EDAF: Early Detection of Atrial Fibrillation from Post-Stroke Brain MRI

Mohammad Javad Shokri[1], Nandakishor Desai[1], Aravinda S. Rao[1], Angelos Sharobeam[2], Bernard Yan[2], and Marimuthu Palaniswami[1]

[1] Electrical and Electronic Engineering, The University of Melbourne, Parkville, Victoria - 3010, Australia
mshokri@student.unimelb.edu.au; nnandakisho@unimelb.edu.au, aravinda.rao@unimelb.edu.au, palani@unimelb.edu.au

[2] Melbourne Brain Center, Royal Melbourne Hospital, Parkville, Victoria - 3052, Australia
angelos.sharobeam@mh.org.au, bernard.yan@mh.org.au

Abstract. Atrial fibrillation (AF) is a common cause of ischemic stroke, accounting for up to one-third of all cases. Untreated AF can increase the risk of stroke by up to five times and make stroke recurrence more likely. Anticoagulation has proven beneficial in reducing stroke risk. However, AF is often paroxysmal and asymptomatic, remaining undetected and undiagnosed in up to 30% of cases. The current methods for AF detection are usually lengthy (cardiac monitoring), expensive (smart devices), or invasive (implantable cardiac monitors), limiting their routine use. We present a novel method to screen for AF by analyzing infarct patterns of stroke patients from brain magnetic resonance imaging (MRI) scans. We propose *EDAF*, a novel method based on the segment anything model (SAM) that leverages the power of a foundational deep learning model to efficiently analyze brain MRI and identify whether the underlying stroke etiology is AF. EDAF is trained and validated using a retrospectively acquired dataset of 235 post-stroke patients, achieving an area under the receiver operating characteristic (AUROC) of $83.08\% \pm 2.96\%$ in identifying the presence of AF. EDAF can achieve optimal solutions with minimal training, highlighting its potential for use in low-resource settings. As MRI is readily available in stroke centers and routinely performed on many patients after a stroke, either during their admission or as an outpatient, the proposed method can effectively identify patients for further AF investigation.

Keywords: Ischemic stroke · Atrial fibrillation (AF) · Magnetic resonance imaging (MRI) · Vision transformer

1 Introduction

Atrial fibrillation (AF), the most common type of heart arrhythmia in the world, characterized by abnormal electrical activity, significantly increases stroke risk and can lead to multiple, including silent, strokes [7]. AF increases stroke risk five-fold compared to the general population and is associated with over two-fold increased risk of recurrent stroke [19,21]. Detecting AF after a first stroke is vital for optimizing treatment with oral anticoagulation, preventing recurrent strokes, uncovering silent AF, and improving long-term outcomes. However, the variability in AF manifestation between patients and even within the same patient over time, along with AF's paroxysmal and asymptomatic nature, makes detection challenging.

Current AF detection techniques are through continuous or intermittent cardiac monitoring. They generally operate by recording electrocardiogram (ECG) or other physiological parameters associated with heart function. These techniques include hospital ECG recording devices, event recorders, Holter monitors, implantable cardiac devices (ICDs), and smartphones. The AF identification accuracy in these devices depends on monitoring length and patient compliance, and they are costly and invasive in some cases [11]. To identify the underlying AF in post-stroke patients, based on the current guidelines, acute ischemic stroke (AIS) patients often undergo a short-term ECG recording in the emergency room.

Prolonged cardiac monitoring is conducted only when evidence indicates a higher likelihood of a cardioembolic stroke. However, the absence of a comprehensive risk stratification framework results in many AF patients only receiving short-term monitoring, which leads to numerous undetected AF cases. This oversight can delay AF treatment or result in incorrect treatment strategies for stroke, causing irreversible consequences such as secondary strokes, further heart complications, and unnecessary patient expenses. To address this issue, a robust risk stratification pipeline must be implemented to identify individuals at elevated risk utilizing the existing modalities in stroke diagnosis and treatment procedures.

Magnetic resonance imaging (MRI), particularly diffusion-weighted imaging (DWI), is commonly used soon after a stroke. Despite its limitations, such as long scanning duration and limited availability, it is essential in the acute phase of stroke to provide detailed diagnostic information, facilitate treatment decisions, and predict outcomes. Ischemic strokes result in distinct infarct patterns visible on DWI images [23][15], which are associated with stroke severity and functional outcomes. We hypothesize that the DWI scans of post-stroke patients can be efficiently analyzed using deep learning models to identify the underlying presence of AF automatically. There are only two published works focusing on analyzing brain MRI scans using deep learning for investigating stroke etiology. In a published abstract [30], a brain MRI dataset of 489 patients was used to identify the underlying atrial fibrillation, achieving an AUC-ROC of 80%. However, the model's reliance on detailed annotation for segmentation and the absence of comprehensive experiments and discussion makes it difficult to assess its specific contributions. Another study [12] utilized two models for stroke subtype classification and infarct segmentation, achieving an accuracy of 81.9%. However, the method is complex and requires costly and time-consuming manual annotation of infarct regions.

Through convolutional neural networks (CNN) and transformers [26,6], deep learning has become the gold standard for challenging computer vision and medical image analysis tasks. Due to the complexity of medical images and often access to small-scale datasets, transfer learning is a common approach for attaining high performance in medical image analysis. The segment anything model (SAM) [13], a large deep learning model, has recently shown exceptional performance in image segmentation. Like most transformer models, SAM's architecture is an auto-encoder in which the encoder extracts features from the input image, and the decoder generates segmentation masks based on these feature maps. After training, the encoder can function separately as the backbone of a classification network. While SAM is primarily designed for segmentation, due to the scale of the dataset used (a dataset containing 11 million images) and the fine-tuned SAM variants already available in medical image analysis [9,10,17,18,22,27,28,29], the rich representations from this pre-trained model can be effectively applied to various downstream tasks, including classification. SAM is highly effective due to its versatile architecture, zero-shot capability, flexible prompting, and extensive training on over 1 billion segmentation masks. This foundation model excels at generalization and has broad applicability across various domains.

We propose an AF detection methodology by developing a novel 3D image classification framework that leverages the power of rich pre-trained model representations of SAM. This framework integrates it with a custom embedding layer to classify the stroke etiology from the 3D DWI scans. Our key contributions include:

- Early AF risk detection without additional procedures and using existing clinical pathways can potentially improve patient outcomes. Our proposed novel EDAF framework stratifies AF risks by utilizing only the DWI-MRI brain images acquired during the acute phase of stroke.
- Developing a robust classification model requires significant model optimization. Our novel approach using SAM [13] achieves high performance with minimal training, making it particularly effective in low-resource clinical settings.
- Unlike existing methods that rely on segmentation masks, we adapt SAM to directly classify DWI brain scans, eliminating labor-intensive intermediate steps.
- Although 3D models provide more detailed information, but they consume higher memory and require longer training times, prohibiting in low-resource clinical settings. Our approach utilizes 3D to 2D projections, maintaining high performance while minimizing computational complexity.
- Model robustness is crucial in clinical applications. Our proposed EDAF framework achieves enhanced generalizability compared to current models.

2 Methodology

2.1 AF and Brain MRI

Short-term cardiac monitoring [24] often misses AF episodes due to limited duration, asymptomatic nature, and infrequent episodes. Longer-term tracking, such as 72-hour ECG or continuous ECG monitoring, significantly increases AF detection rates but faces challenges in patient compliance and resource limitations. On the other hand, an MRI is readily available in stroke centers and routinely performed on many patients after a stroke, either during their admission or as an outpatient. Therefore, an automated MRI-based risk stratification framework can play a critical role in the early identification of AF patients who are more likely to benefit from prolonged cardiac monitoring. In this innovative approach, a deep learning model analyzes the images after a patient undergoes a brain MRI to estimate the likelihood of AF based on infarct patterns. This estimation can be used as a risk stratification tool to identify high-risk patients who would benefit from more extensive cardiac monitoring to diagnose AF definitively.

Figure 1 displays sample brain DW-MRI images, highlighting various infarct patterns, sizes, and locations in AF and large artery atherosclerosis (LAA) categories. These variations present challenges in visually distinguishing the underlying etiology based solely on the images. However, as mentioned previously, clinical studies suggest that these infarct patterns broadly correlate with the cause of stroke. Due to their ability to learn complex patterns and features, we speculate that deep learning models can benefit from these subtle differences and effectively leverage this information to differentiate between these mechanisms, even though it is not visually apparent.

2.2 Dataset

This study was approved by the Royal Melbourne Hospital Ethics Committee (QA 2013.072). The dataset consists of 235 brain DW-MRI images acquired retrospectively from AIS patients. The images were acquired at admission during the acute phase of the stroke. The imaging was performed using various scanners from Siemens Aera, Siemens Prisma Fit, Siemens Skyra, Siemens Magnetom Essenza, and Philips Ingenia. The imaging parameters and protocols are illustrated in Table 1, and sample images of the dataset are shown in Fig. 1.

Two expert neurologists employed the Causative Classification System (CCS) to determine whether the underlying stroke etiology is AF [4]. This system provides a detailed classification to aid in targeted diagnosis and treatment by identifying the underlying causes of strokes. It is based on a comprehensive assessment of clinical features, imaging findings, and other diagnostic criteria to categorize strokes into specific etiological groups. Based on information from various sources such as vascular imaging, cardiac evaluations, and clinical histories, this system categorizes strokes into five main groups: large artery atherosclerosis (LAA), cardioembolism (CE), small artery occlusion (SAO), stroke of other determined cause (OC), and stroke of undetermined cause (UND). We opt for

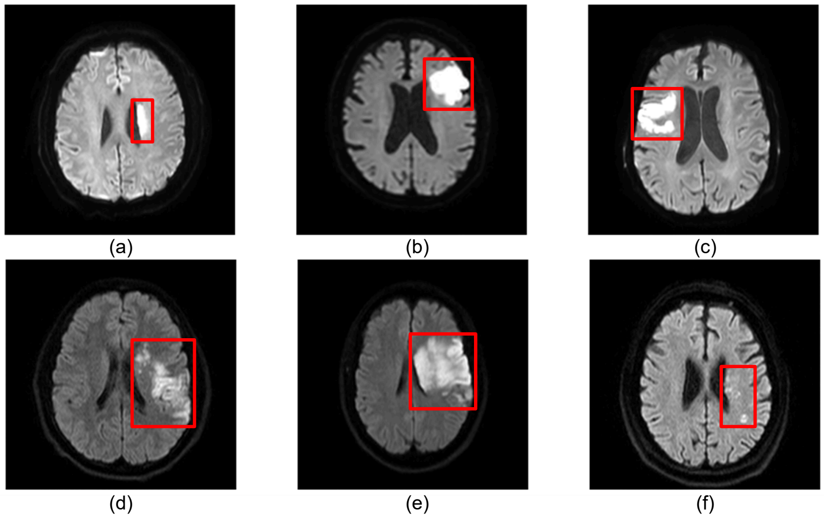

Fig. 1. Samples of the brain DW-MRI images from the dataset. The figure illustrates an axial slice from the MRI scans for three AF patients (a-c) and three LAA patients (d-f). Different infarct patterns, sizes, and locations can be seen within each category, leading to challenges in visually differentiating the underlying etiology solely based on the images. Due to the robust feature extraction capabilities of the deep learning models, the proposed method aims to deploy these tools to identify the underlying AF or LAA.

Table 1. Overview of dataset demographics and imaging parameters, including patient characteristics, imaging protocols, and parameters

Number of patients	235
AF-related strokes	138
Female	83
Age (mean±std)	71.1 ± 14.2
Magnetic field strength	1.5 or 3 Tesla
Repetition time	4100-7920 ms
Echo time	55-104 ms
Flip angle	0-180 degrees
b-values	0 or 1,000 sec/mm^2
Slice thickness	0.256-7.474 mm
Slice spacing	2-7.5 mm
Pixel spacing	0.548-2 mm

the CCS system [4] over the Trial of Org 10172 in Acute Stroke Treatment (TOAST) [1] and the ASCO (atherosclerosis, small-vessel disease, cardiac source, and other cause) [2] systems because CCS offers more precise categorization of stroke causes, with more significant inter-category variability compared to intra-category variability. CCS reassigns 20-40% of cases from the undetermined category in other systems to specific subtypes, providing enhanced discrimination for clinical, imaging, and prognostic characteristics. The unknown category is markedly smaller in CCS (33%) compared to TOAST (53%) and ASCO (42%), highlighting its superior accuracy in categorizing stroke etiologies [3].

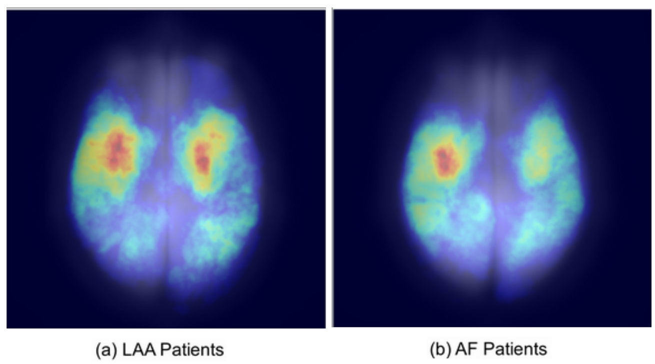

Fig. 2. Heatmaps of the infarct regions in the axial plane for (a) LAA patients and (b) AF patients across the dataset. These maps were generated from the ground-truth segmentation masks. These ground-truth masks were not used in any steps of developing EDAF. These heatmaps reveal notable similarities in spatial patterns across both AF and LAA groups. Despite these common patterns, differences in intensity highlight distinct variations in the distribution of infarct areas between the groups.

Based on CCS, all patients are assigned one of two labels: stroke due to AF or stroke due to large artery atherosclerosis (LAA). For AF-related strokes, cases are identified based on clinical reports or visual analysis from a 12-lead ECG, excluding new AF diagnoses within 30 days after cardiac surgery and strokes caused by other cardioembolic sources. For LAA-related strokes, cases involve infarcts linked to significant parent artery stenosis, with other cardioembolic sources reasonably excluded per CCS criteria.

Our current work focuses on LAA and AF, two of the most common stroke subtypes, while excluding other stroke etiologies. These subtypes are particularly significant in a clinical setting due to their higher rates of recurrence and severity [5,14]. Diagnosing LAA typically involves vascular imaging, whereas cardiac evaluation is essential for identifying AF, contrasting with small artery occlusion (SAO), which is often detectable with just brain MRI [4]. Moreover, targeted preventive measures for high-risk etiologies such as LAA and AF generally yield more significant risk reductions than those for lower-risk etiologies like SAO [3], highlighting their importance for early intervention. Furthermore, the 'other determined cause' category includes a variety of complex stroke mechanisms that are difficult to categorize and thus excluded from our work.

Figure 2 shows heatmaps of the infarct regions in the axial plane for both LAA and AF patients across the dataset. These heatmaps were generated from ground-truth segmentation masks, which were not used in developing EDAF. The heatmaps reveal notable similarities in spatial patterns across both AF and LAA groups. Despite these common patterns, differences in intensity highlight distinct variations in the distribution of infarct areas between the groups. These variations can provide valuable insights for deep learning models to distinguish between AF and LAA mechanisms.

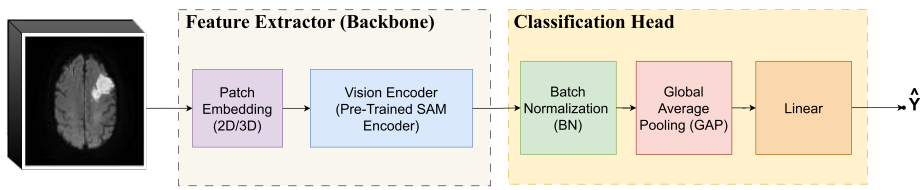

Fig. 3. Overview of the proposed classification network consisting of three components: (1) an embedding layer that projects 3D images onto 2D surfaces, (2) a vision encoder that extracts high-quality feature representations from projected 2D images, and (3) a classification head producing a probabilistic outcome indicating the likelihood of underlying AF presence. The customized patch embedding layer enables the processing of 3D MRI volumes using SAM's 2D vision encoder.

2.3 Network Architecture

Our proposed approach, a deep neural network method to process brain MRIs to detect the presence of AF, is shown in Fig. 3. It consists of three primary components: (1) an embedding layer that projects 3D images onto 2D surfaces, (2) a vision encoder that extracts high-quality feature representations from projected 2D images, and (3) a classification head producing a probabilistic outcome indicating the likelihood of underlying AF presence.

Embedding Layers. Embedding layers, originally introduced for natural language processing (NLP) tasks, transform images or image patches into continuous vectors, simplifying their complexity and effectively encoding the image context for input representations. We investigated two approaches for the embedding layer, focusing on 2D and 3D convolutions.

2D Embedding Layer: First, we use different 3D to 2D projection methods to transform the 3D MRI images into 2D surfaces. Subsequently, the 2D embedding layer of SAM is used to generate embedded input to the vision encoder. Several projection techniques, including maximum intensity projection, mean intensity projection, and sum intensity projection, are employed to convert all the voxels along the depth of the image into a single pixel within the 2D surface. By utilizing these projection techniques, the resulting images possess a depth and number of channels equal to one. We then construct a 3-channel input for the embedding layer by either copying one projection output to all three channels or using each one as a channel of the input.

3D Embedding Layer: We use 3D convolution to convert the 2D embedding layer in SAM to 3D space. We keep the default kernel size and stride in the x and y directions and implement a kernel size and stride of 64 (the input image depth) in the z-direction. The weights of the 2D embedding layer from SAM are copied to the 3D convolutional layer across the depth. This convolutional layer produces an embedding of the input image that can be replicated into three channels for use as an input to the vision encoder.

Vision Encoder. Vision encoders play a crucial role in classification networks by converting raw image data into feature maps, which is essential for deep learning models to make accurate predictions. Vision Transformers (ViTs) [6] have recently emerged as a popular choice for vision encoders in various computer vision tasks. ViTs leverage the transformer architecture, initially designed for natural language processing tasks. The embedding layer output is processed by the transformer layers, where self-attention mechanisms allow the model to weigh the importance of each image patch relative to the others. The output of these transformer layers consists of rich feature embeddings. These embeddings are subsequently reshaped and transformed into feature maps that can be utilized for various downstream tasks, including image classification, by feeding them into a final classification head.

ViTs offer several advantages for vision encoders in classification tasks. One of the primary benefits is their ability to model global context by treating images as sequences of patches, enabling them to capture long-range dependencies within the data. This holistic understanding of the image content often improves performance on complex vision tasks where global context is crucial. Additionally, ViTs are highly scalable; their performance can be enhanced by increasing the number of layers or the embedding dimensions, making them suitable for high-resolution images and large datasets. Moreover, ViTs tend to generalize well across different datasets and tasks, partly due to their capability to leverage large-scale pre-training on diverse image collections. This robustness and versatility make ViTs a powerful choice for vision encoders in modern deep-learning applications.

Our study uses the MedSAM's vision encoder [17] for feature extraction from the input embeddings. MedSAM is a refined version of the base SAM with the same architecture, optimized explicitly for medical image segmentation. Developed on a substantial dataset containing over 1.5 million image-mask pairs across various imaging modalities and pathological conditions, MedSAM demonstrates superior performance in medical tasks. This makes it an excellent choice for our brain MRI classification task, where accurately identifying infarct patterns is crucial. To ensure the robustness of our methodology, we performed comparative analyses using both SAM and MedSAM encoders. By benchmarking MedSAM results against the SAM encoder, we could evaluate the effectiveness of our chosen approach. We conducted experiments with different scales of SAM (base, large, and huge) to determine the most effective configuration for our specific application.

Classification Head. As shown in Fig. 3, the proposed classification head consists of batch normalization, global average pooling (GAP), and a linear layer. Batch normalization stabilizes and accelerates training by normalizing layer inputs. GAP reduces the spatial dimensions of the feature maps, summarizing them into a fixed-size vector. The linear layer then computes class probabilities based on this summarized representation. This structure ensures a focused evaluation of the SAM encoder's performance within the classification framework.

3 Experiments

For a thorough evaluation of our proposed model, we also used two of the frequently used CNNs, ResNet-101 [8] and ConvNeXt [16], to perform 3D image classification and compare them with our presented approach. Different depths of ResNet were experimented with, and ResNet-101 attained the best results, which are included here. Additionally, we employed pre-trained ResNet variants on video datasets due to the lack of publicly available pre-trained 3D CNNs on medical datasets. Transfer learning based on these models led to poor performance compared to training from scratch; hence, the results are not included here.

3.1 Implementation Details

For all experiments, the DWI images were preprocessed following the procedure outlined in [17] and subsequently normalized to account for the variations in imaging equipment. The intensities were clipped to be within the range of 0.95th and 99.5th percentiles, and then, using min-max normalization, they were rescaled to [0,255]. For EDAF experiments, all samples were resized to the shape of (64, 1024, 1024). The images were split into three channels, except for the 3-channel projection, in which each channel was computed using a different projection technique. Since this work aimed to evaluate SAM classification performance without dependence on other factors, data augmentation was not used in fine-tuning SAM. When training the CNNs, the images were resized to (64, 256, 256) to reduce the computational complexity of these models and used as 1-channel inputs. Data augmentation was applied while training the CNNs

Table 2. Data augmentation details for 3D CNNs

Augmentation	Parameters
Flip in all three directions	-
Elastic deformations	deformation_limits = $(0, 0.5)$
90 degrees rotation in the axial plane	-
Gaussian noise	var_limit = $(0, 1)$
Random crop	shape = $(55, 180, 180)$
Random scale	scale_limit = $(0.7, 1.3)$

from scratch to achieve results comparable to our proposed method. Table 2 shows the specifics of the augmentations applied to train these CNNs.

For all experiments, the dataset was split into training (70%, 97 AF, and 68 control), validation (10%, 14 AF, and 10 control), and testing (20%, 27 AF, and 19 control) set for evaluation of the proposed method. To ensure consistency of the distribution of the classes in all subsets, stratified sampling was used for data splitting. The initial EDAF experiments during the model development were performed using the same seed for data splitting. After selecting the final EDAF model, the same experiment with a consistent setup was repeated five times using different seeds to ensure the reliability of the model's performance. All the CNN experiments were repeated five times using the same seeds as the EDAF experiments. The models were implemented using PyTorch [20] and trained on an NVIDIA A100 GPU using the stochastic gradient descent (SGD) optimizer (learning rate for head = 0.1, learning rate for backbone = 0.01, decay rate for step scheduler = 0.1, scheduler step size = 50). The learning rate for the CNN experiments was set to 0.01 for both the backbone and the classification head. The binary cross-entropy (BCE) loss function was selected for our binary classification task. The maximum number of epochs was set to 300, and the batch size to 4. Early stopping (based on the validation loss with a patience of 5 epochs), and L_2 regularization (weight decay = 0.001) were implemented to prevent overfitting. Later on, the final model was evaluated on the unseen test set.

3.2 Evaluation

Comparative Evaluation. We compared our method of identifying AF from brain MRI with similar studies. The comparison results are shown in Table 3. Our strategy was most effective when we used the MedSAM as the encoder, updated only the classification head weights in the network, and used the mean of the intensity values along the depth for 3D to 2D projections. The Area Under the Curve–Receiver Operating Curve (AUC-ROC) value is 83.08 ± 2.96, indicating a powerful discriminating ability.

Our model outperforms [30] even though we used a smaller set of MRI images with 235 samples compared to 489 and did not use radiomic features extracted from segmentation masks. However, there is not enough information in [30] to compare our models regarding computational complexity. In [12], a segmentation model is utilized to enhance the performance of the classification network using two distinct models. Their goal is to classify stroke subtypes, achieving an accuracy of 81.9%, but their approach requires annotating the infarct regions for ad hoc training. On the other hand, our proposed method uses only binary class labels to train a single 2D classification model, demonstrating its efficiency in terms of human labor and computational resources.

EDAF proves to be a superior choice for our brain MRI binary classification task compared to the two CNNs trained from scratch. ResNet-101 [8] and ConvNeXt [16] required extensive data augmentation for generalizability and good performance, but our presented approach outperformed them without additional data augmentation. Its effectiveness is notably remarkable, requiring significantly fewer epochs for fine-tuning (<20 across all experiments), while CNNs required 200-300 epochs to converge to optimal performance.

Moreover, leveraging the pre-trained model benefits from prior knowledge and enhances performance even with limited data. Using MedSAM results in a smaller variation in the AUC-ROC across different data splits compared to CNNs, showcasing that the model's performance is less dependent on data splitting or variations in the data. These factors collectively underscore EDAF's effectiveness and efficiency in tackling the challenges posed by medical image analysis tasks like ours.

Table 3. Comparison with other methods. The experiments demonstrate the efficacy of our methodology, surpassing previous research and 3D CNNs while being more efficient to train.

Method	Accuracy	Precision	Recall	F1 score	AUC-PR	AUC-ROC
Kim et al. [12]	81.9	-	-	-	-	-
Zhang et al. [30]	70	63.8	92.5	75.5	-	79.9
ResNet-101	69.78 ± 6.22	78.28 ± 7.12	68.99 ± 6.17	72.90 ± 2.89	87.71 ± 2.48	79.31 ± 6.28
ConvNeXt	74.46 ± 8.18	83.64 ± 7.80	72.21 ± 8.03	77.07 ± 5.73	88.15 ± 2.03	80.54 ± 5.51
Ours	72.21 ± 7.93	71.90 ± 10.90	**95.91 ± 4.27**	**81.67 ± 6.82**	**92.56 ± 2.58**	**83.08 ± 2.96**

Fine-Tuning SAM Variants. We investigated four variations of SAM with available pre-trained weights. This includes three different sizes of the original SAM and a fine-tuned version of SAM on medical images (MedSAM) [17]. We evaluated different knowledge transfer methods to our target domain and incorporated five strategies. These approaches were evaluated for all four variants of SAM mentioned above. Only the following modules were updated during training in these five strategies, and the rest of the model's parameters were frozen. For example, only the classification head was updated in Strategy 1 (S1). Likewise, the other strategies include S2 (classification head and embedding layer), S3 (classification head and vision encoder), S4 (classification head and vision neck), and S5 (classification head, embedding layer, and vision encoder). For these experiments, the 3D embedding layer was used.

The results of these tests are shown in Fig. 4. The MedSAM, huge SAM, and large SAM achieve the highest AUC-ROC (82.03) following different strategies, whereas MedSAM has the same size as base SAM. Furthermore, the results from different tuning strategies indicate that updating the encoder in parts or its entirety leads to inferior results compared to using a frozen one. Generally, updating the embedding layer improves the classification performance.

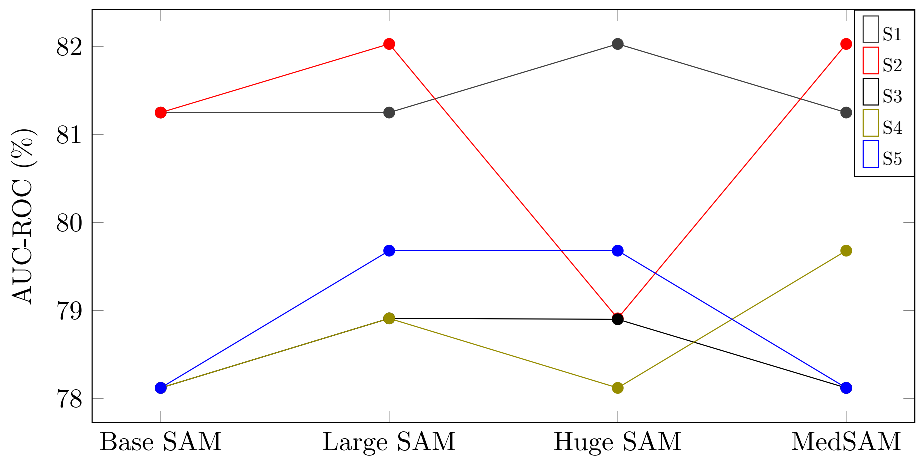

Fig. 4. Results of five fine-tuning strategies (S1-S5) for the four variants of SAM. In each strategy, only the following modules were updated during the training: S1 (classification head), S2 (classification head and embedding layer), S3 (classification head and vision encoder), S4 (classification head and vision neck), and S5 (classification head, embedding layer, and vision encoder). The best-performing models achieved AUC-ROC of 82.03% by updating the classification head and embedding layer for (1) MedSAM and (2) large SAM and updating the classification head for (3) huge SAM.

Embedding Layer. We explored different approaches for integrating the 3D input into an acceptable input for SAM. As discussed in the methodology (Section 2), we used 2D and 3D embedding techniques for our experiments. The MedSAM encoder served as the foundation for these experiments, and only the parameters of the classification head were updated during training. As shown in Table 4, using SAM's default embedding layer along with sum (82.03%) or mean (83.59%) intensity projection approaches achieves better results than using a 3D embedding layer (81.25%) in AF identification. These static approaches proved effective in our application, reaching the highest performance without requiring any changes to the architecture.

In contrast to prior studies such as M2Net [31] that have shown poorer performance with projections compared to methods using interpolated 3D approaches like MMMNA-Net [25], our approach demonstrates improved results due to the unique characteristics of infarct regions in DW-MRI images. In stroke etiologic classification, the decision-making process for the underlying cause is associated with the infarct regions in the brain, which typically exhibit hyperintensity, and projections like mean or sum effectively preserve essential information from these regions. Whether this approach applies to other tasks or modalities should be investigated.

Table 4. Comparing different embedding methods using frozen MedSAM's encoder and updating the classification head: employing SAM's default 2D embedding layer along with sum or mean projection outperforms custom 3D embedding layer. However, all three approaches perform better than 3D CNNs and previous works in the literature.

Embedding method	Accuracy	Precision	Recall	F1 score	AUC-PR	AUC-ROC
3D embedding	66.66	68.18	93.75	78.94	92.23	81.25
Sum projection	66.66	66.66	**100**	80	92.76	82.03
Mean projection	66.66	66.66	**100**	80	**93.37**	**83.59**
Max projection	**75**	**72.72**	**100**	**84.21**	90.13	76.56
3-Channel projection	62.5	68.42	81.25	74.28	88.44	72.65

3.3 Discussion

Due to several key advantages, SAM and its variants were chosen as the encoder for the classification task. SAM was trained on the largest segmentation dataset to date, comprising over 1 billion masks on 11 million images. SAM's training on this large dataset enhances its robustness and generalization capabilities. It generalizes well across diverse visual data, from underwater to microscopy images, which is crucial for applications involving varied data distributions. As a foundation model, SAM is adaptable to a wide range of downstream tasks, extending its utility beyond segmentation. Its strong zero-shot transfer capabilities across multiple tasks, such as edge detection and object proposal generation, highlight its potential effectiveness in classification without extensive task-specific training. Besides these capabilities of SAM, MedSAM has been specifically optimized for medical images using a comprehensive dataset of over 1.5 million images, enhancing its suitability for our task. These inherent strengths of SAM and MedSAM make them compelling choices for classifying brain MRI images. We utilized the rich semantic information from pre-trained SAM and MedSAM models to propose a novel computationally efficient 3D medical image classification pipeline. Our model outperforms existing works in the literature, even with small-scale datasets and without using segmentation masks. Furthermore, we have shown that a large segmentation model (such as SAM) can be used to develop a classification model that is computationally efficient for learning environments with limited resources.

We acknowledge that focusing on LAA and AF limits our study to specific stroke subtypes. Nevertheless, this study marks the first effort to create a fully automated deep classification framework for early identification of AF as the stroke etiology from brain MRI scans. We ensure a solid foundation with reliable ground truth data by beginning with these categories, which offer clear clinical and imaging markers. This approach allows us to develop robust diagnostic models within a manageable scope and, later on, expand our methodology to include other stroke mechanisms to enhance the framework's clinical relevance and utility.

Additionally, this study did not include clinical variables commonly used for AF risk assessment in patients. In the future, we will develop a multi-modal framework to aggregate the imaging and clinical information to create a personalized AF risk assessment pipeline. Moreover, DWI, apparent diffusion coefficient (ADC) maps, fluid-attenuated inversion recovery (FLAIR), and often perfusion imaging are commonly integrated into clinical practices for stroke evaluation. This study uses only DWI images due to the unavailability of ADC maps and other imaging modalities. Future research can incorporate them to enhance stroke etiologic classification and identification of AF-related strokes.

4 Conclusion

Atrial Fibrillation (AF) is a common cause of ischemic stroke. Having AF can increase the risk of stroke by up to five times and the risk of recurrent stroke by 2.1 times. Therefore, identifying AF in the early stages of a stroke is crucial for establishing preventive strategies and optimizing treatment approaches. We presented the EDAF framework, which is a non-invasive, cost-effective screening tool and demonstrated the potential of utilizing post-stroke DWI-MRI images for early detection of underlying AF. The proposed deep learning method effectively harnessed the power of a foundational model trained on millions of samples, demonstrating an efficient pipeline for adapting it to analyze brain MRI.

It should be noted that the purpose of this model is to prioritize patients based on their AF risk. This model does not replace the current methods of AF detection, nor is it a tool for conclusive AF diagnosis. Patients with higher AF risk should undergo extensive cardiac monitoring to diagnose AF and start anticoagulation if clinicians deem it necessary. External validation is critical to assess the generalizability and reliability of the proposed method's performance across diverse datasets and clinical settings.

Acknowledgments. This research was supported by The University of Melbourne's Research Computing Services and the Petascale Campus Initiative.

Disclosure of Interests. The authors have no conflicts of interest to declare that are relevant to the content of this article.

References

1. Adams Jr, H.P., Bendixen, B.H., Kappelle, L.J., Biller, J., Love, B.B., Gordon, D.L., Marsh 3rd, E.: Classification of subtype of acute ischemic stroke. definitions for use in a multicenter clinical trial. toast. trial of org 10172 in acute stroke treatment. stroke **24**(1), 35–41 (1993)
2. Amarenco, P., Bogousslavsky, J., Caplan, L., Donnan, G., Hennerici, M.: New approach to stroke subtyping: the asco (phenotypic) classification of stroke. Cerebrovascular diseases **27**(5), 502–508 (2009)
3. Arsava, E.M., Helenius, J., Avery, R., Sorgun, M.H., Kim, G.M., Pontes-Neto, O.M., Park, K.Y., Rosand, J., Vangel, M., Ay, H.: Assessment of the predictive validity of etiologic stroke classification. JAMA neurology **74**(4), 419–426 (2017)
4. Ay, H., Benner, T., Murat Arsava, E., Furie, K.L., Singhal, A.B., Jensen, M.B., Ayata, C., Towfighi, A., Smith, E.E., Chong, J.Y., et al.: A computerized algorithm for etiologic classification of ischemic stroke: the causative classification of stroke system. Stroke **38**(11), 2979–2984 (2007)
5. Bang, O.Y., Lee, P.H., Joo, S.Y., Lee, J.S., Joo, I.S., Huh, K.: Frequency and mechanisms of stroke recurrence after cryptogenic stroke. Annals of neurology **54**(2), 227–234 (2003)
6. Dosovitskiy, A., Beyer, L., Kolesnikov, A., Weissenborn, D., Zhai, X., Unterthiner, T., Dehghani, M., Minderer, M., Heigold, G., Gelly, S., et al.: An image is worth 16x16 words: Transformers for image recognition at scale. arXiv preprint arXiv:2010.11929 (2020)
7. Fukamizu, S., Hojo, R., Kitamura, T., Kawamura, I., Miyazawa, S., Karashima, J., Nakamura, S., Takeda, K., Yamaoka, K., Arai, T., et al.: Recurrent ischemic stroke in patients with atrial fibrillation ablation and prior stroke: a study based on etiological classification. Journal of Arrhythmia **36**(1), 95–104 (2020)
8. He, K., Zhang, X., Ren, S., Sun, J.: Deep residual learning for image recognition. In: Proceedings of the IEEE conference on computer vision and pattern recognition. pp. 770–778 (2016)
9. He, S., Bao, R., Li, J., Grant, P.E., Ou, Y.: Accuracy of segment-anything model (sam) in medical image segmentation tasks. arXiv preprint arXiv:2304.09324 (2023)
10. Huang, Y., Yang, X., Liu, L., Zhou, H., Chang, A., Zhou, X., Chen, R., Yu, J., Chen, J., Chen, C., et al.: Segment anything model for medical images? Medical Image Analysis **92**, 103061 (2024)
11. Kalarus, Z., Mairesse, G.H., Sokal, A., Boriani, G., Średniawa, B., Casado-Arroyo, R., Wachter, R., Frommeyer, G., Traykov, V., Dagres, N., et al.: Searching for atrial fibrillation: looking harder, looking longer, and in increasingly sophisticated ways. an ehra position paper. Europace **25**(1), 185–198 (2023)
12. Kim, B.K., Park, S., Han, M.K., Hong, J.H., Lee, D.I., Yum, K.S.: Deep learning for prediction of mechanism in acute ischemic stroke using brain diffusion magnetic resonance image. Journal of Neurocritical Care **16**(2), 85–93 (2023)
13. Kirillov, A., Mintun, E., Ravi, N., Mao, H., Rolland, C., Gustafson, L., Xiao, T., Whitehead, S., Berg, A.C., Lo, W.Y., et al.: Segment anything. arXiv preprint arXiv:2304.02643 (2023)
14. Ko, Y., Lee, S., Chung, J.W., Han, M.K., Park, J.M., Kang, K., Park, T.H., Park, S.S., Cho, Y.J., Hong, K.S., et al.: Mri-based algorithm for acute ischemic stroke subtype classification. Journal of stroke **16**(3), 161 (2014)

15. Lee, K.J., Jung, K.H., Byun, J.I., Kim, J.M., Roh, J.K.: Infarct pattern and clinical outcome in acute ischemic stroke following middle cerebral artery occlusion. Cerebrovascular Diseases **38**(1), 31–38 (2014)
16. Liu, Z., Mao, H., Wu, C.Y., Feichtenhofer, C., Darrell, T., Xie, S.: A convnet for the 2020s. In: Proceedings of the IEEE/CVF conference on computer vision and pattern recognition. pp. 11976–11986 (2022)
17. Ma, J., He, Y., Li, F., Han, L., You, C., Wang, B.: Segment anything in medical images. Nature Communications **15**(1), 654 (2024)
18. Mazurowski, M.A., Dong, H., Gu, H., Yang, J., Konz, N., Zhang, Y.: Segment anything model for medical image analysis: an experimental study. Medical Image Analysis **89**, 102918 (2023)
19. McGrath, E.R., Kapral, M.K., Fang, J., Eikelboom, J.W., O'Conghaile, A., Canavan, M., O'Donnell, M.J., of the Ontario Stroke Registry, I., et al.: Association of atrial fibrillation with mortality and disability after ischemic stroke. Neurology **81**(9), 825–832 (2013)
20. Paszke, A., Gross, S., Massa, F., Lerer, A., Bradbury, J., Chanan, G., Killeen, T., Lin, Z., Gimelshein, N., Antiga, L., et al.: Pytorch: An imperative style, high-performance deep learning library. Advances in neural information processing systems **32** (2019)
21. Penado, S., Cano, M., Acha, O., Hernández, J.L., Riancho, J.A.: Atrial fibrillation as a risk factor for stroke recurrence. The American journal of medicine **114**(3), 206–210 (2003)
22. Roy, S., Wald, T., Koehler, G., Rokuss, M.R., Disch, N., Holzschuh, J., Zimmerer, D., Maier-Hein, K.H.: Sam. md: Zero-shot medical image segmentation capabilities of the segment anything model. arXiv preprint arXiv:2304.05396 (2023)
23. Singhal, A.B., Topcuoglu, M.A., Buonanno, F.S.: Acute ischemic stroke patterns in infective and nonbacterial thrombotic endocarditis: a diffusion-weighted magnetic resonance imaging study. Stroke **33**(5), 1267–1273 (2002)
24. Sposato, L.A., Andrade, J.: Prolonged cardiac monitoring for atrial fibrillation detection after stroke: in search of the elusive sweet spot (2022)
25. Tang, W., Zhang, H., Yu, P., Kang, H., Zhang, R.: Mmmna-net for overall survival time prediction of brain tumor patients. In: 2022 44th Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC). pp. 3805–3808. IEEE (2022)
26. Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A.N., Kaiser, Ł., Polosukhin, I.: Attention is all you need. Advances in neural information processing systems **30** (2017)
27. Wu, J., Fu, R., Fang, H., Liu, Y., Wang, Z., Xu, Y., Jin, Y., Arbel, T.: Medical sam adapter: Adapting segment anything model for medical image segmentation. arXiv preprint arXiv:2304.12620 (2023)
28. Zhang, K., Liu, D.: Customized segment anything model for medical image segmentation. arXiv preprint arXiv:2304.13785 (2023)
29. Zhang, Y., Shen, Z., Jiao, R.: Segment anything model for medical image segmentation: Current applications and future directions. Computers in Biology and Medicine p. 108238 (2024)
30. Zhang, Z., Lin, K., Wang, J., Ding, L., Sun, Y., Fu, C., Qian, D., Li, J., Huang, D.: Searching for underlying atrial fibrillation using artificial intelligence-assisted mri images from ischemic stroke patients. European Heart Journal **43**(Supplement_2), ehac544–543 (2022)

31. Zhou, T., Fu, H., Zhang, Y., Zhang, C., Lu, X., Shen, J., Shao, L.: Mˆ 2 net m 2 net: Multi-modal multi-channel network for overall survival time prediction of brain tumor patients. In: Medical Image Computing and Computer Assisted Intervention–MICCAI 2020: 23rd International Conference, Lima, Peru, October 4–8, 2020, Proceedings, Part II 23. pp. 221–231. Springer (2020)