Deep Learning-Based Convolutional Neural Network for Flower Classification

Yangchuan Liu[a]
High School Affiliated to South China Normal University, Guangzhou, China

Keywords: Machine Learning, Convolutional Neural Network, Flower Classification.

Abstract: To tackle the flower classification problem, this study utilizes the Oxford 102 Flowers dataset and develops a machine learning model using a Convolutional Neural Network (CNN). Initially, the preprocessing phase involved the use of grayscale images. However, recognizing the critical role that color plays as a distinctive feature of flowers, the study shifted to using RGB images. To further enhance the model's performance, data augmentation techniques were introduced. These included random adjustments to brightness, saturation, contrast, and hue, which helped diversify the training set and improve the model's generalization ability. To mitigate overfitting, several strategies were employed, such as tuning the number of neurons and incorporating Dropout Layers. These approaches helped the model achieve a validation accuracy of approximately 0.7, which is sufficiently accurate for basic flower classification tasks in everyday applications. This outcome demonstrates the effectiveness of the chosen methods and highlights the potential of CNNs in flower image classification.

1 INTRODUCTION

The flower is the reproductive organ of a seed plant. Composed of corolla, calyx, receptor, and stamen, it has a variety of shapes, colors, and fragrances. There are countless types of flowers worldwide, which have multiple functions in different domains. For example, flowers can be used as decorations in ceremonial events, given to others as gifts on special days, and even processed into food or tea. Because of these uses, there is a considerable flower demand that leads to a large market for the flower flower-growing industry, and technology must be the key to improving the productivity of flowers. However, flower classification is one of the challenging tasks in flower growing. Traditional flower identification methods rely on specialized knowledge and manual classification, which can be time-consuming, labor-intensive, and subject to personal bias. With the rapid advancement of computer vision and machine learning technologies, data-driven automated recognition systems have emerged as a solution to improve productivity and reduce labor costs.

In recent years, Artificial Intelligence (AI) technology has gradually matured, and many fields have been widely used. In the manufacturing industry, image recognition technology and computer vision offer developing opportunities for driverless cars and automatic face recognition systems. In the economy, machine learning provides a more reliable way to predict the future market of certain goods by training with past data. In medicine, neural networking can precisely produce predictions about patients' latent diseases by collecting current body data. More generally, the current Large Language Model (LLM), such as Chat Generative Pre-Trained Transformer (ChatGPT), can even provide help and opinions to people in all fields. There is undoubtedly an astonishing advancement in the AI domain. For instance, back to focus on the planting industry, in 2016, Liu et al. proposed a novel framework based on a Convolutional Neural Network (CNN) that achieves 84.02% flower classification accuracy (Liu, 2016). Based on the study of Knauer et al.in 2019, they found that CNN has a better performance in the tree species classification task than the Random Forest (RF) classifiers (Knauer, 2019). By analyzing the previous relative research and confirming the effectiveness of AI, this research aims to consider the machine learning technology based on CNN to be

[a] https://orcid.org/0009-0000-0082-2061

Liu, Y.
Deep Learning-Based Convolutional Neural Network for Flower Classification.
DOI: 10.5220/0013329800004558

In *Proceedings of the 1st International Conference on Modern Logistics and Supply Chain Management (MLSCM 2024)*, pages 315-319
ISBN: 978-989-758-738-2

feasible and suitable to solve the problem of flower classification.

To achieve the final goal, this paper first decided and downloaded a suitable dataset Oxford 102 Flowers that contains 8, 200 samples of 102 different types of flowers. Then this study standardized the image by resizing all of them to 200-pixel length squares and converting them into grayscale. The model' s neural network consists of 4 convolutional layers and 4 dense layers, and all of them use Rectified Linear Unit (ReLu) as the activation function. The input data is the pixel of an image, and the output is a vector that reflects the possibility of each of the 102 types. In the training process, the Adaptive Moment Estimation (Adam) was used as an optimizer, and the Cross Entropy Error was used as a loss function. After training of 100 epochs, the model accuracy finally reaches about 70%.

2 METHOD

2.1 Dataset Preparation

In the dataset preparation part, Oxford 102 Flower (102 Category Flower Dataset) is chosen as the training dataset in this study (Nilsback, 2008). This dataset consists of 8,200 RGB images divided into 102 flower categories that are commonly occurring in the United Kingdom. Each class contains between 40 and 258 images, and the images have large scale, pose, and light variations. Including these diverse and complicated features, this dataset is suitable for the model that aims to solve the complex flower classification task. When preprocessing the images, the grayscale images were first considered an appropriate method. As the RGB images are translated into the grayscale, image noise is reduced, while texture and structural features are enhanced, thus increasing the efficiency of model processing. Figure 1 clearly shows the difference between the RGB images and grayscale images.

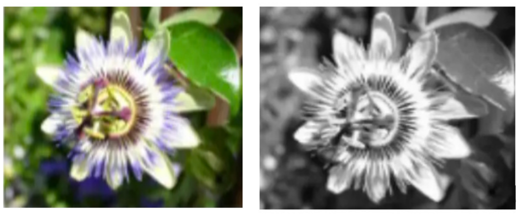

Figure 1: The RGB version(right) and the grayscale version(left) of a flower image (Photo/Picture credit: Original).

However, the study soon found that this preprocessing method leads to a low accuracy because the color is a necessary indicator of the flower categories. As a result, the gray-scale images are not feasible in the flower classification task; instead, considering the importance of the color, random adjustments of image brightness, saturation, contrast, and hue were used to enhance the model's adaptability to color changes. In addition, random up-and-down or side-to-side flips also improve the generalization of the model, and all of the images are resized into 200 pixels length squares, as normalizes the input shape of the model.

2.2 Convolutional Neural Network-Based Prediction

In the step of building up the neural network model, the Convolutional Neural Network (CNN) is first to be considered. CNN is predominantly used to extract the features from the grid-like matrix dataset, such as visual datasets like images or videos (Li, 2021; Gu, 2018; O'Shea, 2015). The components in the CNN include the Input Layer, Convolutional Layer, Activation Layer, Pooling Layer, Flattening, Fully Connected Layer, and Output Layer. In these layers, the Convolutional Layer and Pooling Layer are the essences that make CNN different from other models. Imagine that an image is a cuboid; specifically, the width and length of the cuboid are those of the image while the channels(height) can represent the RGB value of each pixel. By taking a small patch of this cuboid and running a small neural network, which is called a filter, the Convolutional Layer can extract the feature of an image when the filter slides in the cuboid and convert them into another image with different width, height, and channels, which is referred as feature maps. As important as the Convolutional Layer, the Pooling Layer is periodically inserted between the Convolutional Layer to reduce the size of volume which makes the computation fast reduces memory produced by multiple feature maps and prevents overfitting. Because the advantages of CNN include high accuracy of image analysis, robustness to image deformation and rotation, and the need for a large amount of label data, this network is suitable for the flower classification task that meets these conditions.

Considering the complex data contained in each image, the study uses 4 Convolutional Layers, 4 Pooling Layers, and 3 Fully Connected Layers to construct the model. Nevertheless, in this version, the model still does not perform well after 100 epochs. By studying the flaws of the traditional CNN, the

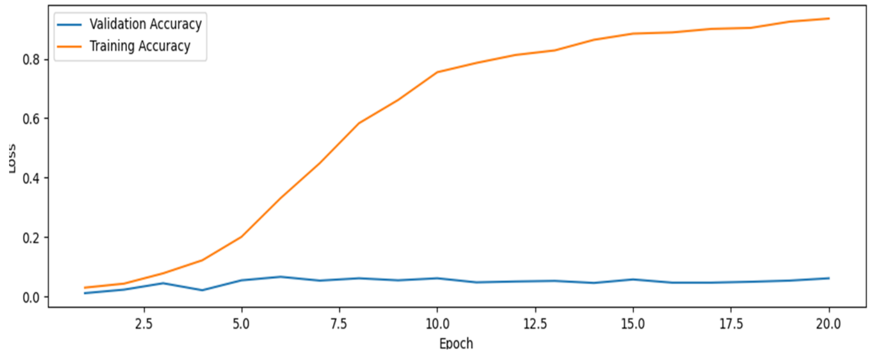

Figure 2: Model accuracy when use the gray-scale images in the preprocessing (Photo/Picture credit: Original).

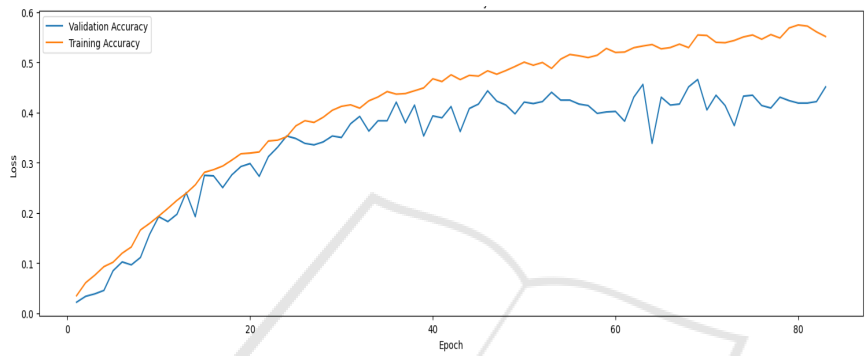

Figure 3: Model accuracy when using the RGB image and data augmentations in the preprocessing (Photo/Picture credit: Original).

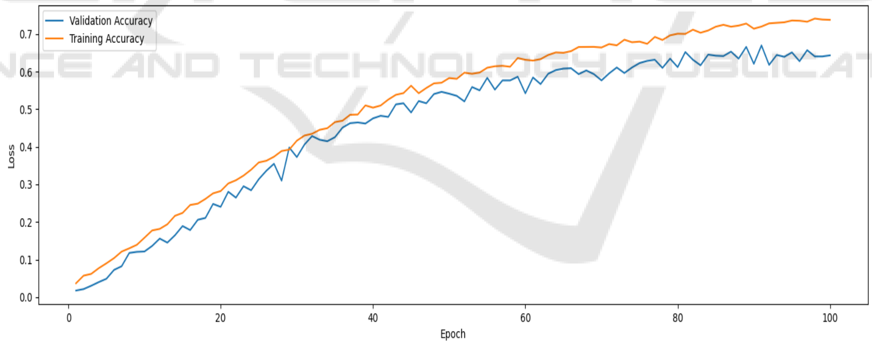

Figure 4: Model accuracy when adding some Dropout Layers in the model (Photo/Picture credit: Original).

study notices that overfitting is the main cause of low performance. Overfitting always happens when fully connected layers have a large number of neurons that extract very similar features from the input data; then, these neurons add more significance to those features for the model, which only work in specific datasets (Srivastava, 2013; Salakhutdinov, 2014). To solve this problem, Dropout Layers -- randomly shut down some fraction of a layer' s neurons by zeroing out the neuron values -- are added between each Fully Connected Layer to reduce the model ' s over-relying on specific neurons.

2.3 Implementation Details

The hyper-parameters of the model are also significant, such as the optimizer, loss function, neuron amount, learning rate, and so on. Because the task is a classification problem, Cross Entropy Error is used as the loss function (Golik, 2013). Based on the complexity of the flower classification task, Adaptive Moment Estimation (Adam) (Tato, 2018; Zhang, 2018), which automatically adjusts the learning rate and occupies lower memory, is adopted as the optimizer. Moreover, to capture abundant features of flowers, the model has a relatively large number of neurons in Fully Connected Layers and filters in Convolutional Layers.

3 RESULTS AND DISCUSSION

In Figure 2, the graph shows the validation accuracy and training accuracy of the model that uses the gray-scale images in the preprocessing. It is clear that the training accuracy increases rapidly during the training process and even reaches over 0.9 after 20 epochs. However, the validation accuracy remains at a shallow level, basically no more than 0.1.

In the Figure 3, RGB images are used instead of gray-scale images; moreover, data augmentations are also added in the preprocessing. The graph shows that the validation accuracy begins to grow with the training accuracy and fits well in the early stage. However, after 40 epochs, the two lines begin to separate: the difference between the training accuracy and validation is greater and greater. At the same time, the final validation accuracy, which stops growing and reaches only about 0.4, is not enough for the expectation of the study.

In the Figure 4, Dropout Layers are added and the neuron amount in some layers is adjusted in the model. In the graph, the two lines fit well in general, and the final validation accuracy makes a huge improvement, which reaches nearly 0.7.

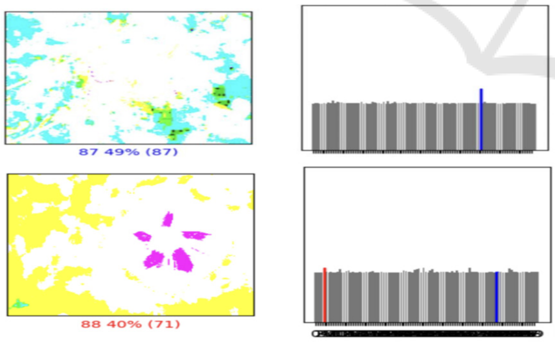

Figure 5: Model' s prediction of two image with data augmentation (Photo/Picture credit: Original).

As shown in Figure 5, both of the images experience data augmentation that randomly changes brightness, saturation, contrast, and hue significantly. The prediction of the first one can still be correct while that of the second one is incorrect.

From the information of the first version model, because it only takes the model 20 epochs to reach the training accuracy over 0.9, it is easy to find the high efficiency of gray-scale preprocessing in image-related tasks. Nevertheless, because flower color as a feature cannot be ignored, the gray-scale image is not feasible. Some inappropriate features, which only belong to the training dataset and do not exist in the validation dataset, are extracted and trained by the model. As a result, although the training accuracy is impressive, the validation accuracy is extremely low, which does not meet the requirement.

In the second version, the RGB images and data augmentations help the model begin to capture and analyze the features of flowers' color. The fitting of training accuracy and validation accuracy proves the opinion that color is non-negligible. However, in this step, the increased complexity due to converting 1-channel grayscale images to 3-channel RGB images is not considered. Therefore, the simple model could not adapt to the abundant features carried by the RGB images and did not perform well, only reaching a validation accuracy of about 0.4 before it stopped growing accuracy.

In the final version, the neuron amount is adjusted to improve the model complexity, while the Dropout Layers are used to decrease redundancy features recorded in neurons to avoid overfitting. Because of the improvement of the complexity, the model can afford and extract the features well and finally reaches its peak validation accuracy of about 0.7 without an obvious overfitting phenomenon. It is undoubtedly precise enough for some basic flower classification tasks in daily life because most flower images in real life have more distinguishing features and do not resemble the images with data augmentations in Figure 4, which could be even harder for human beings to discern.

However, there are still some limitations of the proposed model including lacking training data: the 8,200 images used to classify 102 kinds of flowers are not enough. To achieve higher accuracy, by considering CNN' s reliance on the dataset, more images should be collected as the training dataset of this model.

.4 CONCLUSIONS

In this study, a machine learning model is built and trained to solve the flower classification task. Oxford 102 Flower which consists of 8,200 RGB images divided into 102 flower categories is used as the dataset. For its high performance in image-related tasks, CNN is considered the most suitable network of the model instead of Random Forest (RF); furthermore, to improve its accuracy and robustness, data augmentations, large number of neurons, and Dropout Layers are used. As a result, the accuracy of the model reaches about 0.7, which is impressive and enough for most practical flower classification tasks when it has to classify some images with terrible distortion. Some improvements can be made in the future if use larger and more diverse datasets to improve the accuracy to a higher level.

REFERENCES

Gu, J., Wang, Z., Kuen, J., Ma, L., Shahroudy, A., Shuai, B., ... & Chen, T. 2018. Recent advances in convolutional neural networks. Pattern recognition, 77, 354-377.

Golik, P., Doetsch, P., & Ney, H. 2013, August. Cross-entropy vs. squared error training: a theoretical and experimental comparison. In Interspeech (Vol. 13, pp. 1756-1760).

Knauer, U., von Rekowski, C. S., Stecklina, M., Krokotsch, T., Pham Minh, T., Hauffe, V., ... & Seiffert, U. 2019. Tree species classification based on hybrid ensembles of a convolutional neural network (CNN) and random forest classifiers. Remote Sensing, 11(23), 2788.

Li, Z., Liu, F., Yang, W., Peng, S., & Zhou, J. 2021. A survey of convolutional neural networks: analysis, applications, and prospects. IEEE transactions on neural networks and learning systems, 33(12), 6999-7019.

Liu, Y., Tang, F., Zhou, D., Meng, Y., & Dong, W. 2016, November. Flower classification via convolutional neural network. In 2016 IEEE International Conference on Functional-Structural Plant Growth Modeling, Simulation, Visualization and Applications (FSPMA) (pp. 110-116). IEEE.

Nilsback, M. E., & Zisserman, A. 2008, December. Automated flower classification over a large number of classes. In 2008 Sixth Indian conference on computer vision, graphics & image processing (pp. 722-729). IEEE.

O'Shea, K. 2015. An introduction to convolutional neural networks. arXiv preprint arXiv:1511.08458.

Srivastava, N. 2013. Improving neural networks with dropout. University of Toronto, 182(566), 7.

Srivastava, N., Hinton, G., Krizhevsky, A., Sutskever, I., & Salakhutdinov, R. 2014. Dropout: a simple way to prevent neural networks from overfitting. The journal of machine learning research, 15(1), 1929-1958.

Tato, A., & Nkambou, R. 2018. Improving adam optimizer. Openreview

Zhang, Z. 2018, June. Improved adam optimizer for deep neural networks. In 2018 IEEE/ACM 26th international symposium on quality of service (IWQoS) (pp. 1-2). IEEE.