

# Flower Species Detection Over Large Dataset Using Convolution Neural Networks and Image Processing Techniques

Simriti Koul, Udit Singhania

**Abstract:-** We investigate flower species detection on a large number of classes. In this paper, we try to classify flower species using 102 flower species dataset offered by Oxford. Modern search engines provide methods to visually search for a query image that contains a flower, but it lacks robustness because of the intra-class variation among millions of flower species around the world. So, we use a Deep learning approach using Convolutional Neural Networks (CNN) to recognize flower species with high accuracy. We use the Oxford dataset which was made by the use of electronic items like a built-in camera in mobile phones and also a digital camera. Feature extraction of flower images is performed using a Transfer Learning approach (i.e. extraction of complex features from a pre-trained network). We also use image augmentation and image processing techniques to extract the flower images more efficiently. After the experimental analysis and using different pre-trained models, we achieve an accuracy of 85%. Further advancements can be made by using optimization parameters in the neural nets.

**Keywords:-** Deep Learning, Artificial Intelligence, Convolutional Neural Networks, Transfer Learning, Flower Recognition, Image Processing.

## I. INTRODUCTION

Image classification systems improve effectively when we work on a large number of classes. Recent work [2, 8, 9, 16, 17, 18] has seen much success in this area. It is difficult to recognize between flower species due to great similarities in each of classes, it is not like car, bikes or cats. We use ImageNet pre-trained model weights in VGG19 and then we retrain it on our flower dataset. Previous work on flower classification has dealt with a small number of classes [12, 14] ranging from 10 to 30. In figure 1, we see many images of flower present in the dataset. There are certain factors which are needed to be kept in mind while classifying between flowers like color, size of petals, shape, etc. Working on these features considered for 102 flower species is a cumbersome and tedious task, and outputting an accurate result.

Furthermore, when it comes to flowers, the photographers prefer to take these pictures in natural settings with rich and challenging backgrounds. Even though the background usually provides necessary context, it can at times distract a classification algorithm. For instance, some aspects in the background can be the one more emphasized upon and be highlighted as good discriminators, or some background features may be complemented across various classifications and thus make it hard to differentiate among them (Figure 2). This likely affects the execution of the classification algorithm. One of the ways to solve this problem is by identifying the area not covered by the object, in order to discard the background during classification. It will result in a benefitted outcome if the object can be automatically extracted before the recognition system executes so that it can focus on the important areas of the image.

When provided a sample image, a random person instantly identifies the object of interest, so it is evident that when an expert is required to categorize the image, they will be focused on the most essential foreground region. In this paper, the focus is mainly on automatically extracting the important region of the object of interest before implementing the classification process. The optimization technique used here is the Laplacian propagation, which ensures fast convergence and also leads to a noticeable decrease in the overall run-time. This technique is 5-6 times faster than previously known algorithms in similar cases [2, 13]. Furthermore, the method is not complex and can be applied to a different datasets, unlike many previous researches on flowers [11] or other classifications [13] whose segmentation techniques are specialized to the higher-level classification at present. The experiments in our work portray that the segmentation technique proposed by us, in addition to being less complex and faster in compilation, is a benefactor to the categorization execution. We have executed the algorithm on a well-established flower recognition dataset, the Oxford flower dataset containing 102 species of flowers [11], as well as on our large-scale 578-class flower dataset. The proposed algorithm has proven to improve the basic outcomes for both datasets by at least 4%.







Fig 1:- Oxford Dataset containing different flower species

## II. METHODOLOGY

Our methodology depends on recognizing areas, explicit of the classifications of important features of the sample image at the detection. We then apply the Laplacian-based propagation and segmentation technique to deal with fragment the item in view of lesser-level signs. The Laplacian-based propagation procedure is directed by the first distinguished districts, as they are currently the great markers of the nearness of the conceivable article. The underlying areas are known to be recognized by a specific learning model. The priority instincts of our methodology are that while sectioning an item, perceiving portions of the picture that conceivably has a place with the article can help outline object limits which may not be generally conspicuous. A fragmented article, thus, is helpful for the last acknowledgment, has appeared in our investigations. This is the situation on the grounds that the calculation 1) figures out how to evacuate foundation zones which might mistake for the grouping calculation; 2) it gives limits and shapes data to the article. The calculation is straightforward, a lot quicker than the recently utilized methodologies for division [3, 14], and beats past grouping strategies on standard sets of data. We then go forth with processing the sectioned picture (along with the first picture) by the regular component withdrawal and grouping channel, so both the disposal of the foundation and the extricated outline data can conceivably influence the separated highlights emphatically. While segment division is being utilized in many item acknowledgment settings, our segmentation technique is strong, versatile, and sufficiently productive when it comes to being connected at the analysis time.

## III. RELATED WORK

A huge organization of segmentation work is present [1, 2], with the maximum of it focusing on offline segmentations applied to the training data. For instance [1] discovered methods in which progressively improved segmentations are utilized to study enhanced models for identifying the objects in the database. Likewise, in the co-segmentation body of works [2, 9], better models are

trained by using common appearance aspects in images comprising of the same class of objects. These methods are primarily emphasizing on segmentation while training of the data. Recent works have presented segmentation performed while the categorization time. In [13] the authors present to identify certain particular aspect of the object of interest, and then section the object by deducing from the textures and hues observed. These techniques may at times undergo few assumptions they make that the object has constant texture. Recent work [2] suggested doing segmentation before performing recognition, but they utilized the iterative GrabCut algorithm [14] whose running time is a constraint to its application to offline settings. Another fascinating research on semantic segmentation [3] has comparable runtime. Those techniques, although presented through segmentations, are however slow for the categorization to be done in real-time applications. Lesser than the width of a single column, as major elaborations may change the clarity of your input images and lead to poor duplicate. Therefore, it will lead to better output if the image is slightly bigger, as a slight decrease in proportions must not greatly affect the quality of the image.

## IV. GLOBAL FEATURE DESCRIPTORS

Color, Texture and Shape are the three prime important aspects in the scenario of quantifying the images of flowers.

### ➤ *Detecting areas specific to the object*

We begin our methodology with a exploring for areas remotely appropriate to a given flower in the input image. For less complexity we utilize the super-pixel segmentation technique by Felzenszwalb and Huttenlocher [7] to further section the sample image into smaller rational areas. Each super-pixel area of the image is depicted by the particular set of descriptors that provided with features, namely, mediocre color (R,G,B) of all the pixels within the same area, universal assembling of all HOG features [4] in the region, after using the LLC technique for encoding them [15], outline mask of the output region by standardizing the area's vaulting containment to 6x6 pixels, and dimension and borderline aspects as in [2]. Some of the descriptions

for the features are taken from other segmentation techniques which used high-pixel descriptors: e.g. the use of outline masks and bit-maps symbolizing adjacent to the limit of the areas are proposed in [2]. Unlike previous researches, we use the encoded HOG features here, as we believe they have enhanced capabilities for the simplification and also they fit in our classification technique (Section 5) these aspects are already pre-compiled in the given image and are useful during further researches. Utilizing the aspect demonstration described above, we build a model which aims to differentiate between the lower and the background distractions. We can further apply this model to each area and highlight as well as extract the high buoyancy areas for background as well as foreground.

➤ *Color*

One of the most significant highlights to be considered to perceive flower species is "Color". One of the most solid and basic worldwide element descriptor is the Color Histogram which registers the recurrence of pixel powers happening in a picture. This empowers the descriptor to find out about the dispersion of each hue in a picture. The element vector is taken by connecting the mean each hue. For instance, on the off chance that a histogram of 8-canisters per channel is mulled over, at that point the subsequent element vector will be of  $8 \times 8 \times 8 = 512$ -dfeature vector. Notwithstanding it, straightforward shading Channel measurements, for example, mean and standard deviation could likewise be determined to discover the shading circulation in a picture.

Color attributes of a picture alone are not adequate to measure blossoms on the grounds that in a multi-animal categories condition, at least two species could be of same shading. For instance, Sunflower and Daffodil will have comparative shading content.

➤ *Texture*

Another significant component to be considered to perceive flower species is the "textute" or consistency of examples and hues in a picture. Haralick surfaces, which utilizes the idea of Gray Level Co-occurrence Matrix (GLCM) is generally utilized as the surface descriptor. Haralick et al. [11] portrayed about the 14 measurable highlights that could be processed to evaluate the picture dependent on surface. The subsequent component vector will be 13-d highlight vector overlooking the fourteenth measurement because of high computational time prerequisite.

➤ *Shape*

With regards to common items, for example, blossoms, plants, trees and so forth another significant component to measure such articles in a picture is the "shape". In PC vision research, Hu minutes and Zernike minutes are the two generally utilized worldwide shape descriptors. Minutes depend on the factual desires for an irregular variable. There are seven such minutes which are all things considered called as Hu minutes. First second is the mean, trailed by fluctuation, standard deviation, slant,

kurtosis and other factual boundaries. These seven minutes are joined to frame an element vector of size 7-d. Zernike minutes depend on symmetrical capacities and it was presented by Teague as a shape descriptor. Like Hu minutes, Zernike minutes are likewise used to measure the state of an article in the picture.

➤ *Deep Neural Nets and Convolutional Neural Nets*

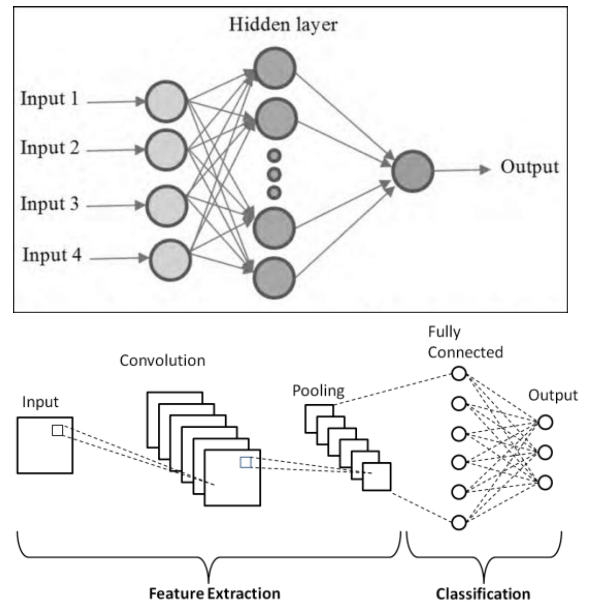


Fig 2:- Artificial Neural Networks (ANN) (top), Convolutional Neural Networks (CNN) (bottom)

In Deep Learning, CNNs are normally applied for Computer Vision applications that include Image Classification and Object Recognition. Flower Species Recognition is a blend of both Object Recognition and Image Classification, as the framework must distinguish a flower in the picture just as perceive which species it has a place with. To perceive the flower species, a keen framework must be prepared with bigger arrangement of pictures, so it could foresee the flower species from its scholarly examples. We use CNNs and max pooling and fully connected layers respectively in order to predict the flower class. We use VGG19 net, which has pre-trained CNNs, padding and pooling layers embedded in the architecture which is applied to the dataset after the segmentation part, this increases the accuracy of the model. We use softmax layer of 102 species.

**V. DATASET AND EXPERIMENTAL PROCEDURE**

In this paper we propose a dataset of 350MB size consisting of 102 flower classes. These are chosen to be flowers usually present in the United Kingdom. Many images were gathered from the internet. A small number of images were included by taking the pictures ourselves. Each class contains a range of images between 40 and 250. Passion flower has the maximum number of images and customs, Mexican aster, celosia, moon orchid, Canterbury bells and primrose have the minimum number of images,

i.e. 40 per class. The images are scaled again and again so that the smallest dimension is 500 pixels. The dataset is divided into a training set, a validation set and a test set. The database can be found at the following URL: <http://www.robots.ox.ac.uk/~vgg/data/flowers/>. The dataset is divided into a training set, a validation set and a test set. The training set and validation set each consist of 10 images per class (totaling 1020 images each). The test set consists of the remaining 6149 images.

**VI. TEST SET RESULT**

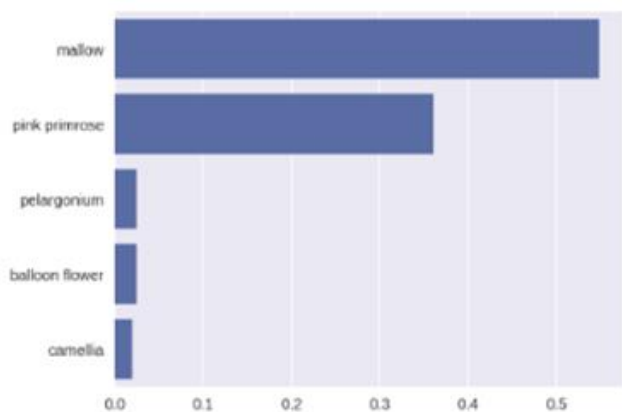
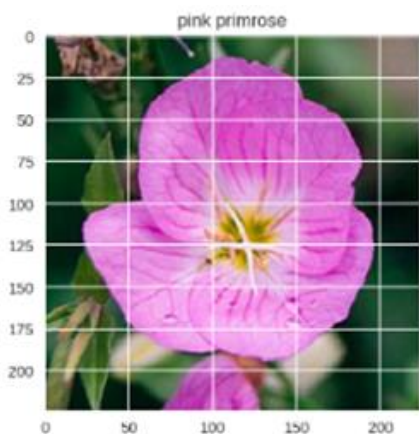


Fig 3:- Output

Depicts the categorization performance for different feature sets. It can be seen that merging all the features results in a far enhanced performance than us and we obtain the accuracy of 86.6%. And we also misclassify the pink primrose class, we predict top 5 classes. Oxford 102 flowers dataset is a settled dataset for sub-classification acknowledgment presented by Nilsback and Zisserman [11]. The dataset contains 102 types of flower and a sum of 8189 pictures, every class consisting somewhere in the range of 40 and 200 pictures. It has set up conventions for preparing and analysis, which we have embraced in our paper as well. Many number of techniques have been tried-and-tested on this dataset [2, 8, 11, 12], including some division based [2, 11]. A portion of the division strategies are intended to be quite certain to the presence of blossoms [11] (with the suspicion that a solitary flower is in the focal point of the picture and takes the greater part of the picture), while others [2] are increasingly broad and can

likewise be connected to different sorts of datasets. Our technique is nearer to the last since we are proposing a general strategy that does not make suspicions about the arrangement of classifications for order or the underlying area or dimension of the items in the picture. In Section 6.2 we experiment the calculation on a lot bigger and progressively differing dataset, in which the flowers are not really in the focal point of the picture, can contain numerous little group blossoms, can fluctuate in scale, and have much more inside class changeability (Figure 1). The execution of our proposed methodology on the given dataset (see Table 1) is 86.66% which beats all recently known techniques in the writing (a few by as much as 4 to 8%) [2, 8, 11, 12].

Method	Accuracy
Our line of reference (No subdivision)	76.7
Chai et. al., Bicos Method [2]	74.8
Nilsback and Zisserman [11]	72.8
Nilsback and Zisserman [12]	76.3
Ito and Cubota [8]	79.4
Angelova , Zhu and Lin [17]	80.66
Ours	86.66

Table 1

**VII. SUMMARY AND FUTURE WORK**

We present an innovative classification algorithm which is strong and scalable to various object forms and backgrounds. Our algorithm utilizes learning to direct the segmentation procedure and is based on the perception that identifying (even imperfectly) some areas of the object can help define its borders and thus section the essential object of interest. We present that the suggested segmentation of objects is very helpful for identification by improving the categorization performance on the Oxford 102 flowers dataset [11]. The enhancements in execution are about 4% for both datasets and are due to the programmed subdivision done at execution time. This is essential as the large-scale datasets contain hundreds of thousands of sample images and no labor-intensive subdivision for them is possible. The algorithm also enhances all other known standard outcomes on the Oxford 102 flower dataset. Our algorithm is less complex and quicker, when it comes to execution, than previously utilized segmentation techniques in similar circumstances, e.g. [3, 14]. It is also more common and vague to the appearance of flowers, so it can theoretically be applied to other kinds of classifications in natural sample images. We can apply Resnet152 neural net for better accuracy and also resnet18 models with the use of parameterized loss functions which may help increase the efficiency of the problem.

**REFERENCES**

- [1]. B. Alexe, T. Deselaers, and V. Ferrari. Classcut for unsupervised class segmentation. ECCV, 2010.
- [2]. Y. Chai, V. Lempitsky, and A. Zisserman. Bicos: A bi-level co-segmentation method for image classification. ICCV, 2011.
- [3]. G. Csurka and F. Perronnin. An efficient approach to semantic segmentation. IJCV, 2011.
- [4]. N. Dalal and B. Triggs. Histograms of oriented gradients for human detection. CVPR, 2005.
- [5]. R.-E. Fan, K.-W. Chang, C.-J. Hsieh, X.-R. Wang, and C.-J. Lin. Liblinear: A library for large linear classification. Journal of Machine Learning Research, 2008.
- [6]. R. Farrell, O. Oza, N. Zhang, V. Morariu, T. Darrell, and L. Davis. Birdlets: Subordinate categorization using volumetric primitives and pose-normalized appearance. ICCV, 2011.
- [7]. P. Felzenszwalb and D. Huttenlocher. Efficient graph-based image segmentation. IJCV, 2004. [8] S. Ito and S. Kubota. Object classification using heterogeneous co-occurrence features. ECCV, 2010.
- [8]. A. Joulin, F. Bach, and J. Ponce. Discriminative clustering for image co-segmentation. CVPR, 2010.
- [9]. Y. Lin, F. Lv, S. Zhu, M. Yang, T. Cour, K. Yu, L. Cao, and T. Huang. Large-scale image classification: fast feature extraction and svm training. CVPR, 2011.
- [10]. M.-E. Nilsback and A. Zisserman. A. automated flower classification over a large number of classes. ICVGIP, 2008.
- [11]. M.-E. Nilsback and A. Zisserman. An automatic visual flora - segmentation and classification of flower images. DPhil Thesis, University of Oxford, UK, 2009.
- [12]. O. Parkhi, A. Vedaldi, C. V. Jawahar, and A. Zisserman. The truth about cats and dogs. ICCV, 2011.
- [13]. C. Rother, V. Kolmogorov, and A. Blake. Grabcut: Interactive foreground extraction using iterated graph cuts. ACM Trans. Graphics, 2004.
- [14]. J. Wang, J. Yang, K. Yu, F. Lv, T. Huang, and Y. Gong. Locality-constrained linear coding for image classification. CVPR, 2010.
- [15]. D. Zhou, O. Bousquet, T. Lal, J. Weston, and B. Scholkopf. Learning with local and global consistency. NIPS, 2004.2013
- [16]. A. Angelova, S. Zhu and Y. Lin. Image segmentation for large-scale subcategory flower recognition. IEEE Workshop on Applications of Computer Vision (WACV)