การเพิ่มสมรรถนะของการตรวจจับใบหน้าด้วยโมเดลจำแนกสองระดับ

Improving Face Detection with Bi-Level Classification Model

Tossapon Boongoen [1] Natthakan Iam-On [2] and Boonlert Undara [3]

[1]กองวิชาคณิตศาสตร์และคอมพิวเตอร์ กองการศึกษา โรงเรียนนายเรืออากาศนวมินทกษัตริยาธิราช

[2]สำนักวิชาเทคโนโลยีสารสนเทศ มหาวิทยาลัยแม่ฟ้าหลวง

[3]กองวิชาวิศวกรรมเครื่องกลและอุตสาหการ กองการศึกษา โรงเรียนนายเรืออากาศนวมินทกษัตริยาธิราช

E-Mail: tossapon_b@rtaf.mi.th, nt.iamon@gmail.com, boonlert_un@rtaf.mi.th

บทคัดย่อ

ปัญหาผู้ก่อความรุนแรงในสามจังหวัดชายแดนภาคใต้ เป็นปัญหาสำคัญด้านความมั่นคงของชาติ ความปลอดภัยในชีวิต รวมถึงทรัพย์สินของข้าราชการและประชาชนในพื้นที่ อีกทั้งยังส่งผลต่อความสูญเสียทางด้านเศรษฐกิจ การลงทุนและการท่องเที่ยว ทั้งนี้รัฐบาลและกระทรวงกลาโหมได้ร่วมกันแก้ไขปัญหาดังกล่าว แต่ได้ผลในขอบเขตที่จำกัด ยังคงมีเหตุการณ์วางระเบิดในที่สาธารณะและซุ่มโจมตีเจ้าหน้าที่ของรัฐอยู่เป็นระยะ แนวทางการแก้ปัญหาในปัจจุบันขาดการประยุกต์ใช้เทคโนโลยีอย่างเป็นรูปธรรม ทั้งระดับผู้ปฏิบัติแลในระดับ การวางแผนบังคับบัญชา อีกทั้งการข่าวที่ยังไม่เป็นเอกภาพ ขาดการบูรณาการระหว่างหน่วยงานต่าง ๆ ที่เกี่ยวข้อง งานวิจัยที่ได้นำเสนอในบทความนี้มีจุดมุ่งหมายที่จะแก้ไขปัญหาข้างต้นด้วยการนำข้อมูลภาพจากกล้องวงจรปิด ไปเป็นพื้นฐานของการปฏิบัติงานเชิงรุก โดยประยุกต์ใช้เทคโนโลยีรู้จำใบหน้าบุคคลเพื่อประโยชน์ในการรักษาความปลอดภัย ความแม่นยำของการรู้จำใบหน้านั้นขึ้นอยู่กับคุณภาพของผลลัพธ์จากขั้นการตรวจจับใบหน้า ที่จะทำการระบุพื้นที่ในภาพที่ศึกษาซึ่งมีคุณลักษณะของใบหน้าบุคคลปรากฏอยู่ ทั้งนี้วิธีการตรวจจับที่เป็นมาตรฐานของงานแขนงนี้ มักจะใช้โมเดลจำแนกประเภทของข้อมูลที่สร้างได้จากภาพตัวอย่างที่ใช่และที่ไม่ใช่ใบหน้า โมเดลดังกล่าวมีประสิทธิภาพในการจำแนกข้อมูลในระดับที่สูง แต่หากระดับความถูกต้องนั้นจะแปรผันตามคุณภาพและจำนวนของภาพตัวอย่างที่ใช้ในการพัฒนาโมเดล อนึ่งโมเดลจะมีความแม่นยำในการตรวจจับต่ำถ้าคุณภาพหรือปริมาณของภาพตัวอย่างไม่ดีหรือน้อยเกินไป เพื่อแก้ไขปัญหาดังกล่าวบทความนี้จึงได้นำเสนอโมเดลจำแนกประเภทข้อมูลแบบสองระดับ ซึ่งเป็นการใช้วิธีการที่มีความถูกต้องสูงมาเสริมการทำงานของโมเดลเดิม นอกจากนั้นโมเดลใหม่ยังสามารถแก้ไขกรณีของภาพใด ๆ ที่ไม่ใช่ใบหน้า แต่ถูกจำแนกว่าเป็นใบหน้า (false positive) ได้อย่างเป็นระบบอีกด้วย

คำสำคัญ : เหตุการณ์ความไม่สงบ, การรู้จำใบหน้า, การตรวจจับใบหน้า, การจำแนกประเภทข้อมูล

Abstract

The event of unrest in the three southern provinces has long been a non-trivial burden to national security. The series of harmful acts causes a high-toll lost in both official and civilian lives, yet damages to personal as well as public possesssions. For some time, the government, espcially the Ministry of Defence, pays a great deal of effort to resolve the situation and restore peace in the area. This proves to be effective only to a certain extend, while car bombs and attacks on officials continue. The overall strategy set for this problem still lacks the synergy of advanced technology made available in academic and research communities. Intelligence acqusition and sharing seems to be the weak link in the ideal problem-solving scheme. This article reports an improvement made towards such end, with face recognition being exploited to identify individuals from CCTV images. The precision of this tool depends highly on the detected image area that is thought to be a human face. A benchmark technique to face detection is prone to errors, due to quality of magnitude of sample face positive and negative samples. To overcome this, a new bi-level model is proposed to improve accuracy and reduce the amount of false positives.

Keywords : event of unrest, face recognition, face detection, classification.

1. Introduction

In this era of security-led society, a number of biometrics tools such as face and fingerprint recognition have been the center of attention, for both academic and industrial perspectives. Specific to the former, it is widely realized as a benchmark for both person verification and identification [17]. The landscape of its applications spans over the area of authentication, security, social network, crime and anti-terrorism. This intelligent has proven useful for several unrest events around the globe. In particular to the case of Thai southern provinces, a prototype is created for the purpose of point-based security, with the research grant of Department of Science and Technology, Ministry of Defence [1]. It can play a major role in filtering likely suspects in various public places. As such, this may allow a terrorist act to be intervened or even prevented. This attempt is perfectly in line with the ambition of Federal Bureau of Investigation (FBI) to launch a face-recognition capability that covers all US residents. It is implemented under the project called

Next Generation Identification (NGI), which was founded in 2013 with an initial budget of one billion US dollars [2].

In general, a face recognition system is composed of two main functional components, face detection and face recognition processes [3]. The former is to automatically find an area within the input image that may represent a human face. The latter searches a repository of face data for a match to that captured by the detection phase. See Fig. 1 for the overview of a face recognition system.

As for the task of recognition, several face recognition models initially exploit a simple geometric template in order to specify individuals. The subject has prospered over the years with a number of advanced representation and matching methods being investigated [18]. One of the pioneer works that has been regarded as a benchmark data-driven approach, is the inclusion of PCA (Principle Component Analysis) to draw significant features from the original matrix of face data [19], [20]. In that, PCA looks for a linear projection that best fits the data under examination. The resulting representation of reduced dimensionality, so-called Eigen face, leads to efficient data management and matching. Despite this, it is inherently not able to extract nonlinear characteristics of the data, which is defined by beyond second order statistics. To this end, different nonlinear algorithms have been put forward and coupled with the problem of face recognition. These include KPCA (Kernel PCA [21]), LLE (Local Linear Embedding [22]) and Isomap [23]. Recently, the use of ensemble clustering for face classification is introduced [4], with the accuracy level being higher than those obtained by PCA and other alike methods mentioned above.

As turn to the task of face detection, a pile of research studies have contributed to this course. These include four groups of detection models [5]: knowledge based, feature invariant, template matching, and appearance based. The last family has received a great deal of attention in the past decade given the fact that it uses machine learning or data mining technology [24]. With the fast growing pace of the two fields, face detection has become an important application, to which advanced techniques are adopted and assessed.

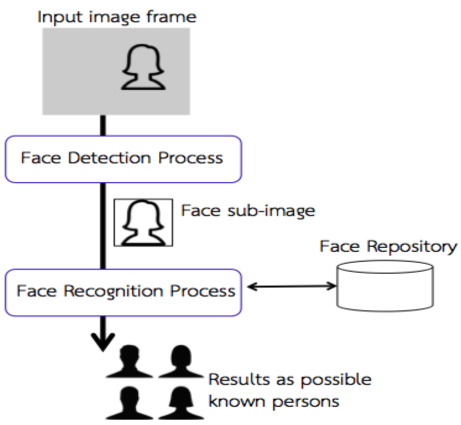

Fig 1 Overview of face recognition system.

The common and widely accepted alternative for an appearance-based method is the Viola-Jones model [6]. It is based on cascaded classifier, in which a number of weak classification models are ordered in a sequence. Based on a specific image-feature template, each model categorizes the input image (i.e., specific area within a given image) as a non-face or a possible face. The ultimate solution as a face is received through all classifiers, one by one. Failing at any hurdle will result in a non-face outcome. In spite of a real-time response, the quality of cascaded classifiers is highly subject to its generation from a set of both positive and negative samples [7]. Both quantity and goodness of these training instances dictate the accuracy of face detection. In order to overcome the limitation, this paper introduces a new bi-level classification model, where an accurate classifier is amended to the existing cascaded framework, such that the resulting is more robust to noise and perform consistently across different training sets.

The rest of this paper is organized as follows. Section 2 introduces concepts related to face detection, especially the Viola-Jones method that is extended in this research. The proposed model is presented in Section 3, with details of the additional classification and ensemble directed data transformation. After that, Section 4 includes experimental design and results. Section 5 provides conclusion with perspective of future work.

2. Face Detection and Related Work

Face detection is considered as the primary stage of face recognition and identification. It plays an important, yet difficult role, given a variety of variables such as face size, position and posture, illumination and emotion. In addition, the desired detection system should be efficient, accurate and suitable for a real-time application. All these requirements have pushed the research forward, with a large number of works introduced in the literature. To set the scene for this research, the section provides both brief reviews on the overview of face detection methods and that of Viola-Jones, respectively.

2.1 Families of face detection methods

According to the study of Yang et al [5], there are four categories of face detection methods. These are

- Knowledge-based methods: are based on human expertise and knowledge regarding face components and their properties. Examples of knowledge-based models [8, 9] focus on the pixel intensity profile both in vertical and horizontal domains. It is assumed that a face should consist of two eyes on the same horizontal level, one nose, one mouth, and the distance measures among these components are relatively constant. However, this set of assumptions may result on only a few candidates being accepted as face images. Yet, reducing those constraints may lead to a large number of false positives.
- Feature-invariant methods: have been developed to seek unique features that can be used to differentiate face images. This approach is motivated by the human capability to recognize objects by properties, which are invariant to position, location and environmental setting. Specific to the investigation of Yow and Cipolla [10], different unique features such as eyebrow, eye, nose, mouth and hair can be captured using an edge detection algorithm. The statistical significant of these extracted features can confirm the discovery of face image. Nevertheless, the aforementioned features may be directly distorted by illumination, noise and obstruction. It is also true that the search for these features can be hard on several cases.
- Template-matching methods: make use of a standard template of face image, which is created manually or by a mathematical function. For instance, Scassellati [11] has introduced the face template as definitions of 16 image areas and 23 relations among those areas. For any input image, its correlation value with the standard template is measured, in order to estimate the similarity (i.e., confidence of being a face image). This method proves to be simple, but its accuracy is not what has been expected. The use of template is prone to errors such as image size, position, image elements, and shape of face.

- Appearance-based methods: make use of the model generated from a collection of positive and negative samples (i.e., face and non-face images). In general, the generation of such a model is obtained using statistical or machine learning techniques. According to the study of Sung and Pogio [12], Gaussian function has been used to derive the density of face and non-face images, which can be employed as references to determine the classification outcome. Most of the appearance-based methods are reported to perform better those from the other groups, especially the so-called Viola-Jones method [13, 14]. Accounting for its efficiency and accuracy, Viola-Jones has been realized in many real-world applications, e.g., image caption in a digital camera, image annotation and management software.

2.2 Viola-Jones method

This method consists of three steps: i) estimation of feature templates using the integral image framework, ii) searching for feature-specific weak classifiers using Adaptive boost (i.e., Adaboost), and iii) develop a cascaded classifier.

- **Step1:** At the initial stage, an input image is divided into a number of sub-windows, each of which will be searched for a possible existence of face. Specific to the Viola-Jones method, the Haar-like features are designed to capture spatial-intensity properties of both face and non-face images. Examples of these area-based features are given in Fig. 2, where they can be of different sizes and details. In the estimation process, every feature is scaled to match the sub-window under examination. Then, the summation of intensity is calculated for each area specified in the feature (i.e., shaded and unshaded areas correspond to those with high and low intensity levels, respectively). With this terminology, a particular feature will be included in the classification model is it is capable of differentiating between face and non-face samples.

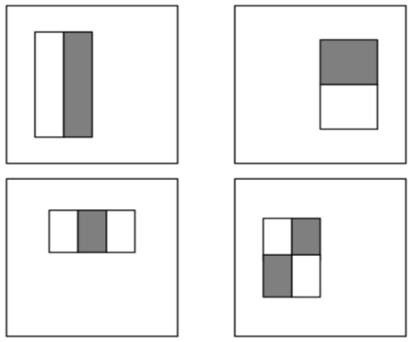

Fig 2 Examples of Haar-like features [1].

- **Step2:** Those pre-specified features are exploited within the Adaboost process, where a set of weak classifiers is created with distinct levels of significance. Supposed that S is a set of n training samples (x_i, d_i), where x_i denotes the i-th sample and d_i is the corresponding flag (d_i = 1 if the sample is face, and d_i = 0, otherwise).

$$S = \{(x_i, d_i), i = 1 \ldots n\} \quad (1)$$

Then, define weak classifiers C_t, t = 1, …, T that are in line with the number of process iteration t = 1, …, T. The weight w_{it} for the pair of t-th classifier and i-th sample is estimated by Eq. 2. Note that m and l refers to the number of face and non-face samples, respectively. In each iteration, find a classifier with the lowest error e_t that can be defined by Eq. 3, provided that $h_t(x_i)$ is the unit function measuring how well the sample x_i fits with the feature of classifier t.

$$w_{it} = \begin{cases} \frac{1}{2m} & if\ d_i = 0 \\ \frac{1}{2l} & if\ d_i = 1 \end{cases} \quad (2)$$

$$e_t = \sum_i w_{it} |h_t(x_i) \cdot d_i| \quad (3)$$

Those classifiers with $e_t = 0$ are desired as the input to the following step. However, if this is not the case, the weights will be adjusted by the end of each iteration such that these error values decrease and reach its possible minimum towards the last round. Fig. 3 shows examples of feature-specific weak classifiers that are obtained from this step. Note that the first makes use of the two-area Haar-like template, which is capable of the unique difference between the horizontal stripe of both eyes and that just below. The second captures the fact that the area between both eyes is brighter than those of the eyes themselves.

Fig 3 Examples of feature-specific weak classifiers [1].

- **Step3:** Having obtained the set of weak classifiers, they are used to form the cascaded classifier that is presented in Fig. 4. Each of the classifiers in the sequence pays attention to a specific feature, and categorizes an input as a face or non-face case. For the former, the input is given to the next classifier down the line. On the other hand, the input is reported as a non-face instance and the detection process is terminated. Only the input that has passed through all classifiers will be identified as a face instance. Find more details of this development in [13, 14].

3. Proposed Bi-Level Classification Model

As mentioned by the end of Section 2 that the generation of cascaded classifier requires a set of training samples. This is composed of both face and non-face instances, which will be employed to validate the relevance of any particular feature to the detection capability. According to [7], the accuracy of this detector depends pretty much on both the magnitude and quality of those training samples. Moreover, it is sensitive to changes in the unknown context that might appear in new images. To response to the challenge and raise the precision, a new bi-level classification model is proposed. Fig. 5 gives the overview of the proposed framework.

With respect to Fig. 5, the proposed method exploits the conventional cascaded classifier of Viola-Jones approach as the first classification level. The sub-windows that are thought to be face images at the end of this primary process will be given to the second classification level. Note that these inputs may contain a number of false positives, even greater when the size and quality of training set is not sufficient. The problem may be resolved by using a highly accurate classifier as the final decision node. To achieve this, the method used in [4] is adopted for the task. The process can be summarized as follows.

- This can be regarded as a binary-class problem, face and non-face. Each sample x_i in the training set of size n can be described with T attributes ($f_1 \ldots f_T$), which correspond to the set of Haar-like features used in the previous stage. The value of attribute f_t for the sample x_i can be estimated with the unit function $h_t(x_i)$. The resulting $n \times T$ matrix can be regarded as the original data for the next phase.
- A cluster ensemble is created from the original matrix by applying k-means for M times. The resulting M data partitions are obtained in each trial using different number of clusters (k). Following [15], this is a random value between 2 to 20.
- The M clustering results are then combined based on the representation of BA (Binary Association) matrix [16]. This is an $n \times P$ matrix, where P denotes the total number of clusters reported in the set of M clusterings. It can be regarded as a transformed data matrix of more informative content than the original, as one may obtain using the attribute reduction techniques such as PCA and KPCA.

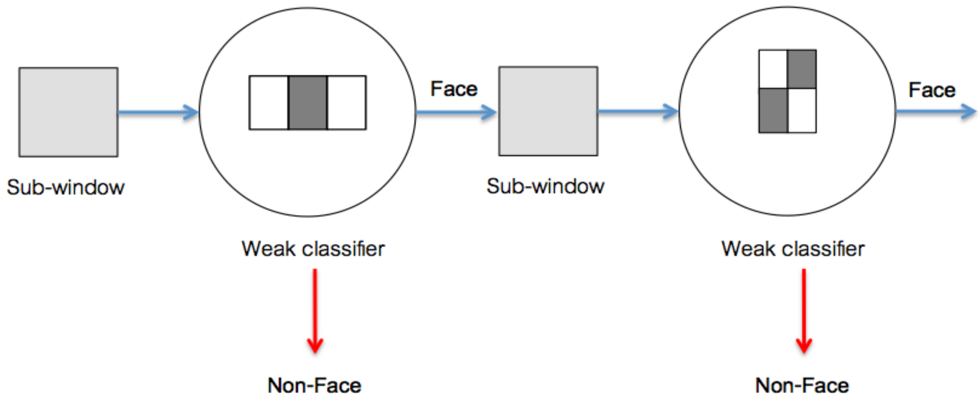

Fig 4 Overview of a cascaded classifier system [1]

- The new matrix is used to develop a classifier using the preferred algorithm. For the present research, C4.5 is employed to deliver the target classification model.

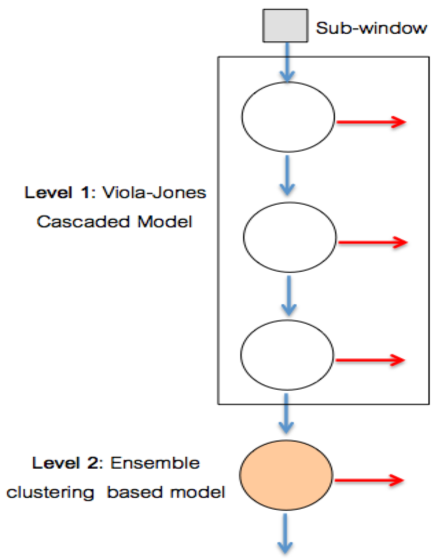

Fig 5 Overview of bi-level classification model.

- For the evaluation with a test or new sample x_j, its T attributes will be first created to form a vector of size T, using the unit functions $h_t(x_j)$, t = 1 … T. Having done this, the sample under question is of the same format as those encoded within the classifier. Then, it is fetched as an input to the classification process, from which either face or non-face outcome is returned.

4. Performance Evaluation

Prior deploying the proposed model in action, it has to be rigorously assessed, such that risk and other useful guidelines may be drawn. This section includes the experimental design and the corresponding results.

4.1 Experimental Design

The experiment is conducted using the 400 samples from the AT&T face repository (or ORL database, www.cl.cam.ac.uk/research/dtg/attarchive/facedatabase.html) as positive or face instances (see Fig. 6 for examples), while the set of non-face samples are randomly extracted from a number of images available in the Internet (see Fig. 7 for an example). A number of test cases specifically designed for this research are elaborated below.

- **Case1:** The training set is composed of 350 positive samples from AT&T and 2,000 non-face instances, while the test set is made up of the other 50 AT&T faces and 50 non-face images that are not included in the training collection.
- **Case2:** Both training and test sets are identical to the first case, but the training samples are modified with 5% of added noise (salt-and-pepper). This is to lower the quality of training data, and observe how robust the classification models become.
- **Case3:** The training set is composed of 50 positive samples from AT&T and 500 non-face instances, while the test set remains the same as the first test case.

Note that the size of ensemble is set to M = 10 for all the test cases. The collection of Haar-like features shown in Fig. 8 are used to generate candidates of weak classifiers. Different sizes of each feature are included in the process. Following several works related to Viola-Jones method, the number of stages or weak classifiers is set to 20. In addition, the classification results are assessed in accordance with the accuracy metric. According to Table 1, the accuracy can be calculated as the ratio between (a + c) and (a + b + c + d).

Fig 6 Examples from AT&T face repository.

4.2 Experimental Results

Based on the setting shown in the previous section, Tables 2 and 3 illustrate the results obtained for the first test case, with the original Viola-Jones cascaded classifier (VJ) and the bi-level model (BI). According to these, the new approach is able to improve the accuracy level from 70% to 76%. In fact, the rate of false positive has declined from 16% to 13%, while the false negative rate also slightly decreases.

Fig 7 An example of image (available in the Internet) from which non-face instances are generated.

Table 1 Confusion matrix, where true positive (TP) is a, true negative (TN) is d, false positive (FP) is b, and false negative (FN) is c, respectively.

		Predicted Class	
		Face	Non-Face
Actual Class	Face	**a**	**c**
	Non-Face	**b**	**d**

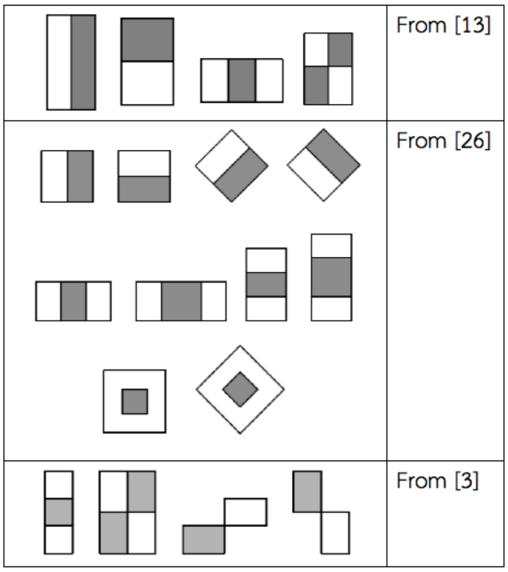

Fig 8 Haar-like features used in this research, which are obtained from the works of [13], [26] and [3].

Table 2 The result of test case 1, with VJ method.

		Predicted Class by VJ	
		Face	Non-Face
Actual Class	Face	36	14
	Non-Face	16	34

Table 3 The result of test case 1, with BI model.

		Predicted Class by BI	
		Face	Non-Face
Actual Class	Face	39	11
	Non-Face	13	37

With respect to test case 2 where noise is added to distort the quality of training data, the results in Tables 4 and 5 lead to a finding that BI is more robust to low-quality samples than the original VJ counterpart. As such, the accuracy obtained by VJ is 55% that is 10% lower than that of BI (at 65%). In addition to these, Tables 6 and 7 show the results with test case 3, where the accuracy of VJ and BI methods are 48% and 56%, respectively. In the event of limited training samples, BI is still able to remain the accuracy above 50%, not the other.

Table 4 The result of test case 2, with VJ method.

		Predicted Class by VJ	
		Face	Non-Face
Actual Class	Face	30	20
	Non-Face	25	25

Table 5 The result of test case 2, with BI model.

		Predicted Class by BI	
		Face	Non-Face
Actual Class	Face	34	16
	Non-Face	19	31

Table 6 The result of test case 3, with VJ method.

		Predicted Class by VJ	
		Face	Non-Face
Actual Class	Face	30	20
	Non-Face	25	25

Table 7 The result of test case 3, with BI model.

		Predicted Class by BI	
		Face	Non-Face
Actual Class	Face	30	20
	Non-Face	24	26

In particular to the bi-level model, Fig. 9 illustrates its accuracy and false positive rates for the three test cases. It is clear from this graph that the performance of BI will decline if the training data is of low quality (case 2) or the number of training samples is small (case 3). Likewise, the similar finding is acquired for the increase of false positives. BI suffers from these conditions just like the VJ method, but appears to be more robust. Hence, the ideal context for BI application is the use a sufficient amount of training data with high quality. As such this model can reach its full potential for the task.

5. Conclusion

This paper has presented a novel bi-level classification method for the problem of face detection. It extends the conventional Viola-Jones model with additional classifier, which proves to be vital for the improved performance. The new classifier is based on the concept of data transformation using ensemble clustering information. This combination is robust to noise and small training data. As shown by the experimental results, the proposed model is consistently more accurate than the existing single-level method. Despite this, there are a few questions left to answer. One regards the effect of the ensemble size to the predictive performance. It is to validate the common intuition that a large ensemble is better than the smaller one. Besides, the other is the use of several other classifiers, instead of C4.5, to generate the classifier in the second level of BI. It would also be interesting to see the coupling of this detection model with a UAV platform [25], such that ground objects can be detected from aerial images.

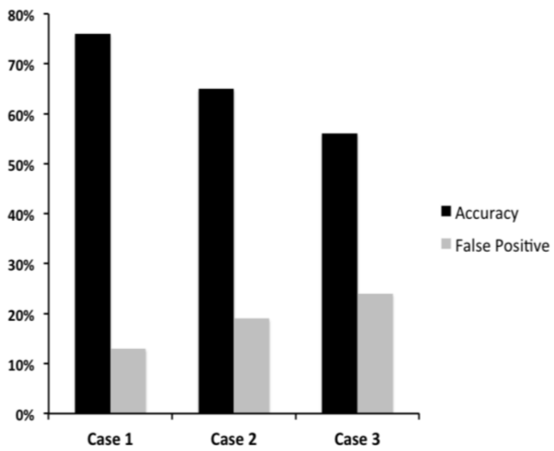

Fig 9 Accuracy and false positive rates of BI.

6. Acknowledgements

The authors would like to thank Department of Science and Technology, Ministry of Defence, as well as Ministry of Science and Technology for the research grants. Also, this project is the great success with the support of Navaminda Kasatriyadhiraj Royal Air Force Academy and that of Mae Fah Luang University.

7. References

[1] T. Boongoen and B. Undara. Image Analysis for Anti-Terrorism and Security Measure. Technical Report, Department of Science and Technology (Ministry of Defence), 2014.

[2] N. Iam-On and T. Boongoen. Improving Face Classification with Multiple--Clustering Induced Feature Reduction. In Proceedings of 49th IEEE International Conference on Security Technology: 241-246, 2015.

[3] T. Mita, T. Kaneko and O. Hori. Joint Haar-like features for face detection. In Proceedings of International Conference on Computer Vision: 1619-1626, 2005.

[4] N. Iam-On and T. Boongoen. Diversity-Driven Generation of Link-Based Cluster Ensemble and Application to Data Classification. Expert Systems with Applications, 42(21): 8259-8273, 2015.

[5] M.H. Yang, D.J. Kriegman and N. Ahuja. Detecting Faces in Images: A survey. IEEE Transactions on Pattern Analysis and Machine Intelligence, 24(1): 34-58, 2002.

[6] Y.Q. Wang. An Analysis of the Viola-Jones Face Detection Algorithm. Image Processing On Line, 4: 128-148, 2014.

[7] S.K. Pavani, D. Delgado and A.F. Fragni. Haar-like features with optimally weighted rectangles for rapid object detection. Pattern Recognition, 43: 160-172, 2010.

[8] G. Yang and T.S. Huang. Human Face Detection in A Complex Back Ground. Pattern Recognition, 27(1): 53-63, 1994.

[9] C. Kotropoulos and I. Pitas. Rule-Based Face Detection in Frontal Views. In Proceedings of International Conference on Acoustics, Speech and Signal Processing, 4: 2537-2540, 1997.

[10] K.C. Yow and R. Cipolla. A Probabilistic Framework for Perceptual Grouping of Features for Human Face Detection. In Proceedings of International Conference on Automatic Face and Gesture Recognition: 16-21, 1998.

[11] B. Scassellati. Eye Finding via Face Detection for a Foevated Active Vision System. In Proceedings of National Conference on Artificial Intelligence, 1998.

[12] K.K. Sung and T. Poggio. Example-Based Learning for View-Based Human Face Detection. IEEE Transactions on Pattern Analysis and Machine Intelligence, 20(1): 39-51, 1998.

[13] P. Viola and M. Jones. Rapid object detection using a Boosted cascade of simple features. In Proceedings of International Conference on Computer Vision and Pattern Recognition: 1-9, 2001.

[14] P. Viola and M. Jones. Robust real-time face detection. International Journal of Computer Vision, 57(2): 137-154, 2004.

[15] N. Iam-on, T. Boongoen and S. Garrett. LCE: A link-based cluster ensemble method for improved gene expression data analysis. Bioinformatics, 26(12): 1513-1519, 2010.

[16] N. Iam-On and T. Boongoen. Comparative Study of Matrix Refinement Approaches for Ensemble Clustering. Machine Learning, 98(1-2): 269-300, 2015.

[17] J.P. Maurya and S. Sharma. A survey on face recognition techniques. Computer Engineering and Intelligent Systems, 4(6): 11–17, 2013.

[18] Y.J. Zhang. Advances in Face Image Analysis: Techniques and Technologies. IGI Global, 2010.

[19] M. Turk and A. Pentland. Eigen faces for recognition. Journal of Cognitive Neuroscience, 3(1): 71–86, 1991.

[20] S.H. Yoo, S.K. Oh and K. Seo. Design of face recognition algorithm using hybrid data preprocessing and polynomial-based RBF neural networks. In Proceedings of International Conference on Advances in Neural Networks: 213–220, 2012.

[21] A. Ng, M. Jordan, and Y. Weiss. On spectral clustering: Analysis and an algorithm. Advances in Neural Information Processing Systems, 14: 849–856, 2001.

[22] S.T.Roweis and L.K.Saul. Nonlinear dimensionality reduction by locally linear embedding. Science, 290(5500): 2323–2326, 2000.

[23] O. Samko abd A.D. Marshall and P.L. Rosin. Selection of the optimal parameter value for the Isomap algorithm. Pattern Recognition Letters, 27(9): 968–979, 2006.

[24] P. Dadech and S. Boonlue. An Applied of Image Processing for Plant Classification on Smartphone. NKRAFA Journal of Science and Technology, 11: 102-110, 2015.

[25] V. Vaswani, S. Loh, B. Tan, K. Singh and S. Srigrarom. Modular Quadrotor MAVs. NKRAFA Journal of Science and Technology, 10: 43-55, 2014.

[26] R. Lienhart and J. Maydt. An extended set of Haar-like features for rapid object detection. In Proceedings of International Conference on Image Processing: 900-903, 2002.