Using transfer learning of a pretrained ResNet50 model and fine tuning for weed classification in Australia

U3220788 – Adrian Ricardo Rodriguez Durand

Bruce, Canberra

***Abstract*—This report summarizes our findings on the usage of Deep Learning algorithms for image classification focused on weed species. The Deep_weeds dataset consists of 17,509 images of 8 different weed species, across Queensland, with highly variable backgrounds (including the soil, withered plants, foreign leaves, etc). The model used methods like transfer learning with ResNet50, pretrained with 'imagenet' dataset, which then fine-tuned its weights using the pictures from deep_weeds to improve the overall model accuracy. The trained model achieved a validation accuracy of 88.64% and a classification accuracy of 88.97% on data that has never seen. This high level of accuracy proves the capabilities of Convolutional Neural Networks for processing pictures in realistic environments.**

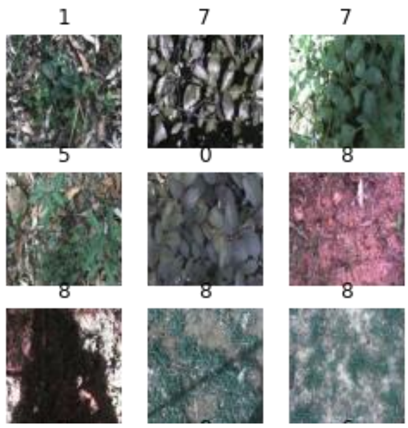
I. INTRODUCTION

Weeds are a big threat not only to native plants and crops as they compete for the resources available (water, minerals from the soil, sunlight, space), but also to the Australian economy. Farmers loose about $3.3 billion to natural weeds [1] and it is estimated that the Australian government spends $4 billion per annum to counter the negative effects of weeds [2]. The developed model aims to provide a quick way for gardening enthusiasts, lawn care businesses and farmers to precisely identify the type of weeds in a backyard or parcel. This will be achieved by training a Convolutional Neural Network (CNN) model, with multiple images of weeds and testing its prediction accuracy. This model is based on an already pretrained model (ResNet50) through Transfer Learning and later improved with fine tuning.

II. METHODOLOGY

Dataset analysis

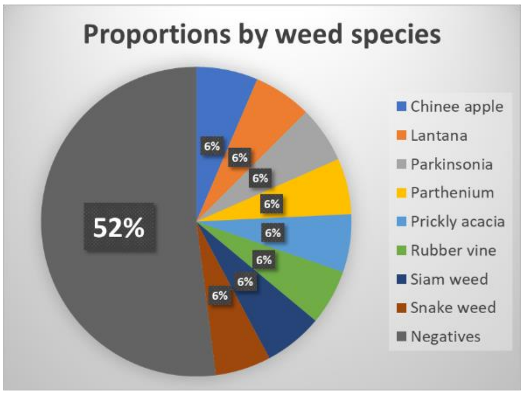
The dataset selected for the model is called "Deep_weeds" provided by Tensorflow Datasets and generated by [3]. The dataset consists of 17,509 images (256 x 256 pixels, RGB) on 8 different Australian weeds in 8 weed-infested locations in Queensland: "Black River", "Charters Towers", "Cluden", "Douglas", "Hervey Range", "Kelso", "McKinlay " and "Paluma". We opted for this dataset because the recorded weeds grow in Australia, our target geographic area. Additionally, it is due to the large number of images provided that will allow us to separate into "training", "testing" and "validation" datasets (with a considerable number of images in each one) necessary to train our network. Finally, most horticultural datasets are under controlled lab conditions or were taken with a solid-colour background; however, the images from Deep_weeds have a realistic and complex background (including the soil, foreign stems, withered plants, etc) which will train the network to differentiate the weeds from its environment.



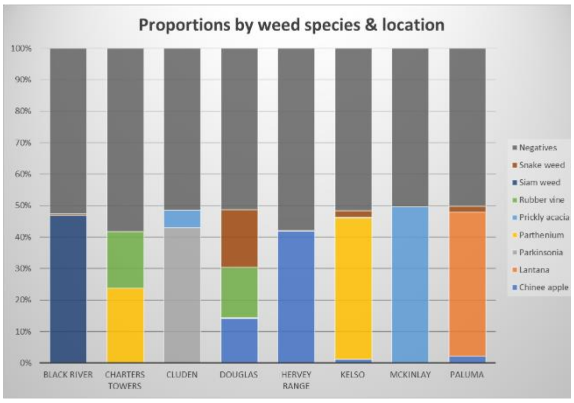
Data visualization

We can notice the disproportionate number of photographs under the “Negatives” label when considering the whole dataset. The creators of “Deep_weeds” explain that the photos are categorized under this label when none of the 8 targeted weed species is visible. Additionally, they explain that the proportion of photos with “positive” categorization and the "Negatives" was prioritized to be approximately 50-50 as it helps to prevent overfitting the model.

Label	Count	Proportion
Chinee apple	1125	0.0643
Lantana	1064	0.0608
Parkinsonia	1031	0.0589
Parthenium	1022	0.0584
Prickly acacia	1062	0.0607
Rubber vine	1009	0.0576
Siam weed	1074	0.0613
Snake weed	1016	0.0580
Negatives	9106	0.5201
Total	**17509**	**1**



Tables and charts showing the proportions between the photos of weeds species and the proportions out of the total dataset considering the “Negatives” label.



Bar chart showing the even proportions between the photos under the “Negatives” label and the pictures under a weed species label by location.

Transfer Learning and Fine Tuning

These are Deep Learning techniques with the objective of improving the accuracy of a model. Transfer learning consists of re-purposing a previously trained model to aid training a new model with a different task. In this case, the ResNet50 model, pretrained with ‘imagenet’ dataset, was chosen. It is important to clarify that the weight and biases of the pretrained model will not be trainable at first. To better adapt the ResNet50 model to classify weeds, we will perform Fine Tuning, which consists of retraining the weights and the biases of the pretrained model to the Deep_weeds dataset.

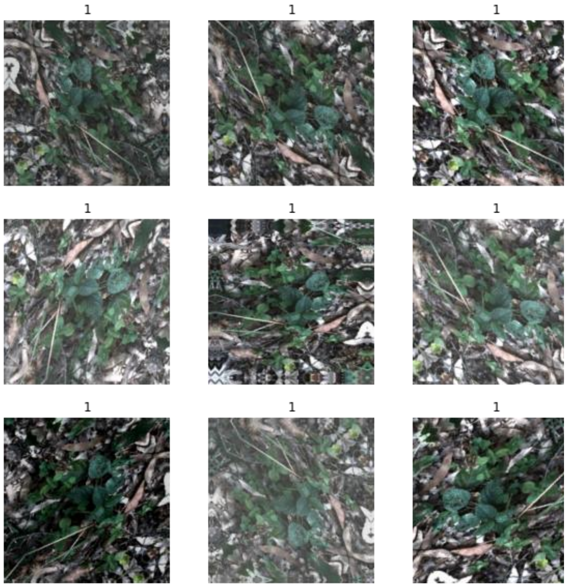
III. EXPERIMENT

In the first run of this project, the entire Deep_weeds dataset was used resulting in running out of the available RAM offered by Google Colab. Although it managed to compile and train the model, it was not possible to re-train the model for fine-tuning as it would require more computing power. Additionally, the model was overfitted due to few data augmentation layers, this would have only increased if we proceeded with fine-tuning.

In order to avoid using all the RAM memory again, the model only considered half of the Deep_weeds dataset in this second run (8,754 images). It was split following the proportion of (5253) 80% for training, (1751) 10% validation and (1750) 10% testing. K-folds cross validation was used for defining the images for training and validation sets being K = 4. Images in all three sets are grouped into batches of size 32.

Data preprocessing and augmentation

For data preprocessing: images are resized to 224 x 224 to match the input of the ResNet50 base model. Initially a rescaling layer was included (making all values go to [0 - 1] instead of [0 - 255]) but it is suspected that this resulted in low accuracy values. For data augmentation: images are randomly flipped horizontally and vertically, randomly rotated between -360° and 360°, the contrast is randomly altered by a factor of 0.3, and brightness is randomly modified by a factor of 0.15. Finally, images are randomly zoomed out by a factor of 0.1. These layers will only run during training of the model.



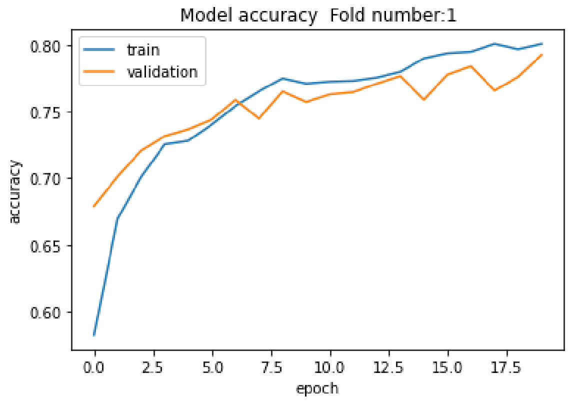
Model

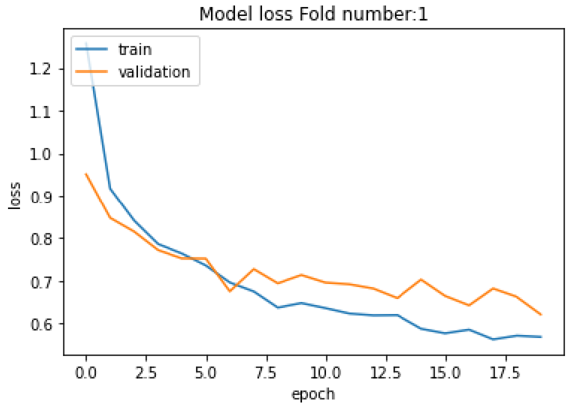
The model consists of the following layers. Two input layers, one for data preprocessing (resize) and one for data augmentation. Next is a “frozen” ResNet50 model pre-trained with Imagenet dataset, this means the weights are not trainable. Finally a flatten layer, a fully connected layer with 512 units and Relu activation function, a 0.5 dropout layer and a final fully connected layer with 9 units (number of labels) and Softmax activation function.

```
Model: "sequential_21"
_________________________________________________________________
 Layer (type)                Output Shape              Param #
=================================================================
 sequential (Sequential)     (None, 224, 224, 3)       0

 sequential_1 (Sequential)   (None, None, None, 3)     0

 resnet50 (Functional)       (None, 2048)              23587712

 flatten_19 (Flatten)        (None, 2048)              0

 dense_38 (Dense)            (None, 512)               1049088

 dropout_21 (Dropout)        (None, 512)               0

 dense_39 (Dense)            (None, 9)                 4617

=================================================================
Total params: 24,641,417
Trainable params: 1,053,705
Non-trainable params: 23,587,712
_________________________________________________________________
```

Best K fold

After running the model using all 4 folds, it was determined that the first fold gets the best values for validation accuracy and loss. The model accuracy graph shows positive trends for both datasets; meanwhile, the model loss graph shows a decreasing trend for both datasets meaning that data augmentation has effectively reduced the overfitting.






```
------------------------------------------------------------------------
Score per fold
------------------------------------------------------------------------
> Fold 1 - Loss: 0.6215155720710754 - Accuracy: 79.26899194717407%
------------------------------------------------------------------------
> Fold 2 - Loss: 0.6218249797821045 - Accuracy: 79.04055118560791%
------------------------------------------------------------------------
> Fold 3 - Loss: 0.6255238652229309 - Accuracy: 78.92632484436035%
------------------------------------------------------------------------
> Fold 4 - Loss: 0.6309394836425781 - Accuracy: 76.52769684791565%
------------------------------------------------------------------------
Average scores for all folds:
> Accuracy: 78.4408912062645 (+- 1.1114519105357152)
> Loss: 0.6249509751796722
------------------------------------------------------------------------
> Fold with best accuracy: 1
```

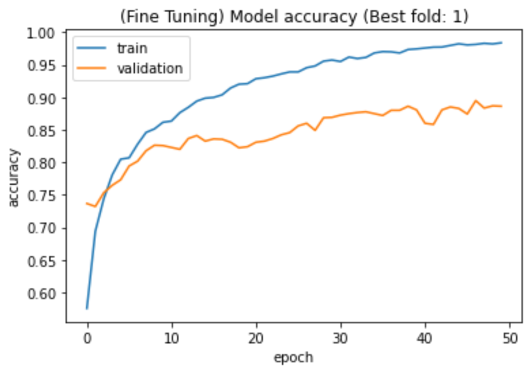
Fine Tuning

Now that the best k fold has been identified and the top layers of the model have been trained its weights and biases, it’s time to unfreeze the Resnet50 model so it can re-train and adapt its weights and biases for weed identification and drastically improve our model accuracy. Notice the steep increase of trainable parameters. When the ResNet50 model was frozen the entire model only had 1,053,705 trainable parameters. Now that it is unfrozen, there are 24,588,297 trainable parameters. This time the model was compile with a learning rate of 1e-5.

```
Model: "sequential_7"
_________________________________________________________________
 Layer (type)                Output Shape              Param #
=================================================================
 sequential_2 (Sequential)   (None, 224, 224, 3)       0

 sequential_3 (Sequential)   (None, None, None, 3)     0

 resnet50 (Functional)       (None, 2048)              23587712

 flatten_3 (Flatten)         (None, 2048)              0

 dense_6 (Dense)             (None, 512)               1049088

 dropout_3 (Dropout)         (None, 512)               0

 dense_7 (Dense)             (None, 9)                 4617

=================================================================
Total params: 24,641,417
Trainable params: 24,588,297
Non-trainable params: 53,120
```

IV. RESULTS

At the 50th epoch the model reached a training accuracy of 98.38%, a validation accuracy of 88.64% and a validation loss 0.6051. The plateaued trend of this last one might indicate that model overfitted. Perhaps this can be further reduced with more data augmentation layers, using the entire dataset or adding more dropout.



The model was presented with completely new pictures (test dataset) and it made predictions about their labels. Considering the realistic complex background of each picture, results show an incredible testing accuracy of 88.97%. In other words, out of 1750 images, 219 were wrongly categorized.

```
Accuracy: 0.8897142857142857
Precision: 0.8901274095033799
Recall: 0.8897142857142857
F1-score: 0.8864231410162803
```

Finally, after obtaining the best model and improving it through fine tuning, the model is saved under the name of ‘DeepWeedsResNet50FineTuning.h5’.

V. DISCUSSION

Ethics

Due to the high level of accuracy obtained by the deep learning model, people without any horticultural knowledge could detect, identify and buy herbicides, for the specific weed infestation, by themselves. It is understandable to believe that this could affect the gardening and lawn care businesses; however, the application of herbicides requires a professional in the field. To attack the species of weeds without damaging or killing the rest of the plants in the surrounding area, knowledge in handling chemicals and experience in applying the herbicide evenly are required. The deep learning model will not affect the work of the businesses, on the contrary, it will allow them to carry out their work more quickly and effectively by supporting them in the detection of weeds.

IV. CONCLUSIONS

Deep_weeds is a highly variable image dataset that is comprised of 17,509 pictures of 8 different weed species taken in 8 locations across Queensland. The developed model uses the methodology of transfer learning, using the ResNet50 model as a base. To further increase the accuracy for this specific application (weed classification), fine tuning was performed. Although the images for testing present a realistic and complex background (including the soil, foreign stems, withered plants, etc), the model achieved a classification performance of 88.97% in its predictions. This high level of accuracy proves the capabilities of Convolutional Neural Networks for processing pictures in realistic environments.

REFERENCES

[1] Chauhan, Bhagirath Singh. "Grand Challenges in Weed Management." *Frontiers in Agronomy*, vol. 1, 22 Jan. 2020, 10.3389/fagro.2019.00003.

[2] Sinden, J., Jones, R., Hester, S., Odom, D., Kalisch, C., James, R. and Cacho, O. (2004). The economic impact of weeds in Australia: Report to the CRC for Australian Weed Management. [online] Weedsaustralia. Adelaide: CRC for Australian Weed Management. Available at: https://weeds.org.au/wp-content/uploads/2020/02/tech_series_8.pdf [Accessed 30 Aug. 2022].

[3] Olsen, Alex, et al. "DeepWeeds: A Multiclass Weed Species Image Dataset for Deep Learning." *Scientific Reports*, vol. 9, no. 1, 14 Feb. 2019, www.nature.com/articles/s41598-018-38343-3, 10.1038/s41598-018-38343-3.