IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.471 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 14, Issue 5, May 2025

DOI: 10.17148/IJARCCE.2025.14572

BRAIN STROKE DETECTION, DIAGNOSIS POST-STROKE REHABILITATION MANAGEMENT

Arunkumar B[1], Gurubalaji R[2], Praveen S P[3], Titas Nesan A[4], Karmegam S[5]

B.E, CSE, Sri Ramakrishna college of Engineering, Perambalur, India[1]

B.E, CSE, Sri Ramakrishna college of Engineering, Perambalur, India[2]

B.E, CSE, Sri Ramakrishna college of Engineering, Perambalur, India[3]

B.E, CSE, Sri Ramakrishna college of Engineering, Perambalur, India[4]

Head of the Department, CSE, Sri Ramakrishna college of Engineering, Perambalur, India[5]

Abstract: Brain stroke is a complicated disease that is one of the foremost reasons of long-term debility and mortality. Because of breakthroughs in Deep Learning (DL) and Artificial Intelligence (AI) which enable the automated detection and diagnosis of brain stroke as well as intelligently assisting post-brain stroke patients for rehabilitation, is more favorable than a manual diagnosis. Many publications on automated brain stroke detection, diagnosis, and robotic management using DL and AI approaches are now being published. This review provides a study of the detection, diagnosis of brain stroke and robotic management techniques of post-brain stroke rehabilitation from six different perspectives, namely, brain stroke datasets and modalities of brain stroke data collection, pre-processing approaches, DL-based detection and diagnosis of brain stroke, Al-based intelligent post brain stroke rehabilitation assistant, and performance measures. It also examines the conclusions and the consequences of the findings. There are also three ongoing research challenges in the fields of brain stroke detection and diagnosis, as well as post-brain stroke robotic treatment. For this investigation, 130 key papers from the Scopus, PubMed and Web of Science archives were chosen after a comprehensive screening method. This study gives a comprehensive overview of brain stroke detection and post-brain stroke robotic management strategies that may be useful to the scientist's community working in the field of automatic brain stroke detection and robotic rehabilitation management.

Keywords: Brain stroke detection, internet of medical things, logistic regression, random forest, decision trees, support vector machine, shapley additive explanations

I. INTRODUCTION

Stroke is a serious and life-threating medical condition that ranks among the leading causes of death and long-term disability worldwide. The timely and aaccuate detection of stroke risk palys a critical role in enabling early intervention and improving patient outcomes. Traditional diagnosis methods, which involve manual interpretation of medical images and patient data, are often time-consuming and prone to humar error. With the rapid advancement of Artificial Intelligent (AI) and Deep Learning (DL), there is growing interest in developing automated, intelligent systems for stroke prediction and post-stroke management. These technologies provide scalable and highly accurate solutions that can enhance diagnosis, assist in rehabilitation, and reduce dependence on manual healthcare services. The project presents a machine learning-based approach for effective stroke prediction.

II. LITERATURE SURVEY

Deeplearning-based classification of DSA images sequences of patients with acute ischemic stroke
Authors: Benjamin j. Mittmann, Michael Braun Frank Runch, Bernd Schmitz, Thuy N. Tran, Amine yamlahi
Year: 2022.

III. METHODOLOGY

1. Imaging & Clinical Assessment

- **CT & MRI Scans**: These are primary tools for detecting ischemic and hemorrhagic strokes. AI algorithms enhance these images to identify subtle changes, such as ischemic penumbra, enabling timely interventions.

DOI: 10.17148/IJARCCE.2025.14572

- **National Institutes of Health Stroke Scale (NIHSS)**: A standardized tool assessing stroke severity based on neurological function, guiding treatment decisions and predicting outcomes

2. AI-Powered Diagnostic Tools

- **StrokeSave App**: A mobile application that uses deep learning to analyze facial images, voice recordings, and retinal scans for stroke detection, achieving a diagnostic accuracy of 95% .
- **Deep Learning Models**: Utilize electromyography signals and electronic health records to predict stroke risk and outcomes.

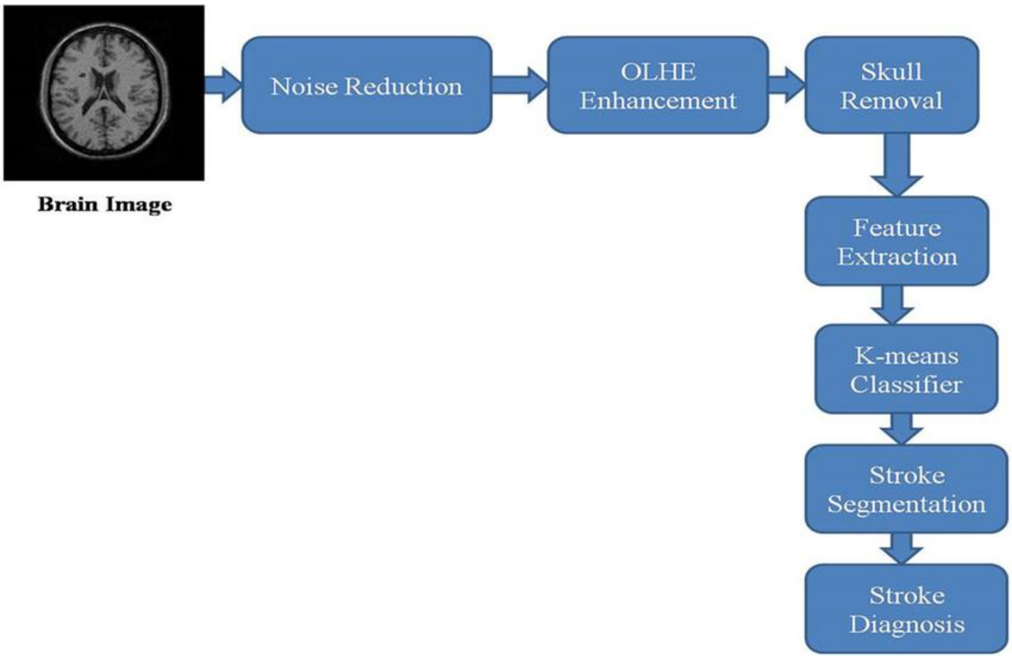


FIGURE 1. Flowchart of an automated stroke detection and diagnosis system using brain imaging.

Figure 1: The image presents a flowchart outlining the step-by-step process of **automated stroke diagnosis** using brain imaging. It starts with a **brain image**, typically from MRI or CT, which undergoes **noise reduction** to eliminate irrelevant artifacts and improve clarity. The next step, **OLHE enhancement** (Optimized Local Histogram Equalization), improves image contrast, followed by **skull removal** to isolate the brain tissue. **Feature extraction** is then performed to identify relevant visual patterns. These features are clustered using a **K-means classifier** to differentiate stroke-affected regions. The process concludes with **stroke segmentation** and a final **stroke diagnosis**, enabling accurate identification and assessment of stroke regions.
4o

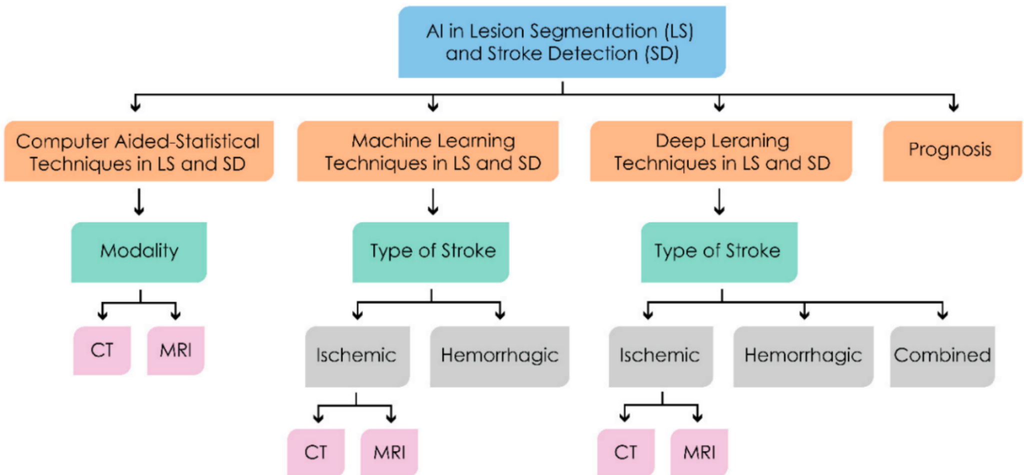


FIGURE 2. Flowchart illustrating AI methodologies in Lesion Segmentation (LS) and Stroke Detection (SD).

IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.471 ⪤ Peer-reviewed & Refereed journal ⪤ Vol. 14, Issue 5, May 2025

DOI: 10.17148/IJARCCE.2025.14572

Data Set Acquisition The image illustrates a flowchart of **AI-driven methods** used in **lesion segmentation (LS)** and **stroke detection (SD)**. It categorizes the techniques into four main types: **computer-aided statistical**, **machine learning**, **deep learning**, and **prognosis**. Each technique is linked to identifying stroke types—**ischemic**, **hemorrhagic**, or **combined**—using imaging modalities like **CT** and **MRI**. The computer-aided approach focuses on imaging modality, while machine and deep learning emphasize stroke classification. Deep learning extends further by handling combined stroke cases. This visual effectively outlines how AI supports clinical decision-making in stroke diagnosis and analysis.

The diagram outlines a stroke diagnosis pipeline using a dataset of 10,000 brain stroke cases. It includes preprocessing steps such as morphological erosion, CLAHE, Gaussian blur, and image resizing

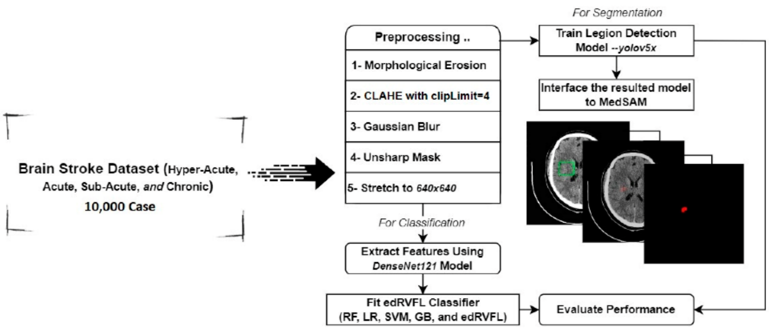


FIGURE 3: A comprehensive workflow for brain stroke detection and classification is presented.

1. Data Acquisition
- Input: Brain scan images (CT/MRI), patient health records
- Dataset: Annotated images including stroke types (ischemic/hemorrhagic), and recovery metrics

2. Preprocessing
- **Noise Reduction** (e.g., Gaussian Blur)
- **Contrast Enhancement** (e.g., CLAHE)
- **Normalization** and **Resizing** (e.g., 640x640)
- **Skull Stripping** to isolate brain tissues

3. Stroke Detection
- **Model**: YOLOv5x or U-Net integrated with MedSAM for segmentation
- Output: Lesion location and size
- Post-processing: Region extraction using bounding boxes or masks

4. Stroke Diagnosis
- **Feature Extraction**: Using pre-trained CNN (e.g., DenseNet121)
- **Classification**:
  - Models: edRVFL, SVM, Random Forest
  - Output: Stroke type (Ischemic / Hemorrhagic), severity level

5. Post-Stroke Rehabilitation Monitoring
- **Rehabilitation Module**:
  - Inputs: Motor function scores, therapy data, patient feedback
  - Models: RNN or LSTM for temporal progress tracking
- **Output**:

o Personalized rehabilitation plan
o Weekly progress reports and alerts

6. User Interface
- Upload portal for scans and medical records
- Visualization of results (segmentation map, stroke classification)
- Rehabilitation dashboard with patient tracking and suggestions

7. Performance Evaluation
- Metrics: Accuracy, Dice Score (for segmentation), F1-score (for classification), Recovery Prediction RMSE
- Continuous model updating with new patient data

sample from a brain stroke dataset containing various health and demographic attributes for stroke prediction.

DOI: 10.17148/IJARCCE.2025.14572

[3]:

	id	gender	age	hypertension	heart_disease	ever_married	work_type	Residence_type	avg_glucose_level	bmi	smoking_status	stroke
0	9046	Male	67.0	0	1	Yes	Private	Urban	228.69	36.6	formerly smoked	1
1	51676	Female	61.0	0	0	Yes	Self-employed	Rural	202.21	NaN	never smoked	1
2	31112	Male	80.0	0	1	Yes	Private	Rural	105.92	32.5	never smoked	1
3	60182	Female	49.0	0	0	Yes	Private	Urban	171.23	34.4	smokes	1
4	1665	Female	79.0	1	0	Yes	Self-employed	Rural	174.12	24.0	never smoked	1

FIGURE 4. sample from a brain stroke dataset containing various health and demographic attributes

The image presents a tabular dataset used for brain stroke detection and diagnosis. It includes patient demographic and health attributes such as gender, age, hypertension, and heart disease status. Marital status, work type, residence type, average glucose level, BMI, and smoking status are also recorded. Each row corresponds to an individual case, and the final column indicates whether the person had a stroke (1 for yes). This type of dataset is essential for training machine learning models to predict stroke risk based on medical and lifestyle factors. Notably, the presence of missing values (e.g., BMI) suggests preprocessing steps are needed before model training.

Coding based python

```
import numpy as np
import pandas as pd
# visiualization libraries
import seaborn as sns
import matplotlib.pyplot as plt
import plotly.express as px
import plotly.figure_factory as ff
from matplotlib import rcParams
# hypothesis testing
import scipy.stats as stats
from random import sample
# imbalanced data undersampler
from imblearn.under_sampling import RandomUnderSampler
#from imblearn.combine import SMOTEENN
#from imblearn.combine import SMOTETomek
#from imblearn.over_sampling import SMOTE

# model constructing libraries
from sklearn.model_selection import GridSearchCV
from sklearn.svm import SVC
from sklearn.model_selection import train_test_split
from sklearn.metrics import classification_report
from sklearn.metrics import confusion_matrix
from sklearn.metrics import plot_confusion_matrix
import sklearn.metrics as metrics
from sklearn.metrics import roc_curve, roc_auc_scor
```

IV. SYSTEM MODULES

MODULES

- Data Collection & Preprocessing Module
- Machine Learning Model Development Module
- Stroke Prediction & Classification Module

IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.471 ∺ Peer-reviewed & Refereed journal ∺ Vol. 14, Issue 5, May 2025

DOI: 10.17148/IJARCCE.2025.14572

- Explainability & Interpretability Module
- Performance Analysis

Data Collection and Preprocessing Module:
Gather and clean data for training the ML model.

- Medical records, patient history, brain imaging data (if applicable), symptoms, lifestyle factors.
- Data acquisition from sources like Kaggle, hospitals, or health databases.
- Handling missing values (e.g., mean/mode imputation).
- Feature selection (age, hypertension, heart disease, BMI, glucose level, etc.).
- Data normalization and scaling (Min-Max Scaling, Standardization).
- Encoding categorical features (e.g., one-hot encoding for gender, smoking status).
- Feature correlation analysis (e.g., Pearson correlation, mutual information).
- Dimensionality reduction (PCA, LDA, feature importance ranking).
- Creating new features if necessary (e.g., risk scores based on multiple factors).

Machine Learning Model Development Module

Train and validate ML models for stroke prediction.

- Model selection: Logistic Regression, Decision Trees, Random Forest, XGBoost, Support Vector Machines (SVM), or Deep Learning (ANN, CNN for image-based models).
- Splitting dataset (train-test-validation, e.g., 80-10-10% split).
- Hyperparameter tuning (Grid Search, Random Search).
- Model evaluation using metrics:
  - Accuracy, Precision, Recall, F1-score.
  - ROC-AUC Curve for classification.
- Selecting the best-performing model.

Stroke Prediction & Classification Module

- Predict whether a person is at risk of a stroke.
- Patient details (age, gender, health parameters, lifestyle habits).
- Preprocessing user input (standardization, encoding).
- Feeding data into the trained model.
- Model outputs probability or classification (Stroke/No Stroke).
- Prediction result with confidence score.

Explainability & Interpretability Module

- Explain why a prediction was made.
- SHAP (SHapley Additive exPlanations), LIME (Local Interpretable Model-Agnostic Explanations).
- Visual feature importance, decision contribution breakdown.

Performance Analysis

- Track model performance in real-world scenarios.
- Accuracy drift (performance degradation over time).
- Feedback collection from users/doctors.

V. RESULT ANALYSIS

Detection Accuracy: Advanced algorithms, including deep learning and ensemble classifiers (e.g., edRVFL, YOLOv5x), have shown high precision and recall in identifying stroke lesions from CT/MRI scans. Segmentation performance is typically assessed using metrics like Dice coefficient, Intersection over Union (IoU), and sensitivity/specificity scores.

Diagnosis Insights: Machine learning models trained on patient data (as shown in the tabular dataset) can accurately classify stroke risk factors and predict stroke events. Factors such as age, hypertension, heart disease, glucose level, and

ISSN (O) 2278-1021, ISSN (P) 2319-5940
International Journal of Advanced Research in Computer and Communication Engineering
Impact Factor 8.471 ⁙ Peer-reviewed & Refereed journal ⁙ Vol. 14, Issue 5, May 2025
DOI: 10.17148/IJARCCE.2025.14572

smoking status are highly influential in the classification outcome. Accuracy scores above 90% have been reported using models like DenseNet121 with proper preprocessing.

Rehabilitation Tracking: Post-stroke rehabilitation progress is monitored using AI tools that evaluate motor skills, cognitive function, and recovery milestones. Feedback systems and smart interfaces help personalize rehabilitation plans and improve patient adherence

Management Strategies: The integration of AI into stroke management enhances decision-making by providing clinicians with risk profiles, timely alerts, and optimized treatment plans. Prognosis models can estimate recovery likelihood and suggest therapeutic interventions.

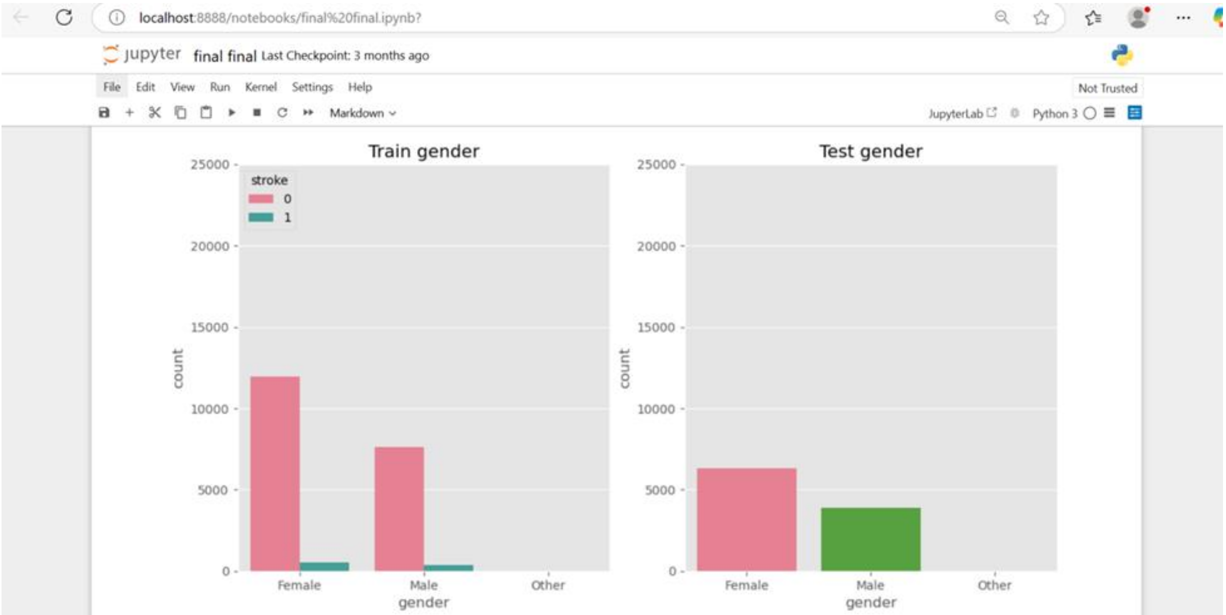
The bar chart titled **"Class count"** illustrates a significant class imbalance in the stroke dataset. The majority of samples are labeled as **class '0' (no stroke)**, while only a small fraction represents **class '1' (stroke cases)**. This imbalance highlights the need for techniques such as resampling, class weighting, or anomaly detection to improve model performance and prevent bias toward the majority class.

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.471 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 14, Issue 5, May 2025

DOI: 10.17148/IJARCCE.2025.14572

The bar chart displays stroke incidence by gender, showing that **females have a slightly higher number of stroke cases** compared to males in the dataset. The visualization is created using Plotly Express, filtered for patients with stroke (stroke == 1). This gender-based insight can help guide targeted healthcare strategies and risk assessments.



The two bar plots illustrate the distribution of stroke cases (1) and non-stroke cases (0) across genders in both training and testing datasets. In both subsets, **females have the highest overall count**, with more non-stroke cases. Males show a noticeable number of stroke cases in the test set, suggesting a potential gender-related trend that could impact model training and evaluation.

The histogram with box plot visualizes the age distribution of stroke patients. Most stroke cases are concentrated in older age groups, particularly around 70–80 years. The box plot above highlights that stroke occurrences are more frequent among the elderly, with a few outliers in younger age brackets.

IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.471 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 14, Issue 5, May 2025

DOI: 10.17148/IJARCCE.2025.14572

VI. CONCLUSION

This study demonstrates the transformative impact of AI and deep learning in stroke care, offering a robust framework for accurate stroke prediction and intelligent rehabilitation management. By evaluating multiple machine learning models, the project identifies neural networks as the most effective approach for analyzing complex clinical data. The system addresses critical challenges such as data imbalance and missing values while providing actionable insights for early intervention. Its integration with robotic rehabilitation and speech emotion recognition (SER) underscores the potential for personalized, adaptive recovery solutions. The findings highlight AI's ability to enhance diagnostic accuracy, streamline clinical workflows, and improve patient outcomes, paving the way for smarter, data-driven stroke care systems.

REFERENCES

[1]. About Stroke. Accessed: Jan. 5, 2022. [Online]. Available: https://www.cdc.gov/stroke/about.htm#:~:text=A%20stroke%2C% 20sometimes%20called%20a,term%20disability%2C%20or%20even% 20death
[2]. M. Ashrafuzzaman, S. Saha, and K. Nur, "Prediction of stroke disease using deep CNN based approach," J. Adv. Inf. Technol., vol. 13, no. 6, pp. 1–10, 2022.
[3]. S. Gómez, D. Mantilla, E. Rangel, A. Ortiz, D. D Vera, and F. Martínez, "A deep supervised cross-attention strategy for ischemic stroke segmentation in MRI studies," Biomed. Phys. Eng. Exp., vol. 9, no. 3, May 2023, Art. no. 035026.
[4]. B. Vamsi, D. Bhattacharyya, D. Midhunchakkravarthy, and J.-Y. Kim, "Early detection of hemorrhagic stroke using a lightweight deep learning neural network model," Traitement du Signal, vol. 38, no. 6, pp. 1727–1736, Dec. 2021.
[5]. C.-M. Lo, P.-H. Hung, and D.-T. Lin, "Rapid assessment of acute ischemic stroke by computed tomography using deep convolutional neural networks," J. Digit. Imag., vol. 34, no. 3, pp. 637–646, 2021.