Segmentation of Bladder Lesions and Detecting Level of Invasion using Mask R-CNN
Cmpe492

Erencan Uysal
Muhammed Göktepe
Department of Computer Engineering
Bogazici University
Istanbul Turkey
erencan.uysal@boun.edu.tr
muhammed.goktepe@boun.edu.tr

June 15, 2022

Acknowledgement

We would like to express our special thanks to Dr. Mustafa Orhan Nalbant and Dr. Rüştü Türkay for their precious efforts in the creation of dataset and sharing their knowledge.

We are grateful to Pınar Yanardağ and Suzan Üsküdarlı for guiding us with their experiences and for the time they have devoted to us.

Abstract

This paper mainly focuses on the segmentation of MRI images and the creation of a usable image dataset for the tumor detection and VI-RADS scoring mechanism. We described the related works on MRI segmentation and tumor detection algorithms with convolutional neural networks. We explained our approach for segmenting MRI's using Mask RCNN and giving VI-RADS scores for lesions in bladder wall.

The main application area of our project will be on detecting bladder cancers and determining VI-RADS scores using MRI images and deep learning techniques.In this paper, we have put forward a general approach and discussed the tools and methods generally used in MRI handling.

Contents

List of Figures

List of Tables

1 Introduction

Bladder cancer is one of the most prevalent cancer types in the world. It has a high recurrence rate, and it has the heaviest treatment cost burden among other cancer types in US [7]. "Magnetic Resonance Imaging (MRI) is a non-invasive imaging technology that produces three-dimensional detailed anatomical images. It is often used for disease detection, diagnosis, and treatment monitoring. [8] MRI is the most commonly used imaging technique for detecting lesions and diagnosing cancer. There are different types of MRI sequences. The T2-weighted scan is the most convenient MRI sequence for bladder cancer detection.

Deep learning methods have been successfully applied to the classification tasks for computer vision. Especially convolutional neural networks enhanced the accuracy of object detection, image classification, and image segmentation. This approach is also frequently used in the medical field. Deep learning is being used more and more frequently in image segmentation. "Accurate segmentation of the bladder wall is of great importance for any computer-aided diagnostic system for bladder cancer (BC) detection and diagnosis." [9] "The automatic delineation of IW and OW in MRI images remains a challenging task due to important bladder shape variations, strong intensity inhomogeneity in urine caused by motion artifacts, weak boundaries and complex background intensity distribution" [3]. There are two different approaches to the segmentation of the bladder wall. The first approach is to segment the inner wall only. Since this approach cannot detect all types of bladder cancer it is not commonly used. The second approach is to segment the inner and outer walls together. In this approach tumors on the bladder wall can be detected. We have provided more detailed information on this subject in the related work section.

In the first phase of this project, we tried to apply a deep learning method to provide an image segmentation for the MRI images that we found on web. After this research, we managed to segmentate MRI images automatically. In this term, we have worked with Dr. Mustafa Orhan Nalbant and Dr. Rüştü Türkay who are expert radiologists. We created a dataset annotated by experts. We tried to develop an approach to automatically classify bladder lesions according to their VIRADS scores. The Vesical Imaging–Reporting and Data System (VI-RADS) is a structured reporting scheme for multiparametric bladder MRI in the evaluation of suspected bladder cancer. [10] We tried to use segmentation masks to reach classified results.

We tried to create a comprehensive and instructive report about MR Imaging, bladder tumors, and deep learning methods used in the segmentation of MRI images of the bladder. In the related work section we discussed:

- **What is MRI and Types of MRI:** In this section, we have explained how the MRI works and the different types of MRIs. It also includes sample MRIs of different MRI types.
- **Converting DICOM file to JPG or Nifti format:** MRI file has .dcm file extension. It includes not only image pixels but also patients' personal information. In deep learning .dcm files are not used frequently instead developers uses PNG and Nifti form of images converted from .dcm file.
- **Bladder Tumor:** We provided some necessary information about bladder cancers and the structure of the bladder.
- **Image Segmentation:** Image segmentation is the main part of our project. We explained the common methodologies of image segmentation that are used in medical images. This part also includes earnings from some research papers about bladder MRI segmentation and tumor classification.
- **Similar projects on Bladder Segmentation:** We tried to examine and work through some similar projects about medical segmentation. There are some repositories on Github that work on this subject. We gained some experience and insights on how to segment a body part and how to detect a tumor on it.
- **Ethical Approval:** Ethics committee application process explained.
- **Further Domain Research on Medical Image Segmentation:** We read more paper about medical image segmentation mainly focuses on bladder segmentation. We obtain some insight about medical image segmentation models and tried to implement them on our case.
- **3D Reconstruction from Medical Images:** We examined papers about 3D reconstruction using 2D MRI images. There are lots of techniques in 3D reconstruction. You can see further information in related work section.

In the future work section, we have explained the steps we are considering for the continuation of the project.

2 Related Work

2.0.1 Magnetic Resonance Imaging

The aim of this research is to detect and identify tumors in bladder. While detecting and classifying tumors we will be use deep learning. At the first phase of this research, we need to make segmentation on bladder to identify the bladder area. To visualize bladder area we are using MRIs. "Magnetic Resonance Imaging (MRI) is a non-invasive imaging technology that produces three dimensional detailed anatomical images. It is often used for disease detection, diagnosis, and treatment monitoring. It is based on sophisticated technology that excites and detects the change in the direction of the rotational axis of protons found in the water that makes up living tissues." [8] We can see an instructive visual in Figure 2.1. MRI contains crossings from all three axes. In cell biology, a vesicle is a structure within or outside a cell, consisting of liquid or cytoplasm enclosed by a lipid bilayer.

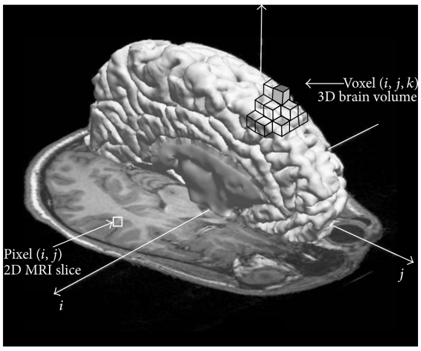

Figure 2.1: "Illustration of image elements in the MRI of the brain. An image pixel (i,j) is represented with the square in the 2D MRI slice and an image voxel (x,y,z) is represented as the cube in 3D space." [1]

MRIs employ powerful magnets which produce a strong magnetic field that forces protons in the body to align with that field. [8] "It differs from other techniques such as computed tomography (CT) by producing excellent soft tissue contrast without harmful ionizing radiation." [11] "The contrast of the image thus depends on the signal intensity (SI) of different tissues. Certain tissues that are rich in free protons, such as water and fat, are very responsive to the RF pulses. Other tissues with fewer free protons, such as cortical bone and air, are less responsive and generate much less signal." [12] "The most common MRI sequences are T1-weighted and T2-weighted scans. T1-weighted images are produced by using short TE and TR times." [13] "Repetition Time (TR) is the amount of time between successive pulse sequences applied to the same slice." [13] "Time to Echo (TE) is the time between the delivery of the RF pulse and the receipt of the echo signal." [13] "A third commonly used sequence is the Fluid Attenuated Inversion Recovery (Flair). The Flair sequence is similar to a T2-weighted image except that the TE and TR times are very long." [13]

- "T1-weighted (T1W) imaging, on which fluid appears dark and fat appears bright." [12]
- "T2-weighted (T2W) imaging, on which both fluid and fat appear bright." [12]

Figure 2.2 shows different types of MRIs.

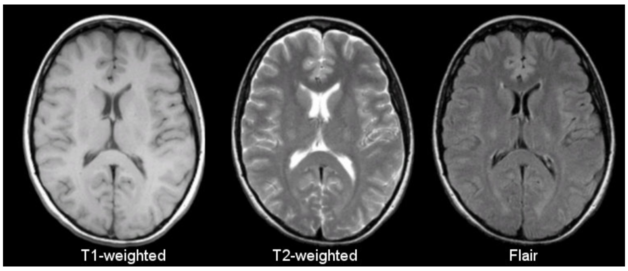

Figure 2.2: Different Magnetic Resonance Imaging Samples, (ref: 2016, Hellerhoff, T2w, FLAIR, T1 mit Kontrastmittel, licensed under CC BY-ND 2.0)

MR images are stored in DICOM files. "A DICOM file consists of a header and image data sets packed into a single file. The information within the header is organized as a constant and standardized series of tags. By extracting data from these tags one can access important information regarding the patient demographics, study parameters, etc." [14] In deep learning developers mostly does not use DICOM files. They convert DICOMs to Nifti file or JPEG format. A DICOM file is a 2D image and it has only a slice of a MRI. Nifti file is sequence of DICOM images. So, nifti is a full representation of a part of the body and it contains all slices that were taken by MR. To convert DICOM files to nifti file we need to install a library named "dicom2nifti". With the following code dicom images creates a image sequence.

```
import dicom2nifti
source = "the␣path␣to␣the␣dicom␣images␣directory"
dest = 'the␣path␣to␣write␣.nii␣file'
dicom2nifti.convert_directory(source, dest)
```

Also you can convert dicom files to JPEG format. Following is a simple script to convert dicom file.

```
import numpy as np
import pydicom
from PIL import Image

# Put the right path of the image
# that you want to convert
ds = pydicom.dcmread('the␣path␣to␣the␣image')
# Convert the values into float
new_image = ds.pixel_array.astype(float)
scaled_image = (np.maximum(new_image, 0)
               / new_image.max()) * 255.0
scaled_image = np.uint8(scaled_image)
final_image = Image.fromarray(scaled_image)
final_image.show() # Display the Image
final_image.save('image.jpg') # Save the image as JPG
final_image.save('image.png') # Save the image as PNG
```

2.0.2 Bladder Tumors

We made some research about bladder tumors and examined the patterns of these tumors in MRI images. Bladder is not a solid or full body organ like brain. Instead, it consists of a hollow muscle layer. Therefore, the place where the tumor can be observed is the muscles that forming bladder wall.

"The bladder wall contains three main layers: the innermost mucosa, the muscularis propria (network of smooth muscle fibers or detrusor muscle), and the outermost serosa with perivesical fat (loose layer of connective tissue)." [7] "Bladder wall layer stratification cannot be entirely defined by MRI. The detrusor muscle appears as a low-signal-intensity line on T2-weighted images and as an intermediate-signal-intensity line on DWI images and apparent diffusion coefficient (ADC) maps, whereas the inner layer, composed of urothelium and lamina propria, is not seen using any of these techniques." [7]

2.0.3 Image Segmentation

Deep learning methods have been successfully applied to the classification tasks for computer vision. Especially convolutional neural networks enhanced the accuracy of object detection, image classification, and image segmentation.

"The goal of image segmentation is to divide an image into a set of semantically meaningful, homogeneous, and nonoverlapping regions of similar attributes such as intensity, depth, color, or texture. The segmentation result is either an image of labels identifying each homogeneous region or a set of contours which describe the region boundaries." [1] "Accurate segmentation of the bladder wall is of great importance for any computer-aided diagnostic system for bladder cancer (BC) detection and diagnosis." [9]

Computers can not see images like humans. If the computer is required to draw conclusions from a picture, it is necessary to submit the pictures to the computer after preprocessing. To illustrate, we can examine figure 2.3. We can clearly see the canrsinoma in the left side of the bladder wall (a). In (b), we see the recovered bladder MRI image of the same patient. Even with an untrained eye, we can understand where the bladder is. But around the bladder, there is a larger area of muscle, fat, and bone. We can see there are lots of difference in the muscle layer. Even in the same person, these layers can change over time with weight gain or loss or the emergence of other diseases. When this image is directly processed by the computer, these layers outside the bladder will not be detected independently and this will affect the result. To avoid the effects of this situation, we are using image segmentation.

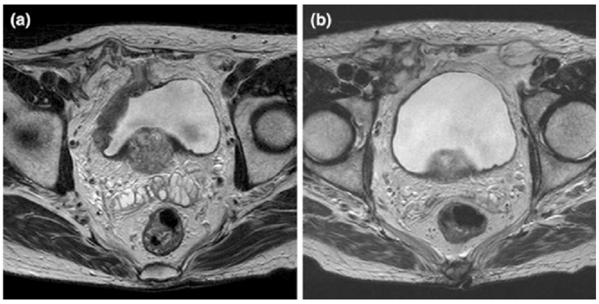

Figure 2.3: MRI (T2-weighted image) revealed a massive tumor at the right wall of urinary bladder with invasion to the abdominal wall (a) at presentation in case 2. After neoadjuvant chemotherapy, MRI (T2-weighted image) demonstrated no residual tumor in the bladder (b) [2]

We found more comprehensive information in an article about the difficulty of processing MRI images. In figure 2.4, you can see these challenges."The automatic delineation of IW and OW in MRI images remains a challenging task due to important bladder shape variations, strong intensity inhomogeneity in urine caused by motion artifacts, weak boundaries and complex background intensity distribution" [3]

Image segmentation can have different meanings in context. First of all, segmentation means "division into separate parts or sections". We can use segmentation for just the purpose of cutting meaningful parts from an image or we can use it with deep learning methods and draw results from this process too. There are some approaches for image segmentation in biomedical imaging field.

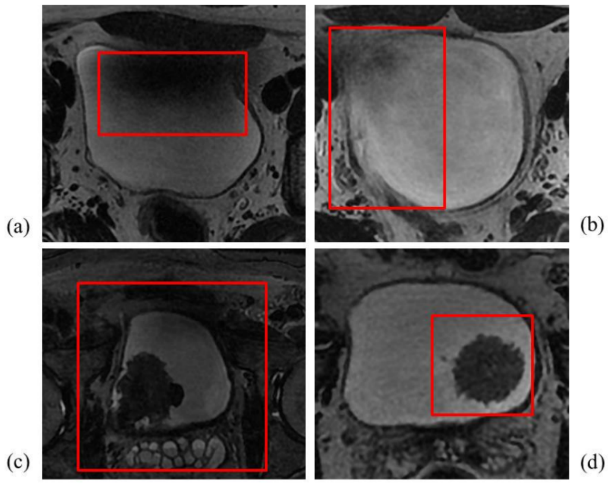

Figure 2.4: Challenges in computer-assisted segmentation of bladder MR images (highlighted in red boxes). Image regions in red boxes represent: (a) intensity inhomogeneity in the lumen, (b) weak wall boundaries, (c) complex background intensity distribution, and (d) disconnected tumor region in the lumen, respectively. [3]

U-Net is a convolutional neural network that was developed for biomedical image segmentation at the Computer Science Department of the University of Freiburg. [15] This model uses “fully convolutional network” architecture. It uses overlap-tile strategy for seamless segmentation of arbitrary large images. [15] Microscopic images and ultrasonography are more convenient subjects for this architecture but we can use this architecture in MRI images too.

Mask RCNN (Region Based CNN) is another great framework for segmentation process. Region Based Convolutional Neural Networks (R-CNN) are machine learning models which are using in computer vision, specifically object detection. Maks R-CNN combines instance segmentation and object detection in a robust and flexible manner. "In principle Mask R-CNN is an intuitive extension of Faster R-CNN, yet constructing the mask branch properly is critical for good results." [4]

"Mask R-CNN is conceptually simple: Faster R-CNN has two outputs for each candidate object, a class label and a bounding-box offset; to this they added a third branch that outputs the object mask." [4] In figure 2.5, mechanism is clearly described. There is a good github repository about Mask R-CNN. We inspected this repository. [16] There are available codes in this page. We also found a project ([6]) which uses Mask R-CNN in brain MRI segmentation. This framework can be implemented to our project.

"Multi-region segmentation of bladder cancer structures in MRI with pro-

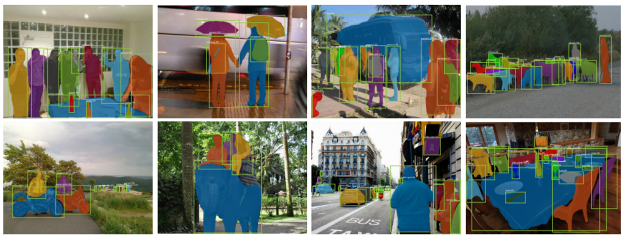

Figure 2.5: Mask R-CNN results on the COCO test set. These results are based on ResNet-101, achieving a mask AP of 35.7 and running at 5 fps. Masks are shown in color, and bounding box, category, and confidences are also shown [4]

gressive dilated convolutional networks" article also stated a detailed comparison between frequently used CNN models. You can see this comparison in figure 2.6. [3] This comparison can help us while selecting the framework to use in our project

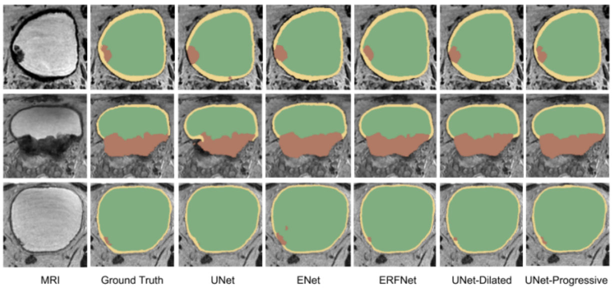

Figure 2.6: Visual results achieved by all the models. While inner and outer walls are respectively represented by green and yellow regions, tumour regions are highlighted in brown. [3]

2.0.4 Similar Projects

We read many articles about bladder segmentation to learn basic approaches to the problem and which methods are used for segmentation. In this section, we reviewed the methods discussed in the following article. [5] Segmentation of the bladder wall (BW) from magnetic resonance imaging (MRI) such as T1- and T2- weighted (T1W-MRI and T2W-MRI) is a key and most crucial first step in bladder image processing and analysis. [5] Segmentation is an important part of the Bladder Cancer (BC) detection. However, it is not straightforward because of the pathological bladder wall. There are two different approaches in segmentation of bladder wall. First approach is segment inner wall only. Since this approach cannot detect all types of bladder cancer it is not commonly used. Second approach is segment inner and outer wall together. In this approach tumors on bladder wall can be detected. Their framework is based on a fully connected convolution neural network (CNN) and also utilizes adaptive shape information and contextual information to enhance the segmentation. [5] In this article they used Random Forest algorithm. A random forest algorithm consists of many decision trees. [17] It predicts by taking the average or mean of the output from various trees. Increasing the number of trees increases the precision of the outcome. [17]

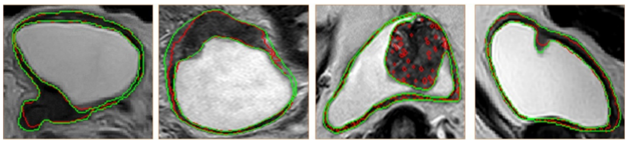

Figure 2.7: Results for four subjects illustrating the segmentation of the inner and outer bladder walls with pathology (tumor) obtained using Random forest (RF) method w.r.t. the segmentation (red) and ground truth (green). [5]

We found a github repository which makes segmentation on brain MRIs using U-Net and Mask-RCNN. [6] It uses lots of libraries like numpy, Scipy, Keras, Tensorflow, OpenCV etc. Firstly they prepared their dataset to use in this project. Their dataset was taken from University of Pennsylvania. [18] Data was in Nifti format. NIFTI format is very high resolution. Therefore, they've converted these images to .png format. They used T1-weighted magnetic resonance imaging type. They used pretrained MASK-RCNN model while detecting tumors. [19] They loaded the model and trained it with BraTS 2019 dataset. Figure 2.8 shows the result of this model.

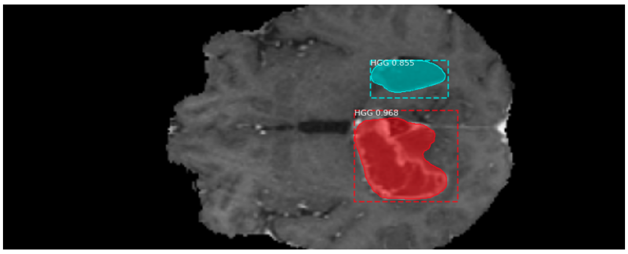

Figure 2.8: Prediction of the solution with Mask R-CNN. [6]

2.0.5 Ethical Approval

In this project we are working with two doctors, Dr. Rüştü Türkay and Dr. Mustafa Nalbant from Istanbul Haseki Training and Research Hospital. We will use MR images of patients in our project. Therefore, before implementing our project we need to get Ethical Approval from Ethical Committee in our university. We filled an application form that we described our method and procedure on how we use these MRI's while segmenting bladders and determining VI-RADS scores of the patients. Also we summarized our project to signify the importance of obtaining a medical data for this project. We stated the participants profile on the application form and possible risk for the patients that will participate in this project. We explained the benefits of the research for the patients and also for the society and the science. We stated that we will get anonymous data so that participants personal information won't be exposed. Also, we prepared a participant information and consent form to document their willingness to participate in this project. Lastly our consultant doctors received an ethical approval from Istanbul Haseki Training and Research Hospital's TUEK institution.

2.0.6 Further Domain Research on Medical Image Segmentation

We read a survey paper [20] about bladder segmentation to compare and explore new methods and obtain some insight about different approaches.

- Multi-region segmentation of bladder cancer structures in MRI with progressive dilated convolutional networks (2018) [3]:
  - They used a deep CNN that builds on UNet.
  - To increase the receptive field spanned by the network, they propose to use a sequence of progressive dilation convolutional layers.
  - In general, all models yield similar performance for the inner wall (IW) and outer wall (OW) regions. However, in the case of tumor regions, the UNet-Progressive model obtained a higher accuracy.
- A Deep Learning-Based Approach for Accurate Segmentation of Bladder Wall Using MR Images (2019) [5]:
  - In order to reduce the BW variability across subjects, the grayscale MR images are co-aligned using an Affine transformation followed by B-splines based transformation to account for global and local motion, respectively. Finally, the obtained transformation parameters for each subject are applied to its labeled data to be used for the construction of the shape prior probability.
  - Due to the limited number of available data sets, they employed a leave-one-subject-out (LOSO) approach for the evaluation of the pipeline.

- To evaluate the accuracy of the pipeline, they used the Dice Similarity coefficient (DSC) to characterize the spatial overlap between segmented and ground truth segmentation as well as the Hausdorff distance (HD) to characterize the closeness of the segmented contours to their ground truth counterparts.

- Affine transformation: Affine transformation is a linear mapping method that preserves points, straight lines, and planes. Sets of parallel lines remain parallel after an affine transformation. The affine transformation technique is typically used to correct for geometric distortions or deformations that occur with non-ideal camera angles. [21]

- LOOCV (leave one out cross-validation): The Leave-One-Out Cross-Validation, or LOOCV, procedure is used to estimate the performance of machine learning algorithms when they are used to make predictions on data not used to train the model. It is a computationally expensive procedure to perform, although it results in a reliable and unbiased estimate of model performance. [22]

- Dice Loss: Dice Loss optimizes networks based on the dice overlap coefficient between the predicted segmentation result and the ground truth annotation, thus can effectively alleviate the imbalance between the foreground and background. [23] A better fit than cross-entropy for imbalanced data. Classes are not balanced in our problem i.e tumor constitutes a modest part of MRI.

- Deepmedic [24]: This project aims to offer easy access to Deep Learning for segmentation of structures of interest in biomedical 3D scans. It is a system that allows the easy creation of a 3D Convolutional Neural Network. The system processes NIFTI images. Dual pathway processes images at multiple scales simultaneously to incorporate both local and larger contextual information. We tried to obtain NIFTI images from T2W scans but since we will use DWI scans we didn't complete this procedure.

2.0.7 3D Image Segmentation and Reconstruction

All the creations lies in the 3D space. However noninvasive imaging techniques are not offering three dimensional visualization. In medical imaging, frequently used applications are "world", "anatomical", and the "medical image coordinate systems". The world coordinate system is a Cartesian coordinate system in which a medical image modality (e.g. an MRI scanner or CT) is positioned. Anatomical space consist of three planes.

- The axial (or transverse) plane is actually when you place your point of view above the patient and look down.

- The sagittal plane basically is a side view.
- The coronal plane is parallel to the face or back of the patient.

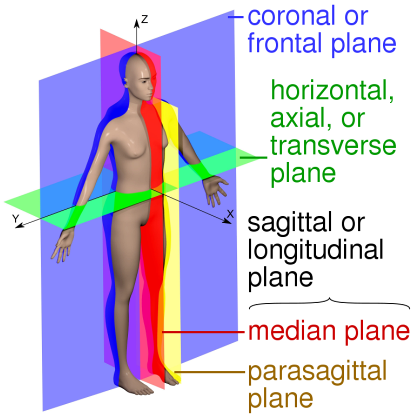

Figure 2.9: Three planes of anatomical space. (ref: David Richfield and Mikael Häggström, M.D. and cmglee, licensed under CC BY-ND 2.0)

As these applications suggests, the three dimensional body space converted to two dimensional 3 axes. Actually this methodology is crucial for the diagnosis phase in the radiology. Contemporary visualization methods are limited to computer screen and the people's perception. But when we use methods such as deep learning and computer vision, we can go beyond the limits of human perception and imaging devices. A computer can have better perception of the objects in the 3D space. Therefore, it may be possible to achieve much better results for segmentation and object detection with three-dimensional inputs that we will create using the two-dimensional images provided by the imaging devices we have. We can consecutively use the images in the same plane to create a stack of lines. These lines represents the borders of the organs and other subjects that we want to examine in the human body. In response to this, we can also create 3D visualization of a body part with using images from different planes. These inputs can be thought of as forming voxels in three-dimensional space.

Voxel is a 3D equivalents of a pixel. The word voxel is actually analogous of the word "pixel" with "vo" representing volume. Voxels are frequently used in the visualization and analysis of medical and scientific data. However there are some difficulties in forming voxels from 3 planar mri data. MRI images can show inhomogeneity inside an image and variety in contrast or elongation between images. However if we can create voxels from 2d images, we can easily reconstruct the 3D form of the body with using mesh creation algorithms from point space.

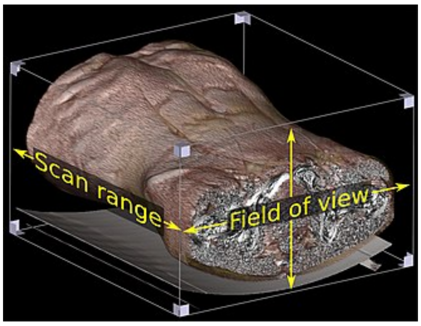

Figure 2.10: Human body imaging in axial plane. (ref: 2019, Mikael Häggström, CT cancer Imaging, licensed under CC BY-ND 2.0

"3-D modeling in medical imaging has been pursued using surface rendering, volume rendering and regularization based methods." [25] There are some methods for 3D modelling from 2D data:

- Consistency based volumetric rendering
- Maximum intensity projection
- Curved planar reformation
- Ray tracing
- Surface shaded display
- Direct volume rendering
- Ray casting
- Splatting
- Texture mapping
- Regularization based surface reconstruction
- 3D reconstruction using convolutional network
- Hierarchial deformable model based reconstruction
- Point spread function based 3D reconstruction

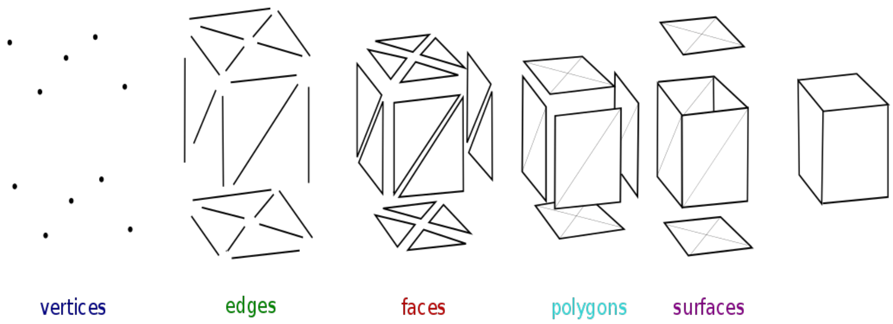

Figure 2.11: Mesh creation from points. (ref: 2009, Rchoetzlein, Mesh Overview, licensed under CC BY-ND 2.0)

Once we create a point space from 2D images, we can create meshes from this point space. Mesh generation is the practice of creating a mesh, a subdivision of a continuous geometric space into discrete geometric and topological cells. We can use points from segmentation step to generate a mesh in 3D format.

3 Approach

In this project we tried to create a system which assists radiologists in determination of VI-RADS scores of the lesions in bladder walls. To accomplish this task we needed to segment bladder by processing MR images. There are multiple types of MR imaging techniques used in bladder imaging. Common types are T1WI, T2WI, DCEI, DWI and ADC imaging techniques. In the diagnosis and VI-RADS scoring of a bladder tumor,usually T2W and DWI scans are used.

T2W scans measure spin–spin relaxation by using long TR and TE times. T2W scans generally has higher signal for more water content and lower signal for paramagnetic substances. A paramagnetic material like Gadolinium should be injected to the patient before imaging process. These substances retained by layers of fat. In this way, lesions and fat layers can be observed with hypointensity. It means tumors are visible in dark colors in T2W scans. [26] These scans generally used to make a first impression about lesions. Before making a diagnosis, a radiologist looks at these scans and has an idea of whether there is a tumor.

DWI scan is a form of MR imaging based upon measuring the random Brownian motion of water molecules within a voxel of tissue. The fundamental idea behind diffusion-weighted imaging is the attenuation of T2* signal based on how easily water molecules are able to diffuse in that region. Highly cellular tissues exhibit lower diffusion coefficients. [27] Cancer staging is made by evaluating the relationship of the bright parts seen in this sequence with the bladder wall.

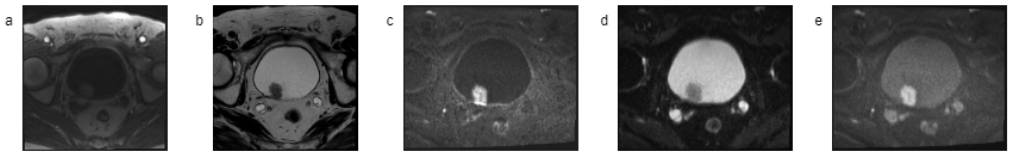

Figure 3.1: Different MRI types. (a) T1W. (b) T2W. (c) DWI. (d) ADCI. (e) DCEI.

The Vesical Imaging–Reporting and Data System (VI-RADS) is a structured reporting scheme for multiparametric bladder MRI in the evaluation of suspected bladder cancer. It proposed by the European Association of Urology in 2018. [10] Each lesion is assigned a score from 1 to 5 indicating the likelihood of clinically significant cancer:

1. Muscle invasion is highly unlikely.

2. Muscle invasion is unlikely to be present.

3. Presence of muscle invasion is equivocal. (Can not be determined confidently)
4. Muscle invasion is likely.
5. Invasion of muscle and beyond the bladder (metastasis) is very likely.

3.0.1 Diagnosis Steps in Bladder Cancer

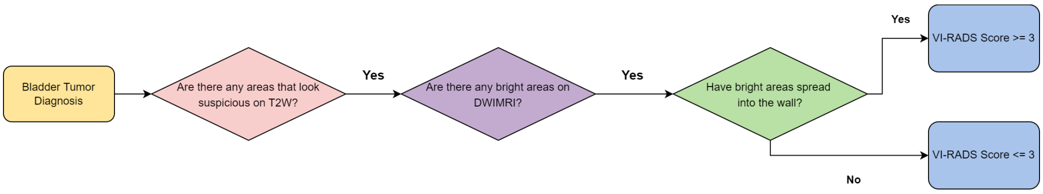

Figure 3.2: Steps for determining VI-RADS level of invasion of a lesion.

Our methodology consists of four steps. These are the following:

1. In the first step of our implementation we are looking for T2W images for suspicious black areas on the bladder wall that may be a lesion.
2. If we detect suspicious areas on the bladder wall we are looking DWI scans. We segmented bladder and tumors to see whether the tumor expanded in the bladder wall.
3. If DWI scans contains continuous bright points along the bladder wall we can say that there is a lesion in that bladder.
4. Lastly, we are looking to the segmented lesion border relation with bladder wall.
    - If segmented lesion mask and the bladder wall borders intersected than we can say that tumor is spread to the wall. It will get the score of VI-RADS >= 3.
    - If segmented lesion mask and the bladder wall borders are not intersected than we can say that tumor isn't spread to the wall. It will get the score of VI-RADS <= 3.

3.0.2 Methodology

Our methodology consists of three steps. These are the following:

1. In the first step, pretrained Mask R-CNN model fed with DWI MRI sequence.

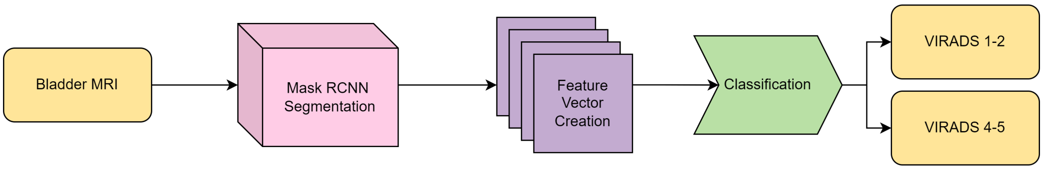

Figure 3.3: Steps for determining VI-RADS level of invasion of a lesion.

2. After getting segmented images, necessary fields are calculated and summed up to derive volumes.

3. The resulting features are used with different classifiers. We implemented a rule-based unsupervised classifier which imitates the diagnosing steps of radiologists.

In the first stage, binary classification was used to increase the success due to the small dataset. After creating a larger dataset, multiclass classification can be used.

3.0.3 Mask R-CNN

Mask R-CNN, or Mask RCNN, is a Convolutional Neural Network (CNN) and state-of-the-art in terms of image segmentation and instance segmentation. Mask R-CNN was developed on top of Faster R-CNN, a Region-Based Convolutional Neural Network. [28] The computer vision task Image Segmentation is the process of partitioning a digital image into multiple segments (sets of pixels, also known as image objects). This segmentation is used to locate objects and boundaries (lines, curves, etc.). [28] There are 2 main types of image segmentation that fall under Mask R-CNN: Semantic Segmentation, Instance Segmentation. Semantic segmentation classifies each pixel into a fixed set of categories without differentiating object instances. Instance Segmentation, or Instance Recognition, deals with the correct detection of all objects in an image while also precisely segmenting each instance. [28] In this project we are using instance segmentation. Mask R-CNN was built using Faster R-CNN. While Faster R-CNN has 2 outputs for each candidate object, a class label and a bounding-box offset, Mask R-CNN is the addition of a third branch that outputs the object mask. The additional mask output is distinct from the class and box outputs, requiring the extraction of a much finer spatial layout of an object. [28]

3.0.4 Segmentation of Bladder Wall and Lumen in 1.5 Tesla T2W MR Images

We used T2W scans for segmenting inner and outer walls of the bladder. The T2-weighted scan is the most convenient MRI sequence for bladder cancer detection. We manually labeled each image in makesense.ai labeling platform. Label type is polygon. Each image contains two labels which are "lumen" and "bladder". Order of labeling lumen and bladder is important because it may reduce performance of our Mask RCNN model. Our model uses vgg annotations so, after we finished labeling we exported vgg.json format of our labels. We used pre-trained Mask R-CNN model. [29] We split our dataset 80% training and 20% testing. We trained our model with 20 epochs. Resulting loss values from our model can bee seen in

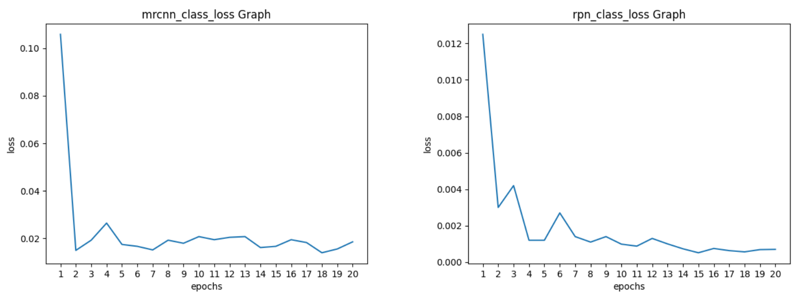

Figure 3.4: Train set mrcnn class loss graph per epoch for T2WI scans.

Figure 3.4. Classification losses reflect how confident the model is when predicting the class labels, or in other words, how close the model is to predicting the correct class. Rpn_class_loss is RPN anchor classifier loss and mrcnn_class_loss is the loss for the classifier head of Mask R-CNN.

3.0.5 Segmentation of Bladder Wall and Lesions in 1.5 Tesla DWI MR Images

We used DWI scans for segmenting bladder walls and lesions in the bladder. Brighter parts on bladder walls represent lesions. DWI scans has lower number of images, so there are fewer images than the T2W dataset. We manually labeled each image in makesense.ai labeling platform. Label type is polygon. Each image contains two types of labels which are "lesion" and "bladder". There are images without lesions too. Mask R-CNN uses vgg annotations. So, after we finished labeling we exported vgg.json format of our labels. We used pre-trained Mask R-CNN model. [29]

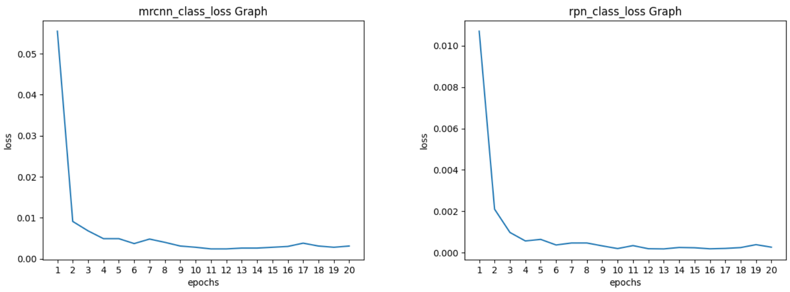

Figure 3.5: Train set mrcnn class loss graph per epoch for DWI scans.

Resulting loss values from our model can bee seen in Figure 3.5. Since our data size is very small, our class losses decreased a lot after the first 2 epochs. In subsequent epochs, however, there was not much change. In the T2W image sequences, the loss values shows more movement in the plots. But in DWI images, loss values are more smooth. This is the result of the nature of these images. Diffusion weighted images are tend to be more smooth and have less contrast. Classification losses reflect how confident the model is when predicting the class labels, or in other words, how close the model is to predicting the correct class. Rpn_class_loss is RPN anchor classifier loss and mrcnn_class_loss is the loss for the classifier head of Mask R-CNN.

3.0.6 Segmentation of Inner Wall, Outer Wall and Lesions in 3 Tesla DWI MR Images

We acquired 3 Tesla MRI images from our advisor MD's Rüştü Türkay and Mustafa Orhan Nalbant. They labelled DWI sequences with Inner wall, Outer wall and Tumor classes. We acquired diffusion weighted MR images of 50 adults who are diagnosed with bladder tumor using Siemens Magnetom Verio 3T. Resolution of the images are 512x512. This dataset contains 1226 images nearly 25 images for each patient. Numbers can deviate because of the nature of MR imaging. Different MRI sequences were used during segmentation. While labeling the bladder wall, wall boundaries were determined from the T2W MRI sequence and marked on the DWI sequence. Wall invasion was evaluated from contrast-enhanced sequences, and then measurements were made on the diffusion sequence, and labeling was done on the DWI sequence. We trained a mask RCNN model with using this dataset. We performed five fold cross validation by dividing the dataset into 5 separate parts, each containing 40 train and 10 test patients. We tried to obtain cross-sections from inner wall, outer wall and the tumor spaces.

We trained our model for 40 epochs with approximately 1220 images.

# of Tumors	# of Patients
1	36
2	7
3	1
4	2
5	1
7	1
9	1
11	1

Table 3.1: You can find information on how many patients have how many lesions.

VIRADS Score	# of Tumors
1	34
2	28
3	22
4	4
5	7

Table 3.2: Table shows how many lesions are at VIRADS 1 2 3 4 5 level.

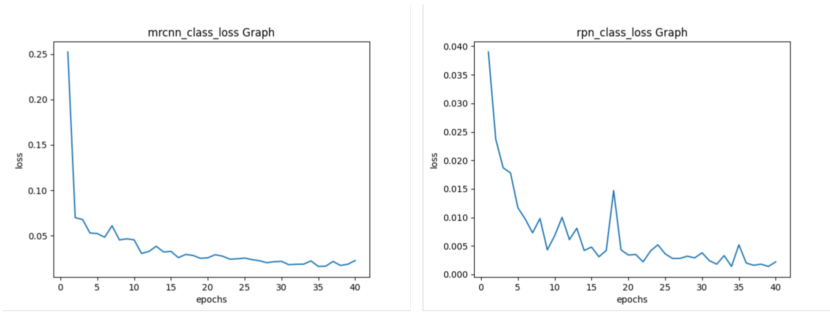

Figure 3.6: Train set mrcnn class loss graph per epoch for DWI scans.

As you can see from the Figure 3.6, our loss values are alternating. We did not see that much zigzags in 1.5 tesla images which has the resolution of 256x256. This situation is a result of high contrast and bigger size of the 3 Tesla images.

3.0.7 Classification

Classification was the most important part of this project. By examining multi-parametric MRI data, radiologists classify the patient's tumor with a scale called the VIRADS score. In the segmentation part, extraction of the necessary area info from MRI for the diagnosis and scaling of bladder lesions was managed. There were wide range of alternative methods for image based and 3D classification. We made an extensive research about 3D classifiers. The results of this research is available in the "Related Work" section.

There are several ways to convert a two-dimensional data into a three-dimensional space data. However all of these methods carry their own difficulties. In medical imaging field, the most restrictive feature is the nature of the human body. MRI images can show inhomogeneity inside an image and variety in contrast or elongation between images. Bladder is not a solid organ. It contains urine inside its wall. So bladder can expand and contract by moving. Breathing or little movements of the MRI patient can cause the results to deteriorate and cause huge differences between the two consecutive images. Even if it is successfully segmented, such defects may cause the result to be corrupted at the classification stage. The 3D reconstruction will not be an efficient solution because of this corruptions. So we decided to use a native classifier instead of 3D reconstruciton.

The most commonly used classification method in sequential image data is to train a pixel-based classifier while preserving ordering property of the data. However this methodology necessitate a relatively large dataset. The dataset that this project based on was only contain 50 patient's data, so we developed a new method tailored for bladder images.

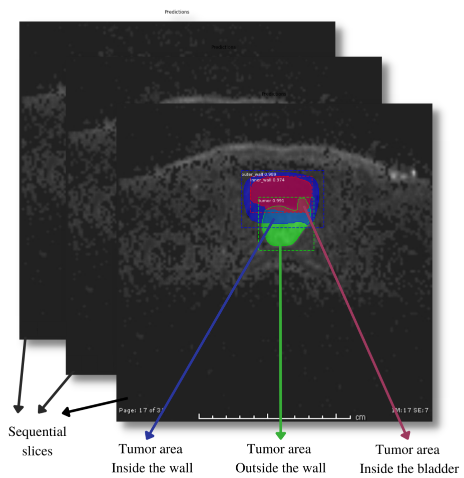

Figure 3.7: Important areas in classification

In VI-RADS scoring, the features that radiologists look for are tumor area growth on the wall and outside of the bladder. Whether there is a muscle invasion or not is determined by examining the spread of the tumor on the bladder wall. If the tumor expanded inside the bladder wall, this patient will probably have VI-RADS 4 from scoring. If it has expanded outside the bladder wall, this will cause VI-RADS to be 5. If it did not expanded inside the wall, the score will probably be less than 3. From these specifications of the scoring scheme, we tried to extract meaningful features from the MRI segmentation areas. In 3.7 a VIRADS-5 lesion is visible. You can see the selected areas for features.

tot : total tumor area, i_b tumor area inside the bladder, i_w : tumor area inside the bladder wall.

$$\frac{1}{\sum tot} \sum [i_w, i_b, o]$$

4 Results

4.0.1 Segmentation Results

4.0.1.1 Segmentation Results of Bladder Wall and Lumen in 1.5 Tesla T2W MR Images

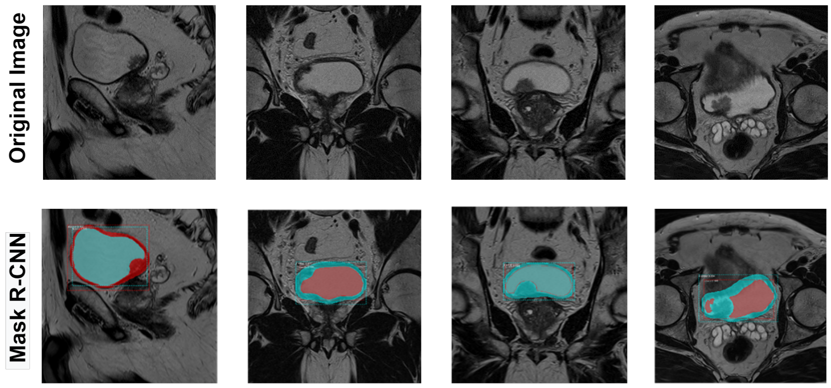

Figure 4.1: Segmented T2W scans with tumor using Mask R-CNN.

In the Figure 4.1, the T2W images contains bladders without lesions. Mask R-CNN results are available in the second row. Overall reliability of our model is good enough to help humans in decision. But in some cases our model fails and it can't segment bladder wall correctly. When lesions are inside the lumen part, it may not detect all lumen and bladder wall. Reasons behind these problems are, we have complicated and insufficient dataset. There are three different MRI planes in our dataset: sagittal, transverse and coronal plane. These planes differ from each other and using these planes together in our model reduces our success rate. For ethical reasons, it is hard to obtain such data from hospitals and patients.

4.0.1.2 Segmentation Results of Bladder Wall and Lesions in 1.5 Tesla DWI MR Images

The Figure 4.2 contains the MRI images with bladder walls invaded by lesions. This cases are more serious and more likely to metastasize than others. Mask R-CNN results are available in the first row. Tumor segmentation in these images are intersected with bladder walls. Actually intersection is not enough to determination of VI-RADS scores. Tumor segmentations should also be protruded beyond the bladder wall boundary. These kind of MRI images will probably get the score of VI-RADS >= 3.

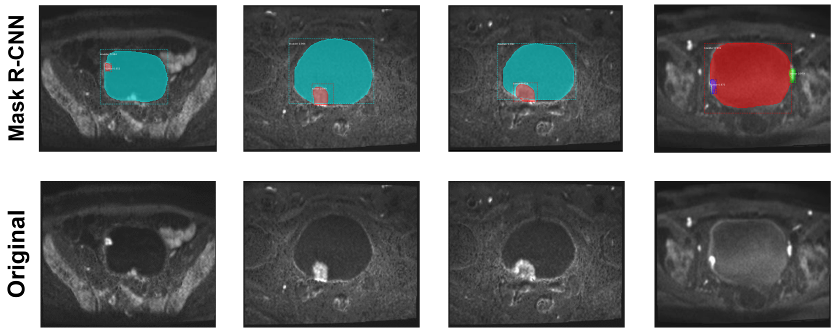

Figure 4.2: Segmented DWI scans with tumor using Mask R-CNN.

4.0.1.3 Segmentation Results of Inner Wall, Outer Wall and Lesions in 3 Tesla DWI MR Images

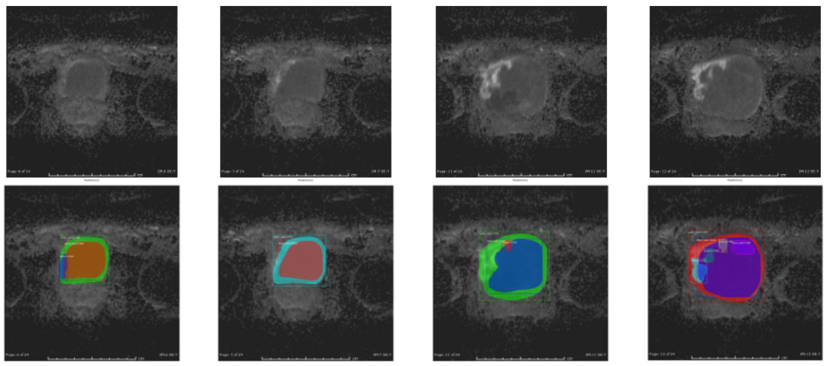

Figure 4.3: Segmented DWI scans with tumor using Mask R-CNN.

The Figure 4.3 contains the 3 Tesla MRI images and the detected masks by Mask RCNN. The results show that concave and angular sections are more difficult to detect by this algorithm. This difficulty may stem from the ambiguous bladder wall borders. In DWI MRI sequence, bladder walls can not be observed clearly but on the other hand we can easily detect tumor areas. In order to segment bladder walls more accurately, we need to work on T2W MRI images. We tried to apply segmentation and classification for both T2W and DWI MRI sequences but we could not achieve to merge the results. Main reason for this is, while taking the MRI of a patient, there is small time difference between sequences. Therefore, there are different number of images in T2W and DWI MRI. Also, because of the time difference MRI images does not overlap each other. Due to this non-overlapping feature of T2W and DWI MRI sequences we could not use them together. So, while segmenting bladder wall in DWI sequence, we can encounter some difficulties. The erroneous segmentation may affect VIRADS classification negatively. Too many errors observed on the detected tumor areas. We hope that we can overcome these errors when we work on a larger dataset.

4.0.2 Classification Results

The K Nearest Neighbors and Multinomial Naive Bayes classifiers were trained using the feature set described in the Section 3. The feature set obtained from the ground truths and segmentation masks of 3 Tesla DWI MRI images. In addition to these methods, a rule-based method was also developed. This method basically classify VIRADS 1-2-3 scores as one class, VIRADS 4 as another class and differentiate VIRADS 5 from them. The algorithm is:

Algorithm 1 Rule-based Algorithm for VIRADS Classification

if $o_i > 0$ **then**
$Class_i \leftarrow 5$
else if $i_b > 0$ **then**
$Class_i \leftarrow 4$
else
$Class_i \leftarrow 1-2-3$

	Our Method			KNN Method			MNB Method		
	Precision	Recall	F1 Score	Precision	Recall	F1 Score	Precision	Recall	F1 Score
Virads 1-2	0.859	1.00	0.854	0.852	1.00	0.856	0.856	1.00	0.852
Virads 4-5	1.00	0.75	0.856	1.00	0.852	0.850	1.00	0.64	0.78
Microavg	0.852	0.852	0.852	0.854	0.854	0.854	0.859	0.859	0.859
Macroavg	0.855	0.858	0.851	0.856	0.851	0.853	0.853	0.852	0.857

Table 4.1: Results of Ground Truth Masks

As you can see from the Table 4.1, the results from ground truths are convincing. Precision and recall rates are really high for both classes. Our method over-performed Naive Bayes classifier and almost has the same accuracy rate as KNN classifier. This shows that our feature extraction methodology is working well on ground truth annotations.

	Our Method			KNN Method			MNB Method		
	Precision	Recall	F1 Score	Precision	Recall	F1 Score	Precision	Recall	F1 Score
Virads 1-2	0.853	0.23	0.36	0.60	0.71	0.65	0.62	0.76	0.68
Virads 4-5	0.35	0.850	0.50	0.00	0.00	0.00	0.00	0.00	0.00
Microavg	0.44	0.44	0.44	0.48	0.48	0.48	0.52	0.52	0.52
Macroavg	0.59	0.56	0.58	0.30	0.36	0.33	0.31	0.38	0.34

Table 4.2: Results of Mask Rcnn Segmentation

However, as seen in Table 4.2 the Mask RCNN segmented images did not resulted well. This situation is the result of wrong segmentation of tumors in VIRADS 1 and 2 classes. They are segmented as if they expanded to the bladder walls. We did not have much data from VIRADS 4 and 5 classes and this may be an another reason of underfitting in 4 and 5 class. As you can see from the Table 3.2 dataset is dominated by the tumors which has VIRADS score 1 and 2. Therefore we acquired erroneous annotations from segmentation such that it dramatically affects the result of the classification. To overcome this problem, the data set must be populated and the classes must be balanced so that large differences between the classification results of different classes can be avoided.

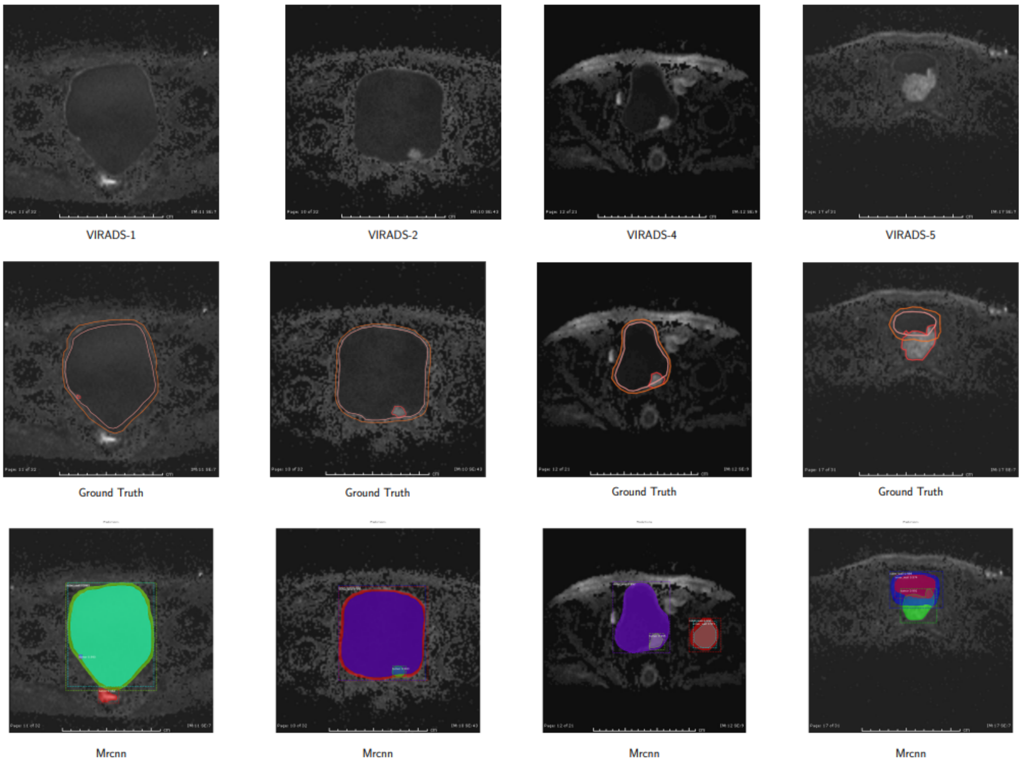

Figure 4.4: Ground Truths and Mask RCNN results of different MRI images.

In 4.4, first row is the raw DWI images. Second row is the ground truths annotated by Dr. Mustafa Orhan Nalbant. Third row contains results of the mask rcnn segmentation.

- In VIRADS -1, muscle invasion is highly unlikely and tumor area does not expand inside the wall borders. Mask RCNN result is success.
- In VIRADS - 2, tumor area should be still inside the bladder. However, mask RCNN segmentation could not exactly detect the wall borders and this cause a tumor expansion inside the wall borders.
- In VIRADS- 4, muscle invasion is likely and tumor area generally expands inside the bladder wall. In this example, Mask RCNN segmentation detected wrong wall borders outside the bladder because there is a field in this MRI really looks like a bladder.
- In VIRADS - 5, tumor area expands outside the bladder and metastasized to other organs. Mask rcnn segmentation is successful in this image.

As seen in these examples, segmentation results are deteriorated due to some difficulties brought by MRI. The successful results of the ground truths show that the methodology is promising and open to progress. Segmentation part should be improved in order to reach better results in classification.

5 Conclusion and Future Work

This project aims to provide early diagnosis and reduce the transition of patients to advanced stages of cancer. In economic terms, thanks to this completely autonomous methodology, each MRI can be uploaded to the system instantly and results can be obtained within seconds. By looking at these results, radiologists can allocate 2 or a third of the time they normally spare for a patient, write the necessary report and start treatment. Thus, both the burden on the doctor is reduced and the number of radiologists to be employed remains constant.

In terms of sustainability, this project can be configured to run in a cloud environment. As the patient sample grows and the amount of data available for each class increases, the model can perform more accurate segmentation and assign a more accurate VIRADS score.

This project can be commercialized by creating a user interface that accepts bladder MRIs with the ethical approval of patients. Any hospital can use this to detect the tumor and assign a VIRADS score to assist their radiologists.

We managed to reach a proper dataset thanks to Dr. Rüştü Türkay and Dr. Mustafa Nalbant. Classification strategy gave promising results in ground truths. Since we have minimal data, the segmentation part did not satisfy the expectations. There must be more data to increase the accuracy of both the segmentation and classification parts. In this way, this model can segment and predict scores more confidently.

Bibliography

[1] Rafael M. Luque-Baena, Ivana Despotović, Bart Goossens, and Wilfried Philips. Mri segmentation of the human brain: Challenges, methods, and applications. *Computational and Mathematical Methods in Medicine*, 2015.

[2] Takuji Hayashi, Go Tanigawa, Kazutoshi Fujita, Ryoichi Imamura, S. Nakazawa, Yoshiyuki Yamamoto, Masahiro Hosomi, Kohki Shimazu, Hiroaki Fushimi, and Seiji Yamaguchi. Two cases of plasmacytoid variant of urothelial carcinoma of urinary bladder: Systemic chemotherapy might be of benefit. *International journal of clinical oncology / Japan Society of Clinical Oncology*, 16:759–62, 07 2011.

[3] Jose Dolz, Xu Xiaopan, Jérôme Rony, Jing Yuan, Yang Liu, Eric Granger, Christian Desrosiers, Xi Zhang, Ismail Ben Ayed, and Hongbing Lu. Multi-region segmentation of bladder cancer structures in mri with progressive dilated convolutional networks, 05 2018.

[4] Kaiming He, Georgia Gkioxari, Piotr Dollár, and Ross Girshick. Mask r-cnn, 2018.

[5] K. Hammouda, F. Khalifa, A. Soliman, M. Ghazal, M. Abou El-Ghar, A. Haddad, M. Elmogy, H. E. Darwish, R. Keynton, and A. El-Baz. A deep learning-based approach for accurate segmentation of bladder wall using mr images. In *2019 IEEE International Conference on Imaging Systems and Techniques (IST)*, pages 1–6, 2019.

[6] Merve Turan. Brain tumor detection and segmentation using deep learning. `https://github.com/mrvturan96/Brain-Tumor-Detection-and-Segmentation-using-Deep-Learning`, 2019.

[7] Martina Pecoraro, Mitsuru Takeuchi, Hebert A. Vargas, Valdair F. Muglia, Stefano Cipollari, Carlo Catalano, and Valeria Panebianco. Overview of vi-rads in bladder cancer. *American Journal of Roentgenology*, 214(6):1259–1268, 2020. PMID: 32286874.

[8] Glossary of terms. In *Web*. National Institute of Biomedical Imaging and Bioengineering, 27 November 2021.

[9] K. Hammouda, F. Khalifa, A. Soliman, M. Ghazal, M. Abou El-Ghar, A. Haddad, M. Elmogy, H. E. Darwish, A. Khalil, A. Elmaghraby, R. Keynton, and A. El-Baz. A cnn-based framework for bladder wall segmentation using mri. In *2019 Fifth International Conference on Advances in Biomedical Engineering (ICABME)*, pages 1–4, 2019.

[10] Valeria Panebianco, Yoshifumi Narumi, Ersan Altun, Bernard H. Bochner, Jason A. Efstathiou, Shaista Hafeez, Robert Huddart, Steve Kennish, Seth Lerner, Rodolfo Montironi, Valdair F. Muglia, Georg Salomon, Stephen Thomas, Hebert Alberto Vargas, J. Alfred Witjes, Mitsuru Takeuchi, Jelle Barentsz, and James W.F. Catto. Multiparametric magnetic resonance imaging for bladder cancer: Development of vi-rads (vesical imaging-reporting and data system). *European Urology*, 74(3):294–306, 2018.

[11] J.O.S.H. Cleary and A.R. Guimarães. Magnetic resonance imaging. In Linda M. McManus and Richard N. Mitchell, editors, *Pathobiology of Human Disease*, pages 3987–4004. Academic Press, San Diego, 2014.

[12] In Steven D. Waldman and Robert S.D. Campbell, editors, *Imaging of Pain*, pages 461–480. W.B. Saunders, Philadelphia, 2011.

[13] MD David C Preston. Magnetic Resonance Imaging (MRI) of the Brain and Spine: Basics, 7 April 2016.

[14] Dandu Ravi Varma. Managing dicom images: Tips and tricks for the radiologist. *Indian Journal of Radiology and Imaging*, February 2012.

[15] Olaf Ronneberger, Philipp Fischer, and Thomas Brox. U-net: Convolutional networks for biomedical image segmentation, 2015.

[16] Waleed Abdulla. Mask r-cnn for object detection and instance segmentation on keras and tensorflow. `https://github.com/matterport/Mask_RCNN`, 2017.

[17] Onesmus Mbaabu. Introduction to Random Forest in Machine Learning, December 11, 2020.

[18] Brats 2019. `https://www.med.upenn.edu/cbica/brats-2019/`, 2019.

[19] U-net weights and mask-rcnn models. `https://drive.google.com/drive/folders/1b5_nN-6a0a_uYxxESsEpDpLf3HGkkOqg`, 2019.

[20] Mark Bandyk, Dheeraj Gopireddy, Chandana Lall, K.C. Balaji, and Jose Dolz. Bladder segmentation based on deep learning approaches: current limitations and lessons, 01 2021.

[21] Linear mapping method using affine transformation. https://www.mathworks.com/discovery/affine-transformation.html#:~:text=Affine%20transformation%20is%20a%20linear,with%20non-ideal%20camera%20angles., note = Accessed: 2022-04-28.

[22] Loocv for evaluating machine learning algorithms. https://machinelearningmastery.com/loocv-for-evaluating-machine-learning-algorithms/#:~:text=The%20Leave-One-Out%20Cross,used%20to%20train%20the%20model., note = Accessed: 2022-04-28.

[23] Rongjian Zhao, Buyue Qian, Xianli Zhang, Yang Li, Rong Wei, Yang Liu, and Yinggang Pan. Rethinking dice loss for medical image segmentation. In *2020 IEEE International Conference on Data Mining (ICDM)*, pages 851–860, 2020.

[24] Deepmedic.

[25] Usman Khan, AmanUllah Yasin, Muhammad Abid, Imran Shafi, and Shoaib Khan. A methodological review of 3d reconstruction techniques in tomographic imaging. *Journal of Medical Systems*, 42:190, 09 2018.

[26] J. Jones and A. Haouimi. T2 weighted image. reference article. In *Radiopaedia.org.*, 2009. Accessed: 23.01.2022.

[27] U. Bashir and Y. Baba. Diffusion-weighted imaging. reference article. In *Radiopaedia.org.*, 2012. Accessed: 23.01.2022.

[28] Elisha Odemakinde. Mask r-cnn: A beginner's guide. https://viso.ai/deep-learning/mask-r-cnn/. Accessed: 23.01.2022.

[29] Dhruvil Shah. Mask rcnn implementation on a custom dataset! https://towardsdatascience.com/mask-rcnn-implementation-on-a-custom-dataset-fd9a878123d4. Accessed: 23.01.2022.