SSRG International Journal of Electrical and Electronics Engineering
ISSN: 2348-8379/ https://doi.org/10.14445/23488379/IJEEE-V11I10P102

Volume 11 Issue 10, 10-21, October 2024

Original Article

Unveiling Precision in Abdominal Organ Segmentation: A Deep Dive into Emerging Deep Learning Paradigms for Single and Multi-modal Image Segmentation

Snehal V. Laddha[1], Rohini S. Ochawar[2]

[1]*Department of Electronics Engineering, School of Electrical and Electronics Engineering, Ramdeobaba University, Maharashtra, India.*
[2]*Department of Electronics and Communication Engineering, School of Electrical and Electronics Engineering, Ramdeobaba University, Maharashtra, India.*

[1]*Corresponding Author : laddhasv2@rknec.edu*

Received: 01 August 2024 Revised: 02 September 2024 Accepted: 03 October 2024 Published: 30 October 2024

Abstract *- Medical imaging is essential for diagnosing and managing diseases that impact human organs. Two widely used imaging techniques, Computed Tomography (CT) and Magnetic Resonance Imaging (MRI), offer valuable information about the structural aspects of organs. However, relying solely on a single imaging modality can sometimes limit disease detection accuracy. To address this limitation, the integration of multiple modalities followed by segmentation has gained traction, offering improved precision in assessing organ health. Accurate segmentation of organs from multi-modal medical images forms the cornerstone of modern healthcare, facilitating precise treatment planning, early disease detection, and personalized medicine. This paper offers an in-depth review of the latest trends and challenges in abdominal organ segmentation using deep learning approaches. It explores the use of attention mechanisms, Graph Neural Networks (GNNs) alongside traditional methods such as Convolutional Neural Networks (CNN), Fully Convolutional Networks (FCN), and Generative Adversarial Networks (GAN) within the scope of abdominal organ segmentation. The paper also addresses key challenges and opportunities in the field, highlighting the importance of continued innovation and collaboration to advance abdominal organ segmentation for improved clinical outcomes.*

Keywords *- Deep learning, Multi-modal images, Segmentation, Abdominal organs, CT, MRI, Liver disease.*

1. Introduction

Medical imaging is a fundamental aspect of modern healthcare, offering critical insights into both the structure and function of human organs. Computed Tomography (CT) and Magnetic Resonance Imaging (MRI) are particularly notable among imaging techniques, providing detailed anatomical data essential for diagnosing diseases and developing treatment plans [1]. However, despite the advancements in imaging technology, accurate segmentation of abdominal organs remains a challenge due to the complex nature of medical images and the inherent variability in organ shapes and sizes.

Abdominal organ segmentation holds immense significance in clinical practice, facilitating precise quantification of organ volumes, early detection of abnormalities, and monitoring of disease progression [2]. Traditional segmentation techniques, based on manually designed features and algorithms, have long been applied in medical image analysis [3]. These methods often struggle to cope with the diverse characteristics of abdominal organs, leading to suboptimal segmentation results, particularly in cases of heterogeneous organs or images with low contrast.

Recently, deep learning techniques have significantly transformed the field of medical image segmentation. Convolutional Neural Networks (CNNs), in particular, have shown exceptional ability to automatically learn complex patterns and features from large datasets, resulting in enhanced segmentation accuracy [5]. Moreover, the flexibility and adaptability of deep learning architectures have enabled the integration of multi-modal imaging data, further enhancing the precision of abdominal organ segmentation [6].

One of the key advancements in deep learning-based segmentation is the integration of attention mechanisms enables models to concentrate on important areas within the input image [5], enabling more precise localization and segmentation of abdominal organs. Additionally, Graph Neural Networks (GNNs) have shown promise in capturing spatial relationships and anatomical structures, particularly in cases where organs exhibit complex interdependencies [8]. Alongside these advanced techniques, conventional models such as CNN, Fully Convolutional Networks (FCNs), and Unet [5,8] continue to play significant roles in abdominal organ segmentation, offering complementary advantages in certain applications.

Despite the remarkable progress achieved with deep learning methods, several challenges persist in abdominal organ segmentation. These include the scarcity of annotated datasets, class imbalance, domain adaptation across different imaging modalities, and the need for interpretable and explainable models for clinical acceptance [9]. Overcoming these challenges demands joint efforts from both researchers and clinicians, along with advancements in data augmentation, transfer learning, and model interpretability techniques.

In light of these considerations, this paper aims to provide a comprehensive review of emerging trends and challenges in abdominal organ segmentation using deep learning techniques. By examining the latest advancements in attention mechanisms and graph neural networks alongside conventional methods such as CNNs, FCNs, and Unet, this paper seeks to elucidate the current landscape of abdominal organ segmentation from single-modality and multi-modal images and chart future directions for research and clinical applications. Through a deeper understanding of these emerging trends and challenges, we can pave the way for more accurate, reliable, and clinically relevant segmentation methods, ultimately enhancing patient care and advancing medical image analysis in abdominal imaging.

2. Fundamentals of Abdominal Organ Segmentation

Abdominal organ segmentation is an essential task in medical imaging, focusing on isolating individual organs from CT or MRI data. Historically, this process has depended on manual or semi-automated methods, which can be time-intensive, subjective, and susceptible to variability between observers. Recently, deep learning-based methods have gained attention as a promising solution for abdominal organ segmentation [11-13].

Figure 1 shows that the process of segmenting abdominal organs from medical images typically follows a structured methodology encompassing various key steps. Initially, the dataset comprising images from MRI, CT, and other modalities is collected and subjected to preprocessing techniques such as normalization, resizing, and noise reduction to ensure consistency and quality. Data augmentation, though optional, may be employed to augment the dataset's diversity, thereby enhancing the model's robustness.

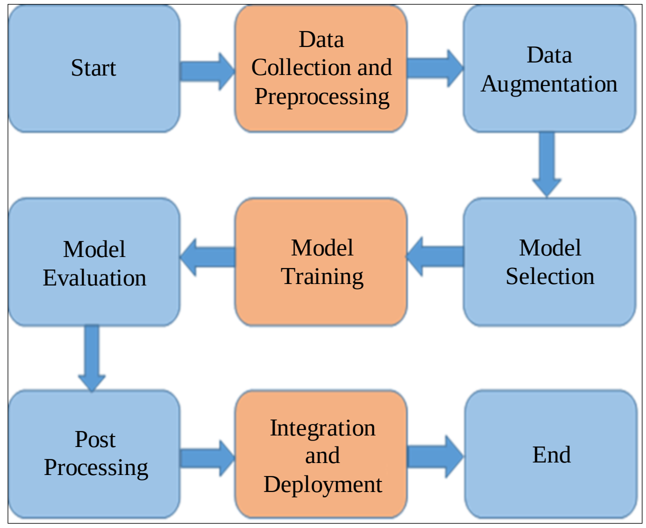

Fig. 1 Steps involved in abdominal organ segmentation using a deep learning approach

Following data preparation, a suitable deep learning architecture is selected based on factors like dataset size and complexity. The selected model is then trained using a portion of the dataset, with the remainder reserved for validation and testing. During training, hyperparameters are fine-tuned to optimize the model's performance. Evaluation of the trained model involves assessing its accuracy and efficacy using metrics like the Dice coefficient and Jaccard index.

Post-processing techniques, such as morphological operations, may be applied to refine the segmented output and improve its quality. Once the model is trained and evaluated, it is integrated into a software pipeline for practical deployment, enabling its use for segmenting new images.

Additionally, a feedback loop may be incorporated to iteratively refine the model's performance based on real-world use cases and feedback from stakeholders. This systematic approach ensures a comprehensive and effective methodology for abdominal organ segmentation using deep learning methods from single-modality and multi-modal images.

Despite the potential of deep learning methods, several challenges exist in abdominal organ segmentation tasks. One major issue is that annotated medical imaging datasets are frequently small and lack diversity, which limits the ability of deep learning models to generalize effectively. To some degree, data augmentation techniques such as geometric transformations and variations in intensity can help mitigate this challenge [15]. Secondly, domain adaptation across different imaging modalities or scanner types remains a significant hurdle, as variations in image acquisition protocols can affect model performance. Transfer learning strategies, where models are pre-trained on large-scale datasets and fine-tuned on smaller target domains, offer a potential solution to this challenge [15].

The selection of model plays an important role in single as well as multi-modal image segmentation; for abdominal organs segmentation, Convolutional Neural Networks (CNNs) have shown remarkable success in learning hierarchical representations from raw imaging data, enabling automatic feature extraction and organ segmentation [16] [17]. Fully Convolutional Networks (FCNs) have extended CNNs to pixel-wise segmentation tasks, allowing for end-to-end learning of organ boundaries without the need for handcrafted features [16].

Furthermore, interpretability and explainability of deep learning models are critical considerations in medical imaging applications. Clinicians require confidence in the decisions made by automated segmentation algorithms, necessitating transparent and interpretable model architectures. Attention mechanisms enable models to concentrate on significant areas within the input image and can enhance interpretability by providing visual explanations for segmentation decisions [19].

DL approaches have demonstrated outstanding performance in single as well as multi-organ segmentation. These methods leverage DL to extract comprehensive features from medical images, capturing intricate structural information pertaining to different organs [18].

3. Basic Deep Learning Models for Abdominal Organ Segmentation

CNN, FCN, and Unet models serve as foundational architectures for abdominal organ segmentation and can be further customized and optimized based on specific requirements and datasets.

3.1. Convolutional Neural Network (CNN)

Convolutional Neural Networks (CNNs) are fundamental tools in medical imaging, particularly for tasks such as abdominal organ segmentation from both single-modality and multi-modal images. CNNs are adept at extracting intricate features and patterns from medical images, allowing for the precise delineation of organ boundaries crucial for diagnosis and treatment planning [10].

In a CNN architecture tailored for abdominal organ segmentation, the hidden layers are tailored to exploit the inherent spatial relationships and structural nuances present in medical images. This adaptation involves layers responsible for convolutions, pooling, activation, and fully connected operations. The arrangement of these layers is crucial for efficient feature extraction, classification, and decision-making stages specific to organ segmentation tasks [12].

The hidden layers of CNNs comprise layers responsible for convolutions. Let I denote the input image or feature map. K represents the convolutional kernel, and O represents the output feature map. The convolution operation is defined as:

$$O[i,j] = \Sigma_m \Sigma_n \mathrm{I[i+,\ j+n].K[m.n]} \qquad (1)$$

Where:
O [i, j] is the value of the pixel at position i,j) in the output feature map.
I [i+m, j+n] represents the pixel values in the input feature map.
K [m, n] denotes the values of the convolutional kernel.
The summation is performed over all possible positions (m,n) of the kernel.

The next operation after convolution is max pooling, which is a downsampling operation that decreases the spatial dimensions of the input feature map by retaining only the maximum value within each pooling window.

Let I be the input feature map, O be the output feature map, and s denote the size of the pooling window. The max pooling operation is defined as:

$$O[i,j] = max_{m,n}\mathrm{[i.s+m,\ j.s+n]} \qquad (2)$$

Where:
O[i,j] is the value of the pixel at position (i, j) in the output feature map.
I [i·s+m, j·s+n] represents the pixel values in the input feature map within the pooling window.
$max_{m,n}$ denotes taking the maximum value over all pixels in the pooling window.

The convolution operation produces a feature map from the input matrix of the layer as the convolution kernel moves across it. This feature map serves as input for the next layer. As the kernel traverses the input matrix, it identifies features pertinent to abdominal organs, which are then utilized in later layers for additional processing. Additionally, pooling layers are strategically employed to downsample feature maps, reducing computational complexity while retaining essential information relevant to segmentation [12]. These layers are followed by others, such as pooling layers, fully connected layers, and normalization layers, as shown in Figure 2.

CNN has proven useful in recognizing images, segmentation, detecting objects, and other areas. Some of the most well-known CNN architectures include Holistically-nested CNN (HNN), MaskR-CNN, DCNN, AlexNet, VGGNet, Residual Networks (ResNet), and DenseNet.

3.2. Fully-Convolutional Neural Network (FCN)

Fully Convolutional Networks (FCNs) are tailored specifically for segmentation tasks and are widely employed in a range of segmentation applications, including medical image segmentation for tasks like abdominal organ segmentation [16].

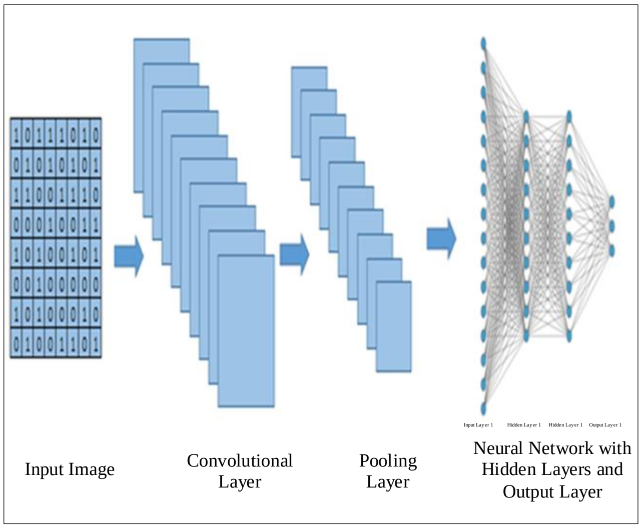

Fig. 2 Convolutional Neural Network (CNN)

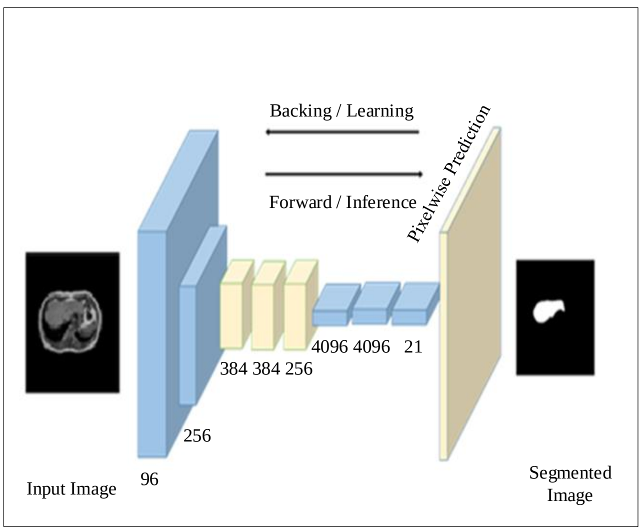

Fig. 3 Fully-Convolutional Neural Network (FCN)

In the case of Fully Convolutional Networks (FCNs) adapted for abdominal organ segmentation, the architecture undergoes modifications to accommodate the segmentation task. FCNs, devoid of fully connected layers, are employed to produce segmentation maps of the same size as the input images.

This is achieved by replacing the fully connected layers in traditional CNNs with convolutional layers. During the segmentation process, the FCN transforms the feature maps from intermediate layers into two-dimensional feature maps of each pixel, retaining spatial information crucial for accurate organ delineation. The FCN utilizes deconvolutional layers to upsample the feature maps, thereby restoring them to the same size as the input images.

This enables pixel-by-pixel classification on the upsampled feature maps, facilitating precise segmentation of abdominal organs. Depending on the desired level of upsampling, FCNs may employ different magnifications, such as FCN-32s, FCN-16s, and FCN-8s, to achieve varying degrees of segmentation granularity [18]. The network architecture of FCN is shown in Figure 3. Some of the most popular FCN models for medical image segmentation are 2D-FCN, 3D-FC, SegNet, etc.

3.3. Generative Adversarial Network (GAN)

GANs hold significant promise in the realm of abdominal organ segmentation, offering avenues for improved accuracy and robustness. Comprising a generator and discriminator network, GANs function in an adversarial setup, where the generator produces synthetic images from random noise, and the discriminator attempts to differentiate between real and generated images.

In the context of abdominal organ segmentation, researchers can adapt this framework by designing a generator network to translate abdominal MRI images into segmented organ masks. Meanwhile, the discriminator is tasked with discerning between genuine organ masks and those generated by the generator, providing feedback to refine the segmentation process. Through combined training, wherein the generator and discriminator networks are optimized concurrently, the model learns to produce realistic organ segmentations. In contrast, the discriminator improves its ability to discern real from fake segmentations, as shown in Figure 4.

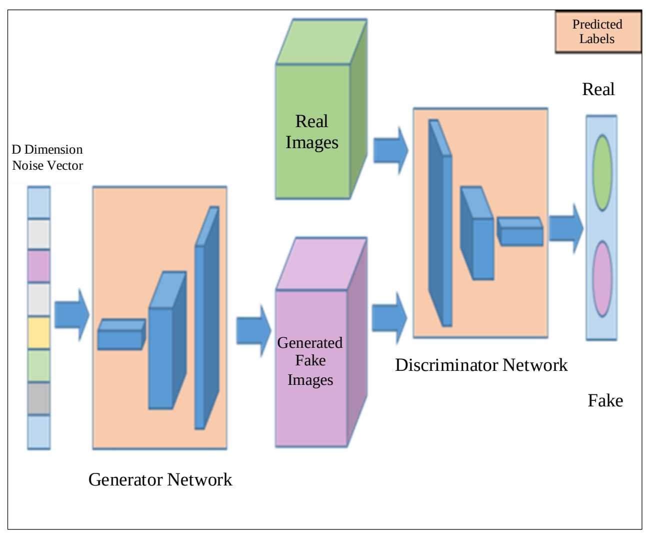

Fig. 4 Generative Adversarial Network (GAN)

Additionally, the integration of convolutional layers for semantic segmentation and the use of multiscale loss functions further enhance the model's capacity to accurately segment the organ structures, mitigating challenges associated with unbalanced pixel categories and anatomical variability commonly encountered in medical imaging datasets. To optimize the segmentation network, many researchers have designed segmentation networks based on the ideas of GAN and a multiscale L1 loss. Some of the popular GAN-based segmentation models are SCAN, PAN, NiftyNet [28], MedGAN, and AsynDGAN.

3.4. Unet

The U-Net architecture, built on the foundation of Fully Convolutional Networks (FCN), was specifically designed for biomedical imaging and became widely adopted for medical image segmentation following its introduction by Ronneberger et al. [8]. The U-Net architecture, as shown in Figure 5, is composed of two main components: a contracting path that captures contextual information and a symmetric expanding path that allows for precise localization.

Convolutional three-fold down-sampling is implemented with an FCN-like architecture for extracting features. The expansion or up-convolution, also known as deconvolution, decreases the number of feature maps while simultaneously enlarging their dimensions. To avoid losing pattern information, feature maps are copied from the down-sampling to the up-sampling parts of the network. A 1x1 convolution is applied to the feature maps in order to generate a segmentation map.

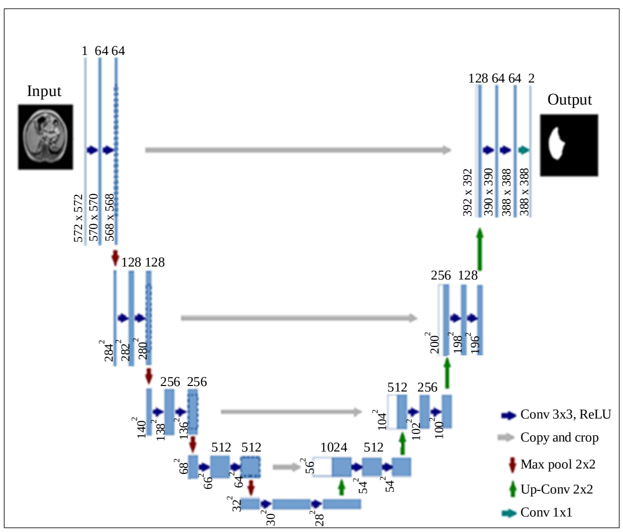

Fig. 5 Unet architecture

Various extensions have been developed for U-Net to support different types of images. A few of these are: Attention Unet, Unet+, Unet++, V-Net, Res U-Net, H-DenseUnet.

4. Attention Mechanisms in Abdominal Organ Segmentation

Attention mechanisms in deep learning have become prominent due to their capability to improve model performance by concentrating on important features and minimizing the influence of less relevant ones. In abdominal organ segmentation, attention mechanisms are essential for enhancing accuracy by enabling the model to focus on the most informative areas of the input image [19].

Several types of attention mechanisms can be integrated into Convolutional Neural Networks (CNNs) and Fully Convolutional Networks (FCNs) for abdominal organ segmentation. One commonly used attention mechanism is the self-attention mechanism, which enables the model to attend to different parts of the input image with varying degrees of importance. This is particularly useful in medical imaging, where certain regions may contain more diagnostically relevant information than others.

Another type of attention mechanism is spatial attention, which focuses on specific spatial locations within the input image. This type of attention is often used in conjunction with self-attention to offer a more thorough understanding of the image context. Some of the popular attention models that have been applied to abdominal organ segmentation are:

4.1. Attention Unet

The Attention U-Net is a Convolutional Neural Network (CNN) architecture employed in semantic segmentation tasks within computer vision. It represents an evolution of the original U-Net model, incorporating attention mechanisms to enhance contextual understanding [19].

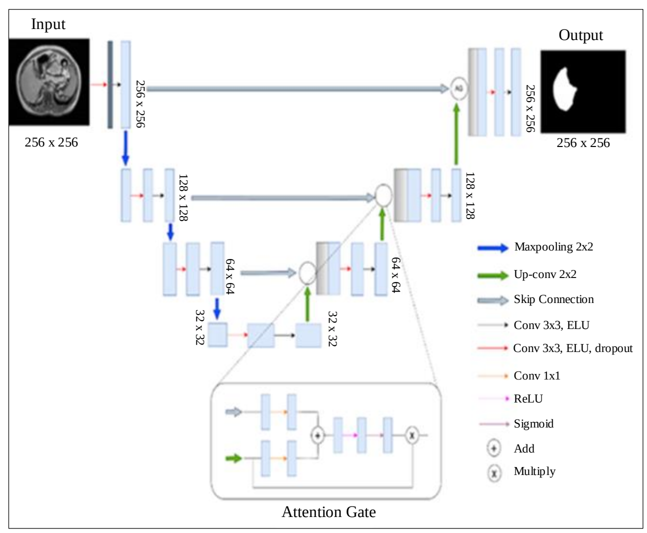

Fig. 6 Attention U-net architecture

Renowned for its precision, even with limited data, the U-Net architecture comprises an encoder network and a decoder network connected by skip links. The encoder gradually reduces spatial resolution, while the decoder employs skip connections to up-sample and merge feature maps from the encoder for the final segmentation.

To improve long-range context comprehension, the Attention U-Net integrates attention mechanisms into the decoder network, as shown in Figure 6. These mechanisms include both spatial attention, which assigns equal importance to all spatial positions within feature maps, and channel-wise attention, which weights each feature map's contribution to the segmentation map [12].

The following formula describes additive attention:

$$q_{att}^{1} = \psi^{T}(\sigma_{1}\left(W_{x}^{T}\left(x_{i}^{l} + W_{g}^{T} g_{i} + b_{i}\right)\right) + b_{\psi} \quad (3)$$

$$\alpha_{i}^{l} = \sigma_{2}(q_{att}^{l}\left(x_{i}^{l}\, , g_{i}\, ;\, \Theta_{att}\right)) \quad (4)$$

Where, g is the gating signal, and x^{l} is the characteristics of the contracting route. The sigmoid function is denoted by the phrase $\sigma_{2}(x_{i,c})$:

$$\sigma_{2}\left(x_{i,c}\right) = \frac{1}{1+\exp(-x_{i,c})} \quad (5)$$

The effectiveness of Attention U-Net extends to abdominal organ segmentation in both single and multi-mode scenarios, where it demonstrates promising performance.

4.2. Residual Attention U- Net

The Residual Attention U-Net is an advanced deep learning architecture that combines the U-Net framework with attention and residual blocks, widely applied in image segmentation tasks [21]. Attention gates within this architecture play a crucial role in diminishing noise and irrelevant background elements in images while emphasizing vital information concerning target objects.

These attention gates utilize spatial information through an attention mechanism, potentially aiding in tasks like early detection, such as identifying small fires. Furthermore, the integration of residual blocks enhances the U-Net's capability to extract finer details from convolutional layers.

Residual blocks, a type of component in deep neural networks, contribute to improving gradient flow and empower the network to learn more intricate information. They incorporate a shortcut link, enabling the network to learn residual mappings by bypassing one or more levels in the network.

The RAUNet architecture integrates the essential components of the UNet, including its contraction and expansion paths, along with an additional bridging element. Within the Residual Attention U-Net framework, attention gates are seamlessly integrated into both encoder and decoder blocks to enable selective focus on crucial features. Through the incorporation of a Residual Attention Module (RAM) at every layer of the encoder and decoder networks, the RAUNet extends the capabilities of the UNet architecture.

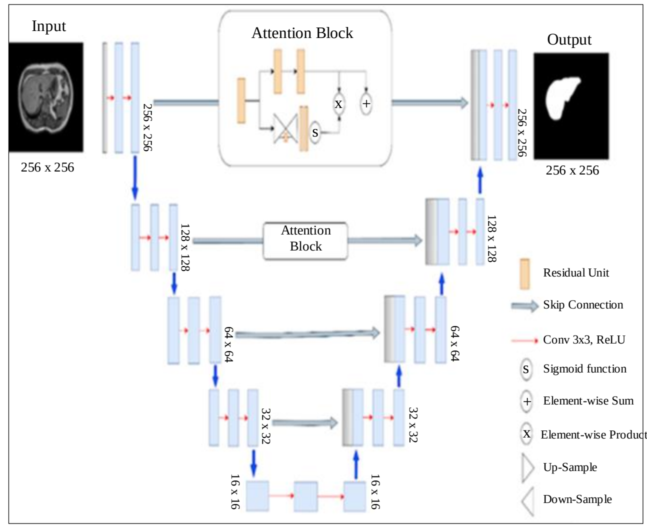

Fig. 7 Residual attention U-net architecture

These RAMs consist of two branches: a mask branch responsible for processing input feature maps to facilitate selective attention and a trunk branch that handles the feature maps themselves. The design of RAMs is aimed at emphasizing important aspects while suppressing irrelevant features, aiding in capturing long-range dependencies within images.

The contraction path in Figure 7 comprises four levels, each housing a residual block containing dilated convolutional layers, batch normalization, and ReLU activation. The bridge layer includes a residual block layer with 512 feature maps of size 16 × 16, facilitating connectivity between the contraction and expansion channels. The output of the bridge layer serves as input for the attention gate on one side while it enters the convolutional transpose layer on the other side. Furthermore, the expansion path consists of four levels, starting with a residual block and concluding with a convolutional transpose layer. Each layer in the expansion path is connected to its corresponding layer in the contraction path through skip connections and attention gates, with the convolutional transpose (2x2) layer employed for upsampling.

4.3. Spatial Attention U-Net

Much like the original U-Net, the Spatial Attention U-Net is built with an encoder-decoder framework, depicted in Figure 8. The encoder component systematically decreases the spatial resolution of the input image using a sequence of convolutional and pooling layers. In contrast, the decoder component upscales the feature maps to generate the final segmentation mask. Skip connections are utilized to combine feature maps from the encoder with the corresponding layers in the decoder. These skip connections promote the transfer of low-level and high-level features across various scales, allowing the network to capture fine-grained details while maintaining contextual information. Spatial attention modules are integrated into the architecture to dynamically adjust the importance of different spatial locations within the feature maps. These attention mechanisms assist the network in concentrating on relevant regions while suppressing irrelevant background information.

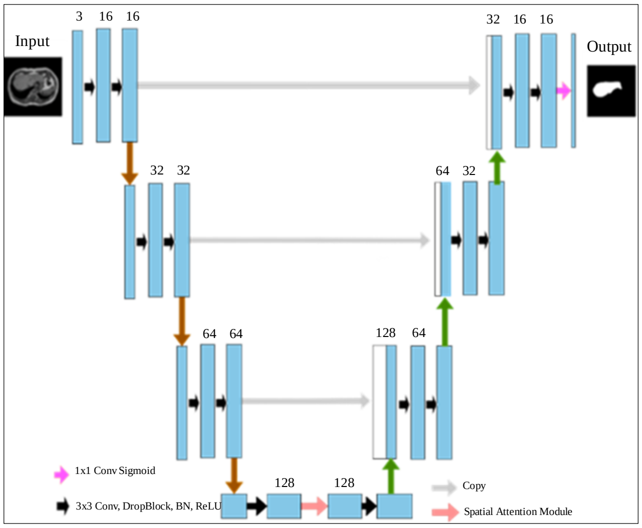

Fig. 8 Spatial attention U-net architecture

In the context of abdominal organ segmentation, spatial attention allows the network to prioritize features within the abdominal region and better delineate organ boundaries. The Spatial Attention U-Net captures multiscale context information by incorporating features from multiple levels of abstraction. This enables the network to effectively handle variations in organ size, shape, and appearance within the abdominal region, improving segmentation performance.

5. Graph Neural Networks (GNN) for Abdominal Organ Segmentation

Graph Neural Networks (GNNs) are a type of neural network uniquely crafted to work with graph-structured data, where entities are depicted as nodes, and the relationships between them are illustrated as edges [22]. GNNs excel in capturing complex relationships and dependencies present in data by leveraging the inherent structure of graphs. Unlike conventional neural networks that function on grid-like structures such as images or sequences, GNNs are capable of effectively modeling non-Euclidean domains, making them ideal for tasks that involve relational data.

GNNs have emerged as a promising approach for segmentation tasks, including abdominal organ segmentation in medical imaging. In the context of segmentation, GNNs leverage the structural information encoded in the graph representation of the image to capture spatial relationships and contextual dependencies between image elements. One of the key advantages of using GNNs for segmentation is their ability to model long-range dependencies and capture global context within the image. Traditional CNNs are limited by their local receptive fields, which may struggle to incorporate contextual information from distant image regions. In contrast, GNNs can propagate information across the entire graph structure, allowing them to capture global relationships and improve segmentation accuracy.

Moreover, GNNs can adaptively aggregate features from neighboring nodes in the graph, enabling them to incorporate both local and global contexts into the segmentation process [23]. This adaptive feature aggregation mechanism allows GNNs to effectively handle variations in organ shape, size, and appearance, making them robust to anatomical differences across different patients and imaging modalities. Some popular GNN models for medical image segmentation are:

5.1. Graph Convolutional Networks (GCNs)

Graph Convolutional Networks (GCNs) are a class of neural networks that operate directly on graph-structured data, enabling efficient feature extraction and propagation across nodes [24]. In the context of abdominal organ segmentation, GCNs offer a powerful framework for leveraging the spatial relationships between different regions of interest within medical images represented as graphs.

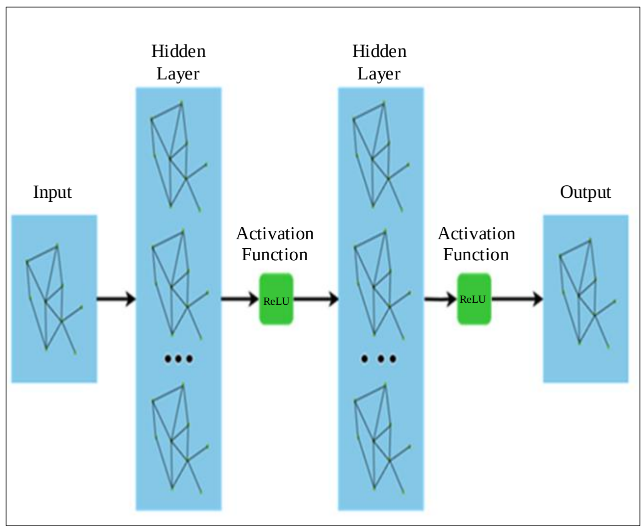

Fig. 9 GCN architecture

Figure 9, at its core, the GCN architecture begins with an input layer that takes in the graph representation of the medical image, where each node represents a pixel or voxel and edges denote spatial relationships. The graph convolutional layer applies convolutional operations directly on the graph, enabling the extraction of informative features from neighboring nodes. These features are subsequently processed through an activation function to introduce non-linearity and improve the network's ability to represent complex patterns. Optionally, pooling layers may be incorporated to downsample the graph and reduce computational complexity while preserving essential information.

Finally, the output layer produces the segmentation map, assigning a label to each pixel or voxel based on the features learned throughout the network. This architecture allows GCNs to effectively capture spatial dependencies within medical images and produce accurate segmentations of abdominal organs, contributing to advancements in medical image analysis and diagnosis.

5.2. Graph U-Net

Graph U-Net represents a significant advancement in the realm of semantic segmentation within medical image analysis. Built upon the foundation of the U-Net architecture, Graph U-Net introduces graph convolutional layers to better leverage the inherent spatial dependencies and structural nuances present in medical images, which are naturally represented as graphs [25]. In this paradigm, nodes correspond to image elements like pixels or voxels, while edges encapsulate the spatial relationships between them. By embracing this graph-based representation, Graph U-Net can effectively capture intricate spatial relationships and contextual cues crucial for accurate segmentation of anatomical structures.

The architecture of Graph U-Net adheres to the encoder-decoder framework commonly seen in segmentation networks. However, unlike traditional approaches that directly process the image, Graph U-Net operates on the graph structure, enabling it to model spatial dependencies among image elements intricately. This novel approach empowers the model to adaptively adjust its segmentation strategy by performing convolution operations directly on the graph representation of the image. These graph convolutional layers enable the collection of information from adjacent nodes, allowing the model to capture both local and global contextual information [25]. As a result, Graph U-Net excels in scenarios where traditional convolutional neural networks struggle to capture complex spatial relationships, leading to more accurate segmentation outcomes across various medical imaging tasks, including organ segmentation, tumor delineation, and structural analysis.

GCNs and Graph U-Net can offer powerful frameworks for abdominal organ segmentation by leveraging graph-based representations to model the spatial relationships and structural dependencies present in medical imaging data. By effectively capturing the complex interactions between different anatomical structures, these models can improve segmentation accuracy and facilitate more reliable diagnosis and treatment planning in clinical practice.

6. Discussion

Various DL techniques based on distinct network architectures have been experimented with for single vs. multi-organs and single vs. multi-mode image segmentation of abdominal organs. The Dice Score is the most widely used metric for assessing the performance of segmentation algorithms. The Dice score is crucial in abdominal organ segmentation as it quantitatively evaluates the algorithm's ability to outline organ boundaries accurately. This score is calculated by comparing the overlap between the predicted segmentation and the ground truth segmentation. It provides a measure of the algorithm's performance in accurately identifying and segmenting different abdominal organs:

Dice Score (DICE): The Dice score is evaluated as

$$\mathrm{DICE}\,(\mathrm{A},\mathrm{B}) = \frac{2\,|\mathrm{A} \cap \mathrm{B}|}{|\mathrm{A}| + |\mathrm{B}|} \tag{6}$$

Where, A is the prediction voxel set, and the ground-truth set is B.

The Dice score ranges from 0 to 1, with a perfect segmentation resulting in a score of 1. It is also called as Dice Similarity Coefficient (DSC) or Dice Coefficient (DC). The Dice score offers a straightforward and intuitive measure of segmentation accuracy. It enables researchers and clinicians to compare and evaluate various segmentation algorithms, aiding in the selection of the most suitable method for specific clinical applications.

We have examined key studies that have made significant contributions to abdominal segmentation and presented in Tables 1 and 2. Analysis of various state-of-the-art deep learning networks experimented by the researchers for image segmentation work on abdominal organs has been summarized in Table 1. In contrast, Table 2 provides an analysis of different deep learning methods for multi-modal abdominal organ segmentation.

The analysis presented in Tables 1 and 2 offers a comprehensive overview of Deep Learning (DL) approaches for abdominal organ segmentation, both in single-modality and multi-modal settings. Single modality segmentation studies encompass a diverse range of DL architectures, from traditional CNNs to more advanced models like Attention-Unet and 3D-FCN. These studies demonstrate varying degrees of segmentation accuracy, with notable achievements such as Hu et al. showcasing remarkably high Dice scores for the liver, spleen, and kidneys using CNN. However, performance variability is evident across different organs and architectures, as seen in the contrasting results reported by Gabriel Efrain et al. for kidney segmentation. The emergence of novel architectures like Attention-Unet and 3D-2D-FCN highlights the ongoing innovation in network design to improve segmentation accuracy and robustness. Moreover, the predominance of CT imaging in single-modality studies underscores its importance and prevalence in abdominal imaging tasks. However, there's a growing interest in MRI-based segmentation, as demonstrated by Valindria et al. and Gross et al.

Table 1. Analysis of DL approaches for single modality abdominal organ segmentation

Author	Year	Network	Abdominal Organ	Modality	Dice Score
Hu [13]	2016	CNN	Multi-Organ	CT	96.0% (Liver), 94.2% (Spleen) and 95.4% (both kidneys)
Gabriel Efrain [14]	2017	CNN	Multi-organ	CT	59% and 58% (kidneys), 83% (liver)
Roth [16]	2018	3D FCN	Multi-Organ	CT	90%
Oktay [19]	2018	Attention-Unet	Pancreas	CT	81.48 %
Rafiei [26]	2018	3D-2D-FCN	Liver	CT	93.52%
Valindria [17]	2018	FCN	Multi-Organ	MRI	90%
Kakeya [27]	2018	3D U-JAPA-net	Multi-Organ	CT	97.1% (liver), 96.9% (spleen), 97.5% (right kidney) & 98.4% (left kidney)
Gibson [28]	2018	NiftyNet	Multi-Organ	CT	95% (liver)
Gibson [29]	2018	DenseVnet	Multi-Organ	CT	81%
M. Ahmad [30]	2019	DBN-DNN	Liver	CT	94.80%
Y. C. Tang [31]	2019	DCNN	Multi-Organ	CT	94%
Chung [32]	2019	FCN	Liver	CT	96%,
Y. Y. Zhou [33]	2019	DMPCT	Multi-Organ	CT	96.15%
Tong [34]	2020	Self-paced DenseNet	Multi-Organ	CT	84%
Lin [5]	2021	Attention U-Net	Multi-Organ	CT	81.3%
Xu, M [35]	2022	nnUNet	Liver	CT	95.8%
Kayhan [36]	2022	3D Unet based on Early Fusion Mechanism	Multi-Organ	CT	95% (spleen), 96% (liver)
Nanyan Shen [20]	2023	Spatial attention and deformable convolution	Multi-Organ	CT	80.46%
Tian [23]	2023	Surface-GCN	Prostate	CT	94.49%
Gross [37]	2024	3D-DCNN	Liver	MRI	97%

In contrast, multi-modal segmentation approaches leverage complementary information from multiple imaging modalities to improve segmentation accuracy and resilience. The integration of CT, MRI, and even EM data allows for a more comprehensive analysis of abdominal organs, leading to improved Dice scores compared to single-modality approaches. Studies such as Valindria et al. and Conze et al. showcase the effectiveness of combining CT and MRI data in achieving higher segmentation accuracy, with Dice scores consistently surpassing 90%. Additionally, the introduction of novel architectures like UNet++ and cGv19pUNet1-1 reflects the evolving landscape of DL models for multi-modal segmentation tasks, promising further advancements in accuracy and clinical applicability.

These findings pave the way for future research directions aimed at enhancing the accuracy, robustness, and clinical utility of DL-based abdominal organ segmentation techniques, ultimately benefiting patients and healthcare practitioners alike.

Table 2. Analysis of DL approaches for multi-modal abdominal organ segmentation

Author	Year	Network	Abdominal Organ	Modality	Dice Score
Valindria [17]	2018	FCN	Liver	Multimodal (CT & MRI)	90.7% (MRI), 94.3% (CT)
Bobo [38]	2018	FCNN	Multi-Organ	Multimodal (CT & MRI)	91.3% (liver)
Mulay S. [39]	2019	Mask R-CNN	Liver	Multimodal (CT & MRI)	94% (CT) 89% (T2-weighted MRI) and 91% (T1-weighted MRI0)
Z. Zhou [40]	2020	UNet++	Liver	Multimodal (CT, MRI & EM)	94.12%
Conze [41]	2020	cGv19pUNet1-1	Multi-Organ	Multimodal (CT & MRI)	97.95%
Elghazy [21]	2023	Triple stream UNET model	Liver	Multimodal (MRI T1 Dual)	96%

7. Conclusion and Future Scope

This research paper presents an in-depth examination of Deep Learning (DL) techniques for abdominal organ segmentation, emphasizing the progress achieved in both single-modality and multi-modal imaging contexts. Through the examination of various DL models such as CNN, FCN, Unet, Attention-Unet, and others, it is evident that DL techniques have significantly enhanced the accuracy and efficiency of abdominal organ segmentation tasks. From achieving high Dice scores in single-modality CT and MRI imaging to effectively handling the complexities of multi-modal imaging, these models demonstrate the versatility and robustness required for real-world clinical applications.

As deep learning continues to evolve, integrating various techniques offers exciting prospects for enhancing abdominal organ segmentation. By combining attention mechanisms, Graph Neural Networks (GNNs), and existing Convolutional Neural Network (CNN) architectures, researchers can unlock new avenues for improving segmentation accuracy and robustness. Attention mechanisms have demonstrated their efficacy in focusing on relevant features during segmentation tasks. By integrating attention modules into CNNs and FCNs, researchers can enhance model interpretability and capture fine-grained details crucial for accurate organ segmentation.

Future research may explore novel attention mechanisms tailored to the specific challenges of abdominal imaging, such as handling anatomical variations and image artifacts. GNNs offer a unique framework for modeling spatial dependencies and structural information in medical images represented as graphs. By treating pixels or voxels as nodes and their spatial relationships as edges, GNNs can effectively capture the underlying anatomical structures and their interactions.

Integrating GNNs into segmentation pipelines enables the exploitation of global context and anatomical priors, leading to more robust segmentation performance across diverse patient populations and imaging modalities. The future of abdominal organ segmentation lies in the development of hybrid architectures that synergistically combine multiple techniques. By integrating attention mechanisms, GNNs, and traditional CNNs, researchers can leverage the complementary strengths of each approach to improve segmentation accuracy and generalization.

Additionally, ensemble learning techniques, which aggregate predictions from multiple models, offer a promising avenue for further enhancing segmentation robustness and reliability in clinical settings. As deep learning techniques mature, the translation of research findings into clinical practice becomes increasingly important.

Future directions should focus on validating proposed segmentation methods on large-scale clinical datasets and conducting rigorous evaluations against ground truth annotations. Moreover, the integration of segmentation algorithms into clinical workflows, such as computer-aided diagnosis systems, can streamline the interpretation of medical images and improve diagnostic accuracy for abdominal pathologies.

Availability of Data and Materials

The datasets, software, and materials used in this research are available upon request to interested parties for purposes of replication, verification, or further scientific inquiry. Requests for access to these resources can be made to the corresponding author.

Authors' Contributions

S.V. Laddha: Conceptualization, Methodology, Investigation, Formal analysis, Data curation, Writing-Original draft preparation. R. S. Ochawar: Supervision, Writing- Reviewing and Editing, project administration

References

[1] Simon R. Cherry, "Multimodality in Vivo Imaging Systems: Twice the Power or Double the Trouble?," *Annual Review of Biomedical Engineering*, vol. 8, pp. 35-62, 2006. [CrossRef] [Google Scholar] [Publisher Link]

[2] Ali Emre Kavur et al., CHAOS - Combined (CT-MR) Healthy Abdominal Organ Segmentation Challenge Data, Zenodo, 2019. [Online]. Available: https://zenodo.org/records/3431873

[3] Yang Lei et al., "Deep Learning in Multi-Organ Segmentation," *arXiv*, 2001. [CrossRef] [Google Scholar] [Publisher Link]

[4] U.S. Food & Drug Administration, Medical Imaging, 2018. [Online]. Available: https://www.fda.gov/radiation-emitting-products/radiation-emitting-products-and-procedures/medical-imaging

[5] Haoneng Lin et al., "Variance-Aware Attention U-Net for Multi-Organ Segmentation," *Medical Physics*, vol. 48, no. 12, pp. 7864-7876, 2021. [CrossRef] [Google Scholar] [Publisher Link]

[6] Zhe Guo et al., "Medical Image Segmentation Based on the Multi-Modal Convolutional Neural Network: Study on Image Fusion Schemes," *2018 IEEE 15th International Symposium on Biomedical Imaging (ISBI 2018)*, Washington, pp. 903-907, 2018. [CrossRef] [Google Scholar] [Publisher Link.]

[7] Vanya V. Valindria et al., "Small Organ Segmentation in Whole-Body MRI Using a Two-Stage FCN and Weighting Schemes," *Machine Learning in Medical Imaging*, pp. 346-354, 2018. [CrossRef] [Google Scholar] [Publisher Link]

[8] Olaf Ronneberger, Philipp Fischer, and Thomas Brox, "U-Net: Convolutional Networks for Biomedical Image Segmentation," *Proceedings of the International Conference on Medical Image Computing and Computer-Assisted Intervention*, Munich, Germany, pp. 234-241, 2015. [CrossRef] [Google Scholar] [Publisher Link]

[9] Holger R. Roth et al., "DeepOrgan: Multi-Level Deep Convolutional Networks for Automated Pancreas Segmentation," *Medical Image Computing and Computer-Assisted Intervention*, pp. 556-564, 2015. [CrossRef] [Google Scholar] [Publisher Link]

[10] William Thong et al., "Convolutional Networks for Kidney Segmentation in Contrast-Enhanced CT Scans," *Computer Methods in Biomechanics and Biomedical Engineering: Imaging & Visualization*, vol. 6, no. 3, 2018. [CrossRef] [Google Scholar] [Publisher Link]

[11] Jinzheng Cai et al., Pancreas Segmentation in MRI Using Graph-Based Decision Fusion on Convolutional Neural Networks, *Medical Image Computing and Computer-Assisted Intervention*, pp. 442-450, 2016. [CrossRef] [Google Scholar] [Publisher Link]

[12] Mans Larsson, Yuhang Zhang, and Fredrik Kahl, "Deepseg: Abdominal Organ Segmentation Using Deep Convolutional Neural Networks," *Swedish Symposium on Image Analysis*, vol. 2016, 2016. [Google Scholar] [Publisher Link]

[13] Peijun Hu et al., 'Automatic Abdominal Multi-Organ Segmentation Using Deep Convolutional Neural Network and Time-Implicit Level Sets," *International Journal of Computer Assisted Radiology and Surgery*, vol. 12, no. 3, pp. 399-411, 2016. [CrossRef] [Google Scholar] [Publisher Link]

[14] Gabriel Efrain Humpire Mamani et al., "Organ Detection in Thorax Abdomen CT Using Multi-Label Convolutional Neural Networks," *Proceedings of the SPIE Medical Imaging 2017: Computer-Aided Diagnosis*, vol. 10134, 2017. [CrossRef] [Google Scholar] [Publisher Link]

[15] Xiaoming Liu et al., "Automatic Organ Segmentation for CT Scans Based on Super-Pixel and Convolutional Neural Networks," *Journal of Digital Imaging*, vol. 31, pp. 748-760, 2016. [CrossRef] [Google Scholar] [Publisher Link]

[16] Holger R. Roth et al., "A Multiscale Pyramid of 3D Fully Convolutional Networks for Abdominal Multi-organ Segmentation," *Medical Image Computing and Computer Assisted Intervention*, pp. 417-425, 2018. [CrossRef] [Google Scholar] [Publisher Link]

[17] Vanya V. Valindria et al., "Multi-Modal Learning from Unpaired Images: Application to Multi-organ Segmentation in CT and MRI," *IEEE Winter Conference on Applications of Computer Vision (WACV)*, USA, pp. 547-556, 2018. [CrossRef] [Google Scholar] [Publisher Link]

[18] Ozan Oktay et al., "Attention U-Net: Learning Where to Look for the Pancreas," *arxiv*, 2018. [CrossRef] [Google Scholar] [Publisher Link]

[19] Nanyan Shen et al., "Multi-Organ Segmentation Network for Abdominal CT Images Based on Spatial Attention and Eformable Convolution," *Expert Systems with Applications*, vol. 211, 2023. [CrossRef] [Google Scholar] [Publisher Link]

[20] Hagar Louye Elghazy, and Mohamed Waleed Fakhr, "Dual-and Triple-Stream RESUNET/UNET Architectures for Multi-Modal Liver Segmentation," *IET Image Processing*, vol. 17, no. 4, pp. 1224-1235, 2023. [CrossRef] [Google Scholar] [Publisher Link]

[21] Amit Aflalo et al., "Deepcut: Unsupervised Segmentation Using Graph Neural Networks Clustering," *Proceedings of the IEEE/CVF International Conference on Computer Vision workshops*, pp. 32-41, 2023. [Google Scholar] [Publisher Link]

[22] Juntao Jiang et al., "ViG-UNet: Vision Graph Neural Networks for Medical Image Segmentation," *IEEE 20th International Symposium on Biomedical Imaging (ISBI)*, Cartagena, Colombia, pp. 1-5, 2023. [CrossRef] [Google Scholar] [Publisher Link]

[23] Si Zhang et al., "Graph Convolutional Networks: A Comprehensive Review," *Computational Social Networks*, vol. 6, 2019. [CrossRef] [Google Scholar] [Publisher Link]

[24] Hongyang Gao, and Shuiwang Ji, "Graph U-Nets," *International Conference on Machine Learning Research*, pp. 2083-2092, 2019. [Google Scholar] [Publisher Link]

[25] Shima Rafiei et al., "Liver Segmentation in CT Images Using Three-Dimensional to Two-Dimensional Fully Convolutional Network," *25th IEEE International Conference on Image Processing (ICIP)*, Athens, Greece, pp. 2067-2071, 2018. [CrossRef] [Google Scholar] [Publisher Link]

[26] Hideki Kakeya, Toshiyuki Okada, and Yukio Oshiro, "3D U-JAPA-Net: Mixture of Convolutional Networks for Abdominal Multi-Organ CT Segmentation," *Medical Image Computing and Computer Assisted Intervention*, pp. 426-433, 2018. [CrossRef] [Google Scholar] [Publisher Link]

[27] Eli Gibson et al., "NiftyNet: A Deep-Learning Platform for Medical Imaging, Computer Methods and Programs in Biomedicine," *Computer Methods and Programs in Biomedicine*, vol. 158, pp. 113-122, 2018. [CrossRef] [Google Scholar] [Publisher Link]

[28] Eli Gibson et al., "Automatic Multi-Organ Segmentation on Abdominal CT with Dense V-Networks," *IEEE Transactions on Medical Imaging*, vol. 37, no. 8, pp. 1822-1834, 2018. [CrossRef] [Google Scholar] [Publisher Link]

[29] Mubashir Ahmad et al., "Deep Belief Network Modeling for Automatic Liver Segmentation," *IEEE Access*, vol. 7, pp. 20585-20595, 2019. [CrossRef] [Google Scholar] [Publisher Link]

[30] Yucheng Tang et al., "Improving Splenomegaly Segmentation by Learning from Heterogeneous Multi-Source Labels," *Proceedings of SPIE - The International Society for Optical Engineering*, 2019. [CrossRef] [Google Scholar] [Publisher Link]

[31] Minyoung Chung et al., "Deeply Self-Supervised Contour Embedded Neural Network Applied to Liver Segmentation," *Computer Methods and Programs in Biomedicine*, vol. 192, 2020. [CrossRef] [Google Scholar] [Publisher Link]

[32] Yuyin Zhou et al., "Semi-Supervised 3D Abdominal Multi-Organ Segmentation via Deep Multi-Planar Co-Training," *IEEE Winter Conference on Applications of Computer Vision (Wacv)*, pp. 121-140, 2019. [CrossRef] [Google Scholar] [Publisher Link]

[33] Nuo Tong, et al., "Self-Paced DenseNet with Boundary Constraint for Automated Multi-Organ Segmentation on Abdominal CT Images," *Physics in Medicine & Biology*, vol. 65, no. 13, 2020. [CrossRef] [Google Scholar] [Publisher Link]

[34] Minfeng Xu et al., "A New Probabilistic V-Net Model with Hierarchical Spatial Feature Transform for Efficient Abdominal Multi-Organ Segmentation," *arXiv*, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[35] Beyza Kayhan, and Sait Ali Uymaz, "Abdominal Multi-Organ Segmentation Using 3D Unet Based on Early Fusion Mechanism," *SSRN*, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[36] Moritz Gross et al., "Automated MRI Liver Segmentation for Anatomical Segmentation, Liver Volumetry, and the Extraction of Radiomics," *European Radiology*, vol. 34, pp. 5056-5065, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[37] Meg F. Bobo et al., "Fully Convolutional Neural Networks Improve Abdominal Organ Segmentation," *Proceedings of the SPIE Medical Imaging 2018: Image Processing*, vol. 10574, 2018. [CrossRef] [Google Scholar] [Publisher Link]

[38] Supriti Mulay et al., "Liver Segmentation from Multi-modal Images Using HED-Mask R-CNN," *Multiscale Multi-modal Medical Imaging*, pp. 68-75, 2019. [CrossRef] [Google Scholar] [Publisher Link]

[39] Zongwei Zhou et al., "UNet++: Redesigning Skip Connections to Exploit Multiscale Features in Image Segmentation," *IEEE Transactions on Medical Imaging*, vol. 39, no. 6, pp. 1856-1867, 2020. [CrossRef] [Google Scholar] [Publisher Link]

[40] Pierre-Henri Conze et al., "Abdominal Multi-Organ Segmentation with Cascaded Convolutional and Adversarial Deep Networks," *Artificial Intelligence in Medicine*, vol. 117, 2021. [CrossRef] [Google Scholar] [Publisher Link]

[41] Pauline Luc et al., "Semantic Segmentation Using Adversarial Networks," *arXiv*, 2016. [CrossRef] [Google Scholar] [Publisher Link]