

# Enhanced AdaBoost Haar Cascade Classifier Model to Detect Partially Occluded Faces in Digital Images

Abdullah Bade\* and Tulasii Sivaraja

*Mathematics, Graphics and Visualization Research Group (MGRAVS), Faculty of Science and Natural Resources, Universiti Malaysia Sabah, 88400 Kota Kinabalu, Sabah, Malaysia*

Face detection is a popular aspect of object detection in the field of image processing. One of the most challenging issue to resolve is the detection of partially occluded faces in digital images. The aim of this research is to improve the detection of partially occluded faces in images. This aim is achieved by modifying boosting parameters heuristically and finding the optimum value of the tree depth during the training process with images found on Google Image database. The database has thousands of images that have various occlusions, lighting, and pose to detect partially occluded faces. Experiments were carried out to find the optimum values of the decision tree as well as the boosting parameters. Once these values were identified, a custom cascade was trained and used to detect partially occluded faces in images. The algorithm was then tested on the Face Detection Database (FDDB) and the performance was evaluated. From the conducted experiment, the accuracy of the enhanced Haar cascade classifier was 29.74% better as compared to the original Haar cascade classifier.

**Keywords:** face detection, cascade training, Haar cascade

## I. INTRODUCTION

The technology of detecting faces has evolved proportional to its usage in various applications over the past few years. In 2001, Viola and Jones introduced a well-known algorithm for face detection. The framework they proposed uses a novel image representation known as integral image and creating a boosted cascade of weak Haar-like feature classifiers to produce real time face detection (Viola & Jones, 2001).

Various researchers have adapted the Viola-Jones framework and made improvements to the original framework in line with the advancements of technology. In 2002, Lienhart and Maydt improved the detection rates when they extended the set of Haar-like features used in the framework. They contributed a new set of Haar-like features by rotating the original features by 45° and

included a centre surround feature (Lienhart & Maydt, 2002).

Face detection has gained popularity in recent years and the applications of it in the industry has grown tremendously. It has an important role in the task of face recognition and biometric authentication now. Furthermore, the usage of face detection is also observed in the software embedded in digital cameras and mobile phone cameras. These gadgets detect faces to ensure that the image captured focuses on people's faces to produce good quality images. Face detection is also used in social networks and social media applications such as Facebook and Snapchat to improve and enhance user experience (Rajawat, Pandey & Rajaput, 2017). Researchers are motivated to optimize and solve the issues surrounding face detection because of the growing usage of face detection in the industry.

---

\*Corresponding author's e-mail: [tulasii.sivaraja@gmail.com](mailto:tulasii.sivaraja@gmail.com)

There are four major challenges that exist in the field of face detection and recognition which are aging, poses, illumination, and expression (Mahalingam, Ricanek, & Albert, 2014). Partial occlusion is also a problem encountered in face detection as well. Researchers began to realize that partial face occlusions affected the overall detection rate of the algorithm. Most face detection algorithms work by finding facial features such as mouth, nose and a set of eyes in images in a joint search. Therefore, partial occlusions from sunglasses and scarves complicates the task of detecting faces which in turn affects the performance of the detection algorithms.

Faces detection in images are made difficult with the limited face visibility due to various factors such as variation in pose and lighting, as well as occlusions caused by hats, scarves, hair and foreign objects. To detect the partially occluded faces in images accurately without using artificial intelligence requires an approach that does not have an expensive trade-off between computation time and false positive rates. This paper addresses the issue of detecting partially occluded faces in images. To achieve good detection rates, we propose training a Haar cascade classifier model using heuristic boosting parameter values and an optimum CART tree depth on a large dataset that contains a variation of occlusions.

The remaining parts of this paper is organised as follows: Section 2 summarizes the existing face detection methods, Section 3 outlines the proposed implementation, Section 4 and Section 5 describe the experiments and results of CART tree depth and boosting parameters of cascade training, Section 6 discusses the Enhanced Face detection Technique and in Section 7, the paper is concluded along with future work recommendations.

## II. LITERATURE REVIEW

Face detection approaches can be generalised as feature-based approaches and image-based approaches. Feature-based approaches in face detection requires the extraction of facial features from an image and comparing it with a knowledge base of face features, whereas image-based approaches attempts to get the best match while comparing training images against testing images (Modi,

Macwan, 2014).

Feature-based methods consist of Active Shape Models (ASM), Low-Level Analysis and Feature Analysis. In 1995, Tim Cootes and Chris Taylor developed Active Shape Model where an object is found in an image based on the statistical model of the shape of the object by iteratively deforming it to fit the new object (Cootes, Baldock, & Graham, 2000). Meanwhile, low-level analysis uses visual features that can be extracted from images such as edges, grey scale levels, colour information, motion, and general measures to differentiate facial structures of a person from its surroundings (Singh, Daniel, & Chaturvedi, 2017). Feature analysis methods such as feature searching and constellation uses the geometrical facial feature information to locate faces without any regards to illumination, pose and angles of the face.

Image based approaches aim to improve the unpredictability of the various feature-based approaches. The strategy this approach utilizes is treating face detection as a pattern recognition problem, whereby the specific application of face knowledge is steered clear from and the issue is tackled as learning to recognize face patterns from examples (Hjelmås & Low, 2001). With this approach, the potential modelling error caused by incomplete face knowledge is eliminated. The fundamental approach of image-based methods is via training procedures that classifies the examples as face and non-face classes. The methods implemented under this approach are linear subspace methods, neural networks, and statistical approaches.

The proposed method aims to create a cascade classifier trained with heuristic boosting parameter values that improves the detection of partially occluded faces in images. Additionally, the implementation of a couple of pre-processing procedures, it is hoped that the computation time is improved.

## III. PROPOSED METHOD

The implementation of the proposed method is shown in Figure 1. There are two distinct phases, the cascade training scheme, and the face detection technique. From

the training phase, a cascade model is created with the heuristic boosting parameters and the optimal tree depth value. This cascade model is then used in the architecture of the enhanced face detection technique to detect partially occluded faces in images. The experiments are tested on a subset of the Wider Face dataset (Yang, Luo, Loy, & Tang, 2016).

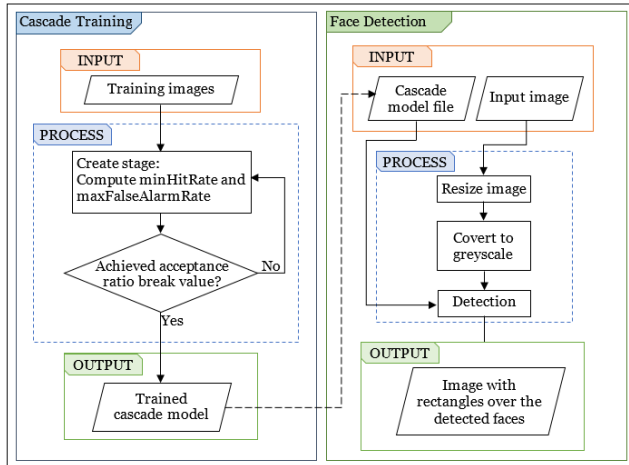


Figure 1. Proposed Method

#### A. Heuristic Cascade Training Scheme

The weights of boosted weak classifiers are computed and the CART of each stage is formed in the process stage of the cascade training. Since the number of features selected to be used for evaluating each stage is influenced by the depth of the decision tree. Therefore, to reduce the number of features evaluated without under fitting the cascade, the optimum value of the tree is determined.

Additionally, two other boosting parameters, minHitRate and max FalseAlarmRate also affects the number of weak classifiers selected in a stage. These parameters also influence the performance of the cascade in terms of detection rates. The acceptance ratio break value is also affected by these two boosting parameters. A cascade is created once the acceptance ratio break value is achieved, otherwise a new stage of cascade is added when the maxFalseAlarmRate is reached until the acceptance ratio break value is reached. The detection rates of partially occluded faces will depend on finding the optimum value of the boosting parameters for training the cascade.

#### B. Enhanced Face Detection Technique

Once the input image and the trained cascade model are provided, the input image undergoes two pre-processing steps, which are resizing and converting the colour space from RGB to grayscale. This will ensure that the best results are obtained for the detection of partially occluded faces in images. During cascade training, grayscale images are used, therefore, when executing the face detection, it is best to use grayscale images in order to reduce the computation time of the detection as the task of pixel comparison requires extensive computation. To accommodate to the very small width and height dimensions of the training images, the size of the input image is resized to be twice as small as the original image. These enhancement to the face detection technique produces faster results as it does not require various resizing when it is compared with the cascade classifier. The trained cascade model is then used to detect faces in the image. An image with rectangles over the detected faces or objects is displayed as the output image in the output phase.

### IV. OPTIMUM CART TREE DEPTH

Six experiments are conducted with different tree depth ranging from [1, 6] to select the optimum value to use for the tree depth. It is important to select the most favourable value of tree depth to reduce the number of features evaluated at each stage since the number of features increases as the tree depth value increases. An optimum value needs to be selected to avoid two cases: over fitting the decision tree if the tree is too deep or under fitting the tree with a shallow tree depth.

The cascade models trained with different tree depth have the same values for the other training parameters. The number of stages is set at 50 and maxWeakCount is 300. Large values are used for these parameters to ensure that the training does not terminate before the acceptance ratio break value of 0.0001 is achieved. Other values of the boosting parameters are minHitRate= 0.999, maxFalseAlarmRate= 0.5 and weightTrimRate= 0.95. The results from the experiments show that the CART with

depth  $n=3$  used the least number of features followed by the depth  $n=2$  and  $n=4$ . Table 1 shows the results of the confusion matrix analysis while Figure 2 is the ROC curves of the cascade models with different CART tree depths.

Table 1. Confusion Matrix analysis of cascade models with different CART tree depth

	CART1	CART2	CART3	CART4	CART5	CART6
Depth	1	2	3	4	5	6
Accuracy	0.60	0.61	0.63	0.61	0.58	0.55
Recall	0.80	0.78	0.76	0.72	0.68	0.60
Precision	0.71	0.74	0.79	0.81	0.80	0.85
F1 Score	0.75	0.76	0.77	0.76	0.73	0.71

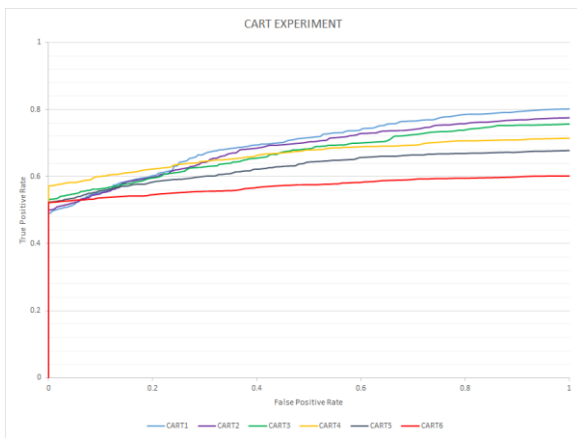


Figure2. ROC curve of cascade models with different tree depths

The output of the experiments conducted show that the tree depth of  $n=3$  and  $n=2$  can be used as the optimum tree depth value. However, the CART decision tree with  $n=3$  was chosen because it used the least amount features in each stage most of the time compared to all the other

cascade models tested. Additionally, at this tree depth, the detection rate of partially occluded faces in images was better compared to a CART with tree depth,  $n=2$ . This is because a deeper tree can preserve feature information better. The F1 score of CART3 was also the highest among all the other cascade models. From the ROC curve, CART3 is observed to be closer to the top left corner of the graph compared to the ROC curve generated by CART2, which makes it the more accurate cascade model. The visual comparison test shows that CART3 performed better with higher true positives confidence measure and the ability to detect partially occluded faces better than CART2. Therefore, the tree depth  $n=3$  is used as a constant value for the heuristic boosting parameter experiments since it produced the best results.

### V. HEURISTIC BOOSTING PARAMETER VALUES

The boosting parameter that are heuristically determined are the AcceptanceRatioBreakValue, MinHitRate, MaxFalseAlarmRate, and weightTrimRate. Setting conservative threshold values may accept lots of negative windows in the early stages of the cascade, however, if the threshold values are set aggressively, the cascade model might be over fitted, reducing the detecting rates of the cascade. Table 2 shows the cascade models trained with different boosting parameter values.

Table 2. Boosting parameter values for cascade models

	EXP1	EXP2	EXP3	EXP4	EXP5	EXP6	EXP7	EXP8	EXP9
Acceptance Ratio	50e-6	75e-7	25e-5	50e-6	50e-6	50e-6	50e-6	50e-6	50e-6
MinHitRate	0.999	0.999	0.999	0.990	0.995	0.999	0.999	0.999	0.999
numStages	50	50	50	50	50	50	50	50	50
maxFalseAlarmRate	0.5	0.5	0.5	0.5	0.5	0.1	0.3	0.5	0.5
weightTrimRate	0.95	0.95	0.95	0.95	0.95	0.95	0.95	0.99	0.90

The cascade models are evaluated in the colour coded groups to ensure that only a certain boosting parameter is compared and not all the trained cascade models

simultaneously. Table 3 is the results of the confusion matrix of the cascade models.

Table 3. Confusion matrix analysis of cascade models trained with different boosting parameter values

	EXP1	EXP2	EXP3	EXP4	EXP5	EXP6	EXP7	EXP8	EXP9
Accuracy	0.6183	0.6223	0.5245	0.5597	0.5901	0.6055	0.6131	0.6150	0.6098
Recall	0.6863	0.7142	0.8066	0.6135	0.6586	0.6704	0.7023	0.6934	0.6672
Precision	0.8618	0.8286	0.6	0.8645	0.8502	0.8621	0.8285	0.8447	0.8763
F1 Score	0.7641	0.7671	0.6881	0.7177	0.7422	0.7543	0.7602	0.7616	0.7576

Based on Table 3, the highest F1 score among the colour coded groups is used to determine the ranges of the boosting parameter values. Additionally, comparing the ROC curves of the models provides further validation of the ranges of values determined from the confusion matrix.

The range of the modified value for the acceptance ratio should fall in [0.00005, 0.0001], minHitRate is narrowed down to [0.995, 0.999], maxFalseAlarmRate in the range of [0.3, 0.5] and the weightTrimRate in [0.95, 0.99]. The minHitRate and weightTrimRate are selected to be the midpoint of range that was suggested. The acceptanceRatioValue and maxFalseAlarmRate are lowered to ensure that sufficient detections of partially occluded faces are used in the training. Based on the results of the experiments, the following values are determined to modify the boosting cascade: Acceptance Ratio BreakValue = 0.000075, MinHitRate = 0.995, MaxFalseAlarmRate = 0.3, and weightTrimRate = 0.95.

## VI. ENHANCED FACE DETECTION TECHNIQUE

Table 4 displays the analysis of the confusion matrix of the enhanced cascade classifier, and two other existing cascade classifiers created by Lienhart.

Table 4. Confusion Matrix Analysis

	Enhanced cascade	Haar-frontalface-default	Haar-frontalface-alt
Accuracy	0.6525	0.5290	0.4843
TPR/ Recall	0.6945	0.5336	0.5199
Precision	0.8683	0.9839	0.8759
F1 Score	0.7717	0.6919	0.6525

The accuracy and true positive rate or recall of the enhanced cascade has the highest value whereas, the precision of Haar-frontalface-default (HF\_def) and Haar-frontalface-alt (HF\_alt) was better than the enhanced cascade. Regardless, the ratio of sensitivity and specificity of the enhanced cascade model was 10.9% better than HF\_def and 16.7% better than HF\_alt, which means that the enhanced cascade surpassed the other two cascades in detecting partially occluded faces. Figure 3 is an example of the detection achieved by the enhanced cascade classifier.



Figure 3. Partially occluded faces detected by heuristic cascade classifier model

Visually comparing the detection of each cascade model, the enhanced cascade model performs better than the other two models in detecting partially occluded faces in digital images. The custom cascade outperforms the Haar\_def and Haar\_alt by 23.66% and 21.7% respectively.

## VII. CONCLUSION

We proposed a heuristically trained Haar cascade classifier model to detect partially occluded faces in digital images. The major challenge we encountered was how other external factors such as facial expressions, poses and poor illumination interfered with the overall performance of the heuristic cascade model. Profile faces were especially unstable to detect. However, the problems faced can be improved upon by increasing the number of images with

more profile faces and occlusions used in training. We successfully determined the optimum value of the tree depth while ensuring the number of features evaluated were at an acceptable range and was still able to perform well in detecting partially occluded faces. The heuristically determined boosting parameter values proved to be optimal as it further enhanced the performance of the trained cascade classifier. The pre-processing steps taken to enhance the face detection further improved the detection time of the heuristic cascade classifier. Future work to improve this method would be to implement several occlusion-specific trained cascades classifiers together in the face detection algorithm which should improve the detection rate of partially occluded faces.

## VIII. REFERENCES

- Cootes, T., Baldock, E. R., & Graham, J. (2000). An introduction to active shape models. *Image processing and analysis*, 223-248.
- Hjelmås, E., & Low, B. K. (2001). Face detection: A survey. *Computer vision and image understanding*, 83(3), 236-274.
- Lienhart, R., & Maydt, J. (2002). An extended set of haar-like features for rapid object detection. In *Proceedings. International Conference on Image Processing* (Vol. 1, pp. I-I). IEEE.
- Modi, M., & Macwan, F. (2014). Face detection approaches: a survey. *Int. J. Innov. Res. Sci. Eng. Technol*, 3.
- Mahalingam, G., Ricanek, K., & Albert, A. M. (2014). Investigating the periocular-based face recognition across gender transformation. *IEEE Transactions on Information Forensics and Security*, 9(12), 2180-2192.
- Rajawat, A., Pandey, M. K., & Rajput, S. S. (2017, February). Low resolution face recognition techniques: A survey. In *2017 3rd International Conference on Computational Intelligence & Communication Technology (CICT)* (pp. 1-4). IEEE.
- Singh, N., Daniel, A. K., & Chaturvedi, P. (2017, February). Template matching for detection & recognition of frontal view of human face through Matlab. In *2017 International Conference on Information Communication and Embedded Systems (ICICES)* (pp. 1-7). IEEE.
- Viola, P., & Jones, M. (2001). Rapid object detection using a boosted cascade of simple features. *CVPR (1)*, 1, 511-518.
- Yang, S., Luo, P., Loy, C. C., & Tang, X. (2016). Wider face: A face detection benchmark. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 5525-5533).