scientific reports

OPEN

Development and validation of a multi-omics hemorrhagic transformation model based on hyperattenuated imaging markers following mechanical thrombectomy

Lina Jiang[1], Guoping Zhu[1], Yue Wang[1], Jiayi Hong[2], Jingjing Fu[2], Jibo Hu[1], Shengxiang Xiao[1], Jiayi Chu[1], Sheng Hu[1]✉ & Wenbo Xiao[3]✉

This study aimed to develop a predictive model integrating clinical, radiomics, and deep learning (DL) features of hyperattenuated imaging markers (HIM) from computed tomography scans immediately following mechanical thrombectomy (MT) to predict hemorrhagic transformation (HT). A total of 239 patients with HIM who underwent MT were enrolled, with 191 patients (80%) in the training cohort and 48 patients (20%) in the validation cohort. Additionally, the model was tested on an internal prospective cohort of 49 patients. A total of 1834 radiomics features and 2048 DL features were extracted from HIM images. Statistical methods, such as analysis of variance, Pearson's correlation coefficient, principal component analysis, and least absolute shrinkage and selection operator, were used to select the most significant features. A K-Nearest Neighbor classifier was employed to develop a combined model integrating clinical, radiomics, and DL features for HT prediction. Model performance was evaluated using metrics such as accuracy, sensitivity, specificity, receiver operating characteristic curves, and area under curve (AUC). In the training, validation, and test cohorts, the combined model achieved AUCs of 0.926, 0.923, and 0.887, respectively, outperforming other models, including clinical, radiomics, and DL models, as well as hybrid models combining subsets of features (Clinical + Radiomics, DL + Radiomics, and Clinical + DL) in predicting HT. The combined model, which integrates clinical, radiomics, and DL features derived from HIM, demonstrated efficacy in noninvasively predicting HT. These findings suggest its potential utility in guiding clinical decision-making for patients with MT.

Keywords Intracranial hemorrhages, Machine learning, Thrombectomy, Acute ischemic stroke, Multi-detector CT

Endovascular mechanical thrombectomy (MT) has become the standard of care for acute ischemic stroke (AIS) due to its proven efficacy in enhancing reperfusion rates and improving clinical outcomes. Despite these advancements, over 50% of patients still experience unfavorable outcomes, with hemorrhagic transformation (HT) emerging as a significant risk factor[1]. HT encompasses intracerebral hemorrhage, which can lead to fatal outcomes, as well as asymptomatic intracerebral hemorrhage and subarachnoid hemorrhage (SAH), both of which can adversely impact recovery[2]. Timely postoperative management plays a pivotal role in improving patient prognosis. Early prediction of HT is essential for guiding therapeutic adjustments, such as initiating blood pressure control and implementing drug therapy to mitigate the risk of re-occlusion caused by endothelial injury.

[1]Department of Radiology, the Fourth Affiliated Hospital of School of Medicine, and International School of Medicine, International Institutes of Medicine, Zhejiang University, Yiwu 322000, China. [2]Department of Neurology, the Fourth Affiliated Hospital of School of Medicine, and International School of Medicine, International Institutes of Medicine, Zhejiang University, Yiwu 322000, China. [3]Department of Radiology, the First Affiliated Hospital, College of Medicine, Zhejiang University, Hangzhou 310000, China. ✉email: 8015135@zju.edu.cn; xiaowenbo@zju.edu.cn

The hyperattenuated imaging marker (HIM) detected immediately on noncontrast computed tomography (NCCT) following MT is the earliest and most readily available imaging marker for predicting postoperative complications[3,4]. HIM appears as areas of hyperattenuation on post-thrombectomy cranial NCCT scans, reflecting extravasated contrast agents and, in some cases, hemorrhage. While extravasated contrast agents are gradually absorbed postoperatively, hemorrhage may develop or worsen over time. Studies have reported that the prevalence of HIM postoperatively ranges from 32.9 to 87.5% in patients[5,6]. The presence of HIM indicates disruption of the blood-brain barrier and blood-cerebrospinal fluid barrier, which is closely linked to bleeding, making it a commonly used imaging marker for predicting postoperative hemorrhage[7,8]. However, its accuracy in predicting hemorrhage based on HIM remains suboptimal.

Artificial intelligence, particularly deep learning (DL), has emerged as a promising tool for predicting HT following MT in patients with stroke, offering high accuracy and potential improvements in diagnostic and therapeutic strategies. Radiomics, which facilitates precise hemorrhage prediction, has gained considerable attention[9]. With the advancements of DL, this approach has demonstrated even greater predictive potential[10]. DL's ability to perform multi-layered feature extraction directly from raw imaging data provides robust predictive capabilities. However, it does not consistently outperform radiomics[11]. Our study aims to integrate DL, radiomics, and clinical features to achieve a more accurate prediction of HT following thrombectomy.

Methods

Patients

Ethical approval of the study protocol

This study was approved by the Ethics Committee of the Fourth Affiliated Hospital, Zhejiang University College of Medicine (Approval Number: KY2024020). The requirement for informed consent was waived by the committee. All clinical studies adhered to the tenets of the Declaration of Helsinki.

Patient selection and study design

Patients diagnosed with AIS due to intracranial large vessel occlusion and treated with endovascular MT between June 2016 and September 2023 at the Fourth Affiliated Hospital, Zhejiang University School of Medicine, were retrospectively reviewed. The indications and contraindications of MT and thrombolysis were based on the most current guidelines available at the time of treatment. General clinical features, laboratory findings, clinical presentations, and imaging data of the patients were collected. The inclusion criteria were as follows: (1) Patients underwent head NCCT or magnetic resonance imaging (MRI) after MT; (2) initial postoperative head NCCT was performed within 30 min after MT; and (3) HIM, defined as an area of hyperattenuation visible in the brain parenchyma and subarachnoid space, was detected on the initial head NCCT after MT. The exclusion criteria were as follows: (1) The follow-up time of head NCCT after MT was less than 19 h[12]; (2) artifacts (for instance, metal artifacts or motion artifacts) impaired the assessment of HIM or hemorrhage; (3) the use of iodinated contrast prior to preoperative CT affected the determination of HT. A flowchart of the study subjects is presented in Fig. 1.

Imaging

All initial postoperative nonenhanced head CT images were obtained within 30 min after thrombectomy from a 64-row spiral CT scanner (Somatom® Definition AS, Siemens Healthineers, Forchheim, Germany) and a 62-row spiral scanner (Optima® CT620, GE Medical Systems, Milwaukee, WI, US). The scanning parameters were as follows: axial mode, tube voltage 120 kV, tube current 250–300 mAs, scanning range from the skull base to the cranial roof, section thickness of 5 mm, and reconstruction with a standard algorithm. Imaging data were evaluated in a randomized order by two neuroradiologists with more than 10 years of experience, who

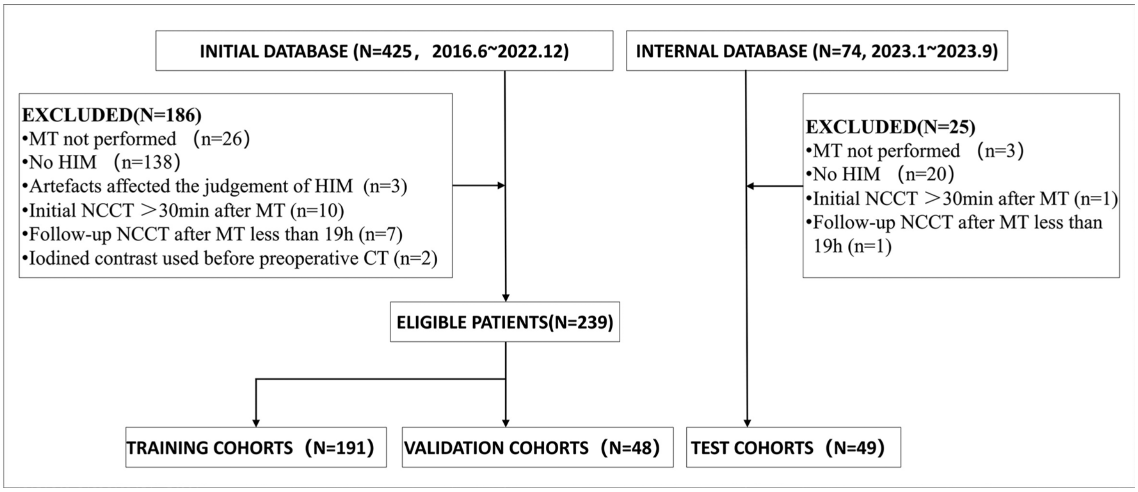

Fig. 1. The patient selection flowchart.

were blinded to the clinical situation. Any conflicting interpretations were resolved through discussion by two evaluators to reach a consensus.

The workflow and global analysis pipeline for the classification model are presented in Fig. 2. HIM was manually contoured on each CT image, and the region of interest (ROI) enclosing the HIM was cropped for input into the predictive model. Inter-rater reliability was assessed using intraclass correlation coefficient (ICC) analysis on a randomly selected subset of 30 cases, with 1,819 features demonstrating excellent consistency (ICC > 0.75) retained as robust features for subsequent analysis. The DL and radiomics features were extracted separately: DL extracted features from the largest cross-sectional ROI, while radiomics extracted features from the entire ROI area. Features selection was performed sequentially using analysis of variance (ANOVA), Pearson's correlation coefficient, and least absolute shrinkage and selection operator (LASSO). K-Nearest Neighbor (KNN) classifier was used to construct various models, including a clinic model, a radiomics model, a DL model, a Clinical + Radiomics model, a DL + Radiomics model, a Clinical + DL model, and a combined model. Model performance was analyzed using receiver operating characteristic (ROC) curves, an area under the curve (AUC), and decision curve analysis (DCA).

Reference standard

HT was classified based on the following criteria: (1) Hyperattenuation remained evident on the 24-h CT scan, and (2) follow-up CT or MRI conducted within 90 days post-surgery revealed hemorrhage within the infarcted area. The maximum Hounsfield unit (HUmax) was defined as the highest HU value within a specified ROI, measured as a 3 × 3 pixel area. Two neuroradiology staff members, each with more than five years of experience in neuroradiology, independently evaluated all training and test cohort images without knowledge of patient outcomes. Any discrepancies were resolved through consensus.

Data collection

In this study, 288 patients were enrolled. A total of 239 patients treated between June 2016 and December 2022 were used for model training and validation. Of these, 191 cases (80%) were randomly assigned to the training cohort, while 48 patients (20%) formed the test cohort. To evaluate the model's clinical applicability, it was prospectively tested on a test cohort from January 2023 to September 2023 at the same hospital using real-world data.

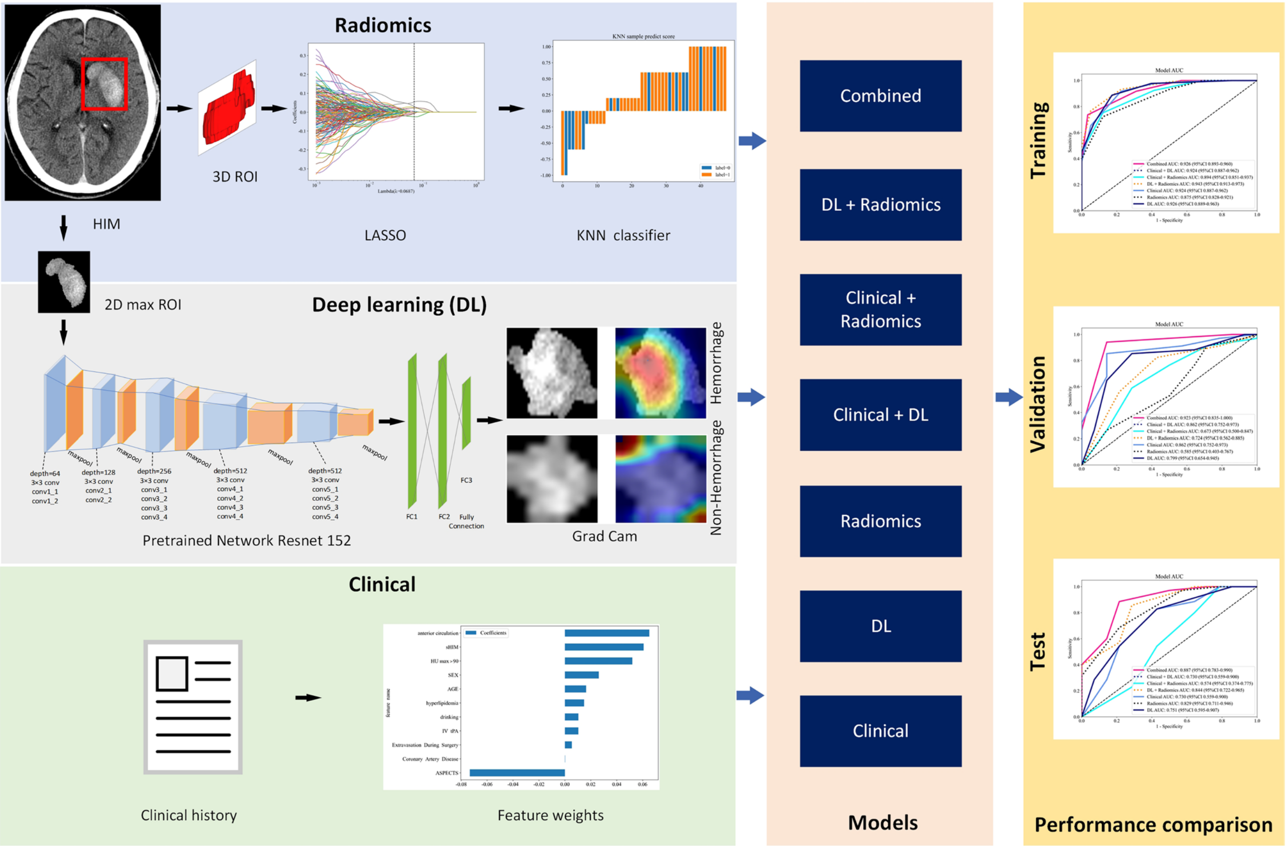

Fig. 2. The overall framework of the proposed model. ROI, region of interest; LASSO, least absolute shrinkage and selection operator; KNN, K-Nearest Neighbor. Grad Cam, Gradient-weighted Class Activation Mapping. AUC, area under curve.

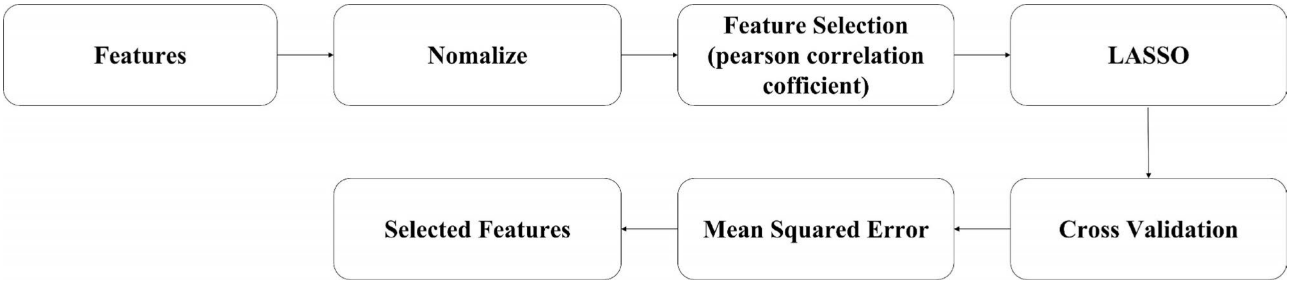

Fig. 3. The process of radiomics extraction. LASSO = least absolute shrinkage and selection operator.

Model name	AUC	95% CI	Sensitivity	Specificity	Accuracy	PPV	NPV	Precision
ResNet152	0.799	0.6538–0.9449	0.647	0.857	0.708	0.917	0.5	0.917
Inception v3	0.738	0.5889–0.8880	0.353	0.929	0.521	0.923	0.371	0.923
VGG 16	0.739	0.5888–0.8902	0.618	0.714	0.646	0.84	0.435	0.84
VGG 19	0.69	0.5207–0.8595	0.441	0.786	0.542	0.833	0.367	0.833
ResNet50	0.683	0.5159–0.8496	0.324	0.857	0.479	0.846	0.343	0.846

Table 1. The prediction performance of different deep learning models. AUC, area under curve; PPV, positive predictive value; NPV, negative predictive value; VGG, Visual Geometry Group.

Baseline and preprocedural variables, including age, sex, hyperlipidemia, hypertension, diabetes mellitus, coronary artery disease, atrial fibrillation, history of drinking or smoking, baseline National Institute of Health Stroke Scale scores, and thrombolysis status, were extracted by reviewing patient charts, procedure notes, and follow-up records. Baseline Alberta Stroke Program Early CT Score for anterior and posterior circulation, subarachnoid HIM, and HUmax values were determined through discussions between two neuroradiologists (each with more than five years of experience) who were blinded to the patient's clinical information. Surgical details, including stent implantation, modified Thrombolysis in Cerebral Infarction scores, and the use of aspiration catheters, were also collected.

Feature extraction and selection

On each initial postoperative CT image, the ROI of the HIM was manually segmented along the HIM contour, incorporating both parenchyma and sulcus. This process was conducted using the ITK-SNAP software (version 3.8.0, http://www.itksnap.org/). Both radiologists were blinded to the clinical information and final outcomes.

The process of radiomics extraction is illustrated in Fig. 3. The methodological workflow consisted of the following key steps: to eliminate any potential variations in CT images obtained using different CT scanners, NCCT images were reconstructed using a voxel size of 1 × 1 × 1 mm3 and gray-scale discretization. Two radiologists independently segmented the images of every HIM and measured them using a method that ensured double anonymization. An intraclass correlation coefficient (ICC) ≥ 0.75 was considered robust. A Pyradiomics in-house feature analysis program (http://pyradiomics.readthedocs.io) was used to extract radiomic features from HIM in NCCT. From the NCCT images, 1834 radiomics features were extracted. Various techniques were used to extract texture features, including first-order statistics, shape-based analysis, gray-level cooccurrence matrix (GLCM), gray-level run length matrix (GLRLM), neighboring gray-tone difference matrix (NGTDM), gray-level size zone matrix (GLSZM), and gray-level dependence matrix (GLDM). All robust radiomics features were selected to develop the radiomics models. A Mann–Whitney U test and feature screening were performed for all features. The features with a p-value < 0.05 were included. Spearman's rank correlation coefficient was calculated for the features exhibiting high repeatability to evaluate the association between features. If the correlation coefficient between any two characteristics > 0.9, one of the characteristics would be preserved. To ensure a comprehensive depiction of the features, a greedy recursive deletion strategy was used for feature filtering. This process includes removing the characteristics with the highest duplication in the existing collection during each cycle. For signature construction, the discovery dataset was subjected to the least absolute shrinkage and selection operator (LASSO) regression model. By applying a regularization weight λ, LASSO reduces the regression coefficients to zero and assigns a value of zero to the coefficients of the irrelevant features. To determine the best λ, a 10-fold cross-validation was applied, and the λ value that resulted in the lowest cross-validation error was selected using the minimum criteria. The included features with coefficients that were not zero were used to fit the regression model and were incorporated into the model. LASSO regression modeling was performed using the Python scikit-learn library.

We extracted features from five DL models to perform modeling, including ResNet152, Inception v3, ResNet50, VGG 16, and VGG 19. Among these, ResNet152 demonstrated the highest performance with an AUC of 0.799, leading us to select it as the DL model for this study (Table 1). ResNet152 was used as a pre-trained model and fine-tuned using the maximum cross-sectional imaging of the HIM ROI. To improve the model's

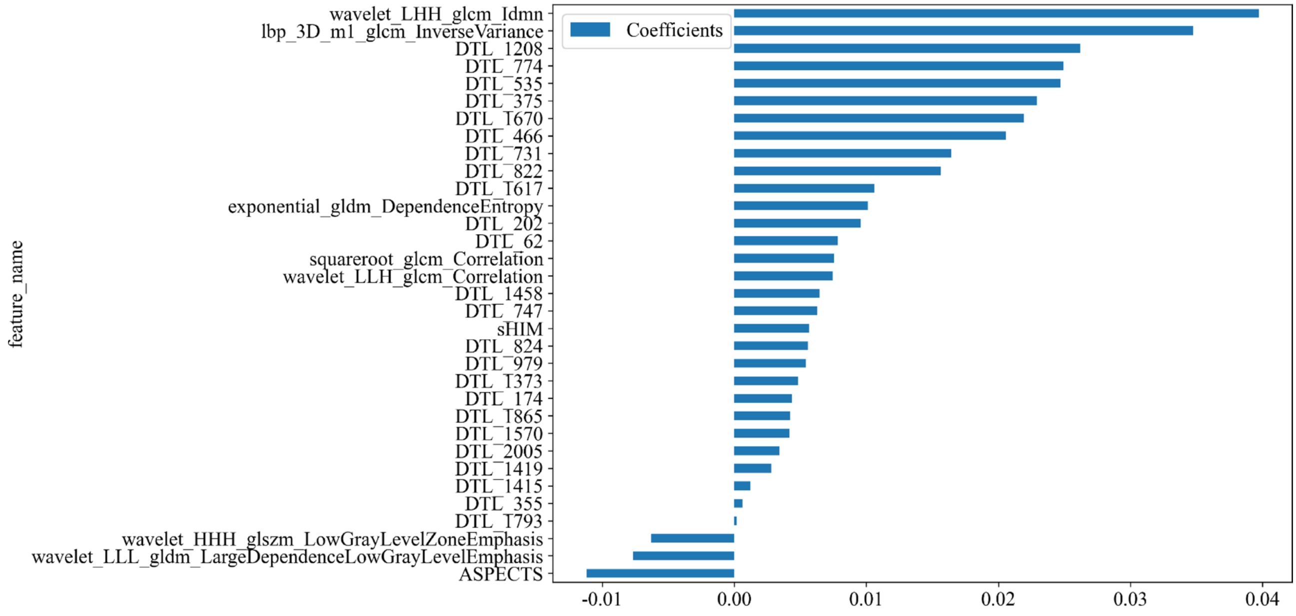

Fig. 4. The histogram of the coefficients of the selected features. LHH = low-high-high-pass filtered image; GLCM = gray-level co-occurrence matrix; DTL = deep transfer learning; GLDM = Gray Level Dependence Matrix; sHIM = Subarachnoid hyperattenuated imaging markers; HHH = high-high-high-pass filtered image; GLSZM = gray level size zone matrix; ASPECTS = Alberta Stroke Program Early CT Score.

Model name	AUC	95% CI	Sensitivity	Specificity	Accuracy	PPV	NPV	Precision	Task
KNN	0.926	0.8927–0.9602	0.383	1	0.571	1	0.414	1	Train
KNN	0.923	0.8348–1.0000	0.618	0.929	0.708	0.955	0.5	0.955	Validation
RandomForest	1	1.0000–1.0000	0.985	1	0.99	1	0.967	1	Train
RandomForest	0.816	0.6834–0.9489	0.794	0.714	0.771	0.871	0.588	0.871	Validation
XGBoost	1	1.0000–1.0000	0.992	1	0.995	1	0.983	1	Train
XGBoost	0.811	0.6848–0.9371	0.647	0.857	0.708	0.917	0.5	0.917	Validation
LightGBM	0.985	0.9714–0.9988	0.947	0.931	0.942	0.969	0.885	0.969	Train
LightGBM	0.784	0.6309–0.9363	0.765	0.714	0.75	0.867	0.556	0.867	Validation
GradientBoosting	0.995	0.9898–1.0000	0.947	0.966	0.953	0.984	0.889	0.984	Train
GradientBoosting	0.803	0.6579–0.9472	0.824	0.643	0.771	0.848	0.6	0.848	Validation
AdaBoost	0.967	0.9460–0.9877	0.85	0.966	0.885	0.983	0.737	0.983	Train
AdaBoost	0.722	0.5574–0.8859	0.647	0.786	0.688	0.88	0.478	0.88	Validation

Table 2. The prediction performance of different machine learning models. AUC, area under curve; PPV, positive predictive value; NPV, negative predictive value; KNN, K-Nearest Neighbor; GBM, Gradient Boosting Machine.

generalization ability, data augmentation techniques such as RandomResizedCrop and RandomHorizontalFlip were applied. The features from the 'avgpool' layer of the ResNet152 convolutional network were used as the DL features. A total of 2048 DL features were extracted from CT images, and 1834 features were selected after dimensionality reduction using principal component analysis (PCA). Ultimately, our multi-omics features comprised clinical data, radiomics features, and DL features. Clinical, radiomics, and DL features were standardized to have a mean of 0 and a variance of 1. To enhance the models' generalization ability and minimize overfitting, feature selection was performed using ANOVA, Pearson's correlation coefficient, PCA, and LASSO. For the clinical model, radiomics model, and DL model, ANOVA, Pearson's correlation coefficient, and LASSO were sequentially applied to complete the feature selection process. For the combined model, PCA, ANOVA, Pearson's correlation coefficient, and LASSO were sequentially used. Ultimately, 33 combined features were selected for constructing the combined model, including two clinical features, seven radiomics features, and 24 DL features. The relative importance of the extracted features is quantitatively demonstrated in Fig. 4, which provides a ranked visualization of the most discriminative radiomics biomarkers identified through our analysis. For LASSO, five-fold cross-validation was applied. The HT risk prediction was carried out using the KNN classifier, which demonstrated the highest performance with an AUC of 0.923 in the validation cohort. Detailed model parameters are presented in Table 2.

Model developed and validation

A combined model integrating clinical, radiomics, and DL features was developed. Additionally, individual models (clinical, radiomics, and DL), as well as combined models (Clinical + Radiomics, DL + Radiomics, and Clinical + DL) were constructed for comparison. The performance of these models was evaluated using accuracy, sensitivity, specificity, ROC curves, and AUC. The training cohort comprised 191 patients (80%), the validation cohort included 48 patients (20%), and the test cohort consisted of 49 patients. The models were trained and tested on the training, validation and test cohorts, respectively. To assess the agreement between the predicted and actual classifications, calibration curves were plotted to evaluate the calibration efficiency of each model. Furthermore, DCA was performed to assess the clinical utility of the predictive models.

Implementation and hardware

The network architecture was implemented using Python and the PyTorch library. The DL model was trained using the stochastic gradient descent algorithm as the optimizer, with a learning rate of 0.001, a mini-batch size of 8, and a binary cross-entropy loss function. Batch normalization was applied after each convolution layer to accelerate the convergence and reduce overfitting. The network was trained for 50 epochs.

Statistical analysis and performance evaluation

Statistical analysis was performed using the open-source ‘SciPy’ package in Python. The independent t-test was used to compare differences between variables. The area under the (ROC) curve and the confusion matrix were calculated to evaluate model performance. A calibration curve was employed to assess the reliability and validity of the model. Furthermore, DCA was used to determine the clinical usefulness of the prediction models. The selection criteria were based on evaluating the performance of machine learning models in predicting HT within the test group, with a focus on prioritizing accuracy and the area under the receiver operating characteristic curve (AUC).

Results

Cohort and clinical characteristics

The study included a training cohort of 191 patients, a validation cohort of 48 patients, and a test cohort of 49 patients. Clinical signatures were selected sequentially using ANOVA, Pearson correlation coefficient, and LASSO. Table 3 provides a detailed summary of the patients’ clinical characteristics. HT occurred in 202 of the

Feature name	Training	Validation	Test	*P* value
Age(year), mean ± SD	67.57 ± 15.10	65.17 ± 11.19	67.33 ± 14.52	0.56
Men, n(%)	118(61.78)	36(75.00)	32(65.31)	0.23
Atrial Fibrillation, n(%)	77(40.31)	17(35.42)	22(44.90)	0.64
Hypertension, n(%)	116(60.73)	24(50.00)	33(67.35)	0.21
Hyperlipidemia, n(%)	17(8.90)	1(2.08)	9(18.37)	0.02*
Diabetes Mellitus, n(%)	33(17.28)	4(8.33)	12(24.49)	0.16
Coronary Artery Disease, n(%)	29(15.18)	4(8.33)	3(6.12)	0.15
Drinking, n(%)	40(20.94)	13(27.08)	8(16.33)	0.43
Smoking, n(%)	42(21.99)	11(22.92)	10(20.41)	0.96
Thrombolysis, n(%)	69(36.13)	23(47.92)	10(20.41)	0.02*
Baseline NIHSS, median(Q1,Q3)	16(12–21)	16(13–19)	16(12–21)	0.30
ASPECTS, median(Q1,Q3)	9(8–10)	9(8–10)	9(8–10)	0.36
HUmax ≥ 90,n(%)	54(28.27)	15(31.25)	13(26.53)	0.87
Anterior Circulation, n(%)	179(93.72)	43(89.58)	44(89.80)	0.48
sHIM, n(%)	75(39.26)	16(33.33)	15(30.61)	0.46
mTICI more than 2b, n(%)	170(89.00)	42(87.50)	47(95.92)	0.30
Stent Type, n(%)				0.11
Solitaire	103(53.93)	33(68.75)	29(59.18)	
Trevo	34(17.80)	8(16.67)	0(0.00)	
Solitaire + Trevo	31(16.23)	5(10.42)	18(36.73)	
Others	23(12.04)	5(4.17)	2(4.08)	
Pass Number, median(Q1,Q3)	1(1–2)	2(1–3)	1(0–2)	0.002*
Aspiration Catheter, n(%)	63(32.98)	14(29.17)	37(75.51)	0.000*
Stent Implantation, n(%)	42(21.99)	6(12.50)	9(18.37)	0.33

Table 3. Characteristics of patients in the train, validation and test cohort. SD, standard deviation; ASPECTS, Alberta Stroke Program Early Computed Tomography Score; Q1, first quartile; Q3, third quartile; NIHSS, National Institutes of Health Stroke Scale; HU, Hounsfield; sHIM, subarachnoid hyperattenuated imaging markers; mTICI, modified thrombolysis in cerebral infarction.

288 patients. Significant differences were observed between the training, validation, and test cohorts in terms of hyperlipidemia, ($p = 0.02$), thrombolysis ($p = 0.02$), pass number ($p = 0.002$), and aspiration catheter ($p = 0.000$).

Performance of the different models in training cohort

In the training cohort, the combined model achieved an AUC of 0.926 (95% confidence interval (CI): 0.893–0.960), with a sensitivity of 0.383, and a specificity of 1.0. The clinical model has an AUC of 0.804 (95% CI: 0.744–0.864), while the radiomics model exhibited an AUC of 0.875 (95% CI: 0.828–0.921). The DL model presented an AUC of 0.926 (95% CI: 0.889–0.9631). The Clinical + Radiomics model depicted an AUC of 0.894 (95% CI: 0.851–0.937). The DL + Radiomics model achieved an AUC of 0.943 (95% CI: 0.913–0.973), and the Clinical + DL model exhibited an AUC of 0.924 (95% CI: 0.887–0.962) (see Fig. 5).

Performance of different models in validation cohort

In the Validation cohort, the combined model achieved an AUC of 0.923 (95% CI: 0.835–1.000), with an a sensitivity of 0.618, and a specificity of 0.929. The clinical model obtained an AUC of 0.714 (95% CI: 0.562–0.867), while the radiomics model depicted an AUC of 0.585 (95% CI: 0.403–0.767). The DL model presented an AUC of 0.799 (95% CI: 0.654–0.945). The Clinical + Radiomics model exhibited an AUC of 0.673 (95% CI: 0.500–0.847). The DL + Radiomics model exhibited an AUC of 0.724 (95% CI: 0.562–0.885). The Clinical + DL model exhibited an AUC of 0.862 (95% CI: 0.7522–0.973) (see Fig. 6).

Performance of different models in the test cohort

The diagnostic accuracy of the combined, clinical, radiomics, DL model, Clinic Rad model, DL + Rad model, and Clinic + DL model is summarized in Table 4. The AUC values in the test cohort were as follows: 0.887 for the combined model (incorporating radiomics, DL, and clinical data), 0.829 for the radiomics model, 0.751 for the DL model, 0.550 for the clinic model, 0.574 for the Clinical + Radiomics model, 0.844 for the DL + Radiomics

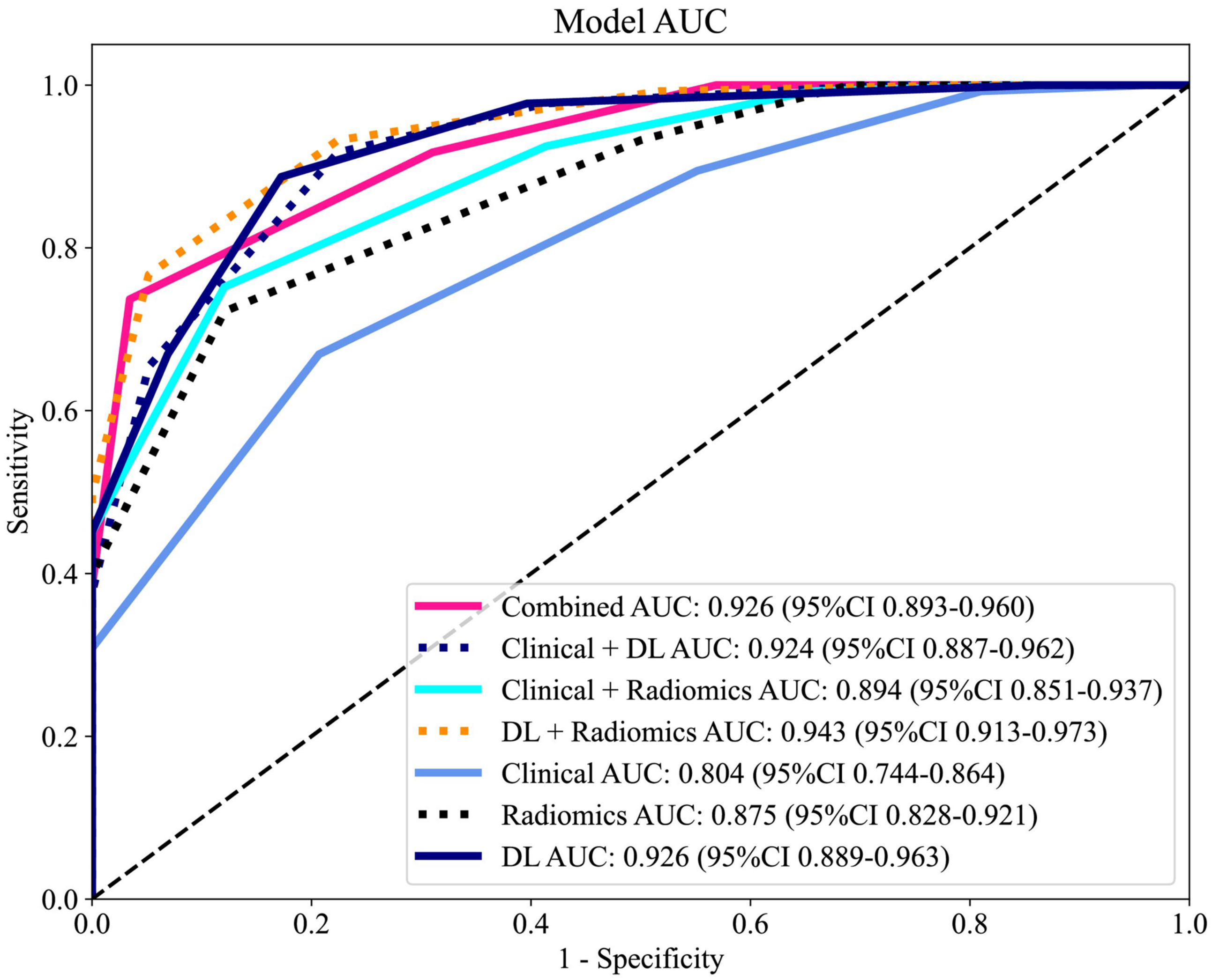

Fig. 5. The ROC curves of all the models in the train cohort. ROC, receiver operating characteristic.

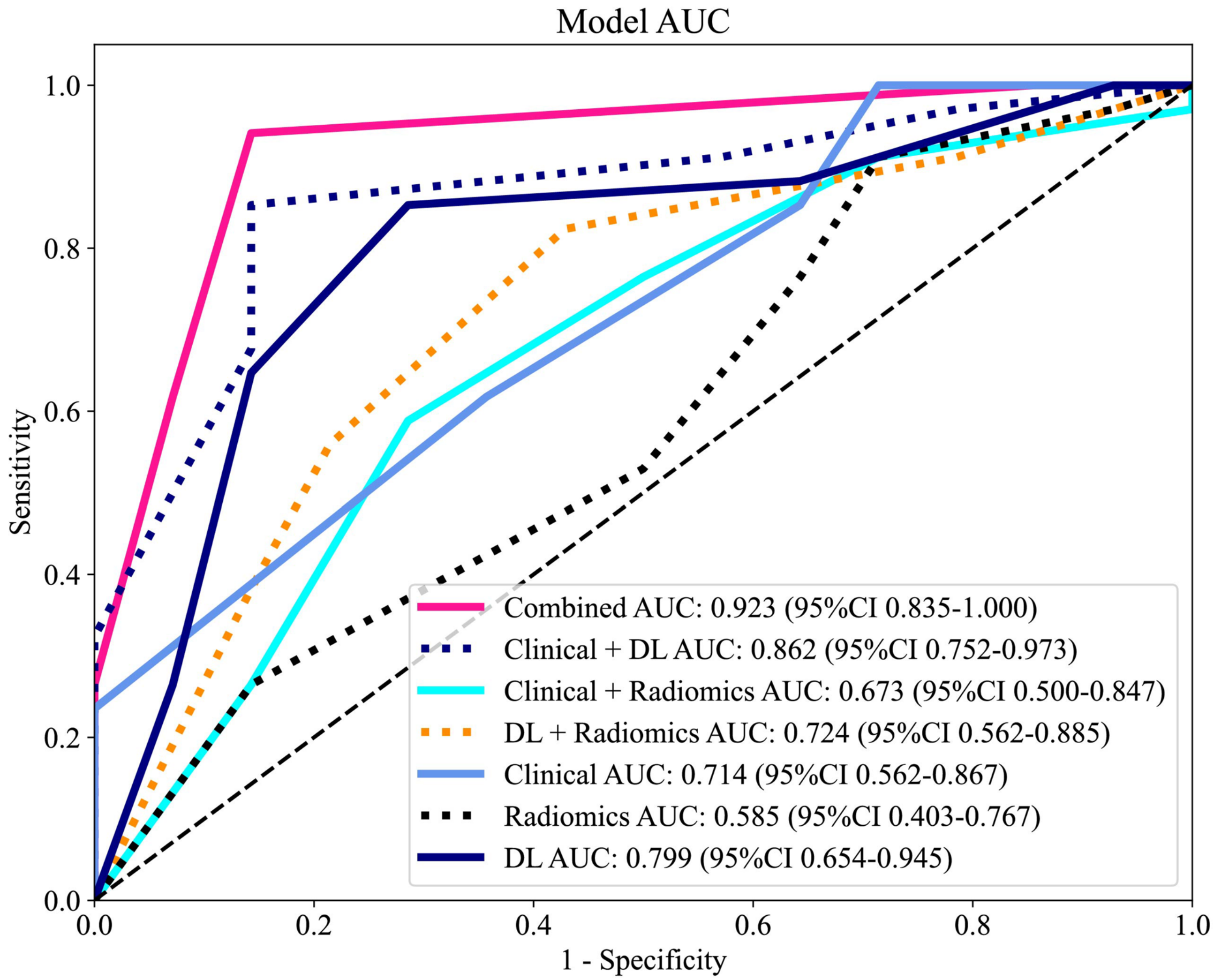

Fig. 6. The ROC curves of all the models in validation cohort. ROC, receiver operating characteristic.

Model	AUC	95% CI	Sensitivity	Specificity	Accuracy	PPV	NPV	Precision
Combined	0.887	0.7832–0.9903	0.6	0.857	0.673	0.913	0.462	0.913
Clinical + DL	0.73	0.5595–0.8997	0.543	0.786	0.612	0.864	0.407	0.864
Clinical + Radiomics	0.574	0.3735–0.7754	0.8	0.357	0.673	0.757	0.417	0.757
DL + Radiomics	0.844	0.7223–0.9654	0.571	0.786	0.633	0.87	0.423	0.87
Clinical	0.55	0.3693–0.7307	0.229	0.786	0.388	0.727	0.289	0.727
Radiomics	0.829	0.7107–0.9464	0.314	1	0.51	1	0.368	1
DL	0.751	0.5952–0.9068	0.543	0.786	0.612	0.864	0.407	0.864

Table 4. The prediction performance of different models. AUC, area under curve, PPV, positive predictive value; NPV, negative predictive value; DL, deep learning.

model, and 0.730 for the Clinical + DL model(see Fig. 7). Table 5 shows the confusion matrix of the combined model.

The calibration curves demonstrated good agreement between the predictions of the models and the actual classifications of HT in the test cohort, with the combined model fitting the data exceptionally well. The DCA further illustrated that, compared to scenarios without any prediction model, all models significantly improved the patient outcome predictions. The combined model indicated the best clinical benefit in automatic HT classification. Figure 8 depicts calibration curves and DCA curves for different models in the test cohort.

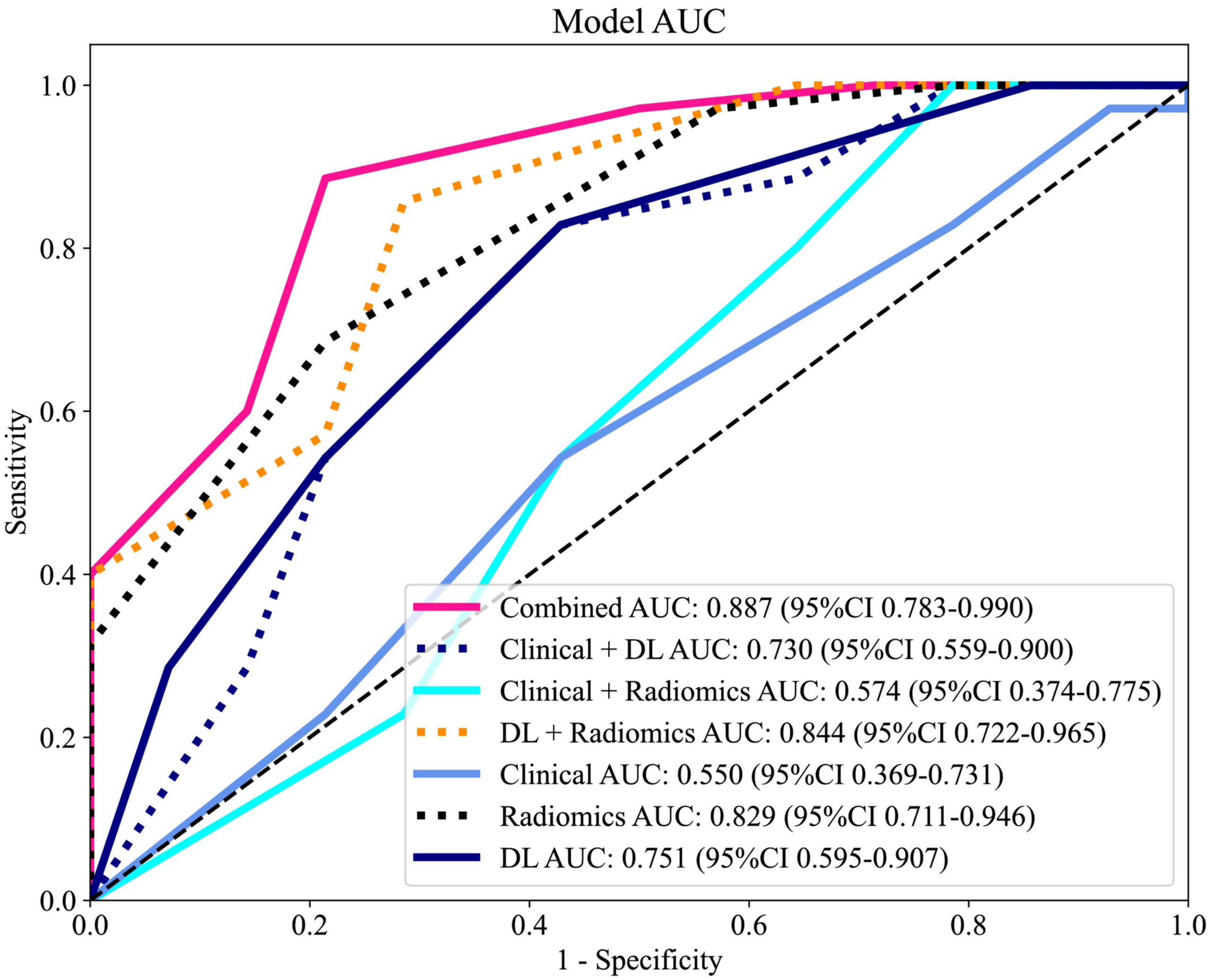

Fig. 7. The ROC curves of all the models in test cohort. ROC, receiver operating characteristic.

	Validation		Test	
	Predicted positive	**Predicted negative**	**Predicted positive**	**Predicted negative**
Actually positive	21	13	21	14
Actually negative	1	13	2	12

Table 5. Confusion matrix of combined model in validation and test.

Discussion

In this study, we extracted significant radiomics and DL features from post-thrombectomy CT images of HIM and developed a combined model integrating multi-omics features to predict the risk of HT. Our combined model achieved an AUC of 0.887 in the internal prospective cohort, outperforming six other models that utilized individual or combined features from two data types. This study represents an advancement by demonstrating the synergistic effect of combining radiomics, DL, and clinical data in the prediction of post-thrombectomy hemorrhage. Our findings highlight the importance of a multi-faceted approach in improving predictive accuracy, particularly given the critical implications of thrombectomy outcomes.

In this study, we observed that 70.14% of patients exhibiting the HIM subsequently developed HT. This finding underscores the potential of HIM as an independent predictor of HT. This finding underscores the potential of HIM as an independent predictor of hemorrhagic transformation. The primary objective of our research was to identify patients at high risk of bleeding at the earliest possible stage. Early identification enables timely diagnosis and intervention, which can improve patient prognosis[13]. To this end, we developed a predictive model that demonstrates high specificity (85.7%) and a positive predictive value (PPV) of 91.3%. These metrics indicate that the model can reliably and confidently identify individuals with an elevated risk of

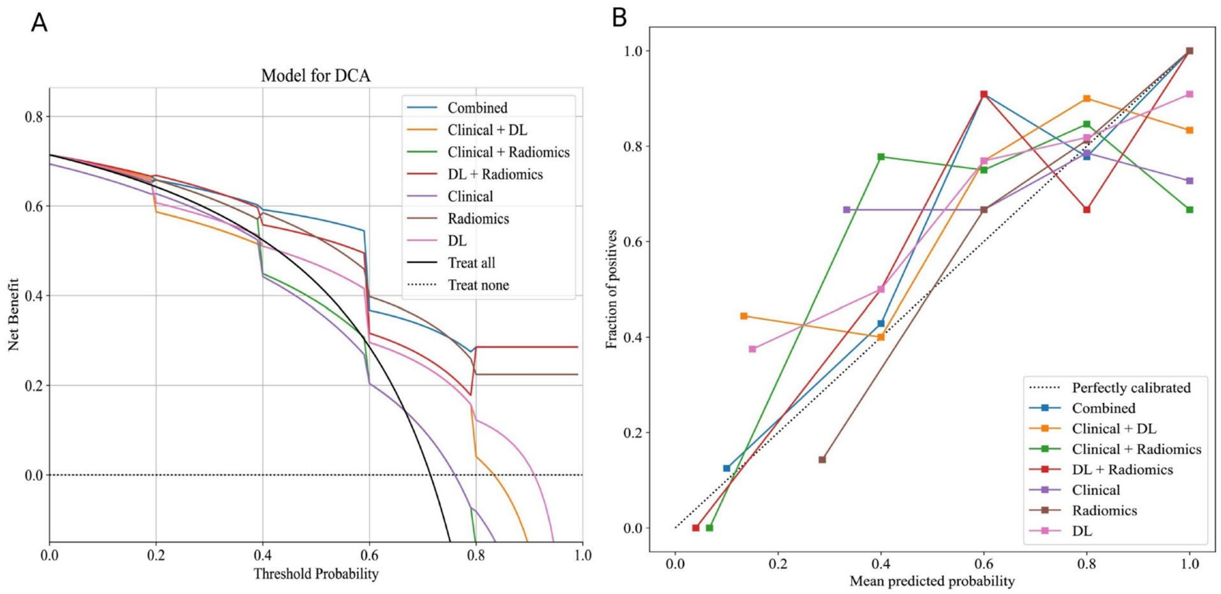

Fig. 8. The DCA curves and calibration curves of all the models in test cohort. (**A**) DCA curves of different models in the test cohort. The y-axis represents the net benefit, and the x-axis represents the threshold probability. (**B**) Calibration curve of different models in the test cohort; the closer the solid line was to the dotted line, the better of the predictive power of the models. DCA, Decision Curve Analysis.

HT. The ability to pinpoint patients at high risk of bleeding is critical for optimizing their clinical management and improving outcomes. For example, in patients identified as high-risk, clinicians might consider delaying antiplatelet therapy to minimize bleeding risk[14]. Additionally, careful monitoring of arterial pressure to prevent hypertensive peaks, the use of sedatives to reduce agitation, and more intensive clinical oversight—such as serial cranial CT scans for selected cases—could be implemented as part of a tailored management strategy[15]. While current therapeutic options for preventing or managing HT remain limited, ongoing research continues to explore promising interventions[16]. This is particularly relevant in the context of expanding indications for thrombectomy. Despite advances in this procedure, the proportion of patients achieving favorable outcomes at 90 days has remained relatively stable[17]. Against this backdrop, the high specificity and positive predictive value of our predictive model offer substantial clinical utility. By accurately identifying a high-risk subgroup, the model supports targeted postoperative interventions. Even in cases where HT is diagnosed more than 24 h after the initial assessment, early therapeutic measures informed by the model may still contribute to improved patient outcomes. However, our model still has some limitations. Its recall rate is only 0.6 on the test set, with a negative predictive value of 0.46. This implies that if the model predicts the absence of hemorrhagic transformation, there is still approximately a 50% probability of its occurrence. Due to this high false-negative rate, clinicians should approach negative predictions cautiously, as the model's ability to reliably rule out HT is limited.

Various indicators are available for predicting post-thrombectomy hemorrhage, with most relying on preoperative CT perfusion assessments[18]. Our study utilized HIM to establish a predictive model, which has been considered a primary indicator for predicting hemorrhage after thrombectomy. HIM offers several advantages for predicting postoperative hemorrhage: It is the earliest examination post-surgery, with most studies utilizing immediate postoperative CT and some even employing intraoperative flat-panel CT scans[19]. Additionally, NCCT is the most accessible examination post-surgery, with CT scanners available in nearly every hospital, making it more broadly applicable than dual-energy CT and MRI[20]. The state of brain tissue is heterogeneously affected by multiple factors from the onset of symptoms to thrombectomy reperfusion, making HIM more consistent in examination results, especially when performed within an hour post-thrombectomy. After ischemia-reperfusion injury, brain tissue undergoes significant pathophysiological changes, which can make preoperative examinations less reflective of the pathophysiological mechanisms occurring during this critical period[21]. The HU in HIM areas primarily reflect the extent of blood-brain barrier disruption[22]. In our research, we have extended the scope to include subarachnoid HIM, which is associated with an increased risk of SAH[23]. Patients experiencing SAH post-thrombectomy face a grave prognosis, which is further compounded by the occurrence of intracerebral hemorrhage[24–26]. The presence of SAH, akin to intracerebral hemorrhage, impacts the timely delivery of postoperative care. Accurate and swift prediction of hemorrhage following thrombectomy is pivotal for the optimal management of patients with AIS. Our study encompassed both intracerebral hemorrhage and SAH as endpoints and demonstrated promising predictive outcomes.

Conventional studies based on HU of HIM to predict HT post-thrombectomy have revealed limited predictive performance[27]. Radiomics based on HIM has demonstrated potential in predicting HT[4,28]; however, previous studies often focused on a smaller range of bleeding, potentially excluding certain instances of HIM or resolved hemorrhage. Our study followed a comprehensive approach by including a wide range of hemorrhage cases within a 90-day window, ensuring that our predictions are relevant to the full spectrum of post-thrombectomy outcomes. This is particularly important for guiding long-term patient management and treatment strategies.

This study is the first to apply machine learning combined with radiomics and DL features extracted from HIM to predict HT. With the advancement of DL, an increasing number of studies are utilizing DL to predict post-thrombectomy hemorrhage, particularly through postoperative CT images[29]. Our study employed DL features of post-thrombectomy HIM, integrating them with other models to enhance predictive performance. The combined model, which incorporated all features, achieved an efficacy of 0.887, demonstrating the best performance. The superior performance of the combined model not only validates our methodological approach but also aligns with the growing body of evidence supporting the utility of multi-modal data integration to improve clinical predictions[30]. The incremental benefits of combining radiomics, DL, and clinical features are evident in our results, highlighting the potential for more personalized and precise patient care.

Our study has several limitations that should be addressed. First, the retrospective nature of the study, its single-center design, and its relatively small sample size may have introduced selection bias and potentially overestimated our findings. Second, the manual segmentation process for lesion identification in our study was time-consuming and complex, especially for lesions with indistinct boundaries. This may have introduced variability and potential bias in the measurements. Third, the small sample size and the imbalance between the two groups in our study may have affected the statistical power and generalizability of the results. Fourth, certain clinically relevant parameters, including antiplatelet/anticoagulant therapy use, tandem occlusions, and various blood pressure metrics, were not systematically collected in the current study, which may limit the comprehensive assessment of predictive factors. We have established protocols to incorporate these variables in future prospective studies. Finally, interpreting the radiomics and DL features used in our study remains a significant challenge, particularly in understanding their underlying biological significance. Addressing these limitations in future research will help strengthen the validity and generalizability of our findings, ultimately enhancing the clinical utility of integrating clinical, radiomics, and DL features for predicting HT following thrombectomy.

Conclusions

In conclusion, our study introduces a predictive model that integrates radiomics, deep learning, and clinical data to enhance the prediction of hemorrhagic transformation (HT) following mechanical thrombectomy (MT). The model exhibits high reliability in identifying patients at risk of HT, thereby aiding in the selection of individuals who may benefit from targeted interventions to improve their prognosis.

Data availability

The datasets used and/or analysed during the current study available from the corresponding author on reasonable request. The source code is available at github website（https://github.com/ChamnLee/Comp2_MCC).

Received: 13 February 2025; Accepted: 12 May 2025

References

1. Morsi, S. et al. Endovascular thrombectomy for DAWN- and DEFUSE-3 ineligible acute ischemic stroke patients: a systematic review and meta-analysis. *J. Neurol.* **271**, 2230–2237 (2024).
2. van der Steen, W. et al. Type of intracranial hemorrhage after endovascular stroke treatment: association with functional outcome. *J. NeuroIntervent Surg.* **15**, 971–976 (2023).
3. Hu, S. et al. An integrated nomogram combining clinical and radiomic features of hyperattenuated imaging markers to predict malignant cerebral edema following endovascular thrombectomy. *Quant. Imaging Med. Surg.* **14**, 4936–4949 (2024).
4. Ma, Y. et al. A CT-based radiomics nomogram for classification of intraparenchymal hyperdense areas in patients with acute ischemic stroke following mechanical thrombectomy treatment. *Front. Neurosci.* **16**, 1061745 (2023).
5. Kim, J-M. et al. The cortical contrast accumulation from brain computed tomography after endovascular treatment predicts symptomatic hemorrhage. *Eur. J. Neurol.* **22**, 1453–1458 (2015).
6. Kau, T. et al. Flat detector Angio-CT following Intra-Arterial therapy of acute ischemic stroke: identification of hemorrhage and distinction from contrast accumulation due to Blood-Brain barrier disruption. *AJNR Am. J. Neuroradiol.* **35**, 1759–1764 (2014).
7. Maïer, B. et al. Association of contrast enhancement after reperfusion with outcomes according to blood pressure Lowering in acute ischemic stroke patients. *Neurology* https://doi.org/10.1212/WNL.0000000000201173 (2022).
8. Pinckaers, F. M. E. et al. Prognostic implications of intracranial haemorrhage on dual-energy CT immediately following endovascular treatment for acute ischemic stroke. *J. Neuroradiol.* (2023). S0150986123002651.
9. Porto-Álvarez, J. et al. How can radiomics help the clinical management of patients with acute ischemic. *Stroke? Appl. Sci.* **13**, 10061 (2023).
10. Li, Y. et al. Molecular subtyping of diffuse gliomas using magnetic resonance imaging: comparison and correlation between radiomics and deep learning. *Eur. Radiol.* **32**, 747–758 (2022).
11. Liu, W. et al. Development and validation of Multi-Omics thymoma risk classification model based on transfer learning. *J. Digit. Imaging* 1–10 (2023).
12. Dekeyzer, S. et al. Distinction between contrast staining and hemorrhage after endovascular stroke treatment: one CT is not enough. *J. NeuroIntervent Surg.* **9**, 394–398 (2017).
13. Lou, Y. et al. Efficacy and safety of very early rehabilitation for acute ischemic stroke: a systematic review and meta-analysis. *Front. Neurol.* **15**, 1423517 (2024).
14. Amaro, S. et al. Risks and benefits of early antithrombotic therapy after thrombolytic treatment in patients with acute stroke. *PLoS ONE.* **8** (8), e71132 (2013).
15. Kovács, K. B. et al. Hemorrhagic transformation of ischemic strokes. *Int. J. Mol.* **24** (18), 14067 (2023).
16. Mazighi, M. et al. Safety and efficacy of platelet glycoprotein VI Inhibition in acute ischaemic stroke (ACTIMIS): a randomised, double-blind, placebo-controlled, phase 1b/2a trial. *Lancet Neurol.* **23** (2), 157–167 (2024).

17. Oliveira, A. J. F., Viana, S. M. N. & Santos, A. S. Mechanical thromb ectomy for acute ischemic stroke: systematic review and meta-analysis. *Einstein (Sao Paulo)*. **20**, eRW6642 (2022).
18. Heo, J. & Sohn, B. Prediction of hemorrhagic transformation in acute ischemic stroke: a never-ending endeavor. *Eur. Radiol.* **34**, 5305–5307 (2024).
19. Knott, M. et al. Can flat-detector CT after successful endovascular treatment predict long-term outcome in patients with large vessel occlusion? An Alberta stroke programme early CT Score–based study. *Neurol. Sci.* **44**, 1193–1200 (2023).
20. Yedavalli, V. & Sammet, S. Contrast extravasation versus hemorrhage after thrombectomy in patients with acute stroke. *J. Neuroimaging*. **27**, 570–576 (2017).
21. Mathias, K. et al. Blood-brain barrier permeability in the ischemic stroke: an update. *Microvasc. Res.* **151**, 104621 (2024).
22. Kim, J-T. et al. Hyperdensity on non-contrast CT immediately after intra-arterial revascularization. *J. Neurol.* **259**, 936–943 (2012).
23. Ogata, A. et al. Subarachnoid iodine leakage on dual-energy computed tomography after mechanical thrombectomy is associated with malignant brain edema. (2024). J NeuroIntervent Surg jnis-2023-021413.
24. Lee, H. et al. Subarachnoid hemorrhage in mechanical thrombectomy for acute ischemic stroke: analysis of the STRATIS registry, systematic review, and Meta-Analysis. (2021). Front Neurol 12.
25. Qureshi, A. I., Saleem, M. A. & Aytac, E. Postprocedure subarachnoid hemorrhage after endovascular treatment for acute ischemic stroke. *J. Neuroimaging*. **27**, 493–498 (2017).
26. Kim, D. Y. et al. Predictors and impact of sulcal SAH after mechanical thrombectomy in patients with isolated M2 occlusion. (2022). AJNR Am J Neuroradiol.
27. Payabvash, S. et al. Middle cerebral artery residual contrast stagnation on Noncontrast CT scan following endovascular treatment in acute ischemic stroke patients. *J. Neuroimaging*. **25**, 946–951 (2015).
28. Chen, X. et al. CT-based radiomics for differentiating intracranial contrast extravasation from intraparenchymal haemorrhage after mechanical thrombectomy. *Eur. Radiol.* **32**, 4771–4779 (2022).
29. Heo, J. et al. *Prediction of Cerebral Hemorrhagic Transformation after Thrombectomy Using a Deep Learning of dual-energy CT* (Eur Radiol, 2023).
30. Jan, Y-T. et al. Machine learning combined with radiomics and deep learning features extracted from CT images: a novel AI model to distinguish benign from malignant ovarian tumors. *Insights Imaging*. **14**, 68 (2023).

Author contributions

L. J. and G.Z. wrote the main manuscript text and prepared figures. Y.W., J.H., J.F. and J.C. curated the data and prepared tables. J.H., S.X., H.S. and W.X. supervised and provided necessary resources. J.F. and S.H. provided fundings. All authors reviewed the manuscript.

Declarations

Competing interests

The authors declare no competing interests.

Additional information

Correspondence and requests for materials should be addressed to S.H. or W.X.

Reprints and permissions information is available at www.nature.com/reprints.