Journal of Electronics, Electromedical Engineering, and Medical Informatics
Homepage: jeeemi.org; Vol. 7, No. 3, July 2025, pp: 850-868 e-ISSN: 2656-8632

RESEARCH ARTICLE OPEN ACCESS

Advanced Deep Learning for Stroke Classification Using Multi-Slice CT Image Analysis

Fouzi Lezzar[1], Seif Eddine Mili[2]

[1] LIRE Laboratory, Faculty of New Technologies of Information and Communication, University of Abdelhamid MEHRI Constantine 2, Constantine, Algeria
[2] Ecole Normale Supérieure, Engineering Laboratory for Complex Systems (LISCO), Constantine, Algeria

Corresponding author: Fouzi Lezzar (e-mail: fouzi.lezzar@univ-constantine2.dz), **Author(s) Email:** Seif Eddine Mili (mili.seifeddine@ensc.dz)

Abstract Brain stroke is a leading cause of mortality and disability globally, necessitating rapid and accurate diagnosis for timely intervention. While Computed Tomography (CT) imaging is the gold standard for stroke detection, manual interpretation is time-consuming, prone to error, and subject to inter-observer variability. Although deep learning models have shown promise in automating stroke detection, many rely on 2D analysis, ignore 3D spatial relationships, or require labour-intensive slice-level annotations, which limits their scalability and clinical applicability. To address these challenges, we propose MedHybridNet, a novel hybrid deep learning architecture that integrates convolutional neural networks (CNNs) for local feature extraction with Transformer-based modules to model global contextual dependencies across volumetric brain scans. Our main contribution is twofold: (1) the SliceAttention mechanism, which dynamically identifies diagnostically relevant slices using only patient-level labels, eliminating the need for costly slice-level annotations while enhancing interpretability through attention maps and Grad-CAM visualizations; and (2) a cGAN-based augmentation strategy that generates high-quality, pathology-informed synthetic CT slices to overcome data scarcity and class imbalance. The framework processes complete 3D brain volumes, leveraging both CNNs and Transformers in a dual-path design, and incorporates hierarchical attention for refined feature selection and classification. Evaluated via patient-wise 5-fold cross-validation on a real-world dataset of 2501 CT scans from 82 patients, MedHybridNet achieves an accuracy of 98.31%, outperforming existing methods under weak supervision. These results demonstrate its robustness, generalization capability, and superior interpretability. By combining architectural innovation with clinically relevant design choices, MedHybridNet advances the integration of Artificial Intelligence (AI) into real-world stroke care, offering a scalable, accurate, and explainable solution that can significantly improve diagnostic efficiency and patient outcomes in routine clinical practice.

Keywords: Stroke Detection, Medical Imaging, Deep Learning, Attention Mechanism, cGAN.

I. Introduction

The accurate and timely diagnosis of brain stroke remains one of the most pressing challenges in modern healthcare, with profound implications for patient survival, long-term recovery, and the efficient use of medical resources. As a leading cause of mortality and disability worldwide [1], stroke demands rapid and precise intervention to minimize neurological damage and improve clinical outcomes [2], [3]. Computed Tomography (CT) imaging has emerged as the gold standard for stroke assessment due to its rapid acquisition, widespread availability, and high sensitivity in detecting acute stroke [4], [5], [6]. However, manual interpretation of CT scans is a complex and time-intensive task that relies heavily on radiologists' expertise [5], [7], [8]. This dependence often leads to variability and error, particularly in under-resourced or high-workload settings [9], [10].

Recent advancements in artificial intelligence (AI), intense learning, have shown immense potential in automating medical image analysis and augmenting diagnostic decision-making [11], [12], [13], [14], [15]. Despite these strides, many existing AI-driven systems for stroke detection face significant limitations [16], [17]. For instance, several models fail to account for the 3D spatial relationships within CT volumes or lack mechanisms to provide interpretable insights, which is

Manuscript received 10 May 2025; Revised 30 June 2025; Accepted 1 July 2025; Available online 5 July 2025
Digital Object Identifier (**DOI**): https://doi.org/10.35882/jeeemi.v7i3.947

a critical requirement for clinical adoption, where transparency and trust are paramount. Additionally, many frameworks rely on labour-intensive slice-level annotations, which hinder scalability and generalizability.

In this study, we propose a novel hybrid deep learning architecture, MedHybridNet, designed explicitly for stroke detection in 3D brain CT imaging. Our approach addresses key limitations of current methods [4], [6], [10], [13], [17] by integrating convolutional neural networks (CNNs) for high-resolution feature extraction with Transformer-based modules for global contextual reasoning. This dual-path design enables the model to capture both fine-grained pathological patterns and broader anatomical dependencies across slices. A key innovation is the SliceAttention mechanism, which dynamically identifies and prioritizes diagnostically relevant slices during training using only patient-level labels (e.g., stroke vs. regular). By eliminating the need for manual slice-level annotation (especially of stroke regions), SliceAttention reduces labelling effort while enhancing interpretability through attention-weight maps and Grad-CAM visualizations that align with clinical observations.

To address the challenges of data scarcity and class imbalance, we introduce a comprehensive data augmentation approach based on a conditional Generative Adversarial Network (cGAN). The proposed cGAN generates realistic synthetic CT slices conditioned on diagnostic labels, ensuring anatomical plausibility while increasing the diversity of the training data. This strategy significantly enhances model robustness and generalization performance. Our main contributions are summarized as follows:

1. A hybrid CNN-Transformer architecture for volumetric medical imaging that models both local features and global context. Skip connections enhance gradient flow and feature fusion, while hierarchical attention at the region- and slice-level boosts diagnostic accuracy.
2. An attention mechanism that identifies diagnostically relevant slices without requiring slice-level annotations, reducing labelling effort and improving model interpretability.
3. A novel application of conditional GANs to generate pathology-informed synthetic CT slices, improving data diversity and model robustness.

By bridging the gap between technical innovation and clinical utility, our work presents a scalable, interpretable, and highly effective solution for automated stroke detection. MedHybridNet not only achieves competitive diagnostic accuracy but also provides clinicians with actionable and explainable outputs, thereby advancing the integration of AI into routine stroke care and contributing meaningfully to improved patient outcomes.

The remainder of this paper is organized as follows: Section 2 presents a comprehensive review of related work in fall detection, highlighting existing methodologies and their limitations. Section 3 details the proposed approach, including dataset description, preprocessing steps, and the architecture of the developed model. Section 4 reports the experimental results, while Section 5 provides an ablation study to evaluate the contribution of each component of the model. Section 6 includes a thorough discussion of the findings and comparison with state-of-the-art methods. Finally, Section 7 concludes the paper and outlines potential directions for future research.

II. Related work

Recent studies have explored various machine learning and deep learning techniques to enhance brain stroke detection, classification, and segmentation using medical imaging modalities, including computed tomography (CT), magnetic resonance imaging (MRI), and microwave imaging. These efforts aim to enhance diagnostic accuracy, reduce processing time, and support early intervention.

Authors in [18] proposed a few-shot learning framework integrated with a self-attention-based CNN to segment ischemic stroke lesions in MRI scans. Their method addresses the challenge of limited labeled data in the medical domain by focusing on lesion-containing slices and employing early fusion of FLAIR and DWI modalities. The system achieves enhanced segmentation accuracy, with a Dice score of 0.68 on the ISLES 2015 SSIS dataset, outperforming several state-of-the-art approaches. Study [19] proposed a soft voting ensemble model that combines Random Forest, Extremely Randomized Trees, and Histogram-Based Gradient Boosting for predicting brain strokes. This approach aggregates class probabilities from individual classifiers, either uniformly or weighted by validation performance, enhancing prediction robustness through ensemble learning.

Another study [20] focused on improving early stroke detection using CT images through a hybrid artificial intelligence framework. It combined a newly designed convolutional neural network architecture, termed OzNet, with traditional machine learning models, including Decision Trees (DT), k-Nearest Neighbours (KNN), Linear Discriminant Analysis (LDA), Naïve Bayes (NB), and Support Vector Machines (SVM). A different approach was taken in [21], where researchers aimed to automate the classification of collateral circulation patterns in ischemic stroke using cone-beam CT (CBCT) images. They employed the VGG11 architecture with an

Manuscript received 10 May 2025; Revised 30 June 2025; Accepted 1 July 2025; Available online 5 July 2025
Digital Object Identifier (DOI): https://doi.org/10.35882/jeeemi.v7i3.947

augmented dataset and preprocessing steps to standardize pixel values. Despite achieving a modest accuracy of 58.32% and an F1 score of 62.10%, this work represents a promising step toward reducing diagnostic delays in stroke management. In [22], deep learning models were applied to classify unenhanced brain CT images such as normal, haemorrhage, infarction, and others. Convolutional neural networks, including CNN-2, VGG-16, and ResNet-50, were evaluated using transfer learning with varying data sizes, batch sizes, and optimizers. While both ResNet-50 and CNN-2 achieved a high accuracy of 98.72%, ResNet-50 came at the cost of increased computational complexity. Study [23] proposed a fast method for generating large datasets for brain stroke classification using microwave imaging. Based on the distorted Born approximation and linearization of the scattering operator, the technique significantly reduced computation time. A classification pipeline involving Support Vector Machines, Multilayer Perceptrons, and k-Nearest Neighbours was tested on simulated datasets with variations in antenna structure and amplitude-only data, showing strong potential for real-time stroke detection. In [24], structured clinical data was used for early and accurate brain stroke prediction via a Multi-Layer Perceptron (MLP) classifier combined with three optimization algorithms: AdaMax, RMSProp, and Adadelta. Among these, RMSProp yielded the best results, achieving 95.8% training accuracy and 94.9% testing accuracy. The integration of multiple optimizers was highlighted as a novel strategy to enhance prediction reliability. Study [25] investigated the prediction of infarct extent and location after reperfusion in acute ischemic stroke using only CT Angiography images.

A modified U-Net model incorporating squeeze-and-excitation blocks was trained on 238 patient cases, with ground truth lesions segmented from 24-hour follow-up CT scans. The model achieved a Dice score of 0.37 and a volume error of 3.9 mL, indicating room for improvement but also potential in image-based infarct prediction. An early stroke detection system based on CT brain images, genetic algorithms, and a bidirectional long short-term memory (BiLSTM) model was developed in [26]. The genetic algorithm was employed for feature selection, followed by classification using BiLSTM, demonstrating a novel combination of evolutionary and sequential modelling techniques. Study [27] explored a hybrid Vision Transformer (ViT) and Long Short-Term Memory (LSTM) model for stroke detection on CT images. ViT was used to extract visual features, while LSTM captured temporal dependencies. With optimizers such as SGD, RMSProp, Adam, and AdamW, the system achieved up to 96.61% accuracy on Kaggle data, demonstrating the effectiveness of combining vision and sequence modelling. In [28], a modified divergence-based deep neural network (DNN), integrating CNN-Walsh matrix feature extraction and a minimum distance network (MDN) classifier, was applied to 6650 authentic CT images for stroke classification into ischemic, haemorrhagic, or non-stroke types. Notably, the model demonstrated robust results with only minor adjustments to the parameters. Study [29] presented an LSTM Generative Adversarial Network (LSTM-GAN) for generating cerebral blood

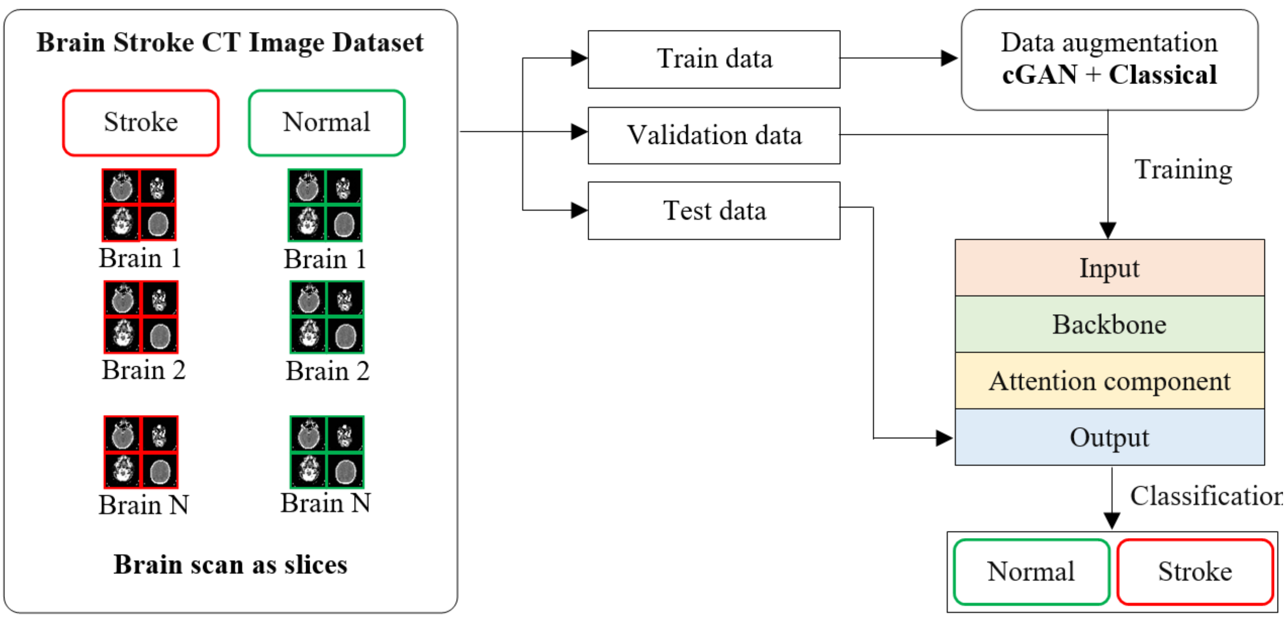


Fig. 1. Proposed architecture for stroke classification using CT images, incorporating data augmentation, a backbone network, attention component, and binary classification.

Manuscript received 10 May 2025; Revised 30 June 2025; Accepted 1 July 2025; Available online 5 July 2025
Digital Object Identifier (**DOI**): https://doi.org/10.35882/jeeemi.v7i3.947

flow (CBF) maps in ischemic stroke diagnosis. By eliminating the need for manual arterial input function (AIF) selection, the model reduced scanning time from 40 to 9 time points, minimizing radiation exposure. Evaluated on the ISLES 2018 dataset, the method achieved an accuracy of 91.4%, offering a more efficient and precise diagnostic approach.

Finally, authors in [30] proposed automated classification and segmentation methods for detecting haemorrhagic and ischemic lesions in non-contrast CT images. Using a U-Net model for segmentation, the study reported a classification precision of 95.06%, along with IoU scores of 92.01% for haemorrhagic and 82.22% for ischemic lesions, confirming the high performance of deep learning in lesion detection.

III. Method

Our approach (Fig. 1) to brain stroke classification leverages a deep learning pipeline that processes 3D CT scan images. Unlike methods that analyse only selected slices [4], [5], [6], [31], [32], [33], our model processes all slices of the brain, providing a comprehensive view and incorporating richer information for classification. The dataset is split into training, validation, and test sets, ensuring that the validation and test sets remain unaltered without data augmentation. To enhance model generalization, we apply a robust data augmentation strategy using a conditional Generative Adversarial Network (cGAN). The augmented data is then used to train a hybrid deep learning model consisting of an input layer, a backbone network, and an attention mechanism (SliceAttention) that improves feature extraction and stroke detection. Finally, the trained model predicts whether a given brain scan corresponds to a normal or stroke case.

A. Dataset

The dataset [34] was collected from Lady Reading Hospital in Peshawar, Khyber Pakhtunkhwa, Pakistan [35]. The authors conducted a descriptive study using this dataset to identify risk factors for brain stroke, including age, gender, smoking, diabetes mellitus, and hypertension. The dataset consists of brain imaging scans Fig. 2 from two classes: Normal and Stroke, with a total of 2501 images across 82 unique individuals. The Normal class contains 1551 images from 51 subjects, with each subject contributing between 23 and 40 slices, averaging 30.41 slices per person. The Stroke class includes 950 images from 31 subjects, with slice counts per individual ranging from 19 to 36, averaging 30.65 slices per person. This distribution reflects a relatively balanced number of slices per person across both classes, supporting consistent training and evaluation for stroke detection in brain imaging.

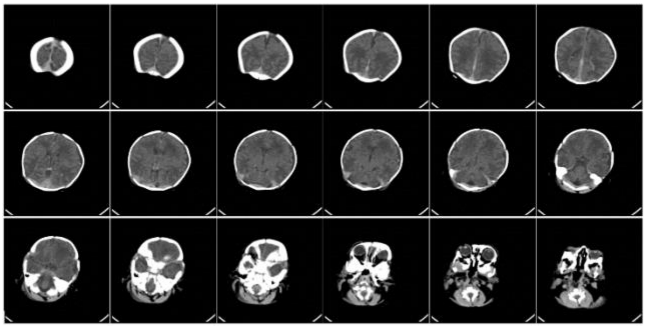


Fig. 2. Visual samples from the original brain CT dataset used for model training.

To ensure consistent and effective input representation for deep learning-based stroke classification, a standardized preprocessing pipeline was applied to all brain CT scans before model training and inference. Each patient's CT scan was acquired as a full 3D volume and resampled to ensure uniformity across subjects. To standardize input dimensions, brain volumes containing more than 30 axial slices were symmetrically truncated from the top and bottom to retain the central brain region, where stroke-related patterns are most commonly observed. For volumes with fewer than 30 slices, we applied linear interpolation along the z-axis (i.e., between adjacent slices) to generate additional slices. This process estimates the content of intermediate slices based on their neighbouring sections, effectively increasing slice count while maintaining spatial coherence and preserving anatomical continuity across the volume.

Each axial slice was independently resized from its original resolution of 650×650 pixels to 224×224 pixels using bicubic interpolation, which better preserves fine anatomical structures compared to bilinear or nearest-neighbour methods. Hounsfield Unit (HU) values were clipped to the range [−100, 100] to focus on soft tissue contrast relevant to stroke detection, followed by min-max normalization to scale pixel intensities to the range [−1, 1], using mean (μ) and standard deviation (σ) computed over the training set. Since the backbone models (e.g., VGG19, ResNet50) were pretrained on RGB images from ImageNet, grayscale CT slices were converted to a three-channel format by replicating intensity values across all channels, thereby enabling effective transfer learning. The final input shape for each brain volume was standardized to input $\in R^{(nSlices=30,\ 224\times224\times3)}$. These preprocessing steps enhance consistency across patients, reduce scanner-related variability, and preserve diagnostic relevance, particularly for identifying hypodensity patterns associated with stroke cases.

Manuscript received 10 May 2025; Revised 30 June 2025; Accepted 1 July 2025; Available online 5 July 2025
Digital Object Identifier (**DOI**): https://doi.org/10.35882/jeeemi.v7i3.947

B. Data augmentation

Data augmentation is critical for addressing the challenges posed by limited and imbalanced medical imaging datasets in stroke detection. It enhances data diversity, mitigates overfitting, and improves model generalization.

Table 1. Detailed architecture of the 3D generator showing input shape, layers, and activation functions used.

Layer type	Input shape + Activation
Input (Noise + Label)	(z_dim + n_classes=2) (-)
Dense	(z_dim + n_classes=2) (RELU)
Reshape	(4×4×4×512) (-)
3D Transposed Conv_1	(4, 4, 4, 512) (RELU)
BatchNorm	(8, 8, 8, 256) (-)
3D Transposed Conv_2	(8, 8, 8, 256) (RELU)
BatchNorm	(16, 16, 16, 128) (-)
3D Transposed Conv_3	(16, 16, 16, 128) (RELU)
BatchNorm	(32, 32, 32, 64) (-)
3D Transposed Conv_4	(32, 32, 32, 64) (RELU)
BatchNorm	(64, 64, 64, 32) (-)
3D Transposed Conv_5	(64, 64, 64, 32) (Tanh)
Resize / Interpolation	(128, 128, 128, 1) → (650, 650, 30, 1)
Output	(650, 650, N=30, 1)

To address dataset imbalance and enhance model generalization, we employ a 3D conditional Generative Adversarial Network (cGAN) specifically designed for volumetric brain CT scans. The generator (Table 1) begins with a latent noise vector $z \in R^{256}$, concatenated with a one-hot encoded class label ("normal" = 1, "stroke" = 0]) at the input layer to guide class-specific image generation. This combined input is passed through dense and reshape layers before being progressively upsampled via 3D transposed convolutional blocks, each followed by batch normalization and ReLU activation. The final output uses the tanh activation to generate high-resolution 3D CT volumes of shape (650×650×30×1), matching the standardized size of real scans.

The discriminator (Table 2) evaluates realism during the training process using 3D convolutional layers with LeakyReLU and batch normalization, gradually reducing the spatial resolution until it reaches a flattened representation. A dropout layer (rate=0.3) prevents overfitting, and a final dense layer with sigmoid activation performs binary classification (real vs. fake), conditioned on class labels.

Table 2. Detailed architecture of the 3D discriminator showing input shape, layers, and activation functions used.

Layer type	Input shape+Activation
Input (Volume + Label)	(650, 650, N=30, 1 + n_classes=2)(-)
3D Conv_1	(650, 650, N=30, 1 + n_classes=2) (LeakyReLU)
BatchNorm	(325, 325, N/2, 32)(-)
3D Conv_2	(325, 325, N/2, 32) (LeakyReLU)
BatchNorm	(163, 163, N/4, 64)(-)
3D Conv_3	(163, 163, N/4,64)(LeakyReLU)
BatchNorm	(82, 82, N/8, 128)(-)
Flatten	(82, 82, N/8, 128)(-)
Dropout = 0.3	(Flattened vector)(-)
Dense	(Flattened vector)(Sigmoid)
Output	1

For stable and realistic image synthesis, we adopt the Wasserstein GAN with gradient penalty (WGAN-GP) (Eq. (1) [36]):

$$L_{WGAN\text{-}GP} = E_{x\sim P_{real}}[D(x)] - E_{z\sim P_{noise}}[D(G(z))] + \lambda_{gp} \cdot E_{\hat{x}\sim P_{\hat{x}}}[(\|\nabla_{\hat{x}} D(\hat{x})\|_2 - 1)^2] \quad (1)$$

In Eq. (1), the first term, $E_{x\sim P_{real}}[D(x)]$, encourages the discriminator D to assign high scores to real samples x drawn from the data distribution P_{real}. The second term, $E_{z\sim P_{noise}}[D(G(z))]$, penalizes the discriminator for assigning high scores to generated samples $G(z)$, where z is sampled from a noise prior P_{noise} . The third term, $\lambda_{gp} \cdot E_{\hat{x}\sim P_{\hat{x}}}[(\|\nabla_{\hat{x}} D(\hat{x})\|_2 - 1)^2]$, imposes a gradient penalty to enforce the 1-Lipschitz continuity condition by penalizing deviations of the gradient norm from 1 for inputs $\hat{x}$ interpolated between real and generated samples.

An optional coherence loss (Eq. (2) [37]) ensures smooth anatomical transitions between adjacent slices:

$$L_{coherence} \ \frac{1}{N-1} \sum_{i=0}^{N-1} \|G(z_i) - G(z_{i+1})\|_p \quad (2)$$

where N is the total number of slices in a volumetric sequence, $G(z_i)$ and $G(z_{i+1})$ denote the generator outputs corresponding to consecutive latent codes z_i and z_{i+1}. $\|\cdot\|_p$ represents the L_p norm, set p=2. This loss encourages the generated images to change smoothly across adjacent slices by penalizing significant differences between successive outputs.

Manuscript received 10 May 2025; Revised 30 June 2025; Accepted 1 July 2025; Available online 5 July 2025
Digital Object Identifier (**DOI**): https://doi.org/10.35882/jeeemi.v7i3.947

The total loss (Eq. (3) [38]) combines these components:

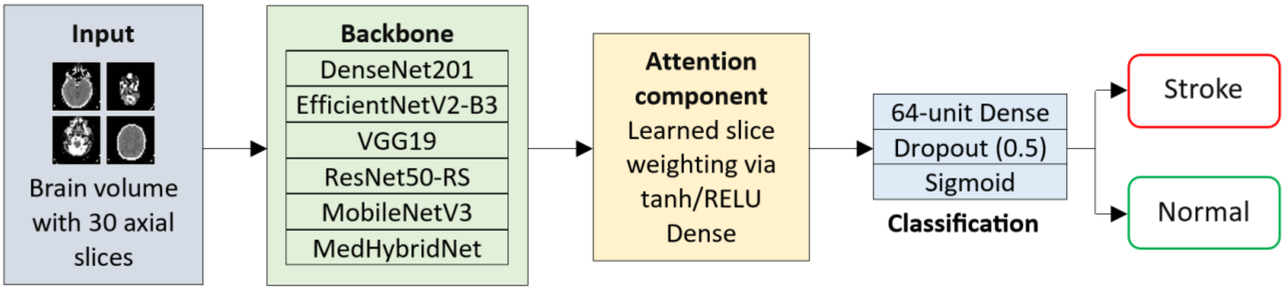


Fig. 4. Deep learning pipeline for stroke classification using CT volumes and attention-based slice weighting.

$$L_{total} = L_{WGAN\text{-}GP} + \lambda_{coh} \cdot L_{coherence} \quad (3)$$

In this third equation, $L_{WGAN\text{-}GP}$ drives the generator to produce visually plausible outputs. At the same time, the additional term $\lambda_{coh} \cdot L_{coherence}$ acts as a regularizer [38] that promotes continuity between sequentially generated slices. The hyperparameter λ_{coh}, set to 10 in our experiments, provides control over the strength of the coherence constraint, enabling a balance between individual image quality and inter-slice smoothness.

The training process follows an alternating update scheme, where the discriminator is updated five times for each generator update. During training, real and synthetic samples are carefully balanced to ensure stable learning. Additionally, a gradient penalty is applied every ten steps on interpolated samples to enforce the Lipschitz constraint.

The proposed 3D conditional GAN (cGAN) was implemented and trained using Google Colab Pro, which provided access to a Tesla T4 or A100 GPU, enabling efficient processing of high-resolution volumetric brain scans. The model was trained for 230 epochs using the Adam optimizer, with learning rates set to lrG=0.0001 for the generator and lrD=0.0004 for the discriminator, along with momentum terms β1=0 and β2=0.99 to ensure stable convergence. A batch size of 4 volumetric samples per iteration was used during training to maintain memory efficiency while preserving gradient stability.

These design choices enable the cGAN to generate high-quality, class-conditioned synthetic CT slices, significantly enhancing dataset diversity and improving model robustness under weak supervision. Fig. 3 presents a subset of brain slices generated by the proposed cGAN. These synthetic slices closely resemble real brain scans, as confirmed by both qualitative and quantitative assessments. The generated images preserve essential anatomical structures and patterns present in the original data. The high visual and statistical similarity between synthetic and real slices demonstrates the effectiveness of the cGAN in generating realistic augmentations for brain stroke classification.

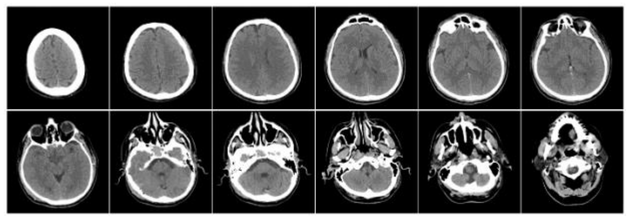


Fig. 3. Representative axial slices from a synthetic brain volume generated by the conditional GAN, illustrating spatial consistency and anatomical plausibility.

C. Backbone

Each patient's CT scan is processed as a sequence of 30 axial brain slices (Fig. 4), with slice-wise feature extraction performed using established convolutional backbones such as VGG19, ResNet50, or EfficientNetV2-B3. The earlier layers of these models are initially frozen during training to leverage their strong low-level feature detection capabilities. In contrast, the final layers are unfrozen to allow fine-tuning on medical imaging data and adapt the model to domain-specific characteristics. This strategy balances transfer learning with specialization, improving classification performance while maintaining generalization across volumetric brain scans.

GlobalAveragePooling2D is applied per slice to condense spatial dimensions into compact feature vectors (e.g., 512-D for VGG19 or 2048-D for ResNet50), enabling efficient downstream processing. To enhance interpretability and classification accuracy, a custom SliceAttention mechanism dynamically computes softmax-normalized weights for each slice through a two-step process. First, a non-linear projection with tanh activation is applied to promote gradient stability. This is followed by a linear scoring step and softmax-based attention normalization, allowing the model to prioritize diagnostically relevant slices.

Manuscript received 10 May 2025; Revised 30 June 2025; Accepted 1 July 2025; Available online 5 July 2025
Digital Object Identifier (**DOI**): https://doi.org/10.35882/jeeemi.v7i3.947

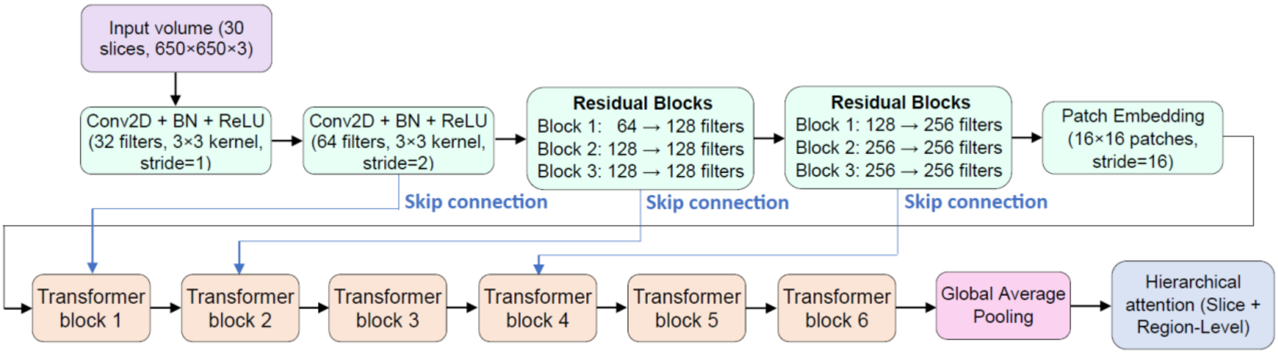


Fig. 5. Proposed MedHybridNet model structure for analysing brain CT volumes in stroke detection.

A multi-head variant is employed to allow the model to focus on both localized lesions and broader contextual patterns. This learned attention mechanism enables the system to prioritize diagnostically meaningful slices without requiring manual slice-level annotations. The resulting attention-weighted feature representation is passed through a fully connected classifier comprising a 64-unit dense layer with L2 regularization to prevent overfitting, followed by dropout at a rate of 0.5 and a final sigmoid activation, which outputs the probability of stroke presence.

We propose a novel advanced backbone architecture, MedHybridNet (Table 3 and Fig. 5), specifically designed to process 3D brain CT scans as sequences of 2D axial slices, enabling both localized feature extraction and global contextual reasoning:

1. MedHybridNet CNN path

Each slice is first passed through two initial convolutional layers:

- Conv2D + BatchNorm + ReLU: 32 filters, 3×3 kernel, stride=1
- Conv2D + BatchNorm + ReLU: 64 filters, 3×3 kernel, stride=2

Table 3. Architecture of the MedHybridNet backbone combining per-slice CNN and Transformer paths with hierarchical attention mechanisms.

Layer type	Configuration	Output
Input	[n_slices=30, 650, 650, 3]	(n_slices, 650, 650, 3)
Per-Slice CNN Path		
Conv2D + BN + ReLU	32 filters, 3×3 kernel, stride=1	(n_slices, 650, 650, 32)
Conv2D + BN + ReLU	64 filters, 3×3 kernel, stride=2	(n_slices, 325, 325, 64)
Residual Block (×3)	64 → 128 filters, 3×3 kernels	(n_slices, 325, 325, 128)
Residual Block (×3)	128 → 256 filters, 3×3 kernels	(n_slices, 325, 325, 256)
Per-Slice Transformer Path		
Patch Embedding	16×16 patches → 768D embedding (stride=16 Conv2D)	(n_slices, 21×21, 768)
Transformer Encoder (×6)	8-head multi-head self-attention, MLP (2048→768 units), GELU activation	(n_slices, 441, 768)
Global Average Pooling	Reduce spatial dimensions	(n_slices, 768)
Skip Connections	CNN-to-Transformer skip connections	(n_slices, 768)
Hierarchical Attention	Multi-head attention at the slice-level and the region-level	(n_slices, 768)

This preprocessing step reduces the spatial dimensions to 325 × 325 pixels, ensuring computational efficiency while preserving diagnostic detail. Subsequently, six residual blocks are organized into two stages. In the first stage, the number of filters increases from 64 to 128, and in the second stage, it is further expanded to 256, enabling progressive feature refinement and deeper representation learning.

Each stage consists of three residual blocks, chosen to provide sufficient depth for hierarchical feature learning while minimizing the risk of overfitting. The number of filters is expanded progressively across stages, following conventional CNN scaling practices such as doubling the filters at each level. Within each block, batch normalization and ReLU activation are employed to enhance gradient flow and ensure stable training. These blocks preserve spatial resolution (325×325) while extracting fine-grained anatomical features, which is particularly important for identifying subtle stroke-related hypodensity patterns in low-contrast regions.

2. MedHybridNet Transformer path

Manuscript received 10 May 2025; Revised 30 June 2025; Accepted 1 July 2025; Available online 5 July 2025
Digital Object Identifier (**DOI**): https://doi.org/10.35882/jeeemi.v7i3.947

Journal of Electronics, Electromedical Engineering, and Medical Informatics
Homepage: jeeemi.org; Vol. 7, No. 3, July 2025, pp: 850-868 e-ISSN: 2656-8632

To transition from local to global modelling, CNN feature maps are projected into 768-dimensional patch embeddings using a stride-16 convolutional layer with a 16×16 patch size, resulting in a sequence of 21×21 = 441 patches per slice. Six Transformer encoder layers then process these embeddings, each containing: 8-head multi-head self-attention for modelling complex inter-slice dependencies and an MLP block (2048 → 768 units) with GELU activation for non-linear transformation.

To improve gradient flow and feature fusion across modalities, skip connections are introduced between corresponding CNN and Transformer layers in MedHybridNet (Fig. 5). These connections transfer feature maps from each CNN residual stage into corresponding Transformer blocks through 1×1 convolutions. This ensures that high-resolution anatomical features are preserved and refined during global contextual modeling. Three strategically placed links were empirically validated based on spatial alignment and performance impact, providing an optimal balance between model accuracy, training stability, and computational efficiency.

Additional connections were tested but found to increase complexity without significant performance gains. This design enhances convergence and ensures that the Transformer path leverages high-resolution CNN features without compromising discriminative patterns due to downsampling or abstraction loss. After the Transformer path, each slice's representation is globally averaged to produce a 768-dimensional feature vector, resulting in a sequence of shape (n_slices = 30, 768). This format preserves slice independence and supports downstream processing via attention mechanisms or classification heads.

3. MedHybridNet hierarchical attention mechanism

In addition to skip connections, a hierarchical attention mechanism operates at two levels: region-level attention identifies diagnostically meaningful areas within individual slices using multi-head self-attention on patch embeddings, and slice-level attention that dynamically prioritizes the most relevant slices across the volume. This dual-level structure enhances diagnostic insight by focusing on both local stroke indicators and global lesion context, supporting robust classification under weak supervision without requiring manual slice-level annotations. For MedHybridNet, the SliceAttention mechanism is part of the hierarchical attention framework, operating at the slice level to dynamically prioritize diagnostically relevant slices. In contrast, other models use it directly after global pooling. In MedHybridNet, skip connections and hierarchical attention work synergistically during feature fusion. Skip connections transfer high-resolution CNN features directly into matching Transformer blocks, preserving spatial detail and improving gradient flow. These features are then refined by hierarchical attention, which operates at two levels: region-level and slice-level (SliceAttention), which dynamically prioritizes the most relevant slices across the volume.

4. MedHybridNet hyperparameter selection

Several key hyperparameters, including dropout rate (for final classification), convolutional kernel size, and attention mechanism dimensions, were determined through empirical validation and established architectural practices (VGG, ResNet, …) during model development. Initial experiments were conducted with varying configurations (e.g., dropout values of 0.3–0.7, kernel sizes of 3×3 vs. 5×5), and performance was evaluated using validation accuracy, training stability, and consistency of interpretability. Based on these trials, final selections were made to ensure optimal convergence and diagnostic relevance. For instance, a dropout rate of 0.5 (for the final classifier) provided the best trade-off between regularization and expressive power, while a hidden dimension of u = d/4 in the attention module improved gradient flow compared to deeper or shallower alternatives. This data-driven approach ensures robustness and supports reproducibility in future implementations. Additionally, a batch size of four brain volumes was selected to balance GPU memory constraints with stable optimization dynamics.

5. MedHybridNet optimization settings

MedHybridNet was trained using the AdamW optimizer (lr = 0.0003, β1 = 0.9, β2 = 0.999), ensuring stable convergence and improved generalization under weak supervision. All backbone models used stratified sampling to maintain balanced batch composition and reduce the risk of overfitting, which is particularly important when working with limited patient-level labels. Stratified sampling significantly enhanced training stability and convergence speed. To optimize classification performance, we incorporate several design choices in the final decision layer:

1. Loss function and class balancing

Training the final classifier is conducted using weighted binary cross-entropy loss to account for class imbalance between stroke and normal cases. Even though we employ a cGAN-based data augmentation strategy that enriches the dataset with realistic synthetic stroke slices, we found it essential to complement this with class weighting during optimization. This ensures that the model remains sensitive to subtle stroke indicators and avoids bias toward the majority class, especially under weak supervision where only patient-level labels are available. To ensure balanced gradient updates during optimization, class weights are computed based on the

Manuscript received 10 May 2025; Revised 30 June 2025; Accepted 1 July 2025; Available online 5 July 2025
Digital Object Identifier (**DOI**): https://doi.org/10.35882/jeeemi.v7i3.947

inverse frequency of each class in the training set (Eq. (4) [39]):

$$W_c = \frac{N}{n_c \times C} \quad (4)$$

where N represents the total number of samples in the training set, n_c the number of samples in class c (stroke or normal), and C the total number of classes. Class weights are computed based on the inverse frequency of each class in the training set and are applied during loss computation to balance gradient updates. In this configuration, the normal class is assigned a weight of 1.00, while the stroke class receives a higher weight of 1.61 to compensate for class imbalance and enhance sensitivity to stroke cases. These fixed weights are integrated into the binary cross-entropy loss function, ensuring fair contribution from both underrepresented stroke patterns and more frequent normal scans.

2. Training stability enhancements

To improve training stability and prevent overfitting, we implemented both early stopping and a stepwise learning rate reduction policy. The early stopping criteria were based on validation accuracy, with a patience parameter of 10 epochs. In other words, training was halted if no improvement in validation performance was observed over 10 consecutive epochs.

For learning rate scheduling, we used the ReduceLROnPlateau method, which reduces the learning rate by a factor of 0.5 if no improvement is observed in the validation loss over five epochs. This ensures smoother convergence after the initial learning phase and helps avoid oscillation around optimal values. These strategies were applied consistently across all backbone models. Training stability enhancements (early stopping, learning rate reduction, stratified sampling, class weighting) are used to both the backbones and the final classifier. Interpretability is further enhanced through Grad-CAM visualizations and attention weight overlays, providing clinicians with transparent insights into the model's decision-making process and highlighting regions of interest within the input scans.

D. SliceAttention algorithm

Each patient's CT scan is processed as a sequence of 30 axial slices, with each slice represented as a feature vector $x_i \in R^d$, where d denotes the embedding dimension (e.g., 512-D for VGG19, 768-D for MedHybridNet). To enable interpretable decision-making under weak supervision, we apply the SliceAttention mechanism, which dynamically computes softmax-normalized weights across slices to highlight diagnostically meaningful regions.

In the SliceAttention mechanism (Algorithm 1), let $X=[x_1, x_2,..., x_{30}]$ represent the input sequence of slice-level feature vectors extracted from a 3D CT scan, where each $x_i \in R^d$ corresponds to the embedding of the i -th axial slice. Each feature vector is first transformed through a non-linear projection layer parameterized by a weight matrix $W_1 \in R^{d\times dh}$ and bias vector $b_1 \in R^{dh}$, producing hidden representations $h_i=\tanh(W_1x_i+b_1)$, where dh=192 denotes the hidden dimension. The value dh=192 corresponds to the dimensionality of the internal feature vector computed for each slice, which the model uses to evaluate its diagnostic relevance. A second linear projection, defined by weight vector $W_2 \in R^{dh\times 1}$ and scalar bias $b_2 \in R$, computes unnormalized attention scores $e_i=W_2h_i+b_2$. These scores are then normalized using the softmax function to yield attention weights $\alpha=\mathrm{softmax}([e_1,...,e_{30}])$, with each $\alpha_i \in [0,1]$ indicating the relative importance of the i-th slice in stroke classification. The final output representation $z \in R^d$ is computed as a weighted sum of all input features: $z=\sum_{i=1}^{30} \alpha_i \cdot x_i$.

Algorithmic 1. Steps of SliceAttention

1. **Input**: A sequence of slice-level feature vectors from all 30 axial slices, denoted as $X=[x_1, x_2, ...,x_{30}]$, where each $x_i \in R^d$ (feature dimension).
   Output
   z: Final weighted representation.
   α: Attention weights indicating diagnostic
2. **Non-linear Projection**:
   Each slice feature vector x_i is projected through a hidden layer with tanh activation:
   $h_i=\tanh(W_1 \cdot x_i+b_1)$
   where $W_1 \in R^{d\times dh}$, $b_1 \in R^{dh}$, and $d_h=192$ (hidden dimension).
3. **Attention Scoring**:
   A linear projection computes unnormalized attention scores:
   $e_i=W_2 \cdot h_i+b_2$
   where $W_2 \in R^{dh\times 1}$, $b_2 \in R$.
4. **Softmax Normalization**:
   The attention weights $\alpha_i \in [0,1]$ are computed via softmax:
   $\alpha=\mathrm{softmax}([e_1,e_2,...,e_{30}])$
   ensuring $\sum_{i=1}^{30} \alpha_i = 1$
5. **Weighted Feature Aggregation**:
   The final representation $z \in R^d$ is obtained by applying attention weights:
   z= ensuring $\sum_{i=1}^{30} \alpha_i \cdot x_i$

Weight matrices W_1 and W_2 are initialized using Xavier (Glorot) uniform initialization, with u=d/4 (e.g., u=128-D when d=512), balancing representational capacity and computational efficiency. Bias terms b_1 and b_2 are initialized to zero and updated during training based on the input statistics.

The process begins with a non-linear projection defined as $h_i=\tanh(W_1x_i+b_1)$, followed by a linear

Manuscript received 10 May 2025; Revised 30 June 2025; Accepted 1 July 2025; Available online 5 July 2025
Digital Object Identifier (**DOI**): https://doi.org/10.35882/jeeemi.v7i3.947

scoring step computed as $e_i = W_2 h_i + b_2$. Finally, softmax normalization is applied to all scores to obtain the attention weights, $\alpha = \text{softmax}([e_1, ..., e_{30}])$. The resulting attention weights α reflect the relative importance of each slice in stroke classification, allowing the model to emphasize diagnostically meaningful sections while suppressing noise or irrelevant anatomy.

In MedHybridNet, the attention mechanism is extended into a hierarchical framework spanning both region and slice levels. A multi-head self-attention module refines patch embeddings extracted through the Transformer path, allowing the model to identify lesion-related patterns within each axial slice. To assess global slice relevance, the model applies a softmax-weighted scoring system that mirrors the diagnostic reasoning processes typically used by radiologists. This dual-level design enhances both local pattern recognition and volumetric context modelling, improving accuracy and interpretability without requiring manual slice-level annotations. The attention weights refine the final feature representation used by the classifier head, which comprises a dense layer (64 units), dropout (rate = 0.5), and sigmoid activation. During training, we use a weighted binary cross-entropy loss (stroke class weight: 1.61, normal class 1.00) and stratified sampling to maintain balanced learning dynamics. This is especially important in situations where weak supervision is present, as only patient-level labels are available. Finally, the resulting attention maps offer clinicians with quantitative importance measures and visual overlays, aligning closely with radiologists' diagnostic processes and enhancing trust in AI-assisted stroke detection.

IV. Results

To ensure robust and reliable results, all models (DenseNet201, EfficientNetV2-B3, VGG19, ResNet50-RS, MobileNetV3, MedHybridNet) were trained across five independent runs on our brain imaging dataset. This allowed us to capture training variability and ensure stable performance. We evaluated average metrics across runs to avoid bias from random initialization. To prevent overfitting, we saved weights at the highest validation accuracy during each run. We applied various scenarios (Fig. 6) to assess their impact on performance. In some experiments, we used curriculum learning to introduce augmented data progressively, enhancing model robustness. The dataset was split into 70% training, 15% validation, and 15% testing sets (with no augmentation applied to the validation and test sets). Early stopping was used based on validation loss and accuracy to ensure optimal generalization. Across runs, models consistently converged to similar performance levels, confirming the reliability of our setup. We conducted three experiments to evaluate the impact of training strategies and augmentation methods. The first served

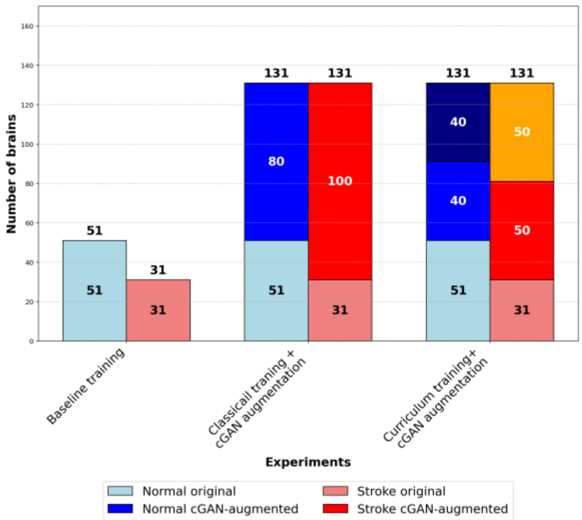


Fig. 6. Overview of data augmentation techniques used across the three experimental scenarios.

as a baseline, with no data augmentation applied. In the second, classical training was combined with data augmentation by introducing synthetic images for both normal and stroke classes. The third experiment incorporated curriculum learning, gradually increasing the model's exposure to augmented data throughout the training process.

A. Quantitative test

In Table 4, MedHybridNet achieves the highest accuracy (93.41%) and AUC (95.00%), indicating strong overall discriminative ability. However, its stroke recall (83.89%) and F1-score (87.41%) reveal limitations in recall under class imbalance. Other models showed lower stroke recall, especially VGG19 and MobileNetV3 (~77–82%), highlighting the challenges of class imbalance and underscoring the need for augmentation to improve balanced performance.

In Table 5, MedHybridNet outperformed other models once again with an accuracy of 96.97% and the highest AUC of 98.93%, maintaining strong stroke recall (94.95%) and precision (95.45%). DenseNet201 (96.17% accuracy, 97.92% AUC) and EfficientNetV2-B3 (96.37% accuracy, 97.31% AUC) showed lower recall, while older models like VGG19 and MobileNetV3 lagged further behind (~91–92% recall). The significant improvement in stroke-class performance highlights the positive impact of cGAN-based data augmentation on model sensitivity and generalization, which is particularly important for minimizing false negatives in clinical stroke detection. The results presented in Table 6, which incorporate curriculum training combined with cGAN-based data augmentation, represent the highest performance achieved across all experimental configurations. MedHybridNet achieves an outstanding accuracy of 98.31% and AUC of 99.60%, solidifying its

Manuscript received 10 May 2025; Revised 30 June 2025; Accepted 1 July 2025; Available online 5 July 2025
Digital Object Identifier (**DOI**): https://doi.org/10.35882/jeeemi.v7i3.947

Table 4. Performance comparison of different backbone networks under normal training without data augmentation.

Backbone	AUC (%)	Accuracy (%)	Normal (%)			Stroke (%)		
			Precision	Recall	F1	Precision	Recall	F1
DenseNet201	93.77	92.53	97.70	96.07	96.88	90.21	82.33	86.09
EfficientNetV2-B3	93.07	92.78	98.03	96.45	97.24	90.72	81.69	85.97
VGG19	89.98	86.28	89.97	89.62	89.79	82.64	78.30	80.41
ResNet50-RS	90.71	88.00	92.26	92.20	92.23	84.72	77.99	81.22
MobileNetV3	89.87	86.29	90.29	89.94	90.12	82.54	77.34	79.86
MedHybridNet	95.00	93.41	98.36	96.39	97.36	91.24	83.89	87.41

Table 5. Performance comparison of backbone networks with cGAN-based data augmentation.

Backbone	AUC (%)	Accuracy (%)	Normal (%)			Stroke (%)		
			Precision	Recall	F1	Precision	Recall	F1
DenseNet201	97.92	96.17	98.56	97.51	98.03	94.90	92.73	93.80
EfficientNetV2-B3	97.31	96.37	98.76	97.76	98.26	95.21	92.67	93.92
VGG19	95.08	92.70	93.66	93.43	93.55	91.06	91.90	91.47
ResNet50-RS	95.52	93.69	95.10	95.06	95.08	92.11	91.47	91.78
MobileNetV3	94.52	92.70	93.87	93.64	93.75	91.03	91.33	91.18
MedHybridNet	98.93	96.97	98.97	97.72	98.34	95.45	94.95	95.20

role as the most effective architecture for stroke classification. DenseNet201 and EfficientNetV2-B3 also perform exceptionally well, with accuracies of 98.08% and 98.15%, respectively. This experiment employs a two-phase curriculum learning strategy to enhance deep learning-based stroke detection in CT scans, utilizing data augmentation. In the first phase, the model trains on original images, establishing robust low-level feature representations of normal and pathological brain anatomy. The second phase introduces cGAN generated synthetic stroke lesions, enabling the model to learn complex pathological patterns while maintaining anatomical coherence. Precision, recall, and F1 scores are well-balanced across both classes, demonstrating the effectiveness of this approach in mitigating class imbalance. These findings highlight the critical role of curriculum learning in maximizing the utility of data augmentation strategies, offering a robust framework for improving stroke detection in imbalanced medical datasets.

To assess the statistical significance of MedHybridNet's performance improvement over baseline models in the final experimental scenario, where all models achieved their best results, we conducted pairwise comparisons using McNemar's test on the test set predictions from five independent training runs. The results (Table 7) indicate that the differences between MedHybridNet and both DenseNet201 (χ^2 = 7.81, p = 0.0052) and EfficientNetV2-B3 (χ^2 = 6.25, p = 0.0124) are statistically significant at the $p < 0.05$ level. These findings support our assertion that the observed improvements are not due to random variation, particularly considering that DenseNet201 and EfficientNetV2-B3 were the top-performing baseline models.

An additional error analysis was performed to understand the limitations of MedHybridNet better and to identify areas for potential improvement. Out of 589 test images, the model misclassified 10 cases (1.69%). These included six false negatives, where stroke cases were incorrectly classified as normal, and four false positives, where normal cases were mistakenly identified as strokes. These errors were reviewed with the help of domain experts to identify recurring patterns. False negatives were primarily associated with early-stage stroke manifestations showing minimal hypodensity, small infarcts in atypical regions such as the brainstem or deep white matter, and scans compromised by low contrast or motion artefacts. In contrast, false positives were often linked to age-related atrophy, chronic white matter disease, post-traumatic changes mimicking acute stroke features, or the presence of image noise and reconstruction artefacts.

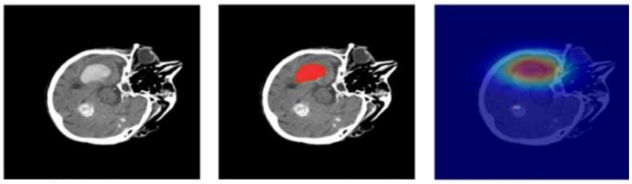
These findings prove that while the model performs robustly under weak supervision and limited annotation, improvements could be made in detecting subtle lesions, distinguishing chronic vs. acute pathology, and handling low-quality scans. This analysis informs our future work on incorporating region-specific attention, semi-supervised refinement, and enhanced data diversity through targeted augmentation.

Manuscript received 10 May 2025; Revised 30 June 2025; Accepted 1 July 2025; Available online 5 July 2025
Digital Object Identifier (**DOI**): https://doi.org/10.35882/jeeemi.v7i3.947

B. Qualitative test

Fig. 7 presents a brain CT scan alongside an expert-annotated lesion (middle) and a Grad-CAM-generated heatmap. The heatmap highlights regions that overlap with the expert-confirmed stroke area, indicating that the model successfully focuses on clinically relevant features. The second image, segmented and validated by a medical expert, supports the relevance of the detected area. This demonstrates the potential of Grad-CAM for providing interpretable insights in stroke detection tasks.

To evaluate the interpretability and clinical relevance of MedHybridNet, we conducted a qualitative assessment with two board-certified radiologists who reviewed both real and synthetic CT scans from the test set (Table 8). Each scan was presented with Grad-CAM visualizations overlaid on axial slices to highlight the regions most influential in the model's decision-making. Radiologists rated the diagnostic alignment of attention maps using an "Average Relevance Rating (Likert)", and the visual quality of anatomical structures using an "Average Realism Rating (Likert)", both measured on a 5-point scale: (1) Not relevant / Not realistic, (2) Somewhat irrelevant / Somewhat unrealistic, (3) Neutral, (4) Relevant / Realistic, (5) Highly relevant / Highly realistic. The radiologists were blinded to whether the images were real or cGAN-generated. They were asked to assess whether the highlighted regions aligned with known stroke patterns, as well as the overall plausibility of the Grad-CAM overlays and input scans. This evaluation provided valuable insights into both the model's diagnostic interpretability and the clinical fidelity of synthetic data, supporting its reliability under weak supervision and the effectiveness of the proposed cGAN-based augmentation strategy.



(a) (b) (c)

Fig. 7. Brain CT slice (a) original image, (b) expert-annotated stroke region (red), and (c) Grad-CAM heatmap highlighting model attention.

Across the test set evaluated by two radiologists, 86% of the Grad-CAM overlays on synthetic images were rated as either "relevant" or "highly relevant," reflecting a strong correspondence between the model's attention focus and stroke-related anatomical regions. Regarding image realism, the synthetic scans consistently received high ratings, indicating that the cGAN-generated images preserved essential anatomical structures and accurately reproduced stroke-like imaging patterns. In several instances, radiologists even reported difficulty distinguishing between real and synthetic images, highlighting the high visual fidelity and diagnostic plausibility of the generated data.

Table 8. Radiologist evaluation of interpretability and realism for real and synthetic brain scans using Likert ratings.

Scan type	Avg. Relevance Rating (Likert)	Avg. Realism Rating (Likert)	% Rated ≥4 (Relevant/Highly Relevant)
Real Scans	4.8	4.9	94%
Synthetic Scans	4.3	4.2	86%

These findings reinforce the claim that MedHybridNet not only achieves high diagnostic accuracy (98.31%) but also generates clinically meaningful and interpretable decision-making pathways. This is an essential characteristic for reliable deployment in real-world clinical settings.

C. Runtime efficiency and hardware requirements

To ensure practical deployment and clinical applicability, we evaluated the computational efficiency of MedHybridNet during both training and inference. During training, the model was optimized using Google Colab Pro, which provided access to a Tesla T4 GPU. As mentioned previously, training was conducted in batches of 4 brain volumes, utilizing early stopping. This optimization strategy significantly reduced total training time while maintaining high classification accuracy. On average, each training epoch took approximately 20–30 minutes, allowing full convergence within two 24-hour sessions (approximately126 epochs).

For inference, MedHybridNet was tested on a computer equipped with an 11th Gen Intel® Core™ i5 processor and an NVIDIA RTX 3050 GPU, representing a cost-effective configuration with moderate computational resources. This setup demonstrates that the model can operate efficiently without requiring high-end or specialized equipment.

On this system, the model processed a full brain volume (30 axial CT slices) in approximately 15 seconds, including preprocessing, feature extraction, attention-based aggregation, and final classification. This execution time demonstrates its suitability for use in time-sensitive applications such as emergency stroke triage, where rapid and reliable decision support

Manuscript received 10 May 2025; Revised 30 June 2025; Accepted 1 July 2025; Available online 5 July 2025
Digital Object Identifier (**DOI**): https://doi.org/10.35882/jeeemi.v7i3.947

Table 9. Ablation study evaluating the impact of key architectural components and training strategies on the performance of MedHybridNet for stroke classification.

Ablation factor	AUC (%)	Acc. (%)	Normal (%)			Stroke (%)		
			Prec.	Recall	F1	Prec.	Recall	F1
MedHybridNet (Full Model)	99.60	98.31	99.52	98.93	99.23	98.18	96.42	97.29
Without transformer path	98.29	95.87	97.68	97.89	94.76	96.30	91.45	95.67
Without hierarchical attention	99.15	97.56	99.34	99.23	96.78	97.99	94.56	98.76
Without stratified sampling	98.20	95.52	97.82	93.57	95.68	94.53	91.00	92.71
With RELU activation in attention	99.42	97.89	99.45	99.12	97.01	98.05	95.12	99.01
Classification on individual slices	98.58	95.98	98.93	97.63	98.28	95.23	90.90	93.02
Data augmentation using classical technique	98.97	96.50	98.70	97.13	97.91	93.80	95.19	94.49
Without skip connections	98.58	97.12	99.12	99.02	96.12	97.55	93.89	98.23

Table 10. Performance comparison of backbone models without the SliceAttention mechanism for stroke classification.

Model without attention	AUC (%)	Acc. (%)	Normal (%)			Stroke (%)		
			Prec.	Recall	F1	Prec.	Recall	F1
DenseNet201	96.22	96.31	97.67	96.37	97.02	96.64	94.41	95.51
EfficientNetV2-B3	95.35	95.37	96.52	96.52	96.52	94.90	93.23	94.06
VGG19	93.61	93.69	94.57	93.90	94.24	93.17	92.89	93.03
ResNet50-RS	94.17	94.26	95.29	94.50	94.89	94.10	92.91	93.50
MobileNetV3	93.68	93.71	94.25	94.21	94.23	93.55	92.57	93.05
MedHybridNet	96.39	96.42	98.75	98.45	95.32	96.86	92.14	96.23

is critical. These results highlight that MedHybridNet not only achieves high diagnostic accuracy but is also well-suited for real-world deployment, particularly in settings where access to cloud infrastructure or high-end computing is limited.

V. Ablation study

The ablation study provides a deeper understanding of how each architectural component contributes to MedHybridNet's performance. As shown in Table 9, removing the Transformer path led to a significant drop in accuracy (from 98.31% to 95.87%) and stroke recall (from 96.42% to 91.45%), confirming that global contextual modelling enhances the model's ability to detect subtle hypodensity patterns across slices. Disabling hierarchical attention also reduced stroke recall by ~1.86% (to 94.56%) and F1-score (to 98.76%), indicating its role in refining both local feature selection and inter-slice prioritization.

Additionally, we evaluated alternative architectures that do not include attention-based slice prioritization (Table 10). When trained without SliceAttention, models such as DenseNet201 and EfficientNetV2-B3 achieved accuracies of only 96.31% and 95.37% , respectively, with corresponding stroke F1-scores of 95.51% and 94.06%. These results are notably lower than MedHybridNet's stroke F1-score of 96.23%. This comparison confirms that volumetric reasoning through inter-slice attention is essential for maintaining high diagnostic reliability, particularly under weak supervision where manual annotations are unavailable.

Furthermore, removing stratified sampling resulted in a more pronounced degradation in stroke-class sensitivity (91.00% recall) despite normal-class precision remaining high, highlighting the importance of balanced batch composition during training. Interestingly, replacing tanh with ReLU activation in attention had only minor effects on performance (97.89% accuracy, 99.01% stroke F1), suggesting that while ReLU introduces slight instability, it does not severely affect classification. Classifying slices individually without inter-slice attention resulted in a substantial decline in diagnostic reliability, underscoring the importance of volumetric reasoning in clinical decision-making. Using classical data augmentation techniques instead of cGAN-based synthesis also reduced performance (94.49% stroke F1-score), demonstrating the importance of realistic synthetic samples in improving generalization under limited data conditions.

Finally, removing skip connections led to a marginal improvement in normal-class recall (99.02%) but reduced stroke-class recall and increased instability during training. This suggests that while the model retains some diagnostic capability through residual feature learning, the lack of direct CNN-to-Transformer feature fusion disrupts gradient flow, leading to less stable convergence and weaker alignment with diagnostically meaningful regions.

Manuscript received 10 May 2025; Revised 30 June 2025; Accepted 1 July 2025; Available online 5 July 2025
Digital Object Identifier (**DOI**): https://doi.org/10.35882/jeeemi.v7i3.947

Table 11. Comparative analysis of the proposed method against existing stroke classification approaches using brain imaging data.

Reference	Accuracy (%)	Classifier	Using the whole brain	Data type
[26]	96.5	Genetic algorithms, CNN, BiLSTM	No	Brain CT
[27]	96.61	ViT, LSTM	No	Brain CT
[30]	95.06	U-Net	No	Brain CT
[32]	99.21	DenseNet121, MobileNetV3, CNN	No	Brain CT
[40]	97	XGBoost	No	Brain CT
[41]	96.5	DCNN	No	Brain CT
[42]	96	DenseNet-121, ResNet-50 and VGG-16	No	MRI
[43]	95	CNN, ML models	No	MRI
[44]	97.15	CNN	No	Brain CT
Our proposition	98.31	CNN. SliceAttention mechanism	Yes	Brain CT

These findings not only validate the hybrid design of MedHybridNet but also offer insights into component synergy: the Transformer path improves global coherence, skip connections stabilize learning, SliceAttention dynamically highlights lesion-related slices, and hierarchical attention ensures interpretability at both regional and volumetric levels. Techniques such as cGAN-based augmentation and stratified sampling provide necessary data diversity and balance, while the Transformer-CNN fusion enables robustness and clinical alignment. Overall, this ablation study confirms that MedHybridNet's superior performance stems from the combined effect of multiple architectural innovations, rather than any single component alone.

VI. Discussion

The early and accurate detection of brain stroke remains one of the most pressing challenges in modern healthcare where timely intervention can significantly influence patient survival and long-term recovery. In this context. we proposed MedHybridNet. a novel hybrid deep learning architecture that combines convolutional feature extraction with Transformer-based global reasoning. guided by a weakly supervised attention mechanism known as SliceAttention. Our experimental results demonstrate that this approach not only achieves high diagnostic accuracy but also enhances model interpretability which is a crucial requirement for real-world deployment in medical imaging.

At the core of our method lies the SliceAttention mechanism, which enables the model to identify diagnostically relevant slices using only patient-level labels dynamically. This eliminates the need for labour-intensive slice-level annotations, making the system more scalable and practical for clinical adoption. The ablation study confirms the importance of this mechanism in prioritizing informative slices and improving classification performance. By assigning higher weights to pathological regions without explicit supervision, SliceAttention aligns with radiologists' decision-making process and offers explainable insights through Grad-CAM visualizations.

Furthermore, the integration of CNN and Transformer pathways within MedHybridNet allows the model to simultaneously capture fine-grained anatomical details and global contextual dependencies across slices. This dual-path design significantly contributes to its robustness, particularly in data-scarce and class-imbalanced conditions.

The use of hierarchical attention at both the regional and slice levels further refine diagnostic inference, reinforcing the model's ability to generalize across variations in anatomy and pathology. Our proposed cGAN-based augmentation strategy plays a pivotal role in addressing dataset imbalance and enhancing model generalization. Unlike traditional augmentation techniques that apply simple geometric or intensity transformations, the cGAN generates realistic, pathology-informed synthetic CT slices conditioned on diagnostic labels. This not only enriches the training data but also improves the model's sensitivity to subtle stroke-related features.

When combined with curriculum learning, the augmentation strategy yields the highest performance, demonstrating the value of structured data exposure in complex classification tasks. When compared to existing methods (Table 11), MedHybridNet achieves competitive accuracy while offering unique advantages such as whole-brain analysis, weak supervision, and interpretability. Several recent approaches have achieved high accuracy by utilizing cropped or segmented regions of interest. For example, [32] reports an accuracy of 99.21%. However, these methods typically rely on extensive preprocessing and manual annotation, which can introduce errors during data manipulation and are often impractical in routine

Manuscript received 10 May 2025; Revised 30 June 2025; Accepted 1 July 2025; Available online 5 July 2025
Digital Object Identifier (DOI): https://doi.org/10.35882/jeeemi.v7i3.947

clinical settings. In contrast, our framework operates directly on full-volume scans without the need for prior segmentation, thereby enhancing its practicality and applicability in real-world clinical environments. Unlike earlier models that rely solely on CNNs or classical machine learning classifiers [26], [30], [44], MedHybridNet introduces a hybrid design that leverages both local and global feature modelling. For instance, while [27] used Vision Transformers and LSTM layers to achieve 96.61% accuracy, their model lacks an interpretable attention mechanism and does not effectively address data scarcity. Similarly, [41] employed DCNNs to reach 96.5% accuracy but lacked mechanisms for weak supervision or interpretability, limiting its clinical relevance. In contrast, our approach not only outperforms many of these methods by achieving 98.31% accuracy but also introduces critical innovations: weak supervision via SliceAttention, a hybrid CNN-Transformer architecture, and cGAN-based augmentation tailored for volumetric medical imaging. Together, these components enhance model reliability, reduce annotation burden, and provide clinically meaningful explanations for each prediction.

Despite its strengths, this study has certain limitations. First, the dataset used in this work was collected from a single institution (Lady Reading Hospital, Peshawar, Pakistan) and includes a relatively small number of patients (n=82). This introduces potential biases related to scanner type, acquisition protocol, and patient demographics, which may affect the model's generalizability when deployed across multi-centre or international clinical environments.

While our use of cGAN-based data synthesis and SliceAttention-guided interpretability helps mitigate some of these constraints by improving feature robustness and reducing overfitting, we acknowledge that scanner-specific artefacts, regional stroke presentation patterns, and imaging variability can still impact performance in unseen or heterogeneous settings. Second, although MedHybridNet achieves strong classification accuracy and interpretability under weak supervision, the model remains sensitive to the quality and consistency of the input. Artefacts such as motion blur, poor contrast, or inconsistent slice thickness could degrade attention maps and reduce diagnostic reliability, especially when applied to scans acquired using different imaging pipelines. Lastly, while SliceAttention provides meaningful insights into which slices are most influential in the model's decision-making process, it computes attention weights in a learned, data-driven manner rather than based on manual annotations. As such, these interpretations should be clinically validated before being used in real-world diagnostic workflows. To address these limitations, we plan to evaluate MedHybridNet on multi-centre CT brain imaging repositories and through collaborative partnerships with other hospitals. These efforts will assess model robustness under varying imaging protocols, scanner types, and patient populations, supporting broader deployment in real-world clinical environments.

In conclusion, MedHybridNet represents a significant step forward in the development of automated, interpretable, and clinically aligned systems for brain stroke detection. By combining architectural innovation, weak supervision, and advanced data synthesis techniques, we present a framework that not only delivers strong diagnostic performance but also addresses practical challenges in deploying AI models in healthcare. Future work will focus on validating the model on multi-centre datasets and integrating it into clinical workflows to assess its impact on diagnostic efficiency and treatment planning.

VII. Conclusion

This study addresses the critical challenge of automated and interpretable brain stroke detection in volumetric CT imaging. Our primary objective was to develop a deep learning framework that not only achieves high diagnostic accuracy but also aligns with clinical requirements by minimizing reliance on labour-intensive annotations, effectively modelling 3D anatomical context, and providing explainable decision-making. To this end, we introduced MedHybridNet, a novel hybrid architecture that combines convolutional networks for local feature extraction with Transformer-based modules for global contextual reasoning. This dual-path design enables the model to capture both fine-grained pathological patterns and inter-slice dependencies, significantly enhancing classification performance. A key innovation of the model is the SliceAttention mechanism, which enables the identification of diagnostically relevant slices using only patient-level labels, eliminating the need for slice-level annotations and offering interpretable visualizations through attention maps and Grad-CAM. We further addressed data scarcity and class imbalance through a cGAN-based augmentation strategy, generating realistic synthetic CT slices conditioned on diagnostic labels. The ablation study confirmed the contribution of each architectural component. At the same time, comparative analysis demonstrated that MedHybridNet achieves competitive accuracy (98.31%) compared to existing methods, while offering unique advantages such as whole-brain processing and weak supervision. Our work advances the field by bridging the gap between technical innovation and clinical utility in medical imaging AI. Unlike many prior approaches that rely on segmented or cropped regions, our system operates directly on

Manuscript received 10 May 2025; Revised 30 June 2025; Accepted 1 July 2025; Available online 5 July 2025
Digital Object Identifier (**DOI**): https://doi.org/10.35882/jeeemi.v7i3.947

full-volume scans, making it more applicable in real-world settings where preprocessing is limited. Moreover, the interpretability features support trust and transparency, which are key prerequisites for clinical deployment.

Future work will focus on validating the model on multi-centre and multi-modal datasets, integrating it into clinical workflows for real-time testing, and extending its application to other neurological conditions detectable via CT imaging. We are currently exploring lightweight versions of MedHybridNet for edge deployment and investigating semi-supervised extensions to reduce labelling effort further. By delivering a scalable, accurate, and interpretable solution for stroke detection, this research makes a meaningful contribution to the integration of AI in emergency neuroimaging, supporting timely, data-driven clinical decisions.

References

[1] The top 10 causes of death. Accessed: May 13, 2024. [Online]. Available: https://www.who.int/news-room/fact-sheets/detail/the-top-10-causes-of-death

[2] G. K. Sharma, S. Kumar, V. Ranga, and M. K. Murmu, "Artificial intelligence in cerebral stroke images classification and segmentation: A comprehensive study," *Multimedia Tools Appl.*, vol. 83, no. 14, pp. 43539–43575, Apr. 2024, doi: doi.org/10.1007/s11042-023-17324-3

[3] A. Tursynova, A. Sakhipov, I. Omirzak, Z. Ikram, S. Smakova, and M. Kutubayeva, "Classification of brain strokes in computed tomography images utilizing deep learning," in *Proc. IEEE 4th Int. Conf. Smart Inf. Syst. Technol. (SIST)*, May 2024, pp. 328–333, doi: 10.1109/SIST61555.2024.10629580

[4] J. Luo, P. Dai, Z. He, Z. Huang, S. Liao, and K. Liu, "Deep learning models for ischemic stroke lesion segmentation in medical images: A survey," *Comput. Biol. Med.*, vol. 177, p. 108509, Aug. 2024, doi: 10.1016/j.compbiomed.2024.108509

[5] C. Huang, J. Wang, S.-H. Wang, and Y.-D. Zhang, "Applicable artificial intelligence for brain disease: A survey," *Neurocomputing*, vol. 504, pp. 223–239, Sep. 2022. doi: doi.org/10.1016/j.neucom.2022.07.005

[6] M. Arabahmadi, R. Farahbakhsh, and J. Rezazadeh, "Deep learning for smart healthcare—A survey on brain tumor detection from medical imaging," *Sensors*, vol. 22, no. 5, p. 1960, Mar. 2022. doi: doi.org/10.3390/s22051960

[7] P. Hu, T. Yan, B. Xiao, H. Shu, Y. Sheng, Y. Wu, et al., "Deep learning-assisted detection and segmentation of intracranial hemorrhage in noncontrast computed tomography scans of acute stroke patients: a systematic review and meta-analysis," *International Journal of Surgery*, vol. 110, no. 6, pp. 3839–3847, 2022, doi: 10.1097/JS9.0000000000001266

[8] S. Ahmed, J. F. Esha, M. S. Rahman, M. S. Kaiser, A. S. M. S. Hosen, D. Ghimire, and M. J. Park, "Exploring deep learning and machine learning approaches for brain hemorrhage detection," *IEEE Access*, vol. 12, pp. 45060–45093, 2024. doi: 10.1109/ACCESS.2024.3376438

[9] Z. Xu and C. Ding, "Combining convolutional attention mechanism and residual deformable transformer for infarct segmentation from CT scans of acute ischemic stroke patients," *Frontiers Neurol.*, vol. 14, Jul. 2023, Art. no. 1178637. doi: 10.3389/fneur.2023.1178637

[10] M. Srikrishna et al., "Deep learning from MRI-derived labels enables automatic brain tissue classification on human brain CT," *NeuroImage*, vol. 244, p. 118606, Dec. 2021, doi: doi.org/10.1016/j.neuroimage.2021.118606

[11] U. Raghavendra, T.-H. Pham, A. Gudigar, V. Vidhya, B. N. Rao, S. Sabut, et al., "Novel and accurate non-linear index for the automated detection of haemorrhagic brain stroke using CT images," *Complex Intell. Syst.*, vol. 7, no. 2, pp. 929–940, Apr. 2021, doi: doi.org/10.1007/s40747-020-00257-x

[12] A. Gautam and B. Raman, "Towards effective classification of brain hemorrhagic and ischemic stroke using CNN," *Biomed. Signal Process. Control*, vol. 63, p. 102178, Jan. 2021, doi: doi.org/10.1016/j.bspc.2020.102178

[13] R. Raj, J. Mathew, S. K. Kannath, and J. Rajan, "StrokeViT with AutoML for brain stroke classification," *Eng. Appl. Artif. Intell.*, vol. 119, p. 105772, Mar. 2023, doi: doi.org/10.1016/j.engappai.2022.105772

[14] S. Yalçın and H. Vural, "Brain stroke classification and segmentation using encoder-decoder based deep convolutional neural networks," *Comput. Biol. Med.*, vol. 149, p. 105941, Oct. 2022, doi: 10.1016/j.compbiomed.2022.105941

[15] O. Shafaat et al., "Leveraging artificial intelligence in ischemic stroke imaging," *J. Neuroradiol.*, vol. 49, no. 4, pp. 343–351, Jul. 2022, doi: 10.1016/j.neurad.2021.05.001

[16] T. Pokorny, D. Vrba, O. Fiser, M. Salucci, and J. Vrba, "Systematic Optimization of Training and Setting of SVM-Based Microwave Stroke Classification: Numerical Simulations for 10 Port System," *IEEE J. Electromagn. RF Microw. Med. Biol.*, vol. 8, no. 3, pp. 273–281, Sep. 2024, doi:

Manuscript received 10 May 2025; Revised 30 June 2025; Accepted 1 July 2025; Available online 5 July 2025
Digital Object Identifier (**DOI**): https://doi.org/10.35882/jeeemi.v7i3.947

10.1109/JERM.2024.3404119

[17] L. de Vries, B. J. Emmer, C. B. Majoie, H. A. Marquering, and E. Gavves, “PerfU-Net: Baseline infarct estimation from CT perfusion source data for acute ischemic stroke,” *Med. Image Anal.*, vol. 85, p. 102749, Apr. 2023, doi: doi.org/10.1016/j.media.2023.102749

[18] M. Srikrishna, J. B. Pereira, R. A. Heckemann, F. Alshehri and G. Muhammad, "A few-shot learning-based ischemic stroke segmentation system using weighted MRI fusion," *Image and Vision Computing*, vol. 140, p. 104865, 2023. doi: 10.1016/j.imavis.2023.104865

[19] A. Srinivas and J. P. Mosiganti, “A brain stroke detection model using soft voting based ensemble machine learning classifier,” *Measurement: Sensors*, vol. 29, p. 100871, Oct. 2023, doi: doi.org/10.1016/j.measen.2023.100871

[20] O. Ozaltin, O. Coskun, O. Yeniay, and A. Subasi, “A Deep Learning Approach for Detecting Stroke from Brain CT Images Using OzNet,” *Bioengineering*, vol. 9, no. 12, p. 783, Dec. 2022, doi: 10.3390/bioengineering9120783

[21] N. H. Ali, A. R. Abdullah, N. M. Saad, A. S. Muda, and E. E. M. Noor, “Automated Classification of Collateral Circulation for Ischemic Stroke in Cone-Beam CT Images Using VGG11: A Deep Learning Approach,” *BioMedInformatics*, vol. 4, no. 3, pp. 1692–1702, Jul. 2024, doi: doi.org/10.3390/biomedinformatics4030091

[22] Y.-T. Chen, Y.-L. Chen, Y.-Y. Chen, Y.-T. Huang, H.-F. Wong, J.-L. Yan, and J.-J. Wang, “Deep Learning–Based Brain Computed Tomography Image Classification with Hyperparameter Optimization through Transfer Learning for Stroke,” *Diagnostics*, vol. 12, no. 4, p. 807, Mar. 2022, doi: 10.3390/diagnostics12040807

[23] V. Mariano, J. A. Tobon Vasquez, M. R. Casu, and F. Vipiana, “Brain Stroke Classification via Machine Learning Algorithms Trained with a Linearized Scattering Operator,” *Diagnostics*, vol. 13, no. 1, p. 23, Dec. 2022, doi: doi.org/10.3390/diagnostics13010023

[24] M. Uppal, D. Gupta, S. Juneja, T. R. Gadekallu, I. El Bayoumy, J. Hussain et al., “Enhancing accuracy in brain stroke detection: Multi-layer perceptron with Adadelta, RMSProp and AdaMax optimizers,” *Front. Bioeng. Biotechnol.*, vol. 11, Sep. 2023, doi: doi.org/10.3389/fbioe.2023.1257591

[25] F. Palsson, N. D. Forkert, L. Meyer, G. Broocks, F. Flottmann, M. E. Maros et al., “Prediction of tissue outcome in acute ischemic stroke based on single-phase CT angiography at admission,” *Front. Neurol.*, vol. 15, Mar. 2024, doi: 10.3389/fneur.2024.1330497

[26] M. A. Saleem, A. Javeed, W. Akarathanawat, A. Chutinet, N. C. Suwanwela, and W. Asdornwised, “Innovations in Stroke Identification: A Machine Learning-Based Diagnostic Model Using Neuroimages,” *IEEE Access*, vol. 12, pp. 35754–35764, 2024, doi: 10.1109/ACCESS.2024.3369673

[27] M. M. Hossain, M. M. Ahmed, A. A. N. Nafi, M. R. Islam, M. S. Ali, and J. Haque, “A novel hybrid ViT-LSTM model with explainable AI for brain stroke detection and classification in CT images: A case study of Rajshahi region,” *Comput. Biol. Med.*, vol. 186, p. 109711, Jan. 2025, doi: doi.org/10.1016/j.compbiomed.2025.109711

[28] Ö. Polat, Z. Dokur, and T. Ölmez, “Classification of brain strokes using divergence-based convolutional neural networks,” *Biomed. Signal Process. Control*, vol. 93, p. 106193, Jul. 2024, doi: doi.org/10.1016/j.bspc.2024.106193

[29] M. Soltanpour, P. Boulanger, and B. Buck, “CT perfusion map synthesis from CTP dynamic images using a learned LSTM generative adversarial network for acute ischemic stroke assessment,” *J. Med. Syst.*, vol. 48, no. 1, p. 37, Apr. 2024, doi: 10.1007/s10916-024-02054-2

[30] B. Kaya and M. Önal, “A CNN transfer learning-based approach for segmentation and classification of brain stroke from noncontrast CT images,” *Int. J. Imaging Syst. Technol.*, vol. 33, no. 4, pp. 1335–1352, Jul. 2023, doi: 10.1002/ima.22864

[31] M. A. Inamdar, A. Gudigar, U. Raghavendra, M. Salvi, R. R. A. B. R. Aman, and N. F. M. Gowdh, “A Dual-Stream Deep Learning Architecture With Adaptive Random Vector Functional Link for Multi-Center Ischemic Stroke Classification,” *IEEE Access*, vol. 13, pp. 46638–46658, 2025, doi: 10.1109/ACCESS.2025.3550344

[32] A. Abumihsan, A. Yousef Owda, M. Owda, M. Abumohsen, L. Stergioulas, and M. Ahmad Abu Amer, “A Novel Hybrid Model for Brain Ischemic Stroke Detection Using Feature Fusion and Convolutional Block Attention Module,” *IEEE Access*, vol. 13, pp. 44466–44483, 2025, doi: 10.1109/ACCESS.2025.3549269

[33] C. H. Patel, D. Undaviya, H. Dave, S. Degadwala, and D. Vyas, “EfficientNetB0 for brain stroke classification on computed tomography scan,” in *Proc. 2nd Int. Conf. Appl. Artif. Intell. Comput. (ICAAIC)*, May 2023, pp. 713–718, doi: 10.1109/ICAAIC56838.2023.10141195

[34] M. A. R. Afridi, “Brain stroke CT image dataset,” Kaggle, 2021. [Online]. Available:

Manuscript received 10 May 2025; Revised 30 June 2025; Accepted 1 July 2025; Available online 5 July 2025
Digital Object Identifier (**DOI**): https://doi.org/10.35882/jeeemi.v7i3.947

https://www.kaggle.com/datasets/afridirahman/brain-stroke-ct-image-dataset [Accessed: December 12, 2024].

[35] M. A. R. Afridi, Z. Ali, R. Muhammad, A. Ahmad, and I. Alam, "Age and gender specific stroke risk factors in a teaching hospital in Khyber Pakhtunkhwa," *J. Postgraduate Med. Inst.*, vol. 29, no. 2, pp. 79–85, 2015.

[36] I. Gulrajani, F. Ahmed, M. Arjovsky, V. Dumoulin, and A. C. Courville, "Improved training of Wasserstein GANs," in *Advances in Neural Information Processing Systems*, vol. 30, 2017.

[37] E. Aksan, M. Kaufmann, and O. Hilliges, "A Spatio-Temporal Transformer for 3D Human Motion Prediction," in *Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV)*, Montreal, QC, Canada, 2021, pp. 11884–11893. doi: 10.1109/3DV53792.2021.00066

[38] Z. Zhao, S. Singh, H. Lee, Z. Zhang, A. Odena, and H. Zhang, "Improved consistency regularization for GANs," in *Proceedings of the AAAI Conference on Artificial Intelligence*, vol. 35, no. 12, pp. 11033–11041, 2021.

[39] S. Deepak and P. M. Ameer, "Brain tumor categorization from imbalanced MRI dataset using weighted loss and deep feature fusion," *Neurocomputing*, vol. 520, pp. 94–102, 2023, doi: 10.1016/j.neucom.2022.11.039.

[40] S. K. UmaMaheswaran, F. Ahmad, R. Hegde, A. M. Alwakeel, and S. R. Zahra, "Enhanced non-contrast computed tomography images for early acute stroke detection using machine learning approach," *Expert Systems with Applications*, vol. 240, p. 122559, 2024, doi: doi.org/10.1016/j.eswa.2023.122559

[41] M. A. Sabir and F. Ashraf, "Development of a novel deep convolutional neural network model for early detection of brain stroke using CT scan images," *Multimedia Tools and Applications*, vol. 2024, pp. 1–25, Apr. 2024, doi: doi.org/10.1007/s11042-024-19001-5

[42] M. Gupta, P. Meghana, K. H. Reddy, and P. Supraja, "Predicting Brain Stroke Using IoT-Enabled Deep Learning and Machine Learning: Advancing Sustainable Healthcare," in *Sustainable Development through Machine Learning, AI and IoT*, P. Whig, N. Silva, A. A. Elngar, N. Aneja, and P. Sharma, Eds., Communications in Computer and Information Science, vol. 1939, 2023, doi: https://doi.org/10.1007/978-3-031-47055-4_10

[43] R. Mena, A. Macas, E. Pelaez, F. Loayza, and H. Franco-Maldonado, "A Pipeline for Segmenting and Classifying Brain Lesions Caused by Stroke: A Machine Learning Approach," in *Information Systems and Technologies*, A. Rocha, H. Adeli, G. Dzemyda, and F. Moreira, Eds., WorldCIST 2022, Lecture Notes in Networks and Systems, vol. 470, 2022, doi: doi.org/10.1007/978-3-031-04829-6_37

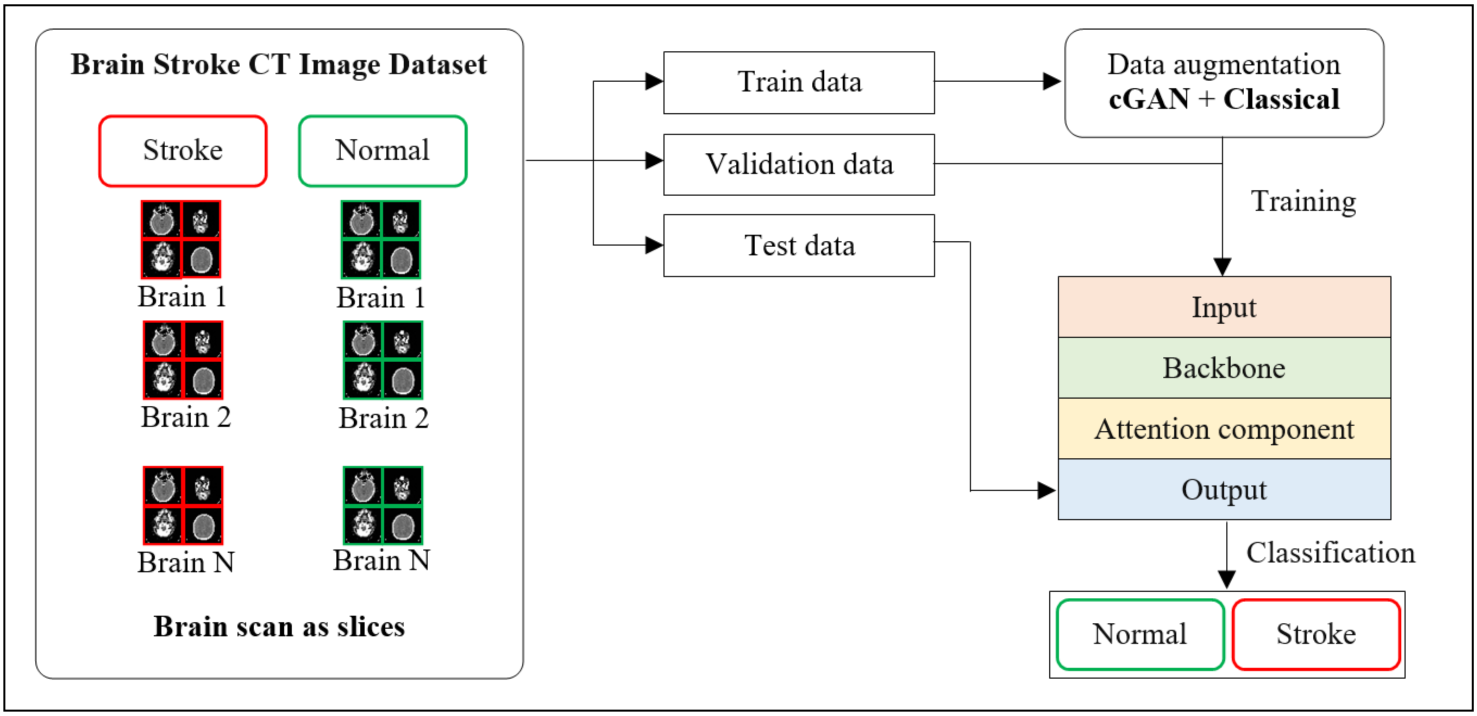
[44] N. A. Chowdhury et al., "A novel approach to detect stroke from 2D images using deep learning," in *Proc. 2nd Int. Conf. Big Data, IoT and Mach. Learn. (BIM 2023)*, M. S. Arefin, M. S. Kaiser, T. Bhuiyan, N. Dey, and M. Mahmud, Eds., Lecture Notes in Networks and Systems, vol. 867, Singapore: Springer, 2024, doi: doi.org/10.1007/978-981-99-8937-9_17

Author Biography

Fouzi Lezzar is currently an Associate Professor in the Department of Software Technologies and Information Systems at the Faculty of New Technologies of Information and Communication, University Constantine 2, Algeria. He earned his PhD in Computer Science from Batna 2 University, Algeria. His research interests include Artificial Intelligence, Internet of Things, smart cities, and e-health. He focuses on the development of intelligent and connected systems that enhance urban services and healthcare. In addition to teaching, he actively supervises master's and doctoral students. He also contributes to scientific publications and serves as a reviewer for academic journals in the field of computing and information systems.

Seif Eddine Mili is currently serving as an Associate Professor in the Department of Computer Science at the École Normale Supérieure, Constantine, Algeria. He holds a Ph.D. in Computer Science from Badji Mokhtar Annaba University, Algeria. He earned his Bachelor's and Master's degrees in Computer Science from Algerian institutions through regular academic programs. With several years of teaching and research experience, he has been actively engaged in mentoring students and leading research activities. His research interests include Artificial Intelligence, Internet of Things (IoT), Web of Things, Multi-Agent Systems, Bio-Inspired Systems, and Model-Driven Architecture. He has participated in various conferences and contributes regularly to academic publications.

Manuscript received 10 May 2025; Revised 30 June 2025; Accepted 1 July 2025; Available online 5 July 2025
Digital Object Identifier (**DOI**): https://doi.org/10.35882/jeeemi.v7i3.947



Manuscript received 10 May 2025; Revised 30 June 2025; Accepted 1 July 2025; Available online 5 July 2025
Digital Object Identifier (**DOI**): https://doi.org/10.35882/jeeemi.v7i3.947