Capstone 2 Summary

Flower Classifier

Springboard Data Science Career Track

Kristopher Simino

Introduction

- Apply transfer learning by using a pretrained CNN
- Pretrained networks trained on 1000 various classes
- New data is 104 similar looking classes.

Data Acquisition and Wrangling

- Kaggle dataset
- Given test set not usable
- New test set taken from training set
- Remaining training set augmented

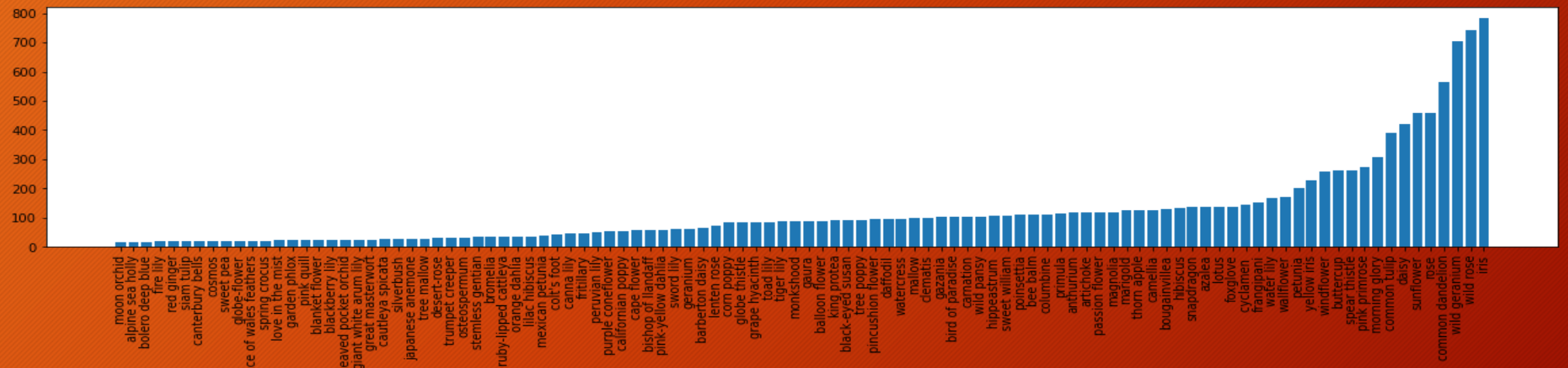
Set	Number
Training	114,912
Validation	3,712
Test	3,177

Dataset

- The images are mostly closeup in a natural setting
- To extend the training set, random images were zoomed, cropped, flipped, and had saturation, brightness and contrasted adjusted

Potential Data Issues

- Distribution highly skewed
- Top five species account for more than 25%
- Bottom five species account for less than 1%



Modeling

VGG

- 16 layers
- 308, 604 Trainable parameters

InceptionNet V3

- 42 layers
- 1,076,604 Trainable Parameters

ResNet50

- 50 layers
- 1,076,604 Trainable Parameters

Each trained for 30 epochs

Used same learning rate scheduler

Results

- InceptionNet outperformed the others
- ResNet far behind the other two

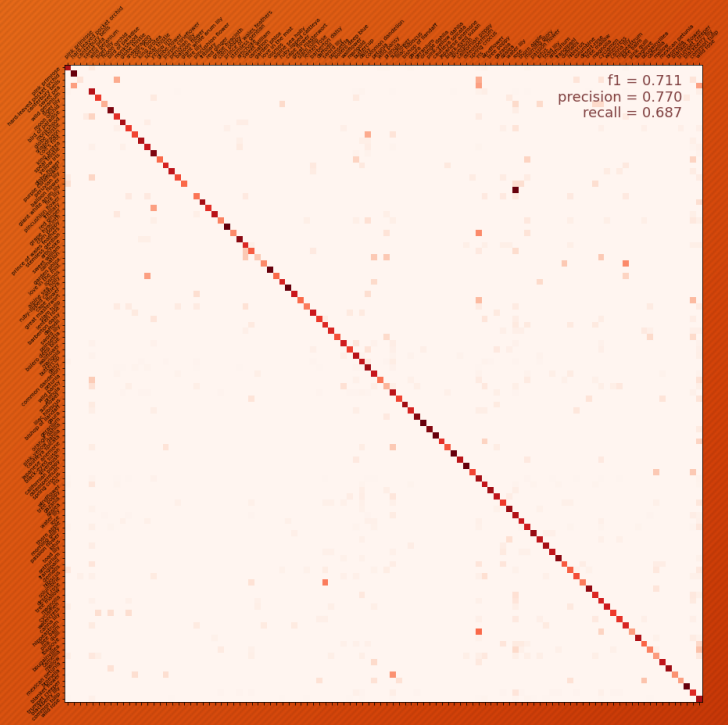
	Accuracy	Precision	Recall	F1 Score
InceptionNet V3	76.1%	77.0%	68.7%	71.1%
VGG	63.7%	62.1%	58.5%	58.7%
ResNet50	12.0%	4.0%	2.5%	1.8%

ResNet50

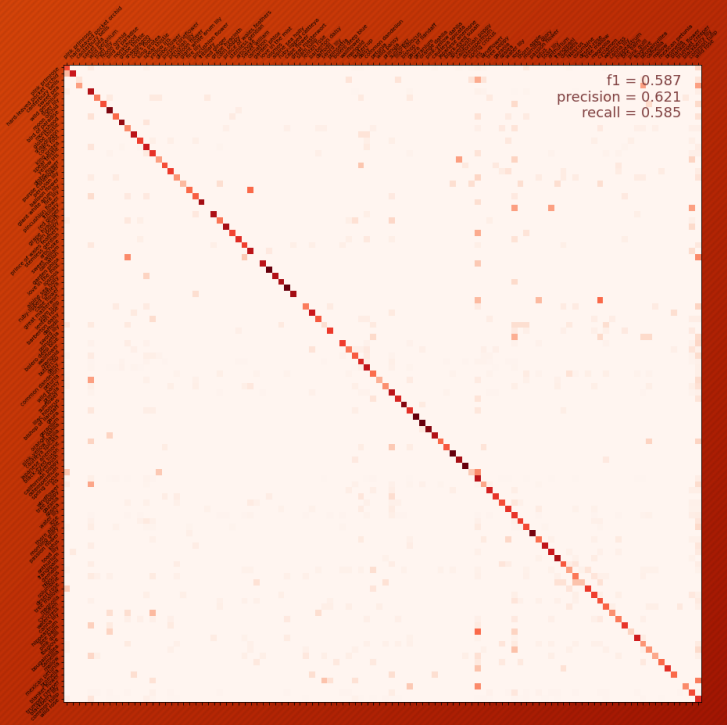
- ResNet had poor results because it mostly put the images into only five classes
- They were the top five largest classes

Similar Results

InceptionNet



VGG



Visual Inspection Example

Sweet Pea

Iris

Conclusion and Future Work

- Pretrained networks appear to produce good results with little adjustment
- Which architecture you choose and the distribution of classes in the data have a great impact on the results.

Better results could be obtained by

- a more balanced distribution of classes. Achieved by collecting more data of the rarer species or by weighting the augmentation process in the preprocessing phase to favor the rarer species.
- a careful selection of hyperparameters could further improve results.
- experimenting with different architectures.