European Journal of Radiology 190 (2025) 112254

Contents lists available at ScienceDirect

European Journal of Radiology

journal homepage: www.elsevier.com/locate/ejrad

Research article

Multimodal deep learning for predicting unsuccessful recanalization in refractory large vessel occlusion

Jesús D. González [a,b,*], Pere Canals [a,b], Marc Rodrigo-Gisbert [a,b], Jordi Mayol [a,b], Alvaro García-Tornel [a], Marc Ribó [a,b]

[a] Stroke Unit, Hospital Vall d'Hebron, Barcelona, Spain
[b] Departament de Medicina, Universitat Autónoma de Barcelona, Barcelona, Spain

ARTICLE INFO

Keywords:
Acute ischemic stroke
Intracranial atherosclerosis disease
Artificial intelligence
Multimodal deep learning
Explainable AI

ABSTRACT

Purpose: This study explores a multi-modal deep learning approach that integrates pre-intervention neuroimaging and clinical data to predict endovascular therapy (EVT) outcomes in acute ischemic stroke patients. To this end, consecutive stroke patients undergoing EVT were included in the study, including patients with suspected Intracranial Atherosclerosis-related Large Vessel Occlusion ICAD-LVO and other refractory occlusions. ***Methods:*** A retrospective, single-center cohort of patients with anterior circulation LVO who underwent EVT between 2017–2023 was analyzed. Refractory LVO (rLVO) defined class, comprised patients who presented any of the following: final angiographic stenosis > 50 %, unsuccessful recanalization (eTICI 0-2a) or required rescue treatments (angioplasty +/- stenting). Neuroimaging data included non-contrast CT and CTA volumes, automated vascular segmentation, and CT perfusion parameters. Clinical data included demographics, comorbidities and stroke severity. Imaging features were encoded using convolutional neural networks and fused with clinical data using a DAFT module. Data were split 80 % for training (with four-fold cross-validation) and 20 % for testing. Explainability methods were used to analyze the contribution of clinical variables and regions of interest in the images. ***Results:*** The final sample comprised 599 patients; 481 for training the model (77, 16.0 % rLVO), and 118 for testing (16, 13.6 % rLVO). The best model predicting rLVO using just imaging achieved an AUC of 0.53 ± 0.02 and F1 of 0.19 ± 0.05 while the proposed multimodal model achieved an AUC of 0.70 ± 0.02 and F1 of 0.39 ± 0.02 in testing. ***Conclusion:*** Combining vascular segmentation, clinical variables, and imaging data improved prediction performance over single-source models. This approach offers an early alert to procedural complexity, potentially guiding more tailored, timely intervention strategies in the EVT workflow.

1. Introduction

Acute ischemic stroke due to large vessel occlusion (LVO) is a severe neurological emergency requiring fast and effective treatment. Endovascular therapy (EVT) is currently the most effective treatment, achieving successful reperfusion in up to 80–85 % of cases [1–3]. However, in patients with intracranial atherosclerotic disease (ICAD), complete recanalization is more difficult to achieve without rescue treatments such as angioplasty and stenting, leading to worse outcomes [4]. Independently from the underlying cause, patients with LVO refractory to frontline EVT might benefit from early adoption of rescue treatments. Predicting failure of recanalization with conventional thrombectomy devices such as stent retrievers or aspiration catheters, before or early into the EVT process could avoid the unsuccessful repetition of sterile attempts that could lead to life-threatening complications. Preprocedural suspicion of ICAD-related LVO (ICAD-LVO), which accounts for 10–30 % (5–10 % in Caucasian populations [4]) of cases, is primarily based on neuroimaging— non-contrast computed tomography (NCCT), computed tomography angiography (CTA), and computed tomography perfusion (CTP)—but these methods are not always definitive or available. Even digital subtraction angiography (DSA), the gold standard for reperfusion assessment can be inconclusive for ICAD diagnosis. Despite recommendations such as the European Stroke Organization's criteria to diagnose ICAD-LVO, which advise to combine clinical and radiological signs (absence of atrial fibrillation, absence of CT hyperdense or MRI susceptibility sign, or presence of

* Corresponding author at: Vall d'Hebron Institut de Recerca, Stroke Research Unit, Pg. de la Vall d'Hebron, 129, Municipality of Horta-Guinardó, 08035 Barcelona, Spain.
E-mail addresses: jesus.gonzalez.riveros@vhir.org, jesusdavid.gonzalez@autonoma.cat (J.D. González).

https://doi.org/10.1016/j.ejrad.2025.112254
Received 28 January 2025; Received in revised form 31 May 2025; Accepted 16 June 2025
Available online 18 June 2025

watershed infarction) [4,6], preoperative ICAD-LVO diagnosis remains challenging [4–6].

Focusing solely and exclusively on ICAD-LVO prediction may be difficult due to its low incidence, uncertain diagnosis, and the multifactorial nature of stroke. Many cases involve overlapping etiologies or conditions associated with resistance to recanalization despite EVT [5,6]. Expanding predictive models to include both ICAD-LVO and other non-recanalizing conditions may create a more practical framework for anticipating complex interventions and improving patient outcomes [7].

Recent studies underscore the importance of early ICAD-LVO identification [8–10]. Chen et al. [11] introduced the ATHE scale, integrating clinical and imaging features for pre-EVT prediction, while Rodrigo-Gisbert et al. [5] highlighted combining clinical variables and CTP parameters. Multimodal approaches, such as those by Wolf et al. [12], merging imaging and clinical data for dementia analysis, and self-supervised frameworks incorporating 3D imaging for stroke risk prediction [13], promising results in improving diagnostic precision. Similar strategies were applied to stroke related predictions [14,15].

This study aims to predict recanalization failure during EVT with a focus not only on ICAD-LVO but with a broader practical approach

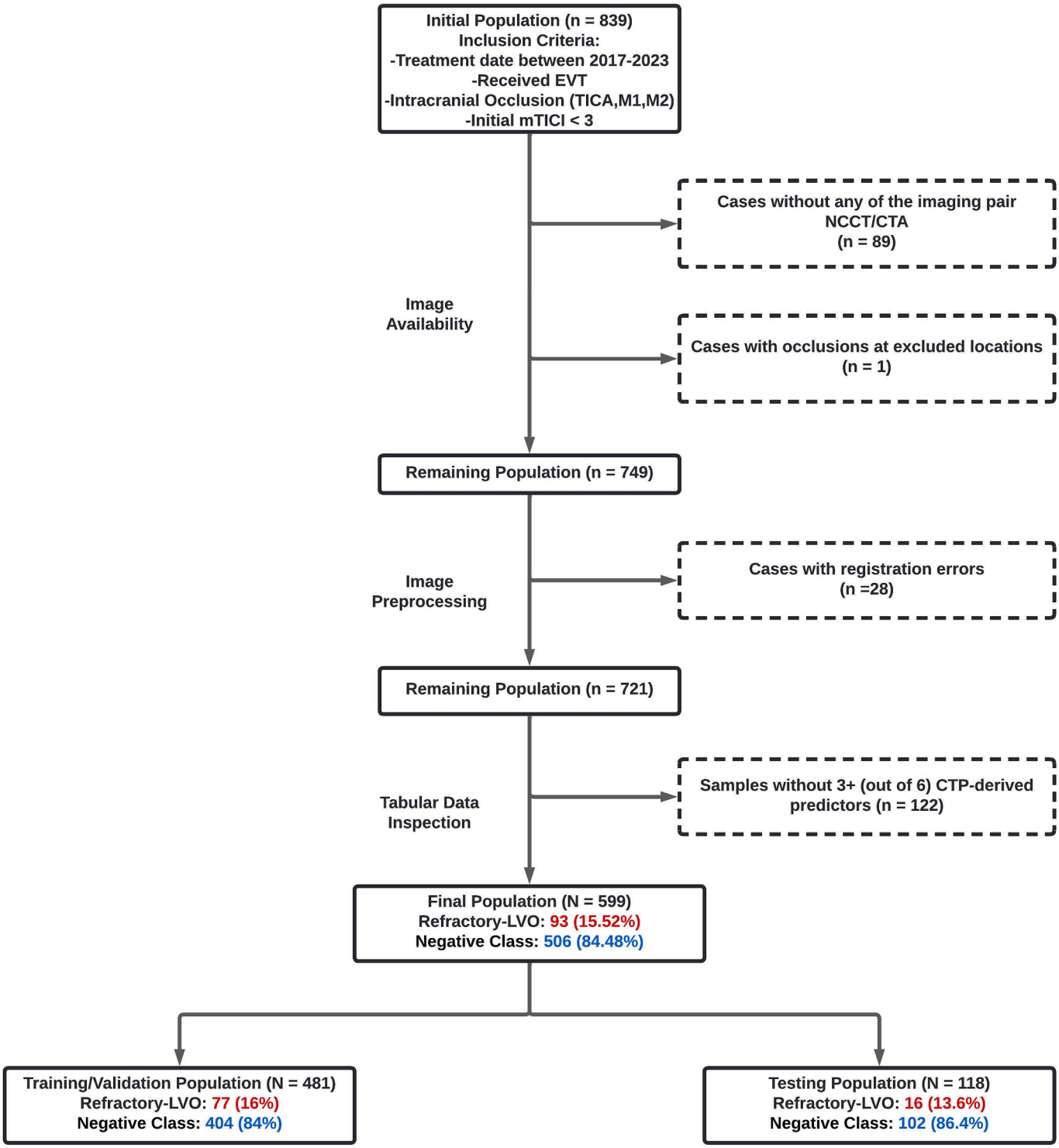

Fig. 1. Overview of Data Selection and Preprocessing Workflow.

aiming to identify any cause of no-recanalization by developing a multi-modal deep learning model. By integrating neuroimaging, clinical data, and vascular segmentation, we aim to maximize pre-procedural prediction of angiographic outcomes. Using the Dynamic Affine Feature Transform (DAFT) [12] framework to fuse data and explainability methods for interpretability, the study evaluates the added value of including clinical data over image-only models.

2. Methods

2.1. Ethics

Due to the retrospective nature of this study and the fact that all data were anonymized, the need for written informed consent was waived by the institutional review board. The study protocol was reviewed and approved by the local ethics committee, approval number (PR(AG)434/2023).

2.2. Study population

This study included patients that received EVT due to anterior circulation LVO in a high-volume comprehensive stroke center (Hospital Universitari Vall d'Hebron) between January 2017 and December 2023. Patient selection flowchart is shown in Fig. 1.

Initial exclusion criteria were lack of occlusions on baseline DSA, multiple synchronous intracranial occlusions, severe imaging artifacts and cases missing any of the required neuroimaging (NCCT, CTA, and CTP).

2.2.1. Imaging data Acquisition and preprocessing

All images were acquired using a Siemens SOMATOM Definition AS+ (Siemens, Erlangen, Germany) 128 slices CT. For CTA, iodinated IV contrast was given in a single bolus to the patient at a rate between 4 and 5 ml/s with an overall volume of 40–80 ml of contrast solution depending on the patient. For standardization purposes, all images were registered to the Montreal Neurology Institute (MNI) space using a resized atlas (246x295x258) voxel spacing (0.43x0.43x0.4 mm) to ensure alignment and uniform size. Image intensity values were clipped per CT modality: 0–100 for NCCT and 0–400 for CTA.

2.2.2. Tabular clinical data selection and preprocessing

Clinical data included demographic, clinical, and CTP-based variables [5,6,8]. Clinical recorded features included age, sex, hypertension, dyslipidemia (DL), diabetes Mellitus, smoking, atrial fibrillation (AF) and stroke severity assessed by the National Institutes of Health Stroke Scale (NIHSS) upon patient arrival, treatment with intravenous thrombolysis, unwitnessed stroke and occlusion side and site. CTP − derived variables were calculated with an automated commercially available software (RAPID® iSchemaView, Menlo Park, CA) and included: Cerebral Blood Flow relative to contralateral hemisphere < 30 %, time to maximum peak concentration > 4 s, > 6 s and > 10 s, hypoperfusion intensity ratio (HIR = Tmax > 10 s/Tmax > 6 s) and Tmax > 4 s/Tmax > 6 s ratio.These parameters have been previously associated to ICAD-LVO and recanalization success [5,17]. Clinical continuous data was normalized to a 0–1 range and preprocessed separately for training and testing to prevent data leakage.

A second exclusion criteria was applied after preprocessing: cases whose registration process failed and those who had more than 3 CTP parameters missing were excluded, to avoid extensive imputation of sensitive data.

2.3. Refractory LVO class definition

Limiting the observation to an ICAD-only cohort would fail to capture the complexity of real-world clinical practice. To ensure a more representative and practical dataset, the target population included not only clinically suspected ICAD-LVO cases but also all unsuccessful recanalization cases. Refractory-LVO (rLVO) was defined as patients with ICAD-LVO (final angiographic stenosis > 50 %) and other etiologies in whom recanalization was not achieved, determined by an expanded Thrombolysis in Cerebral Ischemia (eTICI) score of 0-2a after conventional thrombectomy treatment (aspiration and/or stent retriever). Patients requiring rescue treatments, such as angioplasty and/or stenting, were also categorized as rLVO even if they ultimately reached recanalization. The dataset was split into 80 % for training (with four-fold cross-validation) and 20 % for testing.

2.4. Multi-modal predictive model

A model capable of combining imaging and clinical data was designed for rLVO prediction. For image encoding, we employed MONAI-implemented [18] DenseNet [19] and EfficientNetB0 [20] backbones [21–23]. To integrate clinical data with imaging features, we used the DAFT module [12], which dynamically rescales and shifts the encoded imaging feature maps via fully connected layers that process both global pooled image features and clinical inputs (Fig. 2).

For clinical data alone, experiments using logistic regression were explored, initially using just baseline clinical scores (NIHSS and an additional Alberta Stroke Program Early CT Score, ASPECTS [5,11]) and then multivariate logistic regression using the complete set of available predictors described in section 2.2.2.

For imaging modalities, experiments were conducted using only NCCT, only CTA and BOTH (NCCT + CTA stacked as separate channels of the same image). This stacking approach has been used to capture interactions between modalities that are otherwise difficult to detect [24,25]. We hypothesize that NCCT may provide critical structural information, while CTA offers vascular and occlusion details. Prior vascular segmentation was also included in relevant configurations to enhance the model's focus, further reinforcing attention to clinically relevant areas [24]. Automatic vascular segmentation was obtained from CTA images with ARTERIAL™ a previously validated automated tool [16].

All experiments employed stratified 4-fold cross-validation to account for class imbalance in the dataset. Hyperparameters such as the number of epochs, batch size, and initial learning rate were configurable by the framework user. For consistency, we trained all models for 60 epochs per fold with a batch size of 1, an initial learning rate of $1 \times 10 - 4$ which was dynamically adjusted using the ReduceLROnPlateau scheduler in PyTorch. Training was conducted on an NVIDIA RTX 5000 GPU.

Final predictions were obtained through a weighted ensembled voting method, combining the probabilities of the predictions from the NCCT, CTA, and BOTH models followed by a fold-consensus majority voting approach to leverage information from all training folds to arrive at a single, clinically interpretable output, similar to the ensemble strategy used in multi-fold inference models [28]. Weights were assigned based on model-specific true positive and true negative rates in validation to address class imbalance. Inference thresholds were selected maximizing Youden's index.

2.5. Explainability

Explainability was studied using Grad-CAM both before and after the DAFT block to evaluate how the inclusion of clinical data influenced the model's focus on clinically relevant areas, such as vessel occlusions. The DAFT module integrates clinical and imaging features by dynamically modifying the feature maps through channel-wise scaling (S) and shifting (B) operations computed as follows [12]:

$$F_{new} = F \bullet S + B \quad (1)$$

where F is the feature map from the imaging data (shape: [B, C, D, H, W),

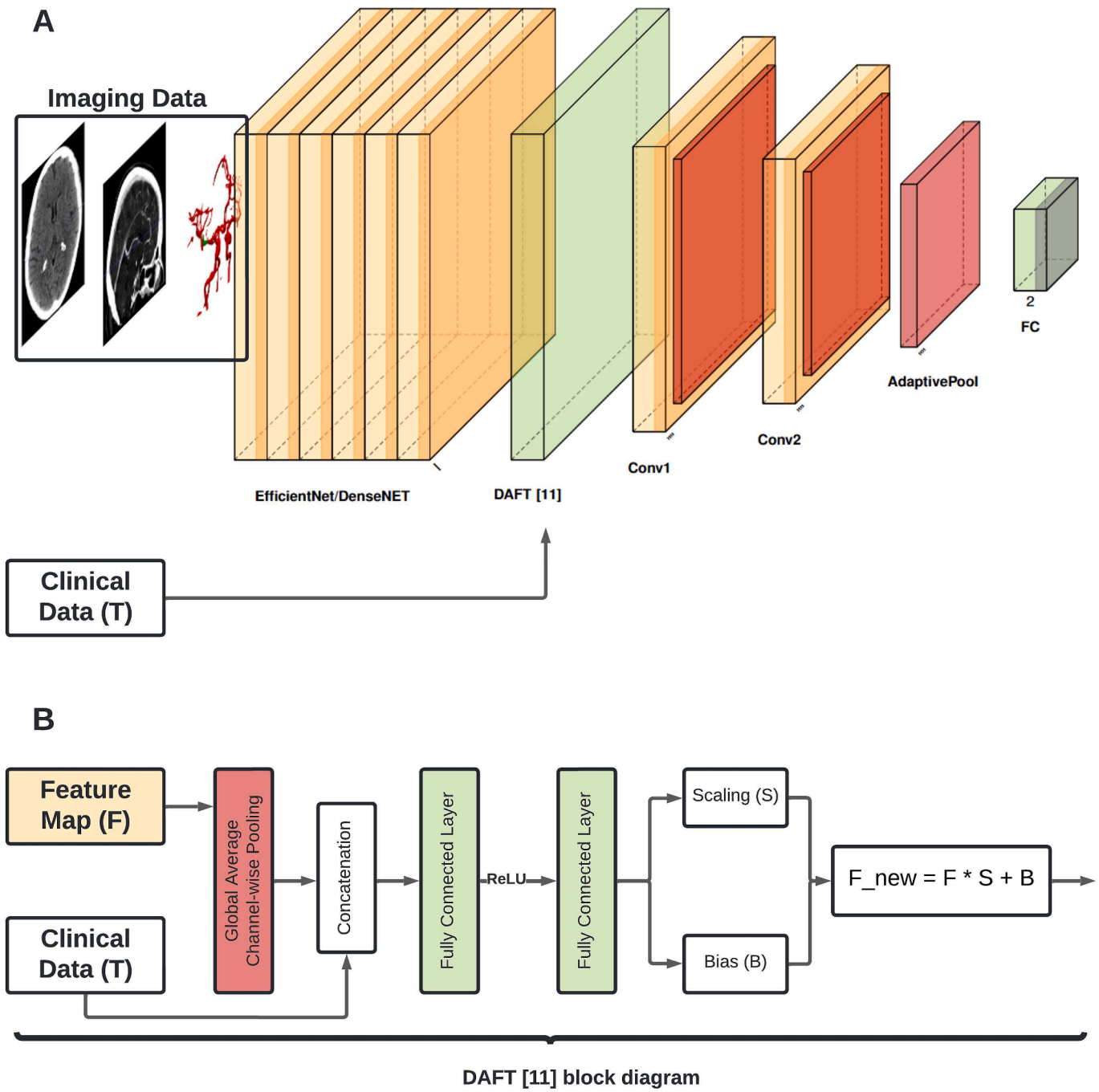

Fig. 2. A: Integration of Imaging Features and Clinical Data in the proposed framework. B: DAFT [12] block.

S represents the scale factor (shape: [B, C]) and B represents the bias (shape: [B, C]). These adjustments aim to refine the imaging-derived features based on patient-specific information provided by clinical data.

To compute S and B, DAFT first concatenates the global average-pooled imaging feature map (P) with the clinical data (T), and passes this combined representation through two fully connected layers:

$$x = ReLU(W_1 \bullet [P;T]), [S;B] = W_2 \bullet x \quad (2)$$

where [P;T] denotes the concatenated pooled feature map (P) and clinical data (T), and W_1, W_2 are the weights of the fully connected layers.

To quantify the indirect contribution of each clinical feature (T_i), we computed the gradients of S and B with respect to the clinical inputs:

$$\partial S/\partial T_i, \partial B/\partial T_i \quad (3)$$

Grad-CAM highlights regions in the imaging feature maps that contribute most to the model's predictions. By comparing Grad-CAM outputs before and after DAFT, we assessed how the inclusion of clinical features changes the model's attention to relevant imaging regions. Together, these analyses—gradient-based attribution for clinical data and spatial explainability from Grad-CAM—provided a comprehensive understanding of how clinical and imaging data were integrated and contributed to the model's predictions on a case-by-case basis.

2.6. Statistical analysis

Model performance was evaluated using the Area Under the Receiver

Operating Characteristic Curve (AUC) and F1 score as primary metrics, providing a balanced assessment of discrimination and predictive power, particularly for imbalanced datasets. Secondary metrics included sensitivity and specificity to assess the model's ability to identify positive and negative cases effectively.

The robustness of the proposed framework was ensured through 4-fold cross-validation during training, with results reported as mean ± standard deviation (SD). Internal testing on a held-out dataset (20 % of the total data) independently validated model performance. To compare the models without the least predictive variables and more recovered data, statistical significance in AUC differences was assessed using the defined primary and secondary metrics along with McNemar's Test to evaluate differences in classification errors.

Grad-CAM outputs were analyzed to compare the model's spatial focus before and after integrating clinical data. In addition to the individual contribution of clinical features, an estimate of the absolute contribution [26] was calculated. Based on these analyses, the least predictive clinical features were removed, which enabled the recovery of part of the data previously excluded due to missing data. Additional experiments on this expanded dataset, including a post-hoc subgroup performance analysis were conducted to assess whether adding more samples could offset the loss of less predictive features, and statistical significance was tested accordingly.

3. Results

Between 2017 and 2023, a total of 839 patients with acute ischemic stroke who underwent EVT for TICA, M1, or M2 occlusions were initially eligible for inclusion. After applying exclusion criteria as detailed in Fig. 1, the final study sample comprised 599 patients. The cohort was divided into a training/validation/test set: 481 patients for training and validation (80.3 % with median age 81, IQR 72–88; 54 % women and 16 % rLVO) and the testing set comprised 118 patients (19.7 % with median age 79, IQR 63–86; 52 % women and 13.6 % rLVO) Table 1 summarizes the baseline characteristics and demographics. Among the rLVO cases, 50.5 % were identified as probable ICAD-LVO, and 49.5 % as cases with other etiologies (cardioembolic or undetermined) in which recanalization was not achieved with conventional EVT techniques. Rescue treatment was performed in 18.2 % of the rLVO cases (of which 12.9 % ended up recanalizing with an eTICI 2b-3 after rescue treatment).

3.1. Baseline models

Multivariate logistic regression was performed on the clinical variables of the study population, to obtain a clinical data baseline (Supplemental Image A1). Obtained variables were known predictors of ICAD-LVO [5]: atrial fibrillation (OR 0.36, 95 % CI 0.19–0.67; p = 0.001), NIHSS (OR 1.04, 95 % CI 1.01–1.07; $p = 0.05$), Tmax > 4 s (OR 1.006, 95 % CI 1.00–1.01; p = 0.01) and DL (OR 0.55, 95 % CI 0.33–0.93; $p = 0.02$).

Table 2 shows the results of the initial experiments with clinical data alone and the different imaging modalities separately: NCCT, CTA or BOTH, as well as the assembling strategy on top of them. Empirically, the assembling approach demonstrated superior performance across all metrics, justifying the use of this approach in the subsequent experiments. Additional experiments with other models can be seen in the supplemental table A1.

3.2. Deep learning models

Results of the ablation study evaluating the impact of adding the automated vessel segmentation, clinical data, and the DAFT block on model performance are summarized in Table 3. Across both validation and internal testing, models that incorporated the automated vessel segmentation and the DAFT block consistently outperformed other configurations.

Table 1
Baseline Characteristics of Training and Test Subsets.

Variable	Training patients (n = 481)	Training rLVO (n = 77)	Test Patients (n = 118)	Test rLVO (n = 16)
Age (years) [median, IQR]	81 (72–88)	78 (67–88)	79 (63–86)	78 (66–85)
Sex (female) [n, %]	261(54.3 %)	41 (53.2 %)	62 (52.5 %)	5 (31.3 %)
Risk Factors [n, %]				
Hypertension	192 (39.9 %)	27 (35.1 %)	35 (29.6 %)	6 (37.5 %)
Dyslipidemia	211 (43.9 %)	27 (35.1 %)	48 (40.7 %)	3 (18.8 %)
Diabetes Mellitus	127 (26.4 %)	15 (19.5 %)	20 (16.9 %)	3 (18.8 %)
Atrial Fibrillation	165 (34.3 %)	13 (16.9 %)	33 (28.0 %)	3 (18.8 %)
Former or current smoker	112 (23.3 %)	18 (23.4 %)	30 (25.4 %)	6 (37.5 %)
Wake-up stroke	176 (36.6 %)	28 (36.3 %)	71 (60.2 %)	7 (43.8 %)
Thrombolysis	179 (37.2 %)	44 (44.1 %)	25 (21.2 %)	4 (25.0 %)
NIHSS baseline (median, IQR)	15 (9–20)	15 (10–21)	17(12–21)	19 (12–34)
ASPECTS (median, IQR)	9 (8–10)	9 (8–10)	7 (6–8)	7 (6–8)
CTP parameters (median, IQR)				
CBF < 30 %	0 (0–15)	0 (0–18)	0 (0–16)	0 (0–21)
Tmax > 10 s	40 (17–82)	44 (20–88)	38 (14–72)	28 (18–80)
Tmax > 6 s	102 (61–155)	115 (71–164)	96 (56–142)	78 (57–145)
Tmax > 4 s	191 (128–261)	204 (149–282)	189 (139–289)	193 (119–276)
HIR	0.43 (0.25–0.57)	0.45 (0.26–0.61)	0.40 (0.26–0.56)	0.46 (0.27–0.55)
Tmax > 4 s / Tmax > 6 s	1.78 (1.49–2.29)	1.84 (1.50–2.38)	1.91 (1.51–2.76)	2.16 (1.55–3.57)

Table 2
Validation Model Performance Across Different Data Modalities: Performance metrics (AUC, F1 score, sensitivity, and specificity) for models trained on baseline clinical and radiological standards (NIHSS, ASPECTS), a set of clinical data, the best result for each imaging mode for brevity: NCCT, CTA, BOTH and the ensemble strategy. The ensembling approach demonstrates superior performance across all metrics, particularly in sensitivity and specificity.

Data Modality	Validation			
	AUC	F1	Sensitivity	Specificity
NIHSS	0.56 ± 0.05	0.31 ± 0.06	0.46 ± 0.24	0.71 ± 0.27
ASPECTS	0.52 ± 0.03	0.29 ± 0.01	0.82 ± 0.08	0.30 ± 0.11
NIHSS + ASPECTS	0.57 ± 0.05	0.31 ± 0.04	0.61 ± 0.21	0.54 ± 0.26
All Clinical	0.58 ± 0.02	0.35 ± 0.01	0.54 ± 0.23	0.72 ± 0.22
NCCT	0.47 ± 0.05	0.13 ± 0.11	0.46 ± 0.41	0.49 ± 0.33
CTA	0.50 ± 0.02	0.22 ± 0.01	0.55 ± 0.11	0.47 ± 0.09
BOTH	0.53 ± 0.02	0.19 ± 0.05	0.31 ± 0.12	0.75 ± 0.12
Ensemble	0.60 ± 0.02	0.29 ± 0.02	0.72 ± 0.12	0.49 ± 0.16

Regarding the saliency maps, an intensity increase in vessel regions and surrounding areas, and an intensity decrease in zones where occlusions are unlikely can be appreciated after the fusion of imaging and clinical data (Fig. 3). Features such as atrial fibrillation, occlusion side, dyslipidemia status, age and sex were among the top contributors, whereas others, such as CTP derived features and IVT status, exhibited relatively lower to moderate influence and a rather minimal impact not being as predictive as expected. These results are seen in the absolute contribution radar plot of each of the variables (Fig. 4).

Table 3
Results of the ablation study for the explored CNN backbones, adding vascular segmentation (prior), the DAFT module or both. Best-performing models for each backbone are highlighted in bold.

Model Name	Validation				Test			
	AUC	F1	Sens.[c]	Spec.[c]	AUC	F1	Sens.[c]	Spec.[c]
EffB0[a]	0.60 ± 0.06	0.31 ± 0.07	0.54 ± 0.28	0.66 ± 0.18	0.60 ± 0.02	0.29 ± 0.02	0.72 ± 0.12	0.49 ± 0.16
EffB0 + Prior	0.56 ± 0.01	0.28 ± 0.02	0.52 ± 0.22	0.60 ± 0.22	0.56 ± 0.02	0.27 ± 0.02	0.69 ± 0.23	0.44 ± 0.30
EffB0 + DAFT	0.59 ± 0.05	0.32 ± 0.04	0.78 ± 0.15	0.39 ± 0.21	0.62 ± 0.02	0.31 ± 0.03	0.60 ± 0.22	0.64 ± 0.22
EffB0 + Prior + DAFT	**0.65 ± 0.04**	**0.36 ± 0.03**	**0.84 ± 0.06**	**0.46 ± 0.09**	**0.64 ± 0.02**	**0.34 ± 0.05**	**0.72 ± 0.25**	**0.55 ± 0.33**
D169[b]	0.57 ± 0.05	0.31 ± 0.03	0.75 ± 0.36	0.38 ± 0.29	0.56 ± 0.04	0.25 ± 0.05	0.43 ± 0.34	0.69 ± 0.39
D169 + Prior	0.56 ± 0.06	0.28 ± 0.08	0.71 ± 0.36	0.42 ± 0.36	0.57 ± 0.04	0.26 ± 0.03	0.66 ± 0.34	0.46 ± 0.37
D169 + DAFT	0.63 ± 0.06	0.36 ± 0.06	0.68 ± 0.12	0.57 ± 0.23	0.63 ± 0.02	0.30 ± 0.01	0.86 ± 0.08	0.40 ± 0.04
D169 + Prior + DAFT	**0.64 ± 0.05**	**0.37 ± 0.05**	**0.74 ± 0.10**	**0.54 ± 0.15**	**0.63 ± 0.02**	**0.31 ± 0.02**	**0.74 ± 0.16**	**0.53 ± 0.18**

[a] EfficientNet B0 Model.
[b] DenseNet 169 Model.
[c] Spec.: Specificity. Sens.:Sensitivity.

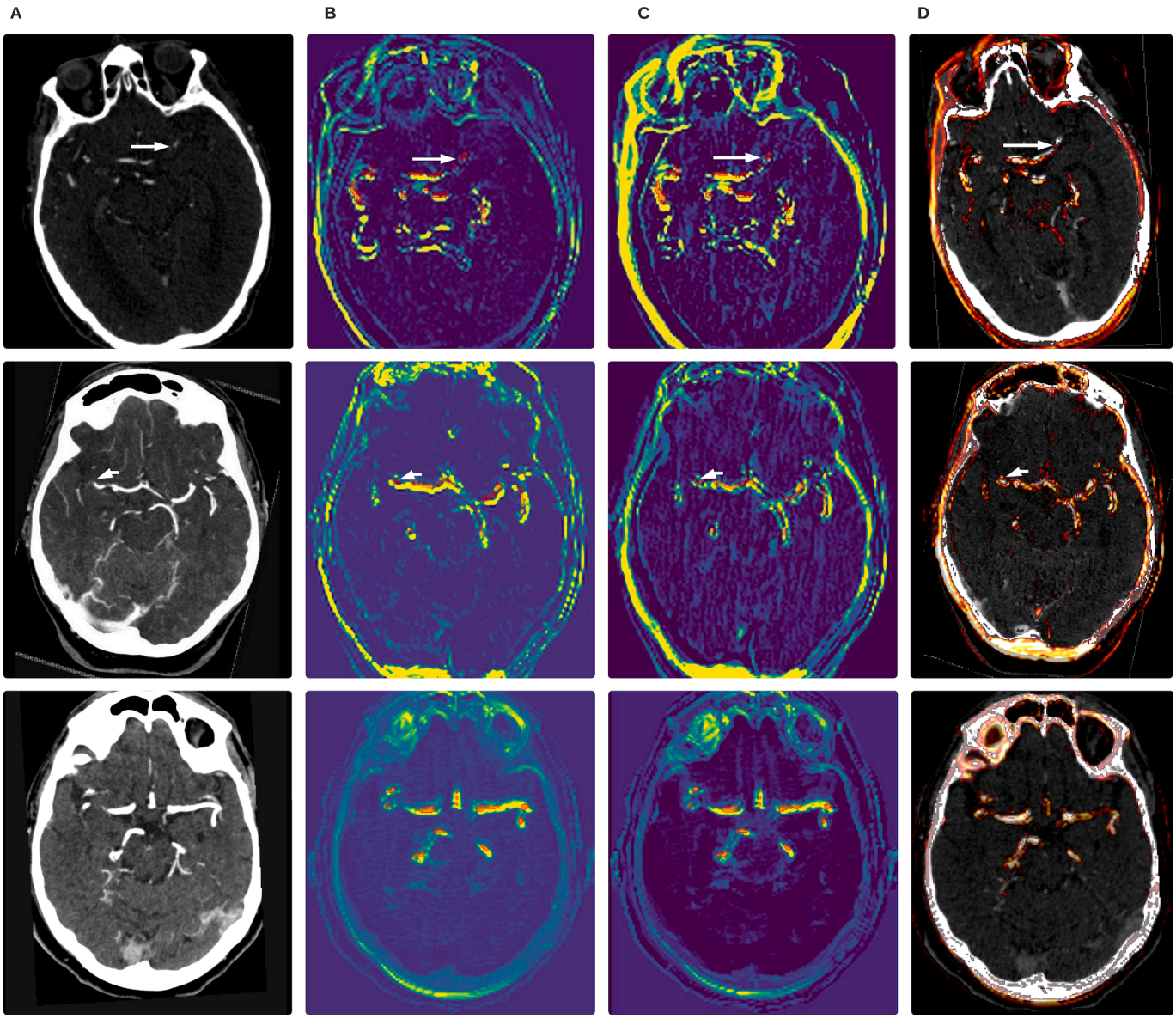

Fig. 3. Saliency Map Analysis Highlighting the Effects of the DAFT Block on Feature Maps. Column A: CTA images pointing to a diagnosed occlusion. Column B: GradCAM before DAFT, with intensity visualized using the viridis color code. Column C: GradCAM after DAFT, with intensity also visualized using the viridis color code. Column D: Overlay of saliency maps on CTA image highlighting the focus on vessel walls. In Columns B and C, red indicates the vessel segmentation. Files 1 and 2 are true positives while file 3 is a false positive case.

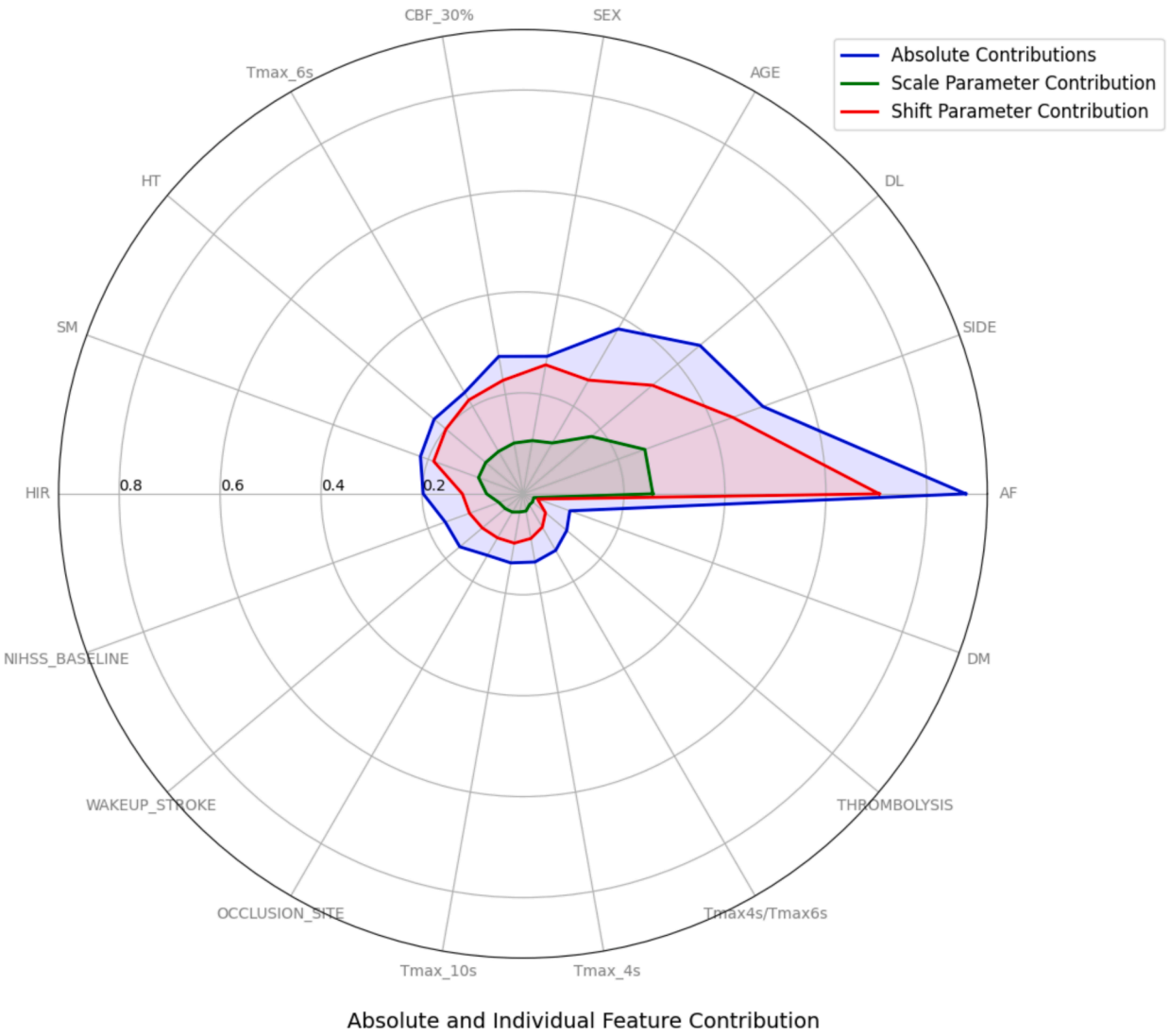

Fig. 4. Absolute and Individual Feature Contribution Estimation ($F_{contribution}$ $|S|*\overline{F}+|T|$) Radar graph displaying the estimated feature contributions, calculated from the mean of all gradient values across the test samples.

The removal of CTP predictive features enabled the inclusion of approximately 13 % (N = 61) additional cases in the training set. The EfficientNet model trained with this expanded dataset and reduced feature set, maintained performance comparable to the original configuration.

The majority voting per-fold strategy improved the best model's AUC from 0.64 (±0.02) to 0.70 (±0.02), and its F1 score increased from 0.34 (±0.05) to 0.39 (±0.02). These gains moved the model's operating point toward higher specificity at the expense of sensitivity. Applying the same assembly method to the model with fewer predictive variables (Model B) and the one with the expanded dataset (Model C) also improved overall performance, in terms of AUC, but reduced recall.

Although the difference in removing clinical data and adding more samples might appear moderate as shown in Table 4, McNemar's Test results revealed differences in error patterns. For example, when comparing the original model (Model A) to Model B, we found more errors in situations where Model A had been correct (McNemar's Statistic: 9.53; p = 0.002). Similarly, Model C appeared comparable to Model B in aggregate metrics, but McNemar's Test (McNemar's Statistic: 11.57; p < 0.001) indicated a meaningful difference. While not completely, model C reduced the performance gap in model B compared to model A highlighting the trade-off of having less predictors but more data. Complete error breakdown can be seen in table A2.

Detailed subgroup analysis revealed that Models A and B had similar detection rates for ICAD and other non-recanalizing causes (Model A identified 63.5 % of ICAD and 63.5 % of other non-recanalizing causes,

Table 4
Impact of reducing clinical variables (model B) and expanding dataset (model C) on model performance compared to baseline model (model A).

Model Name	Validation				Test			
	AUC	F1	Sens.	Spec.	AUC	F1	Sens.	Spec.
Best Study Model (Model A)	0.65 ± 0.04	0.36 ± 0.03	0.84 ± 0.06	0.46 ± 0.09	0.64 ± 0.02	0.34 ± 0.05	0.72 ± 0.25	0.55 ± 0.33
Best Study Model w/o LPV[a] (Model B)	0.62 ± 0.06	0.34 ± 0.05	0.52 ± 0.22	0.60 ± 0.22	0.62 ± 0.03	0.31 ± 0.03	0.69 ± 0.23	0.44 ± 0.30
Ablated Model with more data (Model C)	0.61 ± 0.02	0.33 ± 0.02	0.78 ± 0.15	0.39 ± 0.21	0.64 ± 0.02	0.33 ± 0.03	0.60 ± 0.22	0.64 ± 0.22

[a] Less predictive variables.

and 50 % each in Model B), while Model C identified more non-recanalizing causes (87 %) than ICAD cases (50 %). Additional post hoc analysis showed no statistically significant differences in model performance by sex or age group (see supplemental figure A2), supporting the robustness of predictions across demographic subgroups despite the inherent limitations of a single-center dataset.

4. Discussion

This work demonstrates the feasibility of using multi-modal deep learning to predict recanalization failure during EVT in real-world clinical conditions, using solely data available before EVT. We observed a progressive improvement in predictive performance as we moved from standard baseline models using only clinical or radiological predictors to our final integrated multi-modal approach, showing the importance of data fusion along with interpretability, and practical considerations of data availability.

While the model's positive predictions do not exclusively signify the presence of ICAD-LVO, such predictions can still guide clinical decision-making. A positive prediction of the model indicates that interventionalists should anticipate a more complex course of treatment, potentially requiring multiple passes or rescue therapies. Indeed, emergent angioplasty or stenting has shown promise in achieving better outcomes for ICAD-LVO, suggesting a possible advantage when standard approaches appear insufficient [29]. Thus, even though the model does not yet differentiate purely by stroke etiology, it provides actionable information that can be used to improve workflow, resource allocation, and patient counseling.

This study considered the complexity and heterogeneity of real-world clinical scenarios. By including a broader spectrum of patients with unsuccessful recanalization cases grouped as rLVO, we created a more generalizable model that better reflects the actual patient population encountered in daily practice. The inclusion of demographic, clinical, radiological, and CTP-derived parameters via the DAFT module showed improved prediction metrics along with selected backbone architectures. Grad-CAM analysis suggests that patient-specific clinical variables enhance the model's ability to understand underlying vascular pathology compared to models based solely on image data, guiding it to focus on relevant imaging regions.

Although CTA primarily depicts the intraluminal contrast, the model emphasized in boundary surrounding areas including the volumes comprising the vessel wall (Fig. 4, column D) which probably includes relevant features in a typical place for plaque accumulation [4]. Indeed, this observation aligns with ongoing research using high-resolution vessel wall MRI (HR-VW-MRI) where vessel wall features are critical predictors of ICAD-LVO mechanisms, even before luminal stenosis becomes apparent [4,27]. Furthermore, vessel wall analysis has been increasingly recognized as a key area of investigation for understanding ICAD-LVO and its recurrence risks, as it provides insights into plaque vulnerability and vascular remodeling [4]. In contrast, false positives may occur, as shown in Fig. 4 (file 3), because the model broadly focuses on the entire occluded lumen responding primarily to the generalized lack of contrast filling rather than specific vessel wall features indicative of refractoriness [30]. Such diffuse luminal attention is typical in embolic occlusions, characterized by an abrupt contrast cut-off at arterial bifurcations (branch-site occlusion sign), which typically suggests easier recanalization and lacks the characteristic tapering or wall abnormalities associated with refractory lesions [30].

Time is critical in stroke management, and certain clinical features, particularly CTP-derived variables, may not always be readily available. Our experiments removing less predictive (or not always available) variables illustrate a trade-off: while performance declined marginally, we could include more cases that would otherwise be excluded. Thus, the inclusion of these new samples kept a similar performance, but as seen in the results, balanced more towards identifying more non-recanalizing cases than ICAD cases (model C, Table 3) while model A was more robust in both cases.

In general, predicting ICAD-LVO or other challenging situations preoperatively is a complex task. The subtle imaging and clinical factors that signal underlying stenosis or complex plaque morphology are difficult to isolate, especially before a DSA diagnosis. Our framework's stable performance under various dataset conditions and its ability to integrate multiple data types represents a step forward to AI assistance. Even incremental improvements in early identification can empower interventionalists to consider bailout treatments, or planning additional device passes, potentially improving clinical outcomes.

This study has several limitations: It is based on retrospective, single-center data, which may limit the generalizability of our findings. There is a persistent class imbalance even in the defined rLVO. Certain clinical variables may not be readily available in all emergency settings, potentially restricting the model's applicability. The inclusion of a broad range of etiologies improves realism, and also introduces heterogeneity that may complicate interpretation.

While our model provides early prediction of procedural complexity, we caution against using its output to directly determine rescue treatment strategies such as stenting prior to a first EVT pass. Current evidence like the RESCUE-ICAS registry [31], supports that combining EVT with emergent stenting may improve outcomes compared to thrombectomy alone in select patients with ICAD-LVO. At present, the standard approach is to attempt at least one EVT pass before considering rescue interventions, and this model is intended to support, not replace, case-by-case multidisciplinary clinical judgment.

Instead, the primary clinical utility of our tool lies in alerting stroke teams to anticipated technical challenges, enabling more proactive resource allocation, team preparation, and patient counseling. Additionally, such a predictive algorithm can facilitate more targeted patient selection for future clinical trials evaluating early or upfront rescue strategies such as the ongoing ICARUS trial [32] which aim to clarify the optimal timing of stenting in ICAD-LVO. Once validated in prospective, multicenter cohorts and potentially incorporated into trial design, the algorithm could finally serve as a standardized, objective tool to stratify procedural complexity and guide tailored intervention strategies as seen in Fig. 5.

5. Conclusion

A multi-modal model for rLVO prediction in patients receiving EVT, trained on baseline neuroimaging and clinical variables, was introduced in this study. Our results show that the combination of both data modalities improved imaging-only models, highlighting the advantage of models being able to integrate all available data. The best model achieved an AUC of 0.70, F1 of 0.39, sensitivity 0.69 and specificity 0.70 on a hold-out test set for rLVO prediction. Our approach provides a promising foundation for early identification of complex recanalization scenarios, enhancing decision-making and potentially improving patient selection and outcomes in acute stroke care.

CRediT authorship contribution statement

Jesús D. González: Writing – original draft, Visualization, Validation, Software, Project administration, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Pere Canals:** Writing – review & editing, Visualization, Validation, Supervision, Software, Project administration, Data curation, Conceptualization. **Marc Rodrigo-Gisbert:** Writing – review & editing, Validation, Resources, Investigation, Data curation. **Jordi Mayol:** Writing – review & editing, Investigation, Data curation. **Alvaro García-Tornel:** Writing – review & editing, Validation, Resources, Investigation, Data curation. **Marc Ribó:** Writing – review & editing, Validation, Supervision, Resources, Project administration, Conceptualization.

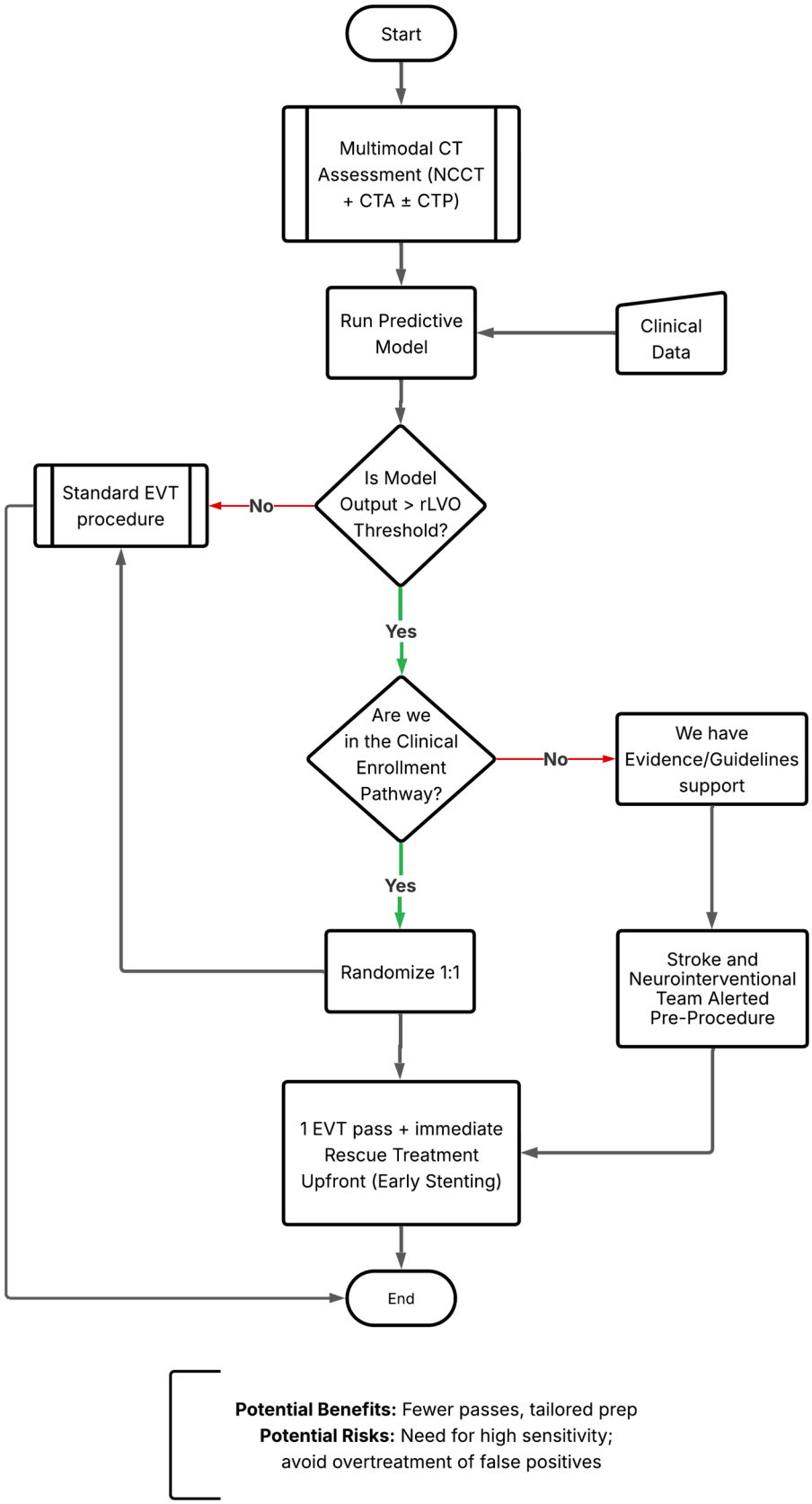

Fig. 5. Advantages and Challenges of an rLVO algorithm and its potential uses as patient selection tool or as a treatment planner.

Declaration of competing interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests: Marc Ribo reports a relationship with Anaconda Biomed S.L. that includes: board membership. Marc Ribo reports a relationship with Nora that includes: board membership. Marc Ribo reports a relationship with Medtronic Inc that includes: consulting or advisory. Marc Ribo reports a relationship with Cerenovus that includes: consulting or advisory. Marc Ribo reports a relationship with Methinks that includes: consulting or advisory. Marc Ribo reports a relationship with Vesalio that includes: consulting or advisory. Marc Ribo reports a relationship with Stryker that includes: consulting or advisory. Marc Ribo reports a relationship with Philips that includes: consulting or advisory. Marc Ribo reports a relationship with Rapid Pulse that includes: consulting or advisory. Marc Ribo reports a relationship with Sensome that includes: consulting or advisory. Marc Ribo reports a relationship with Apta Targets that includes: consulting or advisory. If there are other authors, they declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Appendix A. Supplementary material

Supplementary data to this article can be found online at https://doi.org/10.1016/j.ejrad.2025.112254.

References

[1] M. Goyal, B. Menon, W.H. van Zwam, D.W.J. Dippel, P.J. Mitchell, A.M. Demchuk, Endovascular thrombectomy after large-vessel ischaemic stroke: a meta-analysis of individual patient data from five randomised trials, Lancet 387 (2016) 1723–1731.

[2] D. Xie, J. Huang, S. Fan, C. Guo, W. Sun, Z. Peng, et al., Endovascular Therapy and Outcomes among patients with very Large Ischemic Core Stroke, JAMA Netw. Open 7 (5) (2024) e249298.

[3] C. Lin, J. Saver, B. Obviagele, W. Huang, M. Lee, Endovascular thrombectomy without versus with intravenous thrombolysis in acute ischemic stroke: a non-inferiority meta-analysis of randomized clinical trials, J Neurointervent Surg. 14 (2022) 227–232.

[4] A. de Havenon, O. Zaidat, S. Amin-Hanjani, T. Nguyen, A. Bangdad, M. Abassi, et al., Large vessel occlusion stroke due to intracranial atherosclerotic disease: identification, medical and interventional treatment, and outcomes, Stroke 54 (2023) 6.

[5] M. Rodrigo-Gisbert, A. Garcia-Tornel, M. Requena, I. Vielba-Gómez, S. Bashir, M. Rubiera, et al., Clinico-radiological features of intracranial atherosclerosis-related large vessel occlusion prior to endovascular treatment, Sci. Rep. 14 (2024) 2945.

[6] C. Beaman, S. Yaghi, D. Liebeskind, A decade on: the evolving renaissance in intracranial atherosclerotic disease, Stroke Vasc Interv Neurol. 2 (2022) e000497.

[7] S. Al Kasab, E. Almallouhi, M. Jumaa, V. Inoa, F. Capasso, M. Nahhas, et al., Outcomes of adjunct emergent stenting versus mechanical thrombectomy alone: the RESCUE-ICAS registry, Stroke (2024).

[8] J.G. Kim, S.Y. Ha, Y. Kang, H. Hong, D. Kim, M. Lee, et al., Automated detection of large vessel occlusion using deep learning: a pivotal multicenter study and reader performance study, J Neurointervent Surg. (2024).

[9] J.C. Martinez-Gutierrez, Y. Kim, S. Salazar-Marioni, M.B. Tariq, R. Abdelkhaleg, A. Niktabe, et al., Automated large vessel occlusion detection software and thrombectomy treatment times, JAMA Neurol. 80 (11) (2023) 1182–1190.

[10] N.A. Shlobin, A.A. Baig, M. Waqas, T. Patel, R. Dossani, M. Wilson, et al., Artificial intelligence for large-vessel occlusion stroke: a systematic review, World Neurosurg. 159 (2022) 207–220.

[11] W. Chen, J. Liu, L. Yang, H. Sun, S. Yang, M. Wang, et al., Development and internal-external validation of the ATHE Scale: predicting acute large vessel occlusion due to underlying intracranial atherosclerosis prior to endovascular treatment, J. Neurosurg. 141 (1) (2024) 165–174.

[12] T.N. Wolf, S. Pölsterl, C. Wachinger, DAFT: a universal module to interweave tabular data and 3D images in CNNs, Neuroimage 260 (15) (2022) 119505.

[13] C. Delgrange, O. Demler, S. Mora, B. Menze, E. de la Rosa, N. Davoudi. A Self-Supervised Model for Multi-modal Stroke Risk Prediction. Available from: Doi: 10.48550/arXiv.2411.09822.

[14] A. White, M. Saranti, A.D. Garcez, T.M.H. Hope, C.J. Price, H. Bowman, Predicting recovery following stroke: deep learning, multimodal data and feature selection using explainable AI, NeuroImage Clin. 43 (2024) 103638.

[15] Y. Liu, Y. Yu, J. Ouyang, B. Jiang, G. Yang, S. Ostmeier, et al., Functional outcome prediction in acute ischemic stroke using a fused imaging and clinical deep learning model, Stroke 54 (9) (2023).

[16] P. Canals, S. Balocco, O. Díaz, J. Li, A. García-Tornel, A. Tomasello, et al., A fully automatic method for vascular tortuosity feature extraction in the supra-aortic region: unraveling possibilities in stroke treatment planning, Comput. Med. Imaging Graph. 104 (2023) 102170.

[17] D. Lyndon, M. van de Broek, B. Niu, S. Yip, A. Rohr, F. Settecase, Hypoperfusion intensity ratio correlates with cta collateral status in large-vessel occlusion acute ischemic stroke, AJNR Am. J. Neuroradiol. 42 (2021) 1380–1386.

[18] J. M. Cardoso, W. Li, R. Brown, N. Ma, E. Kerfoot, Y. Wang et al. Monai: An open-source framework for deep learning in healthcare. Available from: Doi: 10.48550/arXiv.2211.02701.

[19] G. Huang, Z. Liu, L. van der Maaten, K.Q. Weinberger, Densely connected convolutional networks, Conf. Comput. Vis Pattern Recognit. (2017) 2261–2269.

[20] M. Tan, Q. V. Le. EfficientNet: rethinking model scaling for convolutional neural networks. Available from: Doi: 10.48550/arXiv.1905.11946.

[21] J.K. Kim, M.C. Chang, D. Park, Deep learning algorithm trained on brain magnetic resonance images and clinical data to predict motor outcomes of patients with corona radiata infarct, Front. Neurosci. 15 (2022) 795553.

[22] T. Zhou, X. Ye, H. Lu, X. Zheng, S. Qiu, Y. Liu, Dense convolutional network and its application in medical image analysis, BioMed. Res. Intl. (2022) 2384830.

[23] K. Lee, C. Liu, D.Y. Chen, C. Weng, H. Chiu, C. Chiang, Automatic detection and vascular territory classification of hyperacute staged ischemic stroke on diffusion weighted image using convolutional neural networks, Sci. Rep. 13 (2023) 404.

[24] H. Jo, C. Kim, D. Gwon, J. Lee, J. Lee, K.M. Park, et al., Combining clinical and imaging data for predicting functional outcomes after acute ischemic stroke: an automated machine learning approach, Sci. Rep. 13 (2023) 16926.

[25] E. de la Rosa, R. Su, M. Reyes. R. Wiest, E. O. Riedel, F. Kofler et al. ISLES'24: Improving final infarct prediction in ischemic stroke using multimodal imaging and clinical data. Available from: Doi: 10.48550/arXiv.2408.10966.

[26] M. T. Ribeiro, S. Singh, C. Guestrin. "Why Should I Trust You?" Explaining the Predictions of Any Classifier. Available from: Doi: 10.48550/arXiv.1602.04938.

[27] J.W. Song, A. Pavlou, J. Xiao, S. Kasner, Z. Fan, S.R. Messé, Vessel wall magnetic resonance imaging biomarkers of symptomatic intracranial atherosclerosis: a meta-analysis, Stroke 52 (2020) 1.

[28] F. Isensee, P.F. Jaeger, S.A.A. Kohl, K.H. Maier-Hein, nnU-net: a self-configuring method for deep learning-based biomedical image segmentation, Nature 18 (2020) 203–211.

[29] W. Li, X. Sui, C. Li, W. Zhao, S. Yuan, S. Dou, et al., Emergency Angioplasty or Stenting for Stroke patients with Intracranial Atherosclerotic Large Vessel Occlusion, J. Atheroscler. Thromb. 30 (2023) 160–169.

[30] J. Li, J. Yang, X. Gao, Q. Han, Y. Wu, Q. Shang, et al., Clot patterns determined by DSA and CTA can help predict intracranial atherosclerotic stenosis in acute ischemic stroke patients, Front. Neurol. 15 (2024) 1395764.

[31] S. Al Kasab, E. Almallouhi, M. Jumaa, V. Inoa, F. Capasso, M. Nahhas, et al., Outcomes of adjunct emergent stenting versus mechanical thrombectomy alone: the RESCUE-ICAS registry, Stroke 56 (2024) 2.

[32] National Library of Medicine (U.S.) (2024, June-) *IntraCranial Atherosclerosis Related Large-vessel Occlusion Treated With Urgent Stenting (ICARUS)*. Identifier NCT06472336. https://clinicaltrials.gov/study/NCT06472336.