Segmentation based Kidney Tumor Classification using Deep Neural Network

by

Md Humaion Kabir Mehedi
17201061
Ehteshamul Haque
18101481
Sameen Yasir Radin
17221003
Md. Abrar Ur Rahman
18101276

A thesis submitted to the Department of Computer Science and Engineering
in partial fulfillment of the requirements for the degree of
B.Sc. in Computer Science and Engineering

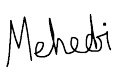
Department of Computer Science and Engineering
Brac University
January, 2022

Declaration

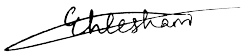
It is hereby declared that

1. The thesis submitted is our own original work while completing degree at Brac University.
2. The thesis does not contain material previously published or written by a third party, except where this is appropriately cited through full and accurate referencing.
3. The thesis does not contain material which has been accepted, or submitted, for any other degree or diploma at a university or other institution.
4. We have acknowledged all main sources of help.

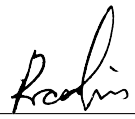
Student's Full Name & Signature:



Md Humaion Kabir Mehedi
17201061



Ehteshamul Haque
18101481



Sameen Yasir Radin
17221003

Md. Abrar Ur Rahman
18101276

Approval

The thesis titled "Segmentation based Kidney Tumor Classification using Deep Neural Network" submitted by

1. Md Humaion Kabir Mehedi (17201061)
2. Ehteshamul Haque (18101481)
3. Sameen Yasir Radin (17221003)
4. Md. Abrar Ur Rahman (18101276)

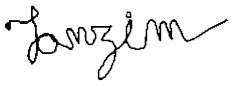
Of Fall, 2021 has been accepted as satisfactory in partial fulfillment of the requirement for the degree of B.Sc. in Computer Science and Engineering on Januray 18, 2022.

Examining Committee:

Supervisor:
(Member)

Dr. Md. Golam Rabiul Alam
Associate Professor
Department of Computer Science and Engineering
Brac University

Co-Supervisor:
(member)



Md Tanzim Reza
Lecturer
Department of Computer Science and Engineering
Brac University

Head of Department:
(Chair)

Sadia Hamid Kazi
Chairperson and Associate Professor
Department of Computer Science and Engineering
Brac University

Abstract

Kidney disease is one of many severe chronic disease that a person can have. Early detection of this disease can be pivotal for proper treatment. Different neural networks have proven to be useful in disease prediction in the progression of modern science. In this paper, we have proposed a segmentation based kidney tumor classification using Deep Neural Network (DNN). We have done our work in two Steps. Firstly, we have segmented kidneys using a manual segmentation technique and trained UNet along with SegNet for kidney segmentation. Then, for the classification task, the modified MobileNetV2, VGG16 and InceptionV3 was trained on the segmented kidney data. CT KIDNEY DATASET: Normal-Cyst-Tumor and Stone dataset(published in Kaggle) was used to train our models. Finally, the classification models MobileNetV2, VGG16, InceptionV3 scored with 95.29%, 99.21% and 97.38% accuracy on test set. We found that the modified VGG16 model has the best accuracy and the highest sensitivity and specificity.

Keywords: Kidney Tumor, Computed Tomography (CT), VGG16, Segmentation, Classification, Deep Neural Network (DNN).

Dedication

It is our genuine gratefulness and warmest regard that we dedicate this work to all our loved ones for their love and inspiration.

Acknowledgement

Firstly, all praise to the Great Allah for whom our thesis have been completed without any major interruption.
Secondly, to our supervisor Dr. Md. Golam Rabiul Alam sir for his kind support and advice in our work. He helped us whenever we needed help.
Thirdly, we would like to convey our gratitude to our co-supervisor Mr.Tanzim Reza for his guidance and handful contribution throughout the whole phase of our thesis work. From the very beginning to the end of the work they have provided us with all kinds of help and inspired us to move forward to our goal.
A special thanks to Md Nazmul Islam and Mehedi Hasan for their dataset and helping in data annotation process.
And finally to our parents without their throughout support it may not be possible. With their kind support and prayer we are now on the verge of our graduation.

Table of Contents

List of Figures

List of Tables

Nomenclature

The next list describes several symbols & abbreviation that will be later used within the body of the document

$Adam$ Adaptive Moment estimation

$AHDCNN$ Adaptive hybridized Deep Convolutional Neural Network

AI Artificial Intelligence

AUC Area Under the Curve

$BiLSTM$ Bidirectional Long Short-Term Memory

CAD Computer Aided Design

CKD Chronic Kidney Disease

CKD Chronic Kidney Disease

CNN Convolutional Neural Network

$ConvNet$ Convolutional Neural Network

CT Computed Tomography

DL Deep learning

$DLNN$ Deep Learning Neural Network

DNN Deep Neural Network

$eGFR$ Estimated Glomerular Filtration Rate

EHR Electronic Health Record

FC Fully connected layer

FPN Feature Pyramid Network

FPR False Positive Rate

$IoMT$ Internet Of Medial Things

IoU Intersection over Union

$KICH$ Kidney Chromophobe

ML Machine Learning

MRI Magnetic Resonance Imaging

NN Neural Network

$PACS$ Picture Archiving and Communication Systems

RCC Renal Cell Carcinoma

$RCNN$ Region Based Convolutional Neural Networks

$ReLU$ The rectified linear activation function

ROI Region of Interest

$ROSNET$ Robust one-stage network

$SCNN$ Segmentation and Classification Convolutional Neural Network

SGD Stochastic Gradient Descent

SNP Single Nucleotide Polymorphisms

SVM Support Vector Machine

TPR True Positive Rate

$VGG16$ Very Deep Convolutional Networks for Large-Scale Image Recognition-16

WHO World Health Organization

Chapter 1

Introduction

Human body due to the complex nature of its functionality, performs hundreds of chemical reactions every second. These reactions happen in the individual cell of the body creating toxins and other wastes which are then discarded into the bloodstream. Kidney is responsible for filtering all the impurities, toxins and extra amount of fluids present in the blood. The pure blood then flows in the main blood system and the waste products are discharged from the body through urination. Prolonged exposure towards unhealthy habits such as nicotine, drug and alcohol abuse increases the chances of kidney disease by many folds. Long term disease such as diabetics and high blood pressure account for many new cases of kidney diseases. Also, people with kidney problems in near relatives are more likely to suffer from the same problem. Considering the nature of the problem, kidney diseases are categorized into 5 parts:

- **Chronic Kidney Diseases (CKD):** It is a long term disease that does not improve over time or treatment. Main reasons causing CKD are high blood pressure and diabetes. Kidney transplant and dialysis prove to be very effective in treating this line of disease [1].
- **Kidney Stones:** The function of a kidney is to filter the blood of impurities. Just as many other materials minerals are also present in the bloodstream. When these minerals crystallize, it is called kidney stones.
- **Glomerulonephritis:** It is a small but very important part of the kidney. It is mainly responsible for the filtration. Glomerulonephritis occurs due to infections, drug abuse and congenital abnormalities which occur during birth.
- **Polycystic Kidney Disease:** It is mainly a genetic disorder which is responsible for the growth of many cysts in the kidney. Not only can it be painful but it can cause kidney failure.
- **Urinary Tract Infections:** These are mainly bacterial infections which can attack any part of the urinary system. They are easily treatable and with proper treatment, do not cause much harm or pain.

Kidney tumor is a very common chronic kidney [2] disease. Not all tumors that form on the kidney are fatal. Tumors found forming on kidneys are sectioned into three: benign, indolent, malignant. A benign tumor is not cancerous; indolent ones

are cancerous but not proliferative; malignant ones are cancerous and proliferative. Several types of kidney cancers include:

- **Renal Cell Carcinoma (RCC):** The most common type of adult kidney cancer comprising 85% of all the cases diagnosed. These grow in proximal renal tubes. They make the kidney filtration system.
- **Urothelial Carcinoma:** It accounts for 5% to 10% of the kidney cancers that forms in the renal pelvis. This is also knows as bladder cancer.
- **Sarcoma:** It forms on the capsule of the kidney which is the soft tissue cover the kidney. It is rare and is treated with surgery.
- **Wilms Tumor:** Consisting of only 1% of the diagnosed kidney cancers it is commonly found in children. Radiation combined with surgery prove to be effective in treating this.
- **Lymphoma:** Enlarges the lymph nodes across the body such as necks, armpits, chest along with enlarged kidneys. The swollen lymph nodes need to be biopsied for the confirmation of cancer.

Renal cancer and related kidney illnesses claim many people's lives each year. A report was published in 2017 [3] which estimated the overall number of deaths related to kidney diseases worldwide from 1990 to 2017. According to the report, more than 130 thousand people died annually on average worldwide. The report also showed that the estimated number of deaths in 2017 was around 10 million. It was almost 17 percent of all deaths in that year. According to the study, Uruguay had the highest cancer age-standardized death rates, while Bangladesh had one of the lowest. Another report published by WHO [4] stated that in 2018, Bangladesh had about 17,000 deaths while the neighboring countries India and Pakistan had mortality of over 250 thousand and 28 thousand. Western Countries such as the United States and United Kingdom had over 66 thousand and 4 thousand deaths. So, we can see that kidney diseases are one of the major global problems, which causes many people's deaths and affects millions of lives worldwide.

1.1 Problem Description

Diagnosing kidney problems is a time-consuming procedure. Before using ultrasound (UT), Magnetic Resonance Imaging (MRI), or CT (Computed Tomography) scans, diagnosis comprised of urine and blood tests. The kidney is responsible for maintaining multiple hormones such as [5] creatinine, albumin, eGFR (estimated glomerular filtration rate) etc. If these hormones are produced at an abnormal level, they should be detected when a urine test is done. But many factors can alter the level of accuracy that is needed for a kidney disease to be diagnosed. Hence, the next reliable test is applied to the patient which is a blood test. Depending on which testing procedure is conducted on the blood sample, it is necessary for the donor patient to fast for at least 12 hours before the blood sample is provided or to adhere to specific limitations concerning food and medication [6]. The more sophisticated the procedure is used to diagnose kidney problems,the greater the chances of human error. To increase the level of accuracy, image-based approaches were taken

for proper diagnosis of kidney diseases which also involved different methods such as UT, MRI, and CT scan.
The first step is a CT scan, performed in a clinic. The report takes time to print. After the patient receives the CT scan report, the doctor and radiologist work together to diagnose the disease. Sometimes, even though it depends on the situation, the total process takes days worth valuable time. The patient's health deteriorates by the day. In certain situations, he or she endures agonizing pain due to the kidney's failure to operate properly, which is an important element in performing any physiological function. In addition to the lengthy process, sometimes kidney tumors are misdiagnosed or recently diagnosed. About 48% of the patients are diagnosed with the disease after it has reached an advanced stage [7]. In this paper, we will detect kidney tumors from CT images which will be cost efficient, time saving and also yields more accuracy.

1.2 Research Objectives

This work aims to use artificial intelligence (AI) to detect the kidney tumor without the supervision of any human eye involved. The accelerated process would reduce the total time needed starting from getting a CT report to getting proper treatment.

The goal of this research is listed below:

- To better understand CT reports and how DNN works.
- To deeply understand how DNN can be applied to segment kidney and classify kidney tumor from CT images.
- To evaluate the applied models.
- To provide recommendations for enhancing the models' performance.

1.3 Thesis Overview and Orientation

In the following segments, we have divided the work we have conducted into chapters and discussed our whole work process in detail.In, chapter 2 we have discussed the previous related research on neural networks and kidney disease prediction done by other researchers and how they have implemented their research in their studies. Then, in chapter 3, we have talked in depth about the model we have proposed to conduct our research. Following that, we have shown our system overview and data preprocessing methods. The chapter also includes our segmentation methods and gives an overview of the CNN models. After that, we have discussed the results and discussion from our research in chapter 4. This chapter also gives us the output for the various training models and graphs from the training. Finally, in chapter 5, we concluded our work and discussed our plans for future work.

Chapter 2

Literature Review

In terms of mortality, cancer is one of the worst illnesses known to humanity. The detection of cancer plays a huge role in not only curing the illness but also to take necessary precautions beforehand. Medical research has come a long way in cancer and lesion diagnosis in recent years. In their research, Yan et al. (2018) [8] discuss how images from the deep lesion dataset can be used to advance research in various sectors. A triplet network has been discussed to model the lesions according to type, location, size. Their dataset consists of 32K lesions. To simulate the close connection in type, position, and size, the lesion embeddings are taught using a triplet network. The framework can also have various uses, including search engine, classifier, and matching tools. After further annotation and processing, the datasets can be used for future research and innovations.

In the paper, Lung et al.(2021) [9] have talked about how to detect lesions from computed tomography scans Using ROSNET. Automatic lesion detection utilizing CT scans has become popular in recent years. However, the vast amount of inconsistent data creates problems during lesion detection. Scale variance, unbalanced data, and prediction stability are examples of these issues. The authors of this work recommended that the ROSNET paradigm be implemented to overcome these issues. The experiments are conducted on a vast dataset showing that the ROSNET outperforms the competition by 3.95 percent at MICCAI 2019. The proposed technique includes a three-part one-stage detector: a nested structural feature pyramid, a realistic data-sensitive re-weighting module, and a shift-variant down-sampling strategy . The layered structure feature pyramid can provide more information about lesions at different scales. To address the problem of unstable prediction in CNN-based models, they developed a detector that can obtain strong semantic information at all levels.

Yan et al.(2018)[10] have discussed how to use a 3D context-based CNN to take data from 3D images generated from 2D images by creating an enhanced region-based CNN named as 3DCE. They have developed an algorithm to find the lesions with one framework. After generating the 3D image data and crossing out the errors, the algorithm can be used for future studies. For their research, the deep lesion dataset has been used, consisting of 32K 2D annotations of various types of lesions. A universal lesion detection algorithm has been developed to find lesions of all sorts with one unified framework. 3DCE is memory-friendly, comprehensive, and effort-

less to use. When used as an initial screening tool, it can transmit the detection result to other systems, which may then be utilized to identify other sorts of lesions.

In their study, Alnazer et al.(2020)[11] have assessed all the techniques and methods to calculate and predict CKD by using different Diagnostic tools like MRI, Ultrasound and CT scans. After discussing the various techniques and tools, they have talked about how AI can be used in Renal Segmentation. Their research has first summarized the various medical imaging modalities to check for CKD. After that, they have shown AI ability to lead renal function assessment from classification to disease diagnosis. Different types of Segmentation methods like, DeepMedic, ScaleNet, VNet and HighRes3dNet etc. can be used by researchers to predict CKD and advance the current knowledge of modern medicine.

Yang et al.(2019)[12] published a study in which they discuss the use of SCNN for the assessment of kidney tumors. After creating the network structure and Pyramid parsing module, it is possible to employ a two way segmentation technique to find out more data. After calculating the Loss function, different sorts of algorithms have been used for classification, such as VGG16, ResNet and so on.

Ren et al.(2018)[13] have talked about predicting multiple diseases using electronic health records. There are typically two parts in an EHR. These are the physical signs as well as the textual description. The authors of the paper have proposed using a neural network model to investigate the topic of predicting renal disease in hypertensive patients. At first, the prediction problem is modelled as a job of binary classification. After that, a hybrid NN can be used to fully capture EHR information.BiLSTM and Autoencoder networks have been used for the aim of predicting kidney disease too.

The data collected from the studies conducted till now and the advancements in technology has paved the way for people of any level to research on their own and present their findings. Such as, Hadjiyski (2020)[14] has specified in his works on how to stage kidney cancer from CT scans using AI and neural networks. By using the cropped images from the CT scan and establishing a method for using DLNN(Deep Learning Neural Network), he has proposed that in the future, doctors will be able to use these techniques and research for easier kidney cancer staging.

Kidney cancer detection and neural networks have been associated together for a while now, and much research has shown that by using these methods, people can detect cancer and therefore prevent it much earlier than in the past. In their research, Chen et al. (2020)[15] have proposed using an AHDCNN for the early detection of kidney disease. With their proposed way of using the IoMT platform, one can use DCNN for advanced detection. To detect renal cancer, which is considered one of the most dangerous and aggressive sorts of cancer, researchers have started using different imaging techniques.

Shehata et al. (2021)[16] have proposed a CAD system that uses the textures from the CT scan of patients, and then with the system, they try to classify tumors into sub-types. Their proposed system has three steps: take the grey area from the CT

images, extract the discriminating features and use a two-stage classification system. Their purpose is to create a high accuracy CAD system in order to Diagnose kidney diseases and prevent further advancement of undetected Kidney Cancer. Their method has showed accuracy on the collected dataset, but they need to test it on a bigger range of tumors to improve it further.

Many research in medical science use algorithms on CT images for lesion and cancer detection. Zhang et al.(2019)[17] have attempted to address the issue of low accuracy and inadequate detection methods in the detection of kidney lesions using CT images. They suggest utilizing a cascaded CNN model to detect kidney abnormalities in CT scans with precision and reliability. Two kinds of morphological operations were used in their research. To generate varied levels of feature maps for location and categorization, a six layer FPN was developed. Lastly, a four IoU(Intersection over Union) threshold cascade RCNN was created to achieve high precision detection. Further study and wider testing should be provided in order to overcome the model's limitation on polycystic kidneys. Furthermore, the model has a number of difficulties with intricate lesions that must be resolved before it can be employed in current detection.

Covid is one of the most dangerous outbreaks in human history. The recent pandemic has caused distress all over the world. Researchers have been working to figure out the cause of the disease, the pattern of its transmission, and how to combat it since the beginning. One of the critical signs of the covid infection is an attack on the lungs. Saood and Hatem(2021)[18] have researched this topic where they talked about a few possible ways to detect infected lung tissue using SegNet and U-NET from Lung CT images. This research can be used for Covid research.but can also be useful for other sectors of research.

Liu et al.(2021)[19]have discussed using artificial intelligence to detect tumors from CT images. AI and CNN are used for categorization of images. Their research proposes an idea of CT Image segmentation which uses Artificial Intelligent medical equipment. Their study collected the data from 120 patients which were divided as Grade A(58 cases) and Grade B(62 cases). The CT scan was found to be more sensitive to liver metastases in their study. Hepatocellular carcinoma is less common than hepatocellular carcinoma, and the KMC algorithm outperforms the RG technique.

In their paper, Skalski and Jakubowski(2016)[20]have discussed a kidney segmentation method For CT scans from kidney cancer patient data using elliptical shape limitations, based on Hybrid Level Set approach. Using RUSBoost and decision trees, they have created a classification of renal regions using an automated method. The effectiveness of their technique in the Dice coefficient was about 0.85 ± 0.04. The proposed categorization model has a 92.1 percent overall accuracy. Their research tries to tackle the challenges faced from kidney segmentation like imbalanced classes, features value range, number of observations etc. They have stated that it is the first solution that allows the division of the kidney into multiple sections including cortex, column, medulla and pelvis.

Again, Muller and Kramer(2021)[21] have proposed using the MIScnn to create a rapid and easy way to set up CNN and deep learning models. The adequate pre-operative plan based on image diagnostic data is the major issue of imaging approaches. In their paper, they have stated that with the use of MIScnn, it may be possible to construct a complete pipeline for preprocessing, information augmentation, patch slicing, and batch creation which can be trained on medical image data. Their research has shown that using the MIScnn, an open-source python library, people can successfully cross-validate kidney image segmentation. Researchers can use it to develop a complete medical image segmentation pipeline with just few lines of code. Thus further research on the topic can create a better system for imaging in medical science.

Many researchers have used deep learning-based methods for medical image segmentation in recent years. In their paper, Mu et al.(2019)[22] have discussed using multi-resolution 3D V-net networks to segment kidney and renal tumors in CT scans. In particular, they have proposed to adopt two resolutions and offer a customized V-Net model for both resolutions dubbed VB-Net. Their strategy reduces the overall load on the GPU and with a high accuracy rate for kidney segmentation in disease detection. Their proposed model can localize the organs in the CT images and accurately define the boundary of each organ. They have used collected the CT scans for training and testing from the KiTS Challenge Dataset for their research. After processing the data, VB-Net was used for accurate organ segmentation. Although their results were positive, there were still problems related to tumors or cysts with uneven densities. To solve the problem, they optimized the post-processing algorithm to make corrections.

In their paper, Marsousi et al. (2014)[23] proposed automated kidney segmentation using Ultrasound imaging of Morison's pouch in three dimensions. They have used a kidney segmentation approach based on form. For their research, they created a probabilistic kidney shape model using previously acquired data. The primary goal of the algorithm they utilized was to find a kidney shape in the binarized model. The primary concept of kidney segmentation is to start with the PKSM. The model voxels on a 3-D grid into a scalar value. They used a quick preprocessing to create binarized volume for the next part. After that, they used PKSM to search for 3D correlation for match position. The probabilistic kidney shape model was created using four manually segmented kidneys. The results from their experiments confirm that it is superior compared to MRF-AC.

Yan et al.(2018)[24]have talked about how to use the massive amount of medical data to create a large-scale image lesion dataset. Their study yielded a dataset of 32,735 lesions on 32,120 CT slices from 10,594 studies involving 4,427 different individuals. Deep lesion has also been utilized to train a universal lesion detector. This allows them to detect a wide range of lesions using a single framework. Using bookmarks in PACS which are already tagged by radiologists, they highlighted crucial picture results that may be used as a reference for future investigations. In their paper, they have created a paradigm to harvest the lesion annotations.

Using a median filter, Obaid. J Ahmed(2020) [25] discusses a kidney tumor de-

tection approach that is very effective. After the filtration, the images are made to be mostly noise-free. Following that, the images are segmented using the K means, clustering them for segmenting data and dividing them into similar classes based on similarity. Principal Component Analysis is further taken, which narrows down the huge number of groups to find the most optimal ones. The images then go through Probabilistic Neural Networks that create output images that detect the tumor. The images used for the experiment are obtained from the tumor dataset. After the testing is completed, the parameters based on the confusion matrix are used to acquire information on the tumor's functionality. After this, the performance matrix is analyzed based on precision, recall, accuracy and sensitivity.

In their paper, Dziekiewicz and Markiewicz(2018)[26] have talked about a novel technique for detecting kidneys on CT images. They have used shape analysis of the object's boundary and figured out the geometrical coefficients for the final detection stage. The accuracy of detecting the renal border is around 84 percent. This study employs Canny's edge identification algorithm, which is based on co-efficient geometric form computation. In addition, their method also has the advantage of identifying other organs. Their method enables a fast and effective way to detect neoplastic changes in kidneys. Their paper talked about Kidney recognition & brightness correction, image segmentation, calculation of geometric coefficients, and complete kidney algorithm recognition. For their testing, they have used 59 CT images collected from a verified source.

Distorted and noise filled images can create problems while conducting research. Many times, the image quality hinders the whole result of the research. To tackle this problem, researchers have tried to develop various ways to reduce the noise and create better quality images using a method like CNN. In their paper, Yu et al.(2019)[27] have talked about a cascading trainable segmentation model known as crossbar-net to complete two goals. The first goal is to create two orthogonal non-squared patches that can acquire both global and local appearance knowledge. The second goal is to train two sub-models in a cascading matter. In the end, the models can complement each other to create self-improvement. They have validated their research with data collected from 94 patients with 3,500 CT images. They have shown with their research that Crossbar-Net has a range of use, even in cardiac and breast mass segmentation.

Yan et al.(2019) [28] have talked about a solution for radiologists to read and annotate medical images deep learning modules. They began by developing a deep learning module to extract relevant semantic components from radiology reports of lesion images. They then proposed utilizing LesaNet, which is built on a multilabel CNN, to learn all of the labels. On LesaNet, they've also added a simple score propagation layer. According to their tests, LesaNet can annotate lesions with an AUC of 0.9344 utilizing an ontology of 171 fine-grained labels. In their paper, they have tried to find a solution for the problem of radiologists. A radiologist could identify and annotate various lesions and other abnormalities from CT scan images using their research. Their paper's main contribution is proposing an automatic learning framework with the least effort required. This algorithm can text-mine relevant labels, Presenting LesaNet, Leveraging Ontology-based medical info to incorporate

label relations in LesaNet.

In their paper, Osowska-Kurczab et al. (2020)[29] have talked about developing a new way of recognition of different variants of renal cancer based on CT scans.To achieve this, they have used two methods of investigation. The first employs texture descriptors to determine diagnostic features, which are then combined with support vector machines for recognition and classification. The second method employs deep learning for various CNN configurations. From the datasets, their method was able to reach 90 percent accuracy. They have offered a complete method for constructing an automatic medical image analysis system as a result of their research. The results of their study could help the studies in medical practice. Implementation of their research can create such a system that improves overall diagnosis process, ultimately leading to better treatment and care. However, the limited variety of tumor cases in the dataset, as well as the tiny size of the lesion region and inter-patient variability, can cause problems with accurate results. Their method can be further improved on by using CNN models to get better information from the images.

Aljouie and Patel(2018) [30] have talked about predicting kidney function with cross-validation and cross-study validation. They gathered data from 110 cases and controls of KICH for kidney chromophobe cell carcinoma for their kidney cancer research. They next did quality checks on the SNPs they had collected and graded them accordingly. Their top-rated SNPs showed support vector machines to get an accuracy of 71%(with 10) and 72% (with 20 SNPs). Their findings indicated that a kidney papillary dataset might predict kidney chromophobe carcinoma better than a random classification. Two kidney cancer datasets were also retrieved from the NCI GDC dataset as part of their study. Their research also shows that using a linear SVM, and SNPs can be used to predict cases and controls of kidney chromophobe carcinoma with a 66 percent accuracy.

In their writing, Simonyan and Zisserman(2015)[31] have talked about large scale image recognition and the impact of the CNN model on image identification accuracy on a huge scale. In their study, they describe how they evaluated increasing-depth networks using a convolution filter architecture with small (3x3) convolution filters. The finding was based on their ImageNet Challenge of 2014. During the testing phase, the CNN was applied to the entire image, therefore there was no need for multiple crops. Simultaneously, using a huge set of crops could improve accuracy. In comparison to a fully convolutional net, the outcome was a finer sampling of the input image. Their application was based on the C++ Caffe toolbox, but they also modified it. Their paper has shown that their representations could be used in datasets as well. For future research, their two best-performing ConvNet models have been made publicly available.

To overcome the limitations of earlier research and make additional progress in the relevant fields, we describe how to employ segmentation approaches to create a more accurate and clearer model for gathering data. This solves the probability of getting wrong information from distorted images. Later on, the segmented image can be used by the classification models to classify the kidney tumors accurately.

Chapter 3

Proposed Model

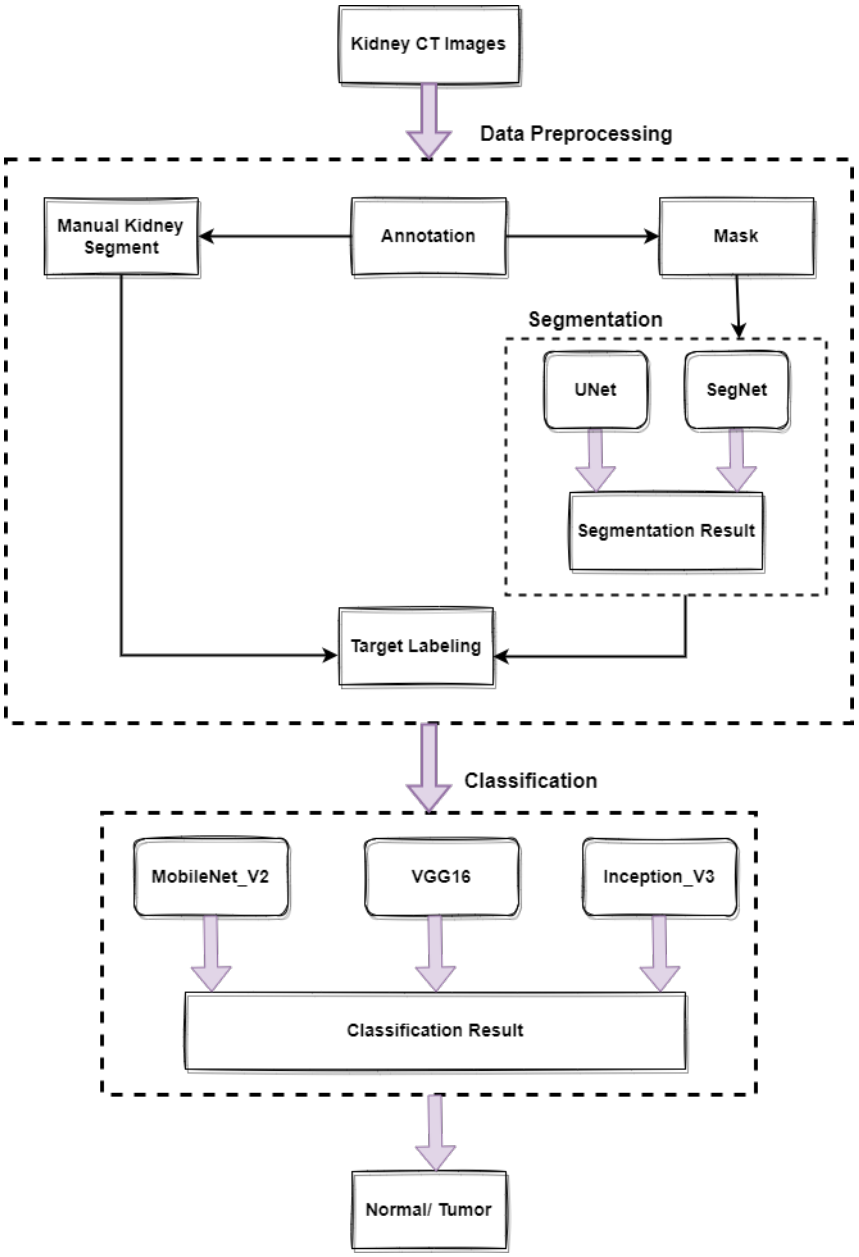


Figure 3.1: System Overview.

This section illustrates how we have used our segmentation based tumor classification model to classify kidney tumors. The above diagram 3.1 shows all the steps that we have used to classified kidney using both segmentation and classification techniques. Firstly, the specified data are gathered from the CT KIDNEY DATASET: "Normal-Cyst-Tumor and Stone" from Kaggle. CT images are typically acquired as jpg images. Then kidneys are annotated from the collected dataset and are fed as the input of the SegNet and UNet for segmentation. After that, the predicted and manually segmented images are fed into the MobileNetV2, VGG16 and InceptionV3 for classification in which random kidney images are used as test subject to get the desired result i.e. whether the kidney has a tumor or not.

3.1 Data Collection

We took our dataset from "CT KIDNEY DATASET: Normal-Cyst-Tumor and Stone" from Kaggle [32]. It contains four labels with images. It contains 12,446 unique data, within which 3,709 are cyst, 5,077 are normal, 1,377 are stones, and 2,283 are tumors, which is shown in the figure 3.2. For our research, we have taken only normal and tumor data from the dataset. All the images in our dataset are in jpg format.

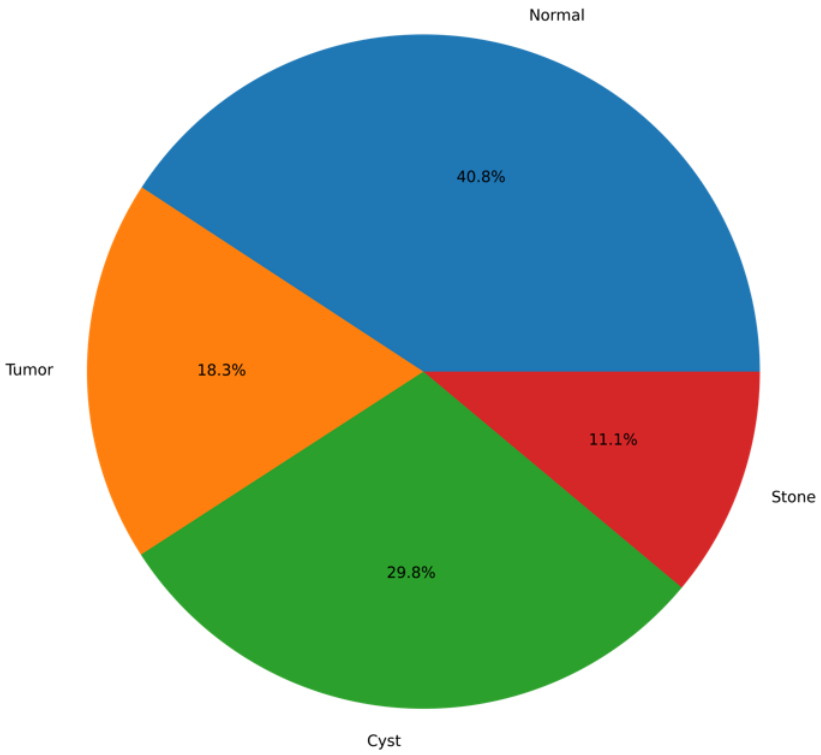


Figure 3.2: Different Labels in Dataset.

3.2 Data Pre-processing

3.2.1 Annotation

Annotation is the act of identifying and describing the subject matter of a picture. We learned the process of annotating the kidney CT images from a radiologist, then performed the annotation independently. We got the annotated images reviewed by the radiologist and only after his approbation we continued with our work. This is done to ensure the integrity of our research.

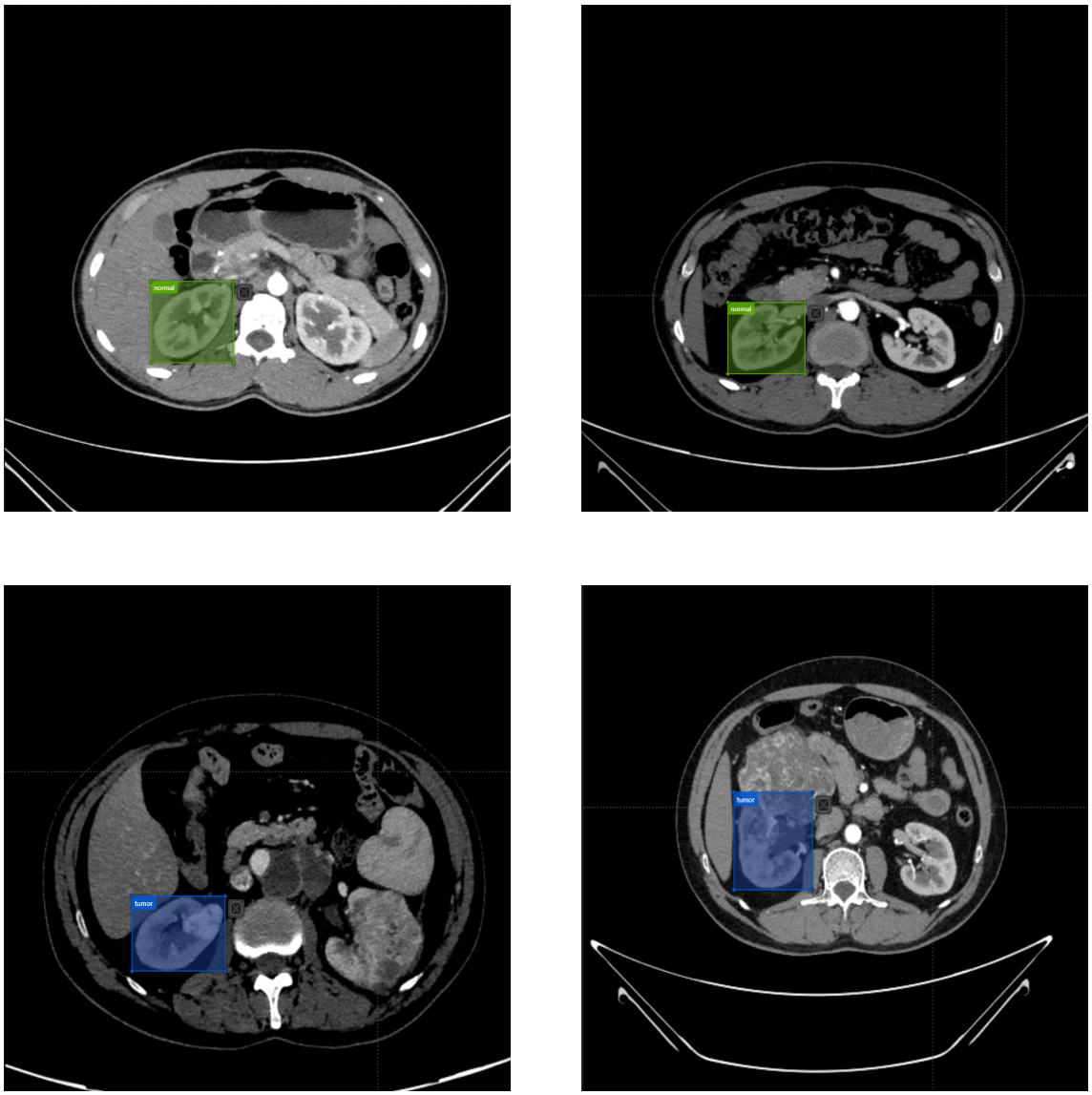


Figure 3.3: Annotated data.

We have used the software VOTT (visual object tracking tool) (version 2.3) from GitHub, an open-source, user-friendly annotation and labeling tool that helped us annotate the images. Rectangular boxes of appropriate shapes, as required, were used as bounding areas to mark the normal and tumor kidneys. Accordingly, we labeled the marked area as either “normal” or “tumor”. Manually, we annotated around 80 percent of the data. VOTT can also export the images with annotation information in CSV (comma-separated values) format. The CSV format is used

because it's easier to read the labels from the CSV spreadsheet file. There are six columns in each row of the CSV file and they are, Filename, Xmin, Ymin, Xmax, Ymax, and label. The figure 3.3 shows annotated data which is done by VOTT tool.

3.2.2 Mask

A mask is a technique for altering a bigger picture by defining a tiny portion of a larger one.

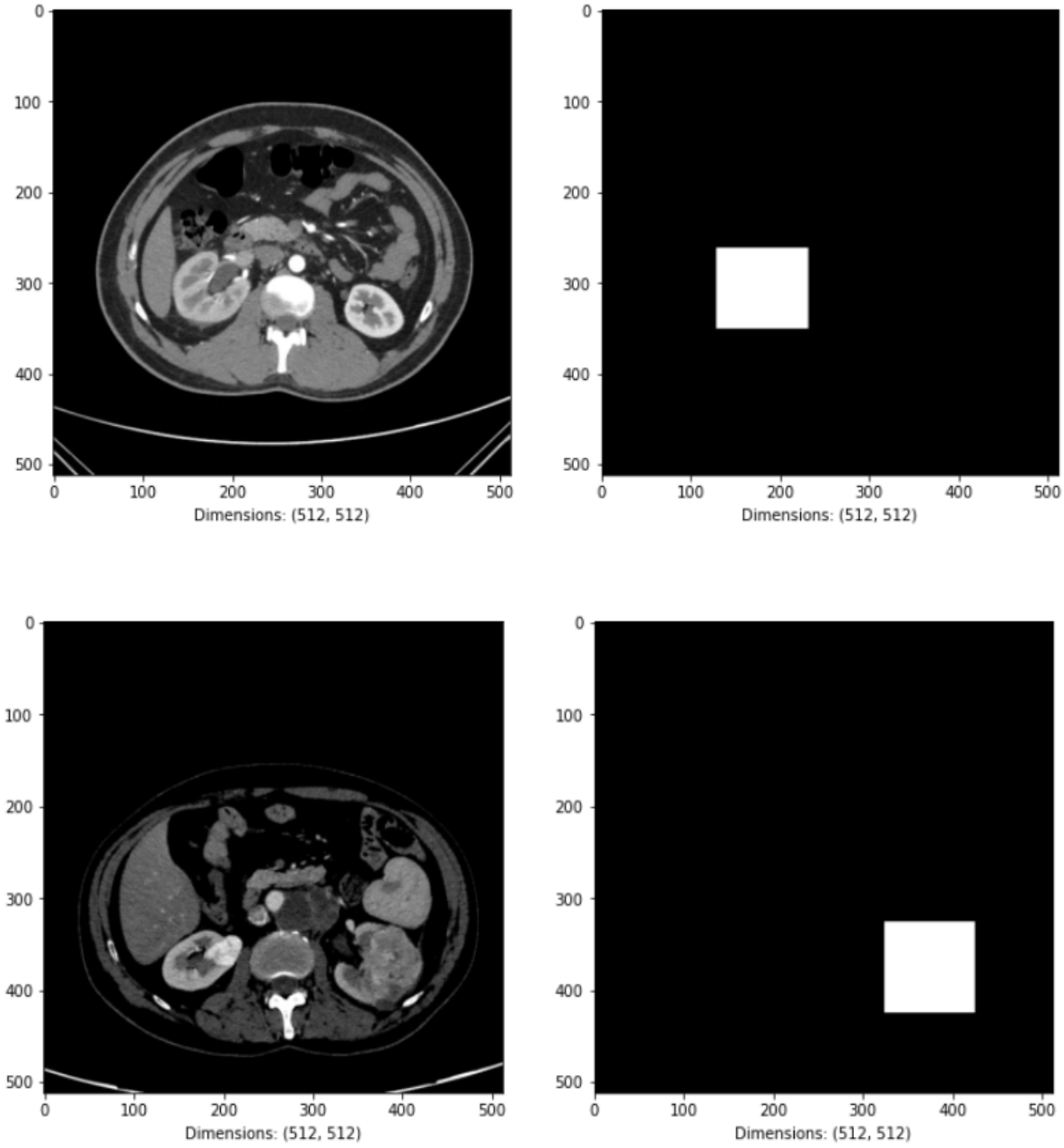


Figure 3.4: Mask with Orginal Image.

Masked images have non-zero and zero intensity values in different parts of the picture. In our research, we have created a python script to make masks of the

grayscaled kidney CT images. The bounding box values of the masks are obtained from the CSV file. For each image, two masks are created for the two co-responding kidneys, where each of the rectangular white areas in the mask represents a kidney area as shown in the figure 3.4. These masks are used to train the segmentation models Segnet and Unet to perform segmentation. We have created a mask of 70 percent of our selected data.

3.2.3 Segmentation

We have used both manual and model based kidney segmentaion to segment kidney from CT images.

Manual Kidney Segmentation

In the first part of segmentation, we have segment kidneys using our annotated bounding box by a python script. The values of bounding boxes are taken from the CSV file. The bounding box has Xmin, Ymin, Xmax, Ymax. We have segment kidneys of all annotated CT images. Figure 3.5 shows manual kidney segmentation result.

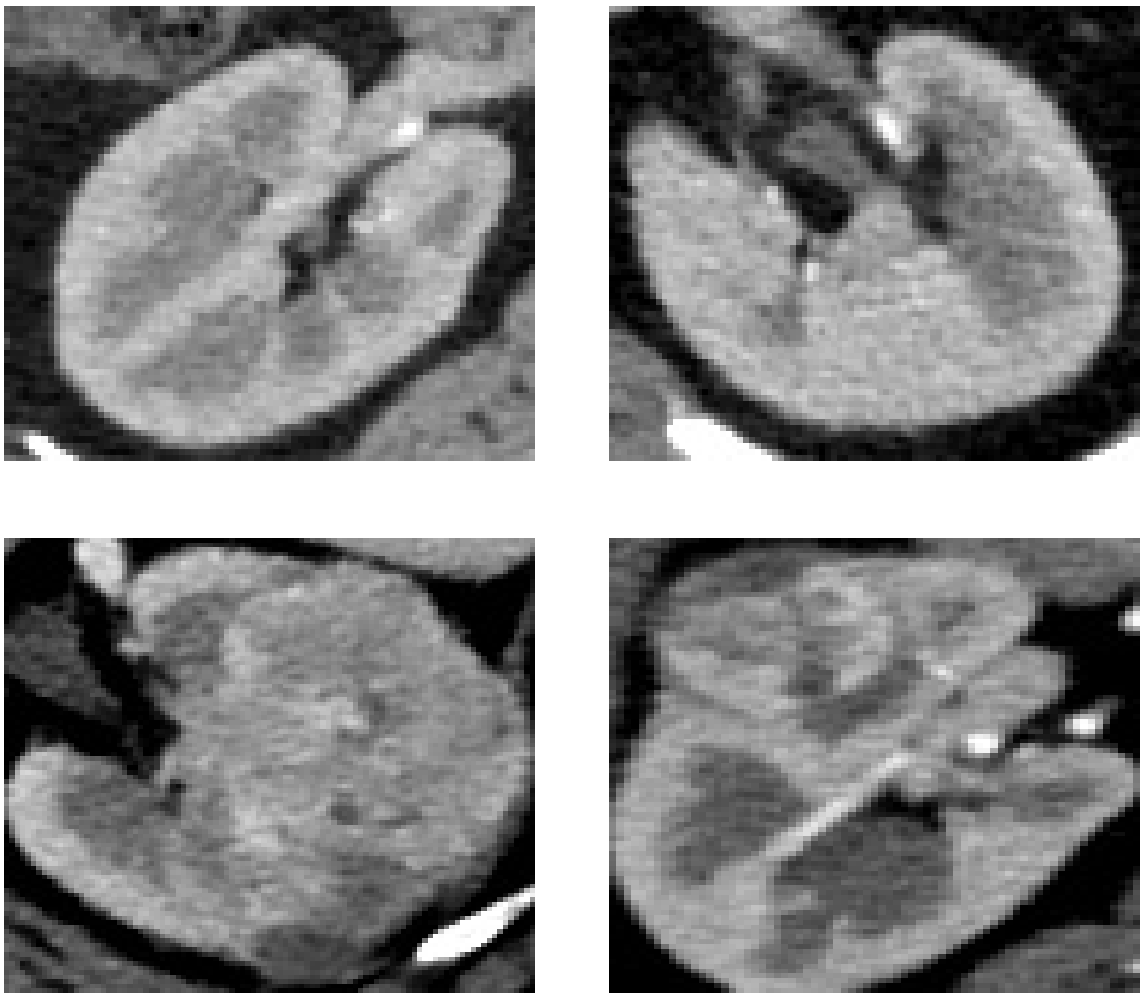


Figure 3.5: Manual Kidney Segmentation.

Kidney Segmentation using Model

In the second part of segmentation, we have used our 70 percent of the greyscaled kidney CT images with their corresponding masks to train both UNet and SegNet models. The output of the segmentation models is a predicted mask. Then we took the best model to create a mask of 30 percent of our remaining data. Using this mask, we have segmented the kidney from kidney CT images properly. We trained the segmentation model in such a way that if we add more data for further kidney segmentation in our future work, it can be done very efficiently and be exactly similar to the way it is done manually.

U-Net

There are many semantic segmentation architecture and U-Net [33] is a popular one among them.

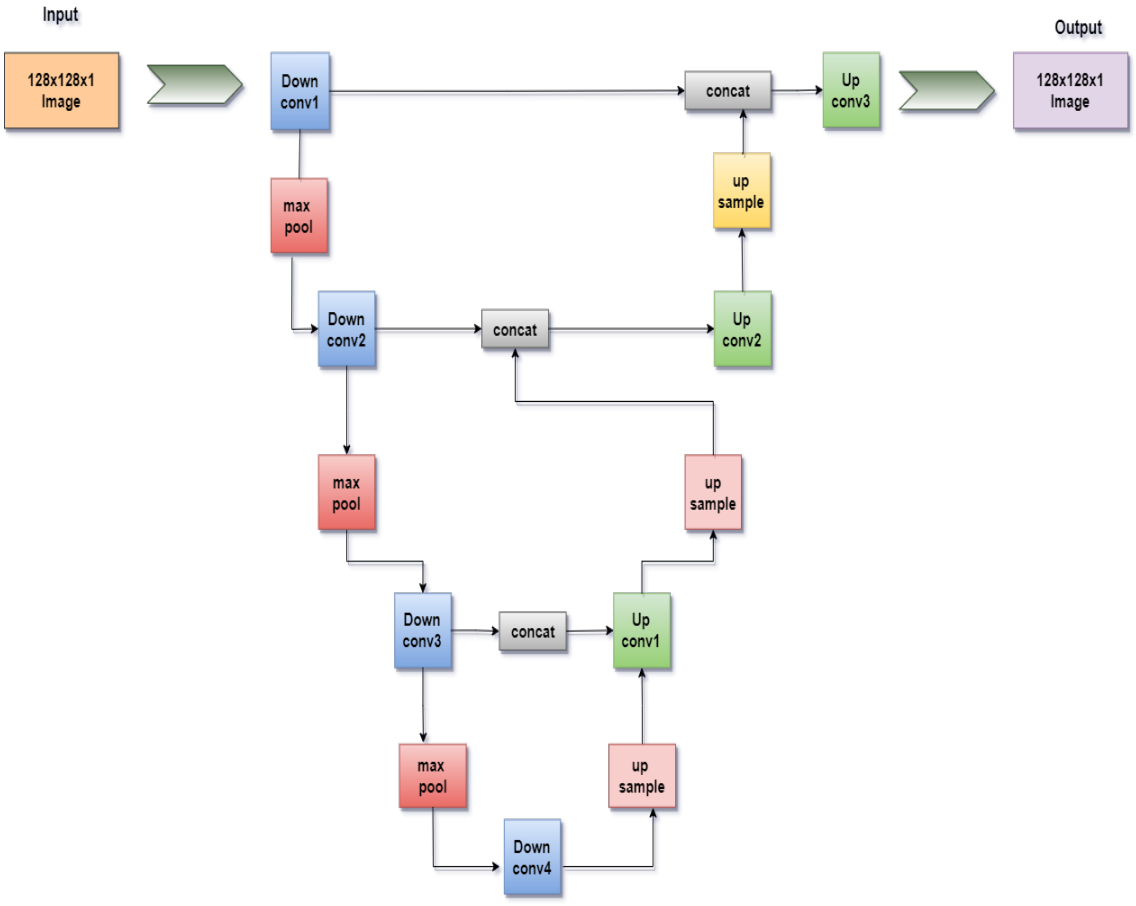


Figure 3.6: UNet Architecture.

The "U", the architecture of UNet is divided into three parts: contraction, bottleneck, and expansion. There are several contraction blocks in the section on contraction. Contracting paths follow the common architecture of a convolutional network, which consists of two 3x3 convolutions, each followed by a rectified linear unit (ReLU), a 2x2 max pooling operation with stride 2 for downsampling, followed by a rectified linear unit (ReLU), a 2x2 max pooling operation with stride 2. The number of feature channels doubles with each downsampling step. A 2x2 convolution

("up-convolution") that divides the number of feature channels in half is applied after each upsampling of the feature map in the expansive path, followed by two 3x3 convolutions, each followed by a ReLU. Then, the correspondingly cropped feature map from the contracting path is added. Every convolution loses pixels from the boundary, therefore the image has to be cropped. As a final step, each of the feature vector's 64 components is mapped to the required number of classes using a 1x1 convolution. In all, the network has 23 convolutional layers. Convolution layers are applied to each input, followed by a 2X2 maximum pooling. For the sake of learning complicated structures, the number of kernels or feature maps doubles. Figure 3.6 represents the architecture view of U-Net where input size is 128x128x1 and predicted output size is 128x128x1.

SegNet

The second semantic segmentation model we have used is SegNet [34]. There is a pixel-by-pixel classification layer, an encoder network, and a decoder network included in this design. VGG-16's 13 convolutional layers are replicated in the encoder network design of this encoder. The decoder comprises a total of 13 layers, with each encoder layer having a corresponding decoder layer. For each pixel, the final decoder output is sent into a softmax classifier that generates class probabilities. At the deepest encoder output, the higher-resolution feature maps are retained by discarding the completely linked layers. Decoder networks convert low-resolution encoder feature maps to full input resolution for per-pixel classification, and this is what they do. While the number of parameters in the SegNet encoder network dramatically decreases (from 134M to 14.7M), the number of parameters in other current designs is not. Using lower-resolution input feature maps, SegNet has a distinct advantage in decoding. It is necessary for the decoder to use the encoder's max-pooling step in order to carry out non-linear upsampling in the decoding process. Figure 3.7 represents the architecture view of SegNet where input size is 128x128x1 and predicted output size is 128x128x1.

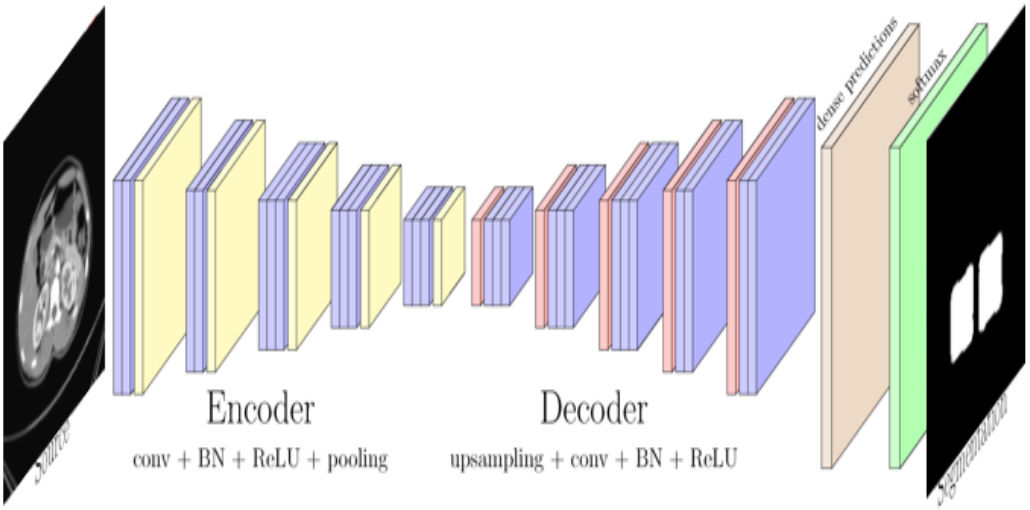


Figure 3.7: SegNet Architecture.

In both the segmentation models we have used 'Sigmoid' as activation function. This function is commonly used mathematical function which has an output similar to an S-shaped curve. The sigmoid function has several variations: logistic function, hyperbolic tangent, and arctangent. Nevertheless, for neural networks, it is defined as an activation function used to limit the output from 0 to 1 regardless of the input data provided into the function. It can also be used in applications where an actual number needs to be converted into a probability. We have applied it at the last layer of our models which can calculate the output as probability score between 0 and 1 that is much easier to work with. Besides this, Stochastic Gradient Descent (SGD) and Adaptive Moment Estimation(Adam) were used for cost function. In addition, We have chosen Binary Cross Entropy as loss function.

3.2.4 Target Labeling

We have labelled the segmented kidney which was found from manual kidney segmentation and segment model. One of the kidney labels is "Normal" and other is "Tumor". After this we stored them in two different directories for two classes.

3.3 Classification

For kidney tumor classification we have used MobileNetV2, VGG16 and InceptionV3.

3.3.1 MobileNetV2

MobileNetV2 [35] is a type of CNN architecture designed to execute efficiently on mobile devices. This model is a predecessor of MobileNetV1 and uses inverted bottleneck blocks and residual connections, which was not present in MobileNetV1. It is about 35 percent faster compared to MobileNetV1. It is created based on an inverted residual architecture in which the bottleneck layers have residual connections in between them. MobileNetV2 is based on the ideas from MobileNetV1 and uses depthwise separable convolution as efficient building blocks. But, Version2 adds 2 new features to the architecture. Bottlenecks in the layers' linear connections, and bottlenecks in the connections between the bottlenecks. the model's capacity to transition from lower-level ideas such as pixels to higher-level descriptors such as picture categories must be encoded, and this must be done by the model's bottleneck and the model's inner layer. Non-linearity is introduced into the model at the intermediate expansion layer by use of lightweight depthwise convolutions. Initial convolution layer with 32 filters is followed by 19 bottleneck layers in the MobileNetV2 architecture. 300 million multiply-adds are required to process the main network (width multiplier 1, 224x224) and 3.4 million parameters. It requires 7 multiply-adds up to 585M MAdds for the network to compute, while the model size ranges from 1.7M to 6.9M parameters.

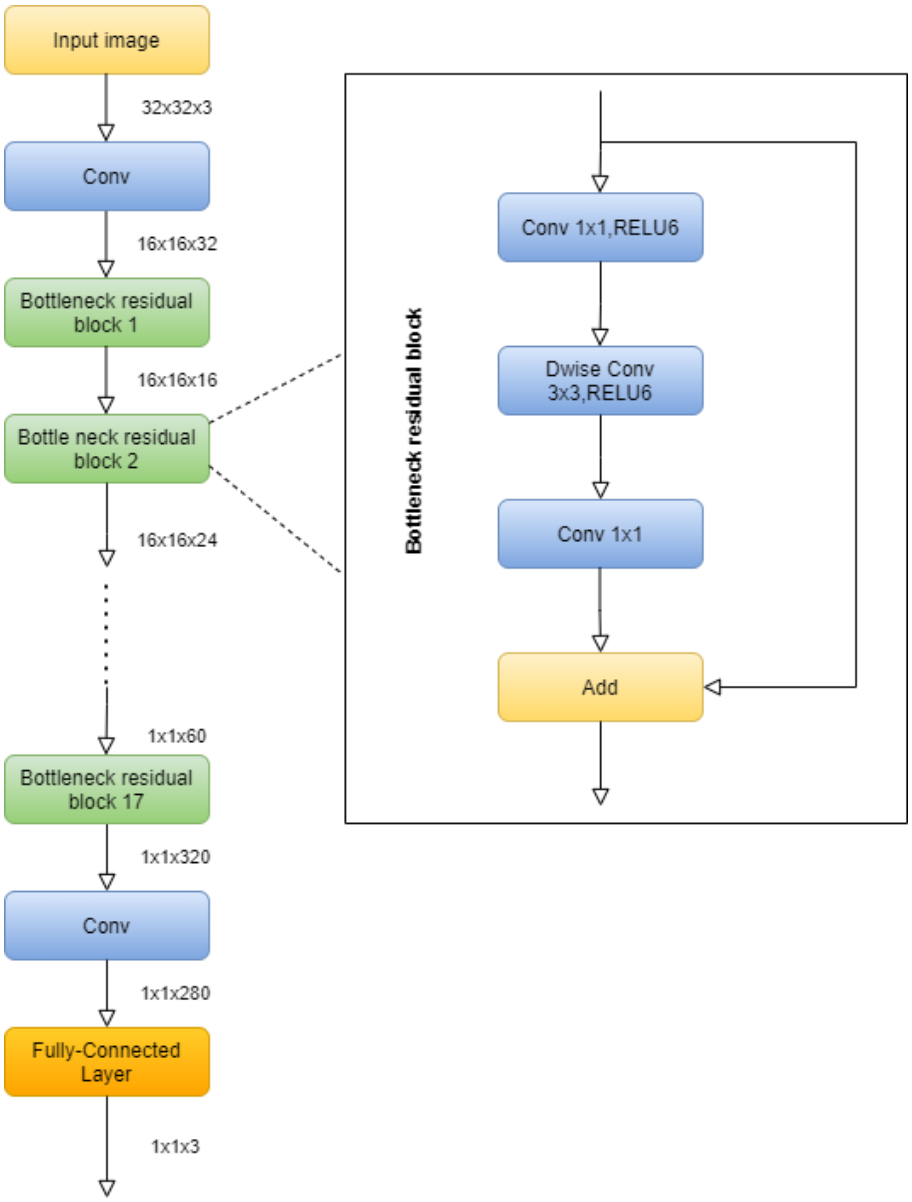


Figure 3.8: MobileNetV2 architecture.

There are two types of blocks: the residual block with a stride of 1 and the block with a stride of 2 used for downsizing. There are three layers for each type of block. The first layer is 1x1 convolution with ReLU6 and the second is the depthwise convolution. The third layer is 1x1 convolution without any non-linearity. Figure 3.8 represents the architecture view of MobileNetV2.

3.3.2 VGG16

In various deep learning approaches for image classification [36], the VGG-16 architecture is a simple, powerful, and widely used Convolutional Neural Network (CNN). It has 16 convolutional which is of uniform architecture.The VGG-16 has 16 layers, which means it can categorize pictures into 1000 different object types. 3x3 filter with stride one, same padding and maxpool layer as 2x2 filter stride 2 used

for convolution layers. Throughout the whole design, the convolution and max pool layers are arranged in this manner. The design concludes with two FC (completely connected layers) and a softmax. There are 138 million parameters in this massive network. Thirteen levels of convolution are included in VGG16, along with five Max Pooling layers and three Dense layers for a total of 21 layers. Still, there are only sixteen weight layers, or learnable parameters, in this algorithm. 224x244 is the input tensor size for VGG16, which has three RGB channels. But, we have modified the input tensor size to 128X128X3.Since it has so many nodes that are all linked, VGG is over 533MB in size. VGG delivery is made more difficult because of this. Figure 3.9 represents the architecture view of VGG16.

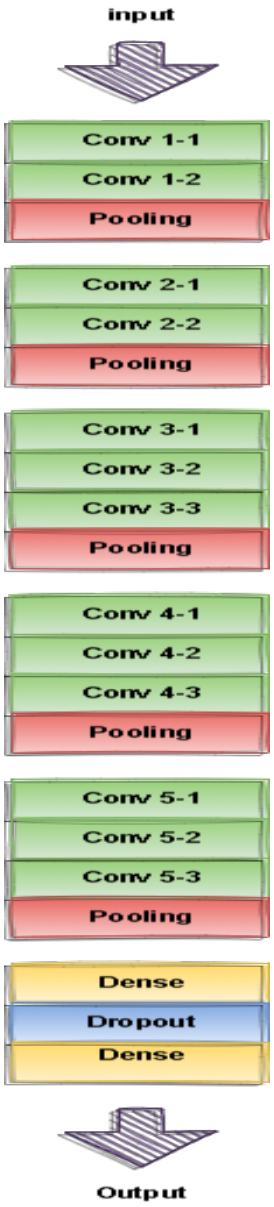


Figure 3.9: VGG16 architecture.

The 16 layers of VGG16:

1. Convolution layer 1 with 64 filters
2. Convolution layer 2 with 64 filters + Max pooling
3. Convolution layer 3 with 128 filters
4. Convolution layer 4 with 128 filters + Max pooling
5. Convolution layer 5 with 256 filters
6. Convolution layer 6 with 256 filters
7. Convolution layer 7 with 256 filters + Max pooling
8. Convolution layer 8 with 512 filters
9. Convolution layer 9 with 512 filters
10. Convolution layer 10 with 512 filters+Max pooling
11. Convolution layer 11 with 512 filters
12. Convolution layer 12 with 512 filters
13. Convolution layer 13 with 512 filters+Max pooling
14. Fully connected layer with 512 nodes
15. A Dropout layer with value of 0.5
16. Output layer with Sigmoid activation with 2 nodes

3.3.3 InceptionV3

This CNN model is used [37] for image analysis and object detection belongs to the Inception family of networks. For example, InceptionV3 includes a batch normalization for layers at the sidehead, which employs label Smoothing, factorized 7x7 convolutions, and an auxiliary classifier to transfer label information down the network. It has a reduced error rate than its predecessors, it has 42 layers. The InceptionV3 is the most recent and most effective model of the InceptionV1. The network of the InceptionV3 model is optimized using a variety of strategies. It's more efficient than the previous version. The InceptionV3 model features a deeper network and is quicker than the V1 and V2 models. It has less computational complexity and uses auxiliary Classifiers as regularizes. The network has been trained to categorize photos into 1000 different item categories.

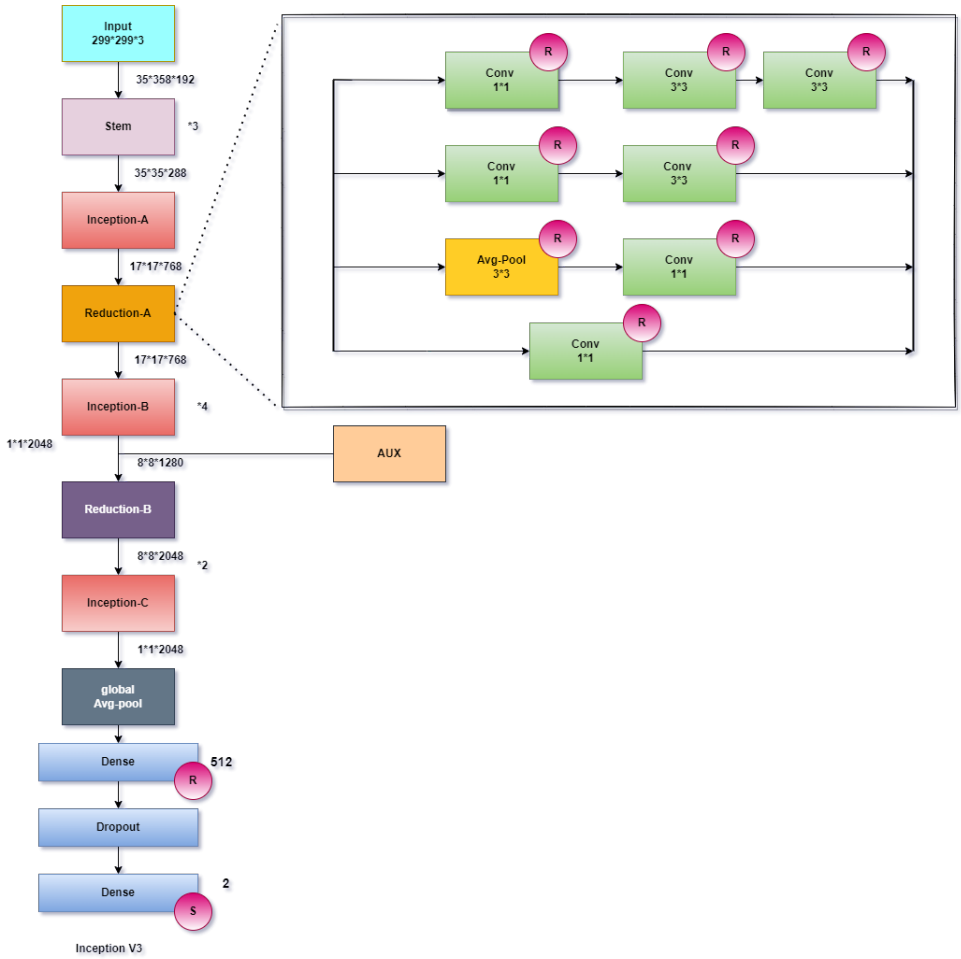


Figure 3.10: InceptionV3 Architecture.

Figure 3.10, The Network uses a 299-by-2 image scale to display its images. "Multi-level feature extraction" is what the first module is attempting to do by calculating 11, 33 and 55 convolutions in a single module of the network. When layered with channel measures, the output from each of these filters is then added to the next layer of the network, which may have recently been fortified with more information.

The role of the activation function in an artificial neural network is to define the output of a node when given an input or set of inputs to that node. Sigmoid function was used in the classification model's last layer so that the output probability score between 0 and 1. We have set lr=0.00001 in the adam optimizer. Moreover, Categorical crossentropy was used for loss function of our classification models.

Chapter 4

Result and Discussion

4.1 Implementation

As a deep neural network-based work that needs a high configuration PC to implement, we have used Nvidia RTX 2060 GPU, 24 GB ram, Ryzen 5 3600x with six core and 3.8 GHZ clock speed. Python version 3.9.7, Tensorflow version 2.7.0 and Keras version 2.7.0.

In order to save time, we decreased the images to 128x128 from their original 512x512 resolution. These CT slices were sent to several CNN architectures for segmentation and classification training after they have been processed. Besides this, the rest CT slices were used to test the accuracy. In the image 4.1 below, a few CT slices are shown.

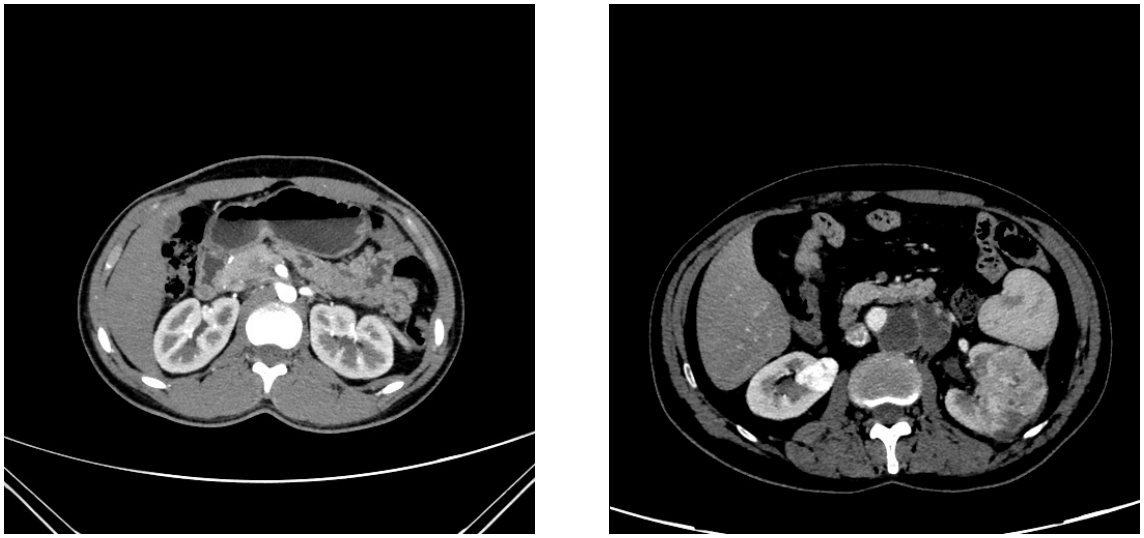


Figure 4.1: CT slies of patiens.

We fed the mask and original data to the segmentation models (UNet, SegNet). After that, we passed the segmented kidney data to the classification models (MobileNetV2, VGG16, InceptionV3) for classification. Using five different CNN architectures, we have been able to get some results and analyze them further.

4.2 Result

We calculated the accuracy using different epochs for different models and took the highest accuracy. For segmentation, we took 70 percent of our annotated data as a train set and 30 percent as a val (validation set). Rest 30 percent of non annotated data as a test set. On the other hand, 75 percent of the total segmented kidneys as a train set, 15 percent as a val and 10 percent as a test set for classification.

4.2.1 UNet

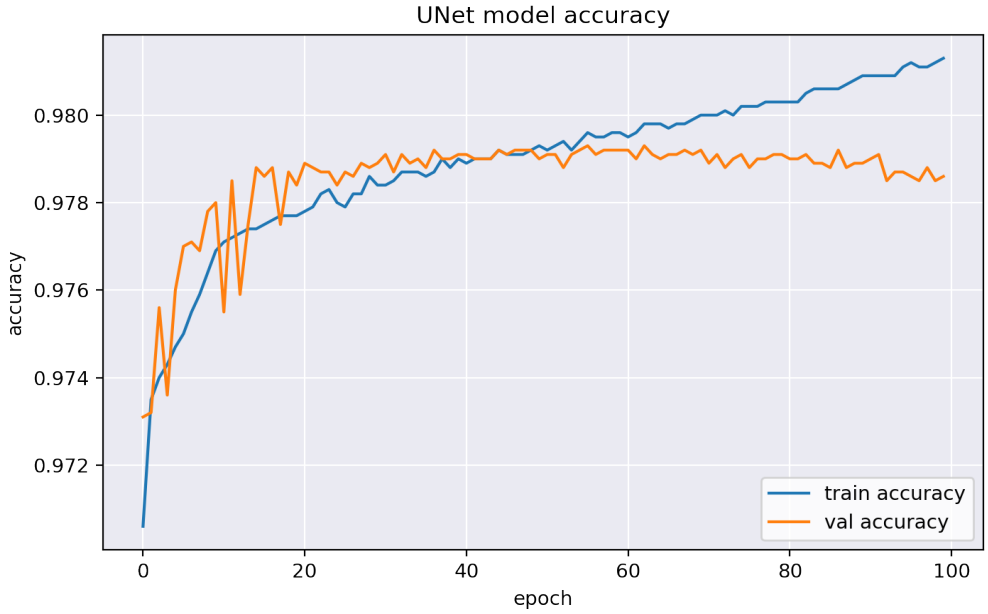


Figure 4.2: UNet Train/Val Accuracy.

In the beginning, training accuracy was about 97% and validation accuracy was also above 82%, as shown in the figure 4.2. Then gradually, train accuracy increased, though val accuracy had ups and downs in the graph. After 42 epochs, the train and val accuracy were the same, around 98% and both the lines intersected each other. At the finishing point, the training accuracy remained slighty above the training accuracy.

4.2.2 SegNet

From the figure 4.3 we can see that at the starting point, the training accuracy was around 94% and validation was 96%. Then gradually both train accuracy and val accuracy increased. At 2 epoch, the train and Val accuracy were almost the same, around 97% . After 11 epochs both lines started to slowly move apart from each other with the train accuracy increasing at a declining rate and val accuracy decreasing at a declining rate. At the finishing point, both reached their highest peak point with training accuracy stopping at 98 percent and val accuracy stopping at 97 percent.

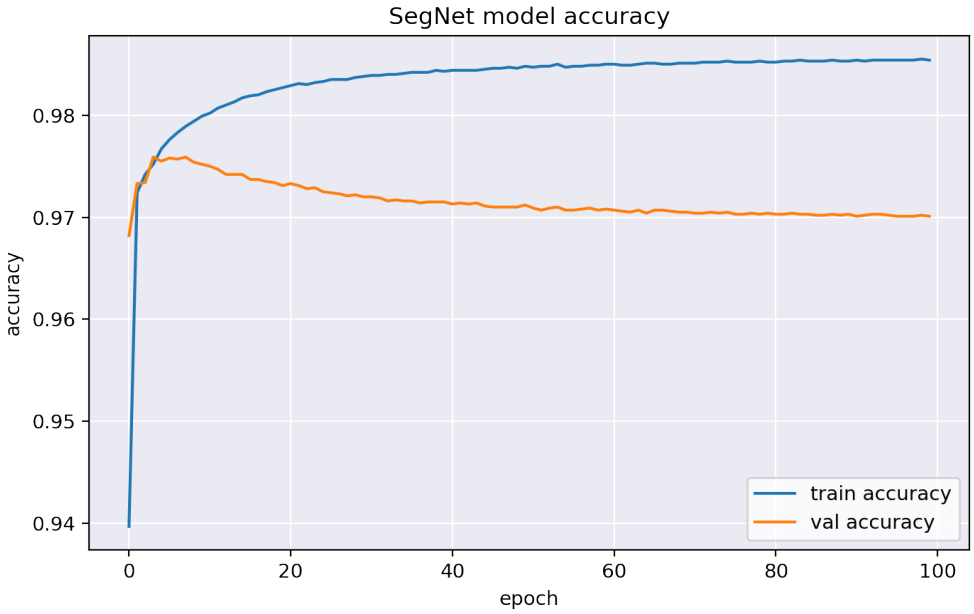


Figure 4.3: SegNet Train/Val Accuracy.

4.2.3 MobileNetV2

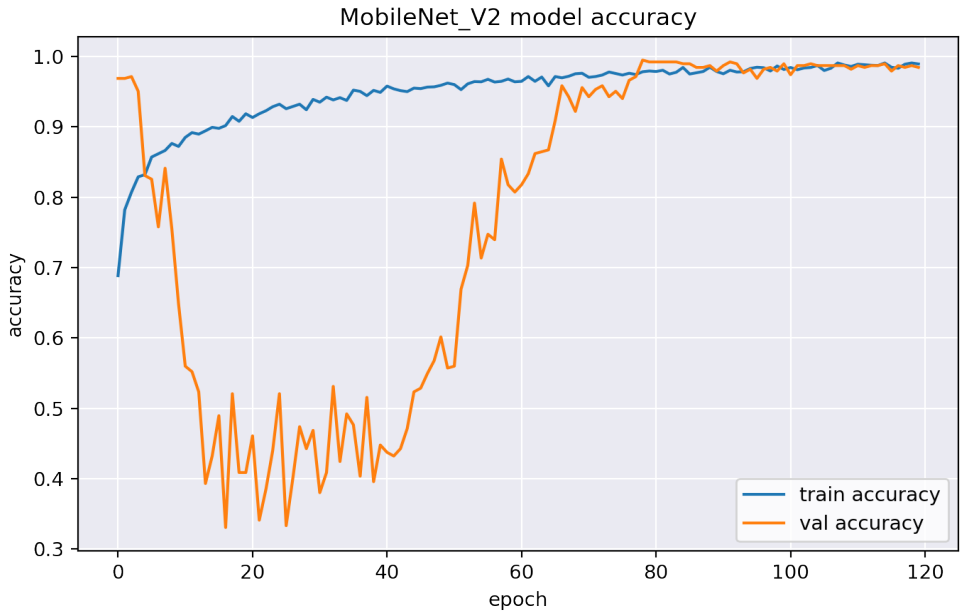


Figure 4.4: MobileNetV2 Train/Val Accuracy.

From the figure 4.4, we can see that the graph produced about 69% accuracy for the training accuracy and about 97% for the validation accuracy at 0 epochs. After 22 epochs, the validation accuracy reached its lowest point with an accuracy of about 34%. On the other hand, the training accuracy rose proportionally with time. After 89 epochs, the accuracy and training validation reached similar results and fell upon

each other on the graph. In the end, both the accuracy lines had similar values of almost 99% that converged with one another on the graph.

The training loss was about 77% and the validation loss was about 12% at the beginning of the training. After going through many epochs, the training loss gradually

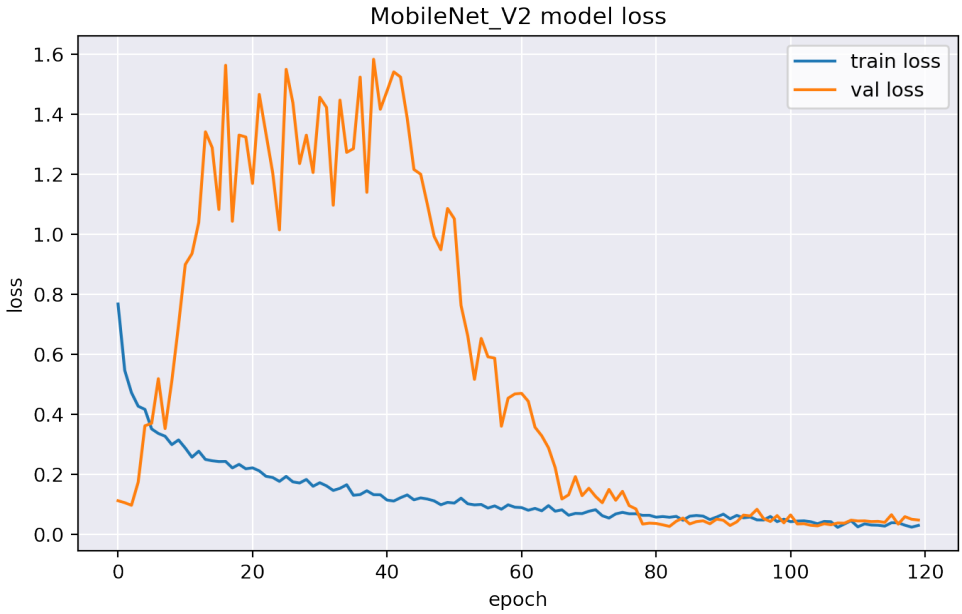


Figure 4.5: MobileNetV2 Train/Val Loss.

fell while the validation loss first rose and then fell. After 17 epochs, the validation loss peaked at about 150%. According to the final results, the training loss was around 3 percent and the validation loss was approximately 4 percent shows in the figure 4.5.

In the figure 4.6 is a confusion matrix which shows that there are two possible classes predicted by the model from the dataset under which the kidneys fall either into the category of " true positive", " true negative", " false positive" and "false negative" where "true negative" indicates kidney that is actually normal and also predicted as normal by the model, "false positive" indicates kidney that is actually normal but predicted as tumorous by the model, "false negative" indicates kidney that is tumorous but predicted as normal by the model and "true negative" indicates kidney that is actually tumorous and also predicted as tumorous by the model. MobileNetV2 made a total of 382 predictions(i.e. category outcome of 382 kidney images are predicted), and out of these 382 predictions, 256 are made from normal kidney image class and 126 are predicted from the class which has tumor category kidney images. Out of the 256 images from the normal kidney image class 252 are predicted as "true negative", while 4 are predicted as "false positive". On the other hand, out of the 126 images from the tumorous kidney class, 112 are predicted as "true positive" and 14 are predicted as "false negative".

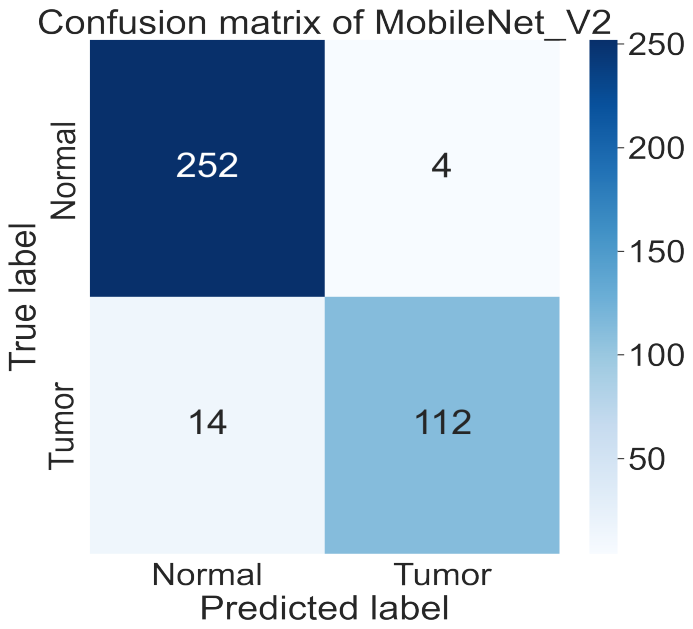


Figure 4.6: MobileNetV2 Confusion Matrix.

4.2.4 VGG16

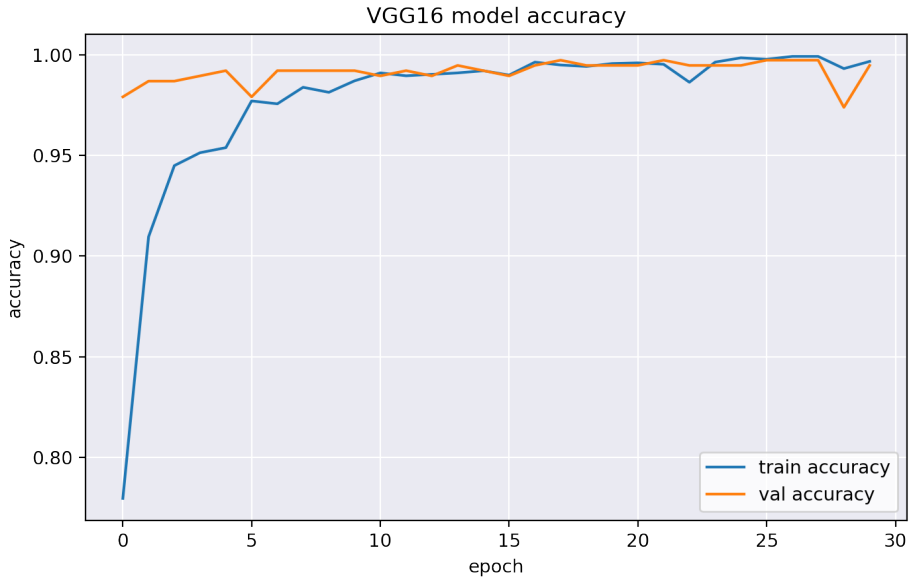


Figure 4.7: VGG16 Train/Val Accuracy.

We notice that the training accuracy was at it's lowest point of about 78% in the start of the training in the figure 4.7. Throughout the training phase, the validation accuracy remained quite constant and was about 98% initially. The validation accuracy gradually rose as the training continued and after 25 epochs, it reached its highest point. At the completion of the training, the training and validation accuracy rates were both over 99 %.

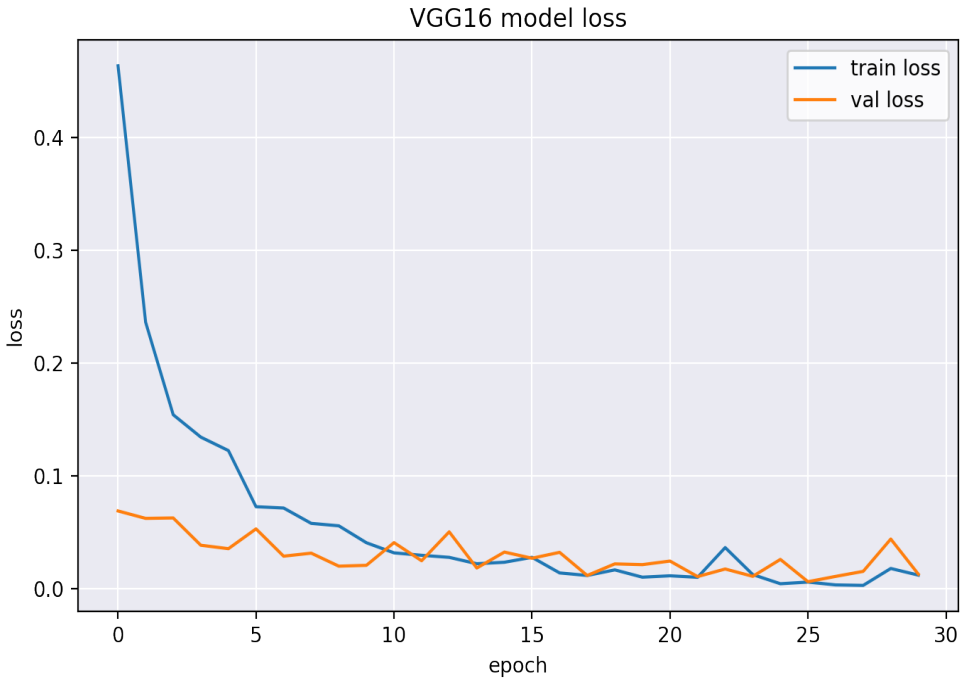


Figure 4.8: VGG16 Train/Val Loss.

The training loss of the model was about 47%, and the validation loss was about 7% at the start of the training, which we found in the figure 4.8. The training loss gradually went down as the training continued, while the validation loss remained almost consistent with slight ups and downs. After 13 epochs, the validation loss peaked at about 5%, and after 23 epochs, it peaked at about 4%. At the end of the training, both training loss and validation loss lines had similar values of about 1%.

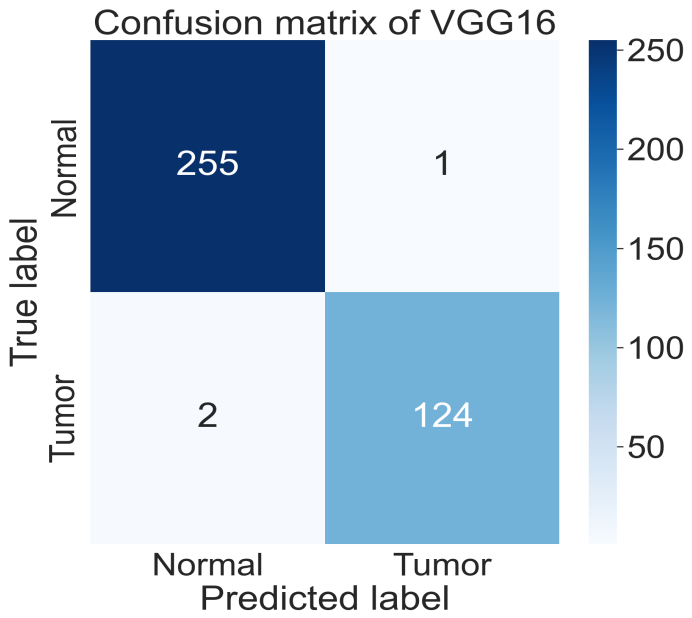


Figure 4.9: VGG16 Confusion Matrix.

The confusion matrix of VGG-16 model which classified kidney images into " true positive", " true negative", " false positive" and "false negative" where "true negative" indicates kidney that is actually normal and also predicted as normal by the model, "false positive" indicates kidney that is actually normal but predicted as tumorous by the model, "false negative" indicates kidney that is tumorous but predicted as normal by the model and "true negative" indicates kidney that is actually tumorous and also predicted as tumorous by the model. It classified 382 images, out of which 256 are from normal kidney image class and 126 are predicted from the class which contains tumor images of the kidney. Among the 256 images from normal kidney image class, 255 are predicted as "true negative"(true normal), while only 1 is predicted as "false positive" (false tumor).Out of the remaining 126 images from the tumorous kidney class 124 are predicted as "true positive"(true tumor), and 2 are predicted as "false negative"(false normal) as shown in figure 4.9.

4.2.5 InceptionV3

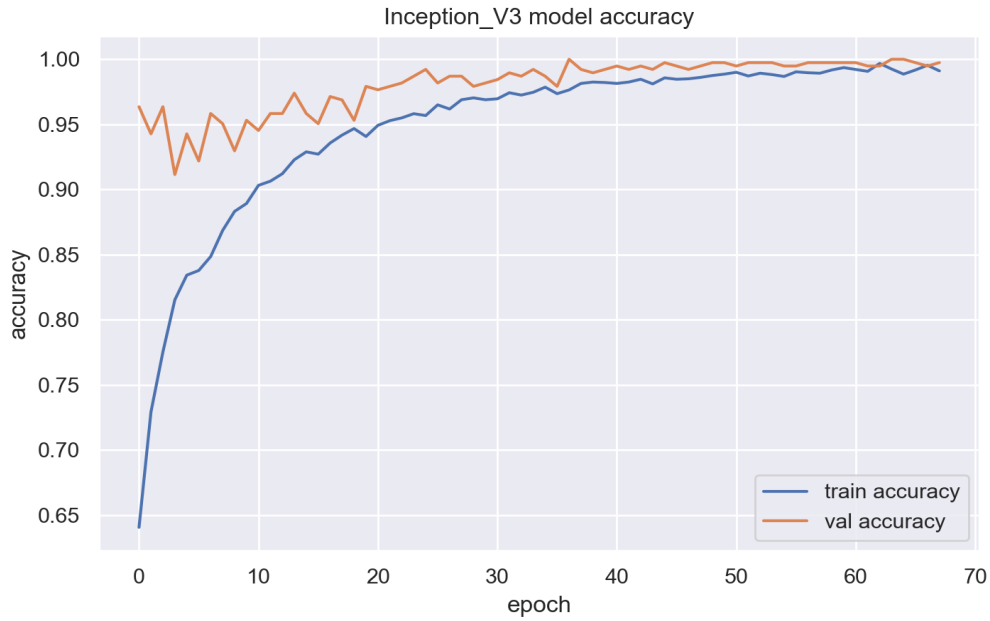


Figure 4.10: InceptionV3 Train/Val Accuracy.

We see in the graph 4.10 that the training accuracy was about 64%, and the validation accuracy was about 96% at the start of the training. Both the training and validation accuracy gradually rose as the training went on. Throughout the training, the validation accuracy was mostly above the training accuracy. After 60 epochs, they start to have similar values, and the lines overlap. In the end, both the training and validation accuracy reaches over 99% accuracy.

Figure 4.11: InceptionV3 Train/Val Loss.

In Figure 4.11, the training loss starts at about 67%, and validation loss is about 18%. Both the training and validation loss goes down as the training continues with slight ups and downs. After 16 epochs, they start to get similar values, and the graph lines overlap. At the end of the training, the validation and training loss get similar values, with the validation loss being almost 4% and training loss being 3%.

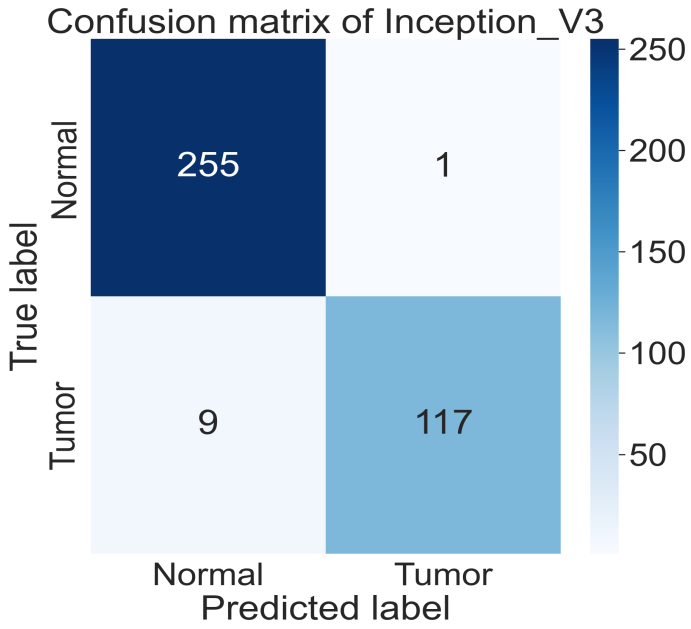


Figure 4.12: InceptionV3 Confusion Matrix.

In figure 4.12 confusion matrix shows the statistics same as the previous two models where kidney images are classified as " true positive", " true negative", " false

positive" and "false negative" by the model from the given kidney image dataset where "true negative" indicates kidney that is actually normal and also predicted as normal by the model, "false positive" indicates kidney that is actually normal but predicted as tumorous by the model, "false negative" indicates kidney that is tumorous but predicted as normal by the model and "true negative" indicates kidney that is actually tumorous and also predicted as tumorous by the model. Total 382 images are classified from which 256 are from normal kidney image class and 126 are from tumor kidney image class. The 255 out of the 256 normal kidney images are classified as true negative and only 1 as false positive. In contrast, the 117 images out of the 126 tumor kidney images are classified as true positive and 9 as false negative.

4.3 Discussion on Segmentation Result

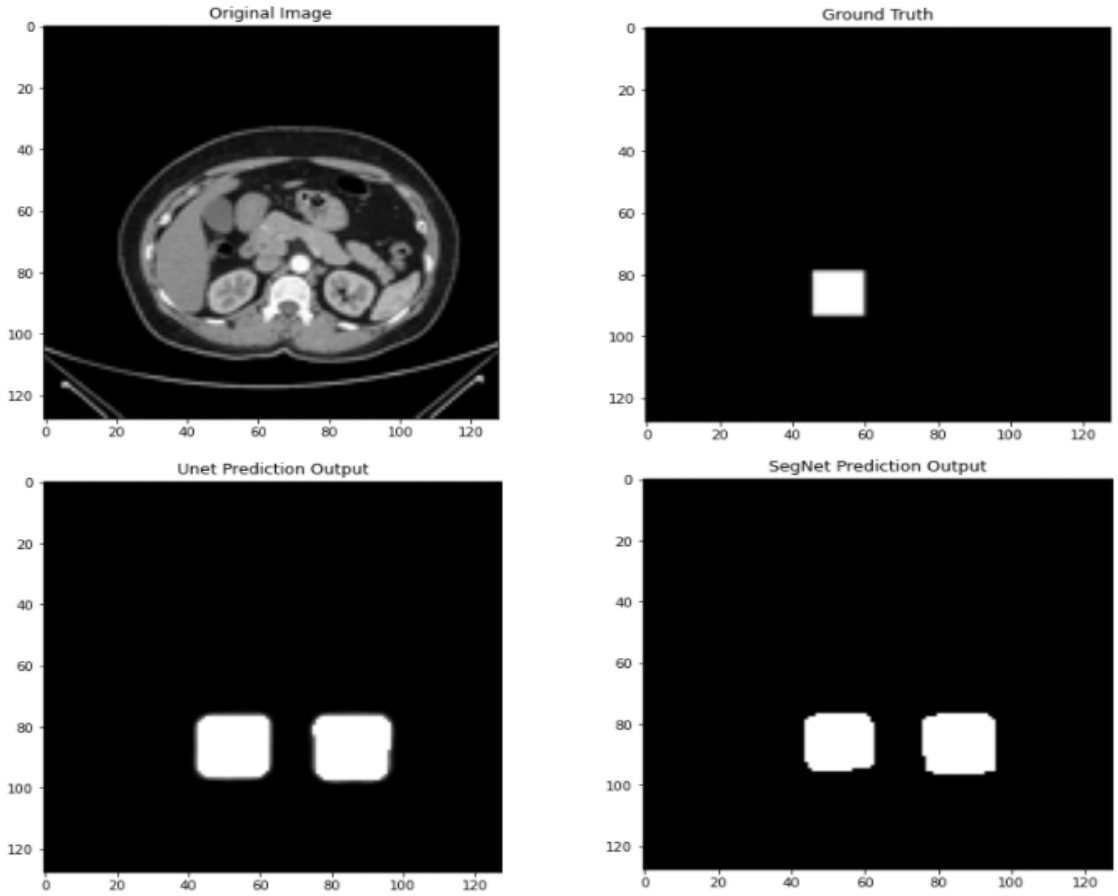


Figure 4.13: Test performance of segmentation models.

For kidney segmentation, we have segmented 70 percent data manually and then for rest of the 30 percent we have used both UNet and SegNet models. Though their performance is almost the same and quite similar with our manual segmentation result, we have chosen UNet as it provides better accuracy than SegNet. The test

accuracy is 97.58% for UNet and 96.38% for SegNet. Here is the figure 4.13 which shows the output for both the models. As we can see, we have created a mask for one kidney, but our trained models created the mask for both kidneys and using this mask we have segmented the rest of the kidneys.

4.4 Discussion on Classification Result

Table 4.1: Train Classification Accuracy of different models

	Train Accuracy		
Epochs	MobileNetV2	VGG16	InceptionV3
1	68.85	77.96	64.06
5	83.21	95.35	83.42
10	87.21	98.71	88.92
15	89.92	99.21	92.89
20	91.85	99.57	94.93

The table 4.1 shows the compared accuracy of three different models that have been used for training accuracy. For each model, a specific epoch value gave the highest accuracy.
In the figure, the contrast of different models yields different accuracy levels. For 1st epoch values, InceptionV3 attains the lowest accuracy of 64.06% and VGG16 achieves the highest accuracy. Nevertheless, the scenario changes for all three models as we increase the epoch values. For epoch value 5, the accuracy of VGG16 jumps to 95.35%. However, InceptionV3 and MobileNetV2 have a close resemblance of 83.42% and 83.21%. As we increment the epoch values, we get more accuracy levels. For epoch 10, MobileNetV2 has 87.21%, VGG16 has 98.71%, and InceptionV3 has 88.92% accuracy. As we can see, in the case of epoch 10, VGG16 also has the upper hand over all other models. Incrementing again, for epoch 15, MobileNetV2 has 89.92%, VGG16 has 99.21%, InceptionV3 has 92.89% of accuracy. And for epoch 20 MobileNetV2, VGG16 and InceptionV3 has 91.85%, 99.57% and 94.93% of accuracy respectively. The highest epoch accuracy has been acquired by VGG16, which is 99.57%.

Table 4.2: Validation Classification Accuracy of different models

	Validation Accuracy		
Epochs	MobileNetV2	VGG16	InceptionV3
1	96.88	97.92	96.35
5	83.07	99.22	94.27
10	64.84	99.22	95.31
15	43.23	99.22	95.83
20	40.89	99.48	97.92

Each model's validation accuracy is shown in Table 4.2, along with its behavior at various epoch values. For just epoch one of all three models give us impressive accuracy results, which are 96.88%, 97.92% and 96.35% respectively. But as the value

of the epoch increases, VGG16 and InceptionV3 have a significant rise in accuracy, but MobileNetV2 has exactly the opposite reaction. For the fifth epoch, VGG16 and InceptionV3 have 99.22% and 94.27% accuracy, but MobileNetV2 has 83.07% accuracy, which is significantly lower than the first epoch. For epoch 10, VGG16 has 99.22%, and InceptionV3 has 95.31% accuracy, but MobileNetV2 has 64.84%. For epoch 15, VGG16 and Inception have 99.22% and 95.83%, but MobileNetV2 has 43.23%. And lastly, for epoch 20, VGG16 gives the highest accuracy, which is 99.48% and the lowest is MobileNetV2, which is 40.89%.

Table 4.3: Test Performance of Different Classification Models

	MobileNetV2	VGG16	InceptionV3
Accuracy	95.29	99.21	97.38
Loss	15.0	4.75	17.03
Sensitivity	88.89	98.41	92.86
Specificity	98.44	99.61	99.61
F1 score	95.0	99.0	97.0

Table 4.3 was used to assess the accuracy, loss, sensitivity, specificity and F1 score of the three models, which were compared to each other using the following criteria. For example, MobileNetV2, VGG16, and InceptionV3 all function differently depending on the situation. Before getting the idea of how these three models have performed in different cases, we need to know what these cases are, which means what is meant by accuracy, loss, sensitivity, specificity and F1 score.

Figure 4.14: Test Performance of Different Classification Models.

The explanation of "accuracy" is straightforward - it is the ratio of correctly predicted classifications to the total number of predictions made [38]. The "Loss" value

indicates how badly a model performs in making the right predictions for a single instance. If the model's prediction is perfect, then the loss is zero [39]. Testing's "sensitivity" measures how well a test can detect the presence of a disease in a given sample of people. Testing's "specificity" is defined as the ability of a test to correctly identify those who do not suffer from the disease. The "F1 score" is generated using the precision and recall of the models, and it indicates how accurate the model is on the kidney dataset. There are four factors involved in calculating the following four terms:

- True Positive (TP): The patient has the tumor, and the model predicts it as positive.
- True Negative (TN): The patient does not have the tumor, and the models prediction is also negative.
- False Positive (FP): The patient does not have the tumor, and the models prediction is positive.
- False Negative (FN): The patient has the tumor and the test is negative [40].

$$\text{Sensitivity=Recall} = \frac{TP}{TP + FN} \tag{4.1}$$

$$\text{Specificity} = \frac{TN}{TN + FP} \tag{4.2}$$

$$\text{F1 Score=Dice Score} = \frac{2.Precision.Recall}{Precision + Recall} \tag{4.3}$$

Among all the three models, VGG16 has the highest accuracy level, which is 99.21%, followed by InceptionV3 97.38% and MobileNetV2 95.29%. Consequently, VGG16 has the least loss, which is 4.75%, then comes MobileNetV2 with 15% and InceptionV3 with 17.03%. Here, we can see a slight discrepancy, as MobileNet has the least accuracy, so its loss function should also be higher than the rest, but IncpetionV3 tops it. The reason is, MobileNetV2 is a lightweight model with fewer parameters. So in the case of predicting loss function, it also has fewer parameters to work with and thus yields a lower score [41]. For sensitivity, VGG16 has the highest score, which is 98.41%. The rest two follow behind with InceptionV3 92.86% and MobileNetV2 88.89%. In calculating specificity, it is a tie between VGG16 and IncpeitonV3, both scoring 99.61% and MobileNetV2 with 98.44%. Even for F1 Score, VGG16 model has the highest score of 99.0% while the MobileNetV2 scores 95.0% and InceptionV3 scores 97.0% respectively. The bar chart 4.14 is the representation of similar quantities.

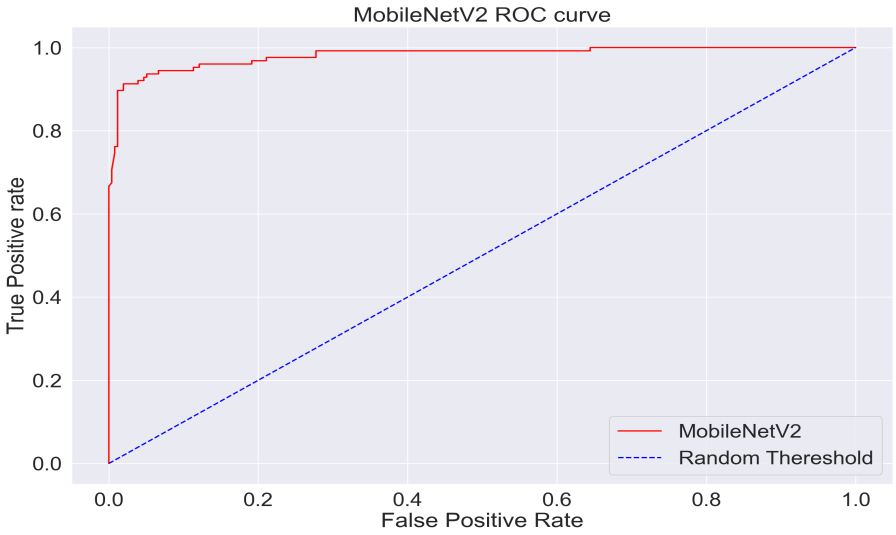


Figure 4.15: ROC MobileNetV2.

The ROC curve shows the performance of the classification model's sensitivity (or TPR) and specificity (1 - FPR). Classifiers that create curves that are closer to the top-left corner indicate better results. There was a Random classifiers which supposed to provide points that fall along the diagonal (FPR = TPR). The ROC space's 45-degree diagonal axis of rotation is closer to being intersected by a curve as a test's accuracy drops. In the calculation we mark tumor as positive and normal as negative.

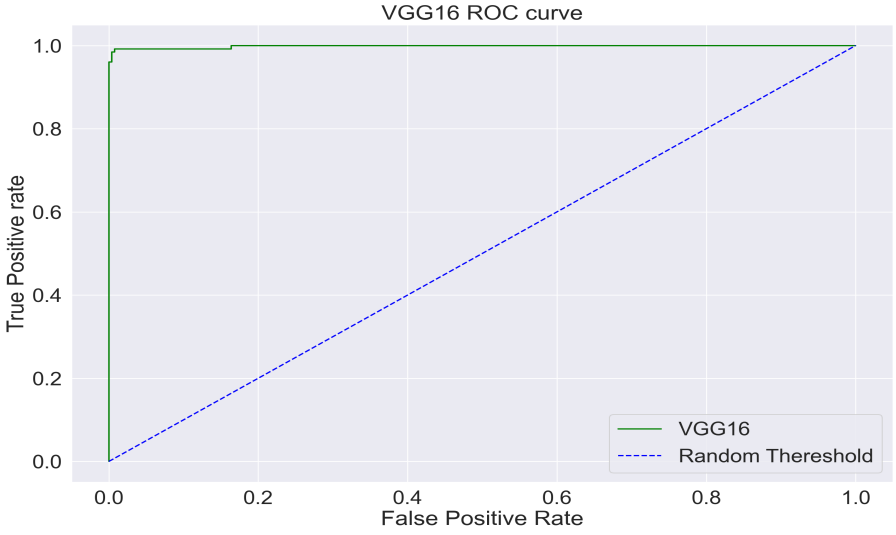


Figure 4.16: ROC of VGG16.

Figure 4.17: ROC of InceptionV3.

Figure 4.15, 4.16 and 4.17 shows that the roc graph behaves differently on different models. Among these three graph, there is a good chance that the VGG16 roc score is near to 0.99. In contrast, the MobileNetV2 roc value is close to 0.90, which is the lowest of all the roc scores. In addition, InceptionV3 holds the middle position and the value is close to 0.93.

Chapter 5

Conclusion

Prediction and diagnosis of renal disease are major topics in scientific research. Machine learning, neural networks, genetic algorithms, and other techniques are being used to improve and speed up data processing. As a result, sifting through the many approaches to discover the most effective one has become critical. We compared the outcomes of several neural networks used to identify renal illness in this paper. Our major objective was to demonstrate which of the networks could deliver the best outcomes with the supplied data. We did this by dividing the kidney dataset into two distinct categories for comparison. A manual segmentation and segmentation models were used to preprocess the CT slices. The segmented kidney are fed into our three classification models: MobileNetV2, VGG16 and InceptionV3. According to the data, VGG16 is the most accurate, with sensitivity and specificity higher than other models.

5.1 Future Work

The results of our proposed model, which looked at the segmented kidney and then identified it with a high accuracy rate, are particularly applicable to real-life circumstances. This is not the end of this work, We will utlize this model with a huge dataset. It is possible to have a better understanding of this approach's position by comparing it to other current techniques based on criteria such as accuracy, efficiency, and practicality. This strategy can be useful to generate computer software which will quickly help doctors in better classifying kidneys than ever before. A larger range of issues allows for more work to be completed and more accurate data to be obtained while discussing this topic. Adding mode data to our model eliminates the need to manually annotate data. Everything will be taken care by our segmentation model. Finally, for future work we will classify the types of tumor using DNN if we can manage a big dataset.

Bibliography

[1] S. Johnson, *Kidney health and kidney disease basics*, Healthline, Jul. 2012. [Online]. Available: https://www.healthline.com/health/kidney-disease#types-and-causes.

[2] *Kidney cancer - introduction*, en, https://www.cancer.net/cancer-types/kidney-cancer/introduction, Accessed: 2021-9-15, Jun. 2012.

[3] S. Safiri, A.-A. Kolahi, M. A. Mansournia, *et al.*, "The burden of kidney cancer and its attributable risk factors in 195 countries and territories, 1990-2017," en, *Sci. Rep.*, vol. 10, no. 1, p. 13 862, Aug. 2020.

[4] *World life expectancy*, World Life Expectancy, 2017. [Online]. Available: https://www.worldlifeexpectancy.com/.

[5] *Kidneys*, en, https://www.yourhormones.info/glands/kidneys/, Accessed: 2021-9-26.

[6] *Blood tests*, en, https://www.nhs.uk/conditions/blood-tests/, Accessed: 2021-9-24.

[7] S. McKee, "Half of all kidney cancer patients initially misdiagnosed," en, Oct. 2018.

[8] K. Yan, X. Wang, L. Lu, *et al.*, "Deep lesion graphs in the wild: Relationship learning and organization of significant radiology image findings in a diverse large-scale lesion database," in *2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2018, pp. 9261–9270. DOI: 10.1109/CVPR.2018.00965.

[9] K.-Y. Lung, C.-R. Chang, S.-E. Weng, H.-S. Lin, H.-H. Shuai, and W.-H. Cheng, "Rosnet: Robust one-stage network for ct lesion detection," *Pattern Recognition Letters*, vol. 144, Jan. 2021. DOI: 10.1016/j.patrec.2021.01.011.

[10] K. Yan, M. Bagheri, and R. Summers, "3d context enhanced region-based convolutional neural network for end-to-end lesion detection," in Sep. 2018, pp. 511–519, ISBN: 978-3-030-00927-4. DOI: 10.1007/978-3-030-00928-1_58.

[11] I. Alnazer, P. Bourdon, T. Urruty, *et al.*, "Recent advances in medical image processing for the evaluation of chronic kidney disease," *Medical Image Analysis*, vol. 69, p. 101 960, Jan. 2021. DOI: 10.1016/j.media.2021.101960.

[12] T. Pan, G. Yang, C. Wang, *et al.*, "A multi-task convolutional neural network for renal tumor segmentation and classification using multi-phasic ct images," in *2019 IEEE International Conference on Image Processing (ICIP)*, 2019, pp. 809–813. DOI: 10.1109/ICIP.2019.8802924.

[13] Y. Ren, H. Fei, X. Liang, D. Ji, and M. Cheng, "A hybrid neural network model for predicting kidney disease in hypertension patients based on electronic health records," *BMC Medical Informatics and Decision Making*, vol. 19, Apr. 2019. DOI: 10.1186/s12911-019-0765-4.

[14] N. Hadjiyski, "Kidney cancer staging: Deep learning neural network based approach," in *2020 International Conference on e-Health and Bioengineering (EHB)*, 2020, pp. 1–4. DOI: 10.1109/EHB50910.2020.9280188.

[15] G. Chen, C. Ding, Y. Li, *et al.*, "Prediction of chronic kidney disease using adaptive hybridized deep convolutional neural network on the internet of medical things platform," *IEEE Access*, vol. 8, pp. 100 497–100 508, 2020. DOI: 10.1109/ACCESS.2020.2995310.

[16] M. Shehata, A. Alksas, R. T. Abouelkheir, *et al.*, "A new computer-aided diagnostic (cad) system for precise identification of renal tumors," in *2021 IEEE 18th International Symposium on Biomedical Imaging (ISBI)*, 2021, pp. 1378–1381. DOI: 10.1109/ISBI48211.2021.9433865.

[17] H. Zhang, Y. Chen, Y. Song, Z. Xiong, Y. Yang, and Q. M. Jonathan Wu, "Automatic kidney lesion detection for ct images using morphological cascade convolutional neural networks," *IEEE Access*, vol. 7, pp. 83 001–83 011, 2019. DOI: 10.1109/ACCESS.2019.2924207.

[18] A. Saood and I. Hatem, "Covid-19 lung ct image segmentation using deep learning methods: U-net versus segnet," *BMC Medical Imaging*, vol. 21, Feb. 2021. DOI: 10.1186/s12880-020-00529-5.

[19] L. Liu, L. Wang, D. Xu, *et al.*, "Ct image segmentation method of liver tumor based on artificial intelligence enabled medical imaging," *Mathematical Problems in Engineering*, vol. 2021, pp. 1–8, May 2021. DOI: 10.1155/2021/9919507.

[20] A. Skalski, J. Jakubowski, and T. Drewniak, "Kidney tumor segmentation and detection on computed tomography data," in *2016 IEEE International Conference on Imaging Systems and Techniques (IST)*, 2016, pp. 238–242. DOI: 10.1109/IST.2016.7738230.

[21] D. Müller and F. Kramer, *Miscnn: A framework for medical image segmentation with convolutional neural networks and deep learning*, 2019. arXiv: 1910.09308 [eess.IV].

[22] G. Mu, Z. Lin, M. Han, G. Yao, and Y. Gao, "Segmentation of kidney tumor by multi-resolution vb-nets," Jan. 2019. DOI: 10.24926/548719.003.

[23] M. Marsousi, K. N. Plataniotis, and S. Stergiopoulos, "Shape-based kidney detection and segmentation in three-dimensional abdominal ultrasound images," in *2014 36th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, 2014, pp. 2890–2894. DOI: 10.1109/EMBC.2014.6944227.

[24] K. Yan, X. Wang, L. Lu, and R. Summers, "Deeplesion: Automated mining of large-scale lesion annotations and universal lesion detection with deep learning," *Journal of Medical Imaging*, vol. 5, p. 1, Jul. 2018. DOI: 10.1117/1.JMI.5.3.036501.

[25] A. Obaid, "An efficient systematized approach for the detection of cancer in kidney," *International Journal of Scientific and Engineering Research*, vol. 7, pp. 1–7, Jan. 2020.

[26] T. Les, T. Markiewicz, M. Dziekiewicz, and M. Lorent, "Automatic recognition of the kidney in ct images," in *19th International Conference Computational Problems of Electrical Engineering*, 2018, pp. 1–4. DOI: 10.1109/CPEE.2018.8506777.

[27] Q. Yu, Y. Shi, J. Sun, Y. Gao, J. Zhu, and Y. Dai, "Crossbar-net: A novel convolutional neural network for kidney tumor segmentation in ct images," *IEEE Transactions on Image Processing*, vol. 28, no. 8, pp. 4060–4074, 2019. DOI: 10.1109/TIP.2019.2905537.

[28] K. Yan, Y. Peng, V. Sandfort, M. Bagheri, Z. Lu, and R. M. Summers, "Holistic and comprehensive annotation of clinically significant findings on diverse ct images: Learning from radiology reports and label ontology," in *2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 2019, pp. 8515–8524. DOI: 10.1109/CVPR.2019.00872.

[29] A. M. Osowska-Kurczab, T. Markiewicz, M. Dziekiewicz, and M. Lorent, "Textural and deep learning methods in recognition of renal cancer types based on ct images," in *2020 International Joint Conference on Neural Networks (IJCNN)*, 2020, pp. 1–8. DOI: 10.1109/IJCNN48605.2020.9206655.

[30] A. Aljouie, N. Patel, and U. Roshan, "Cross-validation and cross-study validation of kidney cancer with machine learning and whole exome sequences from the national cancer institute," in *2018 IEEE Conference on Computational Intelligence in Bioinformatics and Computational Biology (CIBCB)*, 2018, pp. 1–6. DOI: 10.1109/CIBCB.2018.8404967.

[31] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," *arXiv 1409.1556*, Sep. 2014.

[32] M. Islam, "Ct KIDNEY DATASET: Normal-cyst-tumor and stone," *Kaggle. com, 2021. [Online]. Available: https://www.kaggle.com/nazmul0087/ct-kidney-dataset-normal-cyst-tumor-and-stone. [Accessed:*, vol. 02-, 2021.

[33] O. Ronneberger, P. Fischer, and T. Brox, *U-net: Convolutional networks for biomedical image segmentation*, 2015. arXiv: 1505.04597 [cs.CV].

[34] V. Badrinarayanan, A. Kendall, and R. Cipolla, *Segnet: A deep convolutional encoder-decoder architecture for image segmentation*, 2016. arXiv: 1511.00561 [cs.CV].

[35] M. Sandler, A. Howard, M. Zhu, A. Zhmoginov, and L.-C. Chen, *Mobilenetv2: Inverted residuals and linear bottlenecks*, 2019. arXiv: 1801.04381 [cs.CV].

[36] K. Simonyan and A. Zisserman, *Very deep convolutional networks for large-scale image recognition*, 2015. arXiv: 1409.1556 [cs.CV].

[37] C. Szegedy, V. Vanhoucke, S. Ioffe, J. Shlens, and Z. Wojna, *Rethinking the inception architecture for computer vision*, 2015. arXiv: 1512.00567 [cs.CV].

[38] *The value of model accuracy*, en, https://algorithmia.com/blog/the-value-of-model-accuracy., Accessed: 2021-9-28, Jul. 2020.

[39] *Introduction to loss functions*, en, https://algorithmia.com/blog/introduction-to-loss-functions., Accessed: 2021-10-3, Apr. 2018.

[40] K. Steward, *Sensitivity vs specificity*, en, https://www.technologynetworks.com/analysis/articles/sensitivity-vs-specificity-318222, Accessed: 2021-10-3, Apr. 2019.

[41] S.-H. Tsang, *Review: MobileNetV2 — light weight model (image classification)*, en, https://towardsdatascience.com/review-mobilenetv2-light-weight-model-image-classification-8febb490e61c, Accessed: 2021-10-3, May 2019.