SSRG International Journal of Electronics and Communication Engineering
ISSN: 2348-8549/ https://doi.org/10.14445/23488549/IJECE-V11I10P111

Volume 11 Issue 10, 135-148, October 2024

Original Article

Advanced Hybrid Method for Precise Identification and Categorization of Brain Stroke from CT Images

Shahina A.R[1], I. Sowmy[2]

[1]*Department of Electronics and Communication Engineering, Noorul Islam Centre for Higher Education, Tamil Nadu, India.*
[2]*Department of Biomedical Engineering, Noorul Islam Centre for Higher Education, Tamil Nadu, India.*

[1]*Corresponding Author : arshahina@outlook.com*

Received: 11 August 2024 Revised: 12 September 2024 Accepted: 11 October 2024 Published: 30 October 2024

Abstract - *Brain stroke is a serious medical condition that occurs when the brain's blood supply is disrupted, either due to a blockage or the rupture of a blood vessel. This interruption results in a sudden loss of brain function, manifesting as symptoms like difficulty speaking, weakness or paralysis of the limbs, confusion, or altered consciousness. The severity of a stroke is influenced by both the duration of the blood flow disruption and the specific location of the damage within the brain. Immediate medical intervention is crucial for reducing damage and improving the chances of recovery. Key risk factors for stroke include hypertension, diabetes, obesity, a sedentary lifestyle, and smoking. Given the critical need for timely and accurate stroke diagnosis, this study introduces a novel Deep Learning (DL) model for detecting and classifying brain strokes using brain CT images. The proposed method combines DenseNet 201 and Capsule Network (CapsNet) models to enhance classification accuracy. Experimental results demonstrate that the model achieved an accuracy of 93.45%, a precision of 92.18%, a recall of 92.56%, and an F1 score of 92.36%, underscoring its effectiveness in diagnosing and classifying strokes with high accuracy.*

Keywords *- Brain Stroke, DenseNet 201, Capsule network, CT images, Medical imaging, Deep learning.*

1. Introduction

The brain is a highly complex and fascinating organ responsible for intelligence, emotion, memory, and creativity. It is divided into the frontal, occipital, parietal, and temporal lobes, which are connected to the rest of the body through the spinal cord [1]. Stroke, the second leading cause of death globally, demands immediate treatment to reduce the risk of death or severe long-term disability. Stroke occurs due to a rupture or blockage in blood vessels, leading to reduced blood flow to a specific brain area. Strokes are classified into two main types: ischemic and hemorrhagic.

A hemorrhagic stroke, which accounts for 13% of all strokes, is a critical neurological condition caused by the rupture of a blood vessel in the brain. Symptoms typically include a sudden and severe headache, vomiting, weakness on one side of the body or face, and a rapid decline in neurological function or consciousness. This condition is especially dangerous when a Subarachnoid Hemorrhage (SAH) occurs.

On the other hand, Ischemic Stroke (IS), which comprises 87% of strokes, results from an insufficient blood supply due to a clot or blockage in an artery in the neck or brain. This blockage leads to a lack of oxygen reaching the brain. The gold standard treatment for IS, particularly within the first three hours, is administering a thrombolytic or clot-busting drug [2].

Magnetic Resonance Imaging (MRI) and Computed Tomography (CT) are indispensable diagnostic tools for identifying brain tissue damage. In clinical practice, Computer-Aided Diagnosis (CAD) technologies significantly enhance the accuracy of disease detection, interpretation, and decision-making, thereby streamlining the diagnostic process. CT scans, in particular, are critical in emergencies, where their use of multiple X-ray beams and detectors allows for the rapid diagnosis of life-threatening brain abnormalities. Compared to other imaging techniques, CT is superior in trauma situations, offering detailed bone images and greater sensitivity in detecting acute hemorrhages. As a cost-effective option, CT is often favored for early disease screening, especially in patients at high risk of stroke [3]. CT scans are frequently employed to diagnose conditions such as brain strokes, hemorrhages, and skull fractures, providing quicker scanning times and reliable results.

Detecting and segmenting stroke lesions has traditionally relied on medical experts meticulously identifying lesion areas across multiple imaging slices. However, this manual process is often time-consuming, expensive, and prone to variability among different experts, which can compromise reliability and reproducibility. There is a growing demand in the medical field for efficient automated detection and classification methods to address these challenges. These methods can aid radiologists by providing quicker and more accurate stroke diagnoses, ultimately enhancing patient outcomes. This research seeks to meet this demand by introducing a novel deep learning-based model for automatically detecting and classifying brain strokes using CT images. The model's performance is evaluated using accuracy, precision, recall, and F1 score metrics, demonstrating its effectiveness in identifying brain strokes. Key contributions of the proposed work include:

- Developing a deep learning-based CAD system for detecting and classifying brain strokes.
- Enhancing the performance evaluation parameters of the system.
- Assessing the proposed system's performance in comparison to current approaches.

The rest of the paper is organized as follows: Section 2 provides a summary of the literature, highlighting areas that indicate a need for more investigation. In Section 3, the methodology is explained in depth. Section 4 goes into great detail about the results that the suggested strategy produced. Finally, a summary of the findings is included in Section 5, which gives a conclusion to the paper.

2. Related Works

Tursynova et al. (2023) [4] proposed a CAD system to distinguish cerebral strokes using CT images. To enhance categorization accuracy, they employed horizontal flip data augmentation techniques. Utilizing an image data generator, the researchers enhanced the images in real time, incorporating random alterations during training. To mitigate the risk of overfitting, they implemented an early stopping method. Additionally, they developed a Python web application to showcase the findings of their Convolutional Neural Network (CNN) model, utilizing cloud-based development approaches. In their current implementation, the model achieved a commendable accuracy rate of 79% in recognizing the normal class. This automated diagnostic approach shows the potential to assist medical practitioners with the identification and categorization of brain strokes.

Tripura et al. (2023) [5] introduced a hybrid model called BrainNet (BrN), which integrates CNN and Support Vector Machine (SVM) techniques for classifying brain stroke diseases. Their methodology involved designing the BrainNet model through a Deep Neural Network (DNN), encompassing data collection, preprocessing, feature extraction, and subsequent classification using SVM. The dataset used for this classification task was sourced from Kaggle. The proposed BrainNet model demonstrated an impressive accuracy of 91.91%, surpassing the performance of existing models.

Lee et al. (2023) [6] conducted a comprehensive study to evaluate the effectiveness of various CNN models for the detection and classification of Ischemic Strokes (IS) using hyperacute staged Diffusion-Weighted Images (DWI). Their research involved a dataset of 2,119 image slices, categorized into three groups: normal, Acute Cerebral Infarction (ACI), and Posterior Circulation Infarction (PCI). The researchers implemented two CNN models, EfficientNet-b0 and Inception-v3, alongside a self-derived modified LeNet model. Among these, Inception-v3 achieved the highest accuracy of 86.3%, followed closely by LeNet at 85.2%. EfficientNet-b0 demonstrated the lowest performance, with an accuracy of 83.6%. Additionally, Grad-Cam activation maps provided insights into the models' decision-making processes. The study identified key limitations, including the absence of a dedicated test set to comprehensively validate the CNN models' performance and the relatively small size of the dataset.

Peng et al. (2022) [7] aimed to enhance the accuracy of automated stroke identification in CT images by employing a CNN. Their methodology involved preprocessing the CT images to improve tissue clarity, adjusting positions, performing spatial normalization to a CT template, and creating t-score maps for each patient. These t-score maps facilitated the selection of both non-infarcted and infarcted patches. The team then applied data augmentation to generate additional patches for training and testing the CNN. The network achieved a remarkable 93% accuracy in detecting patches on the test set, showcasing the effectiveness of CNNs in analyzing medical images for rapid and precise stroke identification.

Zhang et al. (2022) [8] introduced the AC-YOLOv5, an advanced detection algorithm specifically designed for IS. This algorithm improved the feature detection of IS in Non-Contrast Computed Tomography (NCCT) images and identified the Region of Interest (ROI) using YOLOv5. Tested against other popular detection algorithms, AC-YOLOv5 achieved an impressive accuracy of 91.7%, demonstrating its robustness, accuracy, and generalizability in detecting IS on NCCT images.

Omarov et al. (2022) [9] explored a modified 3D UNet architecture to enhance the segmentation quality of IS in 3D CT images. Utilizing the ISLES 2018 dataset for model training and testing, they employed evaluation metrics such as the Jaccard index and Dice coefficient to assess segmentation accuracy. Their modified architecture achieved a Dice/f1 score similarity coefficient of 58%, surpassing the performance of the standard 3D UNet model. This result highlighted the model's effectiveness in accurately segmenting ischemic stroke lesions in CT images, indicating potential advancements in stroke diagnosis and treatment planning through improved medical imaging.

Gautam and Raman (2021) [10] aimed to categorize brain CT images into three distinct groups: IS, hemorrhagic stroke, and normal categories. They developed a CNN model that combined image fusion techniques with CNN methods, utilizing a recently proposed 13-layer CNN architecture to analyze preprocessed images. This classification process was conducted using a real dataset of CT images collected from the Himalayan Institute of Medical Sciences (HIMS). After performing 10-fold cross-validation on the dataset, they achieved a classification accuracy of 92.22%.

To identify invisible ischemic strokes from NCCT scans, Wu et al. (2021) [11] introduced a two-stage CNN-based approach. Their method featured a cascaded structure with two coordinated networks designed to detect suspicious stroke regions and optimize localization details. In the initial step, an end-to-end U-net with adaptive thresholding combined symmetry, gray texture information, and global positioning to identify suspicious areas. This was followed by a ResNet-based patch classification network aimed at reducing false positives by leveraging deeper image features. Finally, a MAP model optimized results by incorporating spatial constraint information from each patch's classification outcomes. Validated using 277 cases from two hospitals, their model achieved identification accuracies of 91.89%, 87.21%, and 85.71% across three experimental setups. This study highlights the potential of deep learning techniques in enhancing IS detection and localization on NCCT scans, marking significant progress in medical imaging technology.

Lo et al. (2021) [12] proposed a method for the automatic identification of Acute Ischemic Stroke (AIS) using Deep Convolutional Neural Networks (DCNNs) trained on NCCT images. Their dataset comprised grayscale NCCT images from AIS patients and healthy subjects. Utilizing a gold standard for training, they implemented the original AlexNet along with ResNet-101 and Inception-v3 models. Testing results indicated that AlexNet, ResNet-101, and Inception-v3 achieved accuracies of 81.77%, 80.89%, and 85.78%, respectively, demonstrating the potential of DCNNs for AIS identification, with Inception-v3 outperforming the other models.

Aishvarya et al. (2020) [13] tackled the issue of identifying ischemic stroke locations in MRI images, recognizing that the physical identification of lesions can be labor-intensive. Their work aimed to develop an automated stroke detection algorithm employing Machine Learning (ML) across six distinct stages. After preprocessing the images, they utilized Gabor filters for image enhancement and Adaptive Histogram Equalization (AHE) for further improvement. The fuzzy C-means technique was used for image segmentation, and features were extracted with the Gray Level Co-occurrence Matrix (GLCM). The multiclass SVM classifier yielded an accuracy of 90%, demonstrating the effectiveness of their approach.

Maya and Asha (2020) [14] proposed a method using various ML classification algorithms, including Decision Tree (DT), Maximum Expectation, Deep Neural Networks (DNN), Gaussian Naïve Bayesian Classifier, and Random Forest (RF), to assess stroke occurrence. They enhanced efficiency and scalability by employing Principal Component Analysis (PCA) to reduce features. Their results, comparing the DNN classifier to other ML methods regarding accuracy, specificity, and sensitivity, revealed performances of 86.42%, 88.49%, and 74.89%, respectively. However, the absence of MRI image datasets was a noted limitation, indicating opportunities for future research.

Subudhi et al. (2020) [15] introduced a CAD system for detecting ischemic strokes in DWI. Their technique classified brain strokes into three categories: Total Anterior Circulation Stroke (TACS), Partial Anterior Circulation Syndrome (PACS), and Lacunar Syndrome (LACS). They used the Expectation-Maximization (EM) algorithm for lesion segmentation and enhanced detection accuracy with the Fractional-Order Darwinian Particle Swarm Optimization (FODPSO) method. Their evaluation of 192 MRI scans demonstrated that SVM and RF classifiers effectively identified stroke lesions, with the RF classifier showing exceptional efficiency.

Kanchana and Menaka (2020) [16] developed a novel histogram bin-based algorithm for segmenting ischemic stroke lesions in CT images, focusing on optimal feature group selection to differentiate between normal and abnormal brain regions. They extracted features such as gray level run length matrix, first-order statistics, Hu's moments, and GLCM features for lesion characterization. The classification was carried out using Logistic Regression (LR), RF, Neural Network Classifier (NNC), and SVM classifiers. Their approach demonstrated effective IS lesion detection, underscoring the potential of advanced feature extraction and classification techniques in stroke diagnosis.

Gautam & Raman (2020) [17] introduced a new feature extraction method for classifying brain CT scan images, involving a multi-step process that compared neighboring pixels' intensities and computed double gradients within local neighborhoods. Histograms of these computed codes were concatenated into a single feature vector. Nine experiments with various classifiers revealed their method outperformed seven other feature extraction techniques, highlighting its potential to enhance CT scan image analysis in medical diagnostics.

Singh et al. (2019) [18] proposed a technique for differentiating normal tissue from ischemic stroke-affected areas using texture analysis of CT images. They selected five Regions of Interest (ROI) from potentially affected and unaffected areas and calculated 22 texture parameters for classification. The innovative aspect of their study was the use of ratios of five texture features for CT image classification, aiming to assist neurologists in early stroke detection. Preliminary results showed that their algorithm achieved an accuracy of 92%, indicating the potential of texture analysis in enhancing ischemic stroke diagnostics.

Detecting and classifying brain strokes from CT images remains a critical and challenging task due to various factors affecting diagnostic accuracy and timeliness. The potential impact on treatment outcomes underscores the urgency for prompt stroke diagnosis, as delays can significantly affect recovery. The subtlety of stroke symptoms can lead to confusion with other medical conditions, complicating accurate diagnosis. Furthermore, the expertise of medical professionals, access to high-quality imaging technologies like CT and MRI, and variations in stroke presentations further complicate the diagnostic landscape. In low-resource settings, imaging technologies may be less accessible or insufficiently sensitive to detect smaller or less severe strokes. Misdiagnosis of strokes as Transient Ischemic Attacks (TIAs) or other conditions, especially in younger individuals, adds to the challenges. The effectiveness of classification methods, such as Support Vector Machines (SVM), hinges on selecting suitable kernel functions; inappropriate choices can lead to poor performance or overfitting. The task of distinguishing between TIAs and actual strokes introduces additional complexity, underscoring the need for robust diagnostic tools in clinical practice.

3. Materials and Methods

The proposed model presents a robust framework for detecting and classifying brain strokes from CT images by integrating advanced deep-learning techniques. Initially, CT images undergo preprocessing and augmentation to standardize and enhance the data, ensuring its suitability for analysis. These enhanced images are then processed using a DenseNet 201 pre-trained model, leveraging its extensive feature extraction capabilities developed from a comprehensive dataset of CT images. The extracted features capture complex patterns specific to stroke detection, reshaped for compatibility with the subsequent capsule network. This network encodes spatial relationships and hierarchical structures of objects within the images. The output from the capsule network is directed through a dense layer that performs the final classification, determining whether the CT images indicate a stroke or are normal.

For interpretability, the model employs Gradient-weighted Class Activation Mapping (Grad-CAM) visualization, highlighting the regions in the CT images affected by stroke that influence the model's decision. This approach ensures transparency in the classification process, allowing for a clearer understanding of the model's outputs. The workflow for detecting and classifying brain strokes from CT images is illustrated in the block diagram shown in Figure 1.

3.1. Dataset

The dataset for the proposed research is sourced from a public repository, Kaggle [19]. It comprises a total of 2,515 images organized into three primary folders: training, testing, and validation. Each folder is further divided into two classes, namely "normal" and "stroke." Specifically, the dataset contains 1,843 images designated for training, 437 images allocated for testing, and 235 images reserved for validation. Figure 2 presents sample images from the dataset, showcasing examples from both the "normal" and "stroke" classes.

3.2. Data Preprocessing and Augmentation

In classifying brain strokes from CT images, preprocessing and data augmentation are crucial for improving the performance of Deep Learning (DL) models. Preprocessing involves standardizing the input data through various techniques to ensure consistency and enhance model efficiency. For this study, the 'rescale' parameter is used to normalize pixel values to a range between 0 and 1. This normalization step is essential as it adjusts the intensity values of the images, making the data more uniform and easier for the model to process.

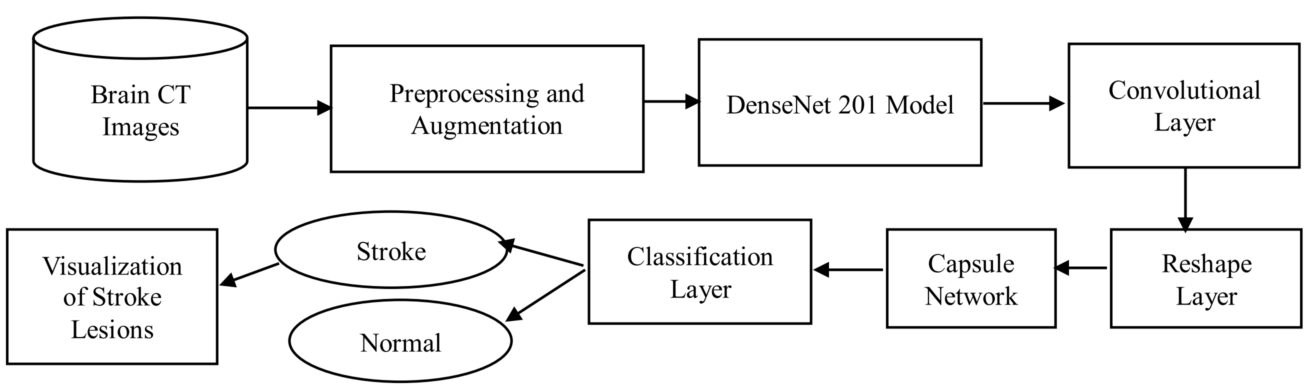
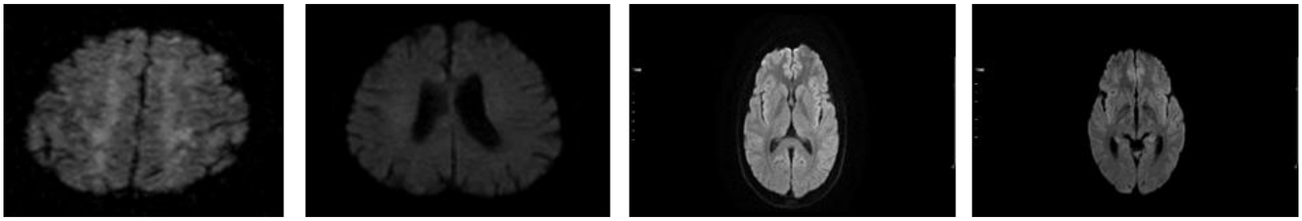
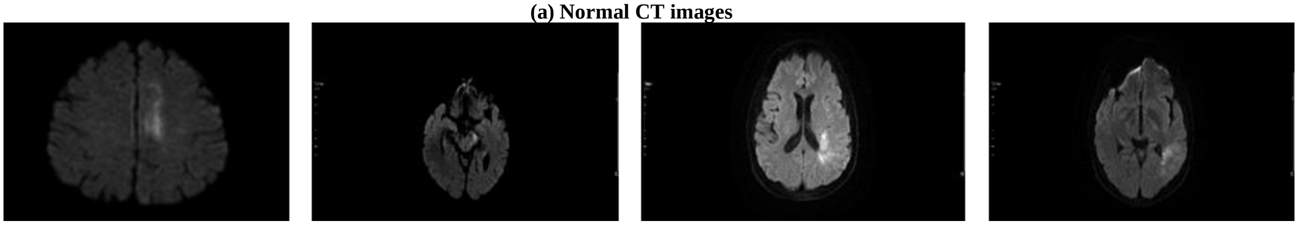


Fig. 1 Block diagram of proposed methodology



(a) Normal CT images



(b) Stroke CT images

Fig. 2 Sample images from the dataset

Data augmentation, on the other hand, introduces variability into the training dataset by applying a range of random transformations. This process includes shearing, zooming, horizontal and vertical flipping, and random rotations up to 30 degrees. These augmentations are applied to the images during the training phase, increasing the dataset's diversity. By simulating different possible variations of the input images, data augmentation helps the DL model to generalize better and become more robust to variations it may encounter in real-world scenarios.

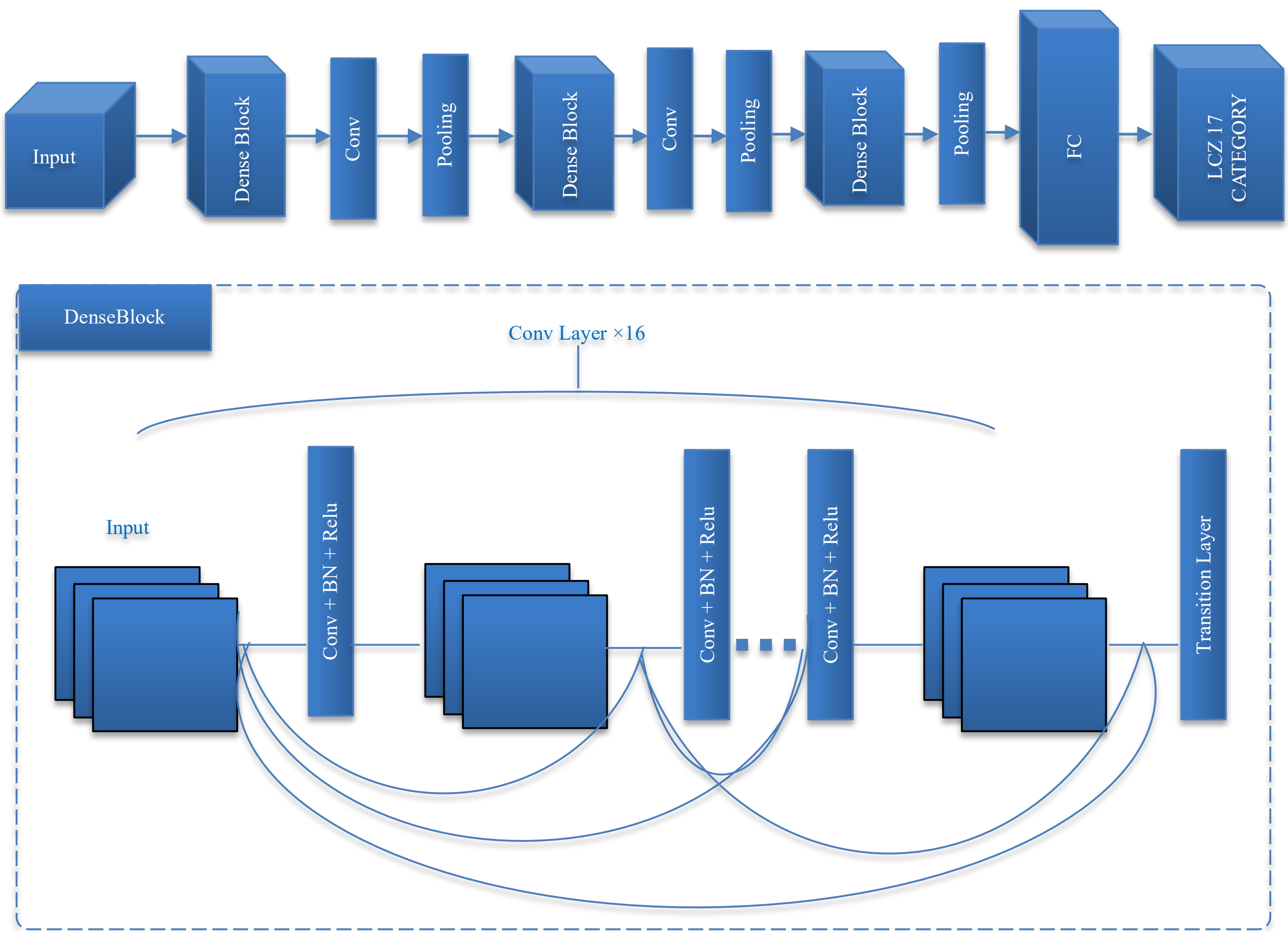


Fig. 3 Basic architecture of DenseNet 201

The dataset is split into training and testing sets using an 80:20 ratio, ensuring a substantial portion of data is reserved for training the model. In contrast, the remainder is used to evaluate its performance. This split is crucial for assessing the model's ability to generalize to new, unseen data.

3.3. Deep Learning Classifier

3.3.1. DenseNet 201

DenseNet is a deep learning architecture designed to improve information flow and gradients throughout the network. Unlike traditional architectures, each layer in DenseNet receives input from all preceding layers and contributes its own feature map to all subsequent layers. This approach enhances feature reuse and facilitates the learning of rich feature representations, which is crucial for tackling complex visual recognition tasks.

The DenseNet-201 model, illustrated in Figure 3, is engineered to optimize feature reuse and gradient flow. It includes several key components: initial layers, dense blocks, transition layers, and classification layers [20]. The dense blocks are integral to the architecture, enabling efficient information propagation and reducing the number of parameters while maintaining high performance. Transition layers help control the dimensionality and smooth the learning process, while the classification layers are responsible for outputting the final prediction. This design makes DenseNet-201 highly effective for a variety of visual recognition challenges.

DenseNet-201 is a sophisticated neural network architecture designed for complex image classification tasks. This architecture features 201 layers, including convolutional, pooling, batch normalization, and activation layers, which collectively enhance its ability to extract and process intricate features from input images.

The network begins with an input layer that accepts image data as pixel values. It then progresses through multiple dense blocks, each composed of several convolutional layers. Unlike traditional networks, each convolutional layer in DenseNet-201 receives input from all preceding layers, allowing for a rich network of feature reuse. This dense connectivity facilitates smooth gradient flow, which enhances the network's ability to transmit information effectively throughout its depth.

Transition layers, which include convolutional and pooling operations, are strategically placed between dense blocks to manage the dimensionality of feature maps. These transitions help maintain computational efficiency while ensuring the network can handle complex features. At the end of the architecture, a global average pooling layer compresses the spatial dimensions of the feature maps into a single vector of feature values, summarizing the learned information before final classification.

The m^{th} layer receives the feature maps from all previous layers. The output feature maps at the m^{th} layer is given by Equqation (1).

$$x_m = H_m\,[x_0, x_1, \dots \dots \dots \dots x_m - 1] \tag{1}$$

Where, $[x_0, x_1, \dots \dots \dots \dots x_m - 1]$ refers to the feature map concatenation. The composite function H_m consist of batch normalization, convolution and Rectified Linear Unit (ReLU).

The batch normalization layer ensures a uniform distribution of activations across the network by normalizing the activations.

$$X_k = \frac{x_k - \mu_b}{\sqrt{\sigma_b^2 + \epsilon}} \tag{2}$$

Where, μ_b is the mean and σ_b^2 is the variance. The training accuracy is improved by applying the ReLU activation function,

$$f(x) = \max(0, x) \tag{3}$$

The convolutional layer extracts features by performing a convolution operation on the input representation using a kernel. This process involves sliding the kernel, or filter, over the input image or feature maps to produce a set of output features. The kernel, a small matrix of weights, multiplies with the local region of the input it is currently covering. The results of these multiplications are then summed to produce a single value in the output feature map.

$$C(s,t) = \sum_{q=-p}^{p} \sum_{r=-p}^{p} I(s-q, t-r).K(q,r) + b \tag{4}$$

Where, b represents the bias term.

Bottleneck layers are employed to address the computational complexity associated with dense layers. These bottleneck layers utilize a 1×1 convolution, effectively reducing the number of parameters and computations required by compressing the feature maps before applying more complex operations.

The hyperparameters of the network, such as kernel size, stride, and padding, are carefully chosen to ensure that the dimensions of the feature maps remain consistent within each dense block. This consistency is crucial for maintaining the integrity of feature representations as they pass through the network.

Dimensionality reduction is managed by transition layers, which play a key role in controlling the size of feature maps between dense blocks. These transition layers typically include a 1×1 convolution to reduce the depth of the feature maps, followed by 2×2 average pooling layers to downsample the spatial dimensions. Additionally, batch normalization is applied within these layers to stabilize and accelerate the training process by normalizing the feature maps.

3.3.2. Capsule Network

Capsule Networks (CapsNet), as illustrated in Figure 4, were developed to overcome some of the inherent limitations of traditional CNNs, particularly their challenges in managing hierarchical relationships among features. CapsNet introduces a dynamic routing mechanism that allows adaptive connections between capsules across different layers, enhancing the network's ability to detect and classify brain tumors from CT scans.

In CapsNet, each capsule in the higher layers predicts the output of lower-layer capsules by assessing the alignment of their spatial configurations. This approach helps the network grasp the spatial hierarchy of features more effectively, which is crucial for accurately interpreting complex patterns in medical images. The highest level of the network integrates digit capsules, with each capsule representing a distinct class of brain tumor or normal tissue. This integration improves the network's capacity to recognize and categorize complex patterns with higher precision and interpretability than traditional CNN architectures, as highlighted by [21].

The dynamic routing algorithm enables the nonlinear mapping between two adjacent capsule layers. The capsule m in layer T attempts to predict the output of capsules n in layer $T+1$. Equation (5) provides the method for obtaining the expected feature vector matrix $\hat{x}(n|m)$ by applying a linear weight to the output of the capsule x_m in layer T. The weighting matrix W_{mn} is obtained through the process of back propagation.

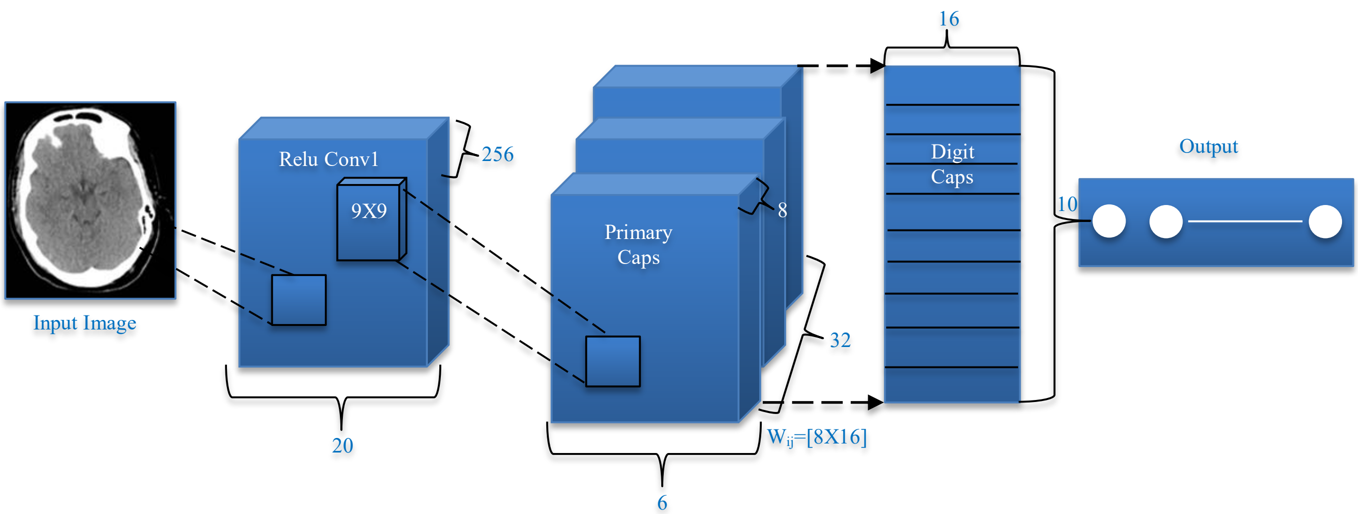


Fig. 4 Basic architecture of capsule network

$$\hat{x}(n|m) = W_{mn} \cdot x_m \quad (5)$$

Equation (6) is used to calculate the output for capsule n in layer $+1$.

$$s_n = \sum_m c_{mn} \cdot \hat{x}(n|m) \quad (6)$$

The coupling coefficient c_{mn} is determined using the softmax function, as outlined in Equation (7). This coefficient quantifies the strength of the connection between capsules in layer T and layer $T+1$. Initially, the variable b_{mn}, which represents this correlation, starts with a value of 0.

$$c_{mn} = \frac{\exp(b_{mn})}{\sum_k \exp(b_{mk})} \quad (7)$$

To refine these correlations, the system iterates through Equation (8) until the iteration conditions are met. This iterative process adjusts b_{mn} to optimize the coupling coefficients c_{mn}, thereby enhancing the effectiveness of the capsule network in capturing complex relationships between different layers and improving its overall performance.

$$b_{mn} = b_{mn} + \hat{x}(n|m) \cdot y_n \quad (8)$$

During each iteration, the output of capsule n is processed by the nonlinear squashing function, represented by Equation (9):

$$y_n = \frac{\|s_n\|^2}{1+\|s_n\|^2} \frac{s_n}{\|s_n\|} \quad (9)$$

Where, $\frac{s_n}{\|s_n\|}$ is a unit vector, that is the original vector length is adjusted proportionally $\frac{\|s_n\|^2}{1+\|s_n\|^2}$. The length of the output vector varies between 0 and 1 and is presented in a probabilistic approach. By scaling the length of the output vectors in this manner, longer vectors are assigned greater significance, while shorter vectors are given less importance. This approach helps emphasize more prominent features in the data, making highlighting key information easier and mitigating the influence of less significant features. The proposed hybrid DL model integrates DenseNet-201 with a Capsule Network to effectively identify and classify brain strokes from CT images. The model is designed to process input images with dimensions of 224x224 pixels. Initially, these images are fed into the DenseNet-201 architecture, which functions as the primary feature extractor, capturing intricate patterns and features from the input data. Following the DenseNet-201 processing, the output feature maps are passed through an additional convolutional layer equipped with 64 filters of size 1x1, coupled with a ReLU activation function. This convolutional layer enhances and refines the extracted features before they are input into the Capsule Network.

The CapsNet is configured with an input shape of 7x7 feature maps with 64 channels, indicating that the preceding convolutional layer's output is a set of 7x7 spatial feature maps with 64 channels. Within the CapsNet are 32 capsules, each with a dimensionality of 8. These capsules are designed to capture hierarchical relationships and spatial configurations present in the input images. The model's final step involves processing the CapsNet output through a Dense layer with a single neuron. This Dense layer applies a sigmoid activation function to produce a binary classification decision, yielding a probability value that indicates whether the input image is classified as "Normal" or "Stroke." The performance of this hybrid framework is assessed using various evaluation metrics, including accuracy, F1 score, recall, and precision. Figure 5 shows the model architecture of the proposed model. The algorithm for the proposed research is shown below.

Algorithm.1. Proposed brain stroke detection

Input: Brain CT image dataset, labels determine Stroke or Normal
Output: Predictions of whether the input image contains Stroke or Normal
Begin:

- ❖ Load and preprocess data:
  1. Collect dataset: C= $\{(A_i, b_i)$, where A_i is a brain CT image and $b_i \in \{0,1\}$ $b_i i \in \{0,1\}$ (0: Normal, 1: Stroke).
  2. Preprocess:
     - Resize: $A_i \rightarrow A'_i \in R^{224\times224}$
     - Normalize: $A'_i \rightarrow \frac{A'_i - \mu}{\sigma}$
     - Data Augmentation: $A'_i \rightarrow \{A''_i\}$(Shear, Zoom, Flipp (horizontal and vertical), Rotation)
- ❖ Define Base Models:
  1. Load Model: DenseNet 201
  2. Input: $224 \times 224 \times 3$
  3. Load Model: CapsNet

     Block 1: Conv2D ((3,3), activation='relu')
     MaxPooling2D (pool size= (2, 2))
     Block 2: Conv2D ((3,3), activation='relu')
     Block 3: Conv2D ((3,3), activation='relu')
     Global Average Pooling 2D ()
     Concatenate
     DenseNet ()
     Conv2D (64, (1,1), activation='relu')
     CapsNet ()
     Dense (1, activation='sigmoid')
- ❖ Model Compilation and Training:
  1. Compile each model M:
     Optimizer=Adam ()
     Loss=binary _crossentropy
     Metrics=[accuracy]
  2. Train: M.fit (X_{train} , y_{train} ,validation_data= (X_{val}, y_{val}))
- ❖ Model Evaluation and Comparison:
  1. Evaluate:
     metrics=M.evaluate(X_{test} , y_{test}), where metrics include accuracy, precision, recall.
- ❖ Save the Model:

End

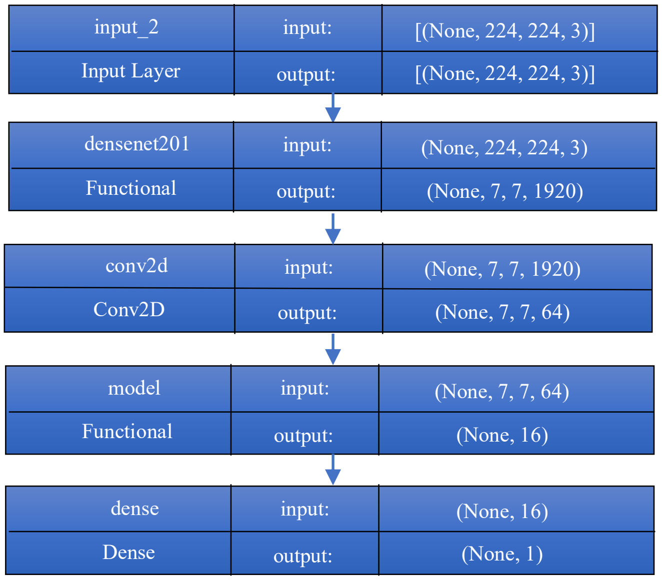


Fig. 5 Architecture of the proposed model

3.4. Software and Hardware Setup

The proposed model was developed and trained using Google Colaboratory, leveraging Python and the Keras framework. The Colab environment was equipped with TensorFlow, a Graphics Processing Unit (GPU), 12.75 gigabytes of Random Access Memory (RAM), 68.50 gigabytes of disk space, and a 64-bit edition of Windows 10. Python's flexibility and advanced capabilities, combined with its user-friendly syntax and extensive libraries, provided a robust foundation for building and training the model.

Its predictions were evaluated on the test dataset to assess the model's performance. The hyperparameters, crucial for optimizing the learning process, were determined through empirical methods. These parameters, detailed in Table 1, significantly influence the model's effectiveness. Various variables were examined and assessed to identify the configuration that yielded the highest classification performance, ensuring the model's efficacy in detecting and classifying brain strokes from CT images.

Table 1. Hyperparameters

Parameters	**Values**
Optimizer	Adam
Loss Function	Binary Crossentropy
Activation Function	ReLu, Sigmoid
Batch Size	32
Class Mode	Binary
Learning Rate	0.001
Number of Epochs	20

4. Results and Discussion

Accuracy and loss plots are essential tools for evaluating the performance of a Machine Learning (ML) model during its training and validation phases. These graphs provide critical insights into the model's ability to distinguish between "normal" and "stroke" conditions from brain CT images.

The accuracy plot, shown in Figure 6, illustrates how the model's accuracy evolves over training and validation epochs. Initially, the hybrid DL model starts with an accuracy of 58.98% on the training set and 68.09% on the validation set. This relatively low starting accuracy is expected due to the random initialization of the model's parameters. As training progresses, the model's accuracy improves substantially, reaching 96.46% on the training set and 84.26% on the validation set. This upward trend indicates that the model is effectively learning and generalizing from the data over time.

The loss plot, depicted in Figure 7, provides a view of the model's loss function changes throughout the training and validation stages. At the beginning of training, the model exhibits a high loss of 0.6697, reflecting the model's early stage of learning and its struggle with significant prediction errors. By the final epoch, the loss decreases markedly to 0.3073, demonstrating that the model has learned to classify the input data with greater precision and reduced error. Overall, these plots confirm that the model improves its performance progressively, indicating successful learning and adaptation to the task of classifying brain strokes from CT images.

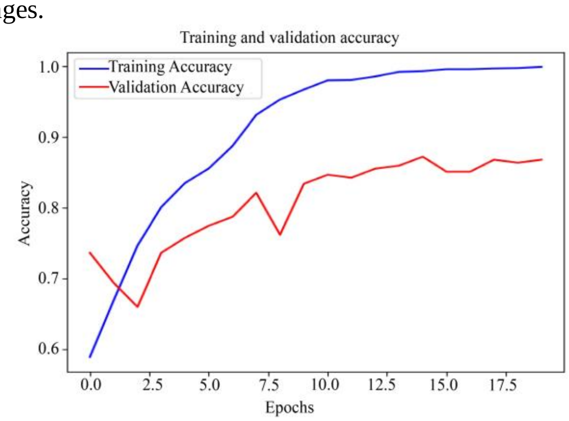


Fig. 6 Accuracy plot

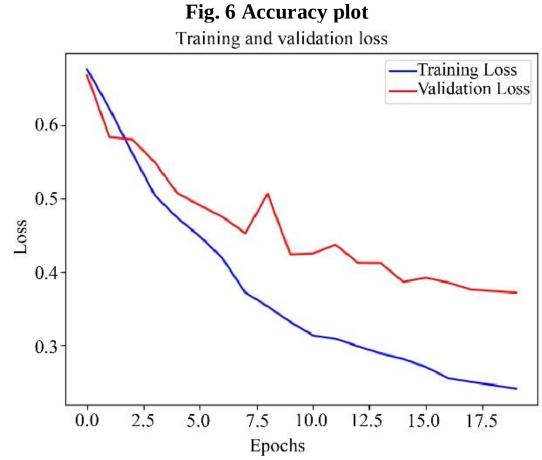


Fig. 7 Loss plot

In order to thoroughly evaluate the efficacy and operational efficiency of the proposed model, the F1-score, accuracy, precision, and recall are the four primary metrics utilized. These measures, which are based on the concepts of False Positive (FP), False Negative (FN), True Negative (TN), and True Positive (TP), are essential for assessing the model's performance. These performance parameters have mathematical formulations that are shown in Equations (10), (11), (12) and (13).

$$Accuracy = (TP + TN) / (TP + TN + FP + FN) \quad (10)$$

$$Precision = TP / (TP + FP) \quad (11)$$

$$Recall = TP / (TP + FN) \quad (12)$$

$$F1\,Score = 2 * (Precision * recall) / (Precision + recall) \quad (13)$$

The model demonstrates exceptional performance, achieving an overall accuracy of 93.45%. This high accuracy reflects its ability to correctly classify the majority of cases within the dataset. Additionally, the model's precision and recall scores are 92.18% and 92.56%, respectively, highlighting its effectiveness in accurately identifying positive instances while minimizing false positives. The F1-score, which balances both precision and recall, is 92.36%, providing a comprehensive measure of the model's reliability and effectiveness. These metrics collectively underscore the model's robust capability in classifying data into "Normal" or "Stroke" categories. Table 2 and Figure 8 shows the performance evaluation of the proposed model.

Table 2. Performance evaluation of the proposed model

Performance Metrics	Obtained Results
Accuracy	93.45%
Precision	92.18%
Recall	92.56%
F1-Score	92.36%

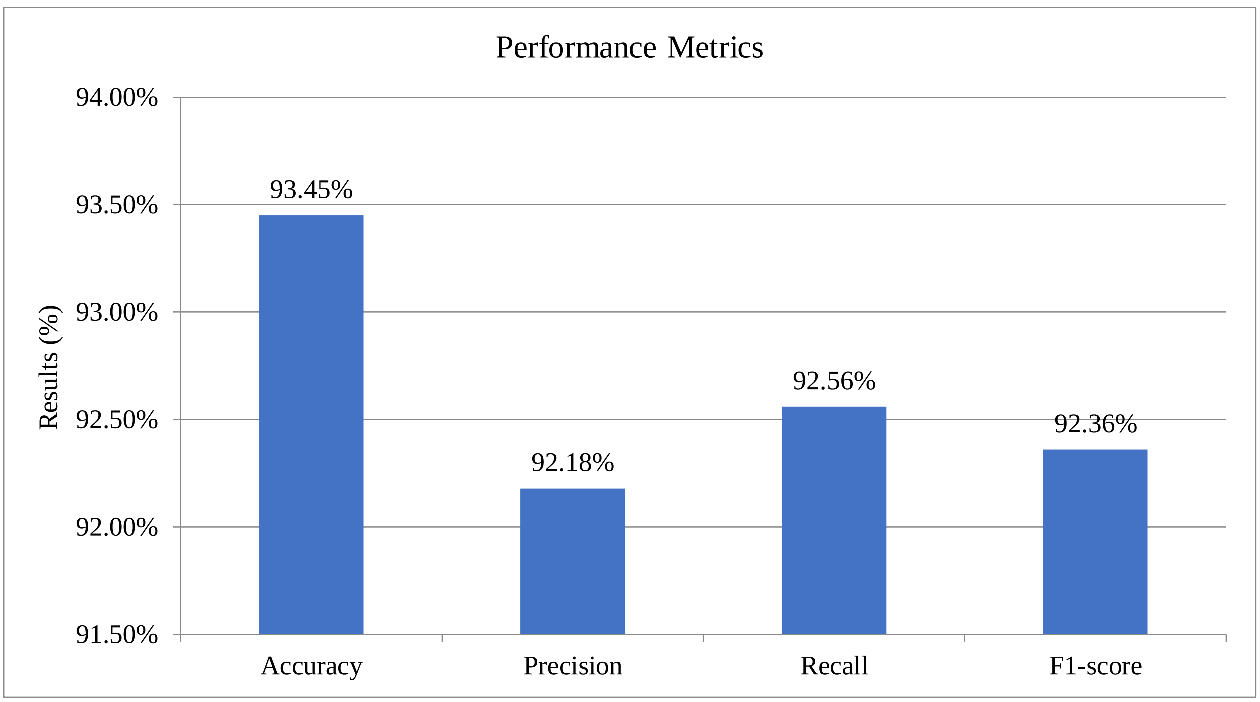


Fig. 8 Performance evaluation of the proposed framework

The effectiveness of a classification algorithm is critically evaluated using a confusion matrix, a tool that provides detailed insights into the model's performance. A confusion matrix is a performance measurement framework that summarizes the results of a classification algorithm. It consists of four key components: true positives, true negatives, false positives, and false negatives. In the context of brain stroke detection, the matrix helps to visualize how well the model has classified images into the correct categories. Figure 9 depicts the confusion matrix used for categorizing brain strokes from CT images, which plays a crucial role in refining the model's accuracy in distinguishing between stroke and normal images. In the given confusion matrix, the model accurately predicted 280 images as normal and 112 images as stroke. However, the matrix also reveals some misclassifications: 18 stroke images were incorrectly labeled as normal, and 27 normal images were mistakenly classified as stroke.

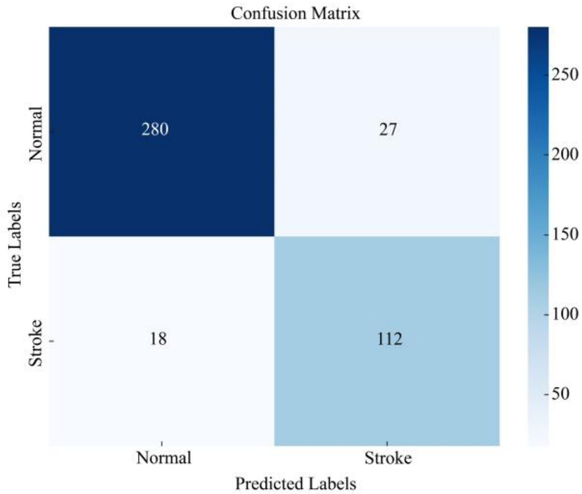


Fig. 9 Confusion matrix

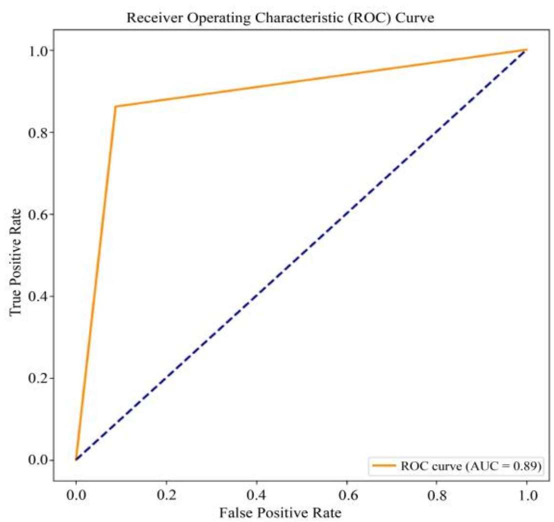


Fig. 10 ROC curve

By analyzing these errors, the confusion matrix aids in identifying patterns of misclassification and provides insights for model refinement. Addressing these inaccuracies is essential for enhancing the diagnostic accuracy of the algorithm, ultimately leading to better performance in stroke detection from brain CT scans.

The Receiver Operating Characteristic (ROC) curve depicted in Figure 10 serves as a visual tool to assess the diagnostic performance of a binary classifier. This curve illustrates the trade-off between the true positive rate (sensitivity) and the false positive rate (1-specificity) across various threshold settings. By plotting these rates against each other, the ROC curve provides a comprehensive view of the model's ability to distinguish between positive cases (stroke) and negative cases (normal).

A crucial aspect of the ROC curve is the Area Under the Curve (AUC), which offers a single numerical value summarizing the classifier's overall effectiveness. An AUC of 1 indicates perfect classification, meaning the model accurately identifies all stroke and normal cases without error. Conversely, an AUC of 0.5 suggests no discriminatory power akin to random guessing. In the context of brain stroke detection, a high AUC value is desirable as it reflects the model's strong capability to correctly differentiate between stroke and normal CT images.

The ROC curve also aids in determining the optimal threshold that balances sensitivity and specificity, which is critical for enhancing the model's clinical relevance. By selecting an appropriate threshold, the model can achieve an ideal trade-off between minimizing false positives (incorrectly classifying a normal image as a stroke) and maximizing true positives (correctly identifying stroke cases). This balance is essential for ensuring that the model's predictions are both accurate and actionable in real-world medical settings.

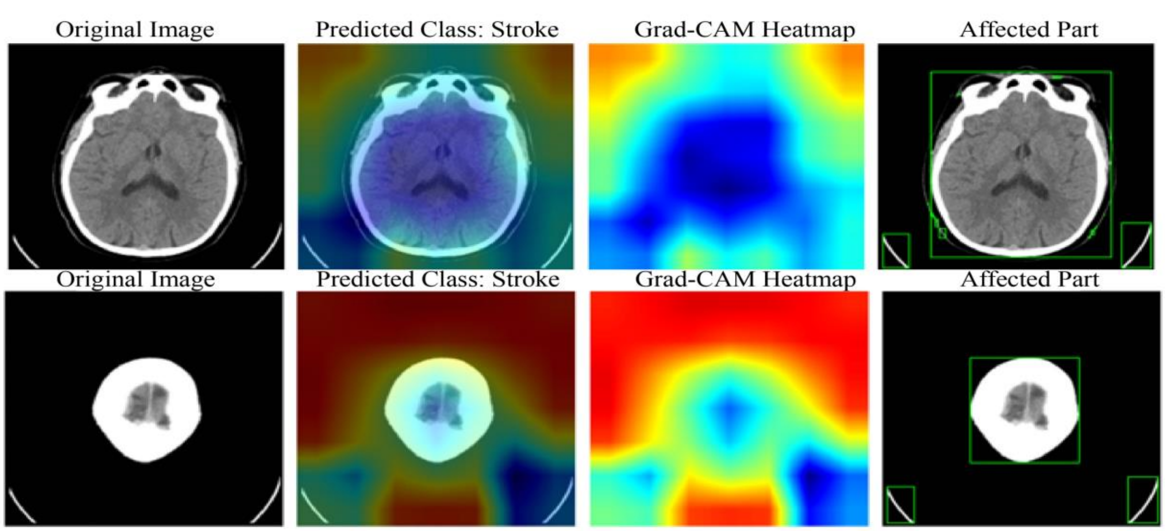


Fig. 11 Visualization using the Grad-CAM approach

The Gradient-weighted Class Activation Mapping (Grad-CAM) technique is a powerful tool for visualizing Regions of Interest (ROI) in brain CT images, particularly for identifying areas indicative of a stroke, as shown in Figure 11. This deep learning method is particularly effective when applied to CNNs, allowing for identifying critical regions within an input image that contribute most significantly to the network's prediction of a particular class.

Grad-CAM is recognized as a class-discriminative localization method that provides visual explanations for CNN-based networks without requiring architectural modifications. It achieves this by examining the gradients that flow into the last convolutional layer of the CNN. These gradients generate a heat map highlighting the image's most important areas for the network's classification decision. Essentially, the technique involves calculating the gradient of the predicted class score with respect to the feature maps of the final convolutional layer. This allows Grad-CAM to identify the significance of each feature map for the given class, thereby pinpointing the regions of the image that are most influential in determining the network's output.

Through this approach, Grad-CAM helps in understanding the decision-making process of deep learning models and enhances their interpretability, making it easier for clinicians and researchers to trust and validate the predictions made by CNNs in critical applications like stroke detection from brain CT images.

When an image is randomly selected from the brain CT dataset, it undergoes classification using the proposed model, determining whether the image falls under the "stroke" or "normal" category. Moreover, Figure 12 presents an example of an image processed by the model, where the prediction outcome is classified as "No Stroke." This highlights the model's ability to correctly identify cases without any stroke indications, further underscoring its accuracy in medical image classification. The consistent performance across various test images reflects the robustness of the model in distinguishing between normal and stroke-affected brain scans, making it a valuable tool in clinical settings for early stroke detection.

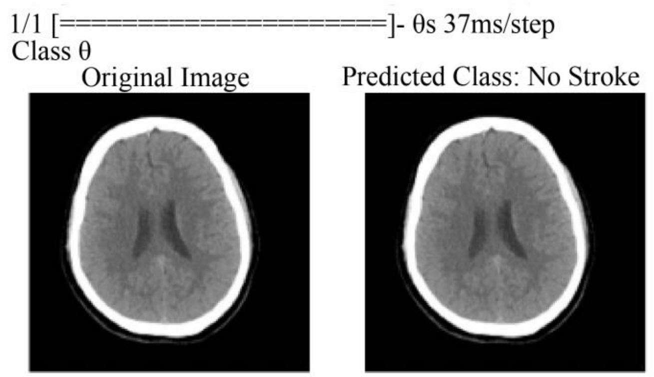


Fig. 12 Prediction output

Table 3 provides a detailed comparison of various methodologies and models used for detecting and classifying brain strokes across different imaging modalities, focusing on the accuracy achieved by each approach. Among these, the proposed hybrid model, which integrates DenseNet-201 and CapsNet, stands out by achieving the highest accuracy of 93.45% when applied to CT images. This result underscores the model's superior performance compared to other existing approaches.

For instance, a Computer-Aided Diagnosis (CAD) system that leverages Convolutional Neural Networks (CNNs) to distinguish between different types of brain strokes attains an accuracy of 79%. On the other hand, the BrainNet model, which combines CNN with Support Vector Machines (SVM), demonstrates a higher accuracy of 91.91% using CT scans, reflecting the effectiveness of hybrid models in stroke detection.

Additionally, assessments of various CNN architectures, including EfficientNet-b0 and Inception V3, reveal that Inception V3 achieves an accuracy of 86.3% when applied to Diffusion-Weighted Imaging (DWI) images. Furthermore, the AC-YOLOv5 model, specifically designed to enhance the detection of Ischemic Stroke (IS) features in Non-Contrast Computed Tomography (NCCT) images, achieves a notable accuracy of 91.7%. Another approach employing Deep Convolutional Neural Networks (DCNNs) for Acute Ischemic Stroke (AIS) detection using NCCT images with the Inception-v3 model reports an accuracy of 85.78%.

Table 3. Performance comparison of the proposed model with other models

Reference	Methodology	Images	Accuracy
[4]	CNN	CT	79%
[5]	BrainNet	CT	91.91%
[6]	Inception V3, EfficientNet-b0	DWI	86.3% (Inception V3)
[8]	AC-YOLOv5	NCCT	91.7%
[12]	AlexNet, Inception-v3, ResNet-101	NCCT	85.78% (Inception-v3)
[13]	Multiclass SVM	MRI	90%
Proposed Hybrid DenseNet 201-CapsNet		**CT**	**93.45%**

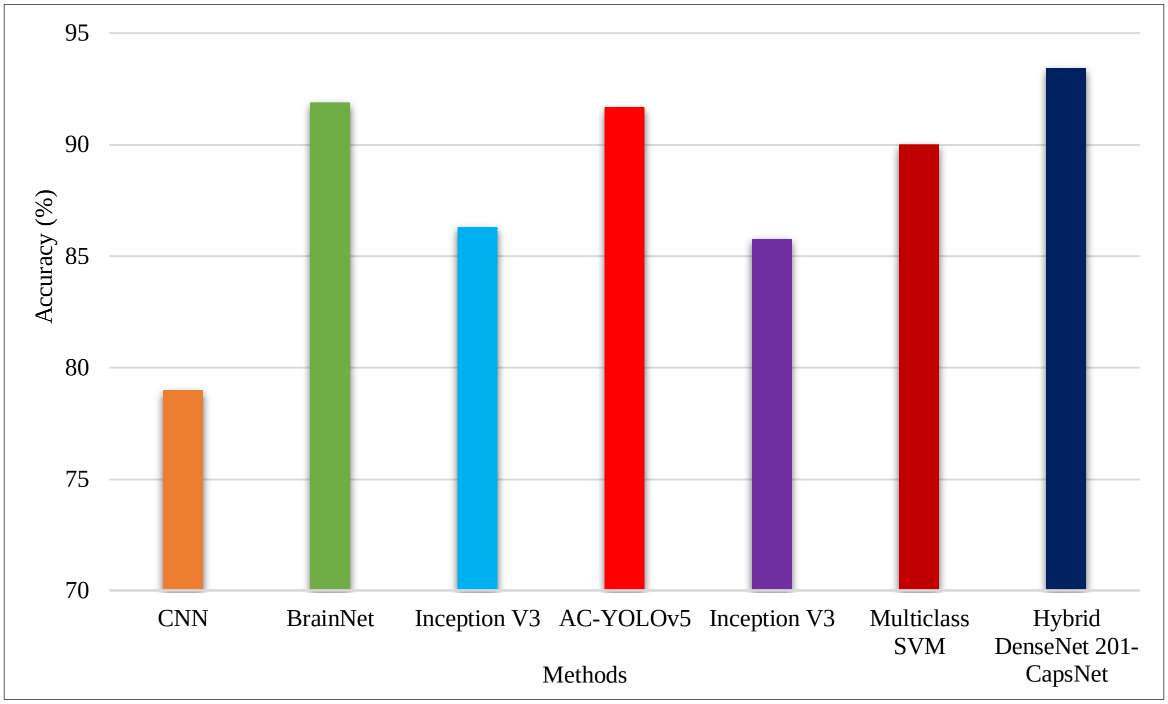


Fig. 13 Performance comparison of the proposed model with other models

Moreover, a computerized stroke detection procedure that utilizes a six-step approach incorporating multiclass SVM on MRI images achieves an accuracy level of 90%. The graphical representation in Figure 13 visually compares the performance of these existing methods with the proposed DenseNet-201 and CapsNet hybrid model, highlighting the advancements made in stroke detection accuracy through the use of this novel approach.

5. Conclusion

Brain stroke is a critical medical condition that develops when the blood supply to the brain is interrupted. Prompt detection is essential to reduce the severity of the condition's effects. Conventional methods for identifying and classifying brain strokes from CT images depend on manual analysis, which causes time consumption and is prone to human error. In this study, the hybrid model, which combines DenseNet-201 and CapsNet, demonstrates promising performance for brain stroke classification from brain CT images. This model attains a notable result with 93.45% accuracy, 92.18% precision, 92.56% recall, and 92.36% F1 score. These metrics indicate the efficiency of the proposed model in precisely identifying stroke cases while minimizing false positives. The hybrid architecture uses DenseNet-201 for feature extraction and CapsNet for hierarchical representation learning. It can capture local and global features, making it easier to correctly classify brain stroke from CT scans. The high performance of this hybrid model highlights its potential as a valuable tool in clinical settings, assisting healthcare professionals with timely and accurate diagnosis and treatment planning for patients with brain stroke. This advanced approach promises to enhance diagnostic accuracy and improve patient outcomes, making it an essential resource in modern medical practice.

Acknowledgements

The author deeply thanks the supervisor for offering guidance and constant encouragement during this study.

References

[1] Lavdim Kurtaj, Vjosa Shatri, and Ilir Limani, "Cerebellar Model Controller with New Model of Granule Cell-Golgi Cell Building Blocks and Two-Phase Learning Acquires Multitude of Generalization Capabilities in Controlling Robot Joint Without Exponential Growth in Complexity," *International Journal of Electrical and Computer Engineering*, vol. 8, no. 6, pp. 4292-4309, 2018. [CrossRef] [Google Scholar] [Publisher Link]

[2] Michael S. Phipps, and Carolyn A. Cronin, "Management of Acute Ischemic Stroke," *BMJ*, vol. 368, 2020. [CrossRef] [Google Scholar] [Publisher Link]

[3] Saleh Salehi Zahabi et al., "Evaluation of Causes of Brain CT Scan in Patients with Minor Trauma," *International Journal of Surgery Open*, vol. 27, pp. 220-224, 2020. [CrossRef] [Google Scholar] [Publisher Link]

[4] Azhar Tursynova et al., "Deep Learning-Enabled Brain Stroke Classification on Computed Tomography Images," *Computers, Materials & Continua*, vol. 75, no. 1, pp. 1431-1446, 2023. [CrossRef] [Google Scholar] [Publisher Link]

[5] Dhonita Tripura et al., "A BrainNet (BrN) Based New Approach to Classify Brain Stroke from CT Scan Images," *2023 International Conference on Advancement in Computation & Computer Technologies (InCACCT)*, Gharuan, India, pp. 1-6, 2023. [CrossRef] [Google Scholar] [Publisher Link]

[6] Kun-Yu Lee et al., "Automatic Detection and Vascular Territory Classification of Hyperacute Staged Ischemic Stroke on Diffusion Weighted Image Using Convolutional Neural Networks," *Scientific Reports*, vol. 13, pp. 1-11, 2023. [CrossRef] [Google Scholar] [Publisher Link]

[7] Syu-Jyun Peng et al., "Automated Cerebral Infarct Detection on Computed Tomography Images Based On Deep Learning," *Biomedicines*, vol. 10, no. 1, pp. 1-14, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[8] Lifeng Zhang et al., "An Improved Detection Algorithm for Ischemic Stroke NCCT Based on YOLOv5," *Diagnostics*, vol. 12, no. 11, pp. 1-10, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[9] Batyrkhan Omarov et al., "Modified UNet Model for Brain Stroke Lesion Segmentation on Computed Tomography Images," *Computers, Materials & Continua*, vol. 71, no. 3, pp. 1-17, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[10] Anjali Gautam, and Balasubramanian Raman, "Towards Effective Classification of Brain Hemorrhagic and Ischemic Stroke Using CNN," *Biomedical Signal Processing and Control*, vol. 63, 2021. [CrossRef] [Google Scholar] [Publisher Link]

[11] Guoqing Wu et al., "Identification of Invisible Ischemic Stroke in Noncontrast CT Based on Novel Two-Stage Convolutional Neural Network Model," *Medical Physics*, vol. 48, no. 3, pp. 1262-1275, 2021. [CrossRef] [Google Scholar] [Publisher Link]

[12] Chung-Ming Lo, Peng-Hsiang Hung, and Daw-Tung Lin, "Rapid Assessment of Acute Ischemic Stroke by Computed Tomography Using Deep Convolutional Neural Networks," *Journal of Digital Imaging*, vol. 34, no. 3, pp. 637-646, 2021. [CrossRef] [Google Scholar] [Publisher Link]

[13] R. Aishvarya et al., "Early Detection of Brain Stroke Using MRI Images," *International Research Journal of Engineering and Technology*, vol. 7, no. 9, pp. 2258-2261, 2020. [Google Scholar] [Publisher Link]

[14] B.S. Maya, and T. Asha, "Predictive Model for Brain Stroke in CT Using Deep Neural Network," *International Journal of Recent Technology and Engineering*, vol. 9, no. 1, pp. 2011-2017, 2020. [CrossRef] [Google Scholar] [Publisher Link]

[15] Asit Subudhi, Manasa Dash, and Sukanta Sabut, "Automated Segmentation and Classification of Brain Stroke Using Expectation-Maximization and Random Forest Classifier," *Biocybernetics and Biomedical Engineering*, vol. 40, no. 1, pp. 277-289, 2020. [CrossRef] [Google Scholar] [Publisher Link]

[16] R. Kanchana, and R. Menaka, "Ischemic Stroke Lesion Detection, Characterization and Classification in CT Images with Optimal Features Selection," *Biomedical Engineering Letters*, vol. 10, pp. 333-344, 2020. [CrossRef] [Google Scholar] [Publisher Link]

[17] Anjali Gautam, and Balasubramanian Raman, "Local Gradient of Gradient Pattern: A Robust Image Descriptor for the Classification of Brain Strokes from Computed Tomography Images," *Pattern Analysis and Applications*, vol. 23, pp. 797-817, 2020. [CrossRef] [Google Scholar] [Publisher Link]

[18] Mandeep Singh, Varinder Garg, and Parmod Bhat, "Early Detection of Stroke Using Texture Analysis," *Indian Journal of Forensic Medicine & Toxicology*, vol. 13, no. 3, pp. 49-52, 2019. [CrossRef] [Google Scholar] [Publisher Link]

[19] Brain Stroke Prediction CT Scan Image Dataset, Kaggle. [Online]. Available: https://www.kaggle.com/datasets/alymaher/brain-stroke-ct-scan-image/data

[20] Poonguzhali Elangovan, "Detection of Glaucoma from Fundus Image Using Pre-Trained Densenet201 Model," *Indian Journal of Radio & Space Physics*, vol. 50, no. 1, pp. 1-7, 2021. [Google Scholar] [Publisher Link]

[21] Tianming Zhao et al., "A Deep Learning Iris Recognition Method Based on Capsule Network Architecture," *IEEE Access*, vol. 7, pp. 49691-49701, 2019. [CrossRef] [Google Scholar] [Publisher Link]