AN2DL HW 2 Angboolers Team

Arnab Mondal 10714338
Massimiliano Sica 10558133
Marco Labate 10491539

December 2020

1 Introduction

For this assignment we were requested to design a convolutional neural network for image segmentation. In particular, we were provided with a dataset containing images referring to two types of crops: haricot and maize. The objective is to segment the image using three classes, background, weeds and the actual crop. To solve this challenge we used a transfer learning approach for the encoding part and a custom decoder.

2 Development

A big part of the time was spent in formatting the data in a proper way. We were requested to upload a json file containing all the segmentation values from the test set. We started performing segmentation over only the Bipbip Haricot images, but soon realized that the data was insufficient to obtain a competitive score on the leaderboard. We then noticed that we could mix all the images together and we performed segmentation over the entire dataset provided, considering every team (Bipbip, Pead, Roseau, Weedelec) and every crop (Haricot, Maize). To create the dataset containing the training and validation images we referred to the CustomDataset function provided by Dr. Lattari and used 70 percent for training and the rest for validation. The other main issue with this challenge was picking the right input image size, since performing a consistent resizing would affect the quality of the images and worsen the segmentation accuracy. After going through [256,256], [512,512],[800,800] we ended up choosing [1024,1024]. We then imported VGG16 and used it as an encoder over which we built a decoder using upsampling and convolutional layers, closed by a convolutional layer with softmax activation. We also tried other architectures like Xception, Inception and ResNet50 but still got the best result with VGG16. In the decoder part we tried to modulate the number of filters as well and ended up with 512 which eventually decrease for every depth. For exporting the data we modified the functions provided in the starting kit.

After training and obtaining results of about 0.6 mean IoU on the validation set, we prepared the json file for submission but obtained very low results on CodaLab. For quite a while we were stuck, not understanding why our training seemed to guarantee decent results but our submission did not. We eventually managed to locate the problem in the size of our predictions; in fact, we were preparing our predictions with a size equal to the reduced one which we used for the training, but the results were evaluated over the full size of the pictures. Taking into account the fact that for every team the size is different, we prepared the predictions accordingly and finally got a result of about 0.45 mean IoU. Two examples of predictions obtained using our model are shown in Figure 1 and 2.

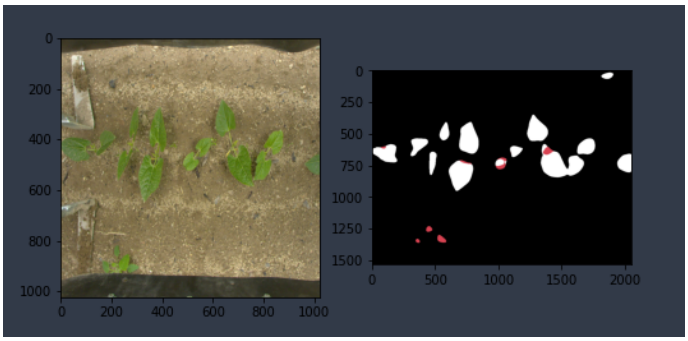

Figure 1: Predicted mask of a Bipbip image

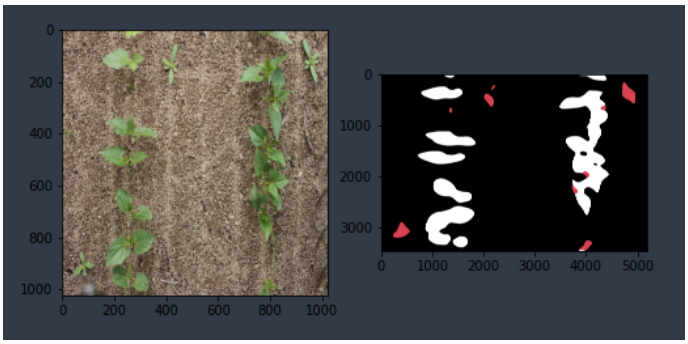

Figure 2: Predicted mask of a Weedelec image

Our best predictions were achieved for Bipbip and Weedelec teams, and for what concerns the different crops, overall we obtained very similar results for both Haricot and Maize.