Liver Segmentation from Combined Tomography (CT) Abdominal Images Data Using Deep Neural Networks

Author

Fatima tuz Zahra Khan

Regn Number

00000206188

Supervisor

Dr. Omer Gilani

NUST SCHOOL OF MECHANICAL AND MANUFACTURING ENGINEERING

NATIONAL UNIVERSITY OF SCIENCES AND TECHNOLOGY,

ISLAMABAD

February 2021

Liver Segmentation from Combined Tomography (CT) Abdominal Images Data Using Deep Neural Networks

Author

Fatima tuz Zahra Khan

Regn number

00000206188

This work is submitted as a Thesis report in partial fulfillment of the requirement for the degree of MS Biomedical Sciences

Thesis Supervisor:

Dr. Omer Gilani

Thesis Supervisor's Signature

School of Mechanical and Manufacturing Engineering

National University of Sciences and Technology

Islamabad, Pakistan

Feburary 2021

Declaration

I certify that this research work titled "*Liver Segmentation from Combined Tomography (CT) Abdominal Images Data Using Deep Neural Networks"* is my own work. The work has not been presented elsewhere for assessment. The content of this document represent my original work, the material is based on my personal research and findings on this topic wherever need be with additional information sources duly cited.. The thesis is submitted with signature of the supervisor and mine and has been proof-read for errors.

Signature of Student

Fatima tuz Zahra Khan

00000206188

Plagiarism Certificate (Turnitin Report)

This thesis has been checked for Plagiarism. Turnitin report endorsed by Supervisor is attached.

Signature of Student

Fatima tuz Zahra Khan

Registration number: 00000206188

Name & Signature of Supervisor

DR. OMER GILANI

Signature: _____

Copyright Statement

FORM TH-4

National University of Sciences & Technology

MASTER THESIS WORK

We hereby recommend that the dissertation prepared under our supervision by: (Student Name & Regn No.) Fatima tuz Zahra Khan Reg no. 00000206188
Titled: Liver Segmentation from Combined Tomography (CT) Abdominal Images data using Deep Neural Networks be accepted in partial fulfillment of the requirements for the award of M.S Biomedical Sciences degree and awarded grade _____ (Initial).

Examination Committee Members

1. Name: Dr. Mohsin Jamil Signature:_____

2. Name: Dr Asim Waris Signature:_____

3. Name: Dr. Umar Ansari Signature:_____

Supervisor's name: Dr. Omer Gilani Signature:_____
Date:_____

_____ _____
Head of Department Date

COUNTERSINGED

Date:_____ _____
Dean/Principal

I Dedicate this thesis to my exceptional parents whose tremendous support and cooperation led me to this wonderful accomplishment.

TABLE OF CONTENTS

List of Acronyms

Computed Tomography	CT
Magnetic Resonance imaging	MRI
Convolutional Residual Network	CRN
Deep Learning	DL
Fully convolutional network FCN	FCN
Convolutional neural networks (CNN)	CNN
Recurrent Neural Network (RNN)	RNN

List of Figures

Abstract

The field of medical image segmentation is rapidly advancing over time. Segmentation of organs helps medical professionals to plan radiotherapies and track the prognosis of disease in general. Image segmentation of abdominal organs can facilitate clinical procedures for example transplantation surgeries and to determine the precise location of organ in aortic abdominal surgeries. The study focuses primarily on segmentation of liver from computed tomography (CT) dataset of healthy abdominal organs. The dataset taken for the study is from Combined (CT-MR) Healthy Abdominal Organ Segmentation (CHAOS) Challenge. The dataset used three-dimensional data to enrich the level of analysis. The dataset used is in DICOM format. All the patient images were of healthy liver, aligned in the same position while they were given contrast injection for enhanced vascular structure. The automated model used for the segmentation of images was U-NET. Two convolutional layers were used and the size of each image was halved. The expansive and contractive parts of the layers were repeated thrice. PYtorch and FastAl libraries were used for training the data set. 15% of the images were used for validation. The learning rate was altered during training to avoid overfitting. The metrics used for calculating results was dice. The calculated dice with our algorithm is 0.968 while the state-of-the-art dice reported by CHAOS for the given challenge was 0.979.

Chapter 1: Introduction

Image segmentation is an evolving and emerging topic in the field of computer vision and image processing. It is a broad field that consists of image compression, video surveillance, robotic perception, augmented reality and medical image segmentation. This field started emerging in the last decade of the 20th century and a lot of work is being done in the development of image segmentation algorithms as well as increasing the application of the field. Substantial amount of work is also being done in the development of deep learning (DL) model due to their efficiency and better results. [1]

Medical image segmentation is a subbranch of this field that is being explored by biomedical scientists in the recent past. This type of segmentation focuses on solving clinical problems to facilitate the medical professionals in multiple diagnostic and surgical procedures. Deep learning model are being widely used for medical image segmentation where the algorithm is capable of automatically learning the image features to segment them [2]

Image Segmentation via Deep learning-An overview of Deep Neural Networks:

From the earliest form of image segmentation models for instance thresholding and watershed, k-mean clustering and histogram-based bundling, the field of image segmentation has evolved onto models like active contours, conditional and Markov random fields and graph cuts. The modern algorithms generally provide better accuracy with less generalizability. These often achieve results close to benchmarks. There is two types of image classification problems, semantic and instance segmentation. The

semantic segmentation is a pixel level classification that labels image into different categories while instance segmentation delineates objects and classifies the image in a more precise manner. While evaluating an algorithm for efficiency different parameters like loss function, training strategies, training data type and the choice of network architecture is kept in consideration.

Following are the major types of deep learning models that are being used currently for image segmentation.

- R-CNN based models
- Convolutional models with graphical models
- Generative models and adversarial training
- Convolutional models with active contour model
- Recurrent neural network based models
- Fully convolutional networks
- Encoder-decoder based models
- Dilated convolutional models and DeepLab family
- Fully convolutional networks
- Multi-scale and pyramid network based models

Training Deep Learning Models:

Deep learning models are constantly evolving. Capsule networks, spatial transformer models and gated recurrent units are being proposed by the researchers to improve the performance of deep learning networks. Interestingly, deep neural networks can be trained on new datasets from scratch. This is done via transfer learning in which a

model that has been trained on one type of dataset is fed with a new type of dataset so that the model adapts to the new type of data and can help in the identification of new features in the dataset. A common method for transfer learning is training the dataset on ImageNet and then feeding new training set t the model. This helps in the recognition of features in dataset with less labeled samples.

Image Segmentation models:

Following are the major types of deep learning models that are most widely used in image segmentation.

R-CNN based Models:

The regional convolutional network is used for instance segmentation. It has three major types i.e., Fast R-CNN, masked R-CNN and Faster R-CNN. These models are widely used for object detection. Some versions are used for object detection and semantic segmentation. R-CNN models are widely used for instance segmentation. He *et al.* **[3]** proposed a model that efficiently detects objects while generating image masks for each instance. This model as three output branches that compute bounding box coordinates, associated coordinates and create binary mass for the images, respectively. The loss function of the model combines the loss function of bounding box coordinates, the segmentation class and the predictions and then trains them jointly. Other efficient models include Path Aggregation network (PANet) (Liu *et al.* [4]), Multitask network for instance-

aware segmentation (Dai *et al.* [5]), Masklab (Chen *et al.* [6]) and Tensormask (Chen *et al.* [7]). All of these models are widely used because of their efficiency.

Fully convolutional networks

Fully convolutional networks (FCN) are widely used, these models are used to produce segmentation map which size equivalent to the size of the image. It was first proposed by Long *et al.* [8] . Over years FCN has evolved as researchers modified networks like GoogLeNet and VGG16. The fully connected layers were replaced by fully convolutional layers. The model achieves state of the art results by up sampling and fusing the feature maps of earlier layers. The model is tested on n PASCAL VOC, NYUDv2, and SIFT Flow to validate the segmentation results achieved. FCNs are widely used for semantic segmentation.

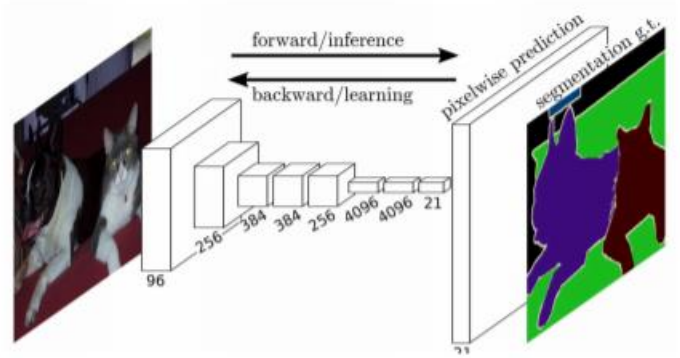

Figure 1: General working of fully convolutional networks

Despite its wide use and good result, FCN has a few drawbacks in comparison to other deep neural networks. First, conventional FCN does not take global context information into account, secondly it is comparatively slow in real time. A few researchers have worked on overcoming these shortcomings for instance ParseNet proposed by Liu *et al.* [9] using the concept of context vector and replaced convolutional layer. Fully convolutional networks are currently being widely used in brain tumor[10] and skin lesion [11] segmentation in the field of medical image segmentation.

Convolutional models with graphical models

Convolutional models which incorporate probabilistic graphical models like Conditional Random Fields and Markov Random Field are more efficient compared to FCNs. Convolutional neural networks (CNN) provide poor localization. Chen *et al.* [12] combined the final convolution layer with Conditional Random Fields that produced better localized results and resulted in less generalizability.

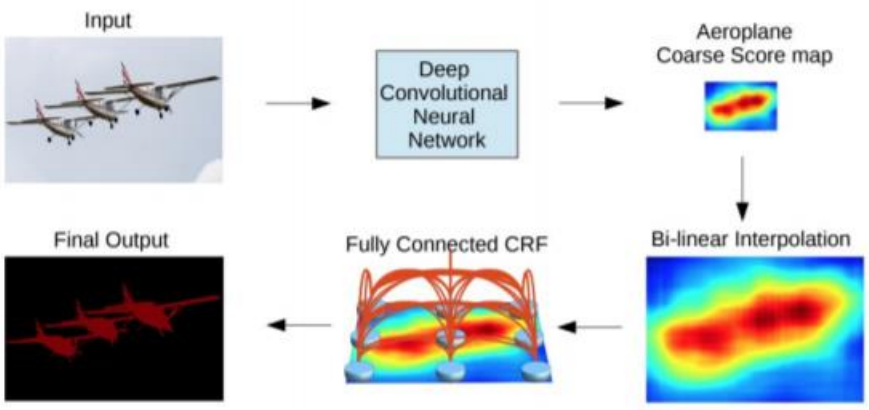

Figure 2: Graphical model (CNN+ Conditional Random Fields) proposed by Chen et al.

Few other models have been proposed for semantic segmentation that include combination of CNN and Conditional Random Fields have produced promising results. Lin *et al.* [13] proposed with concept of deep Conditional Random Fields and patch background to improve semantic segmentation. This model used contextual information. Liu *et al.* proposed model Parsing Network which combines CNN and Markov Random Field [14]

Encoder-decoder based models

Encoder decoder model are of special importance in medical image segmentation due to their wide usage and applicability. U net and V-net are the most commonly used encoder decoder models, these two are primarily used in medical image segmentation but they also have application outside this field. V-net is a type of FCN initially proposed by y Milletari *et al.* [15] It has been widely used for Computed Tomography (CT) and Magnetic resonance imaging (MRI) images. The model is specially used for images where there is a large difference between background and foreground voxels. U-net would be discussed further in the literature review section.

Medical Image Segmentation:

Medical image segmentation is a separate branch of image segmentation. Deep convolutional networks have witnessed immense progress and success in near past in this field. Medical image segmentation is done by traditional and deep learning image segmentation techniques. CT and MRI images are the most used types of images in radiological techniques. These are also most widely used in the medical segmentation. Initial techniques comprised of edge detection and mathematical models but the trend has shifted to mathematical models. Hand crafted features were the primary input of these systems but with the introduction of deep learning techniques since early 2000s, researchers have found deep learning to be more promising. Classification, registration and detection in medical image segmentation techniques hold primary importance.

Network Structures used in Medical Image Segmentation:

There are a number of networks that are widely used in image segmentation. Some of these networks are more successful and efficient compared to others. Given below (Figure 3) is a schematic diagram of these networks and their subtypes. These networks would be discussed one by one with reference to medical image segmentation below.

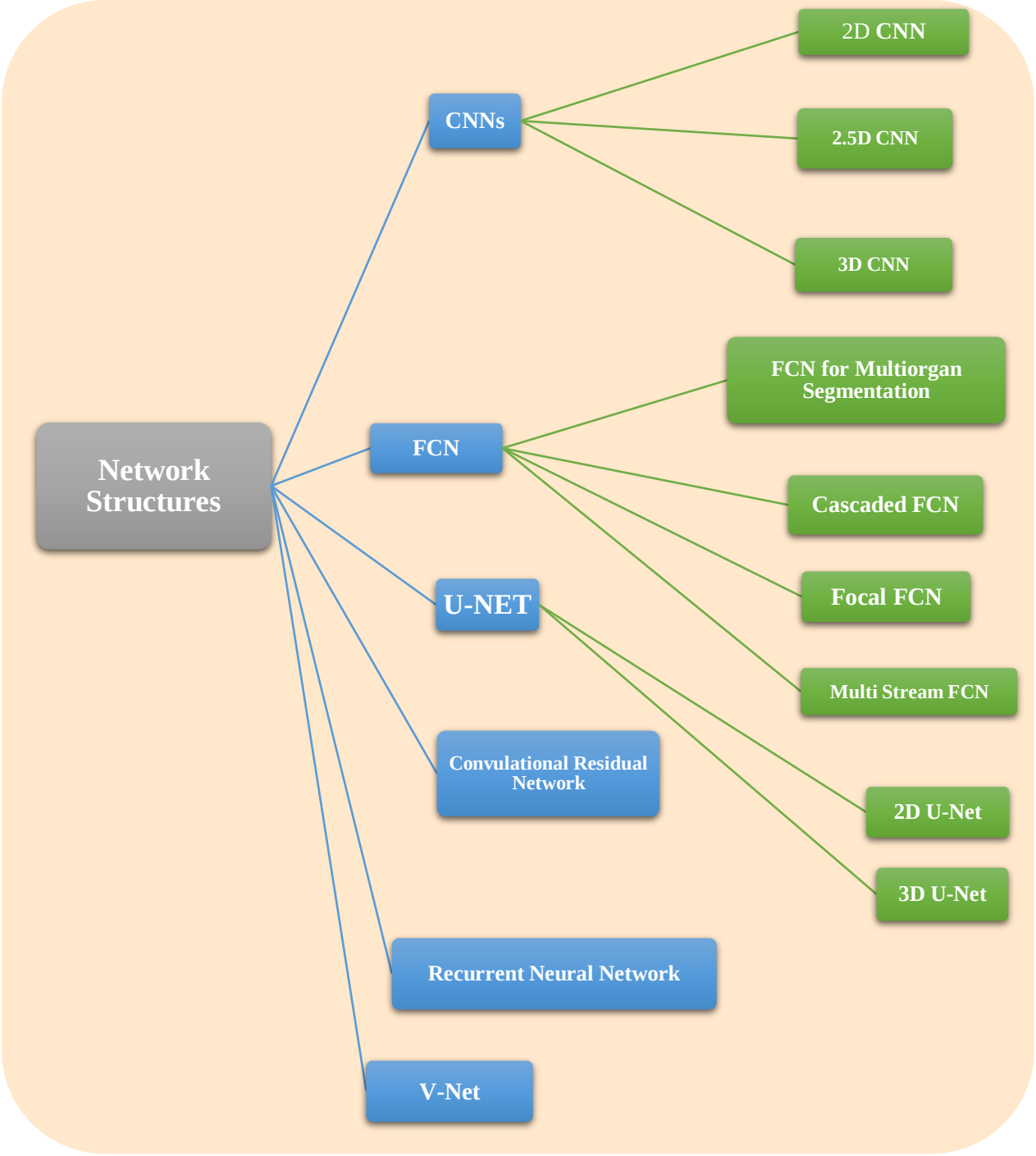

Figure 3:Types of Neutral Networks used in Medical Image Segmentation and their subtypes.

CNNs:

Convolutional neural networks have been discussed above in length. It is the type of neural network where each layer has a specified function for example pooling , l oss calculation etc. The first layer is the input layer, after the first layer is the input layer, after entering this layer the input data goes through a number of layers where the data is convolved. 'Kernels' are the filters present in each layer, they can be of arbitrary sizes and are specified by researchers who develop the CNNs. The output produced by each layer is known as activation map, the activation map basically highlights the impact of each filter or the information that is sorted out after passing through the layer. Activation layer adds no linearity to the output, the next layer is generally the pooling layer. Pooling is done by either average pooling or max pooling depending upon the type of dataset and the choice of researcher. High level abstractions result from fully connected layers. Weights are optimized consistently for kernels and neural connections during the back propagation phase. (Commandeur *et al.*)[16]. There are three types of CNN depending upon the type of image being used to process. 2D CNN are used to process 2D images and same is for 2.5 and 3D CNNs. [17]

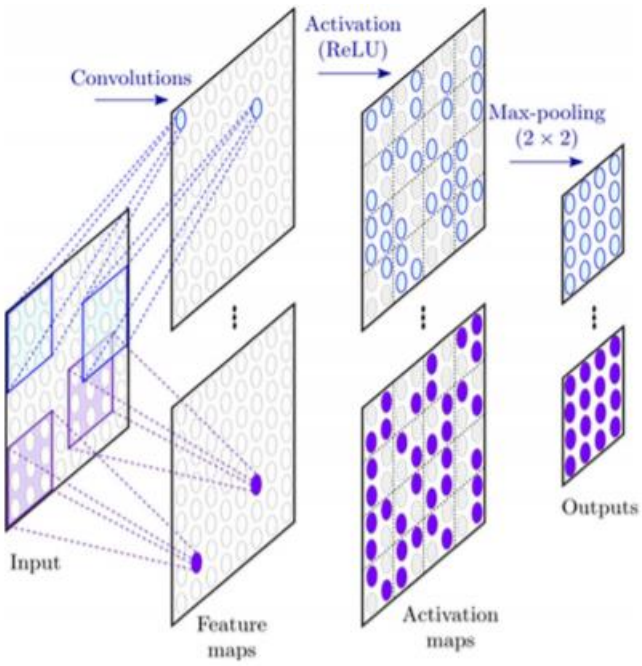

Figure 4: General Structure Of CNN.

Fully Convolutional Network:

The general functioning and mechanism of FCNs have been discussed above. Here the focus would be on the types of FCN specially employed for medical image segmentation.

Multi Organ Segmentation:

These types of FCNs are used for segmenting multiple organs at once. These types of FCN can use 2D, 2.5D and 3D image types. Bigger organs are easier to segment via FCNs. (Gibson at el.) [18]

Cascaded FCN:

This type of FCN uses a stacking technique, where FCNs are stacked. These models are useful for extracting contextual features using prediction maps. Cascaded FCNs provide better segmentation results.

Focal FCN:

This type of FCN apply focal loss to reduce the number of false positives in the results. In this type of segmentation, simple FCN is used for primary segmentation and then focal FCN is used to reduce the false positives. (Zhou e*t al.*)[19]

Multi-Stream FCN:

Multi-stream FCN is applicable on 3D images. It uses multiple form of images of same organ for better segmentation results. This technique improves the robustness of the system due to the wide range of image availability. This type of system uses contextual information for better results. (Zen *et al.*) [20]

Convolutional Residual Network

Convolutional Residual Network (CRN) is a deeper network. This type of networks have better learning capability but as the depth of the network increase, it faces a few problems. First, the accuracy of network degrades due to the vanishing of gradient. This problem is solved by feeding the network with a feature map after every few layers. The residual maps don’t let the features to fade due to the depth of the network

and hence these networks provide better learning capability with good segmentation results. Figure below provides a general overview of the CRNs.

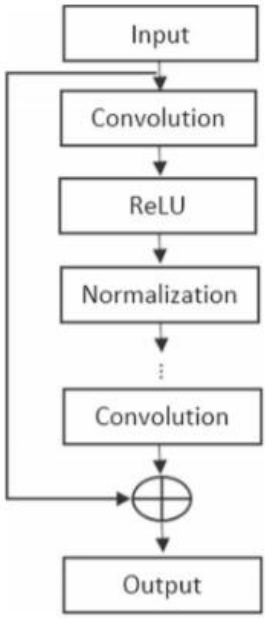

Figure 5: Single Residual Block of CRN.

Recurrent Neural Network

Recurrent Neural Network (RNN) memorize the patterns of the previous inputs. In medical images like CT and MRI, the information is scattered over multiple slices. This type of networks extratct sequential information from adjacent layers. The basic RNN consist of a CNN depending upon the type of dataset for inter-slice extraction of information. The second part is RNN itself for intra-slice information extraction.

V-net

V-Net is also widely used in medical image segmentation. It has been discussed above.

U-Net:

U-Net would be discussed in the literature review chapter in detail.

Types of Training Techniques:

Following are the major training techniques used in deep learning.

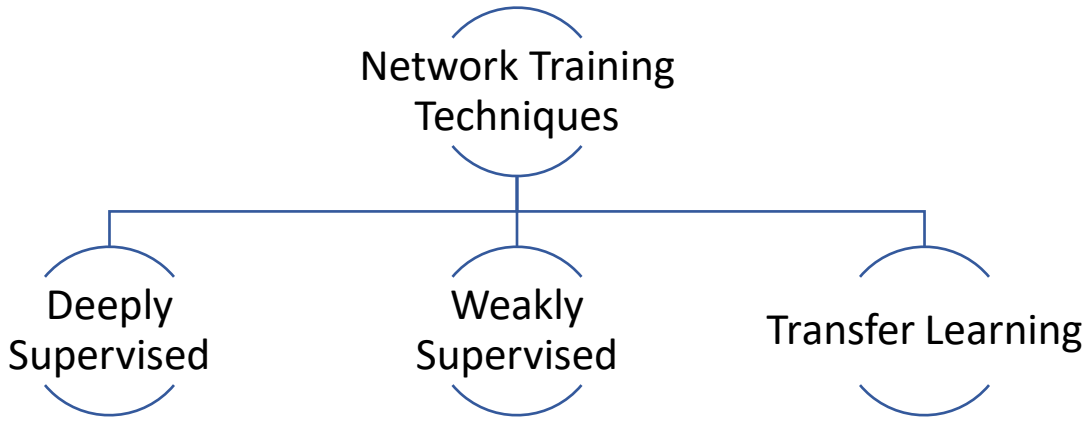

Figure 6: Network Training Techniques.

Challenges in Medical Image Segmentation via deep learning

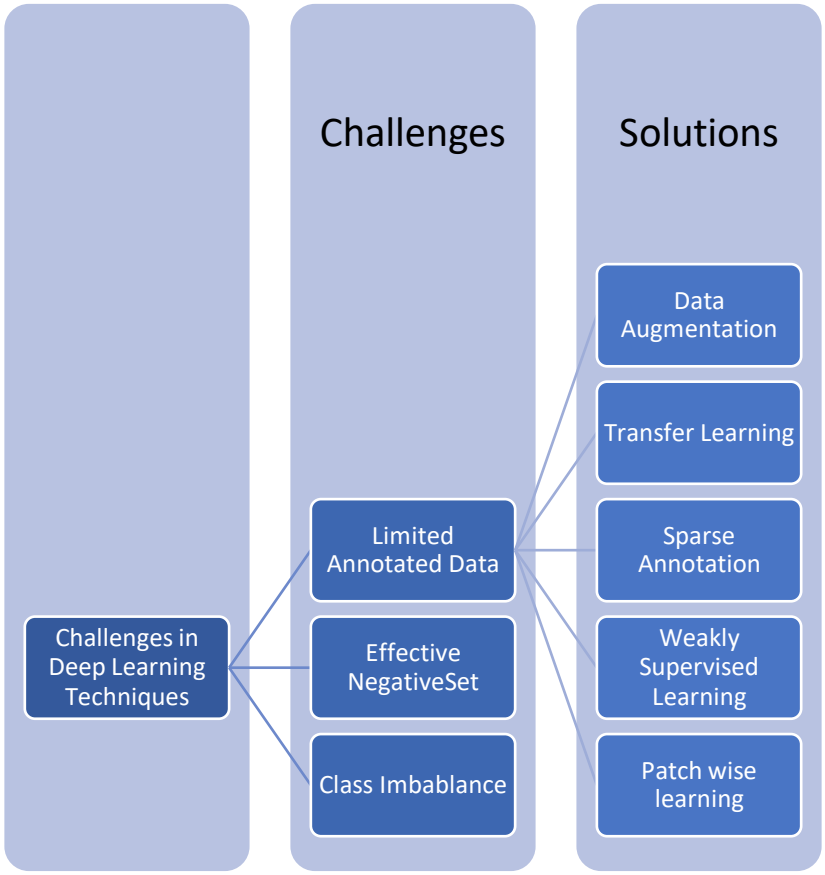

Figure 7: Challenges in Medical Image Segmentation Via Deep Learning.

Challenges in training deep Learning models:

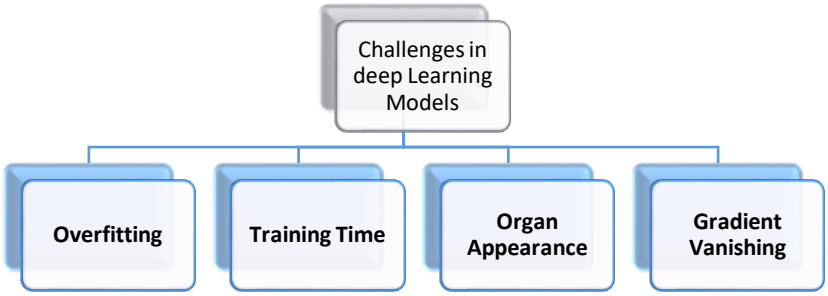

Figure 8: Challenges in Training Deep Learning Models.

Chapter 2: Literature Review

Medical image analysis uses six different image systems for classification and diagnosis. As mentioned above deep learning is a relatively new field in image analysis. Among different systems used in deep learning CNNs are most widely used. CNNs were not efficient because of large dataset sizes for a long time but they have gained efficiency in the past years. U-net is a type of U-net proposed in 2015 by Olaf Ronneberger *et al.*

U-net was used as the neural network for segmentation. U-net is U-net is able to do image localisation by predicting the image pixel by pixel. The network is strong enough to do good prediction based on even few data sets by using excessive data augmentation techniques. Its specially useful in segmentation of sparse regions. It is also particularly useful with small datasets.

Architecture of U-net:

Basic U-net

The basic U-net consists of two fragments: The first segment plays role as a constricting track constituting of a conventional convolutional neural network (CNN) architecture. The constricting track contains block which are composed of two continuous 3x3 convolutions, a max-pooling layer and a ReLU activation unit. This structure is replicated numerous times. The up convolutions (2x2) attribute in the constricting path helps in up sampling the feature map then feature map is cropped and attached as a string onto the upsampled feature map. In the next stage two consecutive convolutions (3x3) and ReLU activation occurs. An additional 1x1 convolution is used to cut the excessive pixel features from the edges of feature map to produce segmented

image. [1] Outcome is formation of a u-shaped network and propagation of contextual information across the network.

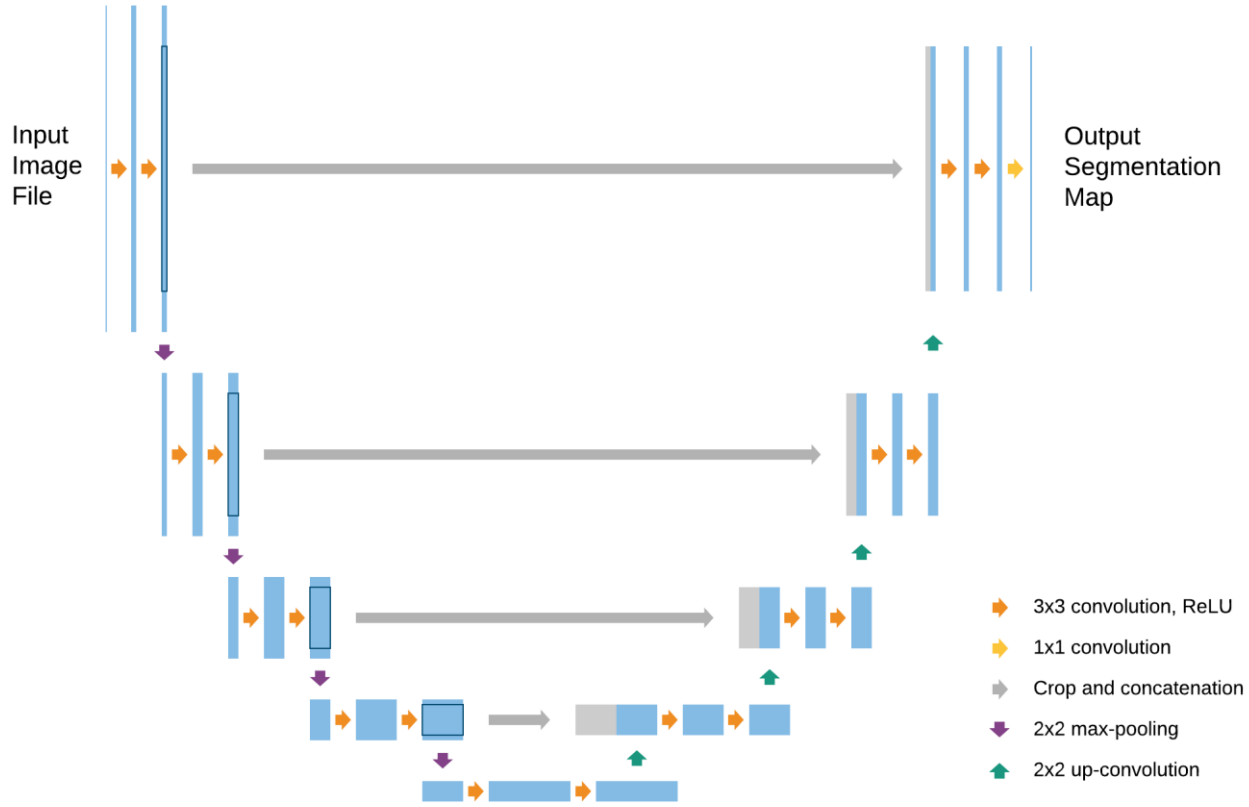

Figure 9: Structural Design of basic U-net. Blue boxes are feature map and cropped feature maps are shown as grey boxes.

The U-net's energy formula is:

$$E = \sum w(x) \log \left(p_{k(x)}(x)\right),$$

pk symbolises pixel-wise SoftMax function

$$p_k(x) = \frac{e^{a_k(x)}}{\sum_{k'=1}^{K} e^{a_k(x)'}}$$

The activation of channel k is represented as ak (x)

3Dimenssional U-net

3-dimenssional segmentation of images can be attained when the 2D operations in the basic U-net are replaced with equivalent 3Doperations. It facilitates 3D volumetric division of images. Extensive labelling of images can be avoided with 3D U-net segmentation because it is easier for the network to analyse a 3D image as it comprises replication of similar shapes and constructs. [2] In this study [3] they designed an image segmentation network that can interpret more details from the given image through multi-level divisions, useful in disease diagnosis. The most cited uses of 3D U-net in biomedical applications are diagnosis of cardiac structures, brain cancer, vertebral column, nasopharyngeal malignancy, lung nodules, head and neck analysis, white matter tract segmentation, and multi-organ dissection through CT and MR scanning applications.

Attention U-net

Attention U-net can focus on targeted important components and ignore the unwanted areas of the given image, this feature makes the network highly desirable for image processing jobs. The use of attention gates in this network is the key enabler of its performance goals which are improved segmentation and least computationally complex network assembly. [4] An attention gate is installed into each layer of expansive path which crops out the unneeded features before the concatenation of corresponding elements from the contracting path with the up-sampled elements in the expansive path.

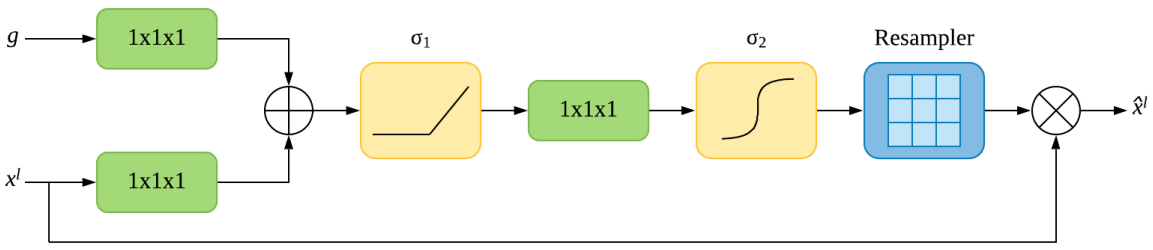

Figure 10: Schematic of Additive Attention Gate.

The attention unit offers focused classification information against generalized classification and can acclimatise to specific parts of an image. The most popular gate owing to its better precision is additive attention. The additive-attention-gate is identified with this equation:

$$q_{att}^{l} = \psi^{T}\left(\sigma_1\left(W_x^{T}x_i^{l} + W_g^{T}g_i + b_g\right)\right) + b_{\psi},$$

$$\alpha_i^{l} = \sigma_2\left(q_{att}^{l}\left(x_i^{l}, g_i; \Theta_{att}\right)\right)$$

Features from the contracting path as x^l. Gating signal as g. Sigmoid function as $\sigma 2(xi,c)$

$$\sigma_2\left(x_{i,c}\right) = \frac{1}{1+\exp\left(-x_{i,c}\right)},$$

The biomedical applications of attention U-net include diagnosis of lung cancer, melanoma, cervical cancer, brain tissue quantification, ophthalmic disease diagnosis, fetus development, and abdominal structure segmentation.

Inception U-net

This network utilizes various sizes of filters on the same layer in the network, then the outputs from various filters are compiled in a string and shifted onto the adjacent layer. Inculcation of

multiple different size filters enables the inception network to analyse images with numerous salient regions effectively. [5] The structural innovations like every 3x3 or larger filter and pooling layers to contain an added 1x1 convolution in front of it in parallel reduces the computational intricacy.

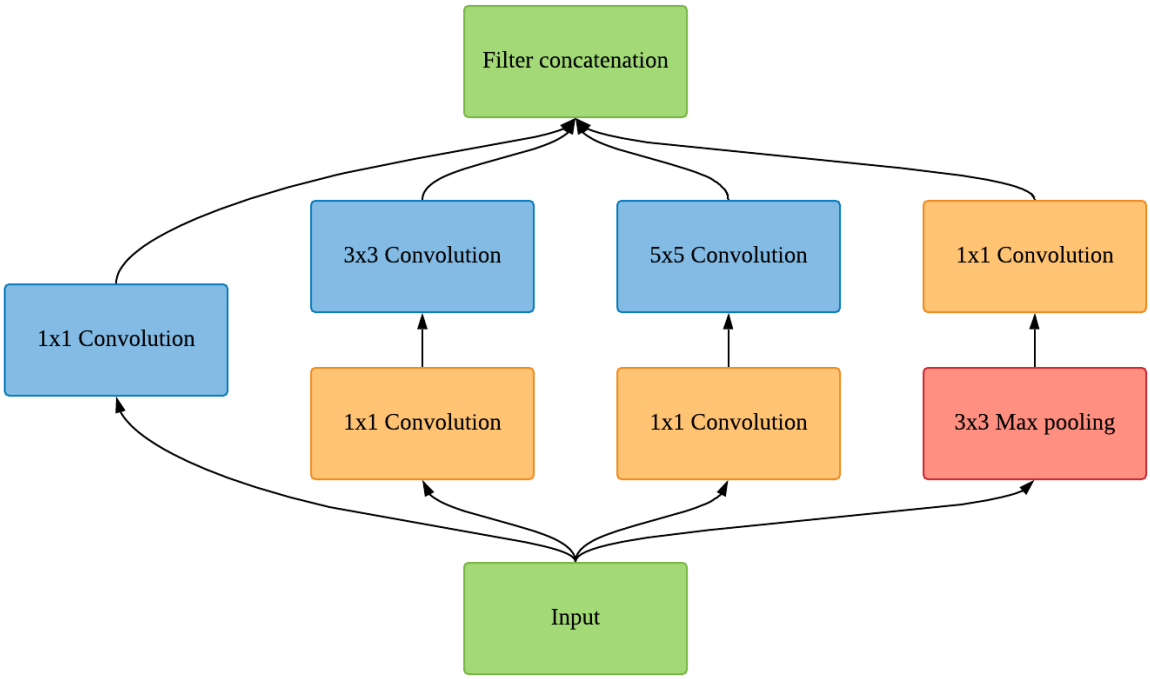

Figure 11: The Novel Block of inception in Google Net.

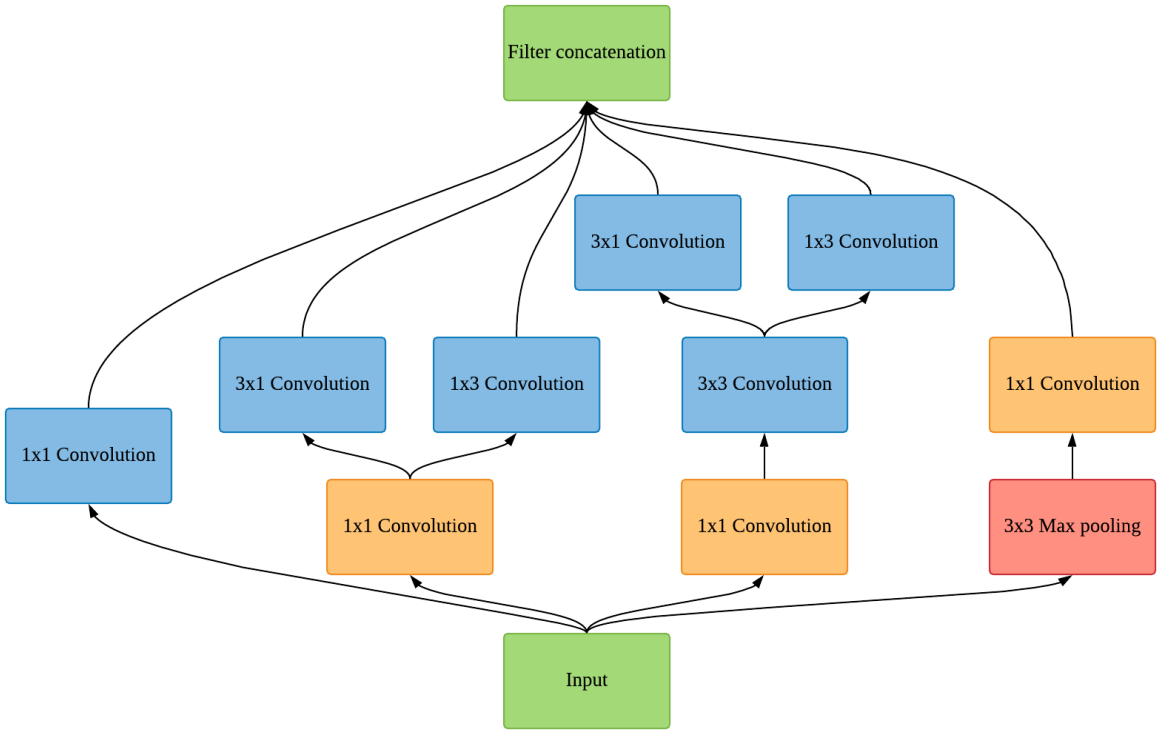

Figure 12 : Addition of Factorizes filters to improvise inception block.

Inception U-net applications include human embryo segmentation, cardiac segmentation, lung nodule detection, brain tissue mapping, ultrasound nerve segmentation and brain tumour detection.

Residual convolutional U-Net

The residual U-net variant is built on the Res-Net structural design to surmount the troubles during training deep-rooted neural networks. Experimental conclusions indicated that addition of more layers leads to overload, and degradation of performance. [6] ResNet minimizes this problem as it takes up the feature map from one layer using the skip connections and add it up into an alternative layer existing deep inside the net. The residual skip connection aids in

relieving the vanishing gradient problem, thus permitting formation of deeper neural U-net networks. Representation of each residual unit in this type of U-net:

$$y_l = h(x_l) + \mathcal{F}(x_l, \mathcal{W}_l),$$

$$x_{l+1} = f(y_l),$$

xl represents input, output of the residual unit *expressed as xl+1*, residual function *as F*(·), function *of activation as f*(·), and *h*(·) is the individuality function for mapping.

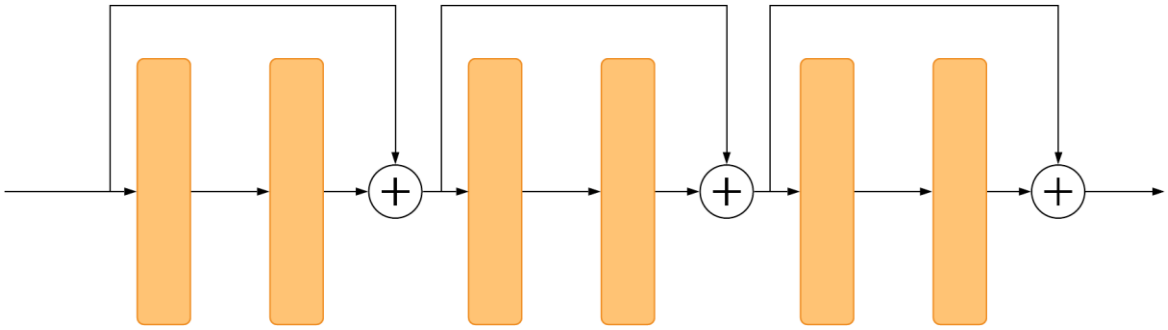

Figure 13: 3 consecutive Res-Net blocks separated by skip networks(connections).

The U-nets of deep residual category are perfect for the analysis of complex images. Deep residual U-nets have following applications in imaging such as retinal vessel segmentation, nuclei segmentation, brain tissue quantification, prostate cancer, brain structure mapping, breast cancer, endoscopy, bone structure analysis and osteosarcoma.

Network of Recurrent Convolutions

This category of networks assesses text or audio data or other sequential data only. The RCNN (recurrent convolutional neural network) includes not the conventional feedforward loops rather it includes in its convolutional layer the recurring feedback loops. Upon completion of both convolution feedback is applied and the activation provides the map (feature-map) produced into the related layer by a filter. [7] The feedback trait provides better efficiency and precision by allowing the units to revise these maps depending on the context provided by the neighbouring elements.

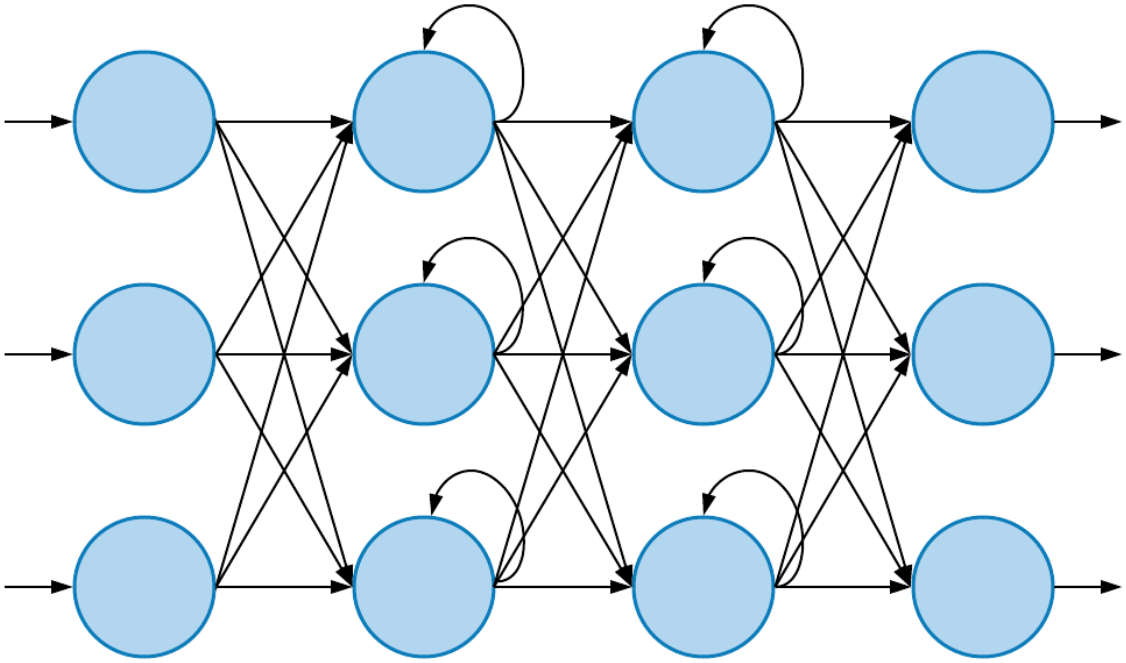

Figure 14: Recurrent neutral network.

The yield ‘y’ of the deep-recurrent U-net:

$$y_{ijk}^{l}(t) = \left(w_k^f\right)^T x_l^{f(i,j)}(t) + (w_k^r)^T x_l^{r(i,j)}(t-1) + b_k,$$

In this expression $x^f(t)$ is the feedforward input and $x^r(t-1)$ is the recurrent input for the l^{th} layer, w^f is the feedforward weight, w^r is the recurrent weight, and b is the bias of the k^{th} feature map.

Dense U-net

This U-net employs Dense-Net blocks instead of the regular layers. Dense-Net is a well-structured design constructed on top of ResNet with two alterations: (1) identity map is collected by each layer from all of its preceding layers (2) the identity maps are combined into tensors via network-wise chain formation. Thus, the identity mapping relies on all the prior layers to each layer in the block which allows protection of all identity maps from prior layers and substantial encouragement of gradient spread in a Dense-Net. Dense-Net allows more accuracy, deep learning, and less computational complexity. [8] The output for every layer should be like this:

$$x_l = H_l([x_0, x_1, x_2, \ldots, x_{l-1}]),$$

dense mapping function = $Hl(\cdot)$ signifies the (batch normalization), function activation = ReLU, and channel-wise chain formation = [].

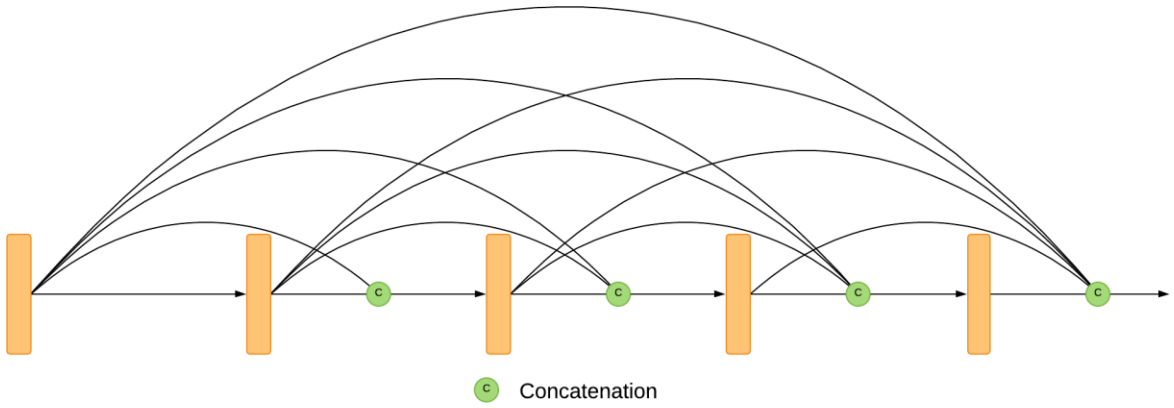

Figure 15: A 5 layered dense block is shown in the figure.

The implementation of dense blocks permits segmentation of components in an image with better distinction. Analysis of cerebral blood vessel segmentation, brain tumors, retinal blood vessel segmentation, lung, and liver cancer using this U-net.

U-net++

U-net++ can segment images more accurately by circulating more semantic information between the contracting and expansive path. It applies skip connections (in a dense form) as a transitional framework between the expansive and contracting paths. U-net++ comprises of several nodes of skip connection among each matching layer every skip connection obtains an up sampled map from its instant lower unit and feature map from all the previous units. This setup enables semantic information transmission among the 2 paths without any loss. [9]The skip connection operates as follows:

$$x^{i,j} = \begin{cases} \mathcal{H}(x^{i-1,j}), & j = 0 \\ \mathcal{H}([[x^{i,k}]_{k=0}^{j-1}, \mathcal{U}(x^{i+1,j-1})]), & j > 0' \end{cases}$$

Here $H(\cdot)$ denotes a convolution and activation operation, $U(\cdot)$ represents the up-sampling operation, and [] signifies a concatenation.

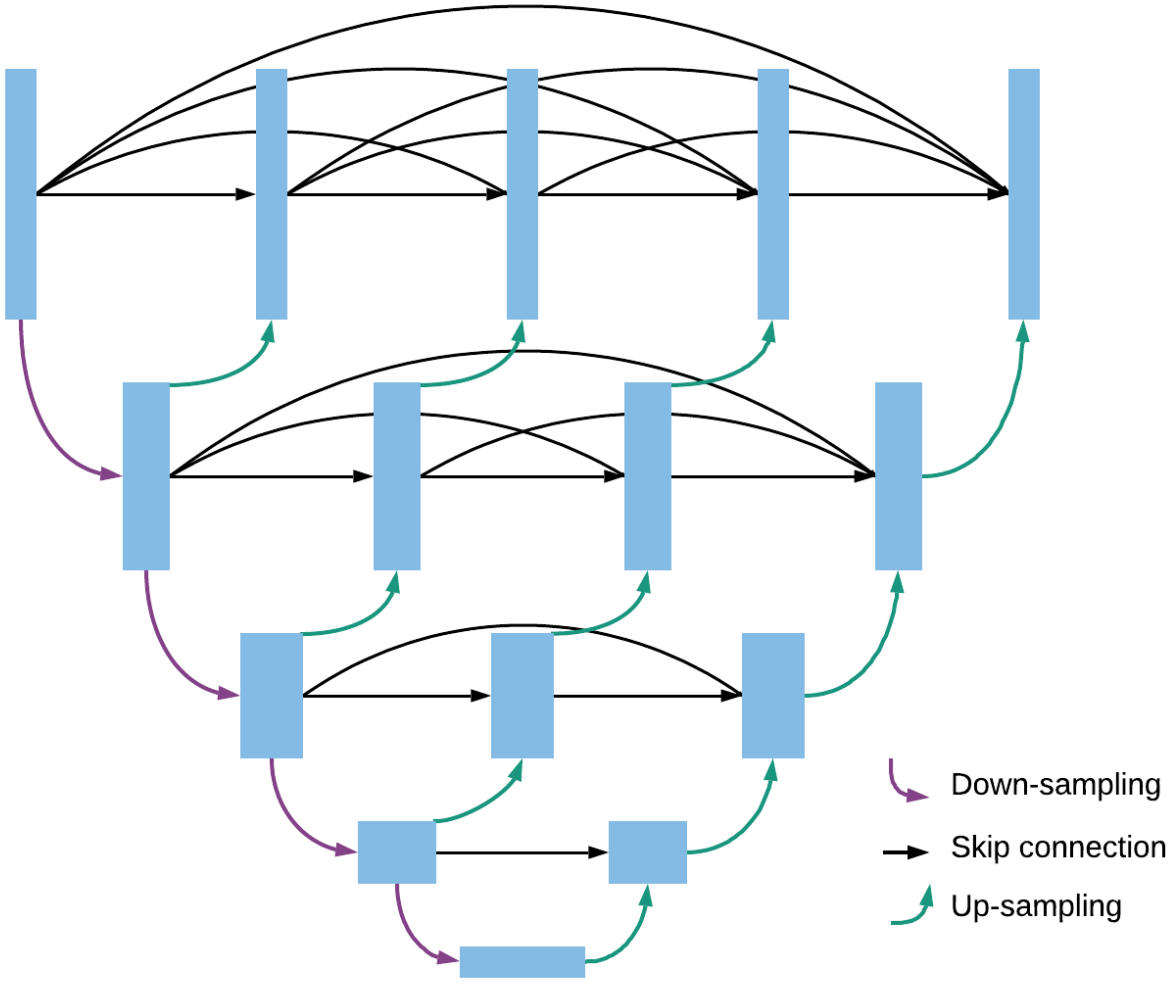

Figure 16: U-net++ schematic representation.

Adversarial U-net

In adversarial U-net a discriminator and a generator network compete against each other to enhance their performing. The discriminator network *D* plays role of an input data set classifier

and the G generates images that are intermittently supplied to the discriminator D and in this way error of producing fake as real data set images is minimized. [10]

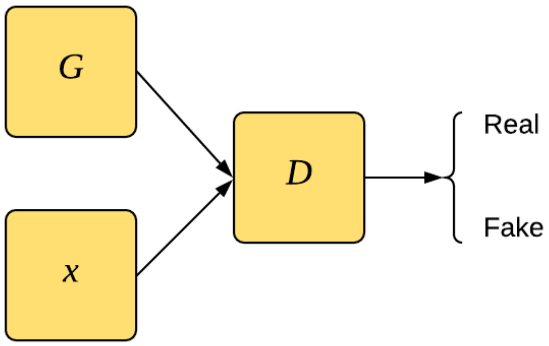

Figure 17: GAN block diagram of adversarial U-net.

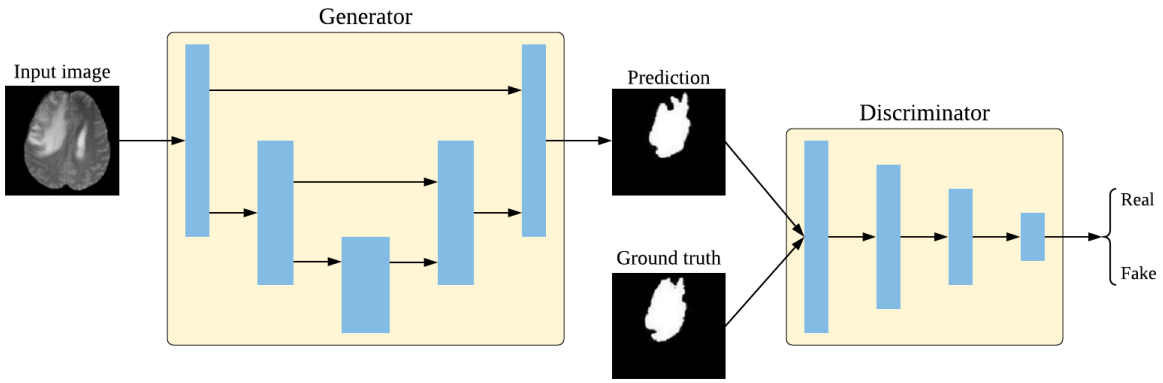

Figure 18: Streamlined diagram of u-net based GAN.

Cascaded arrangement

In this arrangement cascade of 2 or more than 2 U-nets is established the first layer executes upper-level segmentation and then dissection of minor pieces is executed by successive U-net layers. In one study they devised a U-net model which has two cascades in which the 1st U-net isolates the liver from rest of the body organs and the 2nd U-net sections tumours inside liver.

The performance of cascade U-net has been reported to be better than a simple one-unit U-net. [11]

Parallel design U-net

U-nets arranged parallel to each other is another architecture of U-net network. The probable benefit of network with parallel decoders is attainment of diverse levels of separation. In this study organized 2 U-nets in parallel and accumulated outcomes for enhanced separation correctness. [12]

Chapter 3: Methodology

The aim of the project was the segmentation of liver from healthy abdominal CT images. The dataset was obtained from Combined (CT-MR) Healthy Abdominal Organ Segmentation (CHAOS) Challenge held 2019.

Dataset:

The dataset consisted of CT images from 40 patients. The patients were living liver transplantation donors. The data was divided into two parts i.e. 20 patient datasets were used in training while 20 were used as test sets. The dataset used was in the form of DICOM images for three-dimensional analysis and visualization. A contrast injection in portal venous phase was given to the patients to enhance the portal vein in the images. This type of injection is generally used prior to surgery to create contrast between vein and the liver. The patient orientation for all images in the dataset was same. Each data set consists of 16 bit DICOM images with a resolution of 512x512, x-y spacing between 0.7-0.8 mm and having 3 to 3.2 mm inter-slice distance. An average of 90 slices per data set ((i.e. minimum 77, maximum 105 slices). The dataset showed a varying anatomical size of organs among patients. Approximately 4.15% of the liver shape was atypical. The data showed varying Hounsfield ranges across different organs but the Hounsfield range was similar in adjacent organs.

Pre- Processing:

The 3d volumes were converted to 2d images. This allows us to use transfer learning with 2d segmentation models like U-net

Frameworks:

Pytorch:

PyTorch is an open-source machine learning library based on the Torch library, used for applications such as computer vision and natural language processing etc. Pytorch was used as It has faster training speeds, is more efficient, and uses lesser resources. Pytorch is more intuitive as well.

FastAl

FastAI is another library built on top of Pytorch that increases its potential even more. FastAI claims to train imagenet faster on GPU compared to tensorflow on TPU. It has a lot of built in models including U-net. Also its has helper functions for one-cycle-learning, finding optimal learning rate, etc

Training the Model:

15% of the images were used for validation. Transfer learning was used for contracting path of U-net. Resnet34 pretrained on Imagenet was used in contracting path. Initially the learning rate was set to $1e^{-3}$ for the first 85 epochs Later, the learning rate was changed. It was equally distributed in the range $1e^{-5}$ to $1e^{-4}$ for all layers in the network. This slicing helps achieve minima faster

One-cycle policy was used to train where the learning rate first gradually rises linearly and then falls back linearly as well. The rising learning rate also helps as a regularizer avoiding overfitting Apart from this, the weights were decayed (L2 regularizer) by a factor of $1e^{-2}$ in order to regularize the model

Liver Segmentation

Image segmentation is divided into two broad categories, Semantic segmentation and instance segmentation. Semantic segmentation classifies every pixel in the image into given classes. If there are multiple instances of the same class in the image, it does not differentiate between them. On the other hand, instance segmentation not only assigns a class to every pixel, but it also differentiates between different instances of the same class.

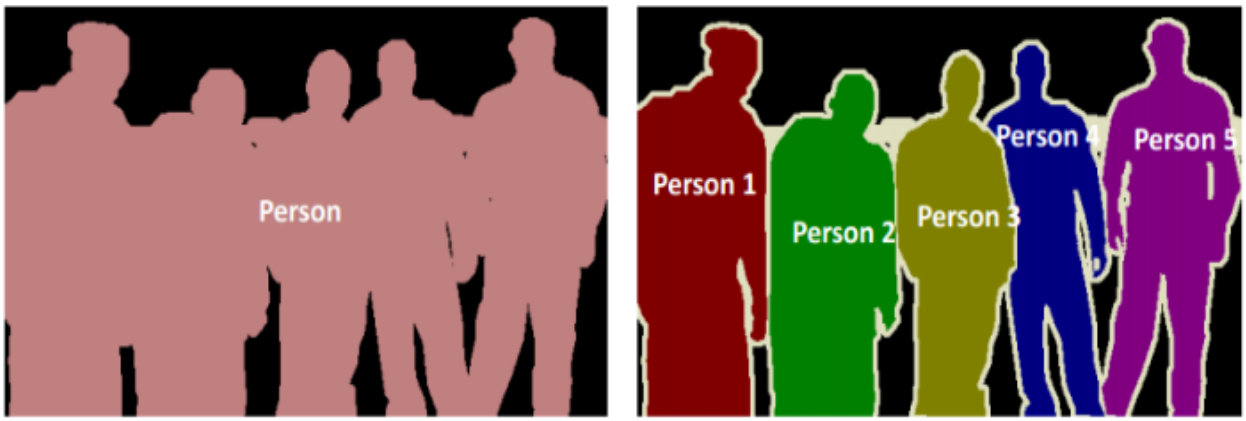

Figure 9: Semantic Segmentation (Left) vs, Instance Segmentation (Right)

In this research, the segmentation of liver in CT images is treated as a semantic segmentation task. This is done since there is only one liver in the image.

While there are multiple segmentation architectures available, one stands out in particular; U-Net [unet]. U-Net has demonstrated its superiority in semantic segmentation of medical datasets, especially on smaller datasets. For the purpose of segmentation of liver, we chose U-Net.

U-Net Architecture

U-Net is built upon the Fully Convolutional Network (FCN) architecture. It resembles an encoder-decoder network. A pictorial representation of U-Net is shared below,

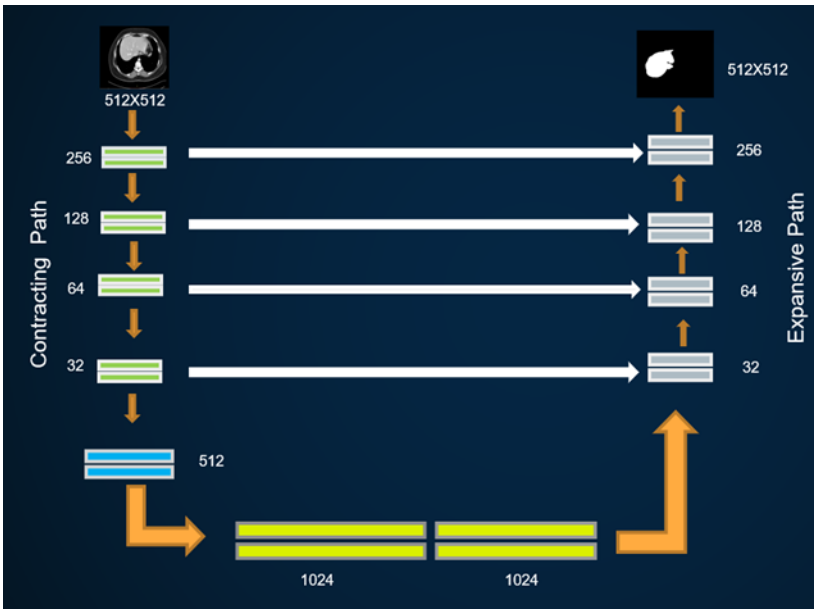

Figure 19: U-Net Architecture.

The network consists of two parts, an encoder (contracting) path and a decoder (expanding) path. In a vanilla U-net, the contracting path consists of convolutional layers with filters of size 3x3. Each convolution is followed by ReLU and 2x2 max pooling. This, in effect, reduces the spatial dimension by a factor of 2. At every successive convolutional layer, the number of filters is doubled while the spatial size gets halved. In the case of this study, instead of this vanilla approach, a residual network was used in the contracting path. The last fully connected layers of the network are stripped off and only convolutional layers remain.

The expanding path of U-Net up-samples the feature maps of the contracting layer. Number of channels in each successive layer of the expanding path are reduced by half while the spatial size is doubled by using “up-convolutions”. This is done as opposite to the contracting path. In the final layer, the spatial size is equal to that of the input and the number of channels is equal to the number of classes.

What makes a U-Net different from a simple encoder-decoder network are its skip connections. These connections concatenate the layers of contracting path to corresponding layers in the expanding path. This serves as a bridge between the two halves of the network and provides the expanding path with additional information that might have been lost in the *encoding* process.

Training U-Net on Liver CT Dataset

We divided the data into 80-20 ratio for training and validation sets respectively. We extracted axial slices to train U-Net. Transfer learning was employed in order to maximize the efficiency. The encoder we chose for our experiments was ResNet34. The network was trained using FastAI framework that is built upon PyTorch.

Cyclic learning rate, or one cycle fitting was used to fit the network on the data. As stated in [cyclic lr reference], a network can minimize the loss more efficiently if the learning rate for one epoch increases for a certain number of iterations and then reaches its initial value again at the end. In our case, we increased the learning rate linearly for 80% of the iterations in an epoch and then decreased back linearly to its initial value in the next 20% of the iterations.

Another novel technique that we used for training the data was the use of layer specific learning rate as described in (Howard *et al.*)[21]. It also helps increase the speed and efficiency of learning/

training. The intuition behind the idea is that since the layers closer to the input have similar, raw features for different datasets (e.g. line detection, color identification, etc.) usually, so they do not need to learn at the same rate as the layers closer to the output end where dataset specific features are learned. FastAI provides a built-in feature for this discriminative learning rate. It divides the layer into three categories: closer to input, closer to output, middle. The learning rate is adjusted on the basis of these groups. First group i.e. closer to input gets the least learning rate, the group of layers close to the output gets the largest learning rate while the middle layers get mean of the two values as the learning rate.

We chose ADAM as the optimizer and categorical cross entropy as our loss function. A summary of the hyper parameters we used is given in the table below.

Hyper parameter	**Value**
Learning rate	$[1e^{-5} – 1e^{-3}]$*
Total number of epochs	100
Batch size	8
Weight decay	0.01

Table 1: Hyperparameters used in the study

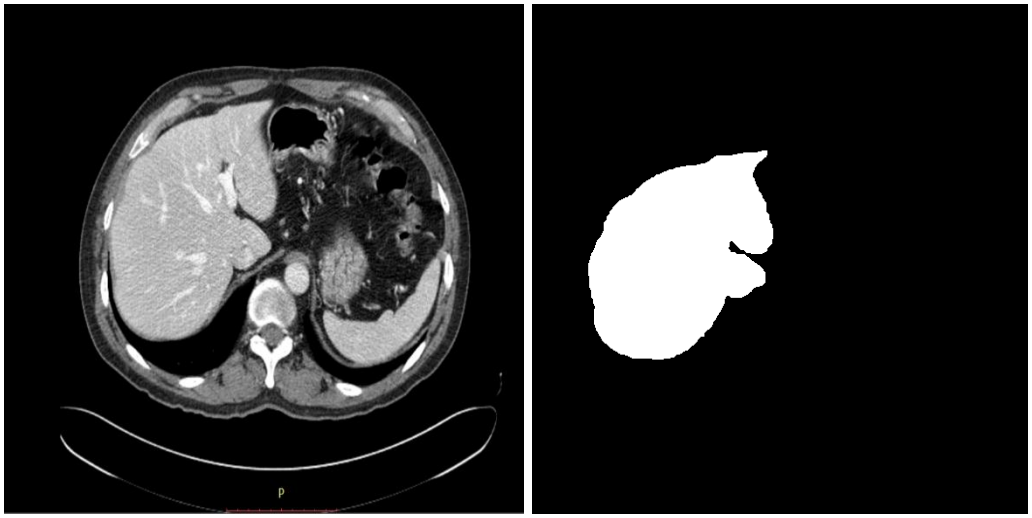

Figure 20: CT Image (Left) , Mask (RIGHT).

The experiments were performed using Amazon Web services (AWS) EC2 instances having 4 nvidia V100 GPUs. After training the network on 2D slices, we created volumes of the original size for evaluation.

Chapter 5: Results

For evaluation of the trained model, Grand Challenge provided a dataset which did not have annotations available for the public. Evaluation was also performed by them with their own scripts. The evaluation took in their own servers. The following metrics were used to evaluate the segmentation results.

- **Sørensen–Dice coefficient:** Provides information about the overlapping parts of segmented and ground truth volumes. (1 for a perfect segmentation, 0 for the worst case). As F1 is to classification tasks, Dice coefficient is to segmentation.
- **Relative absolute volume difference (RAVD):** Also provides information about the differences between volumes between segmented and reference organs but values the differences more than overlap. (0% for a perfect segmentation, 100% for the worst case).
- **Average symmetric surface distance (ASSD):** Determines the average difference between the surface of the segmented object and the reference in 3D. After the border voxels of segmentation and reference are determined, those voxels that have at least one neighbor from a predefined neighborhood that does not belong to the object are collected. For each collected voxel, the closest voxel in the other set is determined and the average of all these distances gives ASSD (0 mm for a perfect segmentation, max distance of image for the worst case).
- **Maximum symmetric surface distance (MSSD):** Similar to ASSD but particularly important for surgical operations (such as living donated liver surgery of database 1) as it determines the maximum margin of error by selecting the biggest of all

calculated distances (0 mm for a perfect segmentation, max distance of image for the worst case).

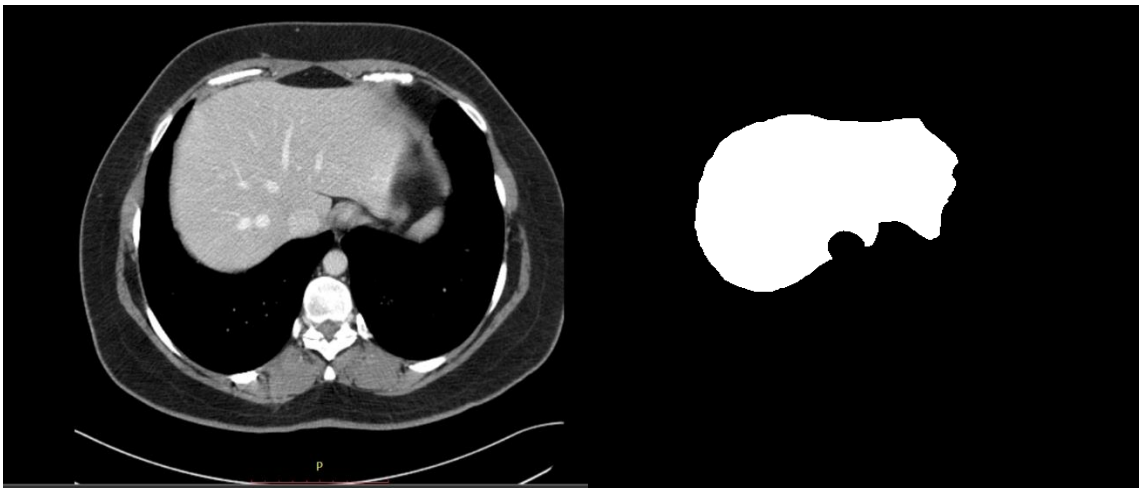

Figure 21: CT Image (Left), Mask (Right).

Our segmentation yielded the following results.

Max	105.40573127739147
Min	0.8251342581711172
Standard deviation	23.132183306406866
25th percentile	1.1084042834454817
50th percentile	2.6647574425258775
75th percentile	6.131154762678959
Mean	9.595963397422938

(Aggregate ASSD evaluation)

Max	0.9786450878383022
Min	0.8489264725346445
Standard deviation	0.029623613189434895
25th percentile	0.9657063860810896
50th percentile	0.9725385305523866
75th percentile	0.9767595847052464
Mean	0.9634216424954047

(Aggregate Dice evaluation)

Max	366.99484347331094
Min	22.158965894469723
Standard deviation	68.05828436150689
25th percentile	107.18110342391869
50th percentile	124.77399064352738
75th percentile	150.2839863308445
Mean	131.97704591814832

(Aggregate MSSD evaluation)

Max	23.63246812785093
Min	0.12859691564256306
Standard deviation	5.2678102586431095
25th percentile	1.0132968845537604
50th percentile	1.3439194354473576
75th percentile	2.200900807395904
Mean	2.926605705805975

(Aggregate RAVD evaluation)

Chapter 06: Plagiarism Report

final thesis report

ORIGINALITY REPORT

12%	9%	7%	3%
SIMILARITY INDEX	INTERNET SOURCES	PUBLICATIONS	STUDENT PAPERS

PRIMARY SOURCES

1	chaos.grand-challenge.org Internet Source	4%
2	medium.com Internet Source	1%
3	www.i-scholar.in Internet Source	1%
4	"ECAI 2020", IOS Press, 2020 Publication	1%
5	Shervin Minaee, Yuri Y. Boykov, Fatih Porikli, Antonio J Plaza, Nasser Kehtarnavaz, Demetri Terzopoulos. "Image Segmentation Using Deep Learning: A Survey", IEEE Transactions on Pattern Analysis and Machine Intelligence, 2021 Publication	1%
6	towardsdatascience.com Internet Source	1%
7	www.slideshare.net Internet Source	<1%

8	en.wikipedia.org Internet Source	<1%
9	"Medical Image Computing and Computer Assisted Intervention – MICCAI 2020", Springer Science and Business Media LLC, 2020 Publication	<1%
10	"Open Semantic Technologies for Intelligent System", Springer Science and Business Media LLC, 2020 Publication	<1%
11	Sara Baber Sial, Muhammad Baber Sial, Yasar Ayaz, Syed Irtiza Ali Shah, Aleksandar Zivanovic. "Interaction of robot with humans by communicating simulated emotional states through expressive movements", Intelligent Service Robotics, 2016 Publication	<1%
12	Chen Li, Yusong Tan, Wei Chen, Xin Luo, Yulin He, Yuanming Gao, Fei Li. "ANU-Net: Attention-based Nested U-Net to exploit full resolution features for medical image segmentation", Computers & Graphics, 2020 Publication	<1%
13	"Medical Image Computing and Computer-Assisted Intervention – MICCAI 2016",	<1%

	Springer Science and Business Media LLC, 2016 Publication	
14	Lei Chen, Hong Song, Qiang Li, Yutao Cui, Jian Yang, Xiaohua Tony Hu. "Liver Segmentation in CT Images Using a Non-Local Fully Convolutional Neural Network", 2019 IEEE International Conference on Bioinformatics and Biomedicine (BIBM), 2019 Publication	<1 %
15	Submitted to The Scientific & Technological Research Council of Turkey (TUBITAK) Student Paper	<1 %
16	journals.sagepub.com Internet Source	<1 %
17	www.ideals.illinois.edu Internet Source	<1 %
18	Submitted to Yonsei University Student Paper	<1 %
19	Tawsifur Rahman, Amith Khandakar, Muhammad Abdul Kadir, Khandaker R. Islam et al. "Reliable Tuberculosis Detection using Chest X-ray with Deep Learning, Segmentation and Visualization", IEEE Access, 2020 Publication	<1 %

20	Shuo Jin, Bo Wang, Haibo Xu, Chuan Luo et al. "AI-assisted CT imaging analysis for COVID-19 screening: Building and deploying a medical AI system in four weeks", Cold Spring Harbor Laboratory, 2020 Publication	<1 %
21	"Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support", Springer Science and Business Media LLC, 2017 Publication	<1 %
22	"Image Analysis and Processing — ICIAP 2015", Springer Science and Business Media LLC, 2015 Publication	<1 %
23	Man Tan, Fa Wu, Dexing Kong, Xiongwei Mao. "Automatic liver segmentation using 3D convolutional neural networks with a hybrid loss function", Medical Physics, 2021 Publication	<1 %
24	Vladislav Panferov, Dmitry Tailakov, Alexander Donets. "Recognition of Rocks Lithology on the Images of Core Samples", 2020 Science and Artificial Intelligence conference (S.A.I.ence), 2020 Publication	<1 %
25	fbr.springeropen.com Internet Source	<1 %

26	"Medical Image Computing and Computer Assisted Intervention – MICCAI 2019", Springer Science and Business Media LLC, 2019 Publication	<1%
27	Ahmed M. Mharib, Abdul Rahman Ramli, Syamsiah Mashohor, Rozi Binti Mahmood. "Survey on liver CT image segmentation methods", Artificial Intelligence Review, 2011 Publication	<1%

Exclude quotes	Off	Exclude matches	Off
Exclude bibliography	Off		

References:

1. Minaee, S., Boykov, Y., Porikli, F., Plaza, A., Kehtarnavaz, N., & Terzopoulos, D. (2020). Image segmentation using deep learning: A survey. arXiv preprint arXiv:2001.05566.

2. Du, G., Cao, X., Liang, J., Chen, X., & Zhan, Y. (2020). Medical image segmentation based on u-net: A review. Journal of Imaging Science and Technology, 64(2), 20508-1.

3. He, K., Gkioxari, G., Dollár, P., & Girshick, R. (2017). Mask r-cnn. In *Proceedings of the IEEE international conference on computer vision* (pp. 2961-2969).

4. Liu, S., Qi, L., Qin, H., Shi, J., & Jia, J. (2018). Path aggregation network for instance segmentation. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 8759-8768).

5. Dai, J., He, K., & Sun, J. (2016). Instance-aware semantic segmentation via multi-task network cascades. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 3150-3158).

6. Chen, L. C., Hermans, A., Papandreou, G., Schroff, F., Wang, P., & Adam, H. (2018). Masklab: Instance segmentation by refining object detection with semantic and direction features. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* (pp. 4013-4022).

7. Chen, X., Girshick, R., He, K., & Dollár, P. (2019). Tensormask: A foundation for dense object segmentation. In *Proceedings of the IEEE/CVF International Conference on Computer Vision* (pp. 2061-2069).

8. Long, J., Shelhamer, E., & Darrell, T. (2015). Fully convolutional networks for semantic segmentation. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 3431-3440).

9. Liu, W., Rabinovich, A., & Berg, A. C. (2015). Parsenet: Looking wider to see better. *arXiv preprint arXiv:1506.04579*.

10. Wang, G., Li, W., Ourselin, S., & Vercauteren, T. (2017, September). Automatic brain tumor segmentation using cascaded anisotropic convolutional neural networks. In *International MICCAI brainlesion workshop* (pp. 178-190). Springer, Cham

11. Yuan, Y., Chao, M., & Lo, Y. C. (2017). Automatic skin lesion segmentation using deep fully convolutional networks with jaccard distance. *IEEE transactions on medical imaging*, *36*(9), 1876-1886.

12. Chen, L. C., Papandreou, G., Kokkinos, I., Murphy, K., & Yuille, A. L. (2014). Semantic image segmentation with deep convolutional nets and fully connected crfs. *arXiv preprint arXiv:1412.7062*.

13. Lin, G., Shen, C., Van Den Hengel, A., & Reid, I. (2016). Efficient piecewise training of deep structured models for semantic segmentation. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 3194-3203).

14. Z. Liu, X. Li, P. Luo, C.-C. Loy, and X. Tang, "Semantic image segmentation via deep parsing network," in Proceedings of the IEEE international conference on computer vision, 2015, pp. 1377–1385.

15. Milletari, F., Navab, N., & 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.

16. Commandeur, F., Goeller, M., Betancur, J., Cadet, S., Doris, M., Chen, X., ... & Dey, D. (2018). Deep learning for quantification of epicardial and thoracic adipose tissue from non-contrast CT. *IEEE transactions on medical imaging*, *37*(8), 1835-1846.

17. Hesamian, M. H., Jia, W., He, X., & Kennedy, P. (2019). Deep learning techniques for medical image segmentation: achievements and challenges. *Journal of digital imaging*, *32*(4), 582-596.

18. Gibson, E., Giganti, F., Hu, Y., Bonmati, E., Bandula, S., Gurusamy, K., ... & Barratt, D. C. (2017, September). Towards image-guided pancreas and biliary endoscopy: Automatic multi-organ segmentation on abdominal CT with dense dilated networks. In *International Conference on Medical Image Computing and Computer-Assisted Intervention* (pp. 728-736). Springer, Cham.

19. Zhou, X. Y., Riga, C., Lee, S. L., & Yang, G. Z. (2018, October). Towards automatic 3D shape instantiation for deployed stent grafts: 2D multiple-class and class-imbalance marker segmentation with equally-weighted focal U-Net. In *2018 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)* (pp. 1261-1267). IEEE.

20. Zeng, G., & Zheng, G. (2018, April). Multi-stream 3D FCN with multi-scale deep supervision for multi-modality isointense infant brain MR image segmentation. In *2018 IEEE 15th International Symposium on Biomedical Imaging (ISBI 2018)* (pp. 136-140). IEEE.

21. Howard, J., & Ruder, S. (2018). Universal language model fine-tuning for text classification. *arXiv preprint arXiv:1801.06146*.

22. O. F. P. &. B. T. Ronneberger, "U-net: Convolutional networks for biomedical image segmentation. In International Conference on Medical image computing and computer-assisted intervention," Springer, pp. 234-241, 2015.

23. Ö. A. A. L. S. S. B. T. &. R. O. Çiçek, "3D U-Net: learning dense volumetric segmentation from sparse annotation.," *Springer,* pp. 424-432, 2016.

24. G. Y. X. L. J. Y. L. H. P. A. &. Z. G. Zeng, "3D U-net with multi-level deep supervision: fully automatic segmentation of proximal femur in 3D MR images.," *Springer,* pp. 274-282, 2017.

25. J. O. O. S. M. H. M. K. B. G. B. &. R. D. Schlemper, " Attention gated networks: Learning to leverage salient regions in medical images.," *Medical image analysis,* vol. 53, pp. 197-207., 2019.

26. C. L. W. J. Y. S. P. R. S. A. D. .. &. R. A. Szegedy, “Going deeper with convolutions.,” *In Proceedings of the IEEE conference on computer vision and pattern recognition,* pp. 1-9, 2015.

27. K. Z. X. R. S. &. S. J. He, “Deep residual learning for image recognition.,” *In Proceedings of the IEEE conference on computer vision and pattern recognition,* pp. 770-778, 2016.

28. M. Liang and X. Hu, “Recurrent convolutional neural network for object recognition,” p. 3367–3375, 2015.

29. S. T. Y. L. H. Z. H. W. Y. &. C. G. Cai, “ Dense-UNet: a novel multiphoton in vivo cellular image segmentation model based on a convolutional neural network.,” *Quantitative imaging in medicine and surgery,* vol. 10, no. 6, p. 1275, 2020.

30. Z. S. M. M. R. T. N. &. L. J. Zhou, “Unet++: A nested u-net architecture for medical image segmentation. In Deep learning in medical image analysis and multimodal learning for clinical decision support,” *Springer,* pp. 3-11, 2018.

31. E. S. B. &. K. A. Schonfeld, “A u-net based discriminator for generative adversarial network,” *In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition ,* pp. 8207-8216, 2020

32. X. W. C. C. S. &. G. L. Feng, “Automatic liver and tumor segmentation of CT based on cascaded U-net.,” *Springer,* pp. 155-164, 2019.

33. M. K. K. A. A. A. A. I. &. H. H. F. Abd-Ellah, “TPUAR-Net: Two parallel U-Net with asymmetric residual-based deep convolutional neural network for brain tumor segmentation.,” *Springer,* pp. 106-116, 2019.