Homework 2 - Image Segmentation

Team Segmentation Fault

Members: Manuel Pedrozo, Tomás Perez Molina

Report

The training and validation split is performed in the “split_dataset.py” script. Here we randomly split the dataset, leaving 20% of the images as validation. It's only run once to prepare the dataset. In order to not lose the shuffle of the dataset, we make sure to set the same random seed for reproducibility.

We started our model based on the one seen in the laboratory classes, and experimented with different techniques to achieve further improvements.

Model
The laboratory model used VGG16 as the encoding phase which feeds into a decoder. This decoder has 5 groups of layers with the following pattern UpSampling2D -> Conv2D -> ReLU with the first convolutional layer containing 256 filters. Finally the last layer is a Conv2D used for prediction of the segmentation.

From this, we implemented skip connections, connecting the pooling layers of VGG to the output of the first 4 convolutional layers of the decoder. In order to be able to implement these connections we had to increase the number of initial filters to 512 so the skip outputs could be added correctly.

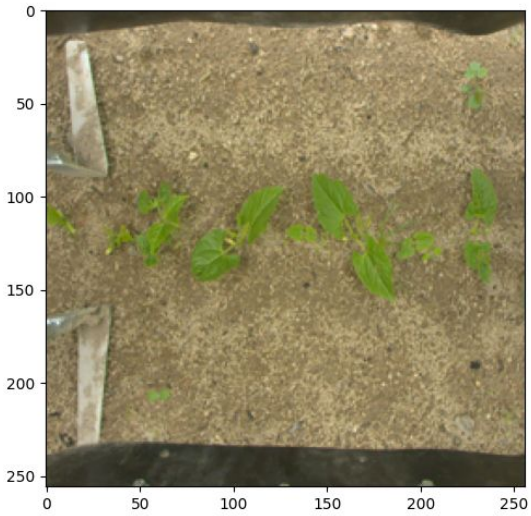
Replacing the UpSampling2D -> Conv2D sections with Conv2DTranspose also helped predictions as it allows the upsampling weights to be trained.

Finally, what improved performance significantly was adding a BatchNormalization layer before the ReLU layer in the decoder, which allowed us to raise the learning rate to 1e-3 and batch size to 32.

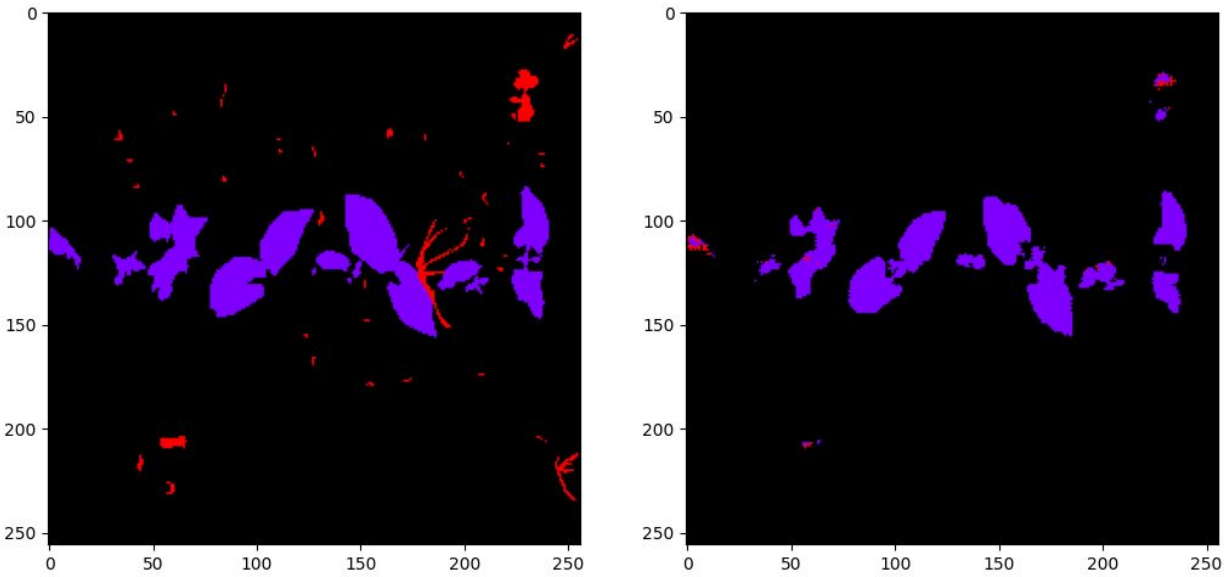
Classification
For the classification of the test dataset, the weights of the best checkpoint of the model in question are first loaded. Then, the images are classified one by one with dimensions 256 by 256 and the output mask is later rescaled back to the original resolution to generate the RLE.

Results
We focused on the Bipbip dataset as we implemented an iterative approach, and running the model for all datasets would take about an entire day in our local hardware.

The following is a prediction of a validation image.



Original image



Target mask (right) and predicted mask (left).

The same model that predicted this segmentation achieved **0.5092 IoU** for Bipbip Haricot and **0.7071 IoU** for the Bipbip Maize test dataset.