Journal of Neonatal Surgery
ISSN(Online): 2226-0439
Vol. 14, Issue 15s (2025)
https://www.jneonatalsurg.com

AI-Driven Stroke Classification and Detection: A Retrospective Study Using MRI Imaging

Nilofer Neshat[1], Shubham Gupta[1*], Nency Unadkat[1], Khush Jain[1], Anil Rathwa[2]

[1]Parul Institute of Paramedical and Health Sciences, Faculty of Medicine, Parul University, Vadodara-391760, Gujarat, India.

[2]Parul Institute Medical Sciences and Research, Faculty of Medicine, Parul University, Vadodara-391760, Gujarat, India.

***Corresponding Author:**

Shubham Gupta

Assistant Professor, Parul Institute of Paramedical and Health Sciences, Faculty of Medicine, Parul University, Vadodara-391760, Gujarat, India.

Email ID: Shubham.gupta5555@gmail.com, Shubham.gupta19175@paruluniversity.ac.in

Cite this paper as: Nilofer Neshat, Shubham Gupta, Nency Unadkat, Khush Jain, Anil Rathwa, (2025) AI-Driven Stroke Classification and Detection: A Retrospective Study Using MRI Imaging. *Journal of Neonatal Surgery*, 14 (15s), 2294-2300.

ABSTRACT

Purpose: This study aims to develop an artificial intelligence (AI)-based method for accurately distinguishing between acute and chronic stroke using MRI imaging, to enhance diagnostic precision, treatment planning, and prognosis evaluation.

Method: A total of 700 patient MRI datasets were utilized, with an 80/20 split for training and testing. The MRI images underwent preprocessing and key feature extraction, followed by classification using a Support Vector Machine (SVM). Model performance was evaluated using standard metrics: precision, recall, F1-score, and support.

Results: The training set, comprising approximately 2,480 data points, demonstrated strong model performance. For non-stroke cases, the model achieved a precision of 0.95 and a recall of 0.92, indicating a low false positive rate. For stroke cases, the precision was 0.95 and the recall was 0.88, reflecting a slightly higher false positive rate. High F1-scores in both categories confirmed a well-balanced performance between precision and recall.

Conclusion: The proposed AI-based classification model effectively distinguishes between stroke and non-stroke MRI images, showing promise in aiding clinical diagnosis and improving patient outcomes. Future research will focus on addressing dataset imbalances and evaluating the performance of alternative machine learning algorithms to further enhance model robustness.

Keywords: *Magnetic Resonance Imaging, Diffusion Weighted Imaging, Apparent Diffusion Coefficient, Acute Stroke, Chronic Stroke, Artificial Intelligence, Machine Learning..*

1. INTRODUCTION

Stroke is a severe neurological condition characterized by sudden loss of brain function due to disrupted blood flow, often resulting from blocked or ruptured arteries.(1) Stroke classification is divided into acute and chronic phases, with types including ischemic, hemorrhagic, and transient ischemic attacks (TIA). Early and accurate diagnosis is crucial for improving patient outcomes and reducing long-term complications.(2,3)

Delays in diagnosis or treatment can lead to severe brain damage, long-term disabilities, or even death. Understanding demographic and risk factors related to strokes is essential for precise diagnosis and effective treatment.(4) Traditional MRI methods, such as T1-weighted imaging, T2-weighted imaging, Fluid Attenuated Inversion Recovery (FLAIR), Magnetic Resonance Angiography, and Diffusion-Weighted Imaging, are used to diagnose and classify strokes. However, manual interpretation of MRI images is time-consuming and can cause delays in diagnosis and treatment. The complex nature of stroke development, diverse image acquisition methods, and variations in tissue signal intensity also present challenges.(4)

Advancements like integrating artificial intelligence (AI) are needed to enhance the efficiency and accuracy of stroke diagnosis from MRI images. Machine Learning (ML) algorithms can identify patterns indicative of ischemic and hemorrhagic strokes, while Deep Learning (DL) uses neural networks to analyze complex MRI data.(5) AI-driven techniques not only speed up diagnosis but also reduce interpretation times, helping clinicians make faster and more accurate decisions. This results in earlier intervention, improved patient outcomes, and reduced treatment delays.(6)

The reviewed studies mainly focused on classifying strokes by pathology, with limited attention to distinguishing acute and chronic strokes. This is a critical research gap, as accurate classification on MRI is crucial for treatment optimization. Acute strokes require immediate interventions, while chronic strokes focus on rehabilitation and long-term care. Acc urate classification aids in imaging interpretation, symptom identification, recurrence prevention, and prognosis.

This research is particularly important as it addresses the rising global burden of stroke and the critical need for faster and more accurate diagnostic tools. Traditional and manual diagnostic methods often fall short in terms of speed and consistency, which can delay treatment and impact patient outcomes. By incorporating artificial intelligence (AI) into MRI-based stroke classification, this study aims to improve the detection and differentiation of acute and chronic strokes. The use of AI enhances diagnostic accuracy and efficiency, offering more consistent results and supporting timely medical decisions. The study's primary focus is to develop AI-driven classification models that can reliably distinguish between different types of strokes, thereby enabling better treatment planning and prognosis evaluation. Additionally, the research explores AI-based data gathering and mining techniques to support informed diagnosis and care, while ensuring ethical standards in healthcare data usage. By enabling early detection and improving image interpretation through AI, the study seeks to transform stroke management and significantly improve patient outcomes.

2. MATERIAL & METHOD

The study design involved evaluating MRI brain data from patients referred to the Department of Radiology between 2023 and 2024, authorized by the Institutional Ethics Committee for Human Research with reference number: PUIECHR/PIMSR/00/081734/6523.

Inclusion Criteria - Individuals over 18 years old with a history of stroke or transient ischemic attack (TIA).

Exclusion Criteria - Individuals with other neurological disorders, such as epilepsy, learning impairments, autism, brain tumors, and cerebral palsy.

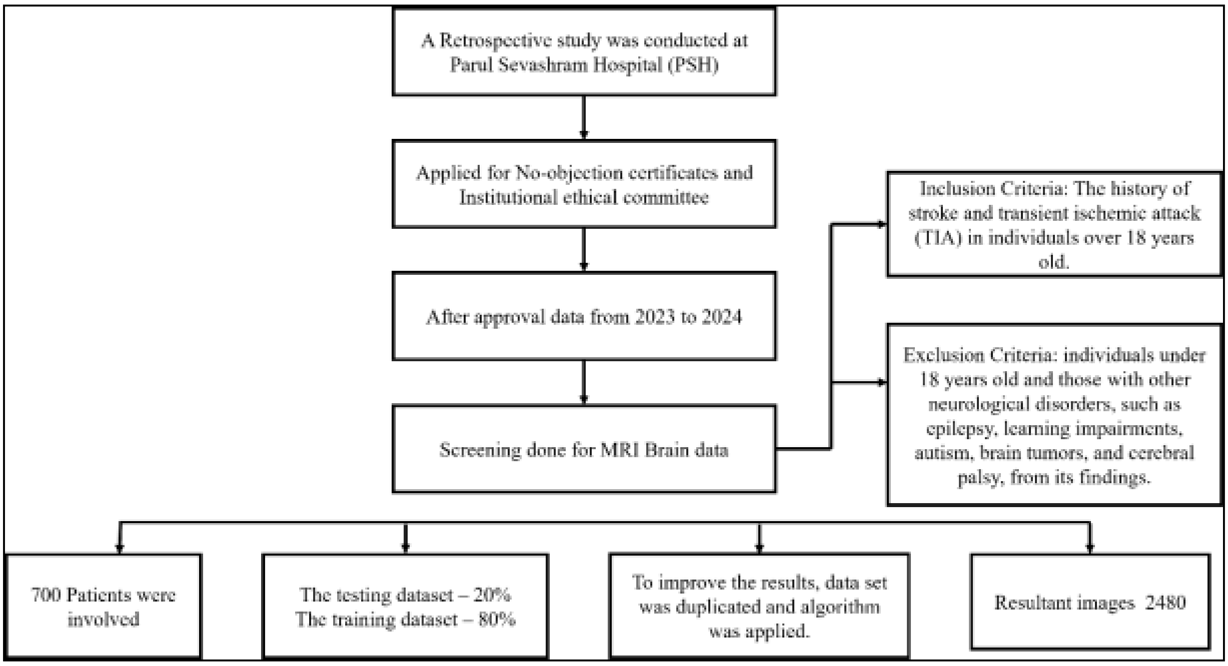

Figure 1: Workflow of the study

Description of Processes

Data Collection –

The study involved 700 patients. The dataset was divided into a training set (80%) and a testing set (20%). The training set was duplicated to improve results, resulting in approximately 2480 data points.

Data Acquisition -

- MRI brain scans were obtained using the Stroke protocol for patients recommended for MRI exams due to stroke episodes or stroke risk factors.
- Data mining software tool - Onis, was used to extract and analyze relevant data obtained from the medical imaging data.
- The mined data was processed and analyzed using data science tools, particularly machine learning with Python

and Anaconda.

- Comprehensive data analysis was conducted to obtain accurate and meaningful outcomes.

Data Preprocessing -

- **Image Normalization**: Pixel values were normalized to a common scale for consistency of the image.
- **Resize Images**: Images were resized to a standard resolution for uniformity.
- **Noise Reduction**: Filters were applied to reduce noise and enhance image clarity.

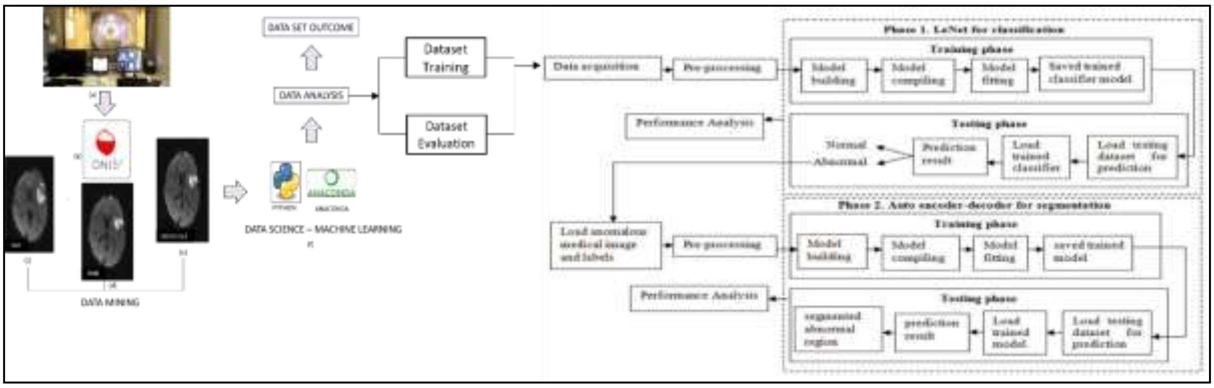

Figure 2: Methodology Workflow for the study (a) MRI Console (b) ONIS 2.5 Software (c) DWI Sequence (d) RGB Image (e) Gray scale Conversion (f) Machine Learning Tools (g) Data Analysis – Data Training and Testing using LeNet for Classification and Auto Encoder Decoder for Segmentation (h) Data set Outcome

Feature Extraction -

- **Region of Interest (ROI) Selection**: Identification and extraction of the region of interest in the brain images, focusing on areas relevant for stroke detection.
- **Edge Detection**: Edge detection algorithms, such as the Canny edge detector, were applied to highlight image boundaries.
- **Texture Analysis**: A texture analysis method, such as the Gray-Level Co-occurrence Matrix (GLCM), was used to capture textural features indicating a stroke.

Algorithm Selection -

- **OpenCV Integration**: Utilizing OpenCV's image processing functions for tasks such as contour analysis, morphology, and image filtration.
- **Machine Learning Model**: Implementing or integrating a machine learning model with external frameworks such as TensorFlow or PyTorch using OpenCV's machine learning module.

Model Training -

- **Train-Test Split**: The datasets were split into training and testing sets for model evaluation.
- **Feature Vector Creation**: For each image, the retrieved features were converted into feature vectors.
- **Model Training**: A suitable algorithm, such as Support Vector Machine (SVM) or Random Forest with cross-validation, was used to train the machine learning model using the training dataset.

Model Evaluation -

- **Testing Set Evaluation**: The trained model was assessed on the testing set to determine performance.
- **Performance Metrics**: Metrics such as recall, accuracy, precision, and F1-score were utilized for evaluation.

Integration and Deployment -

- **OpenCV Implementation**: The trained model was integrated with OpenCV for real-time image recognition.
- **Deployment**: The system was deployed in a healthcare environment, ensuring compatibility with existing infrastructure.

3. RESULTS

Performance metrics, including recall, accuracy, precision, and F1-score, were used to evaluate the machine learning model. A power calculation was not explicitly performed, but the sample size was aimed to ensure robust analysis.

a. Data Acquisition Coding

```
# Input and output folder paths
input_folder = ""
output_folder = ""
# Create the output folder if it doesn't exist
os.makedirs(output_folder, exist_ok=True)
# List all files in the input folder
image_files = os.listdir(input_folder)
# Iterate through each image file in the input folder
for image_file in image_files:
    # Check if the file is an image (you can add more file format checks)
    if image_file.endswith(('.jpg', '.jpeg', '.png', '.bmp')):
        # Read the RGB image
        rgb_image = cv2.imread(os.path.join(input_folder, image_file))
        # Convert to grayscale
        gray_image = cv2.cvtColor(rgb_image, cv2.COLOR_BGR2GRAY)
        # Save the grayscale image in the output folder
        output_path = os.path.join(output_folder, image_file)
        cv2.imwrite(output_path, gray_image)
print("Conversion complete. Grayscale images saved in the output folder.")
```

b. Algorithm

{Sample Algorithm (Pseudocode)}: # Import necessary libraries import cv2 import numpy as np from sklearn.model_selection import train_test_split from sklearn.svm import SVC from sklearn.metrics import accuracy_score

 # Step 1: Data Preprocessing # ...

 # Step 2: Feature Extraction # ...

 # Step 3: Model Training def train_model(features, labels): # Split the dataset into training and testing sets X_train, X_test, y_train, y_test = train_test_split(features, labels, test_size=0.2, random_state=42) # Choose a machine learning algorithm (SVM is used as an example) model = SVC(kernel='linear', C=1) # Train the model model.fit(X_train, y_train) # Predictions on the test set predictions = model.predict(X_test) # Evaluate model performance accuracy = accuracy_score(y_test, predictions) print(f"Model Accuracy: {accuracy}") return model

 # Step 4: Model Integration with OpenCV # ...

 # Step 5: Real-time Image Recognition # ...

 # Step 6: Output/Diagnosis # ... # Main program if __name__ == "__main__": # Execute the steps sequentially features, labels = preprocess_data() # Assuming we have a function for data preprocessing trained_model = train_model(features, labels) integrate_with_opencv(trained_model) # Assuming we have a function for OpenCV integration perform_real_time_recognition() # Assuming we have a function for real-time image recognition generate_output() # Assuming we have a function for generating the final diagnosis.

c. Classification Report

In machine learning, a classification report serves as a performance evaluation statistic, especially for classification issues. It provides a detailed analysis of how well a model is performing by summarizing key performance indicators for each dataset class. An F1 score, precision, recall, and support are the main metrics that are presented in a classification report. Comprehending these metrics is essential for evaluating the effectiveness of a model, identifying its strengths and weaknesses, and guiding further improvements.

		ACTUAL VALUE	
		Present	Absent
PREDICTED VALUE	Present	True Positive (TP)	False Positive (FP)
	Absent	False Negative (FN)	True Negative (TN)

Figure 3: Confusion Matrix calculator for the Stroke Classification

Precision - Positive predictive value, or precision, is a metric used to assess the model's accuracy for positive prediction. It is calculated as the number of true positive predictions divided by the total number of positive predictions (true positives + false positives), illustrated in Equation (1). High precision highlights that the model has a low false positive rate. This is particularly important in medical diagnoses, such as stroke detection, where false positives can lead to unnecessary treatments and anxiety.

$$Precession = \frac{True\ Positive}{(True\ Positive + False\ Positive)}$$

Equation 1: Calculation of confusion metrics by Precession

Recall - Recall, or sensitivity, is a metric that assesses the ability of the model to identify all relevant instances. It is calculated as the number of true positive predictions divided by the total number of actual positive instances (true positives + false negatives) illustrated in Equation 2. High recall highlights that the model has a low false negative rate, suggesting it is less likely to miss positive cases. In the context of stroke detection, high recall is critical to ensure that all potential cases of stroke are identified for further examination.

$$Recall = \frac{True\ Positive}{(True\ Positive\ +\ False\ Negtive)}$$

Equation 2: Calculation of confusion metrics by Recall

F1-Score - The F1-score is a statistic that provides a balance between precision and recall, calculated as the harmonic mean of the two. When there is an unequal distribution of classes or when it is crucial to reduce both false positives and false negatives, it is especially helpful. In medical applications, where accurate and thorough detection is required, a high F1 score denotes a good balance between precision and recall. Illustrated in Equation 3

$$F1 - Score = \frac{2 \times Precession\ \times Recall}{(Precession + Recall)}$$

Equation 3: Calculation of confusion metrics by F1-Score

Support - The term "support" refers to the number of actual occurrences of each dataset class. It provides context for the F1-score, recall, and precision metrics by highlighting how many instances of each class are present. This helps in understanding the reliability of the metrics, especially when dealing with imbalanced datasets.

```
In [19]:  report = classification_report(y_test, y_preds)
           print(report)

                  precision    recall  f1-score   support

               0       0.95      0.92      0.93      1407
               1       0.92      0.95      0.94      1433

        accuracy                           0.93      2840
       macro avg       0.93      0.93      0.93      2840
    weighted avg       0.93      0.93      0.93      2840
```

Figure 4: Classification Report of Stroke i.e., Precision, Recall, F1-score, and Support

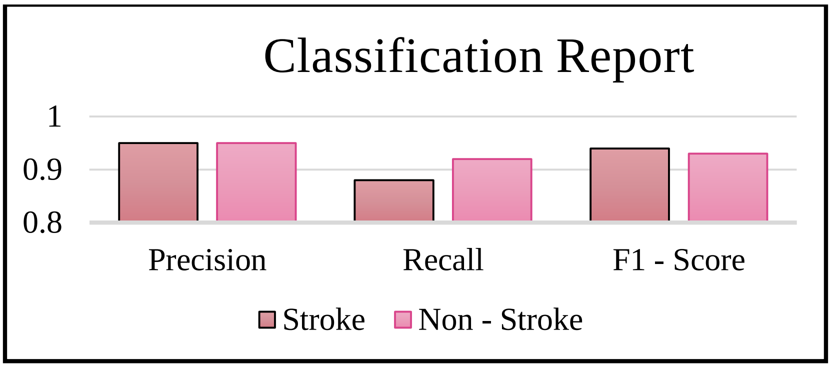

Figure 5: Bar Graph representing the Classification report

- **Precision for "No" (Non-Stroke) Class: 0.95 -** This means that 95% of the instances predicted as "No" (non-stroke) are non-stroke cases. The false positive rate is low.
- **Recall for "No" (Non-Stroke) Class: 0.92 -** This indicates that 92% of actual non-stroke cases are accurately identified by the model. The false negative rate for non-stroke cases is low.
- **F1-Score for "No" (Non-Stroke) Class: 0.93 -** The F1-score of 0.93 indicates an extremely good balance between recall and precision for non-stroke cases.
- **Precision for "Yes" (Stroke) Class: 0.95 -** This means that 95% of the instances predicted as "Yes" (stroke) are stroke cases. There is a moderate rate of false positives.
- **Recall for "Yes" (Stroke) Class: 0.88 -** This indicates that 88% of actual stroke cases are correctly identified. There is a higher rate of false negatives compared to the non-stroke class, which is a critical area for improvement to avoid missing actual stroke cases.
- **F1-Score for "Yes" (Stroke) Class: 0.94 -** The F1-score of 0.94 shows an extremely good balance between recall and precision, but there is room for improvement in capturing more true positive stroke cases without increasing false positives significantly.
- **Support for Each Class -** There are 500 non-stroke instances and 200 stroke instances in the dataset. We are duplicating images for 2840 approx. image dataset for this much accuracy. The higher number of non-stroke cases suggests an imbalanced dataset, which is common in medical diagnostics and needs to be accounted for during model training and evaluation.

The classification report thoroughly analyzes the model's performance, highlighting precision, recall, F1-score, and support for each class. These metrics assess the model's effectiveness in identifying stroke and non-stroke cases, identifying areas for improvement, and guiding further tuning and development. Balancing high precision and recall is critical in stroke detection to ensure accurate and comprehensive identification, minimizing false positives and negatives.

4. DISCUSSION

The classification and detection of stroke using artificial intelligence techniques have led to significant advancements in the field of medical imaging. The study aimed to enhance early detection of stroke using AI-based analysis of MRI scans. When compared to previous studies, our study is specifically tailored to our patient population, improving the accuracy of classification through AI-based feature extraction.

Accurate interpretation of MRI images is crucial in stroke classification and detection, particularly using Diffusion weighted images (DWI) and Apparent diffusion coefficient (ADC) maps. Previous studies, such as Zhang. et al. demonstrated that convolutional neural networks (CNNs) significantly enhance the accuracy of lesion detection in ischemic stroke cases, achieving a precision of 89.77%.(7) Our study, on the other hand, utilized support vector machine (SVM) model, which delivered a higher precision rate of 95% and recall of 92%. This indicates a balanced precision – recall ratio for stroke detection and better categorization for non – stroke cases.

Machine learning has also played in a significant role in determining stroke onset time, which is essential for timely treatment in thrombolysis. According to a study by Lee et al., Machine learning models outperformed human interpretations while determining stroke onset within 4.5 hours using multi-parametric MRI features.(8) While their study focused on logistic regression, support vector machine, and random forest classifiers, our study focused on an enhanced dataset trained using support vector machine. Unlike their methodology, which focused on time-based identification, our approach incorporates various MRI parameters, leading to improved classification accuracy beyond just onset time estimation.

By attaining an accuracy rate of over 85%, Miyamoto et al. highlighted the importance of AI for stroke classification in hospitals lacking specialized healthcare workers. Building on this, our study incorporates a larger dataset and advanced feature extraction techniques, resulting in an increased accuracy rate.(9) Our findings suggest that while both studies utilize multi-model data, our refined dataset structure and feature selection strategies contribute to greater precision.

Traditional stroke identification and detection depend on manual interpretation. Pathan et al. in their study, introduced a rough set-based technique for feature selection in stroke detection with identifying key factors such as age, glucose levels, hypertension, and heart disease.(10) While their work focused on selecting relevant features, our research focuses on incorporating automated AI-based feature extraction within an AI-based classification model, minimizing the variability and enhancing consistency in diagnosis, which makes stroke detection more reliable and efficient.

Overall, our research shows that using AI to diagnose strokes increases classification accuracy, reduces delay, and enhances efficiency. Although previous studies have demonstrated the promise of machine learning in stroke diagnosis, our improved method refines model precision and recall, even in non-stroke cases. To further evaluate effectiveness and advance the stroke detection in various healthcare settings, future research should concentrate on the use of real-time AI models in the clinical environment.

5. CONCLUSION

The described method for brain stroke detection offers a comprehensive approach that integrates various stages of feature extraction, preprocessing of data, algorithm selection, model training, evaluation, and deployment. Each stage plays a critical role in ensuring the accuracy and efficiency of the detection system, ultimately contributing to improved diagnostic capabilities in a healthcare setting. It concludes that, for enhancing stroke identification, focusing on dataset improvement by addressing class imbalances and incorporating diverse imaging conditions will ensure a more representative and robust training set. Concurrently, advancing algorithmic techniques, such as leveraging more sophisticated deep learning models and finetuning existing models, will elevate the system's precision and recall, leading to more reliable and accurate stroke detection.

6. LIMITATION

Due to a smaller number of sample data sets, the algorithm can classify stroke and non–stroke and further give an accuracy for acute and Chronic stroke without defining it in percentage. There is no certainty that these algorithms will work on other large data sets due to their training limitation

REFERENCES

[1] Murphy SJ, Werring DJ. Stroke: causes and clinical features. Medicine (Abingdon). 2020 Sep 1;48(9):561.

[2] What Is an Acute (or Sudden) Stroke? [Internet]. [cited 2025 Mar 22]. Available from: https://www.verywellhealth.com/what-is-an-acute-stroke-3146171

[3] Chronic Stroke: Where in the Recovery Process Is This Phase? [Internet]. [cited 2025 Mar 22]. Available from: https://www.flintrehab.com/chronic-stroke/

[4] WHO EMRO | Stroke, Cerebrovascular accident | Health topics [Internet]. [cited 2025 Mar 22]. Available from: https://www.emro.who.int/health-topics/stroke-cerebrovascular-accident/index.html

[5] Zhang S, Xu S, Tan L, Wang H, Meng J. Stroke Lesion Detection and Analysis in MRI Images Based on Deep Learning. J Healthc Eng. 2021 Jan 1;2021(1):5524769.

[6] Singla R. Artificial intelligence in disease diagnosis: a systematic literature review, synthesizing framework and future research agenda. J Ambient Intell Humaniz Comput.

[7] Zhang S, Xu S, Tan L, Wang H, Meng J. Stroke Lesion Detection and Analysis in MRI Images Based on Deep Learning. J Healthc Eng. 2021;2021.

[8] Lee H, Lee EJ, Ham S, Lee H Bin, Lee JS, Kwon SU, et al. Machine Learning Approach to Identify Stroke Within 4.5 Hours. Stroke. 2020 Mar 1;51(3):860–6.

[9] Miyamoto N, Ueno Y, Yamashiro K, Hira K, Kijima C, Kitora N, et al. Stroke classification and treatment support system artificial intelligence for usefulness of stroke diagnosis. Front Neurol. 2023;14.

[10] Pathan MS, Jianbiao Z, John D, Nag A, Dev S. Identifying Stroke Indicators Using Rough Sets. IEEE Access. 2020;8:210318–27.

..