

Classifying Ischemic and Hemorrhagic Strokes Using Machine Learning To Administer Accurate Treatment

Parimal Yeemani¹

Received November 14, 2025

Accepted March 25, 2026

Electronic access April 15, 2026

“Time is brain”, every second matters when treating stroke patients. Every second, approximately 32,000 neurons are lost after a stroke, equivalent to 3.6 years of aging. Around 137000 people die of a brain stroke in the USA per year. 87% are due to ischemic strokes. Reducing the time to classify a stroke is important for administering accurate treatment. Ischemic and hemorrhagic strokes must be classified at the earliest, as the treatment paths are fundamentally different with respect to the onset time. This paper explains why multiple modalities are required to achieve a faster and more accurate treatment path for a stroke patient. The proposed model aims to develop an automated system and analyze the best classification model. This uses CT images to classify the stroke and MRI to determine the stroke onset time using machine-learning models. During the classification of hemorrhagic and ischemic strokes, CT images were preprocessed and normalized, and HOG features were extracted to train the model. This model achieved a 94.41% accuracy on the testing data. Using a dataset of 1878 patients who had either ischemic or hemorrhagic stroke, the SVM architecture was able to predict the types with an AUC of 0.9754. Post classification, the MRI images DWI-FLAIR mismatch identifies if the stroke onset time was within 4.5 hours. The ResNet-18 model on 250 images with the nibabel library achieved a specificity of 0.9524 in identifying DWI-FLAIR mismatch. This paper shows how multiple modalities are required for classifying different stroke types and how various combinations of imaging types at different stages of treatment can save cost, accelerate stroke classification, assist onset time detection, and also eliminate subjective analysis when humans are involved for more accurate treatment.

Keywords: Strokes, Ischemic, Hemorrhagic, MRI, CT, Support Vector Machine, ResNet-18, CNN, Machine learning, Deep Learning algorithm, thrombectomy, acute stroke

Introduction

According to NIH, ischemic strokes contribute to 87% of total strokes worldwide, while hemorrhagic strokes make up the other 13%^{1,2}. As per the Stroke Awareness Foundation, there are 7 million stroke survivors who live in the USA, and around two-thirds of them are disabled. 795,000 people (which is about half the size of the population of Idaho) suffer from strokes each year, which makes it the second most common illness among humans³. In the US, people have strokes every 45 seconds and die from them every 4 minutes⁴.

Ischemic strokes occur when blood vessels are blocked by the buildup of plaque, preventing blood from reaching the brain, or when a clot develops in the cardiovascular system and travels to the brain. On the other hand, hemorrhagic strokes are ruptures in blood vessels that spill blood into the intracranial area of the brain⁵. In Acute Ischemic Stroke, Vascular Neurologists need to either eliminate plaque or clots inside the arteries to free up the bloodstream to treat strokes. Blood carries oxygen and nutrients to the brain, which keeps all cog-

nitive functions alive. When blood does not reach the relevant areas in the brain, humans are prone to disabilities like impaired speech, physical restrictions, and possibly death⁵. Despite the modern-day advances in medical imaging, accurate and rapid classification of stroke subtypes and stroke timeline remains a huge clinical challenge⁶. AI image analysis conducted on MRI (Magnetic Resonance Imaging) and CT (Computed Tomography Scan) scans can help analyze and treat patients more quickly⁷, reduce brain damage, and improve survival rates. “Time is brain”, a common phrase used to emphasize the rapid and irreparable loss of human nervous tissue as the stroke progresses⁸. Therapeutic interventions should be pursued immediately⁵. As per studies, the CT scans are faster, 5-10 min, compared to MRI, that take 20-60 min. CT is less expensive⁹ than MRI scans, and availability¹⁰ of CT is more than that of MRI.

The major success of stroke treatment lies in 3 prime factors, namely: classification to either hemorrhagic or ischemic stroke, as the treatment and medication depend on this classification. Secondly, determining the last wellness time in ischemic stroke, as a patient arrives before or after a 4.5-

¹ CCIR (Cambridge Centre For International Research)

hour period¹¹. Lastly, the size of the infarct versus penumbra helps neurologists choose eligible patients for thrombectomy¹². Clot-buster drugs are also used for ischemic strokes; patients need to take the drug before a 4 1/2 hour period, or else the treatment can do more harm than good¹³. MRI and CT need to make sure the stroke is not hemorrhagic, or else this drug can lead to the death of the patient. Hemorrhagic strokes are surgically treated differently because doctors reduce bleeding and pressure of excessive fluids in the brain. Surgeons may be able to repair blood vessels to save patients from blood loss. Even though manual interpretation of stroke imaging has been resource-intensive, this method is still time-consuming and is subject to diagnostic error. This study addresses the importance of automation and accurate stroke treatment by deep learning models to support precise and quick decision-making for diagnosis made by radiologists.

One well-known public education message to identify the symptoms of a stroke is F.A.S.T.¹⁴ F stands for face; A stands for Arm; S stands for speech; and T stands for Time to call. Stroke patients will feel the face numbness of drag, and their arms will become weak; speech may also be slurred. When these symptoms arrive during sleep, a patient must try to recall their last time of wellness when they woke up from sleep, so doctors can understand when the stroke happened and calculate their onset time¹⁵. The faster one can get treated, the better the probability of a positive outcome. Even a 2-hour delay can be a huge determinant in the patient's health. The MRI and CT scans need to make sure the stroke isn't hemorrhagic, or else the drug can lead to the death of patients.

Embolic strokes start as small blood clots that move around the bloodstream. These clots start an ischemic stroke in a certain area by blocking arteries, as the clot moves to other areas closer to the brain. To remove the clots, neurosurgeons dissolve them by using medications such as alteplase and Tenecteplase in a surgical method called thrombotic therapy^{13,16}. The people who have the biggest clots are the most in need of mechanical thrombectomy, which is an endovascular technique that removes clots throughout areas using a device that physically removes the clot. Patients with symptoms of stroke should have immediate attention to neuro images with non-contrast computed tomography or magnetic resonance imaging. Although this method is precise, the model is limited by the experiences and availability of staff and the infrastructure of hospitals.

As per existing research a DL model with high sensitivity and fewer false positives can greatly reduce the workload of radiologists^{17,18}. This novice model focuses on achieving high sensitivity. Also, recent studies further demonstrated that deep learning-based assistance systems not only improve diagnostic accuracy, but also significantly reduce reading times¹⁹.

METHODS

This is an experimental research to include multiple modalities to diagnose and determine accurate treatment for stroke patient. In this model, 1878 brain CT scan images of ischemic and hemorrhagic strokes were used for this research. This machine learning model is trained to identify if the stroke the patient is suffering from is Ischemic or hemorrhagic, so that doctors can understand what treatment route to follow. The two datasets that were used were selected because of their public availability and open source medical relevance, along with image quality. In Fig 1, this paper proposes a Proof of Concept to use multimodalities to decide the treatment path.

Dataset 1 uses SVM (Support Vector Machine)²⁰ to classify patients into either having an ischemic or hemorrhagic stroke. The training database CT-scan-image-data set, found on Kaggle, was uploaded by NOSHIN TASNIA²¹. There are an equal number of ischemic and hemorrhagic images to avoid bias in training. Then data is split up into train/test/validation, 60% of the 1878 (1126) ischemic and hemorrhagic images were sent to train, 20% (376) images were sent to test, and 20% (376) images were sent to validation for both ischemic and hemorrhagic. The model preprocessed the image acquisitions to 256 by 256 pixels, to form all images to the same size. SVM is a supervised learning method used for classification, regression, and outlier detection. The model was trained using 4 different feature extraction methods - raw images, deep features, Gabor, and HOG (Histogram of Gradient Radiance). Hyperparameter tuning was performed by using GridsearchCV, with cv=5, param_grid = { 'C': [0.1, 1, 10, 100], 'gamma': [1, 0.1, 0.01, 0.001], 'kernel': ['linear', 'rbf', 'poly', 'sigmoid']}. The best hyperparameter for RAW was identified as {'C': 10, 'gamma': 0.001, 'kernel': 'rbf'}, which produced a validation dataset accuracy of 90.96%. For Deep feature with cv=5 best parameters: {'C': 10, 'gamma': 0.001, 'kernel': 'rbf'} with validation dataset accuracy of 83.51%. Using the Gabor features extraction yielded a validation dataset accuracy of 0.606383. HOG feature extraction resulted in a validation dataset accuracy of 92.55%. For Data Leakage is prevented before HOG feature extraction using a pipeline = Pipeline([("scaler", StandardScaler()), ("pca", PCA(whiten=True)), ("svm", SVC(probability=True))]). The proposed model performs feature normalization and PCA-based dimensionality reduction. Raw image classification used 65536 features, but when feature extraction methods were applied, the model used only relevant features. HOG feature extraction used 50% less features (34596) than raw image training.

Dataset-2: This is a proof of concept to determine DWI-FLAIR mismatch²². A total of 250 images were used with a split of 60% images for training, 20% for validation, and 20% images for testing. Here we use a dual stream archi-

texture to extract features from both the MRI modalities of DWI (Diffusion Weighted Imaging) & FLAIR (Fluid Attenuated Inversion Recovery). As the MRI images are single-channel, we need to modify the images to 3-channel RGB images that can be processed by ResNet-18. Each ResNet outputs a 512 dimension feature-vector. The model learns both modalities together, which is important to identify DWI-FLAIR mismatch. These 2 vectors are fed to the various layers in the model. The Linear layer makes the model learn from the most meaningful 250 patterns. The ReLU layer adds non-linearity. To avoid overfitting, a dropout layer is introduced to disable 50% of neurons randomly. The final linear layer classifies into 2 classes: if there is a mismatch (infarct found) or no-mismatch(no-infarct). The model was trained using cross-entropy loss and optimized with the Adam optimizer at a learning rate of 1×10^{-4} with a batch size of 32 and 4 epochs. Training was performed using shuffled mini-batches, and GPU hardware was used for acceleration when available. The source of the second dataset is the Ischemic Stroke Lesion Segmentation ISLES challenge 2022²³. Since the datasets used were publicly available, there was no ethical approval or patient consent that was required.

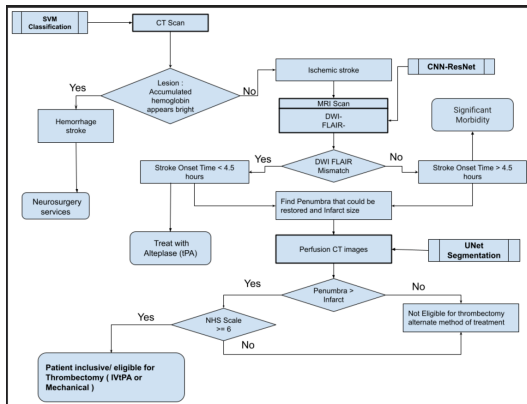


Fig. 1 Model proposed in this research. Decision-making model with ML Algorithms and Multimodal Medical Imaging to accelerate treatment given to Stroke Patients

RESULTS

	VALIDATION DATASET					TEST DATASET				
	ACCURACY	RECALL	PRECISION	F1	AUC	ACCURACY	RECALL	PRECISION	F1	AUC
RAW	0.885638	0.856383	0.909605	0.882192	0.9428	0.885638	0.856383	0.909605	0.882192	0.9428
DEEP FEATURE	0.87766	0.930851	0.841346	0.883838	0.9543	0.87766	0.930851	0.841346	0.883838	0.9543
GABOR	0.601064	0.510638	0.623377	0.561404	0.6434	0.601064	0.510638	0.623377	0.561404	0.6434
HOG	0.944149	0.93617	0.951351	0.9437	0.9754	0.944149	0.93617	0.951351	0.9437	0.9754

Fig. 2 Dataset-1 Validation & Test Dataset Results for SVM Classification.

HOG FEATURE EXTRACTION METHOD				
	ACCURACY	RECALL	PRECISION	F1
TRAIN	0.984014	0.982238	0.985741	0.983986
VALIDATION	0.925532	0.930851	0.921053	0.925926
TEST	0.944149	0.93617	0.951351	0.9437

Fig. 3 Dataset-1 HOG Feature Extraction Results for SVM Classification.

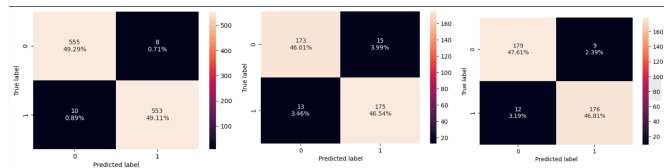


Fig. 4 Confusion Matrix Train, Validation & Test Dataset when HOG feature is applied.

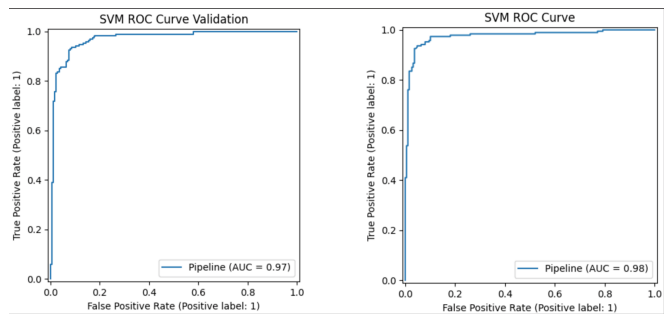


Fig. 5 SVM Dataset 1: Left ROC Curve represents Validation Dataset and Right ROC Curve represents Test Dataset when HOG feature extraction is applied

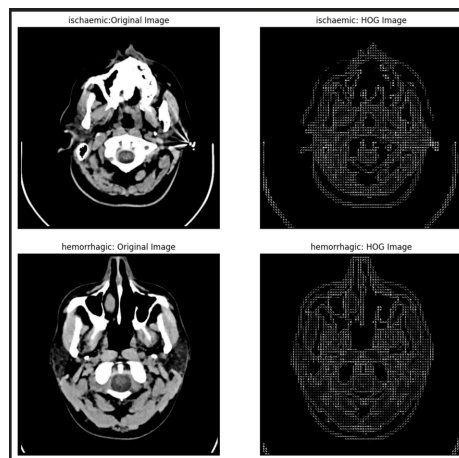


Fig. 6 HOG Image output corresponding to actual CT Scan Images

ResNet - 18 ON MRI IMAGES - CALCULATE DWI FLAIR MISMATCH				
	ACCURACY	SENSITIVITY	SPECIFICITY	AUC
TRAIN	0.9400	1.0000	0.9231	0.9959
VALIDATION	0.9000	0.8462	0.9189	0.9397
TEST	0.9600	1.0000	0.9524	1.0000

Fig. 7 Results for ResNet18 to Calculate DWI FLAIR Mismatch on MRI Scan Image

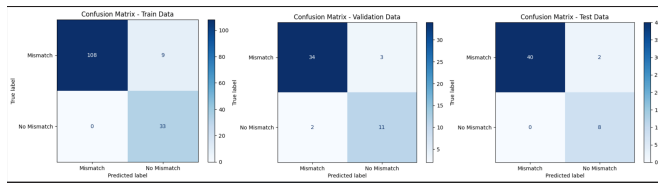


Fig. 8 Confusion Matrix Train, Validation & Test Dataset for DWI FLAIR Mismatch.

DISCUSSION

This study emphasizes the importance of multiple modalities in arriving at an accurate treatment. Different modalities are used at different stages of analysis. The CT scans are enough to identify if the stroke is hemorrhagic or ischemic. The model in this study has compared the results of image analysis using just raw images and compared them to using different feature extraction methods, viz., deep learning, gabor and HOG. The HOG feature extraction and F1 scores used in this model are on-par with state-of-the-art human radiologist predictions²⁴. Based on the results in Fig. 2, the PCA(Principal Component Analysis) SVM is more memory efficient and avoids storing large numbers of training parameters, and when applied with HOG feature extraction, yields a F1 score of 94%. HOG feature extraction used 50% less features than raw image training, hence saving memory. The use of such Machine Learning models will aid the radiologist during high peak hours and night shifts²⁵, when awaiting a decision from one radiologist, which causes a delay in treatments. On average, the radiologist may take 10-20 mins to conclude to classify haemorrhagic stroke patients, while this model may perform it below 2 min²⁶. One of the limitations of CT scan images can be asymmetrical CT brain images. Asymmetrical CT can result of improper head positioning in the CT gantry which can compromise the diagnostic value.²⁷

The main motive of this study is to apply different machine learning algorithms in a single case of stroke classification and treatment. The model proves why different modalities are

needed for a single stroke case. For haemorrhagic classification, CT is sufficient. Once the Ischemic Stroke is predicted, the next step would be to find the Stroke Onset Time. The CT scans in dataset1 only capture structural changes like tissue density, but cannot identify density changes over time. Hence, MRI images are used to identify the stroke onset time. As seen in dataset1, SVM performs better on non-linear CT scans, but stroke onset time needs progressive ischemic changes on non-linear spatial data that changes by time. The ResNet-18 model on dataset2 performs a classification on non-linear data to identify DWI-FLAIR mismatch. The input to this CNN model¹⁶ is an MRI image. The MRI images are not flattened and have high spatial data. The high number of features can be automatically learned by a CNN-based ResNet-18 model. While an SVM classification on a CT scan required a feature extraction method like HOG (Fig. 6) to be used before classifying. For identifying the stroke onset time in dataset-2, based on Fig. 7, the model achieved an accuracy of 96% on test data. The DWI-FLAIR mismatch implies that the stroke onset time occurred before 4.5 hours of arrival at the hospital. The patient, if treated within 4.5 hours, has a better chance of survival and fewer disabilities²⁸. The model implemented detects the mismatch and change between the DWI and FLAIR imaging, which shows that the stroke occurred between the onset time.

The novice model proposed a pathway, as shown in Fig.1, to accurately treat promptly. These findings highlight the accuracy and efficiency that machine learning can offer to aid radiologists in treating stroke patients, which directly impacts patient survival rates. The study strongly suggests that the goal of designing a computational method must include CT and multi-modal MRI imaging in identifying stroke onset time and determining treatment type. The algorithm leads to aiding decision-making for a patient to receive Thrombectomy, alteplase tPA²⁹, or an alternate method³⁰. The results shared are performed on publicly available datasets and not on real patient data. The future of this research will focus on expanding the dataset with more perfusion CT scans^{29,31,32} to improve mapping the lesions of infarct and penumbra more accurately, since they are the critical areas in which the stroke is prominent in the brain. A key limitation of this study is the computation. Currently, the limited GPU, RAM, and memory cause a delay in the output of results, which can be improved with higher-end infrastructure. The model took ~3-4 hours to train. But once trained, the trained SVM model takes 0.8 sec inference time, while the ResNet-18 takes 1.8 sec inference time to test 1 unseen image. The trained model gives results under 2 sec. The actual time to train is proportional to the image data and the hardware. There is scope in implementing advanced deep learning architectures such as ResNet-50, U-Net, or EffeinetNet, one could further improve the segmentation precision and have a reduction in processing time³³. Hence,

this project can be planned to be implemented on larger real datasets, revamped hardware, and alternative and refined CNN algorithms.

References

- 1 Feigin, V.L. (2021). Global, regional, and national burden of stroke and its risk factors, 1990–2019. *The Lancet Neurology*, Vol. 20, No. 10, pp. 795–820. [https://doi.org/10.1016/S1474-4422\(21\)00252-0](https://doi.org/10.1016/S1474-4422(21)00252-0).
- 2 Wang, Y. (2025). Traditional and machine learning models for predicting haemorrhagic transformation in ischaemic stroke: A systematic review and meta-analysis. *Systematic Reviews*, Vol. 14, No. 1, pg. 46. <https://doi.org/10.1186/s13643-025-02771-w>.
- 3 Feigin, V.L., Brainin, M., Norrving, B., et al. (2025). World Stroke Organization: Global stroke fact sheet 2025. *International Journal of Stroke*, Vol. 20, pp. 132–144.
- 4 Virani, S.S. (2023). Heart disease and stroke statistics—2023 update: A report from the American Heart Association. *Circulation*, Vol. 147, pp. 93–621.
- 5 Asch, C.J. (2010). Incidence, case fatality, and functional outcome of intracerebral haemorrhage over time. *The Lancet Neurology*, Vol. 9, No. 2, pp. 167–176. [https://doi.org/10.1016/S1474-4422\(09\)70340-0](https://doi.org/10.1016/S1474-4422(09)70340-0).
- 6 Wang, B., Jiang, B., Liu, D., & Zhu, R. (2025). Early predictive accuracy of machine learning for hemorrhagic transformation in acute ischemic stroke: Systematic review and meta-analysis. *Journal of Medical Internet Research*, Vol. 27, pg. 71654. <https://doi.org/10.2196/71654>.
- 7 Nukovic, J.J. (1908). Neuroimaging modalities used for ischemic stroke diagnosis and monitoring. *Medicina*, Vol. 59, No. 11. <https://doi.org/10.3390/medicina59111908>.
- 8 Saver, J.L. (2006). Time is brain—quantified. *Stroke*, Vol. 37, No. 1, pp. 263–266. <https://doi.org/10.1161/01.str.0000196957.55928.ab>.
- 9 Martinez, G., Katz, J.M., Pandya, A., et al. (2021). Cost-effectiveness study of initial imaging selection in acute ischemic stroke care. *Journal of the American College of Radiology*, Vol. 18, No. 6, pp. 820–833. <https://doi.org/10.1016/j.jacr.2020.12.013>.
- 10 Smith-Bindman, R. (2012). Use of diagnostic imaging studies and associated radiation exposure for patients enrolled in large integrated health care systems. *JAMA*, Vol. 307, No. 22, pp. 2400–2409.
- 11 Ferrone, N.G. (2025). Ten-year trends in last known well to arrival time in acute ischemic stroke patients: 2014–2023. *Stroke*, Vol. 56, pp. 591–602.
- 12 Chen, H., Chaturvedi, S., Gandhi, D., & Colasurdo, M. (2025). Stroke thrombectomy for large infarcts with limited penumbra: Systematic review and meta-analysis of randomized trials. *American Journal of Neuroradiology*, Vol. 46, pp. 915–920.
- 13 Logallo, N. (2017). Tenecteplase versus alteplase for management of acute ischemic stroke (NOR-TEST): A randomized controlled trial. *The Lancet Neurology*, Vol. 16, No. 10, pp. 781–788.
- 14 Kleindorfer, D.O. (2021). 2021 guideline for the prevention of stroke in patients with stroke and transient ischemic attack. *Stroke*, Vol. 52, No. 7, pp. 364–467.
- 15 Powers, W.J. (2019). 2019 guidelines for the early management of patients with acute ischemic stroke. *Stroke*, Vol. 50, No. 12, pp. 344–418.
- 16 Akay, E.M.Z. (2023). A deep learning analysis of stroke onset time prediction and comparison to DWI-FLAIR mismatch. *NeuroImage: Clinical*, Vol. 40, pg. 103544. <https://doi.org/10.1016/j.nicl.2023.103544>.
- 17 Karamian, A., & Seifi, A. (2025). Diagnostic accuracy of deep learning for intracranial hemorrhage detection in non-contrast brain CT scans: A systematic review and meta-analysis. *Journal of Clinical Medicine*, Vol. 14, No. 7, pg. 2377. <https://doi.org/10.3390/jcm14072377>.
- 18 Villringer, K. (2025). An artificial intelligence algorithm integrated into the clinical workflow can ensure high quality acute intracranial hemorrhage CT diagnostic. *Clinical Neuroradiology*, Vol. 35, No. 1, pp. 115–122. <https://doi.org/10.1007/s00062-024-01461-9>.
- 19 Kang, D.W. (2025). Deep learning-assisted detection of intracranial hemorrhage: Validation and impact on reader performance. *Neuroradiology*, Vol. 67, pp. 1511–1519. <https://doi.org/10.1007/s00234-025-03560-x>.
- 20 Cortes, C., & Vapnik, V. (1995). Support-vector networks. *Machine Learning*, Vol. 20, No. 3, pp. 273–297.
- 21 Tasnia, N. (2023). Brain Stroke Prediction CT Scan Image Dataset. Kaggle. <https://www.kaggle.com/datasets/noshintasnia/brain-stroke-prediction-ct-scan-image-dataset>.
- 22 Liu, Y. (2024). Artificial intelligence in ischemic stroke images: Current applications and future directions. *Frontiers in Neurology*, Vol. 15, pg. 1418060. <https://doi.org/10.3389/fneur.2024.1418060>.
- 23 Petzsche, M.R.H. (2022). ISLES 2022: A multi-center magnetic resonance imaging stroke lesion segmentation dataset. Zenodo. <https://doi.org/10.5281/zenodo.7153326>.
- 24 Al-Salahat, A., Pirahanchi, Y., & Dhasakeerthi, T. (2026). Current state and advancements of imaging in acute ischemic stroke: A practical review. *Neurological Sciences*, Vol. 47, pg. 29. <https://doi.org/10.1007/s10072-025-08685-8>.
- 25 D'Angelo, T. (2024). Accuracy and time efficiency of a novel deep learning algorithm for intracranial hemorrhage detection in CT scans. *La Radiologia Medica*, Vol. 129, pp. 1499–1506.
- 26 Khoruzhaya, A.N. (2025). Standalone AI versus AI-assisted radiologists in emergency ICH detection: A prospective multicenter diagnostic accuracy study. *Journal of Clinical Medicine*, Vol. 14, No. 16, pg. 5700. <https://doi.org/10.3390/jcm14165700>.
- 27 Downer, J.J., & Pretorius, P.M. (2009). Symmetry in computed tomography of the brain: The pitfalls. *Clinical Radiology*, Vol. 64, No. 3, pp. 298–306. <https://doi.org/10.1016/j.crad.2008.08.012>.
- 28 Thomalla, G. (2020). Intravenous alteplase for stroke with unknown time of onset guided by advanced imaging: Systematic review and meta-analysis. *The Lancet*, Vol. 396, No. 10262, pp. 1574–1584. [https://doi.org/10.1016/S0140-6736\(20\)32163-2](https://doi.org/10.1016/S0140-6736(20)32163-2).
- 29 Koneru, M. (2024). Early experience with artificial intelligence software to detect intracranial occlusive stroke in trauma patients. *Cureus*. <https://doi.org/10.7759/cureus.57084>.
- 30 Campbell, B.C.V., & Khatri, P. (2020). Stroke. *The Lancet*, Vol. 396, No. 10244, pp. 129–142. [https://doi.org/10.1016/S0140-6736\(20\)31179-X](https://doi.org/10.1016/S0140-6736(20)31179-X).
- 31 Sheth, S.A. (2019). Machine learning-enabled automated determination of acute ischemic core from computed tomography angiography. *Stroke*, Vol. 50, No. 11. <https://doi.org/10.1161/STROKEAHA.119.026189>.
- 32 Ho, K.C., Speier, W., El-Saden, S., & Arnold, C.W. (2018). Classifying acute ischemic stroke onset time using deep imaging features. *AMIA Annual Symposium Proceedings*, pp. 892–901.
- 33 Wang, Y. (2024). Patterns and clinical implications of hemorrhagic transformation after thrombolysis in acute ischemic stroke. *Neurology*, Vol. 103, No. 11.