Segment Anything Model in Medical Image Segmentation: A Survey

Abstract

The Segment Anything Model (SAM) has emerged as a groundbreaking vision foundation model in the field of image segmentation, particularly within medical imaging. This survey paper critically examines SAM's adaptability and performance across diverse medical imaging tasks and modalities. SAM's architecture, leveraging prompt engineering and efficient fine-tuning, significantly enhances segmentation accuracy and operational efficiency, reducing the annotation burden on medical professionals. The model's zero-shot learning capability allows it to perform segmentation tasks without extensive labeled datasets, crucial in medical imaging where data acquisition is costly. Despite its transformative impact, SAM faces challenges such as domain specificity, data scarcity, and the need for further refinement to optimize its performance across different contexts. The integration of advanced models and methods, alongside the development of hybrid architectures, underscores SAM's potential to improve feature extraction and segmentation outcomes. Future research should focus on expanding SAM's capabilities to include multi-modal image segmentation, refining prompting strategies, and developing comprehensive benchmarks to enhance its applicability in clinical settings. By addressing these challenges, SAM can continue to advance medical imaging, providing significant improvements in diagnostic precision and clinical outcomes.

1 Introduction

1.1 Significance of SAM in Image Segmentation

The Segment Anything Model (SAM) represents a groundbreaking advancement in image segmentation, being the first foundation model tailored specifically for this purpose [1]. This innovation marks a pivotal transition in segmentation capabilities, particularly addressing the challenges posed by traditional methods that often falter in complex low-level tasks, especially in medical imaging [2]. SAM's architecture adeptly segments intricate structures, such as brain tumors, thereby enhancing diagnostic accuracy and clinical outcomes [1].

Beyond medical imaging, SAM significantly influences the broader field of image segmentation by improving communication efficiency through compact semantic-level representations [3]. This functionality highlights SAM's role in advancing the efficiency and accuracy of image analysis across various domains. Furthermore, SAM has spurred additional research into foundational models, marking a significant milestone in the evolution of visual perception technologies [4].

Nevertheless, SAM faces challenges in remote sensing instance segmentation due to limited pretraining on diverse datasets [5]. These issues underscore the need for ongoing refinement and adaptation strategies to broaden SAM's applicability [6]. Additionally, initial versions of the model encountered efficiency challenges linked to a cumbersome image encoder and a slow mask decoder, which have since become focal points for optimization [7].

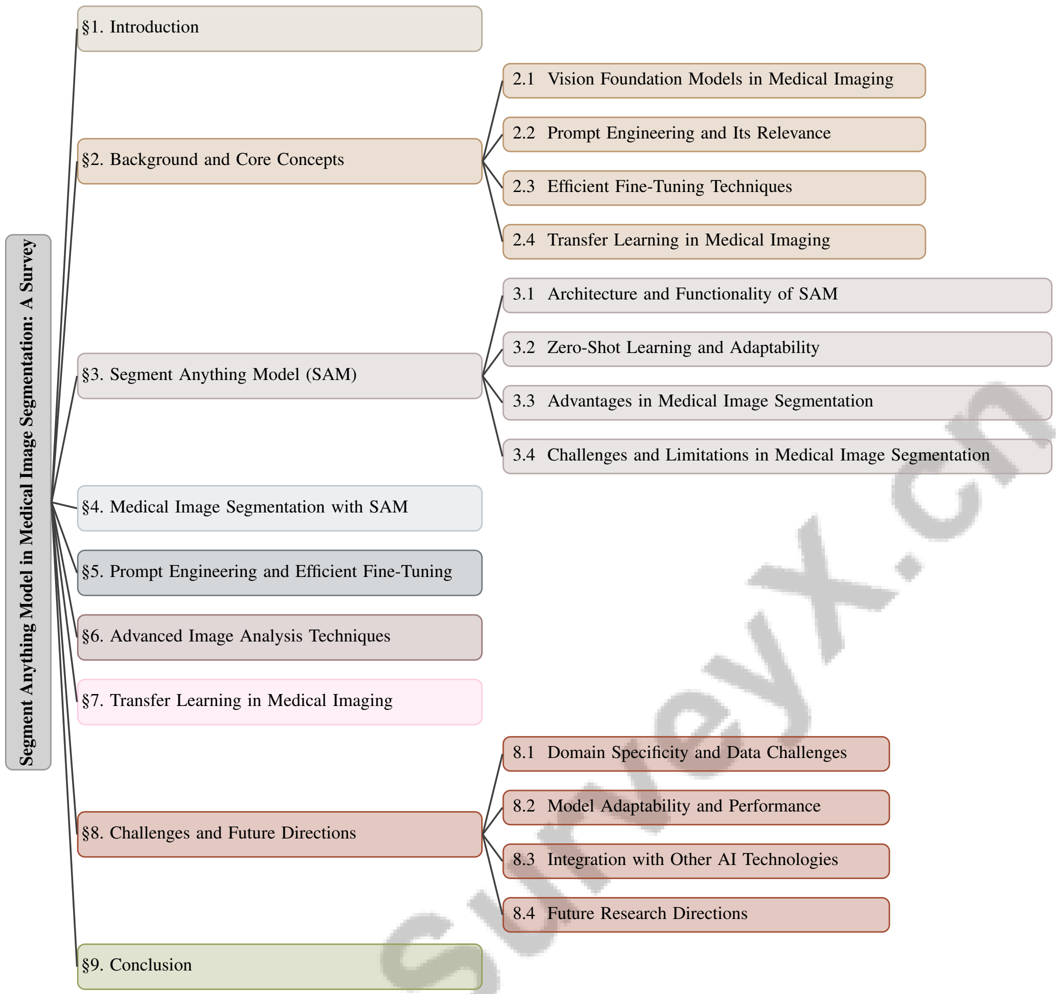


Figure 1: chapter structure

1.2 Motivation for Using SAM in Medical Image Segmentation

SAM has emerged as a transformative tool in medical image segmentation, effectively addressing the unique challenges of this domain. A primary motivation for its adoption is the potential to alleviate the annotation burden on medical professionals by enabling segmentation based solely on 2D interactive feedback, thus streamlining the annotation process and enhancing efficiency [8]. This capability is particularly crucial in medical imaging, where timely and accurate segmentation is paramount.

SAM's versatility, allowing for segmentation of both images and videos, further emphasizes its ability to meet the diverse demands of medical imaging tasks [7]. Its zero-shot learning feature enables segmentation without extensive labeled datasets, a significant advantage in medical imaging where acquiring labeled data is often resource-intensive. Moreover, the incorporation of object-aware prompts in SAM optimizes the mask generation process, enhancing its efficacy in medical applications [7].

Additionally, SAM enhances semi-supervised learning methods in medical image segmentation by leveraging limited labeled data to boost performance. Its innovative combination of automatic and interactive segmentation methodologies further enhances its effectiveness across various medical imaging contexts [9]. However, challenges remain in achieving semantic-aware segmentation due to SAM's generalized approach to grouping image pixels into patches. Despite this, its application in brain tumor segmentation, evaluated through specific benchmarks, underscores its potential to overcome limitations of existing dataset-specific algorithms.

1.3 Objectives of the Survey

This survey aims to critically evaluate the adaptation and performance of SAM within medical image segmentation, focusing on its application across diverse imaging tasks and modalities. A key objective is to assess the methodological adaptations necessary for SAM to effectively manage specific medical segmentation tasks, as evidenced by its performance in various contexts [1]. This includes exploring the potential of MedSAM as a universal foundation model capable of addressing a wide array of segmentation challenges, with its versatility and robustness yielding consistent results across tasks [10].

The survey also investigates novel strategies to enhance SAM's segmentation accuracy, particularly through fine-tuning for specific applications in radiology and pathology. This includes examining the integration of retrieval-augmented techniques to improve segmentation precision in medical imaging [11]. Furthermore, the survey addresses challenges related to SAM's tendency to over-segment instances into multiple patches, complicating the identification of patches belonging to the same instance [12].

A significant focus is placed on exploring federated learning strategies for decentralized SAM training, maintaining performance comparable to centralized methods while addressing data privacy and availability concerns. Additionally, the survey investigates the integration of SAM with semi-supervised learning techniques to tackle the limitations of medical image segmentation using scarce labeled data [13].

Moreover, the survey assesses SAM's zero-shot generalization capability and effectiveness across various prompting strategies in medical imaging [1]. By pursuing these objectives, the survey aims to provide comprehensive insights into SAM's current capabilities and limitations, propose solutions for its effective adaptation in medical image segmentation, and identify potential avenues for future research and development, particularly in scenarios with limited exemplars for fine-tuning [14].

1.4 Structure of the Survey

This survey is structured to comprehensively explore SAM and its application in medical image segmentation. It begins with an introduction that emphasizes SAM's significance in image segmentation and the motivations for its use in the medical field. Following this, the core objectives are outlined, establishing a foundation for subsequent sections.

The second section delves into the background and core concepts, providing an overview of vision foundation models, prompt engineering, efficient fine-tuning, and transfer learning relevant to medical imaging. This groundwork is essential for understanding SAM's role in medical image segmentation.

The third section focuses on SAM itself, detailing its architecture, functionality, and capabilities such as zero-shot learning and adaptability, while also discussing its advantages and limitations in medical contexts.

The fourth section examines SAM's application in medical image segmentation across various modalities and the challenges it addresses, including successful implementations and case studies.

The fifth section analyzes prompt engineering and efficient fine-tuning, exploring their roles in enhancing SAM's performance in medical imaging and discussing techniques that improve segmentation accuracy and adaptability.

The sixth section highlights advanced image analysis techniques used with SAM, showcasing methods that enhance the interpretation of complex medical images and improve segmentation outcomes.

The seventh section is dedicated to transfer learning in medical imaging, analyzing its use with SAM to facilitate adaptation to new tasks and datasets, thereby enhancing generalization and performance.

The eighth section identifies current challenges in applying SAM to medical image segmentation and discusses potential future research directions, including improvements in model adaptability and integration with other AI technologies.

Finally, the survey concludes by summarizing key findings and reflecting on SAM's impact on medical image segmentation, emphasizing the importance of ongoing research and development in this field. By categorizing existing research into performance evaluation and methodological

AI-generated, for reference only.

adaptation, the survey provides a structured framework for understanding SAM's application in medical imaging [15].The following sections are organized as shown in Figure 1.

2 Background and Core Concepts

2.1 Vision Foundation Models in Medical Imaging

Vision foundation models like the Segment Anything Model (SAM) have transformed medical imaging by providing a robust framework for image segmentation, significantly reducing the need for manual annotation, which is often error-prone and labor-intensive [16]. SAM, alongside models such as CLIP, enables comprehensive cross-domain representation crucial for interactive and universal segmentation [17]. These models efficiently handle diverse imaging modalities and complex anatomical structures, with SAM demonstrating robust generalization across anatomical sites, enhancing auto-segmentation efficiency in radiation therapy [18]. Its zero-shot generalizability is validated in robotic surgery benchmarks [19].

The adaptability of SAM is highlighted through applications like the Segment Any Class (SAC) method, which automates class-aware prompt generation for multi-class few-shot segmentation without additional training [20]. Federated Foundation Models for Medical Image Segmentation (FedFMS), including FedSAM and FedMSA, aim to improve segmentation performance and training efficiency in federated settings [21]. Despite their benefits, these models face challenges, such as SAM's dataset inadequacies for medical imaging, leading to suboptimal outcomes [22]. Transitioning to SAM2 has revealed limitations in its auto-mode object detection, necessitating performance benchmarks [23]. Comprehensive frameworks assess SAM's performance with varying prompting strategies compared to existing interactive methods [24]. Continued development and evaluation of these models promise further innovation in medical imaging [25].

2.2 Prompt Engineering and Its Relevance

Prompt engineering is pivotal for optimizing SAM, enhancing its ability to interpret and respond to diverse segmentation tasks. Users specify objects of interest through prompts like points, boxes, and masks, facilitating zero-shot segmentation across domains [26, 27]. This technique improves SAM's adaptability and efficiency, particularly in high granularity scenarios and when explicit prompts are absent. Precision in medical imaging is critical, and prompt engineering effectively manages complex anatomical structures and diverse modalities.

The significance of prompt engineering lies in enhancing SAM's zero-shot learning capabilities, allowing segmentation tasks without extensive retraining on new datasets. This is crucial in medical imaging, where labeled data is scarce and costly. By tailoring prompts for specific tasks, SAM's pre-trained knowledge can achieve high accuracy and robustness, as evidenced by benchmark evaluations [25]. Prompt engineering fosters generalization across medical imaging tasks, enabling SAM to adapt to new challenges with minimal training. This adaptability is vital for applications like brain tumor segmentation, where rapid and precise identification of regions of interest significantly impacts clinical outcomes. Models such as SAM and MedSAM demonstrate enhanced accuracy and efficiency across diverse modalities and cancer types, contributing to improved patient care and personalized treatment strategies [28, 10].

2.3 Efficient Fine-Tuning Techniques

Efficient fine-tuning techniques enhance SAM's adaptability and performance in medical image segmentation, achieving accurate outcomes across various modalities, including challenging tasks like brain tumor and breast cancer segmentation [29, 24, 30]. The S-SAM approach exemplifies efficient fine-tuning, training only 0.4% of SAM's parameters using label names as prompts, significantly reducing computational demands [31]. MedSAM further enhances segmentation accuracy for CT-only and fused multimodal images by optimizing specific components without extensive retraining [32]. The Dynamic Parallel Processing Framework (DPPF) optimizes data processing by dynamically distributing tasks across processors, essential for managing large data volumes in medical imaging [33].

While some frameworks, like ProMISe, utilize SAM without requiring fine-tuning, developing efficient fine-tuning techniques remains critical. These methods enhance SAM's adaptability to new tasks and datasets, ensuring effective application across clinical settings [34].

2.4 Transfer Learning in Medical Imaging

Transfer learning is crucial for applying SAM to diverse medical imaging tasks, addressing domain specificity and data scarcity [35]. In medical imaging, where annotated data is resource-intensive, transfer learning enables SAM to leverage pre-trained knowledge, enhancing performance on specific tasks without comprehensive retraining [36]. This approach allows models trained on natural images to adapt to medical imaging's unique requirements [37].

Integrating SAM with advanced models, such as nnUNet, showcases transfer learning's utility by combining SAM's feature extraction with nnUNet's configuration, ensuring adaptability across tasks [38]. The CC-SAM model merges a frozen CNN with a Vision Transformer (ViT), enhancing adaptability to diverse contexts [39]. Benchmarks highlight SAM's zero-shot segmentation capability across modalities, demonstrating transfer learning's effectiveness in bridging natural image training and medical applications [40]. However, existing benchmarks may not adequately evaluate models in specific contexts, such as polyp segmentation, indicating a need for targeted evaluation metrics [41].

Dynamic processing strategies enhance SAM's application in medical imaging tasks, allowing responsive model adaptation [42]. Adapting SAM for 3D imaging, as explored in ReFSAM3D, leverages 3D spatial information through transfer learning, managing complex 3D data crucial for accurate volumetric segmentation [43]. Transfer learning plays a critical role in SAM's successful application to medical imaging, overcoming domain-specific challenges and data scarcity, contributing to accurate diagnosis and treatment planning [44].

3 Segment Anything Model (SAM)

3.1 Architecture and Functionality of SAM

The Segment Anything Model (SAM) employs an advanced encoder-decoder architecture to address diverse image segmentation tasks, integrating prompt engineering to optimize performance across various domains [3]. The encoder extracts detailed features from input images, while the decoder, utilizing composable prompts, generates precise segmentation masks suitable for semantic, instance, and panoptic segmentation [12]. Enhancements like the SAM2-Adapter, which incorporates task-specific knowledge through a multi-adapter setup, further improve segmentation in specialized fields [2]. The CPC-SAM framework, with its dual-branch design, enhances learning and accuracy by generating prompts and supervisions across two decoder branches [13].

SAM's functionality extends to 3D medical imaging by treating sequential 2D slices as video frames, enabling effective volumetric data segmentation [8]. This versatility is crucial for intricate segmentation tasks. Furthermore, SAM supports the extraction of semantic features, significantly reducing communication overhead in systems like the SAM-based Semantic Communication System, thus optimizing data flow management [3].

3.2 Zero-Shot Learning and Adaptability

SAM exemplifies zero-shot learning by performing segmentation tasks without prior examples or retraining, which is particularly beneficial in medical imaging where annotated datasets are scarce [45]. Leveraging pre-trained knowledge, SAM generalizes effectively across various challenges, enhancing its versatility. In medical imaging, SAM facilitates the segmentation of complex anatomical structures with minimal labeled data. Models like MedSAM enhance this adaptability through zero-shot inference, improving accuracy with minimal manual intervention [32]. The Segment Any Class (SAC) method allows the incorporation of new classes without retraining, preserving learned features and preventing catastrophic forgetting [20].

As illustrated in Figure 2, the hierarchical structure of zero-shot learning and adaptability in medical imaging is depicted, highlighting key models, challenges, and innovative solutions across various imaging tasks. Challenges remain, such as in robotic surgical segmentation, where precise instrument segmentation is difficult [19]. Innovations like the SA3D model refine 2D prompts into comprehensive

3D segmentation masks using radiance fields, enhancing SAM's volumetric data analysis [46]. Adapting 3D MRI slices as video frames extends SAM's video segmentation capabilities to 3D imaging, improving efficiency and accuracy [47]. Specific benchmarks highlight the need for SAM to improve performance on specialized medical images [40]. Adaptations such as the CT-SAM3D model, which encodes 3D spatial prompts, demonstrate SAM's potential to generalize across diverse anatomical structures [48].

SAM's zero-shot learning capabilities and adaptability make it valuable for a wide range of imaging tasks, showing competitive performance across modalities like X-ray, ultrasound, and MRI. Its ability to produce high-quality results with minimal prompting alleviates annotation burdens in clinical settings, while ongoing research suggests that fine-tuning SAM can enhance its effectiveness for specific medical applications [49, 50, 51, 52, 53].

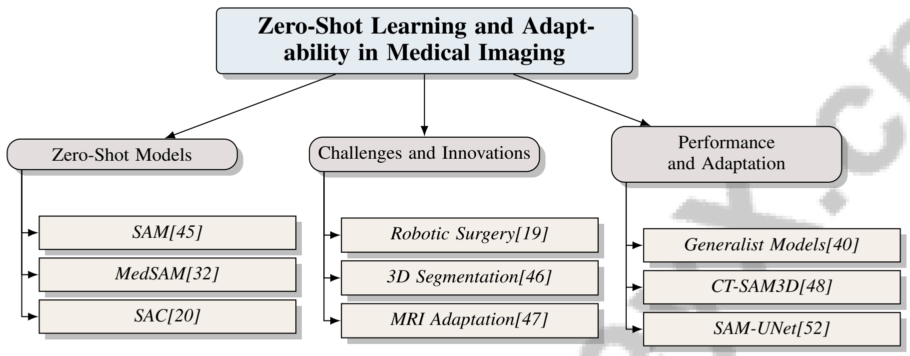


Figure 2: This figure illustrates the hierarchical structure of zero-shot learning and adaptability in medical imaging, highlighting key models, challenges, and innovative solutions across various imaging tasks.

3.3 Advantages in Medical Image Segmentation

SAM offers significant advantages in medical image segmentation, notably enhancing diagnostic accuracy and operational efficiency. Advanced iterations like SAM 2 reduce annotation workloads for medical professionals while maintaining competitive segmentation performance [8]. This efficiency is crucial in clinical environments with limited time and resources. SAM's integration into semantic communication pipelines optimizes resource consumption and negates the need for expert labeling [3]. Combining semantic recognition with instance segmentation enhances accuracy through composable prompts [12]. In data-scarce environments, models like CPC-SAM show significant improvements in semi-supervised segmentation, evidenced by over a 9% increase in Dice score for breast cancer tasks [13].

Fine-tuning strategies consistently outperform traditional methods, especially in complex tasks [9]. Models like AI-SAM, with dual capabilities for automatic and interactive modes, enhance usability and accuracy across medical imaging applications [6]. SAM's adaptability, accuracy, and efficiency position it as a transformative tool in radiology and pathology, contributing to better clinical decision-making and patient care [30, 24, 54].

3.4 Challenges and Limitations in Medical Image Segmentation

Despite advancements, SAM faces challenges in medical image segmentation. Its dependency on the image encoder and decoder's representational power can limit effectiveness, particularly when class labels need predefinition for generating class-specific points [6]. This reliance restricts flexibility in adapting to new classes without manual input. Modality representation imbalance in training datasets can lead to inconsistent performance across imaging types, especially in underrepresented modalities [55]. The labor-intensive nature of data labeling complicates SAM's practical application, as extensive labeled datasets are often necessary for effective fine-tuning [56].

SAM's generalization capabilities are constrained by input data quality and diversity, affecting performance on medical images with significant deviations from training datasets [44]. In low-contrast images or intricate anatomical structures, SAM may struggle without substantial domain-specific knowledge, highlighting the need for refined adaptation strategies [14]. The integration of semantic learning introduces complexity, requiring careful tuning to prevent conflicts between binary and semantic predictions, necessitating more interactions for finer segmentation granularity, especially with incomplete anatomical structures [48]. SAM's limitations are further evident in high-automation tasks like autonomous object detection, where performance may not fully meet application demands [23].

Addressing these challenges is crucial for improving SAM's effectiveness in medical image segmentation, particularly given its varying performance across different datasets and tasks. Enhancements in accuracy and reliability can significantly impact clinical applications, especially in complex areas like radiology and pathology, where precise image segmentation is critical for diagnosis and treatment planning. By refining SAM's capabilities through fine-tuning and adapting it to the unique characteristics of medical images, we can bridge the gap between advanced segmentation techniques and healthcare requirements, unlocking SAM's full potential in automated medical image analysis [57, 15, 24, 30, 53]. This necessitates developing advanced prompt engineering techniques and systematic fine-tuning approaches to optimize performance across diverse medical imaging scenarios.

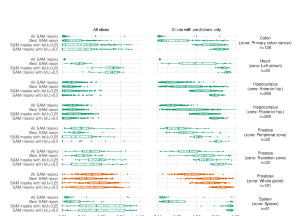
4 Medical Image Segmentation with SAM

4.1 Applications Across Medical Imaging Modalities

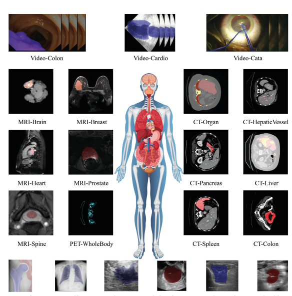
The Segment Anything Model (SAM) demonstrates remarkable versatility and efficacy across various medical imaging modalities, addressing distinct challenges inherent to each. In computed tomography (CT), SAM enhances clinical utility by accurately delineating multi-organ boundaries, exemplified by the CT-SAM3D model's superior performance in volumetric data management [58, 14]. In magnetic resonance imaging (MRI), SAM's zero-shot approach is particularly beneficial for knee segmentation, efficiently treating 3D slices as frames with minimal user interaction [58, 8]. Its broad applicability is evidenced by evaluations on datasets like BraTS2020 and MSD.

Ultrasound imaging benefits from SAM through models like CC-SAM, which effectively tackle segmentation challenges across diverse datasets [14]. In pathology imaging, the Path-SAM2 model, utilizing a dual encoder system, generates precise segmentation masks, highlighting SAM's adaptability [58]. In endoscopic imaging, SAM surpasses baseline methods like U-Net within the ProMISe framework, indicating its potential in this domain [14]. Its application in X-ray imaging, particularly for lung segmentation, underscores its significance where precision and efficiency are crucial [3].

Specialized applications, such as polyp segmentation through Polyp-SAM, demonstrate state-of-the-art performance across multiple datasets [37]. TomoSAM's performance in segmenting tomography datasets within 3D Slicer showcases its capability in complex imaging scenarios, yielding competitive results [14]. SAM's diverse applications across medical imaging modalities highlight its transformative potential in advancing imaging technologies, enabling flexible prompt-based interactions that enhance the segmentation process across various domains [58].



(a) Comparison of Dice scores for different segmentation masks in various anatomical regions[59]



(b) Medical Imaging Techniques and Their Corresponding Body Parts[60]

Algorithm 1: SAM on Zero-Shot Medical Image Segmentation.

Input: Pretrained SAM Model $\theta(\cdot,\cdot)$, contour detector $CD(\cdot)$, midpoint detector $MD(\cdot)$, medical image dataset I with labels of classes C.
Output: Segmentation mask set M.

for $i \in I$ **do**
  for $cls \in C$ **do**
    $i_c \leftarrow (label(i) == cls)$
    $Contours \leftarrow CD(i_c)$
    Initialize image mask m
    for $c \in Contours$ **do**
      $P \leftarrow MD(c)$
      $m_{outs}, scores \leftarrow \theta(i, P)$
      $m \leftarrow Argmax(scores(m_{outs}))$
    end
    M.append(m)
  end
end

(c) Algorithm 1: SAM on Zero-Shot Medical Image Segmentation[40]

Figure 3: Examples of Applications Across Medical Imaging Modalities

AI-generated, for reference only.

As illustrated in Figure 3, medical image segmentation is vital for accurately identifying and analyzing anatomical structures, essential for diagnosis and treatment planning. The figures present diverse applications and effectiveness of segmentation algorithms across various imaging techniques. The first figure compares Dice scores for segmentation masks applied to anatomical regions, showcasing performance in delineating structures such as the colon, heart, and spleen. The second figure links medical imaging techniques, including video colonoscopy, MRI, CT, and ultrasound, to specific body parts, demonstrating the versatility of segmentation algorithms. The third figure details SAM's operational framework for zero-shot medical image segmentation, including a pretrained model and detectors for generating segmentation masks. Together, these figures exemplify the potential of advanced segmentation techniques to enhance medical image analysis across modalities [59, 60, 40].

4.2 Challenges in Complex Anatomical Segmentation

Segmenting complex anatomical structures with SAM involves several challenges critical to its application in medical imaging. The domain gap between dense 2D image data and sparse 3D point cloud data complicates the segmentation of intricate 3D structures, particularly in medical images where transitioning from 2D slices to 3D volumetric data requires sophisticated adaptation strategies [23]. These challenges are intensified in scenarios involving complex anatomical structures, where precise delineation is essential for accurate diagnosis. Figure 4 illustrates the primary challenges in anatomical segmentation, focusing on domain gaps, the balance between texture and shape cues, and innovative approaches for improving segmentation accuracy.

The difficulty of anatomical segmentation is heightened by the need to differentiate between texture and shape cues, crucial for assessing their respective impacts on mask prediction performance. Recent studies indicate that SAM exhibits a bias toward texture, contrasting with human visual perception that prioritizes shape for object recognition. This insight emphasizes the need to disentangle these cues to enhance segmentation effectiveness in both natural and medical imaging contexts, where overlapping structures demand precise delineation for improved clinical outcomes [61, 62, 63, 10]. Addressing this challenge is vital for SAM's ability to interpret the nuanced features of anatomical structures. Furthermore, existing methods often struggle to adapt previously learned features to new tasks, which can degrade performance as the number of classes increases.

Innovative approaches, such as multi-prompt fine-tuning, have been proposed to enhance segmentation accuracy and reduce ambiguity by leveraging multiple prompts. Despite these advancements, benchmarks continue to reveal difficulties in accurately detecting and segmenting visually complex objects, such as transparent and mirror-like structures, due to their reflective and refractive properties. These benchmarks provide a clear framework for evaluating segmentation model performance in critical application areas, facilitating targeted improvements [23].

While SAM shows promise for medical image segmentation, its current limitations in accurately segmenting complex anatomical structures underscore the need for ongoing enhancements in its architectural design and methodological approaches to effectively tackle the unique challenges posed by diverse medical imaging modalities [24, 9, 17, 15, 30]. Developing comprehensive evaluation frameworks and integrating advanced prompting strategies are essential for enhancing SAM's performance across various imaging modalities.

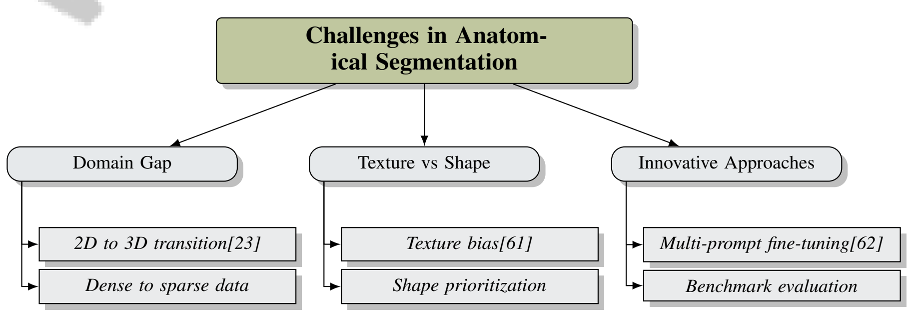


Figure 4: This figure illustrates the primary challenges in anatomical segmentation, focusing on domain gaps, the balance between texture and shape cues, and innovative approaches for improving segmentation accuracy.

AI-generated, for reference only.

4.3 Case Studies and Empirical Evaluations

SAM has undergone extensive empirical evaluations across a variety of medical imaging datasets, demonstrating its adaptability and effectiveness in real-world scenarios. Notable studies have employed public and private datasets, such as the CHAOS and BTCV test sets, to assess SAM's performance against baseline methods like nnUNet. These evaluations indicate SAM's competitive edge in segmenting complex anatomical structures, highlighting its potential in medical imaging applications [64].

The SAM2-Adapter has proven particularly effective in real-world applications, as demonstrated in medical image segmentation tasks using the Kvasir-SEG dataset, where significant improvements in segmentation performance were observed, showcasing the benefits of parameter-efficient fine-tuning [2]. Additionally, SAMAug-C has markedly enhanced medical image classification performance by effectively augmenting raw image inputs and leveraging both augmented and raw images within a novel framework [64].

In whole slide imaging, the WSI-SAM model was evaluated using the dice score metric for zero-shot mask predictions. Comparisons against SAM and MedSAM across different input prompts revealed WSI-SAM's superior performance, particularly in multi-resolution segmentation tasks [44]. Similarly, TomoSAM has shown significant advancements in 3D segmentation, improving both accuracy and efficiency over existing methods [16].

Further evaluations emphasize the enhancements brought by model adaptations like Stable-SAM, which significantly improves segmentation stability and accuracy when dealing with low-quality prompts, outperforming existing methods while maintaining minimal learnable parameters [65]. Additionally, the curriculum prompting approach has demonstrated significant improvements in medical image segmentation performance across multiple datasets, effectively automating prompt generation [66].

These case studies and empirical evaluations underscore SAM's potential to advance medical image segmentation through innovative adaptations and prompt engineering techniques. The model's zero-shot performance is competitive with, and often surpasses, state-of-the-art methods in specific tasks, further solidifying its role as a transformative tool in medical imaging [67].

5 Prompt Engineering and Efficient Fine-Tuning

The optimization of model performance is paramount in AI and machine learning, especially in medical imaging. This section explores the pivotal role of prompt engineering in enhancing the Segment Anything Model (SAM). Strategic prompting techniques are crucial for achieving superior segmentation accuracy, vital for medical diagnostics and treatment planning. The following discussion elucidates the impact of prompt engineering on SAM's performance across various medical imaging scenarios.

5.1 Role of Prompt Engineering in SAM

Prompt engineering significantly enhances SAM's segmentation capabilities in medical imaging. Strategies such as object consistency and boundary preservation losses markedly improve SAM's semantic segmentation performance, leveraging raw outputs for more precise results [68]. This is essential for accurately delineating anatomical structures in medical images. Figure 5 illustrates the role of prompt engineering in enhancing SAM's segmentation capabilities, highlighting strategies for improvement, innovative frameworks, and adaptation techniques in medical imaging.

Adaptive prompt mechanisms, like the ProMISe framework, facilitate effective SAM transfer to medical image segmentation by generating adaptive prompts and modifying segmentation patterns [34]. These strategies enable SAM to meet medical imaging demands, enhancing accuracy and efficiency.

Innovations such as Path-SAM2 automate prompting through modules like the KAN classification module, improving segmentation accuracy while reducing manual intervention [69]. Automation is vital in medical imaging, where manual annotation is resource-intensive.

The emergence of 3D promptable models, such as CT-SAM3D, underscores prompt engineering's significance in augmenting SAM's capabilities. Utilizing comprehensive datasets, these models enhance segmentation accuracy and efficiency, addressing challenges posed by complex anatomical structures [48].

Additionally, fine-tuning SAM with exemplar-guided synthesis and Low-Rank Adaptation (LoRA) allows effective segmentation with minimal labeled data, showcasing prompt engineering's flexibility in optimizing SAM's performance [56]. This approach is beneficial in medical imaging, where labeled datasets are scarce.

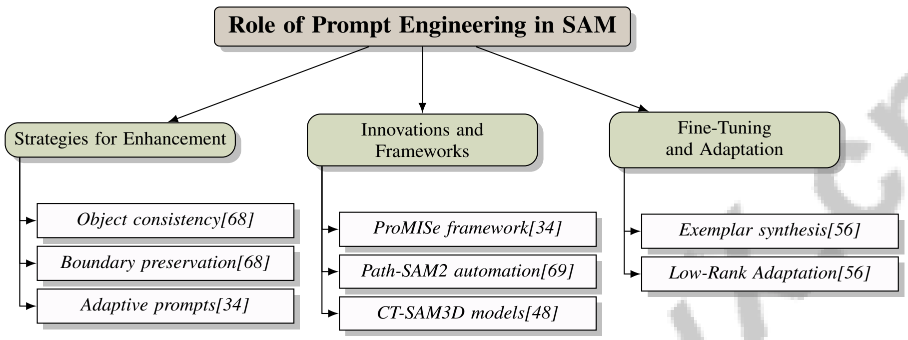


Figure 5: This figure illustrates the role of prompt engineering in enhancing SAM's segmentation capabilities, highlighting strategies for improvement, innovative frameworks, and adaptation techniques in medical imaging.

5.2 Enhancing Segmentation Accuracy

Efficient fine-tuning techniques are crucial for improving SAM's segmentation accuracy in medical imaging. One notable method is Ladder Fine-tuning, integrating a complementary CNN with SAM's architecture, significantly reducing training time and resource costs while enhancing segmentation accuracy across diverse tasks [70].

The AutoProSAM method represents another efficient fine-tuning strategy, focusing on enhancing segmentation accuracy in 3D medical imaging by automating the prompting process, optimizing SAM's performance while minimizing manual effort [71]. Parameter-efficient strategies, as demonstrated in remote sensing, underscore SAM's adaptability to various imaging contexts without compromising accuracy [5].

Novel metrics like the Contour Pixel Rate and Difference of Gini Impurity Deviation provide a robust framework for evaluating segmentation performance, effectively quantifying the tree-likeness and textural separability of segmented regions, offering insights into SAM's capabilities and areas for improvement [72]. The SAMPOT framework evolves prompt locations based on a scoring system predicting segmentation quality, aiming to maximize performance metrics like the Dice coefficient, essential for precise segmentation [73].

RobustSAM enhances segmentation accuracy by extracting features invariant to image degradation, ensuring consistent performance across varying image qualities [74]. This robustness is critical in medical imaging, where image quality can fluctuate. Benchmark evaluations of SAM in robotic surgery provide valuable insights into its performance and limitations, indicating areas for refinement [19].

Future research should prioritize advancing few-shot fine-tuning techniques to overcome existing limitations and further enhance SAM's performance in diverse medical imaging scenarios [75]. By leveraging efficient fine-tuning techniques and innovative evaluation metrics, SAM's segmentation accuracy can be significantly improved, addressing the complexities of medical imaging tasks and contributing to better clinical outcomes.

Category	Feature	Method
Integration of Advanced Models and Methods	Model Integration	nnSAM[38]
Adaptation of SAM for 3D Medical Images	3D Spatial Techniques	SAM2[47], RS3D[43], SA3D[46], CT-SAM3D[48]
Hybrid Architecture for Feature Extraction	Hybrid Extraction Techniques	SAM-SCS[3], CC-SAM[39], CPC-SAM[13]

Table 1: This table provides a comprehensive summary of the integration of advanced models and methods with the Segment Anything Model (SAM) in medical imaging. It categorizes the methods into three key areas: integration of advanced models, adaptation of SAM for 3D medical images, and hybrid architecture for feature extraction, highlighting the specific techniques and models employed in each category.

6 Advanced Image Analysis Techniques

The advancement of image analysis techniques is vital in medical imaging, where precision and efficiency are paramount. This section delves into the integration of sophisticated models and methodologies that enhance existing frameworks, particularly focusing on the Segment Anything Model (SAM). Table 1 presents a detailed overview of the various advanced models and methods integrated with the Segment Anything Model (SAM) to enhance its performance in medical imaging applications. By exploring the synergies between SAM and various advanced models, we gain insights into their contributions to improved segmentation outcomes. The following subsection discusses the integration of these models and methods, demonstrating their impact on SAM's performance in medical image analysis.

6.1 Integration of Advanced Models and Methods

Integrating advanced models with the Segment Anything Model (SAM) represents a significant leap in image analysis, especially for interpreting complex medical images. This integration harnesses the strengths of various segmentation models, enhancing accuracy and efficiency across diverse medical imaging scenarios, while addressing challenges like data annotation scarcity and performance variability across datasets [24, 62].

One approach involves combining SAM with the nnUNet framework, a well-established model in medical image segmentation, which enhances SAM's adaptability and performance through automatic configuration of segmentation tasks, enabling it to effectively address various medical imaging challenges [38]. This synergy facilitates more precise and efficient segmentation outcomes.

The CPC-SAM framework, featuring a dual-branch design, enhances SAM's learning by generating prompts and supervision across two decoder branches, leveraging composable prompts to optimize performance in complex segmentation tasks [13]. Hybrid architectures like the CC-SAM model further demonstrate the potential of integrating advanced models with SAM. By combining a frozen Convolutional Neural Network (CNN) with a Vision Transformer (ViT), CC-SAM improves SAM's adaptability to various medical imaging contexts, enhancing segmentation capabilities [39].

Moreover, integrating SAM with semantic communication systems optimizes segmentation processes by reducing communication overhead and enhancing data flow efficiency, which is critical in medical imaging for resource utilization [3]. The integration of advanced models and methods with SAM significantly enhances its capabilities, offering improved segmentation accuracy and efficiency across a range of medical imaging applications. SAM's transformative potential is evident in critical areas such as brain tumor and breast cancer segmentation. Evaluations across diverse datasets highlight SAM's impressive zero-shot segmentation capabilities, although performance varies with dataset and task. Techniques like Auxiliary Online Learning further enhance SAM's adaptability, bridging the gap between advanced segmentation technology and healthcare diagnostics requirements [30, 24, 76, 53].

6.2 Adaptation of SAM for 3D Medical Images

Adapting the Segment Anything Model (SAM) for 3D medical image analysis signifies a substantial evolution in its application, addressing the complexities of volumetric data. This adaptation involves transforming SAM's architecture to effectively handle 3D spatial information, essential for accurate segmentation in medical contexts [43]. SAM processes sequential 2D slices as video frames, extending its video segmentation capabilities to 3D medical imaging, capturing intricate anatomical details [8].

Models like ReFSAM3D exemplify successful adaptation, leveraging radiance fields to refine 2D prompts into comprehensive 3D segmentation masks, thus improving the analysis of volumetric medical data [46]. This approach enhances SAM's ability to manage complex anatomical structures, boosting segmentation accuracy and efficiency in 3D tasks.

The CT-SAM3D model further demonstrates SAM's adaptation for 3D imaging by encoding 3D spatial prompts and learning from rich annotations, enabling generalization across diverse anatomical structures [48]. This adaptability is crucial for addressing the unique challenges posed by different medical imaging modalities, such as MRI and CT scans.

The benefits of adapting SAM for 3D medical images include improved segmentation accuracy and efficiency, which are vital for clinical decision-making and treatment planning. By leveraging 3D spatial information, SAM provides detailed and accurate segmentation results, contributing to better patient outcomes [47].

The development of the Medical SAM Adapter (Med-SA) highlights SAM's transformative potential in medical imaging. Med-SA integrates domain-specific medical knowledge and employs innovative techniques such as Space-Depth Transpose (SD-Trans) and Hyper-Prompting Adapter (HyP-Adpt) to enhance segmentation precision and efficiency. Comprehensive evaluations across 17 medical image segmentation tasks reveal that Med-SA outperforms several state-of-the-art methods while updating only 2% of SAM's parameters, indicating significant advancements applicable across various clinical applications [77, 24].

6.3 Hybrid Architecture for Feature Extraction

The development of hybrid architectures within the Segment Anything Model (SAM) framework significantly enhances feature extraction capabilities, optimizing performance across diverse medical imaging tasks. These architectures combine various model components, such as SAM and the Mixture of Low-Rank Adaptation Experts (MoE-LoRA), to leverage their strengths in improving segmentation accuracy and efficiency. By integrating techniques like multi-scale feature extraction and adaptive routing for multi-modal data, these models address challenges such as cross-modal inconsistencies and high granularity requirements, resulting in notable gains in segmentation tasks, particularly in scenarios with missing modalities [26, 78].

A notable example is the integration of Convolutional Neural Networks (CNNs) with Vision Transformers (ViTs) within the SAM framework. This combination, as demonstrated in models like CC-SAM, capitalizes on the spatial feature extraction capabilities of CNNs and the global context modeling of ViTs, enhancing SAM's adaptability to various medical imaging contexts [39]. This synergy allows for more precise segmentation outcomes, particularly in complex anatomical structures requiring detailed feature extraction.

The dual-branch designs, such as those in the CPC-SAM framework, enhance learning by generating prompts and supervision across two decoder branches, optimizing information flow between components and improving segmentation accuracy [13]. This approach ensures SAM effectively manages the complexities of medical imaging tasks, delivering robust and reliable segmentation results.

Furthermore, integrating advanced semantic communication systems within SAM optimizes feature extraction by reducing communication overhead and enhancing data flow efficiency [3], a critical advantage in medical imaging for efficient resource utilization.

The integration of hybrid architectures within the Segment Anything Model (SAM) markedly enhances feature extraction capabilities, resulting in improved segmentation accuracy and efficiency across various medical imaging applications, including critical areas like brain tumor and breast cancer segmentation. This advancement is particularly significant given the challenges posed by limited data annotations in medical imaging, allowing SAM to effectively address these issues and outperform comparable segmentation methods in many scenarios [30, 24]. These innovations underscore SAM's potential to transform medical imaging, providing accurate and efficient solutions for complex segmentation challenges.

In recent years, transfer learning has emerged as a pivotal approach in the field of medical imaging, significantly enhancing the performance of various diagnostic tasks. To elucidate this concept, Figure 6 illustrates the hierarchical structure of transfer learning as applied to medical imaging through the Segment Anything Model (SAM). This figure categorizes the enhancements and techniques introduced by SAM, detailing its application to specific medical imaging tasks, such as polyp and brain tumor segmentation. Furthermore, it underscores the model's generalization capabilities across diverse datasets. Notably, the figure highlights task-specific adaptations and emphasizes SAM's integration with other models, such as nnUNet, thereby showcasing its broad applicability and transformative potential in the realm of medical imaging.

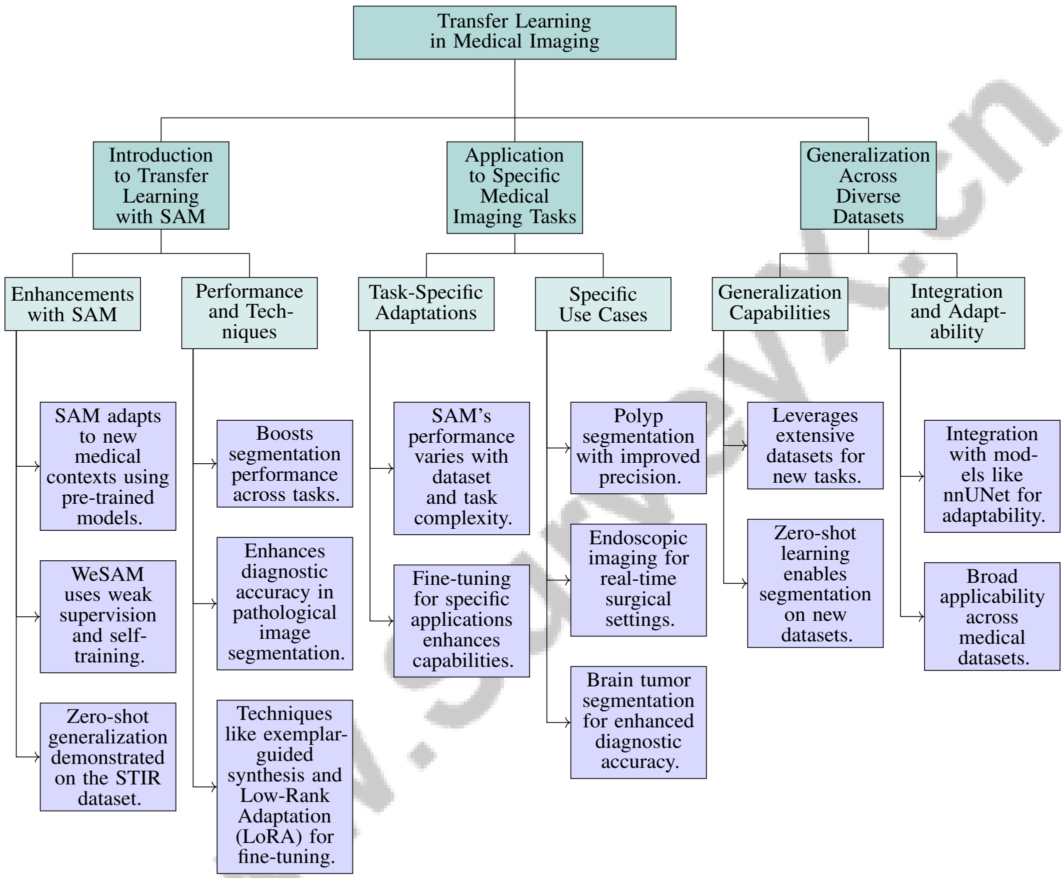


Figure 6: This figure illustrates the hierarchical structure of transfer learning in medical imaging using the Segment Anything Model (SAM). It categorizes the enhancements and techniques in SAM's introduction, its application to specific medical imaging tasks, and its generalization across diverse datasets. The figure highlights task-specific adaptations, such as polyp and brain tumor segmentation, and emphasizes SAM's integration with models like nnUNet, showcasing its broad applicability and transformative potential in medical imaging.

7 Transfer Learning in Medical Imaging

7.1 Introduction to Transfer Learning with SAM

Transfer learning enhances the Segment Anything Model (SAM) in medical imaging, especially where labeled data is scarce. By leveraging pre-trained models on large datasets, SAM adapts effectively to new medical contexts, overcoming traditional segmentation limitations [79]. WeSAM exemplifies this by using weak supervision and self-training to adjust SAM to novel data distributions without relying on source domain data, crucial for maintaining high segmentation accuracy across diverse scenarios [80]. SAM's zero-shot generalization, demonstrated on the label-free STIR dataset,

AI-generated, for reference only.

highlights its potential in real-time surgical environments and other medical imaging applications [81].

Transfer learning with SAM significantly boosts segmentation performance across varied tasks and datasets, addressing domain specificity and data scarcity challenges. It enhances diagnostic accuracy in pathological image segmentation, such as colon polyps and brain tumors. Models like Path-SAM2 and Polyp-SAM achieve state-of-the-art semantic segmentation performance, evidenced by high dice scores across multiple datasets. Techniques such as exemplar-guided synthesis and Low-Rank Adaptation (LoRA) facilitate SAM's fine-tuning with minimal exemplars, making it highly applicable in real-world medical scenarios [69, 37, 56].

7.2 Application to Specific Medical Imaging Tasks

SAM's transfer learning application across 19 medical imaging datasets showcases its adaptability and effectiveness in clinical challenges. Its segmentation performance varies with dataset characteristics and task complexity, achieving an IoU score of 0.8650 for hip X-rays but only 0.1135 for spine MRIs, influenced by object boundary complexity and prompt type. Fine-tuning SAM for specific applications enhances its segmentation capabilities, offering a promising tool for automated medical image analysis [29, 24, 79, 28].

In polyp segmentation, SAM leverages transfer learning to improve precision amid shape and size variability [41]. Similarly, in endoscopic imaging, transfer learning enables accurate segmentation of complex structures in real-time surgical settings [81]. For brain tumor segmentation, SAM adapts pre-trained models to intricate tumor boundaries, enhancing diagnostic accuracy crucial for treatment planning [79].

The integration of SAM with advanced models like nnUNet through transfer learning exemplifies its application in specific tasks, leveraging nnUNet's automatic configuration to improve multi-organ segmentation in CT and MRI scans [38]. By combining strengths, SAM achieves superior outcomes in complex scenarios. This underscores SAM's potential to revolutionize clinical practice by providing accurate segmentation solutions, enhancing diagnostic precision and clinical outcomes, especially for critical conditions like brain tumors and lung abnormalities [30, 24, 79, 76].

7.3 Generalization Across Diverse Datasets

SAM's ability to generalize across diverse datasets enhances its usability in medical imaging. Transfer learning allows SAM to leverage extensive, diverse datasets, improving performance on new tasks [79]. This adaptation is vital in medical imaging, where labeled data is limited and costly. SAM's zero-shot learning enables segmentation on new datasets without prior exposure, supported by its robust architecture and strategic prompt engineering, maintaining high accuracy across modalities and scenarios [81].

SAM's integration with models like nnUNet further illustrates its generalization capability, automatically configuring for various tasks and enhancing adaptability to new datasets [38]. By leveraging multiple models' strengths, SAM excels in complex imaging scenarios, solidifying its transformative role in medical imaging.

SAM's broad applicability across medical datasets underscores its potential in the field. It offers a flexible, interactive segmentation approach, delivering accurate results for well-defined structures. However, performance varies with dataset and object complexity, necessitating careful adaptation in challenging scenarios [29, 24, 62, 17]. Through strategic transfer learning, SAM expands its capabilities, improving diagnostic precision and clinical outcomes across a spectrum of medical applications.

8 Challenges and Future Directions

8.1 Domain Specificity and Data Challenges

The Segment Anything Model (SAM) encounters notable challenges in medical imaging, primarily due to domain specificity and data constraints. A significant issue is the reliance on fully labeled datasets, which are often limited in clinical settings, especially for rare or complex anatomical structures [13]. This scarcity, combined with the structural complexity and variability of medical images compared to natural images, necessitates extensive fine-tuning across different contexts, which can be resource-intensive [9]. The imbalance in modality representation within training datasets further complicates SAM's performance, particularly on less-represented modalities [3].

Challenges related to domain specificity are pronounced in the segmentation of intricate anatomical structures or low-contrast images, where SAM may struggle without predefined bounding boxes or clear anatomical boundaries. This issue is exacerbated by the variability in segmentation tasks across different medical imaging modalities, hindering the model's adaptability without extensive domain-specific knowledge [23]. Existing benchmarks often lack generalizability across diverse datasets and recording setups, leading to suboptimal performance in tasks like gaze estimation, where precise segmentation is essential [45].

Data limitations significantly impede progress, as the scarcity of high-quality labeled medical data complicates the training of supervised models and constrains the effectiveness of few-shot learning approaches. SAM's retrieval processes may falter with low-contrast images and small anatomical structures, where capturing fine details is critical for accurate segmentation. Additionally, the extensive training time required for certain methods underscores the need for improvements in model efficiency to better accommodate the demands of medical imaging [13].

Future research should focus on integrating more unlabeled data with few labeled examples to enhance model performance and exploring advanced algorithms to improve model robustness in non-IID settings. Refining the role of negative prompt points could also clarify benefits in segmentation tasks, further optimizing SAM's performance in medical imaging. By systematically addressing these challenges, researchers can enhance SAM's generalizability across varied patient populations and mitigate biases inherent in training data, ultimately improving its performance in complex scenarios such as robotic surgery and medical imaging [30, 24, 19, 82].

8.2 Model Adaptability and Performance

SAM's adaptability and performance in medical imaging are influenced by several factors that underscore both its potential and limitations. A significant challenge arises from SAM's initial training, which excluded medical images, potentially impacting its segmentation quality for complex anatomical structures. This limitation highlights the necessity for domain-specific adaptations to better align SAM with the unique characteristics of medical imaging modalities, such as MRI and CT scans [55].

Performance issues become especially apparent in challenging conditions, such as low-contrast environments and intricate anatomical features, where SAM's methodologies may not adequately address the complexities involved [24]. The quality of ground-truth masks and the inherent intricacies of medical images further complicate SAM's ability to maintain high segmentation accuracy [83]. Enhancements to SAM's architecture, such as integrating 3D data and leveraging multimodal imaging information, could be explored to address these challenges [1].

Moreover, SAM's adaptability is constrained by the simplicity of its pixel classification strategies in uncertain regions, which may not effectively capture the complexity of certain anatomical structures [37]. This limitation emphasizes the need for more sophisticated annotation techniques and iterative refinement processes to enhance segmentation accuracy and robustness [2]. While the Ladder Fine-tuning approach improves performance, it still encounters challenges in generalizing to other medical datasets, indicating a need for further optimization [5].

The dependency on the quality of bounding box predictions from the object detection model also presents a limitation, as inaccuracies in these predictions can adversely affect overall mask quality [7]. Future research should prioritize refining SAM's generalization capabilities to diverse patient populations and enhancing its robustness against biases in training data [55]. Additionally, addressing segmentation accuracy for split vessels and exploring advanced annotation techniques could significantly enhance SAM's performance in medical imaging [65]. By tackling these areas, future research can optimize SAM's utility in medical imaging, ensuring it meets the demands of varied and complex clinical environments.

8.3 Integration with Other AI Technologies

Integrating SAM with advanced AI technologies offers significant potential to enhance its capabilities in medical image segmentation, particularly in addressing challenges posed by complex anatomical structures, varying object boundaries, and diverse imaging modalities. Recent studies indicate that while SAM demonstrates impressive performance in certain medical imaging tasks, its effectiveness can be augmented through techniques such as fine-tuning and specific prompting strategies, optimizing segmentation accuracy across a broader range of medical datasets [79, 29, 17, 28, 24]. This combined approach could yield more reliable and efficient automated segmentation solutions in clinical applications, thereby supporting improved diagnostic and treatment planning processes.

One promising avenue involves integrating SAM with deep learning models that specialize in specific aspects of image analysis, such as object detection and classification. For instance, combining SAM with convolutional neural networks (CNNs) can enhance its ability to capture fine-grained details, improving the accuracy of segmentation tasks involving complex anatomical structures [38]. This synergy allows SAM to leverage CNNs' strengths in feature extraction, optimizing performance in challenging medical imaging scenarios.

Furthermore, incorporating Vision Transformers (ViTs) alongside SAM can enhance adaptability and efficiency. ViTs excel in capturing global contextual information, crucial for accurately segmenting large and intricate anatomical regions [39]. Integrating ViTs with SAM can lead to a more comprehensive understanding of medical images, resulting in improved segmentation outcomes.

The use of federated learning frameworks in conjunction with SAM presents another opportunity for advancement. Federated learning enables decentralized training of models across multiple institutions, preserving data privacy while enhancing model robustness and generalization [21]. This approach is particularly beneficial in medical imaging, where data privacy and security are paramount.

Additionally, integrating SAM with semantic communication systems can optimize data flow and reduce communication overhead, improving segmentation process efficiency in resource-constrained environments [3]. This capability is essential for real-time applications, such as robotic-assisted surgery, where timely and accurate segmentation is critical.

The integration of SAM with other AI technologies holds promise for enhancing its capabilities in medical image segmentation, particularly in addressing challenges related to limited data annotations and varying performance across different medical imaging datasets. SAM has shown impressive zero-shot segmentation performance for well-defined objects like hip X-rays, while struggling with complex cases such as brain tumors. By fine-tuning SAM and exploring its interactions with additional AI frameworks, researchers can potentially improve segmentation accuracy and reliability in critical areas like radiology and pathology, ultimately advancing automated medical image analysis and diagnosis [24, 30]. Leveraging the strengths of complementary models and techniques can enhance SAM's accuracy, efficiency, and adaptability, ultimately contributing to improved clinical outcomes and advancing the field of medical imaging.

8.4 Future Research Directions

Future research on SAM in medical image segmentation should prioritize developing large-scale medical datasets to enhance model robustness and improve annotation efficiency. This will facilitate more accurate and efficient segmentation outcomes, addressing current limitations in diverse medical imaging contexts [9]. Expanding SAM's capabilities to include multi-modal image segmentation can further enhance its adaptability and effectiveness across diverse imaging scenarios [8]. Additionally, exploring fine-tuning strategies and developing comprehensive benchmarks for medical image segmentation are crucial for optimizing SAM's performance [9].

Refining prompting strategies is another key area for future research, focusing on adapting SAM for specific medical imaging tasks, including 3D segmentation [8]. Extending SAM's capabilities to 3D data and integrating information from multiple imaging modalities will be essential for advancing its application in complex medical imaging scenarios [8]. Moreover, future research should explore extending open-vocabulary segmentation methods to enhance SAM's adaptability to new classes and improve background point generation [6].

Improving SAM's robustness in challenging scenarios and extending benchmarks to cover more diverse datasets are vital steps toward enhancing its applicability in real-world clinical settings

[9]. Developing application-oriented SAM models and enhancing pretraining strategies will further improve its utility across all surveyed domains [7]. Additionally, research should focus on refining SAM2's capabilities in 3D medical image segmentation to enhance generalizability and accuracy [8].

Finally, exploring the scalability of SAM-based methods and their applicability to a wider range of tasks beyond image transmission could significantly broaden SAM's utility in medical imaging and other domains [7]. Future work should also aim to improve patch generation quality from SAM and explore more efficient vision foundation models to enhance SAM-CP performance [12]. By addressing these research directions, SAM can be optimized to meet the demands of complex medical imaging scenarios, ultimately advancing its role in clinical practice.

9 Conclusion

The Segment Anything Model (SAM) represents a significant leap forward in the realm of medical image segmentation, offering notable enhancements in both accuracy and operational efficiency across a wide array of imaging modalities. This survey underscores the necessity for strategic adaptations to fully leverage SAM's capabilities within medical settings, with particular attention to SAM2's role in refining segmentation processes and enhancing the precision of clinical workflows. The integration of systematic fine-tuning and targeted methodologies has yielded substantial advancements in segmentation outcomes, providing valuable insights for future research trajectories.

The development of sophisticated techniques such as AutoProSAM has further expanded SAM's utility, particularly in the domain of 3D multi-organ segmentation, where it has outperformed current leading models. These innovations highlight SAM's transformative impact on medical imaging, offering robust solutions to complex segmentation challenges.

To fully realize SAM's potential in clinical environments, continued research and development are imperative. Future efforts should focus on augmenting SAM's adaptability to diverse imaging tasks, refining fine-tuning processes, and exploring novel methods to optimize its performance. By addressing these areas, SAM can significantly contribute to the advancement of medical imaging, facilitating the creation of more precise and efficient diagnostic tools that enhance clinical decision-making and improve patient outcomes.

References

[1] Peng Zhang and Yaping Wang. Segment anything model for brain tumor segmentation. *arXiv preprint arXiv:2309.08434*, 2023.

[2] Tianrun Chen, Ankang Lu, Lanyun Zhu, Chaotao Ding, Chunan Yu, Deyi Ji, Zejian Li, Lingyun Sun, Papa Mao, and Ying Zang. Sam2-adapter: Evaluating & adapting segment anything 2 in downstream tasks: Camouflage, shadow, medical image segmentation, and more. *arXiv preprint arXiv:2408.04579*, 2024.

[3] Shehbaz Tariq, Brian Estadimas Arfeto, Chaoning Zhang, and Hyundong Shin. Segment anything meets semantic communication, 2023.

[4] Nikhila Ravi, Valentin Gabeur, Yuan-Ting Hu, Ronghang Hu, Chaitanya Ryali, Tengyu Ma, Haitham Khedr, Roman Rädle, Chloe Rolland, Laura Gustafson, Eric Mintun, Junting Pan, Kalyan Vasudev Alwala, Nicolas Carion, Chao-Yuan Wu, Ross Girshick, Piotr Dollár, and Christoph Feichtenhofer. Sam 2: Segment anything in images and videos, 2024.

[5] Linghao Zheng, Xinyang Pu, and Feng Xu. Tuning a sam-based model with multi-cognitive visual adapter to remote sensing instance segmentation, 2024.

[6] Yimu Pan, Sitao Zhang, Alison D. Gernand, Jeffery A. Goldstein, and James Z. Wang. Ai-sam: Automatic and interactive segment anything model, 2023.

[7] Chaoning Zhang, Dongshen Han, Sheng Zheng, Jinwoo Choi, Tae-Ho Kim, and Choong Seon Hong. Mobilesamv2: Faster segment anything to everything, 2023.

[8] Chuyun Shen, Wenhao Li, Yuhang Shi, and Xiangfeng Wang. Interactive 3d medical image segmentation with sam 2, 2025.

[9] Hanxue Gu, Haoyu Dong, Jichen Yang, and Maciej A Mazurowski. How to build the best medical image segmentation algorithm using foundation models: a comprehensive empirical study with segment anything model. *arXiv preprint arXiv:2404.09957*, 2024.

[10] Jun Ma, Yuting He, Feifei Li, Lin Han, Chenyu You, and Bo Wang. Segment anything in medical images. *Nature Communications*, 15(1):654, 2024.

[11] Lin Zhao, Xiao Chen, Eric Z. Chen, Yikang Liu, Terrence Chen, and Shanhui Sun. Retrieval-augmented few-shot medical image segmentation with foundation models, 2024.

[12] Pengfei Chen, Lingxi Xie, Xinyue Huo, Xuehui Yu, Xiaopeng Zhang, Yingfei Sun, Zhenjun Han, and Qi Tian. Sam-cp: Marrying sam with composable prompts for versatile segmentation, 2024.

[13] Juzheng Miao, Cheng Chen, Keli Zhang, Jie Chuai, Quanzheng Li, and Pheng-Ann Heng. Cross prompting consistency with segment anything model for semi-supervised medical image segmentation, 2024.

[14] Wei Ji, Jingjing Li, Qi Bi, Tingwei Liu, Wenbo Li, and Li Cheng. Segment anything is not always perfect: An investigation of sam on different real-world applications, 2023.

[15] Yichi Zhang and Rushi Jiao. Towards segment anything model (sam) for medical image segmentation: A survey, 2023.

[16] Federico Semeraro, Alexandre Quintart, Sergio Fraile Izquierdo, and Joseph C. Ferguson. Tomosam: a 3d slicer extension using sam for tomography segmentation, 2023.

[17] Yichi Zhang and Rushi Jiao. Towards segment anything model (sam) for medical image segmentation: a survey. *arXiv preprint arXiv:2305.03678*, 2023.

[18] Lian Zhang, Zhengliang Liu, Lu Zhang, Zihao Wu, Xiaowei Yu, Jason Holmes, Hongying Feng, Haixing Dai, Xiang Li, Quanzheng Li, Dajiang Zhu, Tianming Liu, and Wei Liu. Segment anything model (sam) for radiation oncology, 2023.

[19] An Wang, Mobarakol Islam, Mengya Xu, Yang Zhang, and Hongliang Ren. Sam meets robotic surgery: An empirical study on generalization, robustness and adaptation, 2023.

[20] Hussni Mohd Zakir and Eric Tatt Wei Ho. Segment any class (sac): Multi-class few-shot semantic segmentation via class region proposals, 2024.

[21] Yuxi Liu, Guibo Luo, and Yuesheng Zhu. Fedfms: Exploring federated foundation models for medical image segmentation, 2024.

[22] Aimee Guo, Grace Fei, Hemanth Pasupuleti, and Jing Wang. Clicksam: Fine-tuning segment anything model using click prompts for ultrasound image segmentation, 2024.

[23] Lv Tang and Bo Li. Evaluating sam2's role in camouflaged object detection: From sam to sam2, 2024.

[24] Maciej A. Mazurowski, Haoyu Dong, Hanxue Gu, Jichen Yang, Nicholas Konz, and Yixin Zhang. Segment anything model for medical image analysis: an experimental study, 2023.

[25] Mehri Mehrnia, Mohamed Elbayumi, and Mohammed S. M. Elbaz. Assessing foundational medical 'segment anything' (med-sam1, med-sam2) deep learning models for left atrial segmentation in 3d lge mri, 2024.

[26] Chaoning Zhang, Joseph Cho, Fachrina Dewi Puspitasari, Sheng Zheng, Chenghao Li, Yu Qiao, Taegoo Kang, Xinru Shan, Chenshuang Zhang, Caiyan Qin, Francois Rameau, Lik-Hang Lee, Sung-Ho Bae, and Choong Seon Hong. A survey on segment anything model (sam): Vision foundation model meets prompt engineering, 2024.

[27] Chaoning Zhang, Fachrina Dewi Puspitasari, Sheng Zheng, Chenghao Li, Yu Qiao, Taegoo Kang, Xinru Shan, Chenshuang Zhang, Caiyan Qin, Francois Rameau, et al. A survey on segment anything model (sam): Vision foundation model meets prompt engineering. *arXiv preprint arXiv:2306.06211*, 2023.

[28] Evan Kellener, Ihina Nath, An Ngo, Thomas Nguyen, Joshua Schuman, Coen Adler, and Arnav Kartikeya. Utilizing segment anything model for assessing localization of grad-cam in medical imaging, 2023.

[29] Yuhao Huang, Xin Yang, Lian Liu, Han Zhou, Ao Chang, Xinrui Zhou, Rusi Chen, Junxuan Yu, Jiongquan Chen, Chaoyu Chen, Sijing Liu, Haozhe Chi, Xindi Hu, Kejuan Yue, Lei Li, Vicente Grau, Deng-Ping Fan, Fajin Dong, and Dong Ni. Segment anything model for medical images?, 2024.

[30] Amin Ranem, Niklas Babendererde, Moritz Fuchs, and Anirban Mukhopadhyay. Exploring sam ablations for enhancing medical segmentation in radiology and pathology, 2023.

[31] Jay N. Paranjape, Shameema Sikder, S. Swaroop Vedula, and Vishal M. Patel. S-sam: Svd-based fine-tuning of segment anything model for medical image segmentation, 2024.

[32] Jintao Ren, Mathis Rasmussen, Jasper Nijkamp, Jesper Grau Eriksen, and Stine Korreman. Segment anything model for head and neck tumor segmentation with ct, pet and mri multi-modality images, 2024.

[33] Yiming Zhao, Tao Zhou, Yunqi Gu, Yi Zhou, Yizhe Zhang, Ye Wu, and Huazhu Fu. Segment anything model-guided collaborative learning network for scribble-supervised polyp segmentation, 2023.

[34] Jinfeng Wang, Sifan Song, Xinkun Wang, Yiyi Wang, Yiyi Miao, Jionglong Su, and S. Kevin Zhou. Promise: Promptable medical image segmentation using sam, 2024.

[35] Jieming Yu, An Wang, Wenzhen Dong, Mengya Xu, Mobarakol Islam, Jie Wang, Long Bai, and Hongliang Ren. Sam 2 in robotic surgery: An empirical evaluation for robustness and generalization in surgical video segmentation, 2024.

[36] Hanxue Gu, Haoyu Dong, Jichen Yang, and Maciej A. Mazurowski. How to build the best medical image segmentation algorithm using foundation models: a comprehensive empirical study with segment anything model, 2024.

[37] Yuheng Li, Mingzhe Hu, and Xiaofeng Yang. Polyp-sam: Transfer sam for polyp segmentation, 2023.

[38] Yunxiang Li, Bowen Jing, Zihan Li, Jing Wang, and You Zhang. nnsam: Plug-and-play segment anything model improves nnunet performance, 2024.

[39] Shreyank N Gowda and David A Clifton. Cc-sam: Sam with cross-feature attention and context for ultrasound image segmentation. In *European Conference on Computer Vision*, pages 108–124. Springer, 2024.

[40] Peilun Shi, Jianing Qiu, Sai Mu Dalike Abaxi, Hao Wei, Frank P. W. Lo, and Wu Yuan. Generalist vision foundation models for medical imaging: A case study of segment anything model on zero-shot medical segmentation, 2023.

[41] Tao Zhou, Yizhe Zhang, Yi Zhou, Ye Wu, and Chen Gong. Can sam segment polyps?, 2023.

[42] Hyung-Il Kim, Kimin Yun, Jun-Seok Yun, and Yuseok Bae. Task-specific adaptation of segmentation foundation model via prompt learning, 2024.

[43] Xiang Gao and Kai Lu. Refsam3d: Adapting sam with cross-modal reference for 3d medical image segmentation, 2024.

[44] Xiaobao Wei, Jiajun Cao, Yizhu Jin, Ming Lu, Guangyu Wang, and Shanghang Zhang. I-medsam: Implicit medical image segmentation with segment anything, 2024.

[45] Virmarie Maquiling, Sean Anthony Byrne, Diederick C. Niehorster, Marcus Nyström, and Enkelejda Kasneci. Zero-shot segmentation of eye features using the segment anything model (sam), 2024.

[46] Jiazhong Cen, Jiemin Fang, Zanwei Zhou, Chen Yang, Lingxi Xie, Xiaopeng Zhang, Wei Shen, and Qi Tian. Segment anything in 3d with radiance fields, 2024.

[47] Andrew Seohwan Yu, Mohsen Hariri, Xuecen Zhang, Mingrui Yang, Vipin Chaudhary, and Xiaojuan Li. Novel adaptation of video segmentation to 3d mri: efficient zero-shot knee segmentation with sam2, 2024.

[48] Heng Guo, Jianfeng Zhang, Jiaxing Huang, Tony C. W. Mok, Dazhou Guo, Ke Yan, Le Lu, Dakai Jin, and Minfeng Xu. Towards a comprehensive, efficient and promptable anatomic structure segmentation model using 3d whole-body ct scans, 2024.

[49] Dongjie Cheng, Ziyuan Qin, Zekun Jiang, Shaoting Zhang, Qicheng Lao, and Kang Li. Sam on medical images: A comprehensive study on three prompt modes, 2023.

[50] Peilun Shi, Jianing Qiu, Sai Mu Dalike Abaxi, Hao Wei, Frank P-W Lo, and Wu Yuan. Generalist vision foundation models for medical imaging: A case study of segment anything model on zero-shot medical segmentation. *Diagnostics*, 13(11):1947, 2023.

[51] Christian Mattjie, Luis Vinicius de Moura, Rafaela Cappelari Ravazio, Lucas Silveira Kupssinskü, Otávio Parraga, Marcelo Mussi Delucis, and Rodrigo Coelho Barros. Zero-shot performance of the segment anything model (sam) in 2d medical imaging: A comprehensive evaluation and practical guidelines, 2023.

[52] Sihan Yang, Haixia Bi, Hai Zhang, and Jian Sun. Sam-unet:enhancing zero-shot segmentation of sam for universal medical images, 2024.

[53] Yichi Zhang and Zhenrong Shen. Unleashing the potential of sam2 for biomedical images and videos: A survey, 2024.

[54] Iira Häkkinen, Iaroslav Melekhov, Erik Englesson, Hossein Azizpour, and Juho Kannala. Medical image segmentation with sam-generated annotations, 2024.

[55] Jun Ma, Yuting He, Feifei Li, Lin Han, Chenyu You, and Bo Wang. Segment anything in medical images, 2024.

[56] Weijia Feng, Lingting Zhu, and Lequan Yu. Cheap lunch for medical image segmentation by fine-tuning sam on few exemplars, 2023.

[57] Ron Keuth, Lasse Hansen, Maren Balks, Ronja Jäger, Anne-Nele Schröder, Ludger Tüshaus, and Mattias Heinrich. Sam carries the burden: A semi-supervised approach refining pseudo labels for medical segmentation, 2024.

[58] Shuangping Huang, Hao Liang, Qingfeng Wang, Chulong Zhong, Zijian Zhou, and Miaojing Shi. Seg-sam: Semantic-guided sam for unified medical image segmentation, 2024.

[59] José Guilherme de Almeida, Nuno M. Rodrigues, Sara Silva, and Nickolas Papanikolaou. Testing the segment anything model on radiology data, 2024.

[60] Haoyu Dong, Hanxue Gu, Yaqian Chen, Jichen Yang, Yuwen Chen, and Maciej A. Mazurowski. Segment anything model 2: an application to 2d and 3d medical images, 2024.

[61] Xiangru Li, Yifei Zhang, and Liang Zhao. Multi-prompt fine-tuning of foundation models for enhanced medical image segmentation, 2023.

[62] Yichi Zhang, Zhenrong Shen, and Rushi Jiao. Segment anything model for medical image segmentation: Current applications and future directions, 2024.

[63] Chaoning Zhang, Yu Qiao, Shehbaz Tariq, Sheng Zheng, Chenshuang Zhang, Chenghao Li, Hyundong Shin, and Choong Seon Hong. Understanding segment anything model: Sam is biased towards texture rather than shape, 2023.

[64] Pengfei Gu, Zihan Zhao, Hongxiao Wang, Yaopeng Peng, Yizhe Zhang, Nishchal Sapkota, Chaoli Wang, and Danny Z. Chen. Boosting medical image classification with segmentation foundation model, 2024.

[65] Qi Fan, Xin Tao, Lei Ke, Mingqiao Ye, Yuan Zhang, Pengfei Wan, Zhongyuan Wang, Yu-Wing Tai, and Chi-Keung Tang. Stable segment anything model, 2023.

[66] Xiuqi Zheng, Yuhang Zhang, Haoran Zhang, Hongrui Liang, Xueqi Bao, Zhuqing Jiang, and Qicheng Lao. Curriculum prompting foundation models for medical image segmentation, 2024.

[67] Yiran Song, Qianyu Zhou, Xuequan Lu, Zhiwen Shao, and Lizhuang Ma. Su-sam: A simple unified framework for adapting segment anything model in underperformed scenes, 2024.

[68] Xianping Ma, Qianqian Wu, Xingyu Zhao, Xiaokang Zhang, Man-On Pun, and Bo Huang. Sam-assisted remote sensing imagery semantic segmentation with object and boundary constraints, 2023.

[69] Mingya Zhang, Liang Wang, Zhihao Chen, Yiyuan Ge, and Xianping Tao. Path-sam2: Transfer sam2 for digital pathology semantic segmentation, 2024.

[70] Shurong Chai, Rahul Kumar Jain, Shiyu Teng, Jiaqing Liu, Yinhao Li, Tomoko Tateyama, and Yen wei Chen. Ladder fine-tuning approach for sam integrating complementary network, 2023.

[71] Chengyin Li, Prashant Khanduri, Yao Qiang, Rafi Ibn Sultan, Indrin Chetty, and Dongxiao Zhu. Autoprosam: Automated prompting sam for 3d multi-organ segmentation, 2024.

[72] Yixin Zhang, Nicholas Konz, Kevin Kramer, and Maciej A. Mazurowski. Quantifying the limits of segment anything model: Analyzing challenges in segmenting tree-like and low-contrast structures, 2024.

[73] Rachana Sathish, Rahul Venkataramani, K S Shriram, and Prasad Sudhakar. Task-driven prompt evolution for foundation models, 2023.

[74] Wei-Ting Chen, Yu-Jiet Vong, Sy-Yen Kuo, Sizhuo Ma, and Jian Wang. Robustsam: Segment anything robustly on degraded images, 2024.

[75] Ruining Deng, Can Cui, Quan Liu, Tianyuan Yao, Lucas W. Remedios, Shunxing Bao, Bennett A. Landman, Lee E. Wheless, Lori A. Coburn, Keith T. Wilson, Yaohong Wang, Shilin Zhao, Agnes B. Fogo, Haichun Yang, Yucheng Tang, and Yuankai Huo. Segment anything model (sam) for digital pathology: Assess zero-shot segmentation on whole slide imaging, 2023.

[76] Tianyu Huang, Tao Zhou, Weidi Xie, Shuo Wang, Qi Dou, and Yizhe Zhang. Improving segment anything on the fly: Auxiliary online learning and adaptive fusion for medical image segmentation, 2024.

[77] Junde Wu, Wei Ji, Yuanpei Liu, Huazhu Fu, Min Xu, Yanwu Xu, and Yueming Jin. Medical sam adapter: Adapting segment anything model for medical image segmentation, 2023.

[78] Chenyang Zhu, Bin Xiao, Lin Shi, Shoukun Xu, and Xu Zheng. Customize segment anything model for multi-modal semantic segmentation with mixture of lora experts, 2024.

[79] Gabriel Bellon de Carvalho and Jurandy Almeida. Exploiting the segment anything model (sam) for lung segmentation in chest x-ray images, 2024.

[80] Haojie Zhang, Yongyi Su, Xun Xu, and Kui Jia. Improving the generalization of segmentation foundation model under distribution shift via weakly supervised adaptation, 2024.

[81] Zijian Wu, Adam Schmidt, Peter Kazanzides, and Septimiu E. Salcudean. Augmenting efficient real-time surgical instrument segmentation in video with point tracking and segment anything, 2024.

[82] Dwith Chenna and Suyash Bhogawar. Segment anything model (sam) for brain extraction in fmri studies, 2024.

[83] Can Cui, Ruining Deng, Quan Liu, Tianyuan Yao, Shunxing Bao, Lucas W. Remedios, Yucheng Tang, and Yuankai Huo. All-in-sam: from weak annotation to pixel-wise nuclei segmentation with prompt-based finetuning, 2023.

Disclaimer:

SurveyX is an AI-powered system designed to automate the generation of surveys. While it aims to produce high-quality, coherent, and comprehensive surveys with accurate citations, the final output is derived from the AI's synthesis of pre-processed materials, which may contain limitations or inaccuracies. As such, the generated content should not be used for academic publication or formal submissions and must be independently reviewed and verified. The developers of SurveyX do not assume responsibility for any errors or consequences arising from the use of the generated surveys.