# Melanoma Detection Using Machine Learning

**Author**

**MUHAMMAD FAROOQ SAEED**

**00000204730**

**Supervisor**

**Dr. Syed Omer Gilani**

**DEPARTMENT OF BIOMEDICAL ENGINEERING AND SCIENCES SCHOOL OF MECHANICAL & MANUFACTURING ENGINEERING NATIONAL UNIVERSITY OF SCIENCES AND TECHNOLOGY**

**ISLAMABAD, PAKISTAN**

**June 2021**

# Melanoma Detection Using Machine Learning

Author

MUHAMMAD FAROOQ SAEED

00000204730

A thesis submitted in partial fulfillment of the requirements for the degree of

MS Biomedical Engineering

Thesis Supervisor:

Dr. Syed Omer Gilani

Thesis Supervisor's Signature: ________________________________

DEPARTMENT OF BIOMEDICAL ENGINEERING AND SCIENCES

SCHOOL OF MECHANICAL & MANUFACTURING ENGINEERING

NATIONAL UNIVERSITY OF SCIENCES AND TECHNOLOGY,

ISLAMABAD, PAKISTAN

JUNE 2021

# THESIS ACCEPTANCE CERTIFICATE

It is certified that the final copy of MS Thesis written by Mr. Muhammad Farooq Saeed (Registration No. 00000204730), of SMME (School of Mechanical and Manufacturing Engineering) has been vetted by undersigned, found complete in all respects as per NUST statutes / regulations, is free of plagiarism, errors and mistakes and is accepted as partial fulfillment for award of MS/MPhil Degree. It is further certified that necessary amendments as pointed out by GEC members of the scholar have also been incorporated in the thesis.

Signature: ______________________________

Name of Supervisor: Dr. Syed Omer Gilani

Date: ________________

Signature (HOD): ________________________

Date: _________________

Signature (Principal): _______________________

Date: __________________

**Nomi IDHAR page ATTACH KARNA HAII**

**Nomi IDHAR page ATTACH KARNA HAII**

**Nomi IDHAR page ATTACH KARNA HAII**

**Nomi IDHAR page ATTACH KARNA HAII**

**Nomi IDHAR page ATTACH KARNA HAII**

**Nomi IDHAR page ATTACH KARNA HAII**

# Declaration

I certify that this research work titled "*Melanoma Detection Using Machine Learning"* is my own work. The work has not been presented elsewhere for assessment. The material that has been used from other sources has been properly acknowledged / referred.

Signature of Student

Muhammad Farooq Saeed

00000204730

## Plagiarism Certificate (Turnitin Report)

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

Signature of Student

Muhammad Farooq Saeed

Registration Number

00000204730

Signature of Supervisor

## **Proposed Certificate for Plagiarism**

It is certified that MS Thesis Titled **Melanoma Detection Using Machine Learning** by **Muhammad Farooq Saeed** has been examined by us. We undertake the follows:

a. Thesis has significant new work/knowledge as compared already published or are under consideration to be published elsewhere. No sentence, equation, diagram, table, paragraph or section has been copied verbatim from previous work unless it is placed under quotation marks and duly referenced.

b. The work presented is original and own work of the author (i.e. there is no plagiarism). No ideas, processes, results or words of others have been presented as Author own work.

c. There is no fabrication of data or results which have been compiled/analyzed.

d. There is no falsification by manipulating research materials, equipment or processes, or changing or omitting data or results such that the research is not accurately represented in the research record.

e. The thesis has been checked using TURNITIN (copy of originality report attached) and found within limits as per HEC plagiarism Policy and instructions issued from time to time.

**Name & Signature of Supervisor**

**__Dr. Syed Omer Gilani____**

Signature**: ________________**

## Copyright Statement

## Acknowledgements

I am grateful to Almighty Allah for granting me with this opportunity and making me able enough to complete this research. I am in debt of my parents and siblings for being the source of this great opportunity.

I would like to express special thanks to my supervisor Dr. Syed Omer Gilani for his continuous support and guidance. Who supervised me every step along the way and helped me in streamlining my approach to get the desired results. I am grateful to have had worked under his wing. I would like to extend my sincere gratitude to him for helping me learn and providing me with time and attention.

I would like to credit Mahad Majeed for the time he invested in my research and my credit courses. And I would like to credit Moiz and Iqra for helping and motivating me during all this effort period.

I would also like to express my sincerest gratitude to Rector NUST, Lieutenant General Javed Mahmood Bukhari, and Principal SMME, Prof. Dr. Javaid Iqbal for providing us with an excellent research environment in SMME, NUST. In the end, I would like to thank all my friends who helped me in achieving this milestone and all my classmates for always being ready to extend their helping hand.

*Dedicated to my parents and teachers for helping me survive in this dreary world which otherwise would have been difficult*

# Abstract

Pulmonary Melanoma is the leading cause of cancer-related deaths. Early detection of lung cancer can significantly reduce the mortality rate. The main goal of this thesis is to segment the boundaries of lesions from Computed Tomographic images to assist the pulmonologists so that decisions can be made easier for them. The dataset is taken from Medical Decathlon. The dataset contains pre-labeled tumor images annotated by medical experts to enrich the level of analysis. The automated models used for the segmentation of images are: Nested U-NET, Mask RCNN, and DeepLabV3, and an Ensemble of these models is created. Pytorch and FastAl libraries were used for training the dataset. 33% of images were used for validation. The learning rate was at least during the data to avoid overfitting. The metric used for calculating results was Dice. The individual Dice scores of Nested U-Net, Mask RCNN, and DeepLabV3 are 0.7487, 0.7811, and .882 respectively and the accuracy for the ensemble is 0.895, and state-of-the-art is 0.77.



***Key Words:*** *Pulmonary melanoma, mortality rate, segmentation*

# Table Of Contents

# List of Figures

# List of Tables

# CHAPTER 1: INTRODUCTION

## 1.1 Lung Pathologies

In the previous few years, Pulmonary Melanoma turned into the major wellbeing sicknesses in the human body. It is very hard to analyze cellular breakdown in the lungs in the beginning phase of it, which may prompt the danger factor of endurance of patients. Correspondingly, the therapy on Lung disease relies on how early this sickness can be detected with the goal that therapy can handle the expanding (in stage) and spreading of Lung Cancer in other parts of the body. It is very conceivable to control Lung Cancer sickness by giving legitimate therapy.

There are different therapies accessible in the field of Medical Science such as Surgery, Chemotherapy *[1]*, and radiography *[2]* as it relies upon the phase of infection, the strength of the patient, and some different elements. The pace of endurance is just 14% of the patients for a very long time. Lung Malignant grows in the respiratory epithelium of the bronchial trees. It is extremely uncommon to recognize the cellular breakdown in the lungs before the age of 45. However, Lung Cancer can be diagnosed at the age of 55 to 70. *[3]*.

### 1.1.1 Lung Cancer

The organs and tissues of the body are composed of small building blocks called cells. Cells in various pieces of the body may look and work diversely yet most repeat themselves similarly. Cells are continually turning out to be old and kicking the bucket, and new cells are delivered to supplant them. Ordinarily, the division and development of cells are efficient and controlled; however, on the off chance that this cycle gains out of power for a few reasons, the cells will proceed to partition and form into an irregularity which is known as a tumor. Tumors can either be malignant *[4]* or benign, as demonstrated in Fig.1.

A malignant tumor gives birth to the term cancer. A cancerous tumor consists of malignant growth cells which can spread past the first site. Whenever left untreated, they may attack furthermore, obliterate encompassing tissues. Some of the time cells split away from the first (essential) malignancy and spread to different organs in the body by going in the circulation system or lymphatic framework. At the point when these cells arrive at another space of the body they may continue partitioning and structure another tumor, frequently alluded to as a "metastasis". It is critical to figure out that malignancy is anything but a solitary infection with a solitary kind of treatment. There are more than 200 various types of malignant growth, each with its own name and treatment.

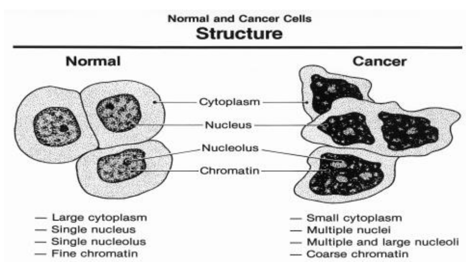



*Figure 1 Normal and Cancer Cells*

Cellular breakdown in the lungs is the harm of lungs that is the main source of malignant growth passing for the two people around the world. Around 90% of all cellular breakdown in the lungs happens in current or previous smokers. According to the American Cancer Society, there are 164,000 new cases of cancer every year in the United States alone and the deaths from these cases are 157,000*[5]*. The quantity of yearly passing nearly rises to the number of new cases. Indications don't show up until the disease is very cutting-edge.

## 1.2 Image Segmentation through Deep learning-An overview:

Earlier types of segmentation models included instance thresholding and watershed, histogram-based bundling, and k-mean clustering etc. Overtime, image segmentation saw diverse changes and it evolved into models like conditional and markov random fields, mean clustering, and graph cuts. These modern models and algorithms provide better accuracy. In image classification, there are primarily two segmentation techniques. In semantic segmentation technique, detection of objects at the pixel level is done. This segmentation technique then groups pixels in a meaningful way which can either be person, bicycle, animal, or anything. In instance segmentation, on the other hand, apart from semantic segmentation, each pixel level is assigned a mask.

## 1.3 Training Deep Learning Models:

Deep learning models are continually developing. Case organizations, spatial transformer models and gated intermittent units are being proposed by the specialists to improve the presentation of deep learning organizations. Curiously, deep neural organizations can be prepared on new datasets without any preparation. This is done through transfer learning in which a model that has been prepared on one kind of dataset is substituted with another sort of dataset so the model adjusts to the new kind of information and can help in the recognizable proof of new highlights in the dataset. A typical strategy for transfer learning is preparing the dataset on ImageNet and afterward feeding information in the new set, basically feeding it according to the previous set in the model. This aids in the acknowledgment of highlights in a dataset with less marked samples.

## 1.4 Image Segmentation models:

Following are the significant sorts of deep learning models that are most generally utilized in image segmentation.

### 1.4.1 R-CNN based Models:

The regional convolutional network is utilized for example segmentation. It has three basic kinds i.e., Faster R-CNN, Fast R-CNN and Masked R-CNN. These models are generally utilized for object discovery. A few versions are utilized for object location and semantic division. R-CNN models are broadly utilized in processes like instance segmentation. He et al. [6] introduced a model that expeditiously distinguishes objects while generating image masks for each instance. This model has three yield branches that process bounding box coordinates, , related coordinates and make double mass for the pictures, individually. The misfortune capacity of the model consolidates the misfortune capacity of jumping box facilitates, the division class and the forecasts and afterward prepares them together. Other effective models incorporate Path Aggregation network (PANet) (Liu et al. [7]), Multitask network for instance-aware segmentation (Dai et al. [8]), Masklab (Chen et al. [9]) and Tensormask (Chen et al. [10]). These models are generally utilized on account of their productivity.

### 1.4.2 Fully convolutional networks

Fully convolutional networks (FCN) are broadly utilized, these models are utilized to deliver segmentation maps which size comparable to the size of the picture. It was first introduced by Long et al. [11]. Over years FCN has developed as analysts changed networks like GoogLeNet and VGG16. The fully associated layers were supplanted by fully convolutional layers. The model accomplishes best in class results by examining and intertwining the element guides of prior layers. The model is tried on n PASCAL VOC, NYUDv2, and SIFT Flow to approve the segmentation results accomplished. FCNs are generally utilized for semantic segmentation.

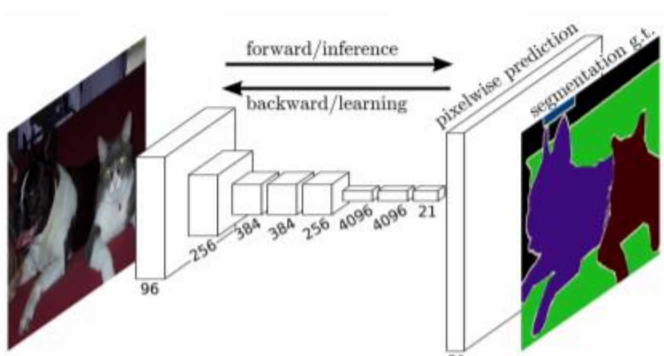



Figure 2 General working of fully convolutional network.

In spite of its wide use and great outcome, FCN has a couple of disadvantages in contrast with other profound neural networks. In the first place, customary FCN doesn't consider worldwide setting data, furthermore it is similarly delayed progressively. A couple of specialists have chipped away at defeating these weaknesses, for example ParseNet introduced by Liu et al. [12] utilizing the idea of setting vector and supplanted convolutional layers. Fully convolutional networks are right now being generally utilized in cerebrum tumor [13] and skin injury [14] segmentation in the field of clinical picture segmentation.

### 1.4.3 Convolutional models with graphical models

Convolutional models which fuse expected graphical models like Conditional Random Fields and Markov Random Field are more productive contrasted with FCNs. Convolutional neural organizations (CNN) give helpless confinement. Chen et al. [12] joined the last convolution layer with Conditional Random Fields that created better confined outcomes and brought about less generalizability.

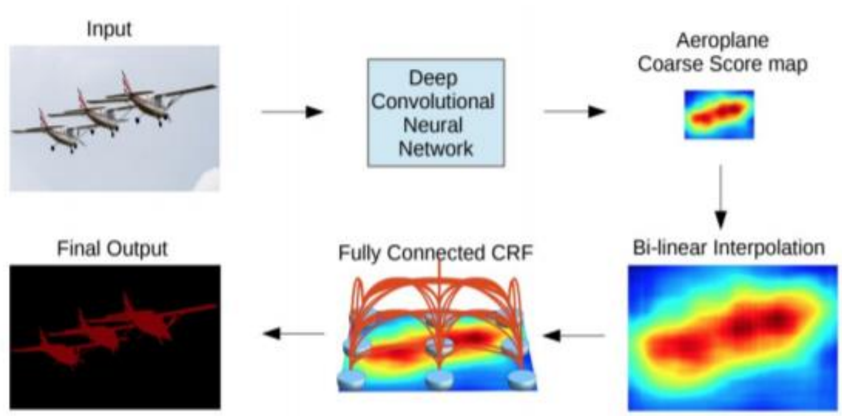



Figure 3: Graphical model (CNN+ Conditional Random Fields) proposed by Chen et al.

Scarcely any different models have been proposed for semantic division that incorporate mix of CNN and Conditional Random Fields have created promising outcomes. Lin et al. [15] proposed with idea of deep Conditional Random Fields and fix foundation to improve semantic division.

This model utilized relevant data. Liu et al. introduced model Parsing Organization which consolidates CNN and Markov Random Field [16]

## 1.5 Medical Image Segmentation:

Medical image segmentation is a distinct branch of image segmentation. Profound convolutional networks have seen huge advancement and accomplishment in close past in this field. Medical image segmentation is finished by customary and profound learning image segmentation strategies. MRI images and CT are the most utilized kinds of images in radiological strategies.

These are additionally most broadly utilized in the medical segmentation. Beginning procedures contained edge discovery and numerical models yet the pattern has moved to numerical models. Handmade highlights were the essential contribution of these frameworks yet with the presentation of profound learning procedures since the mid-2000s, scientists have discovered deep learning techniques figuring out how to be really encouraging. Grouping, enlistment and location in medical image segmentation strategies hold essential significance.

## 1.6 Motivation

To stay protected from the Lung Cancer illness, and for examining the beginning phase of this illness required incredible technology to help the specialists is generally attractive, in specific Machine Learning, Artificial Intelligence Technique and Image Processing can deal with the clinical field information with the guidance of engineering solution to achieve the goal of diagnosing and detection of the Lung Cancer. It is required to prepare and process the clinical field dataset such as Computed Tomography (CT) and scan pictures by applying different Neural Network, Machine Learning procedures to the input dataset *[7]*.

# CHAPTER 2: LITERATURE REVIEW

Computers helped in the diagnosis of lung CT picture have been a noteworthy and progressive advance, in the early and untimely detection of lung abnormalities. The CAD frameworks incorporate systems for 'programmed discovery of anomaly knobs' and '3D remaking of lung' frameworks, which help the radiologists in their final choices. Image processing calculations and procedures are applied to the pictures to explain and improve the picture and afterward to isolate the space of interest from the entire picture. The independently obtained area of interest is then investigated for the location of knobs to analyze the illness *[8]*.

## 2.1 Importance of CT, Segmentation and Machine Learning Technique

Clinical treatment has consistently been finished with indications-based investigation. This implies that patients first have their indications investigated and, if vital, they are sent for a more exact examination (trained professionals and sweeps). These days, the idea of "exact medication" attempts to tackle the issue of the huge yet cracked condition of biomedical information. This is finished utilizing patient-driven arrangements and putting away the computerized information of the patients in shareable online data sets *[9]*. Besides, the European Clinical Association, the World Health Organization, and the United States Association have found that there is a tremendous expansion in a cellular breakdown in the lungs in the United States and Europe, making Lungs Cancer the main source of death in Europe and the US, *[10]*.

There are two general types of newly evolved computational techniques for CT imagery prediction of the lung nodule: one of the Methods focused on objective radiological picture characteristics (QIF. Secondly, Deep learning methods including those based on deep neural networks in convolution (CNN) *[16]*. Radiomics methods typically build a predictive model based on extracted radiological quantitative picture function (2D) or 3D features of lung nodules on a previous know-how of significant features and features. Radiomics methods were developed using information and data collected such as the Lung Image Database Consortium (LIDC/IDRI) and the NLST, or using proprietary data sets mostly small but validated by biopsies or surgical resections [10].

Early identification and detection of lung cancer are the main measures to improve the results. According to the NCCN guidelines, histopathological examination of the biopsies obtained through fiber-optic bronchoscopy should be done for the diagnosis of image suspected tumors. A pathologist's evaluation of biopsy tissue is the main aspect for the diagnosis of lung cancer. However, the precision of the diagnosis was below 80% [4]. Squamous carcinoma, adenocarcinoma, small cell carcinoma and undifferentiated carcinoma are the main histological subtypes with malignant lung diseases. The proper assessment of these biopsy subtypes is crucial for the right choice of care. Nevertheless, in emerging economies such As china, with a significant population of lung cancer patients, the numbers of trained pathologists are too limited to fulfil major clinical demands.

Image Segmentation alludes to the way toward parceling a picture into distinct regions by gathering together neighborhood pixels dependent on some predefined standard. The comparability model can be resolved to utilize explicit properties or highlights of pixels addressing objects in the picture. At the end of the day, the division is a pixel order procedure that permits the development of areas of similitudes in the picture [11].

The division has stayed as a significant apparatus in clinical picture handling, and it has been valuable in numerous applications. The applications incorporate the location of the coronary boundary in angiograms, various sclerosis injury evaluation, medical procedure recreations, careful arranging, estimating tumor volume and its reaction to treatment, practical planning, computerized order of platelets, contemplating mental health, location of miniature calcification on mammograms, picture enlistment, map book coordinating, heart picture extraction from heart cine angiograms, recognition of tumors and so on *[12]*.

In clinical imaging, the division is significant for highlight extraction, picture estimations, and picture show. In a few applications it very well might be helpful to order picture pixels into anatomical areas, like bones, muscles, and blood vessels, while in others into obsessive locales, for example, disease, tissue deformations, and various sclerosis injuries. In a few investigations, the objective is to separate the whole picture into sub-districts like the white matter, dim matter, and cerebrospinal fluid spaces of the mind, while in others explicit design must be removed, for instance, breast malignancy from Magnetic Resonance pictures.

## 2.1 R-CNN based Model:

For segmentation, the regional convolution network has three key categories, i.e. R-CNN fast, R-CNN masked and R-CNN quicker. This models are commonly used for the identification of objects. For target recognition and semantically segmentation several variants are included. For example, segmentation, R-CNN models are commonly used proposed a model that senses objects effectively and generates image masks for each instance. This model consists of three output branches, which compute bounding coordinates, corresponding co-ordinates and construct an image binary mass. The model for automated image segmentation Mask R-CNN. It uses PET imaging for the diagnosis of lung tumour, based on multi-scale region mask – Convolutionary Network (Mask R-CNN). First, for the identification of lung tumour candidates, we developed three versions of mask R-CNN [13]. These models were generated by perfecting the R-CNN mask using some training data consisting of images from 3 various scales. Each data set was comprised of 594 lung tumour slices. These three versions of the R-CNN mask models were combined with a weighted voting technique to reduce the false positive results.

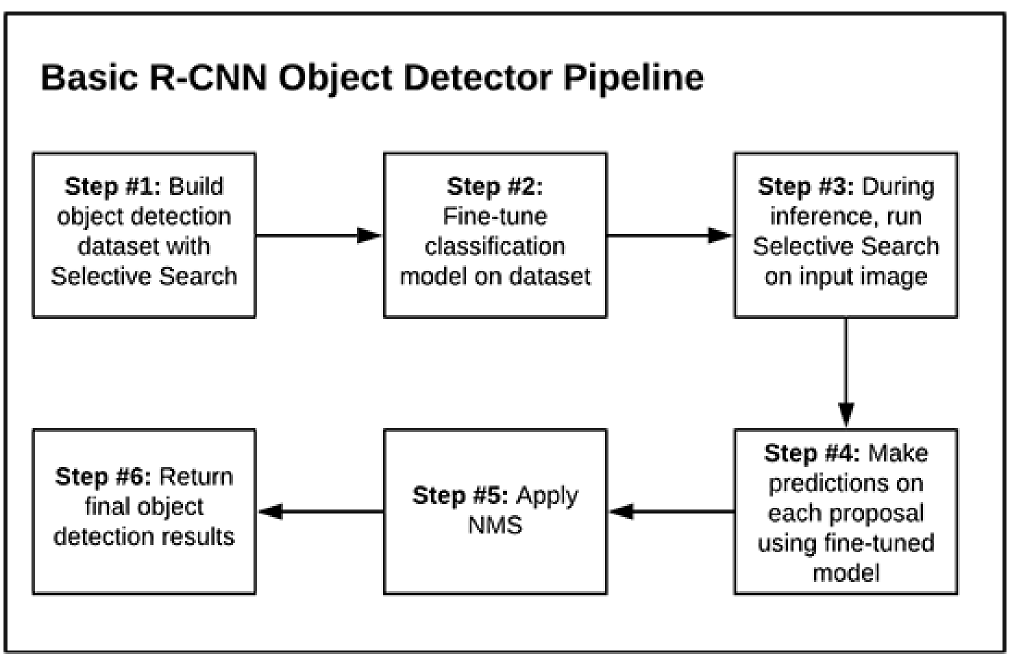



Fig 4: Basic R-CNN Object Detector Pipeline

Motion control is crucial to guarantee the efficacy of radiation therapy for patients with lung cancer. For this purpose, the suggested network, called the motion region-based CNN, consists of four main phases i.e. extraction of functions, position of rough tumor, localization of fine tumors, and categorization within the region of interests of the tumour (ROI). It first fed 4D CT with successive phases into the backbone to extract tumour movement information to allow the network to get rid of unreasonable deleted ROIs, which differ from traditional R-CNN deformation masks, and then used this movement information to estimate the global local deformation vector fields (DVF).

This model is strongly convinced of the efficacy of that approach for the detection of lung tumors, as well as of the potential to distinguish a stable chest pattern and to greatly reduce tumour diagnosis [13]. The model's loss function unites, and then trains, the loss function of bounding box coordinates, segmentation class and projections. Other efficient models include Path Aggregation (PANet), Multitask Network for example Masklab and Tensormask segmentation [14]. Because of the reliability of both of these models.

## 2.2 U-Net network architecture for Segmentation:

Ronneberger launched U-net in which Skip links are the main concept. Multi-sequence images are being integrated into deep-learning approaches and have recently gained increasing interest. Early and late fusion strategies in convolutionary neural networks, i.e. early fusion strategy, were discussed and the various sequences of MRI data were combined in view of the linear connections between each sequence, while all sequences were processed with separate CNN and merge into deeply-rooted levels in late fusion strategies. Although the late fusion strategy segmentation results are better than the early fusion strategy, these strategies are nevertheless functioning on one layer which does not address the non-linearity and complexity of multimodal MR images *[14]*.

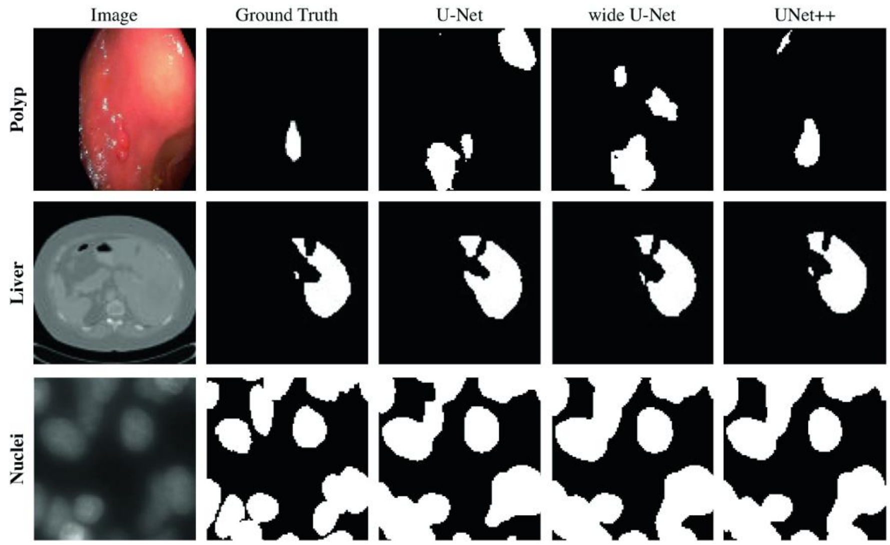



Fig 5: Nested U-Net Architecture for Medical Image Segmentation

This ambiguity was explained by the use of another Hyperdense network that not only connects layer pairs but also the layer along various routes. And this system is called the U-Net model where one direction is encoded and the other is decoded (expanding). This approach is also used for Brain tumour location and segmentation multimodality MR images.

## 2.3 Fully convolutional networks (FCN):

These models are used in large numbers for producing categorization maps that match image dimensions. Fully convolutionary networks (FCNs) are commonly used. Long et al. [8] suggested it first. Over the years the FCN has grown to modify networks such as Google Net and VGG16 as investigators. The completely connected layers have been replaced by completely convolutionary layers. By sequencing and fusing the function maps of earlier sheet, the approach outperforms cutting-edge output. To validate the achieved segmentation performance, the model is evaluated on n PASCAL VOC, NYUDv2, and SIFT Flow. Semantic segmentation is commonly used by FCNs.

In the FCN, up-sampled feature maps are combined with the function maps that the encoder has missed, while U-Net binds them together and adds convolutions and non-linearity between each move. The convolution layers have shown that they can recover the total temporal resolution on the target values, allowing completely convergent methods for semantic segmentation. Li suggested H-dense-net for the segmentation of the liver and liver tums, inspired by Dense-Net's architecture *[15]*. Drozdzalet extensively examined the relevance of skip linkages and added fast skip links even within encoder in this spirit *[14]*.

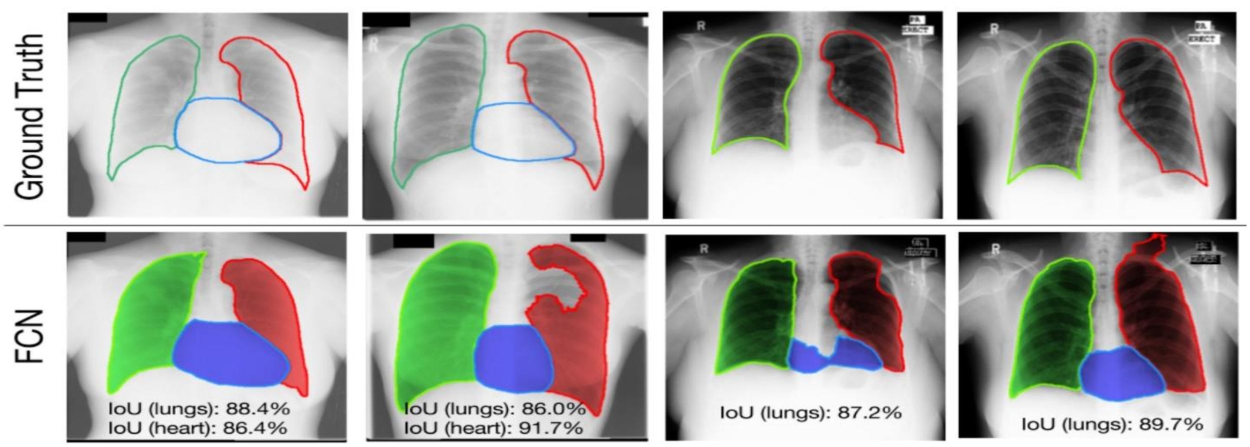



Fig 6: Robust segmentation of lung in chest x-ray through FCN

FCN does have some disadvantages compared to other deep neural networks, including its large usage and strong performance. Firstly, traditional FCN does not account for global context knowledge and, secondly, in real time, it is very sluggish. For example, ParseNet suggested by Liu [9]. using the principle of background vector and substituted convolutional layer has been worked on by a few researchers to overcome these weaknesses. In brain tumor and in skin lesions segmentation in the fields of medical picture segments, entirely convolutional networks are currently commonly used.

### 2.3 DeepLabv3 Architecture:

DeepLabv3 is an architectural semantic categorization which enhances with a number of amendments on DeepLabv2. Modules that use atrocious concoctions in cascading or at the same time as capturing multi-scale contexts at various attractive rates to deal with the issue of multi-scale segmentation. Moreover, DeepLabv2, Atrous Space Pyramid Module has been added to encode global contexts and improve output with image level functionality.*[17]*

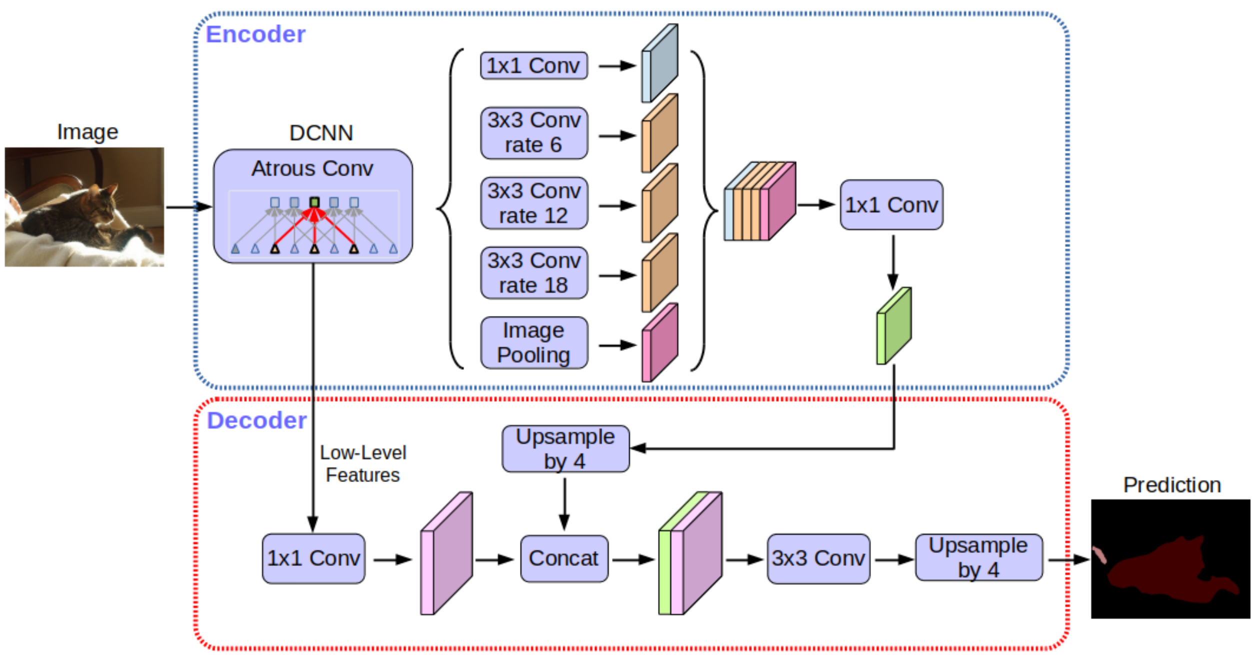



Fig 7: DeepLabv3 semantic segmentation architecture

The modifications to the ASSP framework are to implement a global average grouping on the last function map of the model, feed the resulting image level features into a 1/1 co-ordination with

256 filters (as well as batch standardization) and then upgrade the aspect to the spatial dimensions requested bi-linearly. The enhanced ASPP consists of (a) 1/1 and (a) 3/3 of a range = (6, 12, 18) while the range is 16 (all 255 filter and batch normalization) and (b) the image-level characteristics. Similarly, segmentation technique DeepLabV3 use atrociously convergent solutions in a cascade or simultaneously to catch multi-scale contexts by implementing multitask atrocities are structured to address the issue of multi-scale segmentation of items.

## 2.4 Ensemble models:

Ensemble models is a machine-learning method for combining many other frameworks. These models are known as baseline estimators. The following technical problems are solved in the construction of a shared amenities. A single approach cannot predict a particular data set perfectly. Machine learning algorithms have their drawbacks, and it is difficult to produce a high-precision model. The overall consistency can be improved if we create and mix several models. The combination can be realized by adding the contribution of each model with two aims: the reduction of the model uncertainties and its generalization. Any methods may be used to apply this aggregation.

In Ensemble many models are widely used to optimize the predictability of model enhancement techniques.

- Metamodynamic
- Stimulate
- Package
- Change of the parameter
- Use various predictor sets

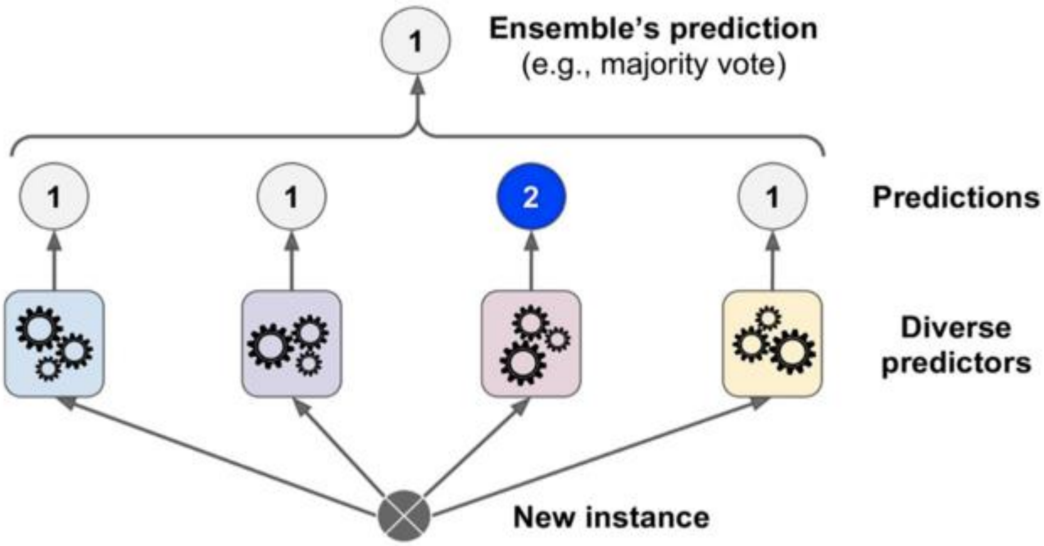



Fig 8: Hard voting classifier predictions

In the early and untimed identification of lung problems, the computers helped to diagnose the lungs of a CT image were a notable and increasing development.

Clinical therapy using indication-based research has regularly been completed. Different models generate the results of medical decathlon challenge according to shape and size of organ.

# CHAPTER 3: METHODOLOGY

## 3.1 Pre- Processing:

The three-dimensional volumes were changed over to two-dimensional pictures. This permits us to use move learning with two-dimensional division models like U-net.

## 3.2 Frameworks:

### 3.2.1 Pytorch:

PyTorch is an open-source machine learning library dependent on the Light library, utilized for applications like PC vision and normal language handling and so on. Pytorch was utilized as it has quicker preparation speeds, is more proficient, and utilizes lesser assets. Pytorch is more instinctive also.

### 3.2.2 FastAl

FastAI is another library. It is based on top of Pytorch that builds its latent capacity much more. FastAI professes to prepare imagenet quicker on GPU contrasted with tensorflow on TPU. It has a great deal of implicit models including U-net. Additionally its partner works for one-cycle-getting the hang of, discovering ideal learning rate, and so on.

## 3.3 Data Training

### 3.3.1 NESTED U-NET

The Dataset was trained on Imagenet Dataset and the backbone employed was Resnet101 as there was no pretrained dataset for nested Unet. Feature maps from each network level are transposed by 1 × 1 × 1 convolutions to create secondary segmentation maps. These are then combined in the following way: First, the segmentation map with the lowest resolution is upsampled with bilinear interpolation to have the same size as the second-lowest resolution segmentation map. The element-wise sum of the two maps is then upsampled and added to the third-lowest segmentation map and so on until we reach the highest resolution level. An illustration is given in the following figure.

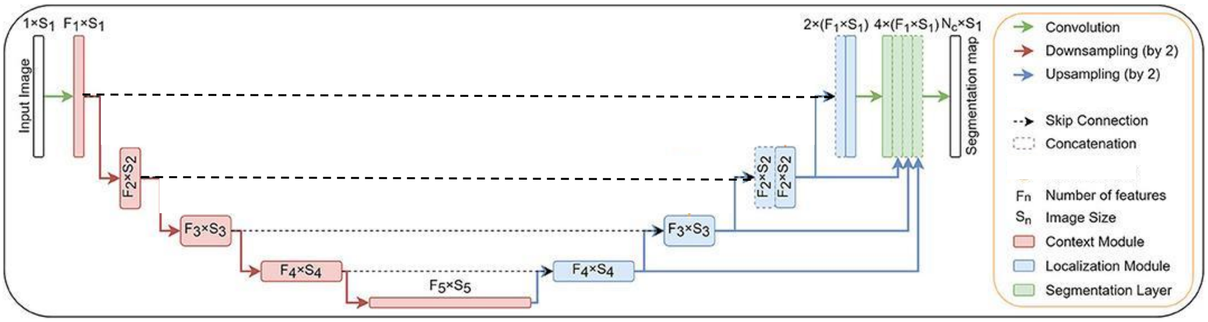



Figure 9: A representation of Nested U-Net architecture [16]

.

| Hyperparameter | Value |
|---|---|
| Epochs | 400 |
| Initial learning rate | 1e-3 |
| Final learning rate | 1e-5 |
| Batch size | 2 |

### 3.3.2 Mask RCNN

Facebook AI Research (FAIR) provides Detectron2 repository that has pretrained Mask RCNN models. The repository was used to train Mask RCNN model on this dataset. The backbone chosen for this purpose was Resnet101 and MRCNN was pretrained on COCO dataset. With ADAM as the optimizer, the network was trained for 200 epochs. For segmentation, categorical crossentropy loss was used while for bounding box regression, mean square error was used. We used cosine annealing scheduler to change the learning rate at every iteration. Other hyperparameters are summarized in the table below.

| Hyperparameter | Value |
|---|---|
| Batch size | 4 |
| Base learrning rate | 0.001 |
| Anchor sizes | 4, 8, 16, 32, 64 |
| Anchor aspect ratios | 0.5, 1, 2 |

The model took around 36 hours to train on nvidia P100 GPU in Google's Colab.

### 3.3.3 DeeplabV3:

| **Hyperparameter** | **Value** |
|---|---|
| Epochs | 200 |
| Optimizer learning rate | .0005 |
| Batch size | 5 |

For training of DeeplabV3, 3D images were used where each channel was basically a duplicate of the gray level image given. Batch size of 5 was used where each image was of size 512. Training was done for 200 epochs.

## 3.4 Ensemble Model:

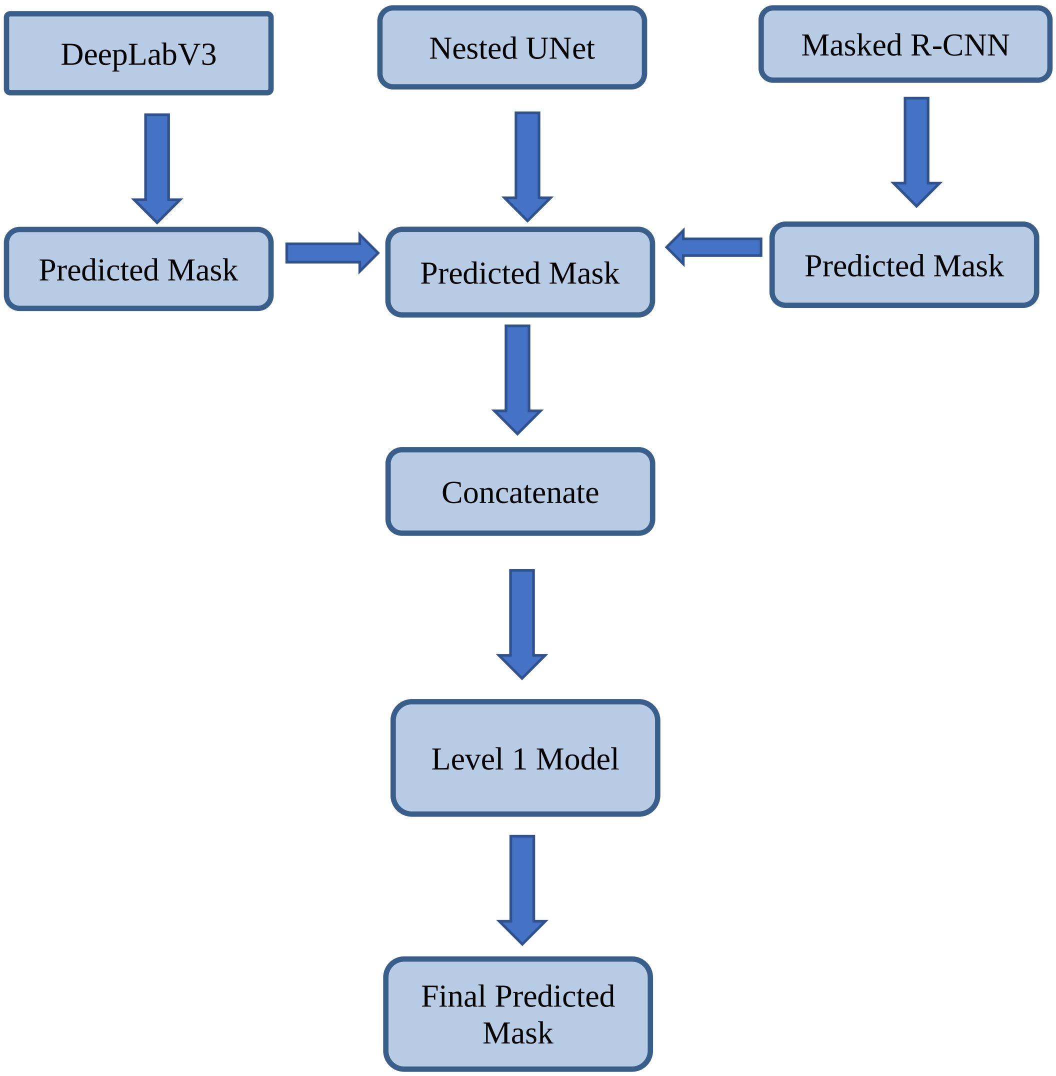
Ensemble Model uses multiple networks for enhanced predictive performance. It consists of multiple base models called as level 0 models. The outputs from these models are then concatenated/ stacked together to get a new matrix which is then used as input to a new model called the level 1 model. In this way, the ensemble model is an optimal network that learns to choose how to best combine the predictions from multiple existing models.

We have used 3 networks for model ensemble, namely:

- DeeplabV3
- Nested UNet
- Detectron (Masked R-CNN)

These 3 models serve as the level 0 models of our Ensemble Model.

Inference was made from each network giving us 3 512x512 mask predictions. These 3 mask predictions were concatenated to get a single matrix of 3x512x512. For ensemble model, a single Conv2D layer was used; it takes this concatenated 3D matrix as input and then gives a final prediction mask as output.




| **Hyperparameter** | **Value** |
|---|---|
| Epochs | 115 |
| Optimizer learning rate | .0005 |
| Batch size | 2 |

Evaluation matrices used: mIOU and mAP

# Chapter 5: Results

For evaluation of the trained model, Medical decatahlon provided the dataset and had the results updated for improving it overtime 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. Grand Challenge provided a dataset which did not have annotations available for the public.

## Sorensen-Dice Coefficient:

The Dice coefficient provides information of the overlapped segments and ground truth values. 1tells us the perfect segmentation while 0 for the worst case.

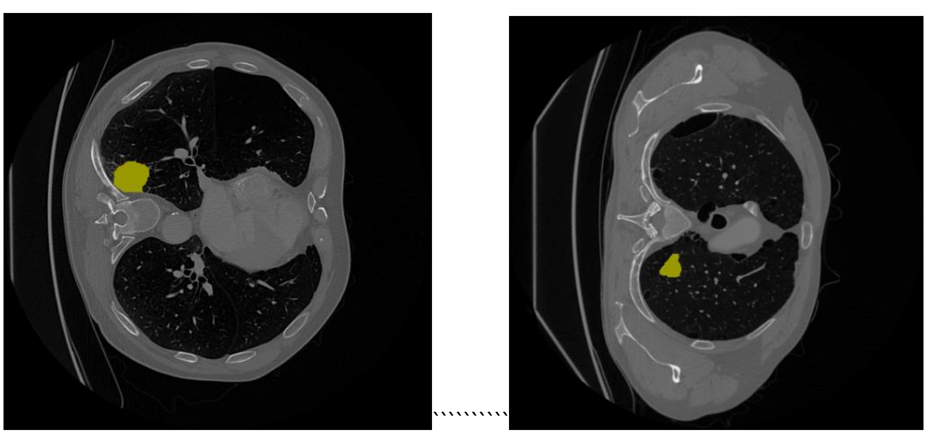


Fig 10: CT images of Lungs

**DICE SCORE FOR MASK RCNN**

| | |
|---|---|
| Max | 0.9743825740144194 |
| Min | 0.0 |
| Standard deviation | 0.28256121098002235 |
| 25th percentile | 0.7527868852459016 |
| 50th percentile | 0.8599766627771295 |
| 75th percentile | 0.9179566563467493 |
| Mean | 0.7811367723895286 |

(Aggregate Dice evaluation)

**DICE SCORE FOR NESTED UNET:**

| | |
|---|---|
| Max | 0.9757071547420965 |
| Min | 0.0 |
| Standard deviation | 0.2951568293816984 |
| 25th percentile | 0.7587131367292225 |
| 50th percentile | 0.8760195758564437 |
| 75th percentile | 0.919908466819222 |
| Mean | 0.7487588413720577 |

(Aggregate Dice evaluation)

**DICE SCORE FOR Deeplab V3**

| | |
|---|---|
| Max | 0.967071547420965 |
| Min | 0.0 |
| Standard deviation | 0.2651568293816984 |
| 25$^{th}$ percentile | 0.7487131367292225 |
| 50$^{th}$ percentile | 0.8760195758564437 |
| 75$^{th}$ percentile | 0.909908466819222 |
| Mean | 0.8887588413720577 |

**SCORE FOR ENSEMBLE**

| | |
|---|---|
| Max | 0.957071547420965 |
| Min | 0.0 |
| Standard deviation | 0.2953674199337415 |
| 25$^{th}$ percentile | 0.7587131367292225 |
| 50$^{th}$ percentile | 0.8760195758564437 |
| 75$^{th}$ percentile | 0.919908466819222 |
| Mean | 0.8929826832943452 |

# Discussion and future work

In this thesis, the main goal to facilitate doctors in their early diagnosis of pulmonary melanoma has been done so that treatment can be started early on because malignancy if diagnosed at early stages help reduce the chances of deaths to a great extent.

So far as the future work is concerned, there is a great potential in this project and although 89% accuracy has been attained, so long as 100% accuracy cannot be acquired, one should keep on working for the betterment of humanity so this plague can be wiped off from the face of this earth.

## References


**1.** 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).

**2.** 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).

**3.** 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).

**4.** 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).

**5.** 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).

**6.** 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).

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

**8.** 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

**9.** 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.

**10.** 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*.

**11.** 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).

12. 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.

**13.** . https://www.webmd.com/lung-cancer/guide/chemotherapy-lung-cancer#:~:text=goals%20and%20preferences.-,When%20Chemotherapy%20Is%20Used,to%20treat%20any%20remaining%20cancer.

**14.** https://www.lung.org/lung-health-diseases/lung-disease-lookup/lung-cancer/patients/treatment/types-of-treatment/radiation-therapy#:~:text=The%20radiation%20is%20aimed%20at,type%20of%20lung%20cancer%20treatment.

**15.** Witschi, H. A short history of lung cancer, Toxicol Sci, 2001. 64(1): p. 4-6.

**16.** Malin Dollinger, "Everyone's Guide to Cancer therapy; How Cancer is Diagnosed, Treated and Managed", 4th Edition, Andrews McMeel Publishing Kansas, USA, 2002.

**17.** R Wicmker Ph.D., P. Rogalla MD, T Blaffert Ph.D., "Aspects of Computer-aided detection (CAD) and volumetry of pulmonary nodules using multislice CT", The British Journal of Radiology (BJR), vol. 78, pp. 46-56, 2005

**18.** Malin Dollinger, "Everyone's Guide to Cancer therapy; How Cancer is Diagnosed, Treated and Managed", 4th Edition, Andrews McMeel Publishing Kansas, USA, 2002.

**19.** Furat, et al., Machine-Learning Techniques for the Segmentations of Tomographic Images Data of Functional Materials, Frontiers in Materials, 2019. 6(145).

**20.** M. A. Khawaja, Muhammed Zaheer Aziz, Nadeem Iqbal, "Effectual Lung Segmentation for CAD Systems Using CT Scan Images", Proceedings of IEEE, INMIC Conference, FAST Lahore, 2004.

**21.** Vallone, S. Luce Report on Lung Cancer-Challenges in Lung Cancer in Europe, Lung Cancer Europe.

**22.** http://www.lungcancereurope.com

**23.** Atam P. Dhawan, "Medical Image Analysis", IEEE press series in Biomedical Engineering, John Wiley & Sons. Inc. Publications, 2003.

**24.** Jadwiga Kogowska, "Overview and Fundamentals of Medical Image Segmentation", Handbook of Medical Imaging, Academic Press, San Diego, pp. 69-85, 2000.

**25.** Yang Lei, Zhen Tian, Tonghe Wang, Justin Roper, Kristin Higgins, Jeffrey D. Bradley, Walter J. Curran, Tian Liu, and Xiaofeng Yang "Mask R-CNN-based tumor localization and segmentation in 4D Lung CT", Proc. SPIE 11600, Medical Imaging 2021: Biomedical Applications in Molecular, Structural, and Functional Imaging, 116000M (15 February 2021)

**26.** Drozdzal, M., Vorontsov, E., Chartrand, G., Kadoury, S., Pal, C.: The importance of skip connections in biomedical image segmentation. In: Carneiro, G., et al. (eds.)

**27.** LABELS/DLMIA -2016. LNCS, vol. 10008, pp. 179–187. Springer, Cham (2016).

**28.** Li, X., Chen, H., Qi, X., Dou, Q., Fu, C.-W., Heng, P.A.: H-DenseUNet: hybrid densely connected UNet for liver and liver tumor segmentation from CT volumes. arXiv preprint arXiv:1709.07330 (2017)

**29.** Liu, Y. Radiological image traits predictive of cancer status in pulmonary nodules. Clin. Can. Res. 23(6), 1442–1449 (2016)

**30.** https://paperswithcode.com/method/deeplabv3