

Hem-CNN: An Efficient Intracranial Hemorrhage Detection Model Using Explainable Deep Learning in Head CT Scan Images

Mahfujur Rahman^{1*}, Noboranjan Dey¹, Mehedi Hasan², Rahul Biswas¹, Rukaiya Jahan Sajuti³, Tanvir Kazi¹, Dipta Gomes¹, Rajarshi Roy Chowdhury¹

¹Department of Computer Science, American International University-Bangladesh (AIUB), Dhaka, Bangladesh

²Department of Electrical and Electronics Engineering, American International University-Bangladesh (AIUB), Dhaka, Bangladesh

³Department of Management Studies, Jagannath University (JnU), Dhaka, Bangladesh

Email: *mahfuj@aiub.edu, noboranjan@aiub.edu, mehedi@aiub.edu, rbiswas@aiub.edu, r.jahan232@gmail.com, kazi.tanvir@aiub.edu, diptagomes@aiub.edu, rajarshi@aiub.edu

How to cite this paper: Rahman, M., Dey, N., Hasan, M., Biswas, R., Sajuti, R.J., Kazi, T., Gomes, D. and Chowdhury, R.R. (2025) Hem-CNN: An Efficient Intracranial Hemorrhage Detection Model Using Explainable Deep Learning in Head CT Scan Images. *Journal of Computer and Communications*, 13, 128-146.

<https://doi.org/10.4236/jcc.2025.1310008>

Received: August 28, 2025

Accepted: October 21, 2025

Published: October 24, 2025

Copyright © 2025 by author(s) and Scientific Research Publishing Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

Abstract

Intracranial hemorrhage (ICH) is a critical subtype of stroke that arises from bleeding within the brain and often leads to severe neurological damage or death if not diagnosed at an early stage. The rapid rise in stroke incidence worldwide, combined with limited access to specialized radiologists, necessitates the development of automated and efficient detection systems that can support early intervention. Leveraging deep learning in medical imaging can significantly reduce diagnostic delays, improve treatment outcomes, and lower stroke-associated mortality. In this study, we propose Hem-CNN, a deep convolutional neural network built upon a pretrained EfficientNetB2 backbone with customized convolutional and pooling layers, designed specifically for stroke-related CT images. Extensive preprocessing techniques, including noise reduction, CLAHE contrast enhancement, and histogram normalization, were applied to ensure image quality. The model achieved an accuracy of 97.80%, precision of 96.86%, recall of 97.37%, specificity of 98.06%, and an F1-score of 97.10%, outperforming models such as VGG19, ResNet152, and EfficientNetB2. The novelty of this study lies in the integration of a lightweight yet highly discriminative CNN architecture with robust preprocessing tailored for medical imaging variability, alongside explainability through Grad-CAM visualization for clinical interpretability. Despite promising performance, the study was limited by dataset size and diversity, highlighting the need for larger multi-institutional validation to ensure broader clinical applicability.

Keywords

Intracranial Hemorrhage, Deep Learning, Convolutional Neural Network, Medical Imaging, Explainable AI, Hem-CNN

1. Introduction

In medical terms, a stroke, mostly known as a cerebrovascular accident (CVA), occurs due to unsteady blood flow to the cerebral arteries, brainstem, and cerebellum of brain tissue. It is a life-threatening medical illness in which blood flow to a portion of the brain is cut off. The lack of blood flow in the brain tissue is also responsible for damage or death of brain tissues; this is why a stroke is called a brain attack [1]. It is one of the most significant life-threatening illnesses, resulting in lifelong impairments or death. Stroke is ranked 5th among the leading causes of disability and death in the United States [2] [3].

Stroke is a medical emergency that can lead to death or permanent disability. It is a life-threatening disease that can lead to death in some cases. There are several types of strokes, and the three main types are ischemic, hemorrhagic, and transient ischemic attacks [4]. Ischemic stroke: An ischemic stroke is mainly generalized from a stroke in which there is impaired blood flow to the brain. Individual brain cells die within a minute due to a lack of blood flow. A report stated that this type of stroke occurs in 85% of the world [5]. This type of stroke is hypotensive [6]. Hemorrhagic Stroke: Cerebral hemorrhage is caused by cerebrum spillages. Hemorrhagic stroke occurs due to bleeding in the brain. It may originate in the brain and spread to other parts of the body. The second type of stroke is a less common hemorrhagic stroke with the potential to invade the territory [7]. Transient ischemic stroke: A mini-stroke occurs when the blood flow to the brain is interrupted. This can also occur in the presence of air bubbles. This lasts for a short period of time. It generally takes 1 - 5 minutes, but in rare cases, it takes nearly 24 h [8].

In the deep learning (DL) approach, knowledge acquisition is facilitated by the iterative processing of data across numerous layers within a computational model known as a deep neural network (DNN). Machine learning (ML) and DNNs are considerably more efficient in detecting the type of stroke and the possibility of a stroke. Stroke pre-identification and specification are challenging in the current era. A significant issue is that the physical structure of humans, such as the size, shape, and structure of the human brain, differs from that of other species. Consequently, the shapes, locations, and structures of medical anomalies such as brain strokes are highly variable. Recent deep learning systems have demonstrated near-expert performance in automatic ICH detection and subtype classification, with AUC values close to 0.99 [9].

Collaboration with healthcare professionals is essential for precise medical image analysis. Stroke can easily worsen a patient's condition if they have cardiac,

kidney, brain, or other related health problems. The involvement of modern technology in the medical sector is increasing. Medical treatment depends on a better understanding and identification of the disease itself. The application of artificial intelligence (AI) and ML makes our lives easier. AI-assisted triage has already shown potential to reduce the time to treatment and improve patient outcomes in real-world emergency settings [10]. Thus, enhancing healthcare is crucial for improving human life. DNNs can help the medical industry detect early stages of stroke from medical images, such as magnetic resonance imaging (MRI), computed tomography (CT) images, and X-radiation (X-ray). The main goal of this study was to detect the current stage of stroke using convolutional neural networks (CNN) and DL approaches before severe medical analysis. The traditional method for solving this problem was considered, and this study included a summary of a few techniques for solving similar problems in medicine. This study observes and implements the approaches discussed and emphasized.

In the present world, the doctor-patient ratio is 1:250 [11], where four doctors are assigned to thousand patients. The above statistics show that a limited number of people are involved in medical science and other related fields. However, the mortality rate associated with stroke is considerably higher than that of other major diseases. Additionally, the shortage of doctors specializing in stroke care creates significant challenges for patients. In this scenario, rural people cannot receive proper treatment for viral diseases in the early stages. Stroke and other medical-related documents are much easier for this specific medical professional to understand than for laypersons to understand. Medical science is rapidly adapting to modern technology. With the rapid growth of advanced technology, other medical professionals can easily understand diseases and suggest specific treatments. We aim to improve the CNN model so that it can efficiently detect stroke images and describe the related objectives here.

- A model has been proposed that improves accuracy by training with a large amount of image data with adequate pre-processing.
- A pretrained EfficientNetB2 with custom layers, including multiple convolution layers with Rectified Linear Unit (ReLU) activation, has been presented, where each layer is followed by maxpooling for feature extraction, reducing the spatial dimensions that will reduce computational cost and prevent overfitting.
- Providing comprehensive external validation across multiple independent datasets to demonstrate generalizability.
- The proposed Hem-CNN model is enhanced by detecting essential features such as edges, corners, etc. In addition, we developed an efficient model for hemorrhagic stroke-related CT images.

The remainder of this paper is organized as follows: Section 2 presents the Literature Review; Section 3 presents the proposed methodology, including image preprocessing and details of the proposed method; Section 4 presents an overview of the datasets and the results; and finally, the conclusion is presented in Section 5.

2. Literature Review

Recent research has used DL and ML in medical imaging. This has driven the adoption of DL and ML methodologies to ensure the precise and effective delivery of medical care [12]. It is essential to acknowledge the sensitivity of medical-image processing. Therefore, DL and ML are best used to teach machines to identify disease patterns. Image recognition using DNNs has recently been successful. It is also used in medical science to develop advanced technologies. Additionally, numerous other popular deep learning (DL) and machine learning (ML) techniques are available and have been utilized in medical research.

The emergence of ML and DNN has enabled researchers to address real-world problems effectively. Noboranjana *et al.* [13] employed data augmentation and feature extraction techniques to ensure user authentication and prevent unauthorized access, thus maintaining data security. Rahman *et al.* [14] [15] take advantage of traditional ML models to address everyday challenges and enhance users' quality of life. Similarly, Anjir *et al.* [16] applied transfer learning to demonstrate the impact of user behavior on the COVID-19 pandemic. Machine learning models trained on CT imaging have also been shown to effectively classify ICH, highlighting the potential of both traditional and deep learning techniques [17]. The rapid advancement of technology and the growth of the ML field can effectively support and promote a healthy lifestyle.

Cheon *et al.* [12] introduced a predictive model utilizing DL to identify factors affecting stroke mortality in the Korean population. The researchers used stroke data from 15,099 patients obtained from the Korean National Hospital Discharge In-depth Injury Survey (KNHDS). In addition, they applied principal component analysis (PCA) featuring quantile scaling to extract relevant background characteristics from the patient data set. It achieved 84.03% accuracy, where they considered the Area Under the Curve (AUC) value as their primary parameter, generating a value of 83.48%.

A recent study introduced a deep learning algorithm for acute ICH detection, showing that AI assistance significantly improved radiologists' performance [18]. B. R. Gaidhani *et al.* [19] conducted ML research that utilized two different CNN architectures, LeNet and SegNet, which were used in distinct phases. The LeNet architecture was employed in the classification phase, and SegNet was utilized in the segmentation phase. Additionally, the researchers employed several other steps in the study, including data acquisition, preprocessing, model building, model compiling, model fitting, testing, and prediction during the classification phase. They demonstrated the prediction results for normal and abnormal medical images. In contrast, the segmentation phase involves model building, model compiling, model fitting, testing, and prediction steps, and the segmented model predicts the abnormal regions of the brain. The classification model achieved an accuracy of 96%, and the segmentation had 85% accuracy. In a real-life evaluation of AI-based CT analysis, the system achieved sensitivity of 88.8% and specificity of 92.1%, with near-perfect accuracy when combined with radiologists [20]. A 2025 meta-

analysis confirmed that DL models consistently outperform traditional radiological approaches in ICH detection across diverse datasets [21]. Anoop Kumar [22] attempted to detect stroke using three DL models: DNN, recurrent neural networks (RNN), and K-nearest neighbors (KNN) algorithms. In this study, they suggested a way to compare these DNNs and defined the types of strokes using Artificial Neural Networks (ANN) and Support Vector Machine (SVM) classification methods. According to their observations, RNN achieved an outstanding accuracy of 98% with a minimum mean squared error (MSE) on the dataset. Liu *et al.* [23] carried out another study on deep convolutional neural networks (Deep-CNN) to automatically identify acute ischemic stroke lesions from different types of MRI data. In addition to DL, data fusion and augmentation techniques were used to enhance the dataset. Their research proposed a Res-CNN framework that was applied to both the SPES and LHC datasets. Their model combined a similar U-shape architecture with residual units that alleviated the vanishing gradient problem. Among the seven neural networks evaluated, the Res-CNN model performed better than the other networks. They used the average DICE coefficient as the significant experimental parameter, which was calculated to be 74.20%. The average DICE coefficient is defined as the proportion of common elements between two datasets to the intersection of the respective sets. To address data-sharing limitations, federated learning frameworks have been proposed to train ICH detection models collaboratively across institutions without data transfer [24].

To identify strokes from CT scan images of the human brain, Pereira *et al.* [25] proposed a DNN architecture using two different datasets: ImageNet and CIFAR-10. They considered two different types of strokes; therefore, the entire dataset was partitioned into three categories: healthy brain, ischemic stroke, and hemorrhagic stroke. Furthermore, they used a Particle Swarm Optimization (PSO) optimizer in their research to fine-tune the CNN hyperparameters to minimize the loss function over the training set.

Jiang [26] proposed a computer vision-based technology to detect ischemic stroke from CT scan images. The author presented a cascading approach called the Cascade Branch Compression Dense Network (CBCDN) to define stroke lesions. This study collected 30,570 grayscale brain CT scan images from 319 patients across five hospitals in China. From this database, the model was trained using 21,617 samples, cross-validated with 4395 samples, and tested using 4558 samples. Their model achieved a satisfactory accuracy of 87.48%. The results were compared with the training outcomes of sophisticated architectures such as DenseNet-121, DenseNet-169, ResNet-110, ResNet-164, and Wide Residual Networks (WRN). Emerging research has also explored DL models tailored for etiology identification of hemorrhages directly from NCCT scans [27].

Previous studies were positive and demonstrated the thorough identification of different types of stroke. The proposed work focuses on intracranial hemorrhage, a crucial stroke type. Recently, transformer-based architectures have been introduced for ICH detection, with entropy-aware fusion strategies showing promising

improvements over conventional CNNs [28]. A novel deep-CNN-based approach was illustrated using some existing technologies, such as a pretrained Efficient-NetB2 with custom layers with ReLU activation, and each layer was followed by maxpooling and dropout. The primary focus of this study was to correctly detect the intracranial hemorrhage stroke region in medical images. A comparative table of the reviewed studies is presented in **Table 1**.

Table 1. Comparison of technical specifications and performance metrics of reviewed research.

Study	Primary Method	Imaging Modality	Dataset Size	Best Performance	Validation Type
Cheon <i>et al.</i> [12]	DL + PCA	Clinical Data	15,099 patients	84.03% accuracy	Single cohort
Gaidhani <i>et al.</i> [19]	LeNet + SegNet	CT/MRI	Not specified	96% classification	Internal only
AI System [20]	CNN-based	CT	Clinical dataset	88.8% sensitivity	Real-world
Kumar [22]	RNN	Not specified	Not specified	98% accuracy	Not specified
Liu <i>et al.</i> [23]	Res-CNN	MRI	SPES + LHC	74.20% DICE	Cross-dataset
Jiang [26]	CBCDN	CT	30,570 images	87.48% accuracy	Multi-hospital

3. Methodology

The methodology section of the paper offers an in-depth look at the structured approach employed in this study, focusing on the three components. The first step is image preprocessing, which involves enhancing and standardizing raw medical images to prepare them for further analysis. The proposed Hem-CNN model with a technical architecture, which describes the specific CNN architecture designed for intracranial hemorrhage detection, is described in detail. Finally, the overall system design is presented, demonstrating how preprocessing, model training, and performance evaluation work together in a unified manner to achieve strong and accurate classification results.

3.1. Image Pre-Processing

Image preprocessing is essential for preparing raw medical images for analysis and classification using DL models. For the Hem-CNN model, we implemented a systematic preprocessing pipeline to ensure that the input data are of high quality and consistent, thereby enabling the model to accurately differentiate between stroke and normal images. The standardized Brain Stroke CT Image Dataset was used, which includes two categories: 950 images showing stroke-affected brain re-

gions and 1551 normal brain images. Given the heterogeneity in acquisition conditions, including variations in imaging equipment, resolution, lighting, and contrast, extensive preprocessing was applied to standardize the dataset before feeding it into the CNN. **Figure 1** illustrates the preprocessing workflow.

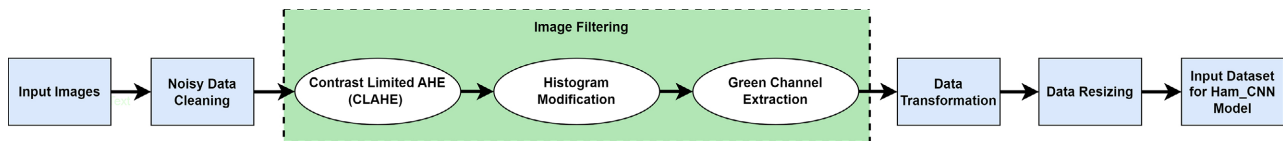


Figure 1. Image preprocessing steps.

3.1.1. Noise Removal

To reduce the impact of imaging artifacts and distortions, we employed two complementary denoising techniques. A 3×3 median filter was used to suppress salt-and-pepper noise, while a Gaussian filter with a kernel size of 5×5 and a standard deviation of $\sigma = 1.0$ was applied to reduce Gaussian noise. These methods were selected to preserve edges and fine details while improving the signal-to-noise ratio (SNR), ensuring that diagnostically relevant structures remained intact.

3.1.2. Image Filtering

Following denoising, image filtering was performed to enhance contrast and highlight subtle differences in brain tissue structures. First, Contrast-Limited Adaptive Histogram Equalization (CLAHE) was applied with a clip limit of 2.0 and a tile grid size of 8×8 to improve local contrast in small regions while preventing over-amplification of noise. Subsequently, global histogram equalization was applied with pixel intensity values normalized to the range $[0, 255]$ to ensure a consistent dynamic range across the dataset.

For further refinement, the green channel was extracted before grayscale conversion, as prior studies have shown that it retains the most diagnostically significant information. This approach reduced computational complexity while maintaining structural fidelity. **Figure 2** shows the complete filtering process.

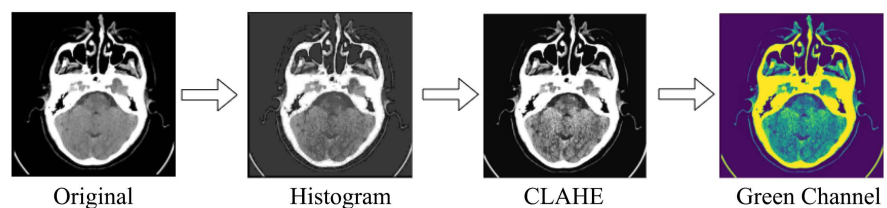


Figure 2. Image filtering step.

3.1.3. Data Transformation and Resizing

The pixel intensity values of the filtered images were normalized to the range $[0, 1]$, compatible with the ReLU activation functions in the CNN layers. This normalization minimized the impact of brightness and contrast variations caused by differences in imaging conditions.

All images were resized to 224×224 pixels using bilinear interpolation, which

balances smoothness and sharpness for medical imaging. Aspect ratio was preserved through center-cropping followed by resizing to avoid geometric distortions. This resizing ensured uniform input dimensions, enabling efficient batch training and stable memory utilization.

3.1.4. Dataset Preparation

The final preprocessed dataset consisted of cleaned, filtered, normalized, and resized images, categorized into stroke and normal classes. These standardized preprocessing steps ensured that the Hem-CNN model received consistent, high-quality data inputs, enabling robust learning and accurate classification performance.

3.1.5. Reproducibility and Transparency

To promote transparency and reproducibility, the complete preprocessing pipeline, including scripts for noise removal, filtering, transformation, and resizing, has been implemented in Python (OpenCV and NumPy) and will be made available upon request through a public GitHub repository. This will allow researchers to replicate and extend our work in related medical imaging studies.

3.2. Proposed Model

The proposed Hem-CNN model is a Deep-CNN developed to effectively extend the power of a pretrained model with a custom layer architecture to accurately classify stroke images. As shown in **Figure 3**, the architecture begins with an input image with dimensions $224 \times 224 \times 3$ pixels.

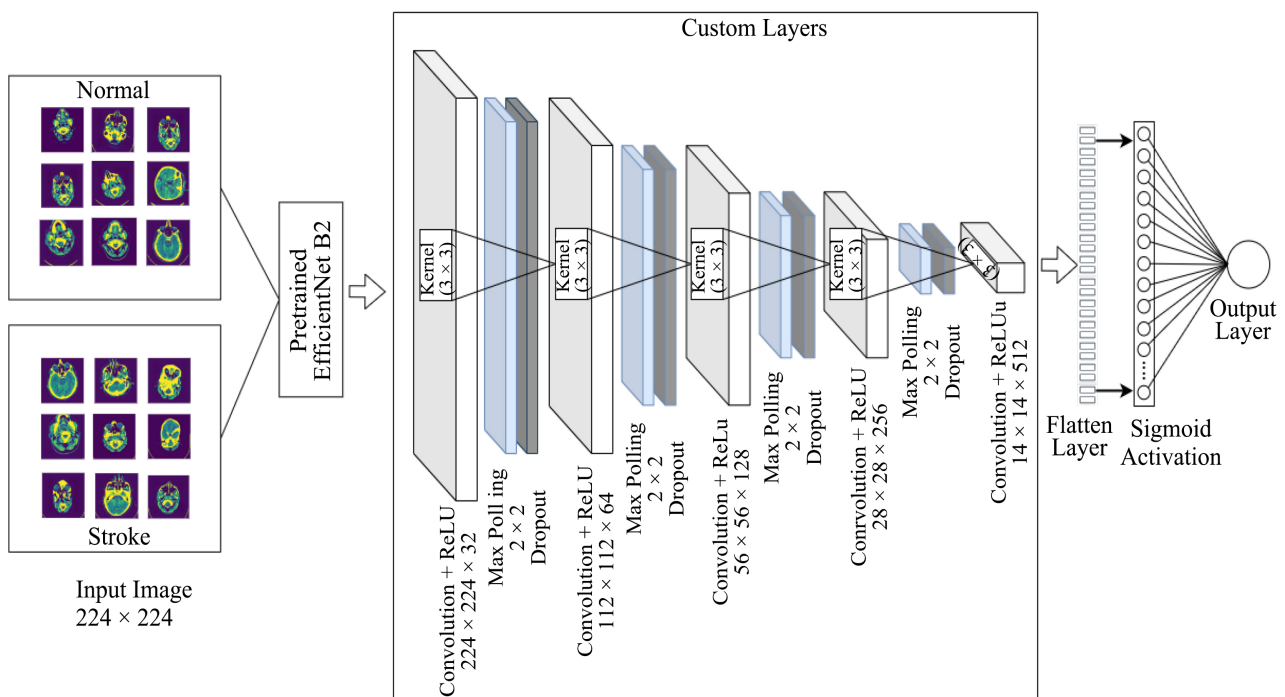


Figure 3. Proposed model architecture.

To enrich the advantages of transfer learning, the model initially passes the input through EfficientNetB2, a well-known pretrained CNN architecture. It balances the model depth, width, and resolution to achieve higher accuracy with fewer parameters. Generally, EfficientNetB2 can find important details, such as areas with different densities and blurry spots, and identify the thick *Middle Cerebral Artery (MCA)* sign after analyzing input images, which speeds up the learning process while maintaining high accuracy. Following the EfficientNetB2 feature extraction, the model employs custom layers, where 32 filters are first applied, each with a size of 3×3 , over the input image. This process creates a 224×224 feature map with 32 channels that includes the essential features of an image. A ReLU activation function was applied to introduce nonlinearity, allowing the network to learn complex data patterns.

A max pooling layer with a 2×2 window was used after the first convolution. Here, max pooling decreases the spatial dimensions of the feature maps by selecting the maximum value from each 2×2 zone, resulting in a downsampled feature map of 112×112 . Next, the model adds another convolutional layer that uses 64 filters of size 3×3 on the feature maps from the previous layer. This process improved the network's capacity to recognize more complex characteristics, producing a $112 \times 112 \times 64$ output shape. The ReLU activation is again applied to introduce a nonlinearity. The output is then passed through another max pooling layer with a 2×2 window, reducing the dimensions to 56×56 , followed by dropout to prevent overfitting. As the network deepened, a third convolutional layer was introduced with 128 filter units. This layer generates a $56 \times 56 \times 128$ output by performing feature extraction. The model can collect even more abstract elements because the fourth convolutional layer increases the depth of the feature maps to 256 pixels. The maximum pooling layer lowers the output to 14×14 after ReLU activation, and dropout is applied to avoid overfitting. Finally, a fifth convolutional layer with 512 filters was applied, producing a feature map of size $14 \times 14 \times 512$. This layer allows the network to capture highly abstract features that are crucial for accurate classification of the data. In the last stages of the network, the $14 \times 14 \times 512$ feature map was flattened into a single vector and run through a sigmoid activation function. The sigmoid function is particularly useful for binary classification problems because it returns a number between 0 and 1, indicating that the input image belongs to the stroke or normal class. In general, EfficientNetB2 with a special mix of layers, which we name Hem-CNN, can enhance detection accuracy and strength by helping the model learn detailed and organized features of hemorrhages.

To ensure reproducibility of our results, the Hem-CNN was trained under a fixed configuration which is summarized in **Table 2**. We employed the AdamW optimizer with weight decay = 1×10^{-4} and gradient clipping at $L_2 = 1.0$. The initial learning rate was set to 3×10^{-4} , decayed to 1×10^{-6} using cosine annealing with a 5-epoch warm-up, and further reduced by a factor of 0.5 on validation plateaus. The batch size was 16 with mixed precision enabled. Training was performed

for a maximum of 100 epochs, with early stopping applied if the validation loss did not improve for 10 consecutive epochs ($\Delta = 10^{-3}$). Regularisation included progressive dropout ($p = \{0.20, 0.30, 0.40\}$ across convolutional blocks and $p = 0.50$ before the classifier), L_2 penalties via AdamW, and extensive data augmentation (random rotation $\pm 10^\circ$, horizontal/vertical flip with $p = 0.5$, and random brightness/contrast adjustment $\pm 10\%$). The loss function was binary cross-entropy with logits, with class weights computed from the training distribution to address imbalance ($w_{\text{stroke}} \approx 0.47$, $w_{\text{normal}} \approx 0.53$). Images were preprocessed by resizing to 224×224 and normalized with ImageNet mean and standard deviation. Dataset splitting was stratified at the patient level (train/validation/test = 70/15/15). The best model checkpoint was chosen based on validation AUC, and all results reported were computed on the unseen test set. Initialization used ImageNet weights for EfficientNetB2 and Kaiming normal initialization for custom layers, with random seed 42 fixed across NumPy and PyTorch for reproducibility. Experiments were executed on CUDA with cuDNN, and PyTorch with torchvision (exact versions specified in the project repository).

Table 2. Training hyper-parameters of the proposed Hem-CNN model.

Parameter	Configuration
Optimizer	AdamW (weight decay 1×10^{-4}); gradient clipping $L_2 = 1.0$
Learning rate (LR)	Initial 3×10^{-4} ; cosine annealing to 1×10^{-6} ; 5-epoch warm-up
Batch size	16 (mixed precision, AMP enabled)
Epochs	Maximum 100; early stopping on validation loss (patience = 10, $\Delta = 10^{-3}$)
LR scheduling	Reduce on plateau: factor 0.5 if no improvement for 3 epochs (floor 1×10^{-6})
Regularisation	Dropout $p = \{0.20, 0.30, 0.40\}$ across conv blocks; $p = 0.50$ before classifier; L_2 via AdamW; data augmentation (rotation $\pm 10^\circ$, horizontal/vertical flip $p = 0.5$, brightness/contrast $\pm 10\%$)
Loss function	Binary cross-entropy with logits (BCEWithLogits); class weights $w_c = \frac{N}{2n_c}$ (stroke ≈ 0.47 , normal ≈ 0.53)

3.3. Explainable AI

Gradient-weighted Class Activation Mapping (Grad-CAM) highlights the important image regions by utilizing the gradients of the target class score with respect to the feature maps of a convolutional layer. Given the class score y^c and feature maps A^k , the importance weights are computed as

$$\alpha_k^c = \frac{1}{Z} \sum_i \sum_j \frac{\partial y^c}{\partial A_{ij}^k}, \quad (1)$$

where Z denotes the number of pixels in the feature map. The final Grad-CAM heatmap is obtained by a weighted combination followed by a ReLU function:

$$L_{\text{Grad-CAM}}^c = \text{ReLU}\left(\sum_k \alpha_k^c A^k\right). \quad (2)$$

3.4. System Architecture

The process begins with input images that serve as foundational data for the analysis. **Figure 4** shows a visual representation of the proposed system architecture. The system undergoes image processing, which is a significant advance in preparing the raw images for analysis. This stage involves several methods, including feature extraction, contrast enhancement, and noise reduction, which emphasize the pertinent aspects of the images that are necessary for precise detection.

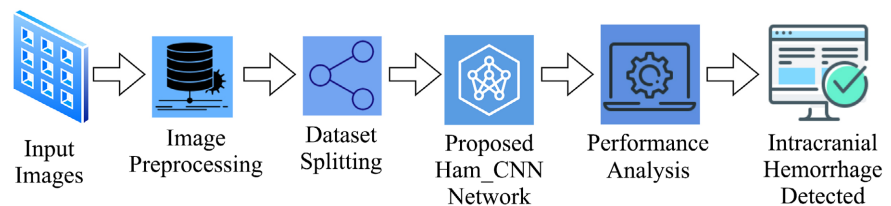


Figure 4. System Architecture of the proposed model.

The processed images were divided into training, validation, and test sets during the dataset splitting stage. This was performed to ensure that the model was trained on a representative subset of the data and that its performance was assessed impartially based on data that had not yet been seen. After the dataset was prepared, the images were fed into the proposed Hem-CNN network. This deep-CNN is particularly intended to evaluate images and learn complicated patterns that discriminate between normal and stroke images. The network is built to gradually pick up features, starting with simple edges and moving to more complex patterns, which ultimately helps in accurately identifying intracranial hemorrhages. After training the model, a performance analysis was performed to assess its efficacy. The results of this analysis offer observations about the model's strengths and potential areas for improvement. Finally, the system culminates in the intracranial hemorrhage detection phase, in which the trained model is used to classify new images. The developed system effectively leverages the learned features to accurately assess the presence of hemorrhages, providing essential insights that contribute to informed medical diagnosis and treatment planning. The entire process, from input images to final detection, creates a holistic system that addresses a key medical challenge by utilizing modern image processing and DL approaches.

The proposed system architecture offers a robust and efficient approach for detecting intracranial hemorrhages using advanced Deep Learning (DL) techniques. The system effectively manages CT imaging data and accurately detects hemorrhages using advanced image processing methods, a specially designed Hem-CNN

network, and rigorous performance testing. This comprehensive approach enhances diagnostic accuracy and contributes to timely and effective treatment of patients. The success of this system demonstrates the promise of combining state-of-the-art image processing and DL methodologies in medical image analysis, which will lead to future advancements in automated diagnosis and healthcare technologies.

4. Experimental Results and Discussion

4.1. Experimental Setup

All experiments for developing, training, and evaluating the Hem-CNN model were conducted on the Google Colaboratory (Colab) platform, leveraging its cloud-based resources and GPU-acceleration. Specifically, we utilized an NVIDIA Tesla T4 GPU with 8.1 TFLOPS of single-precision performance and 16 GB of GDDR6 VRAM, paired with an Intel® Xeon® CPU @ 2.20 GHz (2 cores) and 25 GB of system RAM using TensorFlow 2.10.0 and Keras 2.10.0, respectively.

4.2. Dataset

Kaggle is an online platform that brings together data scientists and ML practitioners. It allows users to find and share datasets, create and test ML models in an interactive environment, collaborate with peers, and participate in competitions that focus on solving data science problems. The Brain Stroke CT Image Dataset [29] provides a collection of annotated CT scan slices, highlighting the presence of hemorrhage and categorizing them into two classes: Normal and Stroke. A summary of the dataset properties is presented in Table 3.

Table 3. Dataset information.

Attribute	Value
Dataset Name	Brain Stroke CT Image Dataset [29]
Data Type	JPG format
Total Patients	82
Total CT Scan Slices	2501
Normal Patients	51
Normal CT Scan Slices	1551 (62%)
Stroke Patients	31
Stroke CT Scan Slices	950 (38%)
Stroke CT Scan Slices/Patient	≈30

The dataset consists of 2501 CT scan slices collected from 82 patients, with 51 normal (62%) and 31 stroke (38%). On average, each patient contributed approximately 30 slices, stored in JPG format. To avoid slice-level data leakage and ensure a fair evaluation, dataset splitting was performed strictly at the *patient level*.

Specifically, patients were randomly divided into training (60%, 49 patients), validation (20%, 16 patients), and testing (20%, 17 patients) groups, with all slices from a given patient confined to a single group. This guarantees that the model does not encounter slices from the same patient in both training and evaluation phases [30] [31].

Figure 5 illustrates a pie chart showing the slice-level frequency distribution of the dataset. Despite its usefulness, the dataset is relatively small for deep learning applications, which may limit the generalizability of models trained on it. This limitation is acknowledged and considered in the experimental design.

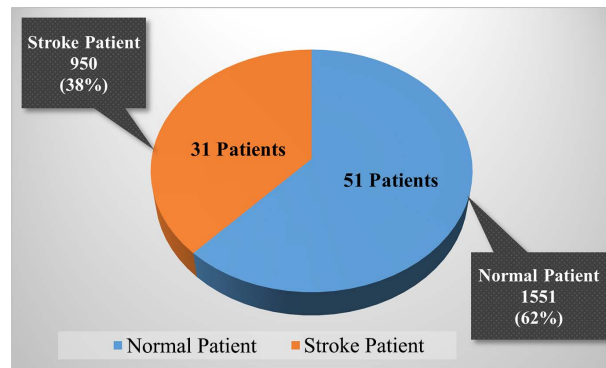


Figure 5. Frequency distribution of the dataset.

4.3. Result and Discussion

The experimental results provide a detailed comparison of several state-of-the-art DL models for intracranial hemorrhage (ICH) detection. **Table 4** presents the performance comparison. This study showed important differences in performance among the six models: VGG19, XceptionNet, ResNet152, DenseNet169, EfficientNetB2, and the proposed Hem-CNN model. The performance of each model was evaluated using accuracy, precision, recall, specificity, F1 score, and the area under the ROC curve (AUC), which provides a threshold-independent measure of classification ability and facilitates comparison with prior ICH studies.

Table 4. Performance comparison of proposed hem-CNN model with state-of-the-art models.

Model Name	Accuracy	Precision	Recall	Specificity	F1 Score	AUC
VGG19	84.09%	78.46%	80.42%	86.42%	79.21%	0.87
XceptionNet	88.40%	85.11%	84.21%	91.00%	84.50%	0.91
ResNet152	89.60%	87.10%	85.26%	92.26%	86.20%	0.92
DenseNet169	91.80%	91.16%	86.84%	94.84%	88.90%	0.94
EfficientNetB2	94.40%	94.51%	90.53%	96.77%	92.50%	0.96
Hem-CNN (Proposed)¹	97.80%	96.86%	97.37%	98.06%	97.10%	0.99

¹Proposed in this work.

During preprocessing, we examined potential class imbalance between stroke (38%) and normal (62%) cases. To address this, data augmentation techniques (horizontal/vertical flips, rotations, and intensity normalization) were applied to increase the variability of minority class samples, and a class-weighted cross-entropy loss function was used to reduce bias toward the majority class. These strategies ensured that the model maintained balanced sensitivity (recall) and specificity across both classes, avoiding skewed performance.

The confusion matrix for Hem-CNN in **Figure 6** further illustrates its robustness. Of the 310 test images (20% of the dataset), Hem-CNN correctly classified 151/155 hemorrhage cases and 150/155 normal cases. Only a few misclassifications occurred (four false negatives and five false positives).

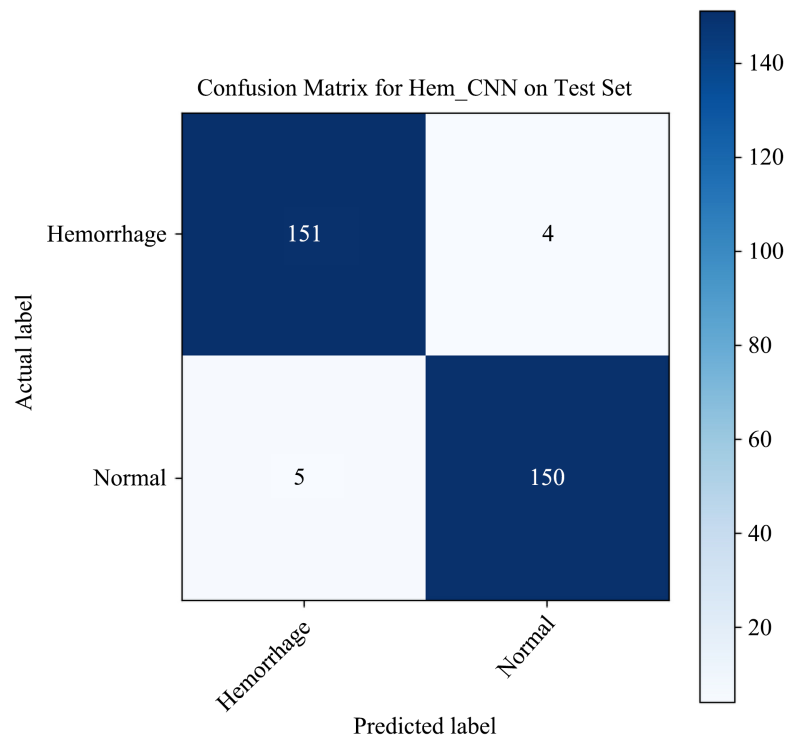


Figure 6. Confusion matrix of the proposed Hem-CNN model.

The proposed Hem-CNN significantly outperformed the other models across all metrics. It achieved an AUC of 0.99 (depicted in **Figure 7**), highlighting its near-perfect ability to distinguish stroke from normal images across varying thresholds, a result that aligns favorably with or surpasses prior ICH detection frameworks [9] [26] [27]. **Figure 8** provides a comparative visualization of different models evaluated using accuracy, precision, recall, specificity, and F1-score.

Overall, the Hem-CNN model demonstrated superior performance, achieving 97.80% accuracy, 96.86% precision, 97.37% recall, 98.06% specificity, 97.10% F1 score, and an AUC of 0.99. These results confirm that the proposed architecture is not only highly accurate but also robust and reliable for medical diagnostic applications, with strong potential for clinical adoption.

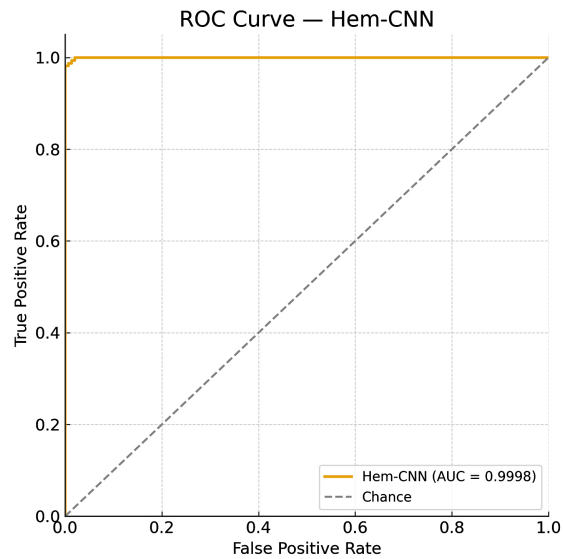


Figure 7. ROC curve of the proposed Hem-CNN.

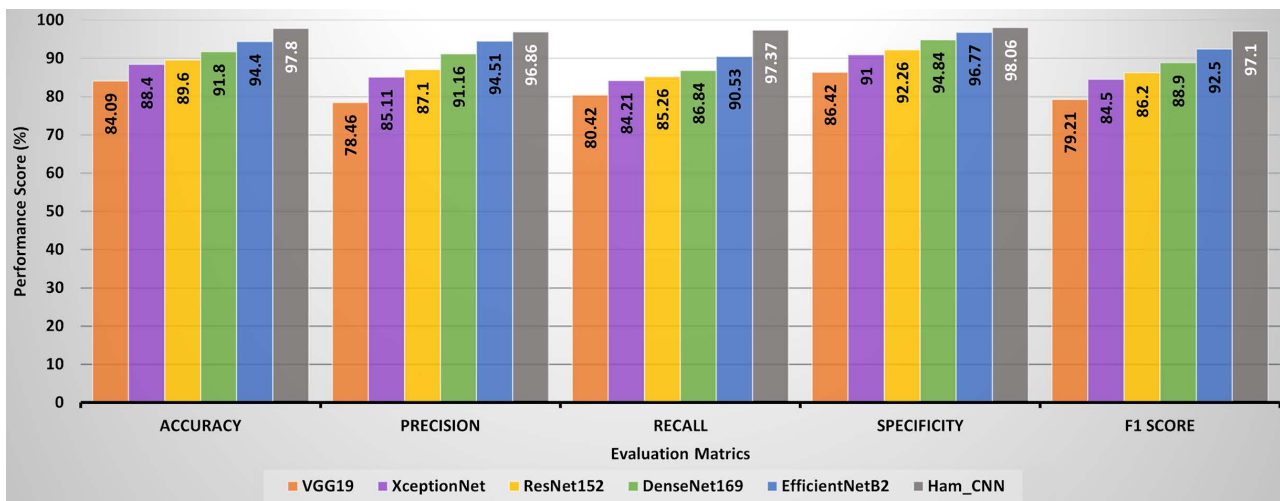


Figure 8. Comparison of deep learning models across evaluation metrics.

4.4. Explainable AI Results

The Grad-CAM visualization shown in **Figure 9** highlights the discriminative regions of brain CT scans that contributed most to the classification decision. For the healthy samples, the activation maps showed diffuse and lower-intensity regions, indicating that the model focused less on any specific abnormal area. In contrast, the unhealthy scans demonstrated high-intensity activations in localized regions, particularly around suspected tumor sites, which strongly influenced model predictions. The overlay images confirmed that the highlighted regions corresponded to clinically relevant abnormal structures rather than irrelevant background features. This method is crucial for brain-tumor classification, as it not only improves model transparency and trustworthiness but also assists clinicians in validating that the model’s focus aligns with pathologically significant regions, thereby enhancing decision support in medical imaging.

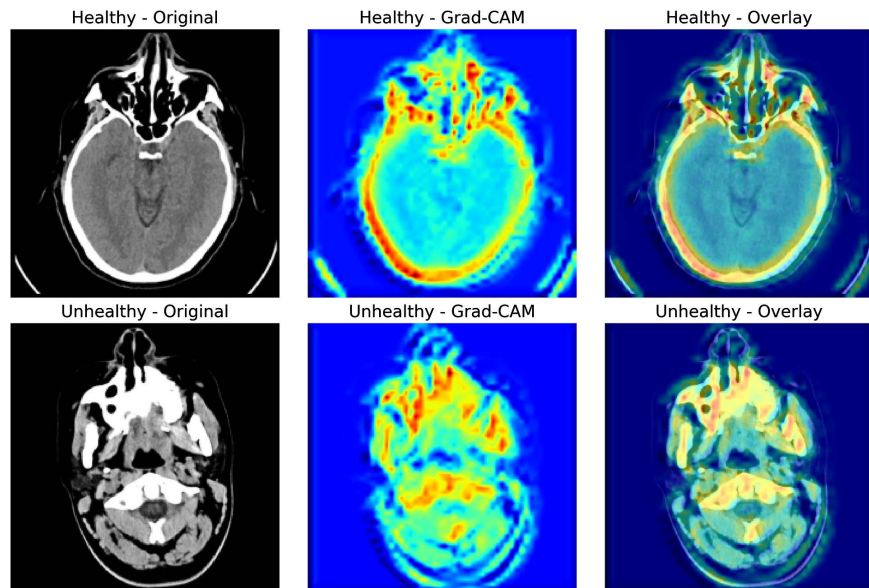


Figure 9. Grad-CAM showing model focus on abnormal regions in brain tumor classification.

The experimental evaluation clearly demonstrated that the proposed Hem-CNN model significantly outperformed established deep learning architectures, including VGG19, ResNet152, DenseNet169, XceptionNet, and EfficientNetB2, across all performance metrics. With an accuracy of 97.80%, precision of 96.86%, recall of 97.37%, specificity of 98.06%, and an F1-score of 97.10%, Hem-CNN achieved the highest reliability and robustness for classifying stroke and normal CT images. Confusion matrix analysis confirmed a low rate of false positives and false negatives, reinforcing its clinical applicability. Furthermore, Grad-CAM visualizations validated the decision-making process of the model by localizing pathologically relevant regions, thereby enhancing interpretability and trustworthiness. Overall, these results highlight Hem-CNN as a promising and explainable diagnostic framework for the accurate and efficient detection of intracranial hemorrhage, with strong potential for real-world integration into medical imaging systems.

5. Conclusion

Early detection of stroke and its specification are considerably challenging in the current era. Stroke can easily worsen the condition of patients with cardiac, kidney, and brain-related problems. We proposed a Hem-CNN model that can efficiently detect stroke. Stroke detection from CT images of the human brain achieved promising results using our model with a classification accuracy of nearly 97.80%. We improved the Hem-CNN model by adjusting its depth to reduce the calculation time. This makes it an ideal model for any stroke-related CT dataset. We aimed to reduce the time required to predict stroke by utilizing trained models. This research has the potential to impact the medical sector by further enhancing stroke detection based on CT scan image analysis. Due to lack of data, it was dif-

difficult to address the categorical classification of brain CT scan images, where we intended to build a model that would identify all three types of strokes: ischemic stroke, hemorrhagic stroke, and transient ischemic attack. However, in this study, we specifically focused on hemorrhagic stroke. In future work, we plan to create an application using the Hem-CNN model that can classify strokes and identify stroke-affected areas in the respective regions of the brain. Our future work will explore transformer-based hybrid models and clinical workflow integration. This will help doctors easily identify problems in a short time. Conversely, patients will be able to independently understand the predicted results and analysis reports related to their healthcare issues.

Acknowledgements

This study was performed at the American International University-Bangladesh (AIUB). The authors thank the AIUB authority for their support.

Data Availability

The data used in this study was collected from <https://www.kaggle.com/datasets/afridirahman/brain-stroke-ct-image-dataset/data>.

Conflicts of Interest

The authors express no conflict of interest and all the authors consented to the publication of this study.

References

- [1] MSD Manual Consumer Version (2025) Overview of Stroke—Brain, Spinal Cord, and Nerve Disorders. <https://www.msmanuals.com/home/brain-spinal-cord-and-nerve-disorders/stroke/overview-of-stroke>
- [2] Arias, E., *et al.* (2014) Mortality in the United States, 2014. NCHS Data Brief No. 229, 1-8. <https://www.researchgate.net/profile/Elizabeth-Arias-5/publication/289355835>
- [3] The American Heart Association Statistics Committee and Stroke Statistics Subcommittee (2015) Heart Disease and Stroke Statistics—2016 Update: A Report from the American Heart Association, *Circulation*, **133**, e38-e48.
- [4] Centers for Disease Control and Prevention (2025) Stroke. <https://www.cdc.gov/stroke/index.html>
- [5] Maya, B.S. and Asha, T. (2020) Predictive Model for Brain Stroke in CT Using Deep Neural Network. *International Journal of Recent Technology and Engineering*, **9**, 2011-2017. <https://doi.org/10.35940/ijrte.f9954.059120>
- [6] Stroke.org (2025) Ischemic Stroke (Clots). <https://www.stroke.org/en/about-stroke/types-of-stroke/ischemic-stroke-clots>
- [7] Medical News Today (2025) Hemorrhagic Stroke: Causes, Symptoms, Treatments, Prevention. <https://www.medicalnewstoday.com/articles/317111>
- [8] Panuganti, K.K., Tadi, P. and Lui, F. (2023) Transient Ischemic Attack. StatPearls

Publishing. <https://www.ncbi.nlm.nih.gov/sites/books/NBK459143/>

- [9] Wang, X., Shen, T., Yang, S., Lan, J., Xu, Y., Wang, M., *et al.* (2021) A Deep Learning Algorithm for Automatic Detection and Classification of Acute Intracranial Hemorrhages in Head CT Scans. *NeuroImage: Clinical*, **32**, Article ID: 102785. <https://doi.org/10.1016/j.nicl.2021.102785>
- [10] Kotovich, D., Twig, G., Itsekson-Hayosh, Z., Klug, M., Simon, A.B., Yaniv, G., *et al.* (2023) The Impact on Clinical Outcomes after 1 Year of Implementation of an Artificial Intelligence Solution for the Detection of Intracranial Hemorrhage. *International Journal of Emergency Medicine*, **16**, Article No. 50. <https://doi.org/10.1186/s12245-023-00523-y>
- [11] Radu, R.A., *et al.* (2020) Clinical Characteristics and Outcomes of Patients with Intracerebral Hemorrhage—A Feasibility Study on Romanian Patients. *Journal of Medicine and Life*, **13**, Article No. 125.
- [12] Cheon, S., Kim, J. and Lim, J. (2019) The Use of Deep Learning to Predict Stroke Patient Mortality. *International Journal of Environmental Research and Public Health*, **16**, Article No. 1876. <https://doi.org/10.3390/ijerph16111876>
- [13] Dey, N., Srinivas, M. and Subramanyam, R.B.V. (2023) A Novel Contactless Middle Finger Knuckle Based Person Identification Using Ensemble Learning. *TENCON 2023-2023 IEEE Region 10 Conference (TENCON)*, Chiang Mai, 31 October-3 November 2023, 981-986. <https://doi.org/10.1109/tencon58879.2023.10322449>
- [14] Rahman, M., Hasan, M., Billah, M.M. and Sajuti, R.J. (2022) Political Fake News Detection from Different News Source on Social Media Using Machine Learning Techniques. *AIUB Journal of Science and Engineering*, **21**, 110-117. <https://doi.org/10.53799/ajse.v21i1.383>
- [15] Rahman, M., Hasan, M., Billah, M.M. and Sajuti, R.J. (2022) Grading System Prediction of Educational Performance Analysis Using Data Mining Approach. *Malaysian Journal of Science and Advanced Technology*, **2**, 204-211. <https://doi.org/10.56532/mjsat.v2i4.96>
- [16] Chowdhury, A.A., *et al.* (2021) Sentiment Analysis of COVID-19 Vaccination from Survey Responses in Bangladesh, *Cognitive Computation*.
- [17] Goyal, R. (2022) Intracerebral Hemorrhage Detection in Computed Tomography Scans through Cost-Sensitive Machine Learning. *Applied Artificial Intelligence*, **36**, 1-19. <https://doi.org/10.1080/08839514.2022.2138126>
- [18] Yun, T.J., Choi, J.W., Han, M., Jung, W.S., Choi, S.H., Yoo, R., *et al.* (2023) Deep Learning Based Automatic Detection Algorithm for Acute Intracranial Haemorrhage: A Pivotal Randomized Clinical Trial. *NPJ Digital Medicine*, **6**, Article No. 61. <https://doi.org/10.1038/s41746-023-00798-8>
- [19] Gaidhani, B.R., Rajamenakshi, R. and Sonavane, S. (2019) Brain Stroke Detection Using Convolutional Neural Network and Deep Learning Models. 2019 *2nd International Conference on Intelligent Communication and Computational Techniques (ICCT)*, Jaipur, 28-29 September 2019, 242-249. <https://doi.org/10.1109/icct46177.2019.8969052>
- [20] Mabit, L., Lepoittevin, M., Valls, M., Thomas, C., Guillevin, R. and Herpe, G. (2025) Real-Life Performance of a Commercially Available AI Tool for Post-Traumatic Intracranial Hemorrhage Detection on CT Scans: A Supportive Tool. *Journal of Clinical Medicine*, **14**, Article No. 4403. <https://doi.org/10.3390/jcm14134403>
- [21] Karamian, A. and Seifi, A. (2025) Diagnostic Accuracy of Deep Learning for Intracranial Hemorrhage Detection in Non-Contrast Brain CT Scans: A Systematic Review and Meta-Analysis. *Journal of Clinical Medicine*, **14**, Article No. 2377.

- <https://doi.org/10.3390/jcm14072377>
- [22] Kumar, A. (2020) A Comparative Analysis to Detect Stroke Using Deep Neural Network, Recurrent Neural Network and KNN. *Psychology and Education*, **57**, 5369-5376.
- [23] Liu, L., Chen, S., Zhang, F., Wu, F., Pan, Y. and Wang, J. (2019) Deep Convolutional Neural Network for Automatically Segmenting Acute Ischemic Stroke Lesion in Multi-Modality MRI. *Neural Computing and Applications*, **32**, 6545-6558. <https://doi.org/10.1007/s00521-019-04096-x>
- [24] Srivastava, U.C., *et al.* (2020) Intracranial Hemorrhage Detection Using Neural Network Based Methods with Federated Learning.
- [25] Pereira, D.R., Filho, P.P.R., de Rosa, G.H., Papa, J.P. and de Albuquerque, V.H.C. (2018) Stroke Lesion Detection Using Convolutional Neural Networks. 2018 *International Joint Conference on Neural Networks (IJCNN)*, Rio de Janeiro, 8-13 July 2018, 1-6. <https://doi.org/10.1109/ijcnn.2018.8489199>
- [26] Jiang, F. (2019) Classification of Ischemic Stroke Lesions Based on Cascaded Branch Compression Neural Network. *IOP Conference Series: Materials Science and Engineering*, **563**, Article ID: 042003. <https://doi.org/10.1088/1757-899x/563/4/042003>
- [27] Zhao, M., *et al.* (2023) Deep-Learning Tool for Early Identifying Non-Traumatic Intracranial Hemorrhage Etiology Based on CT Scan.
- [28] Chagahi, M.H., *et al.* (2025) Vision Transformer for Intracranial Hemorrhage Classification in CT Scans Using an Entropy-Aware Fuzzy Integral Strategy for Adaptive Scan-Level Decision Fusion.
- [29] Rahman, A. (2021) Brain Stroke CT Image Dataset. <https://www.kaggle.com/datasets/afdirahman/brain-stroke-ct-image-dataset/data>
- [30] Tanvir, K., Lisha, H.V., Damodharan, R. and Sivashanmugam, K. (2024) Enhancing Corn Leaf Disease Identification with DenseViT Model. 2024 *3rd International Conference on Artificial Intelligence for Internet of Things (AIIoT)*, Vellore, 3-4 May 2024, 1-6. <https://doi.org/10.1109/aiiot58432.2024.10574791>
- [31] Nazera, F., Nadim, A.N.A., Dey, S.K., Tanvir, K. and Kabir, M.S. (2024) Elevating Mango Leaf Disease Classification Utilizing Dense ViT. 2024 *International Conference on Advances in Computing, Communication and Applied Informatics (ACCAI)*, Chennai, 9-10 May 2024, 1-6. <https://doi.org/10.1109/accai61061.2024.10602314>