






Review

# Transformer and Attention-Based Architectures for Segmentation of Coronary Arterial Walls in Intravascular Ultrasound: A Narrative Review

Vandana Kumari <sup>1</sup>, Alok Katiyar <sup>1</sup>, Mrinalini Bhagawati <sup>2</sup>, Mahesh Maindarkar <sup>3</sup>, Siddharth Gupta <sup>4,5</sup>, Sudip Paul <sup>2</sup>, Tisha Chhabra <sup>6</sup>, Alberto Boi <sup>7</sup>, Ekta Tiwari <sup>8</sup>, Vijay Rathore <sup>5</sup>, Inder M. Singh <sup>5</sup>, Mustafa Al-Maini <sup>9</sup>, Vinod Anand <sup>5</sup>, Luca Saba <sup>7</sup> and Jasjit S. Suri <sup>5,10,11,12,13,\*</sup>

<sup>1</sup> School of Computer Science and Engineering, Galgotias University, Greater Noida 201310, India; vandana.soni80@gmail.com (V.K.); alok\_katiyar@hotmail.com (A.K.)

<sup>2</sup> Department of Biomedical Engineering, North Eastern Hill University, Shillong 793022, India; bhagawatimrinalini07@gmail.com (M.B.); sudip.paul.bhu@gmail.com (S.P.)

<sup>3</sup> School of Bioengineering Research and Sciences, MIT Art, Design and Technology University, Pune 412021, India; mahesh.nehu.333@gmail.com

<sup>4</sup> Department of Computer Science and Engineering, Bharati Vidyapeeth's College of Engineering, New Delhi 110063, India; siddgupta462@gmail.com

<sup>5</sup> Stroke Monitoring and Diagnostic Division, AtheroPoint™, Roseville, CA 95661, USA; rajvivs888@gmail.com (V.R.); drindersingh1@gmail.com (I.M.S.); vinodanand2010@gmail.com (V.A.)

<sup>6</sup> Department of Information Technology, Bharati Vidyapeeth's College of Engineering, New Delhi 110063, India; tishachhabra2702@gmail.com

<sup>7</sup> Department of Cardiology, University of Cagliari, 09124 Cagliari, Italy; albeboi@tiscali.it (A.B.); lucasabamd@gmail.com (L.S.)

<sup>8</sup> Department of Computer Science, Visvesvaraya National Institute of Technology (VNIT), Nagpur 440010, India; ekta.tiwari03@gmail.com

<sup>9</sup> Allergy, Clinical Immunology and Rheumatology Institute, Toronto, ON M5G 1N8, Canada; dralmainioffice@gmail.com

<sup>10</sup> Department of Electrical and Computer Engineering, Idaho State University, Pocatello, ID 83209, USA

<sup>11</sup> Department of Computer Engineering, Graphic Era Deemed to be University, Dehradun 248002, India

<sup>12</sup> Symbiosis Institute of Technology, Nagpur Campus, Symbiosis International (Deemed University), Pune 440008, India

<sup>13</sup> University Centre for Research & Development, Chandigarh University, Mohali 140413, India

\* Correspondence: jasjit.suri@atheropoint.com; Tel.: +1-(916)-749-5628



Academic Editor: Tong Ye

Received: 5 February 2025

Revised: 8 March 2025

Accepted: 20 March 2025

Published: 26 March 2025

**Citation:** Kumari, V.; Katiyar, A.; Bhagawati, M.; Maindarkar, M.; Gupta, S.; Paul, S.; Chhabra, T.; Boi, A.; Tiwari, E.; Rathore, V.; et al. Transformer and Attention-Based Architectures for Segmentation of Coronary Arterial Walls in Intravascular Ultrasound: A Narrative Review. *Diagnostics* **2025**, *15*, 848. <https://doi.org/10.3390/diagnostics15070848>

**Copyright:** © 2025 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

**Abstract: Background:** The leading global cause of death is coronary artery disease (CAD), necessitating early and precise diagnosis. Intravascular ultrasound (IVUS) is a sophisticated imaging technique that provides detailed visualization of coronary arteries. However, the methods for segmenting walls in the IVUS scan into internal wall structures and quantifying plaque are still evolving. This study explores the use of transformers and attention-based models to improve diagnostic accuracy for wall segmentation in IVUS scans. Thus, the objective is to explore the application of transformer models for wall segmentation in IVUS scans to assess their inherent biases in artificial intelligence systems for improving diagnostic accuracy. **Methods:** By employing the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) framework, we pinpointed and examined the top strategies for coronary wall segmentation using transformer-based techniques, assessing their traits, scientific soundness, and clinical relevancy. Coronary artery wall thickness is determined by using the boundaries (inner: lumen-intima and outer: media-adventitia) through cross-sectional IVUS scans. Additionally, it is the first to investigate biases in deep learning (DL) systems that are associated with IVUS scan wall segmentation. Finally, the study incorporates explainable AI (XAI) concepts into the DL structure for IVUS scan wall segmentation. **Findings:** Because of its capacity to automatically extract features at numerous scales in encoders, rebuild segmented pictures via decoders, and fuse variations

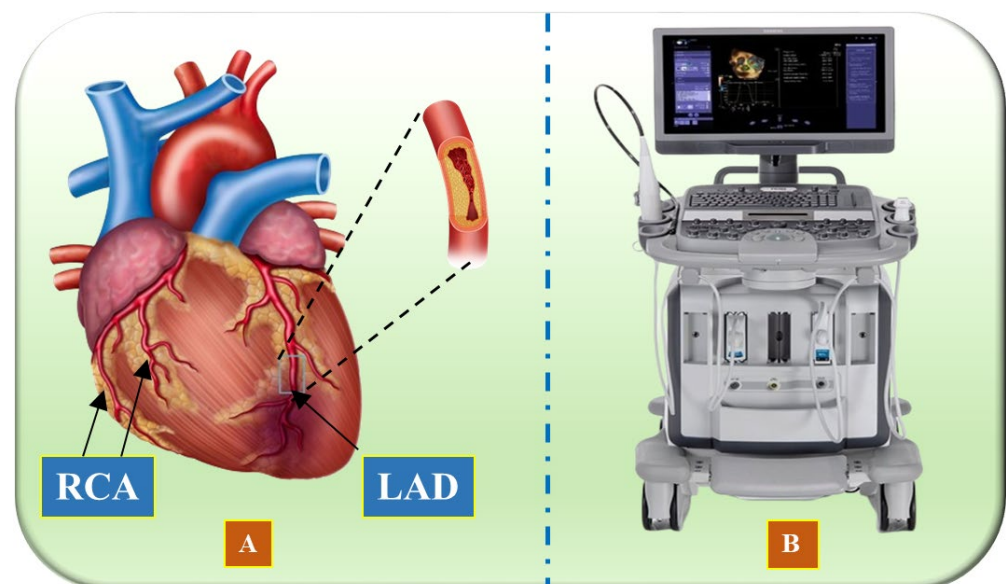
through skip connections, the UNet and transformer-based model stands out as an efficient technique for segmenting coronary walls in IVUS scans. **Conclusions:** The investigation underscores a deficiency in incentives for embracing XAI and pruned AI (PAI) models, with no UNet systems attaining a bias-free configuration. Shifting from theoretical study to practical usage is crucial to bolstering clinical evaluation and deployment.

**Keywords:** coronary artery disease; IVUS; wall segmentation; DL; UNet; transformer; AI bias

## 1. Introduction

Cardiovascular disease (CVD) is a major cause of morbidity globally [1]. Based on data analysis, research shows a forecast of a 96.6% mortality rate over a century in the area of CVD [2]. The two primary sources of mortality are coronary artery disease (CAD) and acute coronary syndrome (ACS) [3]. Due to the accumulation of atherosclerotic plaque inside the coronary walls, the narrowing of arteries causes CAD [3,4]. This build-up of atherosclerotic plaque makes the arterial walls thickened and stiff [5].

Vascular research and therapy have made significant progress over the last several decades, enhancing both the diagnosis and treatment of heart diseases [6–9]. Artificial intelligence (AI) solutions are enabled through the qualitative evaluation of morphology [10]. Several medical imaging modalities exist, namely, intravascular ultrasound (IVUS) [11–13], optical coherence tomography (OCT) [14,15], computed tomography (CT) [16], and magnetic resonance imaging (MRI) [17], which can be used to assist in identification, disease monitoring, and surgical planning [18–20]. An illustration of the coronary artery depicting the left anterior descending artery (LAD) and right coronary artery (RCA) is shown in Figure 1A, while the IVUS acquisition device is shown in Figure 1B.



**Figure 1.** (A) Illustration of the coronary arteries of the heart, highlighting the LAD (left anterior descending coronary artery) and RCA (right coronary artery) (Image courtesy of Atheropoint™, Roseville, CA, USA). (B) IVUS acquisition device (Image courtesy of Dr. Alberto Boi and Luca Saba, University of Cagliari, Cagliari, Italy).

IVUS stands out as a key modality [21], offering a unique perspective by allowing physicians to view coronary arteries internally [22]. Using a small ultrasound probe attached to a catheter, IVUS generates high-resolution, real-time images of the arterial

walls [23]. This enables an accurate assessment of vessel size, plaque composition, and stenosis severity for treatment planning [24]. The process of lumen-intima (LI) and media-adventitia (MA) boundary detection is known as wall segmentation in IVUS scans or IVUS segmentation [25]. Accurate and automatic segmentation of coronary artery borders is vital for developing computerized systems to detect coronary artery stenosis and plaque. However, it faces several challenges. Some branches are too narrow for precise segmentation, and individual variations in the coronary artery tree add complexity. Additionally, similar looking vascular organs near the heart can be mistaken for coronary arteries. The coronary arteries occupy only a small portion of the heart, leading to segmentation imbalances. The limited availability of expertly labelled datasets further complicates the process. Since AI's emergence, significant technological advancements have opened up numerous applications, including addressing these vascular segmentation challenges in medical imaging [26–30].

Machine learning (ML) approaches along with deep learning (DL) have greatly improved coronary artery segmentation by overcoming key challenges [31–34]. These techniques utilize large datasets (so-called big data) to detect complex patterns in IVUS images, enabling accurate and automated wall segmentation [35–37]. They reduce the need for manual intervention and subjective interpretation [38]. Additionally, ML and DL methods can adapt to various datasets and imaging conditions, making them versatile tools for wall segmentation in IVUS scans in different clinical settings [39]. In DL, solo DL (SDL) and hybrid DL (HDL) models both contribute to achieving high accuracy [40]. SDL models rely solely on neural networks trained on large datasets, while HDL models combine DL with other methods like rule-based systems or traditional ML to capitalize on their strengths [41,42]. For example, convolutional neural networks (CNNs) can be paired with conditional random fields (CRFs) [36,43] to enhance spatial segmentation. Hybrid models are effective in handling diverse medical imaging data, and by integrating handcrafted features with DL, they improve generalization and reduce overfitting, making them state-of-the-art [44].

Despite the significant advancements brought by DL models, including CNNs and hybrid approaches, previous state-of-the-art segmentation methods still face notable limitations. CNNs, while highly effective in feature extraction, are constrained by their limited receptive field, making them less capable of capturing long-range spatial dependencies in IVUS images. Additionally, deep CNN architectures require large, annotated datasets for robust training, yet IVUS datasets remain scarce and highly imbalanced, leading to biased segmentation outputs. Existing models also struggle with domain generalization, as variations in IVUS acquisition protocols and scanner types introduce challenges in adapting models across different datasets. Furthermore, computational complexity remains a barrier, particularly in clinical settings where real-time processing is crucial for intraoperative decision-making. Lastly, the black-box nature of many DL architectures raises concerns regarding explainability and clinical trust, making their direct adoption in healthcare settings challenging. These limitations underscore the need for more advanced architectures, such as transformer-based models, which offer superior global context understanding, improved generalization, and enhanced interpretability.

The integration of attention-based mechanisms into transformer architectures has led to significant improvements [45,46]. By harnessing the capabilities of attention and transformer-based models, wall segmentation algorithms in IVUS scans have attained greater precision and robustness [47], thereby enhancing diagnostic accuracy and facilitating more effective treatment strategies.

Consequently, we posit that transformer-based solutions outperform previously released models for segmenting walls in IVUS scans. While research on attention and the integration of transformers in DL architectures has grown steadily, no narrative review emphasizing transformers has been published. The proposed narrative review, utilizing

the PRISMA model to select top publications, is a pioneering effort. This review offers (i) an overview of the primary categories of wall segmentation models in IVUS with illustrative examples, (ii) the architectures of these fundamental categories, and (iii) a thorough statistical analysis considering various AI aspects, such as segmentation processes, performance evaluation metrics, optimizers, loss functions, and learning rates.

The study is structured as follows: Section 2 details the PRISMA model, outlining the process of study selection and statistical analysis of the chosen articles. Section 3 provides a comprehensive background on the introduction to architectures for wall segmentation in IVUS and their classifications. Section 4 engages in a critical discussion, covering principal findings, recommendations, strengths, weaknesses, and potential extensions. Lastly, Section 5 concludes the review, summarizing the key insights and implications for future research.

## 2. Search Strategy and Statistical Distributions

We employed the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) framework to systematically identify and analyze relevant studies in the domain of IVUS segmentation using DL and transformer-based models. The key terms utilized include DL, CVD, and transformer, along with specific phrases such as “IVUS segmentation”, “Coronary artery boundaries segmentation”, “IVUS segmentation using DL models”, “Lumen detection in IVUS scans”, “Media Adventitia detection in IVUS scans”, “Transformers in the coronary artery”, “Bias in DL/AI for CVD risk stratification”, “Attention in segmentation”, “IVUS segmentation using UNet”, and “Transformers in UNet”. To ensure comprehensive coverage, multiple academic databases—including Google Scholar, Science Direct, IEEE Xplore, PubMed, Springer, and Elsevier—were searched.

The PRISMA flowchart (Figure 2) illustrates the study selection process, ensuring a structured and unbiased review. A total of 740 records were obtained after removing duplicates from an initial 298 database searches and 442 additional sources. These records underwent a rigorous screening process, where 358 non-AI-related studies were excluded (E1), followed by 158 non-genomic articles (E2). After full-text assessment, an additional 67 records were removed due to insufficient data (E3), resulting in a final selection of 157 studies for qualitative synthesis.

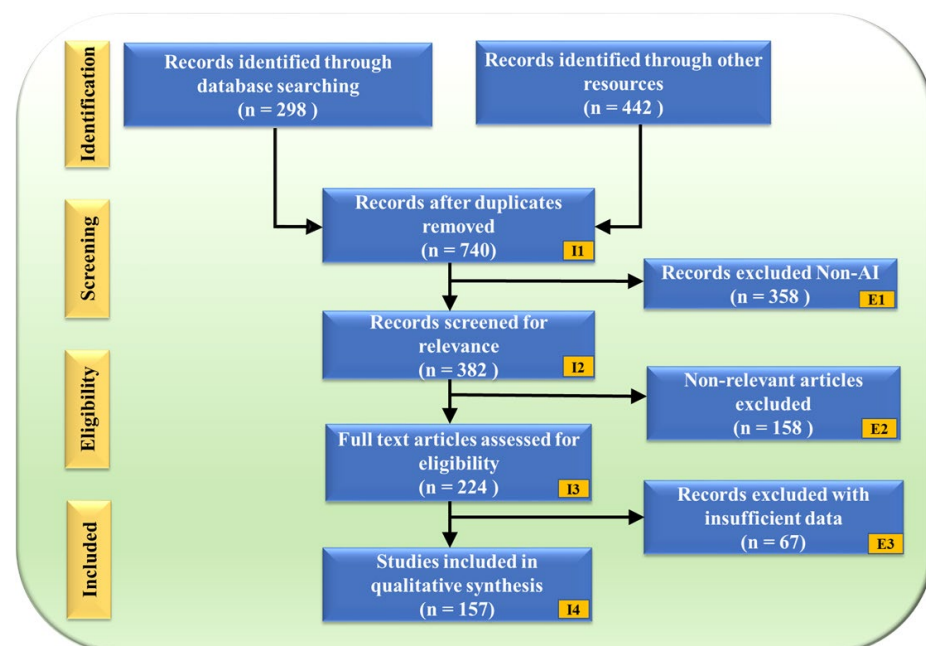
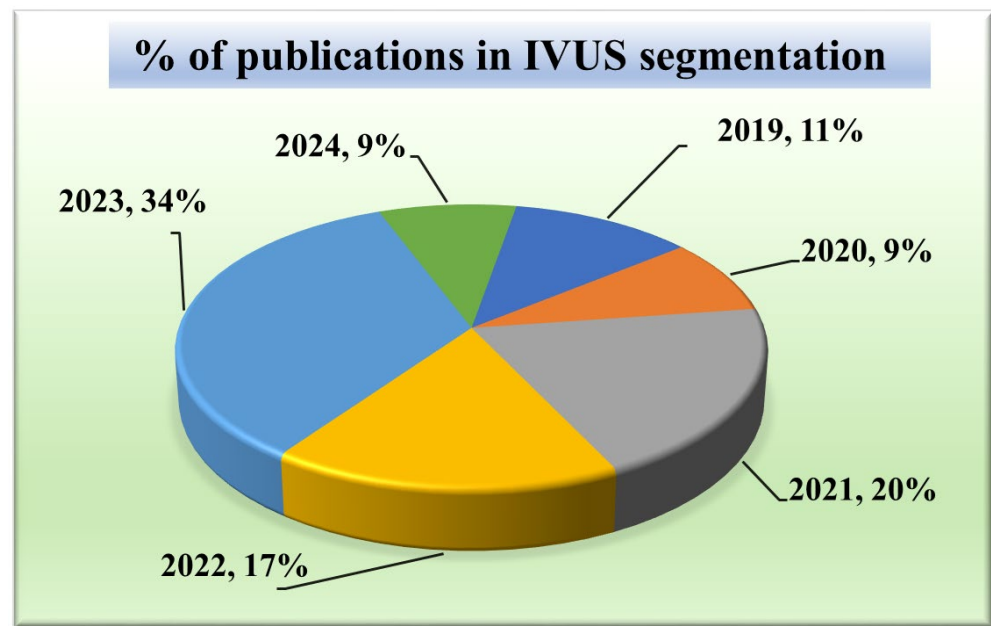


Figure 2. PRISMA model. “I” denotes Inclusion criteria, while “E” denotes Exclusion criteria.

The selected studies were analyzed for their segmentation model architectures, performance metrics, optimization techniques, and biases within AI-driven IVUS segmentation. Additionally, the review explored the integration of transformer models into UNet architectures and examined their impact on segmentation accuracy and clinical relevance. A statistical analysis was conducted to assess the strengths and limitations of the identified approaches, shedding light on bias in AI systems and the role of explainable AI (XAI) in IVUS segmentation.

### 2.1. Statistical Distributions

Given that our research focuses on segmentation techniques in IVUS scans, it is important to comprehend the current trends. Acquiring knowledge of the statistical distribution of the various segmentation techniques and their structures is essential to grasping the current state-of-the-art in this field. Furthermore, tracking the trend in publications over time might provide important insights into how IVUS segmentation techniques have changed and advanced. It can draw attention to how the performance and accuracy of these techniques are affected by technical developments, such as the addition of new algorithms or the integration of new architectures. Such statistical distributions of segmentation methods in IVUS frames are discussed in Figure 3.



**Figure 3.** Publications in IVUS segmentation.

#### 2.1.1. Publication Trend in IVUS Wall Segmentation

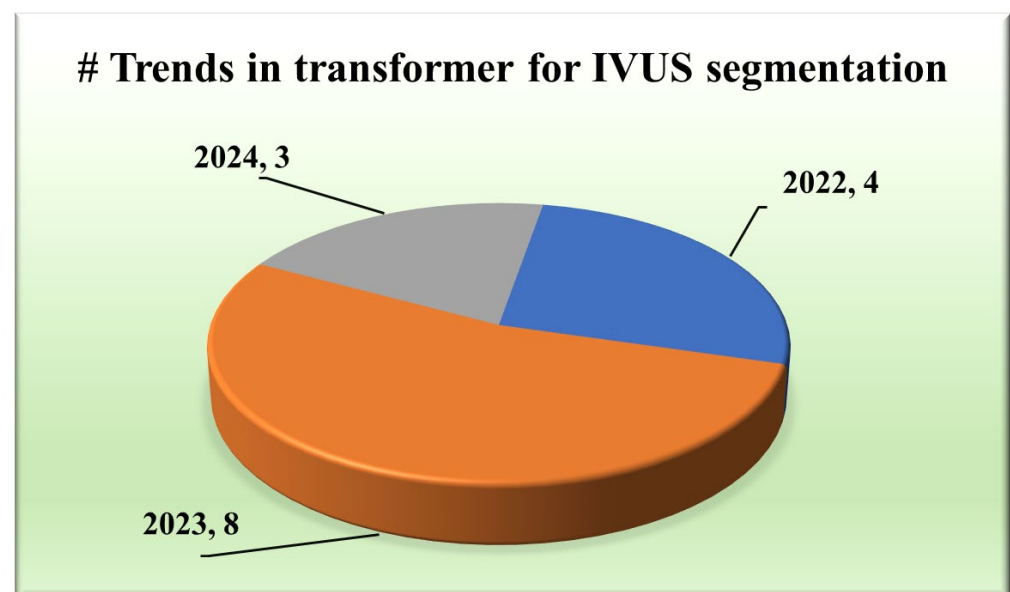
The chart in Figure 3 displays the number of research publications per year in the field of wall segmentation using IVUS from 2002 to 2024. It indicates growing research interest and progress in this area. The trend shows that 2021–2023 is the period where the highest amount of work was conducted in the field of IVUS. These potential spikes may correspond to significant advancements or new methodologies introduced such as attention and transformers in DL architectures.

#### 2.1.2. Publication Trend of Transformers in IVUS Segmentation

The integration of transformers within segmentation architectures [48–54] has emerged as a pivotal stride towards achieving unprecedented accuracy and yielding exclusive results. They use self-attention and hierarchical feature learning to understand the segmentation of IVUS images better, thereby improving the accuracy and reliability of segmentation.

### 2.1.3. Distribution of Type of Transformers

Various types of transformers, including the original transformer model, its variants, and more specialized architectures such as the UNet transformer, have been utilized in segmentation tasks. A gated axial attention layer has been introduced in medical transformers, which is used for building multi-head attention modules, increasing the receptive field of dilated transformers, etc. The introduction of vision transformers (ViTs) by Dosovitskiy et al. [55] in 2020 marked a significant milestone in adapting transformers for image segmentation tasks. However, in wall segmentation in IVUS, some specific architectures, namely the original (introduced in “Attention is All You Need”) [51,54], ViT [52], Cross-Stage Windowed Transformer (CSWin) [50], and Selective Transformer [49], have been found to be effective in capturing intricate features. Figure 4 illustrates this distribution.



**Figure 4.** Trends in transformers for IVUS.

### 2.1.4. Distribution by Publishers

A wide distribution indicates a more holistic approach, considering perspectives from various publishers and potentially mitigating biases inherent in relying solely on a few sources. Knowing the publisher’s distribution will aid in evaluating the credibility and reliability of the research. For that purpose, Figure 5 presents a pie chart showing the distribution of publishers of papers taken into consideration. It depicts that the most common publisher is Elsevier (36%) [49,52,56–65], followed by IEEE (31%) [12,51,54,66–76], Springer (13%) [48,77–81], MDPI (5%) [31,82], SPIE (4%) [83,84], and the others constituted by PLOS [85], Wiley [86], Sage Publications [30], and ACC [87].

### 2.1.5. Distribution of Evaluation Metrics

The AI system has some statistical distributions and they play an important part in designing. These are attributes which serve as evaluation metrics for assessing the performance of various approaches, providing valuable insights into their effectiveness in delineating the structures of IVUS scans. Among the plethora of metrics available, certain trends have emerged regarding their prevalent usage in wall segmentation in the IVUS literature, as seen in Figure 6. Notably, the Hausdorff surface distance (HSD) [30,49,51–54,57–62,67,69–71,73,75,78,79,88], Dice coefficient [48,49,53,58,62–65,67,69–71,73,74,78,87,89], and Jaccard index [30,48,49,51,52,54,58–63,67–71,73,75,79,81,82,87,88,90] (also known as Intersection over Union) have consistently been favored due to their robustness and interpretability. The maximum distance from the

segmented boundary point to the gold standard boundary nearest point is the HSD, providing insights into the surface dissimilarity. On the other hand, the Dice coefficient and Jaccard index measure the overlap between the segmented area and the ground truth area, with higher values indicating better agreement. Some other metrics have also been used namely, the percentage of area difference (PAD) [54,58–61,67,68,75,82,88], precision [48,53,68,74,76,80], and sensitivity [58,74,80,91] etc.

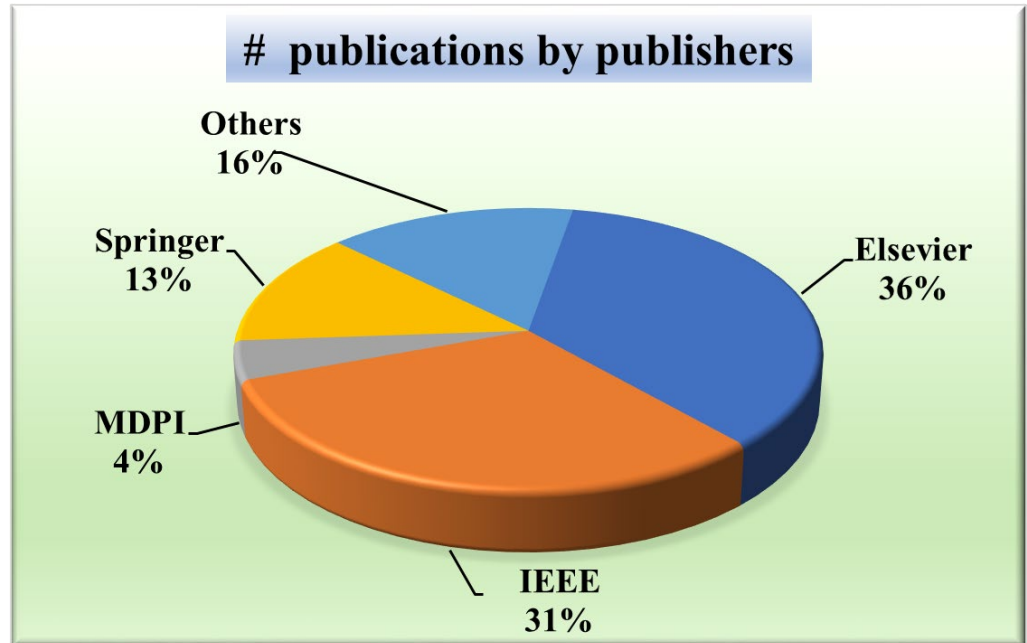


Figure 5. IVUS publications by publishers.

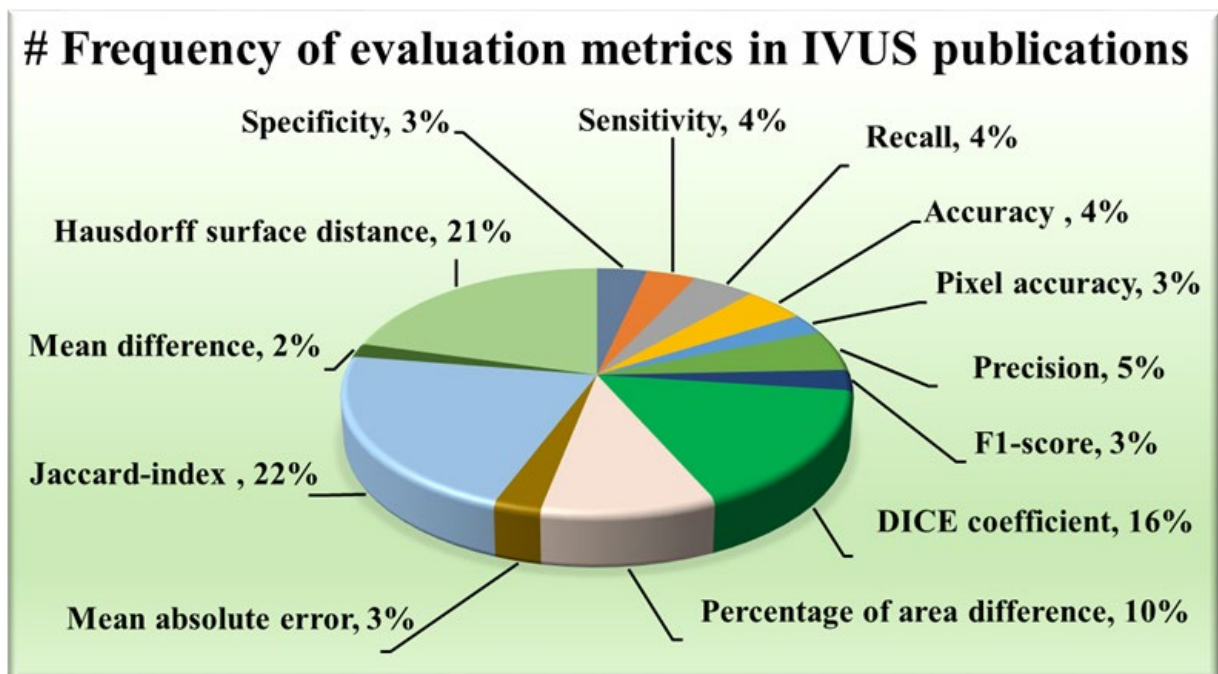


Figure 6. Frequency of evaluation metrics in IVUS publications.

### 3. Classification of IVUS Segmentation Architectures

The walls of a coronary artery can be visualized in a cross-sectional view using IVUS imaging and are crucial to delineate for accurate disease diagnoses. Manual segmentation,

while considered the gold standard, is time-consuming and laborious for cardiologists. Moreover, this method relies entirely on human subjectivity, making it prone to human errors and varying opinions. Therefore, there is a strong desire for the automated detection of intravascular borders, with researchers exploring various approaches in this ongoing research endeavor. Traditional image processing approaches [91], graph search [66], extremal regions selection [59], fuzzy system [92,93], and level set evolution [82,94] have come forward for this task. Wang et al. [88] introduced a feature selection method based on the Fractional-order Darwinian particle swarm optimization (FODPSO) algorithm. It is an intelligent optimization algorithm which provides better global search ability than PSO. These are categorized as conventional methods of segmentation since they typically rely on mathematical techniques rather than explicitly incorporating knowledge-based systems. Various ML algorithms have taken IVUS segmentation forward, offering diverse approaches to tackle its challenges. Supervised training methods [86,87] leverage labelled data to teach algorithms, while unsupervised techniques like K-means clustering [67] provide valuable insights by grouping similar image regions. Ensemble methods combine multiple algorithms [68], leveraging the strengths of each component and producing a powerful component. While traditional and ML techniques laid the groundwork, DL approaches [72] have revolutionized the field by overcoming several limitations. CNNs have demonstrated superior training efficacy compared to traditional ML algorithms due to their ability to process data in their raw form. They employ convolutional layers to extract patterns, eliminating the need for manual feature engineering and pooling layers for dimensionality reduction and abstraction. Furthermore, CNNs offer versatility in segmentation tasks through techniques like transfer learning. Pre-trained models such as InceptionV3 [95], DeepLabV3 [63], and ResNet [62] can be utilized for segmentation tasks. ResNet introduces a residual connection that mitigates the vanishing gradient problem, enabling smooth training of deeper networks. The primary objective of this review is to comprehensively analyze the four main categories of segmentation methods for IVUS boundaries that are instrumental in CVD risk stratification. Section 3 is organized into distinct sections, including (i) Conventional Techniques, (ii) ML Techniques, (iii) DL Techniques, and (iv) Attention and Transformer Methods. The DL techniques section is further subdivided into (i) Non-UNet methods and (ii) UNet and its Variants, providing a detailed exploration of each approach's unique contributions and applications in the context of IVUS boundary segmentation for CVD risk assessment.

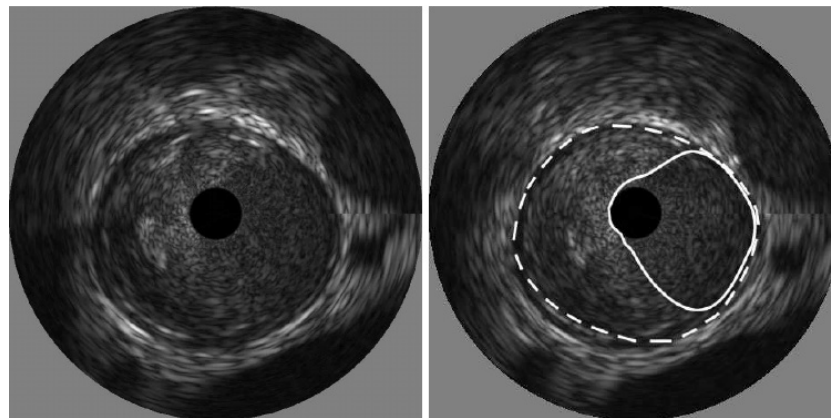
### 3.1. Conventional Methods

Earlier image processing has been a longstanding technique for segmenting structures within IVUS images. This involves applying various algorithms such as thresholding [60,96], active contours [56,58,96], fast marching method (FMM) [57] etc., to analyze the image data, identify edges, and differentiate tissue types based on intensity. Out of these, we have highlighted representative approaches such as thresholding and active contour models for segmenting walls in IVUS scans.

#### 3.1.1. Thresholding

The process of thresholding involves selecting a threshold value and then classifying pixels in the image based on intensity values. In IVUS segmentation, the threshold value is chosen to differentiate between the lumen and the surrounding tissues, such as the vessel wall or plaque deposits. Kermani et al. [60] presented a ground-breaking nonparametric statistical method for precisely identifying media-adventitia (MA) and lumen-intima (LI) borders in ultrasound frames. Unlike traditional techniques, this approach offers an independent border extraction process at each cutting angle. By enhancing the conventional

analysis of variance (ANOVA) criterion with a regularization term customized for IVUS images, the proposed method is designed to be versatile, fully automated, and suitable for parallel implementation. This process involves three steps, beginning with pre-processing the image. Here, a filter tackles the grainy speckle noise common in ultrasound data, while edge enhancement techniques like filtering or gradient calculations highlight the crucial borders. The second step involves dividing the image into small, overlapping windows, where each window pixel's intensities are analyzed, potentially involving calculations like mean, standard deviation, or even more complex statistics. By scrutinizing the statistical information, the algorithm identifies the locations for both the lumen and MA borders within each window. Since initial detections can be noisy, morphological operations like erosion and dilation come into play in the final stage for refinement. The advantages of this method include its ability to effectively remove catheter artefacts, estimate borders sequentially, and provide an ad hoc mechanism for detecting and correcting discontinuous borders, ultimately providing a Jaccard measure of  $0.84 \pm 0.07$  for Lumen, and  $0.82 \pm 0.11$  for MA border. The results of the detected boundaries for seven images are shown in Figure 7 (white bold is lumen border and dotted border is the MA border).



**Figure 7.** Detected boundaries of IVUS scan [70].

### 3.1.2. Active Contours

The active contour technique, often referred to as snakes, is a method used for segmentation tasks. It involves the delineation of boundaries within an image by iteratively deforming a parametric curve to fit the contours. The active contour model aims to minimize an energy functional  $E_{\text{snake}}$ , which is the sum of internal energy  $E_{\text{int}}$  and external energy  $E_{\text{ext}}$ .

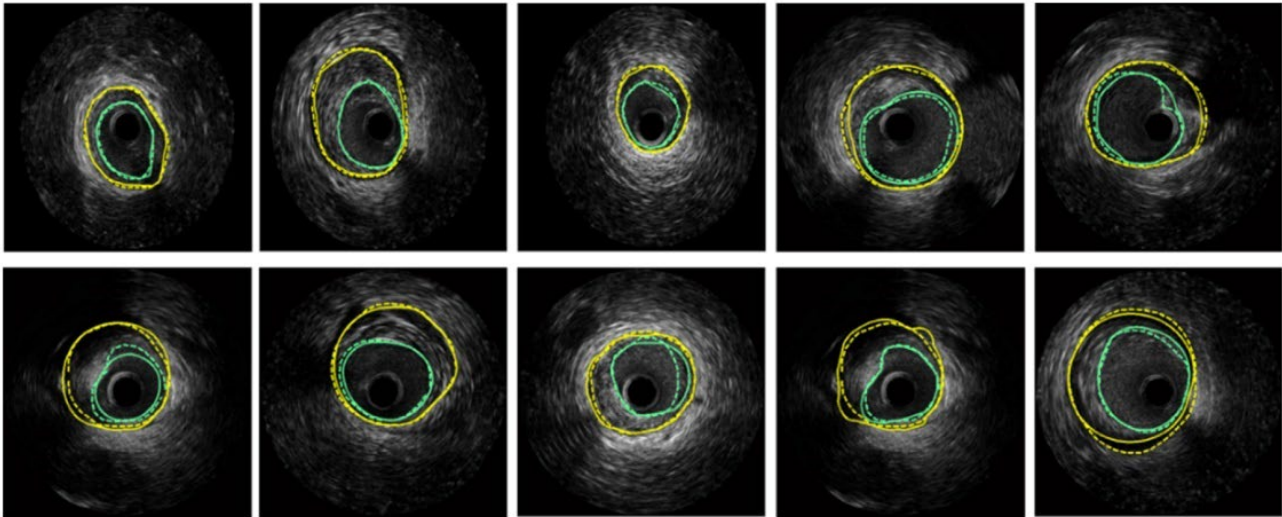
$$E_{\text{snake}} = E_{\text{int}} + E_{\text{ext}}$$

Internal energy represents the energy associated with the configuration of the snake itself. It encourages smoothness and regularity in the shape of the contour. External energy represents external forces acting on the snake. It guides the snake toward features of interest, such as boundaries or image gradients. Initialization is crucial as it provides the starting point for the snake's evolution. The final contour represents the boundaries of the segmented areas.

Giannoglou et al. [56] introduced a fully automated method for segmenting the LI and MA using an active contour model. The model's initialization leverages the inherent morphological characteristics of IVUS images through a two-step process: first, utilizing intensity information and then low-frequency information. This approach capitalizes on the adventitia's representation as a thick echo-dense ring in IVUS images, which appears as a thick bright curve in polar coordinates. The study demonstrated a 96% reduction in analysis time compared to manual segmentation. Wang et al. [96] developed

a framework for detecting the adventitia boundary by enhancing the traditional snake algorithm and incorporating Otsu thresholding, morphological operations, and connected component labeling.

Hammouche et al. [58] achieved significant success with a helical active contour algorithm based on the Rayleigh distribution of grey levels. This innovative algorithm is fast, employs an adaptive, simple, space curve for 3D lumen extraction, and is fully automated. It outperformed various traditional methods, achieving 98.9% accuracy. Figure 8 displays the segmentation results obtained on several IVUS frames, where turquoise color are the LI borders and yellow are the MA borders.



**Figure 8.** Segmentation of several IVUS frames [70], showing the lumen boundary (green) and media-adventitia (MA) boundary (yellow). Solid lines represent ground truth annotations, while dotted lines indicate predicted segmentations from the model.

### 3.2. Machine Learning Techniques

The turning point came when traditional image processing methods struggled to handle the complexity and variability of IVUS images. Machine learning offered a more adaptable approach, allowing algorithms to learn from data and improve segmentation accuracy. Several methods emerged, including probabilistic approaches, contrast features analysis [84], acoustic shadowing detection [97], support vector machines (SVMs) [98], XGBoost (XGB) [99], and more, each contributing uniquely to improving segmentation accuracy and reliability. The following text emphasizes the use cases in Markov random fields (MRFs) [77] and K-means clustering [100,101].

#### 3.2.1. Markov Random Field

Probabilistic approaches employ statistical models to model the relationship between image features and tissue classes. These models assign a probability to each pixel belonging to a specific class [70]. This allows for a more robust segmentation compared to simpler thresholding techniques. Within the realm of probabilistic approaches for IVUS segmentation, MRFs offer a powerful technique for incorporating spatial context. It uses two types of potential functions: unary potentials that consider the observed data (like pixel intensities) to measure how likely a pixel belongs to a particular label, and pairwise potentials that capture spatial dependencies between neighboring pixels.

Optimization methods like graph cuts, belief propagation, or simulated annealing are employed to find the optimal labelling configuration that minimizes the overall potential energy. Further, MRFs were combined with the Rayleigh Mixture Model (RMM)

by Gao et al. [85]. The RMM is utilized for pixel classification based on their gray levels. By leveraging the spatial relationships among neighboring pixels, it effectively clusters pixels. MRFs are employed for the angular place of calcified plaques detection, utilizing prior information derived from RMM results through a curve known as the maximum intensity curve (MIC). Through the MRF, the relationship of observed and hidden variables is computed, enabling the identification of calcified plaques with acoustic shadowing. Araki et al. [77] quantify the volume of calcium deposits within coronary artery walls by employing three different methods for soft computing: fuzzy c-means (FCM), K-means clustering, and hidden Markov random fields (HMRFs). These techniques essentially group pixels in the image based on their characteristics. Nevertheless, designing MRFs and defining appropriate potential functions can be challenging, and the computational cost may be high. As the technology advanced, ML models such as the random forest classifier and SVM captured the industry.

### 3.2.2. Random Forest

Random forest is a ML algorithm that excels by combining the predictions of decision tree numbers. The trees are trained on random training subsets and features separately, which helps mitigate overfitting and enhances model robustness. This inherent randomness allows random forest to effectively capture non-linear relationships between features and segments. Moreover, random forest provides insights into feature importance, indicating which features contribute most significantly to the segmentation process. This feature analysis aids in understanding the underlying characteristics driving segment formation. It handles high-dimensional data well, automatically selecting relevant features for each tree, making it suitable for complex datasets and well-validated in CVD risk stratification using carotid plaque [102,103]. In 2017, Lo Vercio et al. [83] used random forests (RFs) to identify specific morphological structures within vessel walls seen in IVUS images. Later, in 2019, they built upon the concept of structure detection but took a more comprehensive approach by combining SVM with the RF model. First, they trained SVM to classify each pixel in the IVUS image [61]. Recognizing that the presence of structures like bifurcations or calcifications can challenge SVMs, the researchers then introduced RFs. These RFs are trained to identify such morphological structures within the image and refine boundaries. It was chosen based on its ability to deal with multiclass and imbalanced problems in association with randomness under sampling.

### 3.3. Deep Learning Techniques

Traditional methods often struggle with the complexity and variability of IVUS images, which are prone to noise, artefacts, and inconsistencies in tissue appearance due to factors like blood, calcifications, and imaging artefacts. However, the use of DL has overcome these challenges by effectively handling intricate patterns within the data [104]. DL algorithms can automatically learn to recognize and differentiate between subtle variations in tissue characteristics, reducing the impact of noise and artefacts, and improving the accuracy and consistency of image interpretation [105,106].

#### 3.3.1. Non-UNet Paradigms

The significance of CNNs in image segmentation lies in their ability to learn hierarchical representations of image features. They can automatically learn relevant features from the raw IVUS image during the training process by capturing low-level (edges, textures) and higher-level features that represent more complex patterns.

### Scale Mutualized Perception

Existing DL methods struggle with scale-dependent interference due to their top-down aggregation of multi-scale features. To address this, Liu et al. [69] proposed scale mutualized perception, which considers scales that are adjacent together for the conservation of the complimentary outputs. Semantic cues are offered by the small scales for tissue localization and help in perceiving context globally to enhance context in large scales for local representation, and vice versa. By these, similar objects with local features can be distinguished. Detailed information was brought by large scales for vessel boundary refining. The architecture consists of three main components: (i) neighboring-scale interactive learning (NIL), (ii) scale-mutually context complement (SCC), and (iii) densely-connected atrous convolution (DAC). NIL refines blurriness by incorporating complementary features from adjacent scales, avoiding noise interference. SCC distinguishes objects with similar context appearances by transferring context between adjacent scales. It extracts global context from smaller adjacent scales and local context from larger adjacent scales using a non-local method. DAC employs a densely connected structure to reduce the loss of detailed information on vessel borders varying in size, large-scale to small-scale features merging, and enlarging the receptive field.

### Combining Shallow and Deep Networks

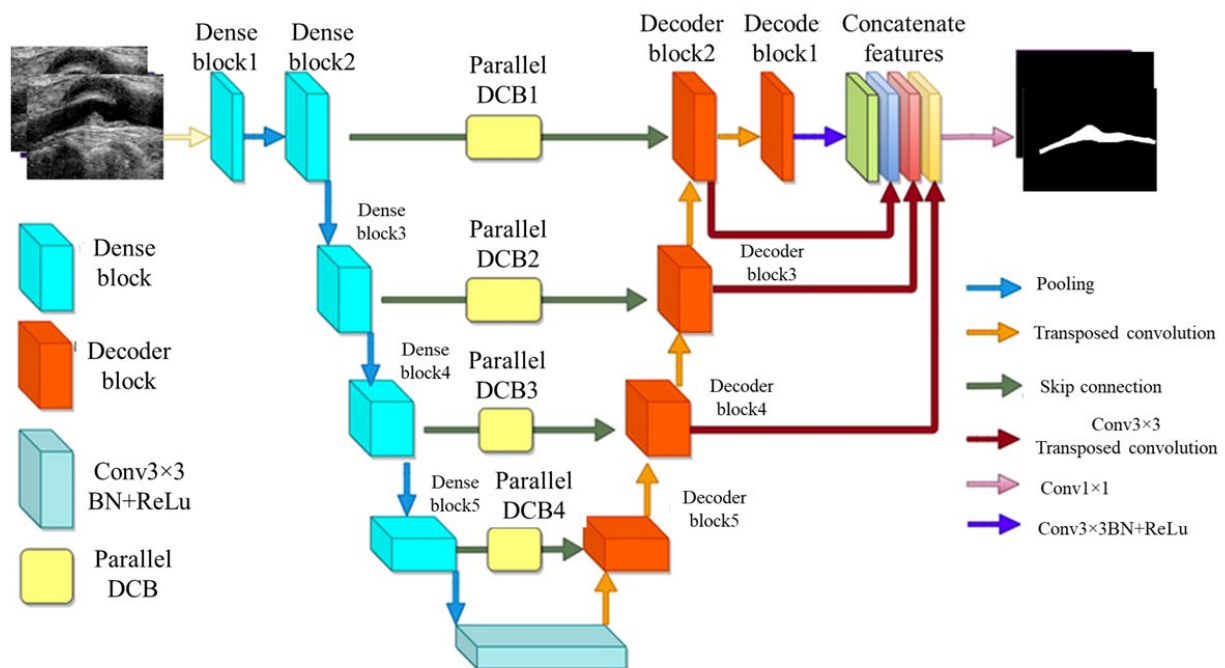
Yuan et al. [73] introduced a two-facet structure called combining shallow and deep networks (CSDN) for better and more accurate segmentation. This framework targets low-level detail extraction with the help of thick channels, and the shallow network is responsible for learning high-level semantics. The framework is designed with five stages for extracting context and semantic information features, providing a large receptive field. The first step is the Stem block, which has two parts for different types of input downsampling. Following batch normalization and the PReLU activation function, the outputs are concatenated and convolution is performed. The Context Block, the last step, provides the maximal receptive field by embedding global contextual information using global average pooling and residual connections. Gather-Expansion Blocks make up the final three phases. The model provides great accuracy and efficiency for real-time segmentation by processing these data independently. A mutually guided fusion module is used to improve and fuse both kinds of feature representations in order to further improve segmentation performance.

#### 3.3.2. UNet Paradigms and Its Variants

UNet is a convolutional neural network architecture introduced in 2015 by Olaf Ronneberger [107] for biomedical image segmentation tasks. It has become the de facto standard for many image segmentation [108] problems due to its ability to capture precise spatial information and efficiency. It comprises an encoder (contract path) that captures context and spatial information through repeated convolutions and max pooling, a decoder (expansion path) that enables precise localization by combining upsampled output with high-resolution features and skip connections providing a seamless fusion of high-resolution features from the encoder with corresponding decoder layers, reserving spatial information. UNet is preferred nowadays as a base architecture because of its ability to effectively retain and combine low-level and high-level features through its skip connections, allowing for precise pixel-wise segmentation even with limited training data [109,110]. Numerous variants and specialized versions have emerged which fuse transfer learning-based UNet [111,112] to further optimize its performance and cater to specific issues [90].

### MFA-UNet

A variant of UNet, known as *multi-scale feature aggregated UNet* (MFA-UNet), integrates a feature aggregation module (FAM) incorporating convolutional bidirectional long short-term memory (BConvLSTM) units. With multi-scale inputs and careful supervision, this feature improves both the encoding and decoding stages by enabling context retrieval from spatial–temporal perspectives. The MFA-UNet’s whole architecture is seen in Figure 9. The FAM component was employed from BCDUNet [113], shown in Figure 9, to improve feature fusion and information retention. The factors, repeated from a relevant layer of encoders (skip connection), include local information of high resolution. In contrast, the factors received by the previous up-convolution layer provide the global semantic information. Information that was lost during cascaded encoding processes is restored by concatenation. ConvLSTM-type recurrent neural networks (RNNs) are capable of handling complicated item distributions, capturing spatiotemporal correlations of sequential input, and remembering previously learned information. Two ConvLSTMs are used to process both forward and backward input. The benefit of this network is that it includes multiscale inputs, the FAM module, and deep supervision of the UNet model to obtain adequate learning with a limited set of detailed IVUS scans. The Focal Tversky loss is used to optimize the MFA-UNet which addresses the issue of data imbalance [114].



**Figure 9.** MFA-UNet’s architecture [115].

### IVUS-UNet++

The UNet++ model’s advanced system is called the IVUS-UNet++ model [116]. In UNet++, the encoder and decoder sub-networks are interconnected through a series of nested, dense convolutional blocks. The redesigned skip pathways aim to minimize the semantic gap between the feature maps of the encoder and decoder sub-networks. UNet++ effectively captures fine-grain details of foreground objects by gradually enriching high-resolution feature maps from the encoder network before fusing them with the corresponding semantically rich feature maps from the decoder network. Experimental results show that UNet++ with deep supervision achieves an average IoU (Intersection over Union score) gain of 3.9 points over the classical UNet [117].

To further enhance the model, a pyramid of factors were integrated into UNet++, resulting in IVUS-UNet++. This addition enables the feature maps utility at different scales, effectively fusing and propagating feature information across various spatial scales throughout the network. Five types of feature maps supervise the feature maps present over the convolutional block (0.5), as illustrated in Figure 10. Upscale operators ensure size compatibility during the fusion of multi-scale features. A voting mechanism with parallel connection is used to generate the final probability map.

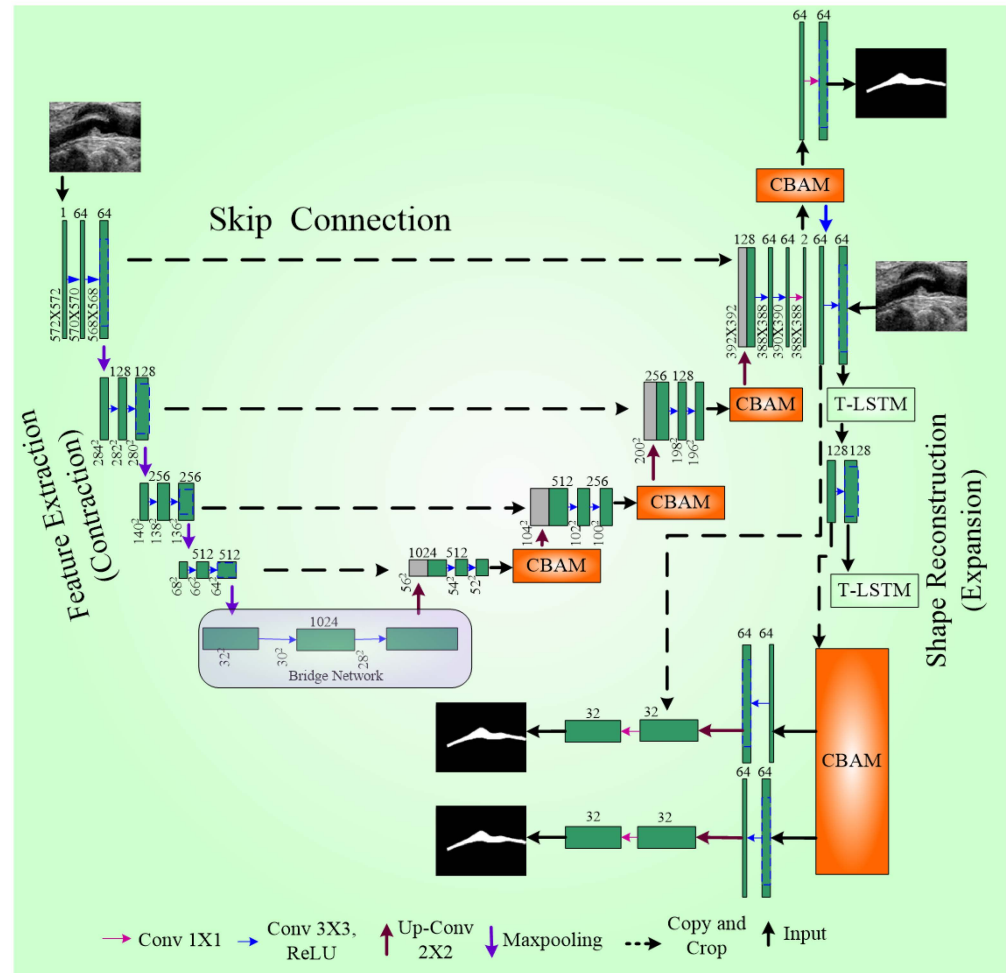


Figure 10. UNet++ architecture [118].

Several optimization strategies were employed to improve boundary detection. First, pre-trained weights were used as the backbone to mitigate the gradient vanishing problem in this deep network. Second, batch normalization (BN) was applied, and third, an activation function called ReLU was used consecutively with the convolution layer for network training improvement. Compared to other models, namely UNet++ [116] and IVUS-Net [79], IVUS-UNet++ achieved the best JM (Jaccard index) and HD (Hausdorff distance) for both the lumen and MA border. A schematic diagram of UNet++ is shown in Figure 10.

### 3.4. Attention and Transformer Methods

The NLP tasks were initially performed through a self-attention mechanism as suggested by Vaswani et al. [119], and the transformer (one form) played a major role in the field, being able to capture the dependencies of long-term. Researchers started integrating attention [27,48,120–122] and transformers [17,46,123–131] into the medical field to enhance

performance. Previously, image segmentation encountered numerous hurdles such as the inability to focus on relevant regions, limited contextual understanding, and challenges in dealing with objects of variable sizes. These shortcomings prompted a quest for innovative solutions, with attention mechanisms emerging as a promising avenue. The transformer model [119] is comprised of multiple encoder–decoder components, with each encoder having two parts: a feed-forward neural network and a self-attention layer. Each sublayer is followed by layer normalization and has a residual connection around it. Patch extraction breaks down the image into separate non-overlapping patches, flattened into a 1D vector. Each patch is then transformed into an embedding vector, potentially capturing features extracted by CNNs. Finally, in positional encoding, spatial information, both absolute and relative, is preserved to maintain the image’s spatial context. These vectors first flow through a self-attention layer, a layer that helps the encoder look at other words in the input sentence as it encodes a specific word. There, it computes the attention-infused representation:

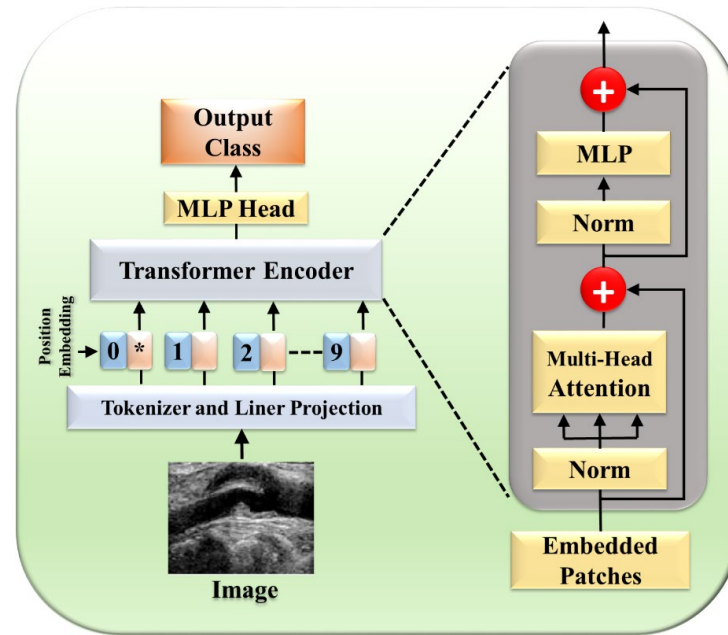
$$\text{Attention}(Q, K, V) = \text{softmax}\left(\frac{QK^T}{\sqrt{d_k}}\right)V$$

A feed-forward neural network receives the self-attention layer’s outputs. An extra attention layer in the decoder aids in concentrating on pertinent input segments. Finally, a linear layer projects the decoder output into a larger logit vector, whose probabilities are determined by a softmax layer. The word with the highest probability becomes the output for that time step. Recently, there has been a fusion of transformers with deep-learning architectures [50], notably combining CNNs and transformers, as exemplified by Tao et al. [51] This integration capitalizes on the complementary strengths of each; transformers capture remote dependencies while CNNs excel at local information extraction. It introduced transformer encoding blocks to model features extracted by CNNs at various scales, thereby enhancing both remote dependencies and local contexts [132]. Furthermore, they proposed a channel enhancement module employing a gated mechanism to enhance feature channels with more pertinent information.

#### 3.4.1. Perceptual Organisation-Aware Selective Transformer Framework

Sangala et al. [41] introduced a novel *perceptual organisation-aware selective transformer* (POST) framework designed to accurately segment vessel walls in IVUS images by mimicking the methods of cardiologists. Motion features with temporal context-based encoders are extracted, while a transformer module captures precise boundary information, guided by a boundary loss. The segmentation results are then made to combine a fusion module and temporal constraining (TF module) to obtain the correct boundaries. This framework has been integrated into software called QCU-CMS1 version 4.69 software (Leiden University Medical Center, Leiden, The Netherlands) for automatic IVUS image segmentation. The model predicts the boundaries that are missing by simulating the perceptual organization principles of human vision.

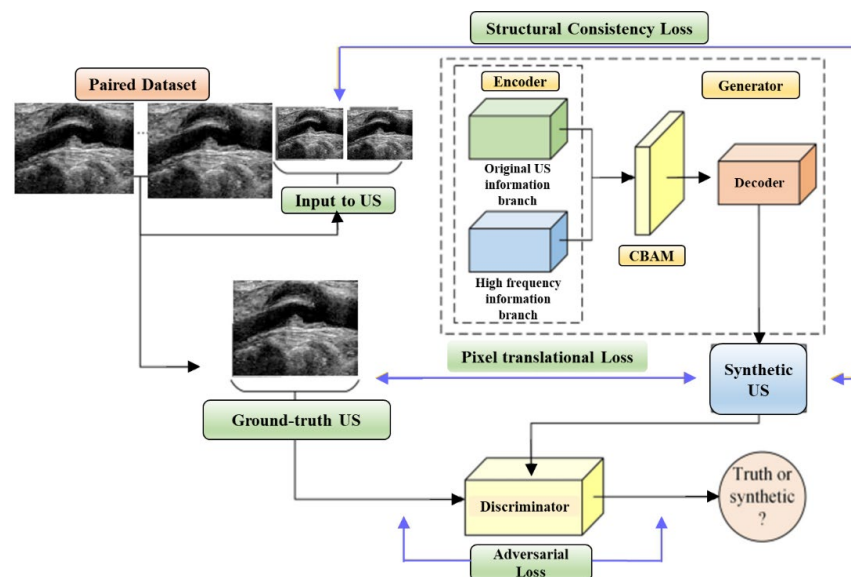
The framework consists of three major components (Figure 11): temporal context-based feature encoders, Selective Transformer Recurrent UNet (STR UNet) with a discriminator, and a TF Module. The temporal context-based feature encoders include encoders such as rotational alignment and visual persistence. The rotational motion of the vessel from the heartbeats was removed by the rotational alignment, while the residual visual factors in sequential order were captured by the visual persistence encoder. STR UNet can infer borders in dark regions by the use of generative adversarial learning (GAN) putting a specific loss function as a penalty during training. A discriminator is added to enhance the inference capability of STR UNet. The TF Module is a post-processing module that provides an improvement to the accuracy of annotation of side-branch and calcium areas.



**Figure 11.** Selective STR UNet. The star (\*) denotes the CLS token, which aggregates global feature information.

### 3.4.2. Multilevel Structure-Preserved Generative Adversarial Network (MSP-GAN)

The limited generalizability of IVUS analysis methods, due to the wide variety of IVUS datasets, is tackled through domain adaptation strategies. Variations in imaging devices, frequencies, and other factors lead to differences in grey distributions, textures, and more among IVUS datasets. Segmentation models designed for the target domain (TD) often fail to generalize to the source domain (SD). Previous approaches required designing a new technique for each SD (re-training DL networks), which is labor-intensive. To address this, domain adaptation is introduced, aiming to translate SD images into realistic TD images (Figure 12). However, current domain adaptation models struggle with preserving intravascular structures due to the complex pathology and low contrast in IVUS images. The challenges of structural preservation in IVUS domain adaptation include (i) the diverse morphologies of vessels, tissues, and artifacts in IVUS image datasets, which can cause the translation process to lose its original status or generate unknown elements.



**Figure 12.** Multilevel Structure-Preserved Generative Adversarial Network (MSP-GAN) architecture.

Therefore, preserving global pathological information is challenging. (ii) Because intravascular structures have limited differentiability and hazy boundaries, it is challenging to guarantee local anatomical consistency between the translated and original pictures. (iii) The absence of corresponding real TD images as reference labels means there is no pixel-wise constraint to prevent distortions of fine structures. Xia et al. [52] introduced MSP-GAN for transferring IVUS domains while preserving intravascular structures. Building upon the generator–discriminator baseline, MSP-GAN incorporates the transformer. Domain adaptation models primarily use generative adversarial networks (GANs), which transfer styles through a loss of adversarial type and preserve content through a content constraint. MSP-GAN enables data-specific IVUS analysis methods to be generalizable across different IVUS datasets. Images adapted using MSP-GAN maintain high anatomical consistency with source images. MSP-GAN employs three innovative techniques for multilevel structural preservation. Table 1 provides a comparative analysis of various IVUS architectures, detailing the advantages, disadvantages, performance metrics, and specific applications of each technique.

**Table 1.** Comparative analysis of various IVUS architectures.

Technique	Advantages	Disadvantages	Performance	Specific Application
Conventional Techniques				
Thresholding	Effective in removing catheter artifacts, sequential border estimation, ad hoc mechanism for discontinuous borders	May not perform well with complex and variable IVUS images	Jaccard measure: $0.84 \pm 0.07$ for Lumen, $0.82 \pm 0.11$ for MA border	Lumen-intima and media-adventitia (MA) borders segmentation
Active Contours	Fully automated, fast, and adaptive to the shape of the object	Sensitive to initialization, may get stuck in local minima	96% reduction in analysis time compared to manual segmentation	Lumen and MA boundary segmentation
Machine Learning Techniques				
Markov Random Field	Incorporates spatial context, robust segmentation	Designing MRF and defining appropriate potential functions can be challenging, high computational cost	Not specified	Calcified plaque detection
Random Forest	Captures non-linear relationships, handles high-dimensional data well, provides feature importance insights	May overfit with noisy data, less interpretable than simpler models	Not specified	Identifying specific morphological structures within vessel walls
Deep Learning Techniques				
Non-UNet—Scale Mutualized Perception	Preserves complementary information from adjacent scales, distinguishes objects with similar local features	Complex architecture, may require large amounts of data for training	Not specified	Vessel boundary segmentation
Non-UNet—CSDN	Efficient segmentation, treats shallow and deep networks separately for high accuracy and efficiency	Complex architecture, may require large amounts of data for training	Not specified	Real-time segmentation

Table 1. Cont.

Technique	Advantages	Disadvantages	Performance	Specific Application
UNet and its variants (MFA-UNet)	Improves feature fusion and information retention, enables context retrieval from spatial-temporal perspectives	Complex architecture, may require large amounts of data for training	Optimized using Focal Tversky loss to address data imbalance	IVUS scan segmentation
UNet and its variants IVUS-UNet++	More effectively captures fine-grained details of the foreground objects, uses feature pyramid network for multi-scale feature utilization	Complex architecture, may require large amounts of data for training	Best JM and HD for both lumen and MA border compared to UNet++ and IVUS-Net	Lumen and MA border segmentation
Attention and Transformer-based Methods				
POST-IVUS	Accurate segmentation of vessel walls in IVUS images, mimics cardiologists' perceptual organization principle	Complex architecture, may require large amounts of data for training	Integrated into QCU-CMS1 software for automatic IVUS image segmentation	IVUS image segmentation
MSP-GAN	Preserves intravascular structures during domain adaptation, uses transformers for global pathology information preservation	Complex architecture, may require large amounts of data for training	Ensures local structures correspondence between source and translated images	IVUS domain adaptation

## 4. Critical Discussion

### 4.1. Principal Findings

IVUS imaging provides essential cross-sectional views of coronary artery walls for an accurate CVD diagnosis. Manual segmentation is time consuming, labor intensive, and prone to errors. This has spurred significant research into automated segmentation methods, categorized into four main approaches. (i) Traditional methods like thresholding, active contours, and graph-based techniques rely on mathematical models to identify tissue boundaries but often struggle with IVUS image complexity and variability [36]. For example, the Kermani et al. [60] approach effectively detects borders accurately, while the Giannoglou et al. [56] model offers automated segmentation. (ii) ML advancements provide robust and adaptable solutions. Methods like MRF incorporate spatial context, while RF and SVMs use statistical models and ensemble learning to improve accuracy and handle variability in IVUS images. (iii) DL, particularly CNNs, has revolutionized IVUS segmentation. CNNs, such as UNet and its variants, enhance performance by learning hierarchical features from raw data. Advanced architectures like MFA-UNet and IVUS-UNet++ improve accuracy and efficiency by incorporating multi-scale feature aggregation. (iv) Borrowed from natural language processing, attention mechanisms and transformers capture long-term dependencies and contextual information, crucial for IVUS segmentation. Models like the POST [41] and MSP-GAN [52] demonstrate significant improvements in segmentation accuracy by incorporating temporal context, generative adversarial learning, and domain adaptation strategies.

#### 4.2. Benchmarking

In the realm of IVUS imaging, various segmentation methods have been developed to enhance the accuracy and efficiency of diagnosing coronary artery disease. Traditional methods, such as conventional techniques proposed by Hammouche et al. [58], utilize a 3D segmentation algorithm with a helical active contour to reduce the effect of ring-down artifacts. Their approach demonstrated high precision with a lumen detection accuracy of 99.42% and a minimal mean absolute error of 0.272 mm as shown in Table 2. Similarly, Giannoglou et al. [56] introduced an active contour model initialized through morphological characteristic analysis. This approach primarily focused on segmenting the lumen-wall and media-adventitia borders, demonstrating a 96% reduction in analysis time compared to manual segmentation. While computationally efficient, this method still depends on well-defined edges, which can be challenging in cases with plaque deposits or noisy IVUS images.

**Table 2.** Benchmarking table. ✓ indicates the presence or use of the specified feature, while ✗ denotes its absence.

C0	C1	C2	C3	C4	C5	C6	C7	C8	C9
SN	Authors	NOF	Type of Data	Architecture Used	Attention	Transformer	#Patients/ #Images	CV	Results
R1	Hammouche et al. [58] (2019)	10	Image	Helical active contour	✗	✗	144/510 497/638	K5	Lumen detection accuracy of 99.42% and a minimal mean absolute error of 0.272 mm.
R2	Giannoglou et al. [56] (2006)	14	Image	Active contour model	✗	✗	97/970	K5	The study demonstrated a 96% reduction in analysis time compared to manual segmentation.
R3	Wang et al. [96] (2021)	17	Image	ML	✗	✗	379, 300	K10	Enhanced traditional snake algorithm, Otsu thresholding, morphological operations, and connected component labeling were incorporated
R4	Vercio et al. [61] (2019)	14	Image	SVM with RF	✓	✗	800	-	High Dice similarity coefficient (DSC) of 0.91 for LI and 0.94 for MA.
R5	Liu et al. [69] (2022)	29	Image	SMP	✓	✓	378	-	DSC of 0.96 for the lumen and 0.97 for the MA
R6	Bargsten et al. [78] (2021)	20	Image	Capsule Network	✓	✓	-	-	Accuracy of 94.59% in lumen segmentation.
R7	Proposed Study	39	Point	Transformers and Attention	✓	✓	500	K5	

Advancing beyond traditional methods, ML-based techniques have incorporated data-driven strategies for improved segmentation performance. Wang et al. [96] refined the traditional snake algorithm by integrating an extended structure tensor for automatic contour extraction. Their method enhanced edge detection accuracy by employing Otsu thresholding, morphological operations, and connected component labeling. Despite these improvements, the method remains dependent on feature engineering, limiting its adaptability to highly complex IVUS datasets. Another significant contribution in ML-based segmentation came from Vercio et al. [61], who combined support vector machines (SVM) with random forests (RF) to detect morphological structures, followed by a deformable contour model to refine the segmentation of the lumen-intima (LI) and media-adventitia (MA) interfaces. This approach achieved a high Dice similarity coefficient (DSC) of 0.91 for LI and 0.94 for MA, showcasing its effectiveness in distinguishing arterial layers. How-

ever, its reliance on feature selection can lead to suboptimal generalization across diverse IVUS datasets.

With the advent of DL, particularly convolutional neural networks CNNs and transformer-based architectures, segmentation techniques have achieved higher accuracy and robustness. Liu et al. [69] introduced a scale-aware multi-path processing (SMP) method, which effectively preserves complementary information across adjacent scales. Their model achieved a DSC of 0.96 for lumen and 0.97 for MA, significantly improving segmentation consistency. The main strength of this approach lies in its ability to address scale-dependent interference, but it requires extensive computational resources, limiting its feasibility in real-time clinical applications. Similarly, Bargsten et al. [78] employed capsule networks, which leverage spatial relationships between image features to enhance performance on small datasets. This method achieved an accuracy of 94.59% in lumen segmentation, outperforming conventional CNNs in terms of spatial awareness. However, capsule networks are computationally intensive, making them challenging to deploy in resource-constrained settings.

The proposed study extends the frontier of IVUS segmentation by employing a transformer- and attention-based architecture. This method capitalizes on the self-attention mechanism, allowing it to capture global contextual dependencies in IVUS images. Unlike CNN-based models, which primarily focus on local features, transformers can effectively model long-range spatial relationships, leading to enhanced segmentation accuracy and robustness. The study was conducted on 39 patients with 500 IVUS scans, demonstrating a promising balance between computational efficiency and segmentation accuracy. However, despite its advantages, transformer models typically require large amounts of training data, and fine-tuning remains essential for domain-specific applications.

Overall, the progression from conventional segmentation techniques to ML- and DL-based approaches reflects a significant leap towards more accurate, automated, and scalable IVUS analysis methods. Each method comes with distinct strengths and limitations, highlighting the continuous need for innovation in medical image segmentation to enhance diagnostic accuracy and clinical applicability in cardiovascular imaging.

#### *4.3. A Special Note on Transformers in Coronary Artery Segmentation*

Transformers have emerged as a powerful tool in coronary artery segmentation, particularly in the analysis of IVUS imaging [133]. Their key innovation, the self-attention mechanism, allows transformers to capture long-range dependencies and relationships between different parts of an image, making them highly effective in handling complex medical imaging tasks. In the context of coronary artery segmentation, this capability is essential for accurately identifying and segmenting intricate structures such as arterial walls, plaques, and other cardiovascular features that are critical for diagnosing and assessing cardiovascular diseases (CVD) [134,135]. One of the standout advantages of transformers is their ability to model both local and global image features simultaneously [109]. This is particularly important in IVUS imaging, where arterial structures can be difficult to distinguish due to noise, variability in image quality, and overlapping features. By utilizing self-attention mechanisms, transformers can focus on the most relevant regions of an image, while still capturing the broader context, which is vital for distinguishing between structures that may appear similar at the local level but have different overall characteristics. This capability enhances the precision of segmentation, particularly in detecting subtle anatomical details critical for treatment planning. Transformers also integrate well with CNNs in hybrid models, combining the strengths of both architectures [136–138]. CNNs are highly effective at extracting detailed local features, while transformers provide a more comprehensive view of the image, handling complex spatial relationships and

dependencies. This hybrid approach has led to significant improvements in segmentation accuracy, particularly for challenging cases such as stenosis or calcified regions within the arteries, where local and global information must be processed together to achieve reliable results [139–141]. Another important application of transformers in IVUS segmentation is their role in domain adaptation, where they help address the variability in datasets from different imaging sources or devices. Transformers, when combined with generative adversarial networks (GANs) [142–146] or other domain adaptation techniques, help preserve the critical structural information of the coronary arteries, ensuring that segmentation models perform well across different data sources. This ability to generalize across datasets is crucial for real-world clinical applications, where variability in imaging quality and protocols is common [141].

In summary, transformers have made a significant impact on the field of coronary artery segmentation, particularly through their ability to model complex, multi-scale structures, and their adaptability in hybrid architectures. Their contributions to improving segmentation accuracy, robustness, and generalizability position them as a key technology in the future of IVUS imaging and cardiovascular disease diagnosis. While it is technically sound, it is however important to see how the economics of AI using attention and transformers will help in the diagnosis and treatment for CAD [147].

#### *4.4. A Special Note on Clinical Relevance and Integration of Transformer-Based IVUS Segmentation into Clinical Practice*

The integration of transformer-based segmentation models into IVUS analysis has the potential to revolutionize cardiovascular diagnostics by enhancing precision, reproducibility, and automation in CAD assessments [148]. Unlike traditional segmentation methods, which are often dependent on manual annotation and prone to inter-observer variability, transformer models consistently delineate LI and MA boundaries with higher accuracy, leading to improved measurements of arterial wall thickness and plaque burden. This level of precision is crucial for early detection of atherosclerosis, facilitating timely intervention before significant stenosis develops. Additionally, by minimizing segmentation errors and refining plaque quantification, transformer-based models contribute to more reliable risk stratification in CAD patients, aiding clinicians in developing personalized treatment strategies. The ability of these models to standardize IVUS interpretation across different clinical settings ensures a more objective and data-driven approach to disease assessment, reducing diagnostic variability and improving patient outcomes.

From a clinical workflow perspective, the adoption of transformer-based IVUS segmentation can streamline decision-making in interventional cardiology, particularly in guiding procedures such as angioplasty and stent placement [118]. By providing automated and real-time segmentation outputs, these models enable faster and more informed procedural planning, reducing dependency on manual delineation and enhancing overall efficiency in catheterization labs. Furthermore, their integration with multi-modal imaging (e.g., optical coherence tomography and coronary computed tomography angiography) can offer a more comprehensive visualization of arterial pathology, improving diagnostic confidence. Additionally, incorporating explainability mechanisms [109], can enhance clinical trust and transparency, ensuring that AI-driven segmentation aligns with expert interpretations. However, despite their potential, challenges remain for real-world deployment, particularly in resource-limited settings where computational infrastructure and access to large, annotated datasets for model training may be constrained. Addressing these barriers through model optimization, cloud-based implementations, and domain adaptation techniques will be critical in ensuring equitable access to AI-driven IVUS segmentation, ultimately advancing precision medicine in cardiovascular care.

#### 4.5. Strengths: Technological Advancements and Integration

The presented manuscript highlights the importance of accurate coronary artery wall segmentation in IVUS imaging for diagnosing CVD. It points out the challenges of manual segmentation, including its time-consuming nature and potential for errors. The overview of automated methods, from traditional image processing to advanced ML and DL techniques, offers a comprehensive look at the current state of IVUS segmentation research. Categorizing methods into conventional, ML, DL, and attention/transformer methods provides a clear structure [12]. The detailed explanations of techniques like thresholding, active contours, MRF, and RF are insightful. Recent advancements in DL and transformer-based methods showcase improvements in accuracy and efficiency [47,109,149]. The dense technical language and numerous citations might be challenging for non-experts. The section could benefit from a more focused discussion on the clinical implications and real-world impact of these techniques. A deeper exploration of the limitations and trade-offs of conventional and deep learning methods would provide a more balanced view.

The incorporation of real-world case studies in clinical settings highlights the practical benefits of automated IVUS segmentation techniques, showcasing their potential to improve diagnostic accuracy and patient outcomes. Integrating multi-modal imaging data [150–153], such as combining IVUS with modalities like optical coherence tomography (OCT) or magnetic resonance imaging (MRI), can enhance segmentation accuracy by leveraging complementary information from different sources. Such approaches demonstrate how diverse data streams can work together to overcome limitations inherent to single-modal imaging.

Hybrid methods that combine traditional image processing, machine learning, and deep learning offer a holistic approach to advancing IVUS segmentation. These methods enable precise delineation of complex anatomical structures, such as vessel walls, plaques, and lumen boundaries, by harnessing the strengths of each computational paradigm. By integrating domain knowledge and algorithmic sophistication, these approaches pave the way for robust and interpretable segmentation systems.

Transformer-based architectures, such as Swin UNet and TransUNet, represent a promising avenue for the future of IVUS segmentation. These models integrate global self-attention mechanisms with hierarchical and local feature extraction capabilities, providing the potential to effectively process high-resolution IVUS images. Swin UNet's shifted window-based self-attention mechanism is well-suited for capturing multi-scale contextual information, while TransUNet combines CNNs for low-level feature extraction with transformers for global dependency modeling. This hybrid architecture may address challenges like artifact interference and enable precise boundary delineation. When used with complementary imaging modalities like OCT or MRI, these models could significantly enhance the robustness and accuracy of IVUS segmentation.

#### 4.6. Weaknesses: Limitations in Data, Generalization, and Practicality

While deep learning and transformer-based architectures have demonstrated substantial improvements in IVUS segmentation [154,155], several challenges remain that hinder their widespread clinical adoption. One of the most significant limitations is the dependency on large, high-quality, annotated datasets, which are scarce in IVUS imaging due to the time-intensive nature of expert annotation. Transformer-based models, in particular, require extensive data to capture spatial and contextual relationships effectively, yet the availability of publicly accessible IVUS datasets remains limited [156]. This data scarcity contributes to issues such as overfitting and poor generalization, particularly when models are trained on small and homogeneous datasets. Moreover, the lack of standardized bench-

mark datasets restricts the ability to fairly compare different segmentation approaches, making it difficult to establish state-of-the-art performance in this domain.

Another critical challenge is the variability across different IVUS scanners and imaging protocols, which leads to inconsistencies in image contrast, resolution, and noise characteristics [157]. Since IVUS devices from different manufacturers employ distinct acquisition settings, deep learning models trained on one dataset often fail to generalize effectively to data from other scanners. Transformer-based architectures, while powerful in capturing long-range dependencies, are highly sensitive to shifts in image distribution, making domain adaptation a crucial but unresolved issue. Although methods such as self-supervised learning and transfer learning offer potential solutions, existing approaches do not fully mitigate the impact of scanner-dependent variations, limiting model applicability in diverse clinical settings.

Furthermore, computational complexity and model interpretability remain major obstacles in deploying transformer-based models for IVUS segmentation in real-time clinical workflows. Compared to CNNs, transformers require significantly higher computational resources, making them impractical for use in resource-constrained environments such as point-of-care applications. Additionally, the lack of transparency in how transformer models process and segment IVUS images poses concerns regarding their reliability in clinical decision-making. Without explainable AI (XAI) mechanisms, it is challenging for clinicians to validate and trust the segmentation outputs, particularly in high-stakes applications like coronary artery disease diagnosis and treatment planning.

#### *4.7. Extensions: Future Research and Clinical Implementation*

Future research should focus on addressing these limitations to develop clinically viable, efficient, and interpretable transformer-based IVUS segmentation models. A key priority is clinical validation through multi-center studies to evaluate model robustness across different patient populations and imaging conditions. Ensuring that AI-driven segmentation methods perform consistently across datasets acquired from various scanners will be essential for regulatory approval and real-world deployment. Additionally, self-supervised learning, domain adaptation, and federated learning should be explored to improve model generalization while minimizing biases introduced by small or imbalanced datasets. These approaches can enable models to adapt dynamically to new imaging environments, reducing performance discrepancies between datasets.

To enhance computational efficiency, lightweight transformer architectures tailored for IVUS segmentation should be developed. Methods such as MobileViT, Swin UNet, and hybrid CNN–transformer models could help reduce computational overhead while maintaining high segmentation accuracy. Furthermore, integrating multi-modal imaging data, including IVUS, OCT, and MRI, can provide complementary anatomical and functional insights, further improving segmentation precision. Another crucial avenue for future research is the incorporation of XAI techniques, such as attention heatmaps, saliency maps, and uncertainty quantification, to enhance the interpretability of transformer-based models. By making segmentation decisions more transparent and explainable, these techniques can help bridge the gap between deep learning advancements and clinical adoption.

Finally, the development of public benchmark datasets and standardized evaluation protocols is imperative to facilitate the reproducibility and fair comparison of segmentation models. Establishing comprehensive datasets with diverse patient populations and scanner types will drive progress in the field by enabling robust performance evaluations. Future work should also focus on defining clinically relevant metrics that go beyond traditional performance measures like Dice similarity and the Jaccard index, incorporating real-world usability factors such as segmentation speed, model confidence, and clinical utility. By

addressing these research directions, transformer-based architectures can be refined into reliable, efficient, and interpretable solutions, ultimately advancing the role of AI in coronary artery disease diagnosis and treatment planning.

## 5. Conclusions

In conclusion, IVUS imaging is crucial for accurate CVD diagnosis, but manual segmentation is time consuming and error prone. Automated segmentation methods have been developed to improve efficiency and accuracy, with traditional techniques providing the foundation. ML advancements, such as MRF, RF, and SVMs, offer robust solutions by incorporating spatial context and ensemble learning. Deep learning, particularly CNNs like UNet and its variants, has revolutionized IVUS segmentation by learning hierarchical features from raw data.

Advanced architectures like MFAUNet and IVUS-UNet++ further enhance accuracy and efficiency. Attention mechanisms and transformers have shown promise in capturing long-term dependencies and contextual information. Models like POST and MSP-GAN demonstrate significant improvements by incorporating temporal context and generative adversarial learning. The evolution of automated segmentation methods has significantly improved IVUS imaging in CVD diagnosis, and future research should continue to refine these methods for clinical applicability and effectiveness.

**Author Contributions:** Conceptualization, V.K., M.B., M.M. and J.S.S.; Methodology, V.K., T.C., M.B. and M.M.; Software, V.K., M.B. and M.M.; Validation, A.K., S.P., T.C., A.B., E.T., V.R., S.G., M.A.-M., V.A., L.S. and J.S.S.; Formal analysis, S.G., M.B. and L.S.; Investigation, S.P., I.M.S., T.C., A.B., E.T., V.R., S.G., M.A.-M., V.A., L.S. and J.S.S.; Resources, S.P. and L.S.; Data curation, V.K., M.B. and M.M.; Writing—original draft, V.K., M.B., T.C., M.M. and J.S.S.; Writing—review and editing, V.K., M.B., M.M., S.G., E.T. and J.S.S.; Visualization, S.P., A.B., E.T., V.R., M.A.-M., V.A., L.S. and J.S.S.; Supervision, A.K., S.P. and J.S.S. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research received no external funding.

**Institutional Review Board Statement:** There are no Institutional Review Board issues applicable to this study.

**Informed Consent Statement:** Not applicable in this study.

**Data Availability Statement:** Data are not available due to the proprietary nature of this study.

**Acknowledgments:** The authors would like to acknowledge AtheroPoint™, Roseville, CA 95661, USA for the medical data and the codes that are proprietary to AtheroPoint™.

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

## Table for Acronyms (SN: Serial Number; Abbn: Abbreviations)

SN	Abbn	Description	SN	Abbn	Description
1	ACS	Acute Coronary Syndrome	43	MCC	Matthews Correlation Coefficient
2	CAD	Coronary Artery Disease	44	AUC	Area Under the Curve
3	CVD	Cardiovascular Disease	45	ROC	Receiver Operating Characteristic
4	CT	Computed Tomography	46	PR	Precision-Recall
5	DL	Deep Learning	47	FPS	Frames Per Second
6	IVUS	Intravascular Ultrasound	48	GPU	Graphics Processing Unit
7	LI	Lumen-Intima	49	CPU	Central Processing Unit
8	MA	Media-Adventitia	50	mAP	Mean Average Precision

9	MRI	Magnetic Resonance Imaging	51	PA	Pixel Accuracy
10	OCT	Optical Coherence Tomography	52	mIoU	Mean Intersection over Union
11	PRISMA	Preferred Reporting Items for Systematic Reviews and Meta-Analyses	53	FWIoU	Frequency Weighted Intersection over Union
12	SDL	Solo Deep Learning	54	TP	True Positive
13	HDL	Hybrid Deep Learning	55	TN	True Negative
14	CNN	Convolutional Neural Network	56	FP	False Positive
15	CRF	Conditional Random Field	57	FN	False Negative
16	FAM	Feature Aggregation Module	58	PPV	Positive Predictive Value
17	BConvLSTM	Bi-directional Convolutional Long Short-Term Memory	59	NPV	Negative Predictive Value
18	MFAUNet	Multi-scale Feature Aggregated UNet	60	FDR	False Discovery Rate
19	UNet++	UNet Plus Plus	61	FOR	False Omission Rate
20	IoU	Intersection over Union	62	MSE	Mean Squared Error
21	JM	Jaccard Measure	63	RMSE	Root Mean Squared Error
22	HD	Hausdorff Distance	64	MAE	Mean Absolute Error
23	GAN	Generative Adversarial Network	65	RAE	Relative Absolute Error
24	MSP-GAN	Multilevel Structure-Preserved Generative Adversarial Network	66	RSE	Relative Squared Error
25	SMC	Super pixel-wise Multi-scale Contrastive Constraint	67	R2	Coefficient of Determination
26	TF	Temporal Constraining and Fusion	68	Adj. R2	Adjusted R-squared
27	STR UNet	Selective Transformer Recurrent UNet	69	AIC	Akaike Information Criterion
28	MIC	Maximum Intensity Curve	70	BIC	Bayesian Information Criterion
29	RMM	Rayleigh Mixture Model	71	LOO-CV	Leave-One-Out Cross-Validation
30	MRF	Markov Random Field	72	K-fold CV	K-fold Cross-Validation
31	FCM	Fuzzy C-Means	73	Lasso	Least Absolute Shrinkage and Selection Operator
32	HMRF	Hidden Markov Random Field	74	Ridge	Ridge Regression
33	SVM	Support Vector Machine	75	Elastic Net	Elastic Net Regression
34	RF	Random Forest	76	PCA	Principal Component Analysis
35	FODPSO	Fractional-order Darwinian Particle Swarm Optimization	77	t-SNE	t-Distributed Stochastic Neighbor Embedding
36	PAD	Percentage of Area Difference	78	UMAP	Uniform Manifold Approximation and Projection
37	HSD	Hausdorff Surface Distance	79	NMF	Non-negative Matrix Factorization
38	Dice	Dice Coefficient	80	ICA	Independent Component Analysis
39	Precision	Precision Score	81	SVD	Singular Value Decomposition
40	Recall	Recall Score	82	LLE	Locally Linear Embedding
41	F1	F1 Score	83	ISOMAP	Isometric Mapping
42	Kappa	Cohen's Kappa Score	84	MDS	Multidimensional Scaling

## References

1. Saba, L.; Maindarkar, M.; Khanna, N.N.; Johri, A.M.; Mantella, L.; Laird, J.R.; Paraskevas, K.I.; Ruzsa, Z.; Kalra, M.K.; Fernandes, J.F.E. A Pharmaceutical Paradigm for Cardiovascular Composite Risk Assessment Using Novel Radiogenomics Risk Predictors in Precision Explainable Artificial Intelligence Framework: Clinical Trial Tool. *Front. Biosci.-Landmark* **2023**, *28*, 248. [[CrossRef](#)] [[PubMed](#)]
2. Gaidai, O.; Xing, Y.; Balakrishna, R.; Sun, J.; Bai, X. Prediction of death rates for cardiovascular diseases and cancers. *Cancer Innov.* **2023**, *2*, 140–147. [[PubMed](#)]
3. Vähätalo, J.; Holmström, L.; Pakanen, L.; Kaikkonen, K.; Perkiömäki, J.; Huikuri, H.; Juntila, J. Coronary artery disease as the cause of sudden cardiac death among victims < 50 years of age. *Am. J. Cardiol.* **2021**, *147*, 33–38.
4. Saba, L.; Maindarkar, M.; Johri, A.M.; Mantella, L.; Laird, J.R.; Khanna, N.N.; Paraskevas, K.I.; Ruzsa, Z.; Kalra, M.K.; Fernandes, J.F.E. UltraAIGenomics: Artificial Intelligence-Based Cardiovascular Disease Risk Assessment by Fusion of Ultrasound-Based Radiomics and Genomics Features for Preventive, Personalized and Precision Medicine: A Narrative Review. *Rev. Cardiovasc. Med.* **2024**, *25*, 184. [[PubMed](#)]
5. Khanna, N.N.; Maindarkar, M.; Puvvula, A.; Paul, S.; Bhagawati, M.; Ahluwalia, P.; Ruzsa, Z.; Sharma, A.; Munjral, S.; Kolluri, R. Vascular implications of COVID-19: Role of radiological imaging, artificial intelligence, and tissue characterization: A special report. *J. Cardiovasc. Dev. Dis.* **2022**, *9*, 268. [[CrossRef](#)]
6. Kandaswamy, E.; Zuo, L. Recent advances in treatment of coronary artery disease: Role of science and technology. *Int. J. Mol. Sci.* **2018**, *19*, 424. [[CrossRef](#)]
7. Suri, J.S.; Kathuria, C.; Molinari, F. *Atherosclerosis Disease Management*; Springer Science & Business Media: Berlin/Heidelberg, Germany, 2010.
8. Gharleghi, R.; Chen, N.; Sowmya, A.; Beier, S. Towards automated coronary artery segmentation: A systematic review. *Comput. Methods Programs Biomed.* **2022**, *225*, 107015. [[CrossRef](#)]
9. Khanna, N.N.; Maindarkar, M.; Saxena, A.; Ahluwalia, P.; Paul, S.; Srivastava, S.K.; Cuadrado-Godia, E.; Sharma, A.; Omerzu, T.; Saba, L. Cardiovascular/stroke risk assessment in patients with erectile dysfunction—A role of carotid wall arterial imaging and plaque tissue characterization using artificial intelligence paradigm: A narrative review. *Diagnostics* **2022**, *12*, 1249. [[CrossRef](#)]
10. Khanna, N.N.; Singh, M.; Maindarkar, M.; Kumar, A.; Johri, A.M.; Mantella, L.; Laird, J.R.; Paraskevas, K.I.; Ruzsa, Z.; Singh, N. Polygenic Risk Score for Cardiovascular Diseases in Artificial Intelligence Paradigm: A Review. *J. Korean Med. Sci.* **2023**, *38*, e395.
11. Garcia-Garcia, H.M.; Gogas, B.D.; Serruys, P.W.; Bruining, N. IVUS-based imaging modalities for tissue characterization: Similarities and differences. *Int. J. Cardiovasc. Imaging* **2011**, *27*, 215–224.
12. Katouzian, A.; Angelini, E.D.; Carlier, S.G.; Suri, J.S.; Navab, N.; Laine, A.F. A state-of-the-art review on segmentation algorithms in intravascular ultrasound (IVUS) images. *IEEE Trans. Inf. Technol. Biomed.* **2012**, *16*, 823–834. [[CrossRef](#)] [[PubMed](#)]
13. Ma, T.; Yu, M.; Chen, Z.; Fei, C.; Shung, K.K.; Zhou, Q. Multi-frequency intravascular ultrasound (IVUS) imaging. *IEEE Trans. Ultrason. Ferroelectr. Freq. Control.* **2015**, *62*, 97–107. [[CrossRef](#)] [[PubMed](#)]
14. Boi, A.; Jamthikar, A.D.; Saba, L.; Gupta, D.; Sharma, A.; Loi, B.; Laird, J.R.; Khanna, N.N.; Suri, J.S. A survey on coronary atherosclerotic plaque tissue characterization in intravascular optical coherence tomography. *Curr. Atheroscler. Rep.* **2018**, *20*, 33. [[CrossRef](#)] [[PubMed](#)]
15. Sinclair, H.; Bourantas, C.; Bagnall, A.; Mintz, G.S.; Kunadian, V. OCT for the identification of vulnerable plaque in acute coronary syndrome. *JACC Cardiovasc. Imaging* **2015**, *8*, 198–209. [[CrossRef](#)]
16. Marano, R.; Rovere, G.; Savino, G.; Flammia, F.C.; Carafa, M.R.P.; Steri, L.; Merlino, B.; Natale, L. CCTA in the diagnosis of coronary artery disease. *La Radiol. Medica* **2020**, *125*, 1102–1113. [[CrossRef](#)]
17. Gao, Y.; Zhou, M.; Metaxas, D.N. Utnet: A hybrid transformer architecture for medical image segmentation. In *Medical Image Computing and Computer Assisted Intervention—MICCAI 2021, Proceedings of the 24th International Conference, Strasbourg, France, 27 September–1 October 2021, Proceedings, Part III 24*; Springer: Berlin/Heidelberg, Germany, 2021.
18. Al-Maini, M.; Maindarkar, M.; Kitas, G.D.; Khanna, N.N.; Misra, D.P.; Johri, A.M.; Mantella, L.; Agarwal, V.; Sharma, A.; Singh, I.M. Artificial intelligence-based preventive, personalized and precision medicine for cardiovascular disease/stroke risk assessment in rheumatoid arthritis patients: A narrative review. *Rheumatol. Int.* **2023**, *43*, 1965–1982. [[CrossRef](#)]
19. Suri, J.S.; Paul, S.; Maindarkar, M.A.; Puvvula, A.; Saxena, S.; Saba, L.; Turk, M.; Laird, J.R.; Khanna, N.N.; Viskovic, K. Cardiovascular/stroke risk stratification in Parkinson’s disease patients using atherosclerosis pathway and artificial intelligence paradigm: A systematic review. *Metabolites* **2022**, *12*, 312. [[CrossRef](#)]
20. Khanna, N.N.; Maindarkar, M.A.; Viswanathan, V.; Puvvula, A.; Paul, S.; Bhagawati, M.; Ahluwalia, P.; Ruzsa, Z.; Sharma, A.; Kolluri, R. Cardiovascular/Stroke Risk Stratification in Diabetic Foot Infection Patients Using Deep Learning-Based Artificial Intelligence: An Investigative Study. *J. Clin. Med.* **2022**, *11*, 6844. [[CrossRef](#)]
21. Bourantas, C.V.; Garg, S.; Naka, K.K.; Thury, A.; Hoye, A.; Michalis, L.K. Focus on the research utility of intravascular ultrasound-comparison with other invasive modalities. *Cardiovasc. Ultrasound* **2011**, *9*, 2. [[CrossRef](#)]

22. Araki, T.; Ikeda, N.; Shukla, D.; Londhe, N.D.; Shrivastava, V.K.; Banchhor, S.K.; Saba, L.; Nicolaides, A.; Shafique, S.; Laird, J.R. A new method for IVUS-based coronary artery disease risk stratification: A link between coronary & carotid ultrasound plaque burdens. *Comput. Methods Programs Biomed.* **2016**, *124*, 161–179.
23. Banchhor, S.K.; Araki, T.; Londhe, N.D.; Ikeda, N.; Radeva, P.; Elbaz, A.; Saba, L.; Nicolaides, A.; Shafique, S.; Laird, J.R. Five multiresolution-based calcium volume measurement techniques from coronary IVUS videos: A comparative approach. *J. Comput. Methods Programs Biomed.* **2016**, *134*, 237–258.
24. Sonoda, S.; Hibi, K.; Okura, H.; Fujii, K.; Node, K.; Kobayashi, Y.; Honye, J. Current clinical use of intravascular ultrasound imaging to guide percutaneous coronary interventions (update). *Cardiovasc. Interv. Ther.* **2023**, *38*, 1–7. [[PubMed](#)]
25. Biswas, M.; Kuppili, V.; Saba, L.; Edla, D.R.; Suri, H.S.; Sharma, A.; Cuadrado-Godia, E.; Laird, J.R.; Nicolaides, A.; Suri, J.S. A deep-learning fully convolutional network for lumen characterization in diabetic patients using carotid ultrasound: A tool for stroke risk. In *Vascular and Intravascular Imaging Trends, Analysis, and Challenges, Volume 2: Plaque Characterization*; IOP Publishing: Bristol, UK, 2019; pp. 5-1–5-33.
26. Jain, P.K.; Sharma, N.; Kalra, M.K.; Johri, A.; Saba, L.; Suri, J.S. Far wall plaque segmentation and area measurement in common and internal carotid artery ultrasound using U-series architectures: An unseen Artificial Intelligence paradigm for stroke risk assessment. *Comput. Biol. Med.* **2022**, *149*, 106017.
27. Jain, P.K.; Dubey, A.; Saba, L.; Khanna, N.N.; Laird, J.R.; Nicolaides, A.; Fouda, M.M.; Suri, J.S.; Sharma, N. Attention-based UNet Deep Learning model for Plaque segmentation in carotid ultrasound for stroke risk stratification: An artificial Intelligence paradigm. *J. Cardiovasc. Dev. Dis.* **2022**, *9*, 326. [[CrossRef](#)]
28. Biswas, M.; Saba, L.; Kalra, M.; Singh, R.; e Fernandes, J.F.; Viswanathan, V.; Laird, J.R.; Mantella, L.E.; Johri, A.M.; Fouda, M.M. MultiNet 2.0: A Lightweight Attention-based Deep Learning Network for Stenosis measurement in Carotid Ultrasound scans and Cardiovascular Risk Assessment. *Comput. Med. Imaging Graph.* **2024**, *117*, 102437.
29. Chen, R.J.; Wang, J.J.; Williamson, D.F.; Chen, T.Y.; Lipkova, J.; Lu, M.Y.; Sahai, S.; Mahmood, F. Algorithmic fairness in artificial intelligence for medicine and healthcare. *Nat. Biomed. Eng.* **2023**, *7*, 719–742. [[CrossRef](#)]
30. Zhu, F.; Gao, Z.; Zhao, C.; Zhu, H.; Nan, J.; Tian, Y.; Dong, Y.; Jiang, J.; Feng, X.; Dai, N. A deep learning-based method to extract lumen and media-adventitia in intravascular ultrasound images. *Ultrason. Imaging* **2022**, *44*, 191–203.
31. Kumari, V.; Kumar, N.; Kumar, K.S.; Kumar, A.; Skandha, S.S.; Saxena, S.; Khanna, N.N.; Laird, J.R.; Singh, N.; Fouda, M.M. Deep Learning Paradigm and Its Bias for Coronary Artery Wall Segmentation in Intravascular Ultrasound Scans: A Closer Look. *J. Cardiovasc. Dev. Dis.* **2023**, *10*, 485. [[CrossRef](#)]
32. Shen, D.; Wu, G.; Suk, H.-I. Deep learning in medical image analysis. *Annu. Rev. Biomed. Eng.* **2017**, *19*, 221–248.
33. Wang, Y.; Yao, Y. Application of artificial intelligence methods in carotid artery segmentation: A review. *IEEE Access* **2023**, *11*, 13846–13858.
34. Tian, F.; Gao, Y.; Fang, Z.; Gu, J. Automatic coronary artery segmentation algorithm based on deep learning and digital image processing. *Appl. Intell.* **2021**, *51*, 8881–8895.
35. El-Baz, A.; Suri, J.S. *Big Data in Multimodal Medical Imaging*; CRC Press: Boca Raton, FL, USA, 2019.
36. El-Baz, A.S.; Suri, J.S. *Cardiovascular and Coronary Artery Imaging: Volume 1*; Academic Press: Cambridge, MA, USA, 2021.
37. Yang, S.; Kweon, J.; Roh, J.-H.; Lee, J.-H.; Kang, H.; Park, L.-J.; Kim, D.J.; Yang, H.; Hur, J.; Kang, D.-Y. Deep learning segmentation of major vessels in X-ray coronary angiography. *Sci. Rep.* **2019**, *9*, 16897.
38. Biswas, M.; Saba, L.; Chakraborty, S.; Khanna, N.N.; Song, H.; Suri, H.S.; Sfikakis, P.P.; Mavrogeni, S.; Viskovic, K.; Laird, J.R. Two-stage artificial intelligence model for jointly measurement of atherosclerotic wall thickness and plaque burden in carotid ultrasound: A screening tool for cardiovascular/stroke risk assessment. *J. Comput. Biol.* **2020**, *123*, 103847.
39. Huang, C.; Wang, J.; Xie, Q.; Zhang, Y.-D. Analysis methods of coronary artery intravascular images: A review. *Neurocomputing* **2022**, *489*, 27–39.
40. Jena, B.; Saxena, S.; Nayak, G.K.; Saba, L.; Sharma, N.; Suri, J.S. Artificial intelligence-based hybrid deep learning models for image classification: The first narrative review. *Comput. Biol. Med.* **2021**, *137*, 104803.
41. Sanagala, S.S.; Nicolaides, A.; Gupta, S.K.; Koppula, V.K.; Saba, L.; Agarwal, S.; Johri, A.M.; Kalra, M.S.; Suri, J.S. Ten Fast Transfer Learning Models for Carotid Ultrasound Plaque Tissue Characterization in Augmentation Framework Embedded with Heatmaps for Stroke Risk Stratification. *J. Diagn.* **2021**, *11*, 2109.
42. Skandha, S.S.; Nicolaides, A.; Gupta, S.K.; Koppula, V.K.; Saba, L.; Johri, A.M.; Kalra, M.S.; Suri, J.S. A hybrid deep learning paradigm for carotid plaque tissue characterization and its validation in multicenter cohorts using a supercomputer framework. *J. Comput. Biol. Med.* **2022**, *141*, 105131.
43. Liu, F.; Lin, G.; Shen, C. CRF learning with CNN features for image segmentation. *Pattern Recognit.* **2015**, *48*, 2983–2992.
44. Jain, P.K.; Sharma, N.; Giannopoulos, A.A.; Saba, L.; Nicolaides, A.; Suri, J.S. Hybrid deep learning segmentation models for atherosclerotic plaque in internal carotid artery B-mode ultrasound. *Comput. Biol. Med.* **2021**, *136*, 104721.
45. Shamshad, F.; Khan, S.; Zamir, S.W.; Khan, M.H.; Hayat, M.; Khan, F.S.; Fu, H. Transformers in medical imaging: A survey. *Med. Image Anal.* **2023**, *88*, 102802.

46. He, K.; Gan, C.; Li, Z.; Rekik, I.; Yin, Z.; Ji, W.; Gao, Y.; Wang, Q.; Zhang, J.; Shen, D. Transformers in medical image analysis. *Intell. Med.* **2023**, *3*, 59–78. [[CrossRef](#)]
47. Strudel, R.; Garcia, R.; Laptev, I.; Schmid, C. Segformer: Transformer for semantic segmentation. In Proceedings of the IEEE/CVF International Conference on Computer Vision, Virtual, 11–17 October 2021.
48. Arora, P.; Singh, P.; Girdhar, A.; Vijayvergiya, R.; Chaudhary, P. CADNet: An advanced architecture for automatic detection of coronary artery calcification and shadow border in intravascular ultrasound (IVUS) images. *Phys. Eng. Sci. Med.* **2023**, *46*, 773–786. [[CrossRef](#)] [[PubMed](#)]
49. Huang, X.; Bajaj, R.; Li, Y.; Ye, X.; Lin, J.; Pugliese, F.; Ramasamy, A.; Gu, Y.; Wang, Y.; Torii, R. POST-IVUS: A perceptual organisation-aware selective transformer framework for intravascular ultrasound segmentation. *Med. Image Anal.* **2023**, *89*, 102922. [[CrossRef](#)]
50. Liu, Y.; Nezami, F.R.; Edelman, E.R. A Transformer-based pyramid network for coronary calcified plaque segmentation in intravascular optical coherence tomography images. *Comput. Med. Imaging Graph.* **2024**, *113*, 102347. [[CrossRef](#)]
51. Tao, X.; Dang, H.; Zhou, X.; Xu, X.; Xiong, D. Automated segmentation of lumen and media-adventitia in intravascular ultrasound images based on deep learning. In Proceedings of the 2023 3rd International Conference on Consumer Electronics and Computer Engineering (ICCECE), Guangzhou, China, 6–8 January 2023; IEEE: New York, NY, USA, 2023.
52. Xia, M.; Yang, H.; Qu, Y.; Guo, Y.; Zhou, G.; Zhang, F.; Wang, Y. Multilevel structure-preserved GAN for domain adaptation in intravascular ultrasound analysis. *Med. Image Anal.* **2022**, *82*, 102614. [[CrossRef](#)]
53. Zhang, L.; Chen, Z.; Zhang, H.; Zaman, F.A.; Wahle, A.; Wu, X.; Sonka, M. Efficient Deep-Learning-Assisted Annotation for Medical Image Segmentation. *Authorea Prepr.* **2023**. [[CrossRef](#)]
54. Zhi, Y.; Zhang, H.; Gao, Z. Vessel contour detection in intracoronary images via bilateral cross-domain adaptation. *IEEE J. Biomed. Health Inform.* **2023**, *27*, 3314–3325. [[CrossRef](#)]
55. Dosovitskiy, A.; Beyer, L.; Kolesnikov, A.; Weissenborn, D.; Zhai, X.; Unterthiner, T.; Dehghani, M.; Minderer, M.; Heigold, G.; Gelly, S. An image is worth 16x16 words: Transformers for image recognition at scale. *arXiv* **2020**, arXiv:2010.11929.
56. Giannoglou, G.D.; Chatzizisis, Y.S.; Koutkias, V.; Kompatsiaris, I.; Papadogiorgaki, M.; Mezaris, V.; Parissi, E.; Diamantopoulos, P.; Strintzis, M.G.; Maglaveras, N. A novel active contour model for fully automated segmentation of intravascular ultrasound images: In vivo validation in human coronary arteries. *Comput. Biol. Med.* **2007**, *37*, 1292–1302. [[CrossRef](#)]
57. Destrepes, F.; Cardinal, M.-H.R.; Allard, L.; Tardif, J.-C.; Cloutier, G. Segmentation method of intravascular ultrasound images of human coronary arteries. *Comput. Med. Imaging Graph.* **2014**, *38*, 91–103. [[CrossRef](#)]
58. Hammouche, A.; Cloutier, G.; Tardif, J.-C.; Hammouche, K.; Meunier, J. Automatic IVUS lumen segmentation using a 3D adaptive helix model. *Comput. Biol. Med.* **2019**, *107*, 58–72. [[PubMed](#)]
59. Huang, Y.; Yan, W.; Xia, M.; Guo, Y.; Zhou, G.; Wang, Y. Vessel membrane segmentation and calcification location in intravascular ultrasound images using a region detector and an effective selection strategy. *Comput. Methods Programs Biomed.* **2020**, *189*, 105339. [[CrossRef](#)] [[PubMed](#)]
60. Kermani, A.; Ayatollahi, A. A new nonparametric statistical approach to detect lumen and Media-Adventitia borders in intravascular ultrasound frames. *Comput. Biol. Med.* **2018**, *104*, 10–28. [[PubMed](#)]
61. Vercio, L.L.; Del Fresno, M.; Larrabide, I. Lumen-intima and media-adventitia segmentation in IVUS images using supervised classifications of arterial layers and morphological structures. *Comput. Methods Programs Biomed.* **2019**, *177*, 113–121. [[CrossRef](#)]
62. Bajaj, R.; Huang, X.; Kilic, Y.; Ramasamy, A.; Jain, A.; Ozkor, M.; Tufaro, V.; Safi, H.; Erdogan, E.; Serruys, P.W. Advanced deep learning methodology for accurate, real-time segmentation of high-resolution intravascular ultrasound images. *Int. J. Cardiol.* **2021**, *339*, 185–191. [[CrossRef](#)]
63. Nishi, T.; Yamashita, R.; Imura, S.; Tateishi, K.; Kitahara, H.; Kobayashi, Y.; Yock, P.G.; Fitzgerald, P.J.; Honda, Y. Deep learning-based intravascular ultrasound segmentation for the assessment of coronary artery disease. *Int. J. Cardiol.* **2021**, *333*, 55–59. [[CrossRef](#)]
64. Balaji, A.; Kelsey, L.J.; Majeed, K.; Schultz, C.J.; Doyle, B.J. Coronary artery segmentation from intravascular optical coherence tomography using deep capsules. *Artif. Intell. Med.* **2021**, *116*, 102072.
65. Meng, L.; Jiang, M.; Zhang, C.; Zhang, J. Deep learning segmentation, classification, and risk prediction of complex vascular lesions on intravascular ultrasound images. *Biomed. Signal Process. Control* **2023**, *82*, 104584. [[CrossRef](#)]
66. Sun, S.; Sonka, M.; Beichel, R.R. Graph-based IVUS segmentation with efficient computer-aided refinement. *IEEE Trans. Med. Imaging* **2013**, *32*, 1536–1549.
67. Jodas, D.S.; Pereira, A.S.; Tavares, J.M.R. Automatic segmentation of the lumen region in intravascular images of the coronary artery. *Med. Image Anal.* **2017**, *40*, 60–79.
68. Onpans, J.; Yookwan, W.; Sangrueng, J.; Srikamdee, S. Intravascular Ultrasound Image Composite Segmentation using Ensemble Gabor-spatial Features. In Proceedings of the 2022 13th International Conference on Information and Communication Technology Convergence (ICTC), Jeju Island, Republic of Korea, 19–21 October 2022; IEEE: New York, NY, USA, 2022.

69. Liu, X.; Feng, T.; Liu, W.; Song, L.; Yuan, Y.; Hau, W.K.; Del Ser, J.; Gao, Z. Scale mutualized perception for vessel border detection in intravascular ultrasound images. *IEEE/ACM Trans. Comput. Biol. Bioinform.* **2024**, *21*, 1060–1071. [[PubMed](#)]
70. Mendizabal-Ruiz, E.G.; Rivera, M.; Kakadiaris, I.A. Segmentation of the luminal border in intravascular ultrasound B-mode images using a probabilistic approach. *Med. Image Anal.* **2013**, *17*, 649–670. [[CrossRef](#)] [[PubMed](#)]
71. Gao, Z.; Chung, J.; Abdelrazek, M.; Leung, S.; Hau, W.K.; Xian, Z.; Zhang, H.; Li, S. Privileged modality distillation for vessel border detection in intracoronary imaging. *IEEE Trans. Med. Imaging* **2019**, *39*, 1524–1534.
72. Mishra, D.; Chaudhury, S.; Sarkar, M.; Sooin, A.S. Ultrasound image segmentation: A deeply supervised network with attention to boundaries. *IEEE Trans. Biomed. Eng.* **2018**, *66*, 1637–1648.
73. Yuan, S.; Yang, F. CSDN: Combining Shallow and Deep Networks for Accurate Real-Time Segmentation of High-Definition Intravascular Ultrasound Images. In Proceedings of the 2023 IEEE 20th International Symposium on Biomedical Imaging (ISBI), Cartagena, Colombia, 18–21 April 2023; IEEE: New York, NY, USA, 2023.
74. Li, Y.-C.; Shen, T.-Y.; Chen, C.-C.; Chang, W.-T.; Lee, P.-Y.; Huang, C.-C.J. Automatic detection of atherosclerotic plaque and calcification from intravascular ultrasound images by using deep convolutional neural networks. *IEEE Trans. Ultrason. Ferroelectr. Freq. Control* **2021**, *68*, 1762–1772.
75. Xia, M.; Yan, W.; Huang, Y.; Guo, Y.; Zhou, G.; Wang, Y. Extracting membrane borders in ivus images using a multi-scale feature aggregated u-net. In Proceedings of the 2020 42nd Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC), Montreal, QC, Canada, 20–24 July 2020; IEEE: New York, NY, USA, 2020.
76. Serrano-Antón, B.; Otero-Cacho, A.; López-Otero, D.; Díaz-Fernández, B.; Bastos-Fernández, M.; Pérez-Muñuzuri, V.; González-Juanatey, J.R.; Muñuzuri, A.P. Coronary artery segmentation based on transfer learning and UNet architecture on computed tomography coronary angiography images. *IEEE Access* **2023**, *11*, 75484–75496.
77. Araki, T.; Banchhor, S.K.; Londhe, N.D.; Ikeda, N.; Radeva, P.; Shukla, D.; Saba, L.; Balestrieri, A.; Nicolaides, A.; Shafique, S. Reliable and accurate calcium volume measurement in coronary artery using intravascular ultrasound videos. *J. Med. Syst.* **2016**, *40*, 51.
78. Bargsten, L.; Raschka, S.; Schlaefer, A. Capsule networks for segmentation of small intravascular ultrasound image datasets. *Int. J. Comput. Assist. Radiol. Surg.* **2021**, *16*, 1243–1254.
79. Yang, J.; Tong, L.; Faraji, M.; Basu, A. IVUS-Net: An intravascular ultrasound segmentation network. In Proceedings of the Smart Multimedia: First International Conference, ICSM 2018, Toulon, France, 24–26 August 2018, Revised Selected Papers 1; Springer: Berlin/Heidelberg, Germany, 2018.
80. Chen, T.; Yu, H.; Jia, H.; Dai, J.; Fang, C.; Ma, L.; Liu, H.; Xu, M.; Yu, B. Automatic assessment of calcified plaque and nodule by optical coherence tomography adopting deep learning model. *Int. J. Cardiovasc. Imaging* **2022**, *38*, 2501–2510.
81. Arora, P.; Singh, P.; Girdhar, A.; Vijayvergiya, R. A state-of-the-art review on coronary artery border segmentation algorithms for intravascular ultrasound (IVUS) images. *Cardiovasc. Eng. Technol.* **2023**, *14*, 264–295.
82. Xia, M.; Yan, W.; Huang, Y.; Guo, Y.; Zhou, G.; Wang, Y. Ivus image segmentation using superpixel-wise fuzzy clustering and level set evolution. *Appl. Sci.* **2019**, *9*, 4967. [[CrossRef](#)]
83. Vercio, L.L.; Del Fresno, M.; Larrabide, I. Detection of morphological structures for vessel wall segmentation in ivus using random forests. In Proceedings of the 12th International Symposium on Medical Information Processing and Analysis, Tandil, Argentina, 5–7 December 2016; SPIE: Bellingham, WA, USA, 2017.
84. Zhu, H.; Liang, Y.; Friedman, M.H. IVUS image segmentation based on contrast. In Proceedings of the Medical Imaging 2002: Image Processing, San Diego, CA, USA, 24–28 February 2002; SPIE: Bellingham, WA, USA, 2002.
85. Gao, Z.; Guo, W.; Liu, X.; Huang, W.; Zhang, H.; Tan, N.; Hau, W.K.; Zhang, Y.-T.; Liu, H. Automated detection framework of the calcified plaque with acoustic shadowing in IVUS images. *PLoS ONE* **2014**, *9*, e109997. [[CrossRef](#)] [[PubMed](#)]
86. Cui, H.; Xia, Y.; Zhang, Y. Supervised machine learning for coronary artery lumen segmentation in intravascular ultrasound images. *Int. J. Numer. Methods Biomed. Eng.* **2020**, *36*, e3348. [[CrossRef](#)] [[PubMed](#)]
87. Matsumura, M.; Mintz, G.S.; Dohi, T.; Li, W.; Shang, A.; Fall, K.; Sato, T.; Sugizaki, Y.; Chatzizisis, Y.S.; Moses, J.W. Accuracy of IVUS-based machine learning segmentation assessment of coronary artery dimensions and balloon sizing. *JACC Adv.* **2023**, *2*, 100564. [[CrossRef](#)]
88. Wang, Y.-Y.; Peng, W.-X.; Qiu, C.-H.; Jiang, J.; Xia, S.-R. Fractional-order Darwinian PSO-based feature selection for media-adventitia border detection in intravascular ultrasound images. *Ultrasonics* **2019**, *92*, 1–7. [[CrossRef](#)]
89. Zhou, R.; Azarpazhooh, M.R.; Spence, J.D.; Hashemi, S.; Ma, W.; Cheng, X.; Gan, H.; Ding, M.; Fenster, A. Deep learning-based carotid plaque segmentation from B-mode ultrasound images. *Ultrasound Med. Biol.* **2021**, *47*, 2723–2733. [[CrossRef](#)]
90. Dong, L.; Jiang, W.; Lu, W.; Jiang, J.; Zhao, Y.; Song, X.; Leng, X.; Zhao, H.; Wang, J.a.; Li, C. Automatic segmentation of coronary lumen and external elastic membrane in intravascular ultrasound images using 8-layer U-Net. *BioMed. Eng. OnLine* **2021**, *20*, 16. [[CrossRef](#)]
91. dos Santos Filho, E.; Yoshizawa, M.; Tanaka, A.; Saijo, Y. A study on intravascular ultrasound image processing. *Rec. Electr. Commun. Eng. Conversazione Tohoku Univ.* **2006**, *74*, 30.

92. Kerre, E.E.; Nachtgeael, M. *Fuzzy Techniques in Image Processing*; Physica: Heidelberg, Germany, 2013; Volume 52.
93. Hu, M.; Zhong, Y.; Xie, S.; Lv, H.; Lv, Z. Fuzzy system based medical image processing for brain disease prediction. *Front. Neurosci.* **2021**, *15*, 714318. [[CrossRef](#)]
94. Li, C.; Xu, C.; Gui, C.; Fox, M.D. Distance regularized level set evolution and its application to image segmentation. *IEEE Trans. Image Process.* **2010**, *19*, 3243–3254. [[CrossRef](#)]
95. Chen, X.-X.; Kong, Z.-X.; Wei, S.-F.; Liang, F.; Feng, T.; Wang, S.-S.; Gao, J.-S. Ultrasound Imaging-vulnerable plaque diagnostics: Automatic carotid plaque segmentation based on deep learning. *J. Radiat. Res. Appl. Sci.* **2023**, *16*, 100598.
96. Wang, Y.; Gao, X.; Wang, Y.; Sun, J. Adventitia segmentation in intravascular ultrasound images based on improved Snake algorithm. *Optik* **2021**, *241*, 167175. [[CrossRef](#)]
97. Hellier, P.; Coupé, P.; Morandi, X.; Collins, D.L. An automatic geometrical and statistical method to detect acoustic shadows in intraoperative ultrasound brain images. *Med. Image Anal.* **2010**, *14*, 195–204. [[PubMed](#)]
98. Hearst, M.A.; Dumais, S.T.; Osuna, E.; Platt, J.; Scholkopf, B. Support vector machines. *IEEE Intell. Syst. Their Appl.* **1998**, *13*, 18–28.
99. Tseng, C.-J.; Tang, C. An optimized XGBoost technique for accurate brain tumor detection using feature selection and image segmentation. *Healthc. Anal.* **2023**, *4*, 100217.
100. Molinari, F.; Zeng, G.; Suri, J.S. Inter-greedy technique for fusion of different segmentation strategies leading to high-performance carotid IMT measurement in ultrasound images. In *Atherosclerosis Disease Management*; Springer: New York, NY, USA, 2011; pp. 253–279.
101. Molinari, F.; Meiburger, K.M.; Saba, L.; Acharya, U.R.; Famiglietti, L.; Georgiou, N.; Nicolaidis, A.; Mamidi, R.S.; Kuper, H.; Suri, J.S. Automated carotid IMT measurement and its validation in low contrast ultrasound database of 885 patient Indian population epidemiological study: Results of AtheroEdge® software. In *Multi-Modality Atherosclerosis Imaging and Diagnosis*; Springer: New York, NY, USA, 2014; pp. 209–219.
102. Jamthikar, A.; Gupta, D.; Khanna, N.N.; Saba, L.; Araki, T.; Viskovic, K.; Suri, H.S.; Gupta, A.; Mavrogeni, S.; Turk, M. A low-cost machine learning-based cardiovascular/stroke risk assessment system: Integration of conventional factors with image phenotypes. *Cardiovasc. Diagn. Ther.* **2019**, *9*, 420. [[CrossRef](#)]
103. Jamthikar, A.; Gupta, D.; Khanna, N.N.; Saba, L.; Laird, J.R.; Suri, J.S. Cardiovascular/stroke risk prevention: A new machine learning framework integrating carotid ultrasound image-based phenotypes and its harmonics with conventional risk factors. *Indian Heart J.* **2020**, *72*, 258–264. [[CrossRef](#)]
104. Saba, L.; Biswas, M.; Kuppili, V.; Godia, E.C.; Suri, H.S.; Edla, D.R.; Omerzu, T.; Laird, J.R.; Khanna, N.N.; Mavrogeni, S. The present and future of deep learning in radiology. *Eur. J. Radiol.* **2019**, *114*, 14–24.
105. Skandha, S.S.; Gupta, S.K.; Saba, L.; Koppula, V.K.; Johri, A.M.; Khanna, N.N.; Mavrogeni, S.; Laird, J.R.; Pareek, G.; Miner, M. 3-D optimized classification and characterization artificial intelligence paradigm for cardiovascular/stroke risk stratification using carotid ultrasound-based delineated plaque: Atheromatic™ 2.0. *Comput. Biol. Med.* **2020**, *125*, 103958.
106. Saba, L.; Sanagala, S.S.; Gupta, S.K.; Koppula, V.K.; Johri, A.M.; Khanna, N.N.; Mavrogeni, S.; Laird, J.R.; Pareek, G.; Miner, M. Multimodality carotid plaque tissue characterization and classification in the artificial intelligence paradigm: A narrative review for stroke application. *Ann. Transl. Med.* **2021**, *9*, 1206.
107. Ronneberger, O.; Fischer, P.; Brox, T. U-net: Convolutional networks for biomedical image segmentation. In *Medical Image Computing and Computer-Assisted Intervention—MICCAI 2015, Proceedings of the 18th International Conference, Munich, Germany, 5–9 October 2015, Proceedings, Part III 18*; Springer: Berlin/Heidelberg, Germany, 2015.
108. Suri, J.S.; Bhagawati, M.; Agarwal, S.; Paul, S.; Pandey, A.; Gupta, S.K.; Saba, L.; Paraskevas, K.I.; Khanna, N.N.; Laird, J.R. UNet deep learning architecture for segmentation of vascular and non-vascular images: A microscopic look at UNet components buffered with pruning, explainable artificial intelligence, and bias. *IEEE Access* **2022**, *11*, 595–645.
109. Gupta, S.; Dubey, A.K.; Singh, R.; Kalra, M.K.; Abraham, A.; Kumari, V.; Laird, J.R.; Al-Maini, M.; Gupta, N.; Singh, I. Four Transformer-Based Deep Learning Classifiers Embedded with an Attention U-Net-Based Lung Segmenter and Layer-Wise Relevance Propagation-Based Heatmaps for COVID-19 X-ray Scans. *Diagnostics* **2024**, *14*, 1534. [[CrossRef](#)] [[PubMed](#)]
110. Gemenaris, M.; Liapi, G.D.; Markides, C.; Constantinou, K.; Loizou, C.P.; Neophytou, M.; Kardoulas, D.; Pattichis, M.S.; Griffin, M.; Nicolaidis, A.; et al. AtheroRisk: An Integrated Computer Software System for Stroke Risk Stratification Utilizing Carotid Plaque Analysis in Ultrasound Videos. *SN Comput. Sci.* **2025**, *6*, 1–16.
111. Suri, J.S.; Agarwal, S.; Chabert, G.L.; Carriero, A.; Paschè, A.; Danna, P.S.; Saba, L.; Mehmedović, A.; Faa, G.; Singh, I.M. COVLIAS 2.0-cXAI: Cloud-based explainable deep learning system for COVID-19 lesion localization in computed tomography scans. *Diagnostics* **2022**, *12*, 1482. [[CrossRef](#)]
112. Suri, J.S.; Agarwal, S.; Chabert, G.L.; Carriero, A.; Paschè, A.; Danna, P.S.; Saba, L.; Mehmedović, A.; Faa, G.; Singh, I.M. COVLIAS 1.0 lesion vs. medseg: An artificial intelligence framework for automated lesion segmentation in COVID-19 lung computed tomography scans. *Diagnostics* **2022**, *12*, 1283. [[CrossRef](#)]

113. Radeva, P.; Suri, J.S. *Vascular and Intravascular Imaging Trends, Analysis, and Challenges, Volume 2: Plaque Characterization*; IOP Publishing: Bristol, UK, 2019.
114. Abraham, N.; Khan, N.M. A novel focal tvrsky loss function with improved attention u-net for lesion segmentation. In Proceedings of the 2019 IEEE 16th International Symposium on Biomedical Imaging (ISBI 2019), Venice, Italy, 8–11 April 2019; IEEE: New York, NY, USA, 2019.
115. Li, D.; Peng, Y.; Guo, Y.; Sun, J. MFAUNet: Multiscale feature attentive U-Net for cardiac MRI structural segmentation. *IET Image Process.* **2022**, *16*, 1227–1242.
116. Zhou, Z.; Rahman Siddiquee, M.M.; Tajbakhsh, N.; Liang, J. Unet++: A nested u-net architecture for medical image segmentation. In *Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support, Proceedings of the 4th International Workshop, DLMIA 2018, and 8th International Workshop, ML-CDS 2018, Held in Conjunction with MICCAI 2018, Granada, Spain, 20 September 2018, Proceedings 4*; Springer: Berlin/Heidelberg, Germany, 2018.
117. Srivastava, S.; Vidyarthi, A.; Jain, S.J. Analytical study of the encoder-decoder models for ultrasound image segmentation. *Serv. Oriented Comput. Appl.* **2024**, *18*, 81–100.
118. Maindarkar, M.; Kumar, A. Artificial intelligence based disease diagnosis using ultrasound imaging. In *Revolutionizing Medical Systems Using Artificial Intelligence*; Academic Press: Cambridge, MA, USA, 2025; pp. 147–161.
119. Vaswani, A.; Shazeer, N.; Parmar, N.; Uszkoreit, J.; Jones, L.; Gomez, A.N.; Kaiser, Ł.; Polosukhin, I. Attention is all you need. *Adv. Neural Inf. Process. Syst.* **2017**, *30*.
120. Oktay, O.; Schlemper, J.; Folgoc, L.L.; Lee, M.; Heinrich, M.; Misawa, K.; Mori, K.; McDonagh, S.; Hammerla, N.Y.; Kainz, B. Attention U-Net: Learning where to look for the pancreas. *arXiv* **2018**, arXiv:1804.03999.
121. Li, Y.; Zou, L.; Xiong, L.; Yu, F.; Jiang, H.; Fan, C.; Cheng, M.; Li, Q. FRDD-Net: Automated carotid plaque ultrasound images segmentation using feature remapping and dense decoding. *Sensors* **2022**, *22*, 887. [[CrossRef](#)]
122. Jávorszky, N.; Homonnay, B.; Gerstenblith, G.; Bluemke, D.; Kiss, P.; Török, M.; Celentano, D.; Lai, H.; Lai, S.; Kolossváry, M. Deep learning-based atherosclerotic coronary plaque segmentation on coronary CT angiography. *Eur. Radiol.* **2022**, *32*, 7217–7226. [[CrossRef](#)] [[PubMed](#)]
123. Chen, J.; Lu, Y.; Yu, Q.; Luo, X.; Adeli, E.; Wang, Y.; Lu, L.; Yuille, A.L.; Zhou, Y. Transunet: Transformers make strong encoders for medical image segmentation. *arXiv* **2021**, arXiv:2102.04306.
124. Jiang, T.; Xing, W.; Yu, M.; Ta, D. A hybrid enhanced attention transformer network for medical ultrasound image segmentation. *Biomed. Signal Process. Control* **2023**, *86*, 105329. [[CrossRef](#)]
125. Shen, J.; Hu, Y.; Zhang, X.; Gong, Y.; Kawasaki, R.; Liu, J. Structure-Oriented Transformer for retinal diseases grading from OCT images. *Comput. Biol. Med.* **2023**, *152*, 106445. [[CrossRef](#)]
126. Kamran, S.A.; Hossain, K.F.; Tavakkoli, A.; Zuckerbrod, S.L.; Baker, S.A. Vtgan: Semi-supervised retinal image synthesis and disease prediction using vision transformers. In Proceedings of the IEEE/CVF International Conference on Computer Vision, Virtual, 11–17 October 2021.
127. Shen, X.; Wang, L.; Zhao, Y.; Liu, R.; Qian, W.; Ma, H. Dilated transformer: Residual axial attention for breast ultrasound image segmentation. *Quant. Imaging Med. Surg.* **2022**, *12*, 4512. [[CrossRef](#)]
128. Wang, Q.; Xu, L.; Wang, L.; Yang, X.; Sun, Y.; Yang, B.; Greenwald, S.E. Automatic coronary artery segmentation of CCTA images using UNet with a local contextual transformer. *Front. Physiol.* **2023**, *14*, 1138257. [[CrossRef](#)]
129. Xu, C.; Li, M.; Wu, X. TransCC: Transformer Network for Coronary Artery CCTA Segmentation. *arXiv* **2023**, arXiv:2310.04779.
130. Cao, H.; Wang, Y.; Chen, J.; Jiang, D.; Zhang, X.; Tian, Q.; Wang, M. Swin-unet: Unet-like pure transformer for medical image segmentation. In Proceedings of the European Conference on Computer Vision, Tel Aviv, Israel, 23–27 October 2022; Springer: Berlin/Heidelberg, Germany, 2022.
131. Shen, X.; Xu, J.; Jia, H.; Fan, P.; Dong, F.; Yu, B.; Ren, S. Self-attentional microvessel segmentation via squeeze-excitation transformer Unet. *Comput. Med. Imaging Graph.* **2022**, *97*, 102055. [[CrossRef](#)]
132. Vafaezadeh, M.; Behnam, H.; Gifani, P. Ultrasound Image Analysis with Vision Transformers. *Diagnostics* **2024**, *14*, 542. [[CrossRef](#)]
133. Saba, L.; Anzidei, M.; Sanfilippo, R.; Montisci, R.; Lucatelli, P.; Catalano, C.; Passariello, R.; Mallarini, G. Imaging of the carotid artery. *Atherosclerosis* **2012**, *220*, 294–309. [[CrossRef](#)]
134. Liu, K.; Suri, J.S. Automatic Vessel Identification for Angiographic Screening. Patent 6845260, 18 January 2005.
135. Acharya, U.R.; Sree, S.V.; Krishnan, M.M.R.; Krishnananda, N.; Ranjan, S.; Umesh, P.; Suri, J.S. Automated classification of patients with coronary artery disease using grayscale features from left ventricle echocardiographic images. *Comput. Methods Programs Biomed.* **2013**, *112*, 624–632. [[PubMed](#)]
136. Ullah, W.; Hussain, T.; Ullah, F.U.M.; Lee, M.Y.; Baik, S.W. TransCNN: Hybrid CNN and transformer mechanism for surveillance anomaly detection. *Eng. Appl. Artif. Intell.* **2023**, *123*, 106173.
137. Li, Z.; Li, D.; Xu, C.; Wang, W.; Hong, Q.; Li, Q.; Tian, J. Tfcns: A cnn-transformer hybrid network for medical image segmentation. In Proceedings of the International Conference on Artificial Neural Networks, Bristol, UK, 6–9 September 2022; Springer: Berlin/Heidelberg, Germany, 2022.

138. He, Q.; Yang, Q.; Xie, M. HCTNet: A hybrid CNN-transformer network for breast ultrasound image segmentation. *Comput. Biol. Med.* **2023**, *155*, 106629. [[CrossRef](#)] [[PubMed](#)]
139. Bhagawati, M.; Paul, S.; Mantella, L.; Johri, A.M.; Gupta, S.; Laird, J.R.; Singh, I.M.; Khanna, N.N.; Al-Maini, M.; Isenovic, E.R. Cardiovascular Disease Risk Stratification Using Hybrid Deep Learning Paradigm: First of Its Kind on Canadian Trial Data. *Diagnostics* **2024**, *14*, 1894. [[CrossRef](#)]
140. Singh, J.; Khanna, N.N.; Rout, R.K.; Singh, N.; Laird, J.R.; Singh, I.M.; Kalra, M.K.; Mantella, L.E.; Johri, A.M.; Isenovic, E.R. GeneAI 3.0: Powerful, novel, generalized hybrid and ensemble deep learning frameworks for miRNA species classification of stationary patterns from nucleotides. *Sci. Rep.* **2024**, *14*, 7154.
141. Gupta, N.; Gupta, S.K.; Pathak, R.K.; Jain, V.; Rashidi, P.; Suri, J.S. Human activity recognition in artificial intelligence framework: A narrative review. *Artif. Intell. Rev.* **2022**, *55*, 4755–4808. [[CrossRef](#)]
142. Lin, C.-H.; Yumer, E.; Wang, O.; Shechtman, E.; Lucey, S. St-gan: Spatial transformer generative adversarial networks for image compositing. In Proceedings of the IEEE conference on computer vision and pattern recognition, Salt Lake City, UT, USA, 18–22 June 2018.
143. Hudson, D.A.; Zitnick, L. Generative adversarial transformers. In Proceedings of the International Conference on Machine Learning, Virtual, 18–24 July 2021; PMLR: Birmingham, UK, 2021.
144. Dubey, S.R.; Singh, S.K. Transformer-based generative adversarial networks in computer vision: A comprehensive survey. *IEEE Trans. Artif. Intell.* **2024**, *5*, 4851–4867.
145. Zhang, B.; Gu, S.; Zhang, B.; Bao, J.; Chen, D.; Wen, F.; Wang, Y.; Guo, B. Styleswin: Transformer-based gan for high-resolution image generation. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, New Orleans, LA, USA, 18–24 June 2022.
146. Luo, Y.; Wang, Y.; Zu, C.; Zhan, B.; Wu, X.; Zhou, J.; Shen, D.; Zhou, L. 3D transformer-GAN for high-quality PET reconstruction. In *Medical Image Computing and Computer Assisted Intervention—MICCAI 2021, Proceedings of the 24th International Conference, Strasbourg, France, 27 September–1 October 2021, Proceedings, Part VI 24*; Springer: Berlin/Heidelberg, Germany, 2021.
147. Khanna, N.N.; Maindarkar, M.A.; Viswanathan, V.; Fernandes, J.F.E.; Paul, S.; Bhagawati, M.; Ahluwalia, P.; Ruzsa, Z.; Sharma, A.; Kolluri, R. Economics of Artificial Intelligence in Healthcare: Diagnosis vs. Treatment. *Healthcare* **2022**, *10*, 2493. [[CrossRef](#)]
148. Singh, M.; Kumar, A.; Khanna, N.N.; Laird, J.R.; Nicolaidis, A.; Faa, G.; Johri, A.M.; Mantella, L.E.; Fernandes, J.F.E.; Teji, J.S. Artificial intelligence for cardiovascular disease risk assessment in personalised framework: A scoping review. *EClinicalMedicine* **2024**, *73*, 102660.
149. Xiao, H.; Li, L.; Liu, Q.; Zhu, X.; Zhang, Q. Transformers in medical image segmentation: A review. *Biomed. Signal Process. Control.* **2023**, *84*, 104791.
150. Milosevic, M.; Jin, Q.; Singh, A.; Amal, S. Applications of AI in multi-modal imaging for cardiovascular disease. *Front. Radiol.* **2024**, *3*, 1294068. [[CrossRef](#)] [[PubMed](#)]
151. Tiwari, P.; Viswanath, S.; Lee, G.; Madabhushi, A. Multi-modal data fusion schemes for integrated classification of imaging and non-imaging biomedical data. In Proceedings of the 2011 IEEE International Symposium on Biomedical Imaging: From Nano to Macro, Chicago, IL, USA, 30 March–2 April 2011; IEEE: New York, NY, USA, 2011.
152. Nie, D.; Zhang, H.; Adeli, E.; Liu, L.; Shen, D. 3D deep learning for multi-modal imaging-guided survival time prediction of brain tumor patients. In *Medical Image Computing and Computer-Assisted Intervention—MICCAI 2016, Proceedings of the 19th International Conference, Athens, Greece, 17–21 October 2016, Proceedings, Part II 19*; Springer: Berlin/Heidelberg, Germany, 2016.
153. Li, J.; Han, X.; Qin, Y.; Tan, F.; Chen, Y.; Wang, Z.; Song, H.; Zhou, X.; Zhang, Y.; Hu, L. Artificial intelligence accelerates multi-modal biomedical process: A Survey. *Neurocomputing* **2023**, *558*, 126720. [[CrossRef](#)]
154. Ho, N.N.; Lee, K.Y.; Noh, J.; Lee, S.-W. Deep Learning-Based Intravascular Ultrasound Images Segmentation in Coronary Artery Disease: A Start Developing the Cornerstone. *Korean Circ. J.* **2024**, *54*, 40–42. [[CrossRef](#)] [[PubMed](#)]
155. Szarski, M.; Chauhan, S. Improved real-time segmentation of intravascular ultrasound images using coordinate-aware fully convolutional networks. *Comput. Med. Imaging Graph.* **2021**, *91*, 101955. [[CrossRef](#)]
156. Hsiao, C.-H.; Chen, K.; Peng, T.-Y.; Huang, W.-C. Federated learning for coronary artery plaque detection in atherosclerosis using ivus imaging: A multi-hospital collaboration. *arXiv* **2024**, arXiv:2412.15307.
157. Dong, L.; Lu, W.; Lu, X.; Leng, X.; Xiang, J.; Li, C. Comparison of deep learning-based image segmentation methods for intravascular ultrasound on retrospective and large image cohort study. *BioMed. Eng. OnLine* **2023**, *22*, 111. [[CrossRef](#)]

**Disclaimer/Publisher’s Note:** The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.