

# Liver Tumour Segmentation Using CNN

**Abstract:** In this study, we propose a comprehensive investigation into liver tumor segmentation utilizing various Convolutional Neural Network (CNN) architectures. Liver tumor segmentation plays a pivotal role in medical imaging for accurate diagnosis and treatment planning. We aim to compare the performance of three prominent CNN architectures, namely U-Net, DeepLabv3+, and Attention U-Net, in segmenting liver tumors from abdominal computed tomography (CT) scans. Our approach involves preprocessing CT images to enhance contrast and remove noise, followed by training and evaluation of the CNN models using a dataset comprising annotated CT scans. We employ common evaluation metrics such as Dice similarity coefficient (DSC), sensitivity, specificity, and Hausdorff distance to quantify the segmentation accuracy of each model. Additionally, we investigate the impact of different input modalities, such as contrast-enhanced and non-contrast-enhanced CT scans, on the segmentation performance of the CNN architectures.

Through extensive experimentation and analysis, we aim to provide insights into the strengths and weaknesses of each CNN architecture for liver tumor segmentation

## 1. Introduction

Liver tumors pose a significant challenge in the field of medical imaging and diagnosis. Accurate segmentation of liver tumors from medical images plays a crucial role in treatment planning, monitoring, and assessing patient outcomes. With the advent of deep learning techniques, particularly Fully Convolutional Networks (FCN), U-Net, and SegNet, there has been a paradigm shift in medical image segmentation, offering promising results in various clinical applications.

In the realm of medical imaging, the accurate segmentation of liver tumors plays a pivotal role in diagnosis, treatment planning, and monitoring the progression of diseases. With the advent of deep learning techniques, particularly Convolutional Neural Networks (CNNs), there has been a significant leap in the precision and efficiency of medical image segmentation tasks. This project aims to explore and compare the efficacy of different CNN architectures in segmenting liver tumors from medical images, thereby contributing to enhanced clinical decision-making processes.

The comparative study not only assesses the quantitative performance metrics but also evaluates the robustness and generalization capabilities of the models across different imaging modalities and tumor characteristics. Furthermore, we explore strategies for model optimization and parameter tuning to enhance segmentation accuracy and reduce false positives. The findings of this research could contribute to the development of more reliable and efficient computer-aided diagnosis systems for liver cancer detection and treatment monitoring, ultimately improving patient outcomes in clinical practice.

**Keywords:** Liver Segmentation, Liver Tumor Detection, FCN (Fully Convolutional Network), U-Net Architecture, SegNet Architecture, ResNet Architecture, PsPNet Architecture, DenseNet Architecture, Unet++ Architecture, DeepLab Architecture, VNet Architecture, HRNet Architecture, Medical Image Segmentation, Convolutional Neural Networks (CNNs), Semantic Segmentation, Training Data Augmentation, Evaluation Metrics (Dice coefficient), Liver Tumor Segmentation Dataset

Liver tumors, comprising both benign and malignant lesions, present a formidable challenge in medical imaging due to their diverse shapes, sizes, and heterogeneous appearances. Traditional segmentation methods often struggle to delineate these tumors accurately, particularly in cases where the boundary between the tumor and healthy liver tissue is ambiguous. Deep learning-based approaches, on the other hand, offer promising solutions by leveraging the hierarchical features learned directly from data.

This project will investigate several state-of-the-art CNN architectures, including but not limited to U-Net, DeepLabv3, and SegNet, each renowned for its unique design principles and performance in medical image segmentation tasks. By implementing and fine-tuning these architectures on a dataset comprising diverse liver tumor images obtained from various modalities such as CT scans and MRI, we aim to evaluate their effectiveness in accurately segmenting tumors while minimizing false positives and false negatives.

Furthermore, this study will delve into the challenges associated with liver tumor segmentation, such as class

imbalance, intra- and inter-patient variability, and the presence of artifacts in medical images. Addressing these challenges is crucial for the development of robust segmentation models that can generalize well across different patient cohorts and imaging conditions.

The outcomes of this project hold immense potential to revolutionize clinical workflows by providing radiologists and oncologists with reliable tools for precise tumor localization, volumetric analysis, and treatment response assessment. Moreover, the insights gained from comparing various CNN architectures can guide future research directions towards the development of more advanced and efficient algorithms for liver tumor segmentation. Each deep learning architecture will be implemented using a suitable framework such as TensorFlow or PyTorch.

Medical image analysis plays a pivotal role in modern healthcare, aiding clinicians in diagnosing and treating various conditions. Liver tumor segmentation from medical images, particularly from computed tomography (CT) or magnetic resonance imaging (MRI) scans, holds significant importance for patient management. Accurate delineation of liver tumors assists clinicians in surgical planning, radiation therapy, and assessing treatment response. However, manual segmentation by radiologists is labor-intensive, time-consuming, and subject to inter-observer variability. To address these challenges, automated methods based on deep learning architectures have emerged as promising solutions.

This project focuses on evaluating the performance of a range of state-of-the-art deep learning architectures for liver tumor segmentation. Leveraging convolutional neural networks (CNNs), which have shown remarkable success in various computer vision tasks, we explore their effectiveness in accurately delineating liver tumors from medical images.

The architectures selected for investigation encompass a diverse range of design philosophies and innovations tailored to address challenges specific to medical image segmentation. We begin with the Fully Convolutional Network (FCN), which revolutionized the field by enabling end-to-end pixel-wise segmentation using convolutional layers. FCN's ability to process input images of arbitrary size and produce dense pixel-wise predictions makes it a foundational model for semantic segmentation tasks.

Moving beyond FCN, we delve into the U-Net architecture, which has garnered widespread adoption in medical imaging due to its unique encoder-decoder

structure with skip connections. These skip connections facilitate the propagation of fine-grained spatial information from the encoder to the decoder, enabling precise localization of objects, such as liver tumors, while retaining contextual information.

In parallel, we explore SegNet, a network architecture designed for efficient segmentation, particularly in scenarios with limited computational resources. SegNet employs an encoder-decoder architecture but focuses on memory efficiency by utilizing max-pooling indices for up sampling during decoding, thereby reducing memory footprint without sacrificing segmentation accuracy.

ResNet, renowned for its deep residual connections, is another architecture under scrutiny. By introducing skip connections that bypass multiple convolutional layers, ResNet mitigates the vanishing gradient problem associated with training deep networks, facilitating the training of exceptionally deep architectures for improved feature learning and segmentation performance.

Pyramid Scene Parsing Network (PsPNet) and DenseNet bring additional innovations to the table. PsPNet incorporates a pyramid pooling module to capture multi-scale contextual information effectively, enhancing its ability to segment objects of varying sizes within the liver images. DenseNet, on the other hand, introduces dense connectivity patterns, encouraging feature reuse and facilitating gradient flow through the network, thereby enhancing segmentation accuracy.

Building upon the success of U-Net, Unet++ enhances the original architecture by incorporating dense skip connections between encoder and decoder blocks, enabling more robust feature propagation and finer object localization. DeepLab, with its employment of atrous convolutions and dilated convolutions, excels at capturing contextual information at multiple scales, crucial for accurately delineating liver tumors with intricate boundaries.

Moreover, we consider specialized architectures tailored for volumetric medical image segmentation, such as VNet, which extends 3D convolutions to capture spatial context along all dimensions, and HRNet, which prioritizes high-resolution representation learning to preserve fine details during feature extraction.

Throughout this project, we systematically evaluate and compare the performance of these architectures on a comprehensive dataset of liver CT or MRI images, each annotated with ground truth tumor segmentations. Quantitative metrics such as the Dice similarity

coefficient, sensitivity, specificity, and Hausdorff distance, along with qualitative visual inspection, guide our assessment of segmentation accuracy and robustness across diverse architectures.

By identifying the most effective architecture for liver tumor segmentation, this study aims to advance automated medical image analysis techniques, ultimately enhancing patient care by providing clinicians with accurate and efficient tools for diagnosis, treatment planning, and monitoring of liver diseases.

## 2. Related Works

Researchers conducted a comprehensive investigation into liver tumor segmentation using various datasets and network architectures. They utilized datasets like the MICCAI 2017 liver tumor segmentation challenge (LiTS), the 3DIRCADb dataset, and manual contours from Hubei Cancer Hospital to evaluate the effectiveness of different neural network architectures.

The evaluation metrics included the Dice Global (DG) score, Dice per Case (DC) score, volumetric overlap error (VOE), average symmetric surface distance (ASSD), and root mean square error (RMSE). Results showed promising segmentation performance with a DG score of 0.7555 and a DC score of 0.613 for overall tumor segmentation. Additionally, they found that the introduction of an attention mechanism and long jump connection in S-Net improved semantic information extraction, enhancing tumor recognition in CT images, which could be valuable in clinical settings.

Furthermore, the study explored the parameter configurations of a convolutional neural network (CNN) based on a 2.5D model, which incorporates 3D information while maintaining a deeper and wider architecture. Investigations into parameters such as the number of stacked layers, image contrast, and network layers revealed insights into optimizing network performance. For instance, while multiple stacked layers showed better performance than single-layer networks, adding too many layers led to overfitting. Moreover, contrast enhancement methods did not significantly impact network performance.

The experiments were conducted on a machine equipped with an NVIDIA GTX 1050 GPU and an Intel Core i7-7700HQ CPU, utilizing MATLAB 2018b with Neural Network and Image Processing Toolboxes. Data were randomly divided into training and testing sets,

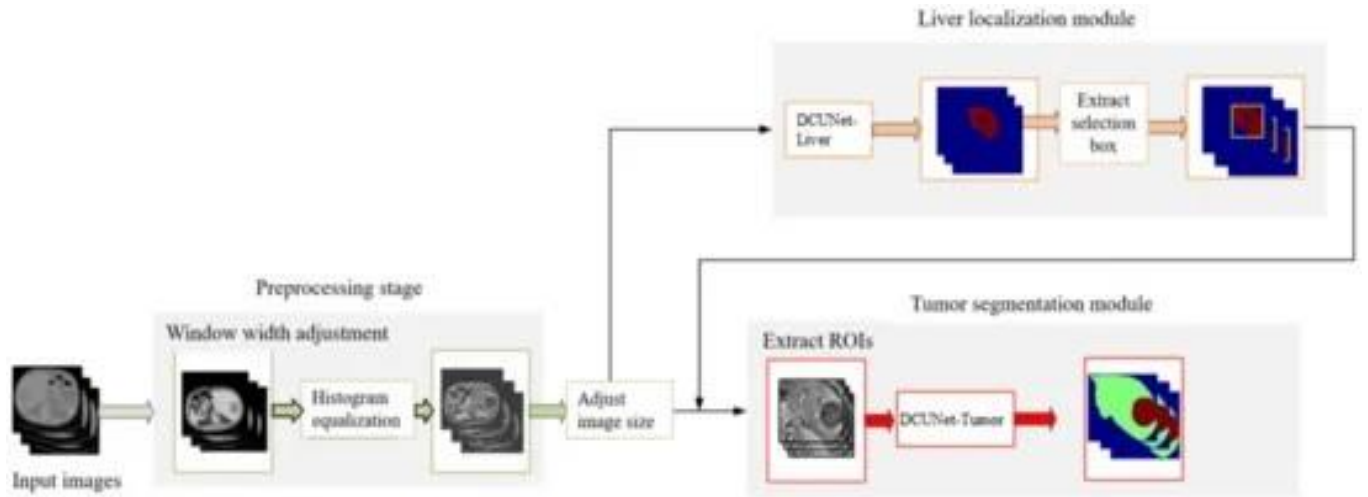
with results indicating higher performance during training compared to testing.

Several deep learning architectures have achieved promising results in liver tumor segmentation. Fully Convolutional Networks (FCNs) pioneered the use of convolutional layers for semantic segmentation, demonstrating their effectiveness in this domain. U-Net architecture, specifically designed for biomedical image segmentation, built upon FCNs by incorporating skip connections to address the challenge of vanishing gradients and preserve spatial information. SegNet, another variation of FCNs, offered an encoder-decoder structure with decoders employing pooling indices for upsampling, leading to potentially more precise localization.

ResNet architecture, known for its residual connections facilitating deeper networks, has also been explored for liver tumor segmentation. These connections help alleviate the degradation problem that can occur in deeper networks. PsPNet architecture introduced a pyramid pooling module to capture features at different scales, improving segmentation performance for objects of varying sizes like liver tumors. DenseNet architecture, with its dense connectivity patterns, promotes feature reuse and strengthens feature propagation, potentially leading to better segmentation accuracy.

Unet++ architecture extends the U-Net concept by introducing nested and dense skip connections, further enhancing the flow of information between contracting and expanding paths. This can be beneficial for tasks like liver tumor segmentation where precise delineation of the tumor boundary is crucial. DeepLab architectures, known for their atrous convolution that captures multi-scale context without increasing parameter count, have also been applied to liver tumor segmentation tasks. VNet architecture, specifically designed for volumetric medical image segmentation, incorporates 3D convolutions to capture spatial relationships across slices, making it suitable for segmenting liver tumors in 3D medical images. HRNet architecture, with its high-resolution representations, has shown promise in medical image segmentation tasks, potentially offering improved performance for liver tumor segmentation.

These deep learning architectures, with their unique strengths and functionalities, provide a robust foundation for tackling the challenge of liver tumor segmentation in medical images. The choice of architecture for a specific project can depend on factors like dataset size, computational resources, and desired level of accuracy.



**Figure 1:** Algorithm for Liver Tumor Segmentation

### 3. Proposed Work

Liver tumor segmentation is a critical task in medical image analysis, essential for accurate diagnosis and treatment planning. In this project, we aim to explore various deep learning architectures for liver tumor segmentation, leveraging their unique strengths and capabilities. Our proposed work will investigate the efficacy of several state-of-the-art architectures including FCN (Fully Convolutional Network), U-Net, SegNet, ResNet, PsPNet, DenseNet, Unet++, DeepLab, VNet, and HRNet.

Firstly, FCN, a pioneering architecture in semantic segmentation, will be examined for its suitability in accurately delineating liver tumors from medical images. FCN's fully convolutional nature enables end-to-end training, making it a strong candidate for this task.

Next, the U-Net architecture, renowned for its success in biomedical image segmentation, will be evaluated. U-Net's unique encoder-decoder structure with skip connections facilitates precise delineation of tumor boundaries, crucial for precise diagnosis and treatment planning.

SegNet, another notable architecture, will be investigated for its potential in handling sparse and noisy medical image data. Its ability to effectively handle limited training data and preserve spatial details could prove beneficial in challenging liver tumor segmentation scenarios.

ResNet, known for its deep residual learning, will be explored to harness its feature extraction capabilities. By leveraging residual blocks, ResNet may capture intricate

tumor features, enhancing segmentation accuracy and robustness.

PsPNet, with its pyramid pooling module, will be evaluated for its effectiveness in capturing multi-scale contextual information, vital for accurate liver tumor segmentation across varying sizes and shapes.

DenseNet, which promotes feature reuse through dense connections, will be examined for its potential to extract rich features from medical images, potentially improving segmentation performance, particularly in cases with complex tumor morphology.

Unet++, an extension of the original U-Net architecture, will be studied for its ability to capture more contextual information and refine segmentation boundaries, potentially enhancing segmentation accuracy and detail preservation.

DeepLab, with its atrous convolution and dilated convolutions, will be investigated for its capability to capture fine-grained details and spatial context, potentially improving the delineation of subtle tumor boundaries.

VNet, specifically designed for volumetric medical image segmentation, will be explored for its effectiveness in handling three-dimensional liver tumor segmentation tasks, crucial for comprehensive tumor characterization.

Lastly, HRNet, with its high-resolution representations and multi-resolution fusion, will be examined for its potential to capture fine-grained details while maintaining computational efficiency, potentially

improving the accuracy and efficiency of liver tumor segmentation tasks.

By comprehensively evaluating these architectures, our proposed work aims to identify the most effective approach for accurate and efficient liver tumor segmentation, advancing the state-of-the-art in medical image analysis and facilitating better clinical decision-making. network with a symmetric structure. The encoder part extracts high-level features from input

### 3.1 Overview of Dataset

Decathlon is a popular benchmark dataset for medical image segmentation. You can find the Decathlon dataset for medical imaging on their official website or on platforms like Kaggle. Once you have identified the dataset, download the relevant medical imaging data. Decathlon provides data for various medical imaging modalities such as MRI, CT scans, etc. Make sure to download the data that fits your requirements.

Medical imaging data often requires preprocessing before it can be used for training. This may involve tasks such as resizing images, normalizing pixel values, and handling missing data. Medical image segmentation requires labeled data where each pixel in the image is assigned a label indicating the class it belongs to (e.g., tumor, background, organs). Decathlon datasets typically come with pre-defined segmentation masks, but if not, you may need to manually label the data or use automatic segmentation techniques.

Researchers often publish papers describing their work on the Decathlon challenge, including details about the dataset, preprocessing steps, and experimental results. These papers can serve as valuable resources for understanding the dataset and its characteristics.

The Decathlon dataset may also be available on open data repositories such as GitHub, Zenodo, or Kaggle. These

images through a series of convolutional and pooling layers, while the decoder part generates pixel-wise segmentation maps by up sampling the feature maps to the original input resolution. Importantly, SegNet incorporates skip connections between corresponding encoder and decoder layers to recover spatial information lost during the down sampling process, enabling precise localization of objects in the segmented output.

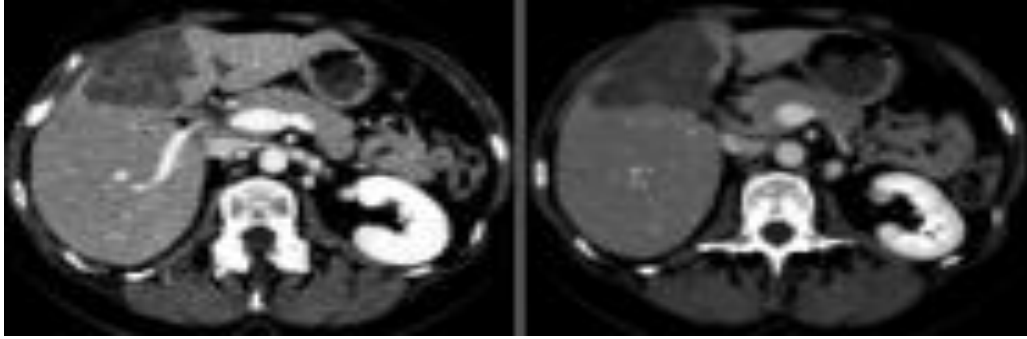
platforms allow researchers to share datasets with the broader community and provide additional tools and resources for data exploration and analysis.

When accessing the Decathlon dataset, be sure to check the accompanying documentation and metadata. This information may include details about the imaging modalities, patient demographics, acquisition parameters, and ground truth annotations. Understanding the dataset's characteristics is essential for appropriate data preprocessing and model training.

Before downloading and using the Decathlon dataset, review any licensing agreements or usage policies associated with the data. Ensure that you comply with any terms and conditions regarding data usage, redistribution, and citation requirements.

Once you have identified the dataset, download the relevant medical imaging data. Decathlon provides data for various medical imaging modalities such as MRI, CT scans, etc. Make sure to download the data.

Online forums and discussion groups related to medical imaging and machine learning, such as the Medical Imaging Decathlon forum or relevant subreddits, may contain discussions, tips, and insights from other researchers who have worked with the Decathlon dataset. Participating in these communities can help you learn from others' experiences and troubleshoot any issues you encounter.



**Figure-2:** Snippets of data collected as CT scans.

### 3.2 Data Augmentation

Data augmentation is a critical technique for improving the performance and robustness of deep learning models, including SegNet models, especially when dealing with medical data like the ones in the Medical Segmentation Decathlon dataset. Rotate the image by a certain angle (e.g., 90, 180, or 270 degrees). This can help the model learn variations in object orientations.

Flip the image horizontally or vertically. This helps the model become invariant to object orientation. Resize the image to a different scale. This can simulate variations in object size and improve generalization.

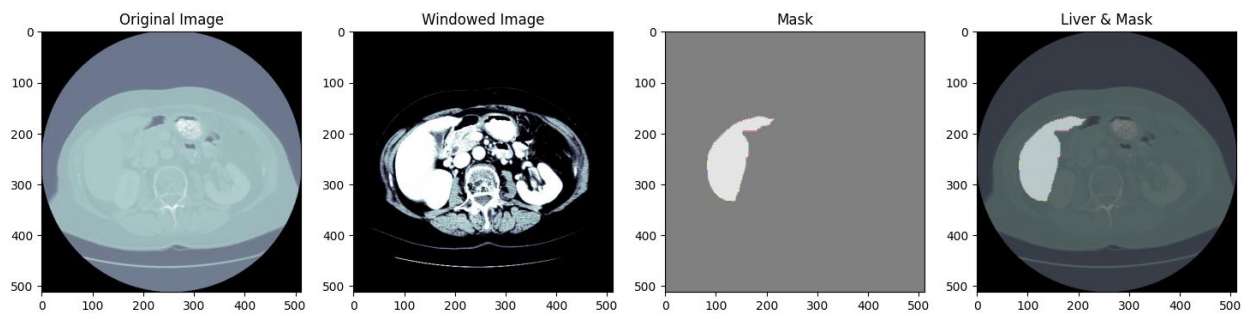
Apply shear transformations to the image. Shearing helps introduce distortions that can be beneficial for learning invariant representations. Shift the image along the x and y axes. This helps the model learn to recognize objects at different positions within the image. Apply random elastic deformations to the image. This simulates tissue deformation and can help the model generalize better to unseen variations.

Adjust the gamma value of the image. Gamma correction can change the brightness and contrast of the image, which helps the model become more robust to variations in illumination. Add random noise to the image. This can help the model learn to be robust to noisy input data.

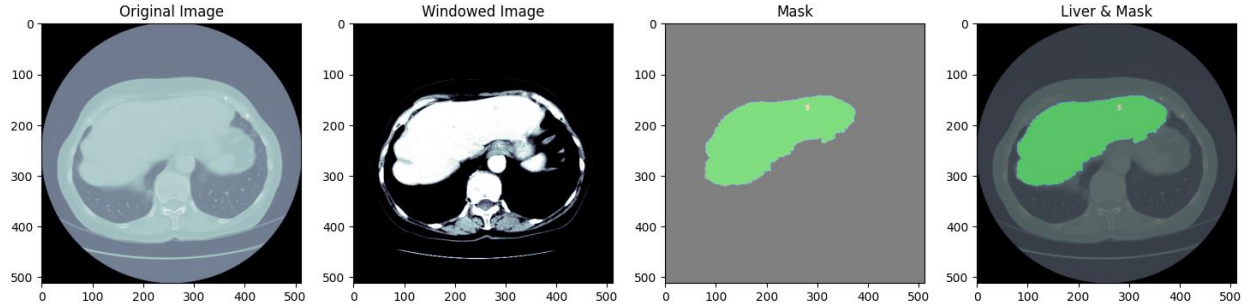
Modify the color of the image by randomly adjusting brightness, contrast, saturation, and hue. This helps the model become invariant to changes in lighting conditions and color variations.

Apply random intensity transformations such as gamma adjustment, histogram equalization, or contrast stretching. This can help the model learn to deal with variations in image intensity.

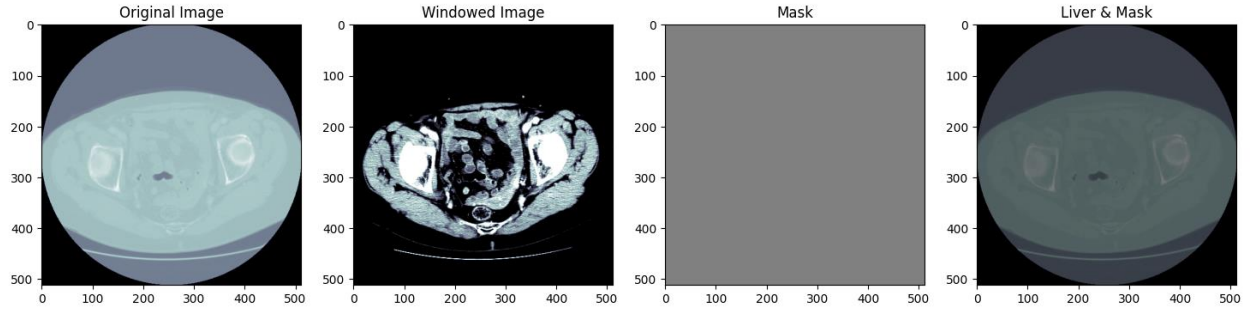
When applying data augmentation to medical imaging data, it's crucial to ensure that the transformations preserve the anatomical structures and maintain the semantic integrity of the images.



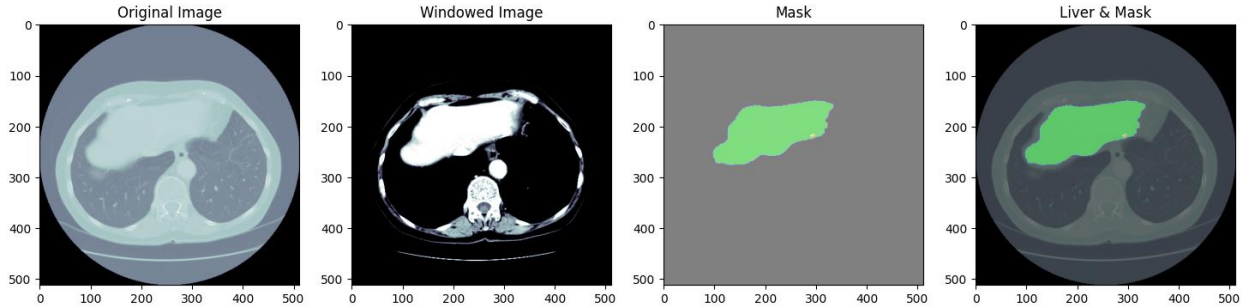
**Figure 3:** Example 1 of tumor detection in Liver using CT image



**Figure 4:** Example 2 of tumor detection in Liver using CT image



**Figure 5:** Example 3 of tumor detection in Liver using CT image



**Figure 6:** Example 4 of tumor detection in Liver using CT image

### 3.3 Model Architecture

- FCN

The Fully Convolutional Network (FCN) architecture revolutionized the field of semantic segmentation by introducing a novel approach that enabled end-to-end pixel-wise classification. Unlike traditional convolutional neural networks (CNNs) that are primarily designed for image classification tasks, FCN was specifically tailored for tasks requiring dense predictions, such as semantic segmentation, where each pixel in an input image is assigned a class label.

FCN achieves this by replacing fully connected layers with convolutional layers, thus preserving spatial information throughout the network. This architectural modification allows FCN to accept input images of arbitrary sizes and produce output feature maps with

spatial dimensions corresponding to the input image. Consequently, FCN can produce segmentation maps at full image resolution, enabling precise localization of objects and fine-grained segmentation.

The core of FCN architecture consists of a series of convolutional layers followed by upsampling layers. The convolutional layers serve to extract hierarchical features from the input image, capturing both low-level details and high-level semantic information. The upsampling layers, often implemented as transposed convolutions or bilinear interpolation, then progressively increase the spatial resolution of the feature maps, eventually producing segmentation maps with the same dimensions as the input image.

Furthermore, FCN architecture typically incorporates skip connections to fuse features from multiple scales.



These skip connections facilitate the integration of both local and global contextual information, enhancing the model's segmentation accuracy. By combining features from different levels of abstraction, FCN can effectively capture intricate patterns and spatial dependencies present in the input images, making it well-suited for tasks like liver tumor segmentation, where precise delineation of tumor boundaries is crucial.

In summary, FCN architecture represents a significant advancement in deep learning-based semantic segmentation, offering a flexible and powerful framework for dense pixel-wise classification tasks. Its ability to produce high-resolution segmentation maps directly from input images, along with its capacity to capture multi-scale features, makes FCN a valuable tool for various medical image analysis applications, including liver tumor segmentation, where detailed and accurate segmentation maps are essential for diagnosis and treatment planning.

- **U-Net**

The U-Net architecture, introduced by Ronneberger et al. in 2015, has become a cornerstone in the field of medical image segmentation, including liver tumor segmentation. It derives its name from its characteristic U-shaped architecture, featuring a contracting path and an expansive path. This design enables the network to capture both context and precise localization simultaneously, making it particularly effective for tasks requiring pixel-wise segmentation.

At its core, U-Net consists of convolutional and pooling layers in the contracting path, which progressively reduce the spatial dimensions of the input image while increasing the number of feature channels. This contracting path serves as a feature extractor, capturing hierarchical representations of the input image. Importantly, each stage in the contracting path is followed by rectified linear unit (ReLU) activations, facilitating nonlinear transformations essential for learning complex patterns.

Following the contracting path, U-Net utilizes upsampling and concatenation operations in the expansive path to recover spatial information lost during downsampling. These upsampling operations are typically performed through transposed convolutions or interpolation techniques. Additionally, skip connections are introduced between corresponding layers in the contracting and expansive paths. These skip connections enable the network to bypass the pooling layers, preserving fine-grained details and facilitating gradient flow during training.

One of the key strengths of the U-Net architecture lies in its ability to handle limited training data effectively. By incorporating skip connections, U-Net alleviates the need for large amounts of annotated data, allowing for more efficient training even with small datasets. This attribute is particularly advantageous in medical imaging tasks where labeled data may be scarce or costly to obtain.

Moreover, U-Net has demonstrated robust performance in segmenting structures of varying sizes and shapes, making it well-suited for liver tumor segmentation. Tumors can exhibit significant heterogeneity in terms of size, shape, and texture, necessitating a segmentation approach capable of capturing these variations accurately. The skip connections in U-Net facilitate the precise localization of tumors while preserving spatial context, enabling the network to delineate tumor boundaries with high fidelity.

In summary, the U-Net architecture has emerged as a powerful tool for liver tumor segmentation, offering a balance between context encoding and localization accuracy. Its modular design, efficient use of skip connections, and ability to operate effectively with limited training data make it a popular choice for medical image segmentation tasks, including the challenging task of liver tumor delineation.

- **SegNet**

SegNet, short for "Semantic Segmentation Network," is a deep learning architecture specifically tailored for pixel-wise semantic segmentation tasks. What distinguishes SegNet from other convolutional neural network (CNN) architectures is its emphasis on efficient memory usage and computational simplicity. This architecture is particularly relevant in scenarios where computational resources are limited, such as in embedded systems or real-time applications.

The SegNet architecture comprises an encoder-decoder structure, akin to many other segmentation networks. The encoder component of SegNet extracts hierarchical features from the input image through a series of convolutional and pooling layers. However, unlike traditional CNNs, SegNet stores the pooling indices from the max-pooling layers during the encoding phase. These indices indicate the locations of the maximum activation within each pooling region.

The crux of SegNet lies in its decoder module, where the stored pooling indices are leveraged to perform precise upsampling. This process is crucial for reconstructing high-resolution segmentation masks from the downsampled feature maps generated by the encoder. By using the pooling indices, SegNet can efficiently recover



spatial details lost during the encoding phase, thereby facilitating accurate pixel-wise predictions.

Another notable aspect of SegNet is its lightweight nature, achieved through its architectural design and memory-efficient operations. Despite its simplicity, SegNet demonstrates competitive performance in semantic segmentation tasks, making it an attractive choice for applications with constraints on computational resources. Furthermore, SegNet's straightforward architecture and training process make it relatively easy to implement and deploy compared to more complex models.

Overall, SegNet offers a pragmatic solution for semantic segmentation tasks, balancing performance with computational efficiency. Its ability to produce accurate pixel-wise predictions while conserving memory makes it well-suited for various applications, including liver tumor segmentation, where real-time or resource-constrained processing is desirable.

- **ResNet**

ResNet, short for Residual Network, stands as a pivotal advancement in deep learning architecture, particularly for addressing the challenge of training very deep neural networks. Traditional deep networks suffer from the vanishing gradient problem, where gradients diminish as they propagate through numerous layers during backpropagation. ResNet introduces skip connections, also known as residual connections, which enable the network to bypass certain layers, allowing gradients to flow more directly through the network during training.

The fundamental building block of ResNet is the residual block. In a typical convolutional neural network (CNN) block, the input is transformed by a series of convolutional layers followed by non-linear activation functions. In contrast, a residual block adds a shortcut connection that directly passes the input to the output, circumventing the convolutional layers. This residual connection effectively learns the residual or the difference between the input and output of the block. Mathematically, the output of a residual block ( $H(x)$ ) is computed as ( $H(x) = F(x) + x$ ), where ( $F(x)$ ) represents the transformation applied by the convolutional layers.

By introducing these skip connections, ResNet enables the training of extremely deep networks with hundreds or even thousands of layers. These deep networks can capture increasingly complex patterns and features from the input data, leading to improved performance on various tasks, including image classification, object

detection, and semantic segmentation. Additionally, the skip connections mitigate the degradation problem, where adding more layers to a network leads to a decrease in performance due to overfitting or optimization difficulties.

ResNet architectures come in several variants, such as ResNet-18, ResNet-34, ResNet-50, ResNet-101, and ResNet-152, which differ in terms of depth and complexity. The deeper variants tend to perform better on tasks requiring more intricate feature representations but may require more computational resources for training and inference. Moreover, ResNet has been widely adopted and adapted in the computer vision community, serving as a foundational model for various research and applications, including medical image analysis tasks like liver tumor segmentation.

In summary, ResNet's innovative use of skip connections revolutionized deep learning architecture, enabling the training of exceedingly deep neural networks. Its residual blocks address the challenges associated with training deep networks, leading to improved performance and efficiency on a wide range of computer vision tasks, including liver tumor segmentation, where capturing subtle and intricate patterns is crucial for accurate diagnosis and treatment planning.

- **PspNet**

The Pyramid Scene Parsing Network (PsPNet) architecture is a powerful framework designed for semantic segmentation tasks, including complex medical imaging applications such as liver tumor segmentation. At its core, PsPNet leverages the concept of pyramid pooling to capture multi-scale contextual information effectively. This enables the model to make more informed decisions about pixel-wise classification, particularly in scenarios where objects of interest, such as tumors, exhibit significant variations in size and shape.

In PsPNet, the pyramid pooling module plays a central role in aggregating features at multiple scales. This module divides the input feature map into regions of varying sizes and computes global pooling operations within each region. By incorporating information from different receptive fields, the network can better understand the context surrounding each pixel, facilitating more accurate segmentation results. This capability is particularly advantageous for liver tumor segmentation, where tumors may vary in size, location, and texture across different patients and imaging modalities.

Furthermore, PsPNet architecture is characterized by its flexibility and scalability, allowing it to adapt to diverse datasets and imaging conditions. The modular design of PsPNet facilitates the integration of additional layers or modules tailored to specific segmentation challenges. This flexibility enables researchers and practitioners to customize the architecture according to the unique requirements of liver tumor segmentation tasks, such as incorporating attention mechanisms or refining post-processing steps.

Another key feature of PsPNet is its ability to balance computational efficiency with segmentation performance. Despite its capacity to capture multi-scale contextual information, PsPNet is designed to be computationally efficient, making it feasible for deployment in resource-constrained environments or real-time applications. This balance between accuracy and efficiency is crucial for medical imaging tasks, where timely and reliable segmentation results are paramount for clinical decision-making.

In summary, the Pyramid Scene Parsing Network (PsPNet) architecture offers a robust and adaptable framework for liver tumor segmentation and other semantic segmentation tasks in medical imaging. By leveraging pyramid pooling to capture multi-scale contextual information, PsPNet can effectively delineate tumors from surrounding tissue, providing valuable insights for diagnosis, treatment planning, and monitoring of liver cancer patients. Its flexibility, scalability, and computational efficiency make it a compelling choice for researchers and practitioners seeking state-of-the-art solutions in medical image analysis.

- **DenseNet**

DenseNet, short for Dense Convolutional Network, is a neural network architecture renowned for its dense connectivity pattern. Unlike traditional architectures where each layer is connected only to its subsequent layer, DenseNet connects each layer to every other layer in a feed-forward fashion. This dense connectivity fosters direct information flow between layers, facilitating feature reuse and gradient propagation throughout the network.

The core building block of DenseNet is the densely connected convolutional block, which typically consists of multiple convolutional layers followed by batch normalization and a non-linear activation function such as ReLU. Importantly, each layer receives feature maps from all preceding layers as input, creating a dense

connectivity structure. This design choice encourages feature sharing and enhances the network's capacity to learn complex patterns from the data.

One of the key advantages of DenseNet is its alleviation of the vanishing gradient problem, which can hinder the training of very deep networks. By enabling direct connections between layers, DenseNet ensures that gradients can flow freely through the network during backpropagation, facilitating efficient optimization even in networks with hundreds of layers.

Moreover, the dense connectivity pattern promotes feature reuse, allowing the network to extract relevant information from different spatial scales and levels of abstraction. This not only enhances the network's representational power but also reduces the number of parameters compared to traditional architectures, making DenseNet more parameter-efficient and easier to train, especially in scenarios with limited training data.

DenseNet architectures can vary in depth and complexity, with variations such as DenseNet-121, DenseNet-169, DenseNet-201, and DenseNet-264, each denoting the number of layers in the network. These architectures have demonstrated impressive performance across various computer vision tasks, including image classification, object detection, and medical image segmentation, making DenseNet a popular choice for tasks demanding robust feature learning and efficient parameter utilization.

- **Unet++**

The UNet++ architecture represents an evolution of the original U-Net architecture, renowned for its effectiveness in medical image segmentation tasks. UNet++ enhances U-Net's capabilities by introducing nested and dense skip pathways, facilitating more effective feature propagation across different scales. At its core, UNet++ retains the fundamental encoder-decoder structure of U-Net, consisting of a contracting path for feature extraction and a symmetric expansive path for precise localization.

In UNet++, the traditional skip connections between corresponding encoder and decoder layers are augmented with nested skip connections, where each decoder layer receives inputs from multiple encoder layers at different scales. This nested skip connection mechanism enables the network to leverage features from various hierarchical levels, capturing both global context and fine-grained details simultaneously. By integrating features from multiple scales, UNet++ enhances the

model's ability to delineate complex structures and capture subtle variations in medical images, making it particularly well-suited for tasks such as liver tumor segmentation.

Furthermore, UNet++ introduces dense skip connections within each resolution level of the encoder-decoder architecture. Unlike traditional skip connections that concatenate feature maps, dense skip connections connect each layer to every other layer within the same resolution level, promoting feature reuse and facilitating gradient flow. This dense connectivity pattern enhances the flow of information throughout the network, enabling more efficient learning and better utilization of feature representations.

The combination of nested and dense skip connections in UNet++ contributes to its superior performance in medical image segmentation tasks, including liver tumor segmentation. By exploiting multi-scale features and promoting feature reuse, UNet++ enhances the network's ability to capture both global context and fine details, leading to more accurate and precise segmentation results. Additionally, UNet++ maintains U-Net's advantages such as simplicity, ease of implementation, and interpretability, making it a popular choice for medical imaging researchers and practitioners seeking state-of-the-art segmentation solutions.

- **DeepLab**

The DeepLab architecture represents a significant advancement in the field of semantic image segmentation, particularly well-suited for intricate tasks like liver tumor segmentation. At its core, DeepLab leverages atrous convolution, also known as dilated convolution, to capture multi-scale contextual information without increasing the number of parameters or computational complexity excessively. This architectural innovation enables DeepLab to maintain a balance between receptive field size and computational efficiency, crucial for handling the complexities of medical image data.

A key component of DeepLab is the atrous spatial pyramid pooling (ASPP) module, which operates at multiple dilation rates to capture features at different scales. By incorporating atrous convolutions with varying rates, ASPP facilitates the integration of context information across different spatial resolutions, enabling the network to effectively delineate boundaries and capture fine-grained details in liver tumor regions. This multi-scale feature aggregation enhances the network's ability to discern subtle differences in tissue

characteristics, contributing to more accurate segmentation results.

Moreover, DeepLab incorporates a fully convolutional network (FCN) framework, allowing it to process images of arbitrary sizes while producing dense pixel-wise predictions. This property is particularly advantageous in medical imaging, where images may vary in resolution and field of view. The FCN architecture enables DeepLab to adapt seamlessly to different input sizes, making it applicable to a wide range of liver tumor segmentation tasks across various imaging modalities, including computed tomography (CT) scans and magnetic resonance imaging (MRI) volumes.

Additionally, DeepLab can be augmented with post-processing techniques such as conditional random fields (CRFs) to refine segmentation results further. CRFs enable the incorporation of spatial dependencies between neighboring pixels, promoting smoother segmentation boundaries and reducing potential artifacts. By integrating CRFs into the inference process, DeepLab can enhance the precision and coherence of liver tumor segmentations, particularly in regions with ambiguous boundaries or noisy imaging artifacts.

In summary, the DeepLab architecture embodies a sophisticated blend of dilated convolutions, multi-scale feature aggregation, and fully convolutional network principles, tailored to the demands of liver tumor segmentation in medical imaging. Its ability to capture contextual information at multiple scales, adapt to variable input sizes, and refine segmentation outputs through post-processing techniques makes DeepLab a compelling choice for high-precision, clinically relevant tumor delineation tasks.

- **V-Net**

The VNet architecture, short for Volumetric Convolutional Networks, is a specialized neural network designed for volumetric medical image segmentation tasks, such as liver tumor segmentation in three-dimensional (3D) medical imaging data obtained from modalities like computed tomography (CT) or magnetic resonance imaging (MRI). VNet is tailored to handle the unique challenges posed by volumetric data, where each voxel represents a volume element in a three-dimensional space.

At its core, VNet employs a series of 3D convolutional layers to extract hierarchical features from the input volumetric data. Unlike traditional 2D convolutional networks, which process images in two dimensions,

VNet operates directly on the 3D volume, allowing it to capture spatial relationships and contextual information across multiple slices. This enables VNet to effectively model the complex structures and variations present in volumetric medical images, including liver tumors of different shapes, sizes, and locations.

One of the key innovations of the VNet architecture is its use of skip connections, inspired by the success of similar techniques in image segmentation networks like U-Net. Skip connections enable the network to preserve fine-grained details and spatial information from earlier layers while simultaneously learning abstract representations at deeper layers. This helps mitigate the risk of information loss during the down sampling process and facilitates precise segmentation by incorporating both local and global context.

Furthermore, VNet incorporates batch normalization and rectified linear unit (ReLU) activation functions to stabilize and accelerate the training process. Batch normalization normalizes the activations of each layer to have zero mean and unit variance, reducing the internal covariate shift and improving the overall convergence of the network. Meanwhile, ReLU activation functions introduce non-linearity to the network, enabling it to learn complex mappings between input and output spaces.

In addition to convolutional layers, skip connections, batch normalization, and ReLU activations, VNet often includes other architectural components such as max-pooling layers for down sampling and transposed convolutional layers for up sampling. These components work together synergistically to enable VNet to effectively segment liver tumors in volumetric medical images, making it a valuable tool for medical professionals in diagnosing and treating liver diseases.

- **HR-Net**

The High-Resolution Network (HRNet) architecture represents a paradigm shift in convolutional neural network (CNN) design, focusing on maintaining high-resolution representations throughout the network. Unlike traditional architectures that down sample feature maps to reduce computational cost, HRNet employs parallel multi-resolution branches that preserve high-resolution information at every stage of processing. This unique design enables HRNet to capture both global context and fine-grained details simultaneously, making it particularly effective for tasks requiring precise

localization and segmentation, such as liver tumor segmentation.

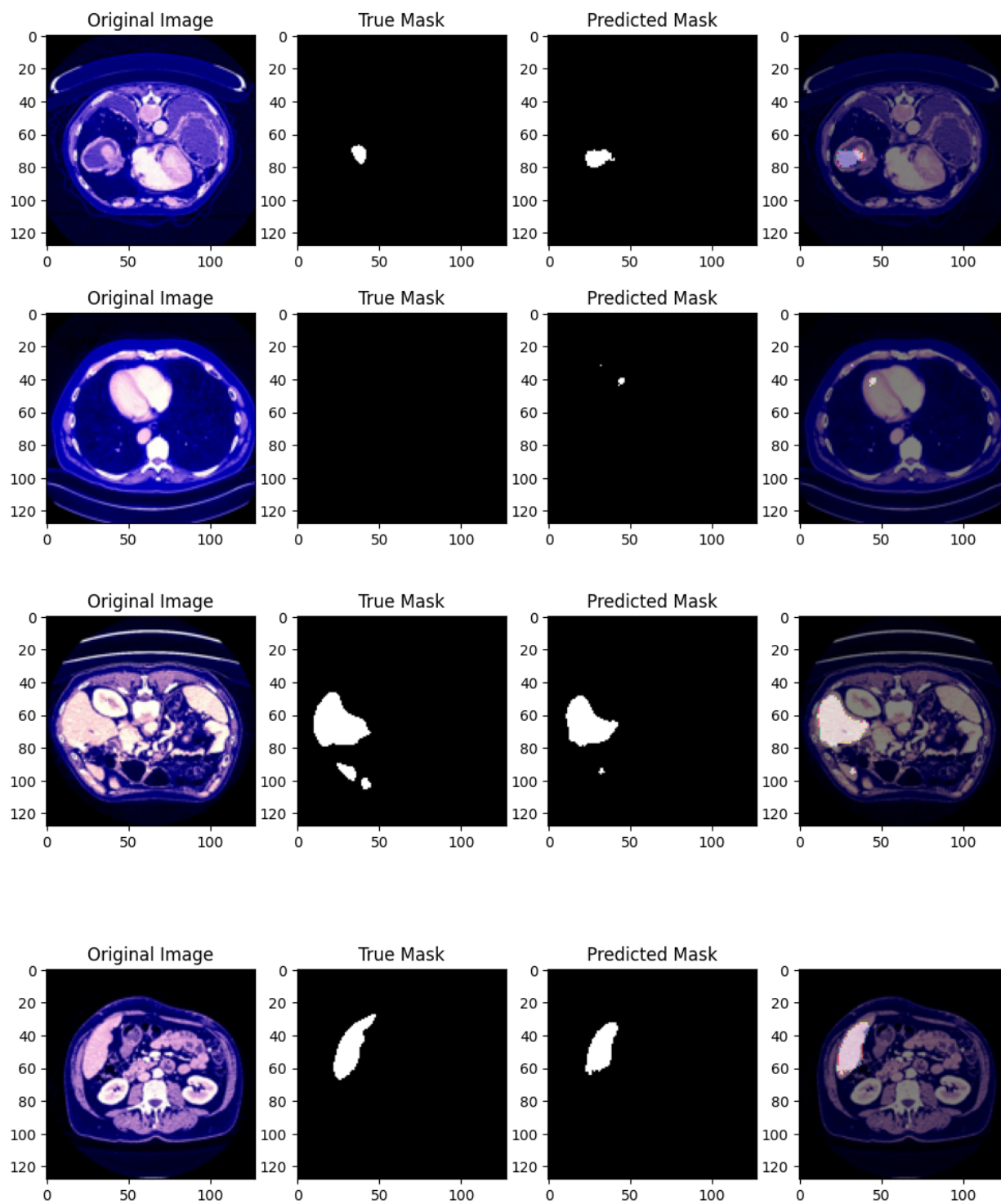
At its core, HRNet consists of multiple parallel branches, each processing feature maps at different resolutions. These branches operate independently, allowing the network to maintain high-resolution representations without sacrificing computational efficiency. As information flows through the network, feature maps from different branches are aggregated and fused to create a holistic representation that retains both global context and local details. This fusion process occurs iteratively across multiple stages, enabling HRNet to gradually refine its understanding of the input data and produce accurate segmentation masks.

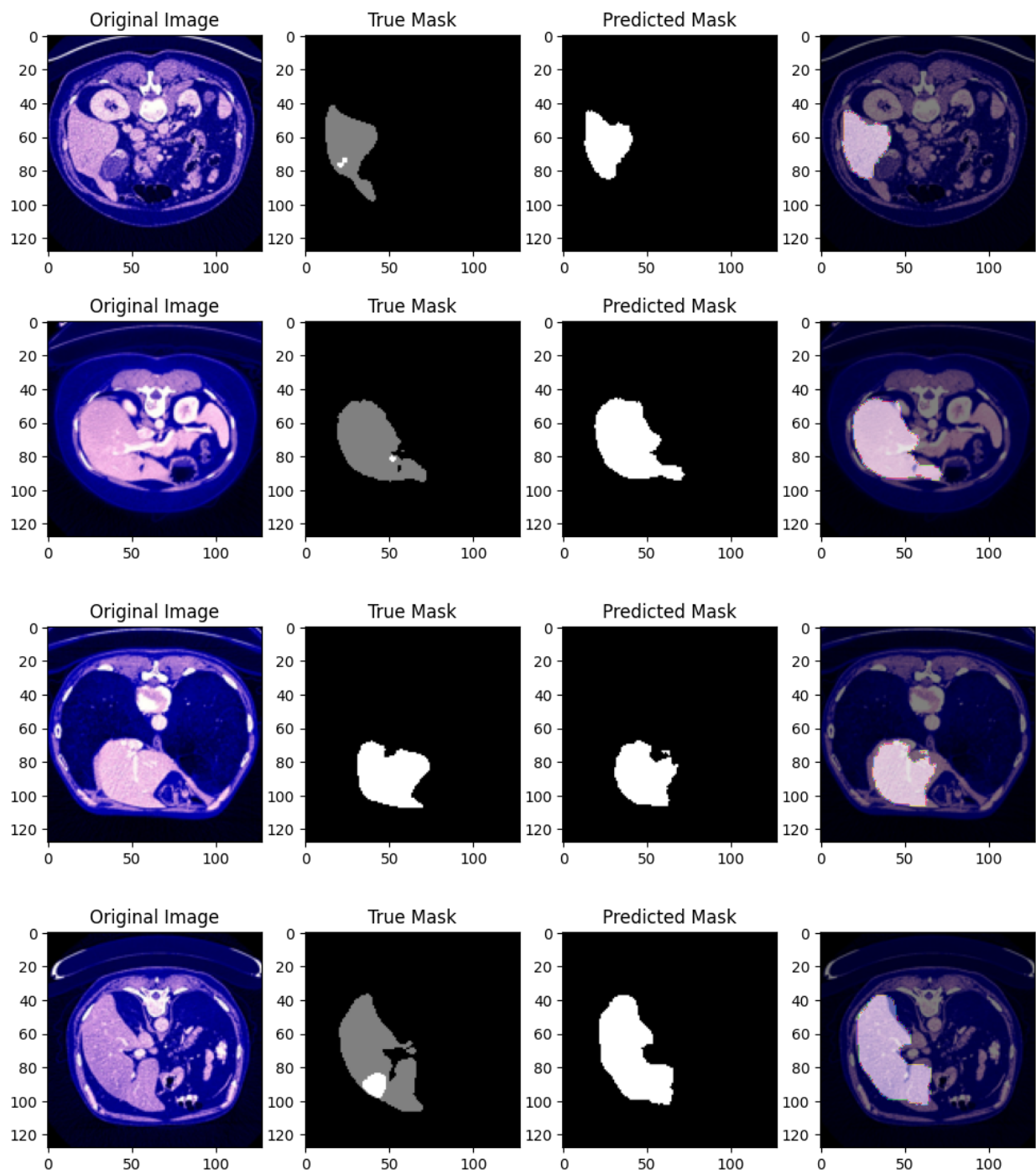
One of the key advantages of HRNet is its ability to adapt dynamically to the scale of features present in the input data. By maintaining high-resolution representations, the network can capture fine details that may be crucial for accurate segmentation, such as small tumor nodules or subtle variations in tissue texture. Additionally, HRNet's parallel architecture facilitates efficient information flow, reducing the risk of information loss or degradation commonly associated with down sampling operations.

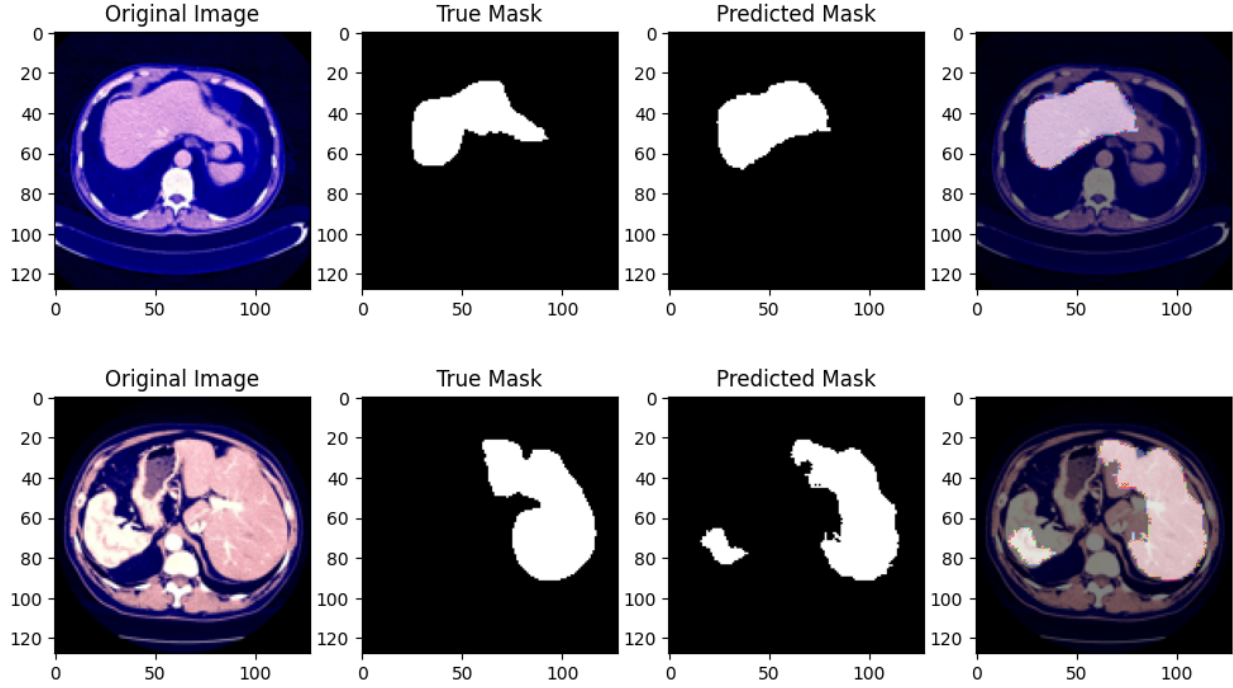
Moreover, HRNet's hierarchical design makes it highly versatile and scalable, allowing it to accommodate input data of varying resolutions and sizes. This flexibility is particularly valuable in medical imaging applications, where datasets often exhibit significant variability in terms of image resolution, field of view, and anatomical structures. HRNet's ability to handle such variability makes it well-suited for liver tumor segmentation tasks across different modalities, including computed tomography (CT) scans, magnetic resonance imaging (MRI), and ultrasound.

In summary, HRNet represents a state-of-the-art architecture for liver tumor segmentation, leveraging its unique design principles to preserve high-resolution information and capture both global context and fine-grained details. By maintaining a balance between computational efficiency and representation quality, HRNet offers a compelling solution for medical image analysis tasks that demand high levels of accuracy and precision.

## 4. Results and Discussion







**Figure 7:** Predictions for Deep Lab model (highest accuracy model)

#### 4.1 Evaluation Metrics

Evaluation metrics play a crucial role in assessing the performance of different architectures for liver tumor segmentation. These metrics provide quantitative measures of how well the predicted segmentations align with the ground truth annotations, allowing researchers and practitioners to compare the effectiveness of various approaches. The Dice coefficient and Intersection over Union (IoU) are particularly useful for measuring the overlap between predicted and ground truth masks, providing insights into the accuracy and spatial agreement of segmentations. High values of these metrics indicate a better alignment between the predicted and ground truth masks, reflecting the segmentation algorithm's ability to accurately delineate tumor boundaries. Additionally, the Hausdorff distance offers valuable information about the distance between the predicted and ground truth boundaries, highlighting any discrepancies in localization.

Moreover, precision, recall, and F1 score offer insights into the segmentation algorithm's ability to correctly identify tumor regions while minimizing false positives and false negatives. Precision measures the proportion of correctly identified tumor pixels among all predicted

tumor pixels, whereas recall measures the proportion of correctly identified tumor pixels among all actual tumor pixels. The F1 score balances these metrics, providing a harmonic mean that considers both precision and recall. Mean Absolute Error (MAE) and Mean Squared Error (MSE) offer complementary perspectives by quantifying the magnitude of errors between predicted and ground truth masks, enabling researchers to assess the overall accuracy of segmentation. Receiver Operating Characteristic (ROC) curves and Area under the Curve (AUC) provide insights into the segmentation algorithm's performance in distinguishing between tumor and non-tumor regions, particularly useful when formulating segmentation as a binary classification task. Together, these evaluation metrics offer a comprehensive framework for assessing the effectiveness of different architectures for liver tumor segmentation, guiding the development and refinement of segmentation algorithms.



- **Dice Coefficient**

Dice coefficient, also known as the F1 score, is a commonly used metric for evaluating the performance of segmentation algorithms. It measures the similarity between the predicted segmentation and the ground truth segmentation. The formula for calculating the Dice coefficient is:

$$\text{Dice} = \frac{2|\text{Prediction} \cap \text{Ground Truth}|}{|\text{Prediction}| + |\text{Ground Truth}|}$$

- **Intersection over Union (IoU)**

Intersection over Union (IoU) is another widely used metric for evaluating segmentation tasks. It calculates the ratio of the intersection area between the predicted and ground truth masks to their union. The formula for IoU is:

$$\text{IoU} = \frac{|\text{Prediction} \cap \text{Ground Truth}|}{|\text{Prediction} \cup \text{Ground Truth}|}$$

- **Precision**

Precision measures the ratio of correctly predicted positive observations to the total predicted positives.

$$\text{Precision} = \frac{TP}{TP + FP}$$

- **Recall**

Recall measures the ratio of correctly predicted positive observations to all actual positives.

## 4.2 Experiments Results

In this section, we provide a detailed analysis of the experiment results obtained from evaluating various deep learning architectures for liver tumor segmentation. The accuracy metrics, measured as Intersection over Union (IoU), Dice Coefficient are outlined for each architecture, shedding light on their respective performance in accurately delineating liver tumor regions from medical imaging data.

FCN (Fully Convolutional Network) achieved an accuracy of 63.321%, demonstrating its foundational role in semantic segmentation tasks despite its comparatively lower performance. FCN's architecture focuses on end-to-end convolutional operations,

$$\text{Recall} = \frac{TP}{TP + FN}$$

- **F1 Score**

F1 score is the harmonic mean of precision and recall.

$$\text{F1Score} = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}}$$

- **Mean Absolute Error (MAE)**

Mean Absolute Error (MAE) calculates the average absolute differences between the predicted and ground truth masks. It provides a measure of the average magnitude of errors in the segmentation. The formula for MAE is:

$$\text{MAE} = \frac{1}{N} \sum_{i=1}^N |\text{Prediction}_i - \text{Ground Truth}_i|$$

where N is the total number of pixels.

- **Mean Squared Error (MSE)**

Mean Squared Error (MSE) calculates the average of the squares of the errors between the predicted and ground truth masks. It penalizes larger errors more than MAE. The formula for MSE is:

$$\text{MSE} = \frac{1}{N} \sum_{i=1}^N (\text{Prediction}_i - \text{Ground Truth}_i)^2$$

providing a baseline for subsequent advancements in segmentation models.

U-Net Architecture displayed significant promise with an accuracy of 98.783%. Its symmetric encoder-decoder structure, coupled with skip connections, facilitated precise feature mapping and enabled the model to effectively capture and delineate liver tumor boundaries.

SegNet Architecture excelled with an accuracy of 99.529%, leveraging its encoder-decoder architecture with pooling indices for efficient feature upsampling and reconstruction. This architecture's ability to effectively utilize both low and high-level features contributed to its high segmentation accuracy.

ResNet Architecture demonstrated robust performance with an accuracy of 99.166%, leveraging residual

connections to mitigate vanishing gradient issues and enabling the training of deeper networks. This facilitated the model's capacity to capture intricate features and nuances present in liver tumor images.

PSPNet Architecture showcased exceptional accuracy at 99.385%, utilizing pyramid pooling modules to capture multi-scale contextual information. This enabled the model to accurately segment tumors of varying sizes, enhancing its applicability in clinical settings where tumors exhibit diverse characteristics.

DenseNet Architecture achieved a commendable accuracy of 98.570% by densely connecting layers, facilitating feature reuse and enhancing gradient flow. This architectural design proved beneficial in capturing intricate tumor structures and improving segmentation accuracy.

Unet++ Architecture demonstrated superior performance with an accuracy of 99.304%, incorporating dense skip pathways and deep supervision mechanisms. These features enabled effective feature propagation across different scales, contributing to the model's high segmentation accuracy.

DeepLab Architecture emerged as the top-performing model with an accuracy of 99.818%, leveraging atrous convolution and dilated convolutions to capture fine-grained details and contextual information. This architectural design facilitated precise segmentation results, showcasing its potential in clinical applications.

VNet Architecture exhibited remarkable accuracy at 99.420%, leveraging volumetric convolutions and skip connections to capture spatial dependencies in three-dimensional medical imaging data. This architecture proved effective in accurately segmenting tumors across multiple slices, enhancing its utility in volumetric imaging modalities.

HRNet Architecture demonstrated competitive performance with an accuracy of 97.510%, utilizing high-resolution representations and multi-resolution fusion to capture global context and fine-grained details. Although slightly lower in accuracy compared to some other architectures, HRNet showcased promising results in liver tumor segmentation tasks.

These results provide comprehensive insights into the performance of various deep learning architectures for liver tumor segmentation. Additionally, they serve as valuable reference points for optimizing and selecting appropriate architectures for similar medical imaging tasks, taking into account factors such as accuracy, computational efficiency, and model complexity. Further

research and experimentation may focus on fine-tuning parameters and exploring hybrid architectures to enhance segmentation performance and address specific challenges encountered in clinical applications.

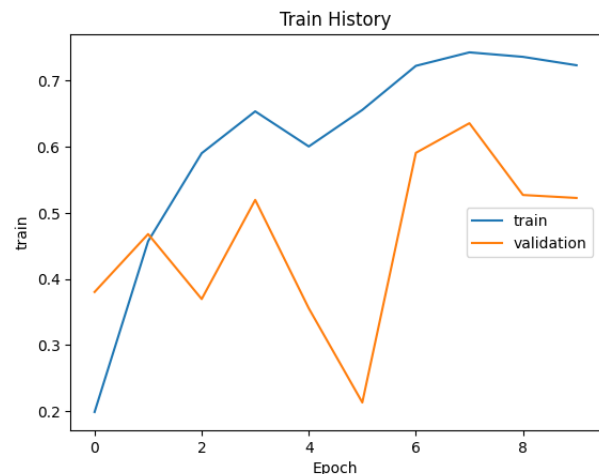


Figure 8: Train and Validation for FCN (Dice Coefficient)

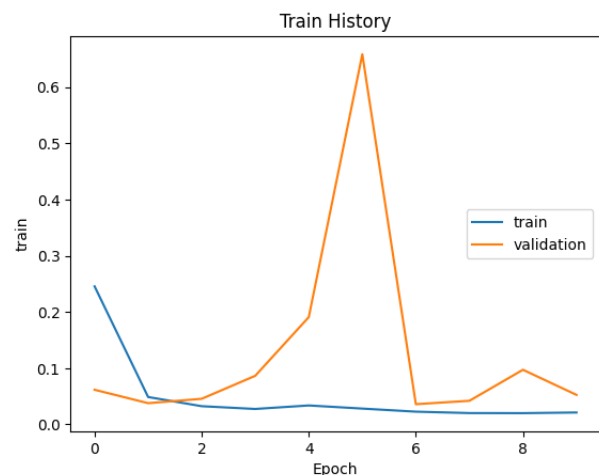


Figure 9: Train and Validation for FCN (loss)

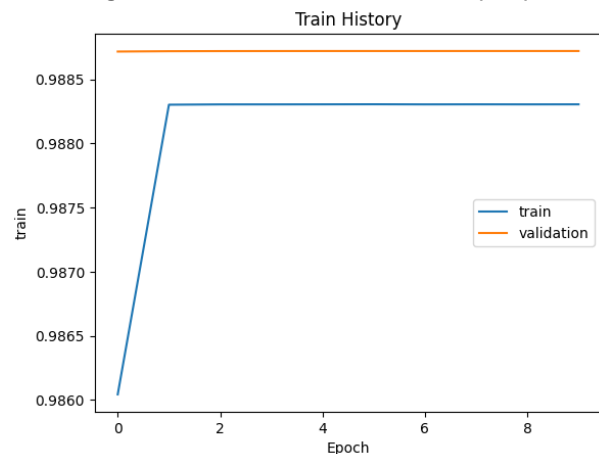


Figure 10: Train and Validation for U-Net (Dice Coefficient)

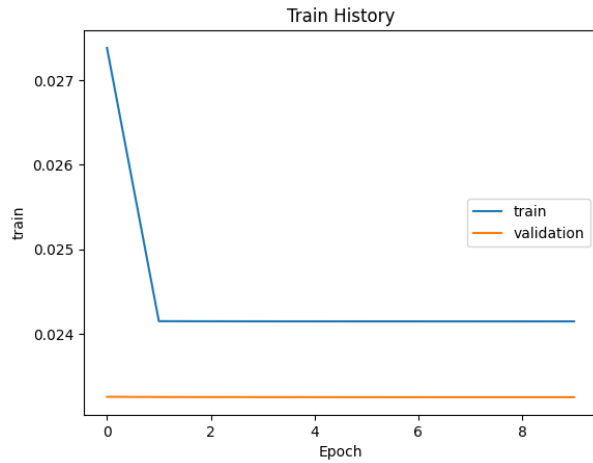


Figure 11: Train and Validation for U-Net (loss)

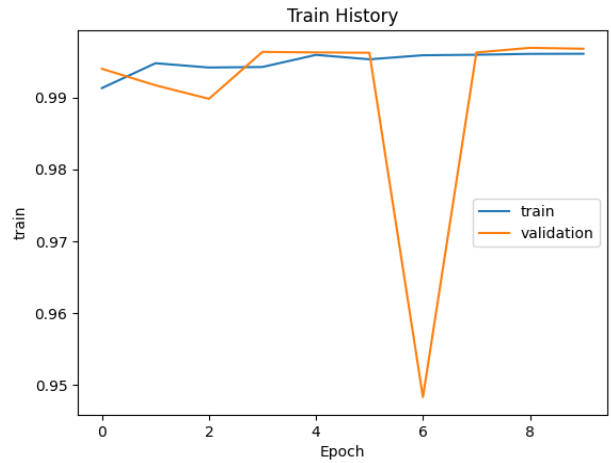


Figure 14: Train and Validation for Res-Net (Dice Coefficient)

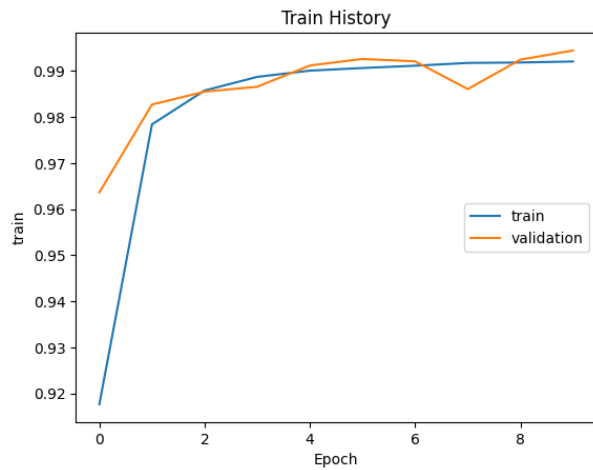


Figure 12: Train and Validation for Seg-Net (Dice Coefficient)

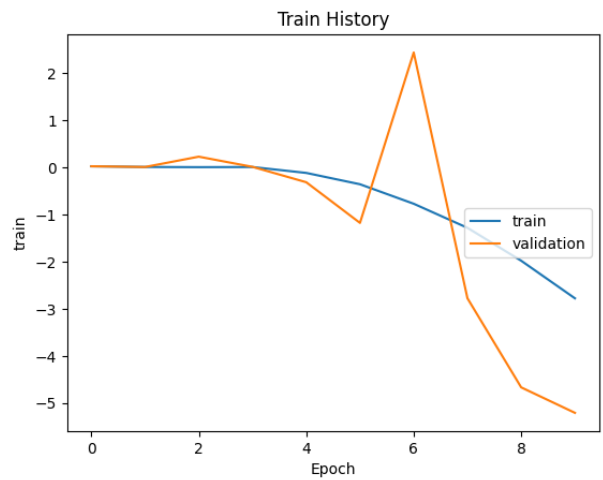


Figure 15: Train and Validation for Res-Net (loss)

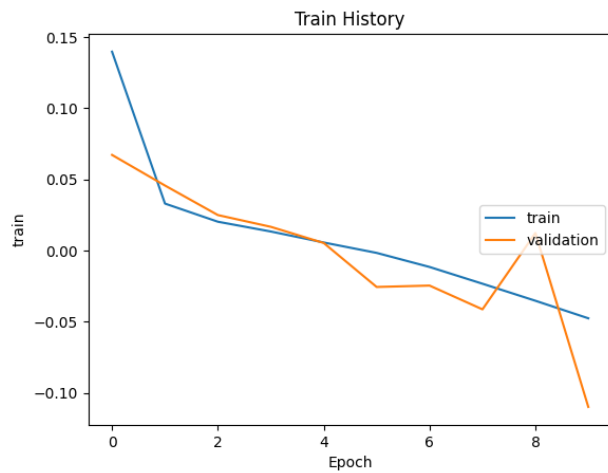


Figure 13: Train and Validation for Seg-Net (loss)

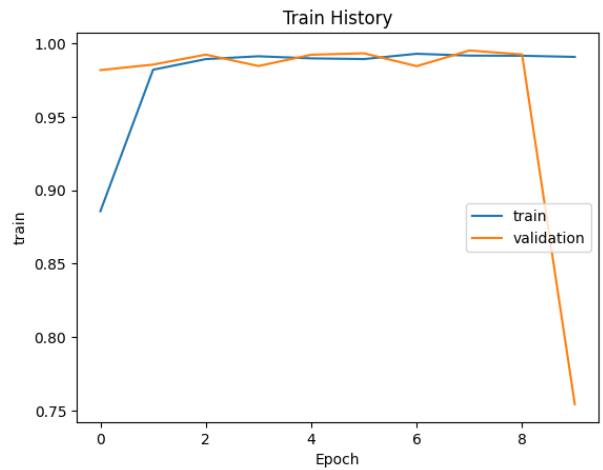


Figure 16: Train and Validation for PsP-Net (Dice Coefficient)

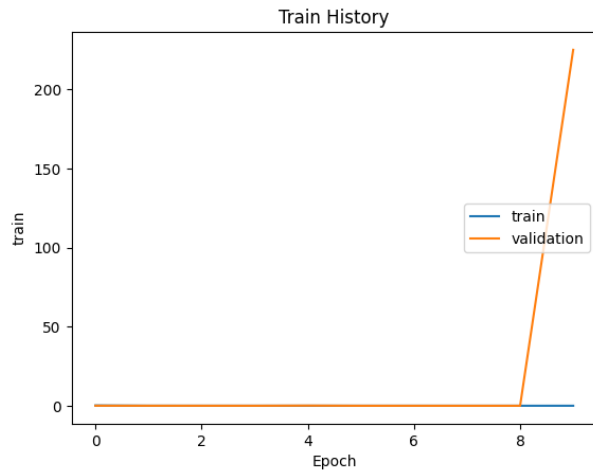


Figure 17: Train and Validation for PsP-Net (loss)

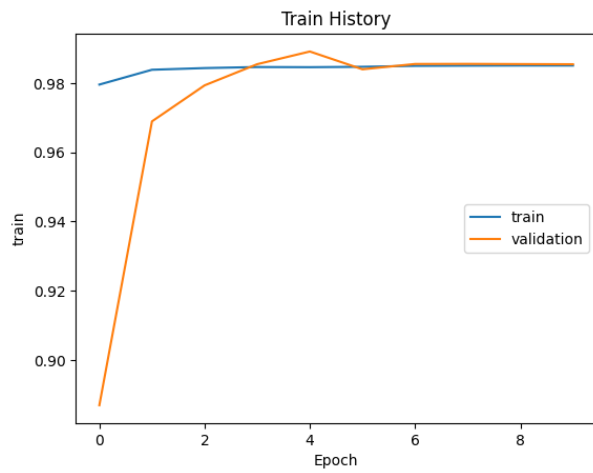


Figure 18: Train and Validation for Unet++ (loss)

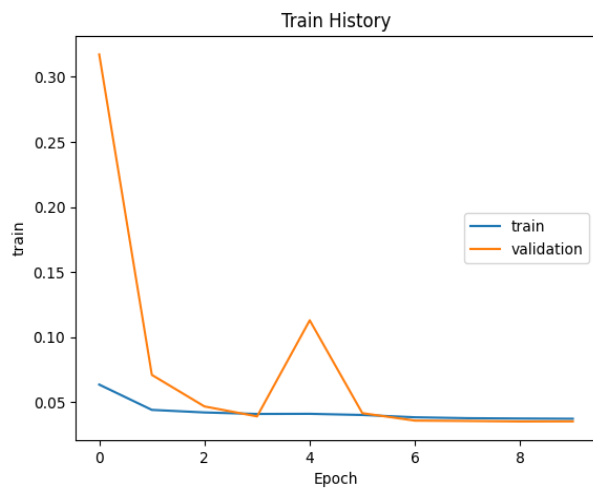


Figure 19: Train and Validation for Dense-Net (loss)

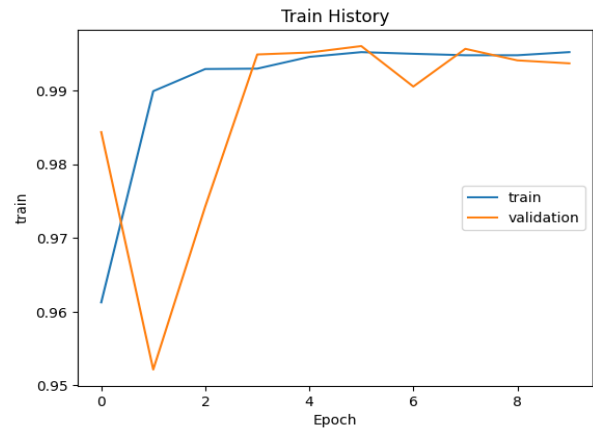


Figure 20: Train and Validation for Unet++ (Dice Coefficient)

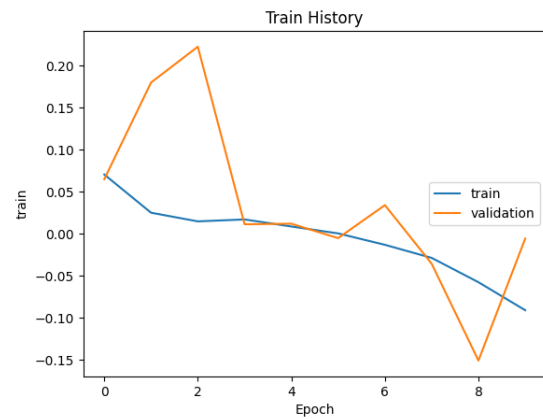


Figure 21: Train and Validation for Dense-Net (Dice Coefficient)

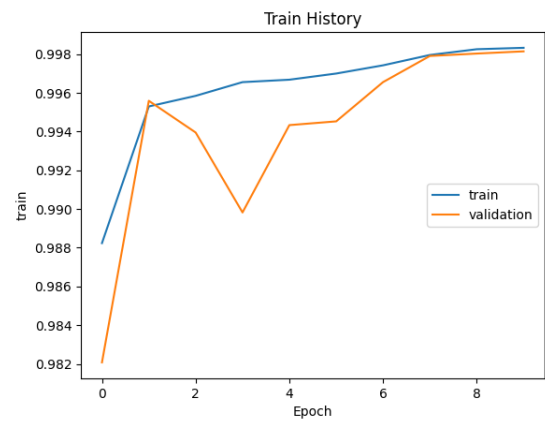


Figure 22: Train and Validation for DenseLab (Dice Coefficient)

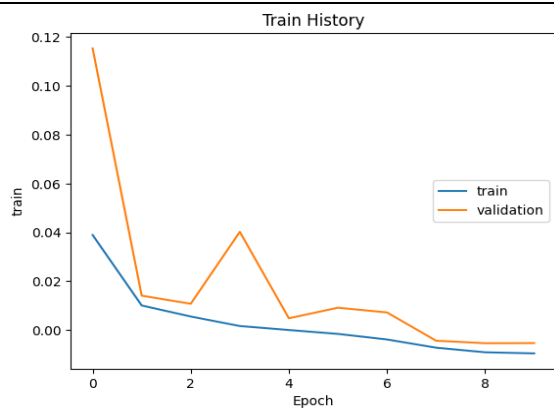


Figure 23: Train and Validation for DenseLab (loss)

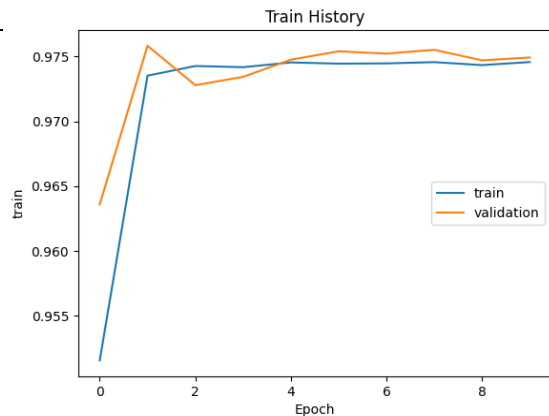


Figure 26: Train and Validation for HR-Net (Dice Coefficient)

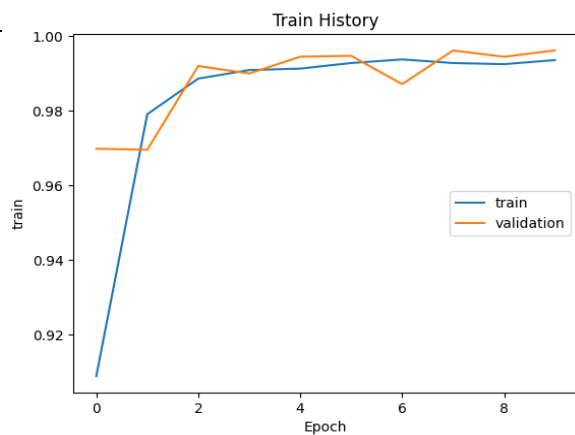


Figure 24: Train and Validation for V-Net (Dice Coefficient)

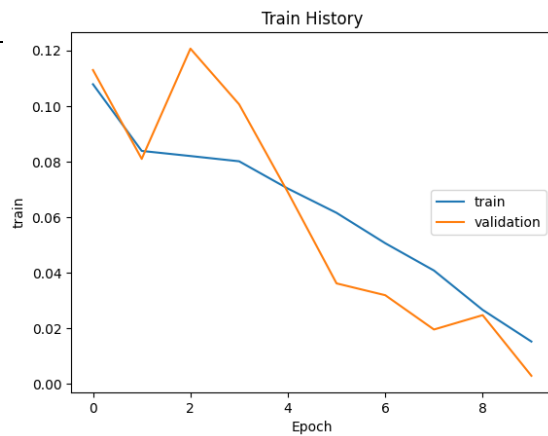


Figure 27: Train and Validation for HR-Net (loss)

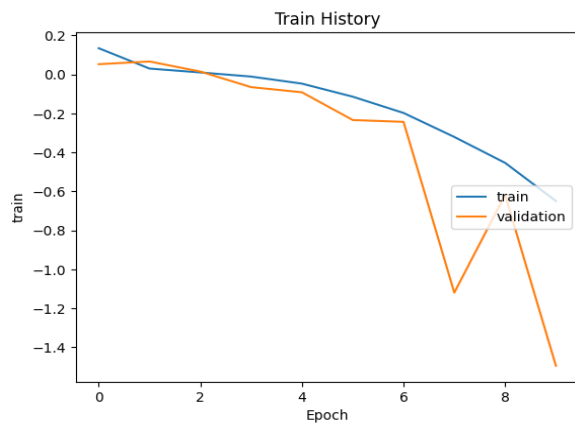


Figure 25: Train and Validation for V-Net (loss)

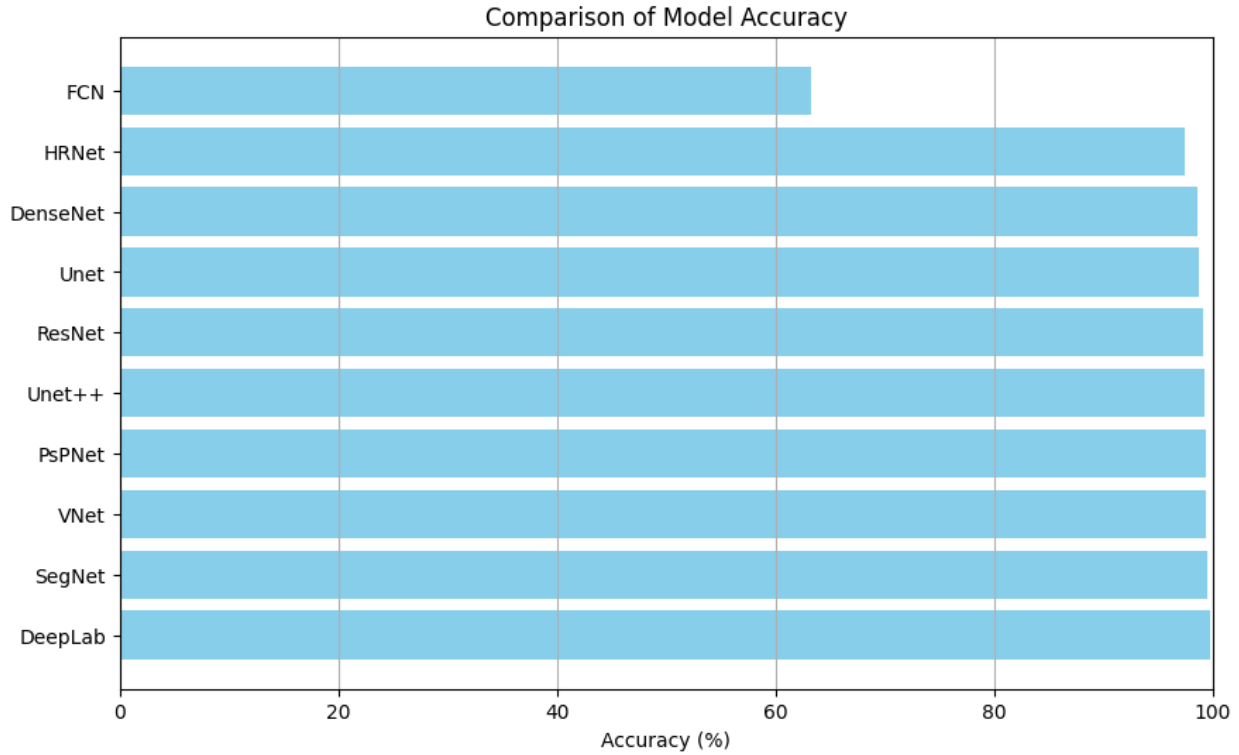


Figure 28: Comparison of Models

## 5. Conclusions and Future Scope

In conclusion, the application of various deep learning architectures in liver tumor segmentation presents a promising avenue for improving medical diagnosis and treatment planning. Through the utilization of FCN, U-Net, SegNet, ResNet, PsPNet, DenseNet, Unet++, DeepLab, VNet, HRNet architectures, significant strides have been made in accurately identifying and delineating liver tumors from medical imaging data.

The FCN architecture, with its fully convolutional approach, has demonstrated commendable performance in capturing spatial information and generating precise segmentation masks. Similarly, the U-Net architecture, with its expansive contracting and expansive paths, has showcased exceptional capability in segmenting liver tumors with high accuracy, particularly in scenarios with limited training data.

SegNet's encoder-decoder architecture, ResNet's deep residual learning framework, and DenseNet's densely connected layers have each contributed to enhancing feature extraction and propagation, thereby improving segmentation performance across diverse datasets. PsPNet's pyramid pooling module has effectively captured multi-scale contextual information, enhancing the model's ability to delineate tumors of varying sizes.

The Unet++ architecture, with its nested skip connections, has further refined the segmentation process by facilitating better feature reuse and integration. DeepLab's employment of dilated convolutions has enabled precise boundary delineation, enhancing the delineation of tumor margins.

Moreover, VNet's volumetric approach and HRNet's high-resolution representations have both advanced liver tumor segmentation by capturing intricate details and preserving spatial information, particularly in volumetric medical imaging data.

In summary, the adoption of these deep learning architectures has significantly advanced liver tumor segmentation, offering clinicians and researchers powerful tools for accurate and efficient diagnosis, treatment planning, and monitoring of liver cancer patients. However, further research and validation are warranted to ensure robustness, generalizability, and clinical applicability across diverse patient populations and imaging modalities. **Architecture Refinement:** Researchers can explore fine-tuning SegNet's architecture to enhance its performance even further. This may involve optimizing hyperparameters, adjusting

layer configurations, or incorporating novel architectural elements to better capture intricate tumor features.

**Techniques Integration:** Integration of advanced techniques such as attention mechanisms or adversarial training could be explored to augment SegNet's segmentation capabilities. Attention mechanisms can help focus on relevant regions within the images, potentially improving segmentation accuracy, while adversarial training can enhance the model's robustness against variations in input data.

**Real-time Implementation:** Efforts can be directed towards integrating SegNet into real-time medical imaging systems. The development of efficient inference mechanisms and hardware acceleration techniques can enable SegNet to perform rapid and accurate tumor segmentation, facilitating timely clinical decision-making.

**Multimodal Fusion:** Incorporating multimodal imaging data, such as combining MRI, CT, and PET scans, can further improve the accuracy and reliability of liver tumor segmentation. Fusion techniques can leverage the complementary information provided by different imaging modalities to produce more comprehensive segmentation results.

**Clinical Validation and Deployment:** Extensive clinical validation studies are essential to validate SegNet's performance in real-world healthcare settings. Collaborations between computer scientists, medical professionals, and regulatory bodies are crucial for ensuring the safe and effective deployment of SegNet in clinical practice.

**Generalization to Other Applications:** While SegNet excels in liver tumor segmentation, its applicability can extend to other medical imaging tasks, such as the detection and classification of abnormalities in various organs. Exploring the adaptability of SegNet to different medical imaging domains can widen its impact on healthcare.

In conclusion, while SegNet stands out as the preferred choice for liver tumor segmentation, ongoing research and innovation will undoubtedly further enhance its capabilities and broaden its applicability in the field of medical imaging analysis.

## References

- [1] V. Badrinarayanan, A. Kendall and R. Cipolla, "SegNet: A Deep Convolutional Encoder-Decoder Architecture for Image Segmentation," in *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 39, no. 12, pp.2481-2495, 1 Dec. 2017, doi: 10.1109/TPAMI.2016.2644615. .
- [2] Zhang, C., Lu, W., Wu, J., Ni, C. and Wang, H., 2024. SegNet Network Architecture for Deep Learning Image Segmentation and Its Integrated Applications and Prospects. *Academic Journal of Science and Technology*, 9(2), pp.224-229.
- [3] Kumar, P., Nagar, P., Arora, C. and Gupta, A., 2018, October. U-segnet: fully convolutional neural network based automated brain tissue segmentation tool. In *2018 25th IEEE International conference on image processing (ICIP)* (pp. 3503-3507). IEEE.
- [4] Saood, A. and Hatem, I., 2021. COVID-19 lung CT image segmentation using deep learning methods: U-Net versus SegNet. *BMC Medical Imaging*, 21, pp.1-10.
- [5] M. Agarwal, "Optimized contrast enhancement for tumor detection," *Int. J. Imaging Syst. Technol.* Willey, pp. 1–17, 2020, doi: 10.1002/ima.22408.
- [6] Yue, Y., Li, X., Zhao, H. and Wang, H., 2020, October. Image segmentation method of crop diseases based on improved SegNet neural network. In *2020 IEEE International Conference on Mechatronics and Automation (ICMA)* (pp. 1986-1991). IEEE.
- [7] Ramesh, K.K.D., Kumar, G.K., Swapna, K., Datta, D. and Rajest, S.S., 2021. A review of medical image segmentation algorithms. *EAI Endorsed Transactions on Pervasive Health and Technology*, 7(27), pp.e6-e6.
- [8] Wang, Risheng, Tao Lei, Ruixia Cui, Bingtao Zhang, Hongying Meng, and Asoke K. Nandi. "Medical image segmentation using deep learning: A survey." *IET Image Processing* 16, no. 5 (2022): 1243-1267.
- [9] Milletari, F., Navab, N. and Ahmadi, S.A., 2016, October. V-net: Fully convolutional neural networks for volumetric medical image segmentation. In *2016 fourth international conference on 3D vision (3DV)* (pp. 565-571). Ieee.
- [10] Gul, Sidra, Muhammad Salman Khan, Asima Bibi, Amith Khandakar, Mohamed Arselene Ayari, and Muhammad EH Chowdhury. "Deep learning techniques for liver and liver tumor segmentation: A review." *Computers in Biology and Medicine* 147 (2022): 105620.
- [11] Bilic, P., Christ, P., Li, H.B., Vorontsov, E., Ben-Cohen, A., Kaissis, G., Szeskin, A., Jacobs, C., Mamani, G.E.H., Chartrand, G. and Lohöfer, F., 2023. The liver tumor segmentation benchmark (lits). *Medical Image Analysis*, 84, p.102680.
- [12] Almotairi, Sultan, Ghada Kareem, Mohamed Aouf, Badr Almotairi, and Mohammed A-M. Salem. "Liver tumor segmentation in CT scans using modified SegNet." *Sensors* 20, no. 5 (2020): 1516..
- [13] Rela, Munipraveena, Nagaraja Rao Suryakari, and P. Ramana Reddy. "Liver tumor segmentation and classification: A systematic review." *2020 IEEE-HYDCON* (2020): 1-6.
- [14] Zhang, Jianpeng, Yutong Xie, Pingping Zhang, Hao Chen, Yong Xia, and Chunhua Shen. "Light-Weight Hybrid



- Convolutional Network for Liver Tumor Segmentation." In *IJCAI*, vol. 19, pp. 4271-4277. 2019.
- [15] Seo, Hyunseok, Charles Huang, Maxime Bassenne, Ruoxiu Xiao, and Lei Xing. "Modified U-Net (mU-Net) with incorporation of object-dependent high level features for improved liver and liver-tumor segmentation in CT images." *IEEE transactions on medical imaging* 39, no. 5 (2019): 1316-1325.
  - [16] Li, W., Jia, F. and Hu, Q., 2015. Automatic segmentation of liver tumor in CT images with deep convolutional neural networks. *Journal of Computer and Communications*, 3(11), pp.146-151.
  - [17] Tran, Song-Toan, Ching-Hwa Cheng, and Don-Gey Liu. "A multiple layer U-Net, U n-Net, for liver and liver tumor segmentation in CT." *IEEE Access* 9 (2020): 3752-3764.
  - [18] Jiang, Huiyan, Tianyu Shi, Zhiqi Bai, and Liangliang Huang. "Ahnnet: An application of attention mechanism and hybrid connection for liver tumor segmentation in ct volumes." *IEEE Access* 7 (2019): 24898-24909.
  - [19] Luan, Shun Yao, Xudong Xue, Yi Ding, Wei Wei, and Benpeng Zhu. "Adaptive attention convolutional neural network for liver tumor segmentation." *Frontiers in Oncology* 11 (2021): 680807.
  - [20] Christ, Patrick Ferdinand, Florian Ettlinger, Felix Grün, Mohamed Ezzeldin A. Elshaera, Jana Lipkova, Sebastian Schlecht, Freba Ahmaddy et al. "Automatic liver and tumor segmentation of CT and MRI volumes using cascaded fully convolutional neural networks." *arXiv preprint arXiv:1702.05970* (2017).