Image Segmentation for Platypuses in Nature

TEAM 18: JONATHAN HARVEY
CVAT: Computer Vision Annotation Tool
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

```python
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

<table>
<thead>
<tr>
<th>Model</th>
<th>Loss</th>
<th>IOU</th>
</tr>
</thead>
<tbody>
<tr>
<td>ResNet-101</td>
<td>0.2893</td>
<td>0.8391</td>
</tr>
<tr>
<td>ResNet-50</td>
<td>0.2921</td>
<td>0.8206</td>
</tr>
<tr>
<td>MobileNet-V3</td>
<td>0.2200</td>
<td>0.8148</td>
</tr>
</tbody>
</table>
Model Comparison – Selected Images

Pre-Trained Model

After Fine Tuning

ResNet-101

ResNet-51

MobileNet-V3

Original/Ground Truth
Model Comparison – Selected Images

Pre-Trained Model

After Fine Tuning

ResNet-101      ResNet-51      MobileNet-V3      Original/Ground Truth

Original/Ground Truth
## Model Comparison – Selected Images

<table>
<thead>
<tr>
<th>Model Type</th>
<th>Pre-Trained Model</th>
<th>After Fine Tuning</th>
<th>Original/Ground Truth</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>ResNet-101</strong></td>
<td><img src="image1" alt="ResNet-101 Pre-Trained" /></td>
<td><img src="image2" alt="ResNet-101 After Fine Tuning" /></td>
<td><img src="image3" alt="ResNet-101 Original" /></td>
</tr>
<tr>
<td><strong>ResNet-51</strong></td>
<td><img src="image4" alt="ResNet-51 Pre-Trained" /></td>
<td><img src="image5" alt="ResNet-51 After Fine Tuning" /></td>
<td><img src="image6" alt="ResNet-51 Original" /></td>
</tr>
<tr>
<td><strong>MobileNet-V3</strong></td>
<td><img src="image7" alt="MobileNet-V3 Pre-Trained" /></td>
<td><img src="image8" alt="MobileNet-V3 After Fine Tuning" /></td>
<td><img src="image9" alt="MobileNet-V3 Original" /></td>
</tr>
</tbody>
</table>
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

