

# Some experiments in automated identification of Australian plants using convolutional neural networks

Tony Boston <sup>a</sup>  and Albert Van Dijk <sup>a</sup> 

<sup>a</sup>*Fenner School of Environment and Society, The Australian National University, Australian Capital Territory*  
Email: [tony.boston@anu.edu.au](mailto:tony.boston@anu.edu.au)

**Abstract:** Accurate plant identification is a skill that generally requires considerable knowledge and advanced training. However, plant identification is useful to a broad range of people within society, from conservationists and farmers to citizen scientists. Access to accurate, widely available knowledge about the identity and distribution of living species is critical for biodiversity conservation and sustainable development.

Automated plant identification has undergone major advances since 2012 with the application of convolutional neural networks (CNNs) from the emerging field of deep learning. This branch of machine learning has shown remarkable accuracy in image classification and visual object recognition when applied to still images through competitions such as the ImageNet Large Scale Visual Recognition Challenge.

This research project used transfer learning to fine-tune pre-trained deep learning CNNs originally developed for the ImageNet challenge, such as Inception and ResNet, which are publicly available through Tensorflow Hub. The models were applied to the automated identification of images of plants extracted from the Australian National Botanic Gardens Australian Plant Image Index and validated using additional images from the Atlas of Living Australia (ALA) and other Internet sources.

A comparison of model performance was undertaken using three different datasets: whole plant images (9,612 images of 392 species with at least 20 images per species), images of flowers (3,384 images of 271 species with at least 10 images per species) and scanning electron microscopy images of liverwort spores from *Fossombronia spp.* (322 images of 12 species with at least 10 images per species).

To decrease the risk of overfitting and extend the training dataset, data augmentation techniques such as scaling and reflection were tested to identify a high performing method, which also improved overall model performance. The best performing model for the All-plants (80.6% accuracy) and Flower datasets (88.4% accuracy) was Inception\_V3 pre-trained on the iNaturalist dataset of plants and animals. For the *Fossombronia spp.* dataset, the best performing model (81.2% accuracy) was ResNet\_V2-50 pre-trained on ImageNet 2012, using the 50-layer implementation of ResNet\_V2. The best performing flower identification model was also shown to have some proficiency in identifying the genus of an unknown species, where the genus but not species was represented in the dataset, with a Top-5 accuracy of 66%. The Flower dataset's best model performance was further tested using 1,000 images (20 images of 50 randomly selected species) downloaded from the Atlas of Living Australia and the Internet which produced a Top-1 accuracy of 85.9%.

Questions that remain to be addressed include further testing of data augmentation approaches and more comprehensive analysis to exclude overfitting. An interesting future extension of this study would be to train the best performing model on a larger dataset of Australian plant images, which could be used to aid scientists and the general public in identifying unknown species through image upload using an online website or phone app.

**Keywords:** *Convolutional neural networks, deep learning, plant identification, transfer learning*

## 1. INTRODUCTION

Plant identification is a skill requiring considerable knowledge, with an associated high barrier to entry for novices interested in gaining taxonomic understanding. Its accurate application is generally the preserve of expert taxonomists, botanists and plant ecologists. However, plant identification is useful for many diverse groups within society, such as farmers, conservationists, foresters, gardeners and citizen scientists (Wäldchen and Mäder, 2018). Bridging the taxonomic gap to build accurate, generally available knowledge of the identity and distribution of living species is critical for biodiversity conservation and sustainable development.

Automated plant identification has undergone major advances since 2012 with the application of convolutional neural networks (CNNs) from the emerging field of deep learning (LeCun et al., 2015). This branch of machine learning, originally from the computer vision discipline, has shown remarkable accuracy in image classification and visual object recognition when applied to still images, through competitions such as the ImageNet Large Scale Visual Recognition Challenge (ILSVRC)<sup>1</sup>. CNNs are designed to process data that comes in the form of multiple arrays and have been applied to processing of many types of 2D images (e.g. medical, biological, remote sensing), as well as 3D video, and 1D speech, audio and text data (LeCun et al., 2015). Use of deep learning CNN models to identify images of flowering plants from Australia was the main focus of this research project.

There are already several apps and associated web sites that support plant identification using deep learning, such as PlantSnap and PictureThis (commercial) and PlantNet, LeafSnap and iNaturalist (free). However, from limited testing, they do not currently support Australian native plants to any significant degree. The Atlas of Living Australia (ALA) has recently signed an agreement with iNaturalist<sup>2</sup> for the automated identification of user-submitted plant images, so this situation is rapidly changing (Andre Zerger, pers. comm., 2019).

Leaf shape is the most studied characteristic for plant identification (Wäldchen and Mäder, 2018). Ghazi et al. (2017) compared the efficacy of different convolutional neural network models in classification using training images of plants from LifeCLEF 2015. The PlantCLEF 2015<sup>3</sup> dataset, one of three components of LifeCLEF 2015, is composed of 113,205 images belonging to 41,794 observations of 1,000 species of trees, herbs and ferns living in Western European regions (Joly et al., 2015). These images are focussed on different plant morphological components such as flowers, fruits, leaves, branches and stems, as well as very detailed leaf scans and images of the entire plant. After data augmentation using the best performing models (GoogLeNet and VGGNet), Ghazi et al. (2017) reported that leaf scans produced the best results (98% accuracy) followed by images of flowers (87% accuracy) and fruit (79% accuracy), with images of the entire plant producing accuracies of 65%.

## 2. METHODOLOGY

The Atlas of Living Australia (ALA), part of the nation's research data infrastructure, compiles and makes available Australian biodiversity data from many sources and includes about 85 million occurrence records for over 121 million species<sup>4</sup>. One component of ALA is the Australian Plant Image Index (APII) sourced from the Australian National Botanic Gardens (ANBG). Using the ANBG interface to APII<sup>5</sup>, it is possible to easily extract subsets of these records including, for example, all images of plants in flower through a search for *Subject Part="Flowers"*.

A download from APII on 6 June 2019, excluding some images such as herbarium specimen records, Flora of Australia illustrations, and National Seed Bank images, yielded 29,563 digital images of plants where the dominant subject was flowers for 9,559 different species. In the download, there were 271 species with 10 images or more, which is the minimum suggested in the literature for training a deep learning model using transfer learning. These 3,385 images of 271 species were used to undertake an experiment in training CNNs to identify flowering plants. In addition to the Flower dataset, a download of all plant images from APII was also undertaken which yielded 78,477 records of 14,295 species, of which 392 had 20 or more images per species. Given the greater degree of difficulty in identification of entire plant images, as shown by Ghazi et al.

---

<sup>1</sup> ImageNet Large Scale Visual Recognition Challenge (ILSVRC). URL: <http://image-net.org/challenges/LSVRC/>. Accessed 24 July 2019.

<sup>2</sup> iNaturalist. URL: <https://www.inaturalist.org/>. Accessed 24 July 2019.

<sup>3</sup> LifeCLEF 2015 Plant task. URL: <https://www.imageclef.org/lifeclef/2015/plant>. Accessed 24 July 2019.

<sup>4</sup> Atlas of Living Australia. URL: <https://www.ala.org.au/>. Accessed 24 July 2019.

<sup>5</sup> Australian Plant Image Index, Export Metadata. URL: <http://anbg.gov.au/cgi-bin/apii-export-metadata>. Accessed 24 July 2019.

(2017), the cut-off for inclusion in this ‘all-plants’ dataset was set at 20 images to improve the chances of building a good model.

A third dataset of interest was identified within the downloaded records. These were scanning electron microscopy (SEM) images of spores from liverworts of *Fossombronia spp.* This dataset comprised 322 images of 12 species with 10 or more images per species. The number of species and images for the three datasets as well as splits between data augmentation, training, validation and test sets are shown in Table 1.

**Table 1.** Datasets used to train convolutional neural network models from Tensorflow Hub\*.

Datasets	All-plants	Flower	<i>Fossombronia spp.</i> SEM
No. species	392	271	12
No. images	9,612	3,385	322
Average no. images per species	24.5	12.5	26.8
No. data augmentation images	9,563	3,384	0
No. training samples	16,039 (7,407)	5,685 (2,494)	251
No. validation samples	1,568 (1,266)	542 (582)	39
No. test samples	1,568 (939)	542 (309)	32

\* (Values in parentheses) indicate the number of images before data augmentation.

TensorFlow Hub<sup>6</sup> is a library for reusable machine learning modules that can be used for purposes such as image, text or video classification. The ImageNet challenge was the original source of training data for the models available for image classification on TensorFlow Hub. This project used several modules from the TensorFlow Hub model store to identify species from the three datasets in Table 1, with results compared to identify the model with the best accuracy.

CNN image classification models generally have hundreds of thousands or millions of parameters (LeCun et al., 2015). Training them requires much labelled training data and extensive computing power (days or weeks and hundreds of GPU-hours computer time). Transfer learning is a technique that saves time and effort by taking the bulk of a model that has already been trained on a different set of images and re-using it in a new application (Donahue et al., 2013). Using transfer learning also greatly reduces the amount of labelled training data that is required to produce models with good accuracy.

This study used the feature extraction capabilities (LeCun et al., 2015) from powerful image classifiers pre-trained on ImageNet (Deng et al., 2009), to train a new classification layer on top using APII data. None of the APII species were in the original ILSVRC 2012 labelled classes on which the Tensorflow Hub models were trained. The 1,000 categories within the benchmark ILSVRC 2012 classification challenge validation set<sup>7</sup> traverse a very broad range of categories for images available on the Internet including, for example, miniature poodles, Siamese cats, catamarans, airliners and castles. The advantage of transfer learning is that only the top layer needs to be trained, which is a relatively rapid and computationally inexpensive process, and the overall model accuracy benefits from features discovered within lower layers trained to distinguish objects of a different type. Utilising transfer learning, a script first loaded the pre-trained module from Tensorflow Hub, then trained a new classifier on top for the flower photos and other datasets downloaded from APII.

TensorFlow Hub models were run on a server with 8 processor 4 core Intel(R) Xeon(R) W-2123 CPUs @ 3.60GHz running a single Titan XP GPU, typically taking an hour to several hours training time. Learning rates of 0.01 and 0.0001 respectively were used for the Gradient descent and Adam optimizers, with a training batch size of 100. Default input image sizes for the models tested ranged from 299x299 pixels for Inception\_V2 and Inception\_ResNet\_V2 models, to 224x224 pixels for Inception\_V1, Inception\_V2 and ResNet\_V2 models but this study used larger input image sizes as outlined below. All models in this study were trained for 10,000 iterations. The best performing models were trained again for 50,000 iterations to observe the effect on test accuracy.

As the plant and flower training datasets include, respectively, only 10+ or 20+ example images for each species, data augmentation was undertaken to increase the number of images available for training the models, with the aim of improving model accuracy. For the small *Fossombronia spp.* dataset, data augmentation was not undertaken as these SEM images are very detailed, clearly distinguishable and cropping, mirroring and other data augmentation techniques were considered to be unnecessary.

<sup>6</sup> TensorFlow Hub. URL: <https://tfhub.dev/>. Accessed 24 July 2019.

<sup>7</sup> ImageNet Large Scale Visual Recognition Challenge 2012 synsets. URL: <http://image-net.org/challenges/LSVRC/2012/browse-synsets>. Accessed 24 July 2019.

APII includes three publicly available images for each record of a species: a low-resolution thumbnail (generally 192 pixels in its longest dimension), a medium resolution image (generally 384 pixels in its longest dimension) and a higher resolution image (generally 750 pixels in its longest dimension). All models in this study were trained using copies of the medium resolution image. The accuracy of deep learning models can be significantly improved through data augmentation, to generate additional training data, which helps prevent overfitting and improves the model’s understanding of dataset features (Perez and Wang, 2017). Examples of data augmentation include random image cropping, rotation, mirroring, translation, elastic distortions, addition of ‘salt and pepper’ noise, random changes to image brightness, contrast, hue, saturation, and conversion to grayscale (Perez and Wang, 2017, Ghazi et al., 2017).

The best CNN models for flower identification were subjected to three additional checks to test the models against real-world data: 1,000 flower images downloaded from ALA and the Internet comprised of 20 examples of 50 randomly selected species from the Flower dataset; 50 images of taxa where the genus but not species was in the Flower dataset; and 50 images of unrelated species not found in the Flower dataset.

### 3. RESULTS AND DISCUSSION

Existing convolutional neural network models from the Tensorflow Hub were trained using data from the ANBG Australian Plant Image Index (APII). Eight models were tested against the three datasets with results shown in Table 2. The best performing model for the All-plants (80.6% accuracy) and Flower datasets (88.4% accuracy) was Inception\_V3 pre-trained on the iNaturalist dataset of plants and animals<sup>8</sup>. Inception\_V3 is a convolutional neural network architecture for image classification, originally published by Szegedy et al. (2015). The weights for this model were obtained through training described by Cui et al. (2018) on the iNaturalist (iNat) 2017 dataset described by Van Horn et al. (2017), after initial training on ILSVRC 2012<sup>7</sup>. The iNat2017 dataset consists of 579,184 training images and 95,986 validation images for 5,089 species obtained from the iNaturalist web site<sup>2</sup>. Images were collected by citizen scientists from around the world using different camera types, have variable image quality and the dataset features a large class imbalance (‘long tail’ distribution) of species images. It is perhaps not unexpected that this dataset, pre-trained to recognise features in hundreds of thousands of images of plants and animals, showed the highest test accuracy when re-trained using APII data. This result is consistent with published studies such as Cui et al. (2018) which observed better transfer learning performance when fine-tuning models from a more similar source, described in terms of their ‘domain similarity’.

**Table 2.** Test accuracies (%) for models obtained from Tensorflow Hub trained using three datasets\*.

Model	All-plants	Flower		<i>Fossombronia spp.</i> SEM
	Optimizer: Adam	Adam	Gradient descent	Gradient descent
Inception_V3	50.2 (36.3)	57.7 (52.1)	52.6 (48.2)	62.5
Inception_V3 iNaturalist	<b>[80.6] 79.8</b> (62.2)	<b>[88.4] 88.0</b> (80.3)	<b>[86.5] 86.2</b> (79.3)	59.4
Inception_ResNet_V2	44.4	55.5	45.8	65.6
Inception_V2	41.3	56.6	49.4	65.6
Inception_V1	42.3	55.9	51.1	75.0
ResNet_V2-50	47.9	55.4	54.1	<b>[81.2] 81.2</b>
ResNet_V2-101	50.0	55.9	52.8	68.8
ResNet_V2-152	46.2	54.4	53.0	59.4

\* (Values in parentheses) indicate results before data augmentation. [Values in brackets] indicate model test accuracies after 50,000 training iterations. All other values are model test accuracies after 10,000 training iterations. Best results highlighted in **bold**.

For the Flower dataset, two optimizers were tested: the Gradient descent optimizer and the Adam optimizer originally developed by Kingma and Ba (2014). The Adam optimizer offers several advantages to the gradient descent optimizer, the foremost being that it uses moving averages of parameters such as momentum in order to converge more quickly on the set of weights for the neural network. However, Adam has recently been criticized for not always converging on the optimal solution and results may not generalise as well as for gradient descent despite have better training performance (Wilson et al., 2017). For the *Fossombronia spp.* dataset, the best performing model (81.2% accuracy) was ResNet\_V2-50<sup>9</sup> pre-trained on ILSVRC 2012<sup>7</sup>. ResNet\_V2 is from a family of convolutional neural network architectures for image classification with a

<sup>8</sup> TensorFlow Hub, inaturalist/inception\_v3/feature\_vector. URL: [https://tfhub.dev/google/inaturalist/inception\\_v3/feature\\_vector/3](https://tfhub.dev/google/inaturalist/inception_v3/feature_vector/3). Accessed 24 July 2019.

<sup>9</sup> TensorFlow Hub, imagenet/resnet\_v2\_50/feature\_vector. URL: [https://tfhub.dev/google/imagenet/resnet\\_v2\\_50/feature\\_vector/3](https://tfhub.dev/google/imagenet/resnet_v2_50/feature_vector/3). Accessed 24 July 2019.

variable number of layers originally developed by He et al. (2016). This model incorporates a 50-layer implementation of ResNet\_V2.

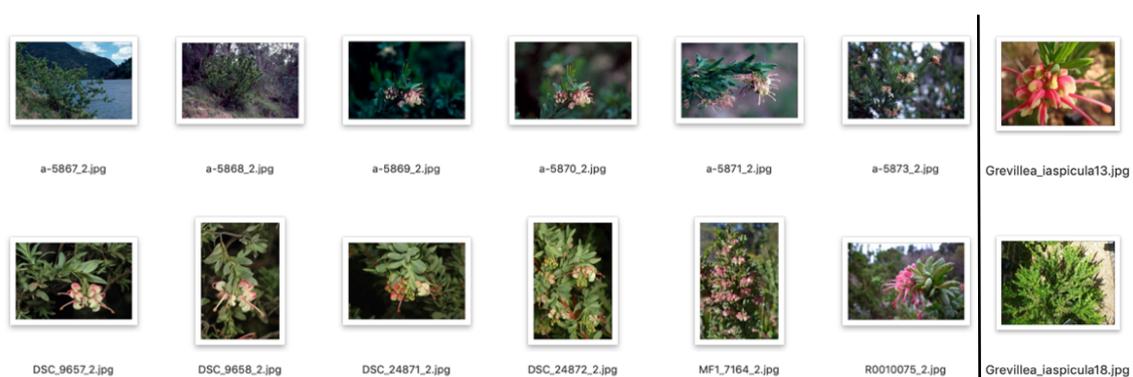
Because the images in APII are highly curated with, for example, approximately 81% of the images in the Flower dataset taken by one individual (Murray Fagg), the image quality is generally good and subject parts are typically clear, centred and captured with good lighting and colour balance. Based on data augmentation tests, a simple central crop of 300x300 pixels from the higher resolution image mirrored vertically was added to the medium resolution images for the flower and all plant datasets. This doubled the number of images available per species and improved flower and all plant model accuracies using the Adam optimizer by 9.6% and 28% respectively as shown in Table 2.

For the first additional test, 50 species of the 271 in the Flower dataset were selected at random. Twenty images for each species were downloaded from the ALA or other Internet sources and run against the best performing models. Care was taken to ensure the downloaded images were not from APII, showed flowers clearly, were from reputable sites where correct species identification was likely, and where image copyright allowed their use for research. Results for this test are in Table 3. Top-1 accuracy was 85.9%, slightly lower than obtained when training the model using the Adam optimizer (88.4%) and Gradient descent optimizer (86.5%), while Top-5 accuracy was 98.1%. It was expected that correct prediction scores (where the predicted species matches the actual species in the image) would be higher on average than incorrect predictions. This is borne out by the results shown in Table 3. Correct predictions have probabilities on average 84% greater for 1<sup>st</sup> prediction (N=859), 148% for 2<sup>nd</sup> prediction (N=76), 122% for 3<sup>rd</sup> prediction (N=30) and 120% for 4<sup>th</sup> prediction (N=10) showing, as expected, that the model produces higher probabilities when correctly identifying images. Interestingly, the best prediction results were obtained for a model trained using the Gradient descent optimizer rather than the Adam optimizer, though the differences were small: 1.6% for Top-1 accuracy and 1.1% for Top-5 accuracy.

**Table 3.** Test of flower identification model (Inception\_V3 iNaturalist 50K - Gradient descent optimizer) using 1,000 flower images (20 images of 50 species) downloaded from ALA and the Internet.

20 images of 50 species test	Prediction #1 correct	Prediction #2 correct	Prediction #3 correct	Prediction #4 correct	Prediction #5 correct	No match	Total
Top-1 accuracy: 85.9%	859	76	30	10	6	19	1000
Top-5 accuracy: 98.1%	85.9%	7.6%	3.0%	1.0%	0.6%	1.9%	100%
Av. prediction non-matches	0.384	0.087	0.0395	0.024	0.0167		
Av. prediction matches	0.706	0.215	0.0876	0.0528	0.027		
Difference	0.322	0.129	0.0481	0.0288	0.0103		
Difference (%)	83.9%	148.3%	121.8%	120.0%	61.7%		

According to Russakovsky et al. (2015), human top-5 classification error rate on the large-scale ImageNet dataset is estimated to be 5.1%. Top-5 error rates for Australian plant identification by experts are not available but are likely to be significantly lower than this, perhaps as low as 1%. Given a Top-5 error rate of 1.9% of the best performing model, this suggests the model may be able to perform with an accuracy close to an expert human observer. Examination of the results for classification of 1,000 images suggested the model will have problems classifying images with which a human observer would also have difficulties.



**Figure 1.** At left: images of *Grevillea iaspicula* from APII in the Flower dataset<sup>10</sup>. Test images at right: *Grevillea\_iaspicula13.jpg*: pred. #1 *Grevillea iaspicula* P=0.621; *Grevillea\_iaspicula18.jpg*: pred. #1

<sup>10</sup> APII top row images, rights owner: Director of National Parks; bottom row images, rights owner: M. Fagg.

Boston and Van Dijk, Some experiments in automated identification of Australian plants using convolutional neural networks

*Rhodanthe pygmaea* P=0.331, #2 *Comesperma ericinum* P=0.0568, #3 *Eragrostis curvula* P=0.0450, #4 *Lechenaultia biloba* P=0.0265 and #5 *Grevillea iaspicula* P=0.0219<sup>11</sup>.

For example, if the image is of the whole plant at a distance rather than individual flowers, the model will struggle to make a correct species identification, as demonstrated in Figure 1. The model identified the *Grevillea iaspicula*18.jpg image with its 5<sup>th</sup> prediction and a low probability (P=0.0219), probably because it was of the whole plant with flowers not clearly visible whereas the model correctly identified the closeup flower image *Grevillea iaspicula*13.jpg with its 1<sup>st</sup> prediction (P=0.621).

The second additional test aimed to check if the best performing model had some ability to predict the genus of an unknown species (where the genus is part of the Flower dataset). This may be the case if flowers of different species within a genus are visually more similar than those of other genera. 50 related taxa were selected at random and images of their flowers run through a new model built using Inception\_V3 iNaturalist for 10,000 iterations where images were grouped by the 197 distinct genera (=> class labels) in the Flower dataset. Results in Table 4 show that this model has some ability (Top-1 34% and Top-5 66% accuracy) in determining the genus for this small sample of 50 images of related species.

**Table 4.** Test of flower identification model based on genera (Inception\_V3 iNaturalist 10K - Gradient descent optimizer) using 50 images of taxa where the genus but not the species occurs in the Flower dataset.

50 related taxa test	Prediction #1 correct	Prediction #2 correct	Prediction #3 correct	Prediction #4 correct	Prediction #5 correct	No match	Total
Top-1 accuracy: 34%	17	8	5	2	1	17	50
Top-5 accuracy: 66%	34.0%	16.0%	10.0%	4.0%	2.0%	34%	100%
Av. prediction non-matches	0.256	0.0996	0.0573	0.0401	0.0295		
Av. prediction matches	0.571	0.1783	0.0664	0.0612	0.0318		
Difference	0.315	0.0787	0.0091	0.0211	0.0023		
Difference (%)	123%	79.0%	15.9%	52.6%	7.8%		

In Table 4 as in Table 3, it appears, from this small sample, that the average prediction scores for matches are higher than for non-matches: 123% greater for 1<sup>st</sup> prediction (N=17), 79% for 2<sup>nd</sup> prediction (N=8), 16% for the 3<sup>rd</sup> prediction (N=5) and 53% for 4<sup>th</sup> prediction (N=2).

The third additional test checked what prediction scores the model would have on average for 50 images of taxa not in the Flower dataset. Unsurprisingly, as shown in Table 5, the average prediction scores are low as the model is not able to recognise these species.

**Table 5.** Test of flower identification model (Inception\_V3 iNaturalist 50K - Gradient descent optimizer) using 50 images of taxa where neither the genus nor species are in the Flower dataset.

Images of 50 species not in Flower dataset test	Prediction #1	Prediction #2	Prediction #3	Prediction #4	Prediction #5	No match
Number of matches	0	0	0	0	0	50
Average prediction scores	0.262	0.108	0.0664	0.0445	0.0332	

For the Flower and All-plants datasets, test accuracies of models built using the Adam optimizer were marginally improved by 0.4% and 0.8% respectively after 50,000 iterations. For the *Fossombronia spp.* model, no improvement occurred when going from 10,000 to 50,000 iterations which suggests the model had already extracted all available information from the images after 10,000 iterations. To further test the sensitivity of the best performing model (Inception\_V3 iNaturalist – Adam optimizer) to the selection of training, validation and test images, 20 datasets were created with randomly selected images making up each subset. Results showed good agreement across all 20 datasets with, for example, average test accuracies for the augmented Flower dataset of 86.3% with a standard deviation of 1.1% for a model built for 10,000 iterations.

#### 4. CONCLUSIONS

This study applied transfer learning to identify plant species from the Australian Plant Image Index using convolutional neural networks. Results are encouraging, especially given the low average number of images per species, and suggest that models which have been pre-trained on related data from the same domain, such as iNaturalist, may generate good prediction results when applied to new species datasets. Results are comparable with published studies (e.g. Ghazi el al., 2017) using the PlantCLEF 2015 dataset. However, the risk of overfitting and a lack of generalisability of the models remains, despite encouraging test results for the Flower dataset using images obtained ‘in the wild’. The best countermeasure to overfitting is obviously training

<sup>11</sup> *Grevillea\_iaspicula*13.jpg rights owner: Janine Hunstone; *Grevillea\_iaspicula*18.jpg rights: Magnus Manske.

Boston and Van Dijk, Some experiments in automated identification of Australian plants using convolutional neural networks

using a greater number and range of images, while the application of more extensive data augmentation and tools such as K-fold cross-validation would also be of use.

Future areas of investigation could include expansion to a larger number of images for a greater range of Australian plants from sources such as ALA or Australian herbaria, exploration of other data augmentation techniques and more comprehensive testing to exclude overfitting. Development of an online version allowing upload and identification of images or an equivalent phone app would be an interesting additional development.

## ACKNOWLEDGMENTS

The primary author wishes to thank Anna Monro of ANBG for assistance in using the Australian Plant Image Index and Andre Zenger and Lee Belbin of the Atlas of Living Australia for suggestions on use of the ALA. During this study, a request was made to gain access to images from Canberra Nature Map<sup>12</sup> for use in testing developed models but this was not possible for copyright reasons and as machine learning-based plant identification is soon to be incorporated in this citizen science website (Michael Mulvaney, pers. comm., 2019).

All APII images in the three datasets are licenced with Creative Commons Attribution<sup>13</sup>. Image creators (number of images) for the All-plants dataset are: Fagg, M (6115), Director of National Parks (2234), Centre for National Biodiversity Research (424), Richardson, R.G. & F.J. (252), Purdie, R.W. (124), Shepherd, R.C.H. (94), Starr, F. & K. (62), Thiele, K.R. (58), Hadobas, P. (34), Jago, B. (24), Sankowsky, G. (21), Dowe, J.L. (18), Burbidge, N.T. (15), Gray, B. and Fuhrer, B.A. (14), Wilson, C.G. (13), Wilson, G.W. (10), Lepp, H. (6), Crisp, M.D., Wrigley, J.W. and Miller, J. (5), Breeden, S., Dewey, S., Fagg, L., Jordan, G. and Old, R. (4), Bodner, T., Lepp, H., Lyne, A.M., Phillips, M.E. and Thompson, H. (3), Brett, M., Curnow, J., Latz, P., Maple, D., Miller, J.H., Mitchell, A., Roche, C., Tenaglia, D., USDA APHIS PPQ Archive, Wright, M. and Wrigley, M. (2), and Brown, S.M., Godwin, M., Harte, M.E., Moorhead, D.J., Purdie, R., Samanek, J., Taylor, R. and Whibley, D. (1).

Image creators (number of images) for the Flower dataset are: Fagg, M (2741), Centre for National Biodiversity Research (287), Director of National Parks (216), Purdie, R.W. (37), Thiele, K.R. (27), Richardson, R.G. & F.J. (16), Jago, B. (14), Hadobas, P. (10), Gray, B. and Wilson, G.W. (6), Breeden, S. and Sankowsky, G. (4), Thompson, H. (3), Phillips, M.E. and Starr, F. & K. (2), and Bower, C., Brown, S.M., Dewey, S., Ford, A. & Goulter, F., Godwin, M., Jordan, G., Molloy, B., Purdie, R., Taylor, R. and USDA APHIS PPQ Archive (1). Image creators (number of images) for the *Fossombronina spp.* dataset are: Centre for National Biodiversity Research (317) and Cargill, C. (5).

## REFERENCES

- Cui, Y., Song, Y., Sun, C., Howard, A., Belongie, S., 2018. Large Scale Fine-Grained Categorization and Domain-Specific Transfer Learning. arXiv:1806.06193 [cs]. <https://arxiv.org/abs/1806.06193>
- Deng, J., Dong, W., Socher, R., Li, L.-J., Kai Li, Li Fei-Fei, 2009. ImageNet: A large-scale hierarchical image database, in: 2009 IEEE Conference on Computer Vision and Pattern Recognition. Presented at the 2009 IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops (CVPR Workshops), IEEE, Miami, FL, pp. 248–255. <https://doi.org/10.1109/CVPR.2009.5206848>
- Donahue, J., Jia, Y., Vinyals, O., Hoffman, J., Zhang, N., Tzeng, E., Darrell, T., 2013. DeCAF: A Deep Convolutional Activation Feature for Generic Visual Recognition. arXiv:1310.1531 [cs]. <http://arxiv.org/abs/1310.1531>
- Ghazi, M., Yanikoglu, B., Aptoula, E., 2017. Plant identification using deep neural networks via optimization of transfer learning parameters. Neurocomputing 235, 228–235. <https://doi.org/10.1016/j.neucom.2017.01.018>
- He, K., Zhang, X., Ren, S., Sun, J., 2016. Identity Mappings in Deep Residual Networks. arXiv:1603.05027 [cs]. <http://arxiv.org/abs/1603.05027>
- Kingma, D.P., Ba, J., 2014. Adam: A Method for Stochastic Optimization. arXiv:1412.6980 [cs]. <http://arxiv.org/abs/1412.6980>
- LeCun, Y., Bengio, Y., Hinton, G., 2015. Deep learning. Nature 521, 436–444. <https://doi.org/10.1038/nature14539>
- Joly, A., Goëau, H., Glotin, H., Spampinato, C., Bonnet, P., Vellinga, W.-P., Planqué, R., Rauber, A., Palazzo, S., Fisher, B., Müller, H., 2015. LifeCLEF 2015: Multimedia Life Species Identification Challenges, in: Mothe, J., Savoy, J., Kamps, J., Pinel-Sauvagnat, K., Jones, G., San Juan, E., Capellato, L., Ferro, N. (Eds.), Experimental IR Meets Multilinguality, Multimodality, and Interaction. Springer International Publishing, pp. 462–483. [https://doi.org/10.1007/978-3-319-44564-9\\_26](https://doi.org/10.1007/978-3-319-44564-9_26)
- Perez, L., Wang, J., 2017. The Effectiveness of Data Augmentation in Image Classification using Deep Learning. arXiv:1712.04621 [cs]. <http://arxiv.org/abs/1712.04621>
- Russakovsky, O., Deng, J., Su, H., Krause, J., Satheesh, S., Ma, S., Huang, Z., Karpathy, A., Khosla, A., Bernstein, M., Berg, A.C., Fei-Fei, L., 2015. ImageNet Large Scale Visual Recognition Challenge. International Journal of Computer Vision 115, 211–252. <https://doi.org/10.1007/s11263-015-0816-y>
- Szegedy, C., Vanhoucke, V., Ioffe, S., Shlens, J., Wojna, Z., 2015. Rethinking the Inception Architecture for Computer Vision. arXiv:1512.00567 [cs]. <https://arxiv.org/abs/1512.00567>
- Van Horn, G., Mac Aodha, O., Song, Y., Cui, Y., Sun, C., Shepard, A., Adam, H., Perona, P., Belongie, S., 2017. The iNaturalist Species Classification and Detection Dataset. arXiv:1707.06642 [cs]. <https://arxiv.org/abs/1707.06642>
- Wäldchen, J., Mäder, P., 2018. Plant Species Identification Using Computer Vision Techniques: A Systematic Literature Review. Arch Computat Methods Eng 25, 507–543. <https://doi.org/10.1007/s11831-016-9206-z>
- Wilson, A.C., Roelofs, R., Stern, M., Srebro, N., Recht, B., 2017. The Marginal Value of Adaptive Gradient Methods in Machine Learning. arXiv:1705.08292 [cs, stat]. <http://arxiv.org/abs/1705.08292>

<sup>12</sup> Canberra Nature Map. URL: <https://canberra.naturemapr.org/>. Accessed 24 July 2019.

<sup>13</sup> Creative Commons Attribution 3.0 Australia (CC BY 3.0 AU) URL: <https://creativecommons.org/licenses/by/3.0/au/>. Accessed 24 July 2019.