ISSN: 0976-9102 (ONLINE)
DOI: 10.21917/ijivp.2020.0306

ASSESSMENT OF SUPPORT VECTOR MACHINE FOR CLASSIFICATION OF SARDINE IMAGES

A. Anushya
Department of Computer Science, St. Jerome's College, India

Abstract

The goal of image classification is to forecast the types of the input image using its features. This research focuses on the classification of sardine fish images via Support Vector Machine. Sardines are abundant in the Pacific and Atlantic Oceans, and are available in all the fish markets around the world. Hence, it is the most common sea food in wide-reaching. The sardine fish has a distinct appearance that sets it apart from other types of fish. So, finding out the best-quality fishes is a task that requires the benefit of classification. The sardine images used for the study are collected from Kanyakumari district, Tamil Nadu, India. Gray-Level Co-Occurrence Matrix (GLCM) is used for the Texture Analysis of the images and to extract the statistical features of images. SVM is applied on data and the categories dates are obtained. Both the algorithms are executed in MATLAB and the experiments are carried out for getting better results.

Keywords:

Classification, Image Classification, Image Processing, Support Vector Machine

1. INTRODUCTION

Nowadays computer science is extensively used in the agricultural and food science sectors. Several techniques like Artificial Intelligence and soft computing are used for image classification and fault detection to provide better-quality food product to the end user [2].

As fishes could be injured while being manually removed from the fishing net, this research strives investigate the sardines which are good (not injured) and the ones that are injured through image classification. Image classification refers to the job of mining information categories from an image. The resultant Raster Image can be classified into two classes based on the interaction between the analyst and the computer during classification. The first type is supervised classification and the second is unsupervised classification [3].

Supervised classification uses training samples to yield classifying the image. Supervised classification includes the creation of training samples that exemplify the classes intended to be mined from raster image with the aid of the image classification tool. Later these classes can be used by the different classification tools for classifying a new image [7].

Unsupervised classification does not depend on analyst interaction to find spectral classes within a multilane image. The process of unsupervised classification is done by making clusters with quality analysis capability by providing access to the classification tools [11].

Image classification refers to a process in computer vision that can classify an image according to its visual content [14]. Image classification is the method of taking an input (like an image) and outputs a class or a probability that the input is a particular class.

The main aim of this paper is to classify the fish image using its features through machine learning techniques. An image recognition model contains three distinct phases: pre-processing, feature extraction, and classification phases. Within the feature extraction phase, a Scale Invariant Feature Transform (SIFT), in addition to shape and color algorithms are used to extract a feature vector for each image. The classification phase uses Support Vector Machine (SVM). Evaluating the recognition model is done by carrying out a series of experiments. The results from implementing these tests reveals that the planned approach can automatically classify the sardines to be either Good or Injured while keeping a high degree of accuracy.

The organization of the paper is as follows: In section 2, we briefly present SVM classification. We describe data source in section 3. Classification of sardines using SVM is depicted in section 4, and section 5 concludes the paper while including argument on finding of the research and on probable future work.

2. RELATED WORKS

A detailed review of the research on fish image classification. But, there is minimum number of study on fish image classification. Consequently, the present system generally introduces the previous work on fish image classification and also produces other image classifications like vegetables and fruits.

Ogunlana et al. [12] contributes an SVM-based fish classification algorithm. Six shape features namely body length, anal fin length, caudal fin length, dorsal fin length, pelvic fin length and pectoral fin length were extracted from 150 fish (divided into 76 training and 74 testing sets) and the extraction forms the basis for the classification. The classification results are compared with other algorithms such as ANN, K-NN and K-Means Clustering. However, the obtained recognition accuracy of 78.59% is better than other algorithms.

Serna et al. [4] uses Artificial Neural Networks for the automatic identification of species. They have worked on a dataset of 697 images achieving an accuracy of 91.65%.

Daramola et al. [13] proposes a system for classifying fish images into distinct classes based on their physical form obtained from Single Value Decomposition (SVD) product extracted from fish block images. Training and testing of the proposed fish classification system are done using Artificial Neural Network (ANN). Experimental test is carried out to determine the species of query fish images. Thirty-six fish images were tested, and 94% correct classification result is recorded.

There are some challenges in vegetable classification. The background of vegetable image is complex, there is similarity between the different species of vegetables, so we cannot just rely on a single feature, such as color, shape or texture to distinguish the species of vegetables [1] [9]. The same species of vegetables will be different because of the shape, scale, viewpoint and so on

[10]. The idea is based on Martin Gorner's Hand-written digit recognition using Tensorflow for Poets [5] [6]. A similar procedure is followed for developing this android based system. The idea is to use computer vision, image processing and convolutional neural networks [8].

3. SUPPORT VECTOR MACHINES FOR CLASSIFICATION

Vladimir developed the original SVM algorithm. Vapnik proposed the current standard incarnation (soft margin). The common task in machine learning is classifying data where some of the given data points that belong to either one of the two classes, and the aim is determine the class to which the new data point belongs. Support vector machines, view a data point as a p-dimensional vector consisting of a list of p numbers. With linear classifier, the goal is to know whether these points can be separated with a (p-1) dimensional hyperplane.

There are many hyperplanes that might classify the data. But there is one rational choice that best represents the largest separation, or margin, between the two classes. The hyperplane is chosen accordingly is such a way that the distance from it to the nearest data point on each side is at the maximum. This hyperplane, if exists, is called the maximum-margin hyper plane and the linear classifier it defines is known as a maximum margin classifier or equivalently, the perceptron of optimal stability.

SVM for Classification is a convenient technique for data classification where a model is produced to predict the target value of data in the testing set. The model is built through training and testing data that involves of set of data instances and each instance contains one target value and several attributes. Classification in SVM is of Supervised Learning type. Recognized labels help in specifying correctness of the performance of the system which gives indication of desired response and validates the accuracy of the system, or helps the system learn to perform correctly. Feature selection or feature extraction is a step in SVM classification which includes identifying which instances are closely connected to the known classes.

Feature selection along with SVM classification can be used to identify key sets which are involved in whatever processes to distinguish the classes even when prediction of unknown samples is not necessary. The goal of a support vector machine is to find the optimal separating hyper plane which maximizes the margin of the training data. The first thing we can see from this definition is that an SVM needs training data. Which means it is a supervised learning algorithm. It is worth stating that SVM is a classification algorithm, in other words we will use it to foretell if some object belongs to a particular class.

4. DATA SOURCE

Totally 225 Sardine fish images are collected from Pillaithoppu village which is coastal area in Kanyakumari District, Tamil Nadu, India for this research to conduct experiments. In this study, to conduct the experiment a dataset is created. This dataset contains 225 images, all are distinct sardines of two classes which are not injured (Good) and injured as shown in Fig.1. Images were taken by Canon - EOS 5D Mark III DSLR Camera with 24-105mm f/4L IS Lens with a uniform red background to insure the accuracy of the extraction features. Each image presents one Sardine only as illustrated in Fig.1. Sardines are also known as pilchards in some places, sardines are small oily fish that belong to the herring (Clupeidae) family. Its body is flat and with large, reflective silver scales (although these scales are not found on their head). In the center of their belly is a set of specialized scales that are jagged and point backwards. These are called scutes. It also has one short dorsal fin. Despite usually being classified as a single species, there are actually 21 different types of fish that fall under the sardine category. Sardinops, sardine, dussumieria and sardinella are some of the most well-known species today. Sardines are generally smaller than other fish, ranging in length from 6 to 12 inches (15 to 30 centimeters), and are soft-boned. These images are classified into two groups such as not injured and injured.

Fig.1. Database Image

5. CLASSIFICATION OF SARDINE IMAGES USING SVM

The simple model for sardine classification essentially consists of four steps:

1. Image acquisition,
2. Image processing,
3. Features are extraction and
4. Classification

In step 1, the images are acquired and a database of the sardines is created. Then, Image processing is done to images in order to improve image quality. After that, desired features are mined and reduced (if needed) and input to the model. Lastly, a classifier is used to apply classification.

We first visually examined the fishes that we used in this experiment and classified them according to their features. The fishes without any scratch or cut were branded as of the best quality, Good (Not Injured). There are two kinds of injured fishes one is the head including eyes and the tail parts are cut or injured as illustrated in Fig.2 and the other is injured on fish surface such

as scratch, texture injury, puddle cutting injury, surface cutting injury, crash injury etc as illustrated in Fig.2.

Fig.2. Samples of Injured Fish Images

For the classification step which is the step that determines which fishes are good and which are not, we used thresholding and edge detection algorithms to separate area. To identify colors, we used up to 12 levels of colors according to Agrexco Carmel Standards. A Hue Saturation Value (HSV) matrix was created and identification was performed on the hue matrix. Then, the WCL feature (weighted color) was calculated by comparing the Agrexco color table with the colors of the screen (100 hue colors of red and green were divided into 20 groups each contains 5 colors). For this color identification, we used the flabbiness and intensity features. The shape and size roundness was identified by calculating the distance from the image edge to the center of the image using the FFT (Fast Fourier Transform) where the more round the object the less frequency of the vector will be. The functions used are embedded in MATLAB which is used as out built classifier.

6. RESULTS AND ANALYSIS

In this section, we discuss results about the experiments we have performed. The database contains 225 images as displayed in Fig.1. There are two different classes of images acting as the data source. The dataset is divided in the ratio 70:30, where 70% is for training and 30% is for testing. We used 157 images in training phase, among these 100 are not injured and 57 are injured. Also 68 images are used in testing phase. 43 are injured and 25 are not injured in testing phase. Among these images 143 are not injured and 82 are injured. The resolutions of images were set to 89×142 pixels for the sake of sensible computational speed. Totally 225 images with 2 values such as injured and not injured are used for experiments. The steps involved using an image in image processing is shown in Fig.3 - Fig.6.

Fig.3. Original image

Fig.4. After background removed

Fig.5. Gray scale image

Fig.6. Segmented image

We used different techniques to extract features under study of the images as in Fig.6(a) and Fig.6(b) for the injured and not eatable group respectively. This method classifies fishes into two classes, injured and not injured in an efficient way. We have used SVM classifier in classification of fishes. We have measured SVM classifier using linear kernel function and found 100% accuracy as illustrated in Table.1, Fig.7(a) and Fig.7(b), respectively.

Table.1. Analysis of SVM on Dates Images

Dataset	Total Number of images	Number of images - not injured	Number of images - injured	Accuracy
Training phase	157	100	57	100%
Testing phase	68	43	25	100%

Fig.7(a). Expected Results with Injured fish images

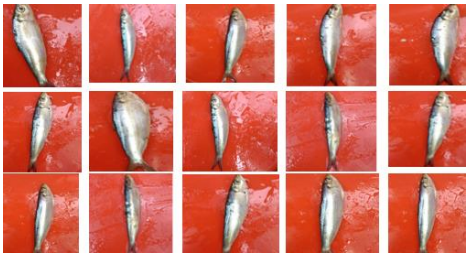

Fig.7(b). Expected Results with Not Injured fish images

7. CONCLUSION

In this paper, we have used SVM to classify the images of sardines as only injured and not injured (Good). The images were collected from Kanyakumari district, Tamil Nadu, India. The method has 100% accuracy which means it is suitable to classify fishes. Moreover, this method can be applied to other objects like seeds, flowers classification, and others, where human involvement exists for classification. In near future, other classification algorithms will be used on the same banana images to evaluate the better accuracy. Also, in the future, after identifying the shapes, some of the shape representation techniques may be introduced for identifying the similarity of shapes.

REFERENCES

[1] Alex Krizhevsky, lIya Sutskever and Geoffrey E. Hinton, "ImageNet Classification with Deep Convolutional Neural Networks", *Communications of the ACM*, Vol. 60, No. 6, pp. 75-79, 2014.

[2] C. Pornpanomchai, B. Lurstwut and P. Leerasakultham, "Shape and Texture based Fish Image Recognition System", *Kasetsart Journal*, Vol. 47, pp. 624-634, 2013.

[3] D.B. Sheng, S.B. Kim and N.T. Hai, "Fish Injured Rate Measurement Using Color Image Segmentation Method Based on K-Means Clustering Algorithm and Otsu's Threshold Algorithm", *Journal of Korean Society for Power System Engineering*, Vol. 20, No. 4, pp. 32-37, 2016.

[4] A.H. Serna and L.F.J. Segura, "Automatic Identification of Species with Neural Networks", *Journal of Life and Environmental Sciences*, Vol. 23, No. 1, pp. 1-8, 2014.

[5] Tensorflow and Deep Learning, Available at: https://github.com/GoogleCloudPlatform/tensorflow-without-a-phd/tree/master/tensorflow-mnist-tutorial.

[6] Le Hoang Thai, Tran Son Hai and Nguyen Thanh Thuy, "Image Classification using Support Vector Machine and Artificial Neural Networks", *International Journal of Information Technology and Computer Science*, Vol. 5, pp. 32-38, 2012.

[7] M.E. Nilsback, and A. Zisserman, "A Visual Vocabulary for Flower Classification", *Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, pp. 1447-1454, 2006.

[8] M.E. Nilsback, and A. Zisserman, "Automated Flower Classification over a Large Number of Classes", *Proceedings of IEEE 6th International Conference on Computer Vision, Graphics and Image Processing*, pp. 722-729, 2008.

[9] N.B.A. Mustafa, K. Arumugam, S.K. Ahmed and Z.A.M. Sharrif, "Classification of Vegetables using Probabilistic Neural Networks-Improvement using Color Features", *Proceedings of IEEE International Conference on Technology, Knowledge and Society*, pp. 122-128, 2011.

[10] Om Patil and Vijay Gaikwand, "Classification of Vegetables using TensorFlow", *International Journal for Research in Applied Science and Engineering Technology*, Vol. 6, No. 4, pp. 2926-2934, 2018.

[11] S.O. Ogunlana, O. Olabode and G.B. Iwasokun, "Fish Image Classification using Support Vector Machine", *African Journal of Computing and ICT*, Vol. 8, No. 2, pp. 1-8, 2015.

[12] D.S. Adebayo and O. Olumide, "Fish Classification Algorithm using Single Value Decomposition", *International Journal of Innovative Research in Science*, Vol. 5, Issue 2, pp. 1621-1626, 2016.

[13] D. Kaeli, P. Mistry, D. Schaa and D.P. Zhang, "*Heterogeneous Computing with OpenCL 2.0*", Morgan Kaufmann Publisher, 2015.

[14] A. Anushya, "Image Clustering using K-Means on Marine Products", *International Journal of Innovative Technology and Exploring Engineering*, Vol. 9, No. 4, pp. 1-13, 2020.