International Journal of Computer Engineering in Research Trends
Multidisciplinary, Open Access, Peer-Reviewed and fully refereed
Research Paper | Volume-5, Issue-2 ,2018 Regular Edition | E-ISSN: 2349-7084

Evaluation of Industrial Based Object Detection Method Using Artificial Neural Network

F. S. Ishaq[1*], I. A. Alhaji[2], Halis Altun[3], Y. Atomsa[4], M. L. Jibrin[5], S. A. Sani[6]

[1]Dept. of Mathematics and Computer Science, Faculty of Science, Federal University, Kashere, Gombe, Nigeria*
[2] Dept. of Physics, Faculty of Science, Federal University, Kashere, Gombe, Nigeria
[3] Dept. of Computer Engineering, Faculty of Engineering, Karatay University, Konya, Turkey
[4]Dept. of Mathematics and Computer Science, Faculty of Science, Federal University, Kashere, Gombe, Nigeria
[5]Dept. of Mathematics and Computer Science, Faculty of Science, Federal University, Kashere, Gombe, Nigeria
[6]ICT Directorate, Federal University, Kashere, Gombe, Nigeria

F. S. Ishaq _ faisalsishaq@gmail.com, I. A. Alhaji _ alhaji259@gmail.com, Y. Atomsa _ au.nlaro@gmail.com, M. L. Jibrin_mljtech@gmail.com, S. A. Sani _ aminushafiu@yahoo.com
**Corresponding Author: F. S. Ishaq1_ faisalsishaq@gmail.com*

Available online at: http://www.ijcert.org

Received: 18/February/2017, Revised: 19/February/2017, Accepted: 23/February/2017, Published: 27/February/2017

Abstract:- The essence of the study is to analyse an algorithm which will provide a robust and computationally light method, which might be suitable to implement in the real-time industrial application such as object detection and recognition. For industrial applications, the primary step in automatic detection and classification of an object is to find the object automatically from an image using features related to its shape. This chore is a very complex one. Therefore, to hit the target Histogram of oriented gradient (HOG) algorithm is selected to extract the image features. Average Magnitude Difference Function AMDF is employed to correct the alignment defect. Finally, Artificial Neural Network (ANN) was employed to detect the type of object in the image efficiently. None the less, a database was generated. The database consists of images of real industrial products which are of different shapes and sizes, captured under different lightning conditions. The outcome of the experiment conducted on the database recorded 98.10% success.

Keywords: 1-D mask, HOG algorithm, AMDF algorithm, k-nearest Neighbours algorithm, Cross-correlation Functions algorithms and MLP algorithm

1. Introduction

For industrial applications, the primary step in automatic detection and classification of an object is to find the object automatically from an image using features related to its shape. Therefore shape detection plays an important role and has an intensive usage in various applications [1] such as in robotics; object classification is needed to recognize a certain object in a cluttered scene [2]. In the industries detecting objects With computers is difficult. This is as a result of some Complications such as Initially, products from industries vary in many forms such as type, size and Shape. Hence, using object detection system in the industries is only possible if the algorithm is robust enough to eliminate these variations. Subsequently, the industrial operation involves moving objects from one place to the other; this might result in objects being placed in different poses. The methods used needs to distinguish the objects with different orientations accurately.

Thirdly, the illumination in the environment varies due to, shadows of other machine or humans. Therefore the algorithms are required to function efficiently under different illumination condition.

Finally, partial occlusion can be a huge problem to the performance of the algorithms. Partial occlusion is inevitable in the industries since many products are processed at a time. To deal with this situation the algorithm needs to recognise an object even if it is partially occluded.

The actual idea behind this work is to utilize an algorithm that is computationally light, but it also shows high performance in detecting and classifying an object in an image. In this study, a comprehensive analysis of the algorithm was carried out using a more complex database and the emphasis will be on the industrial application.

On the one hand, HOG is utilized to extract image feature. However, HOG in not entirely invariant for rotation, therefore AMDF is employed to detect the image pose, which is then removed. Then, the features are passed to the MLP for training. While, on the other hand, after MLP is trained and then used to classify already aligned HOG features obtained from an input image.

2. Previous Studies Related To Object Detection

Yilmaz [3] grouped a lot of current studies related object detection methods. The first group is the segmentation based object detection methods. This is a concept which partitioned images into a conceptually identical region. Examples of such methods are Mean-shift [4], Graph-cut [5] and Active contour [6].

Secondly, the idea of some object detection methods is background based. In these methods, object detection is only possible by creating a representation of the scene known as the background model, and then searching for variation from the model for each incoming frame. Any clear deviation of an image region from the background indicates an object. Some of such methods are Mixture of Gaussian [7], Eigen background [8] and Dynamic texture background [9].

The last category of methods is point-based methods. They are methods that use points which both the horizontal and vertical orientation of their intensity varies rapidly. Such points are called the key points or interest points. Those points are invariant to changes in transformation and illumination. Commonly used detectors which are based on the interest points include Harris interest point detector [10], Scale Invariant Feature Transform (SIFT) [11] and Speed-Up Robust Features (SURF) [12].

The algorithm conceded in this research is a particular case of point-based methods known as the gradient-based method as it is based on the gradient of the pixels, which is related to the intensity of the points especially on the edge and vertex. Pedestrian detection system using the histogram of oriented gradients (HOG) [13] [14] is a well-known approach in this category. But, due to the scale variant and computationally complex nature of HOG, Murat Peker et al. [15] improved the approach by utilizing the whole image as a single cell in contrast to the actual idea which divides the input image into some cells. Then, AMDF is employed to correct the alignment defect. Later, the method was implemented on hardware for robotics education [1] and industrial laser cutting application [16].

3. Methodology

Feature Extraction

The 1-D mask is used to calculate the gradient (i.e. in horizontal and vertical direction) for the entire pixels on the image using equation 1 and 2 respectively. The entire image was considered as a single region instead of segmenting the image into many sub-regions. Hence, this increases the feature extraction throughput. Moreover, equation 3 and 4 were employed to calculate the gradient magnitude and orientation respectively.

Ultimately, the histogram is a vector of length equal to the number of bins, i.e. 360 in this case. Although, each element of the histogram corresponds to a bin (i.e. the category of orientation angle) and the values of the bins are the cumulative weights of gradient magnitude of all the pixels in the image. However, this is based on the values of their respective orientation angle. Although, on each pixel of the image, the value of its orientation angle determines the bin it belongs to. Hence, the magnitude of the pixel gradient is added to the value contained in that bin. But, the threshold value selected determines the pixel gradient to be added. In some instances the study utilized 0.2 gray levels threshold value; therefore, in this case, any pixel value with magnitude less than 0.2 gray levels are rejected*Units*

Alignment Correction

HOG is not completely invariant to changes in orientation, to correct this defect a successful match between the input and reference should be possible irrelevant to the degree of rotation. For this purpose, we propose to use; the Average Magnitude Difference Function (AMDF) algorithm. It is an algorithm used in speech processing to obtain a quantity which indicates

 https://doi.org/10.22362/ijcert/2018/v5/i2/v5i207

the dissimilarity between the speech waveform in a frame and its lagged version. In this study, HOG features are compared to a reference HOG vector using AMDF.
Equation (3) is utilized to obtain the AMDF difference. The AMDF of a signal S_i and it lagging version S_(j-τ) is given in the equation below where N is the number of bins.

$$f_{AMDF_{ij}}(\tau) = \frac{1}{N}\sum_{i=1}^{N}\left|S_i - S_{j-\tau}\right| \quad \tau = 0, 1, \dots N-1 \qquad (3.1)$$

Classification

Classification is a decision-theory for the identification of images. Image classifiers study the numerical properties of image features and arrange the data into classes. For classification algorithms such as Support Vector Machine [17], Neural Network [18] or Adaptive boosting [19], to properly function two phases of processing are involved: training and testing. In training phase, distinct properties of image features are secluded and then used to uniquely describe the classification categories, called training class. In the testing phase, these training classes are used to classify input image features. In this study, to compare the performance of the classifier AMDF, XCF and KNN will be employed to classify the images in the databases, although MLP is required for the complete implementation of the proposed algorithm.

Multi-layer Perceptron (MLP) was found to be one of the most popular among the Neural Network (NN) category. It is a network which comprises of neurons. Each pair of neurons is interconnected with weight, and all the neurons in the network are aligned in groups called layers. Hence, for classification or regression, each MLP constitute of; an input layer, one or more hidden layer and an output layer. In classification, the number of inputs and output is determined by the amount of data and classes involved, respectively. Therefore, the number of neurons in the input and the output layers determine the size of the input and output patterns.

It is a feed forward network where data at the inputs are propagated to the out by passing through all the neurons. That is, the value of any neuron in the hidden and output layers is the weighted linear combination of all the neuron in the previous layer. An activation function is employed to set some constrains on the output of each neuron. Among the activation functions, sigmoid function is well-known and widely used.
Where Zm is the input to the output layer node, Wo is the weighted connecting the hidden layer and the output layer, and m is the number of nodes in the previous layer, the kth output (Yk) is given as;

$$Y_k = \frac{1}{1+\exp(-f(x))} \qquad (3.17)$$

and, the summation function f(x) is;

$$(x) = \sum_{m=1}^{M} Wo_{mk} Z_m \qquad (3.2)$$

The complete network is shown in Figure 2.4.

For MLP to function effectively, training is required. Therefore, to train the algorithm, a set of training data are prepared, which contain inputs and their corresponding targets. All weights are set to initial random values. An input (Xn) is passed to the algorithm, and then the output (Yk) is compared to the corresponding target (tk). The difference between the output and target is the error (E). A cost function computes the error as shown in the equation 2.19, where, S is the amount of pattern in the data set, and L is the total number of the output nodes.

$$E = \frac{1}{2}\sum_{s=1}^{S}\sum_{n=1}^{L}\left(t_k^{(s)} - Y_k^{(s)}\right)^2 \qquad (3.3)$$

To achieve the desired results by properly training the network, back propagation algorithm was utilized. It is an algorithm which minimizes the error between an input and its targeted value by continuously adjusting the weighted connection between them iteratively. Hence, the aim of the training is to optimize the weight by reducing the error to a minimal value.

4. Simulations and Results

A. Experiments with Database I

In the initial stage, a database of images, which imitate the real industrial environment, was created. Hence, it contains image of ten different plane shape (i.e. star a star b, triangle a, triangle b, hexagon, parallelogram, circle and Pentagon. Each of the plane shapes was produced in different three scales. However, some images were rotated and deformed while others were partially occluded.

For the complete evaluation, the algorithm was combined with three different classifiers (i.e. ACF, AMDF and KNN) instead of the MLP classifier to prove the effectiveness of the algorithm with different classifiers. In these experiments, for KNN as a classifier; K = 1. The results obtained was summarized in Table III.

Table 1. The summary of the classifiers comparison on database I

Experiments	Overall Performance
The algorithm with KNN as a classifier	35.93%
The algorithm with AMDF as a	47.35%

classifier	
The algorithm with XCF as a classifier	50.06%

Table III presents the results obtained; it can be observed that on average more than 49% of the images were misclassified in all the experiments. Consequently, the overall performance was poor since only 50.06% success rate was recorded as the maximum.

B. *Experiments with Database I*

The essence of the study is to analyse the algorithm proposed by [Altun et al. 2013] an algorithm which will be used in the industrial environments. Therefore, to complete the evaluation of the algorithm, the research proceeded to test it in the real industrial environment. Hence, a database of some industrial products was generated. Database II consists of images obtained from an industrial scene which has a lot of variation in texture, size and type. Different lightening condition also poses a threat.

The images of the industrial products utilized are categorised into three different classes. Class A contains five different types of circular objects which are also of different size. An elliptical object which is captured in five different poses (i.e. 0, 72, 108, 144 and 180 degrees) formed the content of class B. Moreover, to create more confusion to the algorithm class B object were produced in two sizes. Class C consists of objects with four sides, i.e. a square and a parallelogram. Each of class C objects was rotated from 0 to 180degrees with the step size of 36degrees.

Furthermore, for each class, the images were captured in three different lightening conditions. Class A has 169 images while Class B and C contain 90 and 80 images respectively. Therefore, the database has 339 images in total.

The same sets of experiments were conducted on database II. The pre-processing step was applied but, no threshold value was selected. The summary of the results obtained is presented in Table VI.

Table 2. The summary of the classifiers comparison on database II

Experiments	Overall Performance
The algorithm with XCF as a classifier	10.32%
The algorithm with KNN as a classifier	71.39%
The algorithm with AMDF as a classifier	75.22%
The Complete Algorithm	98.10%

There was an improvement compared to the results obtained with the database I because the algorithm with AMDF and KNN as a classifier recorded 75.22% and 71.39% respectively. The performance of the algorithm with XCF as a classifier was very poor as only 10.32% success rate was recorded.

To compare the performance of the algorithm with that of other feature extraction algorithms, SIFT and SURF feature obtained on database II were classified by MLP. SIFT has 128 key points which were utilized in training MLP while SURF uses 64 key points. Table VII summarises the results obtained.

Table 3 Summary of sifting, surf and the proposed algorithm on database II

Experiments	Overall Performance
SIFT and ANN	53.42%
SURF and ANN	84.93%
The Complete Algorithm	98.10%

As can be seen from the result obtained, the combination of HOG, AMDF and the three classifiers (KNN, XCF and AMDF) implemented on the database I yield poor results, i.e. 35.93%, 47.35% and 50.06% respectively. The reason behind this is that the shape of some objects in the database appears irregular because they were heavily deformed. The edges of some objects were blunt due to the level of noise on some images and the presents of identical objects in different classes. All these pose a severe disturbance to the classifiers because; the corresponding HOG vector would be heavily deteriorated. Nonetheless, the proposed algorithm was able to accurately identify 92.62% of the objects in the database, i.e. 40% above the others. The experiments conducted also prove that the proposed algorithm does not respond well to very heavy image noise as its performance increased to 97.7% when the image noise was reduced to 0.01 on some images while others have 0.07 noise level.

There is an improvement on the results obtained on database II except experiment conducted with XCF as a classifier; this is because database II contains real-world images which are less complicated. Similarly, the results recorded by implementing HOG, AMDF and the three classifiers (ACF, AMDF and KNN) were less when compared with that of the algorithm. Hence, this is as a result of the complications in the database which is as follows: some white objects are captured on a white background under a very bright white light. Due to the low light intensity of some image, a black object is difficult to detect, and the presence shadows in some images also pose a threat.

The algorithm recorded a better performance even when compared with SIFT and SURF because, the database used contains the image with the light texture which causes to detect few, poor key points and descriptors. Therefore, despite it has been shown in some literature that SIFT and SURF perform well for complicated images, the objects in the database II are considered as the simple object with light texture. Hence, this also proves that within the contexts of this study, which tries to mimic real industrial environment, the proposed algorithm is better than SIFT and SURF because it shows higher performance. Table 6.7 gives the summarized comparison of the results.

5. Conclusion

This study aimed at evaluating a newly proposed algorithm that accurately detects objects in an industrial environment. The above-mentioned goal can only be achieved by evaluating the algorithm under some challenges such as partial occlusion, illumination variation etc. Hence this thesis proposed an algorithm that will accurately detect objects in an industrial scene. It is formed by the combination of HOG, AMDF and MLP NN. That is HOG is used to extract features from an image, then AMDF is employed to properly align the image. Finally, MLP NN classifies the features extracted.

For proper evaluation, two databases were generated. The first one contains images of plain shape but, to mimic the real world some object on the image is partially occluded, scaled, rotated and deformed while others are polluted with noise. The second database consists of images of real industrial products which are of different shapes and sizes, captured under different lightening condition.

The proposed algorithm was implemented by using four different classifiers which are Average Magnitude Difference Function (AMDF), K – Nearest Neighbour (KNN), Cross-correlation Function (XCF) and Multi-layer Perceptrons (MLP). Then it was compared with other feature extraction algorithms (i.e. Scale Invariant Feature Transform (SIFT) and Speed-Up Robust Features (SURF)) for further investigation.

The result of the experiments conducted reveals that the proposed algorithm performed well. Since 92.62% and 98.10% success rate were obtained when implemented with the database I & II respectively. Note that, within the context of this study, the proposed algorithm also performed better than SIFT and SURF though 53.42% SIFT and 84.93% SURF features were accurately classified.

References

[1] Demirci, B., Arslan, O., Tunaboylu, N. S., & Altun, H. (2013, May). Implementing HOG & AMDF based shape detection algorithm for computer vision & robotics education using LEGO Mindstorms NXT. In Technological Advances in Electrical, Electronics and Computer Engineering (TAEECE), 2013 International Conference on (pp. 288-293). IEEE.

[2] Ahsan, A. M., & Mohamad, D. B. (2013). Features Extraction for Object Detection Based on Interest Point. TELKOMNIKA Indonesian Journal of Electrical Engineering, (Vol. 11, No. 5, pp. 2716-2722).

[3] Yilmaz, A., Javed, O., & Shah, M. (2006). Object tracking: A survey. Acm computing surveys (CSUR), (Vol. 38, No. 4, pp. 13).

[4] Comaniciu, D., & Meer, P. (1999). Mean shift analysis and applications. In Computer Vision, 1999. The Proceedings of the Seventh IEEE International Conference on (Vol. 2, pp. 1197-1203). IEEE.

[5] Shi, J., & Malik, J. (2000). Normalized cuts and image segmentation. Pattern Analysis and Machine Intelligence, IEEE Transactions on, (Vol. 22, No. 8, pp. 888-905).

[6] Caselles, V., Kimmel, R., & Sapiro, G. (1997). Geodesic active contours. International journal of computer vision, (Vol. 22, No. 1, pp. 61-79).

[7] Stauffer, C., & Grimson, W. E. L. (2000). Learning patterns of activity using real-time tracking. Pattern Analysis and Machine Intelligence, IEEE Transactions on, (Vol. 22, No. 8, pp. 747-757).

[8] Oliver, N. M., Rosario, B., & Pentland, A. P. (2000). A Bayesian computer vision system for modelling human interactions. Pattern Analysis and Machine Intelligence, IEEE Transactions on, (Vol. 22, No. 8, pp. 831-843).

[9] Monnet, A., Mittal, A., Paragios, N., & Ramesh, V. (2003, October). Background modelling and subtraction of dynamic scenes. In Computer Vision, 2003. Proceedings. Ninth IEEE International Conference on (pp. 1305-1312). IEEE.

[10] Harris, C., & Stephens, M. (1988, August). A combined corner and edge detector. In Alvey vision conference (Vol. 15, p. 50).

[11] Lowe, D. G. (2004). Distinctive image features from scale-invariant keypoints. International journal of computer vision, (Vol. 60, No. 2, pp. 91-110).

[12] Bay, H., Ess, A., Tuytelaars, T., & Van Gool, L. (2008). Speeded-up robust features (SURF).

Computer vision and image understanding, (Vol. 110, No. 3, pp. 346-359).

[13] Shashua, A., Gdalyahu, Y., & Hayun, G. (2004, June). Pedestrian detection for driving assistance systems: Single-frame classification and system level performance. In Intelligent Vehicles Symposium, 2004 IEEE (pp. 1-6). IEEE.

[14] Dadal, N. and Triggs, B.: Finding People in Images and Videos, PhD thesis, French National Institute for Research in Computer Science and Control (INRIA), July 2006.

[15] Peker, M., Altun, H., & Karakaya, F. (2012, October). Hardware emulation of HOG and AMDF based scale and rotation invariant robust shape detection. In Engineering and Technology (ICET), 2012 International Conference on (pp. 1-5). IEEE.

[16] [16] Arslan, O., Demirci, B., Altun, H., & Tunaboylu, N. S. (2013, April). A novel rotation-invariant template matching based on HOG and AMDF for industrial laser cutting applications. In Mechatronics and its Applications (ISMA), 2013 9th International Symposium on (pp. 1-5). IEEE.

[17] Papageorgiou, C. P., Oren, M., & Poggio, T. (1998, January). A general framework for object detection. In Computer vision, 1998. Sixth international conference on (pp. 555-562). IEEE.

[18] Rowley, H. A., Baluja, S., & Kanade, T. (1998). Neural network-based face detection. Pattern Analysis and Machine Intelligence, IEEE Transactions on, (Vol. 20, No. 1, pp. 23-38.

[19] Viola, P., Jones, M. J., & Snow, D. (2003, October). Detecting pedestrians using patterns of motion and appearance. In Computer Vision, 2003. Proceedings. Ninth IEEE International Conference on (pp. 734-741). IEEE.

Authors Profile

F. S. Ishaq received his B.Eng. Degree in Computer Engineering from Bayero University, Kano, Nigeria, in 2010 and the M.Sc. degree in Computer Engineering from Mevlana University, Turkey in 2014. He is now an assistant lecturer with Department of Mathematics and Computer Science, Federal University, Kashere, Gombe, Nigeria from 2015 to date. His research interests include Image processing.

Ibrahim Abubakar Alhaji received his B.Sc. degree in Physics from Bayero University, Kano, Nigeria, in 2009 and the M.Sc. degree in Electronics from same University in 2015. He is an assistant lecturer with Department of Physics, Federal University, Kashere, Gombe, Nigeria from 2013 to date. His research interests include Computational Electromagnetics and Microwave and Optical Techniques.