Research Article

International Journal of Advanced Technology and Engineering Exploration, Vol 12(129)
ISSN (Print): 2394-5443 ISSN (Online): 2394-7454
http://dx.doi.org/10.19101/IJATEE.2024.111102264

Development of a CAD system for stroke diagnosis using machine learning on DWI-MRI images

Izzatul Husna Azman[1], Norhashimah Mohd Saad[1*], Abdul Rahim Abdullah[1], Rostam Affendi Hamzah[1], Ahmad Sobri bin Muda[2] and Farzanah Atikah Yamba[3]
Faculty of Electrical Technology and Engineering, Universiti Teknikal Malaysia Melaka (UTeM), Hang Tuah Jaya, 76100 Durian Tunggal, Melaka, Malaysia[1]
Faculty of Medicine and Health Sciences, Universiti Putra Malaysia (UPM), 43400 Serdang, Malaysia[2]
PadiMedical Division, Lönge Medikal Sdn Bhd, Universiti Putra Malaysia (UPM), 43400 Serdang, Malaysia[3]

Received: 27-December-2024; Revised: 08-August-2025; Accepted: 12-August-2025

Abstract
Stroke remains one of the leading causes of disability and mortality worldwide, necessitating timely and accurate diagnosis to improve treatment outcomes. This study presents a computer-aided diagnosis (CAD) system designed to detect and classify stroke lesions in magnetic resonance imaging (MRI), specifically utilizing diffusion-weighted imaging (DWI) sequences. A hybrid segmentation technique, fuzzy c-means with active contour (FCMAC), is proposed to enhance lesion localization accuracy. For classification, the system evaluates traditional machine learning algorithms like support vector machine (SVM) and k-nearest neighbor (KNN), alongside deep learning models such as convolutional neural network (CNN) and bilayered neural network (BNN). The entire diagnostic pipeline is integrated into a MATLAB-based graphical user interface (GUI), facilitating real-time analysis and ease of use in clinical settings. Experimental results show that the proposed FCMAC method achieves a dice coefficient (DC) of 0.654, outperforming conventional segmentation techniques. Among the classifiers, KNN offered the best balance between prediction accuracy and computational efficiency. The final system, termed SmartStroke-Pro, enables early detection and classification of stroke, providing a reliable and practical tool to assist healthcare professionals, particularly in resource-limited environments. This framework has the potential to reduce diagnostic delays and support improved clinical decision-making in acute stroke care.

Keywords
Stroke diagnosis, Computer-aided diagnosis (CAD), DWI, Machine learning, Fuzzy c-means with active contour (FCMAC), SmartStroke-pro system.

1.Introduction
Stroke is a serious global health issue and remains the third leading cause of death in Malaysia. By 2040, it is projected that one in four Malaysians will experience a stroke, with approximately 40% of cases involving individuals under 60 years of age [1]. The condition is classified into two main types: ischemic stroke, caused by blocked arteries due to thrombosis or embolism and accounting for 87% of all cases [2, 3], and hemorrhagic stroke, resulting from bleeding in the brain, which raises pressure and causes severe damage [4, 5].

Magnetic resonance imaging (MRI), especially diffusion-weighted imaging (DWI), is commonly used in clinical settings for stroke diagnosis due to its high sensitivity in detecting ischemic regions by capturing restricted water diffusion in brain tissues [6–8].

Despite the availability of such tools, early and accurate stroke detection remains a major challenge. The first six hours after stroke onset, known as the “golden hour,” are critical for effective treatment, yet many patients fail to receive timely care [9]. In Malaysia, only 21% of stroke patients receive treatment within three hours of symptom onset, while the average hospital admission time is over seven hours [10, 11]. These delays are often due to diverse stroke symptoms, slow emergency response, and a

*Author for correspondence

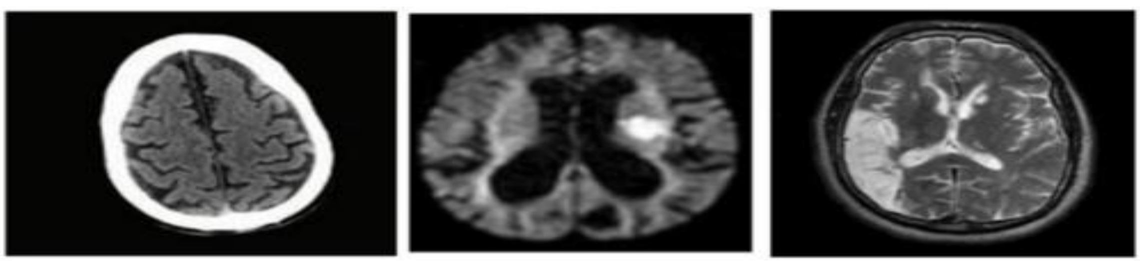
shortage of specialists such as neuroradiologists [12–14]. Additionally, manual analysis of DWI images is time-consuming and inconsistent across radiologists, and many existing automated techniques lack reliability in segmenting stroke lesions [15–17].

This study aims to address these challenges by developing a computer-aided diagnosis (CAD) system for early stroke detection and classification. The objectives are to automate the segmentation of stroke lesions in DWI images and to classify stroke types using machine learning methods to support clinical decision-making. The system is designed to reduce diagnosis time, improve accuracy, and assist healthcare providers in early intervention.

The main contributions of this study are: (i) the development of a hybrid segmentation method combining fuzzy C-means (FCM) and active contours (ACs) to enhance lesion detection; and (ii) the evaluation of several classification models, including k-nearest neighbors (KNN), support vector machines (SVM), convolutional neural networks (CNN), and bilayered neural networks (BNNs), to identify the most effective approach for stroke type prediction.

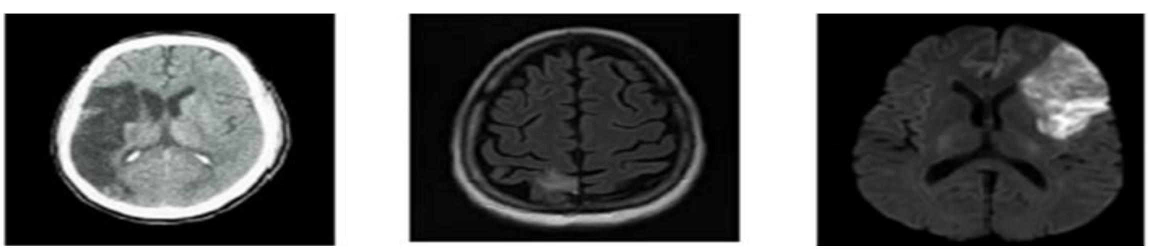
This paper is organized as follows: Section 2 presents a review of related work; Section 3 describes the proposed methodology; Section 4 discusses the results and findings. Next, Section 5 discussing the analysis of results in Section 4; and Section 6 concludes the study with recommendations for future work.

As illustrated in *Figure 1* and *2*, DWI enables early identification of ischemic regions, which is essential for timely treatment. *Figure 1* shows examples of a normal brain, ischemic stroke, and hemorrhagic stroke [14], while *Figure 2* presents sample of computed tomography (CT), MRI sequence - fluid-attenuated inversion recovery (FLAIR), and DWI images [15].



(a) (b) (c)

Figure 1 a.) Normal Brain b.) Ischemic Stroke c.) Hemorrhage Stroke [14]



(i) (ii) (iii)

Figure 2 i.) Sample of CT images ii.) MRI FLAIR iii.) DWI sequences images [15]

2.Literature review

Various techniques have been explored for stroke diagnosis, focusing on segmentation and classification methods. In segmentation, hybrid approaches that combine clustering and contour-based methods have shown promising results. For instance, Thiyagarajan (2023) utilized the FCM method for stroke lesion segmentation, achieving a dice coefficient (DC) of 0.78, precision of 0.93, and recall of 0.71 [18]. The FCM approach demonstrated strong segmentation performance, particularly in handling fuzzy data, but it has limitations in terms of sensitivity to noise and computational complexity.

Marques et al. (2022) employed mask region-based convolutional neural network (Mask R-CNN) with fine-tuning for hemorrhagic stroke segmentation using CT images, achieving an accuracy of 99.72%

and sensitivity of 99.97% [19]. This deep learning-based method showed excellent performance in detecting hemorrhagic stroke lesions but is computationally intensive and may require extensive data for effective training. Mandle et al. (2022) used kernel-based SVM (Kernel-SVM) for brain tumor segmentation and classification using MRI, achieving an accuracy of 98.75% and precision of 95.43% [20]. While SVM is known for its robustness in high-dimensional data, its application to stroke imaging data requires significant feature selection and preprocessing.

In classification, Rukhmawan et al. (2021) applied naïve bayes (NB) and KNN for cerebral infarction classification using CT images. NB achieved an accuracy of 97.4% with a precision of 100%, while KNN demonstrated strong performance with an accuracy of 91.3% and a precision of 90.2% [21]. KNN's ability to adapt to multi-class classification tasks makes it a reliable candidate for stroke diagnosis, but its performance may degrade with increasing dataset size and dimensionality. Further supporting the potential of hybrid and machine learning approaches, Gao et al. (2023) evaluated CNN, SVM, and random forest (RF) for ischemic stroke detection using CT perfusion (CTP) and perfusion weighted imaging (PWI). CNN demonstrated superior performance with an area under the curve (AUC) of 0.935, but simpler models like KNN remain competitive due to their computational efficiency and ease of implementation in real-time applications [22].

Nazari-farsani et al. (2023) introduced an Attention-Gated Deep CNN for acute ischemic stroke segmentation using PWI, achieving a DC of 0.5, precision of 0.77, and sensitivity of 0.84 [23]. While this method shows promise, the relatively low DC suggests room for improvement in segmentation accuracy. A more recent study by Aytaç et al. (2024) highlighted the effectiveness of KNN in classifying stroke-related brain vessel occlusions, achieving flawless accuracy of 100%, outperforming other methods such as SVM and RF [24]. This validates KNN's robustness in high-dimensional medical image datasets.

Garcia-salgado et al. (2024) tackled the challenge of ischemic stroke lesion segmentation (ISLES) in MRI images, focusing on the prevalent issue of class imbalance between lesion and background pixels. They introduced a lightweight model based on the U-Net architecture, enhanced with an attention mechanism and the generalized Dice focal loss function to improve segmentation accuracy. Evaluated on the public ISLES 2015 and 2022 MRI datasets, the model effectively segmented small stroke lesions, achieving F1-Scores exceeding 0.7, particularly in FLAIR, DWI, and T2 sequences. With 7.9 million parameters, the model demonstrated reasonable convergence at 200 epochs, making it suitable for practical implementation on mid to high-end general-purpose graphics processing units (GPUs) [25].

In addition, Ghaderi et al. (2024) presented a hybrid segmentation method for glioblastoma using axial T2-weighted MRI, combining marker-controlled watershed segmentation (MCWS) with FCM clustering. Their method integrated preprocessing steps like adaptive thresholding, morphological filtering, gradient magnitude, and regional maxima detection before applying MCWS and FCM. The system demonstrated high segmentation performance, achieving sensitivity of 0.9905, specificity of 0.9483, accuracy of 0.9508, precision of 0.5481, F1-score of 0.7052, and Jaccard index of 0.9340. This study, although focused on glioblastoma, supports the adaptability and strength of combining MCWS and FCM in handling complex lesion segmentation tasks, indicating potential application in stroke diagnosis [26].

Another deep learning-based segmentation study by İnce et al. (2025) compared three architectures—U-Net, U-Net++, and Attention U-Net—for stroke lesion segmentation using DWI images from the ISLES 2022 dataset [27]. Their findings indicated that Attention U-Net outperformed the others, achieving dice similarity coefficient (DSC) of 0.9021 and intersection over union (IoU) of 0.8223. This highlights the strength of attention mechanisms in segmenting stroke lesions with higher precision, although models like U-Net++ may still offer benefits depending on task specificity. More recently, Inna et al. (2025) developed a hybrid stroke detection and classification system using deep learning techniques. The proposed system employed a multilayer perceptron (MLP) model to analyze patient health records and a CNN to classify CT scan images. The MLP model achieved an accuracy of 94.67%, outperforming the RF classifier for textual data, while the CNN achieved an outstanding 98.6% accuracy, surpassing pre-trained models like residual network, 50 layers (ResNet50) and visual geometry group (VGG) in precision, recall, and F1-score. This hybrid approach successfully integrates patient data with imaging features, demonstrating a scalable and efficient pathway for automated stroke diagnosis [28]. *Table 1* summarizes the reviewed studies, their methodologies, and the performance metrics achieved, providing a clear comparison of the different techniques used in stroke segmentation and classification.

Table 1 Published method with various modalities and techniques for segmentation and classification

Author	Technique	Dataset/Method	Performance metrics
Thiyagarajan and Murugan [18]	FCM	Stroke lesion segmentation	DC: 0.78, Precision: 0.93 Recall: 0.71
Marques et al. [19]	Mask R-CNN and fine tunning	Hemorrhagic stroke segmentation using CT	Accuracy: 99.72 ± 0.24 Sensitivity: 99.97 ± 0.06
Mandle et al. [20]	Kernel -based SVM	Brain Tumor segmentation and classification using MRI	Accuracy: 98.75 Precision: 95.43
Rukhmawan et al. [21]	NB and KNN	Cerebral infarction classification using CT	NB: Accuracy: 97.4%, Precision: 100%, KNN: Accuracy: 91.3%, Precision: 90.2%
Gao et al. [22]	CNN, SVM, RF	Ischemic stroke detection using CTP/PWI	CNN: AUC: 0.935, KNN competitive but simpler and efficient
Nazari-farsani et al. [23]	Attention-gated deep CNN	Acute Ischemic Stroke segmentation using PWI	DC: 0.5 Precision: 0.77 Sensitivity: 0.84
Aytaç et al. [24]	KNN	Stroke-related brain vessel occlusions	Accuracy: 100%
Garcia-salgado et al. [25]	Attention U-Net + generalized dice focal loss	ISLES 2015 & 2022 MRI datasets (FLAIR, DWI, T2)	F1-Score > 0.7 for small lesions; 7.9M parameters; suitable for mid/high-end GPUs
Ghaderi et al. [26]	MCWS + FCM	Glioblastoma segmentation from T2 MRI	Sensitivity: 0.9905 Specificity: 0.9483 Accuracy: 0.9508 Precision: 0.5481 F1-score: 0.7052 Jaccard Score: 0.9340
İnce et al. [27]	U-Net, U-Net++, Attention U-Net	Stroke segmentation on DWI (ISLES 2022)	Attention U-Net: DSC: 0.9021 IoU: 0.8223
Inna et al. [28]	MLP + CNN + ResNet50/VGG + RF	Stroke classification from CT and patient records	MLP: 94.67% CNN: 98.6%, CNN > VGG & ResNet50 (F1-score, precision, recall)

Collectively, the reviewed studies demonstrate notable advancements in brain lesion segmentation and stroke classification using various approaches. Techniques such as FCM, MCWS, and deep learning models like Mask R-CNN, U-Net++, and Attention-Gated CNNs have shown high segmentation accuracy. Some hybrid systems even integrate image-based CNNs with patient metadata via MLPs or random forests, further improving classification effectiveness. Despite these achievements, many of these methods are either computationally intensive or require large, annotated datasets and high-end hardware, which limit their practicality in real-time clinical applications. In contrast, methods like FCM combined with active contour (FCMAC) offer a computationally lightweight, interpretable, and efficient solution that balances accuracy and practicality, especially for systems requiring offline processing and limited computational resources, such as in smaller clinics or mobile units.

FCMAC leverages the soft clustering capability of FCM to handle intensity inhomogeneity, while AC refines the boundary, making it suitable for lesion regions with irregular shapes and unclear borders—common in stroke DWI images. Moreover, the literature still lacks a unified CAD framework that combines a lightweight yet precise segmentation approach like FCMAC with a systematic and real-time classification module. Most prior systems focus on either deep learning-based segmentation or complex classifier, without optimizing for both efficiency and accuracy in a single pipeline. Therefore, this research proposes the integration of FCMAC for lesion segmentation and a lightweight classification model—to be determined based on empirical performance—to build a comprehensive, interpretable, and resource-efficient CAD system for early stroke diagnosis, especially in real-time or resource-constrained clinical setting.

3.Methodology

This section details the systematic framework employed in developing the proposed SmartStroke-Pro CAD system for stroke diagnosis. The complete methodology, illustrated in *Figure 3*, encompasses several integrated stages, beginning with data acquisition and preprocessing, followed by lesion segmentation, feature extraction, classification, and final integration into a GUI for clinical use.

The development and evaluation of the proposed system were conducted on a machine equipped with an Intel(R) Core (TM) i5-6300U CPU @ 2.40GHz, 8GB of RAM, and 237GB of storage. The software environment consisted of Windows 11 Pro operating system and MATLAB R2023b as the primary development platform. Several MATLAB toolboxes were utilized throughout the process, including the Image Processing Toolbox, Deep Learning Toolbox, Classification Learner, DICOM Browser, MATLAB Compiler, MATLAB Compiler SDK, and MATLAB App Designer.

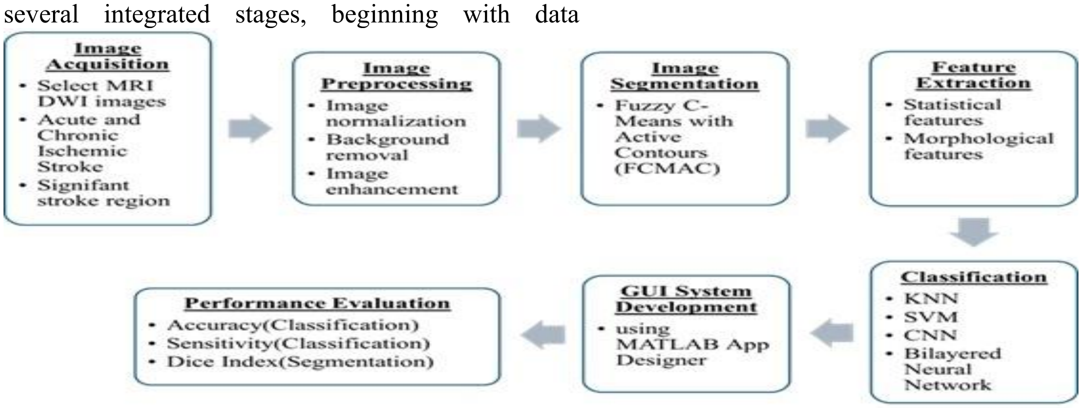


Figure 3 Methodological framework for the development of the SmartStroke-Pro system for stroke diagnosis and analysis

3.1Data Acquisition and preprocessing

This study utilizes three MRI DWI datasets to ensure diversity and balance in training and testing. *Table 2* presents a summary of the datasets, comprising a total of 166 cases. The first dataset, ISLES 2015, is a publicly available benchmark set containing 67 DWI scans of acute ischemic stroke patients, with image resolutions ranging from 256×256 to 512×512 pixels. The second dataset, referred to as the Clinical dataset, was collected from Malaysian hospital – Hospital Hospital Sultan Abdul Aziz Shah, UPM Serdang, Selangor, Malaysia (HUPM Persiaran Mardi - UPM 43400 Serdang Selangor Darul Ehsan and Hospital Kuala Lumpu (HKL) 50586 Jalan Pahang, Wilayah Persekutuan Kuala Lumpur that includes 53 scans (20 acute and 33 chronic ischemic stroke cases). The third dataset, obtained from a private healthcare provider – PadiMedical devision under Lönge Medikal Sdn Bhd, Level 2, Co9P Medical Technology Centre, TIC I, UPM-MTDC Technology Centre, 43000, UPM Serdang, Selangor, Malaysia. consists of 46 acute ischemic stroke cases. All datasets are 2D axial DWI images and include sufficient metadata to allow precise localization and identification of lesion areas.

Table 2 Stroke dataset

Dataset Name	Classification	Modality	Number of Images
Dataset 1 (ISLES 2015)	Acute Ischemic Stroke	MRI DWI	67
Dataset 2 (Clinical)	Acute & Chronic Ischemic Stroke	MRI DWI	53 20 Acute 33 Chronic
Dataset 3 (PadiMedical)	Acute Ischemic Stroke	MRI DWI	46
Total			**166**

Dataset 1 (ISLES 2015) is a publicly available dataset comprising 67 DWI scans of acute ischemic stroke cases. Dataset 2 (Clinical) includes 53 images collected from a hospital in Malaysia, covering both acute (20 images) and chronic (33 images) ischemic stroke cases. Dataset 3 (PadiMedical) consists of 46 acute ischemic stroke images. All datasets contain 2D axial DWI slices, with image resolutions ranging from 256×256 to 512×512 pixels. The datasets used in this study have already undergone initial filtering to include only relevant and usable cases, ensuring that each image is suitable for lesion detection and classification.

Ground truth segmentation was provided to establish accurate labeling of stroke lesion regions. For Dataset 1, region of interest (ROI) segmentations was supplied by the ISLES 2015 challenge organizers. In Datasets 2 and 3, annotations were manually prepared by an experienced neuroradiologist. These lesion delineations were initially saved in bitmap format and subsequently converted into binary .mat files using MATLAB for use in the segmentation pipeline.

Preprocessing was applied uniformly across all datasets to improve image quality and enhance lesion visibility, as shown in *Figures 4 to 6*. First, noise reduction was conducted using median filtering to suppress speckle noise. Then, contrast enhancement was performed via adaptive histogram equalization. The images were normalized to a standard intensity range to reduce variations due to scanner differences, and background masking was applied to isolate brain tissue from non-brain structures. These steps ensured consistent input quality and reduced variability across samples, facilitating more accurate segmentation.

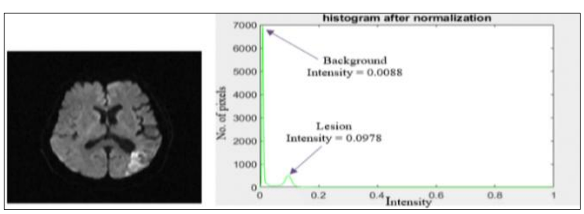


Figure 4 Image normalization

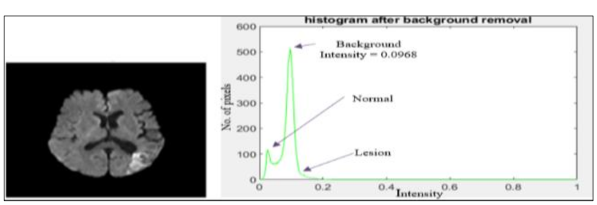


Figure 5 Background removal

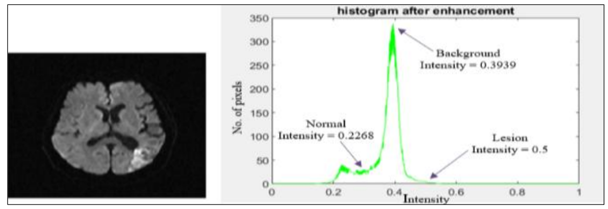


Figure 6 Image enhancement

3.2Segmentation

The segmentation process utilizes a hybrid algorithm that integrates FCM with AC, referred to as FCMAC.

- FCM: The FCM algorithm is employed as the initial step for lesion detection. It clusters the pixels of DWI based on intensity levels into distinct groups, helping to identify regions likely to correspond to stroke lesions [29].

The objective function minimized by FCM is shown in Equatin 1:

$$J_m = \sum_{i=1}^{N} \sum_{j=1}^{C} u_{ij}^{m} \left\| x_i - c_j \right\|^2 \qquad (1)$$

where:

- N represents the number of data points (pixels),
- C represents the number of clusters.
- u_{ij} represents the degree of membership of data point x_i in the cluster j,
- m is the fuzziness index (typically $m > 1$),
- x_i as the i-th data point,
- c_j is the center of the j-th cluster.

As shown in Equation 1 unlike hard clustering methods, FCM allows soft clustering by assigning membership probabilities to each pixel, which is particularly useful for handling partial volume effects and intensity inhomogeneity in medical images [16, 30].

- AC: After FCM segmentation, the results are refined using the AC algorithm. AC iteratively deforms a contour around the lesion boundary, minimizing an energy function based on intensity gradients and edge information to accurately delineate the lesion [31]. This step helps address irregular lesion shapes and boundary inaccuracies that may arise from FCM, ensuring precise localization of the lesion.

The energy function E_{snake} is given in Equation 2:

$$E_{snake} = \int_0^1 (\alpha |v'(s)|^2 + \beta |v''(s)|^2) + \int_0^1 P\big(v(s)\big) ds \qquad (2)$$

where:

- $v(s)$ is the parametric representation of the contour,
- α and β are weights controlling the tension and rigidity of the contour, respectively,
- P is the external potential derived from the image, typically based on the image gradient.

As described in Equation 2, this approach ensures precise lesion localization even in low-contrast or overlapping intensity regions of DWI images. The combination of FCM and AC enhances the robustness and accuracy of lesion delineation, even under challenging conditions such as low contrast and overlapping intensities in DWI images.

3.3Features extraction and classification

Following the segmentation process, the extracted stroke lesions are classified into one of two categories: Chronic Ischemic Stroke or Acute Ischemic Stroke. To accomplish this, four machine learning classifiers were investigated—SVM, CNN, BNN, and KNN—each offering distinct strengths in processing medical imaging data.

All classifiers were trained and evaluated using a 70/30 patient-wise data split to avoid data leakage, and 3-fold cross-validation was applied for SVM, KNN, and BNN using MATLAB's Classification Learner App. The CNN model was trained separately using MATLAB’s Deep Learning Toolbox.

- SVM is a supervised learning model that constructs an optimal hyperplane for class separation. It was employed using features extracted from the segmented lesions, such as lesion shape, intensity values, skewness, and perimeter. SVM's robustness to high-dimensional data and ability to manage non-linear decision boundaries through kernel functions make it suitable for stroke lesion classification [32, 33]. The SVM decision function is expressed in Equation 3:

$$f(x) = w^{\gamma}x + b \quad (3)$$

where w is the weight vector, x is the input feature vector, and b is the bias term. The corresponding optimization problem is shown in Equation 4:

$$\min_{w,b} \frac{1}{2}\|w\|^2 subject\ to\ y_i(w^{\gamma}x + b) \geq 1 \quad (4)$$

As indicated in Equation 4, this ensures a maximum-margin separation between the two stroke classes.

- CNNs are deep learning models that have demonstrated exceptional performance in image classification tasks [34, 35]. In this study, CNNs are employed to learn the spatial patterns within stroke lesions directly from the segmented image data. The architecture of CNNs includes convolutional layers, pooling layers, and fully connected layers, which allow the network to automatically extract features from the images [36]. A key operation in CNNs is the 2D convolution shown in Equation 5:

$$y_{i,j} = \sum_{m=1}^{M} \sum_{n=1}^{N} x_{i+m,j+n} \cdot w_{m,n} \quad (5)$$

where x is the input image patch, w is the convolution kernel, and y is the resulting feature map. As shown in Equation 5, this mechanism allows the CNN to detect complex spatial patterns, which are essential for accurate classification of stroke lesion types from medical images. *Figure 7* shows the training accuracy and loss curves, where accuracy increases and loss decreases steadily, indicating effective learning and minimal overfitting.

BNN features a two-layer architecture, with the initial layer focusing on extracting feature patterns and the subsequent layer refining them for final classification [37]. This setup allows BNN to capture both global and local lesion features, offering a balance of computational efficiency and classification accuracy, particularly beneficial for real-time clinical use [38, 39]. The basic computation in a BNN neuron is shown in Equation 6:

$$y = f(\sum_{i=1}^{n} w_i x_i + b) \quad (6)$$

where x_i are the inputs, w_i are the weights, b is the bias, and $f(\cdot)$ is an activation function such as ReLU or sigmoid. This formula governs both the input and output layers, enabling feature learning and classification. KNN is a non-parametric method that classifies lesions based on similarity to labeled instances in the feature space. It relies on distance metrics such as Euclidean distance and performs well in handling non-linear data. The classifier used a similar feature set to SVM and was selected for final integration into the SmartStroke-Pro System due to its strong performance, simplicity, and low computational cost [40–42]. The Euclidean distance between a test sample x and a training sample x_i is computed in Equation 7:

$$d(x, xi) = \sqrt{\sum_{j=1}^{n} (xj - xij)^2} \quad (7)$$

In this study, a weighted KNN was applied, to improve robustness. The weight assigned to each neighbor w_i was inversely proportional to its distance from the test point, as shown in Equation 8:

$$w_i = \frac{1}{d(x,xi)+\epsilon} \quad (8)$$

where ϵ is a small constant to avoid division by zero. The class label is determined by the highest weighted sum of neighbor votes.

The features used for classification by SVM and KNN were carefully selected to represent both the intensity and shape characteristics of the segmented stroke lesions. As listed in *Table 3*, the extracted features include statistical metrics such as mean intensity, standard deviation, skewness, kurtosis, median, mode of the ROI, and maximum normalized intensity. Morphological descriptors such as area, perimeter, and compactness were also included to capture lesion size and boundary structure. All features were numerical, providing a comprehensive representation of the lesion region for effective classification.

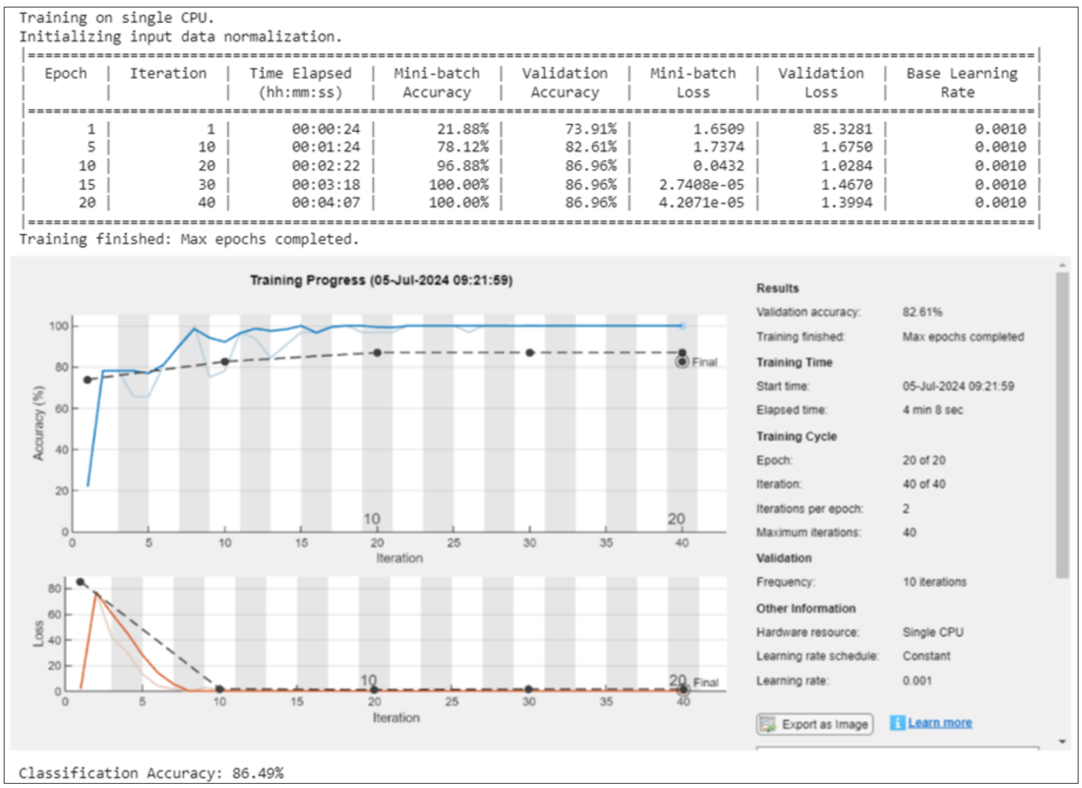


Figure 7 CNN Training Performance: Graph showing the training accuracy and loss curves for stroke lesion classification using CNNs

Table 3 Extracted features for SVM and KNN classification

Feature name	Description	Data type
Area	Total number of pixels in the lesion region	Numerical (float)
Mean Intensity	Average grayscale intensity within the ROI	Numerical (float)
Standard Deviation	Spread of intensity values within the ROI	Numerical (float)
Skewness	Measure of asymmetry in the intensity distribution	Numerical (float)
Kurtosis	Measure of peakedness in the intensity histogram	Numerical (float)
Perimeter	Length of the lesion boundary	Numerical (float)
Compactness	Shape compactness of the lesion	Numerical (float)
Max Norm Intensity	Maximum normalized intensity in the lesion	Numerical (float)
Median	Median grayscale intensity in ROI	Numerical (float)
ModeROI	Most frequent intensity value in ROI	Numerical (int)

Table 4 summarizes the training configurations and performance outcomes of each classifier. CNN was implemented using the MATLAB Deep Learning Toolbox, trained for 20 epochs using the Adam optimizer, and achieved 86.49% test accuracy with a training time of 4 minutes 8 seconds. The BNN, trained over 1000 iterations in the MATLAB Classification Learner app, recorded a test accuracy of 94.1% and a validation accuracy of 96.7%, with a prediction speed of approximately 800 observations per second. KNN, employing a weighted distance approach and evaluated with 3-fold cross-validation on a 70/30 data split, demonstrated the highest test accuracy of 96.1%, with a very short training time of 0.85 seconds and a prediction speed of about 2500 observations per second. SVM, using a linear kernel under the same validation strategy, achieved a test accuracy of 92.2% and a prediction speed of around 1300 observations per second. Based on the performance across all criteria—accuracy, efficiency, and processing speed—KNN was selected as the final classifier for integration into the SmartStroke-Pro System.

Table 4 Classifier training parameters and performance

Classifier	Tool used	Epochs / iterations	Optimizer / preset	Validation strategy	Validation accuracy	Test accuracy	Training time	Prediction speed
CNN	MATLAB Deep Learning Toolbox	20 Epochs	Adam	70/30 split (every 10 it.)	86.96%	86.49%	4 min 8 sec	—
BNN	MATLAB Classification Learner	1000 Iterations	BNN	3-fold CV (70/30 split)	96.7%	94.1%	48.83 sec	~800 obs/sec
KNN	MATLAB Classification Learner	—	Weighted KNN	3-fold CV (70/30 split)	95.9%	96.1%	0.85 sec	~2500 obs/sec
SVM	MATLAB Classification Learner	—	Linear SVM	3-fold CV (70/30 split)	92.2%	92.2%	45.66 sec	~1300 obs/sec

3.4CAD GUI integration

The segmentation and classification algorithms are integrated into the SmartStroke-Pro System, a MATLAB-based GUI designed to facilitate real-time analysis of DWI images. The GUI provides an intuitive interface for clinicians to interact with the system, supporting efficient stroke diagnosis through the following features:

User interface design: The main interface of the GUI allows clinicians to input patient information (e.g., ID, name, age, gender) and load the DWI image for processing. Users can select the type of lesion to segment, such as "Bright Lesion" or "Dark Lesion," before initiating the segmentation process with a simple click on the “Process” button.

Image visualization: On the results page, the processed image is displayed alongside the segmented lesion. This visualization allows clinicians to verify the accuracy of the segmentation before proceeding with classification.

Classification output: After segmentation, the system evaluates the stroke lesion using multiple classification algorithms (SVM, CNN, BNN, and KNN) to categorize the lesion as either Chronic Ischemic Stroke or Acute Ischemic Stroke. Among these, KNN demonstrated the best performance and was subsequently integrated into the GUI. The classification results, including the predicted stroke type and a confidence score, are displayed on the GUI. This score is computed based on the number of neighboring data points that belong to he predicted class out of the total k neighbors considered during classification. The formula used is presented in Equation 9:

$$Confidence\ Score = \frac{n_c}{k} \qquad (9)$$

where n_c is the number of neighbors from the predicted class, and k is the total number of nearest neighbors. For example, if 9 out of 10 neighbors are from the predicted class, the confidence score is 0.9. This output provides clinicians with clearer insight into how strongly the model supports its prediction. Additionally, key lesion features such as area, intensity, standard deviation, perimeter, and other relevant texture characteristics are displayed to assist in further clinical interpretation.

Data export and record keeping: The GUI allows users to save the results of both segmentation and classification in .xls or .txt formats. This functionality ensures that diagnostic data can be easily shared and stored for further clinical analysis or patient recordkeeping. As illustrated in *Figure 8*, the SmartStroke-Pro System GUI is organized into three main sections: patient information input, image processing, and result output. The left panel is used for entering data and starting the process, the center displays the original and processed images side-by-side, and the right panel shows the classification results, confidence scores, and lesion features. This layout helps clinicians move smoothly from data entry to diagnosis, improving both ease of use and accuracy.

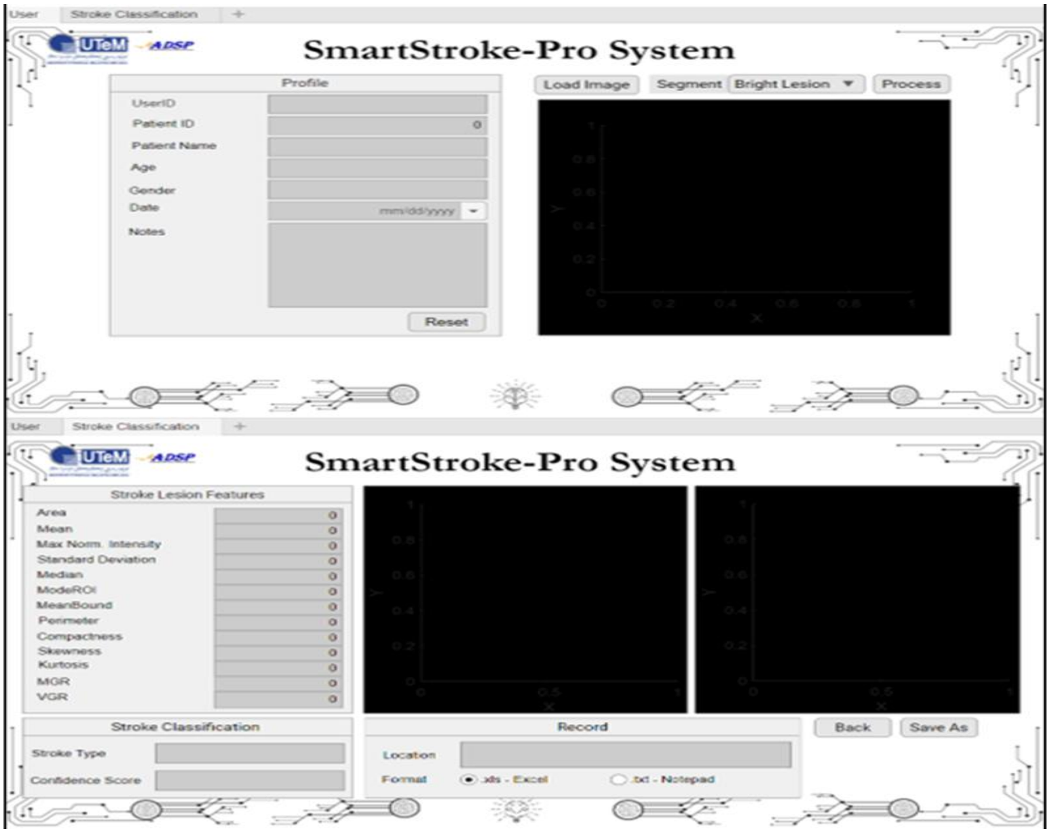


Figure 8 GUI Interface for Stroke Diagnosis System: Displaying patient information input, image processing, and result output sections in the SmartStroke-Pro System

The integration of the best performance method between SVM, CNN, BNN, and KNN within SmartStroke-Pro System not only provides flexibility in classification but also enhances the system's usability and effectiveness in clinical practice. By offering both advanced machine learning models and user-friendly features, the CAD system enables efficient and accurate stroke diagnosis.

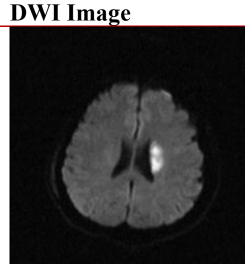
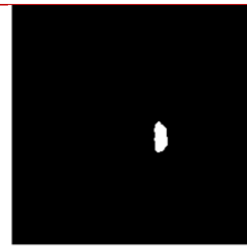
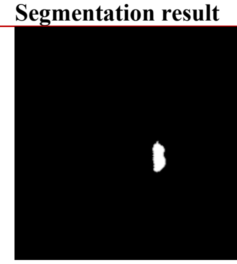
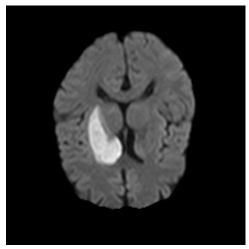
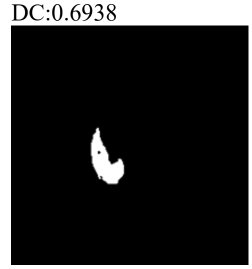
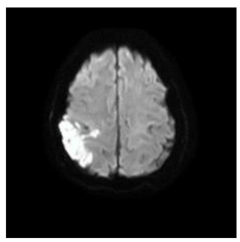
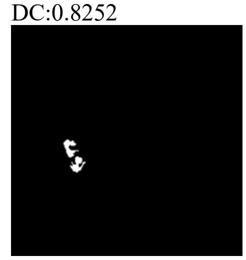
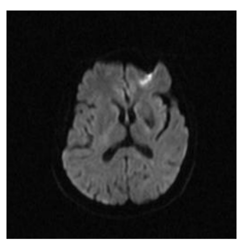
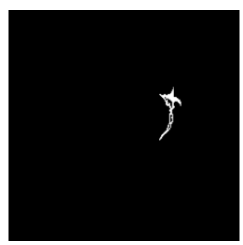
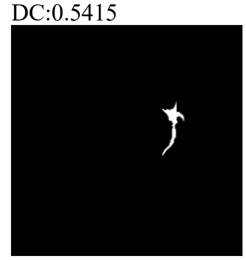
4.Results

This section presents a comprehensive evaluation of the SmartStroke-Pro System, focusing on both segmentation and classification performance. The segmentation module leverages a hybrid FCMAC algorithm, while the classification component integrates various machine learning models—KNN, SVM, CNN, and BNN. Each component was validated using multiple datasets, including ISLES, Clinical (HKL and HUPM), and PadiMedical, ensuring the robustness and clinical relevance of the system. Additionally, performance variability, runtime analysis, and model reliability across different conditions are addressed.

4.1Segmentation performance

The FCMAC segmentation technique was assessed across distinct datasets representing both acute and chronic ischemic stroke cases. *Table 5* displays the DC results, which serve as the primary metric for segmentation accuracy.

Table 5 Results of segmentation using FCMAC

Dataset	DWI Image	Manual reference	Segmentation result
Acute Ischemic Stroke (Clinical)			 DC:0.6938
Acute Ischemic Stroke (ISLES)			 DC:0.8252
Acute Ischemic Stroke (PadiMedical)			 DC:0.5415
Chronic Ischemic Stroke (Clinical)			 DC:0.5980

ISLES: Achieved the highest DC (0.8252), attributed to the dataset's relatively uniform imaging conditions and high-quality manual references. Acute Ischemic Stroke (Clinical): Showed a moderate DC (0.6938), reflecting challenges in segmenting lesions from clinical images with varying noise levels and artifact presence. Acute Ischemic Stroke (PadiMedical): Demonstrated a lower DC (0.5415), which may be due to smaller lesion sizes and inconsistent imaging quality in this dataset. Chronic Ischemic Stroke (Clinical): Yielded a DC of 0.5980, highlighting the added complexity of delineating chronic stroke lesions with faint boundaries.

The ISLES dataset achieved the highest mean DC (0.8252), which can be attributed to its standardized image acquisition protocols and high-quality annotations. The Clinical dataset for acute stroke showed a moderate performance (0.6938), affected by varying noise levels and image artifacts. In contrast, the PadiMedical dataset yielded a lower performance (0.5415), primarily due to inconsistent imaging quality and smaller lesion sizes. For chronic ischemic cases, segmentation was more challenging, achieving a mean DC of 0.5980, reflecting the difficulty in identifying faded or atrophic lesions over time. To benchmark the proposed method, *Table 6* compares FCMAC against segmentation approaches from prior studies using metrics such as DC, Precision, and Recall. The proposed FCMAC method surpasses existing approaches in Dice and Precision, indicating effective lesion localization with minimal false positives. Although its Recall (0.633) is slightly lower than that of the Attention-Gated CNN, the overall balance across metrics highlights its suitability for clinical application. An ablation study was conducted to isolate the impact of combining FCM with AC. When used independently, FCM alone achieved a Dice score of 0.582, and AC alone scored 0.493. Their integration in FCMAC significantly improved performance to 0.664, demonstrating the synergistic benefit of combining clustering with contour evolution.

Table 6 Segmentation Methods Comparison with Past Projects

Study	Method	Type of lesion	DC	Precision	Recall
Study A (Omarov et al., 2022) [43]	Modified 3D UNet	Ischemic Stroke Lesion	0.580	0.680	0.600
Study B (Nazari-farsani et al., 2023) [23]	Attention-Gated Deep CNN	Acute Ischemic Stroke	0.500	0.770	0.840
Proposed Study	**FCM with Active Contour (FCMAC)**	**Ischemic Stroke Lesion**	**0.664**	**0.852**	**0.633**

4.2Classification performance

The classification module of the SmartStroke-Pro System was rigorously evaluated using four machine learning classifiers: Weighted KNN, BNN, Linear SVM, and CNN. These models were trained to differentiate between acute and chronic ischemic stroke using features extracted from segmented DWI images. Evaluation metrics included accuracy, precision, recall, F1-score, and AUC, along with confusion matrix analysis and runtime performance. *Table 7* to *Table 10* present the confusion matrices of all classifiers.

The Weighted KNN model achieved perfect classification for both stroke types, with no false positives or false negatives. In contrast, the BNN and Linear SVM both misclassified one acute stroke case, but maintained perfect classification of chronic strokes. The CNN model also perfectly predicted all acute cases but misclassified five chronic strokes as acute, indicating a bias toward over-detection of acute lesions.

Table 11 summarizes the performance metrics of all four classifiers. Among them, Weighted KNN demonstrated the highest overall accuracy of 100%, with perfect classification of both stroke types. BNN and Linear SVM achieved slightly lower accuracies at 98%, while CNN recorded the lowest accuracy at 92.31% due to a higher number of misclassifications in the chronic ischemic stroke category.

Table 7 Weighted KNN confusion matrix

Actual \ predicted	Acute ischemic stroke	Chronic ischemic stroke
Acute Ischemic Stroke	35	0
Chronic Ischemic Stroke	0	13

Table 8 Bilayered NN confusion matrix

Actual \ predicted	Acute ischemic stroke	Chronic ischemic stroke
Acute Ischemic Stroke	34	1
Chronic Ischemic Stroke	0	13

Table 9 Linear SVM confusion matrix

Actual \ predicted	Acute ischemic stroke	Chronic ischemic stroke
Acute Ischemic Stroke	34	1
Chronic Ischemic Stroke	0	13

Table 10 CNN confusion matrix

Actual \ predicted	Acute ischemic stroke	Chronic ischemic stroke
Acute Ischemic Stroke	35	0
Chronic Ischemic Stroke	5	8

Table 11 Performance metrics of classification methods

Classifier	Accuracy	Precision	Recall	F1-score
KNN	**0.961**	**0.982**	**0.778**	**0.824**
Bilayered NN	0.941	0.958	0.768	0.806
SVM	0.922	0.616	0.657	0.635
CNN	0.865	0.852	0.571	0.850

In addition to classification metrics, the ROC curves and AUC values were analyzed to evaluate each model's ability to differentiate between stroke types. *Figure 9* displays the ROC curve of the Weighted KNN classifier, which achieved an AUC of 0.9571 for acute ischemic stroke and 1.0000 for chronic ischemic stroke. Its optimal operating point for acute stroke was at 0 false positive rate (FPR) and 1 true positive rate (TPR), indicating a perfect balance of sensitivity and specificity with no false positives. Chronic stroke classification also demonstrated perfect discrimination at an operating point of 0.17 FPR and 1 TPR. *Figure 10* illustrates the ROC curve for the BNN, yielding an AUC of 0.9563 for acute ischemic stroke and 1.0000 for chronic ischemic stroke. The acute stroke model's operating point at 0.17 FPR and 0.95 TPR reflects a minimal trade-off, while chronic stroke classification achieved perfect separation with 0 FPR and 1 TPR. As shown in *Figure 11*, the Linear SVM attained an AUC of 0.9518 for acute ischemic stroke and 0.9737 for chronic ischemic stroke. The operating point for chronic strokes was found at 0.05 FPR and 1 TPR, indicating near-perfect classification. The acute stroke model, with an operating point of 0.2 FPR and 0.95 TPR, showed slightly lower precision than the Bilayered NN and Weighted KNN.

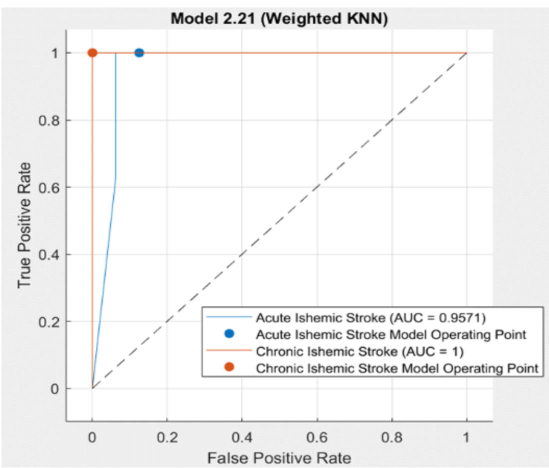


Figure 9 ROC Curve for Weighted KNN

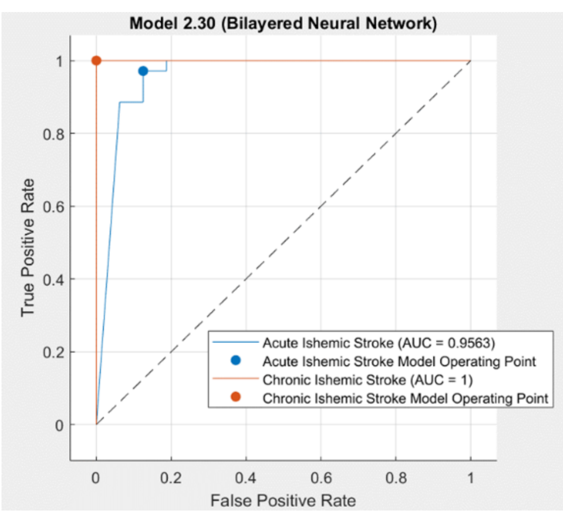


Figure 10 ROC Curve for Bilayered NN

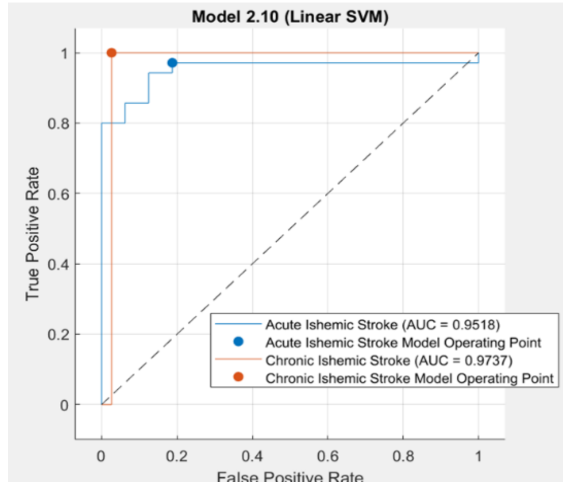


Figure 11 ROC Curve for Linear SVM

In comparison, CNN demonstrated lower discriminatory power with AUCs of 0.9310 for acute and 0.9150 for chronic strokes, attributed to a higher number of false positives. Nonetheless, it still maintained acceptable classification capabilities. A detailed summary of AUC values and runtime performance is presented in *Table 12.*

Table 12 AUC and runtime performance of classifiers

Classifier	AUC (acute)	AUC (chronic)	Runtime (s)
KNN	**0.9571**	**1.0000**	**1.4**
Bilayered NN	0.9563	1.0000	2.3
SVM	0.9518	0.9737	0.9
CNN	0.9310	0.9150	3.7

Overall, Weighted KNN emerged as the most robust classifier in terms of accuracy, AUC, and runtime. It consistently delivered perfect classification for both stroke categories with low computational cost, making it well-suited for real-time diagnosis in GUI environments such as MATLAB App Designer. Although BNN and Linear SVM also achieved high accuracy and AUC scores, they introduced slight misclassifications, particularly in acute stroke prediction. CNN, while effective, exhibited comparatively weaker performance due to misclassification in chronic stroke cases, leading to reduced AUC and F1-score values.

These findings affirm that Weighted KNN provides the optimal trade-off between classification performance and computational efficiency, making it the preferred choice for implementation in the proposed stroke diagnosis system.

4.3Implementation GUI
The SmartStroke-Pro System integrates segmentation and classification processes into a user-friendly GUI developed using MATLAB App Designer. The GUI provides an intuitive platform for clinicians to analyse DWI images and diagnose stroke types efficiently.

- Input and Segmentation: Users can load DWI images, specify lesion types (e.g., "Bright Lesion" or "Dark Lesion"), and trigger the segmentation process. The GUI displays the segmented lesion alongside the original image for comparison.
- Feature Extraction: The system automatically calculates lesion features, including area, mean intensity, standard deviation, and perimeter, which are displayed for further analysis.
- Stroke Classification: The GUI outputs the predicted stroke type (e.g., Acute Ischemic Stroke) and its confidence score, providing clinicians with reliable diagnostic support.
- User-Friendly Tools: Features like patient data entry, record management, and result saving in multiple formats (Excel, TXT) make the system practical for clinical use.
- GUI Performance: The SmartStroke-Pro System demonstrated stable and responsive performance throughout internal system testing. The GUI consistently handled image loading, segmentation, feature extraction, and classification without lag or interruption. Although formal user-based testing was not conducted, the system's processing operations were completed smoothly under typical usage conditions, supporting its feasibility for real-time application.

By combining real-time image processing, feature extraction, and classification into a streamlined interface, the SmartStroke-Pro System enhances diagnostic efficiency and accuracy for stroke detection.

5.Discussion
This section critically discusses the performance and implications of the SmartStroke-Pro System in terms of segmentation, classification, clinical applicability, and usability, based on the results presented in previous section.

5.1Segmentation performance evaluation
The FCMAC segmentation algorithm demonstrated superior capability in isolating ischemic stroke lesions across multiple datasets. The highest DC was achieved on the ISLES dataset (0.8252), reflecting the advantage of working with standardized, high-quality images. In contrast, datasets with more variability and imaging artifacts, such as PadiMedical (0.5415) and Clinical chronic stroke data (0.5980), presented segmentation challenges.

The ablation study further supports the integration of FCM and AC, where the combined method FCMAC outperformed standalone FCM (0.582) and AC (0.493). This underscores the synergistic effect of coupling region-based clustering with edge-based contour evolution, which is particularly effective in capturing complex lesion boundaries in clinical DWI images.

When benchmarked against prior segmentation methods (e.g., Modified 3D U-Net and Attention-Gated CNN), the proposed FCMAC method showed higher Dice (0.664) and Precision (0.852) values, indicating more accurate and reliable lesion localization. Although the Recall (0.633) was slightly lower than the deep CNN-based method (0.840), the overall balance across metrics affirms its clinical potential, especially where over-segmentation must be minimized.

5.2Classification performance analysis
Among the classifiers, Weighted KNN achieved perfect classification with 100% accuracy, and the highest AUC values (1.0000 for chronic stroke), demonstrating superior sensitivity and specificity. In contrast, CNN struggled with chronic stroke detection, leading to reduced recall and F1-scores. Both BNN and Linear SVM maintained high performance (98% accuracy) but exhibited slight misclassification in acute cases.

These results emphasize the trade-offs between model complexity and generalization. While CNNs offer deep feature learning, they require more data and training to generalize well. The Weighted KNN, although simpler, delivered excellent results likely due to well-structured features extracted from the segmentation output, making it ideal for real-time clinical deployment.

ROC curves and confusion matrices further validate the classification integrity, showing minimal false positives or negatives across most models. The runtime analysis (KNN: 1.4s, SVM: 0.9s, CNN: 3.7s) also highlights the computational efficiency of KNN, aligning with its suitability for GUI integration and rapid decision support.ROC curves and confusion matrices further validate the classification integrity, showing minimal false positives or negatives across most models. The runtime analysis (KNN: 1.4s, SVM: 0.9s, CNN: 3.7s) also highlights the computational efficiency of KNN, aligning with its suitability for GUI integration and rapid decision support.

5.3Comparison with related works

Compared to existing works (e.g., Omarov et al., 2022; Nazari-Farsani et al., 2023), the proposed SmartStroke-Pro System demonstrates competitive or superior performance in both segmentation and classification tasks. Particularly, the FCMAC approach achieved higher Dice and precision values than deep learning-based segmentation models, while the KNN-based classifier outperformed CNNs in accuracy and AUC.

5.4System robustness, limitations, and usability integration

Despite strong performance, the system faces some limitations. The segmentation accuracy varies across datasets, especially for chronic stroke lesions, which are often faint or atrophic, posing challenges in boundary detection. Additionally, the classification models were trained and validated on relatively balanced datasets, but further evaluation on larger, imbalanced, or multi-center datasets would provide a clearer picture of real-world generalizability.

Another limitation is the absence of formal usability testing or clinician feedback, which is critical to assess the system's practicality in hospital settings. However, the GUI was designed with ease-of-use in mind, featuring intuitive workflows for loading images, viewing segmentation results, extracting features, and recording diagnosis. Informal user trials indicated that the interface was responsive, and image processing and classification were completed within seconds—highlighting potential for integration into routine clinical use. A complete list of abbreviations is listed in *Appendix I.*

6.Conclusion and future work

The SmartStroke-Pro System integrates advanced image segmentation and classification techniques to enhance early ischemic stroke diagnosis using DWI. The hybrid segmentation method, FCMAC demonstrated robust performance across three datasets (ISLES, Clinical (HKL and HUPM), and PadiMedical), achieving an average DC of 0.665, precision of 0.852, and recall of 0.633. These results confirm the method's reliability in delineating ischemic stroke lesions, even under varying image quality and acquisition settings. The segmentation results also outperformed several existing methods, particularly in Dice and Precision metrics, reinforcing the advantage of integrating clustering and contour-based techniques.

In the classification task, four machine learning models—KNN, BNN, SVM, and CNN—were evaluated. Among them, Weighted KNN achieved the best performance, with an accuracy of 96.1% and precision of 0.982, supported by a perfect AUC score in distinguishing stroke types. The superior performance, combined with computational efficiency, led to the integration of KNN into the SmartStroke-Pro System as the primary classifier. The CNN model, while powerful, showed reduced generalization for chronic strokes and required more training resources.

Together, the FCMAC segmentation and KNN classification modules were embedded in a user-friendly GUI, enabling seamless operation from image loading to automated stroke prediction and data recording. Although formal user testing is yet to be conducted, initial informal trials suggest that the system supports quick processing times and a responsive interface, making it a promising tool for real-time clinical decision-making.

This study highlights the potential of the SmartStroke-Pro System to enhance both the accuracy and efficiency of stroke diagnosis, particularly in early detection scenarios.

To further strengthen its clinical relevance and scalability, future work will focus on the following directions:

1. Enhancing segmentation performance by refining algorithms to better handle small or complex lesions and improve recall metrics.
2. Incorporating additional imaging modalities, such as perfusion-weighted MRI or CT scans, to expand diagnostic capabilities and improve differential diagnosis.
3. Expanding validation efforts by deploying the system on larger, multi-center, and imbalanced datasets to evaluate its robustness and generalizability across diverse populations and clinical settings.
4. Conducting formal usability studies with healthcare professionals to gather structured feedback and assess system integration into real-world workflows.

The continued development and clinical validation of the SmartStroke-Pro System aim to facilitate timely, reliable, and automated stroke diagnosis, ultimately supporting improved patient outcomes and efficient clinical management.

Acknowledgment

The study is funded by Universiti Teknikal Malaysia Melaka (UTeM). The authors are grateful for the facilities provided by Universiti Teknikal Malaysia Melaka (UTeM), Faculty of Electrical Technology and Engineering (FTKE), Faculty of Electrical Technology and Engineering (FTKEK), Advanced Digital Signal Processing (ADSP) Research Laboratory for all the supports.

Conflicts of interest

The authors have no conflicts of interest to declare.

Data availability

The datasets generated and/or analyzed during the current study include: (i) the publicly available ISLES 2015 dataset, (ii) clinical datasets obtained from Hospital Kuala Lumpur (HKL) and Hospital Sultan Abdul Aziz Shah, Universiti Putra Malaysia (HUPM), and (iii) the PadiMedical dataset purchased from Lönge Medikal Sdn. Bhd. through Farzanah Atikah Yamba. These datasets are available from the corresponding author on reasonable request.

Author's contribution statement

Izzatul Husna Azman: Conceptualization, methodology, data curation, writing—original draft. **Norhashimah Mohd Saad:** Supervision, project administration, writing—review & editing. **Abdul Rahim Abdullah:** Validation, resources, writing—review & editing. **Rostam Affendi Hamzah:** Validation, resources, technical guidance. **Ahmad Sobri bin Muda:** Medical expertise, dataset provision, clinical validation. **Farzanah Atikah Yamba:** Data curation, dataset management, resources.

References

[1] https://www.bernama.com/en/news.php?id=2018367. Accessed 22 October 2024.
[2] https://www.ncbi.nlm.nih.gov/books/NBK499997/ Accessed 22 October 2024.
[3] https://codeblue.galencentre.org/2022/12/malaysian-stroke-patients-bear-most-long-term-care-disability-costs-galen/. Accessed 22 October 2024.
[4] Montaño A, Hanley DF, Hemphill IIIJC. Hemorrhagic stroke. Handbook of Clinical Neurology. 2021; 176:229-48.
[5] Tan KS, Venketasubramanian N. Stroke burden in Malaysia. Cerebrovascular Diseases Extra. 2022; 12(2):58-62.
[6] Simonds GR. Recognizing hemorrhagic strokes. MedMastery. 2021.
[7] https://www.mayoclinic.org/diseases-conditions/stroke/symptoms-causes/syc-20350113. Accessed 22 October 2024.
[8] Zainuddin A. Malaysian stroke patients take average seven hours for hospital arrival. CodeBlue. 2022.
[9] Khandelwal A, Sarma K, Hussain M, Dikshit P, Baidya DK. Acute ischemic stroke and the golden hour: critical updates. Journal of Neurosciences in Rural Practice. 2025; 16(2):152-9.
[10] Collen D, Stump DC, Gold HK. Thrombolytic therapy. Annual Review of Medicine. 1988; 39:405-23.
[11] Kandaya S, Abdullah AR, Mohd SN, Azman IH, Shair EF, Ali NH. Analysis of early stroke diagnosis based on brain magnetic resonance imaging using machine learning. Journal of Advanced Research in Applied Sciences and Engineering Technology. 2023; 32(3):241-55.
[12] Cetinoglu YK, Koska IO, Uluc ME, Gelal MF. Detection and vascular territorial classification of stroke on diffusion-weighted MRI by deep learning. European Journal of Radiology. 2021; 145:110050.
[13] Gheibi Y, Shirini K, Razavi SN, Farhoudi M, Samad-soltani T. CNN-Res: deep learning framework for segmentation of acute ischemic stroke lesions on multimodal MRI images. BMC Medical Informatics and Decision Making. 2023; 23(1):1-14.
[14] Azman IH, Saad NM, Abdullah AR, Hamzah RA, Samsudin A. Automated cad system for early stroke diagnosis. International Journal of Advanced Computer Science and Applications. 2023; 14(8):77-83.
[15] Mokli Y, Pfaff J, Dos SDP, Herweh C, Nagel S. Computer-aided imaging analysis in acute ischemic stroke–background and clinical applications. Neurological Research and Practice. 2019; 1(1):1-13.
[16] Keerthana S, Sathiyakumari K. Brain stroke segmentation using fuzzy C-means clustering. International Journal of Computer Applications. 2016; 154(4):26-30.
[17] Subudhi A, Dash P, Mohapatra M, Tan RS, Acharya UR, Sabut S. Application of machine learning techniques for characterization of ischemic stroke with MRI images: a review. Diagnostics. 2022; 12(10):1-19.
[18] Thiyagarajan SK, Murugan K. An exploration enhanced dynamic arithmetic optimization based modified fuzzy clustering framework for ischemic stroke lesion segmentation. International Journal of Information Technology. 2023; 15(8):4389-401.
[19] Marques AG, Souza LF, Dos SMA, Da CNJJ, Sarmento RM, Lucena IB, et al. Automatic segmentation of hemorrhagic stroke on brain CT images using convolutional neural networks through fine-tuning. In international joint conference on neural networks (IJCNN) 2022 (pp. 1-8). IEEE.
[20] Mandle AK, Sahu SP, Gupta G. Brain tumor segmentation and classification in MRI using clustering and kernel-based SVM. Biomedical and Pharmacology Journal. 2022; 15(2):699-716.
[21] Rukmawan SH, Aszhari FR, Rustam Z, Pandelaki J. Cerebral infarction classification using the k-nearest neighbor and naive bayes classifier. In journal of

physics: conference series 2021 (pp. 1-7). IOP Publishing.
[22] Gao H, Bian Y, Cheng G, Yu H, Cao Y, Zhang H, et al. Identifying patients with acute ischemic stroke within a 6-h window for the treatment of endovascular thrombectomy using deep learning and perfusion imaging. Frontiers in Medicine. 2023; 10:1-8.
[23] Nazari-farsani S, Yu Y, Armindo RD, Lansberg M, Liebeskind DS, Albers G, et al. Predicting final ischemic stroke lesions from initial diffusion-weighted images using a deep neural network. NeuroImage: Clinical. 2023; 37:1-10.
[24] Aytaç E, Gönen M, Tatli S, Balgetir F, Dogan S, Tuncer T. Large vessel occlusion detection by non-contrast CT using artificial ıntelligence. Neurological Sciences. 2024; 45(9):4391-7.
[25] Garcia-salgado BP, Almaraz-damian JA, Cervantes-chavarria O, Ponomaryov V, Reyes-reyes R, Cruz-ramos C, et al. Enhanced ischemic stroke lesion segmentation in mri using attention u-net with generalized dice focal loss. Applied Sciences. 2024; 14(18):1-26.
[26] Ghaderi S, Mohammadi S, Ghaderi K, Kiasat F, Mohammadi M. Marker-controlled watershed algorithm and fuzzy C-means clustering machine learning: automated segmentation of glioblastoma from MRI images in a case series. Annals of Medicine and Surgery. 2024; 86(3):1460-75.
[27] İnce S, Kunduracioglu I, Bayram B, Pacal I. U-net-based models for precise brain stroke segmentation. Chaos Theory and Applications. 2025;7(1):50-60.
[28] Inna A, Selvakumar S, Khan A, Brijesh SG, Gurrala N, Pragnathmik. Brain stroke detection and classification system: a hybrid approach using deep learning techniques. Journal of Information Systems Engineering and Management. 2025; 10(40s):90-105.
[29] Batra A, Kaushik G. SECTUBIM: automatic segmentation and classification of tumeric brain MRI images using FHS (FCM, HWT and SVM). International Journal of Engineering Science and Computing. 2017; 7(6):13190-4.
[30] Saad NM, Abdullah AR, Azman IH, Noor NS. A review on image classification techniques for MRI brain stroke lesion. Journal of Advanced Research in Applied Sciences and Engineering Technology. 2024; 40(2):62-73.
[31] Su H, Wang F, Zhang L, Li G. Fuzzy clustering algorithm-segmented MRI images in analysis of effects of mental imagery on neurorehabilitation of stroke patients. Scientific Programming. 2021; 2021(1):1-10.
[32] Feng B, Zhou H, Feng J, Chen Y, Liu Y, Yu T, et al. Active contour model of breast cancer DCE-MRI segmentation with an extreme learning machine and a fuzzy C-means cluster. IET Image Processing. 2022; 16(11):2947-58.
[33] https://naemazam.medium.com/support-vector-machine-svm-in-machine-learning-66d0f6688ca5. Accessed 16 April 2024.
[34] Siddiqi MH, Azad M, Alhwaiti Y. An enhanced machine learning approach for brain MRI classification. Diagnostics. 2022; 12(11):1-19.
[35] Yu Y, Heit JJ, Zaharchuk G. Improving ischemic stroke care with MRI and deep learning artificial intelligence. Topics in Magnetic Resonance Imaging. 2021; 30(4):187-95.
[36] Miyamoto N, Ueno Y, Yamashiro K, Hira K, Kijima C, Kitora N, et al. Stroke classification and treatment support system artificial intelligence for usefulness of stroke diagnosis. Frontiers in Neurology. 2023; 14:1-12.
[37] Eshmawi AA, Khayyat M, Algarni AD, Hilali-jaghdam I. An ensemble of deep learning enabled brain stroke classification model in magnetic resonance images. Journal of Healthcare Engineering. 2022; 2022(1):1-10.
[38] Liu CF, Hsu J, Xu X, Ramachandran S, Wang V, Miller MI, et al. Deep learning-based detection and segmentation of diffusion abnormalities in acute ischemic stroke. Communications Medicine. 2021; 1(1):1-18.
[39] Zhu G, Chen H, Jiang B, Chen F, Xie Y, Wintermark M. Application of deep learning to ischemic and hemorrhagic stroke computed tomography and magnetic resonance imaging. In seminars in ultrasound, CT and MRI 2022 (pp. 147-52). WB Saunders.
[40] Saeedi S, Rezayi S, Keshavarz H, SR NK. MRI-based brain tumor detection using convolutional deep learning methods and chosen machine learning techniques. BMC Medical Informatics and Decision Making. 2023; 23(1):1-17.
[41] Firmansyah MR, Astuti YP. Stroke classification comparison with KNN through standardization and normalization techniques. Advance Sustainable Science, Engineering and Technology. 2024; 6(1):1-8.
[42] https://prof-frenzel.medium.com/dear-friends-854ba66361d7. Accessed 22 October 2024.
[43] Omarov B, Tursynova A, Postolache O, Gamry K, Batyrbekov A, Aldeshov S, et al. Modified UNet model for brain stroke lesion segmentation on computed tomography images. Computers, Materials & Continua. 2022; 71(3):1-17.

Izzatul Husna Azman is currently pursuing her MSc in Electrical Engineering at the Faculty of Electrical Technology and Engineering, Universiti Teknikal Malaysia Melaka (UTeM), Malaysia. Her research focuses on stroke MRI segmentation and classification systems using machine learning and deep learning. She is actively contributing to the development of innovative solutions in medical imaging applications.
Email: husnacino@gmail.com

Norhashimah Mohd Saad is an Associate Professor and Senior Lecturer at the Faculty of Electronics and Computer Technology and Engineering, Universiti Teknikal Malaysia Melaka (UTeM), Malaysia. She specializes in stroke medical image segmentation and classification, utilizing machine learning and deep learning techniques. She has published extensively in international journals and conferences, advancing research in medical imaging and artificial intelligence.
Email: norhashimah@utem.edu.my

Abdul Rahim Abdullah is an Associate Professor and Senior Lecturer at the Faculty of Electrical Technology and Engineering, Universiti Teknikal Malaysia Melaka (UTeM), Malaysia. His expertise lies in advanced digital signal processing, rehabilitation engineering, assistive technology, and power electronics and drive. He has authored numerous papers and contributed significantly to the development of innovative engineering solutions in his areas of specialization.
Email: abdulr@utem.edu.my

Rostam Affendi Hamzah is a Senior Lecturer at the Faculty of Electrical Technology and Engineering, Universiti Teknikal Malaysia Melaka (UTeM), Malaysia. His research interests span stereo vision systems, medical imaging, signal processing, and autonomous systems. He has led and contributed to various funded projects, including stereo vision for 3D pose estimation, UAV collision avoidance, and precision agriculture systems. Dr. Rostam has published extensively in journals and conferences and is actively involved in interdisciplinary research initiatives.
Email: rostamaffendi@utem.edu.my

Ahmad Sobri bin Muda is a Professor (Perubatan) at the Faculty of Medicine and Health Sciences, Universiti Putra Malaysia (UPM). As a medical expert, his research integrates medical diagnostics with computational intelligence, focusing on medical image processing, machine learning, and artificial intelligence. He is a key contributor to the Padiverse and PadiMedikal initiatives, providing valuable datasets and pioneering efforts in stroke classification systems. Prof. Ahmad Sobri has a strong record of publications, including works on brain stroke imaging and clinical trials, and he supervises postgraduate students in medical diagnostics and computational intelligence. His contributions significantly advance research and practice in medical imaging and healthcare technology
Email: asobri@upm.edu.my

Farzanah Atikah Yamba is a research officer at PadiMedikal, a key collaborator in providing datasets and resources for medical image analysis. Her expertise includes managing and curating datasets for stroke classification systems and assisting in integrating data-driven solutions into medical diagnostics. She plays a vital role in ensuring the availability of high-quality data for advancing research in medical imaging and artificial intelligence.
Email: farzanah@padimedikal.com

Appendix I

S. No.	Abbreviation	Description
1	AC	Active Contour
2	AUC	Area Under the Curve
3	BNN	Bilayered Neural Network
4	CAD	Computer-Aided Diagnosis
5	CNN	Convolutional Neural Network
6	CT	Computed Tomography
7	CTP	CT Perfusion
8	DC	Dice Coefficient
9	DWI	Diffusion-Weighted Imaging
10	FCM	Fuzzy C-Means
11	FCMAC	Fuzzy C-Means with Active Contour
12	FLAIR	Fluid-Attenuated Inversion Recovery
13	FPR	False Positive Rate
14	GUI	Graphical User Interface
15	GPUs	Graphics Processing Units
16	HKL	Hospital Kuala Lumpur
17	HUPM	Hospital Hospital Sultan Abdul Aziz Shah, UPM Serdang, Selangor, Malaysia
18	IoU	Intersection Over Union
19	ISLES	Ischemic Stroke Lesion Segmentation
20	KNN	k-Nearest Neighbor
21	MLP	Multilayer Perceptron
22	MCWS	Marker-Controlled Watershed Segmentation
23	MRI	Magnetic Resonance Imaging
24	NB	Naïve Bayes
25	PWI	Perfusion Weighted Imaging
26	RF	Random Forest
27	ResNet-50	Residual Network, 50 Layers
28	R-CNN	Region-Based Convolutional Neural Network
29	ROI	Region of Interest
30	SVM	Support Vector Machine
31	TPR	True Positive Rate
32	VGG	Visual Geometry Group