Preprints.org

Review

Machine Learning in Stroke Lesion Segmentation and Recovery Forecasting: A Review

Simi M S [*] , Sibusiso Mdletshe [*] , Alan Wang

Article

Machine Learning in Stroke Lesion Segmentation and Recovery Forecasting: A Review

Simi M S [1] , **Sibusiso Mdletshe** [1] and **Alan Wang** [1,2,3,4,5,6]

[1]   Department of Anatomy and Medical Imaging, FMHS, University of Auckland, Auckland, New Zealand.
[2]   Auckland Bioengineering Institute, University of Auckland, Auckland, New Zealand.
[3]   Centre for Brain Research, University of Auckland, Auckland, New Zealand.
[4]   Matai Medical Research Institute, Gisborne, New Zealand.
[5]   Medical Imaging Research Centre, University of Auckland, Auckland, New Zealand.
[6]   Centre for Co-Created Ageing Research, University of Auckland, Auckland, New Zealand.
*   Correspondence: simi.m.s@auckland.ac.nz

Abstract

Introduction: Stroke is a leading cause of disability worldwide, and precise identification of stroke lesions is essential for prognosis and rehabilitation planning. Machine learning has emerged as a powerful tool for automating stroke lesion segmentation and predicting post-stroke recovery, though challenges in standardization and generalizability remain. **Methods:** This narrative review synthesizes studies published between 2010 and 2024 on the application of machine learning in stroke lesion segmentation and recovery forecasting. A total of 23 relevant studies were reviewed, including 10 focused on lesion segmentation and 13 on recovery prediction. **Results:** Deep learning models such as U-Net and convolutional neural networks(CNNs) have shown improved segmentation accuracy using the Anatomical Tracings of Lesions After Stroke (ATLAS) dataset, though dataset bias and inconsistent evaluation metrics limit comparability. Recovery prediction studies demonstrated the use of lesion characteristics alongside clinical and demographic features to forecast motor, cognitive, and functional outcomes. However, issues such as small sample sizes, lack of multi-centre validation, and heterogeneous modelling approaches were noted. **Discussion:** The review highlights the interdependence between segmentation and prediction—accurate lesion delineation informs recovery modelling, while recovery prediction may enhance the clinical relevance of segmentation. Integration of both domains is critical to advancing stroke care. **Conclusion:** Machine learning offers promising advances in stroke lesion analysis and recovery forecasting. Future efforts should focus on standardized datasets, multimodal data integration, and robust validation to translate research into clinical practice and support precision rehabilitation.

Keywords: stroke; segmentation; prediction; ATLAS; MRI; machine learning

1. Introduction

A stroke is a medical emergency marked by discontinuing blood flow to the brain, leading to the deficiency of oxygen and nutrition for brain cells [1]. Stroke is a prominent causality of disability worldwide, and precise identification and quantification of stroke lesions is pivotal for diagnosis, prognosis, and treatment planning. There are two major divisions of stroke: ischemic stroke and haemorrhagic stroke [2]. Each has distinct characteristics and causes. Ischemic stroke, constituting almost 80% of all stroke instances, transpires whilst a blockage or plaque accumulation obstructs or constricts a blood artery, resulting in dwindled or overall interruption of blood delivery to some aspects of the brain [3].

Imaging modalities, such as Computed Tomography(CT) and Magnetic Resonance Imaging(MRI), play a pivotal role in identifying and characterising stroke lesions. The development of functional imaging techniques, like perfusion-CT and perfusion-MRI, has allowed for more implications of

stroke's dynamic and evolving nature, shifting the paradigm to a more complex and multifaceted one [4].

Faultless identification and quantification of lesions are vital for treatment planning. Automatic segmentation of stroke lesions from medical imaging data, such as MRI and Computed Tomography Perfusion scans, is challenging due to the variability in lesion size, shape, location, intensity variation, difficulty in capturing complex lesion boundaries precisely and underlying cerebrovascular dynamics [5]. These make it difficult to distinguish the targeted lesions from the surrounding brain tissue, as highlighted in the Figure 1. These drawbacks highlight the necessity of reliable, broadly applicable, and clinically interpretable segmentation methods for stroke imaging.
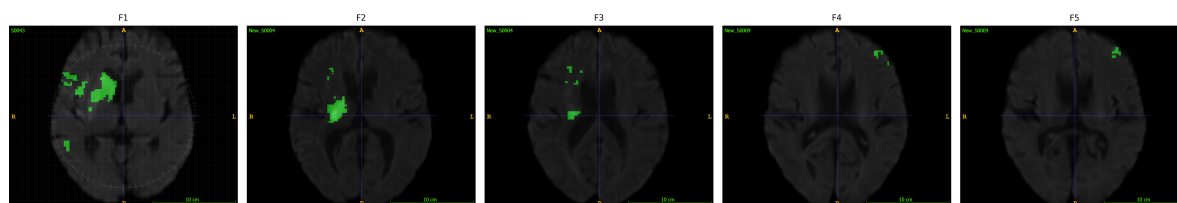

Figure 1. Challenges in stroke lesion segmentation depicted through MRI-DWI examples: F1 demonstrate the complex shapes, and varied textures of lesions; F2 and F3 illustrate low resolution and poor contrast; F4 and F5 highlight the difficulty in differentiating lesions from surrounding tissue .

Machine learning techniques have proven their potential to forecast stroke consequences and predict recovery, making use of lesion facts derived from MRI images [6]. Voxel-based approaches have been widely employed, leveraging global and local contextual information to improve segmentation accuracy [7].

The studies utilise modern imaging modalities such as MRI to get specific information on the stroke lesion's location, size, and other attributes. The dissimilarity in the presence and traits of stroke lesions amongst patients, even in the same clinical stage, is a significant challenge to address the researchers. Researchers have explored several machine-learning techniques to classify and separate stroke lesions, including support vector machine (SVM), random forests, and deep learning models [8,9]. To address the challenge posed by the variability in lesion characteristics, researchers have investigated the use of more sophisticated techniques, such as pyramid pooling and focal loss [7], which aim to capture both global and local contextual information, as well as focus on the more difficult samples during the training process [7].

The science community initiated numerous challenges to support this necessity, ATLAS R2.0 [10] - Stroke Lesion Segmentation, Ischemic Stroke Lesion Segmentation Challenge - ISLES'22, Ischemic Stroke Lesion Segmentation Challenge - ISLES'24. All the dataset challenges have become the de-facto evaluation standard for new algorithms, in particular when adhering to some standards listed on the same web resource: Both training and testing datasets are representative of the task, well described, and large enough to draw significant conclusions from the results; the associated ground truth is created by experts following a clearly defined set of rules; the evaluation metrics were chosen to capture all aspects relevant for the task; and, ideally, challenges remain open for future contestants and serve as an ongoing benchmark for algorithms in the field.

This review aims to conduct a comprehensive analysis and critical appraisal of the existing literature on the utilisation of machine learning on stroke lesion segmentation and recovery forecasting.

2. Methods

This study assessed the current research on machine learning and deep learning methods for segmenting brain stroke lesions and predicting recovery after a stroke. This process involved systematically reviewing and synthesising studies from PubMed, Scopus and Web of Science, focusing on identifying trends, challenges, and future directions for research in these areas.

2.1. Study Search Strategy

This review includes a comprehensive search of various primary research studies from 2010 to September 2024, excluding reviews and surveys from the search. Search term encompassed the following search strings (stroke OR ischemic stroke OR hemorrhagic stroke) AND (lesion* OR scar*) AND (segmentation), (post-stroke OR "post stroke" OR "stroke recovery" OR "stroke rehabilitation") AND (prediction OR predictor* OR forecast* OR prognos*) AND (recovery OR outcome* OR improvement OR rehabilitation OR "functional recovery"), (stroke OR "post-stroke" OR "ischemic stroke" OR "hemorrhagic stroke") AND (recovery OR rehabilitation OR outcome* OR improvement) AND (prediction OR predictor* OR forecast* OR prognos*) AND ("U-Net" OR "U Net" OR "UNet" OR "convolutional neural network" OR "deep learning"). A total of 23 articles were included for analysis in this review; the screening flow chart (Figure 2) details the review process.
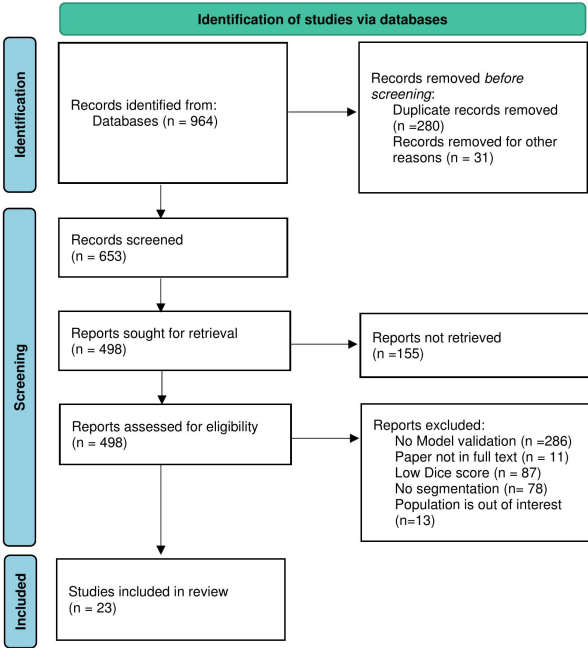

Figure 2. Flow diagram of the literature screening and selection process.

3. Results

This review focuses on two key domains in stroke research: stroke lesion segmentation using the ATLAS dataset and post-stroke recovery prediction through machine learning techniques. Studies that utilised the ATLAS dataset for segmentation tasks were selected to ensure consistency and comparability, as this dataset serves as a benchmark in the field. For post-stroke recovery prediction, the review concentrated on research employing machine learning methods to predict functional outcomes, such as motor recovery, language restoration, and cognitive performance. The review explores the intersection of lesion segmentation and post-stoke recovery prediction by combining this two research and providing insights into how advancements in one domain could inform progress in the other.

3.1. Machine Learning-Driven Stroke Lesion Segmentation with ATLAS Dataset

The key findings and characteristics of ten studies in stroke segmentation using deep learning with ATLAS as the dataset are outlined; seven utilised ATLAS V2.0, while three relied on ATLAS V1. These studies were published between 2020 and 2023.

To take advantage of MRI in stroke lesion segmentation, several deep-learning models with MR images have been proposed. Here, seven studies presented with ATLAS V2.0 dataset, in all of which the preprocessing steps are intensity normalisation, skull stripping, slicing, and cropping [11–14]. One study has a variation in preprocessing; they use Gaussian denoising as a preprocessing step, resulting

in a higher Dice score among all seven studies (94.2%) [15]. Another study used matrix complement as a preprocessing step and was able to achieve a dice coefficient of 69.72 [16].

During the literature gathering, it was observed that many of the researchers adopted the U-Net model as the segmentation technique. Some focused on improving the structure, while others focused on using them as the base of their method. For example, [14], the authors demanded that they apply 3D-Unet, the first hybrid contextual semantic network with k-fold cross-validation and data augmentation for regularised overfitting. [11] focused on standard U-Net variants for both 2D and 3D models, achieved the highest Dice score of 0.583 and 0.504 on the 2D transformer-based model and 3D residual U-Net, respectively.

Similarly, the work proposed by [12] used the nnU-Net framework for lesion segmentation. This method achieved first place on the unseen test data in the 2022 MICCAI (Medical Image Computing and Computer Assisted Intervention) ATLAS Challenge with an average Dice score of 0.6667. They also tried to get an ensemble prediction for better results with effective preprocessing. [15] presented Fuzzy Information Seeded Region Growing (FISRG), integrating fuzzy logic with seeded region growing for enhanced Segmentation accuracy (94.2%). The study employed a final post-processed image where morphological operations refined the segmentation results. [16] Provided a hybrid contextual semantic module (HCSM) in the skip connection layer and residual blocks in the encoder/decoder layers by extending the U-shaped architecture. They argued that it can accurately segment and detect small-size stroke lesions from magnetic resonance images.

[17] They experimented with a simulated quantum mechanics-based joint learning network (SQMLP-net) that simultaneously segments lesions and examines thrombolysis in cerebral infarction (TICI) grade. They concluded that the accuracy of stroke lesion segmentation negatively correlated with the severity of thrombolysis in cerebral infarction. [13] Employed an extensive study of eight variant 2D model architectures, they implemented an ensemble method involving stacking and agreement window for the final enhanced prediction. They used an in-house dataset for training and testing purposes along with ATLAS; they adopted a mixed data approach and intermediate task training approach; in a mixed data approach, VNet outperformed all other 2D models, and in intermediate task training, Fully Connected Densenet outperforms other models. They concluded that the agreement window method super-region-wise generated the best lesion volume.

[18] Proposed an enhanced multi-scale long-range interactive and regional attention network (MLiRA-Net) method, integrating multi-scale hierarchical and local-global features for better segmentation accuracy. They improved the performance of MLiRA-Net by adding skip connections between the output layers of the encoder and the input layers of the decoder. [19] Extended the 3D UNet architecture to accommodate volumetric segmentation of chronic stroke lesions on T1-weighted MRI scans more effectively and accurately to assist neuroradiologists in this task and provide personalised rehabilitation to achieve adequate recovery.

3.2. Post-Stroke Recovery Prediction through Machine Learning Techniques

This section explores studies that apply machine-learning techniques to predict recovery outcomes in stroke patients. These studies have aimed at improving the prediction of recovery trajectories and optimising personalised rehabilitation strategies by leveraging multiple levels of data, including clinical, demographical, neuroimaging, and functional assessment.

Corticospinal tract injury observed in acute stroke imaging has been identified as a key predictor of upper extremity motor recovery, with its predictive accuracy unaffected by the extent of damage to the primary motor and premotor cortices [20]. Factors such as age, sex, obesity, education level, stroke location, stroke severity, comorbidities, smoking and alcohol history, medical complications, and functional assessments may influence the ability to predict functional outcomes such as motor function, mobility, cognition, language, swallowing, and activities of daily living in first-time stroke patients [21].

The Motor-Evoked Potentials (MEP) response in Transcranial Magnetic Stimulation (TMS) and preserved corticospinal tracts (CST) integrity in Diffusion Tensor Imaging (DTI) have been identified

as strong predictors of upper-limb (UL) motor recovery in ischemic stroke patients [22]. A high score on the Fugl-Meyer Assessment (FMA) and Motricity Index (MI) has the potential to refine the outcome of the prediction model further.

[23] has investigated the patterns of recovery from aphasia within the first year after a stroke, focusing on how lesion location and extent influence recovery outcomes. Initial evaluation of speech and language functions were assessed using the Quick Aphasia Battery (QAB) within 5 days of the stroke served as the dependent variable, and lesions were manually delineated based on MRI or CT imaging for the study served as the independent variable of the model. Even though the recovery pattern varies significantly across different speech and language domains, patients with circumscribed frontal lesions recovered well, and those with extensive damage in the middle cerebral artery distribution or temporoparietal regions had a constant shortfall.

Regression was the most common method to develop models among the included studies. Specifically, Multivariable logistic regression analysis was used in four studies [22,24–26] , three studies used linear and ridge and Elastic net regression models [20,27,28]. According to [29], the integrity of the cerebellar tract was identified as the primary biomarker predicting the recovery of upper extremity motor function over three months. The study examined biomarkers for good and poor recovery. They implemented Fractional Anisotropy Analysis from Diffusion Tensor Imaging, Functional Connectivity Analysis from Resting-State functionalMRI, and Statistical Analysis to extract the biomarkers.

Table 1. Summary of ML Approaches in Stroke Lesion Segmentation Using the ATLAS Dataset.

Ref	Preprocessing	Dataset	Segmentation Method	Loss function	Performance Metric	Gap/Limitations
[14]	Intensity normalisation, registration and defacing, brain extraction, skull removing	ATLAS V2.0	Deep Neural Network (DNN) using 3D-UNet with 5-fold cross-validation	NM	DSC =0.65	Although potential biases from the manual segmentation process could influence the model, and the reliance on a subset for testing may limit generalizability, the scarcity of publicly accessible stroke datasets with manual segmentation labels makes independent validation.
[11]	Z-score Norma, Slicing for 2D modality.	ATLAS V2.0	Multiple U Net variants	Proportionate weightage for Dice Loss (DICE) and Binary Cross-Entropy Loss (BCE)	Dice: 0.583	There's a need for better data augmentation, U-Net models, supervised learning, handling small lesions, and decision-making uncertainty.
[12]	Normalisation and Registration	ATLAS V2.0	nnU-Net	compound loss (Dice plus cross-entropy), TopK10 loss	Dice: 0.6667	Small stroke lesions are hard to segment, especially with artifacts or similar intensities to surrounding tissue. Training schemes often predict unconnected lesions as continuous grey matter due to similar intensities.
[15]	Gaussian denoising	ATLAS V2.0	Fuzzy Information Seeded Region Growing (FISRG) algorithm	DiNM	Dice: 0.94	The algorithm struggles with abrupt lesion topology changes, misclassifies regions with similar intensities, and has increased computational time. Intensity-based classification causes errors, especially with variable lesion textures.
[18]	Image Slicing, Cropping, Patch Partitioning, Patch Embedding, Normalization and Augmentation	ATLAS v1	MLiRA-Net (Multi-scale Long-range Interactive and Regional Attention Network)	dice loss + weighted binary cross-entropy loss.	Dice: 0.6119	It comes at the cost of increased computational complexity. Additionally, the current implementation is limited to two-dimensional segmentation.
[16]	Matrix complement and clipping method	ATLAS V2.0	Hybrid Contextual Semantic Network (HCSNet)	Mixing-Loss Function	Dice: 0.69	Segmenting small lesions is challenging due to the heavy reliance on training data quality and quantity. Additionally, the model's ability to generalize to real-world clinical settings needs further validation.
[17]	Intensity Correction, MNI-152 template registration	ATLAS V2.0	Simulated Quantum Mechanics-based joint Learning Network (SQMLP-net)	The joint loss function incorporates the segmentation and classification losses	Dice: 0.7098	Finding the right balance between the trade-offs of multi-task learning weights is crucial for optimising task performance.
[13]	Resampling and Normalisation, Skull Stripping, Slicing and Augmentation	ATLAS V2.0	Transfer Learning and mixed data approaches.	NM	Dice: 0.736	The ensemble methods' accuracy may change by the chosen parameters, which may require further adjustment for different datasets or lesion types. The ensemble method tends to overpredict lesions by approximately 10%.
[19]	Normalisation, Cropping, Zoom-in &out Training Strategy	ATLAS v1	3D U-Net architecture with residual learning	Binary cross entropy (BCE)+Dice Loss	Dice: 0.64	The small dataset limits generalizability, and the model struggles with smaller lesions. Further validation with diverse scanners is needed. Manual tracing introduces variability, and the dataset only focuses on embolic strokes, leaving other types untested.
[30]	Slice Selection, Patch Extraction, Data Augmentation	ATLAS v1	U-Net architecture with Xception as the backbone- XU-Net	BCE & Jaccard Coefficient	Dice: 0.754	The study struggles with accurately segmenting small stroke lesions and reducing false positives. The model's generalizability requires further validation on diverse datasets. Additionally, the approach increases computational complexity, limiting real-time application.

The three studies used Multivariable Logistic Regression analysis with clinical and demographic data to predict post-stroke outcomes: motor impairment, physical functioning, and upper limb functioning [24–26]. They measured performance with the Area Under the Curve (AUC), showing values of 0.833, 0.883, and 0.86. The highest AUC (0.883) was achieved for physical functioning, using data from 717 patients and building models for improved functioning and independence. Variations in AUC reflect differences in data richness. In [26], the logistic model was converted to an integer score using a regression coefficient-based scoring method called the Scoring Rule; the scoring rule helps identify patients at treatment failure risk, aiding in admission decision-making. [24] highlighted Motor-evoked potentials (MEP) status and neglect are crucial for predicting non-normal UL use.

The study by [31] utilised the SVM approach to identify significant inflammatory biomarkers for predicting rehabilitation outcomes in stroke patients. Feature selections were made using the Kernel Density Estimation and Mutual Information (MI) Criterion. They also concluded that the type of stroke is not a vital variable for predicting the discharge data. Some multi-modal Convolutional Neural Networks (CNN) combine MRI data with demographic and clinical characteristics to improve prediction accuracy [32]. The study reveals that left hemisphere lesion size was least important when damage to critical anatomical regions of interest was incorporated, and some of the 2D models outperformed 3D CNNs in terms of mean accuracy. [33] Analysed the key factors influencing independent walking prediction in poststroke patients. Prediction performed by LR and three machine learning (ML) models: eXtreme Gradient Boosting (XGBoost), Support Vector Machines (SVM), and Random Forest (RF); the XGBoost model showed the best overall performance with age, lower limb spasticity, Functional Ambulation Category (FAC) at admission, and Fugl-Meyer Motor Assessment of the Lower Extremity (FMA-LE) as the key factors influencing independent walking.

In this study, we identified recent articles about brain stroke lesion segmentation with machine learning using MRI (70%), CT (10%) images and demographic data (30%). Some papers utilise several modalities in a paper, so the sum of the percentages may be greater than 100%. These papers used Dice score (43%), Accuracy rate (17%), Coefficient of Determination (13%), Area Under the Curve (22%), and correlation (4%) metrics as performance evaluation.

The comparative analysis of different dice scores of stroke lesion segmentation studies is presented in Figure 3. All the methods were compared using the same dataset, i.e., ATLAS. In contrast, the distribution of the dataset into a training set, validation set, and test set was different in different methods. Mean Dice coefficients range from 0.58 to 0.94. Researchers used U-Net (30%), CNN (9%), fuzzy logic (4%), Regression (40%) and deep learning architectures and 17% of them used another model. 74% of papers have dataset more than 200 image scans and 26% have dataset lower than 60 image scans.
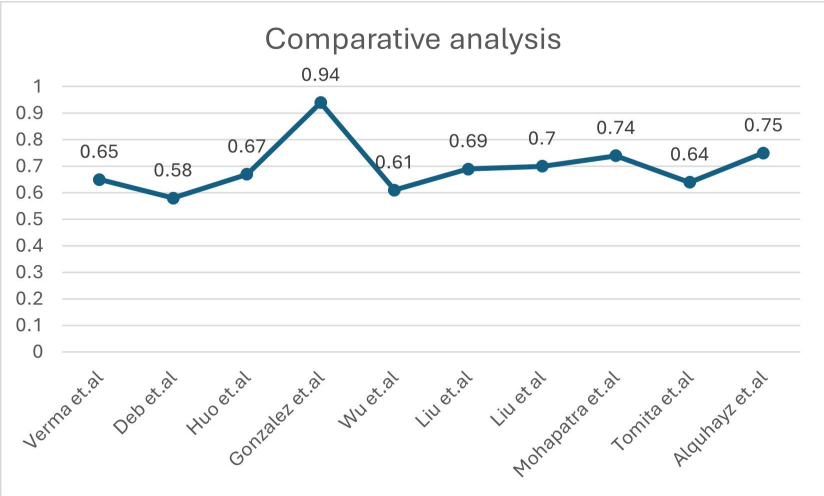

Figure 3. Dice score comparison of stroke lesion segmentation methods on the ATLAS dataset.

Table 2. Study characteristics, model methodology, and predictive performance of the included studies using Other Dataset

Ref	Study Design	Sample Size	Data Modality	Method	Focus Area	Performance Metrics	Gap/Limitations
[27]	Longitudinal design	57 patients	Resting-State Functional MRI, Behavioral Language Measures	Elastic net regression models	Language Recovery	R2 = 0.948	Small sample size & lack of validation set. Use of LOOCV (Leave-One-Out Cross-Validation)-A less robust and reliable model selection method. Recruitment constraints & Minimal fMRI filteringresults in lower accuracy and generalizability
[28]	Longitudinal design	37 patients	T2 and diffusion-weighted MRI, Fugl-Meyer Assessment (FMA) scores	Ridge regression	Motor Recovery	R2 = 0.68	Small sample size & no age-matching, Manual lesion masking, Larger, diverse cohorts required to confirm findings.
[20]	Longitudinal design	48 patients	magnetic resonance diffusion-weighted images & CT (one case)	Logistic and linear regression models	upper extremity motor recovery	AUC=ranging from 0.70 to 0.8	The study sample was younger, predominantly male, and ethnically homogeneous compared to national averages, with a small sample size of 48 participants.
[21]	Longitudinal design	7,858 patients	Demographic and clinical characteristicss	K-means clustering	Functional recovery	ischemic-: 0.926 hemorrhagic-: 0.887	Survivor bias, Design limitations- Absence of functional MRI or dynamic nomograms
[22]	prospective cohort study	29 patients	Transcranial magnetic stimulation or Diffusion tensor imaging parameters or their combinations	Multivariate logistic regression analysis	The upper-limb motor function	predictive ability=93.3%	Rigorous inclusion criteria: Only 29 patients enrolled, limiting generalizability Exclusions: Participants with medical issues were excluded, potentially skewing the results.
[23]	Longitudinal design	334 patients	Demographic factors, MRI or CT imaging	linear model	Language Recovery	QAB overall = 59.5% of the variance	Smaller group sizes and missing data points, Lesions identified through acute clinical imaging may not accurately reflect irreversible tissue damage, QAB limitation: Doesn't assess written word comprehension or writing.
[29]	retrospective cohort study	104, 42 patients	Structural, diffusion, and functional magnetic resonance imaging	multiple linear regression algorithm	Upper Extremity Motor Recovery	R² = 0.853	Small sample size, Single-center study.
[26]	Retrospective cohort study	1265 patients	Data from electronic records	multivariable logistic regression model	motor impairment	AUC=0.833	Retrospective design: May introduce biases, No control group, Dynamic motor changes: Study does not account for rehabilitation-related variations.
[25]	Retrospective cohort study	717 patients	Clinical and Demographic data	multivariable logistic regression model	Physical functioning	AUC=0.883	Retrospective design: Introduces biases, Exclusion of mild stroke patients, and did not assess the prognostic role of neuroimaging.
[24]	Observational prospective cohort design	87 patients	Clinical, Demographic, and Statistical data, Accelerometer Data	multivariate regression model	upper-limb	AUC=0.86	Limited accelerometer use, and visible devices could have led to overestimation.
[31]	Longitudinal cohort design	55 patients	Demographic, Clinical, Biochemical and Hematological Parameters, Health Status Data at Discharge	SVM	Motor and cognitive improvement	Correlation ranged from 0.75 to 0.81	Single hospital data, reliance on common inflammatory biomarkers, which can vary between individuals, and short follow-up period.
[32]	Cross-sectional design	758 patients	T1-weighted structural MRI, demographic and clinical characteristics	CNN	Aphasia	Classification accuracy = 0.854	Restricted to research-quality MRI scanners, and the measure of initial severity is relatively crude. Overfitting risk, and simplifies the Comprehensive Aphasia Test T-scores.
[33]	Retrospective cohort study	778 patients	Demographic, Clinical Information, Functional Scores at Admission.	Machine learning models	Walking independence.	LR:AUC: 0.891, XGBoost:AUC: 0.880, SVM:AUC: 0.659, RF:AUC: 0.713	Retrospective, single-center design, No external validation, Only clinical admission data were used, No long-term outcomes.

4. Discussion

Machine Learning plays a vital role in medical image segmentation, facilitating accurate delineation of areas of interest. Machine learning offers novel approaches to address complex medical challenges. The ability to accurately predict post-stroke recovery by applying ML in quantifiable image features will improve patient outcomes, as it allows healthcare providers to tailor interventions based on individual patient characteristics and projected recovery trajectories [34].

The spectrum of association between different medical parameters and post-stroke recovery still needs to be fully understood. It could shed more light on that by using more precise techniques on medical imaging and clinical data. There is a significant association between behavioural measures and the prediction of Anomia recovery, and resting state functional MRI (rsfMRI) features like fALFF (fractional Amplitude of Low-Frequency Fluctuations) helped predict agrammatism and dysgraphia [27]. They took 27 behavioural measures comprising 11 languages and cognitive assessments. [28] Combined the Fugl-Meyer Assessment (FMA) scores with MRI-based brain connectivity measures to predict motor recovery in stroke patients. Brain Connectivity Measures analysis was done by introducing virtual lesions into healthy connectomes; the study estimates the impact of stroke on the brain's network. Measures like modularity, participation coefficient, and eigenvector centrality are derived from these connectivity analyses.

We suggest that the stroke lesion segmentation researchers share their algorithms' source code, data, and model, enabling other researchers to gain insights from similar studies and provoke a new wave of lesion segmentation advancements. In our included studies, less than 50% of researchers published their source code. Some of them needed to be given access to the data. Machine learning-based methods, such as 3D CNNs and U-Net networks, have shown high accuracy and excellent soft tissue contrast in MRI and CT image segmentation tasks. However, they can be computationally intensive and require extensive training data. Other methods have also shown potential for clinical applications but may need to be more accurate than deep learning-based approaches. Table 1 demonstrates that the fuzzy logic combined with the Seeded Region Growing algorithm is the most accurate for stroke lesion segmentation. In contrast, Table 2 depicts regression models as the best method for accurately segmenting stroke lesions.

5. Conclusions

This study has explored the latest evolutions in brain stroke lesion segmentation using machine learning. The utilisation of machine learning, particularly U-net-based architecture, has shown great assurance in segmenting stroke lesions precisely from different modalities, such as MRI and CT. This adaptability makes them applicable in diverse clinical settings and enhances their potential for widespread adoption in routine clinical practice. Among the advantages, this kind of research should be promoted with more data-driven approaches in clinical practices to reduce the gap between clinical and stroke research areas. For this reason, we suggest additional research with an expanded dataset and external validation.

Abbreviations

ATLAS	Anatomical Tracings of Lesions After Stroke
AUC	Area Under the Curve
BCE	Binary Cross-Entropy Loss
CNN	Convolutional Neural Network
CST	Corticospinal Tracts
CT	Computed Tomography
DICE	Dice Loss
DNN	Deep Neural Network
DTI	Diffusion Tensor Imaging
eXGBoost	eXtreme Gradient Boosting
FAC	Functional Ambulation Category
fALFF	fractional Amplitude of Low-Frequency Fluctuations
FISRG	Fuzzy Information Seeded Region Growing
FMA	Fugl-Meyer Assessment
FMA-LE	Fugl-Meyer Motor Assessment of the Lower Extremity
HCSM	Hybrid Contextual Semantic Module
HCSNet	Hybrid Contextual Semantic Network
ISLES'22	Ischemic Stroke Lesion Segmentation Challenge 2022
ISLES'24	Ischemic Stroke Lesion Segmentation Challenge 2024
LR	Linear Regression
MEP	Motor-Evoked Potentials
MICCAI	Medical Image Computing and Computer-Assisted Intervention
MI	Mutual Information
ML	Machine Learning
MLiRA-Net	Multi-scale Long-range Interactive and Regional Attention Network
MRI	Magnetic Resonance Imaging
QAB	Quick Aphasia Battery

RF	Random Forest
rsfMRI	Resting-State Functional Magnetic Resonance Imaging
SQMLP-net	Simulated Quantum Mechanics-based Joint Learning Network
SVM	Support Vector Machine
TICI	Thrombolysis in Cerebral Infarction
TMS	Transcranial Magnetic Stimulation
UL	Upper Limb
U-Net	U-shaped Convolutional Neural Network
VNet	V-shaped Convolutional Neural Network

Author Contributions: Mrs. M S conceptualized the study, conducted a literature search, synthesized the findings, and wrote the manuscript. Dr.Mdletshe and Dr. Wang contributed to the supervision, review, and editing of the manuscript. All authors read and approved the final manuscript.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: All data generated or analysed during this study are included in this published article.

Conflicts of Interest: The authors declare that they have no competing interests.

References

1. Mackay, J.; Mensah, G.A. *The Atlas of Heart Disease and Stroke*; World Health Organization, 2004.
2. Ling, X.; Zheng, Y.; Tao, J.; Zheng, Z.; Chen, L. Association study of polymorphisms in the ABO gene with ischemic stroke in the Chinese population. *BMC Neurology* **2016**, *16*, 146.
3. Benjamin, E.J.; Muntner, P.; Alonso, A.; Bittencourt, M.S.; Callaway, C.W.; Carson, A.P.; et al.. Heart Disease and Stroke Statistics—2019 Update: A Report From the American Heart Association. *Circulation* **2019**, *139*, e56–e528.
4. Ahuja, C.K.; Gupta, V.; Khandelwal, N. Acute Stroke Imaging: Current Trends. *Annals of the National Academy of Medical Sciences (India)* **2019**, *55*, 193–201.
5. Maier, O.; Schröder, C.; Forkert, N.D.; Martinetz, T.; Handels, H. Classifiers for Ischemic Stroke Lesion Segmentation: A Comparison Study. *PLOS ONE* **2015**, *10*, e0145118.
6. Hope, T.M.H.; Seghier, M.L.; Leff, A.P.; Price, C.J. Predicting outcome and recovery after stroke with lesions extracted from MRI images. *NeuroImage: Clinical* **2013**, *2*, 424–433.
7. Abulnaga, S.M.; Rubin, J. Ischemic Stroke Lesion Segmentation in CT Perfusion Scans using Pyramid Pooling and Focal Loss. arXiv, 2018. Version 1.
8. Halme, H.L.; Korvenoja, A.; Salli, E. ISLES (SISS) Challenge 2015: Segmentation of Stroke Lesions Using Spatial Normalization, Random Forest Classification and Contextual Clustering. In *Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries*; Crimi, A.; Menze, B.; Maier, O.; Reyes, M.; Handels, H., Eds.; Springer International Publishing: Cham, 2016; Vol. 9556, *Lecture Notes in Computer Science*, pp. 211–221.
9. Kamnitsas, K.; Ledig, C.; Newcombe, V.F.J.; Simpson, J.P.; Kane, A.D.; Menon, D.K.; et al.. Efficient multi-scale 3D CNN with fully connected CRF for accurate brain lesion segmentation. *Medical Image Analysis* **2017**, *36*, 61–78.
10. Liew, S.L.; Lo, B.P.; Donnelly, M.R.; Zavaliangos-Petropulu, A.; Jeong, J.N.; Barisano, G.; et al.. A large, curated, open-source stroke neuroimaging dataset to improve lesion segmentation algorithms. *Scientific Data* **2022**, *9*, 320.
11. Deb, P.; Baru, L.B.; Dadi, K.; S, B.R. BeSt-LeS: Benchmarking Stroke Lesion Segmentation using Deep Supervision. arXiv, 2023. Version 1.

12. Huo, J.; Chen, L.; Liu, Y.; Boels, M.; Granados, A.; Ourselin, S.; et al.. MAPPING: Model Average with Post-processing for Stroke Lesion Segmentation. arXiv, 2022. Version 1.

13. Mohapatra, S.; Gosai, A.; Shinde, A.; Rutkovskii, A.; Nouduri, S.; Schlaug, G. Meta-Analysis of Transfer Learning for Segmentation of Brain Lesions. arXiv, 2023, [2306.11714].

14. Verma, K.; Kumar, S.; Paydarfar, D. Automatic Segmentation and Quantitative Assessment of Stroke Lesions on MR Images. *Diagnostics* **2022**, *12*, 2055.

15. González, M.P. Fuzzy Information Seeded Region Growing for Automated Lesions After Stroke Segmentation in MR Brain Images. arXiv, 2023. Version 1.

16. Liu, L.; Chang, J.; Liu, Z.; Zhang, P.; Xu, X.; Shang, H. Hybrid Contextual Semantic Network for Accurate Segmentation and Detection of Small-Size Stroke Lesions From MRI. *IEEE Journal of Biomedical and Health Informatics* **2023**, *27*, 4062–4073.

17. Liu, L.; Chang, J.; Liang, G.; Xiong, S. Simulated Quantum Mechanics-Based Joint Learning Network for Stroke Lesion Segmentation and TICI Grading. *IEEE Journal of Biomedical and Health Informatics* **2023**, *27*, 3372–3383.

18. Wu, Z.; Zhang, X.; Li, F.; Wang, S.; Huang, L. Multi-scale long-range interactive and regional attention network for stroke lesion segmentation. *Computers and Electrical Engineering* **2022**, *103*, 108345.

19. Tomita, N.; Jiang, S.; Maeder, M.E.; Hassanpour, S. Automatic post-stroke lesion segmentation on MR images using 3D residual convolutional neural network. *NeuroImage: Clinical* **2020**, *27*, 102276.

20. Lin, D.J.; Cloutier, A.M.; Erler, K.S.; Cassidy, J.M.; Snider, S.B.; Ranford, J.; et al.. Corticospinal Tract Injury Estimated From Acute Stroke Imaging Predicts Upper Extremity Motor Recovery After Stroke. *Stroke* **2019**, *50*, 3569–3577.

21. Shin, S.; Chang, W.H.; Kim, D.Y.; Lee, J.; Sohn, M.K.; Song, M.K.; et al.. Clustering and prediction of long-term functional recovery patterns in first-time stroke patients. *Frontiers in Neurology* **2023**, *14*, 1130236.

22. Kumar, P.; Prasad, M.; Das, A.; Vibha, D.; Garg, A.; Goyal, V.; et al.. Utility of transcranial magnetic stimulation and diffusion tensor imaging for prediction of upper-limb motor recovery in acute ischemic stroke patients. *Annals of Indian Academy of Neurology* **2022**, *25*, 54.

23. Wilson, S.M.; Entrup, J.L.; Schneck, S.M.; Onuscheck, C.F.; Levy, D.F.; Rahman, M.; et al.. Recovery from aphasia in the first year after stroke. *Brain* **2023**, *146*, 1021–1039.

24. Lundquist, C.B.; Nielsen, J.F.; Brunner, I.C. Prediction of Upper Limb use Three Months after Stroke: A Prospective Longitudinal Study. *Journal of Stroke and Cerebrovascular Diseases* **2021**, *30*, 106025.

25. Scrutinio, D.; Lanzillo, B.; Guida, P.; Mastropasqua, F.; Monitillo, V.; Pusineri, M.; et al.. Development and Validation of a Predictive Model for Functional Outcome After Stroke Rehabilitation: The Maugeri Model. *Stroke* **2017**, *48*, 3308–3315.

26. Scrutinio, D.; Guida, P.; Lanzillo, B.; Ferretti, C.; Loverre, A.; Montrone, N.; et al.. Rehabilitation Outcomes of Patients With Severe Disability Poststroke. *Archives of Physical Medicine and Rehabilitation* **2019**, *100*, 520–529.e3.

27. Iorga, M.; Higgins, J.; Caplan, D.; Zinbarg, R.; Kiran, S.; Thompson, C.K.; et al.. Predicting language recovery in post-stroke aphasia using behavior and functional MRI. *Scientific Reports* **2021**, *11*, 8419.

28. Rivier, C.; Preti, M.G.; Nicolo, P.; Van De Ville, D.; Guggisberg, A.G.; Pirondini, E. Prediction of poststroke motor recovery benefits from measures of sub-acute widespread network damages. *Brain Communications* **2023**, *5*, fcad055.

29. Lee, J.; Kim, H.; Kim, J.; Chang, W.H.; Kim, Y.H. Multimodal Imaging Biomarker-Based Model Using Stratification Strategies for Predicting Upper Extremity Motor Recovery in Severe Stroke Patients. *Neurorehabilitation and Neural Repair* **2022**, *36*, 217–226.

30. Alquhayz, H.; Tufail, H.Z.; Raza, B. The multi-level classification network (MCN) with modified residual U-Net for ischemic stroke lesions segmentation from ATLAS. *Computers in Biology and Medicine* **2022**, *151*, 106332.

31. Sale, P.; Ferriero, G.; Ciabattoni, L.; Cortese, A.M.; Ferracuti, F.; Romeo, L.; et al.. Predicting Motor and Cognitive Improvement Through Machine Learning Algorithm in Human Subject that Underwent a Rehabilitation Treatment in the Early Stage of Stroke. *Journal of Stroke and Cerebrovascular Diseases* **2018**, *27*, 2962–2972.

32. White, A.; Saranti, M.; d'Avila Garcez, A.; Hope, T.M.H.; Price, C.J.; Bowman, H. Predicting recovery following stroke: Deep learning, multimodal data and feature selection using explainable AI. *NeuroImage: Clinical* **2024**, *43*, 103638.

33.   Tang, Z.; Su, W.; Liu, T.; Lu, H.; Liu, Y.; Li, H.; et al.. Prediction of poststroke independent walking using machine learning: a retrospective study. *BMC Neurology* **2024**, *24*, 332.

34.   Jabal, M.S.; Joly, O.; Kallmes, D.; Harston, G.; Rabinstein, A.; Huynh, T.; et al..   Interpretable Machine Learning Modeling for Ischemic Stroke Outcome Prediction. *Frontiers in Neurology* **2022**, *13*, 884693.

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.