Appendix 1

Explanation of Inclusion and Exclusion Criteria

In this study, we exclusively included patients undergoing relevant pharmacological therapy. This decision was made for several reasons. Firstly, pharmacological therapy is the preferred and most commonly used treatment approach for patients with symptomatic intracranial atherosclerotic stenosis (sICAS), as recommended by clinical guidelines. Secondly, we postulated that endovascular treatment may alter the local vascular environment, causing some degree of damage to the vascular wall. Additionally, endovascular procedures can affect hemodynamic conditions postoperatively, and such alterations in the vascular milieu play a crucial role in patient prognosis.

Furthermore, patients with carotid artery stenosis greater than 50% were excluded from the study. Numerous studies have shown that carotid stenosis exceeding this threshold can significantly alter intracranial hemodynamic conditions, which in turn have important implications for the force exerted on intracranial plaques and the progression of plaque changes. Therefore, patients with such hemodynamic alterations were excluded.

Lastly, patients with cardioembolic stroke were also excluded. Cardioembolic stroke is fundamentally different from intracranial atherosclerotic plaques in terms of pathology and physiology. This study focuses on the compositional evolution and risk assessment during plaque progression, which is not applicable to patients with cardioembolic stroke. Moreover, there are significant differences in plaque composition and treatment strategies between cardioembolic stroke and atherosclerotic plaques, justifying the exclusion of cardioembolic stroke patients.

Image Plaque Analysis

The software TS-Vessel·Explore (TSimaging; Healthcare China, Beijing) is used for post-processing and data analysis. The HR-VWI images are imported into the postprocessing software and reconstructed along a cross-section perpendicular to the long axis of the vessel, with a magnification of 400%. A manual tracing model is employed to outline the vessel wall's outer contour and the lumen's inner contour, with the software automatically measuring the corresponding vessel area (VA) and lumen area (LA). The plaque with the narrowest stenosis at the lumen was chosen as the culprit plaque for measurement.

The plaque responsible for the narrowest part of the lumen is selected for measurement. For the reference level, the VA and LA are prioritized from the corresponding lumen section without obvious plaques near it, followed by the related lumen section far away.

The degree of vascular stenosis is calculated using the following formulas: stenosis rate=(1–LAmin/LAreference)×100%; wall area (WA)=VA-LA; plaque area (PA)=WAmin – WAreference; remodeling index (RI)=VAmin/VAreference, with RI≥1.05 indicating positive remodeling and RI≤0.95 indicating negative remodeling

The NWI is calculated as WA/VA. In the enhanced T1W VISTA image, the signal intensity of the pituitary gland is used as the reference, with no change in plaque signal considered grade 0; enhancement lower than the pituitary gland, grade 1; and similar enhancement to the normal pituitary gland, grade 2. Intraplaque hemorrhage (IPH) is defined as T1-weighted imaging (T1WI) signal intensity higher than 150% of the adjacent muscle tissue signal.

Calculation Method for Model Performance

1. Area Under the Curve

The area under the curve (AUC) is associated with the receiver operating characteristic (ROC) curve. The ROC curve plots the true positive rate (Sensitivity) against the false positive rate (1 – Specificity) at various threshold settings of a diagnostic test or predictive model.

ROC Curve: The ROC curve demonstrates the trade-off between sensitivity and specificity for a test as the decision threshold varies. An AUC score quantifies the overall ability of the model to discriminate between positive and negative cases.

AUC Calculation: AUC is the integral of the ROC curve, representing the probability that a randomly chosen positive instance is ranked higher than a randomly chosen negative instance by the model.

https://dx.doi.org/10.21037/qims-24-1723

2. Sensitivity (True Positive Rate or Recall)

Sensitivity is the proportion of actual positive cases that are correctly identified by the test. It reflects the ability of the test to correctly detect those patients who have the condition being tested for.

$$\text{Sensitivity}=\frac{\text{True Positives}\left(\text{TP}\right)}{\text{True Positives}\left(\text{TP}\right)+\text{False Negatives}\left(\text{FN}\right)}$$

Interpretation: Sensitivity ranges from 0 to 1. A sensitivity of 1 (100%) means that the test identifies all actual positive cases, whereas a sensitivity of 0 indicates none of the true positive cases were identified. A highly sensitive test is essential when the goal is to minimize the number of false negatives (e.g., in screening for life-threatening diseases where missing a positive case has serious consequences).

3.Specificity (True Negative Rate)

Specificity is the proportion of actual negative cases that are correctly identified by the test. It quantifies the ability of the test to exclude individuals who do not have the condition.

Formula:

$$\text{Specificity}=\frac{\text{True Negatives}\left(\text{TN}\right)}{\text{True Negatives}\left(\text{TN}\right)+\text{False Positives}\left(\text{FP}\right)}$$

Interpretation: Specificity also ranges from 0 to 1. A test with high specificity ensures that most of the individuals who do not have the disease are correctly classified as negative. Specificity is particularly critical when the aim is to minimize the number of false positives, which is vital in situations where unnecessary treatment or intervention should be avoided.

4. Accuracy

Accuracy represents the overall correctness of the test, reflecting the proportion of true results (both true positives and true negatives) among the total number of cases examined. It provides a general indication of test performance but does not distinguish between sensitivity and specificity.

Formula:

$$\text{Accuracy}=\frac{\text{True Positives}\left(\text{TP}\right)+\text{True Negatives}\left(\text{TN}\right)}{\text{True Positives}\left(\text{TP}\right)+\text{True Negatives}\left(\text{TN}\right)+\text{False Positives}\left(\text{FP}\right)+\text{False Negatives}\left(\text{FN}\right)}$$

Interpretation: Accuracy is expressed as a percentage or a proportion. Although accuracy gives an overall picture of test performance, it may be misleading if the dataset is imbalanced (i.e., when one class is much more prevalent than the other). In such cases, accuracy may remain high even when the model performs poorly for the minority class, necessitating the use of other metrics such as sensitivity, specificity, or the F1 score for more nuanced evaluation.

Table S1 Summary of Imaging Protocol (PHILIPS Ingenia Elition 3.0T)

Sequence	HRVWI		Head MRI			
	3D T1WI	3D TOF	Axial T1WI	Axial T2WI	DWI	FLAIR
FOV (mm)	250×161	200×158	230×189	230×230	230×230	230×190
Matrix	312×201	308×168	288×178	244×244	152×122	308×203
Slice Thickness (mm)	0.6	1.2	5.0	5.0	5.0	5.0
Resolution	0.8×0.8×0.8	0.65×0.94×1.2	0.8×1.05	0.95×0.95	1.5×1.89	0.75×0.91
TR/TE (ms)	600/31	19/3.5	2373/20	2756/105	2194/86	8000/120

TOF, time of flight; DWI, diffusion-weighted imaging; FLAIR, fluid-attenuated inversion recovery; FOV, field of view; TR, repetition time; TE, echo time.

https://dx.doi.org/10.21037/qims-24-1723

Table S2 Summary of Imaging Protocol (SIEMENS MAGNETOM Vida)

Sequence	HRVWI		Head MRI			
	3D T1WI	3D TOF	Axial T1WI	Axial T2WI	DWI	FLAIR
FOV (mm)	220×188	200×181	230×230	230×230	230×230	230×230
Matrix	220×188	200×181	230×230	230×207	230×230	230×184
Slice Thickness (mm)	0.69	0.7	5.0	5.0	5.0	5.0
Resolution	0.7×0.7×0.7	0.3×0.3×0.7	0.8×0.8	0.4×0.4	0.7×0.7	0.8×0.8
TR/TE (ms)	900/28	21/3.42	2000/7.5	5000/92	5000/70	8000/97

TOF, time of flight; DWI, diffusion-weighted imaging; MRI, magnetic resonance imaging; FLAIR, fluid-attenuated inversion recovery; FOV, field of view; TR, repetition time; TE, echo time.

Table S3 Radiomics Features' Selection Results

Radiomic features	LASSO coefficient (β)
gradient_gldm_DependenceNonUniformity	0.016283
gradient_glrlm_ShortRunLowGrayLevelEmphasis	-0.029620
lbp_3D_k_glszm_SizeZoneNonUniformity	0.001732
lbp_3D_m2_firstorder_Range	-0.014540
log_sigma_1_0_mm_3D_glszm_GrayLevelNonUniformityNormalized	-0.015395
log_sigma_1_0_mm_3D_ngtdm_Strength	0.014962
log_sigma_2_0_mm_3D_firstorder_90Percentile	-0.030313
log_sigma_2_0_mm_3D_glszm_SmallAreaEmphasis	0.004545
log_sigma_3_0_mm_3D_ngtdm_Strength	-0.003265
log_sigma_4_0_mm_3D_firstorder_Skewness	-0.042164
log_sigma_5_0_mm_3D_glrlm_LongRunLowGrayLevelEmphasis	-0.003025
original_shape_Maximum2DDiameterColumn	-0.034781
square_ngtdm_Busyness	0.013812
squareroot_glcm_DifferenceVariance	0.035640
wavelet_HHH_firstorder_Median	0.019370
wavelet_HHH_firstorder_Skewness	-0.002553
wavelet_HHH_glcm_ClusterProminence	-0.005991
wavelet_HHH_glszm_LowGrayLevelZoneEmphasis	-0.009503
wavelet_HHL_glcm_Idn	0.004682
wavelet_HLH_firstorder_Mean	0.003729
wavelet_HLH_firstorder_Skewness	0.001049
wavelet_HLH_glcm_MCC	-0.024717
wavelet_HLH_gldm_LargeDependenceEmphasis	0.008992
wavelet_HLL_gldm_LargeDependenceEmphasis	-0.011419
wavelet_HLL_glrlm_RunVariance	-0.000978
wavelet_LHH_firstorder_RobustMeanAbsoluteDevia	-0.001155
wavelet_LHH_gldm_DependenceVariance	0.002380
wavelet_LHL_glszm_ZoneVariance	0.016064
wavelet_LHL_ngtdm_Strength	-0.006098
wavelet_LLH_firstorder_Kurtosis	0.008206
wavelet_LLH_glcm_ClusterShade	-0.021177
wavelet_LLH_glcm_Idn	0.032618
wavelet_LLL_firstorder_RootMeanSquared	-0.040483

LASSO, least absolute shrinkage and selection operator.

https://dx.doi.org/10.21037/qims-24-1723

Table S4 Comparison of Time Consumption between Trans-CNN Model and The Team of Radiologists

	Trans-CNN	The Team of Radiologists	Z	P-value
Time (IQR); s	142.1 (130.1, 186.4)	212.1 (177.3, 256.5)	-4.667	<0.001

IQR, interquartile range.

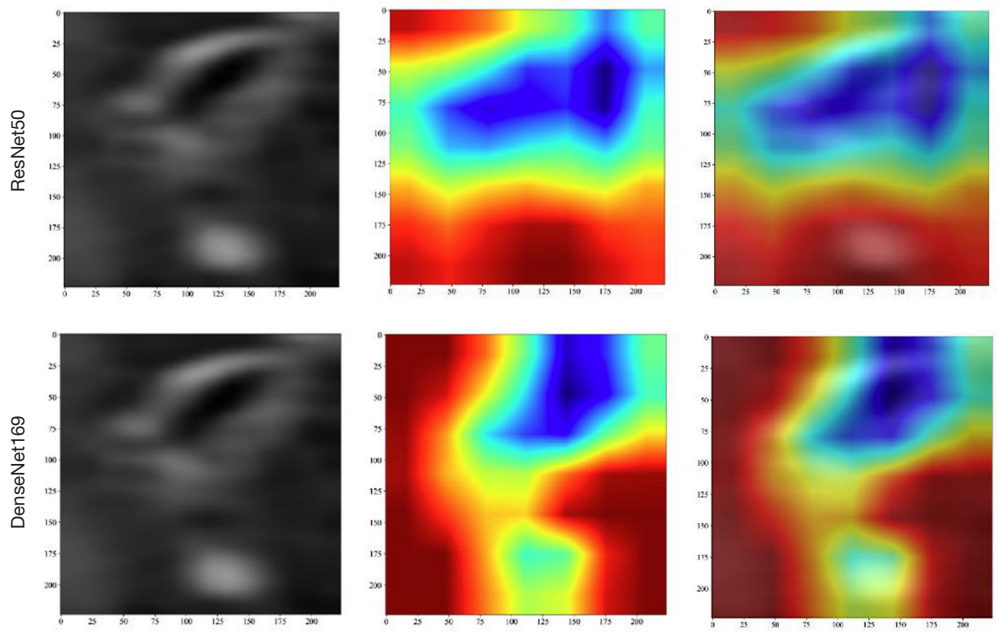

Figure S1 Grad CAM on the last convolutional layer of ResNet50 and DenseNet169 networks. The first column is the T1 enhanced image of HR-VWI, the second column is the output activation map, and the third column is the fused image of the first two columns stacked together. Among them, dark blue represents areas with high learning attention, whereas red represents areas with low learning attention. Grad CAM, Gradient weighted Class Activation Mapping; HR-VWI, High-resolution magnetic resonance vessel wall imaging.

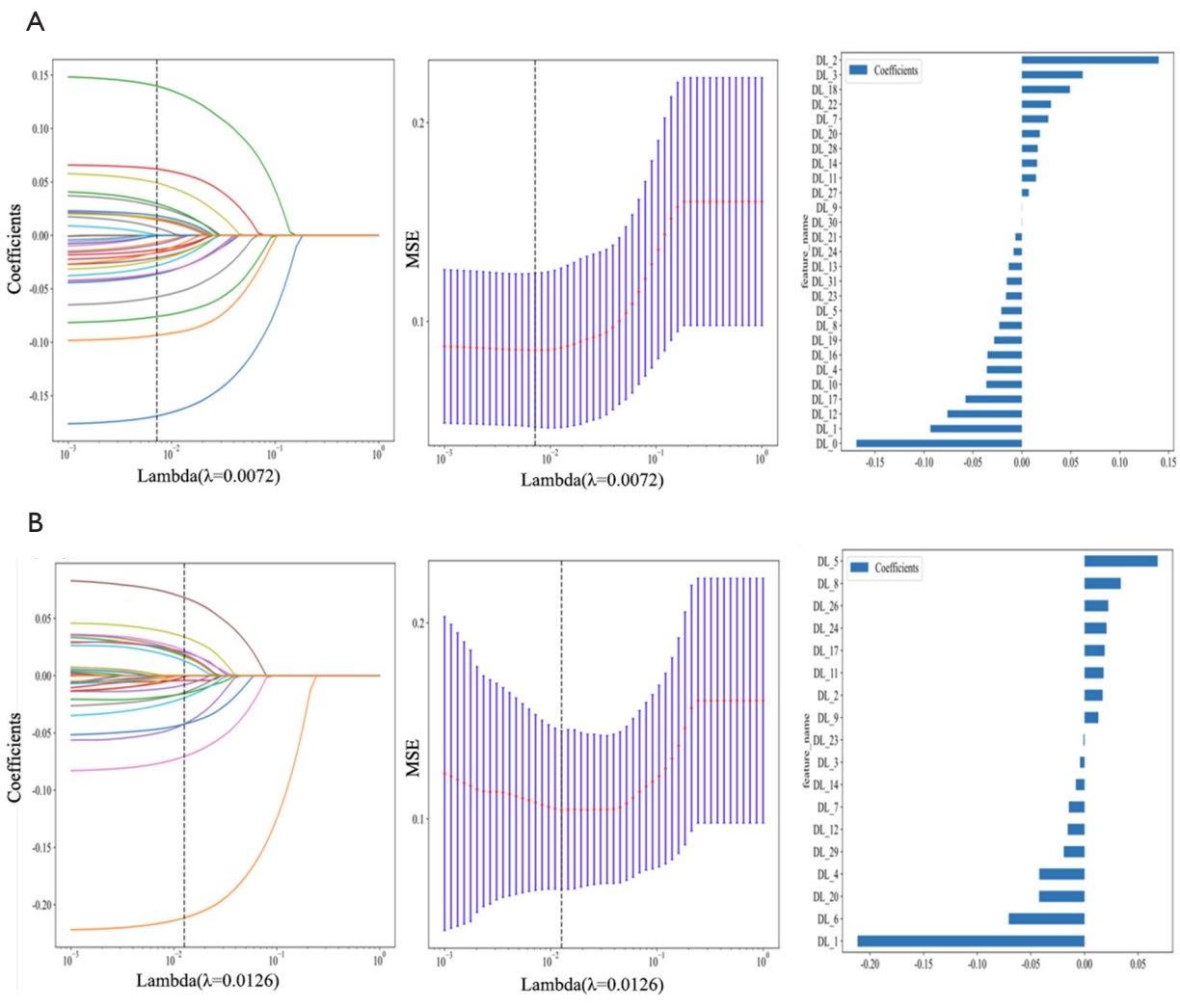

Figure S2 Deep Learning Models Feature Screening. Compressing features of DL models, each DL model ultimately obtains 32 compressed features, which are then eliminated through screening of correlation coefficients, U-test, and Pearson coefficient, and finally screened by LASSO. ResNet50 contains 18 features (B), DenseNet169 contains 27 features (A). We found that in the DenseNet169 model, when λ=0.0072, the Y-axis is minimized, and a total of 27 features are selected. In the RseNet50 model, when λ=0.0126, the Y-axis is minimized, and a total of 18 features are selected. DL, deep learning; LASSO, least absolute shrinkage and selection operator.

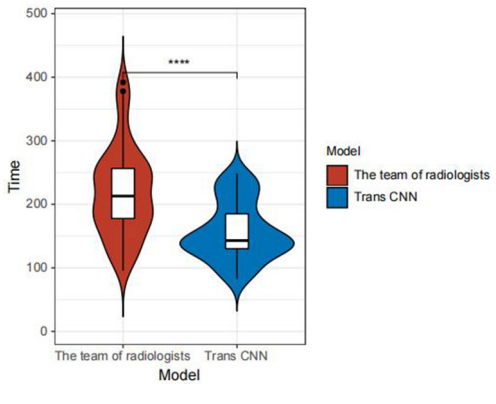

Figure S3 Violin Plot Comparing the Time Consumption of the Trans-CNN model and the team of Radiologists. The box in the figure represents the quartiles of the data, and the horizontal lines inside represent the median of the data. The shape of the violin chart indicates the density of the data distribution. (**** stands for P < 0.001). CNN, convolutional neural network.

https://dx.doi.org/10.21037/qims-24-1723