IMAGE SEGMENTATION: IDENTIFICATION OF SURGICAL TOOLS IN DA VINCI OPERATED GASTROINTESTINAL PROCEDURES

Lola Assad, Adam Cockell, Jackson Reid

INTRODUCTION

Recent advancements in machine vision have revolutionized many distinct fields of study. In particular, the implementation of deep learning techniques in medical contexts has seen much early success. This proposal outlines the central initiative of this project which aims to apply the use of image segmentation models in a medical setting, focusing on the Da Vinci surgical system.

RESEARCH PROBLEM, AND OBJECTIVES

The purpose of the model is to explore image segmentation models with applications in medical imaging. Using a semi-supervised learning approach, the overarching objective is to design a model to distinguish organic tissue from surgical tools in Da Vinci surgical procedures. The relationship between object detection and image segmentation in this setting will also be explored to determine if a workflow starting with object detection to establish a bounding box, then using image segmentation on the reduced area makes for a more efficient algorithm than solely image segmentation.

LITERATURE REVIEW

Key Theories and Concepts

Convolutional Neural Networks[1] are the basis of much of image processing and are therefore used extensively in image segmentation tasks.
Image Segmentation[2] is the partitioning of an image. In the case of this project, partitioning the image into organic tissue and inorganic tissue by segmenting out the surgical tools.
Object Detection[3] is the localization of objects in images (without necessarily pixel partitioning) often through use of bounding boxes.
Fine Tuning a Pretrained Model[4] is the practice of applying a head to a model to specify a pre trained model to a specific task. Most often, the pretrained model is trained on a general dataset, and accuracy for a specific task that has not been seen in the training set is improved through finetuning.

Existing Models

The dataset for this project has been published in the 2022 MICCAI conference[5], and a model was developed for similar purposes in surgical scene segmentation using image segmentation techniques. There are currently other models that have been developed for the purpose of surgical scene segmentation for computer assisted surgery, where our model differs is our teams concentration on surgical tools, specifically the Da Vinci graspers in gastrointestinal procedures.

---

[1] Yamashita, Rikiya, et al. "Convolutional neural networks: an overview and application in radiology." Insights into imaging 9 (2018): 611-629.
[2] Yu, Ying, et al. "Techniques and challenges of image segmentation: A review." Electronics 12.5 (2023): 1199.
[3] Zhao, Zhong-Qiu, et al. "Object detection with deep learning: A review." IEEE transactions on neural networks and learning systems 30.11 (2019): 3212-3232.
[4] Poojary, Ramaprasad, and Akul Pai. "Comparative study of model optimization techniques in fine-tuned CNN models." 2019 International Conference on Electrical and Computing Technologies and Applications (ICECTA). IEEE, 2019.
[5] https://conferences.miccai.org/2022/papers/489-Paper2739.html

METHODS

First our team will annotate a small subset of the data using both bounding boxes and pixel precise outlines. To be clinically relevant as a rule of thumb, 20 cases of each category are recommended to weakly supervise a model based on medical data. Since our data is binary by nature (surgical tool vs organic tissue), we will annotate a minimum of 40 images from the dataset. We will then test different pretrained models with a finetuned head to determine best fit for our purposes, and record the results. Models we are looking to consider include:

- YOLOv5
- SPM: Neuroimaging Segmentation
- SAM: Segment Anything Model
- ImageNet

After a model is selected, we will finetune it with our data to increase its accuracy. Then, we will test the model with various validation subsets of the data. There are approximately 4000 images in our dataset which gives us the ability to test quite rigorously with generous validation sets whilst still having sufficient training and testing sets. It is important to note that our model is not being developed for medical use and is strictly a proof of concept.

EXPERIMENTAL DESIGN

While not exhaustive and potentially subject to change according to the project goals, the following will provide the basis for determining effectiveness: Pixel accuracy, Intersection over union (IoU), and F1 score.

Pixel accuracy describes the number of correctly labeled pixels compared to the total pixels in an image. Intersection over union describes how closely the model labels individual classes when compared to ground truth. F1 score is a more complex metric describing the harmonic mean between two other measures - precision and recall. Precision is the number of correctly labeled pixels compared to the total labeled pixels for a class, and recall is the number of correctly labeled pixels compared to the total ground truth pixels for a class.

EXPECTED RESULTS, AND ANALYSIS

The metrics listed above evaluate both the efficiency and accuracy of a model in its application, allowing for adjustments and optimizations to be made based on these values to improve performance. The model is expected to identify each class in the image with over 80% accuracy. To analyze the model, the ROC curve will be considered along with the F1 score. Additionally, since the false negative rate should be minimized possible, the sensitivity metric will also be analyzed.

DATA COLLECTION AND PREPROCESSING

The dataset for this study consists of images taken from 40 surgical videos of distal gastrectomy for gastric cancer, as well as some synthetic images regarding the same procedure. Each image that we will use contains tissue and one or more operation tools. The goal of this study is to differentiate operation tools from tissue with high accuracy in order to further the development of computer assisted surgery. In order to do this, we will manually label some of our data by coloring the operation tools so that we can use convolutional neural networks in a semi-supervised learning space. If semi-supervised learning does not result in an accurate model, we will label more data and take a fully supervised approach.

TIMELINE

We will begin this study by labeling our data. Our data is binary (tissue or instrument) so we will label a minimum of 40 images to begin, which will take a week unless more data is necessary. Once the data has been labeled, we will experiment with fine tuning pre-trained models in order to have an accurate model with less labeled data than normally required. There are several different pre-trained models to experiment with, and so this will take up a month.

After deciding on a well tuned model, we will test it with the remaining unlabeled data which should take several weeks. Once the model is trained and tested, we will spend several weeks working on the slideshow and report, as well as practicing our delivery of the slideshow.

POTENTIAL CONTRIBUTIONS

The industry of computer assisted surgery has vast potential but has hit major roadblocks with the lack of training data available to those looking to improve upon the existing ways. Although image segmentation technology is rapidly developing, engineers are unable to contribute due to the lack of data to train upon.

This study aims to provide an alternative path to this roadblock by exploring the use and fine tuning of existing models with limited data. Our goal is to create an accurate model with limited training data. If successful with clinical approval, this will open the floodgates for motivated programmers to follow the lead and experiment with fine tuning existing models to solve the medical issues that currently lack the data to be solved.

REFERENCES

1. Ronneberger, Olaf, et al. “U-Net: Convolutional Networks for Biomedical Image Segmentation.” *ArXiv.org*, 18 May 2015, arxiv.org/abs/1505.04597.
2. Yoon, J., & Choi, M. (2022, December 19). Surgical scene segmentation in robotic gastrectomy. Kaggle. https://www.kaggle.com/datasets/yjh4374/sisvse-dataset/data.