Label Ranking Based Semantic Texton Forests for Multi-label Problem

Takaaki Fukuta
Yamaha Motor Co., Ltd.
fukutat@yamaha-motor.co.jp

Hironobu Fujiyoshi
Dept. of Computer Science, Chubu Univ.
hf@cs.chubu.ac.jp
http://www.vision.cs.chubu.ac.jp

1. Introduction

When comprehending scenes from images such as indoor and outdoor landscape photos, it is necessary not only to segment them by color, but also to recognize regions of the same category to segment them. This is called semantic segmentation[2]. Since semantic segmentation has the advantage of making it possible to acquire not just a number of object category names within the image but also the positions of objects, it is an important technique in scene comprehension.

This paper proposes a method of semantic segmentation which handles multiple labels, by expanding two previous methods.

2. Proposed Method

The proposed method is based on Semantic Texton Forets(STFs) [2], which represent a conventional method which deals with single labels, expanded in two ways to implement semantic segmentation which handles multiple labels. The first expansion is to construct LR-STFs that perform training so as to enable separation of multiple labels using label ranking in STFs, for feature extraction. The second expansion is to construct multi-label classifier that are decision tree classifiers corresponding to multiple labels, using a label powerset. This makes it possible to consider multiple labels for a target object.

2.1. Label Ranking based Semantic Texton Forests

Label ranking is an approach for training classifiers that assign rankings to single labels, using multi-label training data [3]. To perform label ranking, it is necessary to have processing that disassembles a multi-label into single labels. That is why we focus on the point of employing information gain in the determination of the optimal split node in the STF training method, and similarly making information gain the standard for disassembly of a multiple label into single labels. We prepare labels that are candidates for disassembly at random from within the labels that have been attached to multi-label samples, to ensure that candidates for the feature and threshold used for creating split nodes are prepared at random. If multiple labels $C_i = \{c_1, c_2, ..., c_n\}$ are attached to a sample $i \in I_n$ utilized in a certain split and a candidate label is assumed to be $\mathcal{C}_i$, $\mathcal{C}_i$ is selected by Equation (1)

$$\mathcal{C}_i = \begin{cases} \mathrm{rand}(C_i) & |C_i| > 2 \\ c_1 & \text{otherwise} \end{cases} \quad (C_i \neq \phi). \tag{1}$$

In this case, the function rand selects one of the elements from the collection C_i at random. c_1 denotes a single label and an element is an element from one collection C_i. This candidate selection is done L times, and the candidate label with the highest information gain from within the L candidates becomes the label for the sample i in that split.

By repeating this process, the same number of labels as the number of splits S is determined, and a history is created as follows: $\{\mathcal{C}_i^1, \mathcal{C}_i^2, ..., \mathcal{C}_i^S\}$. We use the thus-selected frequency to compute the final label $\mathcal{C}_i^{max}$ of the sample i. First of all, we obtain the probability $p(c|i)$ from the frequency of classes selected for the sample i by the history. We then obtain the final label $\mathcal{C}_i^{max}$ from the following equation:

$$\mathcal{C}_i^{max} = \arg\max_c p(c|i). \tag{2}$$

From this processing, the multiple labels C_i attached to the sample i become $\mathcal{C}_i^{max}$, which cannot be disassembled back into single labels. Use of this final label $\mathcal{C}_i^{max}$ enables label ranking, by computing the probability $P(c|l)$ of classes of the leaf nodes l.

2.2. Multi-label Classifier construction

We construct multi-label classifier for semantic segmentation called Segmentation Forest in [2] using Label Powerset [1].

Label powerset (LP) is a simple but effective problem transformation method that works as follows: It considers each unique set of labels that exists in a multi-label training set as one of the classes of a new single-label classification task.

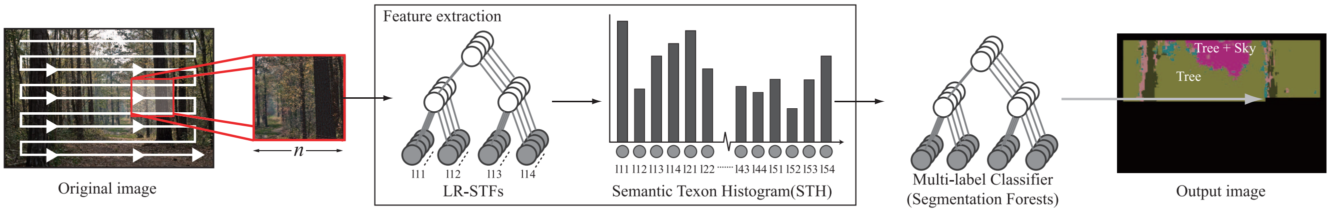
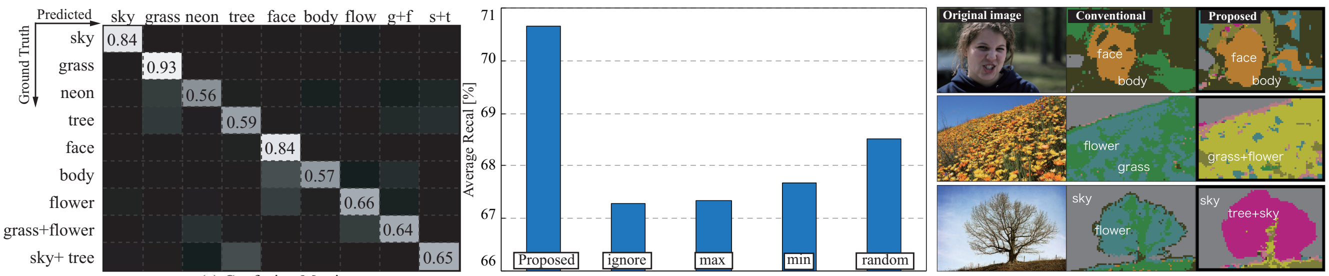


Figure 1. Flow of semantic segmentation



(a) Confusion Matrix (b) Comparison results of label ranking methods (c) Segmentation results

Figure 2. Experimental results

2.3. Semantic segmentation

We performed semantic segmentation of unknown images, using the LR-STFs described in 2.1 and the multi-label classifier described in 2.2. The flow of processing is shown in Figure 1 and described below.

Step1 Set an $n \times n$ ROI on image

Step2 Compute STH[2] by LR-STFs

Step3 Classify a class by multi-label classifier

Step4 Return to Step 1

By performing the above processing on all pixels in the input image, semantic segmentation which handles multiple labels can be implemented.

3. Experiments

We performed evaluation experiments about proposed method, using original dataset.

3.1. Evaluation of multi-label validity

Using the experimental dataset, we performed training in accordance with the proposed method and confirmed the validity of the proposed method. The confusion matrix and segmentation results using the proposed method is shown in Figure 2(a)(c). The paler the diagonal components are, the more correct is the segmentation, and the average recall rate is 70.6%.

From these results, we can say that we have confirmed the validity of the proposed method.

3.2. Evaluation of label ranking method

We performed a comparison with the method of Document [3], to evaluate the validity of the proposed label ranking method. We constructed LR-STFs by using these methods, and compared the average recall rate of semantic segmentation by using experimental data. The results of the comparison are shown in Figure 2(b). From these results, we can say that the validity of the proposed method is shown by an increase in accuracy of approximately 2% in comparison with the Random category, which is the one with the highest accuracy in the conventional method.

4. Conclusion

The proposed method performs semantic segmentation which handles multiple labels using Label Ranking and Label Powerset. From evaluation experiments using a uniquely constructed database, we have demonstrated it the possibility of segmentation with approximately 70% accuracy. In the future, we plan to examine the effects of increasing the numbers of multiple label categories that are combined and the numbers of categories.

References

[1] J. Read. A pruned problem transformation method for multi-label classification. In *Proc. New Zealand Computer Science Research Student Conference*, pages 143–150, 2008. 1

[2] J. Shotton, M. Johnson, and R. Cipolla. Semantic texton forests for image categorization and segmentation. In *Computer Vision and Pattern Recognition*, 2008. 1, 2

[3] G. Tsoumakas, I. Katakis, and I. Vlahavas. Mining multi-label data. *Data Mining and Knowledge Discovery Handbook*, 2010. 1, 2