Techniques for Segmentation of Liver Tumors Using CT Scans

William Kennedy-Moore, Yejin Kim, Abigail Mora Lemus

SF State

Abstract

There are many uses for machine learning in medicine, one of them being the analysis of medical imaging. Segmentation models can be useful in making quicker, more accurate assessments of tumor or unwanted growth location and size than a human can. For our project, we wanted to create a model that as accurate in segmenting tumors in the liver. We got our dataset from Kaggle, a data science competition platform, and downloaded it for preprocessing. The dataset contained CT scans of livers, livers with tumors, as well as masks of the target organ. To process it, we had to enhance the contrast of the CT scans in the dataset as well as alter their size and normalize the images. We decided that the best model to use would be the U-Net architecture. Our final hyperparameter values were a learning rate of 0.0011, batch size of 16, and 5 epochs. We chose to use the Adam optimizer and dice and jaccard loss as our performance metrics. The results from running the model and hyperparameter tuning were varied. Even when the results of the loss curve and dice and jaccard scores were good, we could not load the predictions made by the model. We also could not load the true mask images to compare predictions to. This could be because of issues during preprocessing. When trying to load the predicted masks to compare them to the true masks, not all of them loaded so we could not visually see how the model performed for all the images. This could be because not all the images had a corresponding mask, which could've thrown off the model in training.

Introduction

There can be many applications of machine learning with medical imaging, one being segmentation of tumors. Liver tumors are on the top 5 occurring cancer in women and men. More than half the people who get liver cancer dies from the cancer according to the CDC. Using a dataset from Kaggle of CT scans of livers, we used U-net to help with segmentation and make an efficient way to look for tumors in CT scans. The segmentation algorithms will help segment liver lesions easily in the contrast-enhanced abdominal CT scans since the liver tumors can have a diffusive shape and be homogeneous. The Kaggle data set is built from a variety of clinical sites around the world.
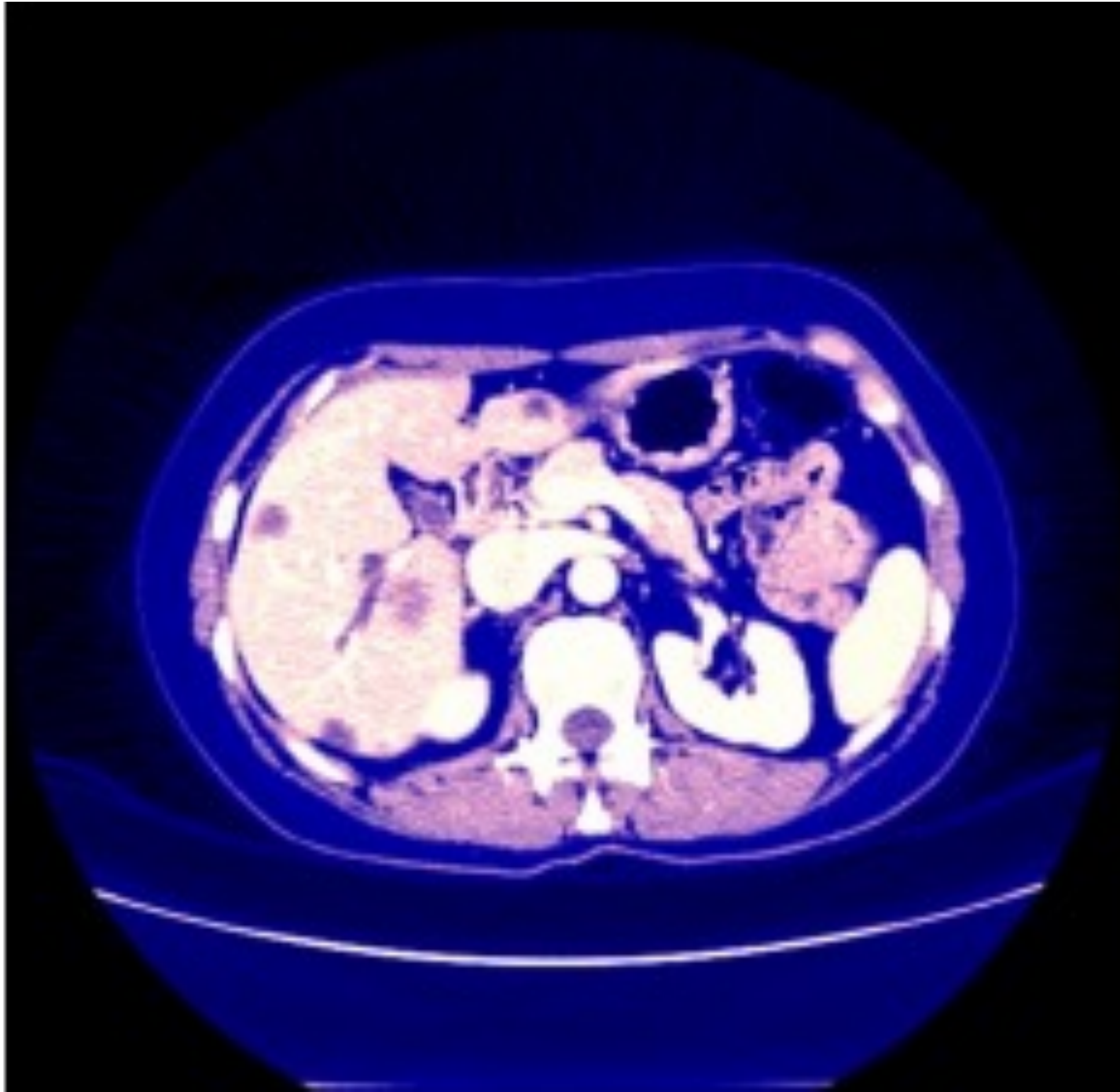

Fig. 1: Contrast-enhance CT scan of liver with tumors. This image shows an axial slice of the abdomen, showing the liver, which is the bright pink mass on the left, and the tumors within it which are seen as the darker spots within the liver. Contrast enhancement is a tool used to improve the quality of the images used in training and testing the model.

Methods

Data: Liver Tumor Segmentation
- This dataset was extracted from LiTS – Liver Tumor Segmentation Challenge (LiTS17) organized in conjunction with ISBI 2017 and MICCAI 2017.
- Features 130 CT images of the liver as well as masks of the liver and tumors.

Data pre-processing:
- Load image files into their own dataframe with labels.
- Enhance contrast or "window" the images.
- Find slices with tumors and slices without.
- Split into test, training, and validation sets.

Best Hyperparameter:
- Epochs: 5
- Learning rate: 0.001
- Smoothing factor: 100
- Batch size: 16
- Optimizer: Adam
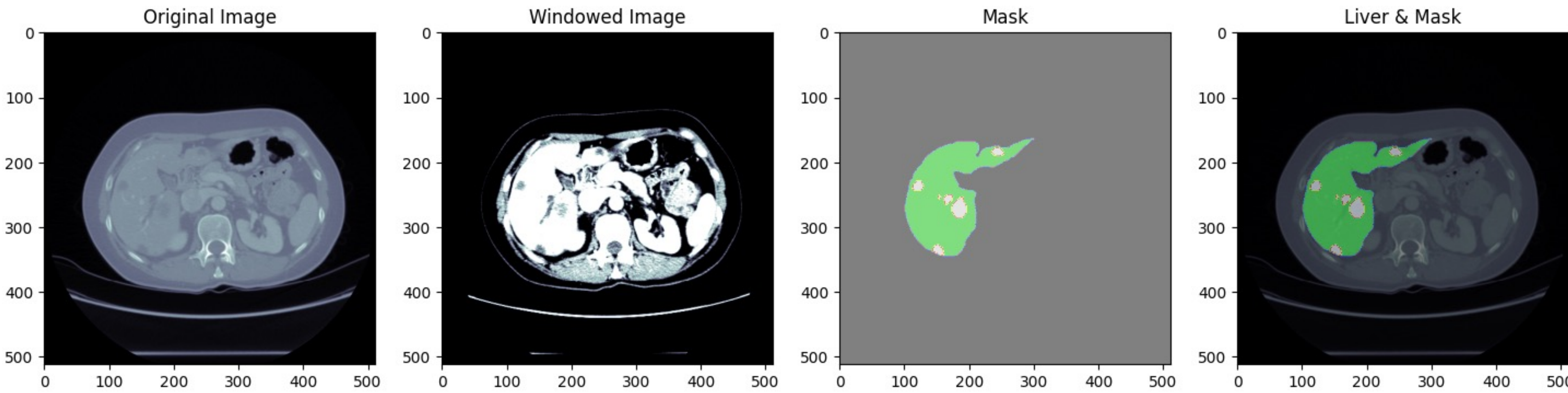- Input shape: (128, 128, 3)

Fig. 2: Comparison of original image from dataset with windowed image, mask of liver, and overlay of mask and original image. The original image from the dataset is hard to see which is why a windowed version of the image is required to run the model. This figure shows a comparison of the original image and the same image after it has been contrast enhanced, along with a mask of the liver (green) and liver tumors (grey). An overlay is included.
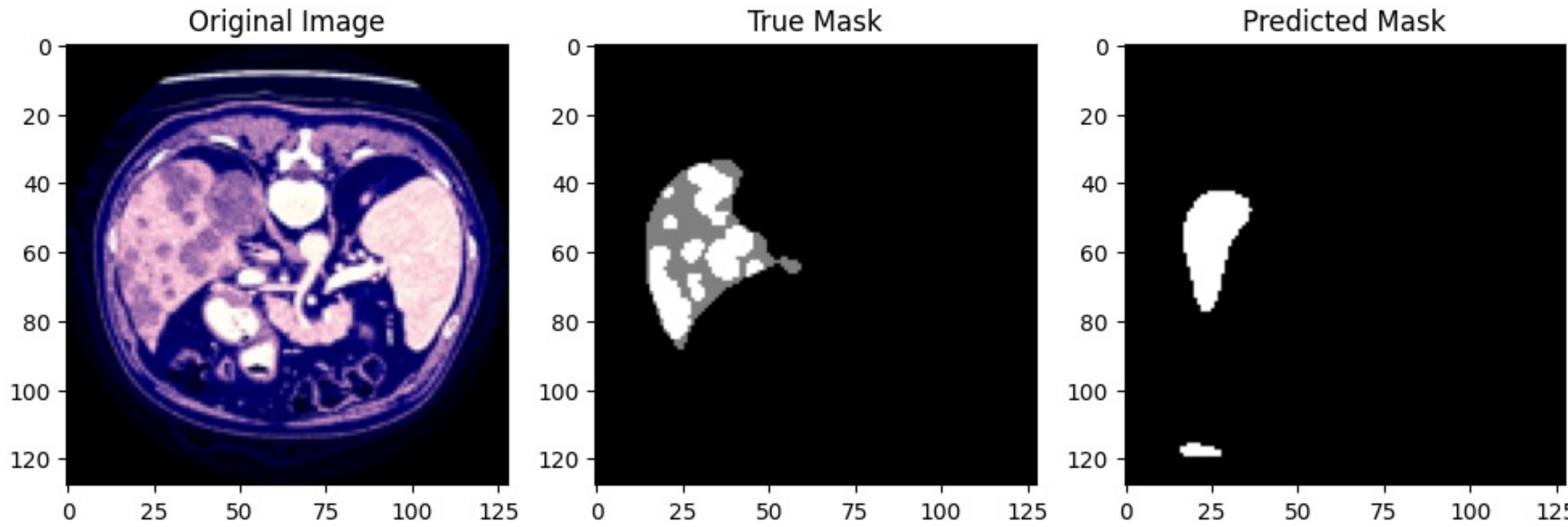
Results

Fig. 3: Input test image with the true mask and model predicted mask. One example of the input image used to test the model, its true mask, and the mask the model predicted.

- Average of all the predictions on test sets: 0.4595
- Average validation loss: 0.0804 | Average validation Jaccard coefficient: 0.4075 | Average validation Dice coefficient: 0.5773
- **Average test loss: 0.0803 | Average test Jaccard coefficient: 0.4150 | Average test Dice coefficient: 0.5831**
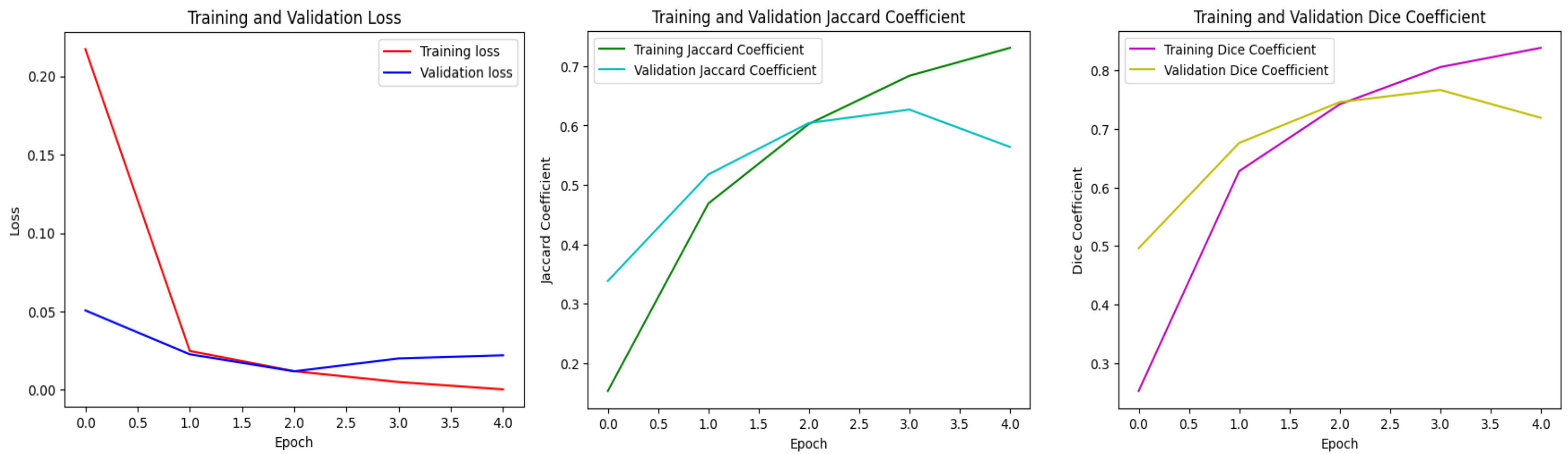

Fig. 4: Learning curve and Jaccard and Dice coefficient results. The jaccard and dice scores for our model with the best hyperparameter tuning. Ideal is closest to 1 as possible.

Conclusions

- The best hyperparameter combination used was a learning rate of 0.001, batch size of 16, optimizer of Adam, input shape of (128, 128, 3) and five epochs.
- Issues loading the true masks and the predicted masks, possibly because of an error in data split pre-processing or an issue with the set up of the dataset.
- The results only segmented the liver itself, and not tumors present in the liver.
- Not every image had a corresponding mask, which could've thrown off the model when it was making predictions.

Future Directions

- Since many of the test scores or metrics scores do not exceed 0.5, we cannot say that the performance is good. With more time we would further tune the hyperparameters of the model to achieve a more accurate model.
- Remove the images with no corresponding mask from the training dataset.

Acknowledgements