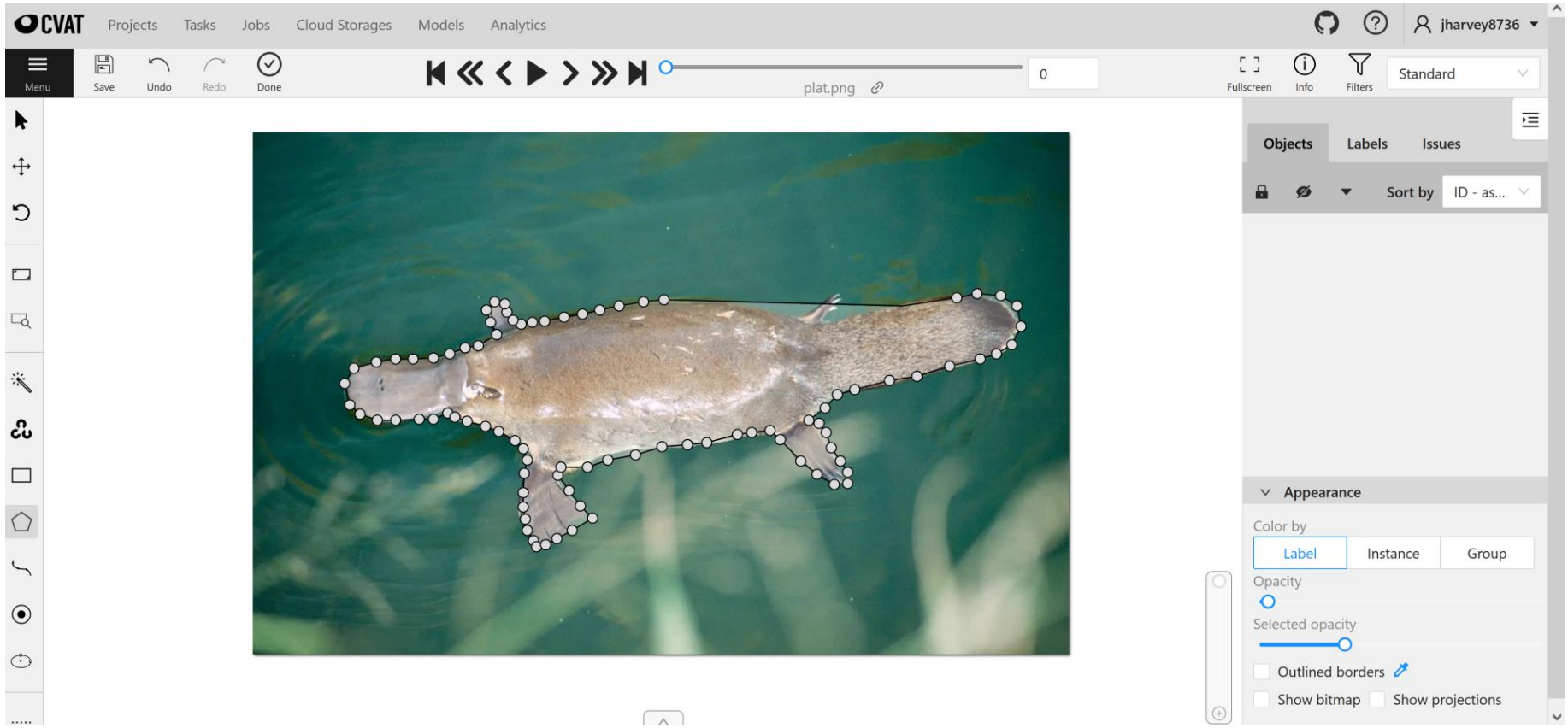
Image Segmentation for Platypuses in Nature

TEAM 18: JONATHAN HARVEY

CVAT: Computer Vision Annotation Tool



Original Image

Ground Truth Image

Problem

There is not enough data to train a good network from scratch.

Solution: Transfer Learning w/ Fine Tuning

Use pre-trained model weights instead of randomly initializing model weights

Train the model on the new data

Use a low learning rate to avoid overfitting since a lot of training has already been done

DeepLabv3 from Google

ResNet-50

ResNet-101

MobileNet-V3

```
model = torch.hub.load('pytorch/vision:v0.10.0', 'deeplabv3_resnet50', pretrained=True)
```

Paper: "Rethinking Atrous Convolution for Semantic Image Segmentation"

Very useful for this domain because DeepLabv3 recognizes platypuses as birds with decent IOU

Model Comparison – Training Time

Using CPU only – 15 epochs, sample size 38, batch size 2

ResNet-101: 132m 6s

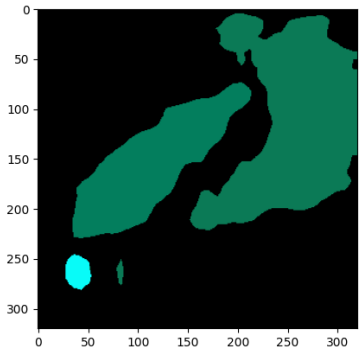
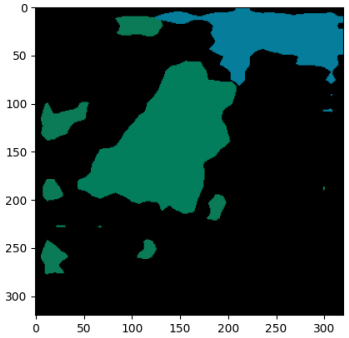
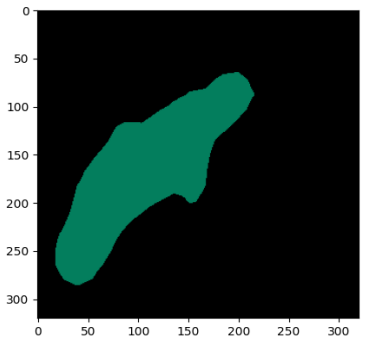
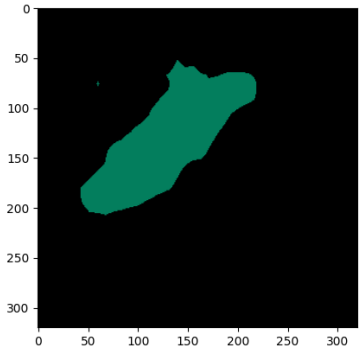
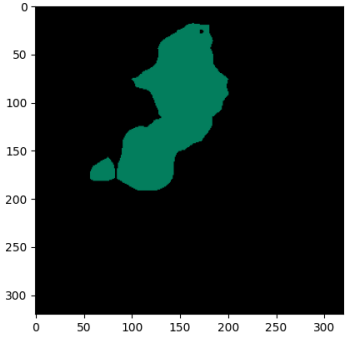
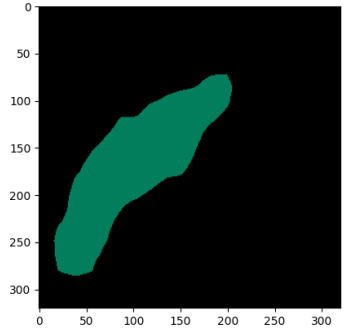
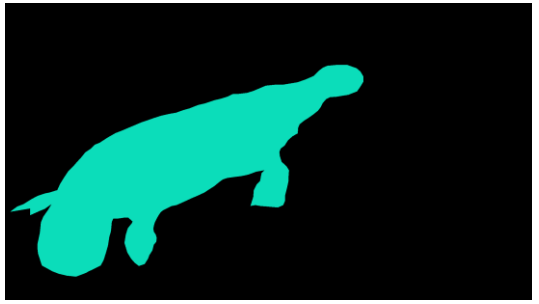
ResNet-50: 110m 18s

MobileNet-V3: 71m 52s

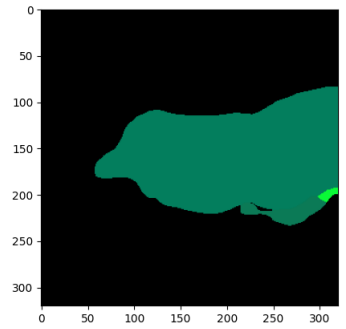
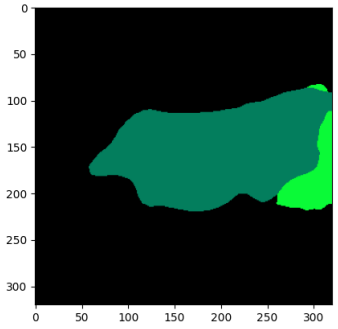
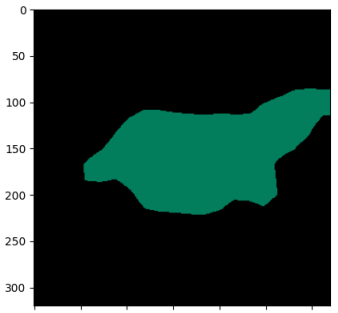
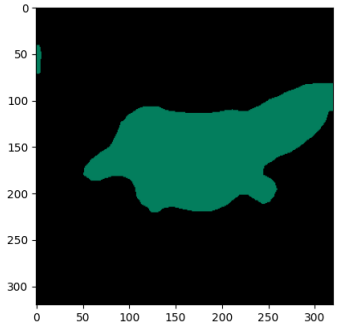
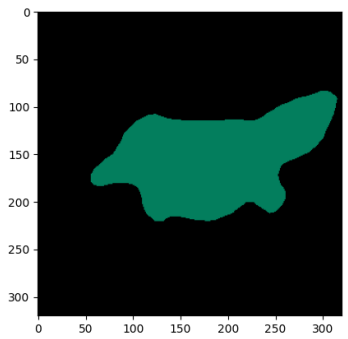
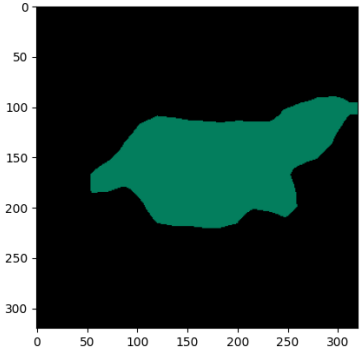
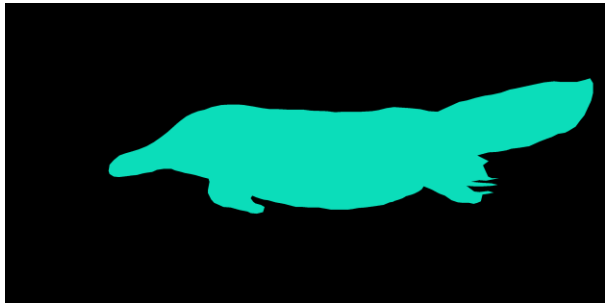
Model Comparison – Loss and IOU

ResNet-101	Loss: 0.2893	IOU: 0.8391
ResNet-50	Loss: 0.2921	IOU: 0.8206
MobileNet-V3	Loss: 0.2200	IOU: 0.8148

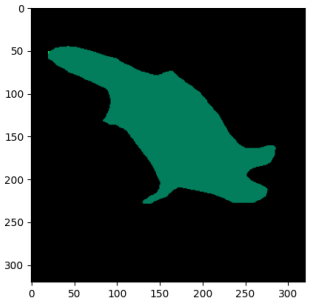
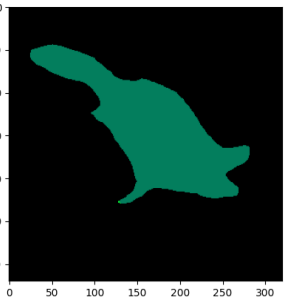
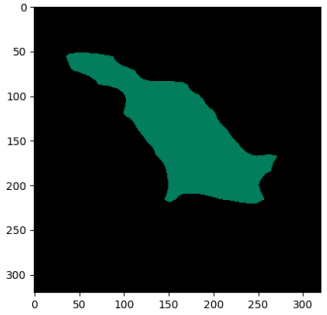
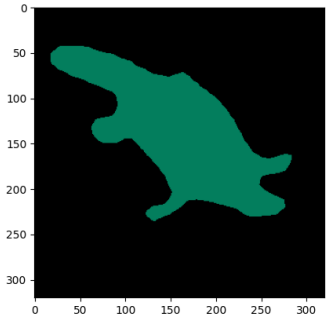
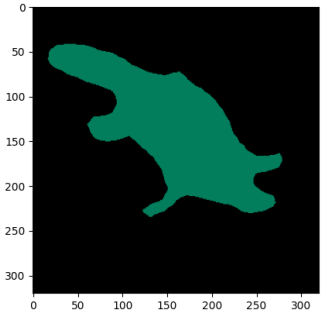
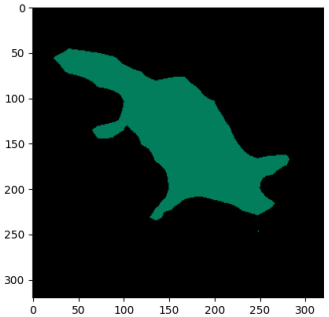
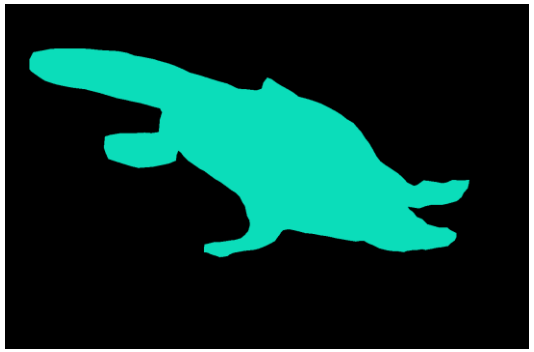
Model Comparison – Selected Images

	ResNet-101	ResNet-51	MobileNet-V3	Original/Ground Truth
Pre-Trained Model				
After Fine Tuning				

Model Comparison – Selected Images

	ResNet-101	ResNet-51	MobileNet-V3	Original/Ground Truth
Pre-Trained Model				
After Fine Tuning				

Model Comparison – Selected Images

	ResNet-101	ResNet-51	MobileNet-V3	Original/Ground Truth
Pre-Trained Model				
After Fine Tuning				

Potential Improvements

Augment the training data for a larger and more diverse sample

Increase sample size manually

Improve training time by using TPU (Tensor Processing Unit) in Google Colab

References

Chen, Liang-Chieh, George Papandreou, Florian Schroff, and Hartwig Adam. "Rethinking atrous convolution for semantic image segmentation." *arXiv preprint arXiv:1706.05587* (2017).

Inkawhich, N., 2022. *Finetuning Torchvision Models — PyTorch Tutorials 1.2.0 documentation*. [online] Pytorch.org. Available at: <https://pytorch.org/tutorials/beginner/finetuning_torchvision_models_tutorial.html>

Cvat.org. 2022. *Computer Vision Annotation Tool*. [online] Available at: <https://cvat.org/>