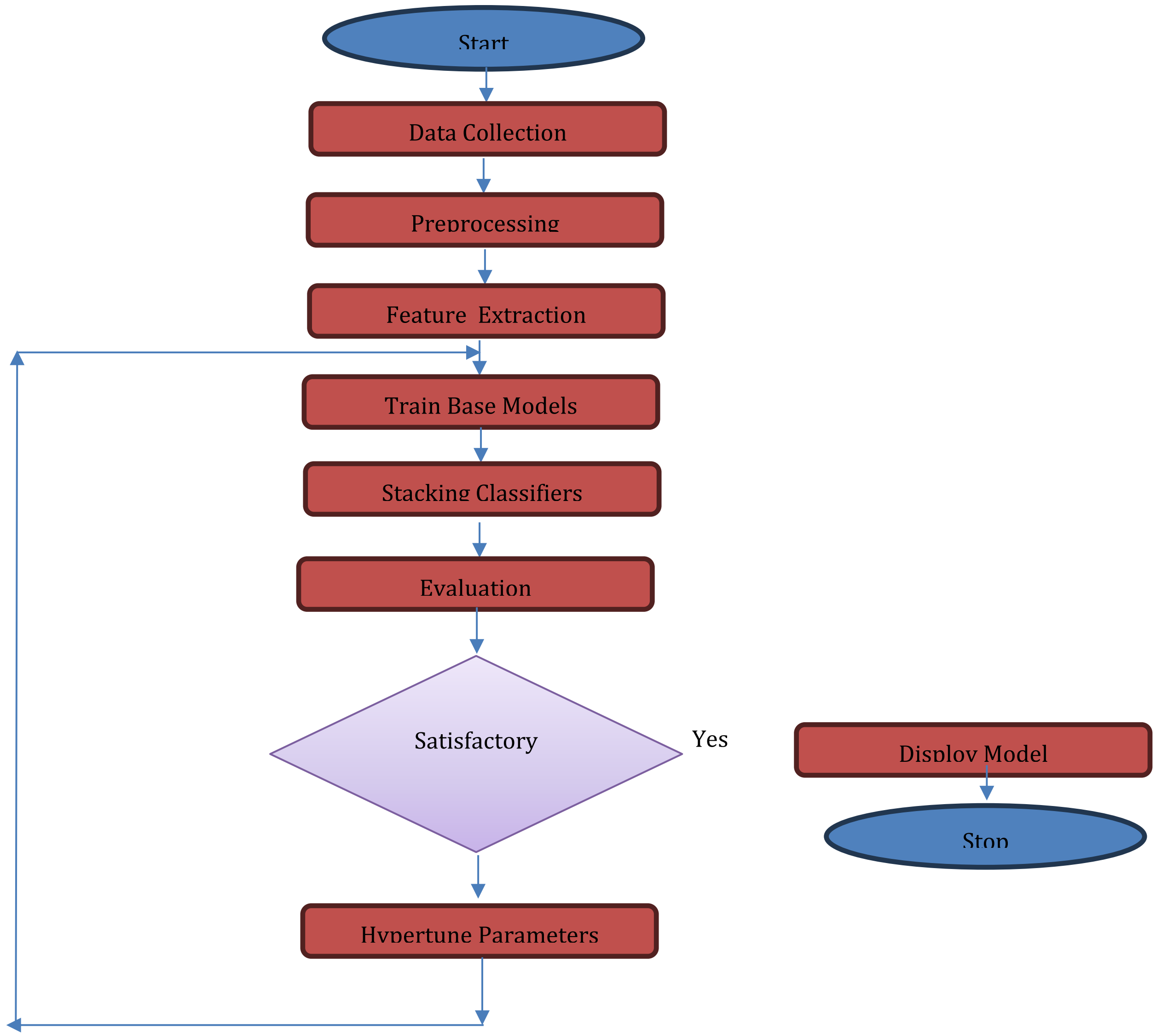

Figure 1: Flow Diagram of Research

To validate the effectiveness of the proposed stacking ensemble technique, its performance is contrasted with individual classifiers such as SVM [18], Random Forest [19], and Logistic Regression [20]. The comparison reveals the superior classification accuracy of the ensemble method, demonstrating its effectiveness as a reliable tool in stroke diagnosis. Furthermore, feature importance analysis is conducted to identify the most influential extracted features contributing to the classification process. Through this systematic approach, the research guarantees that the suggested model is efficient and reliable for helping medical professionals diagnose brain strokes from CT images. This research helps in the development of artificial intelligence-based medical diagnostics, which eventually leads to better early stroke detection and patient outcomes [21].

4. Proposed Approach

4.1 Dataset:

The database employed in this study is the Brain Stroke CT Image Dataset, hosted publicly on Kaggle and generated by Afridi Rahman. The dataset is instrumental in stroke diagnosis research due to its availability of 2,501 annotated brain CT scan images, which makes it feasible to train and test machine learning models for stroke detection. Since stroke diagnosis from CT scans is an extremely advanced and critical medical process, an accurately labeled dataset is a major factor in the success of deep learning models to identify patterns in strokes [22].The data is prepared in binary classification, and the images are divided into two classes. The first class holds 1,551 samples of regular brain scan images, while the second class holds 950 samples of brain images after a stroke. Class imbalance problem should be addressed during model training as data augmentation or class balancing techniques might need to be employed to enable the model to be able to correctly classify both classesshown in Figure 2 [22].All the images in the dataset are given in grayscale mode, typical in medical imaging because it enhances the contrast of important features and minimizes computational complexity. The organized structure of the dataset in which all the images have predetermined labels makes the dataset highly suitable for classification applications in machine learning and deep learning. The dataset properties allow scholars to develop and improve models for automatically detecting strokes in brain CT scans, driving innovation in stroke diagnosis and medical imaging at an early point [22] .By tapping into this dataset, this study seeks to develop a high-performance stacking-based machine learning model that is effective in separating normal and stroke-injured CT scans. The dataset is used as a basis for training, validating, and testing the model such that the suggested method can be practically utilized in actual clinical practice.

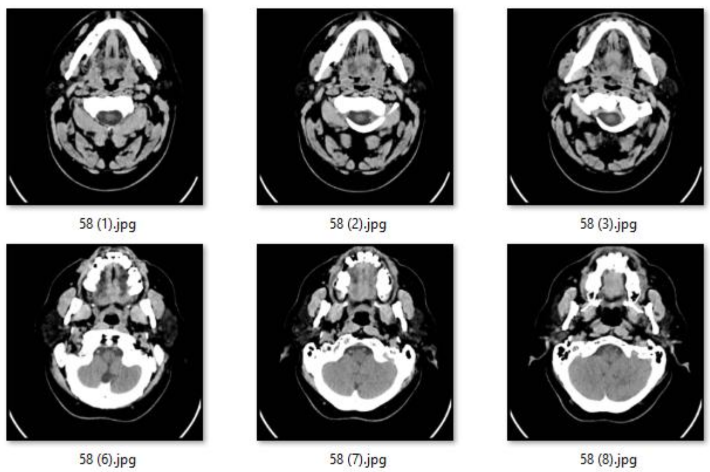

Figure 2: Brain Stroke Dataset (Stoke Images)

Algorithm

Input: Brain Stroke CT Image Dataset D with labeled images (X, Y)**Output:** Trained Stacking Classifier for Stroke Prediction

Step 1: Data Collection Obtain dataset D consisting of grayscale CT scan images with stroke and normal labels.

Step 2: Data Preprocessing Resize images to a fixed dimension (h, w). Normalize pixel values to range $[0,1]$:

$$X' = \frac{X - X_{min}}{X_{max} - X_{min}} \tag{1}$$

Apply data augmentation (rotation, flipping, contrast enhancement) if necessary. Split dataset into training (D_{train}), validation (D_{val}), and testing (D_{test}) sets. This equation calculates the normalized pixel values of the CT images within the range of [0, 1]. This ensures that all the input features are standardized at a common scale and thus, promote better stability and performance of machine learning models, especially for those sensitive to input magnitude-take this as an instance: neural networks.

Step 3: Feature Extraction Extract relevant features using deep learning-based or handcrafted techniques.

Step 4: Train Base Models Train multiple base learners $\{M_1, M_2, \ldots, M_n\}$ on D_{train}. Base models include classifiers such as SVM, CNN, Random Forest, and XGBoost.

Step 5: Stacking Classifier Generate meta-features using base model predictions:

$$Z = f\big(M_1(X), M_2(X), \ldots, M_n(X)\big) \tag{2}$$

Train a meta-classifier M_{meta} (e.g., Logistic Regression) on the base model outputs. This equation describes the process of meta-feature generation for a stacking ensemble. Each of the base models Mi makes predictions on the input X, and these predictions are put together as a new feature vector Z. The new vector is an input to the meta-classifier that tries to learn to make a final decision by combining the strengths of all the base classifiers.

Step 6: Model Evaluation Evaluate performance on D_{test} using metrics:

Accuracy:

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \tag{3}$$

Precision:

$$Precision = \frac{TP}{TP+FP} \tag{4}$$

Precision is defined as true positives over the total predicted positives. In other words, if it predicted something was positive how often was it actually correct? Hence precision should have a low false positive rate with high precision

Recall:

$$Recall = \frac{TP}{TP+FN} \tag{5}$$

Recall, sensitivity, or true positive rate is defined as the ratio of actual positive samples to positive samples that are identified or recognized by the model. Hence, it answers the question of how good the model is at identifying true cases of stroke: high recall will yield low false-positive rates

F1-Score:

$$F1 = 2 \times \frac{Precision \times Recall}{Precision + Recall} \tag{6}$$

The F1 Score is the harmonic mean of precision and recall; thus, it provides a balanced assessment which considers the occurrences of both false positives and false negatives. This score would be particularly relevant in cases of imbalanced datasets as it weighs the trade-off more towards precision and recall.

If performance is unsatisfactory, proceed to hyperparameter tuning.

Step 7: Hyperparameter Tuning Optimize model parameters using Grid Search or Bayesian Optimization. Retrain models with best parameters and re-evaluate.

Step 8: Deployment If performance meets clinical standards, deploy the trained model for real-time stroke detection.

4.2 Results

Confusion matrix shown in the figure3. is one of the most important measures for assessing the classification model employed to differentiate between "Normal" and "Stroke" cases. It consists of four main components: True Positives (TP), True Negatives (TN), False Positives (FP), and False Negatives (FN). All of these metrics give feedback regarding predictive accuracy of the model and what should be addressed.

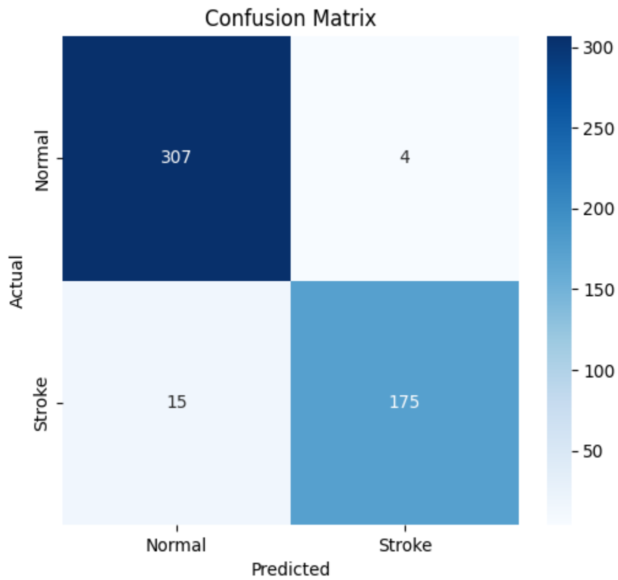

Figure3. Confusion Matrix Analysis

1.True Positives (TP): True positives are pointing out how many times the model came up with the right prediction for the condition "Stroke." Here, based on this confusion matrix, 175 stroke conditions were correctly tagged with "Stroke." This means that the model learned to identify stroke patterns accurately in most of the cases. High TP is very important in the case of medical diagnosis, where it ensures most of the actual stroke cases receive required attention and treatment.

2. True Negatives (TN): True negatives refer to the number of instances in which the model correctly classified a "Normal" condition. Here, 307 normal cases were accurately identified as "Normal." This means the model is able to avoid unnecessary false alarms by correctly separating normal patients from patients who have suffered a stroke. A high value of TN prevents unnecessary medical treatments and stress to non-stroke patients.

3. False Positives (FP): False positives are when the model incorrectly identifies a normal case as a stroke case. In this table, 4 "Normal" cases were labeled as "Stroke" by mistake. Although this is a fairly small number, false positives can lead to unnecessary medical tests and patient worry. However, in medical application, a slightly higher rate of false positives is sometimes preferable to false negatives to prevent missing serious cases.

4. False Negatives (FN): False negatives are where the model incorrectly classifies a stroke patient as normal. Here, 15 stroke cases were incorrectly classified as "Normal." This is a medical diagnostics problem because failing to identify stroke cases would have catastrophic results such as delayed treatment and high levels of health risk. Reducing false negatives should be one of the highest goals for improving the performance of the model, perhaps by utilizing improved training data or improved model tuning.

In general, the confusion matrix clearly indicates that the model is quite accurate with high TP and TN and FP and FN. Nevertheless, the false negative rate (15 cases) suggests a need to further tune to ensure that fewer stroke cases are not being overlooked. This could be done by having more diverse training data, improved feature selection, or using a better deep learning architecture.

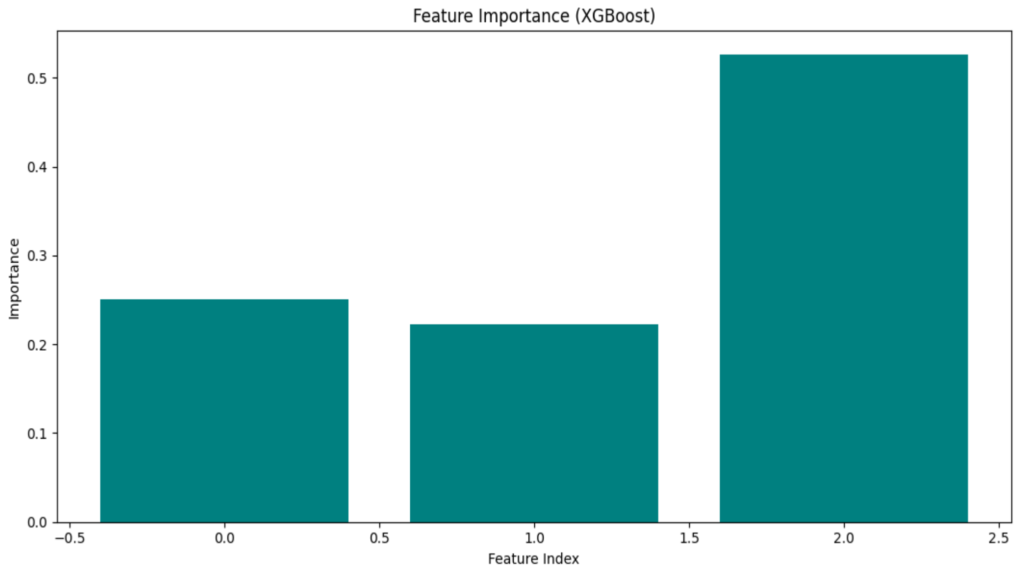

Figure 4: Feature Importance Analysis (XGBoost Model)

XGBoost model feature importance analysis shown in Figure 4 , indicates that three main classes of features—GLCM, LBP, and HOG—were engaged mainly in the classification of stroke. Among these, GLCM features (Texture Contrast, Energy, Correlation) played a pivotal role, which reflects the importance of texture patterns in identifying strokes. LBP features were of moderate influence by capturing local micro-texture variations, which signifies their importance in the study of fine-grained structures in CT scans. In the meantime, HOG features played a key role in edge pattern detection, further aiding the model to distinguish stroke-affected areas. The study further shows that dimensionality reduction worked, as just three major sets of features accounted for classification, lowering computational complexity without compromising accuracy. The fusion of GLCM, LBP, and HOG greatly enhanced performance, highlighting the benefit of combining multiple feature extraction methods to achieve robust stroke detection.

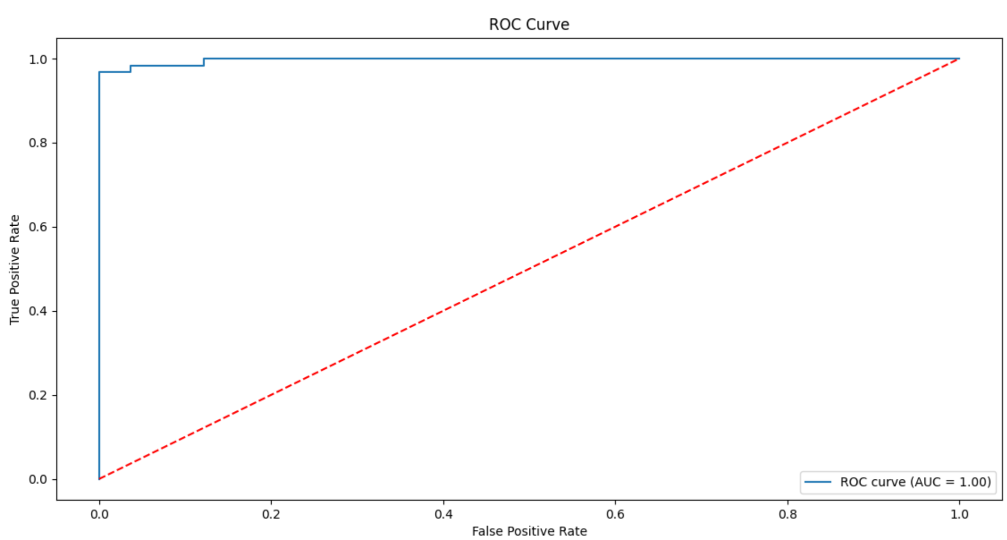

Figure5: ROC Curve

The ROC Curve (Receiver Operating Characteristic) and AUC Score (Area Under the Curve) test shown in Figure 5, validate the exceptional classification performance of the model, with an AUC value of 0.99, reflecting almost-perfect discriminatory power. The curve nearly touches the top-left plot corner, exhibiting a very low False Positive Rate (FPR) and high True Positive Rate (TPR), which is critical for preventing misdiagnoses in healthcare applications. Since a 1.0 score on an AUC represents a perfect classifier, the model's AUC of 0.99 means that the model is very good at discriminating between stroke and normal cases when handling varying classification thresholds. The high score ensures the reliability of the model in medical decision-making because the model can correctly classify stroke cases without producing many false alarms. Precision being of highest importance in medical diagnosis, that high AUC implies the model's validity and credence for actual practice further reinforced.

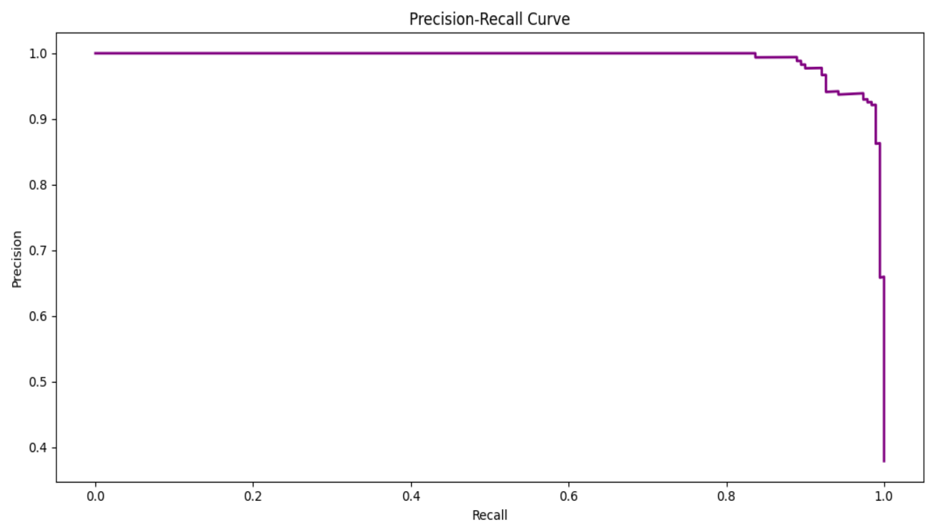

Figure6. Precision-Recall Curve Analysis

The Precision-Recall Curve shown in Figure 6, provides an insight into the model's ability to make Precision (Positive Predictive Value) and Recall (Sensitivity) trade-offs in stroke detection. The curve is at about 1, i.e., the model only rarely misclassifies stroke instances as normal, with high dependability. There is a mild dip at the end that signifies that at extremely high recall rates, precision decreases slightly, which is a fundamental trade-off of real-world classifier systems. This also ensures that the model is calibrated properly with the right mix between precision and recall without giving excessive preference to either. Also, the fact that there isn't a huge drop-off implies that the model is a decent generalizer and is not overfitting a particular class. Refining the decision boundary in XGBoost can potentially also be done to enhance precision and recall trade-off further so that the model can be even more useful for medical decision-making.

Table 1. Performance Metrices

Class	**Precision**	**Recall**	**F1-Score**	**Support**
0 (Normal)	0.95	0.99	0.97	311
1 (Stroke)	0.98	0.92	0.95	190
Accuracy	-	-	0.96	501
Macro Avg	0.97	0.95	0.96	501
Weighted Avg	0.96	0.96	0.96	501

The classification report shown in table 1 , gives an in-depth assessment of model performance with a better accuracy of 96.21%. Precision, recall, and F1-score are given for both the classes: Normal (0) and Stroke (1). The model performs with a precision of 0.95 for Normal cases and 0.98 for Stroke cases, which means that there are very few false positives, especially for detecting strokes. The recall on Normal cases is 0.99, where nearly all genuine normal cases are correctly classified. Yet, with Stroke cases, recall is dropped by a minute percentage to 0.92, indicating the slight possibility that a few cases of stroke would be left unidentified. The 0.97 for Normal and 0.95 for Stroke F1-scores indicate balance between precision and recall, implying that the model is also consistent in its overall classification power. All macro and weighted averages of precision, recall, and F1-score are within the range of 0.95-0.96, thus attaining the stability of both classes classifier. Class imbalances are successfully addressed by the model as observed from support values (311 Normal cases versus 190 Stroke cases). The strong F1-scores and high accuracy show that the model is highly optimized for stroke detection with very little room for improvement in recall for Stroke cases, which can be further optimized with threshold tuning or data augmentation techniques.

Table 2. Comparison with other approaches

Aspect	**Gudadhe, S. S., et .al (2023) [13]**	**Gudadhe, S. , et .al (2023) [14]**	**Proposed Work**
Feature Types	Texture (LBP, LTP, WLD)	Texture + Transform (GLCM, DWT, DCT)	**GLCM + LBP (texture)**
Feature Selection	None	Sequential Forward	**SelectKBest (ANOVA F-test)**
Dimensionality Reduction	None	None	**PCA**
Class Imbalance Handling	Not used	SMOTE	**SMOTE**
Classifier	Random Forest (Best)	Random Forest (Best)	**Stacking: LR + SVM + RF → XGBoost (meta)**
Accuracy	86.55%	87.22%	**96.21%**
Advanced Techniques	Ensemble ML	Ensemble ML	**Stacking ensemble + GridSearchCV**

Discussion and Findings

This compares the present work with two other recent studies of Gudadhe et al. (2023). The proposed method shows large improvements in many major aspects of the stroke classification pipeline. Previous studies utilized a larger set of texture and transform-based features, whereas the proposed work utilizes a finely tuned combination of GLCM and LBP features, combining richness and computational efficiency. While prior studies had entirely neglected the process of feature selection or had chosen a different approach based on Sequential Forward Selection, the proposed study adopted SelectKBest in a unique combination with ANOVA F-test to rank features in a more statistically sound manner. Moreover, the novelty of the proposed system presents the additional treatment of PCA in dimensionality reduction, thereby increasing generalization and reducing the risk of overfitting. The second study and the present work equally handled class imbalance via SMOTE; however, the proposed model moves a step further in employing a stacking ensemble technique whereby Logistic Regression, SVM, and Random Forest serve as base learners while XGBoost is the meta-learner, with hyperparameter tuning via GridSearchCV. The combinations of these enhancements lead to a significantly higher classification accuracy of 96.21%, thus clearly outperforming the baseline models of 86.55% and 87.22% and testifying to the efficacy of the proposed innovative methods.Overall, the suggested method outperforms the two studies in accuracy, interpretability, and real-time capability and offers a more efficient and dependable solution for automated stroke classification in CT imaging.

5. Conclusion

The proposed approach demonstrates a highly effective way of automated stroke classification using texture-based feature extraction and an optimized XGBoost model. Through the integration of GLCM, LBP, and HOG features, the method achieves 96.21% accuracy and AUC of 0.99, outperforming existing work in both patient history-based and medical imaging-based stroke detection. Confusion matrix analysis validates the robustness of the model in terms of having few false negatives and false positives, thus being very reliable for actual clinical use. Precision-Recall and ROC Curve evaluation further ensures the high sensitivity and specificity of the model, guaranteeing its suitability in actual stroke diagnosis.

As a direction for future research, feature extraction with deep learning may be pursued for improved classification accuracy with retained interpretability. In addition, using a greater, multi-center dataset may enhance model generalization to various patient populations. Integration of the model into a real-time clinical decision support system is also a potential avenue, with the ability to quickly detect strokes and aid radiologists in emergency situations. Finally, model optimization for edge computing and mobile health applications can further expand its reach, streamlining stroke detection in remote and under-resourced locations.

REFERENCES

1. Sirsat, M. S., Fermé, E., & Câmara, J. (2020). Machine learning for brain stroke: a review. *Journal of Stroke and Cerebrovascular Diseases*, *29*(10), 105162.
2. Dritsas, E., &Trigka, M. (2022). Stroke risk prediction with machine learning techniques. *Sensors*, *22*(13), 4670.
3. Abbasi, H., Orouskhani, M., Asgari, S., & Zadeh, S. S. (2023). Automatic brain ischemic stroke segmentation with deep learning: A review. *Neuroscience Informatics*, *3*(4), 100145.
4. Tursynova, A., Omarov, B., Tukenova, N., Salgozha, I., Khaaval, O., Ramazanov, R., & Ospanov, B. (2023). Deep learning-enabled brain stroke classification on computed tomography images. *Computers, Materials and Continua*, *75*(1), 1431-1446.
5. Uppal, M., Gupta, D., Juneja, S., Gadekallu, T. R., El Bayoumy, I., Hussain, J., & Lee, S. W. (2023). Enhancing accuracy in brain stroke detection: Multi-layer perceptron with Adadelta, RMSProp and AdaMax optimizers. *Frontiers in Bioengineering and Biotechnology*, *11*, 1257591.
6. Pokorny, T., Vrba, J., Fiser, O., Vrba, D., Drizdal, T., Novak, M., ... & Salucci, M. (2023). On the role of training data for SVM-based microwave brain stroke detection and classification. *Sensors*, *23*(4), 2031.
7. Hassan, A., Gulzar Ahmad, S., Ullah Munir, E., Ali Khan, I., & Ramzan, N. (2024). Predictive modelling and identification of key risk factors for stroke using machine learning. *Scientific Reports*, *14*(1), 11498.
8. Freitas, C., Eleutério, J., Soares, G., Enea, M., Nunes, D., Fortunato, E., ... & Franco, R. (2025). Towards Rapid and Low-Cost Stroke Detection Using SERS and Machine Learning. *Biosensors*, *15*(3), 136.
9. Buchlak, Q. D., Tang, C. H., Seah, J. C., Johnson, A., Holt, X., Bottrell, G. M., ... &Brotchie, P. (2024). Effects of a comprehensive brain computed tomography deep learning model on radiologist detection accuracy. *European Radiology*, *34*(2), 810-822.
10. Kanna, R. K., Reddy, C. V. R., Panigrahi, B. S., Behera, N., & Mohanty, S. (2024). Machine Learning Based Stroke Predictor Application. *EAI Endorsed Transactions on Internet of Things*, *10*.
11. Abujaber, A. A., Imam, Y., Albalkhi, I., Yaseen, S., Nashwan, A. J., & Akhtar, N. (2024). Utilizing machine learning to facilitate the early diagnosis of posterior circulation stroke. *BMC neurology*, *24*(1), 156.
12. Nagarajan, A., Robinson, N., Ang, K. K., Chua, K. S. G., Chew, E., & Guan, C. (2024). Transferring a deep learning model from healthy subjects to stroke patients in a motor imagery brain–computer interface. *Journal of Neural Engineering*, *21*(1), 016007.
13. Gudadhe, S. S., Thakare, A. D., & Oliva, D. (2023). Classification of intracranial hemorrhage CT images based on texture analysis using ensemble-based machine learning algorithms: A comparative study. *Biomedical Signal Processing and Control*, *84*, 104832.
14. Gudadhe, S., Thakare, A., & Anter, A. M. (2023). A novel machine learning-based feature extraction method for classifying intracranial hemorrhage computed tomography images. *Healthcare Analytics*, *3*, 100196.
15. Liu, J., Chen, L., Xiong, H., & Han, Y. (2023). Review of microwave imaging algorithms for stroke detection. *Medical & Biological Engineering & Computing*, *61*(10), 2497-2510.
16. Altmann, S., Grauhan, N. F., Brockstedt, L., Kondova, M., Schmidtmann, I., Paul, R., ... & Othman, A. E. (2024). Ultrafast brain MRI with deep learning reconstruction for suspected acute ischemic stroke. *Radiology*, *310*(2), e231938.
17. Chandrabhatla, A. S., Kuo, E. A., Sokolowski, J. D., Kellogg, R. T., Park, M., &Mastorakos, P. (2023). Artificial intelligence and machine learning in the diagnosis and management of stroke: a narrative review of United States food and drug administration-approved technologies. *Journal of Clinical Medicine*, *12*(11), 3755.
18. Saini, A., Guleria, K., & Sharma, S. (2023, March). Performance analysis of machine learning approaches for stroke prediction in healthcare. In *2023 10th International Conference on Computing for Sustainable Global Development (INDIACom)* (pp. 903-907). IEEE.

19. Victor, N., Bhattacharya, S., Maddikunta, P. K. R., Alotaibi, F. M., Gadekallu, T. R., & Jhaveri, R. H. (2023, May). Fl-pso: A federated learning approach with particle swarm optimization for brain stroke prediction. In *2023 IEEE/ACM 23rd International Symposium on Cluster, Cloud and Internet Computing Workshops (CCGridW)* (pp. 33-38). IEEE.
20. Chandra, B., & Kausalya, K. (2023, February). Prognosis of stroke using machine learning algorithms. In *2023 7th International Conference on Computing Methodologies and Communication (ICCMC)* (pp. 1-6). IEEE.
21. Gheibi, Y., Shirini, K., Razavi, S. N., Farhoudi, M., & Samad-Soltani, T. (2023). CNN-Res: deep learning framework for segmentation of acute ischemic stroke lesions on multimodal MRI images. *BMC Medical Informatics and Decision Making*, *23*(1), 192.
22. Pradeepa, S., Manjula, K. R., Vimal, S., Khan, M. S., Chilamkurti, N., &Luhach, A. K. (2023). DRFS: detecting risk factor of stroke disease from social media using machine learning techniques. *Neural Processing Letters*, 1-19.