



OPEN Noncontrast CT-based deep learning for predicting intracerebral hemorrhage expansion incorporating growth of intraventricular hemorrhage

Youquan Ning^{1,6}, Qiang Yu^{1,6}, Xin Fan^{1,6}, Wenhao Jiang², Xinwei Chen¹, Huan Jiang¹, Kai Xie¹, Rui Liu¹, Yuan Zhou¹, Xiaodi Zhang³, Fajin Lv^{1,5}✉, Xiaoquan Xu^{2,4}✉ & Juan Peng^{1,5}✉

Intracerebral hemorrhage (ICH) is a severe form of stroke with high mortality and disability, where early hematoma expansion (HE) critically influences prognosis. Previous studies suggest that revised hematoma expansion (rHE), defined to include intraventricular hemorrhage (IVH) growth, provides improved prognostic accuracy. Therefore, this study aimed to develop a deep learning model based on noncontrast CT (NCCT) to predict high-risk rHE in ICH patients, enabling timely intervention. A retrospective dataset of 775 spontaneous ICH patients with baseline and follow-up CT scans was collected from two centers and split into training ($n=389$), internal-testing ($n=167$), and external-testing ($n=219$) cohorts. 2D/3D convolutional neural network (CNN) models based on ResNet-101, ResNet-152, DenseNet-121, and DenseNet-201 were separately developed using baseline NCCT images, and the activation areas of the optimal deep learning model were visualized using gradient-weighted class activation mapping (Grad-CAM). Two baseline logistic regression clinical models based on the BRAIN score and independent clinical-radiologic predictors were also developed, along with combined-logistic and combined-SVM models incorporating handcrafted radiomics features and clinical-radiologic factors. Model performance was assessed using the area under the receiver operating characteristic curve (AUC). The 2D-ResNet-101 model outperformed others, with an AUC of 0.777 (95%CI, 0.716–0.830) in the external-testing set, surpassing the baseline clinical-radiologic model and the BRAIN score (AUC increase of 0.087, $p=0.022$; 0.119, $p=0.003$). Compared to the combined-logistic and combined-SVM models, AUC increased by 0.083 ($p=0.029$) and 0.074 ($p<0.058$), respectively. The deep learning model can identify ICH patients with high-risk rHE with favorable predictive performance than traditional baseline models based on clinical-radiologic variables and radiomics features.

Keywords Intracerebral hemorrhage, Revised hematoma expansion, Deep learning, Machine learning, Computed tomography

Abbreviations

AHA/ASA	American heart association/american stroke association
AUC	Area under the curve
BP	Blood pressure
cHE	Conventional hematoma expansion
CTA	CT angiography
CNN	Convolutional neural network

¹Department of Radiology, The First Affiliated Hospital of Chongqing Medical University, Chongqing, China.

²Department of Radiology, The First Affiliated Hospital of Nanjing Medical University, Nanjing, China. ³Philips Healthcare, Chengdu, China. ⁴Department of Radiology, The First Affiliated Hospital of Nanjing Medical University, No. 300 Guangzhou Road, Gulou District, Nanjing 210029, China. ⁵Department of Radiology, The First Affiliated Hospital of Chongqing Medical University, No. 1 Youyi Road, Yuzhong District, Chongqing 400016, China. ⁶These authors contributed equally to the study as co-first authors: Youquan Ning, Qiang Yu and Xin Fan. ✉email: fajinlv@163.com; xiaoquanxu_1987@163.com; pengjuan0717@sina.com

CI	Confidence interval
Grad-CAM	Gradient-weighted class activation mapping
HE	Hematoma expansion
HU	Hounsfield units
ICH	Intracerebral hemorrhage
IVH	Intraventricular hemorrhage
INR	International normalized ratio
ICC	Intraclass correlation coefficient
IQR	Interquartile range
LASSO	Least absolute shrinkage and selection operator
ML	Machine learning
NCCT	Noncontrast computed tomography
NPV	Negative prediction value
PPV	Positive prediction value
rHE	Revised hematoma expansion
ROC	Receiver operating characteristic
SVM	Support vector machine
VOI	Volumes of interest

Hematoma expansion (HE) is an independent predictor of poor prognosis and early neurological deterioration in intracerebral hemorrhage (ICH), approximately one-fifth of whom experience HE within 24 h of acute ICH^{1,2}. Conventional HE (cHE) is defined ICH volume increase (≥ 6 mL or $\geq 33\%$ ratio) within 24–72 h after the initial CT scan³. This definition overlooks another independent factor, intraventricular hemorrhage (IVH) expansion, which is associated with poor prognosis in ICH patients^{4–6}. Some studies have proposed that revised hematoma expansion (rHE), incorporating IVH growth, may enhance the prediction of poor neurological outcomes compared to cHE^{7,8}. Hence, identifying high-risk rHE effectively is crucial to provide timely targeted medical interventions, such as intensive systolic blood pressure (BP) reductions or emergency surgical intervention^{9,10}.

Currently, various predictive scores, including clinical variables, laboratory markers, noncontrast CT (NCCT) markers, and CT-angiography (CTA) spot signs, are used to identify the risk of ICH growth². A clinical prediction score (BRAIN) developed from simple variables achieved good predictions (C-statistic, 0.73) in estimating the probability of ICH growth³. A large meta-analysis revealed that models incorporating additional CTA spot signs had greater overall discriminative ability¹¹ but CTA availability is not standardized in acute settings¹². Although NCCT-based imaging markers may serve as reliable substitutes for CTA spot signs¹³ they are limited by experience dependence, prone to inter- and intrareader variability, and are too time-consuming for routine clinical use. Machine learning (ML) aided diagnosis models have received significant attention in medical imaging^{14,15} and their application in predicting cHE has demonstrated promising preliminary results^{16,17}. Compared with traditional ML, deep learning allows end-to-end prediction and classification by automatically learning the semantic and spatial features from raw input image pixels without relying on human-designed, engineered features¹⁸. Zhong et al.'s study applied an end-to-end convolutional neural network (CNN) model using NCCT images from 266 patients to predict cHE, achieving a high C statistic of 0.80 and demonstrating superiority to NCCT markers alone¹⁹. Recently, Li et al. used a large dataset to develop a two-stage framework for predicting cHE, achieving an AUC of 0.806 in a prospective set²⁰. Moreover, Tran et al.'s study, which was based on a relatively small dataset ($n = 793$), also developed a model to predict cHE, achieving a similar AUC of 0.80 with NCCT images²¹. These studies highlight the ability of deep learning to identify the risk of cHE. However, they did not assess the dynamic evolution of IVH during follow-up, a process may lead to obstructive hydrocephalus requiring external ventricular drainage^{22,23}.

Hence, we applied deep learning models to identify ICH patients at high risk for rHE by incorporating IVH expansion into the cHE definition, evaluating its feasibility for predicting high-risk rHE. This study aimed to develop a deep learning-based prediction tool using baseline NCCT images to efficiently identify high-risk rHE in ICH patients and compare its discriminative ability with traditional baseline ML models, which rely on clinical-radiologic and handcrafted radiomics features.

Materials and methods

Ethics declarations and consent to participate

This study was conducted in accordance with the Declaration of Helsinki principles. Ethical approval was obtained from the institutional review boards of both participating centers (Approval No. K2023-138). The requirement for written informed consent was waived due to the retrospective nature of the study. Additionally, the study was registered with the Research Registry (Registration No. researchregistry10383; <https://www.researchregistry.com/>).

Patient cohort and clinical variables

A total of 775 spontaneous ICH patients who had undergone baseline and follow-up NCCT were consecutively enrolled from the two centers. The exclusion criteria were as follows: (1) age < 18 years; (2) baseline CT obtained more than 24 h after the onset of ICH symptoms or last seen well for baseline CT; (3) follow-up CT obtained > 72 h after baseline CT; (4) primary IVH; (5) previous surgical treatment with external ventricular drain placement or craniotomy; and (6) obvious artifacts observed on the CT images. Finally, 556 patients from center 1 were divided into a training set (389) and an internal-testing set (167) via a 7:3 stratified random sampling method, and all 219 patients from center 2 formed an independent external-testing set. Figure 1 illustrates the process of population selection. Clinical data including sex, age, history of alcohol consumption, smoking status, diabetes

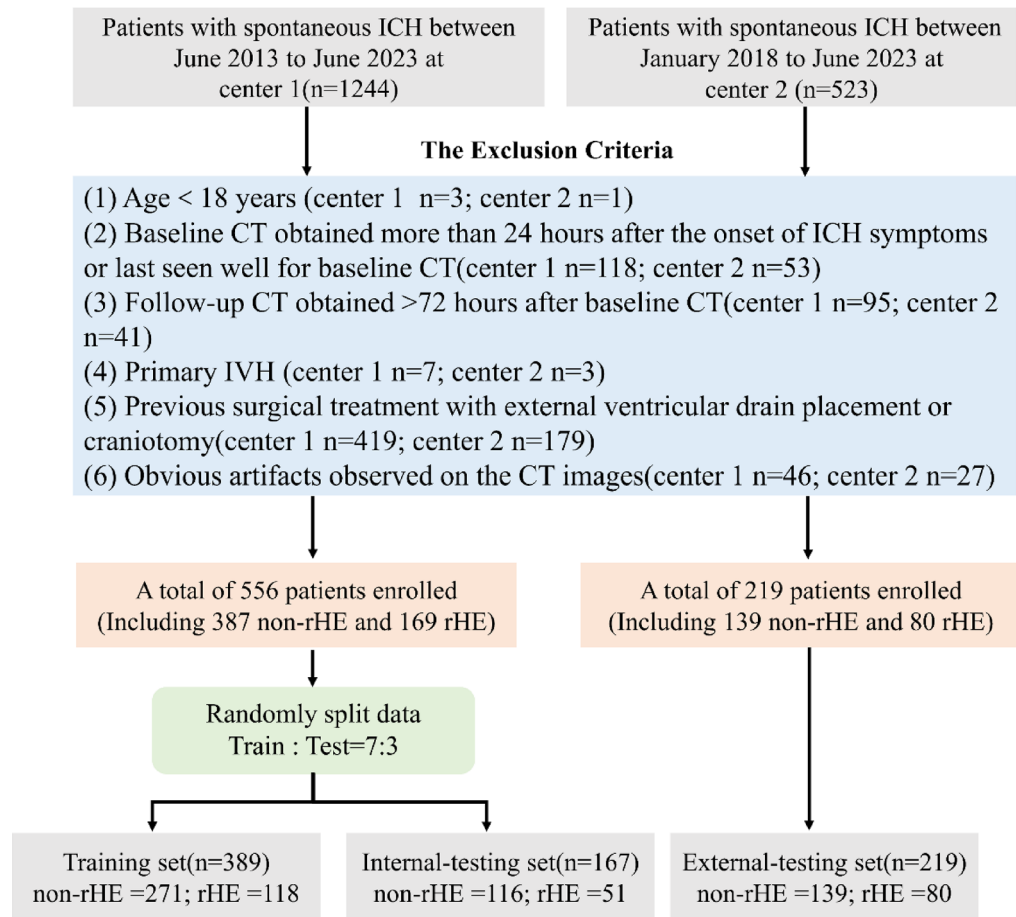


Fig. 1. Flowchart of patient recruitment. *ICH*, intracerebral hemorrhage; *Center 1*, The First Affiliated Hospital of Chongqing Medical University; *Center 2*, The First Affiliated Hospital of Nanjing Medical University; *NCCT*, noncontrast computed tomography; *IVH*, intraventricular hemorrhage; *non-rHE*, non-revised hematoma expansion; *rHE*, revised hematoma expansion.

mellitus, hypertension, Onset to baseline CT time interval, admission systolic/diastolic pressure, international normalized ratio (INR), antithrombotic and anticoagulation treatment, were obtained from the medical records of the patients.

Image acquisition and analysis

Admission and follow-up NCCT scans were performed at each participating center according to the local acquisition protocols, details of the CT scan protocols are provided in Supplementary Table S1. To balance the variations between different CT devices and protocols, we resampled all NCCT images into a uniform voxel spacing of $1.0 \times 1.0 \times 5.0 \text{ mm}^3$ for subsequent analysis. The overall design of this study is presented in Fig. 2. Image analysis was independently performed by two radiologists (A with 4 years and B with 6 years of neuroimaging experience) on the baseline NCCT images without knowledge of the follow-up CT images, patient outcome data, or demographics. Axial NCCT images were analyzed to determine the presence of subarachnoid hemorrhage and IVH, the location of the ICH, and NCCT markers. Nine features of NCCT markers were evaluated and illustrative examples are shown in Fig. 3. The consensus definitions from the 2022 AHA/ASA guidelines were followed throughout the analysis¹⁰. One month later, radiologist A performed a second independent assessment using the same method (see Supplementary Material for more information). Inter- and intra-observer agreement of the subjective CT image features was evaluated using Cohen's kappa coefficient.

The semiautomatic segmentation software of ITK-SNAP 3.8.0 was employed to obtain the initial and follow-up volumes of ICH and IVH. Radiologist A utilized semiautomatic segmentation methods, including thresholding, classification, clustering, and edge attraction to generate rough segmentation masks on the NCCT images. Manual adjustments were subsequently made to refine the segmentation regions and obtain exact 3D hematoma volumes of interest (VOI). Radiologist B independently segmented the images using the same method. To ensure the consistency and repeatability of the segmentations, 60 NCCT images were randomly selected to calculate the intraclass and interclass correlation coefficients (ICC), in which ICC values above 0.80 indicated good agreement. rHE was defined as any of the following: absolute increase in ICH volume $\geq 6 \text{ mL}$ or relative increase $\geq 33\%$, increase in IVH volume $\geq 1 \text{ mL}$, or de novo IVH^{7,8}.

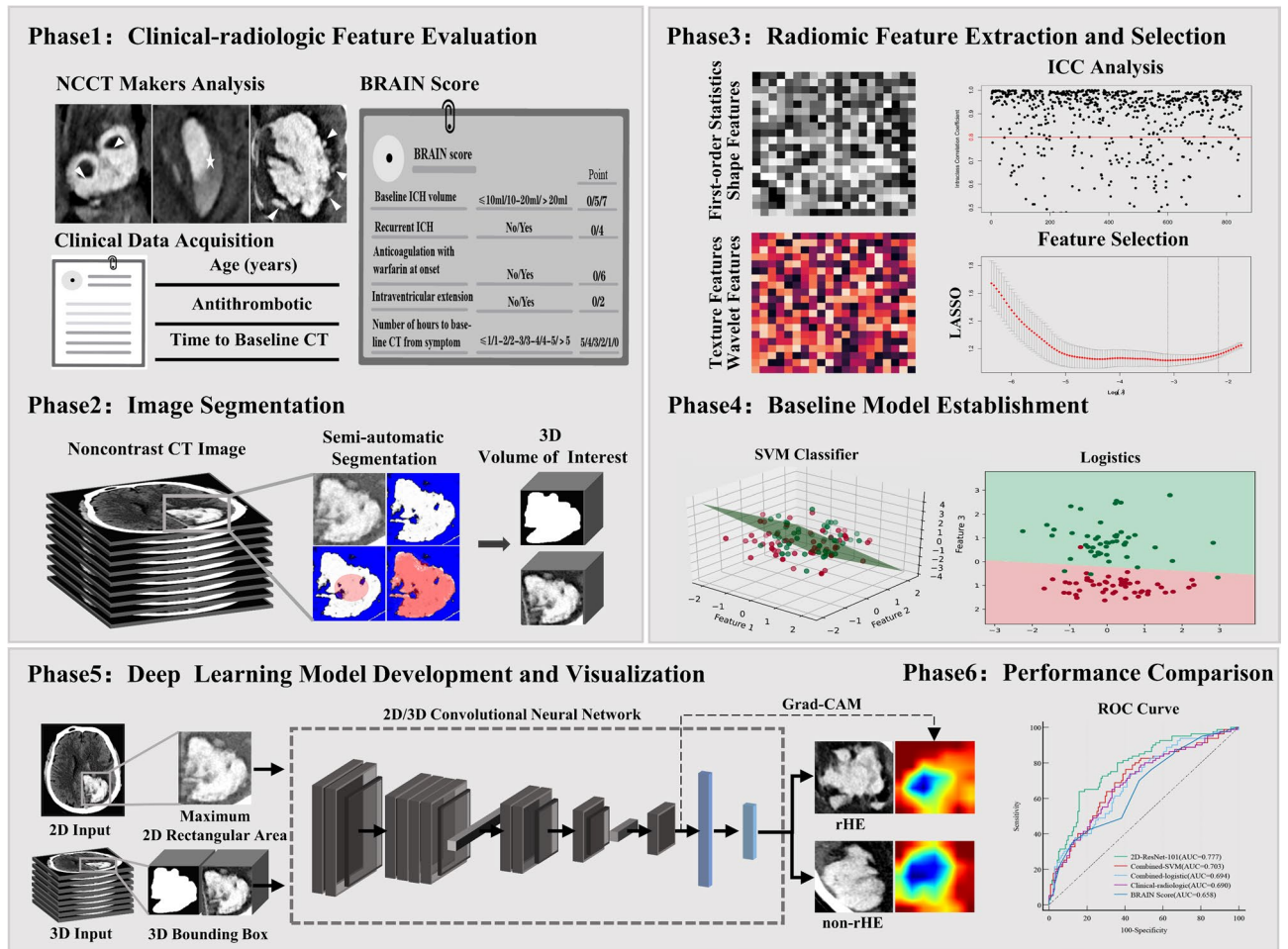


Fig. 2. Overall pipeline of this study. Phase 1: Evaluation of clinical-radiologic features in intracerebral hemorrhage (ICH) patients, including clinical data acquisition, noncontrast computed tomography (NCCT) markers analysis, and BRAIN score assessment. Phase 2: Generation of initial rough segmentation masks on NCCT images with semiautomatic segmentation software followed by manual adjustments to refine the segmentation regions and obtain precise 3D volumes of interest for the hematoma. Phase 3: Extraction and selection of radiomics features with high stability ($ICC > 0.80$), followed by further refinement using the least absolute shrinkage and selection operator (LASSO). Phase 4: Establishment of baseline predictive models with the support vector machine (SVM) algorithm and logistic regression. Phase 5: Development of 2D and 3D convolutional neural network models using NCCT images and gradient-weighted class activation mapping (Grad-CAM) for model visualization. Phase 6: Comparison of the predictive performances of the different models. *ICC*, intraclass correlation coefficient.

Baseline model construction based on clinical-radiologic and radiomics features

Univariate logistic regression analysis was used to assess the associations between clinical-radiologic parameters and rHE. Variables with $p < 0.05$ were included in multivariate logistic regression analysis to identify independent predictors of rHE. The radiomics workflow included 3D VOI segmentation, radiomics feature extraction, selection, and model establishment. PyRadiomics 3.0.1 (<https://pyradiomics.readthedocs.io/en/latest>) was used to extract the radiomics features. The gray values were discretized with a fixed bin width of 25 Hounsfield units (HU) was applied to all NCCT images. A total of 851 features (744 wavelet decomposition features, 75 texture features, 14 shape-based features, and 18 first-order statistics) were extracted from each 3D VOI, and standardized with z-score. Features with high inter- and intragroup stability ($ICC > 0.80$) were retained for subsequent analysis. The optimal radiomics features in the training set were then selected with the least absolute shrinkage and selection operator (LASSO) algorithm.

We developed four baseline ML prediction models: (1) a logistic regression model based on the BRAIN score (B for baseline intracerebral volume, R for recurrent intracerebral hemorrhage, A for anticoagulation treatment, I for IVH hemorrhage, N for hours from onset to CT), calculated using simple clinical variables (see Supplementary Table S2 for scoring criteria); (2) a logistic regression model using independent clinical-radiologic predictors of rHE, named the clinical-radiologic model; (3) a combined-logistic model; and (4) a combined-SVM model, both incorporating radiomics features and independent clinical-radiologic predictors.

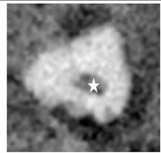
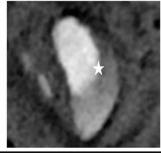

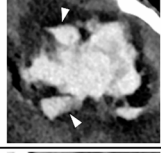
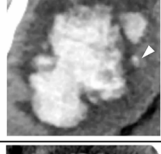
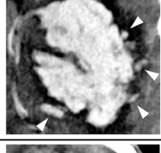
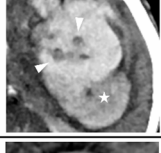
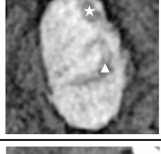
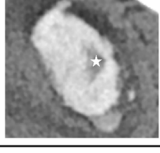
Feature	Definition	Kappa (Intra/Inter)	Example
Black hole sign	Hypoattenuating area encapsulated within the hyperattenuating hematoma with a density difference >28 HU, identifiable border, and any morphology but no connection with the adjacent brain tissue.	0.941/0.846	
Blend sign	Hypoattenuating area adjacent to a hyperattenuating area of the hematoma with a density difference >18 HU and a clear margin between the two areas.	0.938/0.849	
Fluid level	Hypoattenuating area above and hyperattenuating area below a straight line of separation (compared with the brain parenchyma), regardless of its density appearance.	0.959/0.915	
Irregular shape	At least 2 foci of irregularities in the hematoma margin of the largest hematoma region on the axial non-contrast CT slice, meeting the criteria of a Barras scale \geq III.	0.956/0.942	
Satellite sign	A small hematoma with a maximum diameter of 10 mm, distinct from the main hematoma by 1-20 mm separation.	0.942/0.867	
Island sign	At least 3 scattered small hematomas all separate from the main hemorrhage or \geq 4 bubble-like or sprout-like hematomas, some or all of which may connect with the main hemorrhage.	0.937/0.835	
Heterogeneous density	At least 3 foci of low attenuation within the largest hematoma compared with the brain parenchyma, defined with a Barras scale of \geq III.	0.945/0.874	
Swirl sign	Region(s) of hypoattenuation or isoattenuation within the hyperattenuated ICH compared to the brain parenchyma, with any morphology, does not have to be encapsulated in the ICH.	0.952/0.859	
Hypodensity	Any hypodense region with any shape, size and density, which strictly encapsulated within the hemorrhage, without density measurement.	0.973/0.960	

Fig. 3. NCCT markers definitions, illustrative examples, and intra- and inter-observer kappa values. The definitions and illustrative examples of NCCT markers, and the intra- and inter-observer kappa values for radiologists A and B regarding the subjective markers. *HU*, Hounsfield units; *ICH*, intracerebral hemorrhage.

Deep learning model development and interpretation of model visualization

Four CNN models (ResNet-101, ResNet-152, DenseNet-121, DenseNet-201) based on 2D/3D images were developed. Details of these models are provided in the Supplementary Material. According to the semiautomatically generated 3D segmentation masks, the largest 2D rectangular region and smallest 3D bounding box of the hematoma on baseline NCCT images were cropped and input into the 2D-CNN and 3D-CNN

models, respectively. The original CT images were adjusted to a window width of 80 HU and window level of 40 HU during cropping. For the 2D-CNN models, the input size was resized to 224×224 pixels, with a batch size of 128. The model weights were initialized by pretraining on the ImageNet dataset. For the 3D-CNN models, the input size was $112 \times 112 \times 112$ voxels, with a batch size of 16. To mitigate overfitting, batch normalization layers were added after convolutional layers, a dropout layer with a rate of 0.5 was incorporated, and online data augmentation (including random horizontal flipping, cropping, and translation) was applied during training. The Adam optimizer was used with a learning rate of 0.01 and 100 epochs. Models were implemented using Python 3.7.6 and PyTorch 1.8.1.

To visualize the decision-making process of the optimal deep learning model decision-making process, the gradient-weighted class activation mapping (Grad-CAM) was employed to generate 2D attention maps by extracting feature maps from the final CNN layer²⁴ which emphasized the importance of hematoma regions in predicting the risk of rHE. These attention maps were subsequently resized and overlaid onto the original NCCT images to identify crucial regions in the target image for classification.

Statistical analysis

Continuous variables are expressed as the mean \pm standard deviation or median (interquartile range [IQR]), and categorical variables are expressed as frequencies and percentages (%), Student's *t*-test or the Mann-Whitney *U* test was applied to assess differences in continuous variables between groups, and the chi-square test or Fisher's exact test was used to compare categorical variables between groups. The radiologists' interobserver agreement for NCCT image features was assessed with Cohen's kappa coefficient. AUC, Accuracy, Sensitivity, Specificity, Positive Predictive Value (PPV), Negative Predictive Value (NPV) and F1-score, were computed to evaluate model performance. Sensitivity and specificity were calculated based on a predicted probability of 0.5, and differences were evaluated using McNemar's chi-square test. The DeLong test was used to assess the differences among AUCs. A two-sided *p*-value < 0.05 was considered to indicate statistical significance. The analyses were conducted by the MedCalc version 20.009, SPSS version 26.0 and R version 3.6.0.

Results

Clinical-radiologic characteristics and radiomics signature analysis

Table 1 summarizes the baseline clinical-radiologic characteristics of 775 patients with spontaneous ICH. Among them, 118 (mean age, 63.03 years \pm 12.20 [SD]; 87 men) in the training cohort, 51 (mean age, 61.00 years \pm 15.02 [SD]; 41 men) in the internal-testing cohort, and 80 (mean age, 63.89 years \pm 14.58 [SD]; 52 men) in the external-testing cohort developed rHE according to follow-up CT. Intra-observer kappa values for radiologists A and B ranged from 0.937 to 0.973, while inter-observer kappa values ranged from 0.835 to 0.960 (Fig. 3). Univariate and multivariate regression analyses of clinical-radiologic features in the training cohort are shown in Table 2. Multivariate analysis identified onset to baseline CT time interval (OR 0.914; 95% CI, 0.853–0.979), baseline ICH volume (OR 1.040; 95% CI, 1.016–1.064), IVH presence (OR 3.557; 95% CI, 1.663–7.610), and hypodensity (OR 2.328; 95% CI, 1.270–4.267) as independent predictors of rHE. Of 851 radiomics features extracted from each 3D VOI, 488 with ICC > 0.80 were selected, and nine optimal features were identified using LASSO-based screening in the training set (Supplementary Table S3 and Figure S2).

Predictive performance of the baseline prediction models

The baseline BRAIN score and the clinical-radiologic model shown comparable predictive performance. In both internal- and external-testing sets, their AUCs and sensitivities were relatively low (Table 3). In contrast, the combined-logistic and combined-SVM models, which incorporated the four independent clinical predictive factors and the nine radiomics features, demonstrated better performance in identifying high-risk rHE in ICH patients. In the internal-testing, the AUCs were 0.682 (95% CI: 0.605–0.752) and 0.692 (95% CI: 0.616–0.761), with sensitivities of 0.530 and 0.412, specificities of 0.759 and 0.827, respectively. In the external-testing, the AUCs were 0.694 (95% CI: 0.629–0.755) and 0.703 (95% CI: 0.638–0.763), with sensitivities of 0.488 and 0.338, specificities of 0.712 and 0.863, respectively (Table 3; Fig. 4E and F). To some extent, these findings indicate that the radiomics features may have provided additional information for predicting the risk of rHE.

Performance comparison and visualization analysis of the deep learning models

In both test sets, the predictive performance of the 2D-CNN models was significantly greater than that of the 3D-CNN models. All four 3D-CNN models exhibited substantial generalization errors in the two-center datasets (Table 3; Fig. 4A–D). In the external-testing set, the 2D-ResNet-101 had the best overall performance among the eight deep learning models with an AUC of 0.777 (95% CI: 0.716–0.830), accuracy of 0.767, sensitivity of 0.637, specificity of 0.842, PPV of 0.699, NPV of 0.801, F1-score of 0.667, while in the internal-testing set, these values were 0.782 (95% CI: 0.712–0.842), 0.766, 0.667, 0.810, 0.607, 0.847 and 0.636, respectively. Grad-CAM generated heatmaps that visually emphasize the key regions used to classify rHE and non-rHE in ICH. The red regions indicate areas most influential to the 2D-ResNet-101 model's classification, primarily located at the hematoma and its periphery regions (Fig. 5).

The 2D-ResNet-101 model outperformed both the baseline clinical-radiologic model and the BRAIN score. In the internal-testing set, the AUCs increased by 0.094 ($p = 0.048$) and 0.127 ($p = 0.013$), and the sensitivity by 0.216 ($p = 0.013$) and 0.373 ($p < 0.001$), respectively. In the external-testing set, the AUC values increased by 0.087 ($p = 0.022$) and 0.119 ($p = 0.003$), and the sensitivity by 0.174 ($p = 0.009$) and 0.287 ($p < 0.001$), respectively. There were no significant differences in specificity in the test sets. The confusion matrices (Fig. 6) show that the clinical-radiologic model and the BRAIN Score misclassified 28/51 and 36/51 rHE cases as non-rHE in the internal-testing set, and 43/80 and 52/80 cases in the external-testing set, respectively. In contrast, the 2D-ResNet-101 model reduced rHE misclassifications to 17/51 and 29/80 cases in the internal and external test

Clinic-radiologic features	Training set (n = 389)			Internal-testing set (n = 167)			External-testing set (n = 219)		
	non-rHE (n = 271)	rHE (n = 118)	P	non-rHE (n = 116)	rHE (n = 51)	P	non-rHE (n = 139)	rHE (n = 80)	P
Sex*(male)	203(74.9)	87(73.7)	0.806	82(70.7)	41(80.4)	0.190	95(68.3)	52(65.0)	0.612
Age [#] (year)	57.67 (13.81)	63.03 (12.20)	<0.001	58.89 (13.82)	61.00 (15.02)	0.377	60.16 (14.02)	63.89 (14.58)	0.063
Alcohol consumption*	105(38.7)	48(40.7)	0.720	48(41.4)	14(27.5)	0.086	37(26.6)	10(12.5)	0.014
Smoking*	121(44.6)	54(45.8)	0.839	52(44.8)	17(33.3)	0.165	39(28.1)	18(22.5)	0.367
Diabetes mellitus *	36(13.3)	26(22.0)	0.030	21(18.1)	9(17.6)	0.944	22(15.8)	16(20.0)	0.432
Hypertension*	236(87.1)	105(89.0)	0.601	100(86.2)	40(78.4)	0.209	112(80.6)	67(83.8)	0.558
Onset to baseline CT time interval [#] (hour)	5.00 (5.00)	3.00 (4.00)	0.001	4.00 (5.00)	3.00 (3.00)	0.013	4.00 (2.00)	4.00 (3.00)	0.210
Admission SBP [#] (mm Hg)	166.00 (35.00)	170.50 (37.00)	0.045	166.78 (29.78)	172.02 (30.53)	0.300	156.00 (38.00)	150.00 (29.00)	0.124
Admission DBP [#] (mm Hg)	98.00 (22.00)	96.00 (32.00)	0.822	99.7 (18.63)	95.61 (22.57)	0.240	90.00 (19.00)	84.50 (20.00)	0.121
INR [#]	1.01 (0.10)	1.02 (0.09)	0.330	1.01 (0.10)	1.01 (0.08)	0.910	1.01 (0.08)	1.03 (0.12)	0.056
Antithrombotic*	12(4.4)	14(11.9)	0.007	7(6.0)	10(19.6)	0.008	11(7.9)	10(12.5)	0.267
Anticoagulation*	6(2.2)	13(11.0)	<0.001	2(1.7)	4(7.8)	0.132	6(4.3)	6(7.5)	0.491
ICH volume [#] (ml)	12.45 (15.82)	18.95 (19.88)	<0.001	17.43 (23.79)	24.40 (28.13)	0.040	17.39 (26.10)	21.70 (37.87)	0.005
IVH volume [#] (ml)	0.00 (0.00)	0.00 (8.15)	<0.001	0.00 (1.07)	0.00 (3.70)	0.119	0.00 (0.28)	1.82 (15.039)	<0.001
Location*			0.007			0.853			0.486
Lobar	40(14.8)	5(4.2)		4(3.5)	2(3.9)		17(12.2)	6(7.5)	
Deep	177(65.3)	92(78.0)		78(67.2)	32(62.7)		88(63.3)	51(63.8)	
Infratentorial	54(19.9)	21(17.8)		34(29.3)	17(33.3)		34(24.5)	23(28.8)	
IVH*	48(17.7)	55(46.6)	<0.001	36(31.0)	23(45.1)	0.080	35(25.2)	51(63.8)	<0.001
SAH*	24(8.9)	17(14.4)	0.101	19(16.4)	13(25.5)	0.168	18(12.9)	17(21.3)	0.107
Irregular shape*	78(28.8)	51(43.2)	0.005	51(44.0)	24(47.1)	0.711	51(36.7)	39(48.8)	0.081
Satellite sign*	71(26.2)	46(39.0)	0.011	43(37.1)	24(47.1)	0.225	51(36.7)	34(42.5)	0.396
Island sign*	16(5.9)	27(22.9)	<0.001	10(8.6)	14(27.5)	0.001	24(17.3)	15(18.8)	0.782
Black hole sign*	40(14.8)	26(22.0)	0.079	25(21.6)	11(21.6)	0.998	23(16.5)	14(17.5)	0.856
Blend sign*	28(10.3)	20(16.9)	0.068	18(15.5)	11(21.6)	0.342	13(9.4)	7(8.8)	0.882
Fluid level*	3(1.1)	2(1.7)	1.000	3(2.6)	0(0.0)	0.554	4(2.9)	1(1.3)	0.759
Swirl sign*	53(19.6)	39(33.1)	0.004	31(26.7)	16(31.4)	0.538	26(18.7)	13(16.3)	0.647
Hypodensity*	90(33.2)	75(63.6)	<0.001	44(37.9)	27(52.9)	0.071	50(36.0)	30(37.5)	0.821
Heterogeneous density*	25(9.2)	27(22.9)	<0.001	14(12.1)	15(29.4)	0.006	24(17.3)	14(17.5)	0.965

Table 1. Baseline clinical-radiologic features of the datasets. [#]Continuous variables are expressed as mean \pm standard deviation or median (interquartile range [IQR]), *P*-value was calculated with Student's *t*-test or Mann-Whitney *U* test. *Categorical variables are expressed as frequencies and percentages (%), *P*-value was calculated with the chi-square test or Fisher's exact test. *SBP*, systolic blood pressure; *DBP*, diastolic blood pressure; *INR*, international normalized ratio; *ICH*, intracerebral hemorrhage; *IVH*, intraventricular hemorrhage; *SAH*, subarachnoid hemorrhage.

sets, respectively, significantly lowering the risk of missed diagnosis ($p < 0.05$, McNemar's test). Additionally, the 2D-ResNet-101 model was superior to the two baseline combined models, showing improvements in AUC and sensitivity across both test sets (Fig. 7).

Discussion

In this study, we developed 2D/3D CNN models based on NCCT images to predict high-risk rHE in ICH patients and compared its performance with four baseline ML models. The main findings showed that the developed 2D-ResNet-101 model had the optimal predictive performance, demonstrating significant improvement over the BRAIN score and clinical-radiologic model in both the internal- and external-testing sets. Furthermore, it exhibited higher sensitivity and accuracy than the two combined models in the testing sets. These findings suggest that the deep learning model may provide more comprehensive information about hematoma heterogeneity compared to routine clinical predictive indicators and radiomics features alone can, thus more effectively predicting the rHE. This model could allow the identification of patients who may benefit from anti-expansion therapies in the acute ICH settings.

Variable	Univariate OR (95%CI)	P Value	Multivariate OR (95%CI)	P Value
Sex	1.064(0.649–1.742)	0.806		
Age	1.031(1.014–1.049)	< 0.001	1.022(0.999–1.045)	0.064
Alcohol consumption	1.084(0.697–1.685)	0.720		
Smoking	1.046(0.678–1.615)	0.839		
Diabetes mellitus	1.845(1.055–3.227)	0.032	1.553(0.790–3054)	0.202
Hypertension	1.198(0.609–2.357)	0.601		
Onset to baseline CT time interval	0.941(0.889–0.997)	0.041	0.914(0.853–0.979)	0.010
Admission SBP	1.008(1.001–1.016)	0.030	1.006(0.997–1.016)	0.211
Admission DBP	1.003(0.992–1.015)	0.563		
INR	7.150(0.935–54.690)	0.058		
Antithrombotic	2.905(1.300–6.492)	0.009	2.095(0.774–5.674)	0.146
Anticoagulation	5.468(2.025–14.766)	0.001	3.011(0.959–9.458)	0.059
ICH volume	1.043(1.028–1.059)	< 0.001	1.040(1.016–1.064)	0.001
IVH volume	1.061(1.030–1.093)	< 0.001	0.995(0.953–1.039)	0.826
Location				
Lobar	Reference		Reference	
Deep	4.158(1.587–10.895)	0.004	1.677(0.582–4.831)	0.338
Infratentorial	3.111(1.081–8.958)	0.035	0.972(0.271–3.481)	0.965
IVH	4.056(2.516–6.539)	< 0.001	3.557(1.663–7.610)	0.001
SAH	1.732(0.893–3.362)	0.104		
Irregular shape	1.883(1.202–2.952)	0.006	0.786(0.416–1.486)	0.459
Satellite sign	1.800(1.138–2.846)	0.012	0.851(0.433–1.671)	0.639
Island sign	4.729(2.437–9.177)	< 0.001	1.450(0.546–3.850)	0.456
Black hole sign	1.632(0.942–2.828)	0.081		
Blend sign	1.771(0.953–3.292)	0.071		
Fluid level	1.540(0.254–9.340)	0.639		
Swirl sign	2.031(1.248–3.305)	0.004	0.746(0.367–1.518)	0.419
Hypodensity	3.508(2.232–5.513)	< 0.001	2.328(1.270–4.267)	0.006
Heterogeneous density	2.920(1.611–5.292)	< 0.001	0.979(0.411–2.328)	0.961

Table 2. Univariate and multivariate analysis of clinical-radiologic predictors for revised hemorrhage expansion in the training set. *SBP*, systolic blood pressure; *DBP*, diastolic blood pressure; *INR*, international normalized ratio; *ICH*, intracerebral hemorrhage; *IVH*, intraventricular hemorrhage; *SAH*, subarachnoid hemorrhage.

Spontaneous ICH is the deadliest acute stroke type, with high morbidity and mortality^{25,26}. Notably, in real-world clinical scenarios, parenchymal hematomas often extend into the ventricular space²⁷ and the extent of this extension correlates exponentially with patient outcomes²⁸. In the present research, we included IVH expansion in the definition of cHE and explored potential clinical-radiologic factors affecting rHE. Multivariate regression analysis identified significant differences in the onset to baseline CT time interval, ICH volume, and presence of IVH between the groups, with patients who developed rHE showing shorter baseline scan intervals, larger ICH volumes, and a higher likelihood of IVH (Table 2). These findings highlight the urgent need for rapid assessment and intervention to limit ICH growth and improve outcomes, especially for infratentorial hemorrhage. This hemorrhage may disrupt neural pathways related to the Guillain-Morath triangle, a network critical for movement coordination and control, and dysfunction of which can lead to a variety of neurological disorders, such as post-stroke palatal tremor²⁹. According to the 2022 AHA/ASA guidelines, NCCT markers are valuable potential imaging predictors for identifying patients at risk of rHE¹⁰. Our analysis showed that hypodensities were the only independent risk factor among the nine NCCT markers, likely indicating areas of incomplete blood clotting prone to instability and further bleeding^{30,31}. Hypodensities also overlap with other NCCT signs³² and their high prevalence may support their role as a predictor. We also developed the BRAIN score and a clinical-radiologic model based on routinely available clinical variables, but these demonstrated limited predictive performance in the testing sets. The sensitivity of these models ranged from 0.350 to 0.488, suggesting a substantial risk of missing rHE diagnoses, which could lead to delayed treatment and potentially serious consequences. These findings highlight the limitations of clinical-radiologic features in predicting rHE, likely due to their qualitative or semiquantitative nature, which can introduce subjectivity and inconsistency in predictions³³. This was further evidenced by variability in inter- and intra-observer agreement regarding NCCT markers in this study.

Recent studies have shown promising results using traditional machine learning (ML) methods, including radiomics and deep learning, to predict intracerebral hemorrhage (ICH) growth. Feng and Pszczolkowski et al. applied deep learning radiomics or radiomics features derived from NCCT images to predict cHE, achieving

Model		AUC (95%CI)	Accuracy	Sensitivity	Specificity	PPV	NPV	F1-Score
BRAIN score	Training	0.729(0.681–0.772)	0.712	0.254	0.911	0.556	0.737	0.349
	Internal-testing	0.655(0.577–0.727)	0.695	0.294	0.871	0.500	0.737	0.370
	External-testing	0.658(0.591–0.720)	0.685	0.350	0.878	0.622	0.701	0.448
Clinical-radiologic	Training	0.782(0.737–0.822)	0.756	0.407	0.908	0.658	0.778	0.503
	Internal-testing	0.688(0.612–0.757)	0.689	0.451	0.793	0.489	0.767	0.469
	External-testing	0.690(0.625–0.751)	0.662	0.463	0.777	0.544	0.715	0.500
Combined-logistic	Training	0.788(0.744–0.828)	0.761	0.407	0.915	0.676	0.780	0.508
	Internal-testing	0.682(0.605–0.752)	0.689	0.530	0.759	0.491	0.786	0.509
	External-testing	0.694(0.629–0.755)	0.630	0.488	0.712	0.494	0.707	0.490
Combined-SVM	Training	0.855(0.846–0.912)	0.784	0.398	0.952	0.783	0.784	0.528
	Internal-testing	0.692(0.616–0.761)	0.701	0.412	0.827	0.512	0.762	0.457
	External-testing	0.703(0.638–0.763)	0.671	0.338	0.863	0.587	0.694	0.429
2D-ResNet-101	Training	0.882(0.843–0.910)	0.802	0.746	0.827	0.652	0.882	0.696
	Internal-testing	0.782(0.712–0.842)	0.766	0.667	0.810	0.607	0.847	0.636
	External-testing	0.777(0.716–0.830)	0.767	0.637	0.842	0.699	0.801	0.667
2D-ResNet-152	Training	0.879(0.843–0.910)	0.802	0.661	0.863	0.678	0.854	0.670
	Internal-testing	0.795(0.726–0.854)	0.731	0.667	0.759	0.548	0.838	0.602
	External-testing	0.759(0.696–0.814)	0.717	0.700	0.727	0.596	0.808	0.644
2D-DenseNet-121	Training	0.868(0.830–0.900)	0.830	0.627	0.923	0.779	0.850	0.695
	Internal-testing	0.721(0.646–0.787)	0.683	0.627	0.707	0.485	0.812	0.547
	External-testing	0.735(0.671–0.792)	0.689	0.650	0.712	0.565	0.780	0.604
2D-DenseNet-201	Training	0.902(0.868–0.929)	0.823	0.763	0.849	0.687	0.891	0.723
	Internal-testing	0.759(0.687–0.822)	0.737	0.647	0.776	0.559	0.833	0.600
	External-testing	0.740(0.677–0.797)	0.703	0.662	0.727	0.582	0.789	0.620
3D-ResNet-101	Training	0.945(0.902–0.955)	0.882	0.771	0.930	0.827	0.903	0.798
	Internal-testing	0.644(0.567–0.717)	0.605	0.608	0.603	0.403	0.778	0.484
	External-testing	0.615(0.547–0.680)	0.616	0.525	0.669	0.477	0.710	0.500
3D-ResNet-152	Training	0.932(0.902–0.955)	0.851	0.831	0.860	0.721	0.921	0.772
	Internal-testing	0.644(0.566–0.716)	0.623	0.647	0.612	0.423	0.798	0.512
	External-testing	0.577(0.509–0.643)	0.589	0.500	0.641	0.444	0.690	0.471
3D-DenseNet-121	Training	0.952(0.926–0.971)	0.913	0.881	0.926	0.839	0.947	0.860
	Internal-testing	0.645(0.567–0.717)	0.641	0.510	0.698	0.426	0.764	0.464
	External-testing	0.619(0.551–0.683)	0.635	0.537	0.691	0.500	0.722	0.518
3D-DenseNet-201	Training	0.938(0.909–0.959)	0.869	0.822	0.889	0.764	0.920	0.792
	Internal-testing	0.665(0.588–0.736)	0.611	0.608	0.612	0.408	0.780	0.488
	External-testing	0.510(0.442–0.578)	0.562	0.325	0.697	0.382	0.642	0.351

Table 3. Performance comparison of different prediction models. *PPV*, positive predictive value; *NPV*, negative predictive value; *SVM*, support vector machine.

AUCs ranging from 0.693 to 0.820^{16,34}. Xia et al. combined radiomics features with clinical-semantic factors to enhance rHE prediction, achieving an AUC of 0.830 compared to 0.690 for clinical-semantic models alone, though this study had a small sample size³⁵. In our study, with a larger two-center sample, the addition of radiomics features to the clinical-radiologic model improved rHE prediction performance in the external-testing set, consistent with previous findings^{16,35}. However, both combined models exhibited reduced generalizability, likely due to the limited robustness of handcrafted radiomics features, which suffer from low reproducibility across different CT devices and protocols^{36,37}. Furthermore, radiomics features may fail to capture the semantic characteristics of NCCT markers¹⁶. In contrast, deep learning automatically learns complex, discriminative features directly from images through neural network layers, eliminating the need for manual extraction of hard-coded features¹⁸. Most studies have focused on using deep learning models to predict cHE^{20,21,38–40}. In these studies, the follow-up hematoma volume may include both parenchymal hemorrhage and IVH hemorrhage. However, IVH expansion may occur independently of parenchymal hematoma, a factor often overlooked in large dataset studies, such as those by Li²⁰ and Teng⁴⁰ which limits confidence in deep learning's ability to predict rHE risk. Our results demonstrate that 2D CNN models based on baseline NCCT images outperform traditional ML models, suggesting that 2D deep learning may significantly enhance predictive accuracy for rHE.

In this study, we developed eight deep learning models to predict rHE, with the 2D-CNN models outperforming the 3D-CNN models in the testing sets. The differences in performance among the different 3D-CNN or 2D-CNN models may be attributed to the differing internal architectures of each network⁴¹. Previous studies have demonstrated that 3D images, which contain richer 3D spatial information compared

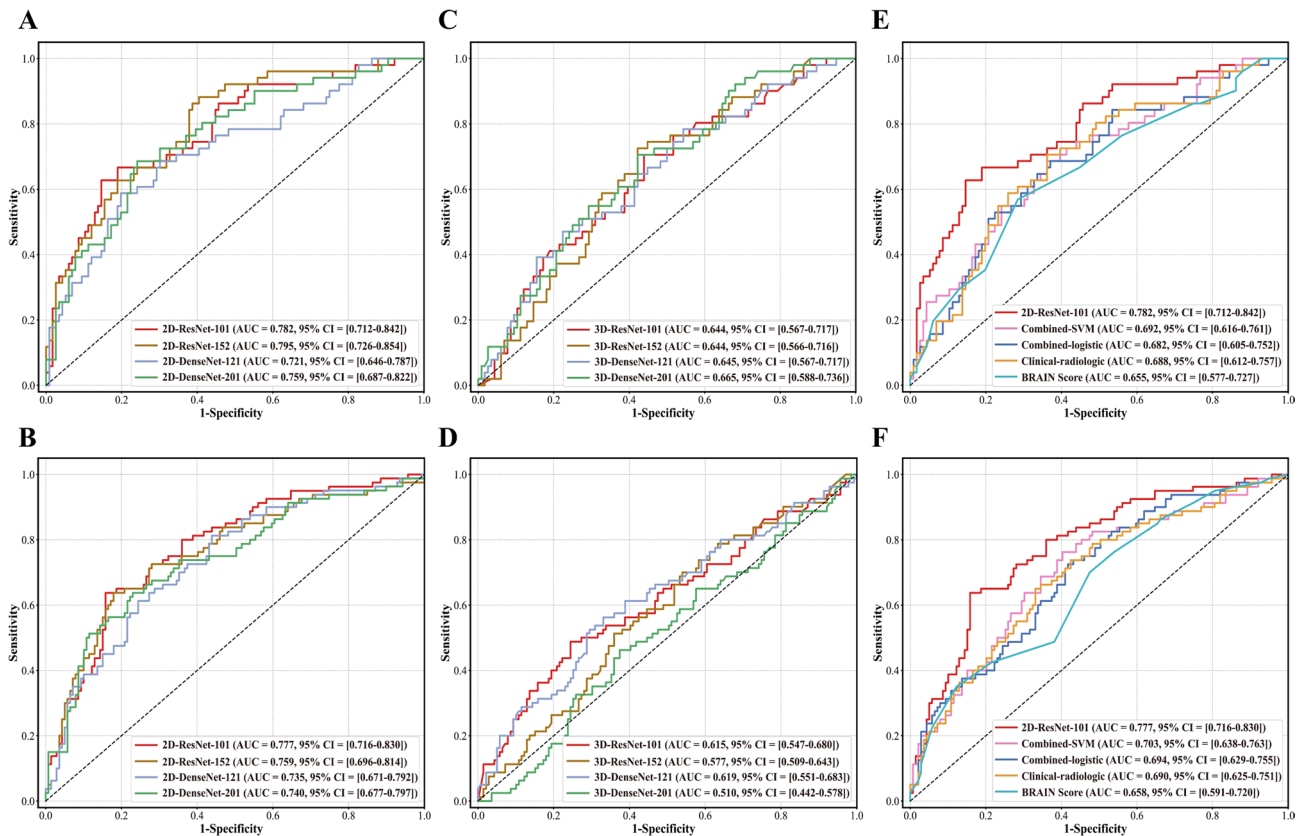


Fig. 4. Receiver operating characteristic (ROC) curves of different prediction models in the testing sets. ROC curves of the 2D-CNN models in the internal-testing (A) and external-testing sets (B); ROC curves of the 3D-CNN models in the internal-testing (C) and external-testing sets (D); performance comparison between the 2D-ResNet-101 model and the four baseline machine learning models in the internal-testing (E) and external-testing (F) sets. SVM, support vector machine.

to 2D images⁴² typically achieve superior performance in disease prediction tasks⁴³. However, in our study, the 3D-CNN models exhibited limited predictive capability, possibly due to their higher complexity and larger number of parameters, which may not be well-suited for small sample sizes of 3D data^{43,44}. Additionally, the lack of pretrained model weights and the low resolution of 3D NCCT images along the z-axis (5 mm slice thickness) could have further hindered their performance^{45,46}. Although the 2D-CNN models achieved relatively high performance, their lack of spatial information may hinder accurate modeling of peri-hematoma structures. An approach that balances the advantages of both 2D and 3D modeling may optimize the trade-off between computational efficiency and model generalizability for limited datasets⁴⁷.

Among the 2D-CNN models in our study, the 2D-ResNet-101, a deep network with 101 layers utilizing residual connections, demonstrated superior predictive performance and improved generalization⁴⁸. While deeper networks can learn more complex representations, increasing depth does not always lead to better model performance due to challenges in gradient descent^{49–51}. This was further supported by our finding that, in most CNN models, greater depth reduced performance on the external-testing set (Table 3). In our study, ResNet outperformed DenseNet, possibly owing to its simpler residual structure and lower memory complexity, which may confer greater robustness under relatively small-sample conditions^{52,53}. Previous studies have shown the effectiveness of deep residual networks in ICH disease classification^{21,54}. Grad-CAM visualizations demonstrated that the 2D-ResNet-101 model primarily focused on the hematoma and its periphery for decision-making, consistent with observations reported by Zhao et al. and Trans et al.^{21,55}. Notably, rHE tends to demonstrate more irregular morphology and internal density heterogeneity compared to non-rHE (Fig. 5A). This peripheral-focused attention pattern may correspond to NCCT markers of active multifocal bleeding, such as irregular shape⁵⁶ (Fig. 5A, Case2). These findings may support for Fisher's 'avalanche model' of HE, which proposes that initial bleeding disrupts adjacent vessels, leading to surrounding secondary hemorrhage⁵⁷. Furthermore, the 2D-ResNet-101 model achieved significantly higher sensitivity than the baseline models, without significant decrease in specificity, indicating that a higher proportion of ICH patients at high risk for rHE can be identified early, thereby helping to ensure that these patients receive timely, early-stage anti-expansion treatments or surgical intervention, as needed.

This study has several limitations. First, due to its retrospective design, some important clinical parameters such as Glasgow Coma Scale scores were unavailable. Therefore, a prospective study is necessary to validate the deep learning model's performance and further explore the relationship between rHE and clinical variables.

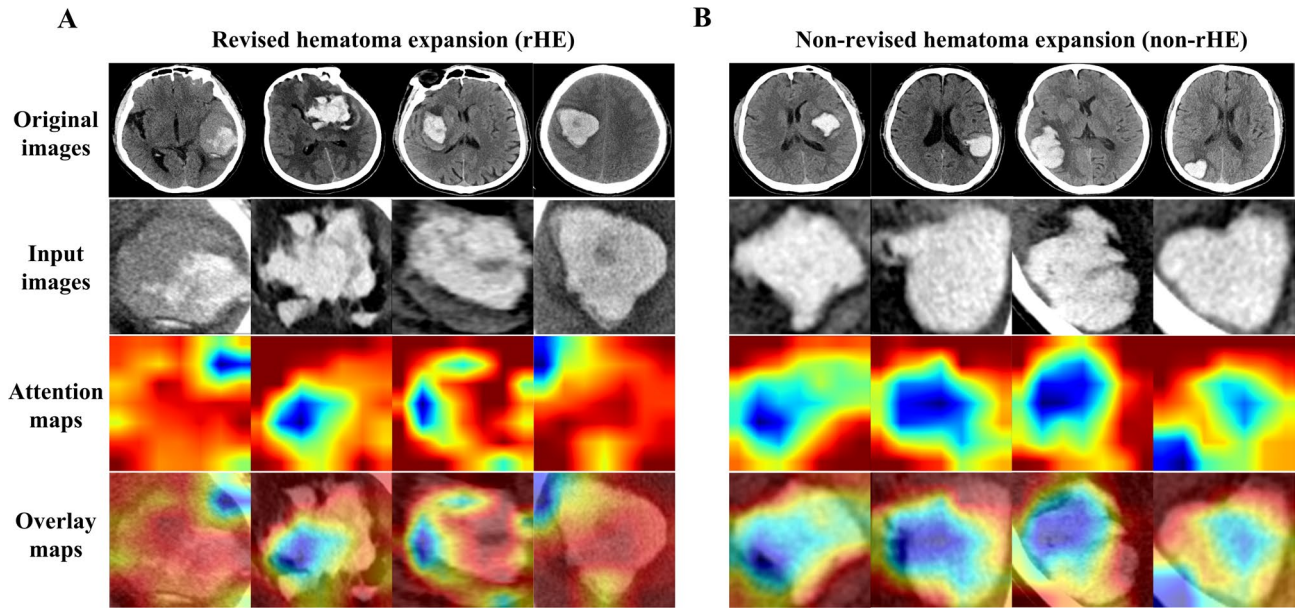


Fig. 5. Visualization of attention maps for the 2D-ResNet-101 model. The gradient-weighted class activation mapping (Grad-CAM) of the 2D-ResNet-101 deep learning model in intracerebral hemorrhage patients with revised hematoma expansion (rHE) (A) and non-rHE (B). To visually verify the decision-making process of the 2D-ResNet-101, Grad-CAM was applied to generate attention maps. These maps were then overlaid onto the input images, producing final overlay attention maps that highlight the hematoma regions most impacting the model predictions. The red regions indicate areas most influential to the model’s classification, primarily located at the hematoma and its periphery regions.

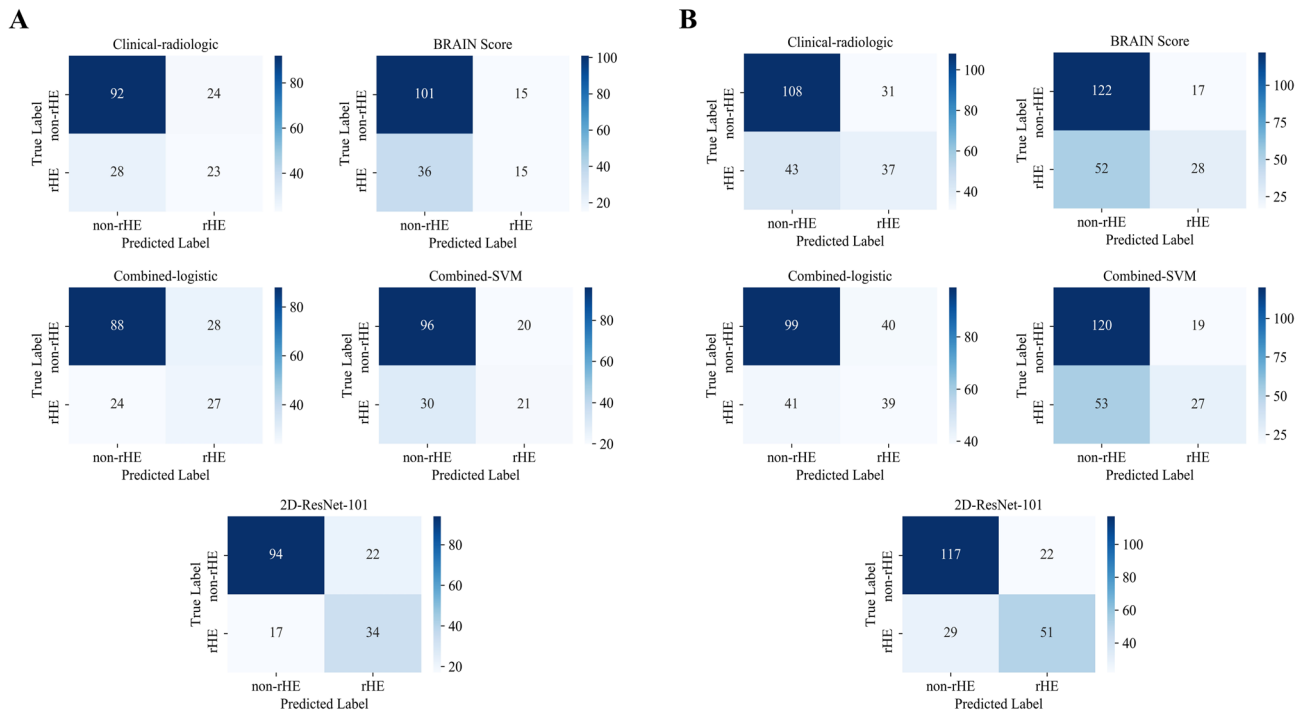


Fig. 6. Confusion matrices of 2D ResNet-101 and baseline models in the internal-testing (A) and external-testing sets (B).

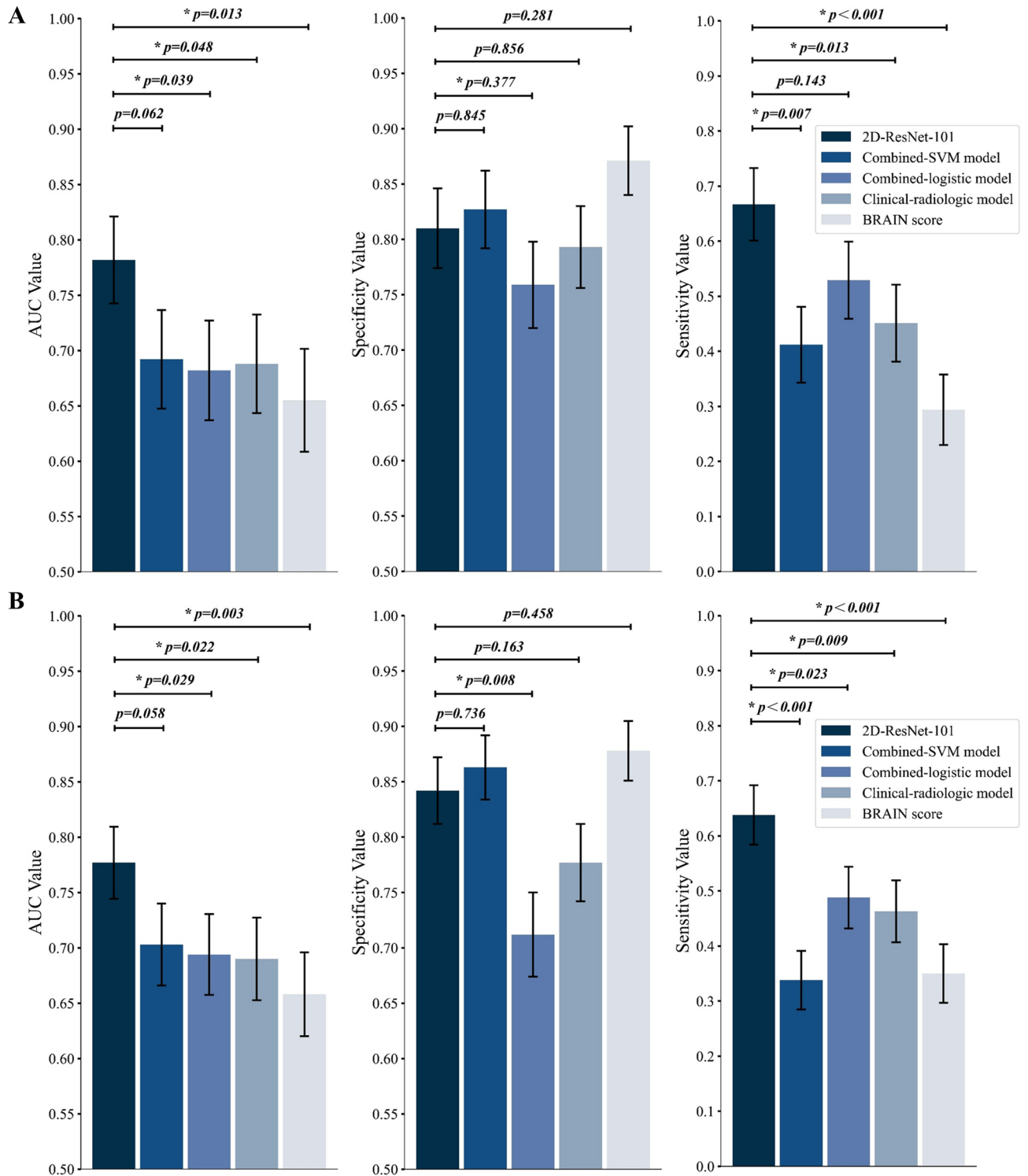


Fig. 7. Comparison of prediction performance between the 2D-ResNet-101 model and baseline machine learning models. Comparison of AUC, sensitivity, and specificity between the 2D-ResNet-101 model and four baseline prediction models in the internal- (A) and external-testing sets (B). McNemar’s chi-squared test was used to compare differences in sensitivity and specificity, while the DeLong test was used to assess the differences between AUC. *Indicate a significant difference ($p < 0.05$).

Second, the relatively small sample size limits the generalizability of the findings. A multi-center trial with larger datasets is essential to assess the model’s applicability in real-world clinical settings. Third, while the current standard for rHE relies on semiautomatic delineation software with manual adjustment, detecting small volume changes, particularly in IVH expansion (≥ 1 mL), can be challenging due to technological limitations.

Implementing fully automated, high-precision IVH delineation may enhance accuracy and reduce human error. Finally, the developed deep learning models primarily focused on image-based predictions without incorporating clinical-radiologic variables. However, medical decisions are multifactorial and not solely based on imaging findings. Future research should aim to integrate these variables to further improve model performance.

Conclusion

In conclusion, the developed 2D deep learning prediction model outperformed traditional baseline ML models in predicting rHE in ICH patients. This model can serve as a potential risk stratification tool by identifying high-risk patients, thereby providing timely targeted medical interventions to streamline decision-making in emergency conditions.

Data availability

The original data that support the findings of this study are available from the corresponding author upon reasonable request.

Received: 5 July 2025; Accepted: 22 August 2025

Published online: 31 August 2025

References

- Dowlatsahi, D. et al. Defining hematoma expansion in intracerebral hemorrhage: relationship with patient outcomes. *Neurology* **76**, 1238–1244. <https://doi.org/10.1212/WNL.0b013e3182143317> (2011).
- Morotti, A. et al. Intracerebral haemorrhage expansion: definitions, predictors, and prevention. *Lancet Neurol.* **22**, 159–171. [https://doi.org/10.1016/S1474-4422\(22\)00338-6](https://doi.org/10.1016/S1474-4422(22)00338-6) (2023).
- Wang, X. et al. Clinical prediction algorithm (BRAIN) to determine risk of hematoma growth in acute intracerebral hemorrhage. *Stroke* **46**, 376–381. <https://doi.org/10.1161/STROKEAHA.114.006910> (2015).
- Witsch, J. et al. Intraventricular hemorrhage expansion in patients with spontaneous intracerebral hemorrhage. *Neurology* **84**, 989–994. <https://doi.org/10.1212/WNL.0000000000001344> (2015).
- Dowlatsahi, D. et al. Do intracerebral hemorrhage nonexpanders actually expand into the ventricular space? *Stroke* **49**, 201–203. <https://doi.org/10.1161/STROKEAHA.117.018716> (2018).
- Roh, D. et al. Intraventricular hemorrhage expansion in the CLEAR III trial: A post hoc exploratory analysis. *Stroke* **53**, 1847–1853. <https://doi.org/10.1161/STROKEAHA.121.037438> (2022).
- Yogendrakumar, V. et al. Redefining hematoma expansion with the inclusion of intraventricular hemorrhage growth. *Stroke* **51**, 1120–1127. <https://doi.org/10.1161/STROKEAHA.119.027451> (2020).
- Li, Q. et al. Intraventricular hemorrhage growth: definition, prevalence and association with hematoma expansion and prognosis. *Neurocrit. Care.* **33**, 732–739. <https://doi.org/10.1007/s12028-020-00958-8> (2020).
- Moullaali, T. J. et al. Early Lowering of blood pressure after acute intracerebral haemorrhage: a systematic review and meta-analysis of individual patient data. *J. Neurol. Neurosurg. Psychiatry.* **93**, 6–13. <https://doi.org/10.1136/jnnp-2021-327195> (2022).
- Greenberg, S. M. et al. 2022 guideline for the management of patients with spontaneous intracerebral hemorrhage: a guideline from the American heart association/american stroke association. *Stroke* **53**, e282–e361. <https://doi.org/10.1161/STR.0000000000000407> (2022).
- Al-Shahi, S. et al. Absolute risk and predictors of the growth of acute spontaneous intracerebral haemorrhage: a systematic review and meta-analysis of individual patient data. *Lancet Neurol.* **17**, 885–894. [https://doi.org/10.1016/S1474-4422\(18\)30253-9](https://doi.org/10.1016/S1474-4422(18)30253-9) (2018).
- Ducroux, C. et al. NCCT markers of intracerebral hemorrhage expansion using revised criteria: an external validation of their predictive accuracy. *AJNR Am. J. Neuroradiol.* **44**, 658–664. <https://doi.org/10.3174/ajnr.A7871> (2023).
- Morotti, A. et al. Standards for detecting, interpreting, and reporting Noncontrast computed tomographic markers of intracerebral hemorrhage expansion. *Ann. Neurol.* **86**, 480–492. <https://doi.org/10.1002/ana.25563> (2019).
- Zhang, H. et al. Artificial intelligence-based classification of breast lesion from contrast enhanced mammography: a multicenter study. *Int. J. Surg.* **110**, 2593–2603. <https://doi.org/10.1097/JS9.0000000000001076> (2024).
- Liu, Y. et al. Machine learning combined with multi-omics to identify immune-related lncRNA signature as biomarkers for predicting breast cancer prognosis. *Sci. Rep.* **15**, 23863. <https://doi.org/10.1038/s41598-025-10186-9> (2025).
- Feng, C. et al. Prediction of early hematoma expansion of spontaneous intracerebral hemorrhage based on deep learning radiomics features of Noncontrast computed tomography. *Eur. Radiol.* **34**, 2908–2920. <https://doi.org/10.1007/s00330-023-10410-y> (2024).
- Yu, B. et al. Predicting hematoma expansion after intracerebral hemorrhage: a comparison of clinician prediction with deep learning radiomics models. *Neurocrit. Care.* <https://doi.org/10.1007/s12028-025-02214-3> (2025). Advance online publication.
- LeCun, Y., Bengio, Y. & Hinton, G. Deep learning. *Nature* **521**, 436–444. <https://doi.org/10.1038/nature14539> (2015).
- Zhong, J. W. et al. Deep learning for automatically predicting early haematoma expansion in Chinese patients. *Stroke Vasc Neurol.* **6**, 610–614. <https://doi.org/10.1136/svn-2020-000647> (2021).
- Li, N. et al. A deep learning-based framework for predicting intracerebral hematoma expansion using head non-contrast CT scan. *Acad. Radiol.* **32**, 347–358. <https://doi.org/10.1016/j.acra.2024.07.039> (2025).
- Tran, A. T. et al. Uncertainty-aware deep-learning model for prediction of supratentorial hematoma expansion from admission non-contrast head computed tomography scan. *NPJ Digit. Med.* **7**, 26. <https://doi.org/10.1038/s41746-024-01007-w> (2024).
- Eslami, V. et al. Influence of intracerebral hemorrhage location on outcomes in patients with severe intraventricular hemorrhage. *Stroke* **50**, 1688–1695. <https://doi.org/10.1161/STROKEAHA.118.024187> (2019).
- Liu, C. et al. Necrostatin-1 decreases necroptosis and inflammatory markers after intraventricular hemorrhage in mice. *Neural Regen Res.* **17**, 2710–2716. <https://doi.org/10.4103/1673-5374.339488> (2022).
- Selvaraju, R. R. et al. Grad-CAM: Visual explanations from deep networks via gradient-based localization. In: *IEEE International Conference on Computer Vision (ICCV)* 618–626 (IEEE, 2017).
- Lioutas, V. A. et al. Assessment of incidence and risk factors of intracerebral hemorrhage among participants in the Framingham heart study between 1948 and 2016. *JAMA Neurol.* **77**, 1252–1260. <https://doi.org/10.1001/jamaneurol.2020.1512> (2020).
- GBD 2019 Stroke Collaborators. Global, regional, and National burden of stroke and its risk factors, 1990–2019: a systematic analysis for the global burden of disease study 2019. *Lancet Neurol.* **20**, 795–820. [https://doi.org/10.1016/S1474-4422\(21\)00252-0](https://doi.org/10.1016/S1474-4422(21)00252-0) (2021).
- Maas, M. B. et al. Delayed intraventricular hemorrhage is common and worsens outcomes in intracerebral hemorrhage. *Neurology* **80**, 1295–1299. <https://doi.org/10.1212/WNL.0b013e31828ab2a7> (2013).
- Yogendrakumar, V. et al. New and expanding ventricular hemorrhage predicts poor outcome in acute intracerebral hemorrhage. *Neurology* **93**, e879–e888. <https://doi.org/10.1212/WNL.0000000000000807> (2019).

29. Ogut, E., Armagan, K. & Tufekci, D. The Guillain-Mollaret triangle: a key player in motor coordination and control with implications for neurological disorders. *Neurosurgical Rev.* **46**, 181. <https://doi.org/10.1007/s10143-023-02086-1> (2023).
30. Morotti, A. et al. Using Noncontrast computed tomography to improve prediction of intracerebral hemorrhage expansion. *Stroke* **54**, 567–574. <https://doi.org/10.1161/STROKEAHA.122.041302> (2023).
31. Boulouis, G. et al. Association between hypodensities detected by computed tomography and hematoma expansion in patients with intracerebral hemorrhage. *JAMA Neurol.* **73**, 961–968. <https://doi.org/10.1001/jamaneurol.2016.1218> (2016).
32. Song, L. et al. Combining non-Contrast CT signs with onset-to-imaging time to predict the evolution of intracerebral hemorrhage. *Korean J. Radiol.* **25**, 166–178. <https://doi.org/10.3348/kjr.2023.0591> (2024).
33. Nehme, A. et al. Non-contrast CT markers of intracerebral hematoma expansion: a reliability study. *Eur. Radiol.* **32**, 6126–6135. <https://doi.org/10.1007/s00330-022-08710-w> (2022).
34. Pszczolkowski, S. et al. Quantitative CT radiomics-based models for prediction of haematoma expansion and poor functional outcome in primary intracerebral haemorrhage. *Eur. Radiol.* **31**, 7945–7959. <https://doi.org/10.1007/s00330-021-07826-9> (2021).
35. Xia, X. et al. Radiomics for predicting revised hematoma expansion with the inclusion of intraventricular hemorrhage growth in patients with supratentorial spontaneous intraparenchymal hematomas. *Ann. Transl. Med.* **10**, 8. <https://doi.org/10.21037/atm-21-6158> (2022).
36. Meyer, M. et al. Reproducibility of CT radiomic features within the same patient: influence of radiation dose and CT reconstruction settings. *Radiology* **293**, 583–591. <https://doi.org/10.1148/radiol.2019190928> (2019).
37. Lee, S. B. et al. Deep learning-based image conversion improves the reproducibility of computed tomography radiomics features: a Phantom study. *Invest. Radiol.* **57**, 308–317. <https://doi.org/10.1097/RLI.0000000000000839> (2022).
38. Yalcin, C. et al. Hematoma expansion prediction in intracerebral hemorrhage patients by using synthesized CT images in an end-to-end deep learning framework. *Comput. Med. Imaging Graph.* **117**, 102430. <https://doi.org/10.1016/j.compmedimag.2024.102430> (2024).
39. Kumar, A. et al. Predicting hematoma expansion using machine learning: an exploratory analysis of the ATACH2 trial. *J. Neurosci.* **46**(1), 123048. <https://doi.org/10.1016/j.jns.2024.123048> (2024).
40. Teng, L. et al. Artificial intelligence can effectively predict early hematoma expansion of intracerebral hemorrhage analyzing Noncontrast computed tomography image. *Front. Aging Neurosci.* **13**, 632138. <https://doi.org/10.3389/fnagi.2021.632138> (2021).
41. Fujima, N. et al. Utility of deep learning for the diagnosis of otosclerosis on Temporal bone CT. *Eur. Radiol.* **31**, 5206–5211. <https://doi.org/10.1007/s00330-020-07568-0> (2021).
42. Saleem, M. A. et al. Comparative analysis of recent architecture of convolutional neural network. *Mathematical Problems in Engineering*, 7313612. (2022). <https://doi.org/10.1155/2022/7313612> (2022).
43. Wang, 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* **67**, 102385. <https://doi.org/10.1016/j.eclinm.2023.102385> (2024).
44. Singh, S. P. et al. 3D deep learning on medical images: a review. *Sensors* **20**, 5097. <https://doi.org/10.3390/s20185097> (2020).
45. Titano, J. J. et al. Automated deep-neural-network surveillance of cranial images for acute neurologic events. *Nat. Med.* **24**, 1337–1341. <https://doi.org/10.1038/s41591-018-0147-y> (2018).
46. Yang, J. et al. AlignShift: Bridging the gap of imaging thickness in 3D anisotropic volumes. In: Martel, A.L., Abolmaesumi, P., Stoyanov, D. et al. (eds.) *Medical Image Computing and Computer Assisted Intervention (MICCAI)* 562–572 (Springer, Cham, 2020).
47. Zhang, Y., Liao, Q., Ding, L. & Zhang, J. Bridging 2D and 3D segmentation networks for computation-efficient volumetric medical image segmentation: an empirical study of 2.5D solutions. *Comput. Med. Imaging Graph.* **99**, 102088. <https://doi.org/10.1016/j.compmedimag.2022.102088> (2022).
48. He, K., Zhang, X., Ren, S. & Sun, J. Deep residual learning for image recognition. In: *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)* 770–778 (IEEE, 2016).
49. Nagpal, P., Bhinge, S. A. & Shitole, A. A comparative analysis of ResNet architectures. In: *International Conference on Smart Generation Computing, Communication and Networking (SMART GENCON)* 1–8 (IEEE, 2022).
50. Gupta, R. & Jindal, R. Impact of too many neural network layers on overfitting. *Int. J. Comput. Sci. Mob. Comput.* **14**, 1–14. <https://doi.org/10.47760/ijcsmc.2025.v14i05.001> (2025).
51. Tan, M., Le, Q. V. & EfficientNet Rethinking model scaling for convolutional neural networks. In: *Proceedings of the 36th International Conference on Machine Learning (ICML)* 6105–6114 (PMLR, 2019).
52. Li, G. Discussion on image recognition under different conditions of ResNet and densenet. *Appl. Comput. Eng.* **4**, 636–641. <https://doi.org/10.54254/2755-2721/4/2023365> (2023).
53. Zhang, C. et al. IEEE., ResNet or DenseNet? Introducing dense shortcuts to ResNet. In: *Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision (WACV)* 3543–3552 (2021).
54. Zhou, Q. et al. Transfer learning of the ResNet-18 and DenseNet-121 model used to diagnose intracranial hemorrhage in CT scanning. *Curr. Pharm. Des.* **28**, 287–295. <https://doi.org/10.2174/1381612827666211213143357> (2022).
55. Zhao, X. et al. Explainable CT-based deep learning model for predicting hematoma expansion including intraventricular hemorrhage growth. *iScience*, 112888. (2025). <https://doi.org/10.1016/j.isci.2025.112888>
56. Barras, C. D. et al. Density and shape as CT predictors of intracerebral hemorrhage growth. *Stroke* **40**, 1325–1331. <https://doi.org/10.1161/STROKEAHA.108.536888> (2009).
57. Fisher, C. M. Pathological observations in hypertensive cerebral hemorrhage. *J. Neuropathol. Exp. Neurol.* **30**, 536–550. <https://doi.org/10.1097/00005072-197107000-00015> (1971).

Acknowledgements

The authors thank the American Journal Experts (AJE) for their assistance with language editing.

Author contributions

Conceived and designed the analysis: J Peng, Q Yu, Y Ning; Provision of study materials or patients: F Lv, X Xu, J Peng; Collected the data: K Xie, H Jiang, Y Zhou, R Liu, W Jiang, X Chen; Contributed data or analysis tools: X Zhang, Q Yu., Y Ning, X Fan; Writing - review and editing: Y Ning, Q Yu, X Fan, J Peng; All Authors Approve the Final Version of the Manuscript.

Funding

This work was supported by the Foundation of Science and Technology Bureau of Yuzhong District, Chongqing, China (Grant No. 20190111), the Natural Science Foundation of Chongqing, China (Grant No. cstc2021j-cyj-msxmX0020).

Declarations

Competing interests

The authors declare no competing interests.

Additional information

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1038/s41598-025-17393-4>.

Correspondence and requests for materials should be addressed to F.L., X.X. or J.P.

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

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>.

© The Author(s) 2025