INDUSTRIAL TRAINING REPORT II

TRAINING ORGANIZATION	**:**	**ZONE 24x7 (PVT) LTD**
PERIOD OF TRAINING	**:**	**FROM 01.08.2022 TO 10.10.2022**
FIELD OF SPECIALIZATION	**:**	**ELECTRICAL & ELECTRONIC ENGINEERING**

P.D.R. FERNANDO

E/16/103

ACKNOWLEDGEMENTS

I had my second industrial training at Zone 24x7 (Pvt) Ltd. During the 10 weeks of training I had the opportunity to gain knowledge on the work carried out by the organization and got a great exposure to the industrial work ethics. Training at an organization like this provide the platform to apply the knowledge which we gained at the undergrad studies at the university.

There are several supervisors I should be thanking Dr.Gerrard Fernando for guiding me through the research project I've been working on at Zone 24x7, Eng. Kanishka Wijayasekara for giving us valuable knowledge on the operations carried out by different engineers throughout a project and Eng. Dinushka Nanayakkara and Eng. Nisal Mihiranga for mentoring me throughout the training. At the same time, I would like to extend my heartfelt gratitude towards the all the staff members of Zone 24x7 for their support and guidance provide throughout the 10 weeks of training.

Eventually, I would like to thank the Industrial Training and Carrier Guidance Unit of the Faculty of Engineering for creating this opportunity for me to have a valuable industrial exposure and for the guidance they provided throughout the training.

CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

US	Ultrasound Scans
NN	Neural Network
CNN	Convolution Neural Network
PCNN	Pulse Coupled Neural Network
CT	Computerised Tomography
KT	Knowledge Transfer
GPU	Graphics Processing Unit
ML	Machine Learning
DL	Deep Learning
CPU	Central Processing Unit
SSH	Secure Shell
CLI	Command Line Interface
PCNN	Pulse Coupled Neural Network

Chapter 1

INTRODUCTION

1.1 INTRODUCTION TO TRAINING ORGANIZATION

Zone 24x7 headquartered in San Jose, California. Zone24x7 is a digital transformation partner to a wide range of organizations from technology startups to Fortune 500 organizations. Established in 2003 by founder Llavan Fernando and CFO Saw-Chin Fernando, the company has successfully served leading Tier 1 retailers and hi-tech companies offering best-in-class technology solutions backed by a strong commitment to customer satisfaction.

In 2004, Llavan and Saw-Chin founded Zone24x7 (Private) Limited, an advanced technology center in Colombo, Sri Lanka as a separate entity. The advanced technology center offers highly skilled engineering specialists and related technology services to Zone24x7 Inc.

Zone24x7 specializes in offering end-to-end technology consulting and engineering services encompassing both hardware and software. The services portfolio includes Enterprise Software Applications, Big Data & Data Science Engineering, Embedded Systems Engineering, Remote Monitoring & IOT, Machine Learning, Cognitive Vision, Robotics, and Innovation Services with Technology Proof of Concept Development. The best part about Zone is it practices a broader range of engineering principles.

At the time of the training the concept that was being followed was work from home due to Covid 19. During the training period Zone 24x7 had some physical fun events to ensure mental needs of their employees. How ever Zone had provided to use it's resources even at home. They also maintain a higher security in their servers to protect their intellectual property.

1.2 SUMMARY OF TRAINING EXPOSURE

During my training at Zone 24x7 I was to embedded engineering team. The title I was designated with was trainee associates electronic engineer. Normally Zone does client projects where they assign an engineer from Zone to the client. But on the other hand Zone 24x7 develops their own products too. I was assigned to a product that Zone develops. Since it is the beginning of the product development I was assigned to do a research. It had a larger team and even out from Zone 24x7. Since it is a medical device the team had a few doctors on board. Since it needs a lot of guidance in computer vision side the team also had academics from UOM. All the experiment were run by the Zone's internal team.

To share the progress of each member the external project participants and Zone's internal members hold a meeting at the end of every week. More-over to be updated about the progress there was a meeting once a week with the project supervisor and the architect. To guide the trainees and give them the momentum initially daily sync ups were taken. Because of this practice everyone was aware of what are other members are doing and it's a great way of learning new things.

One of the best practices I observed was maintaining a document with links to everyone's work. I was also able to contribute to the work practices by introducing google sites where I was given a opportunity to do a presentation on google sites to the team members. Here at Zone interns are treated as engineers and given a real world problem to solve where the intern is fully responsible for his work.

Chapter 2

BACKGROUND STUDY

2.1 PROJECT BACKGROUND

Liver volume is an important factor for a doctor to diagnose diseases in the liver area. Ultrasound is a routine physical examination method, it has many advantages. First of all, it uses ultrasonic to detect human body and forms images by the reflection intensity of ultrasonic. According to these images, the shapes and densities of various organs can be measured. Then, abnormal phenomena and disease locations may be found. Ultrasound, which is no harm to the human body, is a safe inspection method. In contrast to the inspections using a variety of X-rays, such as CT, need to strictly control their doses and have a great harm to the human body. Secondly, the price of ultrasound examination is cheaper than CT and most people can bear it. In addition, the detection accuracy of ultrasound in recent years has been greatly improved. How-ever, ultrasound also has its shortcomings, such as blurry and easy to be affected by noise. It is difficult for non-professional doctors to find the locations of the lesions and so on. So currently ultrasound is often used for physical examination and screening. In order to realize computer aided diagnosis in ultrasound image, there are many methods to deal with ultrasound image in recent years, such as denoising, object extraction and so on.

Based on all the reasons mentioned above Zone is working on a method to estimate liver volume using a very cheap portable US scan device. In this way the user doesn't have to be a professional and even rural hospitals with minimum funding, can afford this device. Not only that even for personal use this can be used.

The proposed approach was to scan the liver area with the portable US probe and then de-noise the scan since it contains lot of noise. Then using ML/DL the liver cross section is segmented. But for the US scan only a portion of liver is captured. Missing part should be extrapolated. Then by comparing the liver cross sections with actual 3D model of the liver we can obtain an estimate of liver volume. Further-more we can obtain a 3D model of the liver. Then Doctors can diagnose the patient without going through lot of trouble. This method has proven to produce 90% accurate estimations. Compared to the proposed method the traditional method which doctors use is only 70% accurate. This method is known as child's method where 3 measurements from liver images are taken and by considering these measurements as dimensions of a cuboid, the volume is estimated through multiplying these dimensions. The most accurate method is estimating using CT scans but unfortunately they are harmful for human body.

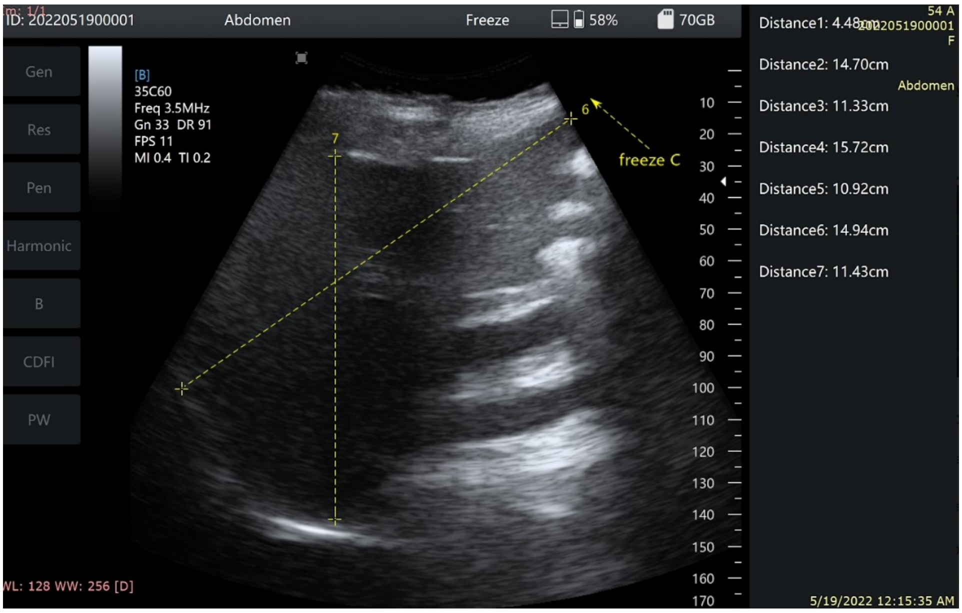
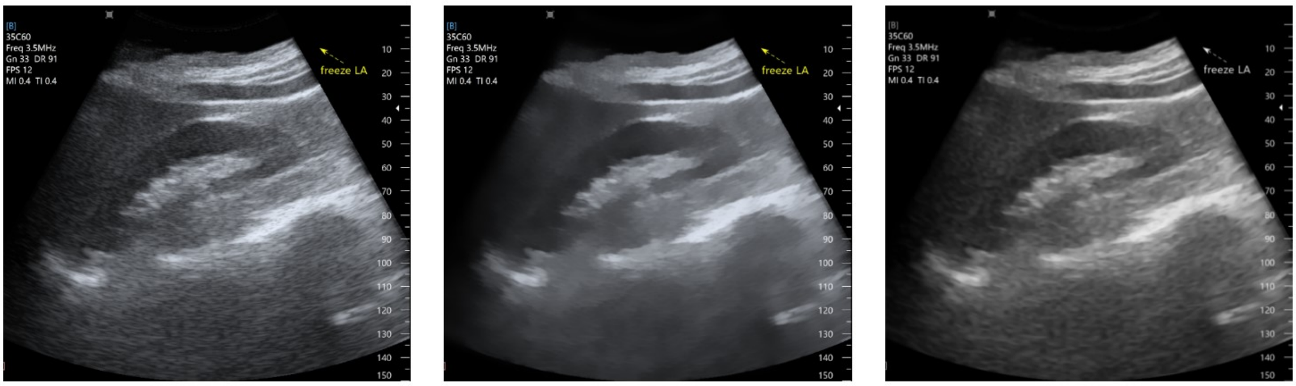


Figure 2.1 - US scan of a person with measurements which are used for child's method

2.2 DE-NOISING ULTRASOUND SCANS

The main issue in US scans is that they consist of speckle noise. Specially in the device it was proposed to use the speckle noise is even higher than common US devices. To find the contour of the liver a professional is needed. But by using ML we can automate the process. Before feeding the scans to a ML algorithm it was proposed it's best to have a clear image where speck noise is reduced and edges of the image are preserved. This part was already done by the previous intern. This is the point of the project where I took on. To denoise the image he has used several deterministic methods to smoothen the content and preserve the edges. Some of the results of these filters can be seen in the Figure 2.2. As we can see form the figure 2.2 non local means method produces the best results.



Original Non Local Means BM3D

Figure 2.2 - Denoised Results

Method	Edge preserving	Overall blur	Preserving/Enhancing Texture Differences	Comments
Non-Local means method				Execution time is high but results are consistent across all images.
Block matching 3D				Performed well for Images from normal probes but not for images from our probes.
Gaussian filter				
Mean Filter				Properties depend on parameters
Median Filter				Properties depend on parameters
Bilateral Filter				
Total variational Filter				
Histogram Equalization				Doesn't have a huge effect on noise
Anisotropic diffusion				
Wavelet filter				

Note - But these could be changed with different parameters.

High	
Medium	
low	

Table 2.2 - Denoised Result Summary

Chapter 3

NEURAL NETWORK TRAINING PROCCESS

3.1 PYTORCH

The task assigned to me was to develop neural network to achieve segmentation task. PyTorch is a machine learning framework based on the Torch library, used for applications such as computer vision and natural language processing, originally developed by Meta AI and now part of the Linux Foundation umbrella. It is free and open-source software released under the modified BSD license. Although the Python interface is more polished and the primary focus of development, PyTorch also has a C++ interface. Since the resources for Python ML are more commonly found and the other part of the project such as web development, denoising are being done using python I used python as the interface for Pytorch. PyTorch provides two high-level features

- Tensor computing (like NumPy) with strong acceleration via graphics processing units (GPU)
- Deep neural networks built on a tape-based automatic differentiation system

3.2 USING COLAB

But GPU capability is low in most of the computers we can get our hands on. Because of that the google service known as colab is used. Colaboratory, or "Colab" for short, is a product from Google Research. Colab allows anybody to write and execute arbitrary python code through the browser, and is especially well suited to machine learning, data analysis and education. More technically, Colab is a hosted Jupyter notebook service that requires no setup to use, while providing access free of charge to computing resources including GPUs. But the GPU is limited for 6 hours per day. Colab notebooks are stored in Google Drive, or can be loaded from GitHub. Colab notebooks can be shared just as you would with Google Docs or Sheets. Code is executed in a virtual machine private to your account. Virtual machines are deleted when idle for a while, and have a maximum lifetime enforced by the Colab service.

3.3 ZONE SERVER

Since the time limitation in google colab training NNs for a longer period of time is not preferred. Another issue is that we can't load bigger datasets to GPU memory and CPU memory hence it becomes limited in the training process. When we're using Colab. One way to overcome loading big data issue is by loading the data to the CPU and only load small data chunks of a defined batch size in the training loop to the GPU. Bbut if the dataset is too big loading it to the CPU is also an issue because the user is only given 12 GB of CPU memory in Colab.

As a solution I was given the chance to use much powerful Zone 24x7 virtual machine. To build the connection SSH can be used. The SSH protocol (also referred to as Secure Shell) is a method for secure remote login from one computer to another. It provides several alternative options for strong authentication, and it protects the communications security and integrity with strong encryption. It is a secure alternative to the non-protected login protocols and insecure file transfer methods.

At the time I was given access to the Zone's virtual machine I was running all the codes on Colab. So transferring files from my google drive to virtual machine was a trouble because I had over 25GB data to be transmitted from google drive to virtual machine. Since internet is limited if I did it the common way I have to sacrifice 50 GB + data for this simple work alone. Hence I used a python library known as "gdown" to directly download drive data to the virtual machine. In that way the data spent on this task is minimized.

Then by running the python script in the virtual machine I was able to train NNs and continue my work smoothly. In fact it enabled me to multi task by running 2 NNs at once. One in Colab and one in Zone's virtual machine.

3.4 NEPTUNE AI

A huge problem we met when we were training NNs was tracking experiments. Initially I used google sites to log my data by manually typing them. Then we found that to come to conclusion using the logged data is hard. And also it take a lot of time. Neptune ai is a MLops tool which lets the user record metadata regarding the NNs. Tools for experiment tracking give you a huge boost for your machine learning projects. You can use them to track hyperparameters, visualize graphs (e.g. metrics) and images (confusion matrix or misclassified samples), store artifacts (model weights) and much more.

Using this is very simple and straight forward. They provide an access token for each project and an id. Once we embed them in the code python will create the bridge between Neptune ai and colab and then the parameters that are need to be saved will be logged accordingly.

3.5 TRANSFER LEARNING

Transfer learning (TL) is a research problem in machine learning (ML) that focuses on storing knowledge gained while solving one problem and applying it to a different but related problem. As I mentioned earlier, Google Colab only give few hours to train NNs using GPU. Hence transfer learning was implemented. Here we train the NN for number of epochs and then the trained model is saved. Then it is again loaded when the GPU is available and continue to train from the last epochs.

Chapter 4

DATASET

4.1 CT DATASET

The US scan dataset is not available in the internet. The dataset which is captured from the handheld probe is very unique and normal US scans cannot resemble the application of our project. Hence a dataset is required to do the training. To prepare the dataset doctors from different hospitals were hired. But during my time at Zone getting my hands on US dataset was not possible. But once the US dataset is available Zone needed an working algorithm to train the neural networks to segment US data. Hence I found a CT scan dataset from internet and I trained the NN according to that. CT scans are very clear and can be easily segmented using NNs. This dataset contained the Ct scan of the liver area and as the ground truth the liver pixels are coloured. Figure 4.2 is a CT scan from the dataset used.

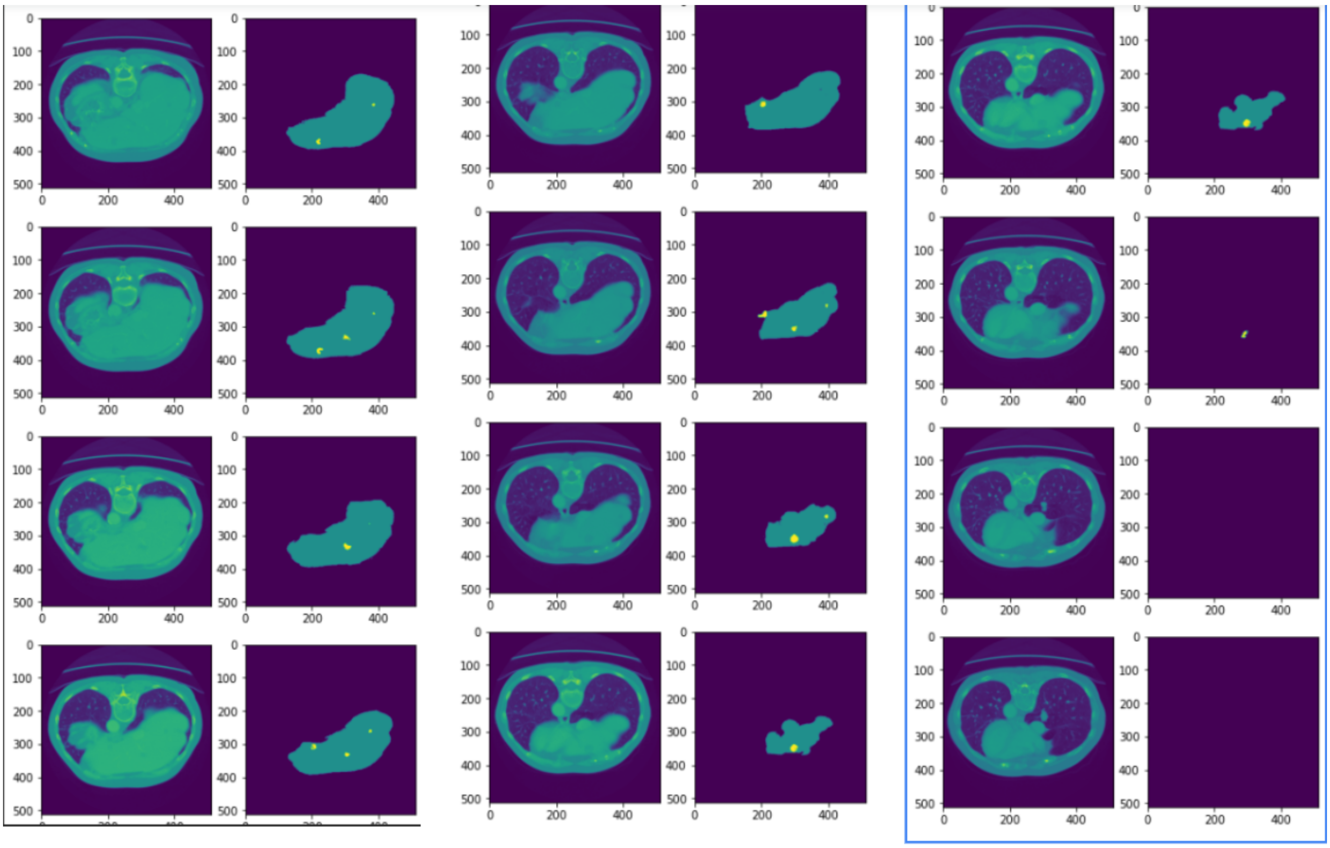


Figure 4.1 – CT scans of a single person

As we can see this a set of scans of a single person and it contain cross section through out the whole liver. Most of the scans did not have a liver. Hence the dataset takes too much space on images that doesn't contribute. To get rid of some of the empty scans, a python script was coded to check the pixel sum of a set of images and iteratively create a dataset where only sum greater than 0 was added to the dataset. This way it was able to reduce the size of the dataset by 30%. Having an optimum and minimalist dataset is important because GPU memory is limited and it affects the performance of the NN.

4.2 US DATASET

US scans are in a format known as Dicom which is native to medical imaging. It doesn't only contain image data but metadata of the patient as well. To create our dataset 3 cross-sections from liver is taken and doctors create the ground truth by coloring the liver area using the ITK snap software. ITK snap is a tool kit for medical annotations. Coloring the entire liver area was time consuming. Hence it was proposed to the doctors to just draw the outline and use a fill algorithm to fill the enclosed area. But at the moment we couldn't find a filling option in ITK snap software. Hence we developed our own filling algorithm.

The US scan looks as the figure 4.2.1 . It contains unnecessary metadata. For an image processing algorithm this is problematic. So to only cut out the scan view I generated a map (figure 4.2.2) manually and then multiplied it with the original scan. The final result can be seen in Figure 4.4.3.

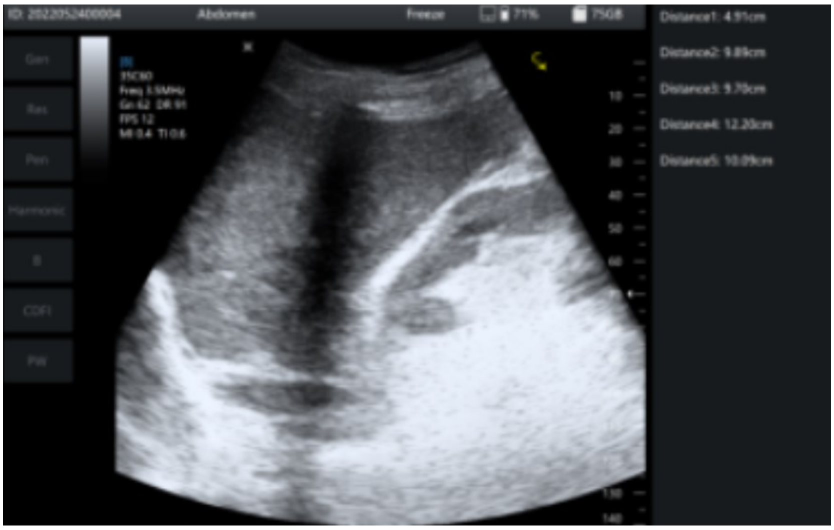


Figure 4.2.1 – US scan of a single person

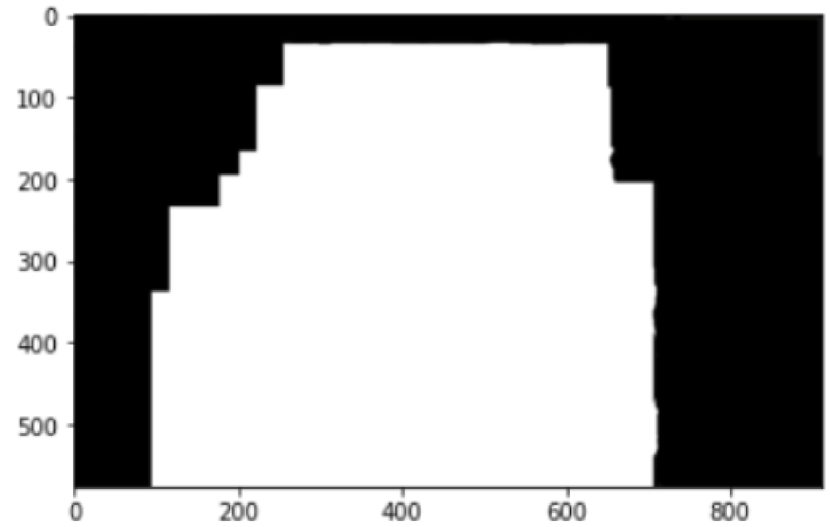


Figure 4.2.2 – Map to separate scan view

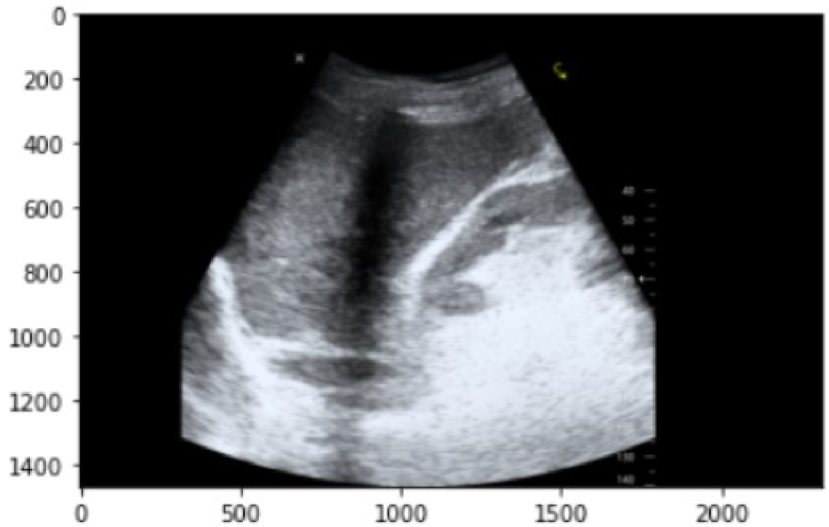


Figure 4.2.3 – US scan fed in to the NN

4.3 FILLING ALGORITHM

For annotation purposes a filling algorithm was required. So I created a simple python script to fill the image when an enclosed contour is given. If we consider a row or a column of an image with closed contour the row of column takes the form where starting from left/top move along the vector then all the pixel values are 0 until the contour pixel is met. Then again it will be 0 and then again a contour pixel is met. Then again it will be 0. If we make all the in between pixels of the contour pixels to 1 and we do this iteratively to all rows and columns we will end up with a filled shape. For more accuracy we separately do for row and column and then take the intersection of both to get rid of the anomalies caused by convex/ concave shapes of the contour.

Input

Filled with the algorithm

Figure 4.3.1 – Fill Algorithm Results

4.4 S3 SERVER DATA ACQUISITION

The project has several doctors contributing from different regions. To gather all the patient data and annotated data Amazon s3 server is used. This server can be accessed through a service known as cyberduck. But after 30 days every data in it automatically deleted. Because of that there was a requirement of automatically downloading the data to Zone's internal server in a monthly basis. I was given the task of creating a linux script that runs automatically following a schedule. To download the data to the internal server cyber duck CLI was used. Then to automate the process linux chron-job facility was used.

Chapter 5

SEGMENTATION

5.1 LITERATURE REVIEW

To achieve segmentation there are many methods. Some of them are deterministic and others are learning based methods. Below is a summary of methods to segment.

1. Adaptive Segmentation Algorithm of Ultrasonic Image Based on Simplified PCNN

   - PCNN (Pulse Coupled Neural Network)
   - 2-D neural network.
   - Each neuron corresponds to a single pixel in the input image. (external stimulus)
   - Each neuron connects with its neighboring neurons.(local stimuli)
   - Both external and internal stimuli are combined with an internal activation system.
   - This stimuli is accumulated until a dynamic threshold is reached.Then it produce a pulse.
   - This neuron activation spreads like a wild fire and clusters similar areas in the image considering the time that these pulses occurred.
   - In order to be able to objectively describe the experimental results, use the segmented region homogeneity, contrast and shape measurement

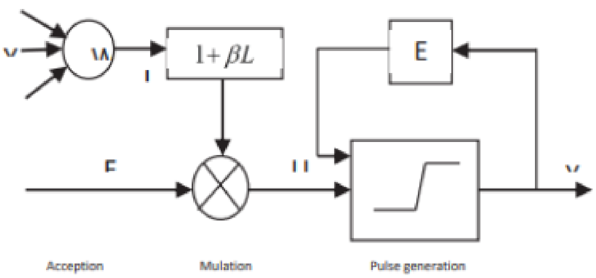


Figure 1 . Simplified model

Figure 5.1.1 – Simplified Model of PCNN

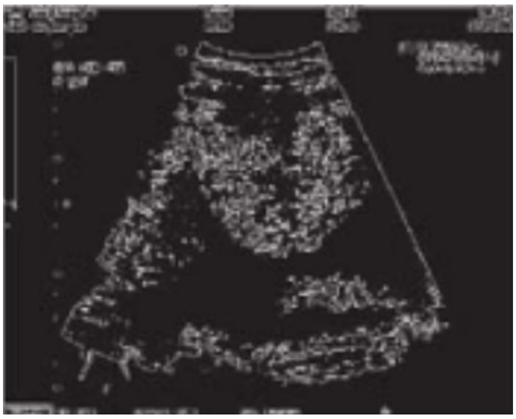


Figure 5.1.2 – Results from PCNN

2. Ultrasound Liver Image Enhancement Using Watershed Segmentation Method

   - Convert image to grayscale and enhance the image contrast2-D neural network.
   - Use the Gradient Magnitude as the Segmentation Function. The gradient magnitude is computed by applying the Sobel edge masks, filtering coefficients, and additional arithmetic. At the borders of the objects, the gradient is high but is low (mostly) inside the objects.
   - Marking the Foreground Objects.
   - Compute Background Markers.
   - Modify the Segmentation Function.
   - Compute the Watershed Transform of the Segmentation Function.
   - Watershed algorithm is based on extracting sure background and foreground and then using markers that will make watershed run and detect the exact boundaries. This algorithm generally helps in detecting touching and overlapping objects in image.

Figure 5.1.3 – Watershed Algorithm Results

3. A Deeply Supervised Network with Attention to Boundaries

- FCNN's were designed for natural images having completely different characteristics from US images. With several artifacts limiting the information content of US training data, deep networks like FCNN and its variants become hard to train. A possible solution to the problem is deep supervision which provides an optimal way of training the network to its potential .
- Deeply supervised nets (DSNs) are well suited for structured input-output applications like image segmentation .
- DSN uses auxiliary branches containing a small number of parameters to supervise the training of intermediate layers of the core network.
- These auxiliary layers try to predict the desired outcome from intermediate layer outputs. In general, the auxiliary layers use the same loss function as the final output layer of the network. Auxiliary predictions are fused at the end to get the final output.
- Manually annotated ground truths for the targeted lumen regions are provided by multiple observers.
- deep supervision of the network. The attention of side layers is steered towards the object boundary definitions whereas the other side layers are focused to identify the regions belonging to the desired object
- Fusion layer (H) is an important part of the proposed network. Instead of an arithmetic or an engineered sum , fusion layer weights are trained to result in a linear combination of the auxiliary outputs. These weights are learned along with other network parameters.

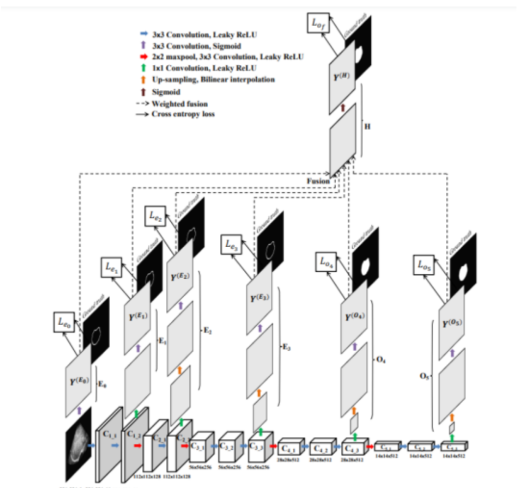


Figure 5.1.4 – A Deeply Supervised Network with Attention to Boundaries

4. IFCM Based Segmentation Method for Liver Ultrasound Images

- Initially A1 matrix contains [Cmin,Cmax] (Cmin/max - minimum/maximum pixel intensity value)
- Run FCM (algo like K-means) to find centeroids C1,C2
- Add it to A1 if maximum difference between any two centroids from A1 > thresh
- A1 = [Cmin , C1 : C1,C2:C2,Cmax]
- if maximum difference between any two centroids from A1 < threshold then it is added to A2.
- Do that until array A1 is empty
- Number of rows in A2 gives number of clusters obtained
- But we do all of this around a known pixel.

5. Graph Cut Liver Segmentation for Interstitial Ultrasound Therapy

- First, a network flow graph is built based on the input image. Then a max-flow algorithm is run on the graph in order to find the min-cut, which produces the optimal segmentation.
- Opposite of dissimilarity (distance between two pixel color intensity values) is affinity.
- From affinity we can obtain the graph.
- We dissect the graph such that the cost of cuts are minimized.
- the user interactively selects a set of voxels so that he considers belonging to the class ”object”, and others , that he considers belonging to the class ”background”. vo and vb have a double role: they will be used as seed points for the segmentation but also their value distribution (histograms) will serve as training basis for the class if related initial weight assignment

6. Segmentation of ultrasound images of liver tumors applying snake algorithms and GVF

- This algorithm we have to define a spline in the area we want to segment. (A set of points)

- Then iteratively it'll deform the spline. (move the points in a way that the energy function is minimized)
- Energy function can be divided in to 2. Internal and external
- Internal energy function deals with the smoothness of the shape, predefined shapes...
- External energy function relates with the edges of the image
- A good fit is when energy is minimized.

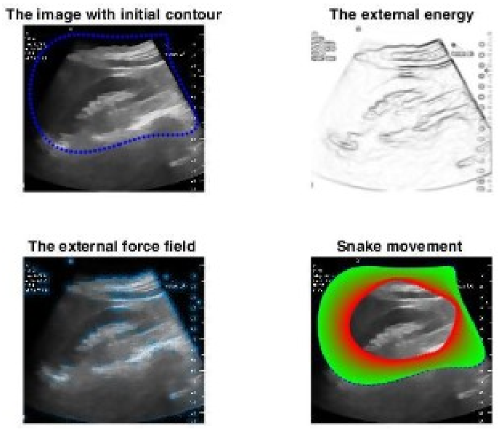


Figure 5.1.5 – GVF Results

7. Liver Ultrasound Image Segmentation Using Region-Difference Filters

- Region difference filters evaluate maximum difference of the average of two regions of the window around the center pixel.
- This image is then converted into binary image and morphologically operated for segmenting the desired lesion from the ultrasound image.
- Initially we select a pixel near the ROI
- Then we get a N x N filter. Then 4 region difference filters are applied to the selected region. RD filters are edge detecting filters oriented in 0,45,90,135 degrees.
- By obtaining the maximum difference given by two of the filter applied image we replace the pixel value with the maximum absolute difference.

$$RD1 = \begin{array}{|c|c|c|c|c|}\hline -1 & -1 & -1 & -1 & -1 \\ \hline -1 & -1 & -1 & -1 & -1 \\ \hline 0 & 0 & 0 & 0 & 0 \\ \hline 1 & 1 & 1 & 1 & 1 \\ \hline 1 & 1 & 1 & 1 & 1 \\ \hline \end{array} \quad RD2 = \begin{array}{|c|c|c|c|c|}\hline -1 & -1 & -1 & -1 & 0 \\ \hline -1 & -1 & -1 & 0 & 1 \\ \hline -1 & -1 & 0 & 1 & 1 \\ \hline -1 & 0 & 1 & 1 & 1 \\ \hline 0 & 1 & 1 & 1 & 1 \\ \hline \end{array}$$

$$RD3 = \begin{array}{|c|c|c|c|c|}\hline -1 & -1 & 0 & 1 & 1 \\ \hline -1 & -1 & 0 & 1 & 1 \\ \hline -1 & -1 & 0 & 1 & 1 \\ \hline -1 & -1 & 0 & 1 & 1 \\ \hline -1 & -1 & 0 & 1 & 1 \\ \hline \end{array} \quad RD4 = \begin{array}{|c|c|c|c|c|}\hline 0 & 1 & 1 & 1 & 1 \\ \hline -1 & 0 & 1 & 1 & 1 \\ \hline -1 & -1 & 0 & 1 & 1 \\ \hline -1 & -1 & -1 & 0 & 1 \\ \hline -1 & -1 & -1 & -1 & 0 \\ \hline \end{array}$$

ig. 1 Windows used for the implementation of the proposed segmentation method

Figure 5.1.6 – Edge Filters

8. Liver Tumor Localization Based on YOLOv3 and 3D-Semantic Segmentation Using Deep Neural Networks

- Major steps of the proposed approach are:
- • The synthetic liver CT images are created with a modified GAN model and fed into the localization part of the model.
- • After synthetic images generation, the YOLOv3-ResNet-50 model is designed for liver and liver tumor localization.
- • In the last step, a modified 3D-semantic segmentation model is presented, where DeepLabv3 serves as the base network for the Inceptionresnetv2.

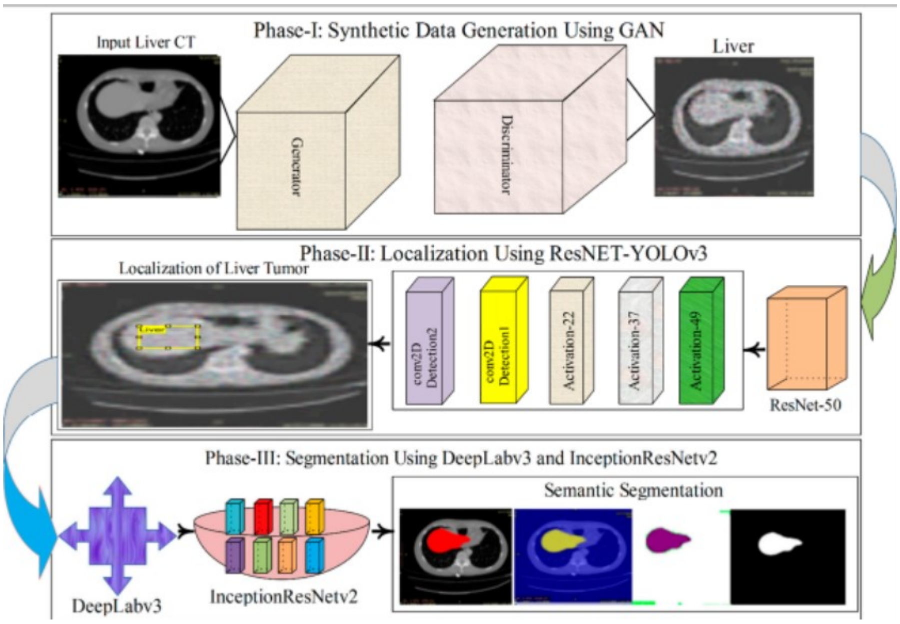


Figure 5.1.7 – DeepLab Architecture

9. Ultrasound prostate segmentation based on multidirectional deeply supervised V-Net

- Similar concept to deeply supervised networks but more like a u-net
- intermediate layers are trained to obtain the final stage quickly
- 3 planes are separated with 3 neural network
- results are incorporated to obtain the best segmentation in a chosen plane

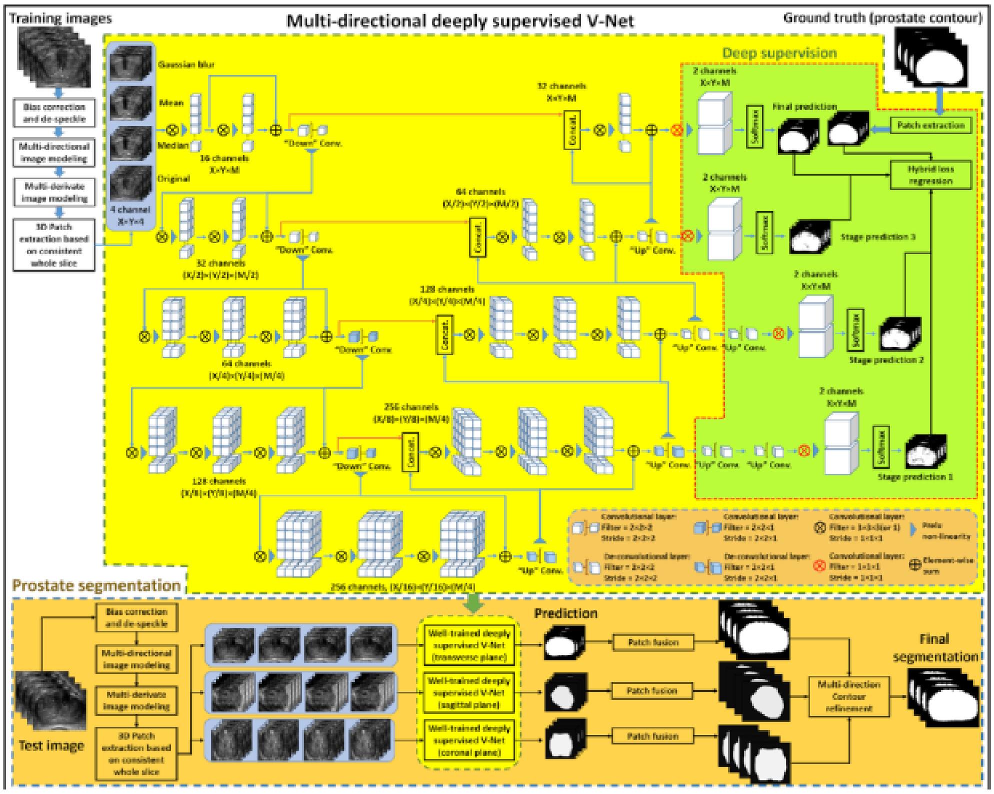


Figure 5.1.8 – V-net Architecture

10. U-Net: Convolutional Networks for Biomedical Image Segmentation

- The most common approach for segmenting today.

5.2 UNET

U-Net is a convolutional neural network that was developed for biomedical image segmentation at the Computer Science Department of the University of Freiburg. The network is based on the fully convolutional network and its architecture was modified and extended to work with fewer training images and to yield more precise segmentations. Segmentation of a 512 × 512 image takes less than a second on a modern GPU.

The network consists of a contracting path and an expansive path, which gives it the u-shaped architecture. The contracting path is a typical convolutional network that consists of repeated application of convolutions, each followed by a rectified linear unit (ReLU) and a max pooling operation. During the contraction, the spatial information is reduced while feature information is increased. The expansive pathway combines the feature and spatial information through a sequence of up-convolutions and concatenations with high-resolution features from the contracting path.

The main idea is to supplement a usual contracting network by successive layers, where pooling operations are replaced by up sampling operators. Hence these layers increase the resolution of the output. A successive convolutional layer can then learn to assemble a precise output based on this information.

One important modification in U-Net is that there are a large number of feature channels in the upsampling part, which allow the network to propagate context information to higher resolution layers. As a consequence, the expansive path is more or less symmetric to the contracting part, and yields a u-shaped architecture. The network only uses the valid part of each convolution without any fully connected layers. To predict the pixels in the border region of the image, the missing context is extrapolated by mirroring the input image. This tiling strategy is important to apply the network to large images, since otherwise the resolution would be limited by the GPU memory.

During the training process several experiments were done with different hyper parameters and varying dataset sizes to segment the liver from CT scans. Since CT scan dataset is very clear U-net performs fairly well even when it is trained using a small dataset. But in the actual case US scan dataset is not clear as CT datasets. Unfortunately I didn’t get the chance to train the NN for the US data. Figure 5.2.2 depicts training result of an experiment. Table 5.2 contains some hyper-parameters used in an experiment. Figure 5.2.3 shows the behaviour of the loss function with the iterations.

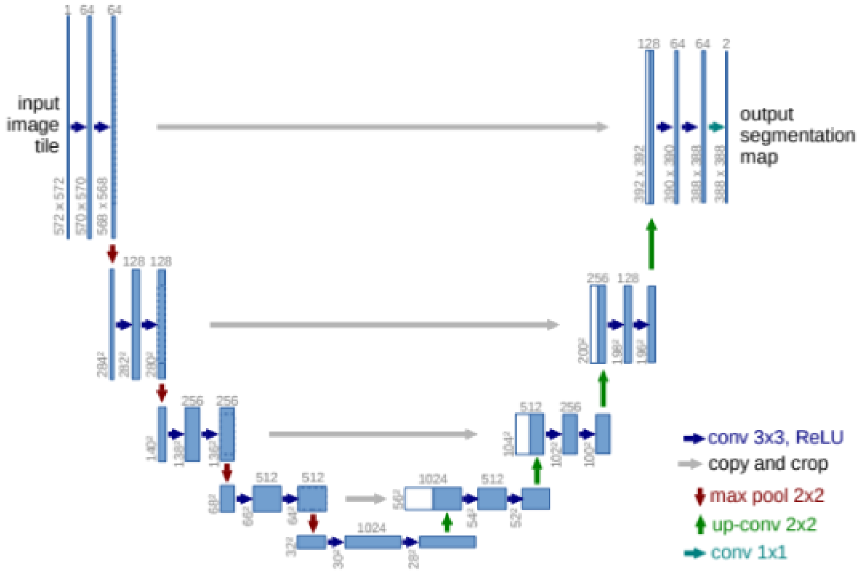


Figure 5.2.1 – U-net Architecture

• epochs	210
• Loss	MSE
• Optimizer	Adam
• Step size	0.001
• Sample Images	200
• Batch size	4

Table 5.2 – Hyper-parameters used in a training session

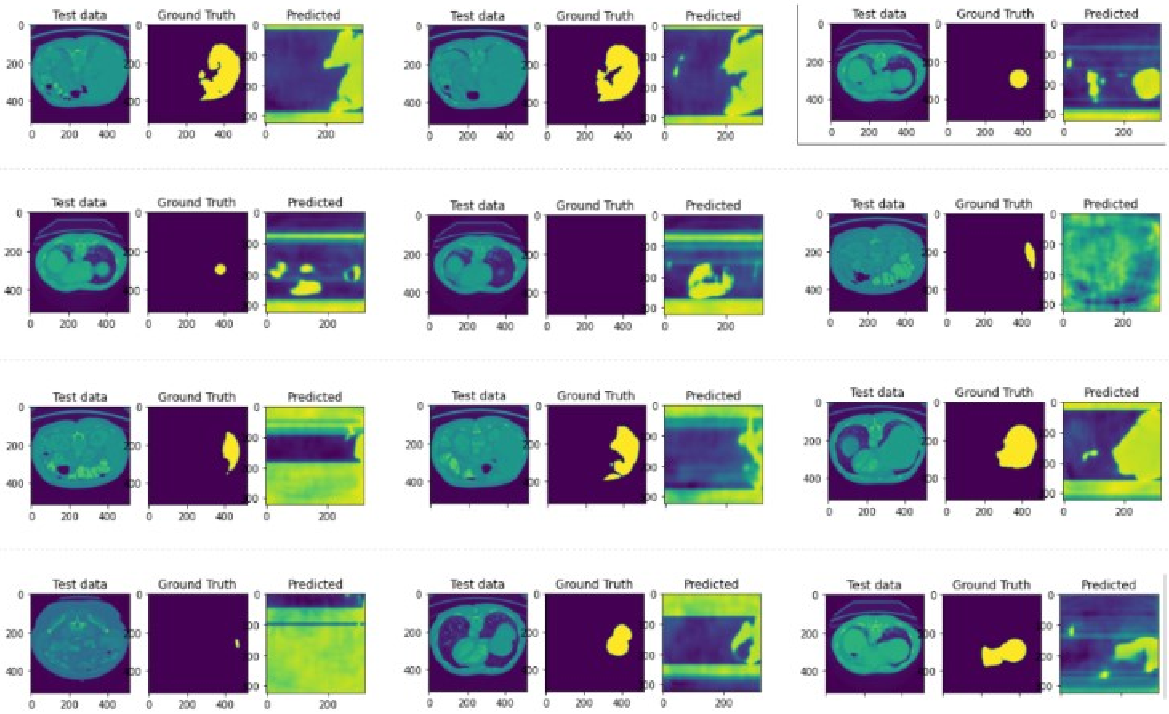


Figure 5.2.2 – U-net Results

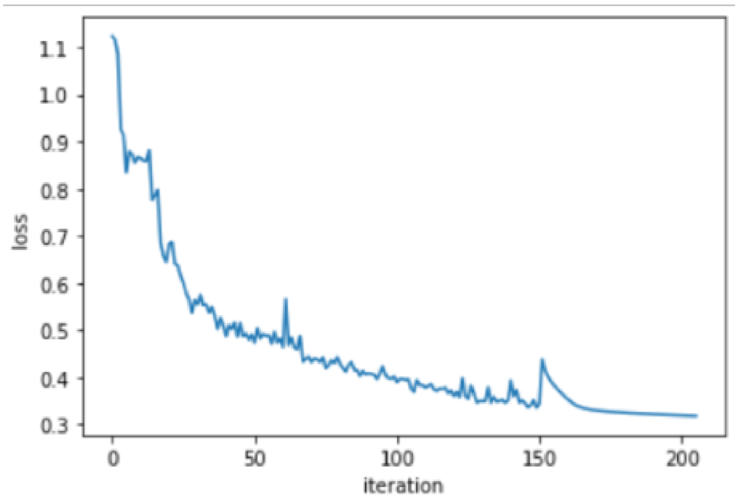


Figure 5.2.3 – Behaviour of the loss function

5.3 DEEPLABV3

For the task of semantic segmentation ,we consider two challenges in applying Deep Convolutional Neural Networks (DCNNs). The first one is the reduced feature resolution caused by consecutive pooling operations or convolution striding, which allows DCNNs to learn increasingly abstract feature representations. However, this invariance to local image transformation may impede dense prediction tasks, where detailed spatial information is desired. To overcome this problem, we advocate the use of atrous convolution , which has been shown to be effective for semantic image segmentation . Atrous convolution, also known as dilated convolution, allows us to repurpose ImageNet pretrained networks to extract denser feature maps by removing the downsampling operations from the last few layers and upsampling the corresponding filter kernels, equivalent to inserting holes between filter weights. With atrous convolution, one is able to control the resolution at which feature responses are computed within DCNNs without requiring learning extra parameters.

One of the main issues developing the deeplab , we faced was that NN is very heavy so only limited amount of data can be loaded into the GPU. Hence this was trained using very small amount of data. The results obtained can be seen in the Figure 5.3.2.

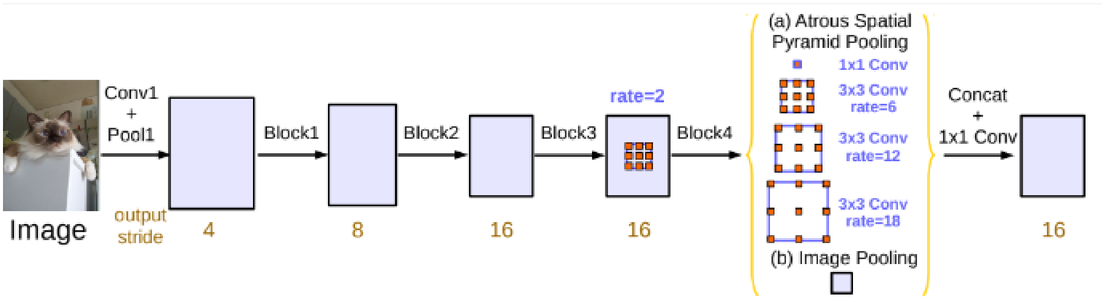


Figure 5.3.1 – Atrous Convolution

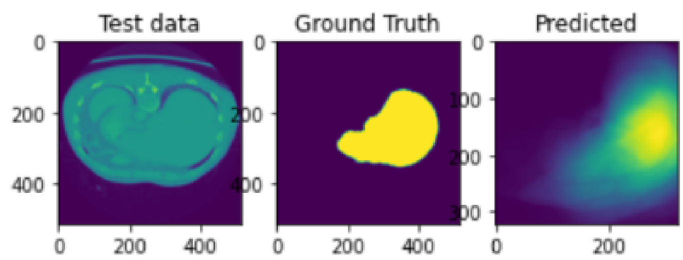


Figure 5.3.2 – Deeplab v3 Results

CONCLUSION

The industrial training session of ten weeks outside the lecture room was extremely valuable. Throughout the training it provided the platform to ap apply the theories learnt at the faculty in a practical environment.

Through the training I gained a better understanding on the organization and its functions. Not only that, I was able to get familiar with practices like problem-based learning, solution-oriented leaning etc. It was also a great opportunity to develop and improve interpersonal skills, problem solving skills, maturity, self-confidence and self-esteem etc. In addition I got the opportunity gain knowledge on stuff out of my field off specialization.

The knowledge gained through the courses at the university was a tremendous support in solving the problems and understanding certain subjects encountered in the training procedure. Finally, I would like to mention that the industrial training was extremely informative and fruitful.

After all Zone24x7 promotes a culture of customer-centric innovation, continuous learning, professionalism, caring for oneself and others, and integrity to create a diverse, inclusive, and thriving workplace for all its associates.