


Computer Vision Project: Segmentation for Medical Images

Huy G. Tong 

Student ID: 21127307

Faculty of Information Technology
HCMC University of Science
tghuy212@clc.fitus.edu.vn

Khoi A. D. Phan

Student ID: 21127325

Faculty of Information Technology
HCMC University of Science
pdakhoi21@clc.fitus.edu.vn

Quan A. Q. Lam

Student ID: 20127291

Faculty of Information Technology
HCMC University of Science
lqaquan20@clc.fitus.edu.vn

Abstract—In the medical field, understanding images is one of the most important processes. Given a photo taken using x-ray or endoscopy, doctors will have to identify regions where the body is damaged. One drawback of this manual process is that it takes time and effort for the doctor to make the appropriate analysis. During rush hours when there are many patients waiting to be diagnosed, the process could also be less accurate which could lead to devastating results. With the technique of segmentation, which is a process of recognizing and separating regions of interest from the original image, it is possible to massively boost the productivity and reduce stress on the medical staff when performing diagnosis. In this project, we investigate some of the widely used techniques for medical image segmentation. In this report, we focus on the prior works of segmentation procedure. Additionally, we present our study on the U-net structure, as well as some experimentation and assessment using the proposed method.

Index Terms—medical imaging, segmentation

I. INTRODUCTION

The field of medical imaging stands at a crossroads of technological innovation and clinical application. By analysing images taken by techniques like Magnetic Resonance Imaging (MRI), Positron Emission Tomography (PET), Computed Tomography (CT), and Ultrasound, doctors and medical workers can have a clearer view of the status of the patients, thus making diagnosis more accurate. However, the manual process of inspecting individual images can be time consuming, and at times, due to heavy workloads, can also be incorrect which can be undesirable. Therefore, there calls for a way to help automate the process whenever possible, which will allow the medical staff to be more productive.

By harvesting the power of technology and computer vision, it is possible to identify regions of interests in medical images, which will help speed up the process of diagnosis. Segmentation, the process of dividing a digital image into distinct segments or pixel sets, is essential and very capable for such task. By extracting valuable information from complex medical images, it enables the detailed isolation of anatomical structures and pathological regions, serving as a foundation for analysis and interpretation across various medical fields.

As technologies for imaging advanced, the demand for sophisticated image segmentation techniques has grown. These methods have evolved from manual to semi-automatic and

fully automatic processes, utilizing the latest in artificial intelligence (AI) and machine learning algorithms. Despite progress, challenges persist due to the complexity of human anatomy, patient variability, diverse imaging modalities, and the fine line between normal and pathological tissues.

This project aims to explore the landscape of segmentation in medical imaging, highlighting its impact on the medical field. We focus on three advanced methods that have significantly influenced medical image analysis: the Model-based method, Threshold-based method, and Classification method. Each method brings unique advantages and is designed for specific clinical scenarios.

In the next sections, we will examine some of the widely used methods for segmentation in order to grasp a better view and understanding of the current technologies. Next, we will highlight a method that best balances accuracy, efficiency, and complexity for our own model. Finally, we will implement the discussed method, train, and evaluate its performance.

II. RELATED WORKS

There have been many surveys conducted to give an overview of the current available algorithms for segmentation, namely [2], [3], and [4]. In this section, we summary the available methods.

Threshold-based method. The threshold approach aims on selecting the pixels based on a certain feature limit (i.e. intensity, orientation, etc.) given by a threshold value. This class of algorithm can be further divided into edge-based and region-based segmentation.

Edge-based methods works with the edges of objects/regions in the image. In theory, the edges detected should separate different regions, thus indicating different segments in the image. The edges are detected using some of the widely used kernels like Sobel, Laplacian of Gaussian, or the Canny detector. A threshold is then applied on the detected edges to further refine the features and to suppress noise. Finally, pixels bounded by the edges are grouped together to produce a segment.

The region-based approaches are constructed based on the idea that pixels belonging to the same part of an object will have similar characteristics. The region growing algorithm harvests this finding by first initializing seed points across the

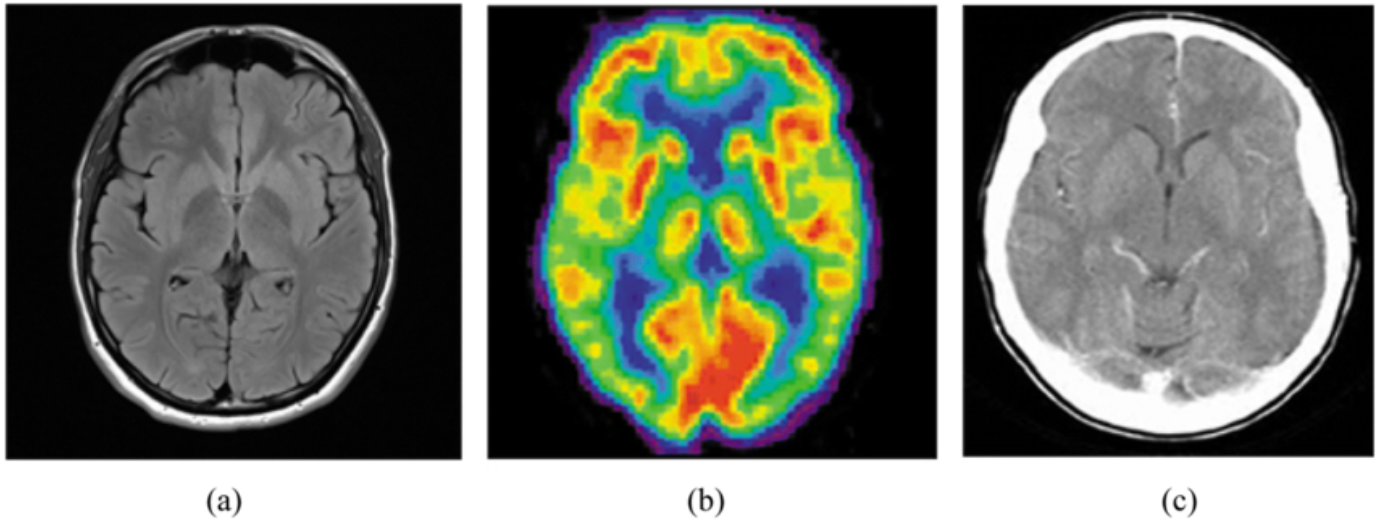


Fig. 1. Image of the brain captured using different techniques: (a) MRI, (b) PET scan, (c) CT scan. Figure taken from [1].

image. These seeds will then gradually grow to neighboring pixels that are inside the threshold range, indicating similar features. The technique of region growing have been applied for segmentation in mammogram images [5].

These methods have now been dubbed the *traditional approach* as they implements feature engineering along with step-by-step algorithms. Some advantages include having a faster runtime and require nearly no training step to produce result. However, for some more complex or noisy data, these approaches are more prone to error. Therefore, other more flexible ways for segmenting images have been researched.

Classification method. Most of the time, disease of the same kind usually have similar characteristics, therefore, classification, or pattern recognition, processes could be used to identify and separate the segment from the image. For *supervised classification*, each pixel in the image can be assigned to one segment or region, then a model will be trained based on those labelled data. Some widely used models for medical image segmentation include artificial neural network [6], support vector machine [7], and active appearance model [8]. More complex structures can include the well-known convolutional neural network [9], U-Net [10], which was initially developed to perform segmentation on medical images, and multimodal deep learning which can learn from a variety of data and imaging types (PET, CT, MRI, etc., see Fig. 1) [11]. One down side of this approach is that the training and testing process requires labelled data, which is not always available. To tackle that problem, unsupervised algorithms are also used.

Different from supervised algorithms, *unsupervised* methods can learn from unlabelled data. Some examples include k-means clustering and hierarchical clustering. In these processes, pixels are gradually grouped together based on the their feature similarities (intensity, color, orientation, etc.).

Model-based method. This segmentation techniques provide flexibility and precision when dealing with complex data.

Unlike machine learning methods, these approaches use pixel intensity, gradients, and other image features to recognize and outline shapes or regions that resemble organs in the body. By employing predefined mathematical or statistical models, they can accurately identify and separate areas of interest within the image.

III. PROPOSAL

The techniques for image segmentation are constantly being developed and improved. More recent methods can have better performance and accuracy than prior approaches, however, the complexity of such algorithms can be significantly higher. Taking into consideration the goal of the project, as well as our current knowledge of the matter, we decide to use a **convolutional neural network model** to further study the detailed process of segmentation, more specifically, we will dive deeper into the **U-net structure** [10]. We believe that with our current and upcoming knowledge from class, as well as the abundance of materials online, this type of model is the most suitable for our team.

IV. METHOD

A. The U-net architecture

U-net [10] is one of the highly regarded structures for image segmentation. Initially developed to work on medical data, it now have many variations and modifications, as well as have seen many usage across different fields, including astronomy [12].

Fig. 2 presents the overall structure of U-net. The architecture contains 2 flows: an encoding path (left side) and a decoding path (right side). Each step of the down sampling encoder path consists of 2 simultaneous convolutions with ReLU activation and 1 max pooling operation. For the decoding process, first the feature map is up sampled and then concatenated with the previous corresponding map from the

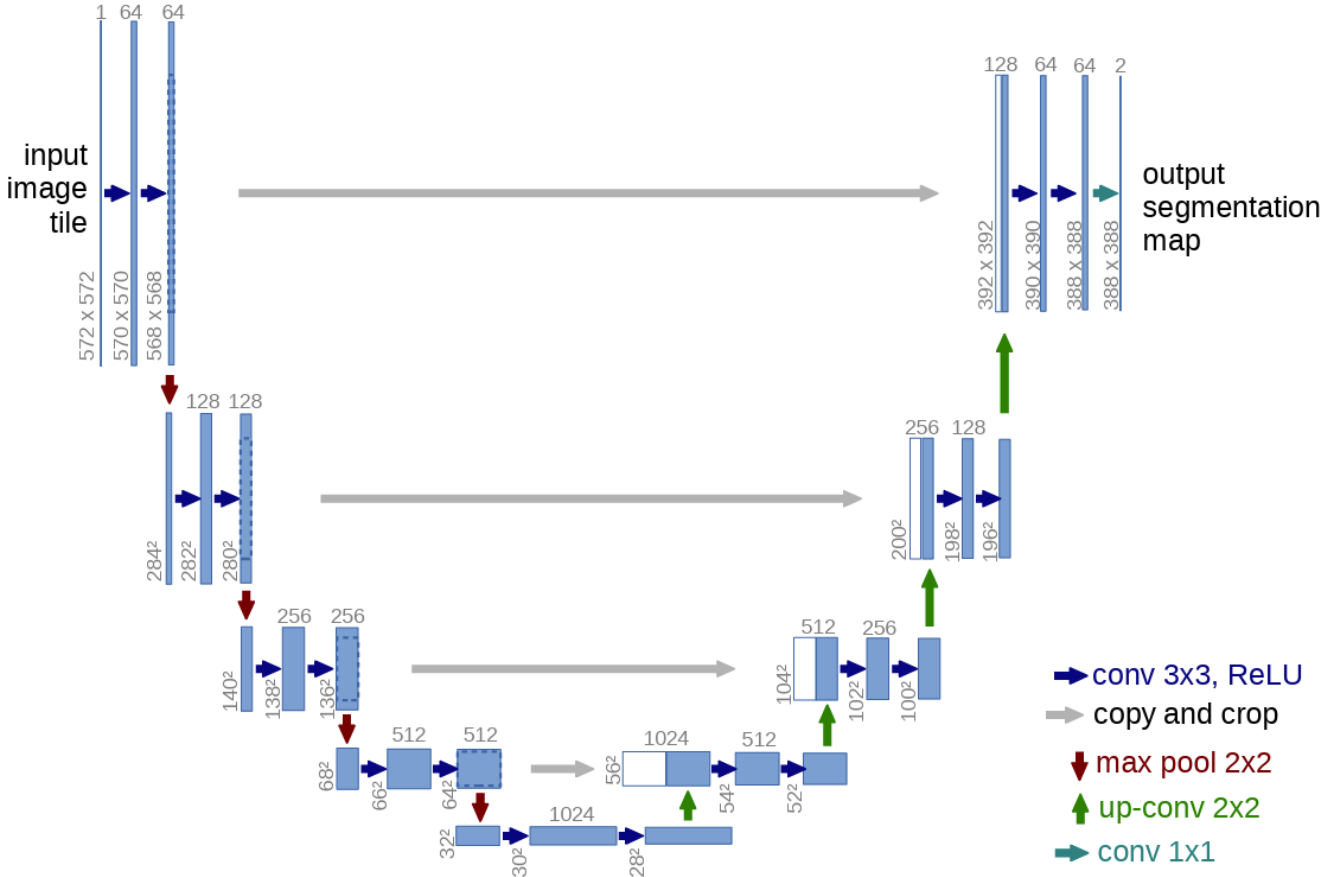


Fig. 2. The U-net architecture. Figure taken from [10].

encoding process, finally, 2 convolutions are applied to the combined map with ReLU activation.

B. Implementation

To compile the U-net model in Python, we follow the implementation of user [nikhilroxtomar](#) on GitHub ([UNet-Segmentation-in-Keras-TensorFlow](#)). During training, we also make use of previous pre-trained weights to make the process more efficient. The pre-trained model is obtained during our previous test runs with 200 epochs, we include that model in our source under the name `pretrained.h5`.

After training and testing, we save the model's weights to a `.h5` file, which can then be used to make real time predictions. In the `Release` folder, we include a simple Python script that loads the trained model along with the user's input image to make a prediction mask. Simply run the script, follow the prompt to input the path to the image, the program will then display a side-by-side comparison of the input and predicted mask. The mask will also be saved to a separate file under the name `prediction.png`.

V. EXPERIMENTS AND DISCUSSION

We trained and tested the model on the Retina Blood Vessel dataset available on [Kaggle](#). The data includes a training set of 80 images and masks, as well as a testing set of 20 samples. In each set, the images and masks are organized into 2 separate folders (`image` and `mask`) so that each image name corresponds to the similar mask and vice versa. This directory structure is incorporated into our source for model training, therefore other re-train attempts with different data would have to follow the similar structure.

We trained the model with maximum 200 epochs. By using the pre-trained weights described earlier, the process only took 23 epochs to be within our early stopping threshold. The final testing accuracy is around 0.8653. Some sample results are presented in Fig. 3.

Overall, the model performed pretty well as it successfully identified the major blood vessels in the retina. Others smaller vessels are more prone to error, as the model sometimes fail to recognize the feature or miss identified the region. To tackle these limitations, newer and more robust variations of U-net have been developed, which introduced new flows of data throughout the network. In the upcoming work, we will further

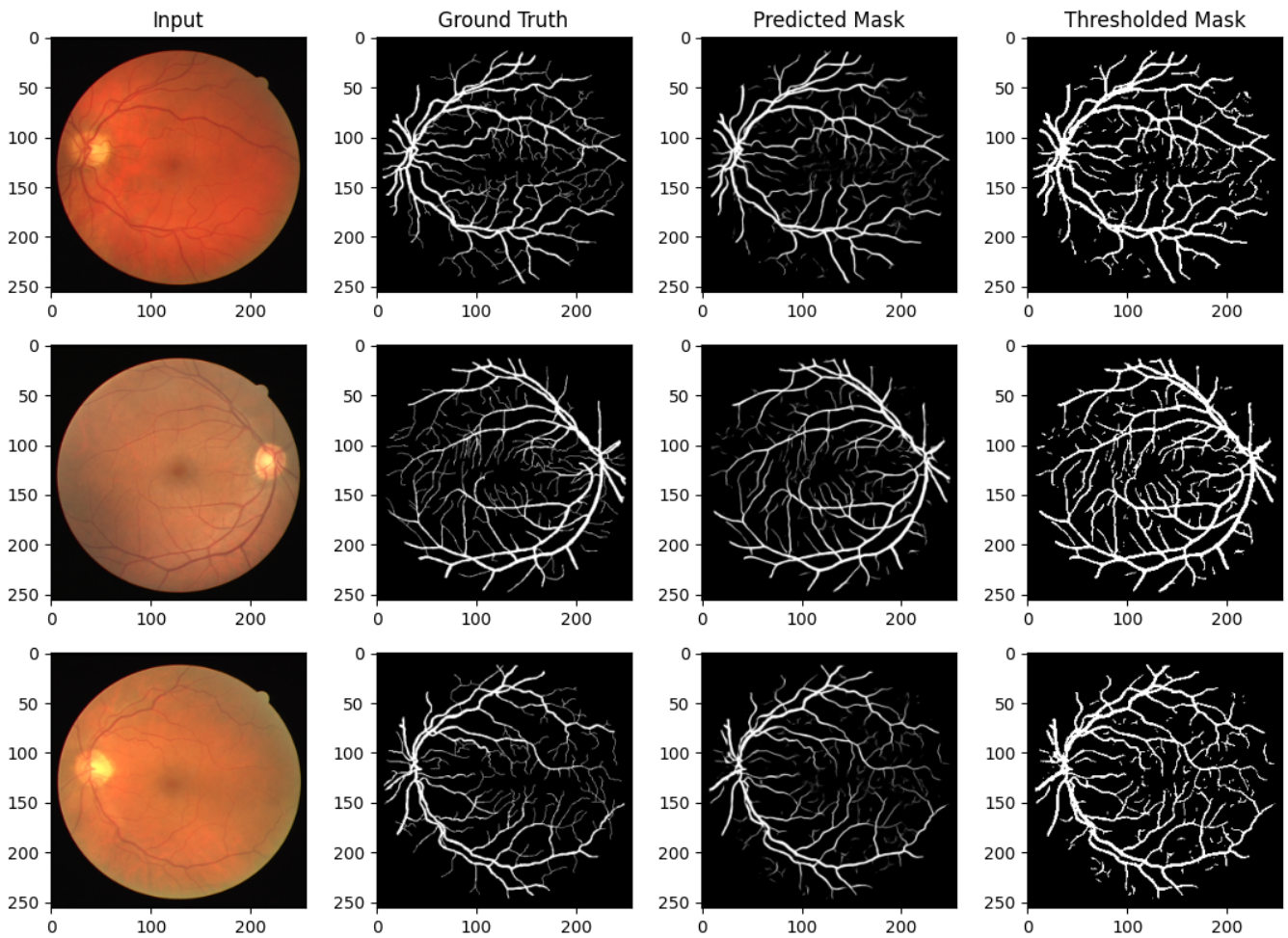


Fig. 3. Sample inputs and outputs of our model.

investigate some of these variations to hopefully improve the accuracy of this experiment.

ACKNOWLEDGMENT

The authors thank the lecturers of the *Computer Vision* course at the Ho Chi Minh University of Science for their instruction in theory and lab session as well as support during the project duration.

REFERENCES

- [1] S. Fouladi, A. A. Safaei, N. I. Arshad, M. J. Ebadi, and A. Ahmadian, "The use of artificial neural networks to diagnose alzheimer's disease from brain images," *Multimedia Tools and Applications*, vol. 81, pp. 37681–37721, Nov 2022.
- [2] K. Ramesh, G. K. Kumar, K. Swapna, D. Datta, and S. S. Rajest, "A review of medical image segmentation algorithms," *EAI Endorsed Transactions on Pervasive Health and Technology*, vol. 7, p. e6, Apr. 2021.
- [3] A. Norouzi, M. S. M. Rahim, A. Altameem, T. Saba, A. E. Rad, A. Rehman, and M. Uddin, "Medical image segmentation methods, algorithms, and applications," *IETE Technical Review*, vol. 31, no. 3, pp. 199–213, 2014.
- [4] Z. Ma, J. M. R. S. Tavares, and R. M. N. Jorge, "A review on the current segmentation algorithms for medical images," in *Proceedings of the First International Conference on Computer Imaging Theory and Applications - Volume 1: IMAGAPP, (VISIGRAPP 2009)*, pp. 135–140, INSTICC, SciTePress, 2009.
- [5] A. Oliver, J. Freixenet, J. Martí, E. Pérez, J. Pont, E. R. Denton, and R. Zwigglelaar, "A review of automatic mass detection and segmentation in mammographic images," *Medical Image Analysis*, vol. 14, no. 2, pp. 87–110, 2010.
- [6] J. Alirezaie, M. Jernigan, and C. Nahmias, "Neural network-based segmentation of magnetic resonance images of the brain," *IEEE Transactions on Nuclear Science*, vol. 44, no. 2, pp. 194–198, 1997.
- [7] S. Wang, W. Zhu, and Z.-P. Liang, "Shape deformation:

- Svm regression and application to medical image segmentation,” in *Proceedings Eighth IEEE International Conference on Computer Vision. ICCV 2001*, vol. 2, pp. 209–216 vol.2, 2001.
- [8] T. Cootes, G. Edwards, and C. Taylor, “Active appearance models,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 23, no. 6, pp. 681–685, 2001.
 - [9] B. Kayalibay, G. Jensen, and P. van der Smagt, “Cnn-based segmentation of medical imaging data,” 2017.
 - [10] O. Ronneberger, P. Fischer, and T. Brox, “U-net: Convolutional networks for biomedical image segmentation,” 2015.
 - [11] Z. Guo, X. Li, H. Huang, N. Guo, and Q. Li, “Deep learning-based image segmentation on multimodal medical imaging,” *IEEE Transactions on Radiation and Plasma Medical Sciences*, vol. 3, no. 2, pp. 162–169, 2019.
 - [12] Bekki, K., “Quantifying the fine structures of disk galaxies with deep learning: Segmentation of spiral arms in different hubble types,” *A&A*, vol. 647, p. A120, 2021.