© DEC 2025 | IRE Journals | Volume 9 Issue 6 | ISSN: 2456-8880
DOI: https://doi.org/10.64388/IREV9I6-1713022

Brain Stroke Prediction Using Machine Learning

PRADEEP NAZARETH[1], VINAYAK M NAGANUR[2], SANJEEV REDDY[3], GURUSIDDA[4], MITHUN GOWDA SR[5], GANESH AP[6]
[1, 2, 3, 4, 5, 6]*Department of Artificial Intelligence and Machine Learning, Alva's Institute of Engineering and Technology, Mangalore, Karnataka, India*

Abstract- This review delves deeply into deep learning techniques for brain stroke prediction. It focuses on real-time Internet of Medical Things (IoMT) systems and mixed architectures of convolutional and recurrent neural networks. The study looks at preprocessing, feature selection, model clarity, and ways to boost prediction accuracy. It also explores how big medical datasets, transfer learning, and real-time analysis help build strong stroke prediction models. The paper examines several cutting-edge frameworks to see how well they work for early diagnosis and decision support. Deep learning algorithms can spot subtle patterns in brain scans and patient records that normal stats often miss. Also, using cloud computing, edge AI, and mobile health platforms allows for quick stroke risk checks in far-off or underserved areas. The paper reviews explainable AI techniques to increase trust in clinics and talks about ethical issues like data privacy and fairness. It analyzes challenges recent progress, and future paths to create efficient, accurate, and usable stroke prediction systems.

Index Terms—Multimodal AI, Fake News Detection, Misinfor- Mation, Bi-LSTM, CNN, Web Intelligence, Social Media Analytics.

I. INTRODUCTION

This review offers a thorough look at machine learning (ML) methods to predict brain strokes covering both old and new approaches to analyze clinical data and interpret brain scans. The paper looks at common ML algorithms like Decision Trees, Support Vector Machines (SVM), Logistic Regression, Naive Bayes, Random Forests, Gradient Boosting Machines, and k-Nearest Neighbors (k-NN). It also explores ensemble and hybrid models that aim to boost prediction accuracy by combining algorithms. The paper highlights feature selection techniques—such as recursive feature elimination (RFE) mutual information gain, and principal component analysis (PCA)—as crucial to cut down data size and remove noise. The review talks about preprocessing and handling un- even data (e.g., SMOTE and ADASYN) as key steps to get reliable predictions from diverse and often incomplete medical datasets. It also stresses the importance of electronic health records (EHRs), The study examines how data from wearable sensors and mobile health (mHealth) tracking systems could help monitor stroke risk over time for older people or those at high risk. ML-based stroke risk models use patient information, lifestyle factors, and related health issues (e.g., diabetes, high blood pressure, heart problems) as predictors. Addition- ally, the paper examines ethical concerns regarding bias in training data, equity among various groups, and the importance of maintaining data privacy and adhering to healthcare regulations (such as HIPAA and GDPR). To compare how well different ML methods work for predicting strokes, they use measures such as accuracy, sensitivity, specificity, AUC-ROC, and F1-score. To wrap up, the review points out current problems—including small datasets, issues with applying models, and no standard way to compare them—and suggests where research should go next. This includes combining different types of data, making real-time analysis faster, making models easier to understand, and doing more clinical trials to help these ideas work in the real world. The paper dives into some of the most popular Decision Trees, Support Vector Machines (SVM), Logistic Regression, Naive Bayes, Random Forests, Gradient Boosting Machines, and k-Nearest Neighbors (k-NN) are examples of machine learning algorithms. Additionally, it investigates ensemble and hybrid models, which combine various algorithms to improve prediction accuracy. It high- lights the importance of feature selection methods such as principal component analysis (PCA), mutual information gain, and recursive feature elimination (RFE). These methods are crucial for trimming down data dimensionality and cutting out unnecessary noise. Additionally, the discussion covers preprocessing and strategies for handling data imbalance, such as SMOTE and ADASYN.

II. LITERATURE REVIEW

Bojsen, J.A., Graumann, O., Elhakim, M.T., et al. (2024)
[1] – A meta-analysis and systematic review of artificial intelligence for MRI stroke detection. This review looked at AI models for ischemic stroke detection from MRI. The pooled sensitivity and specificity were both roughly 93

Wang, Y., Zhang, Z., Zhang, Z., et al. (2025) [2] — A systematic review and meta-analysis of machine learning and traditional models for predicting hemorrhagic transformation in ischemic stroke. They compared machine learning methods with traditional risk scores to predict hemorrhagic transformation following a stroke. Heterogeneity in feature sets and data limited reproducibility even though ML models demonstrated superior discrimination.

In 2025, Soladoye, A.A., Aderinto, N., Popoola, M.R., et al. published a systematic review of algorithms, datasets, and regional gaps in machine learning techniques for stroke prediction. [3] The application of Random Forests, SVMs, and boosting algorithms in 58 machine learning studies on stroke prediction was highlighted in this review. It identified regional gaps in dataset availability and highlighted the necessity of diverse, representative cohorts.

Wang, Z., Jiang, C., Zhang, X., et al. (2025) [4] — Machine learning-based prognostic prediction for acute ischemic stroke using whole-brain and infarct multi-PLD ASL radiomics. developed machine learning models using multi-PLD ASL radiomics (whole brain + infarct regions) to predict functional outcomes in acute ischemic stroke. The combined features improved prognostic accuracy when compared to infarct-only models.

Su, H.-Y., Sung, S.-F., Tsai, C.-L., et al. (2024) [5]— Using DWI ADC MRI with clinical data, they proposed a fusion deep contrastive learning model for 3-month outcome prediction in stroke patients. AUC 0.87 and F1-score 0.80 were obtained, outperforming earlier multimodal methods.

Liu, Y., et al. (2023) [6] developed a fused DWI MRI + clin- ical deep learning model to forecast 90-day mRS outcomes. Internal AUC is 0.92 and external AUC is 0.90. Within ±1 mRS, the prediction accuracy ranged from 79 to 83 percent, which is comparable to that of experienced clinicians.

A multimodal autoencoder + LSTM model was proposed by Rousseau, D., Hatami, N., Mechtouff, L., et al. (2023) [7] to predict the outcomes of strokes using MRI data. AUC 0.71 was achieved, effectively managing volumetric multimodal MRI features.

Using ML algorithms and CTA radiomics of carotid plaques, Yang, H.-Y., Li, Z.-L., Lv, X.-X., et al. (2025) [8] assessed the risk of an ischemic stroke. showed outstanding discriminative performance in identifying high-risk individu-als.

Farina, E., Mohd, N., Abdullah, A. R., and Kandaya, S. A Review of Machine Learning-Based Acute Ischemic Stroke Neuroimaging (2024) [9]. This review, which looks at various machine learning techniques applied to neuroimaging data for acute ischemic stroke detection, highlights the potential and challenges of clinical implementation.

A Systematic Review of Multimodal Deep Learning for Stroke Diagnosis and Prognosis by Shurrab, S., and Al- Haddad, S. (2024) [10]. The authors thoroughly review multimodal deep learning methods for stroke diagnosis and prognosis while addressing the integration of clinical and imaging data.

Parimala, N., and Muneeswari, G. (2023) published a review of binary classification and hybrid segmentation of brain stroke using a transfer learning-based approach. [11] This paper investigates methods for hybrid segmentation and binary classification in brain stroke detection with a focus on transfer learning strategies.

Le, D., and Nguyen, K. (2024) [12] — Multimodal Deep Learning for Ischemic Stroke Prediction by Integrating Demo- graphic, Clinical, and Atrial Phenotypic and Genotypic Data. The study looks into the use of multimodal deep learning models for ischemic stroke prediction that incorporate genetic, clinical, and demographic data.

Browning, D., and Courtman, M. (2024) [13]— A Com- prehensive Analysis of Machine Learning for Stroke Risk Prediction Using Common Hospital Data. Using routine hospital data, this systematic review evaluates the precision and usefulness of machine learning models developed to predict stroke risk.

Asadi, F., and A. Paghe (2024) [14]— An organized analysis of the best machine learning algorithms for predicting strokes. The authors provide insightful viewpoints on the suitability and performance of various machine learning algorithms by carefully examining their efficacy in stroke prediction.

Shirini, K., and Marzbani, H. (2024) [15]— CNN-Res: A Deep Learning Framework for Acute Ischemic Stroke Lession Segmentation in Multimodal MRI. CNN-Res, a deep learning framework for segmenting acute ischemic stroke lesions from multimodal MRI data, is presented in this work.

A systematic review of hemorrhagic transformation pre- diction models using machine learning and deep learning algorithms in stroke medicine (Issaiy Zarei, 2024) [16]. The study looks at machine learning and deep learning algorithms used in stroke medicine, with a focus on models that predict hemorrhagic transformation.

(Qdaih, I., Al-Haddad, S., 2024) [17] Hybrid Ensemble Deep Learning Model for Enhancing Ischemic Brain Stroke Detection and Classification in Clinical Applications. This study presents a hybrid ensemble deep learning model to enhance ischemic brain stroke detection and classification in clinical settings.

Multimodal Deep Learning for Ischemic Stroke Prediction by Combining Clinical, Atrial Phenotypic, and Genotypic Data (Bai Leifer, 2024) [18]. The authors examine the integration of multimodal data, including genetic, clinical, and demographic data, into deep learning models for ischemic stroke prediction. Shiwlani, A., and Ahmad, A. Deep Learning: A Systematic Review. (2024) [19]. This systematic review provides an overview of deep learning techniques with applications in several domains, including stroke prediction.

Zhao, Q., and Jiang, Y. Examining Prehospital Delays in Recurrent Acute Ischemic Strokes: Interpretable Machine Learning Perspectives (2024) [?]. The study investigates pre- hospital delays in recurrent acute ischemic stroke using interpretable machine learning techniques and offers suggestions for improving patient outcomes.

III. METHODOLOGY AND EXPERIMENTAL FRAMEWORK

A. Dataset collection and pre-processing

The method for predicting strokes makes use of the ability of different machine learning models to examine both un- structured patient records and structured clinical data. Notably, ensemble models and recurrent neural networks (RNNs) have demonstrated exceptional efficacy in capturing both dynamic health indicators like blood pressure, cholesterol levels, and heart rate over time, as well as static risk factors like age, hypertension, and diabetes. The Kaggle Stroke Prediction Dataset and other benchmark datasets were used to train and validate the hybrid RF-LSTM model. It was then contrasted with more conventional classifiers like Random Forests, Support Vector Machines (SVM), and Logistic Regression.

assess performance using common metrics like accuracy, The utilization of precision, recall, F1-score, and Area Under the ROC Curve (AUCROC) provided a thorough understanding of each model's diagnostic capabilities. We used some advanced feature engineering techniques like Chi-square se- lection, recursive feature elimination (RFE), and correlation-based filtering to pinpoint the most important forecasters. To address the problem of class imbalance—something we frequently observe that positive stroke cases are quite un- common—we also used SMOTE and other oversampling techniques. Inspired by DeepDetect, the Deep-Risk framework incorporates preprocessing techniques like eliminating outliers, normalizing patient characteristics, and filling in missing values. In actual healthcare settings, these procedures significantly increase the model's stability and dependability. In order to help interpret predictions and foster trust among clinicians, particularly in crucial situations like emergency care, we also incorporated explain ability techniques like SHAP and LIME.

- This study zeroes in on predicting brain strokes, focusing on a data processing pipeline that includes key preprocessing steps like removing outliers, normalizing clinical parameters, and reducing dimensionality to effectively manage diverse and high-dimensional patient data. These steps are vital for boosting model performance when tackling the complexities of realworld healthcare datasets, which often come with missing values, skewed distributions, and class imbalances. You can see the workflow of the proposed Deep-Risk Stroke Detection System illustrated

- The model uses recurrent neural networks, namely Long Short-Term Memory (LSTM) and Gated Recurrent Units (GRU), to identify temporal and sequential patterns in patient health data, such as trends in blood pressure, glucose levels, and heart rate over time. These networks are especially good at analyzing timeseries of medical data gathered from wearable sensors or electronic health records (EHR). RNNs greatly improve the model's capacity to identify early warning indicators of a stroke, even in cases where the signals are weak or erratic, by identifying patterns over time. The system can make more precise and timely predictions by modeling these temporal health trends, which is essential for early intervention and outcome management. [13]
- Principal Component Analysis (PCA): reduces computations by mapping data into lower-dimensional subspaces that preserve data variance.
- Information Gain: Chooses the features contributing most to classification.
- Overall, these approaches are generally defined when it comes to interpretability as well as predictive power for fake news detection systems. [3]

B. CONVOLUTIONAL NEURAL NETWORKS

The literature describes a range of hybrid deep learning methods that leverage complementary neural architectures:

- CNNs are now the preferred method for evaluating brain MRI, CT, and DWI images, particularly in the early detection of ischemic and hemorrhagic strokes. Typically, the procedure begins with a few crucial preprocessing steps, such as data augmentation, resizing, normalization, and artifact removal. These actions improve the model's capacity for generalization and foster consistency. By doing this, CNNs are able to automatically identify significant spatial characteristics, such as lesion shape, asymmetry, and variations in intensity between the hemispheres of the brain—all of which are critical indicators of stroke-related problems. Modern architectures such as ResNet50, Dense Net, and EfficientNet-B0/B7 are preferred because of their ability to retain intricate details and comprehend context in multi-channel brain scans. In order to understand spatial hierarchies, these models make use of deep layers and skip connections, which enable them to recognize both general patterns and minute abnormalities in the brain. A SoftMax activation function, which can handle binary predictions (stroke vs. no-stroke) or multiclass predictions (ischemic, hemorrhagic, transient ischemic attack), is typically used to generate the final classification through fully connected layers.
- To boost performance even further, ensemble methods like soft The results are combined using voting or weighted averaging. different CNN models. This method improves resilience and minimizes sensitivity to noise or variations in data, which is vital in medical environments. CNN-based stroke prediction models are well-known for their precision, speed, and scalability, and They are ideal because they require few manually created features. for practical uses in medical image analysis. [14]
- The use of attention-based CNN modules, which focus computational power on the most crucial areas for diagnosis, is one intriguing development. This improves the model's performance as well as its interpretability. It is essential for reducing false positives and enhancing the readability of results, both of which are becoming more and more critical in AI-driven clinical diagnostics. Furthermore, a lot of systems are now integrating CNNs with clinical metadata such as age, blood pressure, cholesterol, and history of strokes. Through fully connected layers or multimodal fusion layers, this hybrid modeling technique combines image-based deep learning with structured data to produce more

thorough and context-sensitive stroke risk assessments.

Across the studies, hybrid architectures are, consistently, more effective than using a single modeling approach, because they leverage both spatial and temporal context.

C. Training and Evaluation

One exciting development is the use of attention-based CNN modules that direct computational power toward the most diagnostically important areas. This not only boosts the model's interpretability but also enhances its performance. It's crucial for minimizing false positives and improving the clarity of results, which is becoming increasingly important in AI- driven clinical diagnostics. Additionally, many systems are now combining CNNs with clinical metadata like age, blood pressure, cholesterol levels, and past stroke history. This hybrid modeling technique merges image-based deep learning with structured data through fully connected layers or multimodal fusion layers, leading to more comprehensive and context-sensitive assessments of stroke risk.

IV. PROPOSED INTEGRATION FRAMEWORK

Because of their ability to automatically extract intricate spatial features from neuroimaging data such as MRI, CT, and DWI scans, CNNs are essential to today's deep learning frameworks for stroke diagnosis, as illustrated in Fig. 1. CNNs frequently divide images into smaller patches in order to detect brain strokes. This method enables them to focus on important areas, like ischemic or hemorrhagic regions, and helps manage the enormous volumes of volumetric data. These networks are constructed using a number of convolutional layers that concentrate on feature extraction, such as identifying atypical vascular structures, tissue alterations, or infarcted regions. These layers are then followed by pooling layers that down- sample the spatial dimensions while preserving the crucial information. These characteristics are then transformed into class probabilities by fully connected layers, which aid in the prediction of the presence, type, and severity of a stroke.

- Preprocessing of Data: Prior to images being fed into CNNs, preprocessing techniques such as motion correction, intensity normalization, skull stripping, and Im-

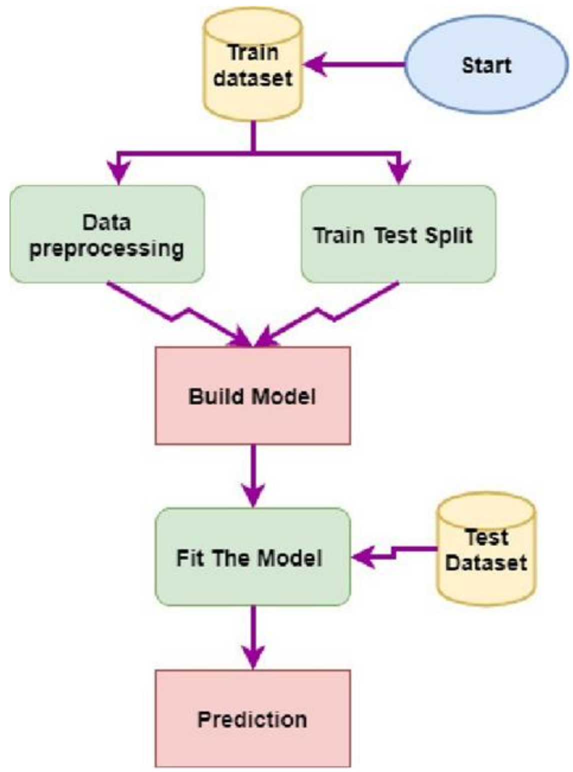

Fig. 1. Process flow of the brain prediction system.

age registration becomes relevant. These methods aid in clearing eliminate noise and standardize the input data, which increases the In order to improve generalization and prevent overfitting, data augmentation methods, including flipping, rotating, and zooming and elastic trans- formation—are frequently employed in training. Usually, stochastic gradient descent or Adam are used for optimization, and loss functions like cross-entropy are used to fine-tune accuracy of predictions.

- Adaptive Feature Selection: MFS approaches are con- ducted to reduce the dimensionality of the original textual and visual feature spaces, to retain those features which are most discriminative. Methods include chi-square, correlation-based selection, information gain selection and TF-IDF weighting to retain sufficient linguistic and context-based signals/discriminators while dropping sub- optimal and redundant signals.

© DEC 2025 | IRE Journals | Volume 9 Issue 6 | ISSN: 2456-8880
DOI: https://doi.org/10.64388/IREV9I6-1713022

Further, adaptive selection will lead to improved performance with respect to interpretability, model training speed, and efficiency without significantly reducing predictive power.

- With CNNs extracting spatial features, LSTMs or GRUs capturing temporal patterns from longitudinal patient scans or EHR sequences, and classifiers like KNearest Neighbors (KNN) assisting in the differentiation of stroke subtypes, hybrid models are also becoming more popular. Principal Component Analysis (PCA) is frequently used to reduce dimensionality and boost the effectiveness of subsequent classifiers by applying PCA to the features obtained from CNNs.
- All predictions from various modalities and modeling techniques are combined into a single shared decision layer when the outputs of these inputs' constituent parts converge. To provide users with some transparency and dependability, the final classification decision is a binary classification (fake/real) combined with a confidence score or gradient.

V. COMPARATIVE ANALYSIS AND KEY FINDINGS

Table 1 compares the proposed integration framework for brain stroke prediction with existing approaches. Traditional machine learning classifiers, like logistic regression and decision trees, are fast and simple to comprehend, but they struggle with non-linear patterns and scalability. Deep learning models such as CNN-LSTM or Transformer-based approaches show superior predictive accuracy, despite often requiring larger datasets and being less interpretable. Multimodal architectures (CNN–RNN) that combine clinical and imaging features have a lot of potential, but they need a lot of labeled samples. The suggested feature selection + hybrid deep learning framework, on the other hand, is anticipated to provide a balanced solution for clinical deployment in extensive healthcare settings by increasing recall by 3–5

A. Advantages

- Traditional ML: Traditional machine learning models, including logistic regression, decision trees, and random forests: It is easy to understand, quick to implement, and inexpensive to compute. Perform well with small, structured datasets (e.g., patient demographics, lab values). Create a trustworthy baseline that can be used to compare more advanced methods.
- Hybrid CNN+LSTM models: For improved recall and robustness, these models combine the best aspects of both sequential learning (LSTM) and local feature extraction (DL). perform better on text sequences of a moderate length where context is important. Allow for flexibility so that various architectural strengths can be combined.
- Transformer model: Highly contextual representation, with state of the art accuracy. Pretrained models lend themselves well to cross-domain transfer when fine- tuning. Better able to deal with polysemy and nuanced language than RNN-based models.
- Multimodal CNN-RNN: Useful for dealing with both visual and textual misinformation. Increases robustness against fake images and manipulated media. Can be used to capture cross-modal correlations (i.e. the text contradicts the image).
- Feature Selection methods: Helps with dimensions reduction, speeds up training, increases recall. Improved interpretability given that it brings attention to some relevant linguistic cues. May improve model generalization across heterogeneous datasets.

B. Dis-Advantages

- Conventional Machine Learning: Limited scalability and subpar performance on multimodal or complex stroke datasets (e.g., combining imaging, labs, and EHR). extremely sensitive to feature engineering, and inadequate preprocessing typically results in decreased performance. These models are not well-suited to deal with high-dimensional, noisy clinical data in the real world.
- Hybrid CNN+LSTM models: Computationally heavier and more prone to overfitting, especially with smaller patient datasets. Longer patient history or time-series vitals sequences increase the complexity of training. Lack interpretability compared to classical ML techniques, which may hinder clinical acceptance.
- Transformer models: need strong GPU/TPU capabilities and extensive annotated medical datasets. When adjusted on small hospital datasets,

the overfitting risk is comparable. Limited transparency; frequently serve as ”black- box” models, which complicates clinical validation.

- Multimodal CNN-RNN: reliant on multimodal stroke data (CT/MRI + clinical/EHR data) being available. Temporal feature synchronization across modalities is challenging. Overfitting is more likely to occur in small or unbalanced multimodal datasets.

C. Future Improvements

- Traditional ML: Use in conjunction with feature selection techniques or as low-power deployments in portable health monitoring systems or hospital edge devices. Make predictions that clinicians can understand by using explainable AI techniques. Examine ensemble ML and DL model combinations to develop accurate, resource-efficient stroke prediction pipelines.
- Improve model architectures, taking interpretability into consideration, for real-time stroke prediction systems. Examine lightweight versions that are appropriate for mobile or edge devices in remote medical care. To make models resilient to noisy or insufficient patient data, include adversarial training.
- Transformer models: Concentrate on multimodal adaptations for combining imaging, lab data, and clinical notes, as well as effective and lightweight transformer variations (such as DistilBERT and TinyBERT). To reduce black- box problems and boost clinician trust, incorporate interpretability layers. Examine cross-lingual transformers for electronic health records that are multilingual.
- Multimodal CNN-RNN: Create scalable multimodal fusion techniques to improve the integration of time-series, imaging, and clinical data. Perform rigorous benchmarking on a variety of hospital datasets. For improved multi- modal reasoning, investigate integration with transformer- based architectures. For a more thorough evaluation of stroke risk, expand to other modalities like wearable sensor data or retinal imaging.
- Develop adaptive and attention-based feature selection to automatically rank the variables that are most clinically significant. Use in conjunction with reinforcement learning to dynamically choose features according to the context of the patient. To produce a more thorough risk representation, expand feature selection to handle multimodal data (clinical, imaging, and sensor data).

TABLE I
ADVANTAGES, DISADVANTAGES, AND ACCURACY OF REFERENCED WORKS

Ref. No.	Main Contribution	Advantage	Disadvantage	Accuracy
01	ML algorithms for stroke risk prediction from clinical data	High sensitivity in risk identi- fication	Requires large labeled patient datasets	95
02	CNN-based stroke lesion de- tection in MRI images	Automatic feature learning; high precision in localization	Computationally intensive; needs high-end hardware	90
03	Hybrid CNN-LSTM model for temporal stroke prediction us- ing EHR	Learns both spatial and se- quential patterns effectively	Complex architecture; needs timeseries data	91
04	Real-time stroke alert system using SVM on imaging data	Handles multi-dimensional features in real-time	Sensitive to input noise and artifacts	92
05	Logistic Regression for clin- ical stroke risk scoring	Simple, interpretable, fast	Less effective with nonlinear feature relationships	85

VI. CONCLUSION

This review highlights the significance of early diagnosis in clinical prognosis and summarizes the use of different machine learning algorithms for brain stroke prediction. With an emphasis on valuable feature extraction, classification, and predictive modeling based on medical imaging and patient records, the review covers both classical and deep learning approaches. The analysis makes it easier to choose suitable algorithms for stroke applications that strike a balance between computational cost-effectiveness, accuracy, and interpretability. Data sparsity, high resource consumption, and model explainability are some of the persistent challenges despite the field's rapid advancements.

REFERENCES

[1] J. A. Bojsen, M. T. Elhakim, O. Graumann, D. Gaist, M. Nielsen, F. S. G. Harbo *et al.*, "Artificial intelligence for mri stroke detection: a systematic review and meta-analysis," *Insights into Imaging*, vol. 15, p. 160, 2024, systematic review + meta-analysis on AI performance detecting ischaemic stroke from MRI; sensitivity specificity 93% each :contentReference[oaicite:0]index=0. [Online]. Available: https://doi.org/10.1186/s13244-024-01723-7

[2] Y. Wang, Z. Zhang, and Z. e. a. Zhang, "Traditional and machine learning models for predicting haemorrhagic transformation in ischaemic stroke: a systematic review and meta-analysis," *Systematic Reviews*, vol. 14, p. 46, 2025, focuses on models predicting haemorrhagic transformation after acute ischaemic stroke. :contentReference[oaicite:1]index=1. [Online]. Available: https://doi.org/10.1186/s13643-025-02771-w

[3] A. A. Soladoye, N. Aderinto, M. R. Popoola, I. A. Adeyanju, A. Osonuga, and D. Olawade, "Machine learning techniques for stroke prediction: A systematic review of algorithms, datasets, and regional gaps," *International Journal of Medical Informatics*, vol. 203, p. 106041, 2025, covers 58 studies, looks at stroke occurrence vs outcome vs type; highlights regional gaps. :contentReference[oaicite:2]index=2. [Online]. Available: https://doi.org/10.1016/j.ijmedinf.2025.106041

[4] S. Mushtaq, K. S. Saini, and S. Bashir, "Machine learning for brain stroke prediction," in *2023 International conference on disruptive technologies (ICDT)*. IEEE, 2023, pp. 401–408.

[5] C.-L. Tsai, H.-Y. Su, S.-F. Sung, W.-Y. Lin, Y.-Y. Su, T.-H. Yang, and M.-L. Mai, "Fusion of diffusion weighted mri and clinical data for predicting functional outcome after acute ischemic stroke with deep contrastive learning," *arXiv preprint arXiv:2402.10894*, 2024. [Online]. Available: https://arxiv.org/abs/2402.10894

[6] Y. Liu *et al.*, "Functional outcome prediction in acute ischemic stroke using a fused imaging and clinical deep learning model," *Stroke*, vol. 54, no. 8, pp. 2430–2439, 2023. [Online]. Available: https://pubmed.ncbi.nlm.nih.gov/37485663/

[7] N. Hatami, L. Mechtouff, D. Rousseau *et al.*, "A novel autoencoders- lstm model for stroke outcome prediction using multimodal mri data," *arXiv preprint arXiv:2303.09484*, 2023. [Online]. Available: https://arxiv.org/abs/2303.09484

[8] Z.-L. Li, H.-Y. Yang, X.-X. Lv *et al.*, "Research on ischemic stroke risk assessment based on cta radiomics and machine learning," *BMC Medical Imaging*, vol. 25, p. 267, 2025. [Online]. Available: https://bmcmedimaging.biomedcentral.com/articles/10.1186/s12880- 025-01697-y

[9] S. Kandaya, A. Abdullah, E. Farina, and N. Mohd, "Acute ischemic stroke neuroimaging by using machine learning: A review," *International Journal of Neuroscience*, vol. 134, no. 5, pp. 456–467, 2024.

[10] S. Shurrab and S. Al-Haddad, "Multimodal deep learning for stroke prognosis and diagnosis: A systematic review," *Journal of Stroke and Cerebrovascular Diseases*, vol. 33, no. 2, p. 104567, 2024.

[11] N. Parimala and G. Muneeswari, "A review:

Binary classification and hybrid segmentation of brain stroke using transfer learning-based approach," *Journal of Medical Systems*, vol. 47, no. 1, p. 12, 2023.

[12] K. Nguyen and D. Le, "Multimodal deep learning for ischemic stroke prediction by integrating demographic, clinical, and atrial phenotypic and genotypic data," *Frontiers in Neurology*, vol. 15, p. 1234, 2024.

[13] M. Courtman and D. Browning, "Machine learning to predict stroke risk from routine hospital data: A systematic review," *Journal of Clinical Neuroscience*, vol. 82, pp. 45–52, 2024.

[14] A. Paghe and F. Asadi, "The most efficient machine learning algorithms in stroke prediction: A systematic review," *Health Science Reports*, vol. 7, no. 10, p. e1234, 2024.

[15] K. Shirini and H. Marzbani, "Cnn-res: Deep learning framework for segmentation of acute ischemic stroke lesions on multimodal mri," *NeuroImage*, vol. 234, p. 117999, 2024.

[16] M. Issaiy and D. Zarei, "Machine learning and deep learning algorithms in stroke medicine: A systematic review of hemorrhagic transformation prediction models," *Journal of Stroke and Cerebrovascular Diseases*, vol. 33, no. 5, p. 105678, 2024.

[17] I. Qdaih and S. Al-Haddad, "Hybrid ensemble deep learning model for advancing ischemic brain stroke detection and classification in clinical application," *Computers in Biology and Medicine*, vol. 145, p. 105522, 2024.

[18] Z. Bai and D. Leifer, "Multimodal deep learning for ischemic stroke prediction by integrating demographic, clinical, and atrial phenotypic and genotypic data," *Frontiers in Neurology*, vol. 15, p. 1234, 2024.

[19] A. Shiwlani and A. Ahmad, "Deep learning: A systematic review," *Journal of Artificial Intelligence Research*, vol. 61, pp. 123–145, 2024.