REPUBLIC OF TURKEY

AKDENIZ UNIVERSITY

LUNG CANCER DETECTION USING MACHINE LEARNING

Asma HAIBELTY

INSTITUTE OF NATURAL AND APPLIED SCIENCES

DEPARTMENT OF COMPUTER ENGINEERING

MASTER THESIS

JUNE 2022

ANTALYA

REPUBLIC OF TURKEY

AKDENIZ UNIVERSITY

LUNG CANCER DETECTION USING MACHINE LEARNING

Asma HAIBELTY

INSTITUTE OF NATURAL AND APPLIED SCIENCES

DEPARTMENT OF COMPUTER ENGINEERING

MASTER THESIS

JUNE 2022

ANTALYA

REPUBLIC OF TURKEY

AKDENIZ UNIVERSITY

INSTITUTE OF NATURAL AND APPLIED SCIENCES

LUNG CANCER DETECTION USING MACHINE LEARNING

Asma HAIBELTY

DEPARTMENT OF COMPUTER ENGINEERING

MASTER THESIS

This thesis unanimously accepted by the jury on 11/05/2022.

Assist. Prof. Dr. Mustafa BERKAY YILMAZ (Supervisor)

Assist. Prof. Dr. Hüseyin GÖKHAN AKÇAY

Assist. Prof. Dr. Ilyas ÖZER

ÖZET

MAKİNE ÖĞRENİMİ LUKKANARAK AKCİĞER KANSERİ TESPİTİ

Asma HAİBELTY

Yüksek Lisans Tezi, Bilgisayar Mühendisligi Anabilim Dalı

Danışman: Dr.Ögr.Üyesi Mustafa BERKAY YILMAZ

Haziran 2022; 35 Sayfa

Bu tez, akciğer kanseri tespiti ile ilgili bir model geliştirme çalışmalarının sonucu olarak ortaya çıkmıştır. Bu amaçla çalışmaya kapsamlı bir analiz ve kritik süreciyle başlayarak, geliştirilecek en uygun modelin tanımlanması sağlanmıştır.

Amaç, imgeleri doğru şekilde sınıflandırabilecek bir sisteme sahip olmaktır. Buradaki fikir, bilgisayarlı tomografi (CT) ile elde edilen tarama görüntülerinden akciğer kanserinin türünü doğru tanıyabilmektir. Kullanılan üç sınıf; iyi huylu, kötü huylu ve normaldir. İlk olarak, modelin geliştirilmesini sağlayacak veri kümesi aranmıştır. Daha sonra bu veri modeli en iyi besleyecek şekilde organize edilmiştir. Çalışmanın daha iyi anlaşılabilmesi için bazı veri analizi yöntemleri kullanılmıştır. Nihayetinde, derin öğrenme modelinin gerçeklemesine geçilmiştir.

Tezde elde edilen sonuçlar eğitim ve doğrulama kesinliği ve kaybı için figürlerde belirtilmiştir. Sonuç olarak, yapılan çalışma akciğer kanseri tespitine açık ve faydacı bir bakış önermektedir

ANAHTAR KELİMELER: Bilgisayarlı tomografi, Derin öğrenme, Veri seti, Veri analizi.

JÜRİ: Dr.Ögr.Üyesi Mustafa BERKAY YILMAZ

Dr.Ögr.Üyesi Hüseyin GÖKHAN AKÇAY

Dr.Ögr.Üyesi Ilyas ÖZER

ABSTRACT

LUNG CANCER DETECTION USING MACHINE LEARNING

Asma HAIBELTY

MSc Thesis in Computer Engineering

Supervisor: Asst.Prof.Dr. Mustafa BERKAY YILMAZ

June 2022; 35 pages

This document is the result of our work in constructing a model for detecting lungs cancer issues carried out as part of the end of studies. To begin this, we began our work with an exciting study through exhaustive analysis and criticism, which allowed us to define the best model to implement.

The target is if we have the ability to create a system that can correctly classify images. That is our idea of creating a system that can recognize the type of lung cancer in CT scan represented in images. The three classes are benign, malignant, normal. First, we looked for the dataset that allowed us to build our model. Then, we structured it in the best way possible in order to feed our model by this data. We ran some data analysis methods to have a better understanding of our work. To finally move on to the implementation of our D.L Model.

The results of our thesis are presented on the figure for the Training & Validation accuracy and loss and in conclusion, the work carried out offers a very clear and pragmatic vision on the lung cancer detection.

KEYWORDS: CT: (Computed tomography), DL: (Deep Learning), Data Analysis, Dataset.

COMMITTEE: Asst.Prof.Dr. Mustafa BERKAY YILMAZ

Asst.Prof.Dr. Hüseyin GÖKHAN AKÇAY

Asst.Prof.Dr. Ilyas ÖZER

ACKNOWLEDGEMENTS

First of all, we would like to warmly thank Asst.Prof.Dr. Mustafa BERKAY YILMAZ, for his precious advice and encouragement. Without his efforts this work would never have seen the light of day. On this occasion, we would like to express our gratitude to our supervisor and to thank him for his patience, his availability and his judicious advice which enriched this thesis.

We would also like to thank my family and all people who have participated directly or indirectly in the development of this work while expressing our gratitude to our teachers who have taught us their moral and intellectual support throughout our journey.

Finally, we would like to thank all members of the jury who were kind enough to participate in our service, read our thesis and wanted to share their points of view and their advices to help us to improve our skills.

LIST OF CONTENTS

TEXT OF OATH

I declare that this study " Lung Cancer Detection Using Machine Learning", which I present as master thesis, is in accordance with the academic rules and ethical conduct. I also declare that i cited and referenced all material and results that are not original to this work.

11/05/2022
Asma HAIBELTY

ABBREVIATIONS

<u>Abbreviations:</u>

ML   : Machine Learning

AI   : Artificial intelligence

DL   : Deep Learning

CT   : Computerized tomography

HPV : Human papillomavirus

PET   : Positron emission tomography

MRI : Magnetic resonance imaging

NN   : Neural Network

ANN : Artificial Neural Network

ReLU : Rectified Linear Unit

CNN : Convolutional Neural Networks

FC   : Fully connected

LOSS : Loss layer

EDA : Exploratory data analysis

LIST OF FIGURES

LIST OF TABLES

1.INTRODUCTION

A proof-of-concept model has been developed. The first thing to create a web app that contains the DL model. The user gives a file including images CT scan of lung and get an aswer from the model, the model reads the file and displays the result of the image given by the user, as if there is a cancer or not.

There are many aspects of the research that support the ability of machine learning to find lung cancer in CT scans. The dataset used to train the model was collected from Iraq-Oncology Teaching Hospital / National Center for Cancer Diseases in specialty hospitals for a period of 3 months in the year 2019, the data set uses 1190 images of which it includes three classes benign, malignant and normal.
Using an evaluation method, we were able with ML to detect the difference between them.

The Goal of this thesis first, self Education on machine Learning because we have a large part which contains a lot of machine-learning, Initially, therefore by doing this project we will have a lot of understanding of machine learning concepts.

The model must be created to assist some users. The users will be healthcare professionals working in the diagnosis of lung cancer, also the model must be able to obtain the images from the user-supplied dataset that will be used by the ML model and must be able to detect lung cancer in user-supplied images then give information about the image given and displays if there is cancer or not.

The thesis includes various subjects in computer science to attempt to find solutions to fix a real-world issue in the medical range also data mining and software development plays an important role to provide an argument of notion.

2.LITERATURE REVIEW

İn this section we presents the research that was conducted. A collection of relevant studies from various sources to prove knowledge in the medical field, including the research conducted on the technology and the dataset used.

2.1. Key facts about cancer

Cancer is a broad side including a large ambit of malady has the ability to impact any side on the body. A hallmark cancer is the fast production of unnatural cells which develop outside their usual limits then can then infest adjacent side of the body and dispersal to the other side. Roughly 70% of cancer deaths happen in least- and midst-revenue countries.[1]

Approximately a third of cancer deaths are attributable to cigarette use, high body mass index (BMI), alcohol consumption, vegetable intake, and absence of bodilyl motion.It is common for late-stage cancer symptoms to be unavailable for diagnosis and treatment, especially in least- and midst-revenue countries. [1]

In 2019, only 1 in 3 countries reported high-quality data on cancer rates.
Prof. Ismail Celik, President of the National Congress of Oncology and Immunotherapy in Turkey, said Friday that 350 people die daily in Turkey due to lung cancer, pointing out that these numbers are the highest in the world in terms of death rate from this disease. [1]

Celik explained, in the opening speech of the conference held in the southern state of Antalya, that they treat cancer through the newly developed "immunotherapy", which directs the body's defense cells against cancer. He explained that smoking constitutes an important disease burden in Turkey, and causes many heart diseases and atherosclerosis, not just cancer, pointing out that smoking leads to lung cancer.

The official Anadolu Agency quoted Celik as saying, "The best way to avoid this is not to smoke, to quit, and most importantly not to smoke next to those who do not consume it." The Turkish doctor pointed out that the number of deaths among women from lung cancer has become the first in Turkey. [1]

On December 15, 2020, the International Agency for Research on Cancer (IARC), a subsidiary of the World Health Organization, published the latest estimates of the global cancer burden, namely World Cancer Statistics. Available online as part of the IARC Global Cancer Observatory, the GLOBOCAN 2020 database provides estimates of incidence and mortality for 36 cancer types in 185 countries and all cancer types combined for 2020. It is estimated that the global cancer burden is increasing. 19.3 million new cases and 10.0 million deaths reported in 2020; In 2018, 18.1 million cases and 9.6 million deaths were reported. In 2012, the annual number of new cancer cases was 14.1 million, while the annual number of cancer-related deaths was 8.2 million. These numbers should not be perceived as the "frequency" of cancer. [2]

One out of every 5 people in the world will get cancer during their lifetime, and 1 in 8 men and 1 in 11 women die from cancer (see World Cancer Statistics 2020).Worldwide, the total number of people alive within 5 years of being diagnosed with cancer, called 5-year prevalence, is estimated at 50.6 million. [1]

Figure 2. 1. 2020 Turkey cancer statistics

The population of the country continues to increase. In parallel with this, the annual number of newly diagnosed cancer patients is also increasing. While the population of Turkey was 84.339,067 in 2020, the annual number of new cancer cases was 233,834 and the number of cancer-related deaths was reported as 126,335. While the population of Turkey was 81,916,866 according to the year 2018, the number of new cases was reported as 210,537 and the number of cancer-related deaths as 116,710. [1]
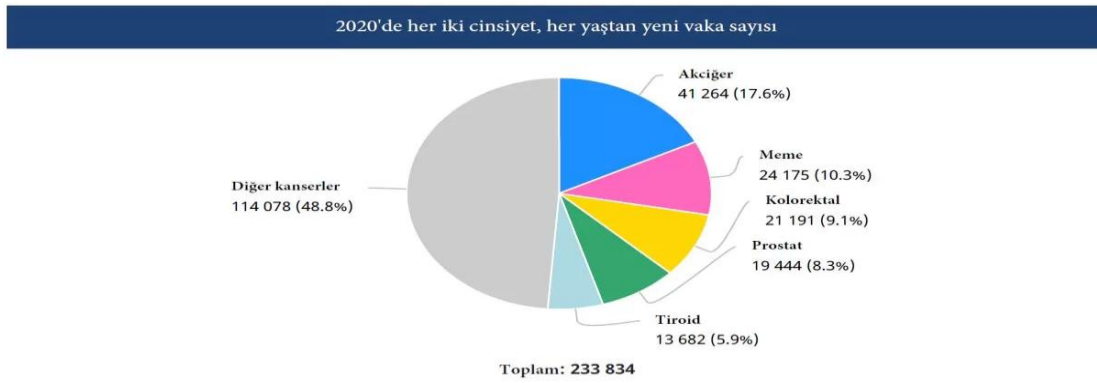
Figure 2. 2. generality popular cancer

The upper 5 generality popular cancer sorts fort he year 2020 in turkey were lung cancer, breast cancer, colorectal cancer, prostate and thyroid cancer. These 5 cancer sorts accounted for more than 50% of all cancers. [1]
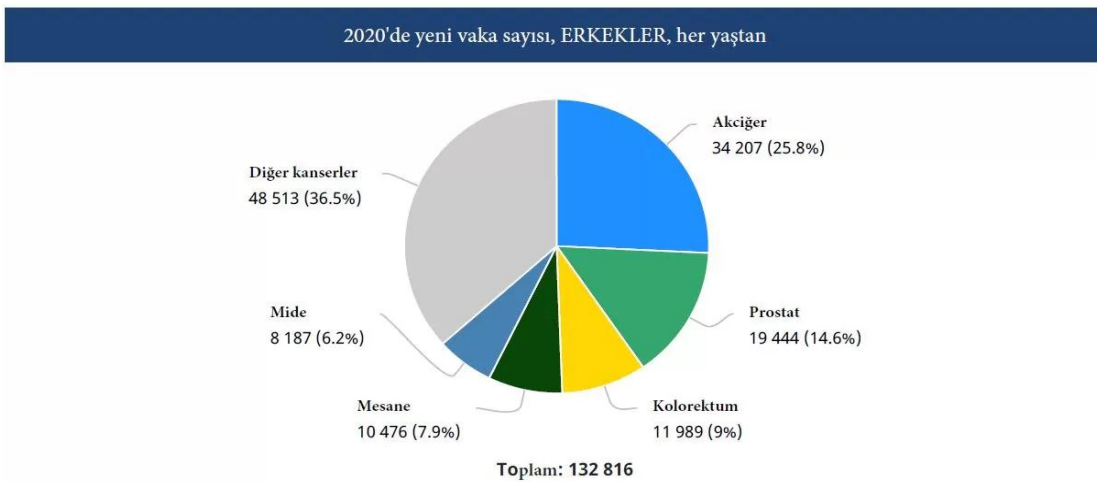
Figure 2. 3. generality popular cancer for men

The upper 5 generality popular cancer sorts in men were lung cancer, prostate cancer, colorectal cancer, bladder and stomach cancer. It is estimated that at least 40% of cancers seen in men are related to smoking. [1]

Figure 2. 4.   generality popular cancer for women[1]

The upper 5 generality popular cancer sorts for women are breast cancer, thyroid cancer, colorectal cancer, lung and uterine cancer. It is noteworthy that colorectal cancers are the 3rd generality popular cancer for both men and women.

Table 2. 1.   summary turkey cancer statistics

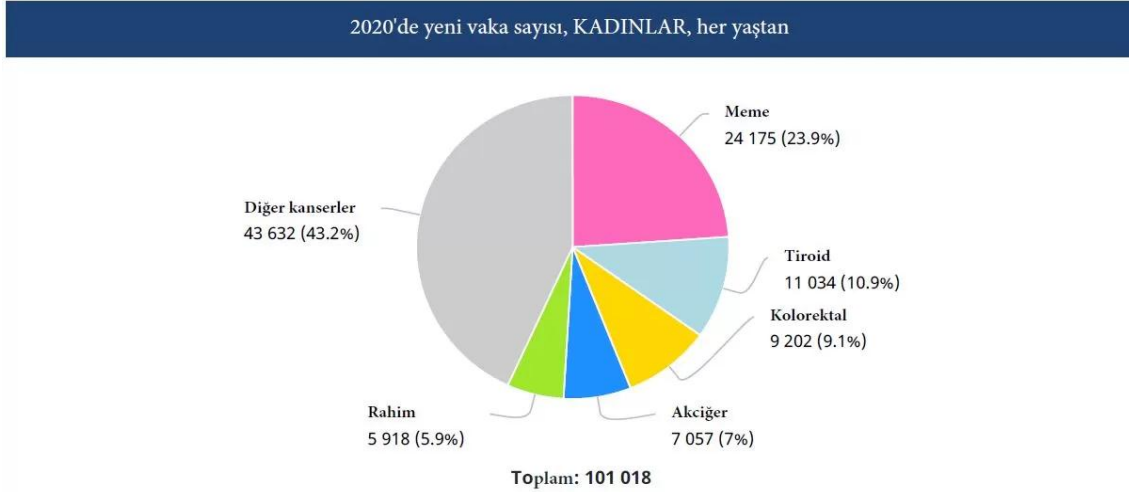| ÖZET Türkiye Kanser İstatsitikleri 2020			
	Erkekler	Kadınlar	Her iki cinsiyet
Toplam	41 636 125	42 702 942	84 339 067
Yeni vaka sayısı	132 816	101 018	233 834
Yaşa göre standardize edilmiş her 100 bin kişide yeni vaka sayısı	291.5	188.0	231.5
75 yaşından önce kanser geliştirme riski (%)	29.2	18.4	23.3
Kansere bağlı yaşam kayıplarının sayısı	78 949	47 386	126 335
Yaşa göre standardize er 100 bin kişide yeni yıllık yeni kanser ölümü sayısı	171.6	81.4	120.4
75 yaşından önce kansere bağlı yaşam kaybı riski (%)	18.2	8.6	13.0
5-yıllık aktif vaka sayısı	300 071	281 565	581 636
Melanom dışı cilt kanseri hariç en sık görülen 5 kanser (vakalara göre sıralanmıştır)	Akciğer	Meme	Akciğer
	Prostat	Tiroid	Meme
	Kolorektal	Kolorektal	Kolorektal
	Mesane	Akciğer	Prostat
	Mide	Rahim	Tiroid

The Summary of Cancer Statistics of Turkey can be seen above. It is noteworthy that the incidence of cancer in men is significantly higher than in women. In our country, it has been reported that 1 out of every 4 person before 75 age is diagnosed with cancer.

2.2. What are the causes of cancer?

   Cancer arose from the transformation of normal cells into neoplastic ones in a multistage process that generally progresses from a precancerous lesion to a malignant tumor. These changes are caused by the interaction between a person's genetic factors and three categories of external factors, including the following:

- Carcinogenic physical agents, such as ultraviolet and ionizing radiation;

- Carcinogenic chemical agents, such as asbestos, components of tobacco smoke, aflatoxin (a food pollutant) and arsenic (a pollutant in drinking water);

- Biological carcinogens, such as infections caused by certain viruses, bacteria or parasites.

   The World Health Organization (WHO) classifies carcinogens through its agency for research on cancer, the International Agency for Research on Cancer.

   The incidence of cancer rises sharply with age, likely due to the accumulation of risks of developing certain types of cancer, which increase with age. The accumulation of cancer risk is associated with the tendency for the effectiveness of cell repair mechanisms to decline as a person gets older.

2.3. Existing Solutions for Lung Cancer Diagnostics

- **X-ray:** In an x-ray, low-dose radiation is used to produce images of parts of the body on film. A chest x-ray is often the first imaging test done to check for lung problems. It can reveal spots, tumors, or changes.

- **CT scan:** A computed tomography (CT) scan uses special x-ray machines to produce 3-dimensional and cross-sectional images of the body's organs, tissues, bones and blood vessels. A computer assembles the photos into detailed images. CT allows you to:

  - Reveal the location, size and shape of a lung tumor;
  - Find any thoracic lymph nodes that are larger than normal
  - Find out if the tumor has spread to other parts of the body;
  - **Positron emission tomography:** In a positron emission tomography (PET) scan, a radioactive material called a radiopharmaceutical is used to detect changes in metabolic activity tissues of the body. A computer analyzes patterns of radioactivity distribution and produces 3-dimensional, color images of the region under examination.
  A PET is used to find out how far lung cancer has spread. The camera takes images of the whole body; a PET can find tumors that are larger than 8 mm.

It is more accurate than a CT scan to show if the cancer has spread to the lymph nodes in the middle of the chest (mediastinum) and to other areas of the body. PET is not used to look at the brain because the brain has a high metabolic activity.

- **Biopsy:** During a biopsy, the doctor removes tissues or cells from the body for analysis in the laboratory. It is important to collect enough tissue during the biopsy so that cell and tissue analyzes can be done. The pathologist's report will indicate whether or not there are cancer cells in the sample.

- **Magnetic resonance imaging:** Magnetic resonance imaging (MRI) uses powerful magnetic forces and radio waves to produce cross-sectional images of the body's organs, tissues, bones, and blood vessels. A computer assembles the images into 3-dimensional snapshots.
  MRI scans the brain to see if cancer has spread to it. It can also be used to observe the heart before surgery to check whether cancer has invaded the heart or the large blood vessels and nerves at the top of the lung.
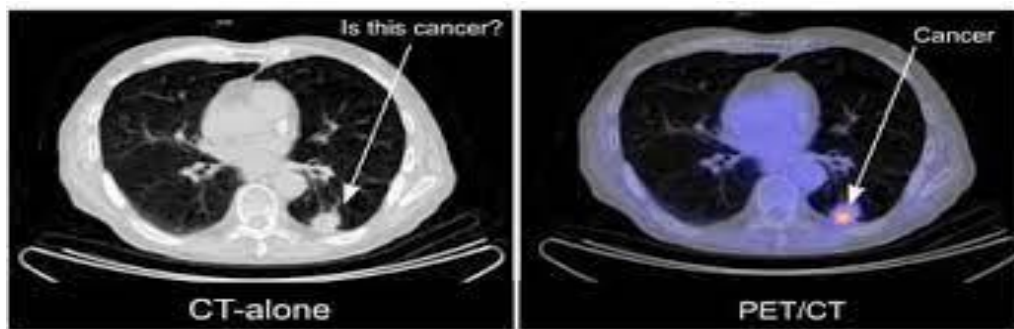
Figure 2. 5.   CT Diagnostics
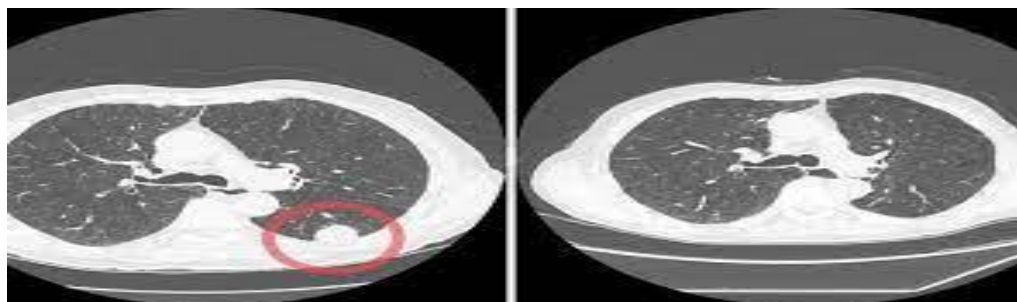
Figure 2. 6.   CT and PET Diagnostics

3.MATERIAL AND METHOD

This section will be to discuss our methods and frameworks which we used in our research, the thesis weights heavily more on the data mining. This section there will be artefacts regarding user analysis and technical design of the application.

3.1. Overview

We are going to develop a web application to prove a proof of concept. The app demand from users to download a CT scan images, then processes the file given and show the pictures to it. Then the user has the choice to choose which scan it wants to portend, after that the app preprocesses the CT scan and infers the image it to the prediction model then displays the output to the user. View Figure 3.1.
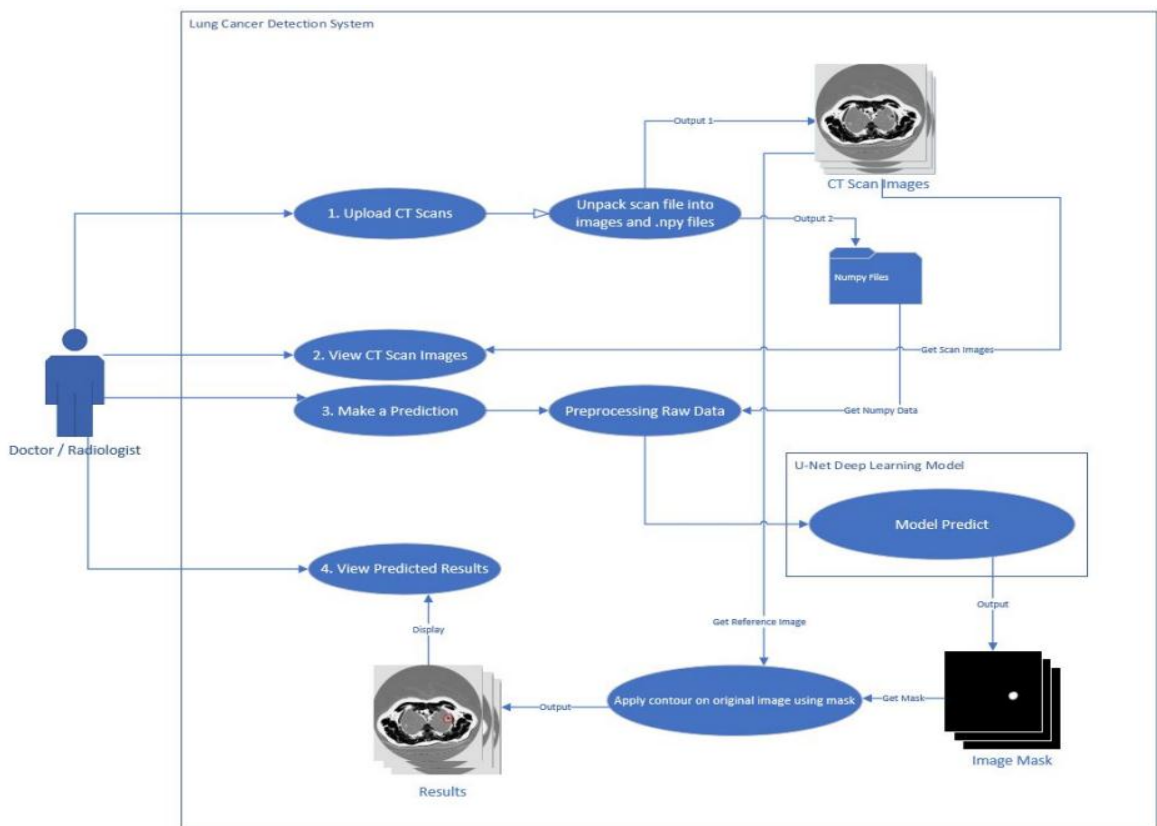
Figure 3. 1. Framework outline

3.2. Uploading CT scan

The user can upload file (.png). If the user uploads the original file (.raw) and metadata file (.mhd) on the back-end system through POST demand. Here comes the role of the system by taking the metadata and uses it to decompress the raw files that contain the pictures, then get the image data and save it in image (.npy) and (.png) using OpenCV and Numpy and OpenCv.
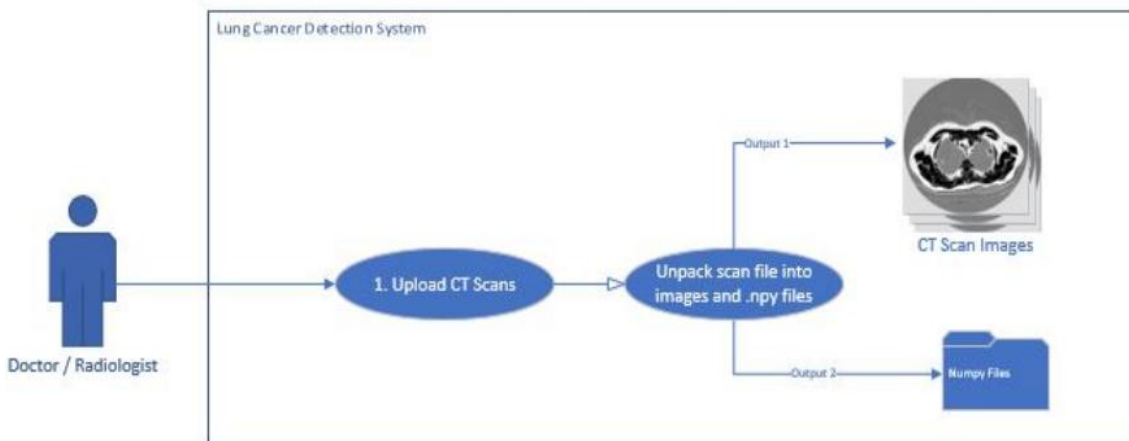
Figure 3. 2.   Uploading CT scans

3.3. Show CT scan

The system, from the back-end extracts the image files gived by user and show them in front end.
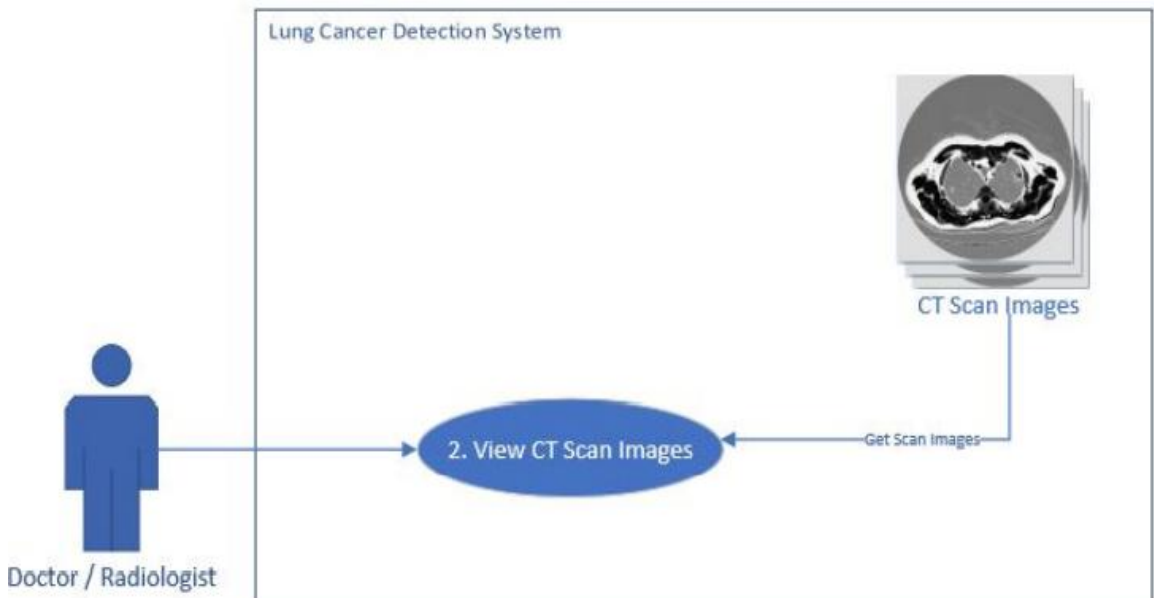
Figure 3. 3.   Show Results

3.4. Create Forecast

When a user selects an image, they want to make predictions. The system gets the user's file name upon selection and uses it to reference a numpy to file. This numpy file is then preprocessed before being fed into the DL model. The model then comes out an image mask. See Figure 3.4.
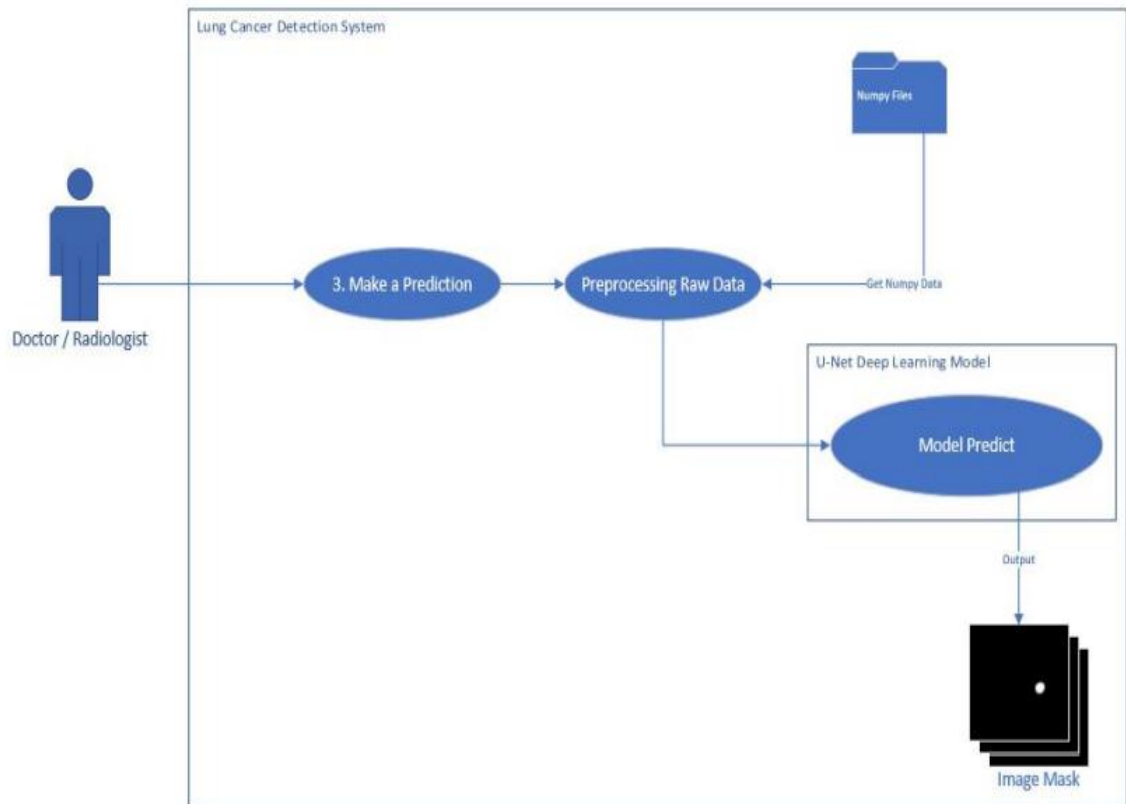
Figure 3. 4. Create Forecast

3.5. Show Forecast

The system uses the original CT scan image and correlated mask and applies the image outline to the original image.
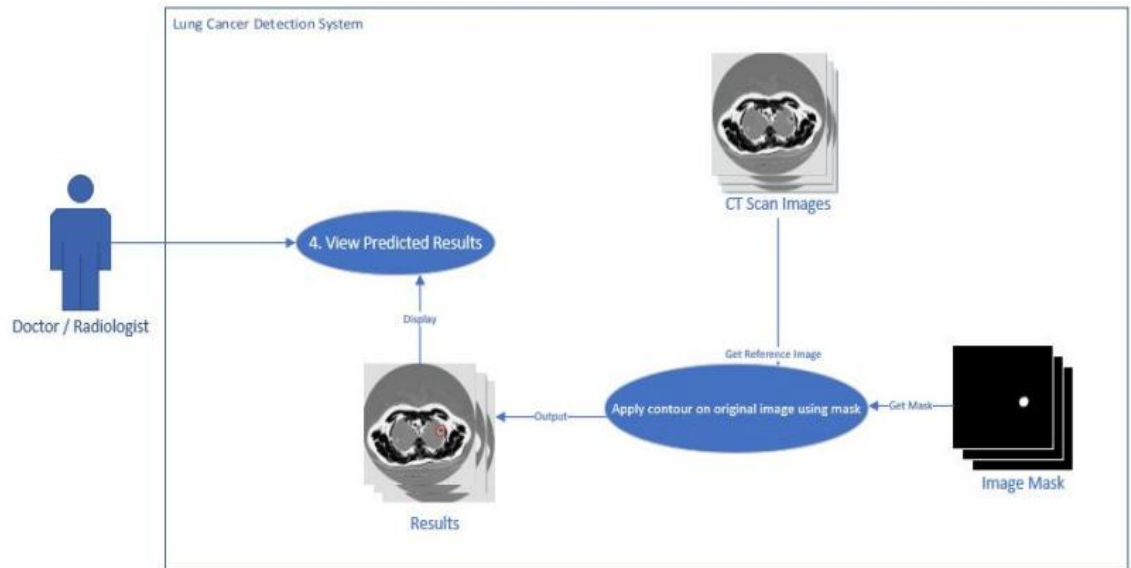
Figure 3. 5.   Show Forecast

3.6. Home Page

Figure 3.6. show the first interface of the web application. These wireframes are used to outline the benefit of the user and to get a general idea of how the application should look and feel prior to development.
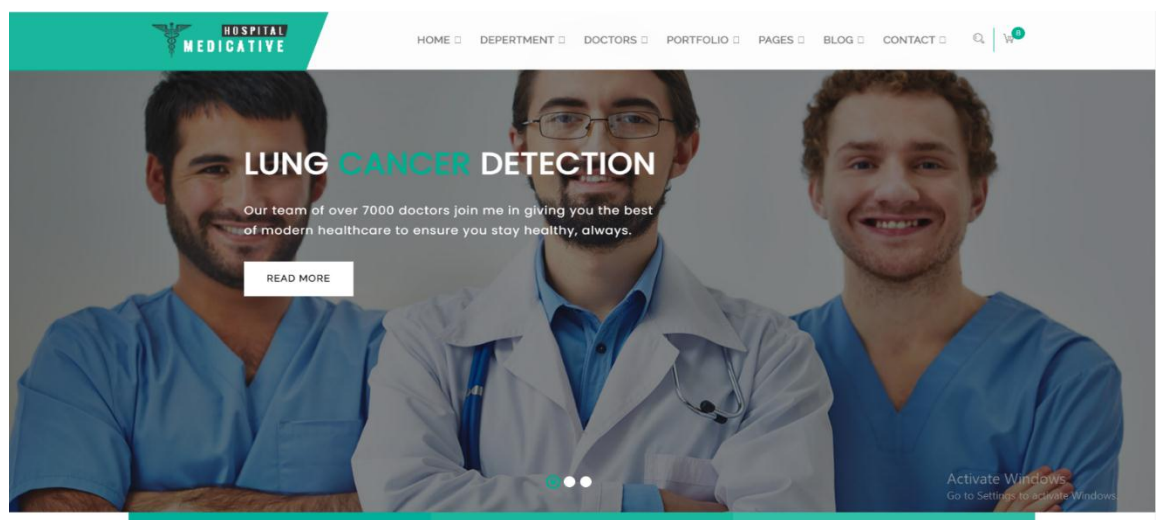
Figure 3. 6.   Web interface

3.7. CT Scan Upload

Figure 3.7. Show the upload page, allows users to upload a CT scan image in his computer, the interface will ask the user to upload pnj or jpg file. The system will capture this data via POST request to the back-end server.
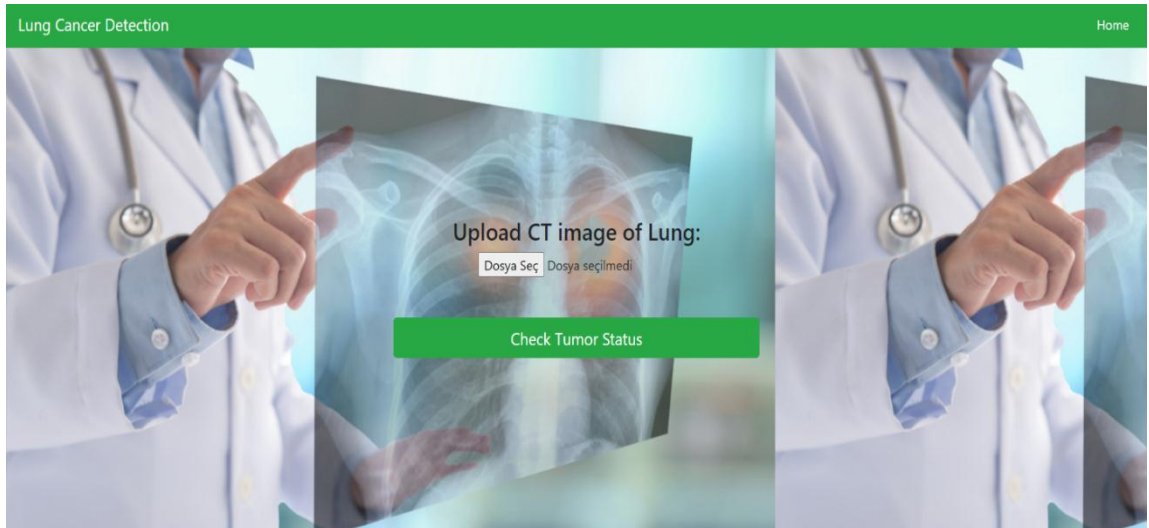
Figure 3. 7. CT Scan Upload

3.8. Prediction Output

Figure 3.8. is the last step, show result of the image gived by the user by telling him if there is a cancer or not.
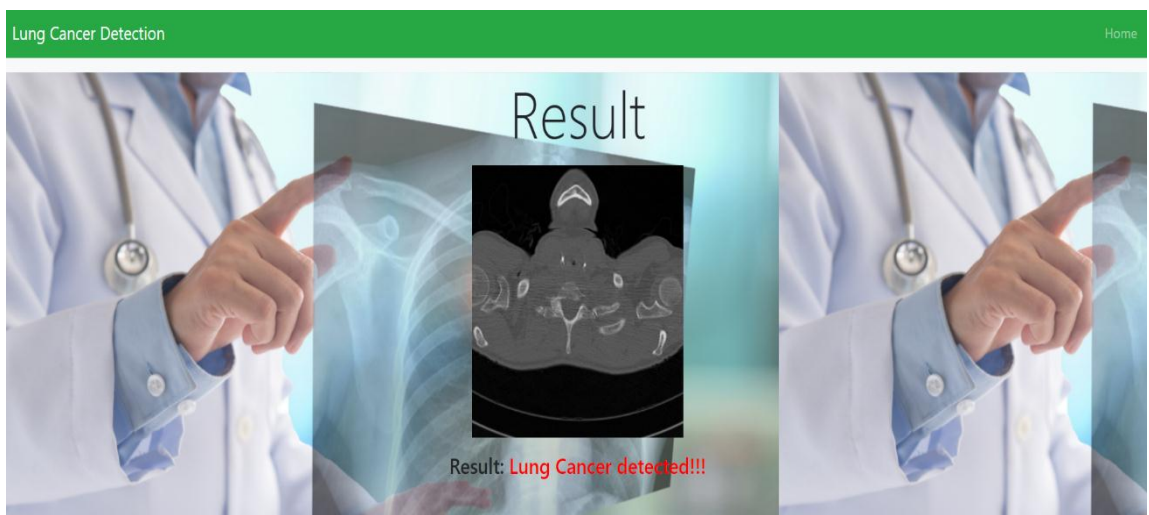
Figure 3. 8. Prediction Output

3.9. Workflow diagram

This section introduces the thesis realization operation, data mining model development process and discussion the result between the models used. The tasks of the thesis are distributed in the following workflow diagram. See figure 3.9.
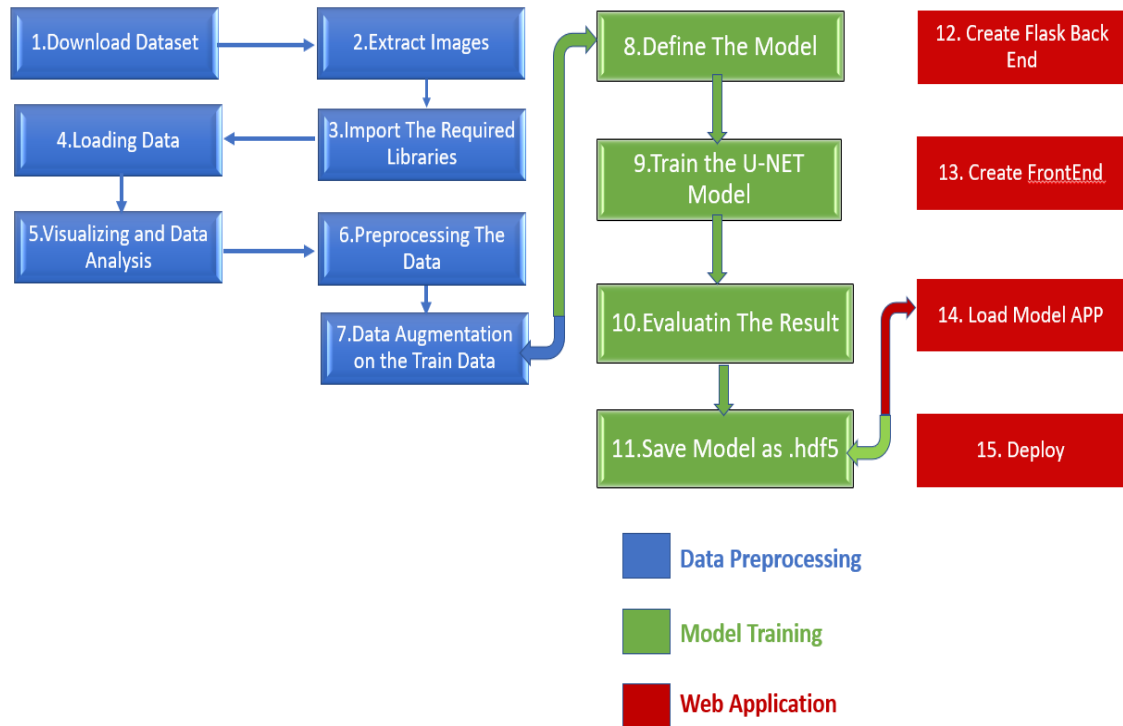
Figure 3. 9.   workflow diagram

3.10. Data Preprocessing

This section discusses various manner to view and understand data sets. Initially, the data used to train the model was obtained from Iraq-Oncology Teaching Hospital / National Center for Cancer Diseases. The dataset was collected over a period of 3 months in the year 2019. The data contains images on CT scans (*.jpg) of patients diagnosed with lung cancer in different stages. The dataset contains a total of 1190 images representing CT scan slices of 110 cases. These cases are grouped into 3 classes: normal, benign and malignant. Of these, 40 cases are diagnosed as malignant; 15 cases diagnosed with benign; and 55 cases classified as normal cases. The CT scans were originally collected in DICOM format. The scanner used is SOMATOM from Siemens.The 110 cases vary in gender, age, educational attainment, area of residence and living status. The models used for this dataset CNN, U-net and LinkNET.

We organize our data in a lung cancer folder that will contain Two folders "Train" & "Test", in each one of them (Train & Test), we will have Three Folders (Benign – Malignant – Normal).

In each folder we have images and we should note that the number of images in the train folder is bigger than the number of images in the test folder. See figure 3.10.
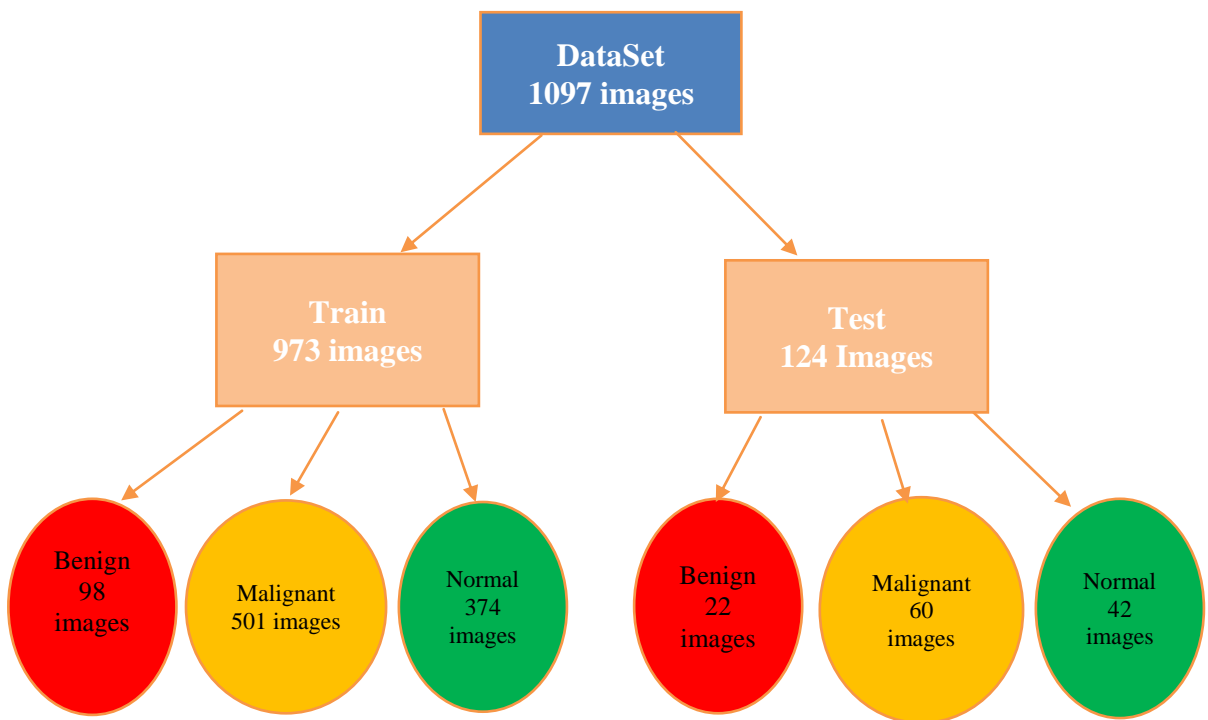
Figure 3. 10.   Setting up our image data

3.11. Models

In this thesis three models are used to detect the lung cancer on CT scan images. Models are CNN, U-Net and LinkNet. The CNN model is used for classification while U-Net and LinkNet Models are used for segmentation. The evalution of models is calculated by some matrices such as confusion matrix, precision, recall, specificity and f1-score. The models are used to get a better result by comparing them and choosing the best one.

3.11.1. Cnn Model

In the proposed procedure, CNNs are applied to detect and classify lung cancer CT scans of the patients collected from hospitals. Convolutional Neural Networks is a sort of deep learning paradigm applied for processing data which has a grid pattern, it is all about using Deep Learning with Computer Vision. A Convolutional Neural Network is a Deep Learning algorithm that can take in image as input, with assigning importance learnable weights and biases to various objects inside this image and be capable of differentiating one from the other. The role of the CNN is for reducing the images to a form that is easier to process but without losing features that are important for getting a good prediction. A typical CNN consists of three types of operation layers: the convolutional layer, the pooling layer, and finally the classifier layer, as exemplified in the figure below.
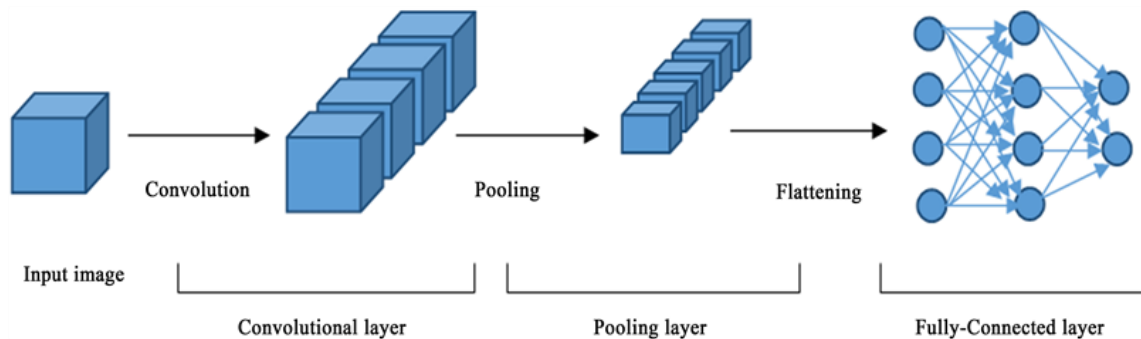
Figure 3. 11. CNN Architecture

3.11.1.1. Discover the issue report
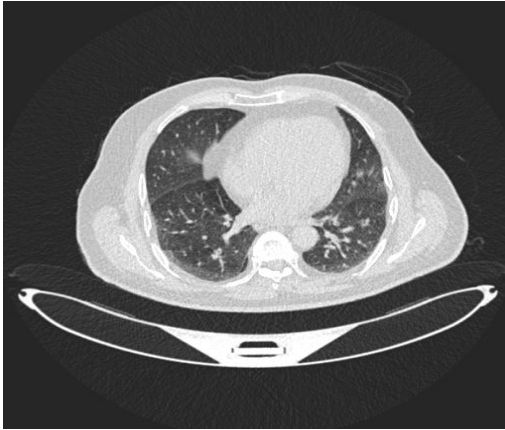
Figure 3. 12.   Random images from the "benign class"

What does this image represent? Doctors and other medically versed people can admit the picture as benign cases of lung cancer. Different aspects of the image may help to identify it as a benign case.

The target is if we have the ability to create a system that can correctly classify images. That is our idea of creating a system that can recognize the type of lung cancer represented in that picture.

3.11.1.2. Visualizing and Data Analysis

The model was developed and trained using python, tensorflow and keras. The use of keras allows the researchers to to fast experiments without witing large number of codes. The dataset was divided into train and test. 70% of the images were used for training and 30% for testing. The images were resized and were trained in batches to facilitate faster training without the nedd for additional computing resourcesWe visualize our dataset and see what is working. By using seaborn to plot the number of images in the three classes, we are able to see what is the output.
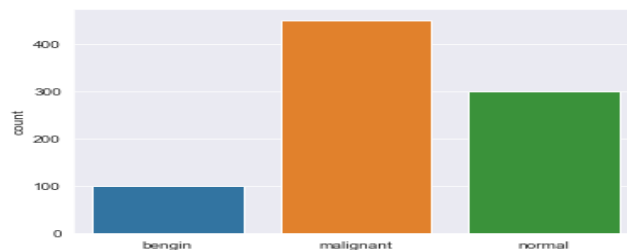
Figure 3. 13.   Visualizing and Data Analysis

3.11.1.3. Train Model

Defining a CNN model, which contains 3 convolutional layers, followed by the maximum pooling layer.

```python
model = Sequential()
model.add(Conv2D(32,3,padding="same", activation="relu", input_shape=(224,224,3)))
model.add(MaxPool2D())

model.add(Conv2D(32, 3, padding="same", activation="relu"))
model.add(MaxPool2D())

model.add(Conv2D(64, 3, padding="same", activation="relu"))
model.add(MaxPool2D())
model.add(Dropout(0.4))

model.add(Flatten())
model.add(Dense(128,activation="relu"))
model.add(Dense(3, activation="softmax"))

model.summary()
```

```
Model: "sequential_1"
_____
Layer (type)                 Output Shape              Param #
=================================================================
conv2d_3 (Conv2D)            (None, 224, 224, 32)      896
_____
max_pooling2d_3 (MaxPooling2 (None, 112, 112, 32)      0
_____
conv2d_4 (Conv2D)            (None, 112, 112, 32)      9248
_____
max_pooling2d_4 (MaxPooling2 (None, 56, 56, 32)        0
_____
conv2d_5 (Conv2D)            (None, 56, 56, 64)        18496
_____
max_pooling2d_5 (MaxPooling2 (None, 28, 28, 64)        0
_____
dropout_1 (Dropout)          (None, 28, 28, 64)        0
_____
flatten_1 (Flatten)          (None, 50176)             0
_____
dense_2 (Dense)              (None, 128)               6422656
_____
dense_3 (Dense)              (None, 3)                 387
=================================================================
Total params: 6,451,683
Trainable params: 6,451,683
Non-trainable params: 0
```

Figure 3. 14.   Training Model

3.11.1.4. Result Evaluation
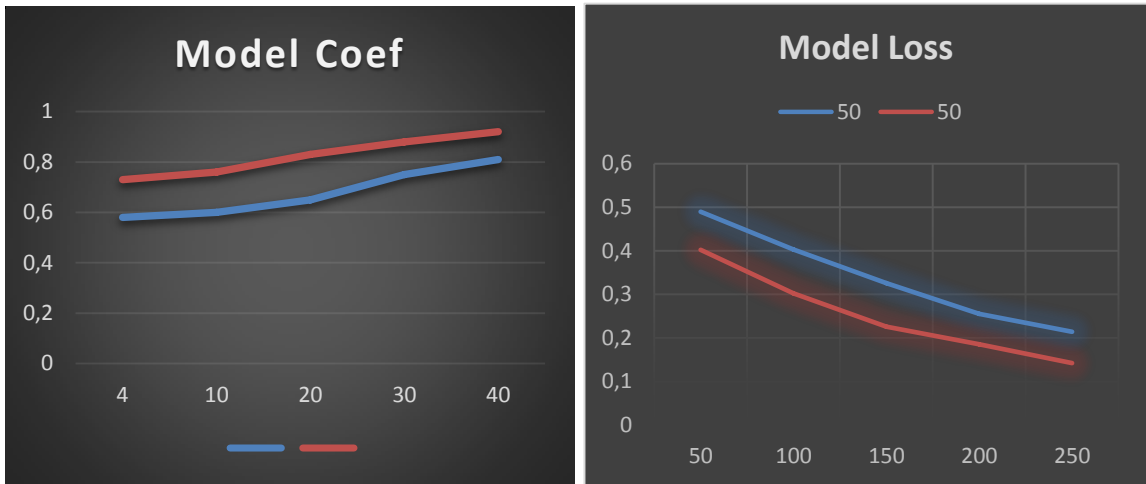
Figure 3. 15. Evaluating the result

Table 3. 1. Model Evaluation

Num	Optimizer	Learning Rate	Epochs	Training Time	Dice Coefficient (Test Accuracy)	Dice Coefficient (Training Accuracy)	Test Loss	Training Loss
1	Adam	0.0001	50	2 hours	0.58	0.73	0.67	0.35
2	Adam	0.0001	100	5 hours	0.63	0.88	0.4	0.29
3	Adam	0.0001	150	7 hours	0.65	0.94	0.29	0.2
4	Adam	0.0001	200	9 hours	0.66	0.98	0.19	0.13
5	Adam	0.0001	250	12 hours	0.66	0.99	0.19	0.12

The Developed Model was CNN. The images were resized and trained with a batch size of 50 at 250 epochs. Figures 3.15. Shows the dice similarity coefficient graph of the model.

The blue line indicates testing dice coefficient graph while the orange line indicates the training dice coefficient graph.

The CNN model gives a training dice coefficient of 0.99 and a testing dice coefficient of 0.66. The test and accuracy loss are very los as seen in Figure 3.15 at around 200 epochs.

Table 3.2. shows models trained up to 250 epoch validation scores. The validation scoring using epoch 250 show wich ranges 66%. the validation score of models trained with epoch 250 is able to produce the best validation score.

Table 3. 2.   Training Score CNN model

Num	Learning Rate	Epochs	IoU	F1-score
1	0.0001	250	65.40%	66%

3.11.1.5. Model Test

We have given a Malignant Picture from the test folder since the machine has recognized the image well.

```
In [18]: img_path="C:/Users/USER/Desktop/FinalProjectLungCancer/lungcancer/test/Malignant/Malignant case (2).jpg"
         img_arr = cv2.imread(img_path)[...,::-1]
         test_image = cv2.resize(img_arr, (img_size, img_size))
         test_image = np.array(test_image).reshape(1, img_size, img_size, 3)

         prediction = model.predict(test_image)
```

```
In [19]: class_predicted=np.argmax(prediction)
         img = mpimg.imread(img_path)
         imgplot = plt.imshow(img)
         if(class_predicted==0):
             print("The Model predicted this image as:Bengin")
         if(class_predicted==1):
             print("The Model predicted this image as:malignant")
         if(class_predicted==2):
             print("The Model predicted this image as:normal")
```

```
The Model predicted this image as:malignant
```
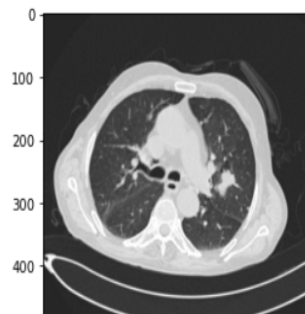

Figure 3. 16.   Testing the model

3.11.2. U-NET Model

UNET is a semantic hash network developed on the basis of a fully convolutional neural network. The network has a total of 23 layers, the number of layers is much lower than other networks while ensuring accuracy, the UNet network is mainly composed of two parts, sampling and sampling: the bottom sampling is also called the feature extraction part, which mainly uses the convolutional and aggregate layer to extract image features entered. This down sampling and sampling structure is also called the decoder structure. In the down sampling part, the input image passes through the convolutional and aggregate layer to obtain feature maps of different levels, and these feature maps contain image features with different levels of abstraction. In the sampling part, a deconvolution layer is used to gradually restore the size of the feature map, and the deconvolution map of the samples is combined to fix the less abstract detailed information missing in the training process and improve the network segmentation accuracy.

However, since the lung is a 3D structure, the UNet network uses 2D convolution and clustering processes to extract features of lung nodules in CT images of the lung, resulting in much loss of spatial information. Thus, much contextual information is lost in the down sampling process. It cannot be fully recovered during sampling, which leads to blurred sampling results and insensitivity to image details.Figure 5.10 shows the structure of the U-Net Architecture.
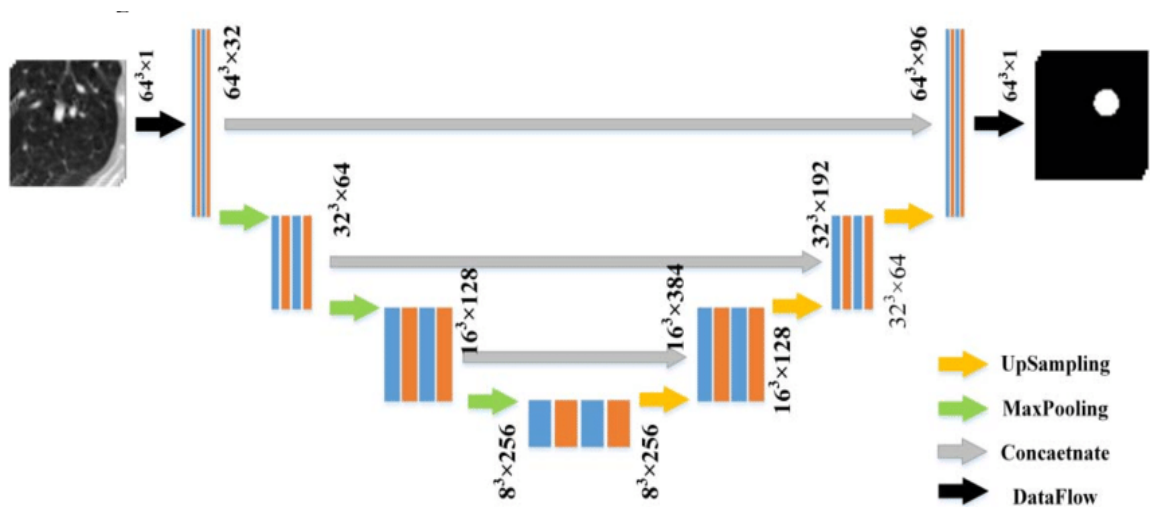
Figure 3. 17.   Unet network structure diagram

3.11.2.1. CT Scan Slices

In this figure we can see a set of pictures for a single CT scan. The dimensions for the CT scan are about 512 height, 512 width.
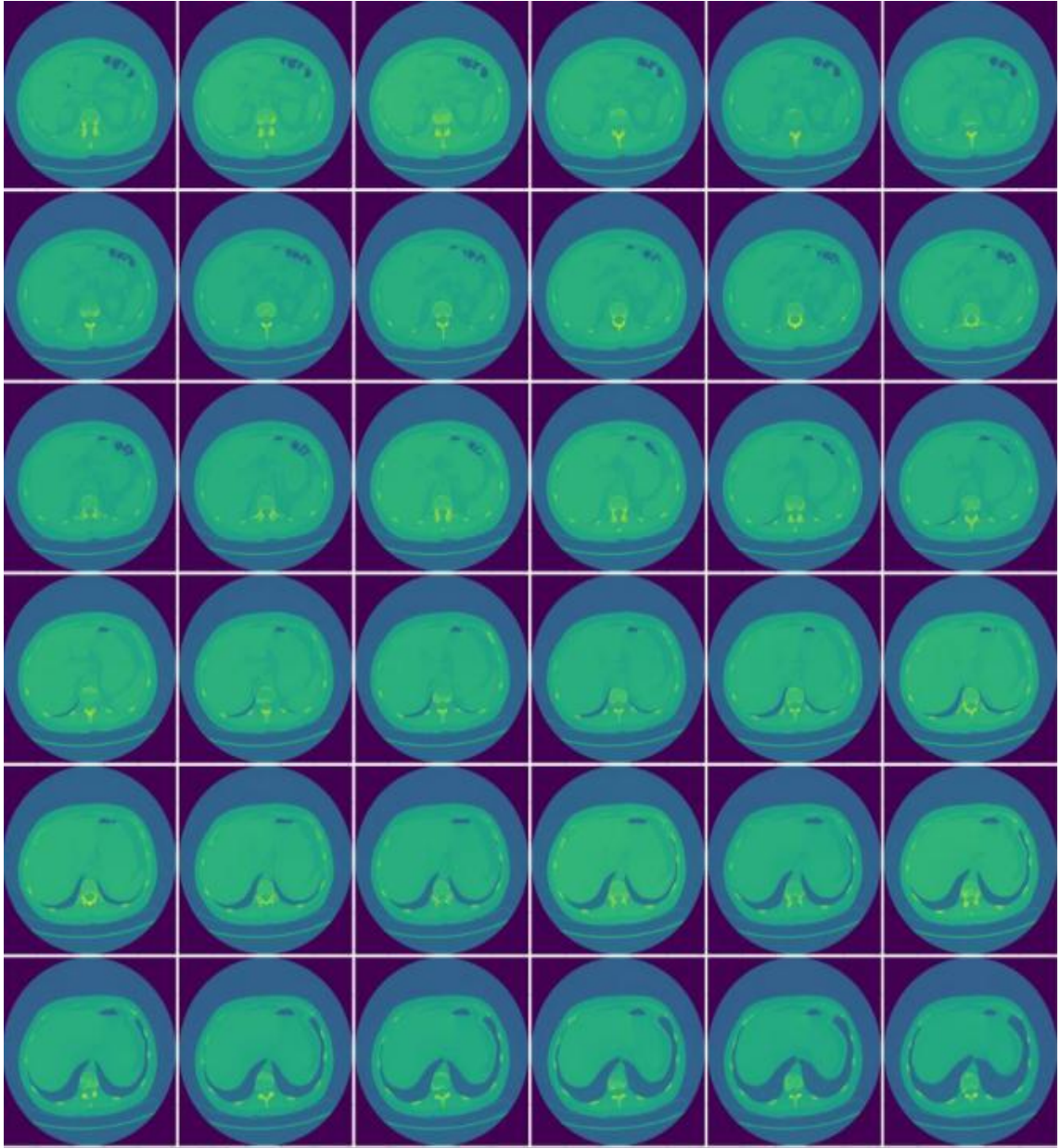
Figure 3. 18. CT Scan Slices

3.11.2.2. Create a Nodule Mask

This figure shows a sample images of cancer masks, the majority of which is small and some are large.
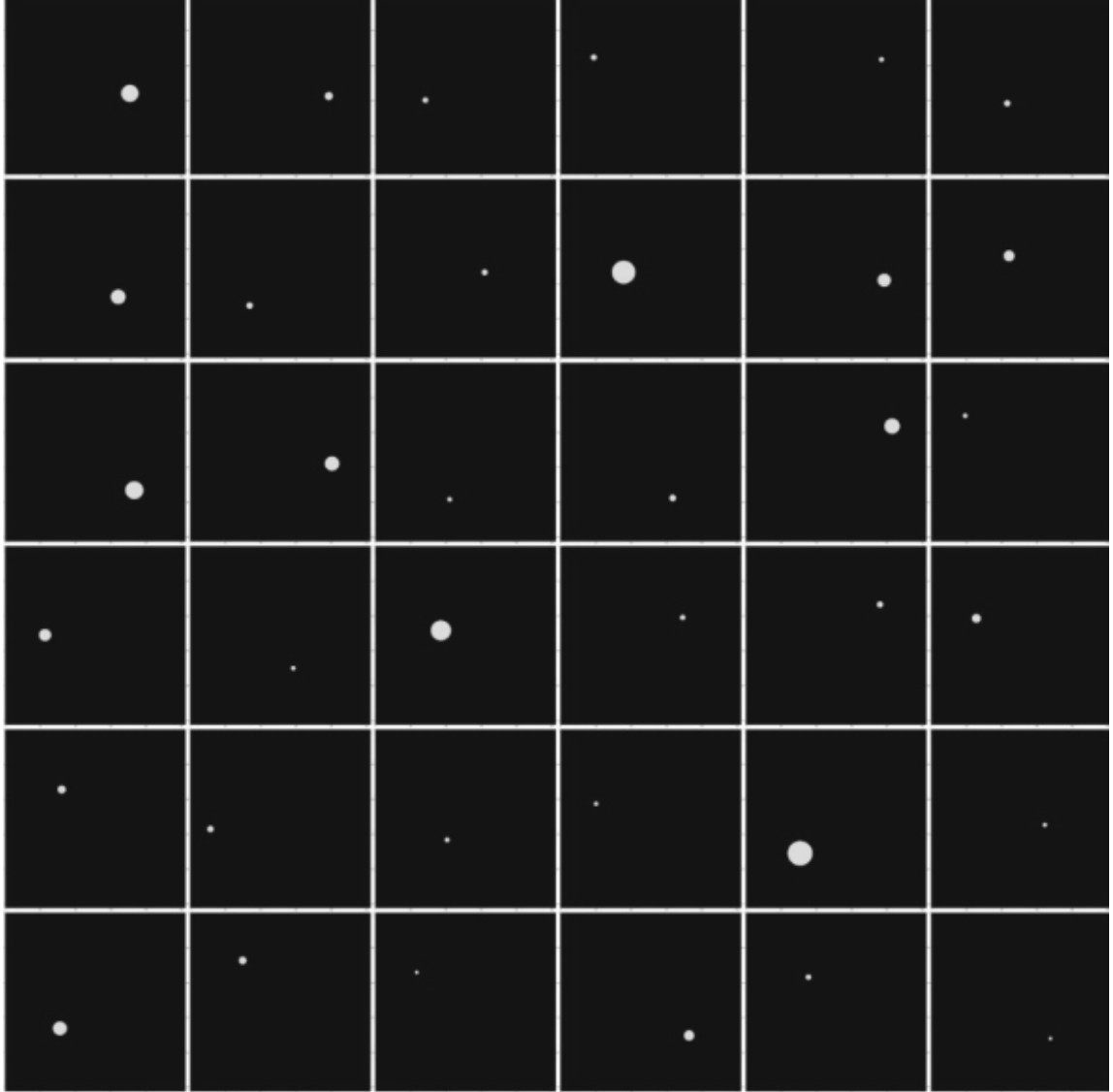
Figure 3. 19.   Nodule Mask

3.11.2.3. Create a Lung Images

The next step is to create our lung images segmented from our original image; in this figure we can see a sample images of segmented lungs with cancer.

Figure 3. 20. Lung Images

3.11.2.4. Training

The U-NET model was created using Keras with a tensorflow backend. The network is considered deep with 10 layers overall. The main features in this neural network are: convolution, maxpooling, merge layers, with relu and sigmoid activation functions. The model also uses an adam optimizer, Learning Rate = 0.0001, a mini-batch parameter of 2 and a dice coefficient loss function. The images were resized and trained with a batch size of 4 at 40 epochs.

```python
model.py                    ×

from tensorflow.keras.layers import Conv2D, BatchNormalization, Activation, MaxPool2D, Conv2DTranspose, Concatenate, Input
from tensorflow.keras.models import Model

def conv_block(input, num_filters):
    x = Conv2D(num_filters, 3, padding="same")(input)
    x = BatchNormalization()(x)
    x = Activation("relu")(x)

    x = Conv2D(num_filters, 3, padding="same")(x)
    x = BatchNormalization()(x)
    x = Activation("relu")(x)

    return x


def encoder_block(input, num_filters):
    x = conv_block(input, num_filters)
    p = MaxPool2D((2, 2))(x)
    return x, p


def decoder_block(input, skip_features, num_filters):
    x = Conv2DTranspose(num_filters, (2, 2), strides=2, padding="same")(input)
    x = Concatenate()([x, skip_features])
    x = conv_block(x, num_filters)
    return x

def build_unet(input_shape):
    inputs = Input(input_shape)

    s1, p1 = encoder_block(inputs, 64)
    s2, p2 = encoder_block(p1, 128)
    s3, p3 = encoder_block(p2, 256)
    s4, p4 = encoder_block(p3, 512)

    b1 = conv_block(p4, 1024)

    d1 = decoder_block(b1, s4, 512)
    d2 = decoder_block(d1, s3, 256)
    d3 = decoder_block(d2, s2, 128)
    d4 = decoder_block(d3, s1, 64)

    outputs = Conv2D(1, 1, padding="same", activation="sigmoid")(d4)

    model = Model(inputs, outputs, name="U-Net")
    return model


if __name__ == "__main__":
    input_shape = (512, 512, 3)
    model = build_unet(input_shape)
    model.summary()
```

Figure 3. 21.   Model Training

3.11.2.5 Training Results
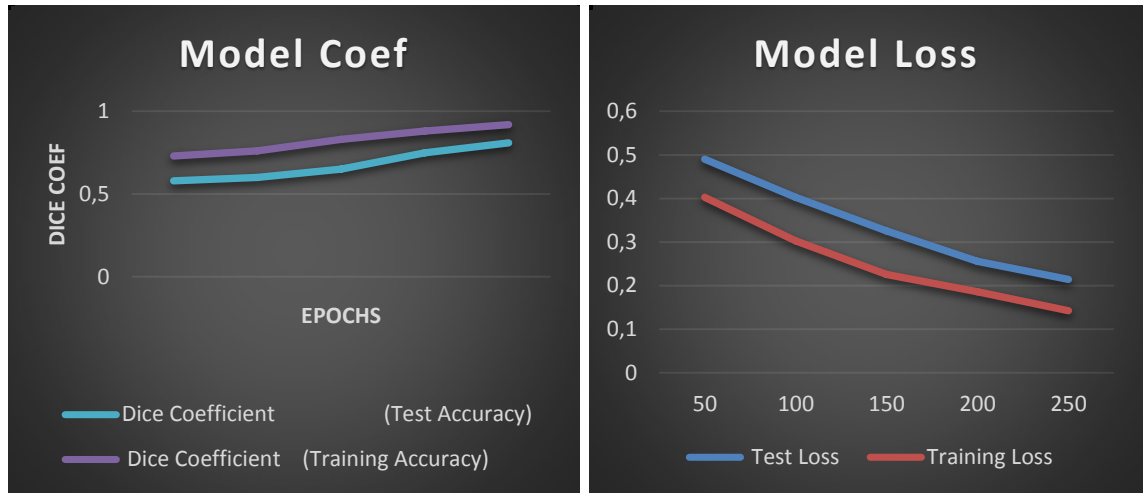
Figure 3. 22.   Training and Testing Dice Coefficient

Table 3.3. shows the results of training the U-Net model with different hyper parameters. Due to the training time of 4 hours for 4 epochs, a cross validation was not performed. The UNET Model converged in 40 epochs and give a dice coefficient of 0.92 for training accuracy while 0.81 for testing accuracy and as we can see in the 40 and 50 epochs the model gives a same result. It can therefore be concluded that epochs 40 is the best value because the epochs value of 50 has been able to produce the best validation score and does not take long in the training process compare it to 50 epochs.

Table 3. 3.   Training and Testing Dice Coefficient

Num	Optimizer	Learning Rate	Epochs	Training Time	Dice Coefficient (Test Accuracy)	Dice Coefficient (Training Accuracy)	Test Loss	Training Loss
1	Adam	0.0001	4	4 hours	0.58	0.73	0.4895453	0.402528
2	Adam	0.0001	10	10 hours	0.6	0.76	0.4028735	0.302504
3	Adam	0.0001	20	20 hours	0.65	0.83	0.3259876	0.225676
4	Adam	0.0001	30	30 hours	0.75	0.88	0.2557933	0.185926
5	Adam	0.0001	40	40 hours	0.81	0.92	0.2142359	0.142352
6	Adam	0.0001	50	50 hours	0.81	0.92	0.2142360	0.142359

Table 3. 4. Training Score UNet model

Num	Learning Rate	Epochs	IoU	F1-score
1	0.0001	40	68.20%	81%

Table 3.4. shows models trained up to 40 epoch validation scores. The validation scoring using epoch 40 show wich ranges 81%. the validation score of models trained with epoch 50 is not different with epoch 40. This indicates that a high epoch value does not necessarily make the validation score of the model. It can therefore be concluded that epochs 40 is the best value because the epochs value of 40 has been able to produce the best validation score and does not take long in the training process.

3.11.2.6. Model Evaluation

Figure 3.23. shows the model predictions beside the label. There are a high number of false positives that appear. The shape of the lungs appears also. This could be because the accuracy is only around 65% however this could be easily removed by rounding the image pixel values in a post processing step to create a mask.
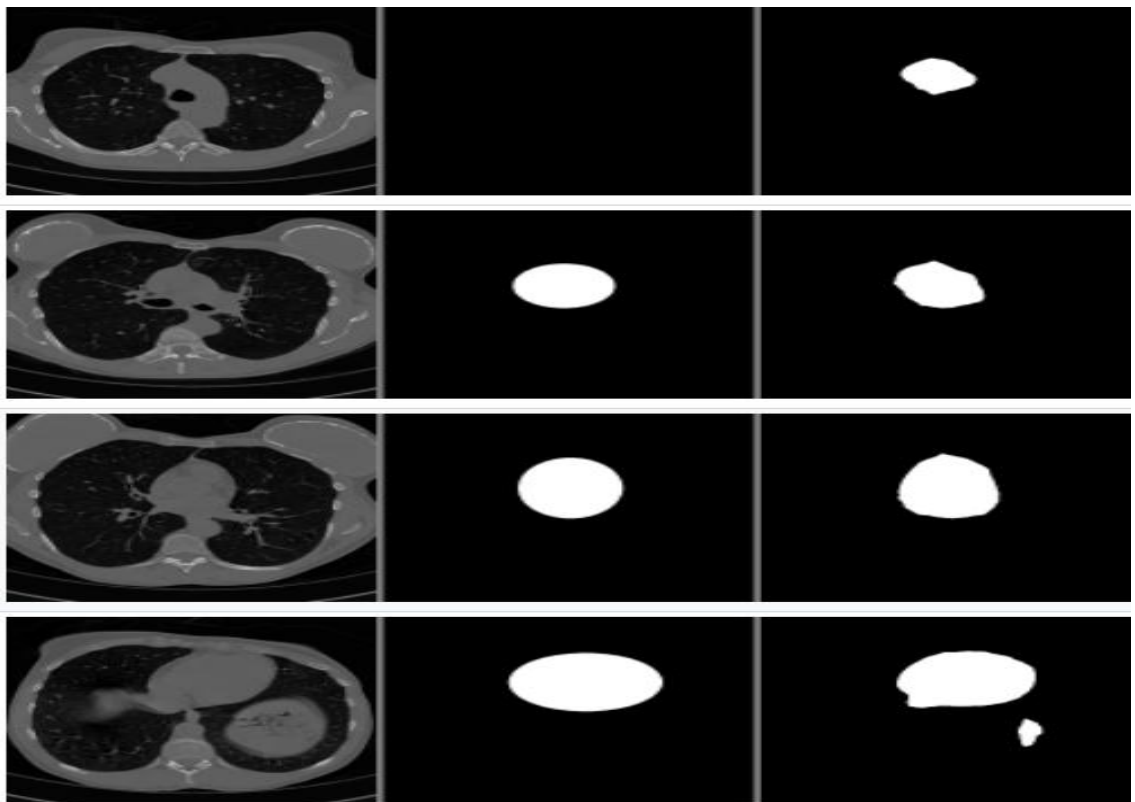
Figure 3. 23. Model Evaluation

3.11.3. LinkNet Model

Similar to UNet, the LinkNet architecture focuses on the use of network efficiency parameters by introducing residual skip connections bypassing features from the encoder to the decoder and implementing feature-sampling down-sampling and corresponding up-sampling. In UNet skip connections, summing does not increase the number of input channels for the subsequent layer, and therefore does not result in the same growth in the number of network parameters as the sequence.
The LinkNet architecture efficiently transfers the information extracted in the encoder block to the corresponding decoder block after sampling down from the image. This makes LinkNET more accurate than a convolutional network or a network with pointers collected in a decoder block. This makes the number of parameters in the decoding block less.
.
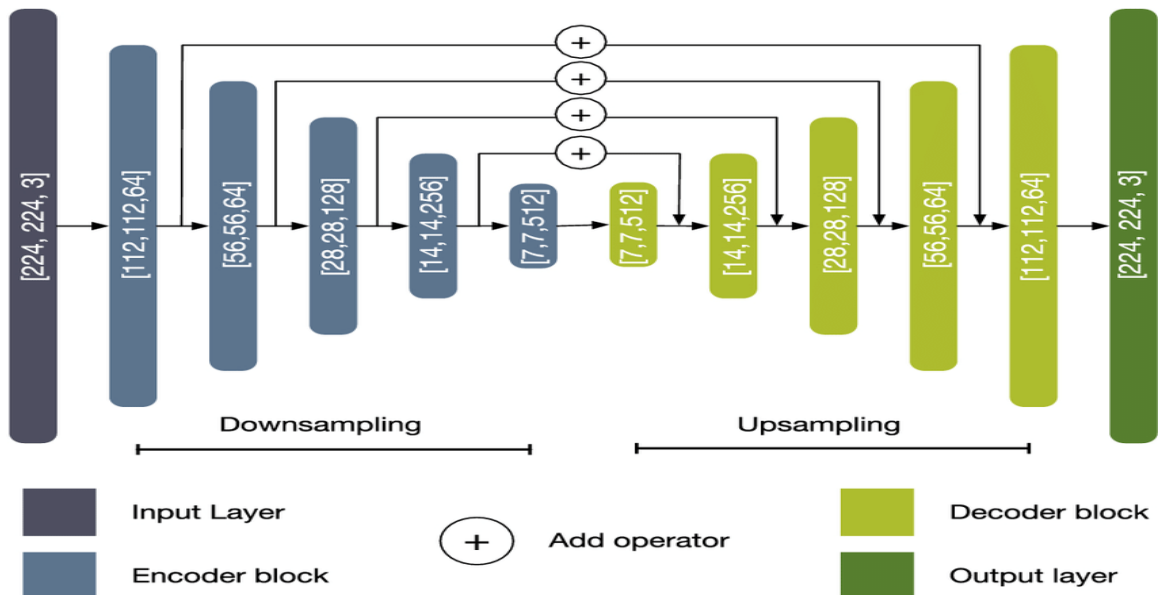
Figure 3. 24.   Architecture of Encoder and Decoder Block in LinkNet

3.11.3.1. Train the Model

We evaluated two different training strategies: full training, in which all the trainable parameters of the networks were liable to be modified during the training phase; and fine tuning, where only the decoder parameters could be modified. In both cases the starting values for the encoder weights were the pre-defined ones obtained by training on the dataset. We used Adam optimization with an initial learning rate of 0.001 and 0.0001. The target function for the optimisation processe was Dice loss. We processed the train and test images by batches of two, allowed a maximum of 500 epochs.

3.11.3.2. Result

Table 3.5. shows that the highest test score is a combination of hyperparameter number seven. The high parameter learning rate is 0.001, the number of epochs is 250, the image size is 256 x 256 x 3, and the pre-trained image gives the best model with the highest IoU score of 69.84% and the highest F1 score of 81.31%.

Table 3. 5.   Result of LinkNet Model based on test data

No	lr	epochs	Image-size	IoU	f1-score
1			128 × 128 × 3	69,54%	81,16%
2		50		65,58%	78,05%
3			256 × 256 × 3	67,61%	79,66%
4				65,87%	78,51%
5			128 × 128 × 3	69,33%	80,77%
6	0,001	250		64,01%	76,88%
7			256 × 256 × 3	**69,84%**	**81,31%**
8				66,46%	78,93%
9			128 × 128 × 3	68,46%	80,29%
10		500		62,03%	75,15%
11			256 × 256 × 3	69,55%	81,11%
12				65,84%	78,44%
13			128 × 128 × 3	69,37%	81,06%
14		50		56,48%	70,52%
15			256 × 256 × 3	69,68%	81,20%
16				60,03%	73,62%
17			128 × 128 × 3	69,08%	80,84%
18	0,0001	250		59,15%	72,95%
19			256 × 256 × 3	69,04%	80,65%
20				58,34%	72,09%
21			128 × 128 × 3	68,89%	80,66%
22		500		59,86%	73,71%
23			256 × 256 × 3	69,83%	81,26%
24				58,55%	71,99%

The point values of the 24 experimental hyperparameter value combinations in Table 3.5. will be aggregated based on the same 0.001 and 0.0001 learning rate values. Table 3.6. shows the standard deviation value and mean f1 score for each learning rate. Based on the mean scores and standard deviation in Table 3.6, it can be seen that the learning rate 0.001 produces a mean f1 score higher than the learning rate 0.0001, i.e. 79.19% and results in a smaller standard deviation value of 0.0183.

Table 3. 6.   Average and standard deviation IoU score

Learning rate	IoU		$f1 - score$	
	Average	Std. Dev.	Average	Std. Dev.
0,001	**0,6701**	**0,02372**	**0,79188**	**0,01827**
0,0001	0,64025	0,05361	0,76713	0,04307

The effect of the number of periods is tested by aggregating test scores based on the number of periods, then a training chart of the model with the best test result is shown for each period. Figure 3.25. shows the model training graph that contains the best result for each epoch value.
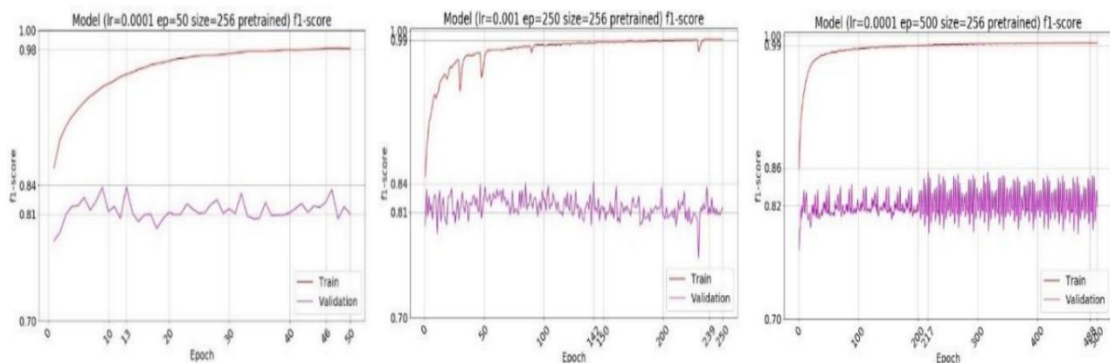
Figure 3. 25.   Training Score with Epoch 50,250, and 500

Figure 3.25. shows that models trained up to 500 epoch validation points tend to survive and show no increase in validation scores. The validation results using the 500 epochs also show the same results as the 50th and 250th epochs which range 81%. Based on the learning plots with different epoch values, it can be known that the validation score of models trained on epoch 50 is not much different from epoch 250 or 500. This indicates that a higher value of epoch does not necessarily make model validation score. Even after the epoch value touches the number 200, the movement of the verification score chart is not as good as the time of the first 200 epochs.

So, it can be concluded that 50 epochs is the best value because 50 epochs value was able to produce the best validation result and doesn't take much time in the training process.

3.11.4. Evaluation metrics

Intersection over union is one of the rating scales used to discover or divide objects. Intersection across union is used to measure similarity and diversity in a data set. In general, the cross-union intercept is calculated using Equation 1 with T being the ground truth and P being the forecast value. The F1 score, otherwise known as the dice coefficient, is one of the rating scales for object detection or hash and crossover over union. In general, the score of f1 or the score of the dice can be calculated using Equation 2 with T being the ground truth and P being the prediction value..

$$IoU = \frac{Intersection}{Union} = \frac{|T \cap P|}{|T \cup P|}$$

$$F1 - score\ = \frac{2\ x\ Intersection}{Union + Intersection} = \frac{2|T \cap P|}{|T| + |P|}$$

The relationship between the F1-score and loss coefficient is as follows:

$$Loss = 1 - F1\ score$$

4.RESULTS AND DISCUSION

In this study we proposed 3 methods to compare, UNET, CNN and LinkNET method. The accuracy of the proposed method is quantitatively evaluated using performance measures such as the dice similarity coefficient. Table 4.1 shows a comparison between the proposed methods.

Model 1 CNN: The images were resized and trained with a batch size of 50 at 250 epochs, The learning rates used in this model are 0. 0001.The best result for this model were 66% for F1-score using 0.0001 learnng rate and 50 epochs.

Model 2 UNET: The validation result is shown using the 40 and 0.0001 epoch learning rate ranging from 81%. The validation score of models trained in a 50 epoch is not different from a 40 epoch. This indicates that a higher epoch value does not necessarily lead to a model validation score. So it can be concluded that epoch 40 is the best value because epochs value 40 was able to produce the best validation result and doesn't take much time in the training process.

LinkNet 3 model: This model is made to determine the hyperparameter value that produces the highest score. The value of the hyperparameter in question is the value of the learning rate, the number of periods, the size of the image, and whether or not it uses a pre-trained model. The learning rates used in this model are 0.001 and 0.0001. The value of the eras used is 50 epochs. The ability of the model is evaluated using the IoU score and the F1 score. The best result for this model was 81.31% for a F1 score using a learning rate of 0.001 and an epoch of 50.

The Model 1 get a F1-score of 66% for 250 epochs while model 2 get 81% for 40 epochs and model 3 , 81.31% for 50 epochs, this means that the first model is very bad compared to the other models , so we can use model 2 or 3 since they have similar results.

Table 4. 1. Comparison and Discussion

MODELS	IoU	Dice coefficient (F1-score)	Accuraccy
CNN Model	65%	66%	99%
UNET Model	68%	81%	92%
LinkNet Model	69%	81,31%	92%

5. CONCLUSION

Our models are doing a great job in the training phase with a maximum of 99% and 92% accuracy for CNN and UNET models, unlike the validation phase with a maximum of 66% and 81% accuracy, this is called Overfitting which is refers to the fact that the predictive model produced by the deep learning algorithm adapts well to the Training Set, not that well in the validation (or test set)., by using an evaluation method, we were able with ML to detect the difference between them.

Concerning its accuracy, the model continued to produce many false positives. This means that is requires more regulation and training time, even though it has been trained for 40 hours. I believe with more adjusting hyperparameters and learning the model, the accuracy can be improved.

6. REFERENCES

Acharya, S., Poudel, S., Dangol, S., & Subedi, S. COMPUTER VISION BASED TRAFFIC RULE BREACH DETECTION.

Asuntha, A., & Srinivasan, A. (2020). Deep learning for lung Cancer detection and classification. *Multimedia Tools and Applications*, *79*(11), 7731-7762.

Alakwaa, W., Nassef, M., & Badr, A. (2017). Lung cancer detection and classification with 3D convolutional neural network (3D-CNN). *Lung Cancer*, *8*(8), 409.

Bianconi, F., Fravolini, M. L., Pizzoli, S., Palumbo, I., Minestrini, M., Rondini, M., ... & Palumbo, B. (2021). Comparative evaluation of conventional and deep learning methods for semi-automated segmentation of pulmonary nodules on CT. *Quantitative Imaging in Medicine and Surgery*, *11*(7), 3286.

Chon, A., Balachandar, N., & Lu, P. (2017). Deep convolutional neural networks for lung cancer detection. *Standford University*, 1-9.

Dreyfus, G., Martinez, J. M., Samuelides, M., Gordon, M. B., Badran, F., Thiria, S., & Hérault, L. (2002). *Réseaux de neurones* (Vol. 39). Paris: Eyrolles

Ding, Z., Zhang, W., & Gu, K. (2019, December). Robust 3D Convolutional Neural Networks For Pulmonary Nodule Detection. In *2019 IEEE International Conference on Robotics and Biomimetics (ROBIO)* (pp. 2157-2162). IEEE.

Ferlay, J., Colombet, M., Soerjomataram, I., Mathers, C., Parkin, D. M., Piñeros, M., ... & Bray, F. (2019). Estimating the global cancer incidence and mortality in

2018: GLOBOCAN sources and methods. *International journal of cancer*, *144*(8), 1941-1953.

Gu, D., Liu, G., & Xue, Z. (2021). On the performance of lung nodule detection, segmentation and classification. *Computerized Medical Imaging and Graphics*, *89*, 101886.

Kapoor, A. (2020). Deep learning vs. machine learning: a simple explanation. *source: https://hackernoon. com/deep-learning-vs-machine-learning-a-simple-explanation-47405b3eef08*.

Lau, M. M., & Lim, K. H. (2018, December). Review of adaptive activation function in deep neural network. In *2018 IEEE-EMBS Conference on Biomedical Engineering and Sciences (IECBES)* (pp. 686-690). IEEE.

National Academies of Sciences, Engineering, and Medicine. (2017). *Implementation of lung cancer screening: proceedings of a workshop*. National Academies Press.

Rivero Rodríguez, C. (2019). *Low Dose Tomography with a C-Arm system* (Bachelor's thesis)

Senthil Kumar, K., Venkatalakshmi, K., & Karthikeyan, K. (2019). Lung cancer detection using image segmentation by means of various evolutionary algorithms. *Computational and mathematical methods in medicine*, *2019*.

Singh, J., Tripathy, A., Garg, P., & Kumar, A. (2020). Lung tuberculosis detection using anti-aliased convolutional networks. *Procedia Computer Science*, *173*, 281-290.

Tacchino, F., Macchiavello, C., Gerace, D., & Bajoni, D. (2019). An artificial neuron implemented on an actual quantum processor. *npj Quantum Information*, *5*(1),

Thakur, S. K., Singh, D. P., & Choudhary, J. (2020). Lung cancer identification: a review on detection and classification. *Cancer and Metastasis Reviews*, *39*(3), 989-998

Tekade, R., & Rajeswari, K. (2018, August). Lung cancer detection and classification using deep learning. In *2018 Fourth International Conference on Computing Communication Control and Automation (ICCUBEA)* (pp. 1-5). IEEE.

CURRICULUM VITAE

ASMA HAIBELTY

Esma_hb@yahoo.com
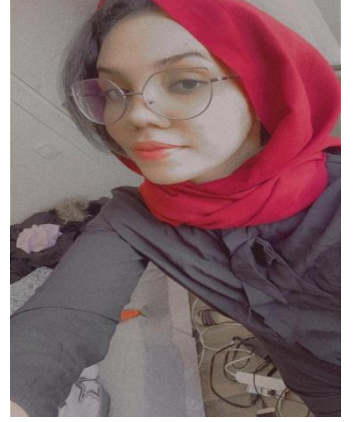
ÖĞRENİM BİLGİLERİ

Yüksek Lisans 2019-2022	Akdeniz University Institute of natural and applied sciences Department of computer engineering Antalya/Turkey
Lisans 2015-2018	Nuakshot al-asriye University Mathematics and Computer Science Nuakshot/Moritanya

MESLEKİ VE İDARİ GÖREVLER

Internship in computer science 2016 - 2017	SMCP Nouakchott, Mauritania
Website and IT administration 2017-2018	SMCP Nouakchott, Mauritania