



OPEN ACCESS

EDITED BY

Andrea Bianconi,
University of Genoa, Italy

REVIEWED BY

Adria Arboix,
Sacred Heart University Hospital, Spain
Federico Iaccarino,
University of Modena and Reggio Emilia, Italy

*CORRESPONDENCE

Jingjing Fu
✉ fujingjing1985@zju.edu.cn

RECEIVED 30 September 2025

REVISED 23 October 2025

ACCEPTED 17 November 2025

PUBLISHED 27 November 2025

CITATION

Hong J, Fu J, Liu F, Chen Y, Shen Y, Li Y, Hu S and Fu J (2025) Development of a multimodal model combining radiomics and deep learning to predict malignant cerebral edema after endovascular thrombectomy. *Front. Neurol.* 16:1716443. doi: 10.3389/fneur.2025.1716443

COPYRIGHT

© 2025 Hong, Fu, Liu, Chen, Shen, Li, Hu and Fu. This is an open-access article distributed under the terms of the [Creative Commons Attribution License \(CC BY\)](https://creativecommons.org/licenses/by/4.0/). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

Development of a multimodal model combining radiomics and deep learning to predict malignant cerebral edema after endovascular thrombectomy

Jiayi Hong¹, Jiahong Fu¹, Feifan Liu¹, Yuhan Chen¹, Yujie Shen¹, Yan Li¹, Sheng Hu² and Jingjing Fu^{1*}

¹Department of Neurology, The Fourth Affiliated Hospital of School of Medicine, and International School of Medicine, International Institutes of Medicine, Zhejiang University, Yiwu, Zhejiang, China,

²Department of Radiology, The Fourth Affiliated Hospital of School of Medicine, and International School of Medicine, International Institutes of Medicine, Zhejiang University, Yiwu, Zhejiang, China

Background: Malignant cerebral edema (MCE) represents a severe complication after endovascular thrombectomy (EVT) in treating acute ischemic stroke. This study aimed to develop and validate a multimodal predictive model integrating clinical data, radiomics features, and deep learning (DL)-derived features to improve the accuracy of MCE risk prediction following EVT.

Methods: A total of 290 patients were included, comprising 189 in the training, 47 in the validation, and 54 in the internal test cohorts. A fusion model was developed by integrating clinical variables, radiomics, and DL features. Separate models based on clinical data, radiomics, and DL features were also constructed for comparison. Model training and evaluation were performed on training, validation, and test cohorts. The predictive performance of the combined model was compared with the ACORNS grading scale using an area under the curve (AUC) analysis to assess clinical effectiveness.

Results: The combined model exhibited the best predictive performance. Analysis of the receiver operating characteristic curve revealed an AUC of 0.927 [95% confidence interval (CI): 0.849–1.000] for predicting MCE in the validation group and an AUC of 0.924 (95% CI: 0.846–1.000) in the test group. Additionally, the fusion model consistently demonstrated higher net benefits across all threshold probabilities than the ACORNS grading scale.

Conclusions: This study integrated clinical data, radiomics, and DL features to develop a multimodal predictive model with a strong discriminative ability to predict MCE after EVT.

KEYWORDS

thrombectomy, acute ischemic stroke, deep learning, radiomics, multimodal fusion

1 Introduction

Endovascular thrombectomy (EVT) has proven to be an effective intervention for acute ischemic stroke (AIS) resulting from large vessel occlusion in the anterior circulation (1). However, this treatment carries significant risks, particularly the development of malignant cerebral edema (MCE), which can undermine the clinical benefits of the procedure and may lead to death (2). Current evidence indicates that early decompressive craniectomy considerably reduces mortality rates (2–4). Consequently, developing reliable methods for

the early detection and precise prediction of post-EVT MCE is essential for optimizing treatment strategies and improving clinical outcomes (5).

Existing predictive models for MCE following EVT primarily include the ACORNS grading scale (6) and clinical nomograms (7–9), all of which have demonstrated favorable outcomes. It should be noted that in a hospital-based clinical series, decreased level of consciousness, nausea or vomiting, and heavy smoking were identified as the main predictive clinical factors associated with malignant middle cerebral artery infarction (10). However, many existing models often depend on parameters susceptible to subjective interpretation and technical variability, undermining their predictive accuracy and clinical reliability. Recent studies have emphasized the significance of hyperattenuated imaging markers (HIM) as objective imaging biomarkers in predicting MCE after EVT (11, 12). HIM refers to areas of abnormal hyperdensity observed on noncontrast head computed tomography (NCCT) immediately post-EVT, detectable in up to 87% of patients' CT images (13). Emerging evidence suggests this radiological phenomenon may indicate blood-brain barrier disruption following reperfusion injury (14). Previous studies have confirmed that the extent and severity of HIM are effective predictors of MCE after EVT in patients with a stroke (11, 12).

Artificial intelligence (AI) technology has recently made significant advancements in predicting MCE following EVT in patients with AIS (14–16). A Previous study (17) reported strong predictive performance (AUC = 0.879 in the test set) using a combined CT-based radiomics and clinical model. Notably, multimodal fusion strategies integrating clinical data, radiomics, and deep learning (DL) techniques have demonstrated considerable success in oncology. Studies reported that multimodal models demonstrate significantly higher accuracy than single-modality models in differentiating ovarian tumors from thymomas (18, 19). Therefore, we hypothesized that multimodal fusion strategies may enhance the precision of predictive systems for MCE after EVT.

This study developed a multimodal model that integrates clinical data, radiomics features, and DL techniques to improve the accuracy of MCE risk prediction following EVT.

2 Materials and methods

2.1 Ethical approval of the study protocol

The study received approval from the Ethics Committee of the Fourth Affiliated Hospital of the School of Medicine and International School of Medicine, International Institutes of Medicine, Zhejiang University (Approval Number: K2024139). The committee exempted informed consent. All clinical studies complied with the principles of the Declaration of Helsinki.

2.2 Patient cohort

A dataset of post-EVT cases for large-vessel occlusive stroke in the anterior circulation was used, collected from the Fourth Affiliated Hospital of the School of Medicine and the International School of Medicine, International Institutes of Medicine, Zhejiang

University, between September 2016 and March 2023. The model was validated in a test cohort from April 2023 to January 2024 at the same hospital to evaluate clinical applicability and performance with real-world data.

We enrolled patients who (1) had a diagnosis of AIS confirmed by diffusion-weighted imaging or CT; (2) received EVT as treatment; (3) underwent a head NCCT within 1 h following EVT; (4) demonstrated HIM, characterized by increased attenuation in the brain parenchyma and/or subarachnoid space. The exclusion criteria were as follows: (1) Follow-up NCCT performed outside the 2–5 day window after EVT, (2) artifacts, such as metal or motion artifacts, that interfered with the assessment of HIM or MCE, and (3) patients with posterior circulation stroke. Figure 1 presents a flowchart detailing participant inclusion in this study.

2.3 Clinical data

Data on general clinical features, laboratory examinations, clinical presentations were analyzed. Patient charts, procedure notes, and follow-up records were reviewed to extract baseline and preprocedural factors, including age, sex, history of alcohol consumption or smoking, hypertension, hyperlipidemia, diabetes mellitus, atrial fibrillation, coronary artery disease, and baseline National Institute of Health Stroke Scale (NIHSS) score. Additionally, modified thrombolysis in cerebral infarction scores and various pre-interventional parameters, including intravenous thrombolysis therapy, were collected. Interventional parameters were also recorded, such as the percentage of stent implantation, stent pass number, and type of stent used.

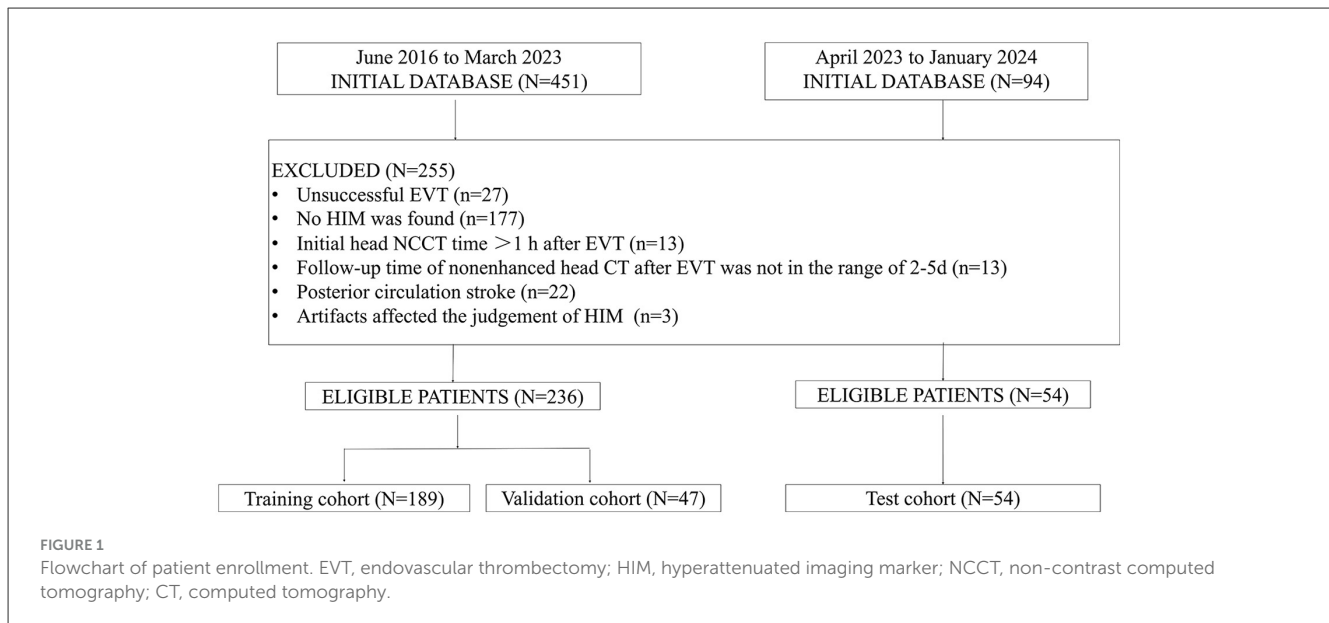
2.4 Data preprocessing

A total of 290 patients were enrolled in this study. Of these, 236 patients were randomly assigned to the training cohort (189 cases) and the validation cohort (47 cases). An additional 54 patients were included in the internal test cohort.

Images of the NCCT were obtained using a 64-row spiral CT scanner (SOMATOM Definition AS, Siemens Healthineers, Erlangen, Germany) and a 62-row spiral scanner (Optima CT620, GE HealthCare, Chicago, IL, USA). The scanning parameters included axial mode, tube voltage of 120 kV, tube current of 250–300 mAs, scanning range from the skull base to the cranial roof, section thickness of 5 mm, and reconstruction using the standard algorithm. Imaging data were evaluated independently by two trained junior neuroradiologists blinded to clinical outcome and demographic data. All imaging data were evaluated with subsequent consensus reading of one senior neuroradiologist and analyzed over a consecutive period of 12 weeks.

2.5 Image analysis

Neuroradiologists independently assessed the Alberta Stroke Program Early CT Score (ASPECTS) on baseline NCCT images. HIM were evaluated on NCCT images within 1 h after EVT,



characterized by increased attenuation in the brain parenchyma and/or subarachnoid space. The neuroradiologists analyzed the maximum Hounsfield unit (HU_{max}) and the volume of the HIM within the region of interest (ROI). The HU_{max} was defined as the highest HU value within a specified ROI, measuring 3×3 pixels.

2.6 Outcome measures

A review of NCCT images obtained after EVT was conducted, with the development of MCE was defined as the occurrence of midline brain shift, specifically a displacement of the septum pellucidum of ≥ 5 mm, within the first 5 days after admission. Two neuroradiologists independently evaluated all datasets without knowledge of patient outcomes. Any discrepancies were resolved by consensus.

2.7 Feature extraction and selection

Radiomic features were extracted from HIM in NCCT using the Pyradiomics in-house feature analysis program (<http://pyradiomics.readthedocs.io>). A total of 1,834 radiomic features were extracted, including shape features, first-order features, gray co-occurrence matrix features, gray region size matrix features, gray stroke matrix features, neighborhood gray difference matrix features, gray correlation matrix features, and wavelet features.

DL features based on the HIM region of interest (ROI) were extracted from five commonly used DL models for modeling, including ResNet152, Inception v3, ResNet50, Visual Geometry Group (VGG) 16, and VGG19. We compared and selected the best DL model and used it as a pre-trained model. Data augmentation techniques, such as RandomResizedCrop and RandomHorizontalFlip, were applied to enhance the model's generalization ability. DL features were extracted from CT

images and selected after dimensionality reduction via principal component analysis (PCA).

2.8 Development and validation of the model

Ultimately, the multi-omics features consisted of clinical data, radiomics, and DL features. These features were standardized to have a mean of 0 and a variance of 1. Feature selection was performed using analysis of variance (ANOVA), Pearson's correlation coefficient, PCA, and Least Absolute Shrinkage and Selection Operator (LASSO) to improve the model's generalization ability and minimize overfitting. For clinical, radiomics, and DL models, ANOVA, Pearson's correlation coefficient, and LASSO were sequentially applied for feature selection. PCA, ANOVA, Pearson's correlation coefficient, and LASSO were sequentially utilized for the combined model. Additionally, individual models based on clinical data, radiomics, and DL features were also constructed for comparison. MCE risk prediction was performed using seven different classifiers. The optimal classifier was selected based on predictive performance in validation and test cohorts. The performance of these models was evaluated using accuracy, sensitivity, specificity, receiver operating characteristic (ROC) curves, area under the curve (AUC), and decision curve analysis (DCA). Calibration curves were plotted to assess the agreement between the predicted and actual classifications, thereby evaluating the calibration efficiency of each model.

2.9 Comparison of the model with the ACORNS grading scale

Huang et al. (6) developed the ACORNS grading score based on clinical predictors, including hypertension history, baseline

NIHSS score, ASPECT score, glycemia, collateral circulation status, occlusion location, thrombolysis prior to thrombectomy, and reperfusion status. This score demonstrated strong predictive performance. The same variables were collected to compare with the AI model, and an external validation of the ACORNS grading score was conducted. The workflow and global analysis pipeline for the classification model are presented in [Figure 2](#).

2.10 Statistical analysis

Statistical analysis was conducted using the Statistical Package for the Social Sciences software (version 26.0, IBM Corp., Armonk, NY, USA). The normality of the evaluated variables was assessed using the Kolmogorov–Smirnov and Shapiro–Wilk tests. Continuous variables are presented as mean \pm standard deviation, while categorical variables are reported as frequency counts and percentages. Categorical variables were compared using the chi-square or Fisher's exact test, and continuous variables were compared using the Mann–Whitney *U*-test or independent *t*-test. Feature extraction, screening, and model construction were performed using Python software (version 3.11.1, Python Software Foundation, Wilmington, DE, USA). Model performance was evaluated using AUC with 95% confidence interval (CI), accuracy, sensitivity, and specificity. A *p*-value of <0.05 was considered statistically significant for all tests.

3 Results

3.1 Patient baseline characteristics

This study included 290 patients, with 189 in the training, 47 in the validation, and 54 in the test cohorts. [Table 1](#) summarizes the baseline demographic, clinical, and imaging characteristics of the patients. Significant differences were observed between the training, validation, and test cohorts regarding thrombolysis ($p = 0.006$) and aspiration catheter ($p = 0.000$). MCE occurred in 58 of the 290 patients. The overall in-hospital mortality was 6.9% (20/290). The hospital mortality rate was significantly higher in the MCE subgroup (24.1%, 14/58) than in the non-MCE subgroup (2.6%, 6/232; Fisher's exact test, $p < 0.001$). The presence of MCE was associated with an 11.83-fold increase in the odds of death (OR = 11.83, 95% CI: 4.01–39.71).

3.2 Performance of different models

The fusion model was conducted using the ResNet152 model and the MLP classifier, which demonstrated the highest performance across the training, validation, and test cohorts. In the training cohort, the fusion model achieved an AUC of 0.978 (95% CI: 0.961–0.995), with a sensitivity of 0.950 and a specificity of 0.872. In the validation cohort, the fusion model achieved an AUC of 0.927 (95% CI: 0.849–1.000), with a sensitivity of 0.700 and a specificity of 0.919. In the test cohort, the fusion model achieved an AUC of 0.924 (95% CI: 0.846–1.000), with a sensitivity of 0.750 and

a specificity of 0.913. Predictive performances of each classifier in test cohort are provided in [Table 2](#).

The predictive performance was also compared across clinical, radiomics, and DL models in the training, validation, and test cohorts. In the test cohort, AUC values were as follows: 0.924 for the fusion model (incorporating radiomics, DL, and clinical data), 0.721 for the clinic model, 0.902 for the radiomics model, and 0.872 for the DL model. Detailed model parameters in test cohort are provided in [Table 3](#). [Figure 3](#) presents a comparison of ROC curves and AUC values of the different models in the training, validation, and test cohorts. [Figure 4](#) illustrates DCA and calibration curves for various models in the test cohort. The calibration curves revealed strong agreement between the model predictions and the actual MCE classifications in the test cohort, with the fusion model demonstrating the best fit. The DCA further highlighted that all models notably improved patient outcome predictions compared to scenarios without any prediction model. The fusion model provided the most substantial clinical benefit for automatic MCE classification.

3.3 Comparison of the fusion model with the ACORNS grading scale

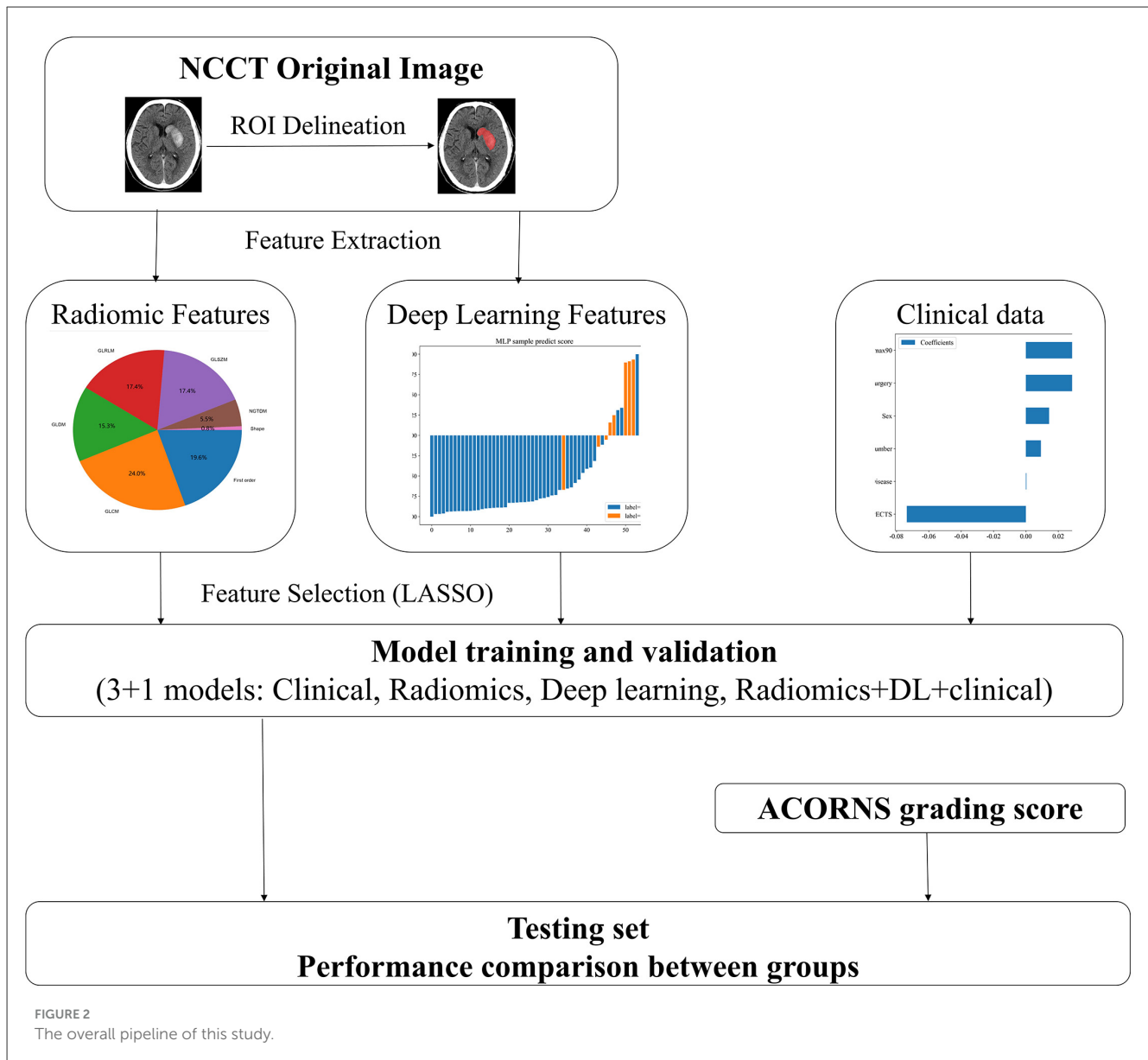
The ACORNS grading scale's clinical utility and discriminatory performance were evaluated and compared with the fusion model using DCA and ROC curves. In the test cohort, the AUC for the ACORNS grading scale was 0.641 (95% CI: 0.447–0.836). Additionally, the clinical effectiveness of the ACORNS grading scale and the combined model were compared using a DCA curve, revealing that the fusion model consistently provided higher net benefits across all threshold probabilities than the ACORNS grading scale. [Figure 5](#) depicts a comparison of AUC values, DCA curves, and calibration curves for the two models in the test cohort.

4 Discussion

This study developed a fusion AI model based on the HIM framework, integrating clinical data, radiomic signatures, and DL-derived features using an ensemble machine learning classifier. Internal validation demonstrated strong discriminative performance, achieving AUC values of 0.978 in the training cohorts and 0.927 in the validation cohort. A subsequent 10-month evaluation confirmed improved predictive capability (AUC = 0.924), surpassing the conventional ACORNS grading scale (AUC = 0.641) in prognostic accuracy and clinical relevance.

In this study, comprehensive screening identified ASPECTS score and H_Umax ≥ 90 as the two clinical features most strongly associated with MCE, aligning with previous research findings (20, 21). The ASPECTS score quantifies early ischemic changes on CT imaging and serves as an indicator of stroke severity; lower scores indicate a higher risk of MCE (20). Additionally, H_Umax ≥ 90 reflects significant blood-brain barrier disruption (21), which is associated with a higher risk of MCE.

Radiomics and DL enable quantitative extraction of subvisual pathological information through advanced image feature analysis compared to subjective radiological assessments (19).



Extensive evidence supports the prognostic value of radiomics in predicting cerebral edema following EVT, with Wen et al.’s (17) clinical-radiomics model achieving an AUC of 0.879 for MCE detection. A subsequently developed HIM-incorporated nomogram (14) demonstrated enhanced predictive performance. However, the limited robustness of radiomic features to spatial transformation restricts clinical generalizability, while DL architectures provide automated global feature representation, highlighting methodological complementarity. Moreover, DL features may exhibit sensitivity to global translation, rotation, and scaling of images, whereas radiomics features remain unaffected by such variations (22). Transfer learning has emerged as an effective strategy to improve cross-domain feature transferability and address the challenge of limited medical imaging data scarcity (23). A previous study (24) employed ResNeXt-101 to analyze HIM on NCCT following EVT, yielding promising results in predicting MCE. These findings demonstrated that HIM provides a robust

dataset for AI applications and validated the technical feasibility of predicting post-EVT MCE by integrating radiomic features with DL architectures. This evidence establishes a critical foundation for developing multimodal predictive systems in neurocritical care.

Radiomics, DL features, and clinical data capture distinct pathophysiological characteristics of MCE, leading to the hypothesis that multimodal fusion strategies can enhance predictive accuracy for post-EVT MCE. Successful cross-domain applications support this theoretical framework; Jan et al. (19) developed a CT-based paradigm that integrated radiomics and DL features to differentiate benign from malignant ovarian tumors, while Wang et al. (22) combined DL, radiomics, and clinical data to predict lymph node metastasis in laryngeal squamous cell carcinoma. All these studies achieved favorable results. Despite these advancements, the synergistic integration of radiomics and DL for post-EVT MCE prediction remains unexplored. This study constructed independent models based on clinical data, radiomics,

TABLE 1 Characteristics of the training, validation and test sets.

Feature name	Training	Validation	Test	<i>p</i> -value
Age (years), mean \pm SD	66.20 \pm 15.00	70.62 \pm 11.14	66.17 \pm 18.04	0.185
Males, <i>n</i> (%)	119 (62.96)	29 (61.70)	37 (68.52)	0.718
Atrial fibrillation, <i>n</i> (%)	75 (39.68)	23 (48.94)	23 (42.59)	0.513
Hypertension, <i>n</i> (%)	111 (58.73)	28 (59.57)	31 (57.41)	0.975
Hyperlipidemia, <i>n</i> (%)	15 (7.94)	4 (8.51)	9 (16.67)	0.154
Diabetes mellitus, <i>n</i> (%)	28 (14.81)	7 (14.89)	9 (16.67)	0.945
Coronary artery disease, <i>n</i> (%)	21 (11.11)	8 (17.02)	3 (5.56)	0.187
Drinking, <i>n</i> (%)	34 (18.00)	13 (27.66)	9 (16.67)	0.281
Smoking, <i>n</i> (%)	43 (22.75)	8 (17.02)	12 (22.22)	0.694
Thrombolysis, <i>n</i> (%)	70 (37.04)	21 (44.68)	9 (16.67)	0.006*
Baseline NIHSS, median (Q1, Q3)	15 (12–19)	18 (15–21)	15 (10–20)	0.084
ASPECTS, median (Q1, Q3)	9 (8–10)	9 (8–10)	9 (9–10)	0.280
HUmax \geq 90, <i>n</i> (%)	54 (28.57)	15 (31.91)	16 (29.63)	0.903
mTICI beyond 2b, <i>n</i> (%)	173 (91.53)	42 (89.36)	51 (94.44)	0.646
Stent type, <i>n</i> (%)				0.526
Solitaire	111 (58.73)	28 (59.57)	32 (59.26)	
Trevo	24 (12.70)	11 (23.40)	0 (0.00)	
Solitaire+trevo	32 (16.93)	3 (6.38)	16 (29.63)	
Others	22 (11.64)	5 (10.64)	6 (11.11)	
Pass number, median (Q1, Q3)	2 (1–3)	2 (1–3)	2 (0–2)	0.364
Aspiration catheter, <i>n</i> (%)	71 (37.57)	9 (19.15)	45 (83.33)	0.000*
Extravasation during surgery, <i>n</i> (%)	10 (5.29)	1 (2.13)	3 (5.56)	0.641
Stent implantation, <i>n</i> (%)	36 (19.05)	10 (21.28)	7 (12.96)	0.505

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; mTICI, modified thrombolysis in cerebral infarction. **P* < 0.05.

and DL features. The multimodal fusion model demonstrated superior performance, achieving stable accuracy rates of 0.927 in the validation and 0.924 in the test cohorts, significantly surpassing unimodal approaches. This enhanced predictive capability supports the hypothesis that multimodal integration of clinical, radiomic, and DL biomarkers provides complementary insights for MCE, resulting in a more effective synergistic effect for predicting post-EVT MCE.

The ACORNS grading scale (6), which incorporates eight features, including the NIHSS score and the ASPECTS, predicts the risk of MCE in patients with stroke and large vessel occlusion following EVT. It demonstrated robust discriminative performance, with AUC values of 0.850, 0.874, and 0.785 in the training, internal validation, and external validation cohorts. This study further evaluated the predictive performance of the ACORNS grading scale through an external validation cohort (AUC = 0.641), reaffirming its predictive capability for MCE. However, the proposed model integrates clinical data, radiomic signatures, and DL features. These features can quantify disease characteristics and detect subclinical alterations undetectable by conventional clinical assessments, resulting in superior predictive performance.

An interesting yet contrasting finding of our study was the identification of male gender as an independent risk factor for MCE, which diverges from some previous reports, including a meta-analysis suggesting female gender as an independent risk factor (25). This discrepancy may be attributable to several factors specific to our study. Firstly, the high prevalence of risk factors among male stroke patients in China, such as smoking, alcohol consumption (26), may contribute to a greater burden of pre-existing cerebrovascular disease, potentially increasing susceptibility to MCE. Secondly, potential biases in our study must be considered. The predominance of males in our single-center cohort (63.79%, 185/290), which is susceptible to referral bias as a comprehensive stroke center, may reflect a population with more severe strokes, influencing the observed association. Furthermore, our cohort was exclusively composed of patients exhibiting HIM on post-EVT CT scans, which may introduce selection bias, limiting the generalizability of our gender-specific finding to the broader EVT population. Finally, despite multivariate adjustment, residual confounding by unmeasured variables, such as collateral status, genetic factors, cannot be ruled out. Therefore, the relationship between gender and MCE risk

TABLE 2 Predictive performance of each machine learning model in test cohort.

Model name	AUC (95% CI)	Sensitivity*	Specificity*	Accuracy	PPV	NPV
NaiveBayes	0.887 (0.802–0.973)	0.875	0.826	0.833	0.467	0.974
SVM	0.902 (0.809–0.996)	0.75	0.826	0.815	0.429	0.95
RandomForest	0.908 (0.814–1.000)	0.75	0.783	0.778	0.375	0.947
ExtraTrees	0.864 (0.741–0.988)	0.625	0.891	0.852	0.5	0.932
XGBoost	0.856 (0.698–1.000)	0.625	0.913	0.87	0.556	0.933
LightGBM	0.838 (0.636–1.000)	0.5	0.978	0.907	0.8	0.918
MLP	0.924 (0.846–1.000)	0.75	0.913	0.889	0.6	0.955

*Balanced sensitivity and specificity at the cutoff yielding the largest Youden index value.

AUC, area under the receiver operating characteristic curve; CI, confidence interval; PPV, positive predictive value; NPV, negative predictive value.

TABLE 3 Predictive performance of each model in test cohort.

Model name	AUC (95% CI)	Sensitivity*	Specificity*	Accuracy	PPV	NPV
Clinical model	0.721 (0.508–0.935)	0.750	0.652	0.667	0.273	0.937
Radiomics model	0.902 (0.817–0.987)	0.750	0.870	0.852	0.500	0.952
Deep learning model	0.872 (0.755–0.990)	0.750	0.848	0.833	0.462	0.951
Fusion model	0.924 (0.846–1.000)	0.750	0.913	0.889	0.600	0.955

*Balanced sensitivity and specificity at the cutoff yielding the largest Youden index value.

AUC, area under the receiver operating characteristic curve; CI, confidence interval; PPV, positive predictive value; NPV, negative predictive value.

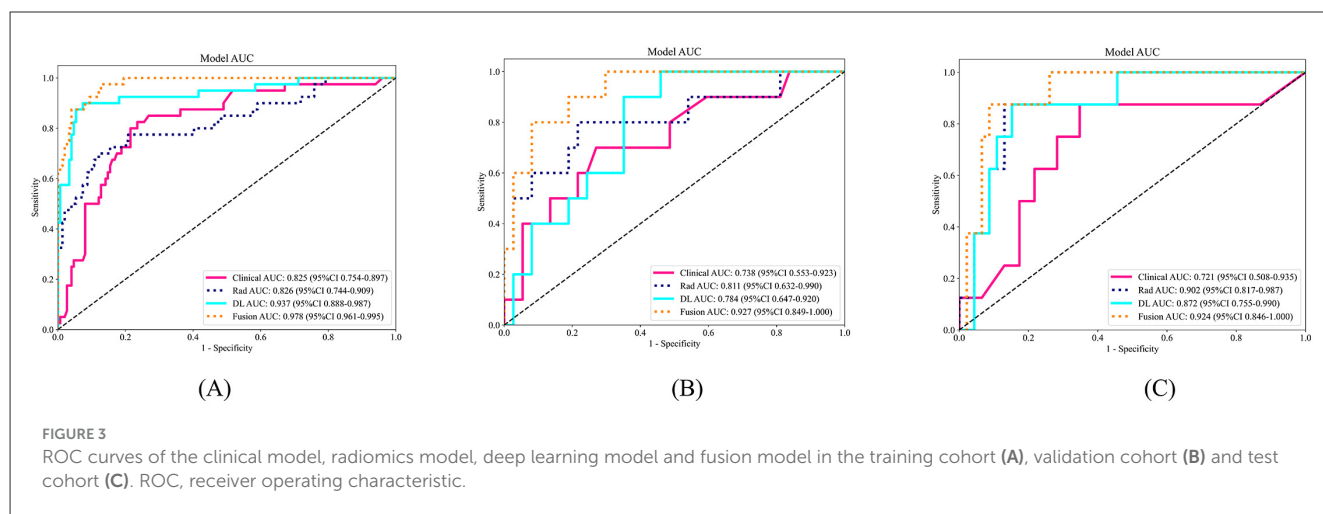


FIGURE 3

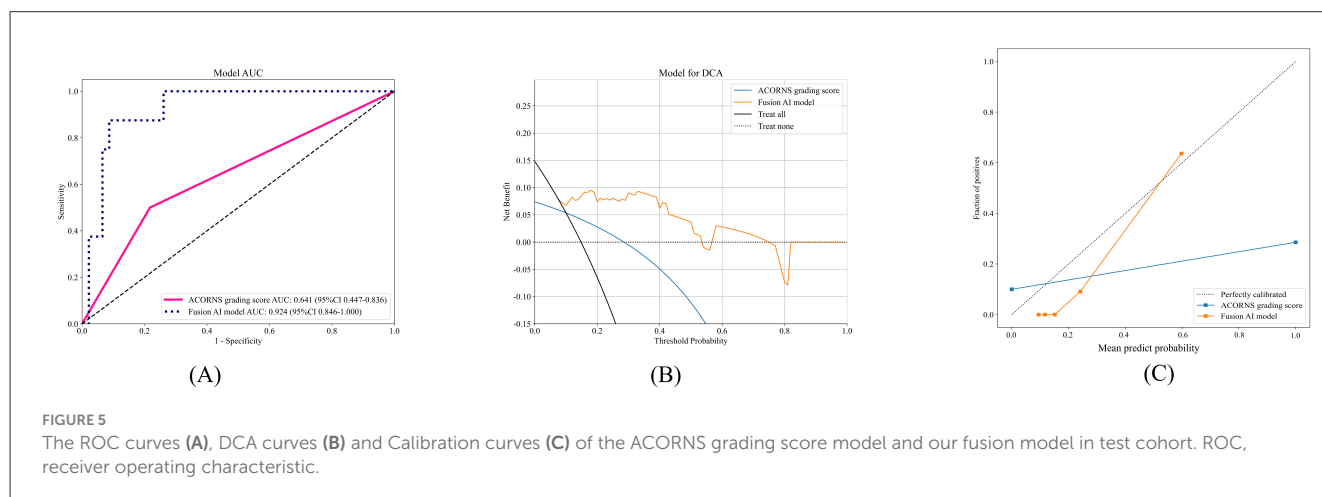
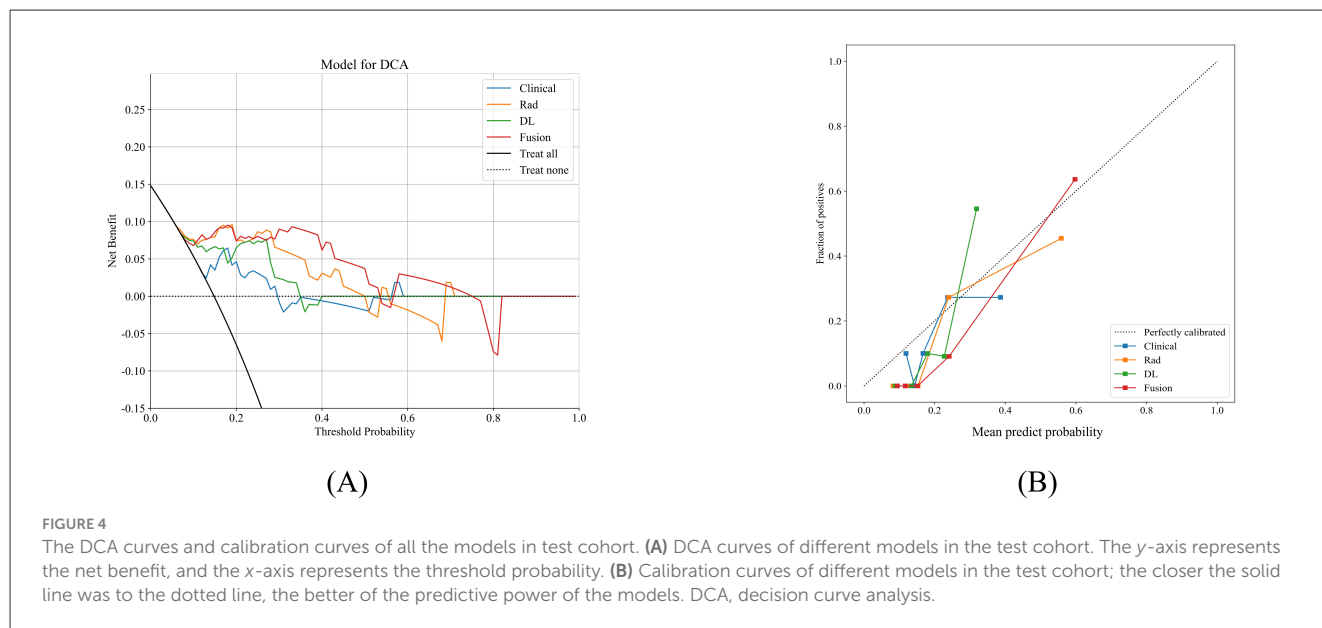
ROC curves of the clinical model, radiomics model, deep learning model and fusion model in the training cohort (A), validation cohort (B) and test cohort (C). ROC, receiver operating characteristic.

warrants further investigation in larger, prospective, and multi-center studies.

This study has several limitations that should be acknowledged. First, our study cohort exclusively included patients who exhibited HIM on post-EVT NCCT. While this selection criterion was essential for model development targeting this specific imaging phenotype, it inevitably introduces a potential selection bias. Consequently, the applicability of our proposed model is currently restricted to this patient subgroup, and its predictive performance in the broader, unselected EVT population remains uncertain and warrants future investigation. Second, the single-center design with a modest sample size ($n = 290$) and the lack of external validation cohorts may limit the model's generalizability and increase the risk of overfitting. Third, an important technical limitation pertains to the CT data acquisition. The scans were acquired using scanners

from different vendors (Siemens and GE) with heterogeneous acquisition and reconstruction parameters. As no harmonization techniques (e.g., ComBat) or intensity normalization procedures were applied, the inherent scanner-related variability could have influenced the stability of extracted radiomic features, potentially affecting the model's performance and robustness. Fourth, the ROI for the HIM was delineated manually, which was time-consuming and may introduce interobserver variability. Fifth, although the strategy of our feature reduction pipeline may improve internal predictive performance, it introduces complexity that raises concerns about reproducibility. Finally, the interpretability of the model is limited, as it is unclear which features contribute most to prediction.

Based on the findings and limitations identified in this study, several strategic directions for future research are



recommended. First, external validation through large-scale, prospective, multicenter cohorts is essential to verify the model's generalizability and robustness. Second, the clinical workflow could be streamlined by developing deep learning-based algorithms for the fully automated segmentation of HIM along with using more advanced approaches such as habitat analysis to enable finer-grained examination of lesion characteristics. In parallel, the interpretability of the model must be improved using techniques, such as SHAP (SHapley Additive exPlanations) and Grad-CAM (Gradient-weighted Class Activation Mapping), to clarify the contribution of key features. Lastly, future work could focus on constructing dynamic prediction models that integrate serial imaging data, such as follow-up CT or MRI, to monitor edema progression in real time, which may facilitate preemptive clinical interventions.

In conclusion, this study introduces a novel and comprehensive approach that integrates clinical, radiomic,

and DL models, offering a rapid and effective method to predict MCE after EVT. With its robust predictive capability, the fusion model has significant potential as a practical tool for clinicians.

Data availability statement

The datasets presented in this article are not readily available because all data are required for further research. Requests to access the datasets should be directed to Jingjing Fu, fujingjing1985@zju.edu.cn.

Ethics statement

The studies involving humans were approved by the Ethics Committee of the Fourth Affiliated Hospital of School of

Medicine, and International School of Medicine, International Institutes of Medicine, Zhejiang University (Approval number: K2024139). The studies were conducted in accordance with the local legislation and institutional requirements. Written informed consent from the patients/participants or patients/participants' legal guardian/next of kin was not required to participate in this study in accordance with the national legislation and the institutional requirements.

Author contributions

JH: Methodology, Validation, Writing – original draft, Writing – review & editing. JiaF: Data curation, Validation, Writing – review & editing. FL: Data curation, Methodology, Writing – review & editing. YC: Data curation, Formal analysis, Writing – review & editing. YS: Data curation, Validation, Writing – review & editing. YL: Data curation, Writing – original draft. SH: Conceptualization, Funding acquisition, Methodology, Writing – review & editing. JinF: Conceptualization, Funding acquisition, Supervision, Writing – review & editing.

Funding

The author(s) declare that financial support was received for the research and/or publication of this article. This work was supported by Medical Scientific Research Foundation of Zhejiang Province, China, (Grant number: 2025KY090); the Fundamental Research Funds for the Central Universities (Grant number: 226202400226).

References

- Sahoo A, Abdalkader M, Yamagami H, Huo X, Sun D, Jia B, et al. Endovascular therapy for acute stroke: new evidence and indications. *Noushinkei Kekkannai Tiryou*. (2023) 17:232–42. doi: 10.5797/jnet.ra.2023-0047
- Lin J, Frontera JA. Decompressive hemicraniectomy for large hemispheric strokes. *Stroke*. (2021) 52:1500–10. doi: 10.1161/STROKEAHA.120.032359
- Bem Junior LS, Silva ACV, de Menezes MD, Galvão MJTC, Ferreira Neto ODC, de Alencar Neto JF, et al. Decompressive craniectomy: comparative analysis between surgical time and better prognosis. *Front Neurol*. (2022) 13:1041947. doi: 10.3389/fneur.2022.1041947
- Dower A, Mulcahy M, Maharaj M, Chen H, Lim CED, Li Y, et al. Surgical decompression for malignant cerebral edema after ischemic stroke: cochrane review. *Stroke*. (2023) 54:e500–2. doi: 10.1161/STROKEAHA.122.042260
- Wu S, Anderson CS. Precision management of brain oedema after acute ischaemic stroke. *Precis Clin Med*. (2022) 5:pbac019. doi: 10.1093/pmedi/pbac019
- Huang X, Chen C, Wang H, Cai Q, Li Z, Xu J, et al. The ACORNS grading scale: a novel tool for the prediction of malignant brain edema after endovascular thrombectomy. *J NeuroIntervent Surg*. (2023) 15:e190–7. doi: 10.1136/jnis-2022-019404
- Wang H, Xu C, Xiao Y, Shen P, Guo S, Shang Y, et al. A dynamic nomogram for predict individual risk of malignant brain edema after endovascular thrombectomy in acute ischemic stroke. *Sci Rep*. (2024) 14:9529. doi: 10.1038/s41598-024-60083-w
- Xie W, Ma X, Xu G, Wang Y, Huang W, Liu M, et al. Development and validation of a nomogram for the risk prediction of malignant cerebral edema after acute large hemispheric infarction involving the anterior circulation. *Front Neurol*. (2023) 14:1221879. doi: 10.3389/fneur.2023.1221879
- Du M, Huang X, Li S, Xu L, Yan B, Zhang Y, et al. A nomogram model to predict malignant cerebral edema in ischemic stroke patients treated with endovascular thrombectomy: an observational study. *NDT*. (2020) 16:2913–20. doi: 10.2147/NDT.S279303
- Arboix A, García-Eroles L, Oliveres M, Comes E, Sánchez MJ, Massons J. Malignant middle cerebral artery infarction: a clinical study of 32 patients. *Rev Invest Clin*. (2015) 67:64–70. doi: 10.1016/S0034-8376(25)00223-2
- Song SY, Ahn SY, Rhee JJ, Lee JW, Hur JW, Lee HK. Extent of contrast enhancement on non-enhanced computed tomography after intra-arterial thrombectomy for acute infarction on anterior circulation: as a predictive value for malignant brain edema. *J Korean Neurosurg Soc*. (2015) 58:321. doi: 10.3340/jkns.2015.58.4.321
- Wang C, Zhu Q, Cui T, Wang L, Yang T, Hao Z, et al. Early prediction of malignant edema after successful recanalization in patients with acute ischemic stroke. *Neurocrit Care*. (2022) 36:822–30. doi: 10.1007/s12028-021-01380-4
- Kim J-M, Park K-Y, Lee WJ, Byun JS, Kim JK, Park M-S, et al. The cortical contrast accumulation from brain computed tomography after endovascular treatment predicts symptomatic hemorrhage. *Euro J Neurol*. (2015) 22:1453–8. doi: 10.1111/ene.12764
- Hu S, Hong J, Liu F, Wang Z, Li N, Wang 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*. (2024) 14:4936–49. doi: 10.21037/qims-24-99
- Zeng W, Li W, Huang K, Lin Z, Dai H, He Z, et al. Predicting futile recanalization, malignant cerebral edema, and cerebral herniation using intelligible ensemble machine learning following mechanical thrombectomy for acute ischemic stroke. *Front Neurol*. (2022) 13:982783. doi: 10.3389/fneur.2022.982783

Acknowledgments

Some of our experiments were carried out on the Onekey AI platform. We would like to thank Onekey AI and its developers for helping in this scientific study.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Generative AI statement

The author(s) declare that no Gen AI was used in the creation of this manuscript.

Any alternative text (alt text) provided alongside figures in this article has been generated by Frontiers with the support of artificial intelligence and reasonable efforts have been made to ensure accuracy, including review by the authors wherever possible. If you identify any issues, please contact us.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

16. Wen X, Li Y, He X, Xu Y, Shu Z, Hu X, et al. Prediction of malignant acute middle cerebral artery infarction via computed tomography radiomics. *Front Neurosci.* (2020) 14:708. doi: 10.3389/fnins.2020.00708
17. Wen X, Hu X, Xiao Y, Chen J. Radiomics analysis for predicting malignant cerebral edema in patients undergoing endovascular treatment for acute ischemic stroke. *Diagn Interv Radiol.* (2023) 29:402–9. doi: 10.4274/dir.2023.221764
18. Liu W, Wang W, Zhang H, Guo M, Xu Y, Liu X. Development and validation of multi-omics thymoma risk classification model based on transfer learning. *J Digit Imaging.* (2023) 36:2015–24. doi: 10.1007/s10278-023-00855-4
19. Jan YT, Tsai PS, Huang WH, Chou LY, Huang SC, Wang JZ, 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.* (2023) 14:68. doi: 10.1186/s13244-023-01412-x
20. Guo W, Xu J, Zhao W, Zhang M, Ma J, Chen J, et al. A nomogram for predicting malignant cerebral artery infarction in the modern thrombectomy era. *Front Neurol.* (2022) 13:934051. doi: 10.3389/fneur.2022.934051
21. Cai J, Zhou Y, Zhao Y, Xu C, Yan S, Ding X, et al. Comparison of various reconstructions derived from dual-energy CT immediately after endovascular treatment of acute ischemic stroke in predicting hemorrhage. *Eur Radiol.* (2021) 31:4419–27. doi: 10.1007/s00330-020-07574-2
22. Wang W, Liang H, Zhang Z, Xu C, Wei D, Li W, et al. Comparing three-dimensional and two-dimensional deep-learning, radiomics, and fusion models for predicting occult lymph node metastasis in laryngeal squamous cell carcinoma based on CT imaging: a multicentre, retrospective, diagnostic study. *eClinicalMedicine.* (2024) 67:102385. doi: 10.1016/j.eclinm.2023.102385
23. Ayana G, Dese K, Abagaro AM, Jeong KC, Yoon SD, Choe Sw. Multistage transfer learning for medical images. *Artif Intell Rev.* (2024) 57:232. doi: 10.1007/s10462-024-10855-7
24. Song Y, Hong J, Liu F, Liu J, Chen Y, Li Z, et al. Deep learning-assisted diagnosis of malignant cerebral edema following endovascular thrombectomy. *Acad Radiol.* (2025) 32:3237–47. doi: 10.1016/j.acra.2025.02.021
25. Miao J, Song X, Sun W, Qiu X, Lan Y, Zhu Z. Predictors of malignant cerebral edema in cerebral artery infarction: a meta-analysis. *J Neurol Sci.* (2020) 409:116607. doi: 10.1016/j.jns.2019.116607
26. Wu S, Wu B, Liu M, Chen Z, Wang W, Anderson CS, et al. Stroke in China: advances and challenges in epidemiology, prevention, and management. *Lancet Neurol.* (2019) 18:394–405. doi: 10.1016/S1474-4422(18)30500-3