VOL. 12, NO. 8, APRIL 2017 ISSN 1819-6608
ARPN Journal of Engineering and Applied Sciences

www.arpnjournals.com

DETECTION OF MULTIPLE MANGOES USING HISTOGRAM OF ORIENTED GRADIENT TECHNIQUE IN AERIAL MONITORING

Nursabillilah Mohd Ali[1,2], Mohd Safirin Karis[1,2], Nur Maisarah Mohd Sobran[1,2], Mohd Bazli Bahar[1,2], Oh Kok Ken[1,2], Masrullizam Mat Ibrahim[3] and Nurul Fatiha Johan[1,2]
[1]Faculty of Electrical Engineering, Universiti Teknikal Malaysia Melaka, Durian Tunggal, Melaka, Malaysia
[2]Robotics and Industrial Application Automation Research Laboratory, Durian Tunggal, Melaka, Malaysia
[3]Faculty of Electronics and Computer Engineering, Universiti Teknikal Malaysia Melaka, Melaka, Malaysia
E-Mail: nursabilillah@utem.edu.my

ABSTRACT

The project uses shape identification algorithm and Histogram of Oriented Gradient principle to detect and count the total number of mango on its tree using a quad copter with an attachable webcam. The traditional method in harvesting mango has its limitation which leads to the degradation of harvested mango's quality. As a result, the rate of production and the structure of the tree will be dampening. Hence, usage of image processing algorithm could be a solution for a better and more precise mango's pre-harvesting process. It differentiates the mango and its leaf based on the images captured on real scene and thus forecast the growth rate of the mango tree for time being. Tallness of the mango tree and location of mango would not affect farmer's capability to inspect the mango as the drone hovers according to user's intention. It is expected to provide an alternate review for the mango grower, agricultural developer and investor.

Keywords: harvesting, histogram of oriented gradient, image processing, shape.

INTRODUCTION

Harvesting is one of the main processes in agriculture sector. In harvesting a mango, ripeness of the fruit will determine the output taste either sours, sweet, creamy or mild bitterness. The normal inspection of mangoes ripeness usually done based on manual inspection by the farmer. However, the height of a mango tree, abundant of branches and leaves might hindrance the inspection process. With this motivation, we apply the Histogram of Oriented Gradient method in detecting multiple mangoes at the tree branches and implementing the quad copter as a mechanism in reaching the height of the tree.

Researches in harvesting mango in agriculture cover wide area. Major researches are in detecting the level of ripeness of the mango in [1, 2]. After that, issue in grading or classifying the mangos according to the detected ripeness as mention in [3] and [4]. Other than that, problems arise in plucking mango in [5] and peeling numbers of mango in [6]. There is also large scale of research done in [7] for monitoring the green house plant for Harum Manis Mango. This project focus more on detecting mangoes at branches for monitoring process before the harvesting were done. The idea is to provide early observation to the person in charge before decide whether or not to pluck the mango, while covering the restricted sight that the person in charge had due to mango trees height.

In [8] detect mango in the presence of leaves and branches using Randomized Hough Transform and Backpropagation Neural Network, the same method applied by [9]. In [10] implementing the Euclidean distance based classifier in detecting apples at the branch, whereas in [11] used centroid based detection in recognizing orange at the tree. This project proposed histogram of oriented gradient (HoG) based mango detection as algorithm in finding mangoes in the captured images.

DETECTION USING HISTOGRAM OF ORIENTED GRADIENT (HoG)

Histogram of Oriented Gradient descriptor is actually a feature descriptor widely used in machine vision field for the purpose of detection of object. It calculates the number of occurrences of gradient orientation in localized parts of an image. The main point of HoG is the distribution of intensity gradient and edge directions. It could be done by dividing the image into small connected regions, each compiled with histogram of gradient direction and edge orientations for the pixels involved. Hence, the histograms merged to become a descriptor. The performance could also be improved by contrast-normalizing the local histogram by computing a measure of intensity across a region within the image (a block). The block would normalize the smaller connected region cells within itself. Shadowing or imperfect illumination would probably be reduced by the normalization. The next descriptions of HoF technique are based on [12, 13].

In order to obtain the feature descriptor, the first step taken is filtering the grayscale image I with the following kernels as in Equation. (1) and (2).

$$D_x = \begin{bmatrix} -1 & 0 & 1 \end{bmatrix} \tag{1}$$

$$D_y = \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix} \tag{2}$$

Figure-1. Example of grayscale image and its corresponding filtered images [12].

Next, computing the x and y derivatives using a convolution operation with:

$$I_x = I \times D_x \tag{3}$$

$$I_y = I \times D_y \tag{4}$$

The magnitude of the gradient:

$$|G| = \sqrt{I_x^{\,2} + I_y^{\,2}} \tag{5}$$

The orientation of the gradient:

$$\theta = \tan^{-1} \frac{I_y}{I_x} \tag{6}$$

Each pixel carries a directional vote for an orientated-based histogram channel based on the values found in the calculation. These cells are rectangular and the channels points in 0 to 180 degrees or 0 to 360 degrees. The author found that unsigned gradients used with 9 histogram channels would offer the best result.

Figure-2. Example of orientated-based histogram channel [12].

The gradient strengths must be locally normalized. The cells have to group together in a larger connected block. Hence, the HoG descriptor is the main vector component. The overlapping blocks determine the final descriptors. The descriptor blocks represented by a rectangular type and circular type. The final way is to normalize the blocks and execute similarity measure between featured vectors.

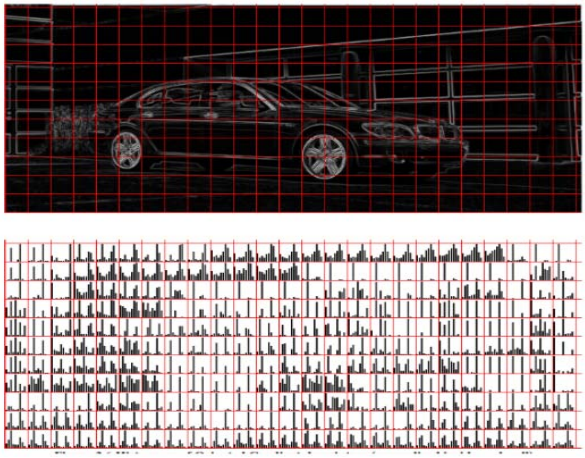

Figure-3. Example of rectangular type blocks [12].

METHODOLOGY

Under variable light intensity condition, the fabricated prototype of the quadcopter undergone outdoor flight test to capture relevant images around a mango tree. The main purpose of this experiment is to conduct shape analysis and detection under non-fixed light intensity environment. The prototype consists of bundle of pre-programmed Raspberry Pi 2, battery, plastic casing, Logitech Webcam C310 and the quadcopter. Cable tights were used to bundle the prototype together. The experiment could proceed whenever the targeted mango tree is chosen. Upon connected the Wi-Fi, user can now watch the real live scene on PC. Images could be captured and saved into the PC [14-18].

Figure-4. Image capturing process with a quadcopter.

The captured images would undergo a process named supervised learning which operates based on the principle of feature extraction with Histogram of Oriented Gradient (HoG). The process firstly required a collection of positive image samples which comprises of mango in different sizes. User has to label the corrected or needed region of interest, which is significantly different from its background. The application with complete GUI is available in MATLAB library known as Training Image Labeler. Figure-5 shows training using Matlab Labeler.

ARPN Journal of Engineering and Applied Sciences

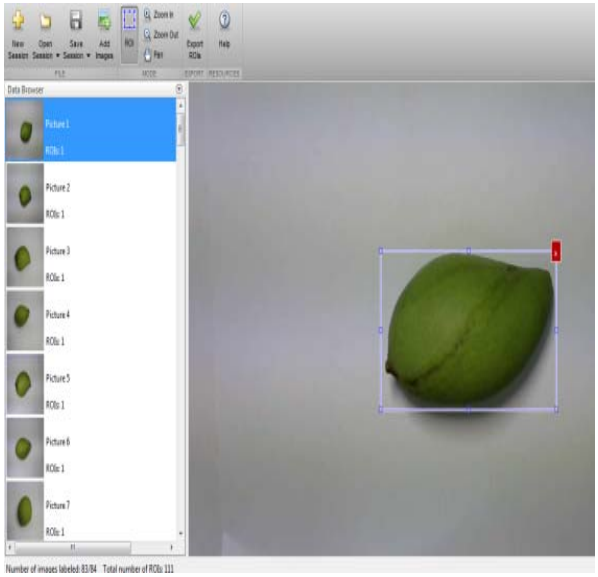

Figure-5. Training image labeler in MATLAB.

Next, the collection of region of interest files would be exported to the main process beforehand. User has to prepare a collection of negative samples to assist the system to classify the mango properly. Within this sample of negative images, it contained no similar object like mango. For instances, the images were bus, motorcycles, human frontal face, watch, key and rubber duck.

In order to execute the process to train the weak learners, user has to predefine number of training stages and its false positive rate as well. In the process, cascaded classifier consists of stages where each stage is an ensemble of weak learners. In addition, each stage is trained with boosting which provides the ability to train a highly accurate classifier by taking a weighted average of the decision made by the weak learners. Hence, the stages were designed to reject negative samples in soonest possible manner.

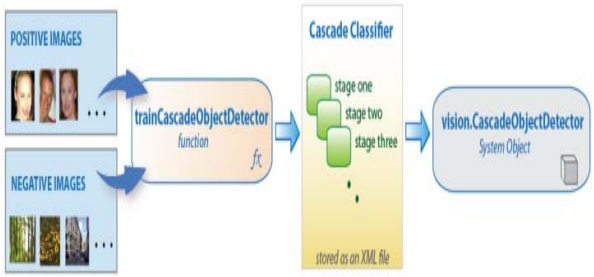

Figure-6. Cascaded training with different number of stages in MATLAB IDE.

Firstly, the function calculates the amount of positive samples needed before undergone training. The stage one would normally computes the number of positive samples that is lesser than the total number of provided by the user. In order to balance the amount of positive samples used, the almost double the number of negative samples would be generated [19-20].

Once it has passed stage one, the second stage would enable the weak learner to identify the positive samples from previous stage [21-25]. It classifies all positive samples, whereas negative samples would be discarded. The process carried on for the next stages until the last is done. The precaution step is that small amount of positive and negative images would somehow lead to termination of training. For example, the pre-defined stage is seven; however the training halted at stage four. There was inadequate amount of input images to compromise with the training required.

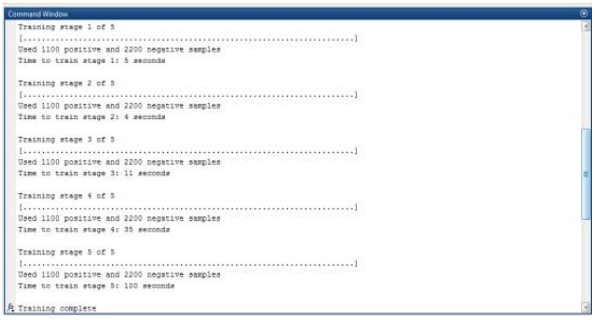

Figure-7. Training complete.

Based on the output result, it is expected to foresee three possible groups of regions. These three are true positive region, false positive region and a false negative region. True positive indicate a positive sample (mango) is classified correctly. In contrast, false positive happens when a negative sample is mistakenly classified as positive sample. The last non-wanted condition is where a positive sample is mistakenly labelled as negative sample.

The test is conducted with different number of training stages undergo, false positive rate used and the number of positive samples used with a fixed number of negative samples. Each of the tests performed is supposed to obtain the detection rate and number of detected false positive region.

MATERIALS AND METHODS

The number of region of interest used is roughly 600 positive samples. As the number of training cascaded stages increases, the detection rate would show decrement. Meanwhile, misclassification rate had shown increment.

ARPN Journal of Engineering and Applied Sciences

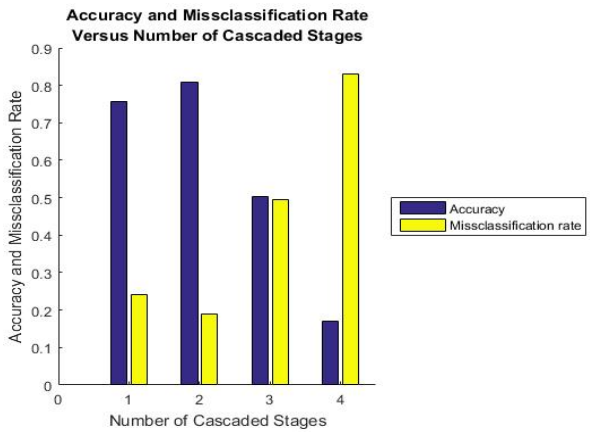

Figure-8. Accuracy and misclassification rate versus number of cascaded stages.

The number of false positive detection reduces as the number of training cascading stages increases. The process had shown a nearly zero false positive detection for four stages. The highest detection rate (80.95%) has recorded when only two training cascade stages are used.

At the maximum usage of four training stages, only 17.00% detection rate is accounted.

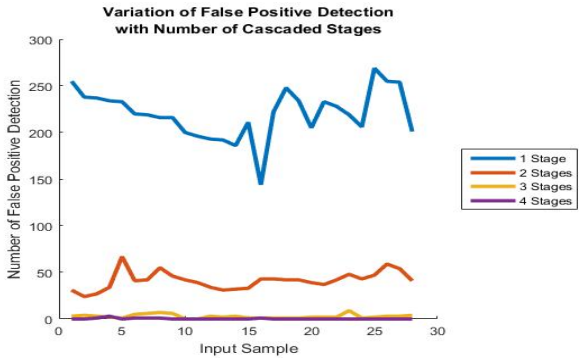

Figure-9. Variation of false positive detection with number of cascaded stages.

The detection rate increases and misclassification rate increases as false positive rate increases in the range of 0.01 to 0.10. Misclassification rate shows the opposite way.

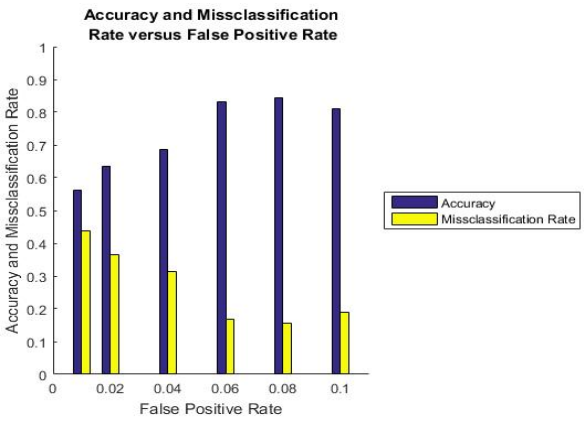

Figure-10. Accuracy and misclassification rate versus false positive rate.

As false positive rate increases, the number of false positive detection increases.

The increasing number of stages would increases the false negative rate or decreases true positive rate (detection rate). The most significant reason is because it provokes a higher chance of rejecting a positive sample by mistake. The training time also would increase as the number of cascaded stages rises. Generally, a positive result from the first classifier triggers the evaluation of a second classifier which has also a high level of detection rate. Thus, the loops carried on for the next subsequent stages. Thus, a negative outcome would bring immediate rejection of the process.

The graph in Figure-11 shows the variation of number of false positive detection with the number of cascaded stages. In the experiments, 25 of the best samples have used to illustrate the differences. The number of false positive detection is the highest only happened if the system uses only one training stage to classify them. In contrast, the number of false positive detection somehow goes to a lower level (approximately zero) as the number of cascaded training increases to four. Whereas, Figure-12 shows output image with false positive rate for multiple mangoes.

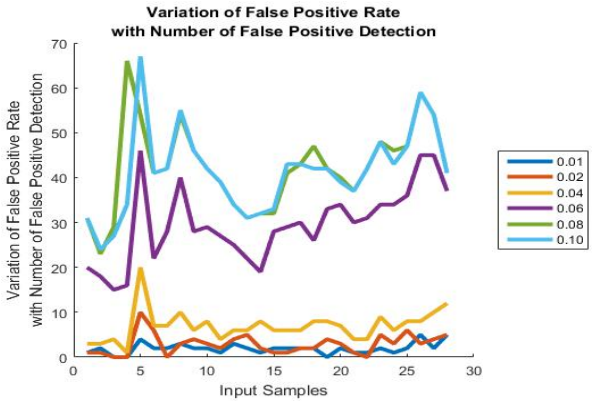

Figure-11. Variation of false positive rate with number of false positive detection.

No. Training Stages	False Positive Rate	Output Result
1	0.1	
2	0.1	
3	0.1	
4	0.1	

Figure-12. Output images.

False positive rate determines how high the possibility of detecting an incorrect object based on the training. A low false positive rate disallows the learner to mistakenly classify the incorrect object. For instance, the system has disregards a mango as positive at first stage because of low false positive rate. In contrast, a relatively higher false positive rate permits the learner to commit more erroneous than usual case. In this case, higher false positive rate deemed to be a more secure way of not neglecting any detectable object. Number of positive rate will be increased if the lighting of the environment is changing. It is illustrated in Figure-13.

No. Training Stages	False Positive Rate	Output Result
3	0.1	
3	0.2	
3	0.3	
3	0.4	

Figure-13. Output images.

In short, a large training set (more than thousand) should have sufficient positive samples to undergo more training stages. The reason is more negative images would be generated to compensate the amount of positive images for each stages of training. Besides, higher false positive rate should be implemented into the training [26-29]. A higher false positive rate permits more false detection (mango leaf) to be labeled as true detection (mango) in each stage. As the training stages gradually increases, false detection would reduce drastically. It means the process is recoverable for the first few training stages before reaching the last stage.

In few trials, there is a chance of missing an object which is strictly undesirable. For instance, there were mangoes ignored into calculation in stage one or two. Hence, the solution is to increase the true positive rate. A high true positive rate prohibits from achieving the desired false positive rate per stage, thus causing the detector more likely to predict false detection [30-31]. The scenario is equivalent to increasing false positive rate. However the obvious drawback is number of false detection would increase tremendously. User then should fine-tune the system by increases the number of training stages to overcome the effect.

CONCLUSIONS

The merging of machine vision technology into agricultural sector has resulted in improvement of harvested crop's quality. One of the most obvious advantages of conducting technical research over the sampled object is capable to understand its feature properly before making comparison with upcoming experiment. With the usage of MATLAB, the Image Processing Toolbox and Computer Vision Toolbox assist in the feature extraction as well as classification wise. The result indicates the potential usage of this monitoring tool in object (mango) detection from a given static image. Quality of result is much dependent on the surrounding tree leaves and intensity of light. The greater the number of samples used in first stage of feature extraction and segmentation would lead to higher performance of the system. For instance, the number of false positive value would be lessened and true positive value would increase. Thus, accuracy of the system is expected to rise after there is an increment of input samples. The greater the sample inputs for Training Classifier in Computer Vision System Toolbox, the higher the accuracy of the system.

ACKNOWLEDGEMENT

The authors would like to thank to the Robotics and Industrial Automation (RIA) research group under Center of Robotics and Industrial Automation (CeRIA) and Centre for Research and Innovation Management (CRIM) of Universiti Teknikal Malaysia for providing tools to complete this project. This work was funded by RAGS/1/2014/TK03/FKE/B00054.

REFERENCES

[1] Salunkhe R. P. and Patil A. A. 2015. Image processing for mango ripening stage detection: RGB and HSV method. In: 3rd International Conference on Image Information Processing. pp. 362-365.

[2] Naik S., Patel B. and Pandey R. 2015. Shape, size and maturity features extraction with fuzzy classifier for non-destructive mango (Mangifera Indica L., cv. Kesar) grading. In: IEEE Technological Innovation in ICT for Agriculture and Rural Development. pp. 1-7.

[3] Pandey R., Gamit N. and Naik S. 2014. A novel non-destructive grading method for Mango (Mangifera Indica L.) using fuzzy expert system. In: International Conference on Advances in Computing, Communications and Informatics. pp. 1087-1094.

[4] C. S. Nandi, B. Tudu and C. Koley. 2014. A machine vision-based maturity prediction system for sorting of harvested mangoes. IEEE Transactions on Instrumentation and Measurement. 63(7): 1722-1730.

[5] Konam S. 2014. Agricultural aid for mango cutting (AAM). In International Conference on Advances in Computing, Communications and Informatics. pp. 1520-1524.

[6] Donado R., Tan L. H., Fernandez D. M., Calilung E. J., Francia D. A. and Dadios E. P. 2015. Design, fabrication and testing of a semi-auto green mango peeling machine. In: International Conference on Humanoid, Nanotechnology, Information Technology, Communication and Control, Environment and Management. pp. 1-5.

[7] Saad S. M., Kamarudin L. M., Kamarudin K., Nooriman W. M., Mamduh S. M., Zakaria A., Shakaff A. Y. M. and Jaafar M. N. 2014. A real-time greenhouse monitoring system for mango with Wireless Sensor Network (WSN).In: 2nd International Conference on Electronic Design. pp. 521-526.

[8] Nanaa K., Rizon M., Rahman M. N. A., Ibrahim Y. and Aziz A. Z. A. 2014. Detecting mango fruits by using randomized Hough transform and backpropagation neural network. In: 18th International Conference on Information Visualisation. pp. 388-391.

[9] Yacob Y. M., Shaiful A. R. A. M., Husin Z., Farook R. S. and Aziz A. H. A. 2008. Harum Manis Mango weevil infestation classification using backpropagation neural network. In: International Conference on Electronic Design. pp. 1-6.

[10] Syal A., Garg D. and Sharma S. 2014. Apple fruit detection and counting using computer vision techniques. In: IEEE International Conference on Computational Intelligence and Computing Research. pp. 1-6.

[11] Çakır Y., Kırcı M., Güneş E. O. and Üstündağ B. B. 2013. Detection of oranges in outdoor conditions. In: 2nd International Conference on Agro-Geoinformatics. pp. 500-503.

[12] Bertozzi M., Broggi A. and Castelluccio S. 1997. A real-time oriented system for vehicle detection. Journal of Systems Architecture. 43(1-5): 317-325.

[13] Dalal N. and Triggs B. 2005. Histogram of oriented gradient for human detection. In: IEEE Computer Society Conference on Computer Vision and Pattern Recognition. pp. 886-893.

[14] V. Singh, V. Shokeen and B. Singh. 2013. Comparison of feature extraction algorithms for gender classification from faces. International Journal of Engineering Research and Technology. 2(5): 1313-1318.

[15] MathWorks. 2015b. Quick start to MathWork's MATLAB and Simulink. MathWorks Inc., Massachusetts, USA.

[16] Rafael C. and Gonzalez R. E. 2011. Digital image processing. Pearson Education, London, UK.

[17] Belbachir A. N. 2009. Smart camera. Springer Science and Business Media, Berlin, Germany.

[18] Forsyth D. A. and Ponce J. 2012. Computer vision: A modern approach. Pearson, London, UK.

[19] MathWorks. 2014. Image processing user's guide. MathWorks Inc., Massachusetts, USA.

[20] Shapiro L. and Linda G. 2001. Computer vision. Prentice-Hall, New Jersey, USA.

[21] da Fona C. L. and Cesar J. R. M. 2009. Shape classification and analysis: Theory and practice. CRC Press, Florida, USA.

[22] Ramesh J., Rangachar K. and Brian G. S. 1995. Machine vision: Chapter 4-Image filtering. McGraw-Hill, New York, USA.

[23] Rastislav L. and Plataniotis K. N. 2006. Color image processing: Methods and applications. CRC Press, Florida, USA.

[24] Umbaugh S. E. 2010. Digital image processing and analysis: Human and computer vision applications with CVIP tools. CRC Press, Florida, USA.

[25] MathWorks. 2015b. Computer vision user's guide. MathWorks Inc., Massachusetts, USA.

[26] N. M. Ali, M. S. Karis, A. F. Z. Abidin, B. Bakri and N. R. A. Razif. 2015. Traffic sign detection and recognition: Review and analysis. Journal Teknologi 77(20): 107-113.

[27] N. M. Ali, Y. M. Mustafah and N. K. A. M. Rashid. 2013. Performance analysis of robust road sign identification. IOP Conference Series: Materials Science and Engineering. 53(1): 1-6.

[28] Ali N. M., Karis M. S. and Safei J. 2014. Hidden nodes of neural network: useful application in traffic sign recognition. In: IEEE Conference on Smart Instrumentation, Measurement and Applications. pp. 1-4.

[29] N. M. Ali, N. K. A. Md Rashid and Y. M. Mustafah.2013. Performance comparison between RGB and HSV color segmentations for road sign detection. Applied Mechanics and Materials. 393: 550-555.

[30] Ali N. M., Jun S. W., Karis M. S., Ghazaly M. M. and Aras M. S. M. 2016. Object classification and recognition using bag-of-words (BOW) model. In: IEEE 12th International Colloquium on Signal Processing and Its Applications. pp. 216-220.

[31] Karis M. S., Razif N. R. A., Ali N. M., Rosli MA., Aras M. S. M. and Ghazaly M. M. 2016. Local binary pattern (LBP) with application to variant object detection: A survey and method. In: IEEE 12thInternational Colloquium on Signal Processing and Its Applications. pp. 221-226.