

1 Application of reject option on flower images

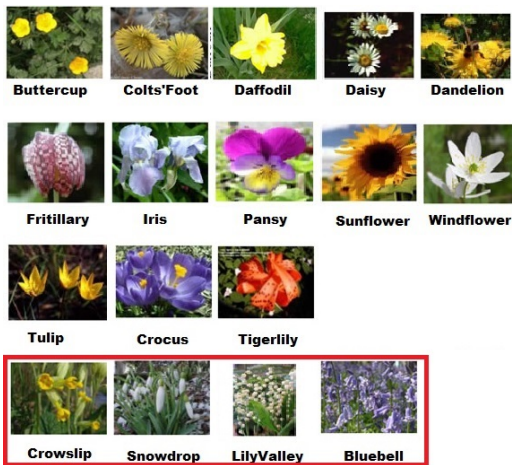


Fig. 1 Flower dataset of common flowers in the UK. Four classes (snowdrops, lily of the valleys, cowslips and bluebells, as marked within the red box) are not segmented due to their tiny size of foreground objects.

We applied the proposed rejection algorithm on a popular dataset: the Oxford flower datasets with 17 classes of common flowers in the UK [1] (as shown in Figure 1). This task is also difficult because the images have large scale, pose and light variations. Some classes are quite similar to others and they both have enormous variations. The authors, Nilsback and Zisserman, used a visual vocabulary method for the flower classification and they produced an accuracy of 81.4% over all samples with 3-fold cross validation. In [2], the authors applied a segmentation algorithm to 13 categories of flowers (753 flower images), while the other 4 classes (snowdrops, lily of the valley, cowslips and bluebells) are omitted because their foreground objects are too tiny for segmentation. We exploited the segmentation results and used the same feature extraction and hierarchical classification of [3]. We used the BGOT method as described in previous sections. We trained a 13-class BGOT tree and it achieves an accuracy score of 83.2%, which is better than a flat SVM with forward sequential feature selection (82.0%). Note, the visual vocabulary method is based on all 17 classes while our training classes cover the 13-class subset which have the segmentation results.

To evaluate the performance of the reject option for the new classes, we choose another 7399 samples of 90 different classes from an extended flower dataset which is provided by the same authors [4]. This dataset consists of 102 categories of flowers and we exclude 12 classes which already exist in the training set. We repeated our

TRs (known class)		TRs (new class)	
rate(%)	number	rate(%)	number
21.5	7	37.4	2764
True Positives		False Rejections	
rate(%)	number	rate(%)	number
83.2	158	4.0	6

Table 1 Rejection performance of classification result, averaged by 3-fold cross validation. (TR=True Rejection)

proposed rejection algorithm after the classification and calculated the *posterior* probability of these results by using a GMM for each class (the same as we did for the fish dataset). Each GMM is trained on a selected subset of features where the feature selection algorithm maximizes the accuracy of classifying the given class from all other classes. The distributions of *posterior* probability of three different groups (True Positives, False Positives, New classes) are shown in Figure 2. We set a small threshold (*i.e.* 0.01) and reject all test samples whose *posterior* probabilities are below the threshold. As a result, the proposed method filters out a significant portion of True Rejections (misclassified samples, either False Positives or samples from new classes, shown as the scores of 21.5% & 37.4%, respectively) with a small cost (4.0%) of the False Rejections (correctly classified samples but falsely rejected), as shown in Table 1. This task is challenging since the trained GMM has no *prior* knowledge about any of the new classes. The proposed rejection method has rejected more than one third (37.4%) of the test samples from the new classes, at the cost of a slightly deduction of accuracy (4% True Positives are falsely rejected).

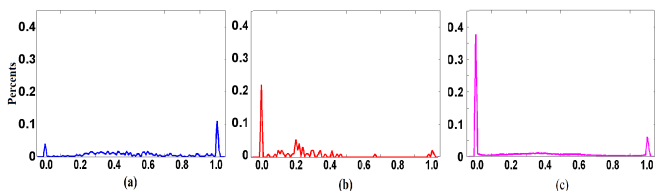


Fig. 2 *Posterior* probability of the samples of True Positives (a), False Positives (b), test samples from new classes (c). The average *posterior* probabilities of both the False Positives and test samples from new classes are lower than the True Positives. We set a small threshold (*i.e.* 0.01) and reject a significant portion of misclassified samples (the rear peaks in b & c).

References

1. M.-E. Nilsback and A. Zisserman, “A visual vocabulary for flower classification,” in *Proceedings of the IEEE Confer-*

- ence on *Computer Vision and Pattern Recognition*, vol. 2, pp. 1447–C1454, 2006.
2. M.-E. Nilsback and A. Zisserman, “Delving into the whorl of flower segmentation,” in *Proceedings of the BMVC*, vol. 1, pp. 570–579, 2007.
 3. P. X. Huang, B. J. Boom, and R. B. Fisher, “Underwater live fish recognition using a balance-guaranteed optimized tree,” in *Computer Vision–ACCV 2012*, pp. 422–433, 2013.
 4. M.-E. Nilsback and A. Zisserman, “Automated flower classification over a large number of classes,” in *Proceedings of the Indian Conference on Computer Vision, Graphics and Image Processing*, 2008.