Classify Flower Species Using Convolutional Neural Network Transfer Learning

Hong Xu
Georgetown University
hx67@georgetown.edu

Tianjing Cai
Georgetown University
tc856@georgetown.edu

Zihao Li
Georgetown University
zl225@georgetown.edu

Lufei Wang
Georgetown University
lw738@georgetown.edu

Abstract

In this project, my team applied different Convolutional Neural Network (CNN) models, including VGG16, VGG19 ResNet50, ResNet152, and Inception V3 on the flower classification problem. The dataset is provided by the University of Oxford and contains 102 flower categories. All the models are tested on the dataset and the results shows that Inception v3 has the highest accuracy. The team also tested the effect of image augment on the result of the classification with the ResNet50 model and find that augment could increase the accuracy by around 3%.

1 Introduction

Deep learning is currently a remarkably active research area in machine learning and artificial intelligence and has been extensively and successfully applied in numerous fields. Essentially, it is a class of machine learning techniques that exploit many layers of non-linear information processing for supervised or unsupervised feature extraction and transformation, and for pattern analysis and classification (Deng and Yu, 2014). Deep learning techniques have been widely applied in many sectors of the world, such as Business, agriculture, automotive industry, in object detection and image classification (Mohanty et al., 2016; Sladojevic et al., 2016; Dyrmann et al., 2016; Reyes et al., 2015). There have been breakthroughs for image classification through the deep Convolution Neural Network (CNN).

Recently, many modifications of CNN architecture have been proposed with a gradual increase in the number of layers. Some of the architectures include: GoogLeNet Inception V3 (Szegedy et al., 2015), VGG net (Simonyan and Zisserman, 2015), and Microsoft ResNet (He et al., 2016). These deep networks have difficulties and challenges such as exploding and vanishing gradients and degradation in the training process. Most deeper networks suffer from the degradation problem, where there is a reduction of accuracy when the depth of the network exceeds maximum. Another challenge is the internal covariate shift which is the change of the distribution of the input data to a layer during training. To solve these problems, these CNN pre-trained models could be combined with optimization techniques, including skip connections (He et al., 2016), transfer learning (Pan and Fellow, 2009), initialization strategies (Mishkin and Matas, 2016), Optimization strategies (Le et al., 2011), batch Normalization (Szegedy and Com, 2015) and data augmentation (Wong, Sebastien, et al, 2016). This paper focus on test the performance of different CNN pre-trained models and the influence of data augmentation on the image classification on the flower categories classification.

The dataset used in this project is the flower image dataset from the University of Oxford. To evaluate the performance of each model, the team uses top one accuracy, top three accuracy, and top five accuracy, which represent that the rate for the categories which have the top one, top three, and top five highest predicted possibilities contains the right answer.

2 Related Work

The University of Oxford provides two versions of the flower categories dataset, one contains 17 categories and the other one contains 102 categories. There is a blog which tests the performance of the models on the smaller dataset. The result in the blog is as follows:

Model	Image size	Top-1 accuracy	Top-5 accuracy
VGG16	224 x 224	0.715	0.901
VGG19	224 x 224	0.727	0.91
ResNet50	224 x 224	0.759	0.929
InceptionV3	299 x 299	0.788	0.944

Table 1: The Result for Smaller Dataset

According to the blog, the Inception v3 ended up with the best top one and top five accuracies. In this project, the team uses larger dataset. In addition, the results will include the top three accuracy in the result and take the effect of image augmentation into account. The team will also compare the performance of the models on the smaller dataset and the larger dataset classification problem.

3 Dataset

The Department of Engineering Science at the University of Oxford have created a flower dataset, consisting of 102 flower categories. The flowers chosen to be flower commonly occurring in the United Kingdom. Each class consists of between 40 and 258 images. The images have large scale, pose and light variations. The pixel resolution of each image is different. In addition, there are categories that have large variations within the category and several very similar categories.

Image 1: Sample Images for Dataset

For this project, our team use 6149 images as training set, 1020 images as validation set, and 1020 images as test set. To ensure the balance in the validation set and test set, the team randomly choose 10 images in each flower category for the validation set and test set respectively.
As shown in the Related Work section, each model requires different image size. In the pre-processing step, all images are resized based on the input requirements.

4 Method

This paper focus on the effect of data augmentation and comparison between the performance of different CNN pre-trained models on flower images classification.

4.1 Data Augmentation

Image augmentation artificially creates training images through different ways of processing or combination of multiple processing, such as random rotation, shifts, shear and flips, etc. In this project, the team use the ImageDataGenerator API in Keras to generate the additional images. The parameters chosen are as follows:

Parameter	Value
rotation_range	90
width_shift_range	0.2
height_shift_range	0.2
rescale	1./255,
shear_range	0.2
zoom_range	0.2
horizontal_flip	TRUE
fill_mode	nearest

Table 2: The Parameter Used for Image augmentation

The sample for Image augmentation is as follow:

Image 2: Sample Result for Image augmentation

The team test the impact of Image augmentation With ResNet50 model, which means both the data with and without Image augmentation are used to train this model and the results for the test set will be used to evaluate the influence.

4.2 Transfer Learning

The team include the VGG16, VGG19, ResNet50, ResNet152, Inception v3 in the paper. As mentioned above, all images are resized to the required size for each model before the training process. To prevent over-fitting, the team adds the early stop to the training process and the maximum epoch is 150. VGG models need around 70 epochs to converge, while the ResNet models used 40 epochs. The training time for Inception v3 is longest and the training process ends when reaching the maximum epoch allowed.

5 Result

The result are as follows:

	Top1 Accuracy	Top3 Accuracy	Top5 Accuracy
ResNet50	0.5951	0.7902	0.9324
ResNet50 (NO Image Augmentation)	0.5656	0.7617	0.9019
VGG16	0.6549	0.8088	0.9176
Inception_v3	0.9078	0.9608	0.9922
ResNet152	0.6326	0.7924	0.9421
VGG19	0.7268	0.8552	0.9125

Table 3: Result of Transfer Learning

According to the result, the team find that the Image Augmentation could improve the prediction accuracy by 3%. The possible reasons might be that Image Augmentation could help to enlarge the training dataset and prevent over-fitting. With the image augmentation, we can find that all models could predict the flower categories with at least around 60% accuracy. Among all models tested with the test set, Inception v3 has the highest accuracy and the Top 5 accuracy is around 99.22%.

6 Discussion of Results

In this project, the teams use a larger dataset with more flower categories than the previous study so that the team could compare the performance of different CNN models on a classification task on larger dataset. According to the comparison between our result and the blog mentioned in the related work, we can find that the top one accuracy for all models except Inception v3 decrease. The possible reason is that there are much more categories in our dataset and the difficulty for classification increase. Besides, the top five accuracy are very close which means that the performance of these models on a dataset with a larger size remains promising. The accuracy of Inception v3 increase and it outperform other models which indicates that this model has great advantage on deal with larger dataset.

In addition, the result for this project suggest that Image Augmentation has significant impact on the classification accuracy, which is a supplement to the previous study.

7 Conclusions

Based on the result, we can conclude that Image Augmentation has positive influence on the image classification task, especially when there are more categories and images in the dataset. Besides, Inception v3 has the best performance among all CNN pre-trained models on a larger dataset. In the future, the team will focus on finding out the reasons for the best performance of the Inception model and testing other models, such as AlexNet (Krizhevsky et al., 2012), Inception V4 (Szegedy et al., 2016), and DenseNets (Huang et al., 2016).

References

Deng, L., Yu, D., 2014. Deep Learning: Methods and Applications. Foundations and Trends® in Signal Processing. pp. 3–4

Dyrmann, M., Karstoft, H., Midtiby, H.S., 2016. Plant species classification using deep convolutional neural network. Biosyst. Eng. 151 (2005), 72–80

He, K., Zhang, X., Ren, S., Sun, J., 2016. Identity Mappings in Deep Residual Networks. pp. 1–15.

Huang, G., Weinberger, K.Q., Van Der Maaten, L., 2016. Densely Connected Convolutional Networks.

Krizhevsky, A., Sutskever, I., Geoffrey, E.H., 2012. ImageNet classification with deep convolutional neural networks. Adv. Neural Inf. Process. Syst. 25 (NIPS2012), 1–9. http://dx.doi.org/10.1109/5.726791.

Le, Q.V., Coates, A., Prochnow, B., Ng, A.Y., 2011. On Optimization Methods for Deep Learning,. In: Proc. 28th Int. Conf. Mach. Learn., pp. 265–272.

Mishkin, D., Matas, J., 2016. All you need is a good init, pp. 1–13.

Mohanty, S.P., Hughes, D.P., Salathé, M., 2016. Using deep learning for image-based plant disease detection. Front. Plant Sci. 7 (September), 1–7.

Reyes, A.K., Caicedo, J.C., Camargo, J.E., 2015. Fine-tuning deep convolutional networks for plant recognition. CEUR Workshop Proc. 1391.

Simonyan, K., Zisserman, A., 2015. Very deep convolutiona

Szegedy, C., Com, S.G., 2015. Batch Normalization : Accelerating Deep Network Training by Reducing Internal Covariate Shift, vol. 37. Szegedy, C., Vanhoucke, V., Shlens, J., 2015. Rethinking the Inception Architecture for Computer Vision. Szegedy, C., Ioffe, S., Vanhoucke, V., Alemi, A., 2016. Inception-v4, Inception-ResNet and the Impact of Residual Connections on Learning.

Sladojevic, S., Arsenovic, M., Anderla, A., Culibrk, D., Stefanovic, D., 2016. Deep neural networks based recognition of plant diseases by leaf image classification. Comput. Intell. Neurosci. 2016 (June).

Wong, Sebastien, et al. 2016“Understanding Data Augmentation for Classification: When to Warp?” arXiv.org, Cornell University Library, arXiv.org, Nov., http://search.proquest.com/docview/2080891105/.