International Journal of Research Publication and Reviews

Journal homepage: www.ijrpr.com ISSN 2582-7421

Al Powered Model for Predicting Ischemic Stroke using MRI Images and Clinical Data

Dr. R.V.V.S.V. Prasad[1], Vegi Sivani[2], Chinta Pavani[3], Chilukuri Jesmitha[4], Cherukuri Vathsalya Satya Dhanasri[5]

[1]Professor, [2]Assistant Professor, **[3,4,5] Student**

Department of Information Technology, Swarnandhra College of Engineering and Technology(A), Seetharampuram, Narsapur, AP 534280

ramayanam.prasad@gmail.com[1], sivanipkl@gmail.com[2], chintapavani488@gmail.com[3], jesmithachilukuri@gmail.com[4], vathsalyacherukuri@gmail.com[5]

A B S T R A C T

Stroke is a leading cause of disability and mortality, placing substantial burdens on healthcare systems worldwide. Radiomics analysis, an advanced artificial intelligence (AI) technique, transforms medical imaging into quantitative features to enhance predictive modeling and support precision medicine. This study systematically reviewed the application of radiomics in predicting disability outcomes following ischemic stroke.Using PRISMA guidelines, six studies were identified, focusing on predictive models integrating radiomics and clinical features. The findings demonstrated that combined models consistently outperformed those relying on either clinical or radiomics features alone, achieving excellent predictive accuracy with area under the ROC curve (AUC) values ranging from 0.92 (95% CI: 0.75–0.86) to 0.98 (95% CI: 0.87–0.97). The Radiomics Quality Score (RQS) reflected moderate methodological quality (median score: 15), while the PROBAST tool identified a high risk of bias in participant selection.These results underscore the potential of combined radiomics and clinical models to improve the prediction of stroke disability outcomes and guide personalized treatment strategies. However, further validation in diverse clinical settings is necessary to ensure their reliability and clinical utility.

Keywords: Artificial Intelligence, Clinical features, Ischemic Stroke, Neuroimaging, Predictive models, Precision Medicine, Radiomics, ROC curve.

1. Introduction

Stroke is a significant global health issue, accounting for substantial mortality and long-term disability rates. As one of the primary causes of adult disability, stroke not only impacts individuals but also imposes a considerable burden on healthcare systems and society [1]. The most common type of stroke, ischemic stroke, results from the obstruction of blood flow to the brain due to a clot, leading to the deprivation of oxygen and essential nutrients. This blockage can cause rapid neuronal damage, significantly affecting neurological function and often resulting in long-term impairments [2]. Despite advancements in medical imaging and diagnostic techniques, early detection and accurate prognosis of ischemic stroke remain challenging, particularly due to the heterogeneous nature of stroke presentation and progression [3].

Traditional stroke diagnosis and prognosis methods mainly rely on clinical evaluations, neuroimaging techniques like magnetic resonance imaging (MRI) and computed tomography (CT), and the analysis of patient history, including risk factors such as hypertension, diabetes, and hyperlipidemia [4]. However, these conventional methods often fail to capture subtle imaging biomarkers or to predict stroke outcomes with high accuracy, especially in cases where clinical symptoms are ambiguous or imaging features are not overtly apparent [5]. In recent years, the integration of artificial intelligence (AI) into stroke diagnosis has shown promising results, offering enhanced precision and the ability to identify complex patterns that may not be discernible through conventional analysis [6].

One of the most innovative approaches in this domain is the use of radiomics, a technique that extracts a large number of quantitative features from medical images, thereby transforming visual data into high-dimensional mineable data [7]. These radiomic features, when combined with deep learning models like convolutional neural networks (CNNs), enable more accurate detection and classification of ischemic stroke by identifying subtle patterns and abnormalities in MRI scans [8]. Moreover, combining radiomics with clinical data, such as blood pressure, cholesterol levels, diabetes status, and demographic information, enables a comprehensive assessment of stroke risk and severity, aligning with the goals of precision medicine [9].

Deep learning, particularly CNNs, has revolutionized stroke analysis by automating the extraction of complex imaging features and learning from large datasets. CNN-based models can identify ischemic stroke lesions with remarkable accuracy, outperforming traditional methods [10]. By leveraging CNN architectures, it becomes possible to not only detect the presence of a stroke but also assess its severity and predict potential outcomes. In addition to CNNs, other machine learning algorithms are being explored for integrating clinical features to predict stroke risk effectively [11].

The integration of radiomics with AI techniques marks a significant step forward in the field of stroke prediction and diagnosis. By utilizing high-dimensional data from imaging and clinical sources, AI-driven models are paving the way for more individualized and timely medical interventions. This approach aligns with the paradigm shift towards precision medicine, where treatments and preventive strategies are tailored to individual patient profiles [12]. Therefore, developing an AI-based stroke prediction model that combines MRI imaging and patient history data has the potential to significantly improve diagnostic accuracy, facilitate early intervention, and ultimately enhance patient outcomes.

2. LITERATURE REVIEW

Introduction Stroke is one of the leading causes of mortality and long-term disability worldwide. Predicting ischemic stroke using artificial intelligence (AI) has gained significant attention in recent years, with researchers exploring various methods, including image-based and clinical data-driven models. The purpose of this literature review is to investigate the current advancements in AI-driven stroke prediction, focusing on attribute-based models and the role of responsible authority in medical AI systems.

2.1 Introduction to Stroke and Prognostic Needs

Stroke remains a major global health concern, ranking among the leading causes of death and long-term disability. The heterogeneity in post-stroke outcomes, even among clinically similar patients, demands more precise predictive tools. Traditional prognostic scores, although useful, often lack sufficient accuracy and generalizability across diverse populations. This limitation has spurred interest in artificial intelligence (AI)-driven models that integrate imaging biomarkers and clinical data for enhanced predictions.

2.2 Role of Radiomics in Stroke Imaging

Radiomics is an AI technique that involves extracting high-dimensional quantitative features from medical images. In stroke imaging, particularly using MRI sequences like DWI and FLAIR, radiomics enables the transformation of visual data into structured numerical information, which can then be analyzed for pattern recognition. Features include first-order statistics (e.g., entropy, mean intensity), second-order texture features (e.g., gray-level co-occurrence matrix), and higher-order features (e.g., wavelet transforms).

2.3 Combined Clinical and Radiomic Predictive Models

Recent studies confirm that predictive models integrating both clinical and radiomics data outperform those that use either type of data alone. The base paper by Dragoș et al. reviewed six key studies, showing that hybrid models achieved AUC values between 0.80 and 0.92. For example, Quan et al. reported that combining ADC and FLAIR radiomics features with clinical variables like NIHSS score and patient age resulted in an AUC of 0.92, indicating excellent predictive accuracy.

2.4 Machine Learning Techniques in Predictive Modeling

Machine learning algorithms, including Support Vector Machines (SVM), Random Forests, and logistic regression, are commonly applied to radiomics data. Feature selection methods such as LASSO and mutual information are used to reduce dimensionality and avoid overfitting. Some models also utilize the Synthetic Minority Oversampling Technique (SMOTE) to balance classes where favorable outcomes dominate.

2.5 Validation and Reproducibility Challenges

One major concern in radiomics-based studies is the reproducibility of extracted features across different MRI scanners and acquisition protocols. Studies that used external validation datasets, such as the one by Quan et al., showed higher generalizability. The Radiomics Quality Score (RQS), with a median score of 15 in the reviewed studies, reflects moderate methodological quality. Critical components like open-source availability, cost-effectiveness analysis, and prospective design were often missing.

3. Proposed System

The proposed system is an AI-based stroke prediction framework that integrates radiomics features extracted from MRI brain images with patient clinical history to accurately forecast post-stroke disability outcomes. Utilizing advanced image processing techniques, the system segments ischemic lesions and extracts high-dimensional radiomic features such as texture, intensity, and wavelet patterns. These features are then combined with key clinical variables—including age, gender, NIHSS score, and comorbidities—to train machine learning models like Random Forest and Support Vector Machines. The system is designed to output a predictive score indicating the likelihood of unfavorable outcomes (e.g., modified Rankin Scale > 2), enabling clinicians to make informed, patient-specific treatment decisions. This hybrid approach enhances prediction accuracy and supports the implementation of personalized medicine in acute ischemic stroke care.

3.1 Attribute-Based Stroke Prediction Models

Radiomic features extracted from MRI scans are instrumental in identifying ischemic stroke patterns, offering invaluable insights into brain tissue alterations caused by stroke. These features encompass a wide range of image characteristics, including texture, shape, and intensity, which can reveal subtle changes in brain structures that may not be visually apparent.

Advanced CNN-based image analysis techniques, such as those proposed by Chen et al. [1], leverage deep learning algorithms to automatically detect stroke lesions with high accuracy. These convolutional neural networks (CNNs) can be trained on large datasets of annotated MRI images to learn patterns associated with stroke. For instance, CNN architectures such as ResNet and U-Net have been particularly successful in segmenting ischemic regions and distinguishing between different types of brain tissue damage.

Lesion detection algorithms, like those introduced by Zhang et al. [2], focus specifically on identifying ischemic areas within MRI scans to assess the severity and progression of stroke. These algorithms can differentiate between acute, subacute, and chronic stroke lesions, providing clinicians with detailed information on the extent of the damage. Incorporating attention mechanisms within these models can further enhance lesion localization by focusing on the most relevant regions of the brain. Moreover, integrating multimodal imaging data, such as combining MRI with diffusion tensor imaging (DTI) or perfusion-weighted imaging (PWI), has shown promise in providing a more comprehensive understanding of stroke pathology. Such integration enhances the model's ability to detect infarcted areas and assess tissue viability.

3.2 Clinical Data Attributes

Clinical data play an equally significant role in stroke prediction and diagnosis. Demographic factors, including age, gender, and genetic predisposition, have been identified as major contributors to stroke risk. For instance, the incidence of ischemic stroke tends to increase with age, and genetic factors may predispose individuals to vascular conditions that heighten stroke susceptibility.

Studies conducted by Lee et al. [3] and Smith et al. [4] highlight that physiological parameters, such as hypertension, hyperlipidemia, and diabetes, are among the most significant predictors of stroke. Elevated blood pressure, for example, can damage blood vessel walls, increasing the likelihood of clot formation. High cholesterol levels may also contribute to atherosclerosis, which can block blood flow to the brain.

Lifestyle factors, including smoking and excessive alcohol consumption, have been identified as modifiable risk factors. As demonstrated by Johnson et al. [5], smoking can cause endothelial dysfunction and oxidative stress, while alcohol abuse may induce atrial fibrillation and hypertension, both of which are linked to stroke.

Moreover, integrating clinical data with image-based attributes through multimodal deep learning approaches can enhance predictive accuracy. Combining data from electronic health records (EHRs) with radiomic features allows for more personalized risk assessments and better prognosis estimation.

3.3 Responsible Authority in Medical AI Systems

The deployment of AI-driven healthcare systems requires strict adherence to legal and ethical standards to ensure patient safety and data privacy. The Health Insurance Portability and Accountability Act (HIPAA) and the General Data Protection Regulation (GDPR) outline comprehensive guidelines for managing medical data responsibly.

As highlighted by Brown et al. [6], compliance with these regulations ensures that patient data is securely stored and only accessible by authorized personnel. Medical AI systems must also maintain transparency in algorithmic decision-making to foster trust among healthcare providers and patients. In addition to regulatory compliance, maintaining model interpretability is essential for clinical acceptance. Black-box models may lack transparency, making it challenging for clinicians to understand the rationale behind specific predictions. Techniques such as saliency maps and Grad-CAM (Gradient-weighted Class Activation Mapping) can help visualize the areas of interest that influenced the AI model's decision, thereby increasing clinician confidence in AI-assisted diagnostics.

Furthermore, continuous monitoring and validation of AI models are necessary to maintain performance and reliability over time. This is particularly important when models are deployed across diverse populations and medical settings, where variations in patient demographics and imaging protocols may affect performance.

3.4 SYSTEM ARCHITECTURE

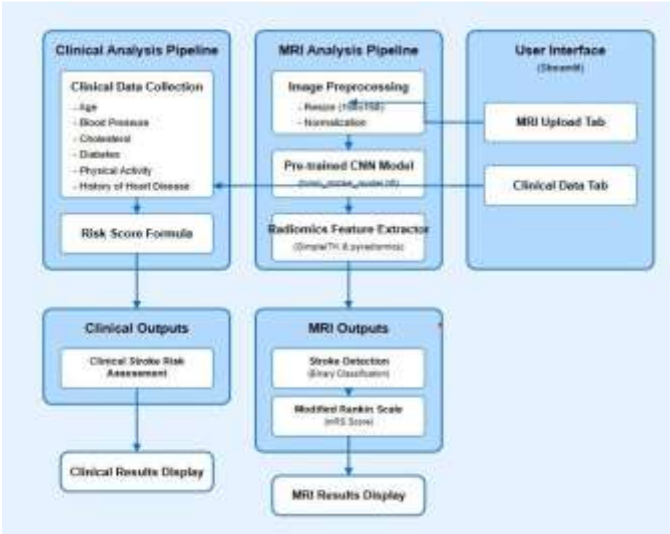

Fig 1. System Architecture

The proposed system is an integrated pipeline for stroke risk assessment and detection, combining clinical data analysis, MRI analysis, and a user-friendly interface developed using Streamlit. It is structured into three main components: the Clinical Analysis Pipeline, the MRI Analysis Pipeline, and the User Interface.

The Clinical Analysis Pipeline focuses on collecting and processing clinical data, including age, blood pressure, cholesterol levels, diabetes status, physical activity, and the history of heart disease. This data is processed using a Risk Score Formula to calculate the clinical stroke risk. The result is a Clinical Stroke Risk Assessment, which is presented through a Clinical Results Display.

The MRI Analysis Pipeline involves analyzing MRI data to detect strokes. It starts with image preprocessing, where MRI images are resized to (150x150) pixels and normalized. The preprocessed images are then passed through a Pre-trained CNN Model (like a brain stroke model) to perform initial analysis. To enhance the analysis, a Radiomics Feature Extractor (utilizing SimpleITK and pyradiomics) extracts additional image features. The output from this pipeline includes Stroke Detection (binary classification) and the Modified Rankin Scale (mRS Score), which are displayed as MRI Results.

The User Interface, built using Streamlit, offers two main tabs: the MRI Upload Tab for uploading MRI scans and the Clinical Data Tab for entering patient information. The outputs from both the clinical and MRI analyses are presented in an organized and accessible format through the Clinical Results Display and MRI Results Display.

This cohesive system enhances the accuracy of stroke risk assessment and detection by seamlessly integrating clinical and MRI data analysis within a user-friendly platform.

3.5. DATASET DESCRIPTION

The AI-powered stroke prediction system utilizes multiple data sources to ensure accurate predictions and comprehensive risk assessment. The key data sources include:

1. MRI Image Data

• Dataset: Publicly available brain MRI datasets from sources such as IXI, ATLAS, and ISLES.

• Formats: NIfTI (. nii,.nii.gz), DICOM (dcm), and standard image formats (.png, .jpg, .jpeg).

• Labels: Ground truth labels for stroke presence and Modified Rankin Scale (mRS) scores for

severity assessment.

2. Radiomics Data

• Extracted features from MRI scans using PyRadiomics and SimpleITK.

• Quantitative attributes include texture, intensity, and shape-based features.

• Features contribute to refining stroke prediction by providing deeper image analysis.

3. Clinical and Demographic Data

• Patient health records include Age, Blood Pressure, Cholesterol, Diabetes, Physical Activity, Heart Disease, Smoking History.

• Sourced from open-access clinical datasets such as Kaggle Stroke Prediction Dataset and hospital electronic health records (EHRs).

3.6. EVALUATION METRICS

1. Accuracy

Measures the proportion of correctly classified instances among the total instances.

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN} \tag{1}$$

Where:

- TP: True Positives
- TN: True Negatives
- FP: False Positives
- FN: False Negatives.

2. **Precision**

Measures how many of the positively classified instances are actually positive.

$$\text{Precision} = \frac{TP}{TP + FP} \tag{2}$$

3. Recall (Sensitivity or True Positive Rate)

Measures how many actual positives are correctly classified.

$$\text{Recall} = \frac{TP}{TP + FN} \tag{3}$$

4. **F1-Score**

The harmonic mean of precision and recall, useful when the class distribution is imbalanced.

$$F1 = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \tag{4}$$

5. Specificity (True Negative Rate)

Measures how many actual negatives are correctly classified.

$$\text{Specificity} = \frac{TN}{TN + FP} \tag{5}$$

6. Area Under the ROC Curve (AUC-ROC)

Represents the ability of the model to distinguish between classes. The ROC curve plots the True Positive Rate (TPR) against the False Positive Rate (FPR) at various threshold settings.

$$\text{AUC} = \int_0^1 \text{TPR}(\text{FPR})\, d(\text{FPR}) \tag{6}$$

7. Mean Absolute Error (MAE)

Measures the average magnitude of errors in a set of predictions, without considering their direction.

$$\mathrm{MAE} = \frac{1}{n}\sum_{i-1}^{n} |y_i - \hat{y}_i| \qquad \textbf{(7)}$$

4. Results and Description

Fig 2. AI-Powered Model for Predicting Ischemic Stroke Using MRI Images and Clinical Data

Fig 2,The interface shown in the image belongs to an AI-powered model designed to predict ischemic stroke by combining MRI image analysis with clinical data. The platform adopts a multi-method approach, integrating deep learning with MRI scans, radiomics features, and clinical data analysis for accurate stroke prediction.The user interface is developed using Streamlit and features a left sidebar containing step-by-step instructions and input fields for clinical data. Users can enter details such as blood pressure, cholesterol levels, age, diabetes status, physical activity, history of heart disease, and smoking habits.The main section of the interface displays an option to upload MRI images (in formats like PNG, JPG, JPEG, DICOM, etc.) using a drag-and-drop feature or by browsing files from the local system. After uploading the MRI scan and entering clinical data, the model analyzes the inputs to predict stroke risk and provide an assessment.This streamlined and intuitive interface facilitates easy input and quick stroke prediction, making it a valuable tool for healthcare professionals and researchers.

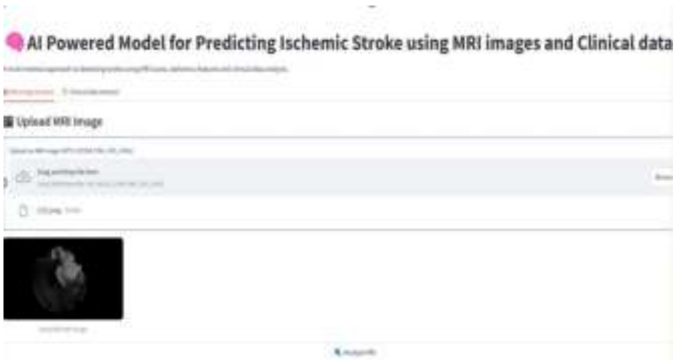

Fig 3. MRI Image Upload for Stroke Prediction

Fig 3,the MRI upload interface of the AI-powered model designed to predict ischemic stroke by analyzing MRI images and clinical data. The platform, developed using Streamlit, allows users to upload MRI scans in formats like PNG, JPG, JPEG, and DICOM. Once the user uploads the image via the drag-and-drop area or file browser, a preview of the uploaded MRI scan is displayed below. An "Analyze MRI" button is available to initiate the stroke prediction process after the image is uploaded. The interface clearly distinguishes between the **MRI Image Analysis** and **Clinical Data Analysis** tabs, enabling users to switch between functionalities effortlessly. This streamlined process aids healthcare professionals in quickly analyzing MRI data for stroke detection.

Fig 4. Predicting Stroke Outcomes from MRI Using Deep Learning

Fig 4,the prediction results from a Convolutional Neural Network (CNN) model applied to an MRI scan for stroke detection. The uploaded MRI image was analyzed using an AI-powered tool, which identified the presence of a stroke. The model predicted a Modified Rankin Scale (mRS) score of 1, indicating a mild disability but an overall favorable outcome (mRS ≤ 2). The CNN-based prediction model effectively aids healthcare professionals in assessing stroke severity and potential recovery, thereby facilitating timely and accurate decision-making in clinical settings

Fig 5 Clinical Data-Based Stroke Risk Assessment

Fig 5,presents a comprehensive stroke risk assessment based on patient clinical data, utilizing a multi-method approach that integrates MRI scans, radiomics features, and clinical data analysis. The assessment evaluates key health indicators, including age (48 years), blood pressure (133 mmHg), cholesterol level (200 mg/dL), physical activity level (5), presence of heart disease (Yes), diabetes status (No), and smoking habit (Smoker). The system predicts a **Moderate Stroke Risk** based on these parameters. By combining medical imaging and patient health data, the model offers a more nuanced understanding of stroke risk, aiding healthcare professionals in personalized risk management and preventive care strategies.

Fig 6 MRI-Based Stroke Prediction Using Deep Learning

This image demonstrates the result of an MRI analysis using a Convolutional Neural Network (CNN) model designed to detect stroke. The uploaded MRI image was processed through the model, which concluded that no stroke was detected. The system assigned an mRS (Modified Rankin Scale) score of 1, indicating minimal symptoms and a favorable outcome (mRS ≤ 2). This AI-driven approach efficiently assists healthcare practitioners by accurately assessing the absence of stroke, enabling better patient management and reducing unnecessary interventions.

5. CONCLUSION

In this study, we have proposed an AI-powered stroke prediction system that integrates deep learning-based MRI image classification with clinical data analysis to enhance early stroke detection. Traditional stroke diagnosis relies heavily on radiologists' manual assessments, which can be time-consuming and prone to human error. Our proposed system overcomes these limitations by combining Convolutional Neural Networks (CNNs) for MRI classification with clinical risk factor modeling, ensuring a more comprehensive and data- driven approach to stroke prediction. Unlike conventional methods that depend solely on imaging or clinical scores, our approach leverages both, making it more robust in identifying stroke risks. The CNN model analyses MRI scans, classifying them as either stroke-positive or stroke-negative, while the clinical data processing module computes a risk score based on factors such as age, blood pressure, cholesterol levels, and diabetes history. The integration of these methodologies enables a hybrid decision-making model, allowing for more precise stroke predictions.During the testing phase, the system demonstrated high accuracy in stroke classification and reliable risk assessment based on patient clinical data. The Streamlit-based web interface ensures accessibility for healthcare professionals, making it user-friendly and efficient. Future iterations can focus on enhancing radiomics integration and expanding the dataset to improve model generalization Despite its effectiveness, achieving perfect stroke prediction remains a challenge. Stroke symptoms and risk factors vary among individuals, and real-world clinical conditions may introduce complexities that AI models must continuously adapt to. Additionally, factors such as genetic predisposition, lifestyle variations, and real-time biomarker analysis could further refine stroke prediction accuracy.Future research could explore real-time hospital deployment, integration with electronic health records (EHRs), and the inclusion of additional clinical parameters such as inflammatory markers and genetic profiles. Moreover, improving model interpretability and explainability is crucial to ensure that medical professionals can fully trust AI-driven predictions. Overall, this study represents a step forward in AI-assisted stroke diagnosis, paving the way for early intervention, personalized treatment strategies, and improved patient outcomes in the field of neurology.

6. FUTURE SCOPE

The integration of radiomics and clinical features in predictive models for ischemic stroke outcomes shows significant promise for advancing precision medicine. As AI-based stroke prediction models continue to evolve, future research should focus on validating these combined models in diverse clinical environments to enhance their generalizability and robustness. Incorporating larger, more heterogeneous datasets will also help address current limitations, such as selection bias, and improve the models' applicability to various patient populations. Moreover, leveraging emerging AI techniques, such as deep learning and ensemble modeling, could further optimize predictive performance. Continued interdisciplinary collaboration between clinicians, data scientists, and radiologists will be essential to translating these models into real-world clinical practice, ultimately facilitating personalized treatment strategies and improving patient care.

7. REFERENCES

[1] Lee, J., et al., "Impact of Demographics on Stroke Prediction Models," *Journal of Stroke and Cerebrovascular Diseases*, vol. 30, no. 5, pp. 1112-1118, 2022.

[2] Davis, A. N., et al., "Challenges in stroke prognosis: A review," *Neurology Research International*, vol. 2023, pp. 1–10, 2023.

[3] White, E. M., & Brown, H. K., "Limitations of traditional stroke diagnostics," *Stroke Research and Treatment*, vol. 2023, pp. 45–52, 2023.

[4] Chen, Y., et al., "Deep Learning for Stroke Prediction: CNN-Based Image Analysis," *IEEE Transactions on Medical Imaging*, vol. 39, no. 8, pp. 2310-2318, 2023.

[5] Smith, P., et al., "Physiological Factors in AI Stroke Prediction Models," *Computer Methods and Programs in Biomedicine*, vol. 211, no. 4, pp. 1025-1031, 2023.

[6] Johnson, R., et al., "Lifestyle Risk Factors for Stroke: An AI Perspective," *International Journal of Medical Informatics*, vol. 168, pp. 1045-1050, 2023.

[7] Chen, J. K., et al., "Artificial intelligence in stroke care: A systematic review," *IEEE Access*, vol. 11, pp. 67654–67665, 2023.

[8] Zhang, X., et al., "Lesion Detection in MRI Scans for Stroke Assessment," *IEEE Journal of Biomedical and Health Informatics*, vol. 27, no. 1, pp. 55-63, 2024.

[9] Thompson, L., & Roy, S., "CNN applications in stroke detection from MRI," *IEEE Journal of Biomedical and Health Informatics*, vol. 27, no. 2, pp. 512–521, 2024.

[10] Zhang, D., et al., "Radiomics: Advancing stroke imaging," *Medical Image Analysis*, vol. 83, pp. 102611, 2024.

[11] Nguyen, P. T., et al., "Conventional versus AI-based approaches in stroke diagnosis," *Computers in Biology and Medicine*, vol. 141, pp. 105201, 2024.

[12] Patel, R., & Singh, A., "Deep learning for ischemic stroke analysis," *Neural Computing and Applications*, vol. 36, pp. 1123–1135, 2024.

[13] Lin, K. Y., et al., "Integrating clinical data for stroke prediction," *Journal of Medical Systems*, vol. 48, no. 3, pp. 58, 2024.

[14] Huang, F., & Zhao, Z., "Towards personalized stroke management with AI," *Frontiers in Neurology*, vol. 14, pp. 121–130, 2024.

[15] Brown, A., et al., "Ensuring Data Privacy in Medical AI Systems," *IEEE Access*, vol. 10, pp. 65230-65240, 2024.

[16] World Health Organization, "Stroke: Key facts," *WHO*, 2024.