Venous Bifurcation detection for Automated IV Insertion

Final Report

Polycarpos Yiorkadjis, Olisaemeka Mbanefo, Sandip Pal

December 2022

1 Introduction

The aim of this project is to create an ML model which will assist a surgical robot in detecting the optimal insertion point for IV's. Misaligned IV insertion can cause a host of issues, such as infiltration, extravasation, etc. These issues can lead to tissue and nerve damage and other avoidable problems. Our main motivation for this project is to reduce the risk involved in IV insertion by speeding up the process of detecting the bifurcation points.

Our model takes in a greyscale image captured with a Near Infrared (NIR) stereo camera and LED Grid that produces uniquely accurate, high-resolution, high-contrast, and 3D vein point-clouds as input, we then use a Convolutional Neural Network (YOLOV3) to output the same image with an overlay of bounding boxes centered around the detected bifurcation points along with the (x,y) coordinates of the box.

2 Related Work

There have been some very different approaches to problems similar to this. Most of these were attempting to segment vessels and detect bifurcations in angiographic volumes such as the human brain. Some aimed to collect geometric features of the bifurcations by utilizing dimensionality reduction [3]. DeepVesselNet [2] is a popular network that has been used for this purpose, however most of the data used was artificially generated and focused on cerebral bifurcations which is quite different from our approach of using images of real human hands.

3 Dataset

For this problem, we had 2 datasets. One original dataset of 69 images taken with the original NIR camera and another with 498 images taken with a different version of the NIR camera. Our best results were achieved on the original 69 images as the new 498 images were extremely dark.

Our dataset consists of 142 training images, 20 test images post data augumentation. Some preprocessing steps applied to our images were, increasing the brightness by a factor of 2.5 due to the darkness of some of the images, it was difficult to identify bifurcation points for labeling. After increasing the brightness of the image we applied some augmentation including horizontal flipping, bounding box rotation between -15 degrees to +15 degrees, brightness adjustment between -20% to +20%.

Our dataset was obtained manually by PhD student Reuben D. Brewer by taking photos of random human hands with the NIR camera.

4 Method

To solve this problem we had initially proposed two object detection algorithms, YOLOV3 and Faster-RCNN. To make a final decision on the algorithm to move forward we leverage recent results on a pill identification problem from this paper [1]. From the study it was found that Faster-RCNN had the highest mean average precision (MAP) of 87.69% while YOLOV3 had a MAP of 80.17% however, YOLOV3 significantly outperformed Faster-RCNN in detection speed with frames per second (FPS) more than 8 times that of Faster-RCNN. From the above data, we decided on YOLOV3 for our framework.

The YOLO framework takes the entire image in a single instance and predicts bounding box coordinates and class probabilities. It takes in the input image and divides it into grids, image classification and localization are then applied to each grid. For this project we had just one class which was bifurcation point, therefore, our label y is a 6 dimensional vector:

Table 1: YOLO output vector

Vector	Vector Parameters
y	pc
	bx
	by
	bh
	bw
	c1

- pc represents the probability that an object is present in the grid
- bx, by, bh and bw specify the dimensions of the bounding box if an object is present..
- c1 represents the class, so if a bifurcation point is present, c1 will be 1 else 0.

For making predictions, there are two major methods that YOLO uses, intersection over union (IOU) and non-max suppression (NMS). IOU helps us to decide if the predicted outcome is a good one or a bad one using the formula below:

$$IOU = \frac{Areaofintersection}{Areaofunion}$$

Object detection algorithms do not just detect an object once. They may detect the same object multiple times in a single image, this is where NMS helps to concatenate the detections. It looks at the probabilities of all the predicted boxes and selects the box with the highest probability. Then it looks at the rest of the boxes in the image and the ones with high IOU will be suppressed. NMS will repeat this step until the boxes have either been selected or suppressed and we are left with our final bounding box. For training we used the original YOLOV3 loss function below:

$$\lambda_{coord} \sum_{i=0}^{S^2} \sum_{j=0}^{B} 1_{ij}^{obj} [(x_i - \hat{x_i})^2 + (y_i - \hat{y_i})^2]$$

$$+\lambda_{coord} \sum_{i=0}^{S^2} \sum_{j=0}^{B} 1_{ij}^{obj} [(\sqrt{w_i} - \sqrt{\hat{w_i}})^2 + [(\sqrt{h_i} - \sqrt{\hat{h_i}})^2]$$

$$+\sum_{i=0}^{S^2} \sum_{j=0}^{B} 1_{ij}^{obj} (C_i - \hat{C_i})^2 + \lambda_{noobj} \sum_{i=0}^{S^2} \sum_{j=0}^{B} 1_{ij}^{noobj} (C_i - \hat{C_i})^2$$

$$+\sum_{i=0}^{S^2} 1_{i}^{obj} \sum_{c \in classes} (p_i(c) - \hat{p_i}(c))^2$$

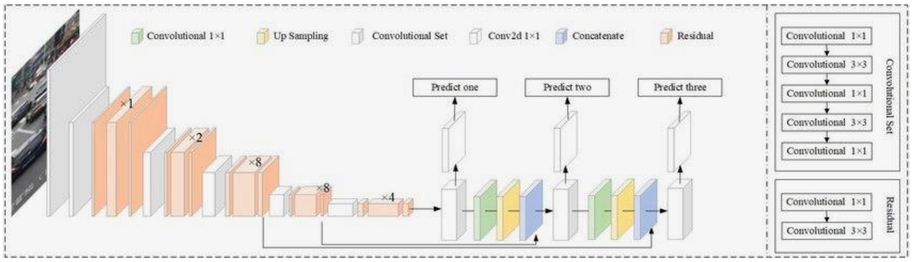


Figure 1: YOLO-v3 Architecture

5 Experiments

The initial experiment was run with just the 69 initial images from our dataset. The data set was augmented using horizontal flip, brightness and +-15 degree rotations. The new dataset was delivered to the project team around 11/10/2022 and was not labeled. The team used “Roboflow” to label the bifurcations for approximately 498 images.

5.1 Final work

Based on three experiments that the team performed A. using 69 images and no Yolo parameters changed **(mAP score 0.21)**, B. New dataset (69 + 498) and only Yolo final layer anchor change**(mAP score 0.24)** and C. New dataset (69 + 498) with high contrast pre-processing**(mAP score 0.12)** , we concluded that the new set of 498 images were not clear enough for the vein bifurcations to be properly captured by the neural network. Further investigations in to the data revealed that as compared to the first 69 images almost 50% of the 498 images had very faint vein edges.
We went back to the old dataset and ran a training on the dataset (69 initial clear images and a few clean images from the second dataset) and all all three Yolo anchor list changed to fit the bifurcation object

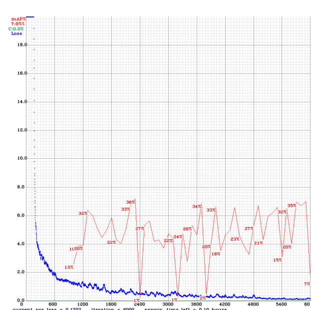
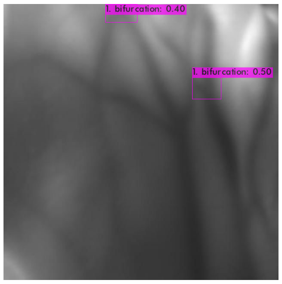
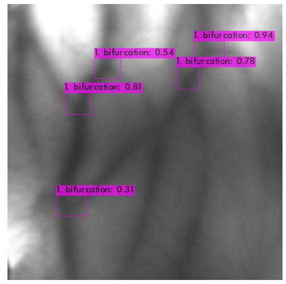


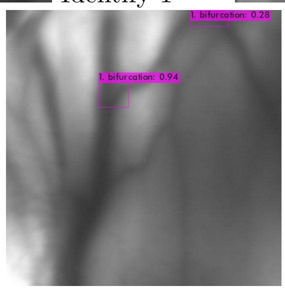
Figure 2: mAP and training error chart experiment 4



Identify 1



Identify 2



Identify 3

A best **mAP score of 0.36** was recorded. Below are a few identifications using the trained weights

6 Future Work

The overall conclusion of this research points to the fact that more data is required to obtain a good mAP score. We will further work on getting more data from the camera and increase the train and test dataset.
For further improvement in future work, a u-net segmentation algorithm could be added to the output of YOLO. This will further identify the optimal bifurcation point as not all bifurcations are made the same. Some may be too thin to insert an IV needle in, some may be oriented towards the fingers rather than the body which may slow down the flow of the medicine etc. There are multiple factors to consider to select the optimal insertion point.

7 Code

https://github.com/CS230BSP/VeinBifurcation

Jupyter Notebook : src/vein_detection.ipynb

Data folder : data/[sub directory for various dataset]

Config folder: cfg/[sub directory for various configurations]

Weights folder: weights/[sub directory for various training runs]

8 Contributions

Polycarpos Yiorkadjis:

- Hand labeling the dataset using bounding boxes.
- Curating hand vein images from internet to augment the data
- Running experiments using darknet to train a Yolo model
- Darknet configuration changes to enable data to be fed to the training binary.

Sandip Pal:

- Hand labeling a chunk of the internet images
- Setting up github source tree directories
- Running darknet predictions on test data
- Enabling appropriate batch sizes to run Yolo training using Nvidia cuda enabled gpu
- Contributed to writing code to augment data using numpy libraries.

Olisaemeka Mbanefo:

- Hand labeling the entire dataset from the data provider
- Using roboflow to automate data augmentation and text label generation
- Cleaning up the data for eventual consumption for training
- Yolo loss function research and feature tuning to increase accuracy
- Research on Faster-RCNN and U-NET Segmentation for future work
- Research on similar bifurcation projects

References

- [1] Tan, L., Huangfu , T., Wu , L., Chen , W. (n.d.). Comparison of yolo v3, faster R-CNN, and SSD for real time pill identification - research square. Retrieved December 10, 2022, from https://assets.researchsquare.com/files/rs-668895/v1$_{c}overed.pdf?c = 1631875157$.
- [2] Tetteh, Giles, et al. "DeepVesselNet: Vessel Segmentation, Centerline Prediction, and Bifurcation Detection in 3-D Angiographic Volumes." Frontiers in Neuroscience, U.S. National Library of Medicine, https://pubmed.ncbi.nlm.nih.gov/33363452/
- [3] Essadik, Ibtissam, et al. Automatic Classification of the Cerebral Vascular Bifurcations Using Dimensionality Reduction and Machine Learning. 29 Sept. 2022, https://www.sciencedirect.com/science/article/pii/S277252862200070X.