Plant species recognition using Bag-Of-Word with SVM classifier in the context of the LifeCLEF challenge

Issolah Mohamed, Lingrand Diane, and Precioso Frédéric

Univ. Nice Sophia Antipolis
Laboratory I3S, UMR 7271 UNS-CNRS
06900 Sophia Antipolis, France
`isso.moh@gmail.com`,
`precioso@unice.fr`,
`lingrand@i3s.unice.fr`

Abstract. For the plant task of the LifeCLEF challenge, we adopted the reference Bag-of-Word framework (BoW) with local soft assignment. The points of interests (POI) are both detected and described with the SIFT and OpponentColor SIFT. Parameters of the bag of word are optimized through cross-validation and we present the results of different experimentations. A Support Vector Machine is trained with different strategies according to the organs and species of plants.

1 Method

For this 2014 participation in the LifeCLEF challenge [4] and more specifically to the plant identification task [3], we build an image processing chain based on the reference Bag-of-Word framework (BoW). We study the results obtained with this framework when optimizing parameters with respect to the different organs. The goal of the plant identification task is to determine plant species from *plant observations* that may consist on one or more images and associated meta-data.

As a first step, we focus on plant species recognition from a single image with the organ type as metadata. We have considered images of each organ separately. There are 7 categories of organs to be considered this year: leaf with natural background, leaf with uniform background, flower, fruit, branch, stem and entire. From these categories of organs, we build 7 quite similar but independent processing chains.

We extract Points-of-Interest (PoI) using the SIFT detector in every image and describe each local feature with the SIFT descriptor or Opponent Color SIFT descriptor [5]. The visual dictionary is built with a K-means algorithm on the local features. Each image is then represented by its histogram onto the dictionary using local soft assignment strategy [1, 2]. We classify the images with as many binary one-against-all Support Vector Machines as the number of plant

classes per organ types. Considering 7 categories of organs by 500 species leads to almost 3500 SVMs (not all organs are available for every species).

This year participation differs from our previous one by using local soft assignment instead of hard assignment, color processing for flowers and the optimization of parameters in the clustering and SVM classification:

1. The number of cluster for the K-means.
2. C which represents the sum of error distances for SVM.

We now detail the different steps and discuss the results of experimentations.

1.1 Feature Extraction and Image Description

For each images, Points Of Interest (POIs) are extracted using the SIFT detector. They are described using Opponent Color SIFT for flower and standard SIFT for other organs.

About 1000 points are extracted in each images, with standard settings:

1. Number of Layers per Octave $= 3$
2. The minimum threshold to consider a point as POI $= 0.04$
3. Sigma of Gaussian $= 1.6$

From these POIs, visual dictionaries (one specific per organ) are computed using a K-means algorithm. K, the number of clusters, is cross-validated to be set to different values: 4000 for leaf with uniform background, 2000 for leaf with natural background and 500 for the other organs.

Finally, image features are encoded with a local soft assignment onto the dictionary: each local features is participating for its 5 nearest clusters in the BOW.

1.2 Training

For each category (i.e. each organ type), linear binary Support Vector Machine (SVM) are learned on training data, in a one-against-all strategy, in order to predict the different plant species. The C parameter is configured according to the cross-validation results: 100 for leaf (uniform and natural background) and 0.5 for the other categories of organs.

The SVMs are organized into 7 vectors according to the categories of organs (see equation 1). Since images for the 7 categories of organs are not available for all species, size of SVMs could be less than 500.

$$\begin{pmatrix} class_{o_1 1} \\ class_{o_1 2} \\ class_{o_1 3} \\ \vdots \\ class_{o_1 n_1} \end{pmatrix} \begin{pmatrix} class_{o_2 1} \\ class_{o_2 2} \\ class_{o_2 3} \\ \vdots \\ class_{o_2 n_2} \end{pmatrix} \dots \begin{pmatrix} class_{o_7 1} \\ class_{o_7 2} \\ class_{o_7 3} \\ \vdots \\ class_{o_7 n_7} \end{pmatrix} \qquad (1)$$

1.3   Query Image

When all SVMs are computed on the training data, each test image has to be analyzed with the following steps:

1. get the category of organ from the XML file,
2. extract and describe points of interest with SIFT or OpponentColor SIFT, depending on the organ category,
3. generate the BOW using the vocabulary specific to the considered organ,
4. test all the SVM(s) of this organ (≤ 500) and get a list of confidence on the prediction of species.

1.4   Generation of runs

The confidence scores (d) obtained for the test images represent the distance of the vector to the margin corresponding to each SVM. Thus, it is not possible to compare scores from different SVMs. In order to overcome this problem, confidence scores of each species are normalized in order to be compared each other. This step will project each confidence in the interval of [0,1].

All confidence scores concerning the same observation plant are gathered. Let denote S_d this set of confidence scores. All values are normalized using:

$$Sn = \{dno_n \mid dno_n = \frac{d - Min_{S_d}}{Max_{S_d} - Min_{S_d}}\} \tag{2}$$

with:
S_n: new set of confidence normalized of the nth observation plant.
$dnor_n$: normalized confidence score.
d: confidence score obtained by the SVM.
Min_{S_d}: minimum value in S_d.
Max_{S_d}: maximum value in S_d.

After the normalization, confidence scores obtained for all single images corresponding to the same observation plant are merged to generate the final run. Two ways of merging have been tested:

run1 : confidence scores S_n have been sorted in descending order. Scores of the same class are summed up.
run2 : confidence scores S_n have been sorted in descending order and keep the uttermost.

Merging confidence scores obtained with individual images for observation plant increases the whole score.

2 Experiments for the Optimization

In order to tune our process, different experiments have been done on the 2013 ImageCLEF challenge dataset. First of all, parameters have been optimized: number of clusters K for the K-means clustering algorithm and C for SVM.

This study has been focused on leaf with uniform or natural background and flower. Different values of the number of clusters K (100, 200, 500, 1000, 2000, 4000) and C (0.1, 0.2, 0.5, 1, 2, 5, 10, 20, 50, 100) have been tested and chosen with respect to the maximal value of score, computed from 2013 ImageCLEF challenge rules. Comparison has also been made with default values ($K = 100$ and $C = 100$) that were used in our 2013 submissions.

For the experimentations, train data from 2013 ImageCLEF were divided according to table 1 and cross validation was performed.

Table 1. Percentage of training and testing data for each organ

	Leaf with uniform background	Leaf with natural background	Flower
Training data	5061 (52%)	2107 (60%)	2167 (62%)
Test data	4719 (48%)	1396 (40%)	1355 (38%)

We now detail experiments for each organ category.

Tuning K and C for leaf with uniform background For all tested values of K and C, scores have been reported in table 2. Blue colored values correspond to the best scores and the maximum is reached with $K = 4000$ and $C = 100$. Variations of scores in the neighborhood of the best score with respect to K and C have been plotted in figure 1 and show that increasing K and C leads to better results.

Tuning K and C lead to an increase of 27%.

The table 2 suggest that we can have better result if we increase the parameters: C and more specifically K. For this paper, we have tested a pre-defined set of parameters for all organs. High value of K are costly in term of computation, this is why higher than 4000 values of K have not been tested in this paper but need to be in a further study. Refining quantification of parameters in neighborhood of optimal values should also be examined.

Tuning K and C for leaf with natural background Scores are reported in table 3 and the maximal score is reached for $K = 2000$ and $C = 100$. Increasing C value leads to higher score while K should not be increased over 4000 (see figure 2.

Table 2. Tuning K and C for leaf with uniform background

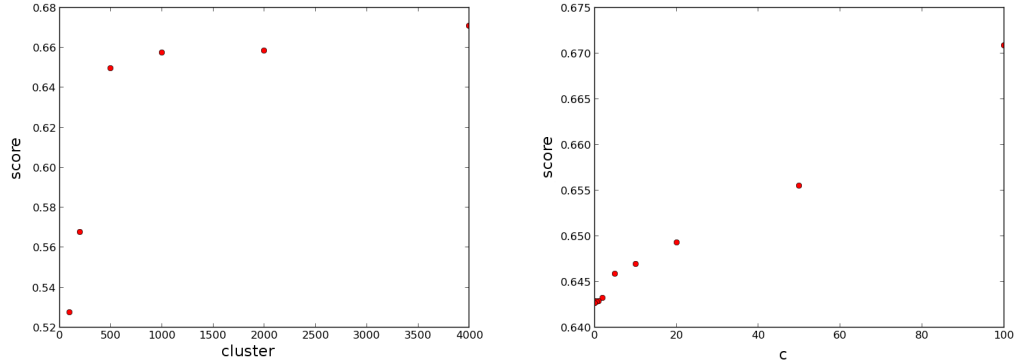| C	number of clusters					
	100	**200**	**500**	**1000**	**2000**	**4000**
0.1	0.512	0.547	0.608	0.617	0.632	0.642
0.2	0.513	0.548	0.608	0.617	0.632	0.642
0.5	0.511	0.547	0.604	0.616	0.632	0.642
1	0.511	0.547	0.604	0.616	0.632	0.642
2	0.511	0.547	0.604	0.617	0.632	0.643
5	0.509	0.546	0.608	0.617	0.633	0.645
10	0.512	0.547	0.605	0.619	0.635	0.646
20	0.512	0.549	0.611	0.627	0.636	0.649
50	0.525	0.563	0.618	0.640	0.645	0.655
100	0.527	0.567	0.649	0.657	0.658	0.670

Fig. 1. Leaf with uniform background. Left: variations of score with $C = 100$; Right: variations of score with $K = 4000$. Increasing C and K increases the score.

Tuning K and C lead to an increase of 76%. However, the maximal score (0.368) is almost half the maximal score of leaf with uniform background. Segmenting the leaf should significantly improve the performances by removing noise introduced by the background.

Tuning K and C for flower Scores are reported in table 4 and present a maximal value for $K = 500$ and $C = 0.5$. Variations of score in the neighborhood of the maximal value are not similar to the ones observed for the 2 categories of organs associated to leafs. Even if the discretization that is performed on K and C parameters may lead to miss global maximal values, we expect to be close enough to this maximal value. Tuning K and C lead to an increase of 43%.

Tuning K and C parameters improves significantly the performance. However, the impact of the C parameter is less important than the K parameter.

Table 3. Tuning K and C for leaf with natural background

C	number of clusters					
	100	200	500	1000	2000	4000
0.1	0.192	0.276	0.303	0.313	0.363	0.334
0.2	0.202	0.282	0.308	0.315	0.365	0.337
0.5	0.202	0.284	0.313	0.315	0.365	0.337
1	0.206	0.287	0.313	0.314	0.365	0.337
2	0.206	0.281	0.314	0.314	0.366	0.338
5	0.206	0.286	0.314	0.314	0.366	0.338
10	0.215	0.279	0.308	0.314	0.366	0.338
20	0.208	0.297	0.308	0.313	0.366	0.337
50	0.192	0.283	0.308	0.309	0.367	0.337
100	0.209	0.294	0.316	0.319	0.368	0.338

Table 4. Tuning K and C for flower

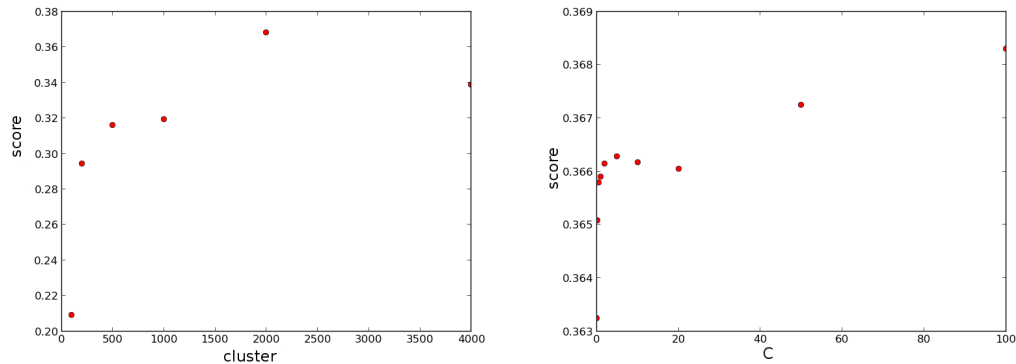| C	number of clusters					
	100	200	500	1000	2000	4000
0.1	0.207	0.282	0.305	0.272	0.300	0.300
0.2	0.216	0.283	0.305	0.273	0.302	0.298
0.5	0.218	0.283	0.312	0.277	0.302	0.298
1	0.209	0.284	0.299	0.275	0.302	0.299
2	0.209	0.284	0.305	0.276	0.301	0.299
5	0.209	0.284	0.305	0.276	0.302	0.299
10	0.208	0.283	0.305	0.276	0.302	0.299
20	0.213	0.284	0.305	0.276	0.301	0.299
50	0.214	0.284	0.305	0.276	0.302	0.299
100	0.219	0.284	0.305	0.276	0.302	0.299

Fig. 2. Leaf with natural background. Left: variations of score with $C = 100$; Right: variations of score with $K = 2000$. Increasing C increases the score but score with respect to K reaches a maximum value for $K = 2000$.

Tuning has be done using a set of predefined values that could be extended and also refined by reducing discretization steps of the different parameters in the neighborhood of optimal values.

Further experiment should be done in order to refine optimal K and C parameters for these organ categories but also for other categories. These experiments are costly in term of computations and have to be planned for a long period of time.

Description of points of interest for flower Two different descriptors of points of interest have been tested for the organ category flower: SIFT and OpponentColor SIFT in order to take the color into account. Using optimal parameters $K = 500$ and $C = 0.5$, the score increases from 0.31 to 0.49 (+58%): not really surprisingly, color has to be taken into account for flowers.

Local soft assignment versus hard assignment Different species may present organs that are visually similar, such as leafs for instance. In order to consider different species for one image to be tested, local soft assignment has been compared on the leaf organ category with uniform background. Optimal parameter were $K = 4000$ and $C = 100$. The score increases from 0.67 to 0.74 (+10%). This assignment has been used on all categories of organs.

3 Result Obtained

Submission has been done with the K and C parameters tuned on 2013 data, SIFT for all categories of organs except flowers (OpponentColor SIFT) and local soft assignment. Training has been done on the 2014 training data. The score
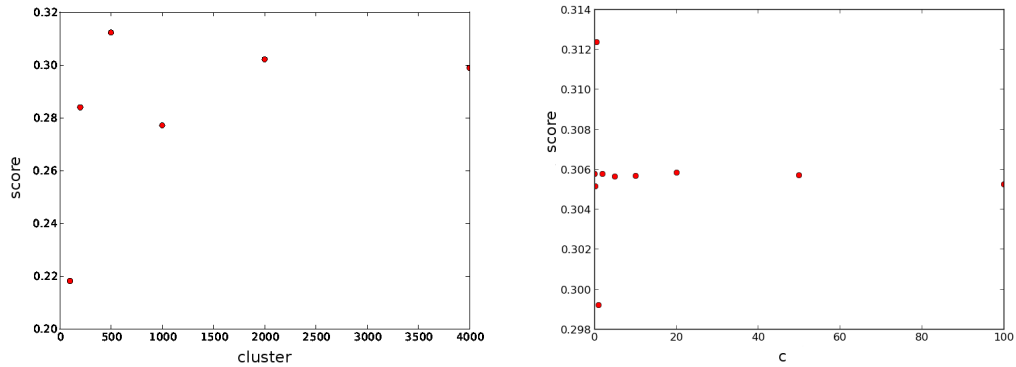
Fig. 3. Flower. Left: variations of score with $C = 0.5$; Right: variation of score with $K = 500$. Profiles of variations are different from the leaf organ.

obtained are: 0.091 for run1 and 0.089 for run2 (see figure 4). Scores for run1 are detailed according to organ categories:

Table 5. Result obtained

branch	entire	flower	fruit	leaf nat.	leaf uniform	stem
0.041	0.023	0.04	0.04	0.035	0.089	0.086

Compared to 2013 results, results of 2014 have been improved. However, we were expecting better results from what we had obtained on 2013 data sets, especially on the leaf and flower categories. Parameters have been tuned and scores have been computed on the 2013 dataset which was smaller with half number of species.

4 Conclusion

Tuning parameters is not an instinctive task and its computation is time consuming. However, it increases a lot the performance of the recognition. Using local hard assignment can be benefit for the problem where we need more discrimination. Further studies will focus on refining our tuning process and taking into account metadata.
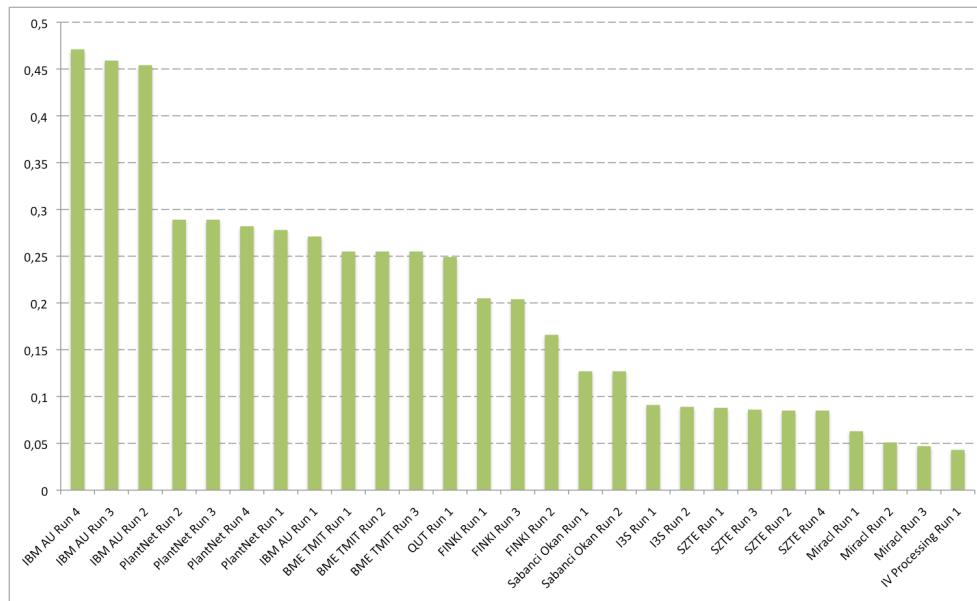
Fig. 4. LifeCLEF 2014: results of the plant task for all participants (results for our team are labeled "I3S Run1" and "I3S Run2"). Results for leafs and flower are below what we were expecting.

References

1. Gemert, J.C., Geusebroek, J.M., Veenman, C.J., Smeulders, A.W.: Kernel code-books for scene categorization. In: Proceedings of the 10th European Conference on Computer Vision: Part III. pp. 696–709. ECCV 2008, Springer-Verlag, Berlin, Heidelberg (2008)
2. van Gemert, J.C., Veenman, C.J., Smeulders, A.W.M., Geusebroek, J.M.: Visual word ambiguity. IEEE Trans. Pattern Anal. Mach. Intell. 32(7), 1271–1283 (Jul 2010)
3. Goëau, H., Joly, A., Bonnet, P., Molino, J.F., Barthélémy, D., Boujemaa, N.: Lifeclef plant identification task 2014. In: CLEF working notes 2014 (2014)
4. Joly, A., Müller, H., Goëau, H., Glotin, H., Spampinato, C., Rauber, A., Bonnet, P., Vellinga, W.P., Fisher, B.: Lifeclef 2014: multimedia life species identification challenges. In: Proceedings of CLEF 2014 (2014)
5. van de Sande, K.E.A., Gevers, T., Snoek, C.G.M.: Evaluation of color descriptors for object and scene recognition. In: Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition (June 2008)