Plankton Identification from Underwater Imagery

Marmar Orooji, Kazim Sekeroglu, Nhon Ai Tran

***Abstract*—We apply image processing techniques to extract features from the underwater images for identifying different plankton classes in the oceans. The goal of the project is to learn a SVM classifier with the extracted global and local features to classify plankton automatically and accurately. We applied principal component analysis as a dimension reduction technique to reduce the overall computation time during model training phase.**

Keywords*—*plankton identification, computer vision, SVM, machine learning.

I. INTRODUCTION

Plankton are plants and tiny animals that drift with the ocean currents. They provide a important source of food for many larger animals in the bodies of water. The loss of plankton population would have big impact on the ocean ecosystems. Due to this reason, marine researchers are interested in measuring and monitoring the population level of plankton to keep an eye on the health of the oceans and ecosystems.

Marine research has advanced from the traditional methods for measuring and monitoring plankton population to using towed, underwater camera system that captures microscopic, high-resolution images for large areas of study. However, this advancement introduced another challenge. Manual analysis of the images captured by the system has shown to be impractical. Thus, there is a need for a system for analyzing the images and automatically identifying object captured in the images.

In this project, we proposed an approach that achieves high recognition rate of different plankton classes using raw pixel values plus five global and two local features. We used Support Vector Machines (SVM) classifier for learning the model. Our approach also includes principal component analysis (PCA) for reducing the feature dimensions to improve training time of the classifier.

The rest of the paper is organized as follows. Section 2 presents the methodology for computing the features. We discuss the empirical experiments of the proposed method in Section 3 and provide analysis of the results in Section 4. Finally, we conclude and provide a summary of the project findings in Section 5.

II. METHODOLOGY

In this project, plankton classification from underwater imagery was done by using the global and local image features. Since there are many challenges such as high variation in the same class and small image size, features extracted from the images have to be very robust to be able to make a good classification. To overcome these issues not only the global features but also the robust local features such as SIFT, SURF and ORB were used for classification. Before extracting the image features, images were pre-processed and region of interest were segmented out. Then features were extracted from the region of interest. Global and local features that were used in this project are discussed in section II-A and II-B.

A. Global Image Features

1) PCA Applied Raw Pixel Intensity Values: Based on our experiment we have seen that the raw pixel intensity values were playing an important role in this application. Although the dimension of the images were small (25x25 pixels), classifying a huge number of images with 625 dimensional feature vector is very time consuming. To reduce the dimension of the feature vector obtained from 25x25 raw pixel intensities, Principal Component Analysis (PCA) method was used and the 625 dimensional feature vector is reduced to 50 dimensions.

2) Axis Ratio: After extracting the region of interest, an ellipse was fitted and the ratio of the major and the minor axis were used as a feature. As seen in Figure 1, axis ratio is a good discriminative feature for describing the shape of the object.

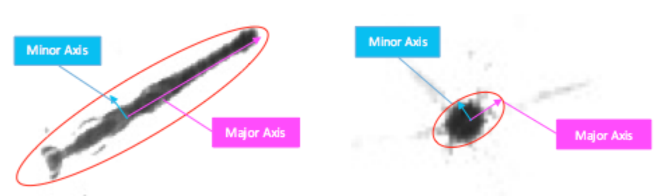


Fig. 1. Ellipse fitted to sample plankton images.

3) Compactness: Compactness basically gives an idea about the shape of the object such as whether the object is spread over the image or it is compact like a circle. Compactness can be computed by using the perimeter and the area of the region. Perimeter is the number of pixels located on the contour of the object and the area is the number of pixels within the object [2]. The equation to calculate the compactness is given below.

$$Compactness = \frac{perimeter^2}{area} \quad (1)$$

Figure 2 shows the intuition behind the usage of compactness as a feature.

4) Moments: Image moment is a certain particular weighted average of the image pixels' intensities, and they are usually used to have some attractive property or interpretation about the images. Moments are giving some fundamental geometric properties such as area (or total intensity), centroid, and information about the orientation of the object. In this project, Hus invariant moments were used to extract the global features of the images. General moment equation is given below.

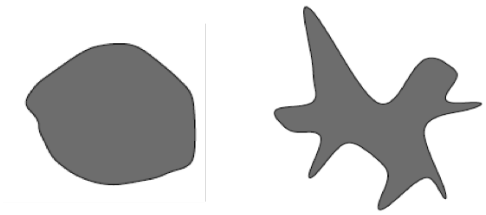


Fig. 2. The left object is clearly the more compact of the two.

$$m_{pg} = \sum_{x} \sum_{y} x^p y^q f(x, y) \tag{2}$$

B. Local Features

Three of the most famous and robust local feature detectors SIFT, SURF and ORB were experimented in this project and among all these there local feature detection algorithm SIFT was performed the best based on our experiments. Hence, SIFT was used to extract the local image features in this application. In addition to SIFT features, corners were also used. Shi-Tomasi corner detection algorithm was used to detect the number of corners in the region of interest.

C. Combination of Global and Local Features

As mentioned earlier in the method section, combination of global and local features were used to get a better classification performance. Table 1 shows the method names and their abbreviations used in this application.

TABLE I. FEATURES AND CORRESPONDING ABBREVIATIONS

Features		Abbreviation
PCA applied on raw pixels		PCA
Global	Axes ratio	AR
	Filled Area	FA
	Perimeter	P
	Compactness	C
	Moments	M
Local	SIFT	S
	$Shi-Tomasi$ corner detector	CD

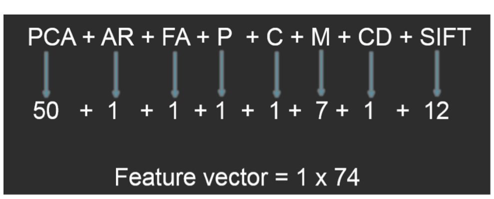


Fig. 3. Feature vector obtained by combining the global and local features for each object.

III. Experiment

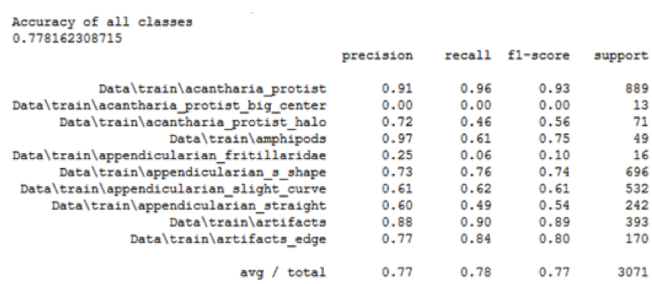
The data set used in the project was obtained from Kaggle website [1]. It was the data set provided for the National Data Science Bowl competition. It includes a total of 30,336 images for 121 plankton classes. The images were in various sizes. However, we re-sized the images to 25x25 to perform image segmentation and smoothing prior to computing the global and local features.

In our experiment, we reduced the number of classes from 121 to 10, since our machines could not handle with this huge data set. We tried to define a set of discriminant features which best recognize and classify plankton images. As a baseline, we considered raw pixel intensity values as our features. The images were 25 by 25 pixels, thus making the baseline feature vector size to be 625. Then, we examined different combination of global and local features, with or without raw pixel values, in order to get improvement in classification accuracy.

After defining each set of features, we used SVM classification model. Of course, SVM has different parameters that affect the result and needed to be tuned to avoid over-fitting and under-fitting and get the best generalization. In our experiment, we used RBF kernel for SVM, which is usually the best [3]. However, we tuned the other parameters to get the best 10-fold cross validation accuracy. Thus, by using 10-fold cross validation, we could evaluate the performance of our model with the specific features as the input and compare the results to find the best set of features.

IV. Results

In addition to 10-fold cross validation accuracy, we obtained precision, recall and f1-score for each class, so that we can see the performance on each class separately. Below is the baseline result of SVM by having just raw pixel intensity values as the image features. As shown in Figure 4, the accuracy is 77.81%, which is surprisingly high. Because we usually do not expect raw pixels as good features.



Accuracy of all classes
0.778162308715

	precision	recall	f1-score	support
Data\train\acantharia_protist	0.91	0.96	0.93	889
Data\train\acantharia_protist_big_center	0.00	0.00	0.00	13
Data\train\acantharia_protist_halo	0.72	0.46	0.56	71
Data\train\amphipods	0.97	0.61	0.75	49
Data\train\appendicularian_fritillaridae	0.25	0.06	0.10	16
Data\train\appendicularian_s_shape	0.73	0.76	0.74	696
Data\train\appendicularian_slight_curve	0.61	0.62	0.61	532
Data\train\appendicularian_straight	0.60	0.49	0.54	242
Data\train\artifacts	0.88	0.90	0.89	393
Data\train\artifacts_edge	0.77	0.84	0.80	170
avg / total	0.77	0.78	0.77	3071

Fig. 4. Classification result of the baseline, just having raw pixel intensity values as the features.

Therefore, we tried to substitute raw pixels with better features. So we used a set of global and local features instead and found that the accuracy decrease to 76%, Figure 5.

Hence, we came up with the idea of keeping raw pixel values inside our set of features. It seems that in our image set, raw pixel values are representative. So in the next step, we added raw pixel values to our set of local and global features and fortunately the accuracy got improved to 81.33%, Figure 6. However, this set of feature is computationally expensive and time consuming. We applied that to 35 classes and it just took about 1 hour. So, reducing dimensionality is necessary here.

Accuracy of all classes
0.760305321752

	precision	recall	f1-score	support
Data\train\acantharia_protist	0.86	0.94	0.90	889
Data\train\acantharia_protist_big_center	0.54	0.54	0.54	13
Data\train\acantharia_protist_halo	0.78	0.54	0.63	71
Data\train\amphipods	0.68	0.73	0.71	49
Data\train\appendicularian_fritillaridae	0.00	0.00	0.00	16
Data\train\appendicularian_s_shape	0.72	0.74	0.73	696
Data\train\appendicularian_slight_curve	0.53	0.56	0.55	532
Data\train\appendicularian_straight	0.65	0.38	0.48	242
Data\train\artifacts	0.89	0.87	0.88	393
Data\train\artifacts_edge	0.96	0.96	0.96	170
avg / total	0.75	0.76	0.75	3071

Fig. 5. Classification result by having combination of local and global features, without raw pixel values.

Thus, in the next step, we applied PCA on our raw pixels to reduce it from 625 to 50 features and the result can be seen in Figure 7. Actually it did not improve the accuracy a lot, from 81.33% to 82%, but it speeds up the performance tremendously.

0.813380913285

	precision	recall	f1-score	support
Data\train\acantharia_protist	0.92	0.97	0.94	889
Data\train\acantharia_protist_big_center	0.62	0.38	0.48	13
Data\train\acantharia_protist_halo	0.81	0.68	0.74	71
~~Data\train\amphipods~~	~~0.81~~	~~0.80~~	~~0.80~~	~~49~~
$\boxed{\text{Data\textbackslash train\textbackslash appendicularian\_fritillaridae}}$	0.00	0.00	0.00	16
Data\train\appendicularian_s_shape	0.75	0.78	0.76	696
Data\train\appendicularian_slight_curve	0.64	0.64	0.64	532
Data\train\appendicularian_straight	0.67	0.58	0.62	242
Data\train\artifacts	0.94	0.93	0.93	393
Data\train\artifacts_edge	0.98	0.96	0.97	170
avg / total	0.81	0.81	0.81	3071

Fig. 6. Classification result by having combination of local and global features, plus raw pixel values.

Accuracy of all classes
0.820213525365

	precision	recall	f1-score	support
Data\train\acantharia_protist	0.92	0.97	0.94	889
Data\train\acantharia_protist_big_center	0.62	0.38	0.48	13
Data\train\acantharia_protist_halo	0.81	0.68	0.74	71
Data\train\amphipods	0.85	0.82	0.83	49
Data\train\appendicularian_fritillaridae	0.00	0.00	0.00	16
Data\train\appendicularian_s_shape	0.76	0.79	0.77	696
Data\train\appendicularian_slight_curve	0.65	0.66	0.65	532
Data\train\appendicularian_straight	0.68	0.57	0.62	242
Data\train\artifacts	0.94	0.93	0.93	393
Data\train\artifacts_edge	0.97	0.96	0.96	170
avg / total	0.81	0.82	0.82	3071

Fig. 7. Classification result by having combination of local and global features, plus PCA on raw pixel values.

In the results, it is shown that the performance of the fifth class in all the steps is really bad, specified by a red box in Figure 6. Meaning that no instance in the fifth class can be classified correctly, regardless of the feature selection. This happens because of two reasons. First, the total number of instances in this class is just 16, which is really small. Secondly, the images in this class, shown in Figure 8, have a high variations in a way that we can hardly claim that these belong to the same class.

Table II shows a summary of the cross validation accuracy with different set of features.

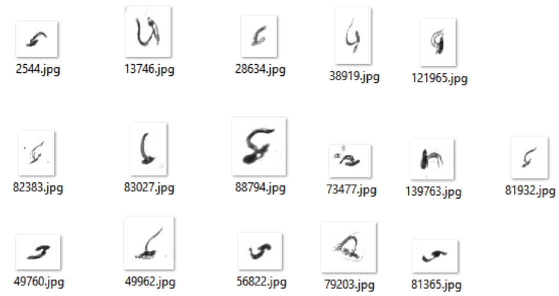


Fig. 8. Sample images of a high variation class, the Appendicularian Fritillaridae class.

TABLE II. SUMMARY OF THE CROSS-VALIDATION ACCURACY

Features	Accuracy (%)
Raw pixel values (Baseline)	77.81
$Global + Local$ (without raw pixels)	76.03
Raw pixel values $+Global + Local$	81.33
$PCA + Global + Local$	82.02

V. CONCLUSION

Firstly, due to the structure of our images, raw pixel intensity values need to present in our feature vector. However, it is necessary to apply PCA to raw pixels, as we have 625 pixels, in order to reduce the dimensionality. In addition, we used some local features like SIFT that are known as the best features in image processing field, since they are robust to scale, rotation and pose variations. However, in our experiment, local features, alone, could not classify plankton images very well due to small image size and high variations between images within a class. Therefore, we found that the presence of features that could globally represent the image, is necessary. Thus, the best set of features would be the one that has PCA plus combination of local and global features.

REFERENCES

[1] Kaggle.com, *National Data Science Bowl Competition*, https://www.kaggle.com/c/datasciencebowl, 2015.

[2] M. Sonka, V. Hlavac, and R. Boyle, *Image Processing, Analysis, and Machine Learning*, CL Engineering, 2007.

[3] T. Hastie, R. Tibshirani, and J. Friedman. *The Elements of Statistical Learning.* Springer, 2009. 2nd edition.