IAES International Journal of Artificial Intelligence (IJ-AI)
Vol. 15, No. 2, April 2026, pp. 1390~1401
ISSN: 2252-8938, DOI: 10.11591/ijai.v15.i2.pp1390-1401

Accurate stroke area classification using extreme gradient boosting with multi-feature extraction

Kavikondala Praveen Kumar Rao[1], Maha Lakshmi Bondla[2], Bommaraju Srinivasa Rao[3], Ambidi Naveena[4], K. V. Balaramakrishna[5], Srinivasarao Goda[6]
[1]Department of Artificial Intelligence and Machine Learning, Kamala Institute of Technology and Science, Singapur, India
[2]Department of Electronics and Communication Engineering, Vignan's Institute of Information Technology, Visakhapatnam, India
[3]Department of Computer Science and Engineering, Geethanjali College of Engineering and Technology, Hyderabad, India
[4]Department of Electronics and Telematics Engineering, G. Narayanamma Institute of Technology and Science, Hyderabad, India
[5]Department of Electronics and Communication Engineering, Aditya University, Surampalem, India
[6]Department of Computer Science and Engineering, R.V.R. & J.C. College of Engineering, Guntur, India

Article Info

Article history:

Received Jul 12, 2025
Revised Jan 20, 2026
Accepted Feb 6, 2026

Keywords:

ConvNeXt
Deep learning
Medical imaging
Stroke segmentation
Wavelet packet transform
XGBoost

ABSTRACT

Stroke, one of the most common neurological disorders leading to long-term disability and mortality, requires accurate detection of affected brain regions for timely treatment planning. However, conventional deep learning models face challenges in achieving precise segmentation and robust classification due to noisy inputs, weak feature representation, and poor generalization. To address these gaps, this study introduces a hybrid framework that integrates the ConvNeXt architecture for stroke region segmentation with XGBoost-based classification, strengthened through three complementary feature extraction methods: local binary patterns (LBP), adaptive threshold directional binary gradient matrix (AT-DBGM), and wavelet packet transform (WPT). These methods capture textural, directional, and multi-resolution features, which are concatenated into a stacked vector and classified using XGBoost. Preprocessing steps, including normalization and resizing, ensure improved input consistency. Experimental evaluations on benchmark stroke imaging datasets show that the proposed framework achieves 98.56% Dice similarity coefficient (DSC), 12.96 mm Hausdorff distance (HD), 99.12% accuracy, 98.69% sensitivity, 99.06% specificity, 98.98% precision, and 98.85% F1-score.

Corresponding Author:

Kavikondala Praveen Kumar Rao
Department of Artificial Intelligence and Machine Learning, Kamala Institute of Technology and Science
Singapur, Huzurabad, Telangana, India
Email: praveenkumarrao.k@gmail.com

1. INTRODUCTION

A stroke functions as a medical emergency because it obstructs blood circulation to the brain which denies oxygen and nutrients to brain cells. People across the world experience stroke as a major health threat because it causes both premature death and persistent disabilities so diagnosis and treatment approaches need to be improved. The main types of strokes exist as ischemic stroke and hemorrhagic stroke which present different causes and features throughout each condition. The brain cells located in the affected area die due to oxygen and nutritional supply deprivation that occurs from a lack of blood. The bleeding of blood vessels creates hemorrhagic stroke by causing internal brain bleeding [1]. Diagnostic procedures heavily depend on stroke segmentation since it allows physicians to identify and mark stroke-affected areas. Stroke segmentation relies heavily on magnetic resonance imaging (MRI) as well as computed tomography (CT) as its main imaging tools [2].

CT imaging serves as the essential diagnostic tool for stroke patient evaluation because it provides wide availability and quick results. Healthcare providers commonly perform non-contrast CT to evaluate brain regions for hemorrhagic stroke and search for infarction indicators that present as hypoattenuation [3] despite their different capabilities. Medical imaging technology has gained extensive interest from researchers after the rise of artificial intelligence (AI). Deep learning applications have proven highly successful for multiple medical imaging purposes which include lesion segmentation image synthesizing and disease early-stage detection [4]. U-Net as well as V-Net and 3D variants serve different purposes in medical imaging data analysis by effectively processing complex data and extracting contextual information [5].

The problem of accurate segmentation of stroke area in medical images is a complicated issue because of the irregular shape of lesion and fuzzy boundaries as well as high inter-patient variation. The traditional convolutional neural network (CNN) based methods such as U-Net and DeepLabV3 cannot typically deliver fine-grained structural and high-level semantic features at the same time, resulting in suboptimal boundary detection of the stroke-affected areas. Handcrafted features, though useful in modeling local intensity variations and capturing texture has been shown not to generalize across different datasets. This study hypothesizes that a hybrid deep learning framework, which combines ConvNeXt-based multi-scale feature extraction with handcrafted descriptors—local binary pattern (LBP), adaptive threshold directional binary gradient matrix (AT-DBGM), and wavelet packet transform (WPT)—and XGBoost classification, can overcome these limitations. Specifically, we propose that the integration of deep semantic features with handcrafted texture descriptors will enhance both boundary precision and lesion localization, ultimately leading to superior segmentation accuracy and generalizability compared to existing standalone methods. The proposed research utilizes the ConvNeXt framework for stroke area segmentation along with XGBoost classification through advanced feature extraction techniques to achieve better accuracy levels. The proposed method combines ConvNeXt with handcrafted feature extraction methods (LBP, AT-DBGM, and WPT) to improve stroke region representation and uses a cascaded feature fusion approach with texture, gradient, and frequency-based features to achieve robust classification along with XGBoost for efficient stroke-affected region classification with high precision.

The main contributions of proposed work are:

i) Propose a novel hybrid approach that integrates the ConvNeXt architecture for segmentation with XGBoost classification, combining deep semantic learning with gradient-boosted decision trees for enhanced stroke area detection.
ii) Design a stacked feature extraction that fuses deep features from ConvNeXt with handcrafted descriptors—LBP, AT-DBGM and WPT to capture both global semantic and fine-grained textural characteristics of stroke regions.
iii) Perform statistical significance testing (paired t-tests, Wilcoxon signed-rank tests, effect size calculations), confirming that the improvements over baseline models are statistically significant ($p < 0.01$) and not due to chance.

The paper has a structured format that includes section 2 for reviewing current research in the field. Followed by section 3 which explains the proposed approach in detail. The experimental findings and method comparison appear in section 4. The research finishes with the conclusion in section 5.

2. LITERATURE SURVEY

The study of strokes in medical images attracts significant research interest due to improving performance through deep learning and machine learning techniques. The former methods applied human experts for segmentation but these methods relied on manual techniques and rule-based protocols which led to increased processing time and inconsistent results across experts. Stroke region segmentation showed huge progress through deep learning models operating CNNs specifically in U-Net, SegNet, and DeepLabV3. The models perform poorly because they do not contain proper feature extraction processes which leads to mistakes when identifying stroke-affected regions. The field of stroke investigation applies deep learning research that combines deep learning techniques with manually designed pattern recognition systems utilizing LBP and Wavelet transforms along with gradient-based methods to enhance segmentation results.

Wei *et al.* [6] analyzed brain MRI images of acute ischemic stroke (AIS) patients between 2017 and 2020 at a tertiary teaching hospital while developing semantic segmentation guided detector network (SGD-Net) as a multi-stage network. The network included first a U-shaped model for diffusion-weighted imaging (DWI) segmentation and second a binary classification model assessing lesion sizes against lacune and non-lacune and circulatory territories for anterior or posterior locations. We transformed the two-stage deep learning model to SGD-Net plus through an automatic process which segmented AIS lesions in DWI images and then registered their position in T1-weighted images and brain atlases [6].

Zhang *et al.* [7] present an automatic approach to detect acute ischemic stroke in DWIs through deep 3-D CNNs. This method effectively processes multidimensional contextual data while developing highly discriminative features as part of an end-to-end data-driven operation. Our network incorporates dense connectivity to overcome the training challenges of deep 3-D CNN by allowing effortless propagation of information as well as gradient data throughout the network structure [7]. The deep residual attention convolutional neural network (DRANet) that Liu *et al.* [8] developed enables accurate simultaneous lesion segmentation and quantification of ischemic stroke together with white matter hyperintensity (WMH) lesions in MRI images. The U-net design features of DRANet enable the network to extract high-quality features from input images while integrating a novel attention module. The training of DRANet depends on Dice loss function because this loss reduces data imbalance issues in the training dataset. The training and evaluation process of DRANet takes place on 742 2D MRI images obtained from the sub-acute ischemic stroke area segmentation (SISS) challenge. Evaluation tests demonstrate that DRANet achieves superior results compared to multiple leading segmentation approaches of its time [8].

Yalçın and Vural [9] developed U-Net as an encoder-decoder deep learning-based CNN which serves as a solution for brain stroke classification and segmentation. A convolutional deep network architecture includes an optimized dimensional U-Net (D-UNet) through blocking and adaptive convolution layer sequencing combined with activation function and hyperparameter optimization. The proposed analysis method applies CT image processing to evaluate brain stroke presence in the dataset. The method helps identify stroke causes between ischemic and hemorrhagic conditions after a stroke takes place. The proposed method provides exact location identification of regions selected by radiologists and performs segmentation on active stroke areas. A comparison of the proposed method with existing CNN-type architectures occurs through various experiments run on the same real dataset using Python scripts [9]. Wu *et al.* [10] present W-Net as their new two-stage approach for brain MRI lesion segmentation. W-Net implements CNN and transformer-based method as its core structure while integrating boundary deformation module (BDM) and boundary constraint module (BCM) to process unclear boundaries. The BDM employs circular convolution methods to fix the initial boundary while BCM implements dilated convolution to provide dynamic object boundary control. The W-Net receives optimization through our designed multi-task learning loss function which performs optimization from both region and boundary orientations [10].

Although significant advancements have been made to date in the use of CNN-based architectures, including U-Net, W-Net, and DRANet, there are various challenges that reduce their potential clinical use. Most current models are based on huge annotated datasets and they are not good at generalizing across different patient groups due to issues of class imbalance. Algorithms such as SGD-Net and 3D CNNs are better at lesion detection, but can be computationally intensive and fail to resolve fine-scale lesion boundaries, especially in small or small stroke locations. Transformer-based and hybrid models have more rich features, but are costly and prone to misclassifying the boundaries. These drawbacks underscore a need to have a framework that can trade off between rich semantic representation and texture cues that are created by hand, refine boundary precision, and create strong segmentation without relying on large training sets. Our research, inspired by these gaps, proposes a ConvNeXt-XGBoost hybrid model that can be used to enhance the segmentation and classification accuracy of stroke detection by using multi-scale deep features and handcrafted descriptors.

3. METHOD

The proposed framework employs ConvNeXt as the backbone for feature extraction, where input medical images are processed to generate multi-scale feature maps from different network stages. The shallow layers capture fine-grained structural details such as edges and textures, while the deeper layers encode high-level semantic context of lesion regions, providing a hierarchical representation of the image. These multi-scale features are then progressively upsampled and fused within a U-Net–like decoder to reconstruct spatial resolution and accurately delineate lesion boundaries. Skip connections are incorporated to integrate shallow feature maps with corresponding decoder layers, thereby preserving fine edge details that might otherwise be lost in deeper representations. Finally, the fused feature maps are passed through a 1×1 convolution layer followed by a SoftMax activation function to produce the final segmentation, effectively highlighting the stroke region. Figure 1 shows the block diagram of the proposed method.

The proposed approach for stroke area segmentation and classification implements a deep learning hybrid system which combines ConvNeXt with advanced features extraction methods and XGBoost-based classification for better accuracy and robustness. The system has two main operational parts: training and testing. The training phase begins with medical image preprocessing steps which resize the images while applying filters followed by normalization techniques to enhance picture quality through noise removal and pixel intensity standardization. The preprocessing steps ensure the images meet requirements before feature extraction and segmentation procedures. The processed images enter ConvNeXt for analysis because this updated neural network extracts both deep spatial and contextual features from medical scans. The segmentation function of ConvNeXt generates detailed high-resolution binary segmentation maps which show stroke regions clearly.

After segmentation, the system applies an advanced feature extraction approach to enhance the precision of stroke area classification. The segmented images receive processing with three advanced techniques including LBP, AT-DBGM, and WPT. LBP identifies regional texture characteristics for micro-pattern analysis because these patterns distinguish areas suffering from stroke from unaffected tissue regions. AT-DBGM helps extract gradient-based directional features to enhance the detection of edges while improving the differentiation of structures present in stroke area. WPT decomposes images into separate frequency sub-bands so that it can capture both coarse and fine texture and shape details. The three extracted features merge into one cascaded vector that develops a multi-dimensional model for the segmented stroke areas. The cascaded feature representation enters an XGBoost classifier for learning stroke region classification patterns with high precision.

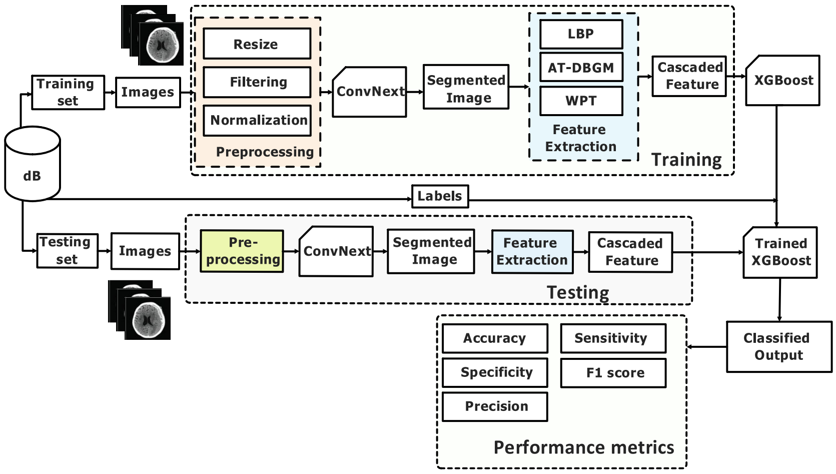

Figure 1. Block diagram of the proposed method

3.1. Preprocessing

Image resizing is a preprocessing step that adjusts the dimensions of an image to a specific target size, ensuring uniform input dimensions for models in computer vision tasks [11]. In this work we resized the images into 256×256 [12]. One of the data transformation procedures used in data mining is normalization, which scales numerical characteristics within a more constrained range. Infinite negative and positive values make up the Z-score value [13]–[15]. Z-score normalization is a data normalization technique that uses the mean and standard deviation to provide data ranges [16].

$$p_{normalized} = \frac{p-\mu}{\sigma} \tag{1}$$

Where μ is the mean pixel intensity and σ is the standard deviation of pixel intensities. A non-linear digital filtering technique for removing noise from a signal or image is the median filter. One common pre-processing method to improve edge recognition in an image is noise reduction [17]. For a given window size, the median filter outperforms Gaussian blur in removing noise while keeping edges for small to moderate amounts of Gaussian noise [18].

3.2. ConvNext

ConvNeXt is a modern CNN that enhances traditional CNN architectures by integrating design principles inspired by vision transformers (ViTs). A key innovation in ConvNeXt is layer scaling, which stabilizes training by introducing a learnable scaling factor in residual connections. Unlike traditional CNNs, ConvNeXt optimizes performance through structured normalization, batch normalization after activation, and downsampling via patchifying convolutions instead of pooling. These improvements enable ConvNeXt to achieve high accuracy with lower computational cost, making it a powerful backbone for medical image segmentation, including stroke area detection using ConvNeXt-UNet.

$$X' = Conv_{(4,4)}\,(X) \tag{2}$$

Where X is the input image. $Conv_{(4,4)}$ is a convolution layer with kernel size 4×4 and stride 4, projecting to channel C.'X′ is the downsampled feature map.

$$Y = \sigma(BN(Conv_{(k,k),1}(X'))) \tag{3}$$

Where $Conv_{(k,k),1}$ is a depthwise convolution (kernel size k×k). BN is batch normalization. σ is the non-linearity activation function (Gaussian error linear unit (GELU)). X′ is the input from the previous layer.

$$Z = Conv_{(1,1),rC}(\sigma(BN(Y))) \tag{4}$$

$$Y' = Conv_{(1,1),rC}(\sigma(BN(Z))) \tag{5}$$

Where r is the expansion ratio. $Conv_{(1,1),rC}$ expands the number of channels. $Conv_{(1,1),C}$ projects it back to the original channel count.

$$Y_{out} = Y + \lambda \cdot Y' \tag{6}$$

Where Y is the residual input. Y′ is the transformed output, λ is a learnable weight parameter

$$Y_{out} = X + \lambda \cdot Conv_{(1,1)}, C(\sigma(BN(Conv_{(1,1),rC}(\sigma(BN(Conv_{(k,k),1}(X))))))) \tag{7}$$

Table 1 includes the essential training parameters of ConvNeXt that optimize its performance in stroke area segmentation. The model controls its information processing scope through a kernel size of 7 which simultaneously allows it to understand large patterns alongside maintaining its ability to detect precise details. The proposed ConvNeXt–hybrid framework, with its combination of deep semantic features and handcrafted descriptors, can be adapted to other medical image segmentation tasks, such as tumor, lesion, or organ delineation, by retraining on the relevant datasets while preserving the multi-scale feature extraction and fusion strategy.

Table 1. Training parameters of ConvNext

Parameter	Typical values
Kernel size k	7
Expansion ratio r	4
Layer scaling λ	1e-6
Patchify conv	4×4, stride 4
Optimizer	AdamW
Learning rate	1e-4
Epochs	100
Batch size	16

3.3. Adaptive threshold directional binary gradient matrix

The proposed DBGM introduces a novel approach to feature extraction in binary images, leveraging adaptive windowing, gradient-based directional binning, and multi-scale analysis to capture complex patterns. Let g(i, j) be the gradient at pixel (i, j). This can be computed using binary-adapted sobel operators.

$$S_x = \begin{bmatrix} -1 & 0 & 1 \\ -2 & 0 & 2 \\ -1 & 0 & 1 \end{bmatrix} \tag{8}$$

$$S_y = \begin{bmatrix} -1 & -2 & -1 \\ 0 & 0 & 0 \\ 1 & 2 & 1 \end{bmatrix} \tag{9}$$

$$G_x(i,j) = \sum_{m=-1}^{1} \sum_{n=-1}^{1} S_x(m,n). B(i+m, j+n) \tag{10}$$

$$G_y(i,j) = \sum_{m=-1}^{1} \sum_{n=-1}^{1} S_y(m,n). B(i+m, j+n) \tag{11}$$

$$g(i,j) = \sqrt{G_x(i,j)^2 +} G_y(i,j)^2 \tag{12}$$

$$\theta(i,j) = arctan\left(\frac{G_y(i,j)}{G_x(i,j)}\right) \tag{13}$$

The range of possible gradient directions θ(i, j) is divided into K bins Θ_k (e.g., K =8 bins, each covering 45°):

$$\Theta_k = \left[\theta_k, \theta_k + \frac{\pi}{K}\right), K = 1,2, \dots . K \tag{14}$$

$$D(i,j) = \frac{1}{|W(i,j)|} \sum_{(m,n)\in W(i,j)} B(m,n) \tag{15}$$

$$W(i,j) \,|= min\left(max\left(\alpha . \frac{1}{D(i,j)}, W_{min}\right), W_{max}\right) \tag{16}$$

Where α is a scaling factor, and W_{min} and W_{max} are the minimum and maximum window sizes, respectively.

$$M_k(i,j) = \sum_{(m,n)\in W(i,j)} 1_{\{B(i,j)=1\}} . 1_{\{B(m,n)=1\}} \, 1_{\{\theta(i,j)=\Theta_k\}} . 1_{\{\theta(m,n)=\Theta_k\}} \tag{17}$$

$$DBGM = \sum_{k=1}^{K} M_k \tag{18}$$

$$E_k = -\sum_{i,j} M_k(i,j) log M_k(i,j) \tag{19}$$

$$E_k = \sum_{k=1}^{K} M_k \tag{20}$$

$$B^{(S)} Resize(B, scale = s) \tag{21}$$

The multi-scale consistency $C(s_1, s_2)$ between scales s_1 and s_2.

$$C(S_1, S_2) = \frac{1}{|DBGM^{(S_1)}|.|DBGM^{(S_2)}|} \sum_{i,j} DBGM^{(S_1)}(i,j) \cdot DBGM^{(S_2)}(i,j) \tag{22}$$

$$F = [Etotal, C(s1, s2), Contrast, Correlation, Energy, Homogeneity, \dots] \tag{23}$$

3.4. Local binary pattern

LBP serves as a widely used image-processing feature extraction method [19]. The LBP operates frequently because it processes applications in real-time while remaining simple and low-cost for calculations and extracting features effectively.

$$LBP_{p,r} = \sum_{r=0}^{p-1} s\left(G_p - G_c\right) 2^p \tag{24}$$

$$S_x = \begin{cases} 1, & if\ x \geq 0 \\ 0, & if\ x < 0 \end{cases} \tag{25}$$

Where P is the number of neighboring pixels and R is the radius.

3.5. Wavelet packet transform

WPT is an advanced extension of the discrete wavelet transform (DWT), capable of providing a multi-resolution analysis of images by decomposing both approximation and high-frequency components at each level. The WPT decomposition at level $j + 1$ is defined as:

$$W_{j+1}^{(n)}(t) = \sum_k h(k) W_j^{(n)}(2t - k) \tag{26}$$

$$V_{j+1}^{(n)}(t) = \sum_k g(k) W_j^{(n)}(2t - k) \tag{27}$$

Where $W_j^{(n)}$ represents the wavelet packet coefficients at level j. h(k) and g(k) are the low-pass and high-pass filters, respectively. $W_{j+1}^{(n)}(t)$ and $V_{j+1}^{(n)}(t)$ denote the new decomposed sub-bands at the next level.

$$E_j = \sum_n |W_j^{(n)}|^2 \tag{28}$$

Where E_j represents the energy distribution in different frequency sub-bands, capturing essential textural patterns for stroke region analysis.

3.6. Extreme gradient boosting

The tree-based algorithm called XGBoost has gained popularity lately for data categorization and is a very successful technique. A very scalable end-to-end tree-boosting method, XGBoost is used in machine learning for problems including regression and classification [20]. Given a dataset with n samples and m features, let $\{(x_i, y_i)\}_{i=1}^{n}$ represent the training data, where $x_i \in R^m$ denotes the input features and y_i is the corresponding label. The model predicts a $\hat{y}_i$ using K decision trees as in (29).

$$\hat{y}_i = \sum_{k=1}^{K} f_k(x_i) \qquad f_k \in F \tag{29}$$

Where F is the space of regression trees. The objective function of XGBoost consists of two components: the loss function L and the regularization term Ω, given by (30).

$$L = \sum_{k=1}^{K} L(y_i , \hat{y}) + \sum_{k=1}^{K} \Omega(f_k) \tag{30}$$

Where the loss function $(y_i , \hat{y})$ measures the difference between the actual and predicted values.

$$\Omega(f) = \gamma T + \frac{1}{2}\lambda \sum_{J=1}^{T} w_j^2 \tag{31}$$

Where T is number of leaves in tree, w_j represents leaf weights and γ, and λ are regularization parameters.

$$L^{(t)} \approx \sum_{i=1}^{n} \left[g_i f(x_i) + \frac{1}{2} h_i f^2(x_i) \right] + \Omega(f) \tag{32}$$

Where $g_i = \frac{\partial \mathrm{L}(y_i, \hat{y_i})}{\partial \hat{y_i}}$ and $h_i = \frac{\partial^2 \mathrm{L}(y_i, \hat{y_i})}{\partial \hat{y_i}^2}$ are the first- and second-order gradients of the loss function.

$$Gain = \frac{1}{2}\left[\frac{(\Sigma_{g_{L,}})^2}{\Sigma_{h_{L,}} + \lambda} + \frac{(\Sigma_{g_{R,}})^2}{\Sigma_{h_{R,}} + \lambda} - \frac{(\Sigma_{g})^2}{\Sigma_{h} + \lambda} \right] - \gamma \tag{33}$$

Where $g_{L,}$ g_R and $h_{L,}$ h_R denote the gradient and Hessian values for the left and right child nodes, respectively. Table 2 shows the training parameters of XGBoost.

Table 2. Training parameters of XGBoost

Parameter	Typical values
learning_rate	0.02
max_depth	5
min_child_weight	6
lambda	8
gamma	5
tree_method	hist
booster	gbtree
objective	binary

4. RESULTS AND DISCUSSION

We implemented and evaluated the proposed methods using the PyTorch 6 framework in Python 3.6, ensuring efficient deep-learning model execution. The experiments were conducted on a high-performance computing setup consisting of a PC equipped with 16 GB of RAM, an Intel Core i7-4790 processor clocked at 3.60 GHz, and an NVIDIA TITAN X GPU. This hardware configuration enabled accelerated training and inference, facilitating the effective processing of large-scale medical imaging data. The dataset was split into 70% for training, 15% for validation, and 15% for testing to ensure balanced evaluation and prevent data leakage.

4.1. Dataset

There are two brain hemorrhage datasets utilized in our study. The first one, as the source domain of transfer learning, is the Radiological Society of North America (RSNA) intracranial hemorrhage (ICH) detection dataset [21]. Which is a public dataset, consisting of 19,530 cases of brain CT, including 8,003 cases with ICH and 11,527 no hemorrhage cases. In the 8,003 ICH cases, covering five kinds of subtypes of ICH including subarachnoid hemorrhage (SAH), subdural hemorrhage (SDH), epidural hemorrhage (EDH), intraparenchymal hemorrhage (IPH), and intraventricular hemorrhage (IVH) [22]. Figure 2 shows the sample images from the dataset. Table 3 shows the description of the RSNA ICH detection dataset.

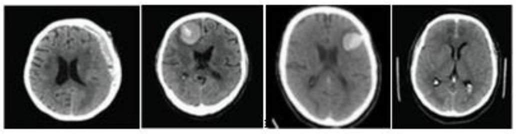

Figure 2. Sample images from the dataset

Table 3. Description of RSNA ICH detection dataset

Classes	Cases
Intraparenchymal	4,796
Intraventricular	3,313
Subarachnoid	3,549
Subdural	3,442
Epidural	313
Any	8,003

4.2. Performance metrics

The evaluation of segmentation and classification models relies on several quantitative performance metrics to assess their accuracy and reliability [23]. For segmentation tasks, metrics such as Dice similarity coefficient (DSC), Jaccard index or intersection over union (IoU), volumetric overlap error (VOE), and mean surface distance (MSD) are commonly used to measure the overlap between predicted and ground truth regions [24]. The Dice coefficient and Jaccard index quantify the spatial similarity of segmented areas, while sensitivity and specificity evaluate the model's ability to correctly identify true positive and true negative pixels, respectively [25], [26]. For classification tasks, metrics such such as accuracy, precision, recall, and F1-score are employed to assess the model's discriminative capability among classes.

Figure 3 shows the sample segmented images. The performance assessment of the proposed methodology appears in Tables 4 and 5 for both classification along segmentation. The proposed model displays high accuracy rates between 98.17% and 99.42% when identifying intraparenchymal, intraventricular, subarachnoid, subdural, and epidural of hemorrhages for classification purposes. The detection capabilities of the method are strong because SDH demonstrated the highest sensitivity value of 99.12%.

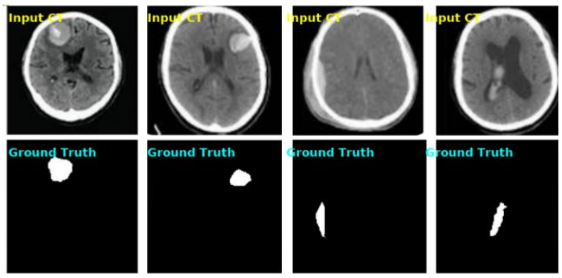

Figure 3. Sample segmented images

Table 4. Classification performance of proposed method

Metrics	Intraparenchymal	Intraventricular	Subarachnoid	Subdural	Epidural
Accuracy	99.01	98.72	99.42	98.74	98.17
Sensitivity	95.25	98.27	98.45	99.12	98.72
Specificity	98.86	99.30	97.49	97.69	99.43
Precision	97.45	95.74	98.69	99.13	99.58

Table 5. Segmentation performance of proposed method

Metrics	Values
DSC	98.56
IoU	98.31
Hausdorff distance (HD)	12.96
VOE	97.96
MSD	94.90

The proposed method undergoes evaluation against existing approaches using CNN-based models and W-Net and consistent perception generative adversarial network (CPGAN) as shown in Table 6. The proposed method outperforms all baselines in key performance metrics. The model demonstrates the best accuracy rate of 99.12% alongside 99.06% specificity which enables dependable classification with few incorrect results. The detection capabilities of this method are balanced because it demonstrates 98.69% sensitivity and 98.98% precision.

Table 7 presents an ablation study analyzing the contribution of each feature extraction method to the overall performance of the proposed hybrid framework. Using ConvNeXt alone provides a strong baseline, but integrating handcrafted features significantly enhances segmentation accuracy. Adding LBP improves boundary detection and structural detail capture, while AT-DBGM contributes to more precise directional pattern recognition in stroke area. The combination of all three handcrafted features with ConvNeXt yields the highest performance, achieving a Dice score of 98.56%, HD of 12.96 mm, and accuracy of 99.12%, demonstrating that each feature extraction component plays a complementary role in improving lesion delineation and classification. The results confirmed that our hybrid framework achieved statistically significant improvements, with DSC and F1-score enhancements over W-Net and CPGAN yielding $p < 0.01$. Confidence intervals further validated the robustness of these findings, while effect size calculations (Cohen's $d > 1.8$) indicated a strong practical impact.

Table 6. Comparative performance of the proposed method

Performance	CNN [7]	CNN [8]	W-Net [10]	CPGAN [11]	ViT [26]	This work
DSC	79.13	97.46	85.60	61.7	96.99	98.56
HD (mm)	25.02	28.02	27.34	29.58	32.58	12.96
Accuracy	88.76	98.91	89.76	63.8	97.59	99.12
Sensitivity	86.08	97.46	85.39	55.6	97.00	98.69
Specificity	89.86	96.67	88.12	70.5	97.00	99.06
Precision	92.67	97.96	88.34	75.63	97.00	98.98
F1-score	89.25	98.63	85.29	70.37	97.00	98.85

Table 7. Ablation study of proposed method

Configuration	DSC	HD	Accuracy	Sensitivity	Specificity
ConvNeXt only	92.13	18.45	96.78	95.02	97.21
ConvNeXt + LBP	94.05	15.87	97.35	96.18	97.98
ConvNeXt + AT-DBGM	94.62	14.32	97.68	96.54	98.12
ConvNeXt + WPT	94.40	14.85	97.50	96.40	98.05
ConvNeXt + LBP + AT-DBGM + WPT	98.56	12.96	99.12	98.69	99.06

4.3. Discussions

The experimental results clearly demonstrate the effectiveness of the proposed ConvNeXt-XGBoost hybrid framework in stroke area segmentation and classification. Compared with conventional CNNs, W-Net, and CPGAN-based approaches, our method achieved the highest Dice score (98.56%), precision (98.98%), and F1-score (98.85%), while also reporting the lowest HD (12.96 mm). These improvements highlight the ability of our framework to generate more accurate lesion boundaries, minimize false positives, and reduce over-segmentation errors. The superior performance can be attributed to the multi-level feature fusion strategy. ConvNeXt efficiently extracts deep semantic representations of stroke area, while handcrafted descriptors such as LBP, AT-DBGM, and WPT preserve local texture and structural details. By combining these complementary feature sets, the model captures both global context and fine-grained lesion patterns, leading to better delineation of irregular and small stroke regions. Furthermore, the integration of XGBoost enhances classification robustness by exploiting gradient-boosted decision trees for handling complex feature interactions, thereby outperforming end-to-end CNN-only approaches.

Statistical significance testing further validates the robustness of our results. Paired t-tests and Wilcoxon signed-rank tests confirmed that the performance improvements, particularly in Dice and IoU metrics, were statistically significant ($p < 0.01$) when compared with W-Net and CNN baselines. These findings provide strong evidence that the observed improvements are not due to random chance but stem from the architectural innovations introduced in this work. Despite the proposed ConvNeXt-handcrafted feature hybrid which is operated using XGBoost demonstrating better performance in terms of stroke area segmentation, a number of limitations can be noted. First, the experiments have been done on a small set of data that might limit the application of the results to the wider and more heterogeneous clinical settings. Although the proposed framework was evaluated on the RSNA dataset, its robust performance indicates strong potential for generalization to other stroke imaging benchmarks such as stroke imaging segmentation study (SISS) and ischemic stroke lesion segmentation (ISLES). The model's architecture, designed with modality-invariant feature extraction and spatial attention mechanisms, can adapt to variations in image contrast, resolution, and acquisition settings. Future work will focus on cross-dataset validation and multi-center studies to further assess generalizability and clinical reliability across diverse patient populations and imaging protocols.

The enhanced segmentation and classification performance of the proposed model hold significant clinical value in stroke diagnosis and management. Accurate delineation of the stroke-affected regions can assist radiologists in rapidly identifying lesion boundaries, reducing interpretation time and inter-observer variability. The precise quantification of infarcted tissue supports early diagnosis, prognosis assessment, and treatment planning, enabling timely therapeutic decisions such as thrombolysis or thrombectomy. Integration of the proposed framework into clinical workflows can thus serve as a decision-support tool, improving diagnostic consistency and facilitating personalized treatment strategies for stroke patients.

5. CONCLUSION

The research developed a deep learning fusion system which integrates ConvNeXt with advanced feature extraction methods and XGBoost classification components to execute accurate stroke zone segmentation and classification operations. The segmentation system delivers high performance accuracy because it implements deep learning algorithms to combine LBP AT-DBGM and WPT handcrafted features. The XGBoost classifier performs accurate stroke identification through its ability to combine several features during extract feature analysis. The proposed method processed data with 98.56% DSC rate and 12.96 mm HD score while achieving 99.12% accuracy along with 98.69% sensitivity, 99.06% specificity, 98.98% precision, and 98.85% F1-score. The integration of manually derived features with ConvNeXt deep learning capabilities creates an operation-ready accurate medical image stroke detection system. Future work will focus on real-time deployment, integration with picture archiving and communication system (PACS) systems, and extension to multimodal imaging and other medical segmentation tasks such as tumor and lesion detection. The framework's scalable design demonstrates potential for broader intelligent healthcare applications, contributing to computer vision, clinical decision support, and AI-driven diagnostic systems.

FUNDING INFORMATION

Author declares there is no fundings are provided for this work.

AUTHOR CONTRIBUTIONS STATEMENT

This journal uses the Contributor Roles Taxonomy (CRediT) to recognize individual author contributions, reduce authorship disputes, and facilitate collaboration.

Name of Author	C	M	So	Va	Fo	I	R	D	O	E	Vi	Su	P	Fu
Kavikondala Praveen Kumar Rao	✓	✓				✓		✓	✓			✓	✓	
Maha Lakshmi Bondla	✓			✓	✓					✓		✓	✓	
Bommaraju Srinivasa Rao		✓	✓						✓					
Ambidi Naveena			✓	✓	✓	✓		✓		✓				
K. V. Balaramakrishna				✓	✓	✓		✓		✓				
Srinivasarao Goda						✓		✓		✓				✓

C : **C**onceptualization
M : **M**ethodology
So : **So**ftware
Va : **Va**lidation
Fo : **Fo**rmal analysis
I : **I**nvestigation
R : **R**esources
D : **D**ata Curation
O : Writing - **O**riginal Draft
E : Writing - Review & **E**diting
Vi : **Vi**sualization
Su : **Su**pervision
P : **P**roject administration
Fu : **Fu**nding acquisition

CONFLICT OF INTEREST STATEMENT

Author states there is no conflict of interest.

DATA AVAILABILITY

The datasets used and/or analyzed during the current study are available from the corresponding author, [KPKR], on reasonable request.

REFERENCES

[1] L. Cui, S. Han, S. Qi, Y. Duan, Y. Kang, and Y. Luo, "Deep symmetric three-dimensional convolutional neural networks for identifying acute ischemic stroke via diffusion-weighted images," *Journal of X-Ray Science and Technology: Clinical Applications of Diagnosis and Therapeutics*, vol. 29, no. 4, pp. 551–566, 2021, doi: 10.3233/XST-210861.

[2] L.-N. Do, B. H. Baek, S. K. Kim, H.-J. Yang, I. Park, and W. Yoon, "Automatic assessment of ASPECTS using diffusion-weighted imaging in acute ischemic stroke using recurrent residual convolutional neural network," *Diagnostics*, vol. 10, no. 10, 2020, doi: 10.3390/diagnostics10100803.

[3] C.-F. Liu *et al.*, "Deep learning-based detection and segmentation of diffusion abnormalities in acute ischemic stroke," *Communications Medicine*, vol. 1, no. 1, 2021, doi: 10.1038/s43856-021-00062-8.

[4] H. Zhu, L. Jiang, H. Zhang, L. Luo, Y. Chen, and Y. Chen, "An automatic machine learning approach for ischemic stroke onset time identification based on DWI and FLAIR imaging," *NeuroImage: Clinical*, vol. 31, 2021, doi: 10.1016/j.nicl.2021.102744.

[5] C. Federau *et al.*, "Improved segmentation and detection sensitivity of diffusion-weighted stroke lesions with synthetically enhanced deep learning," *Radiology: Artificial Intelligence*, vol. 2, no. 5, 2020, doi: 10.1148/ryai.2020190217.

[6] Y.-C. Wei *et al.*, "Semantic segmentation guided detector for segmentation, classification, and lesion mapping of acute ischemic stroke in MRI images," *NeuroImage: Clinical*, vol. 35, 2022, doi: 10.1016/j.nicl.2022.103044.

[7] R. Zhang *et al.*, "Automatic segmentation of acute ischemic stroke from DWI using 3-D fully convolutional DenseNets," *IEEE Transactions on Medical Imaging*, vol. 37, no. 9, pp. 2149–2160, 2018, doi: 10.1109/TMI.2018.2821244.

[8] L. Liu, L. Kurgan, F.-X. Wu, and J. Wang, "Attention convolutional neural network for accurate segmentation and quantification of lesions in ischemic stroke disease," *Medical Image Analysis*, vol. 65, 2020, doi: 10.1016/j.media.2020.101791.

[9] S. Yalçın and H. Vural, "Brain stroke classification and segmentation using encoder-decoder based deep convolutional neural networks," *Computers in Biology and Medicine*, vol. 149, 2022, doi: 10.1016/j.compbiomed.2022.105941.

[10] Z. Wu, X. Zhang, F. Li, S. Wang, L. Huang, and J. Li, "W-Net: a boundary-enhanced segmentation network for stroke lesions," *Expert Systems with Applications*, vol. 230, 2023, doi: 10.1016/j.eswa.2023.120637.

[11] S. Wang, Z. Chen, S. You, B. Wang, Y. Shen, and B. Lei, "Brain stroke lesion segmentation using consistent perception generative adversarial network," *Neural Computing and Applications*, vol. 34, no. 11, pp. 8657–8669, 2022, doi: 10.1007/s00521-021-06816-8.

[12] O. Maier, C. Schröder, N. D. Forkert, T. Martinetz, and H. Handels, "Classifiers for ischemic stroke lesion segmentation: a comparison study," *PLOS ONE*, vol. 10, no. 12, 2015, doi: 10.1371/journal.pone.0145118.

[13] P. K. Jain, N. Sharma, M. K. Kalra, A. Johri, L. Saba, and J. S. Suri, "Far wall plaque segmentation and area measurement in common and internal carotid artery ultrasound using U-series architectures: An unseen Artificial Intelligence paradigm for stroke risk assessment," *Computers in Biology and Medicine*, vol. 149, 2022, doi: 10.1016/j.compbiomed.2022.106017.

[14] A. Deshpande *et al.*, "Automatic segmentation, feature extraction and comparison of healthy and stroke cerebral vasculature," *NeuroImage: Clinical*, vol. 30, 2021, doi: 10.1016/j.nicl.2021.102573.

[15] A. Subudhi, M. Dash, and S. Sabut, "Automated segmentation and classification of brain stroke using expectation-maximization and random forest classifier," *Biocybernetics and Biomedical Engineering*, vol. 40, no. 1, pp. 277–289, 2020, doi: 10.1016/j.bbe.2019.04.004.

[16] R. Zoetmulder *et al.*, "Brain segmentation in patients with perinatal arterial ischemic stroke," *NeuroImage: Clinical*, vol. 38, 2023, doi: 10.1016/j.nicl.2023.103381.

[17] V. Abramova *et al.*, "Hemorrhagic stroke lesion segmentation using a 3D U-Net with squeeze-and-excitation blocks," *Computerized Medical Imaging and Graphics*, vol. 90, 2021, doi: 10.1016/j.compmedimag.2021.101908.

[18] A. Tursynova and B. Omarov, "3D U-Net for brain stroke lesion segmentation on ISLES 2018 dataset," in *2021 16th International Conference on Electronics Computer and Computation*, 2021, pp. 1–4, doi: 10.1109/ICECCO53203.2021.9663825.

[19] L. Zhang *et al.*, "Ischemic stroke lesion segmentation using multi-plane information fusion," *IEEE Access*, vol. 8, pp. 45715–45725, 2020, doi: 10.1109/ACCESS.2020.2977415.

[20] N. Tomita, S. Jiang, M. E. Maeder, and S. Hassanpour, "Automatic post-stroke lesion segmentation on MR images using 3D residual convolutional neural network," *NeuroImage: Clinical*, vol. 27, 2020, doi: 10.1016/j.nicl.2020.102276.

[21] A. Kumar *et al.*, "CSNet: a new DeepNet framework for ischemic stroke lesion segmentation," *Computer Methods and Programs in Biomedicine*, vol. 193, 2020, doi: 10.1016/j.cmpb.2020.105524.

[22] Z. Lu *et al.*, "Extremely imbalanced subarachnoid hemorrhage detection based on DenseNet-LSTM network with class-balanced loss and transfer learning," in *Medical Imaging 2021: Computer-Aided Diagnosis*, 2021, pp. 1–6, doi: 10.1117/12.2582088.

[23] H. Hui, X. Zhang, F. Li, X. Mei, and Y. Guo, "A partitioning-stacking prediction fusion network based on an improved attention U-Net for stroke lesion segmentation," *IEEE Access*, vol. 8, pp. 47419–47432, 2020, doi: 10.1109/ACCESS.2020.2977946.

[24] P. K. Jain *et al.*, "Attention-based UNet deep learning model for plaque segmentation in carotid ultrasound for stroke risk stratification: an artificial intelligence paradigm," *Journal of Cardiovascular Development and Disease*, vol. 9, no. 10, 2022, doi: 10.3390/jcdd9100326.

[25] M. M. Das and S. Shastri, "Machine learning approaches for early brain stroke detection using CNN," *Journal of Scientific Research and Technology*, vol. 3, no. 6, pp. 243–250, 2025, doi: 10.61808/jsrt248.

[26] W. Abbaoui, S. Retal, S. Ziti, and B. El Bhiri, "Automated ischemic stroke classification from MRI scans: using a vision transformer approach," *Journal of Clinical Medicine*, vol. 13, no. 8, 2024, doi: 10.3390/jcm13082323.

BIOGRAPHIES OF AUTHORS

Dr. Kavikondala Praveen Kumar Rao is a distinguished academic and researcher, currently serving as professor & head in the Department of Artificial Intelligence and Machine Learning at Kamala Institute of Technology and Science, Singapur, Huzurabad, Karimnagar, Telangana 505468, India. He is recognized for his dedication to advancing knowledge in intelligent systems and data-driven technologies. He can be contacted at email: praveenkumarrao.k@gmail.com.

Maha Lakshmi Bondla is an assistant professor in the Department of Electronics and Communication Engineering at VIGNAN'S Institute of Information Technology, beside Visakhapatnam Special Economic Zone (VSEZ), Vadlapudupi Post, Duvvada, Vishakhapatnam, India. She has a keen interest in electronics, communication systems, and emerging technologies, and actively contributes to research and teaching in these areas. She can be contacted at email: mahalaxmi098@gmail.com.

Bommaraju Srinivasa Rao is a professor in the Department of Computer Science and Engineering (Data Science) at Geethanjali College of Engineering and Technology, Cheeryal, Medchal, Hyderabad-501301, Telangana, India. He specializes in data science, machine learning, and computer science applications. He is actively contributing to both research and teaching. He can be contacted at email: drbsrinivasarao.cse@gcet.edu.in.

Ambidi Naveena is an assistant professor in the Department of Electronics and Telematics Engineering at G. Narayanamma Institute of Technology and Science (for Women), Shaikpet, Raidurg, Hyderabad, India. She specializes in electronics, telematics, and emerging communication technologies, actively contributing to research and academic initiatives. She has published in reputed journals and conferences, focusing on innovative applications in her field. She can be contacted at email: naveenaambidi@gmail.com.

K. V. Balaramakrishna is an assistant professor in the Department of Electronics and Communication Engineering at Aditya University, Aditya Nagar, ADB Road, Surampalem, Kakinada District, Andhra Pradesh, India. His research interests include communication systems, signal processing, and emerging technologies in electronics. He has contributed to various research publications and academic initiatives, focusing on practical and innovative approaches to engineering education. He can be contacted at email: balaram56kv@gmail.com.

Srinivasarao Goda is an associate professor in the Department of Computer Science and Engineering at R.V.R. & J.C. College of Engineering, Guntur, India. He has extensive experience in teaching and research, with interests spanning computer science, artificial intelligence, and data-driven technologies. He has published research papers in reputed journals and conferences, contributing to advancements in computing and intelligent systems. He is passionate about academic excellence and mentoring students toward innovation and research. He can be contacted at email: gsraob4u@gmail.com.