IJARSCT

ISSN (Online) 2581-9429

International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)

Impact Factor: 6.252

Volume 2, Issue 2, March 2022

A Review Paper on Visual Vocabularies for Image Flower Classification

Rachana C[1] and Vasudev Shahapur[2]

Student, Department of Computer Science and Engineering[1]

Senior Associate Professor, Department of Computer Science and Engineering[2]

Alva's Institute of Engineering and Technology, Tenkamijar, Karnataka, India

Abstract: *Detecting the existence of objects in photos is a difficult challenge for a machine, but it has improved dramatically in recent years. In particular, a text analytics representation known as bag-of-words has gained a lot of traction and has been successfully applied to a variety of problems in the flowers domain. Local floral features are computed in the first step of this approach, then clustered into K visual words, and ultimately each flower is represented as a K-dimensional histogram over the visual words. A classifier uses this fixed-size vector representation of flowers as input. The major purpose of this project is to look into ways to improve the visual vocabulary (the collection of all visual words) in the bag-of-words method. As a result of the aggregation and fast parallel processing of tiny class-specific vocabularies, this approach allows for the construction of huge vocabularies relatively quickly.*

Keywords: Visual Vocabularies

I. INTRODUCTION

In computer vision, automatic plant classification is a hot topic. Typical methods include consulting catalogues to determine the plant's species. However, because of the vast amount of data that must be processed, they are often difficult to use. Furthermore, they are described using a botanical vocabulary that is difficult to comprehend even for experts. Content-based picture indexing approaches may now be used to evaluate and characterise photographs based on their visual content, thanks to technical advancements. These strategies can give the necessary tools for describing the visual appearance of plants, such as colour, shape, and texture aspects. There have been a number of earlier studies on plant image categorization.

Some of them were primarily concerned with leaf classification. The authors of (Krishna Singh, 2010) extracted twelve morphological features (leaf perimeter, aspect ratio, rectangularity, etc.) to represent the shape of the leaf and used and compared three plant classification techniques: Binary SVM Decision Tree (SVM-BDT), probabilistic neural networks (PNN), and Fourier moment technique. Kadir also suggested a leaf categorization approach. They start by employing an adaptive thresholding method to distinguish the leaf from its background. Then they extract characteristics such as leaf shape, colour, venation, and texture. Finally, they used the PNN to classify leaf images.

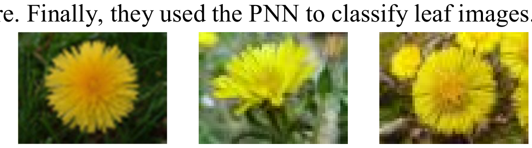

Figure 1: There are three photographs in this set, each from a different category. Dandelions are depicted on both the left and right sides. A colt's foot is in the middle. The intra-class variation between the two dandelions photos is bigger than the inter-class variation between the left dandelion and the image of the colts' foot.

The classification of floral images is the focus of this paper. It's a form. Attribute features are the outputs of these classifiers. Then, for flower retrieval, we use distances between attribute characteristics. The 17 Category Flower Dataset was used to test this approach. In terms of mean average precision, experimental data suggest that attribute-based representation outperforms low-level features. Because natural items are slightly different for each individual, it is impossible to identify them with a single feature, object classification for categories with a considerable visual resemblance is a tough challenge. As a result, numerous characteristics are required to classify them.

IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
Impact Factor: 6.252
Volume 2, Issue 2, March 2022

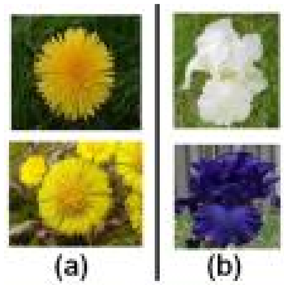

Fig. 2:(a) Two visually dissimilar flowers from the same species. (b) Two visually similar flowers from different species.

MKL has recently gotten a lot of attention as a method of combining multiple features. We use colour, shape, and texture aspects in this study. We use MKL to classify the floral photos and look at the recognition rate. categorise things based on multiple distinguishing characteristics Color, form, and texture were the ones they used. Each feature goes through a separate procedure, but they all end up as vectors. The vector is combined (normalised) at the end of the process, and nearest neighbour is employed for classification.

II. Proposed Methods

The proposed method for flower image classification is divided into two phases: training and testing. The goal of the training phase is to create a model from a subset of images known as training images. Those images must first be segmented and the features extracted. Then, for each characteristic, a visual vocabulary is computed. Finally, we compute a histogram for each image in the training set, counting the occurrence of each visual vocabulary word. These histograms are fed into the SVM classifier as input. The purpose of the testing phase is to determine the class of the flower contained in an image using the generated model and the histogram of visual words. The many steps of our suggested schema are described in depth in following subsections.

2.1 Segmentation

In the case of image analysis, segmentation is a crucial step. Flowers, in general, live in similar environments, and as a result, their backgrounds are frequently similar. To improve categorization, the segmentation seeks to separate the region containing the flower (foreground) from the backdrop. Several papers (Nilsback and Zisserman, 2007) and (Najjar and Zagrouba, 2012) have proposed segmentation algorithms in the literature. Because the authors' method was tested and validated on the same flower dataset that we would use to evaluate our classification schema, we adopted their segmentation results (Najjar and Zagrouba, 2012). On Lab colour space, the proposed segmentation approach is achieved using the OTSU thresholding technique.

2.2. Feature Extraction

Flowers from the same species might seem very different, yet flowers from different species can also look very similar. Furthermore, certain flowers may be identified by their colour, while others have a particular texture or shape. The most difficult aspect of classification is identifying appropriate attributes to characterise the visual content of a flower image and developing a classifier that can distinguish across species. Three characteristics are used to represent different aspects of the flower in this paper: Both the foreground portion of the flower and its boundary, as well as the Lab values, were sampled using SURF.

SURF (Bay et al., 2008) is an interest point detector and descriptor. First, it detect interest points based on approximation of the Hessian matrix determinant. Then, around each interest point, a window is divided into 16 sub-regions a calculated from each sub-region using the integral images. The resulting SURF descriptor is a vector of length 64 describing the neighborhood intensity

IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
Impact Factor: 6.252
Volume 2, Issue 2, March 2022

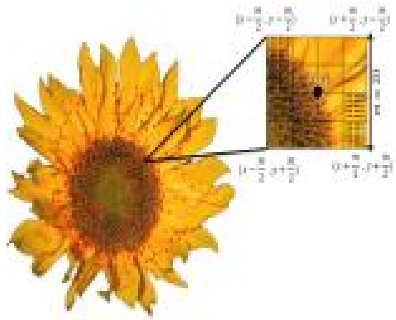

Figure 3: Color feature extraction.

Comparison between SURF and SIFT: SURF is inspired by the SIFT descriptor (Low, 2004) It is, however, faster and more resilient to picture modifications than SIFT. (Juan and Gwun, 2009) (Bay et al., 2008).Although SIFT features performed well in many applications such as object recognition, it has a high computation cost. SURF uses integral pictures to detect and describe interest points in order to reduce feature calculation time, moreover, SURF is smaller than SIFT since it only utilises 64 bits of information to represent the interest point, whereas SIFT uses 128 bits.

For those reasons, we chose the SURF to extract features from flower images. In fact, the set of SURF interest points detected in the image is divided into two subsets: the first one, denoted E SURF includes the interest points sampled on the foreground region. The second subset, denoted E SURF contains the interest points computed on the boundary of the flower. SURF on the Foreground Region: By computing SURF features over the foreground flower region, we can describe not only the local shape of the flower (for example thin petal structure, flower corolla), but also its texture. SURF on the Foreground Boundary: Flowers can deform in different ways, and consequently the difficulty of describing the flower shape is increased by its natural deformations. Also, the petals are often flexible and can twist, bend, which changes the appearance of the flower shape. By computing SURF features on this area, we give more emphasis to the local shape of the flower boundary. In fact, to extract the boundary from an image, we converted it, first, into binary image. Then, we perform erosion operation. Finally, we subtract the binary image from the eroded one and the boundary is extracted.

III. Overview and Performance Measure

In the rest of this paper we develop a nearest neighbour classifier. The classifier involves a number of stages, starting with representing the three aspects as histograms of occurrences of visual words (a separate vocabulary is developed for each aspect) then combining the histograms (section 3) into a single vocabulary. SIFT descriptors on a regular grid are used to describe shape, HSV-values to describe colour, and MR-filters to describe texture. Each is vector quantized to provide the visual words for that aspect. Each stage is separately optimized.

Since we are mainly interested in being able to retrieve a short list of correct matches we optimize a performance cost to reflect this. Given a test image Itest i , the classifier returns a ranked list of training images Itrain j , j = 1, 2, ..., M with j = 1 being the highest ranked. Suppose the highest ranked correct classification is at j = p, then the performance score for Itest i is wp if p ≤ S 0 otherwise where S is the length of the shortlist (here S = 5), and wi is a weight which can be chosen to penalize lower ranks. If wi = 1 ∀ i then the rank of the correctly classified image in the shortlist is irrelevant. We use a gentle fall off, of the form wi = 100 − 20 i−1 S−1 , so that higher ranked images are rewarded slightly (w1 = 100, w5 = 80 for S = 5).

Suppose the classifier is specified by a set of parameters θ, then the performance score over all test images is: f(θ) = 1.In essence, this is our utility/loss function, and we seek to maximize f(θ) over θ. This optimization is carried out over a validation set in each of the following classification sections. The performance of the developed classifier is compared to that of a baseline algorithm using colour histograms.

3.1 Datasets

There are species that have a very unique visual appearance, for example fritillaries and tigerlilies, as well as species with very similar appearance, for example dandelions and colts feet. There are large viewpoint, scale, and illumination variations. The large intra-class variability and the sometimes small inter-class variability makes this dataset very challenging. The flower categories are deliberately chosen to have some ambiguity on each aspect. For example, some classes cannot be distinguished on colour alone (e.g. dandelion and buttercup), others cannot be distinguished on shape alone (e.g. daffodils and windflower). The flower images were retrieved from various websites, with some supplementary images from our own photographs. Consistent viewpoint set: For the running example of the various stages of the classifier we do not use the full dataset, but instead consider only a subset. This consists of 10 species with 40 images of each. For each class the 40 images selected are somewhat easier than those of the full set, e.g. the flowers occupy more of the foreground or are orientated in a more consistent pose. We randomly select 3 splits into 20 training, 10 validation and 10 test images. The parameters are optimized on the validation set and tested on the test set. All images are resized so that the smallest dimension is 500 pixels. Both the full and consistent viewpoint sets are available

3.2 Creating a Flower Vocabular

Like botanists we need to be able to answer certain questions in order to classify flowers correctly. The more similar the flowers, the more questions that need to be answered. The flowering parts of a flower can be either petals, tepals or sepals. For simplicity we will refer to these as petals. The petals give crucial information about the species of a flower. Some flowers have petals with very distinctive shape, some have very distinctive colour, some have very characteristic texture patterns, and some are characterized by a combination of these properties. We want to create a vocabulary that gives an accurate representation of each of these properties.

Flowers in images are often surrounded by greenery in the background. Hence, the background regions in images of two different flowers can be very similar. In order to avoid matching the green background region, rather than the desired foreground region, the image is segmented. The foreground and background RGB colour distributions are determined by labelling pixels in a subset of the training images as foreground (i.e. part of the flower), or background (i.e. part of the greenery). Given these foreground and background distributions, all images are automatically binary segmented using the contrast dependent prior MRF cost function of [2], optimized using graph cuts. Note, these distributions are common across all categories, rather than being particular to a species or image. This procedure produces clean segmentations in most cases. For the vocabulary optimization for colour and shape we compare the performance for both segmented and non-segmented images. The top row shows the original images and the bottom the segmentation obtained. The flowers in the first and third column are almost perfectly segmented out from the background greenery. The middle column shows an example where part of the flower is missing from the segmentation –problem occurs in less than 6% of the images.

A. Colour Vocabulary

We want to create a vocabulary to represent the colour of a flower. Some flowers exist in a wide variety of colours, but many have a distinctive colour. The colour of a flower can help narrow down the possible species, but it will not enable us to determine the exact species of the flower. For example, if a flower is yellow then it could be a daffodil or a dandelion, but it could not be a bluebell.

Figure 4: Segmented images. The top row shows the original im ages and the bottom the segmentation obtained. The flowers in the first and third column are almost perfectly segmented out from the background greenery. The middle column shows an example where part of the flower is missing from the segmentation – this problem occurs in less than 6% of the images.

Images of flowers are often taken in natural outdoor scenes where the lighting varies with the weather and time of day. In addition, flowers are often more or less transparent, and specular highlights can make the flower appear lighter or even white. These environmental factors cause large variations in the measured colour, which in turn leads to confusion between classes.

One way to reduce the effect of illumination variations is to use a colour space which is less sensitive to it. Hence, we describe the colour using the HSV colour space. In order to obtain a good generalization, the HSV values for each pixel in the training images are clustered using k-means clustering. The results shown are averaged over three random permutation of training, validation and test sets. Best results are obtained with non segmented images and 200 clusters, although the performance does not change much with the number of clusters, represented by a Vc dimensional normalized frequency histogram n(wc|Ij). A novel test image is classified using a nearest neighbour classifier on the frequency histograms. Results are presented for both segmented and non-segmented images. Perhaps surprisingly, the non-segmented images show better performance. This is because members of a flower species usually exist in similar habitats, thus making the background similar, and positively supporting the classification of the non-segmented images. However, in the full data set (as opposed to the rather restricted development set) this is not always the case and it is therefore better to segment the images. The best result using the segmented images is obtained with 500 clusters. The overall recognition rate is 55.3% for the first hypothesis and 84.3% for the fifth hypothesis (i.e. the flower is deemed correctly classified if one of the images in the top five retrieved has the correct classification).

B. Shape Vocabulary

Individual petal shapes, as well as their configuration and the overall shape of the flower, can all be utilised to differentiate between flowers. Figure 5 shows that, while the windflower (left) and buttercup (centre) have similar general shapes, the windflower's petals are more pointed. The petals of the daffodil (right) are more similar to those of the windflower, but the overall form is significantly different due to the tubular shape of the daffodil's corolla in the middle. The natural deformations of a flower make it more difficult to describe the shape. The petals are often exceedingly soft and flexible, allowing them to bend, curl, twist, and so on, resulting in a variety of floral shapes. The shape of a flower changes as it gets older, and the petals may even fall off. For these reasons, the form representation has been designed to be redundant; for example, rather than expressing each petal only once, each petal is represented by numerous visual words (attempting to count petals). This redundancy protects against misclassification, obscured or missing petals, and other issues.

In the same way, we want to describe the shape of each petal of a flower. As a result, we require a rotation invariant descriptor. On a normal grid, we compute SIFT descriptors and optimise three parameters: grid spacing M, which ranges from 10 to 70 pixels; the support region for SIFT computation, which has a radius R of 10 to 70 pixels; and finally, the number of clusters. We obtain the images by vector quantization and classify them similarly to the colour features.

The best performance for the segmented images is obtained with 1000 words, a 25 pixel radius and a stepsize of 20 pixels. Note that the performance is highly dependent on the radius of the descriptor. The recognition rate for the first hypothesis is 82.7% and for the fifth hypothesis is 98.3%. This intra-image grouping has some similarities to the Epitome representation of Jojic et al. [8], where an image is represented by a set of overlapping patches.

C. Texture Vocabulary

On the petals of some flowers, there are distinctive patterns. These patterns can be more recognisable, such as the stripes of the pansy, the checks of the fritillary, or the dots of the tiger-lily. The same word is represented by all circles of the same colour. Petal intersections are represented by the blue/dashed word, whereas rounded petal ends are represented by the red word. It's worth noting that the words may detect comparable petal sections inside the same image as well as different flowers (intra-image grouping). Due to lighting circumstances, faint patterns might be difficult to detect - a problem that also impacts the appearance of more distinct patterns.

The filter bank includes filters in a variety of orientations. The maximal response across orientations is chosen to provide rotation invariance. We optimise over filters with s = 3 19 pixel square support regions. The frequency histograms n are obtained after clustering the descriptors and creating a lexicon. The classification is done in the same way that the colour features are classified. The best results are obtained with 700 clusters and a size 11 filter. The first hypothesis is recognised 56.0 percent of the time, whereas the fifth hypothesis is recognised 84.3 percent of the time.

3.3 Combined Vocabulary

For different species, the discriminative power of each aspect varies. The confusion matrices for the different aspects of the consistent viewpoint set are shown in Table 1. Not surprisingly, it demonstrates that some flowers may be differentiated by shape, such as daisies, by colour, such as fritillaries, and by texture, such as colts' feet and fritillaries. It also demonstrates that certain elements of particular flowers are too close, such as the colour difference between buttercups and daffodils, the form difference between colts' feet and dandelions, and the texture difference between buttercups and irises. It is possible to achieve improved performance by combining the various aspects in a flexible manner. To create a joint frequency histogram n(w|I), we combine the vocabularies for each facet into a single flower vocabulary. However, we have some leeway because we don't have to assign each feature equal weight - for example, if one aspect contained many more words than the other, the one with more words would, on average, dominate the distance in nearest neighbour comparisons. Due to the normalisation of the final histogram, there are only two independent parameters that reflect two of the ratios in s: c: t. By maximising the performance score, here f(), on the validation set, we learn the weights, on the consistent viewpoint set. The test set is used to evaluate the performance.

IV. Conclusion

We may have handled the challenge of flower classification by creating flow specialized descriptors with specially written descriptors. Indeed, descriptors for categorising based on scanned leaf form have already been created. We have proven that more general purpose descriptors are sufficient - at least for a database of this complexity level. Tuning the vocabulary and merging vocabularies for various features resulted in a considerable performance improvement, with the final classifier outperforming each of the individual ones. The main issues today are dealing with considerable scale variations and a varied number of occurrences, where a test image could have a single flower or 10 or more.

References

[1]. Abdul Kadir, Lukito Edi Nugroho, A. S. P. I. S. (2011). Leaf classification using shape, color, and texture features. International Journal of Computer Trends and Technology, 2:225–230.

[2]. Bay, H., Ess, A., Tuytelaars, T., and Gool, L. V. (2008).Surf: Speeded up robust features. Computer Vision and Image Understanding (CVIU), 110:346–359.

[3]. Chai, Y., Lempitsky, V., and Zisserman, A. (2011). Bicos: A bi-level co-segmentation method for image classification. In IEEE International Conference on Computer Vision, pages 2579–2586.

[4]. Guru, D. S., Sharath, Y. H., and Manjunath, S. (2010). Texture features and knn in classification of flower images. IJCA,Special Issue on RTIPPR, pages 21–29.

[5]. Juan, L. and Gwun, O. (2009). A comparison of sift, pca-sift and surf. International Journal of Image Processing,3:143–152.

[6]. Krishna Singh, Indra Gupta, S. G. (2010). Svm-bdt pnn and fourier moment technique for of leaf shape. International Journal of Signal Processing, Image Processing and Pattern Recognition, 3:67–78.

[7]. Louradour, J., Daoudi, K., and Bach, F. (2007). Feature space mahalanobis sequence kernels: Application to svm speaker verification. IEEE Transactions on Audio, Speech and Language Processing, 15:2465–2475.

[8]. Low, D. (2004). Distinctive image features from scale invariant keypoints. International Journal of Computer Vision, 60:91–110.

[9]. Najjar, A. and Zagrouba, E. (2012). Flower image segmentation based on color analysis and a supervised evaluation. In International Conference on Communications and Information Technology (ICCIT), volume 2, pages 397–401.