TISSUE DETECTION IN ULTRASOUND IMAGES USING REAL-TIME IMAGE SEGMENTATION MODEL

INDEPENDENT ENGINEERING PROJECT

Zain Shariff
Student
Curtis Junior High School
University Place, WA 98467
zainshariff6506@gmail.com

March 11, 2024

ABSTRACT

Ultrasounds are utilized in the medical field to visualize organ structures in the body. A physician viewing ultrasound scans has to identify the contours of the structures and differences in the brightness and texture of the tissue to further make interpretation of the image. YOLOv8 is a program that can be used in a real-time image segmentation model, which when trained, can trace contours of an object in an image. This study focuses on the development of a real-time segmentation model to identify tissue in ultrasound images. Ultrasound images were obtained and spliced into frames, which were manually annotated in Roboflow. Command Line Interface was used to train an extra large segmentation model with YOLO. Training was conducted in three sessions, and showed improvement of the model throughout training with lower loss, and increase in metrics. The best model was obtained at the 86th epoch. Its mAP50-95 rating was 0.85115 and F1 Score was measured at 0.91 at 0.521 confidence. The model was precisely able to identify tissue in the 60 test images assessment done manually. The model predicted images in 300 ms or less reliably. This study developed a model that can be utilized in real-time clinical practice such as the automated demarcation of tissue with three dimensional modeling scans, which was previously manually done. The application of knowledge in ultrasound scans of in-vivo organ structures would be helpful in clinical practice.

Keywords Ultrasound · Image Segmentation

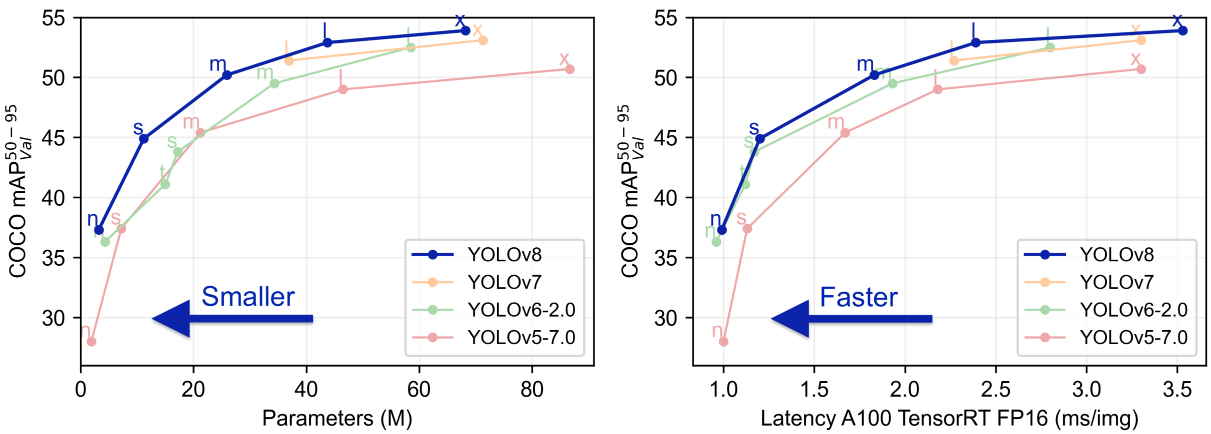
1 Introduction

1.1 Ultrasounds

Ultrasounds are sound waves with frequencies higher than the upper audible limit of human hearing (over 20KHz). Like normal "audible" sound, ultrasounds bounce off of structures and can be utilized to provide information about the structure they reflect off of. Piezoelectric crystals in the ultrasound probe emit waves and also receive information of the reflected waves which are translated electrical signals and represented as brightness on the images. The signals received have information about the distance travels and the ultrasound reflection/transmission of the tissue they interact with. Studying the images represented provides information on the tissue and identifying and significant structural abnormality. McMahon (2024).

1.2 YOLOv8

Ultralytics YOLOv8 (You Only Look Once) is a real time object detection and image segmentation model built on cutting edge advancements in deep learning and computer vision. YOLOv8 is suitable for a variety of applications and a variety of computer components. YOLOv8 offers instance segmentation, which goes beyond a simple bounding box image detection, able to display masks and contours that outline each object in an image. Jocher et al. (2023) Recently,



artificial intelligence is being increasingly used to help address tasks in a very efficient and accurate manner. Using YOLOv8, we can automate the process of identifying tissue present in ultrasound scans to assist physicians in clinical care. The goal of this study is to train and develop a machine learning model to reliably detect tissue in ultrasound scans. We also addressed if this developed protocol could be applicable in real time obtained images.

2 Materials and Methods

2.1 Materials

2 pieces of tissue (1 chicken muscular tissue and 1 lamb muscular tissue) were studied. A 3.5 MHz 2D Ultrasound was used to collect ultrasound images. A USB drive was used to move files from the Ultrasound Device to my computer system. For machine learning training, I used Ultralytics YOLOv8.1.0 through Command Line Interface. Roboflow was used to annotate the images.

2.2 Methods

A wet lab was set up, with a container filled with distilled water, and the experiment tissues submerged in the container. Ultrasound scans were obtained by placing the convex array ultrasound probe about 2 cm away from the material. Ultrasound scans of 60 frames per second in 3 seconds bursts, and 8 videos per piece of tissue was obtained. These scans were recorded on a USB drive and imported onto a computer system. 60 images were set to the side and compiled into a video. On the system I executed Python code to splice the videos selected into frames, which were stored in a folder. I proceeded to import these frames onto Roboflow for inspection and annotation. Manually, I inspected the images and removed images of low quality from the dataset. I annotated the tissue by tracing the tissue and labeling it. Images were manually demarcated for the borders. Images were refitted to 800x600 px. A correlogram and annotation collage describing my annotations can be found in **Figure 1**. Ultralytics YOLOv8.1.0 was installed onto the system. I exported the annotated images from Roboflow to the system. To start a session of training, I issued a Command Line Interface command to train a YOLO extra large segmentation model (dubbed yolov8x-seg.yaml) from scratch. At the end of the session, YOLO provided the best model from the whole session, which was trained again the next session. The best model was obtained, in which I made the model run predictions on the video consisting of 60 images, in which it ran image segmentation, and I can manually evaluate the performance of the model. A .csv file was an output summarizing the training including pivotal data such as box loss, segmentation loss, CLS and DFL loss as well as metrics on performance.

3 Results

3.1 Overview

Two pieces of tissue, lamb muscular tissue and chicken muscular tissue were analyzed in this study. A total of 1595 images were studied. It took three sessions to determine an optimal model, in total was 111 epochs. An epoch consisted of a training session on the model to utilize its performance with the images set for training and of a validation session

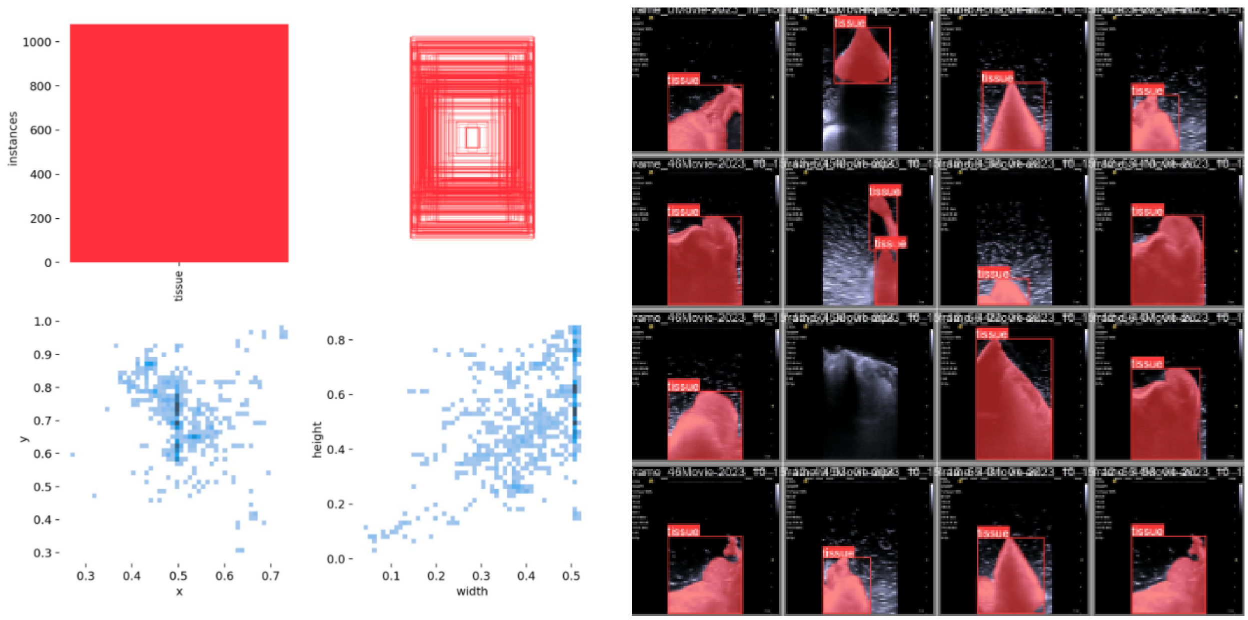


Figure 1: Left: Correlogram of annotation of the tissue in ultrasound images (A- amount of annotation of tissue in the ultrasound data set. B- bounding boxes of the annotation. C -x and y correlation of the annotation in the images. D - height and width of the bounding box annotation.). Right: a collage of image annotations.

to evaluate the model's performance with different images outside of the training set. In a YOLO prediction, an overall segmented subject, such as tissue in this study, is circumscribed by a square, also defined as a bounding box. Both the bounding box and segmented subject is utilized in training and validation, as well as predictions manually ordered by a user with a Command Line Interface command or a Python script. YOLO predictions were consistently finished in less than 300ms, and the best YOLO model was determined at the 86th epoch.

3.2 Confusion Chart

After every session, YOLO creates a confusion chart to show the amount of times the best model predicted an object and what it actually was. Generally, you want to see little to no mistakes in which the model predicted correctly, which was demonstrated in **Figure 2**.

3.3 Training

3.3.1 Loss Throughout Training

Throughout training, YOLO outputted the loss metrics of the best model in every epoch. There are four loss metrics YOLO measures: bounding box loss (box loss), segmentation loss (seg loss), classification loss (CLS loss), and Distribution Focal Loss (DFL). All across the board, when loss approaches zero, the model tends to be more accurate. Bounding Box Loss and Segmentation Loss measure the issues a model made with incorrectly picking out an object. Bounding Box Loss is determined by the area of the circumscribed "box" of the tissue, that the model incorrectly predicted to be the tissue or background. Segmentation Loss is defined as the segmented area that the model incorrectly predicted to be the tissue or background. CLS Loss determines overall incorrect predictions made by model. DFL works to help class imbalance and directly optimizes a distribution of bounding box borders. These loss metrics are demonstrated in **Figure 3** for regular training and **Figure 3** for validation. Overall, throughout training, the model got better, with bounding box, segmentation loss, CLS loss, and DFL decreasing until a plateau and an increase at around the 100th epoch.

3.3.2 Precision and Recall

Precision and Recall are two principal and inverse metrics that one should study when training an artificial intelligence model. Precision is determined by accuracy of the segmented mask. It requires it to be specific and precise. Recall is defined as the accuracy of the model to recognize and object present in an image. If your model has high recall

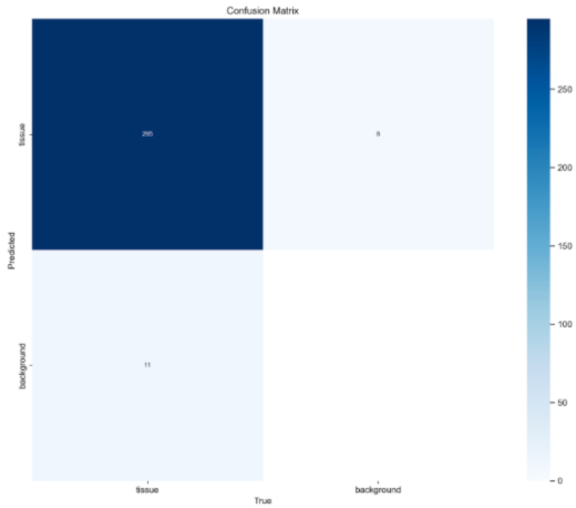


Figure 2: Confusion Matrix, displays the most common predictions of the model, generally you want the Predicted and True to match, displaying dark squares in the diagonal of the grid from top left to bottom right.

then the precision of the segmentation will be low, but if your model has high precision then it will be specific about conditions defining parts of an object and will most likely not recall the objects. This is illustrated in **Figure 4**. YOLO outputs performance metrics regarding Precision and Recall with Bounding Boxes and Segmentation Masks. All of these performance metrics were increasing, shown in **Figure 3**, determining an improvement in both traits.

3.3.3 mAP50 and mAP50-95

YOLO also outputs general performance metrics as an indicator of model prediction improvement. Generally, as the metrics approach 1, the model is effective. Two overall general and primary metrics of success of the model are mAP50, and mAP50-95, which are further divided into Bounding Box type and Mask type. Bounding Box type deals with the circumscribed bounding box, and Mask type deals with the actual segmented prediction. mAP50 integrates (area under curve) the standard Precision-Recall curve and compares it with various classes and backgrounds, with a scale from 0 to 1. mAP50-95 goes even further and measures with confidence thresholds involved, with a scale from 0 to 0.95. Overall as training continues, mAP50 and mAP50-95 values increased, showing an improvement in model prediction accuracy, shown in **Figure 3**. The model produced in the 86th epoch, which was the best overall, had an mAP50-95 score of 0.85115.

3.3.4 F1 Confidence

Confidence thresholds are vital in developing image segmentation models. A confidence threshold is an amount in confidence that the model rates an object that determines if it should consider it an object, or remove it as an incorrectly detected object. A high confidence means a high precision but a low recall. A low confidence means a low precision but a high recall. To balance and find an optimal, an F1 score and curve are used to find the optimal confidence threshold for the trained model. An F1 curve can help in indicating if precision and recall are balanced, when the F1 score is indicated with a confidence close to 0.5 The F1 score was 0.91 at 0.521 confidence.

3.3.5 Test Images

The 60 image video was used to evaluate the model's performance manually. After running a prediction with the best model, it was found that all 60 images had the model precisely mark out the tissues, as seen in **Figure 5**. All these results satisfy the criteria set for an optimal model to trace contours and identify tissue in ultrasound scans with precise image segmentation.

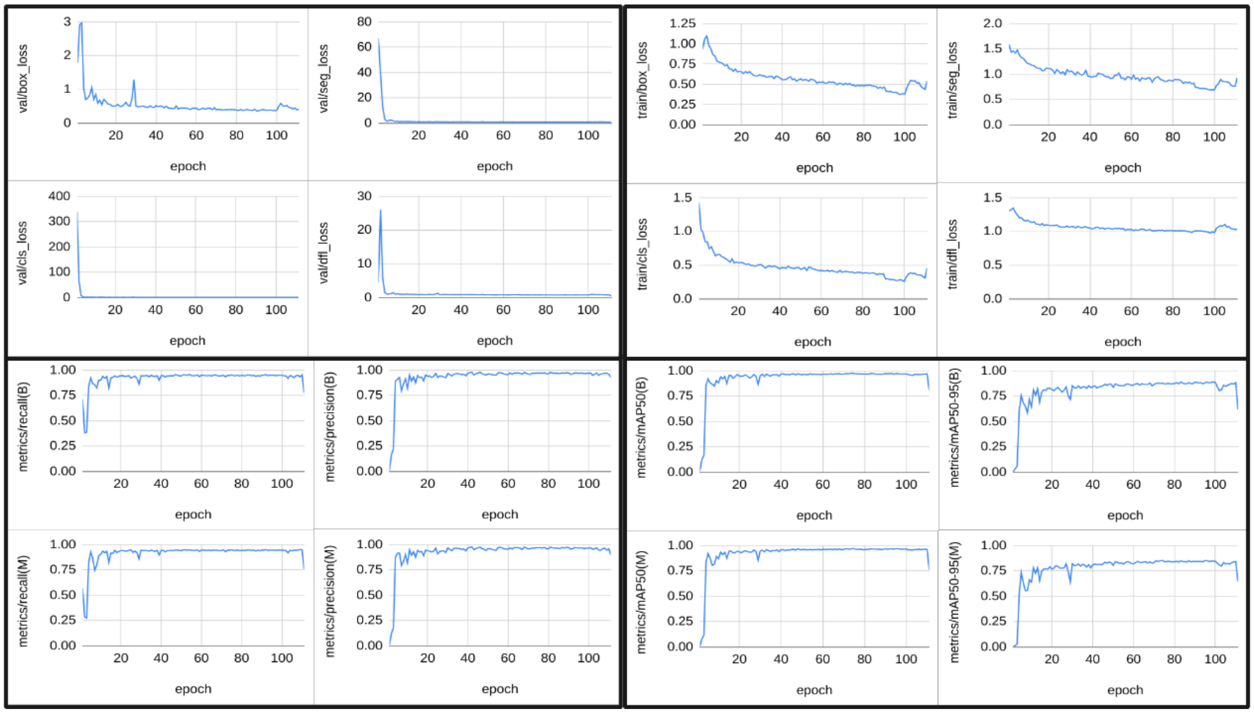


Figure 3: Training Data, Top Left: Validation Loss, best when close to 0. Top Right: Training Loss, best when close to 0. Bottom Left: Metric-based recall and precision, best when close to 1. Bottom Right: mAP50 on Left and mAP50-95 on Right. Model shows best performance when mAP50 approaches 1 and mAP50-95 approaches 0.95

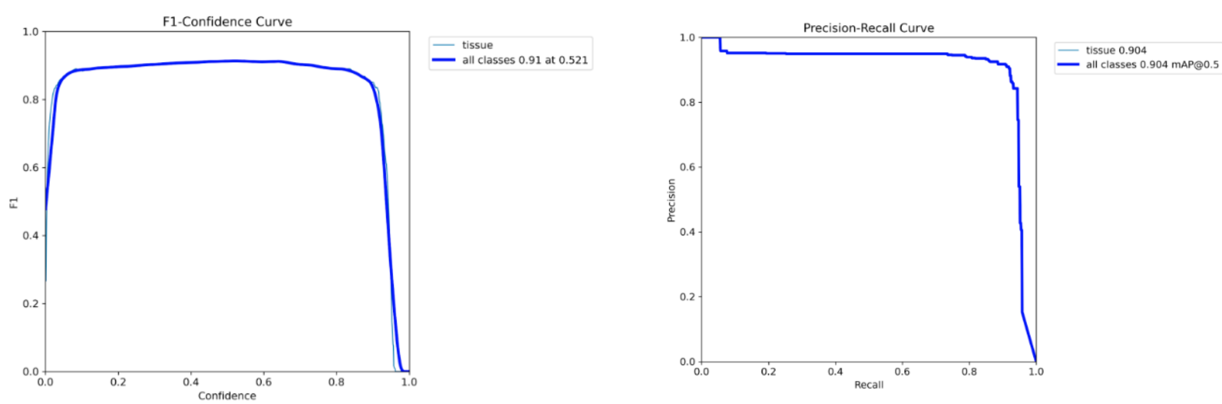


Figure 4: F1 Confidence Curve and Precision Recall Curve

4 Discussion

4.1 Other Related Research

Multiple two-dimensional slices at various imaging planes are obtained in the chamber of interest. The endocardial border from each image is traced either manually or using edge-detection software. These images are integrated by the CARTOSOUND software to build into an online 3D shell. Mapping information can then be applied to this anatomic reconstruction Enriquez et al. (2018). Microwave or radiofrequency ablations are critical in destroying abnormal electrical cells in the human heart that can cause life threatening issues such as bradycardias and tachycardias. A standard measurement or imaging technique to assist physicians in cardiac ablations does not exist. My last study found changes in ultrasound scans when exposing organic tissue to microwaves, especially where the surface of the tissue reflected more ultrasound waves the longer the duration of exposure to microwaves was and processed a monochromatic

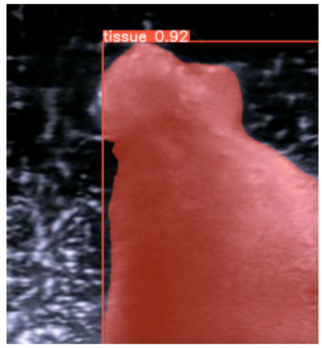


Figure 5: Example Prediction done by the best model

ultrasound imaging into spectral colors for easier identification of these patterns. In this study, I trained a machine learning program with YOLOv8 to run a segmentation algorithm on images. The best model obtained from the training can be used in real-time to display the contours of a detected tissue in ultrasound scans. The model is also saved and can be applicable in real world procedures regarding ultrasounds. A research article in 2011 developed a small-use deep learning model to perform segmentation on ultrasound scans of the left ventricle for possible three dimensional imaging and for diagnosis of certain conditions of patientsCarneiro et al. (2012). Another examination article devolved into training through supervised dictionary learning to diagnose certain heart conditions such as congenital heart disease using ultrasound scansEnriquez et al. (2018). A 2016 research created a neural network based artificial intelligence for classification of focal liver lesions using ultrasound scans Hwang et al. (2015).

4.2 Application of Study

The findings of this study, with extensive research in clinical practice, can aid physicians performing microwave or radiofrequency ablations as a real-time assessment to pinpoint the necessary amount of power and duration of ablation. Most complications from atrial fibrillations (cardiac ablation procedure) are due to inadvertent lesions or excessive exposure harming normal functioning cells

4.3 Future

The data from this scientific report is not meant to be used in clinical practice in anyway. The report focuses on animal tissue, though similar, is not identical to abnormal electrical cells in the human heart. It is known that abnormal electrical cells may possess different properties to animal tissue, normal cardiac electrical tissue, or even other abnormal electrical tissue. Furthermore, this is an ex-vivo study. During microwave ablation, tissue undergoes the inflammatory response and increases in mass. These properties may have a significant impact on the model's accuracy to detect tissue. A fairly small size sample was investigated, a larger sample size may provide a reliable result with little chances of error. Further randomized evaluations are required to test the model's accuracy. Further studies are required to investigate if the findings of this study apply in in-real-time clinical practice. Larger sample sizes are fundamental to support the reliability of the findings. The implementation of image machine learning models may be very useful to pinpointing the location of ablation lesions and extensively listing the state of the abnormal tissue.

4.4 Conclusion

This study was able to develop a model that identified tissue and its contours present in ultrasound scan diligently and accurately and displayed it's results for implementation in clinical use.

References

Carneiro, G., Nascimento, J. C., and Freitas, A. (2012). The Segmentation of the Left Ventricle of the Heart From Ultrasound Data Using Deep Learning Architectures and Derivative-Based Search Methods. *IEEE Transactions on Image Processing*, 21(3):968–982. Conference Name: IEEE Transactions on Image Processing.

Enriquez, A., Saenz, L. C., Rosso, R., Silvestry, F. E., Callans, D., Marchlinski, F. E., and Garcia, F. (2018). Use of Intracardiac Echocardiography in Interventional Cardiology. *Circulation*, 137(21):2278–2294. Publisher: American Heart Association.

Hwang, Y. N., Lee, J. H., Kim, G. Y., Jiang, Y. Y., and Kim, S. M. (2015). Classification of focal liver lesions on ultrasound images by extracting hybrid textural features and using an artificial neural network. *Bio-Medical Materials and Engineering*, 26(s1):S1599–S1611. Publisher: IOS Press.

Jocher, G., Chaurasia, A., and Qiu, J. (2023). Ultralytics YOLO. original-date: 2022-09-11T16:39:45Z.

McMahon, M. (2024). What is Cardiac Ultrasound? (with pictures).