Journal of Engineering and Applied Sciences 13 (20): 8506-8510, 2018
ISSN: 1816-949X

An Intelligent Image Classifier Based on Histogram of Oriented Gradients Features

Methaq Talib Gaata and Sahar Muneam
Department of Computer Science, College of Science, Mustansiriyah University, Baghdad, Iraq

Abstract: This study presents an intelligent classifier for images classification based on Artificial Neural Network (ANN). The Histogram of Oriented Gradients (HOG) techinque has been used in order to extract features from image. The ANN supervised feed-forward scaled conjugate gradient algorithm used to bulid the proposed classifier. The input image is processed directly to extract vector features regardless of size or colour map. The architecture of ANN is selected to be simple and appropriate to carry out the classification process with high accuracy. This work is performed on the Caltech dataset. Four classes of image are used to test and evaluate the performance of the proposed classifier (96 images for all category), the testing images consists of 192 images (48 images for each category). Experimental results showed that the classification rate was 93.23%.

Key words: Image classification, feature extraction, ANN, (HOG), extract vector, high accuracy

INTRODUCTION

According the diversity of images type, sizes and its textures information, the process of images classification represent one of the most important approaches in images processing field (Chen *et al.*, 2015). Also, the images classification subject represent an main stage in different analyticals systems that takes an images as based data for preparing final decision likes biometric applications such as detecting face, recognizing an optical character or a finger print, classifying the abnormalities in medical images and in any other surveillance purposes, industrial visual inspection, vehicle navigation, robot control, remote sensing and many computer vision applications (Lehmann *et al.*, 2005). Though, the manual process of classifying images is characterized by being time-consuming and repetitive and not always is reliable. Therefore, the automatic process of image classification still represents an important need (Tong *et al.*, 2017). Research on image classification has grown quickly in the field of machine learning. However, though various methods have been used in this domain, like Artificial Neural Networks (ANN), Supper Vector Machines (SVM), fuzzy logic, Genetic Algorithms, etc. between all these methods the best results have been reported by ANN techniques (Zhou *et al.*, 2010).

The neural network process the information in a way similar the human mind does. The networks is consists of a big number of interconnected processing element (neurons) that work in parallel to solve a particular problem through exploited their significant ability to obtain meaning from difficult or imprecise data (Rouhani and Ravasan, 2013). The trained neural network can be consideration of as an (expert) in category of information it have been specified to analyze. This expert is able then to given prediction of new situation (Gao and Liu, 2008).

In recent years, some of research works performed which are concerned to image classification based on the intelligent approaches. Chitaliya and Trivedi (2010) used discrete wavelet transform to produce the element pictures from individual wavelet sub groups. The element pictures developed from wavelet coefficients were employed as a component vector for the further process. Additional to that the PCA used to lessen the element's dimensionality vector. The lessened component vector has been used for further recognition. This proposition is given 93.3% recognition rate which is preferable think about over separation classifier. Raja and Shanmugam, (2011), classified a natural war scene by extracting wavelet features using ANN and SVM. Then, the results obtained from both methods were compared and the best classified scene was determined. Later, the extracted features will be normalized to maintain the data and improve the performance of the classifier. After, the normalized features will be sent to ANN and SVM as an input. Once received, ANN and SVM will start classifying the image using backward-propagation algorithm and radial basis kernel function with p = 5, respectively. The rate of SVM classification was 59% whereas that of ANN was 72.5%. This result proves that ANN is better than SVM with regard to the process of let wavfeature extraction. Tiay*et al.* (2014) proposed flowers recognition system based on K-NN. The characteristics of edges and colors of flowers images take in account in flowers classification. Also, Hu's seven moment algorithm was also applied to obtain edge characteristics. In addition, the following

Corresponding Author: Methaq Talib Gaata,Department of Computer Science, College of Science, Mustansiriyah University, Baghdad, Iraq

characteristics, reds, greens, blue, hue and saturation were derived using histograms. During the classificatory stage all values will be classified using K-NN algorithm. The accuracy of this system soared up to 80% and more than that.

In this study, the proposed classifier will be focusing on the task of classification images by ANN into different categories.

MATERIALS AND METHODS

Proposed intelligent classifier: The aim of this study is how to make the classifier has the ability to take a decision in a way that is similar to what a human does to perform classification of real images with different texture and pattern based on visual content of image. Figure 1 illustrates a typical image classification structure.

As shown in Fig. 1 at started, the images data set is initialized and then divided within two equal parts are training and testing data. Second, the features extraction process using Histogram of Oriented Gradients (HOG) algorithm applied on images in order to produce the vector of features. This vector will be used as input in classification process. Third, the intelligent classifier will be design and implement based on the ANN. The ANN that is designed will learn through the training data and learning signal until the appropriate weight values are obtained. Fourth, the proposed intelligent classifier tested with testing data. Finally, the obtainded results evaluated using statistical measurements in order to determine the accuracy level.

Preparing dataset

Image loading: The proposed method is applied on specific image dataset called Caltech which is downloading free from internet. The input image has JPG format. Four categories of object image are used. These images include airplanes, cars, spider and paper shredder. The image size in different categories of the dataset is not the same. As an example, four image samples from four categories in this dataset are viewed in Fig. 2.

In addition to the variety of image size in all categories of Caltech dataset that are shown in previous figure, the morphological structure of main object and colour map are different in each category. Thus, any designed classifier for this image dataset must be independent on these issues (size, morphological structure, colour map) in order to generate significant percent of accuracy. There are different numbers of images in each category of the Caltech dataset but the number of images in the selected categories must be the same for all categories that are used for implementation.

Divided dataset: In general, any classification system using specific technique of intelligent system needs enough number of input data for training and testing process within the intelligent system. The complete set of input data is divided in two groups. First group of data is used for training the intelligent system to discover potentially predictive relationships. The samples of

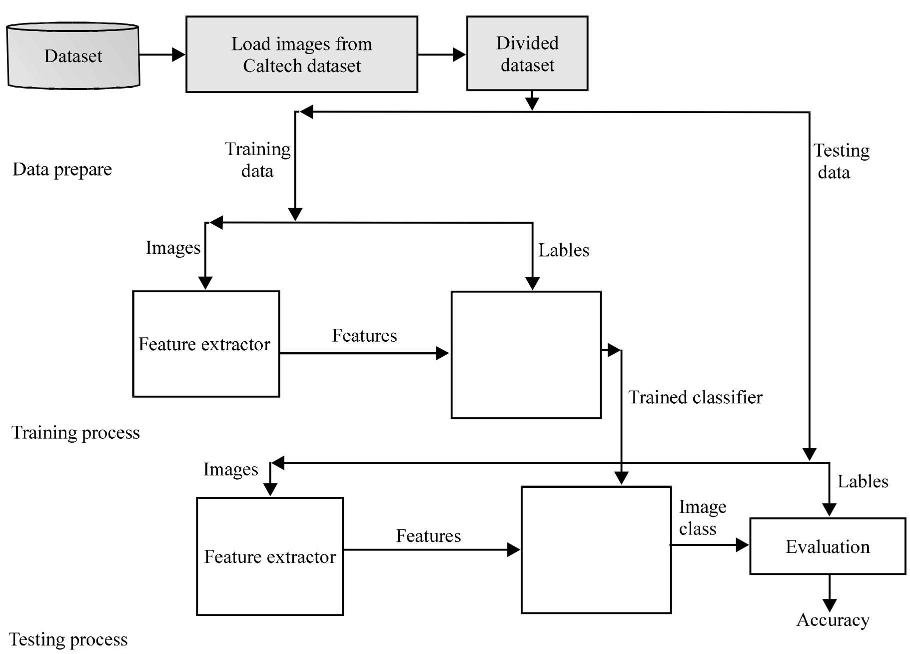

Fig. 1: The block diagram of intelligent classifier

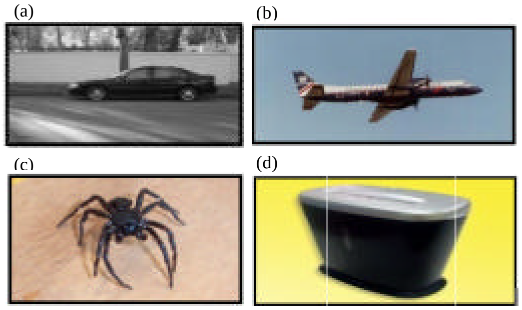

Fig. 2: Image samples from the four categories of Caltech dataset: a) The size is 300×197; b) The size is 209×378; c) The size is 365×246 and d) The size is 180×180

training set must be selected randomly but must includes all possible variations of texture, morphologies, etc. in the complete set of input data. In our works in making size of training set is equal to the testing set 50%.

The feature extraction method: In this step, operations are applied after preparing dataset from the previous step to extract some represented features. The computation of these features is done by extracting histogram of oriented gradients feature are features descriptors that are compute by counting the occurrence of gradient orientation in localize parts of an images. For compute these features, the image is separated into cells and histograms of gradient direction are created for each of cells. This histogram forms the descriptor. HOG features has been effectively implemented for other application such as human detections (Dalal and Triggs, 2005), pedestrian detections (Junior *et al*., 2009).

In this method, the image was separated into 9 overlap rectangular cells and for every of these cells, gradients directions were compute. Based on the gradients directions all pixels within a cell cast a weighted vote to forms an orientation based on histogram channel of 9 bins. The gradient strength of all cell was normalize according to L1-norm. The 9 histograms with 9 bins are concatenate to form 81 dimension features vector is feed as inputs to classifier.

Artificial neural network classifier: The feed-forward (SCG) algorithm is selected to be the supervised classifier. This classifier is constructing and trained with following organization and parameters:

- Three layers architecture, i.e. one hidden layers
- Numbers of nodes in the input layers equal to the length of input vectors (81 input neurons)
- Number of neurons in the hidden layers = 10 (best NN architecture)

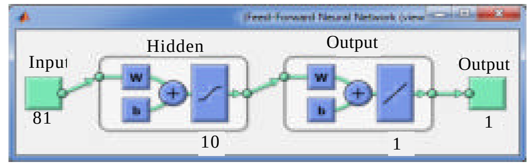

Fig. 3: The proposed feed forward neural network

- One neuron in output layer to aggregate the class no of the input image is found as the optimal model to image classification
- Kind of activation function among input and hidden layersis hyperbolic tangent sigmoid transfer function, the function between hidden and output layer is linear transfers function
- The maximum epochs for training = 40000.
- Performance goal is mean square error = 1e-30

In order to simplify the architecture of the ANN and due to the large amount of input data to the classifier, each single element from HOG feature vector goes to single input neuron. Figure 3 is show the proposed feed-forward neural network.

After determined NN architecture, training process will be start in this process the classifier learns its own classification rules from a training set. The goal of the learning process is to make ANN have the ability to think like human to perform classification of real images with different texture and pattern. As the proposed neural network is planned to classify multi categories, the trained data is selected from different categories of the Caltech dataset. Thus, the network is trained.

After learning an ANN can memorize the characteristics of the training examples and predict a new output due to its prediction capability. Testing dataset was used to check the classification performance and its accuracy.

RESULTS AND DISCUSSION

The performance of ANN along all epochs is shown in Fig. 4. In this Fig. 4, the green line represents the resulted image category (validated category) while the red line represents the test image category. The behavior for both curves is the same and the best error found at epoch 1450 with (0.069471) value as shown in the upper enlarge slice segment in Fig. 4. Furthermore, the blue line in the same figure represents the flow of the training process along all epochs and the steady state value as shown in the Fig. 4. Also, the value of error (1.51e-24) itself is too low. These results beside the similarity of tested and validated line prove the robustness of the proposed ANN

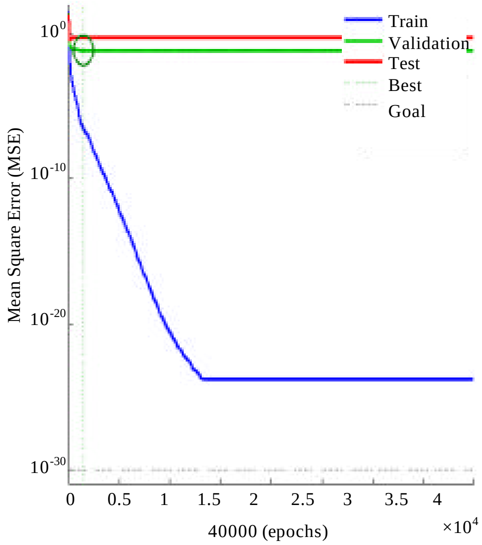

Fig. 4: Performance of ANN (Best validation performance is 0.069471 at epoch 1450)

Table 1: The percentage of classification accuracy for the in the caltech dataset

Classes	Results: Training process (%)	Testing process (%)
Airplane	100	93.75
Car	100	95.83
Spider	100	91.68
Paper-shredder	100	91.66

approach using (TRAINSCG) training algorithm to produce significant percent of accuracy. And its accuracy that is given by:

$$\text{Accuracy} = \frac{\text{Number of correct classified images}}{\text{Total number of images}} \times 100\%$$

The classification result showed that the accuracy of the ANN classifier in Table 1. As noted results of classification in Table 1, clear that for the "Car class" the accuracy have the highest dgree. The second best result is obtained with "Airplane class" followed by "Spider class". The "Paper-shredder class" have poorest result.

CONCLUSION

The main conclusions drawn from the implementation of the proposed system are in this research, HOG descriptor was used to extract features from image. Although, there are several methods of extracting the features they were used HOG for possibility us to extract the features without need to preprocessing the image as it has the ability to deal with any image regardless of size and color because as it is mainly focus on local shape and edge information is similar to human visual processing system.

Neural networks represent a viable choice during the process of image classification. This is because such networks have proven their usefulness in solving problems with nonlinear solutions. However, the method is not void of shortcomings. The first drawback can be found during the training stage. Through this stage, the system passes through many iterations; Causing as a result much time consumption; Especially, if the computers being used have older hardware. Another shortcoming is the one related to the difficulty in selection the correct parameters (such as learning rule and number of nodes in hidden layer). Generally, choosing the correct parameters helps having a massive difference in the results it further decreases error rate percentage. However, the difficulty in maintaining balance and the wrong choices can both long training time and produce inaccurate results.

In this job, classifiers are tested and performance is then evaluated for these classifiers. Results have shown that the accuracy of the process of classifying 192 testing images of 48 per category was 93.23%. This performance of the classification process is satisfactory since, the application of the feature extraction process was done on only one single descriptor-type. Experimental results shown that the feature extraction and classification models in this thesis work provide robust performance. This system returns best results.

REFERENCES

Chen, Q., Z. Song, J. Dong, Z. Huang and Y. Hua *et al*., 2015. Contextualizing object detection and classification. IEEE. Trans. Pattern Anal. Mach. Intell., 37: 13-27.

Chitaliya, N.G. and A.I. Trivedi, 2010. Feature extraction using wavelet-PCA and neural network for application of object classification and face recognition. Proceedings of the 2010 2nd International Conference on Computer Engineering and Applications (ICCEA) Vol. 1, March 19-21, 2010, IEEE, Bali Island, Indonesia, ISBN:978-1-4244-6079-3, pp: 510-514.

Dalal, N. and B. Triggs, 2005. Histograms of oriented gradients for human detection. Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition CVPR Vol. 1, June 20-25, 2005, IEEE, San Diego, California, USA., pp: 886-893.

Gao, Y. and J. Liu, 2008. A modified differential evolution algorithm and its application in the training of BP neural network. Proceedings of the IEEE/ASME International Conference on Advanced Intelligent Mechatronics AIM, July 2-5, 2008, IEEE, Xian, China, ISBN:978-1-4244-2494-8, pp: 1373-1377.

Junior, O.L., D. Delgado, V. Gonçalves and U. Nunes, 2009. Trainable classifier-fusion schemes: An application to pedestrian detection. Proceedings of the 12th International IEEE Conference on Intelligent Transportation Systems ITSC'09, October 4-7, 2009, IEEE, St. Louis, Missouri, USA., ISBN:978-1-4244-5519-5, pp: 1-6.

Lehmann, T.M., M.O. Guld, T. Deselaers, D. Keysers and H. Schubert *et al*., 2005. Automatic categorization of medical images for content-based retrieval and data mining. Comput. Med. Imag. Graphics, 29: 143-155.

Raja, S. and A. Shanmugam, 2011. ANN and SVM based war scene classification using wavelet features: A comparative study. J. Comput. Inf. Syst., 86: 1402-1411.

Rouhani, S. and A.Z. Ravasan, 2013. ERP success prediction: An artificial neural network approach. Sci. Iran., 20: 992-1001.

Tiay, T., P. Benyaphaichit and P. Riyamongkol, 2014. Flower recognition system based on image processing. Proceedings of the 2014 3rd International Conference on Student Project ICT (ICT-ISPC), March 26-27, 2014, IEEE, Nakhon Pathom, Thailand, ISBN:978-1-4799-5572-5, pp: 99-102.

Tong, H., H. Ng, T. Yap, W. Ahmad and M. Fauzi, 2017. Evaluation of feature extraction and selection techniques for the classification of wooddefect images. J. Eng. A ppl. Sci., 12: 602-608.

Zhou, X., K. Yu, T. Zhang and T.S. Huang, 2010. Image Classification using Super-Vector Coding of Local Image Descriptors. In: Computer Vision, Daniilidis, K., P. Maragos and N. Paragios (Eds.). Springer, Berlin, Germany, ISBN:978-3-642-15554-3, pp: 141-154.