Leveraging Unsupervised Segmentation for Semi-supervised Renal Calculi and Carcinoma Segmentation and Classification

by

Farhan Faruk
20301137

H.M. Sarwer Alam
20301224

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

Department of Computer Science and Engineering
Brac University
June 2024

Declaration

It is hereby declared that

1. The thesis submitted is my/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:

Farhan Faruk	H.M. Sarwer Alam
20301137	20301224

Approval

The thesis/project titled "Leveraging Unsupervised Segmentation for Semi-supervised Renal Calculi and Carcinoma Segmentation and Classification" submitted by

1. Farhan Faruk (20301137)
2. H.M. Sarwer Alam (20301224)

Of Spring, 2024 has been accepted as satisfactory in partial fulfillment of the requirement for the degree of B.Sc. in Computer Science on June 28, 2024.

Examining Committee:

Supervisor:
(Member)

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

Co-Supervisor:
(Member)

Rafeed Rahman
Lecturer
Department of Computer Science and Engineering
Brac University

Program Coordinator:
(Member)

Md. Golam Rabiul Alam
Professor
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

Indeed it became crucial to develop an AI-driven system for detecting Renal illnesses spontaneously due to a healthcare issue of Renal failure. The global shortage of nephrologists is the core reason for this. This research concerns two crucial diseases: Renal Calculi and Carcinoma by using semi-supervised localization and unsupervised segmentation of a total 5477 CT images of axial point of view in order to use it in commonly used models. The gathered CT images are prepared by resizing and balancing via augmentation and the analysis shows that the mean color distribution of images was the same across all classes. We have used YoloV8 for kidney localization, quick shift and label merge for segmentation. Also, we manually annotated all the images with the supervision of a medical expert. Furthermore, we compare both datasets using VGG19, AlexNet, ResNet50. We got better output in Semi-Supervised localized and unsupervised segmented images rather than the annotated images via experts in terms of classification. We believe that reducing and focusing on the region of interest can easily achieve good output in commonly used models which will be very time efficient and cost effective as well.

Keywords: Renal; Calculi; Carcinoma; CT Image; Segmentation; Classification.

Acknowledgement

We are delighted to express sincere gratitude to our supervisor, Md Golam Rabiul Alam, Professor at the School of Data Science, BRAC University. Whenever we needed assistance and support with our research, his door was always open for us. Moreover, he let us take the lead on this paper, but offered guidance whenever it was required.

We are also indebted to Md Tanzim Reza, Lecturer at BRAC University, and Nafiz Imtiaz Rafin, Research Assistant at BRAC University, for their insights which helped us during the implementation phase.

We extend our heartfelt thanks to our parents for their unwavering support throughout our journey and it is their sacrifices that made our education possible.

To end, we would like to convey our gratitude to the Department of Computer Science and Engineering within the School of Data Science at BRAC University, as well as our respected faculties for their immense support throughout our undergrad journey.

Table of Contents

List of Figures

List of Tables

Chapter 1

Introduction

Chronic Kidney disease has become a major global public health problem which is on the rise [13]. It is expected to become the fifth typical cause of death by overtaking heart disease by the year 2040 [28] and currently, it impacts over ten percent of the global population [14]. Formation of kidney stones known as nephrolithiasis, and kidney tumors called kidney carcinoma. Kidney stone disease affects approximately twelve percent of the global population [27]. Also, kidney disease affects the human body's fluid balance, blood purification, and blood pressure. For instance, vomiting, back pain, fever, and difficulty in urination are few possible symptoms of kidney disease. In fact, 70 percent of the kidney disease cases involve diabetes and hypertension [18][31][16].

Pathological examinations are widely utilized with the help of technology such as CT scans, X-rays or MRI scans to ascertain and diagnose kidney issues. X-ray beams are used in CT scans to create a cross-sectional view, which offers useful 3D anatomical information and they provide precise slice-by-slice images [18]. The scarcity of nephrologists and radiologists, particularly in South Asia poses a significant challenge as there is an inadequate ratio of nephrologists like one per every million individuals, as compared to 25.3 in Europe [37]. Moreover, it is necessary to improve kidney disease diagnostics and a potential remedy is provided by deep learning research in computer vision which can assist medical professionals.

Beyond the research, we believe that the combination of semi-supervised localization and super-pixel-based unsupervised segmentation on the dataset has the potential to improve clinical decision-making using low cost model and enable individualized treatment regimens as well as improved patient care.

1.1 Research Problem

The CT scan images contain the human body's several organs from different points of view. If we want to classify the kidney problem our ROI should be only the kidney not other organs. Using the whole CT scan images required a heavy weight model which needs to run higher configured GPU and CPU's to get a good result. Our aim is to reduce ROI and then ready the images for classification.

1.2 Preliminary

1.2.1 Kidney Stone

Kidney stones have been a very common urinary tract problem for such long time. Preventing the disease from coming back is a significant concern for public health since it is closely associated with an increased susceptibility to long-term kidney dysfunction, kidney malfunction, elevated blood pressure, cardiovascular disorders, and metabolic disorders such as diabetes [24][11][7][6][3]. Also, it might be a symptom of a systemic illness related to other metabolic syndromes [25].

Figure 1.1: Graphical Instance of Kidney Stone.

Kidney stones vary in form, chemical content, and as well as in size depending on urine composition abnormalities. They are categorized into five types based on the mineral they contain as well as the formation of these stones [17][2]. These are-

1. Calcification Renal Stones
2. Renal Stones composed of Ammonium Magnesium Phosphate
3. Urate Renal Stones
4. Cystinuria Renal Stones
5. Renal Stones caused by Medication

Indeed the complex mechanism of renal stone formation is affected by different numbers of internal as well as external variables that include genetic factors, age, sex, dietary habits, geographical location, mineral composition, climate conditions, and hydration levels [4].

1.2.2 Kidney Tumor

Renal Tumor is basically an aberrant proliferation of cells within the renal organ, which can potentially be life-threatening. [1]. Abnormal DNA mutations in kidney cells are the core reason for starting a tumor that causes cells to proliferate quickly and perhaps easily spread into other body organs [10]. Moreover, it can develop inside the kidney or spread to other organs like the lungs [5]. Obviously, patients' regular lives are also impacted by the symptoms of kidney tumors, which might include appetite loss and unexplained weight loss, etc [12]. Over 62,000 new patients with kidney tumors are diagnosed each year in the United States that making up

3.7 percent of all tumors. The risk of this disease increased with age and men are more at risk than women in terms of these diseases [1][46].

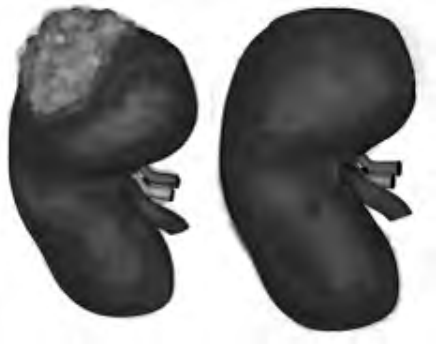


Figure 1.2: Graphical Instance of Tumorous Kidneys VS Healthy Kidneys.

Renal tumors are diseases that include a wide variety of both malignant and benign growths [32]. These diseases usually emerge from different nephron areas and depend on distinctive genetic characteristics, genomic features, and to some extent, clinical symptoms. Various forms of renal tumors are shown in Figure 1.4. Among the major two types of renal tumors, the benign tumor grows slowly and does not contact surrounding organs which means it does not spread to other cells easily [32]. Since benign tumors do not have a serious threat to life, surgical excision often avoid recurrence. Nevertheless, some noncancerous tumor forms like benign [32] can potentially become malignant. So, it is important to pay attention to their existence and get the proper treatment within time.

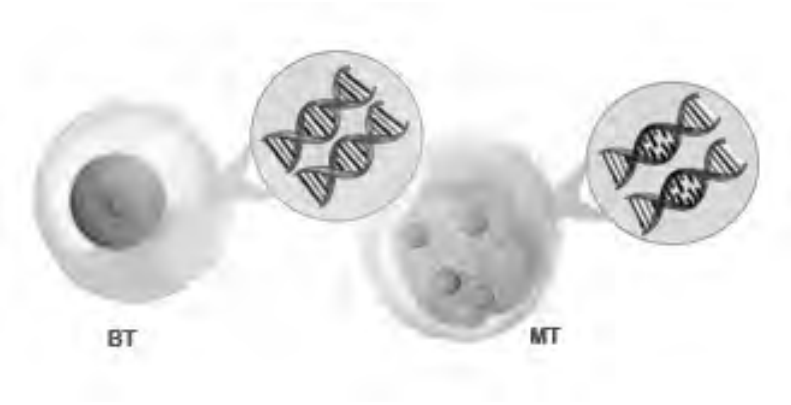


Figure 1.3: Graphical Instance of kidney tumor types.

Tumors often originated from different nephron regions and these tumors exhibit distinct genetic and clinical characteristics, with the majority of these tumors being either malignant or benign [32]. Benign tumors are often removed surgically since their growth is comparatively slow, and chances are less to spread or infect nearby tissues [32]. Despite the fact that there are many different kinds of benign tumors and if they are not treated properly, they pose little risk to life and this delay can even cause the development of cancer in kidney tissue.

A malignant tumor usually originates in the kidney and this tumor has the ability to metastasize because the carcinoma cells infiltrate neighboring tissues and contin-

uously initiate new tumors in various parts of the body with the help of the human bloodstream or the lymphatic system. Metastasis is also a major contributor to the increasing mortality rate due to cancerous tumors [47]. Surgical operation is typically the initial treatment for carcinomas along with the use of radiotherapy to eliminate any residual cells which can be malignant, sometimes they just use radiotherapy because if it is not the initial stages of the tumor then the surgical removal of the tumor becomes challenging.

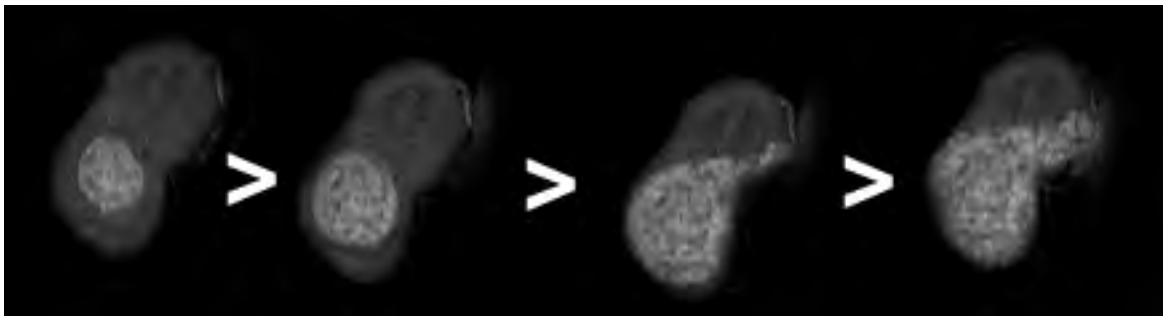


Figure 1.4: Graphical Instance of kidney tumor stages.

Malignant tumors (MTs) are malignant growths that begin in the kidney and spread or infect nearby tissues which can be the reason for forming new tumors known as metastasis or secondary tumors and these are the significant factor in cancer-related death [47]. Surgery is usually the first step in treating carcinomas, and if there are any leftover cancerous cells within the kidney then these are removed later using radiation or chemotherapy in an advanced condition.

1.2.3 CT Image

A non-invasive procedure known as radiography imaging helps doctors in terms of making accurate diagnoses and providing care to enhance the quality of life as well as treatment for patients. This radiography imaging method is frequently used to detect tumors [46][20]. The diagnosis of kidney tumors must be made precise and fast. However, doing this only based on clinical characteristics is insufficient in this regard. For these reasons, medical personnel uses radiography imaging techniques in terms of identifying kidney tumors [2]. There are a lot of radiological tools used in the medical sector. However, we are only using CT scan images in our research.

Table 1.1: Benefit and Limitation Of CT Scan Image.

CT Images	
Applications	Identify the type of tumor, and can provide images of infected tissues, organs, and skeletal structure
Benefits	No need of doing a biopsy, since it increases the chance of spreading the tumor
Drawbacks	Quite complex in terms of assessing the kidney vessels

1.3 Research Contributions

In this research, we focus on reducing the ROI that can effectively classify kidney stones and tumors with better output in lightweight models than the clinically annotated data on heavyweight models. Using semi-supervised localization and unsupervised segmentation methods based on super-pixels, we can precisely recognize and delineate these abnormalities with the help of CT images. The major contributions of this research paper are summed up as follows:

- We introduce a novel approach of kidney localization on CT images rather than analyzing the whole CT image. This approach shows that reducing the ROI can comfortably give better output even in lightweight classifying models.
- We explored the unique relationship between the Aorta vessel characteristics in the CT images of stone-infected, tumor-infected, and normal kidneys.
- We initiated a modified version of quick-shift segmentation by doing label merge, which reduces the creation of cluster in the images during unsupervised segmentation.
- We created a full dataset of manually annotated kidney stones and tumors under the supervision of medical experts, and we demonstrated the classification output using this supervised segmentation without reducing the ROI.
- This manually annotated or supervised segmented dataset that we created can be helpful to many researchers in this field in the future.

1.4 Organization of the paper

The format of the following paper: Chapter 2 contains the background study and literature review of the research, and in Chapter 3 the method of every process is reviewed and described briefly. Next, Chapter 4 has the findings and results of our thesis and analysis. In the end, the conclusion, limitations, and future work are mentioned in Chapter 5.

Chapter 2

Literature Review

To understand how we need to process our work further and get more knowledge about our topic we reviewed some papers. We started to wonder how efficient Machine learning (ML), Deep Learning (DL) which is a part of machine learning will help us to identify kidney cysts, stones, and tumors from radiography images. Then we tried to see how they conduct those research and what kind of steps made their paper unique.To understand the hope of ML and DL in kidney diseases detection we have come up with a table which reflects the previous works and their success rate (accuracy) on this sector-

Table 2.1: ML and DL work on kidney disease

Reference	Year	Topic	Dataset	Model	Accuracy
Parakh et al. [30]	2019	Renal stone detection	CT Images	CNN	0.9500
Yildirim et al. [38]	2021	Renal stone detection	CT Images	Swin transformers	0.9682
Islam et al. [43]	2022	Renal stone detection	CT and urogram images	Swin transformers	0.9930
Li et al. [44]	2022	Renal stone detection	CT images	Res U-Net	0.9995
Caglayan et al. [41]	2022	Renal stone detection	CT images	XResNet50	0.9300
Our work	2022	Renal stone and Tumor detection	CT images	ResNet50	0.9865

Here are some paper review has been given below -
From the paper of Gharaibeh et al.[42] we got more information about some works in recent years of Renal Tumor detection and those are - there was 4 information which based on CT Images database and their applied system and accuracy, and working years are respectively ML(SURF, HOG ,LBP , MMD) - 0.9500 - 2015, DL (InceptionV3, DAG-SVM, CNN) - 0.9300 - 2019, DL (GoogleNet, CNN) - 0.9700 -

2019 , DL (CNN(Modified) , 3 cross folds) - 0.9000 to 0.9900 - 2021. After that they showed information of a US (Ultrasonography) based DL model of 2015 which used ROI, PCA and NN with a performance of 0.9700. After that there was another study which was conducted in 2018 on the same paper where the NCA and LSTM model of ML were used on MiRNA and CT images where the results were 0.9500. Lastly, a study of 2020 based on MRI images was held on model GBM and RF of ML which got the result of 0.8120.

From the table 2.1 and above para, it is clearly reflected that machine learning and deep learning bringed a huge change in the kidney disease detection sector which can solve the shortage of nephrologist problems which we have discussed in chapter one. Now, lets see some paper review which can give us more detail ideas about how they conducted their works and what are the main points that is why they got almost high level accuracy also is there lackings -

Deep learning algorithms were used in research by Ghalib et al.[19] to detect kidney tumors with the help of CT scan images. The dataset (CT scans) of this research were disclosed and they collected those data from several hospitals which are located in Bangalore. To remove the noise they used contrast-limited adaptive histogram and a hybrid type filter during data preprocessing. For clustering, features were retrieved using a developing a Map which is Self-Organized in short SOM. They built the model using an ANN (Artificial Neural Network) technique, which determined categorization based on visual appearance patterns. This model performed well in categorizing tumors as normal or abnormal, with an average execution time of 0.85 seconds.

Liu et al. [22] used machine learning approaches to detect exophytic kidney tumors with the help of CT images in their research. CT colonography images were utilized in 141 instances with the help of exophytic renal lesions, another 38 cases based on endophytic renal lesions, and last 71 cases they did not use renal abnormalities. For picture segmentation, a belief propagation strategy was used, with a special focus on the of kidney lesions's search region. The kidneys were segmented on both the right and left sides. Feature extraction used a feature descriptor based on manifold diffusion to search for protrusions due to the lesion. According to the experimental data, the suggested model performed admirably, with a sensitivity rate of 95% for exophytic lesion detection and 80% for endophytic lesion detection.

Uhlig et al.[34] used machine learning algorithms on the images through CT scans to categorize kidney tumor subtypes in research. Their dataset included 97 scans of patients from prior research that has kidney tumors as well as 107 scans of new patients collected from a hospital which name is University Medical Center Goettingen and the subtypes are AML, clear cell, chromophobe, and papillary. Manual segmentation using 3D Slicer and PyRadiomics was used to extract those features. This model used the random forest for multi-class type classification, and it performed well, with a 78% AUC (Area Under the Curve) when the oncocytomas subtype was excluded.

Akram et al.[35] used medical image segmentation and algorithms which are based

on deep learning to segment kidney images and predicted kidney tumors in the research. The number of data is 300 CT images(KITS19 dataset) from patients with kidney tumors they have used in their datasets and used preprocessing techniques, for example- morphological filtering. For classification, a type of modified Convolutional Neural Network (CNN) was applied in their study, with segmentation accuracy ranging from 90% to 99%. This model also predicted tumor detection with a 70% to 100% accuracy level.

By using 4 VNet models (FusionVNet, VNet, Hybrid VNet mode, ET-Net) in a paper which was conducted by Türk et al.[33] proposed a model which type is hybrid by using VNet models which have more flexible structure and adaptability. This model performed an excellent result in accuracy score with Dice coefficient (average) of 0.9770 for the renal segmentation and for tumor segmentation the number is 0.8650. From 210 datasets of KiTS19 they prepared the clinical features, the data imaging, the renal borders, and tumor (borders) with the help of manual segmentation method and the CT images were converted into 16*256*256 (Pixel value splitted by 255 with 0-1 standardization). In addition, this model needs five days to compute its result with the help of Graphic Processing Unit (GPU) which is the NVIDIATeslaV100 (Graphics Size-32 GB, NVLink). The more beneficial part of this model is it is capable of determining the kidney tumor thickness, width, and location.

Recently, a good number of deep learning algorithms and models have been implemented on kidney disease detection but there is always a problem of rigorous data training process and high memory cost. Bhandari et al.[45] discussed this problem and came up with a solution in their paper. In their research they proposed a lightweight CNN and demonstrated a DL model with AI (XAI) to detect kidney cysts, stones, and tumors. This study used a publicly available dataset that contained information on patients who had, on average, 3709 cysts, 5077 normal cells, 1377 stones, and 2283 malignancies.Their working plan was like they input CT images, then pre-processed the data basically, they resize the images into 150x150x3 and normalize them to fit (0,1). After that the data was trained and tested in Light Weighted CNN. After getting the result they performed analysis based on accuracy, precision, recall, F1 Score and Roc. Lastly, they used explainable AI to lime and sharp. Here is a list of their radiological result and Impressions-

1. "Right renal pelvis simple cystic lesion" which gives the impression is "Parapelvic cyst on the right."
2. The impression of "The size of both kidneys seems to be normal, and the parenchyma has a normal width and shape." is "Bilateral kidneys in the indicated area appear normal."
3. From "Right kidney revealed a radiopaque lesion." the impression was created like - "kidney revealed a radiopaque lesion. Kidney calculi on the right."
4. Last the impression created by "Enhanced lesions that are heterogeneously distributed most likely came from the left kidney." is "Tumor growing from the left kidney."

Bhandari et al.[45] also included, in terms of training, validation, and testing, the model's accuracy was 99.30 +/- 0.18, 99.39 +/- 0.99%, and 99.52 +/- 0.84%. The

outcomes produced by this XAI algorithm were verified by radiologists. Despite having a simpler architecture, our model fared better when identifying kidney illness using CT scans.

Next, Bayram et al.[40] used a deep learning-based xAI-assisted CAD system to detect kidney disorders. The main technique or method used in this study was explainable artificial intelligence (xAI) and YOLO architecture designs for finding kidney stones and cysts. In this research, this model's performance has been assessed using CT scans of 192 kidneys in good health, 394 kidneys with stones, and 72 kidneys with cysts. Finally, YOLOv7 performed better than YOLOv7 Tiny, achieving a mAP50 of 0.85, accuracy, sensitivity, and F1 scores of 0.829, 0.882, and 0.854 respectively. Let's compare their results of YOLOv7 and YOLOv7 Tiny for better understanding. For the classes of Kidney, Kidney Stone, Kidney Cyst the mAP50's value of YOLOv7 is respectively 0.973, 0.849, 0.727 on the other hand in YOLOv7 Tiny the values are 0.969, 0.703, 0.181. If we average the value then for YOLOv7 the value is 0.850 which is 0.232 points more than the YOLOv7 tiny. Next, for PRE the values are 0.936, 0.819, 0.892 of kidney, kidney stone and kidney cyst of YOLOv7 but in YOLOv7 tiny the values are 0.912, 0.789, 0.473. Here, YOLOv7 is also ahead with 0.882 points of average score. In STV if we see the average score of points then YOLOv7 has 0.829 but YOLOvs tiny has only 0.654. In F1 Score, YOLOvs tiny's score for kidney, kidney stone and kidney cyst is 0.946, 0.747 and 0.687 but the scores in YOLOv7 are 0.966, 0.966, 0.740. Here, also YOLOv7 beats YOLOV7 tiny by 0.158 point. Moreover, it is notable that YOLOv7 scored 1 in the kidney score of STV. So from this data it is clear that YOLOv7 is better than YOLOV7 tiny.

Alzu'bi et al.[39] proposed three new 2D CNN models with CNN-6 (6 layers), a ResNet50 (50 layers), and a VGG16 (16 layers). Their research model used 8400 images of 120 adults as the dataset whose accuracy is respectively 97%, 96%, and 60%. Moreover, their newly developed CNN-based network model, VGG16, performance for the top-5 test accuracy of ImageNet was 0.927. Let's understand how CNN architecture helps to detect and classify the kidney tumor, For kidney tumor detection, First they took the input data. Then they normalized (branch) those data. After that they converted that data (basically images file) to 2D images. Then they performed a max polling 2D system on that. After that they dropout all the layers, later which layers were flattened and nodes were created. In a diagram they showed some nodes are hidden. On the output node one is normal and the other is tumor. This is how CNN worked to identify kidney tumor. For kidney tumor classification, Like before first the data was imputed and after that images were converted into 2D layers. After that Max Pooling 2D and flattened where performed. Here also some nodes are hidden and in output nodes some are bening and others are malignant. By this technique cnn classify the kidney tumors.

Islam, M.N. et al.[43] worked on 6 models which are from Vision transformers and deep learning models. From their findings they found the Swin transformer is not only more accurate with 0.993 but also it trains quickly than others. The VGG16 model is better than Resnet50 and Inceptionv3 by exhibiting the necessary anatomical defects in detail, according to the study's attempt to uncover the blackbox of these models.

Chapter 3

Methodology

Our research involves processing the CT images twice for the classification. Initially, CT images undergo manual annotation as pre-processing. Later the original image again processed with semi-supervised localization using YOLOv8 and followed by unsupervised segmentation via Quick Shift and Level Merge. After that the processed images are cropped, resized, and both the prepared datasets are split into training, testing, and validation sets. Data augmentation is applied only to the training sets to address the imbalance between the classes before feeding them into classification models like ResNet50, AlexNet, and VGG19. Lastly, the paper analyze the classification results for both approaches.

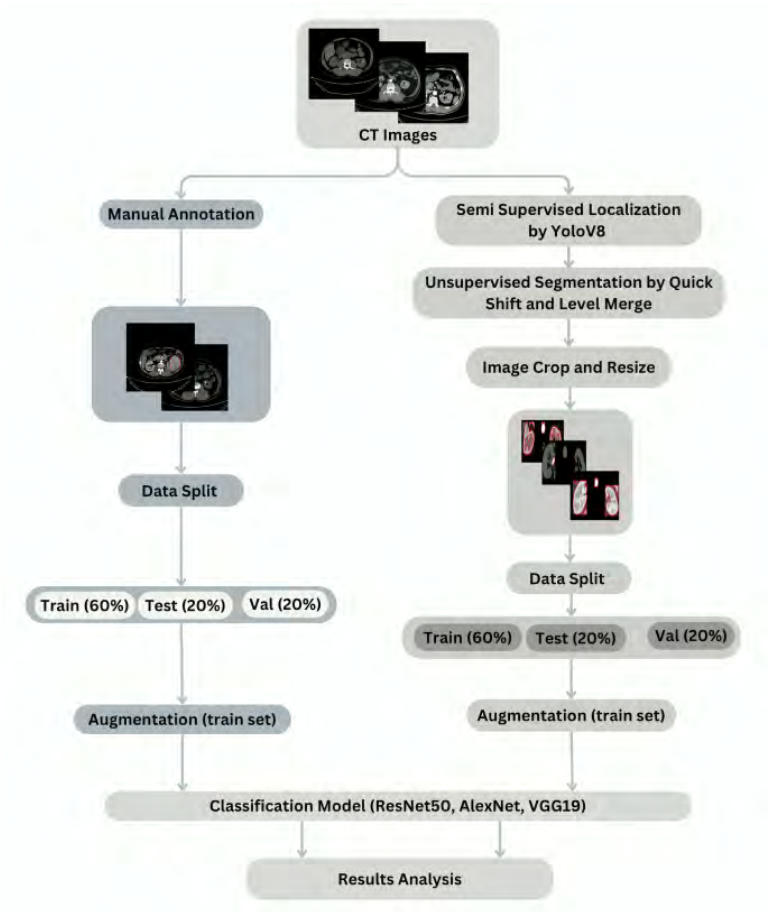


Figure 3.1: Top Level Overview of the Proposed System.

3.1 Data Description

The "CT Kidney Dataset - Normal, Tumor, Cyst, and Stone" is an invaluable resource available on Kaggle, carefully maintained by Nazmul0087, and designed for cutting-edge research in kidney health utilizing Computed Tomography (CT) images. This huge collection includes a wide range of DICOM format images that give a thorough perspective of kidney disorders, including normal anatomical features, possibly malignant or benign tumors, and the presence of solid mineral deposits known as kidney stones. Along with these images, the dataset is strengthened with painstakingly created metadata that provides crucial patient specific and scan-related information.

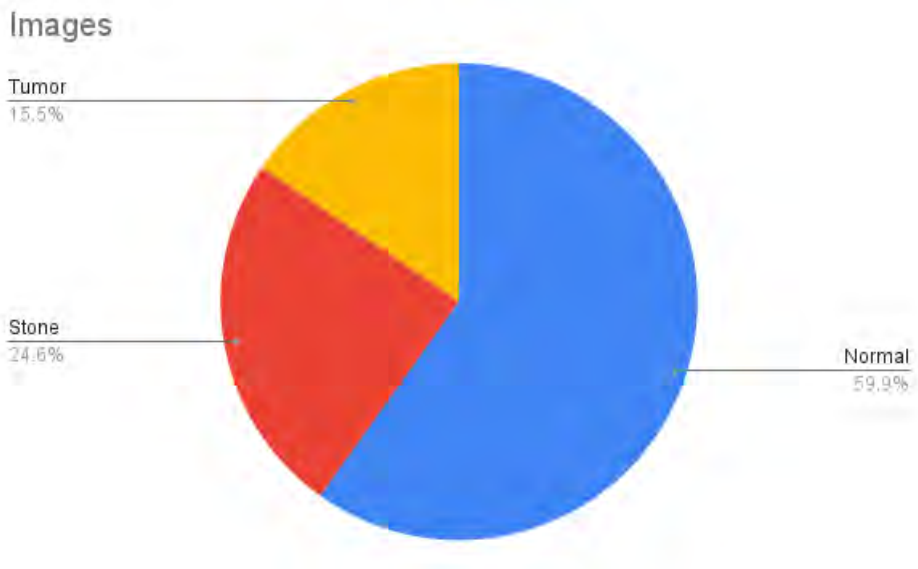


Figure 3.2: Visual Representation of Dataset.

This dataset serves as a cornerstone for research aiming to automate the detection and categorization of kidney disorders using advanced radiographic image processing. It has a wide range of potential uses, including machine learning-driven classification, accurate segmentation of regions of interest within kidney scans, and complex feature extraction for high-level diagnostic studies. Researchers who utilize this dataset are urged to properly credit Nazmul0087 as the inspirational curator and cite the source in accordance with Kaggle's policies. The fact that Nazmul0087 was so generous and committed to making this information available for research and improvements in kidney health diagnoses, is what makes it such a tremendous gift to the field of medical imaging research.

3.2 Data Preprocessing

3.2.1 Manual Annotation

As a comparative study in our research, we want to see how supervised segmented CT image works in our classification models. We went to a renowned medical college along with some specialized permission and we have manually annotated all the 848 CT images of stone and 1350 CT image of tumor under the supervision of medical expert and that's how we have created a different dataset of 2198 supervised segmented images along with non-annotated 3279 images using the same original dataset.

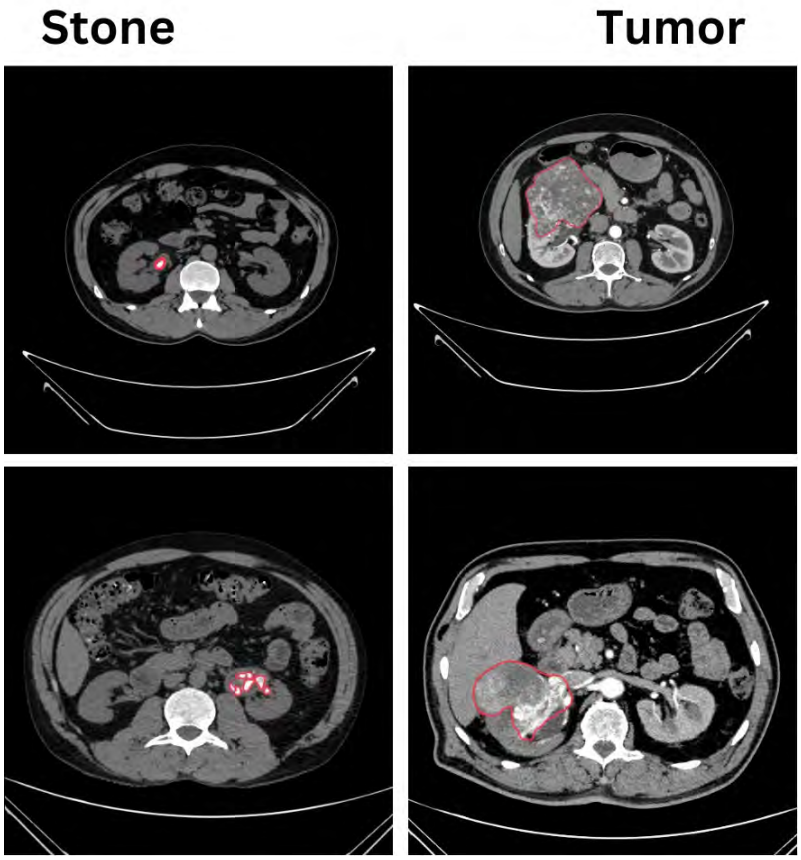


Figure 3.3: Visual Representation of Manual Annotation.

3.3 Proposed Unsupervised Segmentation Method

3.3.1 Semi-Supervised Kidney Localization

We have used the YoloV8 model to do object localization in the CT image. Firstly, we extract the most unique sets of 20% from the original dataset of 5477 CT images. Then, we upload those 1095 CT images of Tumor, Stone and Normal kidney in Roboflow framework. After that, by using a bounded box we annotate the Left Kidney, Right Kidney and Aorta Vessels in those CT Image. Then, we trained the YoloV8 model by using those annotated images and finally we tested the Yolov8

object detection model on the original dataset of 5477 CT images to generate the annotated output using the bounding box. Moreover, we vanish every other organ of the CT images without those three bounded box regions, however we did change the image size or ratio. This background removal technique helped us to reduce ROI a lot. Because we just have to analyze only the region inside the three bounded box for our research.

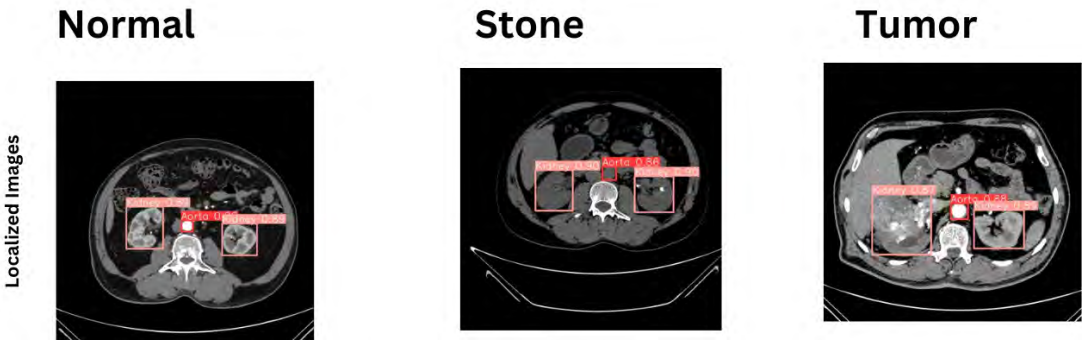


Figure 3.4: Visual Representation of Kidney Localization.

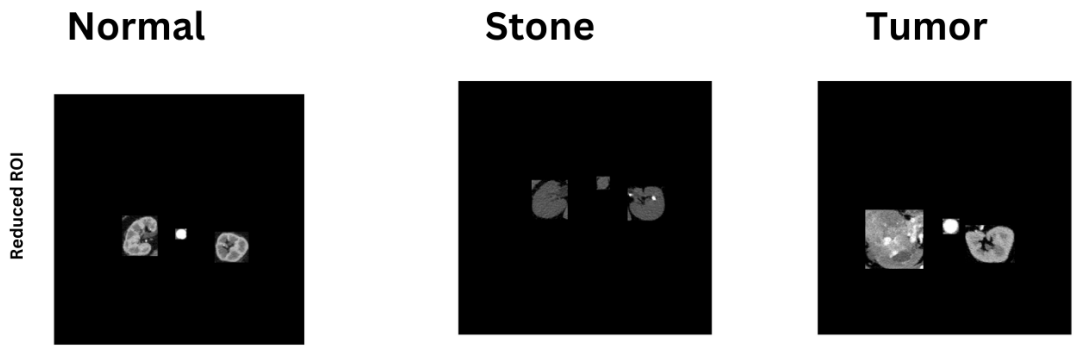


Figure 3.5: Visual Representation of Reduced ROI.

3.3.2 Unsupervised Segmentation

Image segmentation is an undoubtedly crucial method that separates an image into perceptual sections based on comparable properties like color, intensity, texture, etc. It has been widely used in terms of identifying the object [15], tracking an object [8], and semantic segmentation [23]. This image segmentation is difficult as well as a fundamental issue in computer vision. Moreover, an image processing method known as super-pixel-based unsupervised segmentation divides an image into coherent sections by grouping pixels based on visual similarity. Again, this method is also used for object identification and scene comprehension to give that particular image a more meaningful representation.

In our research we used the Quick-Shift algorithm which is an unsupervised image segmentation technique that clusters pixels by identifying density peaks in feature space. In these clusters, each pixel is represented by its spatial coordinates and color values. Moreover, the algorithm estimates the local density of pixels and computes

a shift vector for each pixel, directing it towards regions of higher density. Pixels are translated in both dimensions until they reach the mode or the local density peak. Segments are formed by pixels that overlap the same mode. Some of the parameters of interest include the kernel size which defines the range of proximity for density estimation, and the max distance that regulates the segmentation resolution. Lastly, to reduce the number of clusters and improve segmentation quality, label merging is applied along with Quick-Shift, combining similar segments to form larger, more coherent regions. also , Quick-Shift with label merging is effective but quite sensitive to parameter settings.

Figure 3.6: Visual Representation of Unsupervised Segmentation.

3.3.3 Data Resizing

In our research, we have used images of axial point of view data from "CT Kidney Dataset - Normal, Tumor, and Stone". Firstly, we crop the CT image to zoom and reduce the ROI. Then, we determine the maximum, average, minimum dimensions of all the CT images since those images have a variety of dimensions. Moreover, we resize the cropped CT images into 224x224x3 aspect ratio to make them more suitable for classification models. In short, we have turned all the training, validation and testing sets into these specific dimensions as input for classifying models using the OpenCV library of python.

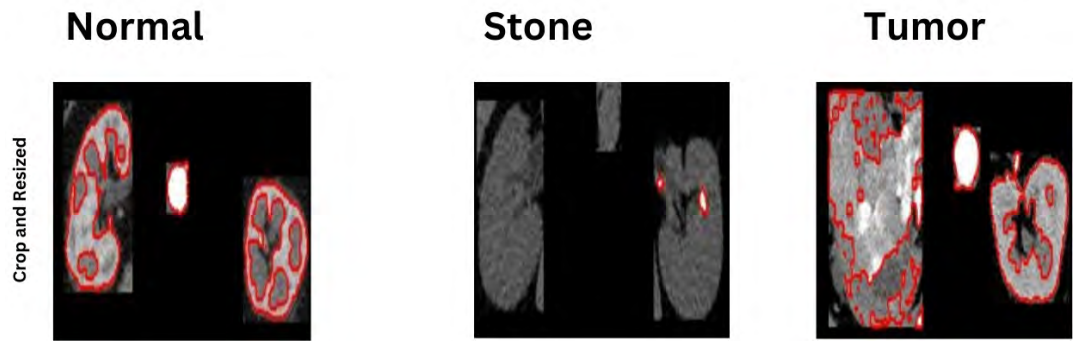


Figure 3.7: Visual Representation of Resized Data.

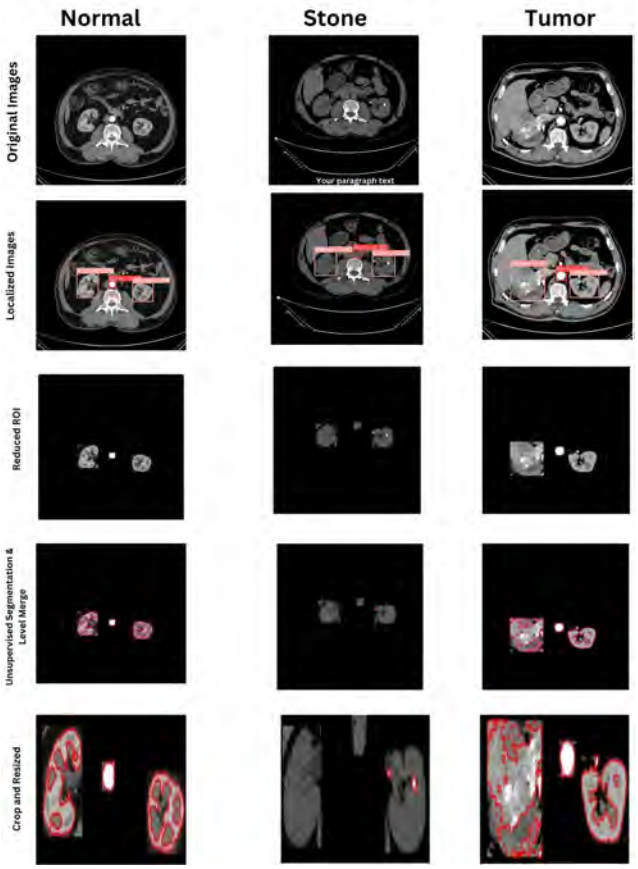


Figure 3.8: Workflow of Image Modification.

3.4 Data Spitting

We split Both the unsupervised segmented dataset and supervised segmented dataset in the data preprocessing part. Since both the dataset contains a total of 5477 CT kidney images, we split all the images of both dataset into three pairs of sets across the three different classes: Train that contains 60% and Test that contains 20% of CT images and lastly the validation set that that contains again 20% of CT images. The final distribution of training, validation and testing set is given below:

Table 3.1: Dataset Distribution

Class	**Train**	**Validation**	**Test**	**Total**
NORMAL	1967	655	657	3279
STONE	508	169	171	848
TUMOR	810	270	270	1350
Total	3285	1094	1098	5477

3.5 Data Augmentation

The technique of artificially generating new data from pre-existing data is known as data augmentation. We already mentioned that we have used a total 5477 CT images including three classes - Normal, Stone, Tumor. The dataset contains 848 CT images of Stone, 1350 CT images of Tumor and 3279 images of Normal kidney. This

shows a clear imbalance between the three classes of both the pre-processed dataset and to balance the dataset we have used an augmentation method in the train set which results in the same 1967 images for all the three classes in the train set. While augmenting the images, we have used horizontal flip, vertical flip, rotation, random brightness contrast, gaussian blur and random gamma function.

Table 3.2: Augmentation Distribution

Class	Before Augmentation	Augmentations Needed	After Augmentation
NORMAL	1967	0	1967
STONE	508	1459	1967
TUMOR	810	1157	1967

The motive of this augmentation method was to balance the dataset by increasing the number of samples in the "STONE" and "TUMOR" classes to match the number of samples in the "NORMAL" class in the Train set. This ensures that each class has an equal number of 1967 images which is highly likely to improve the performance of a ML Classification model trained on this updated train set which clearly prevents class imbalance issues.

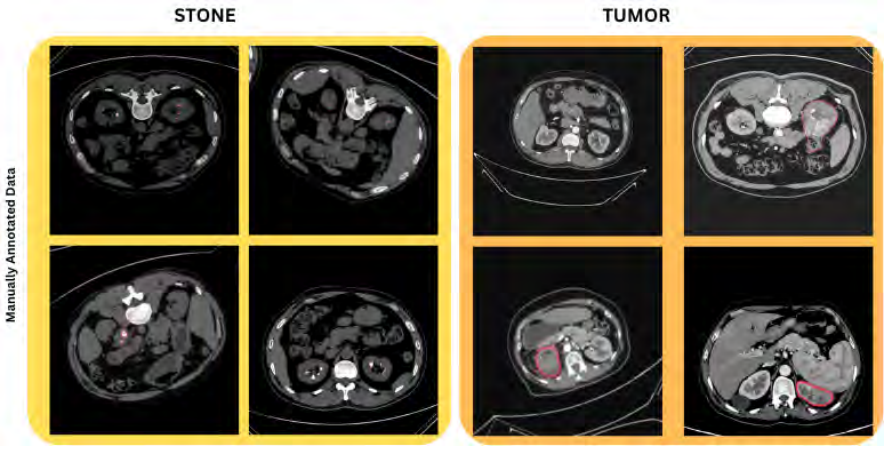


Figure 3.9: Visual Representation of Augmentation - Manual Annotation.

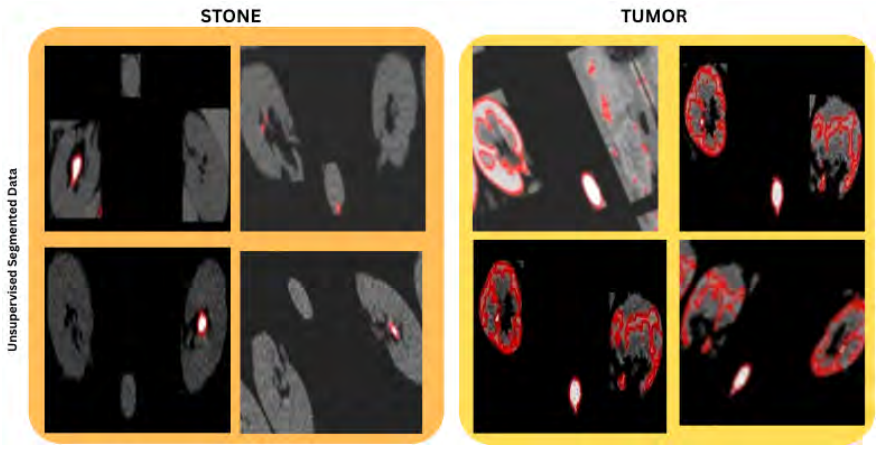


Figure 3.10: Visual Representation of Augmentation - Proposed method.

3.6 Data Normalization

Normalization means organizing the dataset within a relational dataset on the basis of some predefined rules, known as normal forms. This technique aims to enhance data integrity and minimize the data redundancy by grouping similar data in the same column which prevents data clustering. In our case, for both the pre-processed datasets, we normalized all 5477 CT images twice or scaled their pixel values by dividing each pixel value by 255, since the maximum grayscale value of an image is 255. After normalization, all the pixel values in our CT images fall within the range of 0 to 1.

3.7 Model Specification

After performing all the necessary steps of data preprocessing for Both "semi-supervised localized unsupervised segmented dataset" and "manually annotated supervised segmented dataset", we have done our classification part by train and testing both datasets on commonly used models such VGG19, AlexNet and ResNet50 for our stone-tumor-normal classification.

3.7.1 Transfer Learning

Transfer learning is a ML technique that involves using a model that is initially trained on a large dataset, such as ImageNet. Later, it is reused and fine-tuned on a smaller dataset to perform a specific task. Although finding large datasets like ImageNet can be challenging, they are crucial for training models effectively. To overcome this issue, we are using transfer learning leveraged models that have already been trained on extensive datasets like ImageNet. The core idea is to achieve good performance on different and smaller datasets by utilizing the common features learned during the training stage on a large dataset [29]. A key advantage of deep learning is its ability to transfer learned generic features across various domain tasks, even with small datasets [29].

In our research, we employed the fixed feature extraction method where we kept all layers frozen to retain the learned weights, discarded the final fully connected layer(s), and replaced them with new fully connected layers tailored to our specific classification problem. Basically, we remove the final fully connected layer(s) to transform the network into a fixed feature extractor after a network is pretrained on a large dataset. The remaining parts of the network are left unchanged to serve as the fixed feature extractor. This process is illustrated below-

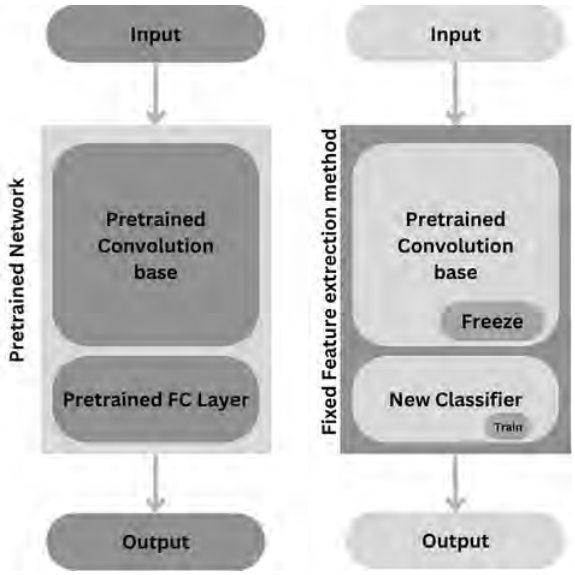


Figure 3.11: Visual Representation Transfer Learning.

For classification task, we utilized transfer learning models that include VGG19, AlexNet, and ResNet50.

ResNet50

Talking about the ResNet50 model, ResNet stands for Residual Network, which is a specific type of CNN introduced in 2015 [26]. ResNet50 is a variant that has been trained on over 14 million ImageNet dataset's images. This model is 50 layers deep and consists of approximately 25.6 million parameters. In terms of Residual Network, the learning process focuses on residuals, which are the distinction between the input features and the features learned by the layer [36]. The graphical representation of ResNet50 architecture is given below:

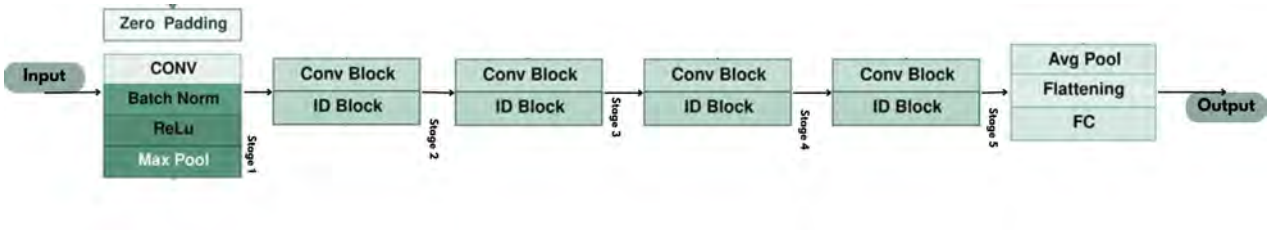


Figure 3.12: Architecture of ResNet50.

VGG19

Talking about the VGG19 model, it is another CNN model which was introduced by the Visual Geometry Group known as VGG at the University of Oxford in 2014 [21]. Again, VGG19 is a variant that has been trained on ImageNet's large dataset that contains over 14 million images across 1,000 classes. This network consists of 19 deep layers and incorporates around 143.7 million parameters. It uses very small (3x3) convolution filters. The network's depth and the small size of the convolution filters

enable it to capture intricate patterns in the data while maintaining computational efficiency. The graphical representation of VGG19 architecture is given below:

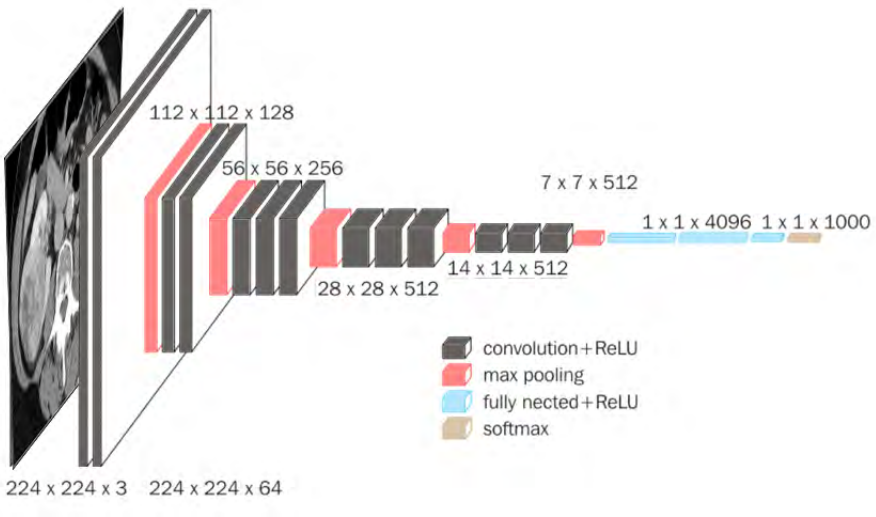


Figure 3.13: Architecture of VGG19.

AlexNet

Talking about the AlexNet model, it is also a CNN that was introduced in 2012 that marked as a significant breakthrough in the field of computer vision [9]. Again it was also trained on ImageNet's large dataset. This network consists of 8 layers, with 5 convolutional layers followed by 3 fully connected layers, and includes around 60 million parameters [9]. AlexNet played a pivotal role in demonstrating the potential of deep learning in image recognition and classification tasks by significantly outperforming previous models. The graphical representation of AlexNet architecture is given below:

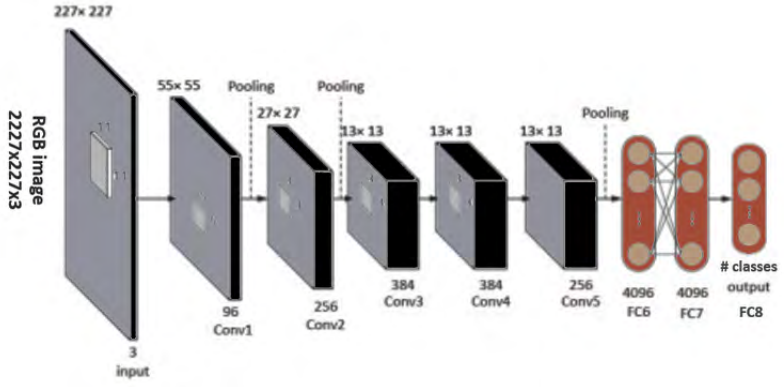


Figure 3.14: Architecture of AlexNet.

Chapter 4

Result Analysis

4.1 Experimental Study

In our work we have used standard performance measures like classification - (Precision, accuracy, recall, f1-score and model history to see our results. We also saw the confusion matrix to understand the model performance-

4.1.1 Accuracy

Accuracy means how the model actually performed, normally it represents as percentage or value from 0 to 1. Basically, it compares the number of correctly categorized samples to the number of samples tested. The four parameter of classification are-
1. True Positive (TP): Correctly predicted positive.
2. True Negative (TN): Correctly predicted negative
3. False Positive (FP): Incorrectly predicted positive.
4. False Negative (FN): Incorrectly predicted negative.

Assume - TP = a, TN = b, FP =c, FN = d
Formula-

$$\text{Accuracy} = \frac{a+b}{a+b+c+d} \tag{4.1}$$

4.1.2 Loss Early Stop Epochs

In this method after running all the epochs the model calculates the training and validation accuracy and loss. This part was handled by the model, we set early epoch stop 10. If in the 10 epochs the validation loss does not increase then the best checkpoint will be returned.

4.1.3 Recall - Precision - F1 Score - Dice - IoU

Sensitivity (Recall)
Sensitivity is also called recall and it means that the rate at which the model captures positive actual diagnosis cases. The fourth metric aids the measurement of the model's potential to classify actual positives. Statistical sensitivity is the formula that is used to measure the percentage of positive cases that are detected by the

model to be reduced to the lowest levels possible including false-negative cases.

$$\text{Recall} = \frac{a}{a+d} \tag{4.2}$$

Precision
Sensitivity is the rate at which the population of positive samples that were predicted to be positives is calculated as the number of correctly predicted positive samples divided by the total number of positive samples. It is the metric that is defined as the comparison between the total number of true positives and the actual positives according to the model's predicted cases. The average accuracy shows the propensity of the model to make a high number of true results so that the false positives are low.

$$\text{Precision} = \frac{a}{a+c} \tag{4.3}$$

F1 Score
F1 score as its name suggests is an indicator that expresses both the precision and the sensitivity of the algorithm used. It possesses a valuable and balanced measure of a model performance and it is especially useful when there is an imbalance in your class distribution. The F1 score is a measure that can be used to reduce the number of False Positives or False Negatives of the F1 score.

$$F1 = 2 \cdot \frac{\text{Precision} \cdot \text{Recall}}{\text{Precision} + \text{Recall}} \tag{4.4}$$

Dice
The Dice Coefficient is also known as the Sørensen–Dice index which is a statistical tool to measure the similarity between two sets of data. In the case of image segmentation, it compares the predicted segmentation mask with the ground truth mask.
Let us assume-

- A means the set of pixels in the predicted segmentation.
- B means the set of pixels in the ground truth segmentation.
- $|A \cap B|$ means the number of pixels common to both sets - predicted and ground truth (the intersection).

So the formula is-

$$\text{Dice Coefficient} = \frac{2 \times |A \cap B|}{|A| + |B|} \tag{4.5}$$

IoU
Intersection over Union which is also known as the Jaccard Index, measures the overlap between the predicted and ground truth segmentation masks. The calculation method of Iou is dividing the area of overlap by the area of union of the two sets.
Let us assume-

- A = Set of pixels in the predicted segmentation.
- B = Set of pixels in the ground truth segmentation.
- $|A \cap B|$ = Number of pixels common to both sets - predicted and ground truth (Intersection).
- $|A \cup B|$ = Number of pixels in either set (Union).

Formula-

$$\text{IoU} = \frac{|A \cap B|}{|A \cup B|} \tag{4.6}$$

4.2 Implementation

For our work we have used two Mid Range Config PCs- 1. Nvidia GTX 1050 Ti, Intel Core i5 7th Gen, 2. Nvidia GTX 1650, Core i5, 13th Gen. In both PCs the python version was - Python 3.12.3, TensorFlow Version 2.16.1, Keras Version 3.3.3 and Cuda- 12.1. In order to save time we used 224x224 images. We are using the YoloV8, ResNet50, VGG19, AlexNet model in our whole process.

We tested the both dataset (Manually annotated data and kidney localized with unsupervised segmented dataset) using ResNet50, AlexNet and VGG19. Also we tried to compare the segmentation matrix of stone images. The results so far we got from the models are mentioned in 4.3.

4.2.1 Outcome Evaluation - I

Kidney Localize - YoloV8

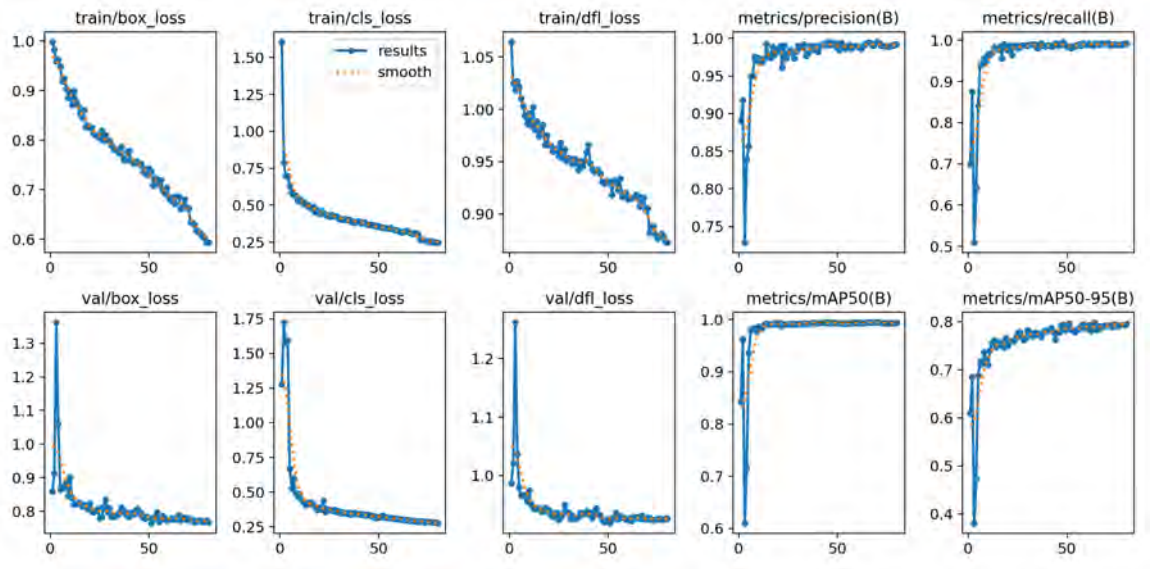


Figure 4.1: Visual Representation of YoloV8 result Analysis.

The graphs plot the performance for the YoloV8 during training and validation across several epochs (Total 80 epochs). The train/box loss graph demonstrates the

general trend toward the decreasing loss, which means that the model learns how to improve bounding box predictions during training. Likewise, the train/classification loss plot shows a pattern of decreasing loss which can be attributed to the model's ability to correctly classify objects in the dataset. Another graph provided, namely the train/distributional focal loss (DFL) graph, is also decreasing, indicating that the model is getting better at learning to select the right part of the dataset. The precision metric stays at a high level and does not fluctuate significantly from test to test showing that the model is able to identify true positives among predicted positives. Precision takes off from a high starting point and continues to grow, indicating the model learns to correctly identify many true positives.

Box loss on validation data decreases, but at a slower rate in comparison with training which means that even though generalization works good enough, there is still room for improvement. The other validation loss classification has a similar trend in as losses decrease and stabilize showing that performance held on the unseen data. The validation DFL also follows the general trend of the correspondent training DFL, and continues increasing throughout the experiment. The Average Precision at 50% (AP50) also stays high – used to measure the performance of an object detector at an overlap threshold of 50 percent of the entire image. Moreover, the mAP of various IOU (50% to 95%) also increases and becomes stable showing the performance of various detections' levels. The graphs show that there is efficient learning through continuously decreasing losses and high precision and recall for validation metrics being close to the training metrics which indicates good generalization to new data.

Table 4.1: Confusion Matrix YoloV8

Actual / Predicted	**Aorta**	**Kidney**	**Background**
Aorta	178	0	0
Kidney	0	320	2
Background	0	4	0

The performance of the YoloV8 is also shown by the confusion matrix. It classifies 320 kidney as kidney (320 KY) and classifies 178 aorta as aorta (178 A) with only 3 other kidney instances misclassified as Aorta. It is also able to pinpoint correctly 2 occurrences of Background. This indicates that the model is effective, the performance in Kidneys also suggests that the model is effective in classification and the low number of misclassifications infers that it is possible for the model to perform better in distinguishing between Aorta and Kidney. This analysis should demonstrate that the classification accuracy for different categories requires further improvement.

Here is a demo of label and prediction during batch 1, by seeing them side by side we will be able to unserstand how perfectly the model is predecting the kidney localization-

(a) Label (b) Predicted

Figure 4.2: Comparison between segmentation

So, we can say here our trained YoloV8 model can almost perfectly localize the kidney easily.

Classification - ResNet50, AlexNet, VGG19

CT images are considered as critical images from the point of view of Medical Images. As we used only Axial point of view it is little bit easier for the model to extract shapes, objects, texture, edges, and pattern as features. So, we can expect good result from commonly used models. Also, in the case of our reduced ROI and segmented pictures the model can extract only important feature, so we can expect better accuracy from the model like AlexNet. Let us see the classification report to understand it better-

Here, N = Normal, S = Stone, T = Tumor, P = Precision, R = Recall, F1 = F1 Score, Acc = Accuracy

Table 4.2: Classification Report of All Models using Unsupervised Dataset

	ResNet50			**AlexNet**			**VGG19**		
	N	**S**	**T**	**N**	**S**	**T**	**N**	**S**	**T**
P	0.99	0.98	0.98	0.97	0.99	0.98	0.99	0.99	0.99
R	0.98	0.98	0.99	0.99	0.98	0.99	0.99	0.97	0.99
F1	0.96	0.97	0.97	0.97	0.99	0.98	0.98	0.99	0.99
Acc	0.9865			0.99			0.9957		

Table 4.3: Classification Report of All Models using Manually Annotated Dataset

	ResNet50			**AlexNet**			**VGG19**		
	N	**S**	**T**	**N**	**S**	**T**	**N**	**S**	**T**
P	0.98	0.98	0.96	0.0	0.16	0.0	0.81	0.99	0.94
R	0.97	0.97	0.99	0.0	1.0	0.0	0.97	0.82	0.92
F1	0.95	0.95	0.97	0.0	0.27	0.0	0.88	0.89	0.93
Acc	0.97			0.16			0.9		

In table 4.2, it is clear that for the reduced region of interest, all the model performed well and also, we got good result from the AlexNet. The reason is though AlexNet has only 8 layers, as we made the image easier for the model so only with 8 layers it performed so well. On the other hand, in table 4.3, we can see that ResNet50 and VGG19 performed well as they have enough layers and extract complex feature from images but in the case of AlexNet the performance is not good. The main reason here as in the manually annotated image though we marked the stone and tumor but the model is still confused for it's lower number of layers and simple architecture.

Now let's see the confusion matrix of all the models-
Here, TN = True Class Normal, TS = True Class STONE, TT = True Class TUMOR, PN = Predicted Class Normal, PS = Predicted Class STONE, PT = Predicted Class TUMOR.

Table 4.4: ResNet50 (Unsupervised)

TN	642	6	7
TS	3	166	0
TT	0	3	267
	PN	PS	PT

Table 4.5: AlexNet (Unsupervised)

TN	648	2	5
TS	20	166	2
TT	7	3	267
	PN	PS	PT

Table 4.6: VGG19 (Unsupervised)

TN	648	7	0
TS	7	164	3
TT	0	5	267
	PN	PS	PT

Table 4.7: ResNet50 (Annotated)

TN	635	3	17
TS	6	164	3
TT	14	3	253
	PN	PS	PT

Table 4.8: AlexNet (Annotated)

TN	0	655	0
TS	0	169	0
TT	0	270	0
	PN	PS	PT

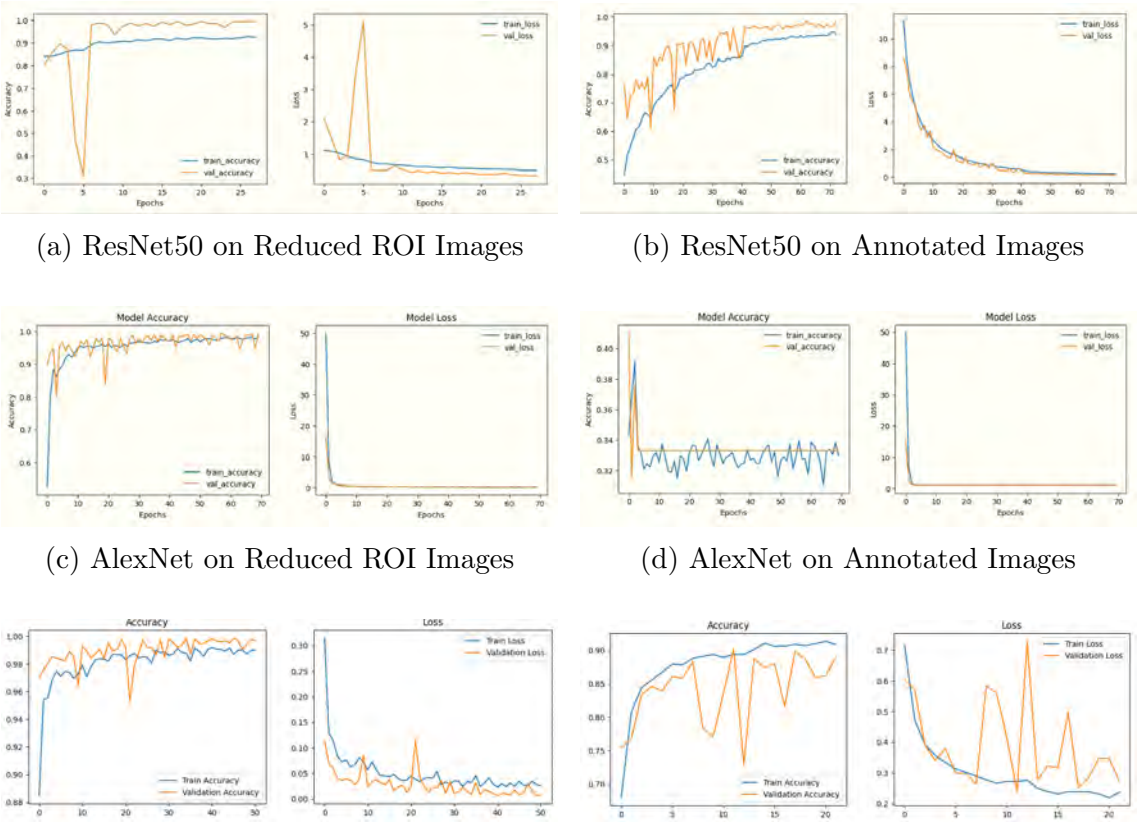
Table 4.9: VGG19 (Annotated)

TN	635	1	19
TS	148	139	16
TT	20	30	248
	PN	PS	PT

From all the confusion matrix table of Unsupervised dataset we can easily see that for the reduced ROI the models perform well and could predict the classes well. There are some missconfusion but overall result we can say good. In the case of manual annotation for complex architecture and better feacure extraction the ResNet50 and VGG19 could predict the picture well but the AlexNet could not perform well, even it could not predict the normal and tumor, it always predicted all the images as

stone. So, we can say based on the confusion matrix our unsupervised segmentation and Reduced ROI dataset performed well.

Next let us analyze the model history-



(a) ResNet50 on Reduced ROI Images

(b) ResNet50 on Annotated Images

(c) AlexNet on Reduced ROI Images

(d) AlexNet on Annotated Images

(e) VGG19 on Reduced ROI Images

(f) VGG19 on Annotated Images

Figure 4.3: Model History Comparison on Different Models

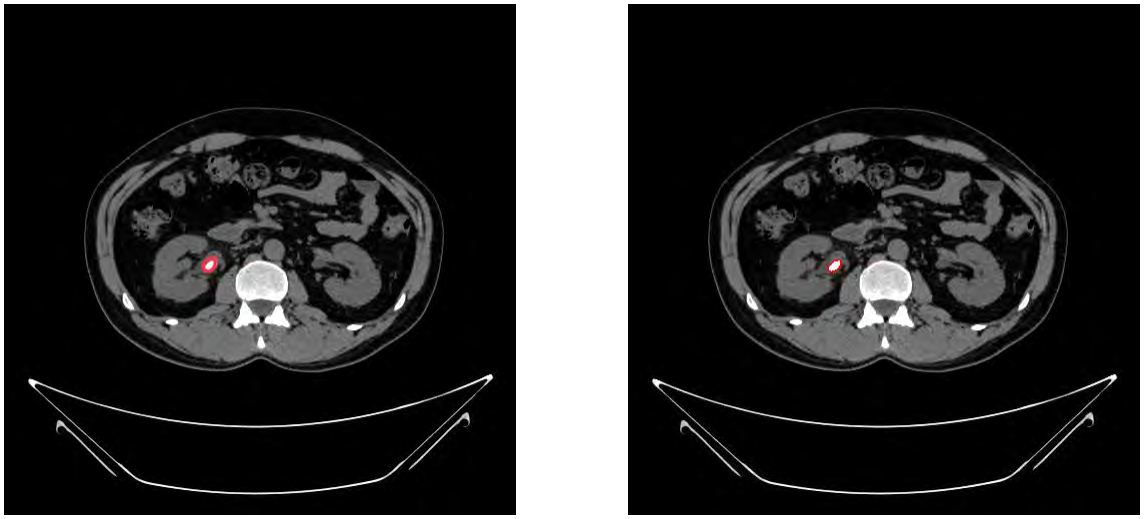
From those model history (Figure- 4.3) we can easily see that, the model history looks better in the Reduced ROI and Unsupervised Segmented dataset. Then over all fluctuation of carve both in the case of Training accuracy Vs Validation accuracy and Training loss Vs Validation loss, the reduced ROI and unsupervised Segmented dataset performed well and there are less fluctuation. The reason is for the less ROI and for the segmentation the model could easily detect the objects, patterns, edges, shapes and texture, so they performed really well. But in the case of Manual Annotation as the ROI is not reduced so model got confused in some region so it flactuated a lot.

So, based on all the result related to classification we can say if we reduce ROI of CT images we can expect good output easily and faster rather than working on the full images. This technique can help us not only in the case of Kidney disease but also for the disease of other organs. Also, as manual annotation takes a lot of time, our technique, Kidney Localization to Reduce ROI and Unsupervised Segmentation, is faster than manual annotation.

4.2.2 Outcome Evaluation - II

Segmentation Performance Metrix - STONE

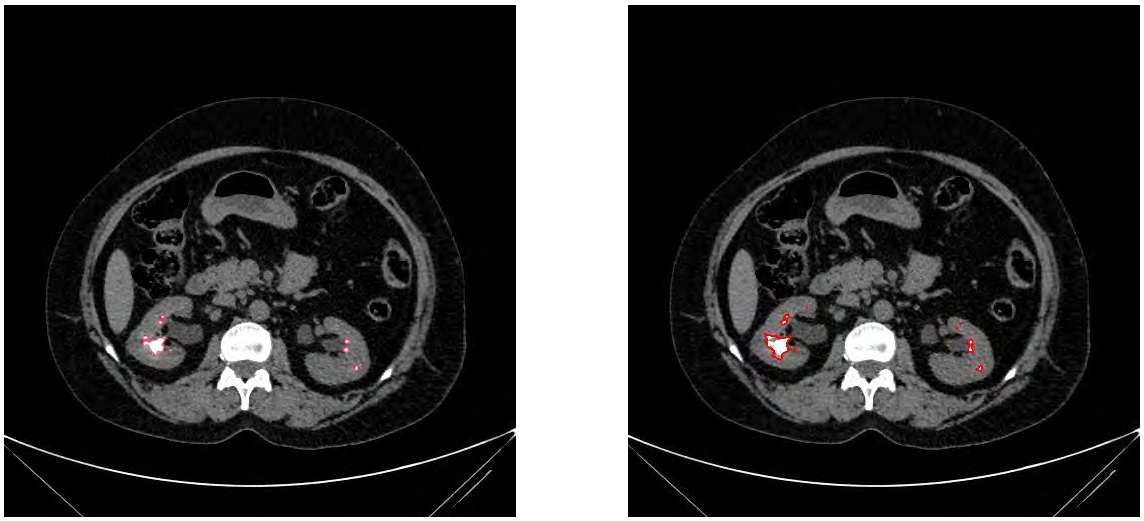
We used the picture of Kidney stones to measure the Segmentation Performance Metrix as they are easier to measure and we got best pictures from our unsupervised segmentation form stone images. For checking Segmentation Performance Metrix we did not reduced the ROI as both pic will be compared for the output



(a) Manual Anotation (b) Unsupervised Segmentation

Figure 4.4: Comparison between segmentation

In Figure 4.4 we can clearly see that unsupervised segmentation looks more perfect than the manual annotation, the reason is we human can be imperfect but the algorithm is mostly perfect. For this pictures we got the result of IoU, Precision, Recall, Dice are respectively- 0.8999, 0.9512, 0.9435, 0.9473.



(a) Manual Anotation (b) Unsupervised Segmentation

Figure 4.5: Comparison between segmentation

In Figure 4.5 we can clearly see the same result like before that unsupervised segmentation looks more perfect than the manual annotation, also if we observe closely we can see that unintentionally we missed small white spot but we can see those spots marked on the unsupervised segmentation. For this pictures we got the result of IoU, Precision, Recall, Dice are respectively- 0.8303, 0.9200, 0.8949, 0.9073.

We compared all the pictures of 848 stone pictures and we got the average result-

Table 4.10: Average Scores for Evaluation Metrics

	IoU	Precision	Recall	Dice
Average	0.8719	0.9369	0.9234	0.9301

So, based on the result we can say the segmentation for stone is overall look perfect and close to manual annotation.

Chapter 5

Conclusion

5.1 Conclusion

As we discussed, kidney disease is one of the biggest problems and also, there is a shortage of nephrolithiasis in the world nowadays. Moreover, identifying the disease is also a costly process. As a major part of the population in the world lives below the poverty line, it's really challenging for them to identify kidney disease and get proper treatment at a low cost. We hope that our research outcome will help the general public to reduce their hassle. At least they will be able to identify kidney disease from CT images at a low cost and without any specialist. As a result, they can take treatment at the right time before their disease becomes operational and save their life as well as hard-earned money. Finally, we believe that reducing and focusing on the region of interest can easily achieve good output in lightweight models which will be very time efficient and cost effective that can afford our general people and this research can be a game-changing element for the future of Medical Science.

5.2 Limitation

Our research has some limitations like we only analyzed axial views of CT scans among 3 points of view, missing insights from other views like coronal and sagittal. Therefore, including these views in future work can help which will help our model to detect kidney disease better.

5.3 Future Work

We will also include coronal and sagittal views of CT scans to improve our model's robustness and accuracy. We will experiment with modified mid-weight models for a balance between complexity and computational efficiency. To improve segmentation precision, we will introduce new metrics and focus on collecting more data to create a balanced dataset.

Bibliography

[1] N. J. Vogelzang and W. M. Stadler, "Kidney cancer," *The Lancet*, vol. 352, no. 9141, pp. 1691–1696, 1998.

[2] C. Barbas, A. Garcıa, L. Saavedra, and M. Muros, "Urinary analysis of nephrolithiasis markers," *Journal of Chromatography B*, vol. 781, no. 1-2, pp. 433–455, 2002.

[3] E. N. Taylor, M. J. Stampfer, and G. C. Curhan, "Obesity, weight gain, and the risk of kidney stones," *Jama*, vol. 293, no. 4, pp. 455–462, 2005.

[4] O. W. Moe, "Kidney stones: Pathophysiology and medical management," *The lancet*, vol. 367, no. 9507, pp. 333–344, 2006.

[5] R. Ricketts, P. Tamboli, B. Czerniak, and C. C. Guo, "Tumor-to-tumor metastasis: Report of 2 cases of metastatic carcinoma to angiomyolipoma of the kidney," *Archives of pathology & laboratory medicine*, vol. 132, no. 6, pp. 1016–1020, 2008.

[6] E. M. Worcester and F. L. Coe, "Nephrolithiasis," *Primary Care: Clinics in Office Practice*, vol. 35, no. 2, pp. 369–391, 2008.

[7] A. D. Rule, V. L. Roger, L. J. Melton, *et al.*, "Kidney stones associate with increased risk for myocardial infarction," *Journal of the American Society of Nephrology*, vol. 21, no. 10, pp. 1641–1644, 2010.

[8] S. Wang, H. Lu, F. Yang, and M.-H. Yang, "Superpixel tracking," in *2011 International Conference on Computer Vision*, IEEE, 2011, pp. 1323–1330.

[9] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "Imagenet classification with deep convolutional neural networks," in *Advances in neural information processing systems*, 2012, pp. 1097–1105.

[10] W. M. Linehan, "Genetic basis of kidney cancer: Role of genomics for the development of disease-based therapeutics," *Genome research*, vol. 22, no. 11, pp. 2089–2100, 2012.

[11] Z. M. El-Zoghby, J. C. Lieske, R. N. Foley, *et al.*, "Urolithiasis and the risk of esrd," *Clinical Journal of the American Society of Nephrology*, vol. 7, no. 9, pp. 1409–1415, 2012.

[12] S. DR, *Mcaninch jw, lue tf. smith & tanagho's general urology*, 2013.

[13] S. Jacobson, "Chronic kidney disease–a public health problem?" *Lakartidningen*, vol. 110, no. 21, pp. 1018–1020, 2013.

[14] V. Jha, G. Garcia-Garcia, K. Iseki, *et al.*, "Chronic kidney disease: Global dimension and perspectives," *The Lancet*, vol. 382, no. 9888, pp. 260–272, 2013.

[15] M. Juneja, A. Vedaldi, C. Jawahar, and A. Zisserman, "Blocks that shout: Distinctive parts for scene classification," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2013, pp. 923–930.

[16] J. L. Pirkle and B. I. Freedman, "Hypertension and chronic kidney disease: Controversies in pathogenesis and treatment," *Minerva urologica e nefrologica= The Italian journal of urology and nephrology*, vol. 65, no. 1, p. 37, 2013.

[17] N. Chhiber, M. Sharma, T. Kaur, S. Singla, *et al.*, "Mineralization in health and mechanism of kidney stone formation," *International Journal of Pharmaceutical Science Invention*, vol. 3, pp. 25–31, 2014.

[18] S. B. Ghaderian and S. S. Beladi-Mousavi, "The role of diabetes mellitus and hypertension in chronic kidney disease," *Journal of renal injury prevention*, vol. 3, no. 4, p. 109, 2014.

[19] M. R. Ghalib, S. Bhatnagar, S. Jayapoorani, and U. Pande, "Artificial neural network based detection of renal tumors using ct scan image processing," *International Journal of Engineering and Technology*, vol. 6, no. 1, pp. 28–35, 2014.

[20] E. Quaia, *Radiological imaging of the kidney*. Springer, 2014.

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

[22] J. Liu, S. Wang, M. G. Linguraru, J. Yao, and R. M. Summers, "Computer-aided detection of exophytic renal lesions on non-contrast ct images," *Medical image analysis*, vol. 19, no. 1, pp. 15–29, 2015.

[23] J. Long, E. Shelhamer, and T. Darrell, "Fully convolutional networks for semantic segmentation," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2015, pp. 3431–3440.

[24] V. K. Sigurjonsdottir, H. L. Runolfsdottir, O. S. Indridason, R. Palsson, and V. O. Edvardsson, "Impact of nephrolithiasis on kidney function," *BMC nephrology*, vol. 16, pp. 1–7, 2015.

[25] M. Courbebaisse, C. Prot-Bertoye, J. Bertocchio, *et al.*, "Nephrolithiasis of adult: From mechanisms to preventive medical treatment," *La Revue de medecine interne*, vol. 38, no. 1, pp. 44–52, 2016.

[26] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2016, pp. 770–778.

[27] T. Alelign and B. Petros, "Kidney stone disease: An update on current concepts," *Advances in urology*, vol. 2018, 2018.

[28] K. J. Foreman, N. Marquez, A. Dolgert, *et al.*, "Forecasting life expectancy, years of life lost, and all-cause and cause-specific mortality for 250 causes of death: Reference and alternative scenarios for 2016–40 for 195 countries and territories," *The Lancet*, vol. 392, no. 10159, pp. 2052–2090, 2018.

[29] R. Yamashita, M. Nishio, R. K. G. Do, and K. Togashi, "Convolutional neural networks: An overview and application in radiology," *Insights into imaging*, vol. 9, pp. 611–629, 2018.

[30] A. Parakh, H. Lee, J. H. Lee, B. H. Eisner, D. V. Sahani, and S. Do, "Urinary stone detection on ct images using deep convolutional neural networks: Evaluation of model performance and generalization," *Radiology: Artificial Intelligence*, vol. 1, no. 4, e180066, 2019.

[31] M. Akhtar, N. M. Taha, A. Nauman, I. B. Mujeeb, and A. D. M. Al-Nabet, "Diabetic kidney disease: Past and present," *Advances in anatomic pathology*, vol. 27, no. 2, pp. 87–97, 2020.

[32] M. L. Stoller, "Urinary stone disease," in *Smith amp; Tanagho's General Urology, 19e*, J. W. McAninch and T. F. Lue, Eds. New York, NY: McGraw Hill, 2020. [Online]. Available: accessmedicine.mhmedical.com/content.aspx?aid=1171179946.

[33] F. Türk, M. Lüy, and N. Barışçı, "Kidney and renal tumor segmentation using a hybrid v-net-based model," *Mathematics*, vol. 8, no. 10, p. 1772, 2020.

[34] J. Uhlig, A. Leha, L. M. Delonge, *et al.*, "Radiomic features and machine learning for the discrimination of renal tumor histological subtypes: A pragmatic study using clinical-routine computed tomography," *Cancers*, vol. 12, no. 10, p. 3010, 2020.

[35] Z. Akram, "Cancerous tumor segmentation of kidney images and prediction of tumor using medical image segmentation and deep learning techniques," *Clin. Oncol*, vol. 4, pp. 1–9, 2021.

[36] H. A. Khan, W. Jue, M. Mushtaq, and M. U. Mushtaq, "Brain tumor classification in mri image using convolutional neural network," *Mathematical Biosciences and Engineering*, 2021.

[37] S. M. Sozio, K. A. Pivert, F. J. Caskey, and A. Levin, "The state of the global nephrology workforce: A joint asn–era-edta–isn investigation," *Kidney international*, vol. 100, no. 5, pp. 995–1000, 2021.

[38] K. Yildirim, P. G. Bozdag, M. Talo, O. Yildirim, M. Karabatak, and U. R. Acharya, "Deep learning model for automated kidney stone detection using coronal ct images," *Computers in biology and medicine*, vol. 135, p. 104 569, 2021.

[39] D. Alzu'bi, M. Abdullah, I. Hmeidi, *et al.*, "Kidney tumor detection and classification based on deep learning approaches: A new dataset in ct scans," *Journal of Healthcare Engineering*, vol. 2022, 2022.

[40] A. F. Bayram, C. Gurkan, A. Budak, and H. KARATAŞ, "A detection and prediction model based on deep learning assisted by explainable artificial intelligence for kidney diseases," *Avrupa Bilim ve Teknoloji Dergisi*, no. 40, pp. 67–74, 2022.

[41] A. Caglayan, M. O. Horsanali, K. Kocadurdu, E. Ismailoglu, and S. Guneyli, "Deep learning model-assisted detection of kidney stones on computed tomography," *International braz j urol*, vol. 48, pp. 830–839, 2022.

[42] M. Gharaibeh, D. Alzu'bi, M. Abdullah, *et al.*, "Radiology imaging scans for early diagnosis of kidney tumors: A review of data analytics-based machine learning and deep learning approaches," *Big Data and Cognitive Computing*, vol. 6, no. 1, p. 29, 2022.

[43] M. N. Islam, M. Hasan, M. K. Hossain, M. G. R. Alam, M. Z. Uddin, and A. Soylu, "Vision transformer and explainable transfer learning models for auto detection of kidney cyst, stone and tumor from ct-radiography," *Scientific Reports*, vol. 12, no. 1, p. 11 440, 2022.

[44] D. Li, C. Xiao, Y. Liu, *et al.*, "Deep segmentation networks for segmenting kidneys and detecting kidney stones in unenhanced abdominal ct images," *Diagnostics*, vol. 12, no. 8, p. 1788, 2022.

[45] M. Bhandari, P. Yogarajah, M. S. Kavitha, and J. Condell, "Exploring the capabilities of a lightweight cnn model in accurately identifying renal abnormalities: Cysts, stones, and tumors, using lime and shap," *Applied Sciences*, vol. 13, no. 5, p. 3125, 2023.

[46] *Home - Kidney Cancer Association — kidneycancer.org*, https://www.kidneycancer.org/, [Accessed 21-May-2023].

[47] *Home — thecardiologyadvisor.com*, https://www.thecardiologyadvisor.com/, [Accessed 21-May-2023].