

## Task Description

The project is to create classifier for fruits and vegetables. There is given subset of fruits and vegetables: orange, apple, watermelon, cauliflower, broccoli. Additionally there is one more class describing not known objects. Classifier will give an answer of the type of the object on the input and will also be able to tell if it is an object not from the given list of learned objects.

## Proposed solution

In order to solve the given problem the following solution was proposed: Features will be extracted from images via SIFT algorithm. Later features will be added to Bag of Words. Finally images will be classified using Support Vector Machine.

## Final program

The proposed solution was implemented exactly as it was proposed in the preliminary report. From the given subset of fruits watermelon was removed, because this class was completely not recognized correctly. For the other classes it is visible that classifier works, even though the accuracy is far from perfection.

## Testing Images Selection

Due to lack of prepared dataset of already tagged fruits and vegetables imaged, images were gathered from [www.image-net.org](http://www.image-net.org) and [www.google.com](http://www.google.com) and thoroughly checked. Then they were divided into six classes: orange, apple, cauliflower, broccoli, other. In total only 307 images were qualified to be train data, and 146 images for test purposes. Most of the data is the class "other" which were taken from MIT Objects 101 dataset. **Number of images per category:** Train data: 'apple': 74, 'cauliflower': 60, 'broccoli': 38, 'orange': 51, 'other': 84

**Test data:** 'apple': 20, 'cauliflower': 11, 'broccoli': 9, 'orange': 8, 'other': 98

Class	Apple	Cauliflower	Broccoli	Orange	Other
Precision	0.52	0.34	0.28	0.25	0.89
Recall	0.6	0.81	0.44	0.75	0.54
F score	0.55	0.48	0.34	0.37	0.67
True sum	20	11	9	8	98

## Example Results

Figures from 1 to 5 show example of correct and incorrect classification of each class.

## Result Analysis

ROC curve



(a) Apple classified as apple



(b) Apple classified as cauliflower

Figure 1: Example of apple classification.



(a) Broccoli classified as broccoli



(b) Broccoli classified as cauliflower

Figure 2: Example of broccoli classification.



(a) Cauliflower classified as cauliflower



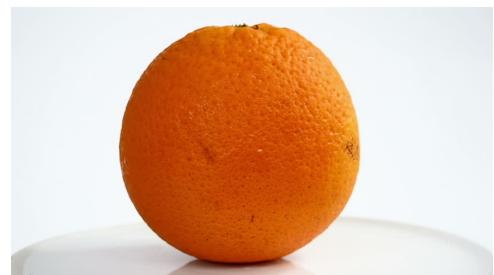
© Bozena\_Fulawka / iStock

(b) Cauliflower classified as other

Figure 3: Example of cauliflower classification.



(a) Orange classified as orange



(b) Orange as classified apple

Figure 4: Example of orange classification.



(a) Other classified as other



(b) Other as classified apple

Figure 5: Example of other classification.

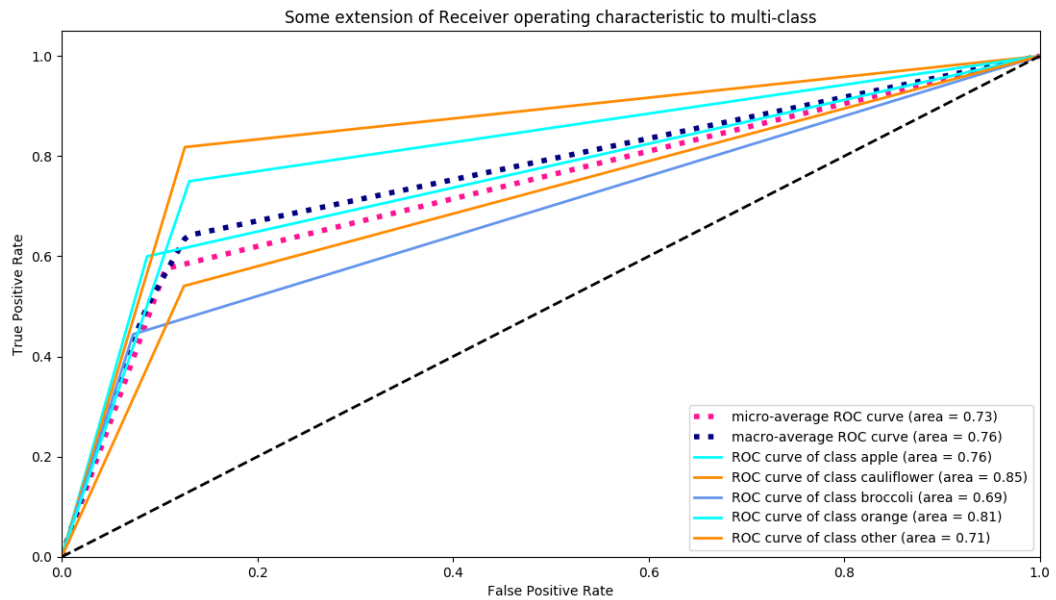


Figure 6: ROC Curve

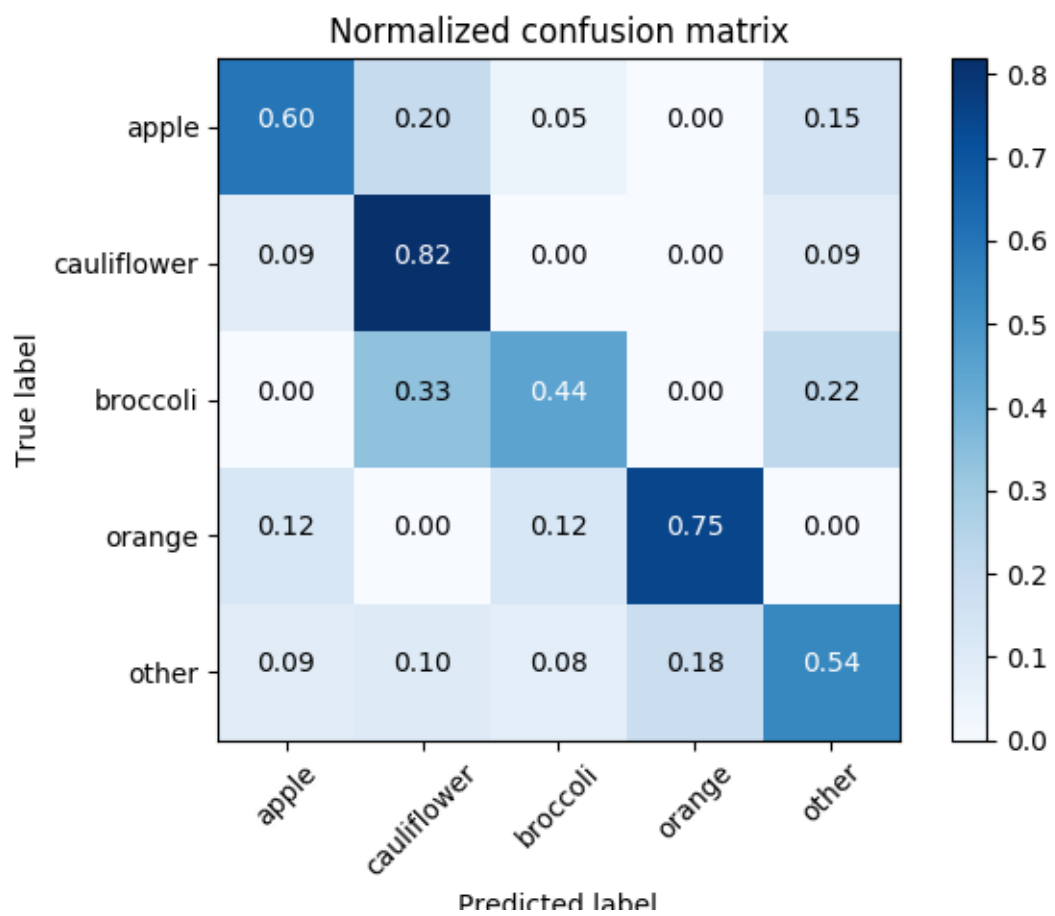


Figure 7: Confusion Matrix

CMC curve is not possible to be created do due to the fact, that OpenCV's SVM implementation does not produce probabilities but only final result.

As it is visible, classifier works with average precision at 73%, recall 0.57% and fscore 60%. Unfortunately the overall accuracy is too low to be used in real application. It is possible that more train data with better quality would make classifier work better.

## **Additional Remarks**

The program was executed on two machines producing different results. In both cases classifier worked, however on the other machine results were somewhat worse, especially for orange class, for which classifying accuracy was around 10% only. The code and data were the same and the same OpenCV version (3.3.0) was used on both computers. The reason of the difference is unknown.