

GI-Tract-Image-Segmentation

Yasien Essam Mohamed
Cairo University

Andrew Muhsen Paramedis
Cairo University

Omar Essam Mohamed
Cairo University

Mohamed Alaa
Cairo University

Abstract

In this report, we present our approaches to the Kaggle GI Tract Image Segmentation challenge. Since radiation oncologists try to deliver high doses of radiation using X-ray beams pointed to tumors while avoiding the stomach and intestines, the goal of the challenge is to effectively segment the stomach and intestines in MRI scans in order to improve the cancer treatment, circumventing the need for the time-consuming and labor intensive process in which radiation oncologists must manually outline the position of the stomach and intestines. We apply U-Net method to segment the organ areas. Our best U-Net model achieves a Jaccard Index of 0.96 on the validation set..

1. Introduction

Each year, around 5 million people are diagnosed with gastrointestinal cancer worldwide, with about half of them eligible for radiation therapy. This type of treatment uses X-ray beams to deliver high doses of radiation to the tumor, while trying to avoid the stomach and intestines. With the help of newer technologies like MR-Linacs (which combine magnetic resonance imaging and linear accelerator systems), doctors can visualize the daily position of the tumor and the stomach and intestines, allowing them to adjust the direction of the X-ray beams as needed. However, this process requires radiation oncologists to manually outline the position of the stomach and intestines, which can be time-consuming and labor-intensive. It can also extend treatment time from 15 minutes to an hour, which can be difficult for patients to tolerate. Using deep learning to automate the segmentation process (i.e., the process of identifying and outlining the stomach and intestines) could make treatments faster and more effective, potentially allowing more patients to receive this type of care.

We are developing a model to appropriately segment the stomach and intestines using MRI images for this research. The model is fed MRI scans from cancer patients who had 1-5 MRI scans on different days throughout their radiation therapy at the UW-Madison Carbone Cancer Center. The U-Net algorithm is then used to forecast segmented sections of patients' MRI data with labels for "stomach," "big bowel," and "small bowel."

2. Related Work

2.1. U-Net Type Networks

Image segmentation is a common task in computer vision, and it has many applications in the field of biomedicine. In 2015, Ronneberger *et al.* [1] proposed a new network architecture called U-Net as shown in fig[1], which was specifically designed to address the challenges of biomedical image segmentation, where training data is often limited. U-Net was initially developed for the task of segmenting neuronal structures in electron microscopy images, but the authors demonstrated that it was effective for a wide range of other biomedical image segmentation tasks as well. They also introduced data augmentation techniques to improve the performance of U-Net in these tasks.

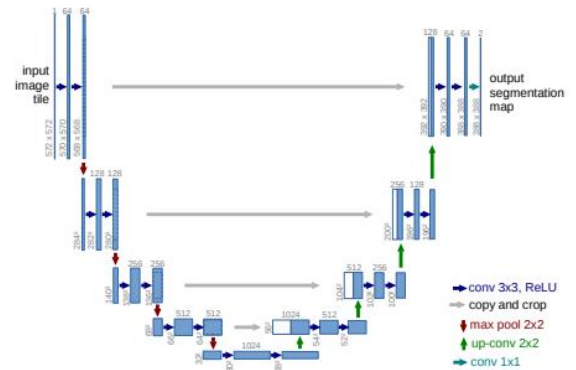


Fig. 1. U-net architecture (example for 32x32 pixels in the lowest resolution). Each blue box corresponds to a multi-channel feature map.

One of the key features of U-Net is its fully convolutional network structure, which allows it to perform precise localization, or the assignment of class labels to each pixel in an image. This enables U-Net to achieve image segmentation. U-Net also includes upsampling layers, which help to increase the resolution of the output by replacing traditional pooling layers. These design choices enable U-Net to effectively identify and segment objects of interest in images, even when the training data is limited.

There are several network architectures that have been proposed for the task of medical image segmentation. In 2016, Milletari et al.[2] introduced V-Net, a network architecture similar to U-Net, but designed specifically for segmenting volumetric images, such as those obtained from MRI scans. V-Net is fully convolutional and uses 3D convolutions to learn the volumetric structure of image stacks. The authors also applied non-linear transformations and histogram matching to the training data to improve performance.

In 2018, Zhou et al.[3] proposed a new architecture called U-Net++, which builds upon the original U-Net architecture by using nested U-Net structures and redesigned connectivity to make optimization easier. Another important model for medical image segmentation is CE-Net, which was proposed by Gu et al. in 2019[4]. CE-Net is built on the approach of U-Net and is capable of preserving spatial information that is lost in the consecutive pooling and strided convolutional operations of U-Net. CE-Net has three major components: a feature encoder module, a context extractor, and a feature decoder module. It has been applied to various 2D medical image segmentation tasks and has achieved significantly better results compared to U-Net and other state-of-the-art methods. U-Net-type networks have also been used in a variety of biomedical vision tasks, such as pancreas segmentation and cell counting and detection. This demonstrates the versatility and effectiveness of the U-Net architecture in biomedical applications.

3. Dataset and Features

3.1. Dataset Details

The data we use is anonymized MRIs of patients treated with MRI guided radiotherapy provided by the UWMadison Carbone Cancer Center. Specifically, the dataset contains 85 cases with 38496 scan slices of organs represented in 16-bit grayscale PNG format, and the annotations are provided in a csv format with the segmented areas represented as RLE-encoded masks. Sample csv annotations are shown in Table [1]. An empty segmentation entry represents no mask presented for the class in the MRI scan slice. All the data is available as part of the Kaggle challenge[5]

| ID | | Class | Segmentation |
|------------|-------|-------------|---------------------|
| case123 | day20 | large bowel | 17746 6 18010 9 ... |
| slice 0081 | | | |
| case123 | day20 | small bowel | |
| slice 0081 | | | |
| case123 | day20 | stomach | 17481 4 17746 7 ... |
| slice 0081 | | | |

Table 1. MRI Scan Dataset CSV Annotation Sample

3.2. Preprocessing

Upon examining the dataset, we noticed that some images are ambiguous and do not contain much visual information. In order to improve the performance of our model, we preprocessed the dataset by selecting only the images that have sufficient visual information. This helps to ensure that our model is able to learn meaningful features from the data and make accurate predictions.

Then we split the class label into three categories instead of one to help us segment each one separately. In order to create masks for segmented areas, we also have to ensure the sizes of images are the same, so we reshape all images to 128 x 128.

.Since the training labels in the dataset are in the form of 16-bit RLE encoded masks, which cannot be easily understood or analyzed by simply looking at them., we convert them into pixel values and visualize the labeled areas by highlighting the masks as in fig[2]

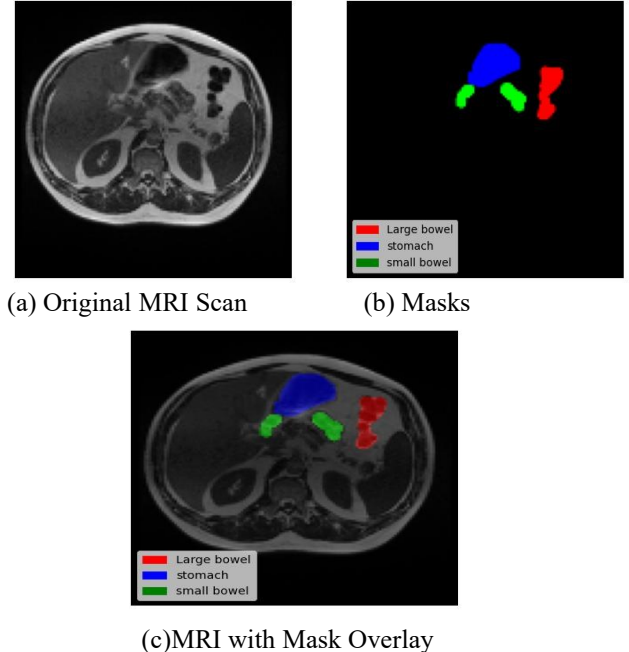


Figure2. Example Preprocessing Mask Visualization

Finally, we split up the dataset to training, validation, and testing sets by 80%, 10%, and 10% respectively.

4. Methods

4.1. Loss Function

We are using Categorical Cross-Entropy loss as shown below since we modified our true labels to one hot encoded

$$L_{CE} = - \sum_{i=1}^n t_i \log(p_i), \text{ for } n \text{ classes,}$$

Where t_i is the truth label and p_i is the softmax probability for the i^{th} class.

4.2. Baseline Method

We saved the images after the preprocessing as NumPy arrays in a npy file so we could exceed them in our model notebook, we took half the images of the dataset so we could save resources, we then split our data and supply our model with 80% of the images to train them and we get a jaccard_coef of 0.9705 and loss of 0.0287 and the The remaining 20% we used for test and validation

4.3. U-Net

We used the U-Net Model on our dataset provided by Sreenivasa Reddy[6]. The network consists of The network consists of two main parts, the down sampling part and the up sampling part (thus the name U-Net). The down sampling part has 4 steps, and we will call each step a step down layer. The 4 step down layers get exponentially larger, i.e. the first one has a double convolutional layer of type (16, 32), the second has a double convolutional layer of type (32, 64), the third (64, 128), and the fourth (128, 256). The MaxPool layer always has kernel size 2 and stride 2.

The up sampling part has 4 corresponding steps, and we will call each step a step up layer. A step up layer consists of a upsample layer and a double convolutional layer. The 4 step up layers get exponentially smaller, i.e. the first one has a double convolutional layer of type (256, 128), the second has a double convolutional layer of type (128, 64), the third (64, 32), and the fourth (32, 16).

5. Experiments/Results/Discussion

5.1. Evaluation method

We finally evaluate our model by using the categorical Cross-Entropy loss (between the predicted mask and the true mask) as described above.

We had 2 approach one was using simple U-Net and the other one was to apply transfer learning from open source implementation of segmentation model[7] with pretrained backbones but the accuracy drop.

5.2. U-Net

For the U-Net model, our implementation mainly follow the open source implementation by Sreenivasa[6]. We train the U-Net model on our dataset from scratch, and we experiment with multiple hyperparameter combinations as well as tune on the validation set.

We use categorical_crossentropy as our loss function, Adam as our optimizer and our accuracy coefficient was jaccard. Our final model has a total parameter of 1,940,868.

And this the final result of our U-net model that we manage to achieve as shown in fig 3.

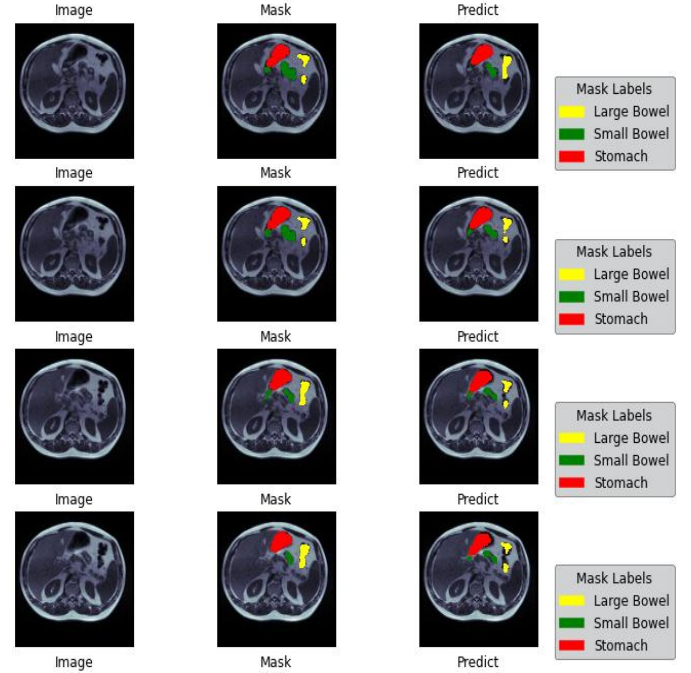
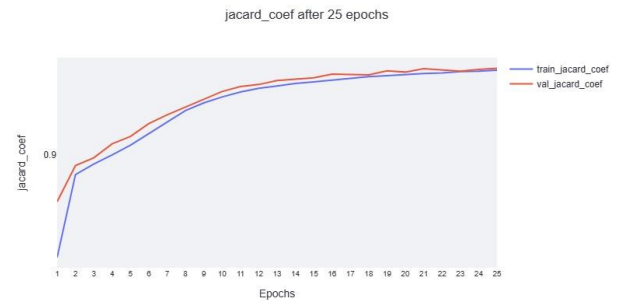


Fig 3. Prediction with mask overlay Example Results

And we manage to achieve a jaccard score of 0.9626 on the test data and loss of 0.0362 as shown in fig 4.



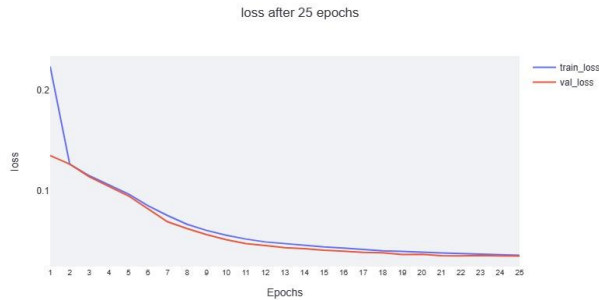


fig 4. Best Performing U-Net Results

6. Conclusion/Future Work

In this project, we implemented and modified the UNet model, and applied them to the GI tract image segmentation problem. Our U-Net based model achieves a jaccard score of 0.54 on the training set and 0.5134 on the validation set. We can see from the high score of our model that he might overfit to our train data but he acts well on the test data, so our hypothesis is that our model may not act that well if we test it on images from different distrubtion, So our future work is:

1. Formulate a more sophisticated loss function for the U-Net model. Considering the current behavior of our U-Net model. including a term that punishes overpredicting might increase the performance.
2. Make our model expose to more images from different datasets so it can predict well

7. Contributions and Acknowledgements

Omar and yasien have done the U-net model and mohamed alaa done the U-net model with transfer learning and Andrew has done the data visulization and the literature review, and we all contrirube to make the final report.

References

- [1] Zhou, Z., Siddiquee, M. M. R., Tajbakhsh, N., & Liang, J. (2018). Unet++: A nested u-net architecture for medical image segmentation. In Deep learning in medical image analysis and multimodal learning for clinical decision support (pp. 3-11). Springer.
- [2] Fausto Milletari, Nassir Navab, and Seyed-Ahmad Ahmadi. V-net: Fully convolutional neural networks for volumetric medical image segmentation. In 2016 fourth international conference on 3D vision (3DV), pages 565–571. IEEE, 2016..
- [3] Zongwei Zhou, Md Mahfuzur Rahman Siddiquee, Nima Tajbakhsh, and Jianming Liang. Unet++: A nested u-net architecture for medical image segmentation. In Deep learning in medical image analysis and multimodal learning for clinical decision support, pages 3–11. Springer, 2018.

- [4] aiwang Gu, Jun Cheng, Huazhu Fu, Kang Zhou, Huaying Hao, Yitian Zhao, Tianyang Zhang, Shenghua Gao, and Jiang Liu. Ce-net: Context encoder network for 2d medical image segmentation. CoRR, abs/1903.02740, 2019
- [5] [UW-Madison GI Tract Image Segmentation | Kaggle](#)
- [6] B. Sreenivasa Reddy. Semantic segmentation of aerial imagery using U-Net.Retrieved from ([Github](#)) (2020, December 22).
- [7] [qubvel/segmentation_models: Segmentation models with pretrained backbones. Keras and TensorFlow Keras. \(github.com\)](#)