Deep Learning for Image Segmentation in Computer Vision: A Survey

www.surveyx.cn

Abstract

The integration of deep learning, segmentation, computer vision, neural networks, and image processing is transforming artificial intelligence by enabling machines to interpret visual data with unprecedented precision and efficiency. This survey explores the synergy between these interconnected fields, highlighting the revolution brought about by convolutional neural networks (CNNs) in image segmentation tasks, such as semantic segmentation critical in biomedical imaging for diagnosis and treatment planning. Challenges like annotated dataset scarcity and diverse material segmentation in outdoor environments necessitate innovative solutions, including the use of multi-view satellite imagery and deep learning in agricultural applications for automated crop disease detection. The survey addresses the importance of explainability in image segmentation through saliency maps and hierarchical graph-based methods, enhancing machine understanding of complex visual information. The survey also examines the pivotal role of deep learning in advancing AI and ML, emphasizing its impact on medical imaging by automating segmentation tasks, thus improving consistency and efficiency in clinical evaluations. The nnU-Net framework exemplifies this by democratizing access to advanced segmentation capabilities. Moreover, the survey discusses deep learning's role in agriculture, enhancing disease detection efficiency, and the challenges of deploying models on low-power devices due to computational complexities. The survey aims to provide a comprehensive overview of deep learning techniques for image segmentation, exploring universal vision models, challenges in instance segmentation, and the evaluation of CNN architectures in medical diagnostics. It concludes by reflecting on the implications of these advancements, offering insights into the future potential of image segmentation in deep learning.

1 Introduction

1.1 Interconnected Fields

The convergence of deep learning, segmentation, computer vision, neural networks, and image processing is foundational to contemporary artificial intelligence, empowering machines to analyze visual data with remarkable accuracy. Deep learning, particularly via convolutional neural networks (CNNs), has transformed image segmentation tasks, notably semantic segmentation, which partitions images into meaningful segments [1]. This is especially critical in biomedical imaging, where lesion and anatomical structure segmentation is essential for diagnosis and treatment planning.

Material segmentation faces challenges such as limited annotated datasets and the variety of materials in outdoor settings, prompting the need for innovative solutions [2]. For example, multi-view satellite imagery leverages deep learning to improve segmentation accuracy by utilizing multi-angle reflectance data [3]. Additionally, the integration of deep learning with image processing has proven effective in agriculture, where automated systems identify crop diseases, illustrating the practical interrelations among these fields [4].

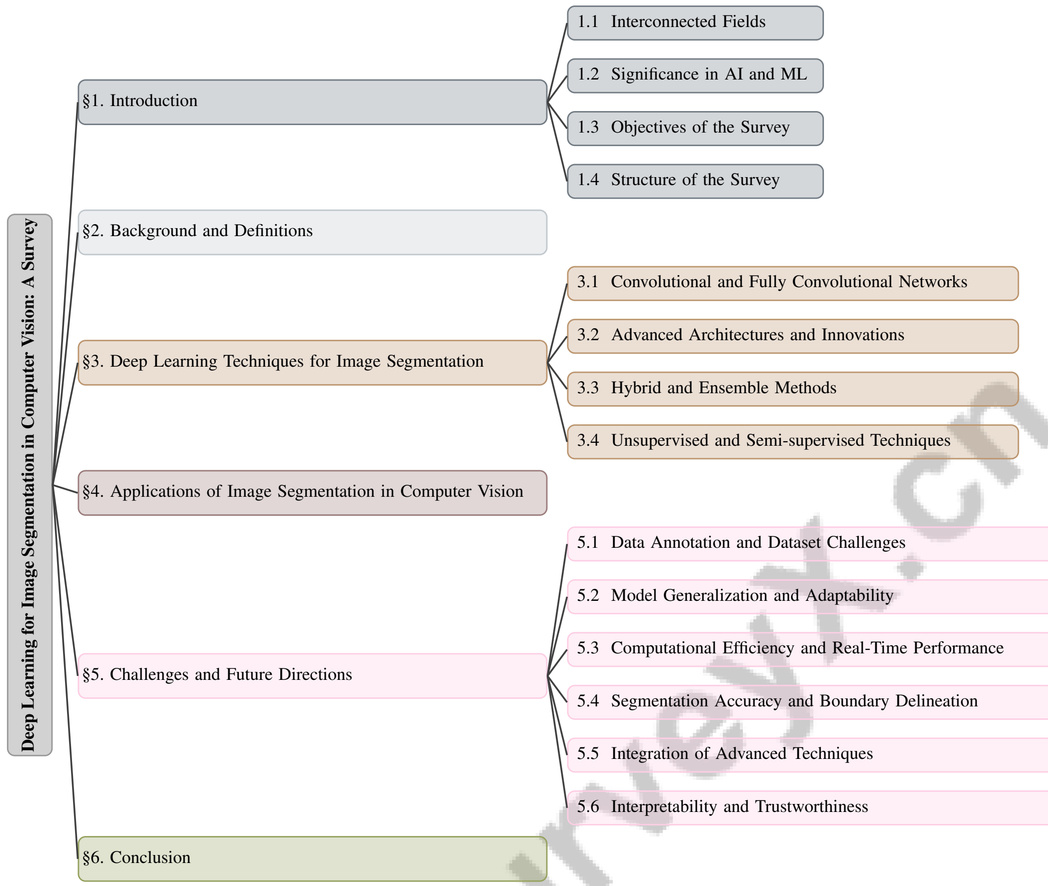


Figure 1: chapter structure

The necessity for explainability in image segmentation is addressed through saliency maps, which clarify the decision-making of neural networks [5]. This is particularly vital in medical applications, such as melanoma detection, where lesion segmentation precedes feature extraction and classification [6]. Hierarchical graph-based segmentation methods further enhance image partitioning by enabling scalability across different observation scales [7].

In biomedical imaging, CNNs exhibit effectiveness in semantic segmentation even with small datasets, highlighting the adaptability of deep learning in resource-limited contexts [8]. As these technologies evolve, the interplay between deep learning, segmentation, and image processing will continue to propel advancements in computer vision, enhancing machines' capabilities to process complex visual information.

1.2 Significance in AI and ML

The integration of deep learning, image segmentation, computer vision, neural networks, and image processing is crucial for the advancement of artificial intelligence and machine learning, significantly improving machines' ability to analyze complex visual data. In medical imaging, deep learning has automated segmentation tasks, mitigating issues such as high inter-rater variability and the time-consuming nature of manual segmentation [9]. Automated methods are essential for consistency and efficiency, particularly in glioma management and cardiac function evaluation, where manual delineation is labor-intensive and requires expertise [10]. The nnU-Net framework exemplifies this by achieving high-quality segmentation without necessitating expert knowledge, thus democratizing advanced segmentation capabilities [11].

In healthcare, precise segmentation of tumor regions in histopathological images is vital for patient-specific microdosimetry and radiobiological modeling, underscoring the significance of segmentation in medical diagnostics [12]. Automated systems for diabetic retinopathy detection, particularly for identifying microaneurysms, emphasize the role of image segmentation in early disease detection and management [13]. The potential for early malignant melanoma detection through dermoscopy images and advanced image processing techniques is increasingly important due to rising incidence rates [6].

In agriculture, deep learning enhances disease detection efficiency and accuracy in critical crops like tomatoes and corn, reflecting its positive impact on productivity and sustainability [4]. The demand for energy-efficient, real-time performance on embedded devices is crucial for environmental monitoring applications, with deep learning addressing various challenges through image processing and data analysis, thus improving efficiency and precision [14]. However, deploying deep learning models on low-power edge devices remains challenging due to architectural and computational complexities, highlighting the need for memory-efficient networks suitable for devices with limited resources, such as mobile phones and IoT devices [15].

In computer vision, image segmentation is essential for advancing AI and ML, with hierarchical image segmentation offering a region-oriented scale-space to tackle significant challenges in existing algorithms [7]. The search for new approaches to achieve desired accuracy, particularly where existing methods fall short, underscores the ongoing evolution of deep learning techniques [16]. Object-centric representation learning enhances deep neural networks' ability to understand and reason about visual scenes by modeling them as compositions of objects and their relationships [17]. Current methods often encounter limitations due to task-specific architectures that do not generalize well across different segmentation tasks, resulting in inefficiencies and suboptimal performance [18].

Addressing inefficiencies in processing large datasets is vital for the continued evolution and applicability of deep learning techniques across diverse fields. Developing quantitative evaluation metrics for color image segmentation methods is crucial for selecting appropriate metrics and understanding their impacts on segmentation results [19]. These advancements underscore the versatility and predictive power of deep learning across various domains, including image processing, natural language processing, and healthcare diagnostics. The reliance on substantial labeled data and computational resources highlights the necessity for innovative methods that can bypass these requirements, broadening AI and ML's applicability in resource-constrained environments [20].

1.3 Objectives of the Survey

This survey aims to provide a comprehensive overview of deep learning techniques for image segmentation, emphasizing the integration of neural networks and image processing within computer vision [21]. By addressing rapid advancements in this field, the survey seeks to offer a structured understanding of various semantic segmentation methods, which is essential given the diversity of approaches that complicate comprehension [22]. A key objective is to explore universal vision models capable of effectively addressing multiple vision tasks, including image classification, object detection, and image segmentation, thereby unifying these tasks to enhance performance.

The survey also examines the challenges and advancements in instance segmentation, providing insights into its evolution and the difficulties associated with delineating semantic boundaries in AI-generated content [23]. Furthermore, it evaluates the effectiveness of automatic segmentation methods using CNNs, particularly in improving the accuracy and efficiency of left ventricle segmentation in CCTA scans, which is crucial for medical diagnostics [24].

By analyzing the performance of various CNN architectures in segmenting MS lesions and brain structures, the survey seeks to establish benchmarks for assessing segmentation performance [10]. Additionally, it addresses the complexities involved in the design process due to the high dimensional space of hyperparameters and dataset properties, as evidenced in the development of automated deep learning solutions [11].

In the realm of medical imaging, the survey explores deep learning techniques for glaucoma detection, highlighting the limitations of traditional methods and the potential of automated approaches [25]. It also discusses the early diagnosis of diabetic retinopathy through microaneurysm detection in retinal images, emphasizing the importance of segmentation in early disease management [13].

The survey aims to enhance the understanding of deep learning-based image segmentation by providing a comprehensive review of current literature, encompassing a wide array of advanced techniques such as fully convolutional networks, encoder-decoder architectures, and generative adversarial models. It highlights applications of image segmentation in critical fields like medical image analysis, autonomous vehicles, and augmented reality. Additionally, the survey identifies the strengths and challenges of various deep learning models, evaluates their performance using widely utilized datasets, and outlines promising directions for future research and development in this rapidly evolving area of computer vision [26, 21, 27, 28].

1.4 Structure of the Survey

The survey is meticulously organized to provide a comprehensive exploration of deep learning techniques for image segmentation, structured into several key sections that systematically address the field's complexities and advancements. Initially, the survey introduces the interconnected fields of deep learning, segmentation, computer vision, neural networks, and image processing, emphasizing their significance in artificial intelligence and machine learning. This foundational understanding sets the stage for subsequent sections.

In the background and definitions section, core concepts such as deep learning, neural networks, and image segmentation are defined and interrelated, providing readers with the necessary theoretical grounding [28]. The survey then delves into deep learning techniques for image segmentation, covering various approaches including convolutional neural networks (CNNs) and fully convolutional networks (FCNs), as well as advanced architectures and innovations like the CLUSTERFORMER method [29]. The exploration extends to hybrid and ensemble methods, unsupervised and semi-supervised techniques, and their applications in enhancing segmentation performance.

The applications of image segmentation in computer vision are extensively discussed, highlighting its impact on diverse domains such as medical imaging, autonomous vehicles, agriculture, and industrial applications. The survey also addresses the challenges and future directions in image segmentation, focusing on data annotation, model generalization, computational efficiency, segmentation accuracy, and the integration of advanced techniques [23].

Finally, the conclusion synthesizes the key findings and reflects on the implications of deep learning advancements in image segmentation for the future of computer vision and AI. Throughout the survey, a clear distinction is made between deep learning-based techniques and traditional methods, ensuring a focused and relevant discourse [21]. The structure is designed to guide readers through a logical progression of topics, facilitating a comprehensive understanding of the current landscape and future potential of image segmentation in deep learning.The following sections are organized as shown in Figure 1.

2 Background and Definitions

2.1 Core Concepts in Image Segmentation

Image segmentation is a crucial component in computer vision, facilitating the division of images into distinct regions for detailed analysis. This process assigns labels to each pixel, indicating its class, which is essential for machines to interpret visual information effectively [21]. Segmentation types include semantic, instance, and panoptic. Semantic segmentation assigns class labels to pixels, a task complicated by scenes with multiple categories [1]. Instance segmentation advances this by providing object class labels and pixel-specific instance masks, which are vital for distinguishing between objects of the same class [30].

In medical imaging, segmentation is indispensable for delineating organs and lesions, critical for accurate diagnosis and treatment planning, such as left ventricle segmentation for cardiac assessment [24]. Challenges like over-segmentation from coarse outputs in fully convolutional networks can lead to inaccuracies in delineating lesion borders. Effective segmentation, especially with limited annotated data, is crucial, as seen in microaneurysm detection for early diabetic retinopathy diagnosis [13]. Conditional random fields (CRFs) are utilized to enhance voxel classification results, improving segmentation accuracy [31].

Deep learning models have significantly improved segmentation accuracy and efficiency by partitioning images into homogeneous regions [32]. In agriculture, segmenting diseased leaves is vital for effective disease detection, reflecting segmentation's practical applications in crop health management

[4]. Segmenting mirrors, where reflections resemble surroundings, underscores the need for robust techniques [33].

The transition from manual to automated feature extraction using deep learning has been pivotal in overcoming segmentation challenges across domains, including glaucoma detection [25]. Embedding visual prior knowledge in instance segmentation, especially with limited data and resources, remains challenging [30]. Segmenting and tracking multiple moving objects in videos, particularly under occlusion, highlights the complexity and need for advanced methods [34].

Image segmentation is integral to computer vision, enhancing machines' ability to interpret visual data by dividing images into meaningful segments. It underpins applications in medical imaging, face recognition, and object detection, supported by algorithms like region-based segmentation, edge detection, and clustering methods. The extensive research body underscores its significance across fields like healthcare and agriculture, aiding in tasks like disease detection in plants and crop monitoring [23, 35, 26]. Ongoing advancements continue to drive research and innovation, leading to sophisticated and robust techniques critical for diverse applications.

2.2 Deep Learning and Neural Networks

Deep learning, particularly through neural networks, has revolutionized image segmentation by enhancing precision and efficiency. Convolutional Neural Networks (CNNs) are central to these advancements, acting as essential tools for feature extraction and segmentation due to their hierarchical structure that captures spatial hierarchies within images. This capability is particularly beneficial in medical imaging tasks, such as skin lesion and stroke lesion segmentation, where CNNs improve accuracy by learning spatial dependencies. Fully Convolutional Networks (FCNs) extend CNNs' utility by enabling pixel-wise segmentation outputs from images of arbitrary sizes, facilitating diverse applications [24].

Advancements in neural network architectures, including deep convolutional and recurrent networks, have refined segmentation capabilities by effectively learning spatial dependencies [36]. Modular neural network approaches, combining modular learning for classification with self-training for data labeling, utilize both labeled and unlabeled data, addressing the challenge of limited annotated datasets [37]. The Segmenter approach employs a transformer architecture for semantic segmentation, leveraging global context without relying on convolutional operations [38].

Integrating high-level constraints into segmentation, methods like using a Conditional Restricted Boltzmann Machine (CRBM) enhance deep CNN outputs, showcasing the evolution of deep learning techniques in semantic segmentation [1]. The hybridization of traditional image processing with deep learning, as seen in microaneurysm detection in retinal images, exemplifies approaches that improve outcomes by combining methodologies [13].

The robustness of neural networks in brain tumor segmentation is evaluated through benchmarks ensuring models generalize effectively beyond specific datasets [39]. Bayesian convolutional neural networks for active learning on high-dimensional data model uncertainty and select informative data points for labeling, enhancing training efficiency [40].

Hierarchical graph-based image segmentation (HGBIS) methods compute a hierarchy of segmentations using edge-weighted graphs and dissimilarity measures, enabling stable segmentations across scales [7]. Additionally, integrating Conditional Random Fields (CRFs) as a post-processing method enhances voxel classification results, further boosting accuracy [31].

As deep learning techniques evolve, they promise to refine the precision and applicability of image segmentation across domains, from medical imaging to autonomous systems. The rapid advancements in neural network architectures highlight the transformative potential of deep learning, revolutionizing the automatic recognition of complex visual patterns and achieving accuracy that often surpasses human capabilities. This evolution enhances machines' ability to interpret intricate visual data, broadening applications across diverse fields, including medical image analysis, financial forecasting, and robotic perception, addressing challenges previously considered intractable [41, 21, 42, 28, 43].

3 Deep Learning Techniques for Image Segmentation

The integration of deep learning techniques has significantly transformed image segmentation, providing more precise and efficient methods for delineating objects within images. Foundational architectures, such as Convolutional Neural Networks (CNNs) and Fully Convolutional Networks (FCNs), have been instrumental in this evolution. These architectures form the backbone of contemporary segmentation strategies, offering robust frameworks for feature extraction and pixel-wise classification. Table 2 offers a comprehensive comparison of different methods in convolutional and fully convolutional networks, advanced architectures, and hybrid and ensemble methods, illustrating their application domains, key innovations, and the challenges they address in the context of image segmentation.

Figure 2 illustrates the hierarchical structure of deep learning techniques in image segmentation, categorized into foundational networks, advanced architectures, hybrid methods, and unsupervised/semi-supervised techniques. Each category highlights key applications, challenges, and innovations, emphasizing their contributions to improving segmentation accuracy, efficiency, and adaptability across various domains. The following subsections delve into the characteristics and advancements of CNNs and FCNs, underscoring their critical roles in enhancing segmentation performance across various applications.

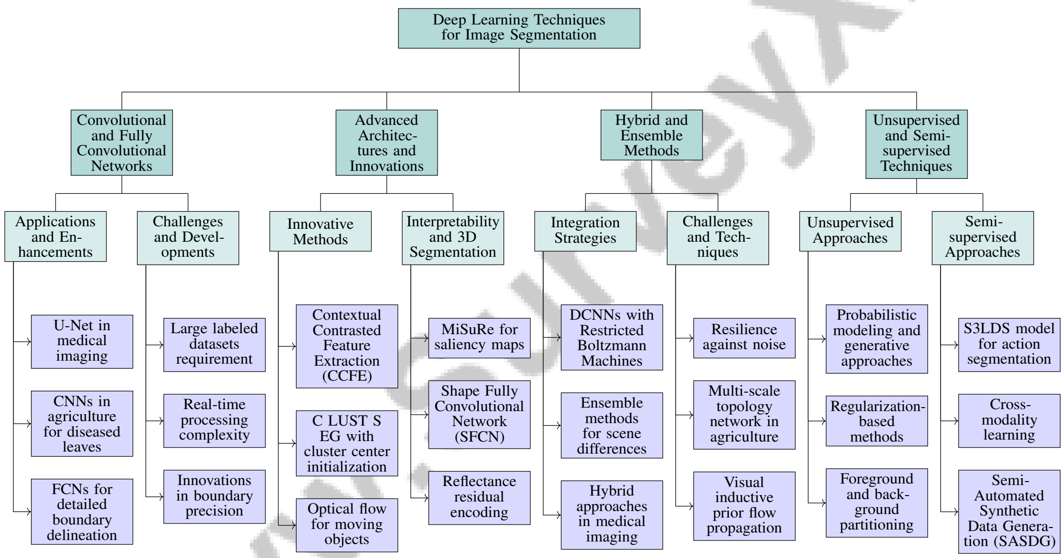


Figure 2: This figure illustrates the hierarchical structure of deep learning techniques in image segmentation, categorized into foundational networks, advanced architectures, hybrid methods, and unsupervised/semi-supervised techniques. Each category highlights key applications, challenges, and innovations, emphasizing their contributions to improving segmentation accuracy, efficiency, and adaptability across various domains.

3.1 Convolutional and Fully Convolutional Networks

Method Name	Architectural Features	Application Domains	Challenges and Solutions
CNN-SS[8]	Hierarchical Architecture Capability	Biomedical Images Segmentation	Small Datasets Imbalance
DL-PDD[4]	Multi-scale Topology	Agriculture	Large Datasets
RRE[3]	Segmentation Architectures	Urban Planning	Noisy Predictions
PCRF[31]	3D Unet	Medical Imaging	Posterior-CRF Integration
HGBIS[7]	Hierarchical Architecture	Medical Imaging, Agriculture	Large Datasets Requirement

Table 1: Comparison of Convolutional and Fully Convolutional Network Methods: This table presents a detailed analysis of various methods within the domain of Convolutional and Fully Convolutional Networks, highlighting their architectural features, application domains, and the specific challenges they address. The comparison underscores the versatility and adaptability of these networks across different fields, such as biomedical imaging and agriculture.

Table 1 provides a comprehensive comparison of different Convolutional and Fully Convolutional Network methods, illustrating their architectural features, application domains, and the challenges they address.

Convolutional Neural Networks (CNNs) and Fully Convolutional Networks (FCNs) are pivotal in image segmentation, providing robust frameworks for feature extraction and pixel-wise classification. CNNs, with their hierarchical architecture, capture spatial hierarchies within images, enhancing segmentation accuracy and efficiency. The U-Net architecture exemplifies this in medical imaging, segmenting images into meaningful regions and improving classification accuracy [8]. In agriculture, CNNs effectively classify images, aiding in segmenting diseased leaves in crops like tomatoes and corn [4].

FCNs extend CNN capabilities by enabling dense, pixel-wise predictions, crucial for tasks requiring detailed boundary delineation, such as material segmentation in satellite imagery [3]. Enhancements like integrating Posterior-CRF optimize CNN and CRF parameters during training, improving segmentation accuracy by incorporating posterior probabilities from CNNs [31].

Despite their strengths, CNNs and FCNs face challenges, such as the need for large labeled datasets and real-time processing complexity. Hierarchical graph-based image segmentation (HGBIS) addresses these by constructing a minimum spanning tree from the image's adjacency graph, achieving stable segmentations across scales [7]. CNNs applied to small biomedical datasets demonstrate that architectures like U-Net yield superior results, highlighting CNN adaptability in resource-constrained environments [8].

Ongoing CNN and FCN development continues to refine performance and expand applicability across diverse segmentation tasks. Innovations overcoming limitations like boundary precision and dataset dependency enhance utility in complex scenarios. As deep learning advances, CNNs and FCNs play crucial roles in advancing image segmentation, particularly in biomedical analysis, scene understanding, and augmented reality, achieving efficient and effective segmentation processes [44, 21].

3.2 Advanced Architectures and Innovations

Recent deep learning advancements have significantly enhanced image segmentation capabilities, introducing innovative methods that improve accuracy and efficiency. The Contextual Contrasted Feature Extraction (CCFE) module exemplifies this progress by modeling semantic and low-level discontinuities, enhancing mirror segmentation [33]. This approach integrates diverse architectural components to leverage individual strengths.

The C LUST S EG method introduces novel cluster center initialization and a recurrent cross-attention mechanism, refining pixel assignments and cluster centers, showcasing improved segmentation through innovative clustering [18]. Using optical flow as the sole input for segmenting moving objects allows for discovering and segmenting multiple objects while inferring mutual occlusions, bypassing traditional appearance-based methods [34].

In interpretability, the MiSuRe approach generates sufficient and minimally sufficient saliency maps, enhancing model interpretability by clarifying neural network decision-making [5]. This focus is crucial in applications where understanding the model's rationale is as important as the outcome.

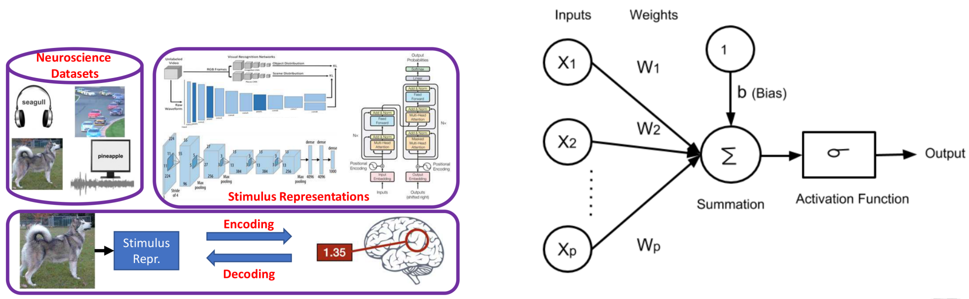
The Shape Fully Convolutional Network (SFCN) incorporates novel graph convolution and pooling operations for effective 3D shape segmentation, extending techniques into three-dimensional domains [45]. Reflectance residual encoding combines reflectance measurements with non-uniform sampling angles, enhancing material segmentation in satellite imagery [3].

Exploring multiple CNN architectures and hyper-parameters for small biomedical datasets underscores CNN adaptability in resource-constrained environments, ensuring effective segmentation with limited data [8]. The Posterior-CRF method leverages CNN-generated feature maps for better class separation and stability, enhancing segmentation accuracy [31].

These advancements underscore the rapidly evolving landscape of deep learning research in image segmentation. Researchers explore and refine techniques—including fully convolutional networks, encoder-decoder architectures, and generative adversarial models—expanding capabilities and applications in medical analysis, autonomous vehicles, and surveillance [21, 27, 28]. By addressing

AI-generated, for reference only.

traditional limitations and integrating novel techniques, these innovations pave the way for more accurate, efficient, and versatile segmentation methods across domains.



(a) Neuroscience Datasets and Stimulus Representations in Brain-Computer Interfaces[46]

(b) A Neural Network Architecture[43]

Figure 3: Examples of Advanced Architectures and Innovations

In Figure 3, advanced architectures and innovations in deep learning for image segmentation are highlighted. The first image illustrates a brain-computer interface (BCI) system integrating diverse neuroscience datasets, crucial for understanding how stimuli are represented and processed in the brain. The second image presents a simplified schematic of a neural network, showcasing essential components like inputs, weights, and activation functions. Together, these images underscore the innovative architectures driving advancements in deep learning for image segmentation, offering insights into theoretical and practical applications [46, 43].

3.3 Hybrid and Ensemble Methods

Hybrid and ensemble methods are powerful strategies in image segmentation, capitalizing on various models' strengths to enhance accuracy, robustness, and generalization. By integrating diverse techniques, these methods mitigate individual model limitations, leading to superior segmentation outcomes. For instance, integrating deep convolutional neural network (DCNN) outputs with Restricted Boltzmann Machines (RBM) exemplifies a hybrid approach leveraging high-level information for enhanced accuracy [1].

Ensemble methods, like combining predictions from multiple learners, refine results through an adjustment factor enhancing performance based on scene differences [47]. In medical imaging, a hybrid approach combining three CNNs for localization with a dedicated CNN for voxel classification significantly improves accuracy, particularly in left ventricle segmentation [24]. This strategy highlights integrating multiple CNN architectures for precise medical analysis.

Innovative techniques like the MiSuRe method, involving a two-stage process for sufficient and minimally sufficient regions, enhance interpretability by providing coarse and fine explanations [5]. Using a multi-scale topology network and attention mechanisms in agriculture further illustrates how hybrid methods enhance feature extraction and improve segmentation in tasks like disease detection in crops [4].

In brain tumor segmentation, hybrid methods demonstrate resilience against artifacts like noise and distribution shifts, although performance can be challenged under such conditions [39]. Integrating intrinsic prior knowledge through a visual inductive prior flow propagation framework exemplifies hybrid methods' potential in addressing complex dataset challenges [30].

Overall, hybrid and ensemble methods advance image segmentation by integrating diverse techniques and leveraging different models' strengths. These approaches aim to improve performance across fields, especially in medical imaging, by addressing limitations and leveraging techniques like Auxiliary Online Learning and adaptive fusion. Integrating online learning with rectified annotations enhances model accuracy, making them more effective in real-world applications. Exploring diverse prompt types in interactive segmentation and developing faster alternatives demonstrate a commitment to overcoming challenges and achieving superior results [48, 49, 28, 50, 26].

3.4 Unsupervised and Semi-supervised Techniques

Unsupervised and semi-supervised learning techniques are pivotal in advancing image segmentation, particularly where acquiring extensive labeled datasets is challenging. These approaches leverage labeled and unlabeled data to significantly improve accuracy and efficiency, overcoming traditional supervised methods' challenges that heavily depend on annotated datasets. For instance, one method enhances training data quality by identifying and correcting unreliable pixel-level annotations through aleatoric uncertainty assessment. Another approach optimizes annotation efforts by suggesting the most informative instances for annotation, reducing workload while achieving state-of-the-art performance. An auxiliary online learning technique refines segmentation outputs in real-time, using rectified annotations to boost existing models during medical analysis. Together, these strategies address conventional methods' limitations and pave the way for robust and efficient processes [50, 44, 51].

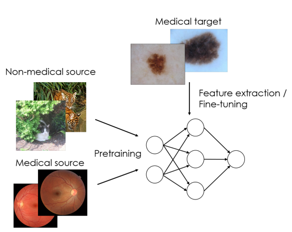
Unsupervised techniques, like those employing probabilistic modeling and generative approaches, partition foreground and background through frameworks like Expectation-Maximization, as in the DRC method [52]. This integration delineates image components effectively without labeled data reliance. Similarly, regularization-based approaches iteratively refine segmentation by classifying superpixels and enforcing spatial consistency, showcasing unsupervised learning's potential in enhancing performance [53].

Semi-supervised techniques leverage weak and strong annotations to improve outcomes. The Semi-Supervised Switching Linear Dynamical System (S3LDS) model incorporates temporal dynamics with labeled and unlabeled data for action segmentation, underscoring semi-supervised approaches' importance in dynamic environments [54]. Cross-modality learning exploits prior knowledge from an assistant modality to enhance target modality performance, addressing annotated data scarcity [55].

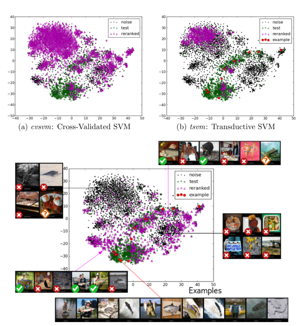
Innovative methods like Semi-Automated Synthetic Data Generation (SASDG) generate synthetic images by placing animals from labeled images into target backgrounds, significantly improving performance by augmenting datasets with realistic variations [56]. Uncertainty-based detection methods differentiate benign samples and adversarial attacks in semantic models, highlighting uncertainty estimation's role in refining accuracy [57].

Incorporating weakly labeled data from web images enhances convolutional networks' representation learning, illustrating leveraging readily available but imperfect data sources to improve models [58]. Streamlined, memory-efficient frameworks incorporating visual priors exemplify instance segmentation advancements, achieving better performance with reduced computational demands [30].

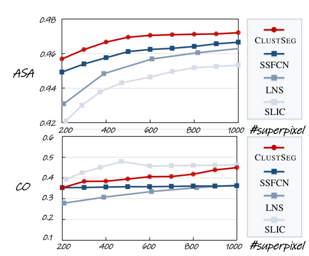
Overall, unsupervised and semi-supervised techniques contribute significantly to advancing image segmentation by reducing extensive labeled dataset dependency and enhancing model adaptability and accuracy. These approaches drive innovation, offering promising solutions to data scarcity and annotation costs, particularly in decentralized data scenarios like vertical federated learning [59].



(a) Medical Image Classification Using Pre-trained Networks[60]



(b) Comparison of Cross-Validated and Transductive SVMs in Image Classification[58]



(c) Comparison of Superpixel Segmentation Algorithms[18]

Figure 4: Examples of Unsupervised and Semi-supervised Techniques

In Figure 4, deep learning techniques in image segmentation are highlighted, especially unsupervised and semi-supervised approaches. The first example showcases pre-trained networks for medical classification, leveraging non-medical and medical images for enhanced feature extraction. The second example compares cross-validated and transductive SVMs, providing insights into their performance in classification tasks. Lastly, the third example evaluates superpixel segmentation algorithms, determining efficiency and precision in segmenting meaningful regions. Together, these examples illuminate unsupervised and semi-supervised techniques' potential in advancing segmentation [60, 58, 18].

Feature	Convolutional and Fully Convolutional Networks	Advanced Architectures and Innovations	Hybrid and Ensemble Methods
Application Domain	Medical Imaging	Autonomous Vehicles	Medical Imaging
Key Innovation	Dense Predictions	Semantic Discontinuities	Model Integration
Challenge Addressed	Large Dataset Need	Boundary Precision	Model Limitations

Table 2: This table provides a comparative analysis of various deep learning methods used in image segmentation, focusing on convolutional and fully convolutional networks, advanced architectures, and hybrid and ensemble methods. It highlights the application domains, key innovations, and challenges addressed by each method, offering insights into their specific contributions to the field of image segmentation.

4 Applications of Image Segmentation in Computer Vision

Image segmentation is a cornerstone of computer vision, enabling precise delineation and classification of visual data across numerous applications. It is indispensable in areas ranging from medical diagnostics to autonomous navigation, where accurate visual interpretation is critical. This section delves into the diverse applications of image segmentation, beginning with its pivotal role in medical imaging, where it aids in identifying anatomical structures and pathological regions, significantly influencing patient care and treatment strategies.

4.1 Medical Imaging

Image segmentation is crucial in medical imaging for accurately delineating anatomical structures and pathological regions, which are vital for diagnosis, treatment planning, and surgical interventions. The integration of deep learning, especially convolutional neural networks (CNNs), has markedly improved automated segmentation, achieving high accuracy in complex medical scenarios. The nnU-Net framework exemplifies this advancement, setting benchmarks in biomedical imaging through automated method design [11].

Advanced segmentation techniques address challenges such as the hole and shrink problems in melanoma segmentation, with methods like the Complementary Network with Adaptive Receptive Fields achieving a dice coefficient of 86.4% [61]. The use of multimodal datasets like MCubeS highlights the necessity of diverse imaging protocols for robust applications [2].

Weakly supervised learning methods have shown promise in reducing annotation efforts while maintaining competitive performance in medical imaging. The clinical evaluation of deep learning methods for stereotactic radiosurgery underscores the applicability of these technologies in diverse medical contexts [9].

Recent innovations such as the Mutual Inclusion Mechanism Precise Convolutional Network (MIPC-Net) enhance segmentation accuracy by combining channel and position attention, particularly at boundaries and abnormal regions. This approach, along with the GL-MIPC-Residue module, significantly reduces Hausdorff Distance, demonstrating superior performance across benchmarks [62, 63, 21]. The WaveMix neural network further exemplifies efficiency in resource-constrained environments.

Multimodal fusion enhances segmentation accuracy and robustness, as seen in geographic atrophy and retinal blood vessel segmentation, integrating diverse modalities to improve diagnostic precision and streamline workflows [64, 55, 65, 66].

These advancements in image segmentation technologies continue to enhance precision, reduce manual effort, and improve patient care in medical imaging. The HASA framework and MM-UNet exemplify progress in segmentation methodologies, offering robust solutions for precise ophthalmic image segmentation [67, 68, 69, 63, 26].

4.2 Autonomous Vehicles

Image segmentation is vital for enhancing the functionality and safety of autonomous vehicles by enabling precise scene understanding and object detection. Panoptic segmentation techniques unify semantic and instance segmentation, enhancing situational awareness by categorizing the environment at a pixel level, crucial for safety and reliability in autonomous navigation [70, 71, 49, 26]. CNNs and vision transformers have advanced real-time processing and accuracy in dynamic environments.

Segmentation in autonomous vehicles involves partitioning visual scenes into distinct regions, aiding in object classification and localization. High precision is essential for tasks like lane detection, obstacle avoidance, and traffic sign recognition, which rely on accurate situational awareness [71, 26]. Advanced segmentation methods enhance vehicle perception systems, enabling interpretation of complex urban environments.

Advancements in adversarial robustness, such as the Adversarial Robustness Augmentation (ARA) method, improve segmentation model reliability by addressing vulnerabilities to adversarial attacks, ensuring functionality in diverse conditions [72].

As autonomous vehicles evolve, image segmentation remains integral to their development. Innovations in panoptic segmentation and machine learning algorithms enhance navigation systems' accuracy and efficiency, improving object identification and categorization at a pixel level while addressing real-world noise challenges. These methods enhance visual perception systems, improving decision-making in assisted and automated driving functions [50, 73, 71, 26].

4.3 Agriculture and Environmental Monitoring

Image segmentation is crucial in agriculture and environmental monitoring, enhancing precision and efficiency. In agriculture, segmentation aids tasks like selective weeding, optimizing yield and reducing herbicide use. Model-based approaches improve crop and weed detection, addressing precision agriculture challenges [74].

Advanced methods like VoteNet significantly improve farmland boundary detection in high-resolution imagery, enhancing land management and resource allocation [75]. The U-Net based approach for tree crown delineation supports precision agriculture by accurately identifying tree health and growth [14].

In environmental monitoring, segmentation enhances urban planning and resource management. Multi-angle reflectance information integration improves material segmentation accuracy, providing insights into land cover and material distribution [3].

Overall, image segmentation in agriculture and environmental monitoring evolves with a focus on high-resolution aerial imagery, enhancing farmland boundary delineation and plant health identification. Recent developments demonstrate a 94.34

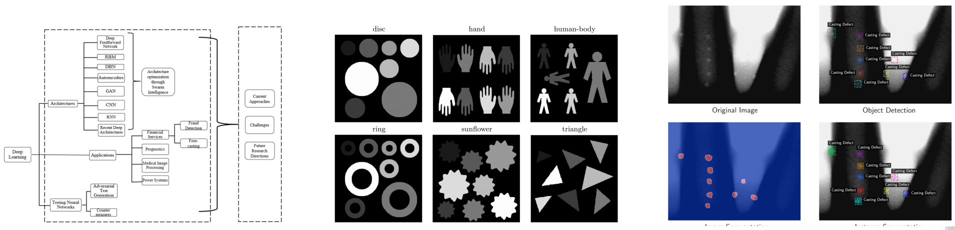
4.4 Industrial and Manufacturing Applications

Image segmentation is essential in industrial and manufacturing settings, enhancing automation, precision, and efficiency. It facilitates object detection, localization, and classification, optimizing workflows and reducing manual intervention. In engineering drawings, advanced segmentation automates complex design interpretation, streamlining manufacturing tasks [76, 23, 28, 35, 26].

Automated vectorization and semantic segmentation of raster engineering drawings outperform traditional methods, reducing manual intervention and enhancing precision [76]. In manufacturing, segmentation aids quality assurance by detecting defects and anomalies, ensuring consistent product quality [77, 42, 78, 26].

Segmentation in robotic vision systems facilitates automated assembly and material handling, enhancing manufacturing efficiency. Recent deep learning frameworks, like graph convolutional networks, improve component identification from engineering drawings, optimizing production [76, 79, 80, 49, 50].

The advancement of sophisticated image segmentation techniques propels innovation across industrial and manufacturing sectors, enhancing object detection accuracy and streamlining processes [50, 44, 76, 26]. These techniques promise to improve accuracy, efficiency, and automation, increasing competitiveness and sustainability in global manufacturing.



(a) Deep Learning: Architectures, Applications, and Challenges[43]

(b) Shapes and Objects[32]

(c) Object Detection and Instance Segmentation on a Radiographic Image[80]

Figure 5: Examples of Industrial and Manufacturing Applications

As shown in Figure 5, image segmentation plays a pivotal role across various industrial and manufacturing applications, enhancing precision and efficiency in processes relying on visual data interpretation. The examples demonstrate segmentation's integration with deep learning architectures, its application in identifying shapes and objects, and its efficacy in medical imaging. The first example illustrates deep learning's role in segmentation tasks, the second highlights segmentation of geometric shapes and objects, and the third showcases segmentation in medical imaging, emphasizing its impact on diagnostic accuracy and quality control [43, 32, 80].

4.5 3D and Video Segmentation

3D and video segmentation are critical in computer vision, enabling dynamic scene analysis and interpretation of complex structures. These techniques extract meaningful information from 3D data and video sequences, essential for applications like augmented reality, autonomous driving, and robotics. Partitioning 3D models or video frames into segments allows for accurate object identification and tracking, even with occlusions and appearance changes, while minimizing manual annotations [70, 81, 82, 34].

3D segmentation challenges include delineating objects within volumetric data characterized by high dimensionality. Hierarchical graph-based segmentation computes segmentation hierarchies using edge-weighted graphs for stable segmentations across scales [7]. The Shape Fully Convolutional Network (SFCN) introduces novel graph convolution and pooling operations, extending segmentation into three-dimensional domains [45].

Video segmentation involves continuous object segmentation within video streams, requiring spatial and temporal consistency. Using optical flow for segmenting moving objects allows for object discovery and segmentation without traditional appearance-based methods, emphasizing temporal dynamics' importance [34].

Despite advancements, challenges remain in real-time performance and maintaining segmentation accuracy amid occlusions, lighting variations, and complex motion patterns. Deep learning integration with traditional image processing continues to enhance 3D and video segmentation robustness and efficiency. As deep learning algorithms and Transformers advance, they enhance computer vision systems' capabilities, improving object detection, localization, and recognition in diverse environments [28, 78, 26].

5 Challenges and Future Directions

Advancements in image segmentation technologies encounter several complex challenges. This section examines the obstacles related to data annotation, dataset quality, model generalization, computational efficiency, segmentation accuracy, and the integration of advanced techniques, which are pivotal for future research and development.

5.1 Data Annotation and Dataset Challenges

Benchmark	Size	Domain	Task Format	Metric
MVSeg[83]	2,500	Vineyard Segmentation	Semantic Segmentation	F1-score
COLosSAL[84]	1,000	3D Medical Image Segmentation	Segmentation	Dice, HD95
SAM-2[85]	900	Surgical Video Segmentation	Segmentation	DSC, NSD
CIS[86]	1,000	Image Segmentation	Color Image Segmentation	MSE, PSNR
BandNet[87]	1,446	Water Segmentation	Image Segmentation	mIoU
SAM[88]	15,361	Medical Imaging	Image Segmentation	Dice, IoU
UMT-SSD[89]	5,000	Computer Vision	Joint Semantic Segmentation And Depth Estimation	mIoU, ECE
WSSS-Bench[90]	171,125	Histopathology	Semantic Segmentation	mIoU

Table 3: The table presents a comprehensive overview of various benchmarks utilized in image segmentation research, detailing their respective dataset sizes, application domains, task formats, and evaluation metrics. This information underscores the diversity and complexity inherent in segmentation tasks across different fields, highlighting the reliance on specific metrics for performance evaluation.

Benchmark	Size	Domain	Task Format	Metric
MVSeg[83]	2,500	Vineyard Segmentation	Semantic Segmentation	F1-score
COLosSAL[84]	1,000	3D Medical Image Segmentation	Segmentation	Dice, HD95
SAM-2[85]	900	Surgical Video Segmentation	Segmentation	DSC, NSD
CIS[86]	1,000	Image Segmentation	Color Image Segmentation	MSE, PSNR
BandNet[87]	1,446	Water Segmentation	Image Segmentation	mIoU
SAM[88]	15,361	Medical Imaging	Image Segmentation	Dice, IoU
UMT-SSD[89]	5,000	Computer Vision	Joint Semantic Segmentation And Depth Estimation	mIoU, ECE
WSSS-Bench[90]	171,125	Histopathology	Semantic Segmentation	mIoU

Table 4: The table presents a comprehensive overview of various benchmarks utilized in image segmentation research, detailing their respective dataset sizes, application domains, task formats, and evaluation metrics. This information underscores the diversity and complexity inherent in segmentation tasks across different fields, highlighting the reliance on specific metrics for performance evaluation.

Table 4 provides a detailed examination of the benchmarks employed in image segmentation studies, illustrating the challenges and variations in dataset characteristics, task formats, and evaluation criteria. The reliance on large volumes of annotated data remains a significant barrier to advancing image segmentation technologies. This is particularly evident in domains such as agricultural disease detection, where extensive labeled data is crucial for optimal model performance [4]. Class imbalance within datasets can skew accuracy metrics, affecting overall performance evaluations [8]. Accurate data annotation is further complicated by the need to model multimodal label distributions in high-dimensional output spaces, especially given limited annotation data. Existing methods often rely on low-quality intensity information, resulting in ineffective segmentation due to noise and overlapping intensity values [31]. Moreover, parameter tuning for hierarchical segmentation can lead to contour instability [7].

The scarcity of high-quality annotated data hampers segmentation accuracy, notably in specialized fields like biomedical image analysis, where precise pixel-level annotations are vital. Inadequate time and expertise during annotation processes can lead to incorrect labeling and boundary inaccuracies, adversely affecting deep convolutional neural networks (DCNNs), which may memorize random labels, resulting in poor outcomes. Innovative approaches such as uncertainty-based methods and active learning frameworks have been proposed to optimize annotation processes, facilitating effective training on reduced datasets while maintaining high performance [26, 44, 51]. Developing efficient annotation techniques and adaptive algorithms to manage diverse datasets is essential for enhancing segmentation model accuracy and applicability.

5.2 Model Generalization and Adaptability

Model generalization and adaptability are particularly challenging when deploying deep learning models across diverse domains. A significant hurdle is the dependence on large annotated datasets, which can impede a model's ability to generalize effectively across varying imaging protocols and environments. This issue is pronounced in medical imaging, where variations in modalities can lead to performance discrepancies; models trained on one data type may not perform well on another due to modality-specific differences [11]. For instance, the nnU-Net framework may require further optimization when applied to datasets with different properties [11].

In 3D shape segmentation, the irregular data structure of 3D shapes complicates the application of convolutional neural network techniques effective for 2D images [45]. CNN-based models, which primarily capture local information, struggle with diverse target shapes and textures, complicating generalization [29]. Class imbalance can lead to misleading accuracy metrics, impacting segmentation performance evaluations [8]. Adaptive algorithms that dynamically adjust to new data distributions are essential. Techniques such as ensemble learning can enhance classification performance and model generalization by leveraging the strengths of multiple models [47]. Improving model adaptability and generalization is crucial for extending the applicability of segmentation technologies across a broader range of real-world scenarios.

5.3 Computational Efficiency and Real-Time Performance

Balancing high segmentation accuracy with computational efficiency is essential for practical deployment across various domains. Recent advancements demonstrate potential for efficient solutions. Energy-efficient training methods have shown satisfactory performance in fetal brain segmentation while significantly reducing energy consumption, emphasizing the optimization of training processes for better resource management [91]. The VoteNet method enhances accuracy by integrating contextual information and performing majority voting, effectively reducing noise and improving boundary delineation [75].

Operating effectively with small datasets and on low-power devices is crucial for real-time processing, as illustrated by efficient U-Net-based approaches in agriculture [14]. This underscores the potential for deploying models in resource-constrained environments, ensuring rapid and accurate disease detection [4]. The CLUSTSEG method exemplifies unifying multiple segmentation tasks under a single framework, improving performance and contributing to efficiency [18]. Despite advancements, challenges persist in achieving real-time performance, particularly in complex scenarios where models must adapt dynamically. The integration of Divisive Normalization into networks presents a novel approach to enhancing adaptability across varying conditions [92]. Efficient algorithms requiring less supervision are critical for advancing computational efficiency, with approaches like deep Bayesian active learning optimizing the learning process by capturing and utilizing model uncertainty [40].

5.4 Segmentation Accuracy and Boundary Delineation

Achieving high segmentation accuracy and precise boundary delineation remains challenging, particularly with complex datasets. Sensitivity to variations in input data quality and manual interventions, such as clinician sketches, can introduce variability in outcomes [93]. This is compounded in scenarios involving small objects or intricate structures, like vascular networks, where models often struggle to maintain connectivity [94]. Segmenting smaller vessels while maintaining connectivity highlights model limitations in complex scenarios where fine boundary delineation is critical. Pixel density of classes can affect accuracy, particularly for under-represented classes [95].

In medical imaging, accurately delineating lesions is complicated by variability in lesion characteristics and overlap with normal tissues, as seen in multiple sclerosis lesion segmentation [96]. Advanced algorithms integrating spatial context information enhance accuracy, as demonstrated in abdominal multi-organ segmentation [97]. However, reliance on initial user input for semi-automatic methods can affect reconstruction quality, indicating a need for more autonomous solutions [82]. Model limitations when dealing with out-of-distribution data can lead to boundary delineation inaccuracies [98]. Addressing these challenges requires innovative techniques that adapt to varying conditions while maintaining high accuracy across applications.

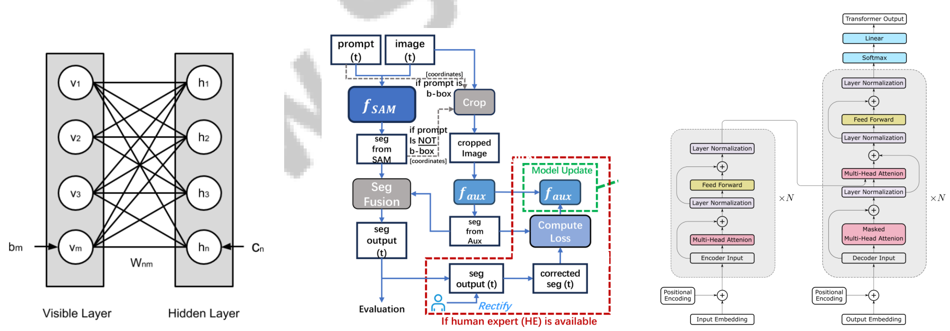
The proposed method for liver segmentation, achieving a Dice similarity coefficient (DSC) of 93.14%, exemplifies improvements in accuracy by addressing noise sensitivity and intensity overlap [99]. However, limitations like overfitting due to limited data, class imbalance, and computational demands persist [64]. The method's dense representation of visual prompts is more computationally expensive than sparse representations, indicating challenges in achieving high accuracy [48]. Additionally, reliance on the BraTS 2018 dataset may not fully capture clinical variability, highlighting the need for comprehensive datasets [100]. Accurate shape learning is critical, as inaccuracies can introduce performance loss [101]. K-Net improves performance, but mask boundaries may still require refinement [102]. Unanswered questions include the need for effective unsupervised and weakly supervised learning approaches [28]. The method's reliance on cropped lesion images for classification rather than processing entire ultrasound images limits applicability [103]. The CCDM's ability to produce diverse segmentation outputs accurately reflects variability in expert annotations, offering an advantage over deterministic methods [104]. Achieving high accuracy and boundary delineation remains a challenge in classifying diseases from chest X-rays [105].

5.5 Integration of Advanced Techniques

Integrating advanced techniques, such as Auxiliary Online Learning and adaptive fusion methods, is crucial for improving segmentation accuracy, addressing domain variations, and enhancing model generalization in complex fields like medical imaging. This approach optimizes performance by leveraging real-time corrections from expert annotations and facilitates deep learning applications across various modalities [26, 28, 50]. Using Random Feature Embedding in Fully Convolutional Networks for semi-supervised fine-tuning leverages unlabeled datasets to enhance generalization and address domain discrepancies [106]. Domain randomization to generate synthetic training data offers a solution to data scarcity, enabling effective learning without extensive real-world labeled datasets [74].

Incorporating domain knowledge into architectures, such as through wavelet families, can further enhance performance by integrating domain-specific insights. Developing segmentation-aware networks that optimize mechanisms for other dense prediction tasks presents intriguing future exploration avenues [62]. The CLUSTERFORMER method exemplifies optimizing clustering processes to improve performance and universality across tasks. Future research could focus on integrating advanced techniques to enhance clustering robustness and model adaptability [7].

Integrating advanced techniques promises to elevate image segmentation capabilities, enabling more precise, efficient, and adaptable solutions across applications. As research progresses, innovations in this field are poised to tackle challenges, enhance accuracy, and broaden applications across domains such as healthcare, agriculture, and artificial vision. Recent developments in frameworks are improving annotation efficiency, allowing high-performance segmentation with reduced training data requirements, thus expanding potential implementations [44, 26].



(a) A Neural Network Architecture[43]

(b) The image depicts a flowchart illustrating a machine learning process for image segmentation.[50]

(c) Transformer Model Architecture[78]

Figure 6: Examples of Integration of Advanced Techniques

As shown in Figure 6, integrating advanced techniques is pivotal for pushing the boundaries of machine learning and AI capabilities. The examples illustrate diverse methodologies employed to enhance AI systems. The first example, a neural network architecture, highlights the multilayer perceptron (MLP) structure, showcasing interactions between visible and hidden layers through weighted connections, crucial for pattern recognition and data classification. The second example delves into image segmentation through a flowchart, emphasizing the role of supervised attention modules in feature extraction, essential for accurately interpreting visual data. Lastly, the transformer model architecture exemplifies sophisticated mechanisms of modern AI, with its encoder-decoder framework leveraging multi-head attention and layer normalization to handle complex tasks. Together, these examples underscore the challenges and future directions in AI as researchers continue to integrate and refine techniques to tackle complex problems [43, 50, 78].

5.6 Interpretability and Trustworthiness

The interpretability and trustworthiness of deep learning models in image segmentation are critical for adoption in precision-demanding domains like medical imaging, autonomous driving, and video surveillance. Understanding decision-making processes ensures reliable predictions that align with human cognitive frameworks, especially in applications requiring precise uncertainty quantification. The MiSuRe method exemplifies the importance of providing both coarse and fine explanations for predictions, enhancing interpretability [5].

Future research should emphasize developing efficient algorithms that enhance output interpretability, particularly through integrating semi-supervised and unsupervised learning approaches [77]. Exploring unsupervised pretraining and the impact of dataset diversity on transfer learning can further improve robustness and interpretability, ensuring consistent performance across conditions [70]. Developing explainers that provide insights while ensuring fairness and reducing biases highlights the need for transparent AI systems [41].

Interdisciplinary collaboration among imaging specialists, data scientists, and domain experts is essential for advancing segmentation research and developing robust, interpretable models [33]. Enhancing model robustness against unknown occluding objects and integrating temporal information for video instance segmentation are areas for future exploration, impacting trustworthiness in dynamic environments [107].

The study underscores the necessity of tailored approaches for specific applications, suggesting future models incorporate cross-domain generalizability to enhance performance and reliability [6]. Refining network architectures to improve performance on challenging datasets and exploring applicability to other domains are critical for advancing interpretability and trustworthiness [3].

Advancing interpretability and trustworthiness in deep learning models for segmentation is crucial for broader adoption. By focusing on machine learning, deep learning architectures, and AI system interpretability, future research can enhance the development of reliable and comprehensible AI technologies addressing ethical concerns. This will ultimately promote greater trust in AI applications across diverse sectors, ensuring transparency and accountability in decision-making processes [41, 46, 77, 43, 108].

6 Conclusion

Deep learning has revolutionized image segmentation, significantly impacting computer vision and artificial intelligence. Frameworks like nnU-Net have achieved segmentation accuracy comparable to human experts, especially in medical diagnostics such as MS lesion analysis, highlighting their clinical potential. In 3D shape segmentation, advancements such as the Shape Fully Convolutional Network (SFCN) have set new performance standards, showcasing the ability of deep learning models to handle intricate three-dimensional data.

The development of hybrid architectures and sophisticated techniques, such as the Posterior-CRF method, has notably improved segmentation precision in complex scenarios like white matter hyperintensities. These innovations underscore the benefits of integrating diverse features and employing end-to-end training for optimal segmentation results. Additionally, the exploration of unsupervised and semi-supervised methodologies offers promising solutions to reduce reliance on large labeled datasets, thereby expanding the applicability of segmentation technologies in resource-limited settings.

Universal models, exemplified by CLUSTERFORMER, represent a significant shift by addressing multiple vision tasks, including image segmentation, with superior performance compared to specialized models. This is complemented by the rise of efficient hierarchical segmentation methods, which ensure stability and versatility across various applications. In agriculture, deep learning has demonstrated its efficacy by achieving high validation accuracies, thereby enhancing precision and efficiency in detecting crop diseases.

The continuous advancement of deep learning techniques promises to further enhance image segmentation capabilities, driving innovation across multiple domains. These improvements are poised to play a crucial role in the future of computer vision and AI, providing more precise, efficient, and adaptable solutions for interpreting complex visual data. As deep learning methodologies evolve, their integration into diverse applications is expected to lead to significant gains in accuracy, efficiency, and scalability, paving the way for continued research and development in this dynamic field.

References

[1] S. Tsogkas, I. Kokkinos, G. Papandreou, and A. Vedaldi. Deep learning for semantic part segmentation with high-level guidance, 2015.

[2] Yupeng Liang, Ryosuke Wakaki, Shohei Nobuhara, and Ko Nishino. Multimodal material segmentation. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 19800–19808, 2022.

[3] Matthew Purri, Jia Xue, Kristin Dana, Matthew Leotta, Dan Lipsa, Zhixin Li, Bo Xu, and Jie Shan. Material segmentation of multi-view satellite imagery, 2019.

[4] Affan Yasin and Rubia Fatima. On the image-based detection of tomato and corn leaves diseases : An in-depth comparative experiments, 2023.

[5] Syed Nouman Hasany, Fabrice Mériaudeau, and Caroline Petitjean. Misure is all you need to explain your image segmentation, 2024.

[6] Nabin K. Mishra and M. Emre Celebi. An overview of melanoma detection in dermoscopy images using image processing and machine learning, 2016.

[7] Silvio Jamil F. Guimarães, Jean Cousty, Yukiko Kenmochi, and Laurent Najman. An efficient hierarchical graph based image segmentation, 2012.

[8] Vitaly Nikolaev. Convolution neural networks for semantic segmentation: Application to small datasets of biomedical images, 2020.

[9] Boris Shirokikh, Alexandra Dalechina, Alexey Shevtsov, Egor Krivov, Valery Kostjuchenko, Amayak Durgaryan, Mikhail Galkin, Andrey Golanov, and Mikhail Belyaev. Systematic clinical evaluation of a deep learning method for medical image segmentation: Radiosurgery application, 2021.

[10] Richard McKinley, Rik Wepfer, Fabian Aschwanden, Lorenz Grunder, Raphaela Muri, Christian Rummel, Rajeev Verma, Christian Weisstanner, Mauricio Reyes, Anke Salmen, Andrew Chan, Franca Wagner, and Roland Wiest. Simultaneous lesion and neuroanatomy segmentation in multiple sclerosis using deep neural networks, 2020.

[11] Fabian Isensee, Paul F. Jäger, Simon A. A. Kohl, Jens Petersen, and Klaus H. Maier-Hein. Automated design of deep learning methods for biomedical image segmentation, 2020.

[12] Luca L. Weishaupt, Jose Torres, Sophie Camilleri-Broët, Roni F. Rayes, Jonathan D. Spicer, Sabrina Côté Maldonado, and Shirin A. Enger. Deep learning-based tumor segmentation on digital images of histopathology slides for microdosimetry applications, 2021.

[13] Meysam Tavakoli and Mahdieh Nazar. Comparison different vessel segmentation methods in automated microaneurysms detection in retinal images using convolutional neural networks, 2020.

[14] Kostas Blekos, Stavros Nousias, and Aris S Lalos. Efficient automated u-net based tree crown delineation using uav multi-spectral imagery on embedded devices, 2021.

[15] Caglar Aytekin, Xingyang Ni, Francesco Cricri, Lixin Fan, and Emre Aksu. Memory-efficient deep salient object segmentation networks on gridized superpixels, 2018.

[16] Tomer Amit, Tal Shaharbany, Eliya Nachmani, and Lior Wolf. Segdiff: Image segmentation with diffusion probabilistic models, 2022.

[17] Guillermo Puebla and Jeffrey S. Bowers. Visual reasoning in object-centric deep neural networks: A comparative cognition approach, 2024.

[18] James Liang, Tianfei Zhou, Dongfang Liu, and Wenguan Wang. Clustseg: Clustering for universal segmentation. *arXiv preprint arXiv:2305.02187*, 2023.

[19] Majid Harouni and Hadi Yazdani Baghmaleki. Color image segmentation metrics, 2020.

[20] Tendai Mapungwana Chikake, Boris Goldengorin, and Alexey Samosyuk. Pseudo-boolean polynomials approach to edge detection and image segmentation, 2023.

[21] Shervin Minaee, Yuri Boykov, Fatih Porikli, Antonio Plaza, Nasser Kehtarnavaz, and Demetri Terzopoulos. Image segmentation using deep learning: A survey, 2020.

[22] An abstraction model for semantic segmentation algorithms.

[23] Abdul Mueed Hafiz and Ghulam Mohiuddin Bhat. A survey on instance segmentation: State of the art, 2020.

[24] Majd Zreik, Tim Leiner, Bob D. de Vos, Robbert W. van Hamersvelt, Max A. Viergever, and Ivana Isgum. Automatic segmentation of the left ventricle in cardiac ct angiography using convolutional neural network, 2017.

[25] Aized Amin Soofi and Fazal e Amin. Exploring deep learning techniques for glaucoma detection: A comprehensive review, 2023.

[26] Farid Garcia-Lamont, Jair Cervantes, Asdrúbal López, and Lisbeth Rodriguez. Segmentation of images by color features: A survey. *Neurocomputing*, 292:1–27, 2018.

[27] Shervin Minaee, Yuri Boykov, Fatih Porikli, Antonio Plaza, Nasser Kehtarnavaz, and Demetri Terzopoulos. Image segmentation using deep learning: A survey. *IEEE transactions on pattern analysis and machine intelligence*, 44(7):3523–3542, 2021.

[28] Swarnendu Ghosh, Nibaran Das, Ishita Das, and Ujjwal Maulik. Understanding deep learning techniques for image segmentation. *ACM computing surveys (CSUR)*, 52(4):1–35, 2019.

[29] James C. Liang, Yiming Cui, Qifan Wang, Tong Geng, Wenguan Wang, and Dongfang Liu. Clusterformer: Clustering as a universal visual learner, 2023.

[30] Chih-Chung Hsu and Chia-Ming Lee. Miss: Memory-efficient instance segmentation framework by visual inductive priors flow propagation, 2024.

[31] Shuai Chen and Marleen de Bruijne. An end-to-end approach to semantic segmentation with 3d cnn and posterior-crf in medical images, 2018.

[32] Alexander Aue and Thomas C. M. Lee. On image segmentation using information theoretic criteria, 2012.

[33] Xin Yang, Haiyang Mei, Ke Xu, Xiaopeng Wei, Baocai Yin, and Rynson W. H. Lau. Where is my mirror?, 2019.

[34] Junyu Xie, Weidi Xie, and Andrew Zisserman. Segmenting moving objects via an object-centric layered representation, 2022.

[35] Song Yuheng and Yan Hao. Image segmentation algorithms overview. *arXiv preprint arXiv:1707.02051*, 2017.

[36] M. Attia, M. Hossny, S. Nahavandi, and A. Yazdabadi. Spatially aware melanoma segmentation using hybrid deep learning techniques, 2017.

[37] Nosseiba Ben Salem, Younes Bennani, Joseph Karkazan, Abir Barbara, Charles Dacheux, and Thomas Gregory. Modular neural network approaches for surgical image recognition, 2023.

[38] Robin Strudel, Ricardo Garcia, Ivan Laptev, and Cordelia Schmid. Segmenter: Transformer for semantic segmentation. In *Proceedings of the IEEE/CVF international conference on computer vision*, pages 7262–7272, 2021.

[39] Sabine Müller, Joachim Weickert, and Norbert Graf. Robustness of brain tumor segmentation, 2020.

[40] Yarin Gal, Riashat Islam, and Zoubin Ghahramani. Deep bayesian active learning with image data, 2017.

[41] Vanessa Buhrmester, David Münch, and Michael Arens. Analysis of explainers of black box deep neural networks for computer vision: A survey. *Machine Learning and Knowledge Extraction*, 3(4):966–989, 2021.

[42] Nicholas G. Polson and Vadim O. Sokolov. Deep learning, 2018.

[43] Saptarshi Sengupta, Sanchita Basak, Pallabi Saikia, Sayak Paul, Vasilios Tsalavoutis, Frederick Atiah, Vadlamani Ravi, and Alan Peters. A review of deep learning with special emphasis on architectures, applications and recent trends, 2019.

[44] Lin Yang, Yizhe Zhang, Jianxu Chen, Siyuan Zhang, and Danny Z. Chen. Suggestive annotation: A deep active learning framework for biomedical image segmentation, 2017.

[45] Pengyu Wang, Yuan Gan, Panpan Shui, Fenggen Yu, Yan Zhang, Songle Chen, and Zhengxing Sun. 3d shape segmentation via shape fully convolutional networks, 2018.

[46] Subba Reddy Oota, Zijiao Chen, Manish Gupta, Raju S. Bapi, Gael Jobard, Frederic Alexandre, and Xavier Hinaut. Deep neural networks and brain alignment: Brain encoding and decoding (survey), 2024.

[47] Chunbo Lang, Gong Cheng, Binfei Tu, and Junwei Han. Learning what not to segment: A new perspective on few-shot segmentation, 2022.

[48] Qin Liu, Jaemin Cho, Mohit Bansal, and Marc Niethammer. Rethinking interactive image segmentation with low latency, high quality, and diverse prompts, 2024.

[49] Xu Zhao, Wenchao Ding, Yongqi An, Yinglong Du, Tao Yu, Min Li, Ming Tang, and Jinqiao Wang. Fast segment anything, 2023.

[50] 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.

[51] Ekaterina Redekop and Alexey Chernyavskiy. Uncertainty-based method for improving poorly labeled segmentation datasets, 2021.

[52] Peiyu Yu, Sirui Xie, Xiaojian Ma, Yixin Zhu, Ying Nian Wu, and Song-Chun Zhu. Unsupervised foreground extraction via deep region competition, 2023.

[53] Aleksandar Dimitriev and Matej Kristan. A regularization-based approach for unsupervised image segmentation, 2016.

[54] Ari Blau, Evan S Schaffer, Neeli Mishra, Nathaniel J Miska, The International Brain Laboratory, Liam Paninski, and Matthew R Whiteway. A study of animal action segmentation algorithms across supervised, unsupervised, and semi-supervised learning paradigms, 2024.

[55] Kang Li, Lequan Yu, Shujun Wang, and Pheng-Ann Heng. Towards cross-modality medical image segmentation with online mutual knowledge distillation, 2020.

[56] Abhineet Singh, Marcin Pietrasik, Gabriell Natha, Nehla Ghouaiel, Ken Brizel, and Nilanjan Ray. Animal detection in man-made environments, 2020.

[57] Kira Maag, Roman Resner, and Asja Fischer. Detecting adversarial attacks in semantic segmentation via uncertainty estimation: A deep analysis, 2024.

[58] Phong D. Vo, Alexandru Ginsca, Hervé Le Borgne, and Adrian Popescu. On deep representation learning from noisy web images, 2016.

[59] Paul K. Mandal and Cole Leo. Vertical federated image segmentation, 2024.

[60] Veronika Cheplygina. Cats or cat scans: transfer learning from natural or medical image source datasets?, 2019.

[61] Xiaoqing Guo, Zhen Chen, and Yixuan Yuan. Complementary network with adaptive receptive fields for melanoma segmentation, 2020.

[62] Adam W. Harley, Konstantinos G. Derpanis, and Iasonas Kokkinos. Segmentation-aware convolutional networks using local attention masks, 2017.

[63] Yizhi Pan, Junyi Xin, Tianhua Yang, Teeradaj Racharak, Le-Minh Nguyen, and Guanqun Sun. A mutual inclusion mechanism for precise boundary segmentation in medical images, 2024.

[64] Tongxue Zhou, Su Ruan, and Stéphane Canu. A review: Deep learning for medical image segmentation using multi-modality fusion, 2020.

[65] Yihao Li, Mostafa El Habib Daho, Pierre-Henri Conze, Rachid Zeghlache, Hugo Le Boité, Ramin Tadayoni, Béatrice Cochener, Mathieu Lamard, and Gwenolé Quellec. A review of deep learning-based information fusion techniques for multimodal medical image classification, 2024.

[66] José Morano, Guilherme Aresta, Christoph Grechenig, Ursula Schmidt-Erfurth, and Hrvoje Bogunović. Deep multimodal fusion of data with heterogeneous dimensionality via projective networks, 2024.

[67] Nico Catalano, Alessandro Maranelli, Agnese Chiatti, and Matteo Matteucci. More than the sum of its parts: Ensembling backbone networks for few-shot segmentation, 2024.

[68] Zunjie Xiao, Xiaoqing Zhang, Risa Higashita, and Jiang Liu. Mm-unet: A mixed mlp architecture for improved ophthalmic image segmentation, 2024.

[69] Weilin Fu, Katharina Breininger, Zhaoya Pan, and Andreas Maier. What do we really need? degenerating u-net on retinal vessel segmentation, 2019.

[70] Alexander Kirillov, Kaiming He, Ross Girshick, Carsten Rother, and Piotr Dollár. Panoptic segmentation. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pages 9404–9413, 2019.

[71] Yiting Wang, Haonan Zhao, Daniel Gummadi, Mehrdad Dianati, Kurt Debattista, and Valentina Donzella. Benchmarking the robustness of panoptic segmentation for automated driving, 2024.

[72] Ping Li, Yu Zhang, Li Yuan, Jian Zhao, Xianghua Xu, and Xiaoqin Zhang. Adversarial attacks on video object segmentation with hard region discovery, 2023.

[73] Tirumalapudi Raviteja and Rajay Vedaraj . I. S. Global image segmentation process using machine learning algorithm convolution neural network method for self- driving vehicles, 2020.

[74] Maurilio Di Cicco, Ciro Potena, Giorgio Grisetti, and Alberto Pretto. Automatic model based dataset generation for fast and accurate crop and weeds detection, 2017.

[75] Abolfazl Meyarian and Xiaohui Yuan. Voting network for contour levee farmland segmentation and classification, 2023.

[76] Wentai Zhang, Joe Joseph, Yue Yin, Liuyue Xie, Tomotake Furuhata, Soji Yamakawa, Kenji Shimada, and Levent Burak Kara. Component segmentation of engineering drawings using graph convolutional networks, 2023.

[77] Hao Wu. What is learning? a primary discussion about information and representation, 2015.

[78] Kelei He, Chen Gan, Zhuoyuan Li, Islem Rekik, Zihao Yin, Wen Ji, Yang Gao, Qian Wang, Junfeng Zhang, and Dinggang Shen. Transformers in medical image analysis: A review, 2022.

[79] Christian Janiesch, Patrick Zschech, and Kai Heinrich. Machine learning and deep learning. *Electronic Markets*, 31(3):685–695, 2021.

[80] Max Ferguson, Ronay Ak, Yung-Tsun Tina Lee, and Kincho H. Law. Detection and segmentation of manufacturing defects with convolutional neural networks and transfer learning, 2018.

[81] Thomas Tsao and Doris Y. Tsao. A topological solution to object segmentation and tracking, 2021.

[82] Gemmechu Hassena, Jonathan Moon, Ryan Fujii, Andrew Yuen, Noah Snavely, Steve Marschner, and Bharath Hariharan. Objectcarver: Semi-automatic segmentation, reconstruction and separation of 3d objects, 2024.

[83] T. Barros, P. Conde, G. Gonçalves, C. Premebida, M. Monteiro, C. S. S. Ferreira, and U. J. Nunes. Multispectral vineyard segmentation: A deep learning approach, 2022.

[84] Han Liu, Hao Li, Xing Yao, Yubo Fan, Dewei Hu, Benoit Dawant, Vishwesh Nath, Zhoubing Xu, and Ipek Oguz. Colossal: A benchmark for cold-start active learning for 3d medical image segmentation, 2023.

[85] Yiqing Shen, Hao Ding, Xinyuan Shao, and Mathias Unberath. Performance and non-adversarial robustness of the segment anything model 2 in surgical video segmentation, 2024.

[86] Dibya Jyoti Bora, Anil Kumar Gupta, and Fayaz Ahmad Khan. Comparing the performance of l*a*b* and hsv color spaces with respect to color image segmentation, 2015.

[87] Shubham Gupta, Uma D., and Ramachandra Hebbar. Analysis and application of multispectral data for water segmentation using machine learning, 2022.

[88] 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.

[89] Steven Landgraf, Markus Hillemann, Theodor Kapler, and Markus Ulrich. A comparative study on multi-task uncertainty quantification in semantic segmentation and monocular depth estimation, 2025.

[90] Lyndon Chan, Mahdi S. Hosseini, and Konstantinos N. Plataniotis. A comprehensive analysis of weakly-supervised semantic segmentation in different image domains, 2020.

[91] Szymon Mazurek, Monika Pytlarz, Sylwia Malec, and Alessandro Crimi. Investigation of energy-efficient ai model architectures and compression techniques for "green" fetal brain segmentation, 2024.

[92] Pablo Hernández-Cámara, Jorge Vila-Tomás, Paula Dauden-Oliver, Nuria Alabau-Bosque, Valero Laparra, and Jesús Malo. Image segmentation via divisive normalization: dealing with environmental diversity, 2024.

[93] Haipeng Li, Dingrui Liu, Yu Zeng, Shuaicheng Liu, Tao Gan, Nini Rao, Jinlin Yang, and Bing Zeng. Single-image-based deep learning for segmentation of early esophageal cancer lesions, 2023.

[94] Ekin Yagis, Shahab Aslani, Yashvardhan Jain, Yang Zhou, Shahrokh Rahmani, Joseph Brunet, Alexandre Bellier, Christopher Werlein, Maximilian Ackermann, Danny Jonigk, Paul Tafforeau, Peter D Lee, and Claire Walsh. Deep learning for vascular segmentation and applications in phase contrast tomography imaging, 2023.

[95] Timo Sämann, Karl Amende, Stefan Milz, Christian Witt, Martin Simon, and Johannes Petzold. Efficient semantic segmentation for visual bird's-eye view interpretation, 2018.

[96] Atif Shah, Maged S. Al-Shaibani, Moataz Ahmad, and Reem Bunyan. A bibliography of multiple sclerosis lesions detection methods using brain mris, 2023.

[97] Yuhan Song, Armagan Elibol, and Nak Young Chong. Abdominal multi-organ segmentation based on feature pyramid network and spatial recurrent neural network, 2023.

[98] Adrian Galdran. Polyp and surgical instrument segmentation with double encoder-decoder networks, 2024.

[99] Khaled Humady, Yasmeen Al-Saeed, Nabila Eladawi, Ahmed Elgarayhi, Mohammed Elmogy, and Mohammed Sallah. Efficient liver segmentation with 3d cnn using computed tomography scans, 2022.

[100] Investigating certain choices of cnn configurations for brain lesion segmentation.

AI-generated, for reference only.

[101] Jinzheng Cai, Yingda Xia, Dong Yang, Daguang Xu, Lin Yang, and Holger Roth. End-to-end adversarial shape learning for abdomen organ deep segmentation, 2019.

[102] Wenwei Zhang, Jiangmiao Pang, Kai Chen, and Chen Change Loy. K-net: Towards unified image segmentation. *Advances in Neural Information Processing Systems*, 34:10326–10338, 2021.

[103] Jikuan Qian, Rui Li, Xin Yang, Yuhao Huang, Mingyuan Luo, Zehui Lin, Wenhui Hong, Ruobing Huang, Haining Fan, Dong Ni, and Jun Cheng. Hasa: Hybrid architecture search with aggregation strategy for echinococcosis classification and ovary segmentation in ultrasound images, 2022.

[104] Lukas Zbinden, Lars Doorenbos, Theodoros Pissas, Adrian Thomas Huber, Raphael Sznitman, and Pablo Márquez-Neila. Stochastic segmentation with conditional categorical diffusion models, 2023.

[105] Ivo M. Baltruschat, Leonhard Steinmeister, Harald Ittrich, Gerhard Adam, Hannes Nickisch, Axel Saalbach, Jens von Berg, Michael Grass, and Tobias Knopp. When does bone suppression and lung field segmentation improve chest x-ray disease classification?, 2018.

[106] Gustav Grund Pihlgren, Konstantina Nikolaidou, Prakash Chandra Chhipa, Nosheen Abid, Rajkumar Saini, Fredrik Sandin, and Marcus Liwicki. A systematic performance analysis of deep perceptual loss networks: Breaking transfer learning conventions, 2024.

[107] Xuebin Qin, Deng-Ping Fan, Chenyang Huang, Cyril Diagne, Zichen Zhang, Adrià Cabeza Sant'Anna, Albert Suàrez, Martin Jagersand, and Ling Shao. Boundary-aware segmentation network for mobile and web applications, 2021.

[108] Iqbal H Sarker. Machine learning: Algorithms, real-world applications and research directions. *SN computer science*, 2(3):160, 2021.

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.

AI-generated, for reference only.