IMAGE BASED FLOWER SPECIES CLASSIFICATION USING CNN

Santosh Giri[*1]

ABSTRACT

Deep learning is one of the essential parts of machine learning. Applications such as image classification, text recognition, object detection etc. used deep learning architectures. In this paper neural network model was designed for image classification. A NN classifier with one fully connected layer and one softmax layer was designed and feature extraction part of inception v3 model was reused to calculate the feature value of each images. And by using these feature values the NN classifier was trained. By adopting transfer learning mechanism NN classifier was trained with 17 classes of oxford 17 flower image dataset. The system provided final training accuracy of 99 %. After training, system was evaluated with testing dataset images. The mean testing accuracy was 86.4%.

Index Terms *Image Classification, Flower Image, Neural Network, Deep Learning, Transfer Learning.*

I. INTRODUCTION

The concept of neural network was first introduced by W Pitts and W Mcculloch in 1943 [1], the model was based on neural network that tries to simulate the human brain. The model was a simple neural network with electrical circuits. In 1960s H Kelly [2] and A Bryson [3] was given credit for developing continues back propagation, a supervised learning algorithm for training an Artificial Neural Network. And the practical illustration of back propagation was presented by Yann LeCun [4] at Bell lab in 1989. He has combined convolutional neural network and back propagation for handwritten digit recognition [4]. His work is still referred in todays research work. In 1995, Support Vector Machine [5], a system used for analyzing and recognizing similar data was developed by C Cortes and V Vapnik. Also in 1997 Long Short Term Memory

[6] was developed by S Hochreiter, J Schmidhuber, which is used in Recurrent Neural Network. These systems are still used frequently in today machine learning research. Furthermore, development of Graphical Processing Unit in 1999, introduced faster processing of data. During this time neural network started complete with support vector machine. Neural network was slower compared to SVM but offer better accuracy on same data. Neural network have advantage of obtaining better result if more training data is added. With increased computing speed, deep learning provided significant improvement on efficiency and processing speed. One example is AlexNet [7]

1   *Department of Computer and Electronics Engineering Thapathali Campus, Kathmandu, Nepal*

*   *Corresponding author*
     girisantos016@gmail.com

Basanta Joshi[*2]

a Convolutional neural network that won many international competition during 2011 and 2012 AD.

II. RELATED WORKS In [8], unsupervised segmentation on training sets of flower image is applied to separate an image into foreground and background in order to improve image classification performance. To achieve this new scalable, alternation-based algorithm called BiCoS [8] is used for co-segmentation. In preprocessing Grabcut algorithm [9] a tool for image segmentation is applied at pixel level to each image independently. Grabcut [9] combines a binary label random field defined on image pixels and a generative background/ foreground classifier. It uses Gaussian mixtures to estimate foreground/ background probabilities given that foreground/background pixel labels. Each super pixel is then assigned to background and foreground and mapped into a descriptor space using super pixel descriptors. Then linear SVM [5] was used to classify those super pixels into foreground and background. During classification, the images were classified with shared background patterns for all categories. In [10], Nearest neighbor classifier [11] was applied for flower image classification. The classification involved number of stages, in first stage a separate vocabulary is developed for color, shape and texture. SIFT descriptor [12] was used to describe the shape, HSV [13] values to describe the color and

MR filters [14] to describe texture of the flower image. Those three aspects are then combined into single vocabulary.
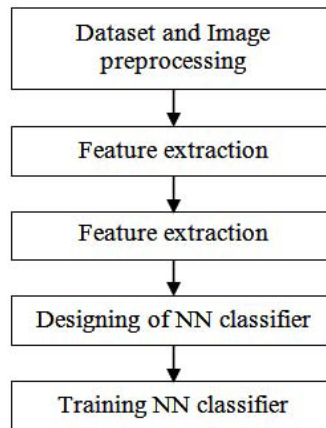
III. METHODOLOGY

Fig. 1. Research methodology followed during this experiment.

[2]   *Thapathali Campus,*

[2]   *Corresponding author*
       nepalbasanta@ioe.edu.np

A. Image Pre-processing

In order to train and test the neural network classifier, Oxford 17 flower dataset [10] was used. This dataset consist of 17 categories of flower images with 80 images for each class. The flowers are some common flowers in UK. In preprocessing, the dataset is divided into training set and testing set. Training consists of 40 images per class and testing sets consist of 10 images per class. Each images from the flower dataset [10] are then resized into 299X299X3.

B. Feature vector extraction

The feature extraction capability of pre-trained Inception v3 model [15] was used here to calculate the feature value from an image. Inception model is 42 layers deep, which takes input image as 299 x 299 of 3 channels. The inception model has two parts: a feature extraction and a classification layer. We make use of feature extraction part of inception model to calculate the feature value of each image. Feature value is 1xN array of float value where N=2048. We input the custom images of size 299x299x3 into the features extraction part of CNN model, and then pre-trained model converts the image into feature vectors consisting of 2048 float values for each image, representing the features of the image in an abstract manner. The features vectors for each image during preprocessing is calculated and then read cached values repeatedly during training the NN classifier.[3]

[1] W. S. McCulloch and W. Pitts, "A logical calculus of the ideas immanent in nervous activity," **The bulletin of mathematical biophysics**, vol. 5, no. 4, pp. 115–133, 1943.

[2] H. J. Kelley, "Gradient theory of optimal flight paths," **Ars Journal**, vol. 30, no. 10, pp. 947–954, 1960.

[3] A. E. Bryson, "A gradient method for optimizing multi-stage allocation processes," in **Proc. Harvard Univ. Symposium on digital computers and their applications**, vol. 72, 1961.

[4] Y. LeCun, B. Boser, J. S. Denker, D. Henderson, R. E. Howard,

W. Hubbard, and L. D. Jackel, "Backpropagation applied to handwritten zip code recognition," **Neural computation**, vol. 1, no. 4, pp. 541–551, 1989.

[5] C. Cortes and V. Vapnik, "Support-vector networks," **Machine learning**, vol. 20, no. 3, pp. 273–297, 1995.

[6] S. Hochreiter and J. Schmidhuber, "Long short-term memory," **Neural computation**, vol. 9, no. 8, pp. 1735–1780, 1997.

[7] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "Imagenet classification with deep convolutional neural networks," in **Advances in neural information processing systems**, 2012, pp. 1097–1105.

[8] Y. Chai, V. Lempitsky, and A. Zisserman, "Bicos: A bi-level cosegmentation method for image classification," 2011.

[9] C. Rother, V. Kolmogorov, and A. Blake, "Grabcut: Interactive foreground extraction using iterated graph cuts," in **ACM transactions on graphics (TOG)**, vol. 23, no. 3. ACM, 2004, pp. 309–314.

[10] M.-E. Nilsback and A. Zisserman, "A visual vocabulary for flower classification," in **Computer Vision and Pattern Recognition, 2006 IEEE Computer Society Conference on**, vol. 2. IEEE, 2006, pp. 1447–1454.

[11] J. M. Keller, M. R. Gray, and J. A. Givens, "A fuzzy k-nearest neighbor algorithm," **IEEE transactions on systems, man, and cybernetics**, no. 4, pp. 580–585, 1985.

[12] D. G. Lowe, "Distinctive image features from scale-invariant keypoints," **International journal of computer vision**, vol. 60, no. 2, pp. 91–110, 2004.

[13] A. R. Smith, "Color gamut transform pairs," **ACM Siggraph Computer Graphics**, vol. 12, no. 3, pp. 12–19, 1978.

[14] M. Varma and A. Zisserman, "Classifying images of materials: Achieving viewpoint and illumination independence," in **European Conference on Computer Vision**. Springer, 2002, pp. 255–271.

C. Designing of NN classifier

Fig. 2. Neural Network classifier with one fully connected layer and one softmax classifier.
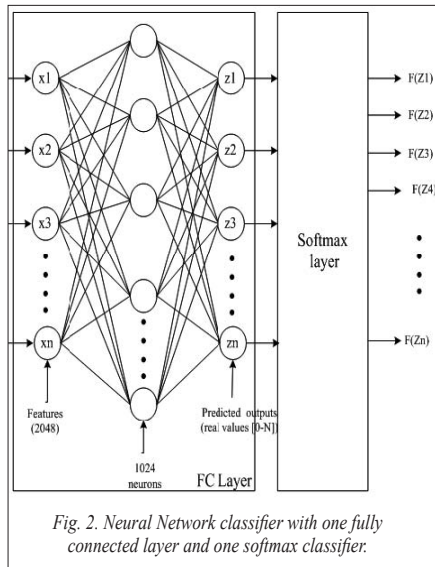
1) Fully connected layer: The fully connected layer of the classification part consists of 1024 neurons or nodes. The fully connected layer takes feature vector of 2048 float values from feature extraction part of pre-trained CNN [15] as input and produce linear value ZL for each class. The linear value is calculated as

$$Z_L = W_L * A_{L-1} + b_L \qquad \text{where,}$$

$$Z_L = \text{Linear Value}$$
$$W_L = \text{weight value}$$
$$A_{L-1} = \text{feature value}$$

2) Softmax layer: The soft max layer of NN classifier consists of with 17 nodes for flower dataset [10]. The softmax layer takes linear value i.e. Z value as input and calculate probability for each class using softmax function. The probability of class 'j' is calculated as follow:

$$P(C_j) = \frac{e^{Z_j}}{\sum_{k=1}^{j} e^{Z_j}} \qquad \text{where,}$$

$$C_j = j^{th} \text{ class}$$
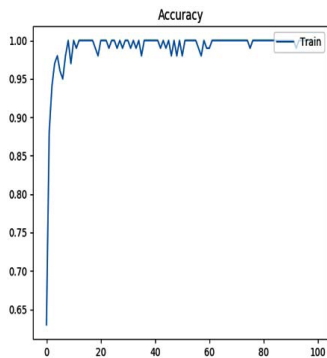$$Z_j = \text{Arbitory real value of class j}$$

D. Training NN classifier

Fig. 3. Training accuracy graph on Oxford 17 flower dataset [10].The parameters used for training the model was, 2500 iterations and learning rate of 0.01. The x label represents the total no. of iterations and y label represents training accuracy value.
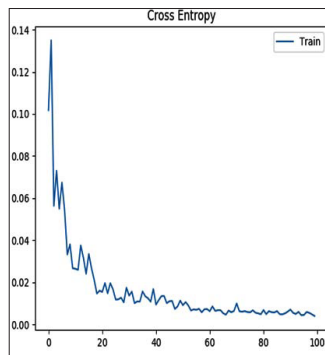

Fig. 4. Cross entropy error graph of training set images on Oxford 17 flower dataset [10]. The x label represents the total no. of iterations and y label represents the value of cross entropy value.

After Designing NN classifier, the system was trained with Oxford 17 flower image dataset [10]. The inputs to NN classifier are feature vectors that were calculated by features extraction part of the pre-trained CNN. To train the new classification layer Back propagation algorithm [4] was used.

The training accuracy and cross entropy graph were given in Fig. 3 and Fig. 4.

[15] C. Szegedy, V. Vanhoucke, S. Ioffe, J. Shlens, and Z. Wojna, "Rethinking the inception architecture for computer vision," in **Proceedings of the IEEE conference on computer vision and pattern recognition**, 2016, pp. 2818–2826.

蟲椀最「.. ....
....
....

IV. RESULTS AND DISCUSSION

The testing results of the model on Oxford 17 flower image dataset [10] is represented in graph given in fig. 5. 10 flower images were used from each class to test the model. True Positive in graph represents correctly classified images from each class. And performance comparison of our system with other experiment on flower dataset [10] is given in Table I.
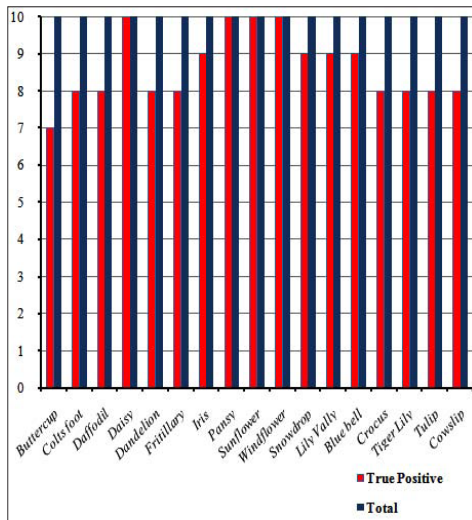

Fig. 5. performance evolution chart of the model on Oxford 17 flower dataset [10].

TABLE I PERFORMANCE COMPARISON OF OUR SYSTEM WITH OTHER EXPERIMENT ON FLOWER DATASET [10].

S.N.	Method	Mean Precision
1	Color features [10]	73.7 %
2	Shape features [10]	71.8 %
3	Our system	86.4 %

V. CONCLUSIONS

In this paper we represented an approach to better classify the flower images. We reused the feature extraction capability of Inception v3, a pre-trained CNN and trained our NN classifier. The system gives classification accuracy of 86.4%.