Boosting the Accuracy of AdaBoost for Object Detection and Recognition

Zahid Mahmood[1], Muhammad Nazeer[2], Muhammad Arif[1], Imran Shahzad[1], Fahad Khan[1], Mazhar Ali [3], Uzair Khan[1], and Samee U. Khan[4]

[1] Department of Electrical Engineering, COMSATS Institute of Information Technology, Abbottabad, Pakistan.
[1] {zahid0987, marif, imranshahzad, fahadkhan, uzairkhan}@ciit.net.pk
[2] Department of Mathematics, COMSATS Institute of Information Technology, Wah Cantt, Pakistan.
[2] nazeermuhammad@ciitwah.edu.pk
[3] Department Computer Science, COMSATS Institute of Information Technology, Abbottabad, Pakistan.
[4] mazhar@ciit.net.pk
[4] Department of Electrical and Computer Engineering, North Dakota State University, USA.
[5] same.khan@ndsu.edu

***Abstract*—Recently, large number of object detection and recognition algorithms are playing a key role in several applications, such as security and surveillance. Although, these algorithms perform exceptionally well under normal lighting conditions, however their detection and recognition accuracy abruptly degrades under non-uniform illuminations, such as strong sunlight and bad lighting conditions. In this paper, we apply our own developed Multi-Scale Retinex (MSR) algorithm as a pre-processing module to boost the accuracy of the AdaBoost algorithm, which is considered to be state-of-the-art algorithm for robust object detection and recognition. Simulation results show that the MSR can be reliably and effectively used under non-uniform illuminations to boost the accuracy of the AdaBoost.**

***Index Terms*—**AdaBoost, Multi-Scale Retinex, Object Detection and Recognition

I. Introduction

Detecting and identifying objects is a crucial methodology for analyzing objects under various situations. Automatic object detection and recognition systems are extensively used in large number of real life applications, such as displaying players' statistics in live sports [1]. Real-time objects, such as vehicle detection continuously notifies drivers with crash warning for safe driving. During past two decades, several algorithms, for example, [2], [3], [4], and [5] have been proposed to detect and recognize various objects. Among them is the AdaBoost algorithm that has been successfully applied in various sports entertainment and security applications [4]. Aforementioned papers report the accuracy of AdaBoost under normal lighting conditions only. Our main contributions in this paper are highlighted below:

- We apply our own developed MSR based image enhancement algorithm as a pre-processing module to boost the accuracy of AdaBoost algorithm under non-uniform lighting conditions. The earlier works [1], [2], [4], [7], [8], [9], and [10] evaluate the AdaBoost algorithm under normal lighting conditions only.
- We perform experiments on real life challenging images that contain severe dark contrast when captured under strong sunlight.
- We show experiments on 5312 images to demonstrate the usefulness of the application of the MSR as pre-processing module to significantly boost the detection and recognition accuracy.

This paper is organized as follows. Section II briefly discusses the MSR based image enhancement algorithm. Sections III and IV briefly describe the AdaBoost algorithm to detect and recognize objects, respectively. Simulations are presented in Section V. Finally, Section VI concludes the discussion and hints possible future research. For each of the Sections, Table 1 shows the common terms and their meanings used in the paper.

II. The MSR Image Enhancement Algorithm

Recently large numbers of image enhancement algorithms have been published in the literature. Among them is the MSR based algorithms that proved to be very effective in enhancing images degraded due to bad lighting conditions. Our developed image enhancement algorithm has several separate steps [6]. **First** step utilizes the PCA to process the input image to extract the luminance channel. **Second** step enhances luminance channel using three scales of the MSR. This step also utilizes the DFT to speed up the enhancement procedure. In the **third** step, the luminance ratio is computed from the original luminance. **Fourth** step calculates new Red Green Blue pixel values by multiplying the luminance ratio with input RGB values. Finally, Contrast Stretching (CS) is applied on each RGB channel separately to obtain the enhanced output image. The CS is applied according to the following expression:

$$\begin{cases} new\ pixel = \frac{old\ pixel - low}{high - low} \times 255 \\ Low = 5 \le low \le 26 \\ High = 255 - low \end{cases} \tag{1}$$

TABLE 1: NOTATIONS WITH THEIR MEANINGS

Notations	Meanings
AdaBoost	Adaptive Boosting Algorithm
CS	Contrast Stretching
DFT	Discrete Fourier Transform
MSR	Multi-Scale Retinex
PCA	Principal Component Analysis
RGB	Red Green Blue Channels in colored image

For 3×3 sliding window in the input image, we set coordinates $(1,2)$ as high pixel, $(2,1)$ as old pixel, $(3,2)$ as low pixel, and $(2,2)$ as new pixel. The MSR based player image enhancement algorithm is extensively tested on images captured under non-uniform lighting conditions. Further details of the MSR can be seen in [6].

III. OBJECT DETECTION USING ADABOOST

The AdaBoost algorithm as shown in Algorithm 1 was originally introduced by Viola and Jones [7]. The AdaBoost generates a robust final classifier by linearly linking large number of weak classifiers. For a set of training data, Algorithm 1 picks numerous weak classifiers from many Haar features and pools them into a resilient classifier as described by Eq. (2).

$$h(x) = \begin{cases} 1 & \sum_{k=1}^{k_{max}} \alpha_k h_k(x) \geq \theta \\ 0 & otherwise \end{cases} \qquad (2)$$

where x is the input image, $h_k(x)$ are weak classifiers, θ is the threshold, and $h(x)$ is the final strong classifier. Algorithm 1 explains the pseudo code of the AdaBoost algorithm. The AdaBoost has been successfully used to detect various objects, such as, face, players in a stadium, and vehicles on highways and parking-lots [1], [7], [8], and [9].

IV. OBJECT RECOGNITION USING ADABOOST

The AdaBoost has been successfully used to recognize objects, such as face in diverse situations [10], [11]. To recognize a face using the AdaBoost, the performance of LDA based methodology is enhanced by including it in the boosting structure. Each step in boosting simplifies a new LDA subspace that concentrates on face samples that are incorrectly classified in the preceding LDA subspace. Final classifier as shown in Eq (3) is an ensemble of different specific LDA solutions. Such type of collective approach takes advantage of both Boosting and LDA (BLDA) and outperforms the traditional LDA based schemes in face recognition tasks.

$$h_f(\mathbf{z}) = \underset{y \epsilon \mathbb{Y}}{\arg max} \sum_{t=1}^{T} \left(\log \frac{1}{\beta_t}\right) h_t(\mathbf{z}, y) \qquad (3)$$

Where $h_t(z, y)$ are weak classifiers generated during boosting procedure and $h_f(z)$ is a final strong classifier. Further details of this algorithm can be seen in [10].

ALGORITHM 1: THE ADABOOST ALGORITHM

Given example images $(x_1, y_1), ..., (x_n, y_n)$ where $y_i = 0, 1$ for negative & positive examples respectively.
Initialize weights $w_{1,i} = \frac{1}{2m}, \frac{1}{2l}$ for $y_i = 0, 1$, where m and l are the number of negatives & positives respectively.
for $t = 1;\ t < T;\ t++$ **do**
    Normalize the weights,

$$w_{t,i} \leftarrow \frac{w_{t,i}}{\sum_{j=1}^{n} w_{t,j}}$$

    to make w_t as a probability distribution.
    For each feature j , train a classifier h_j which is restricted to using a single feature. The error is evaluated with respect to

$$w_t, \epsilon_j = \sum_i w_i |h_i(x_i) - y_i|.$$

    Choose the classifier, h_t, with the lowest error ϵ_t.
    Update the weights:

$$w_{t+1,i} = w_{t,i}\beta_t^{1-e_i}$$

    where $e_i = 0$ if example x_i is classified correctly, $e_i = 1$ otherwise, and $\beta_t = \frac{\epsilon_t}{1-\epsilon_t}$.
end
The final strong classifier is:

$$h(x) = \begin{cases} 1 & \sum_{t=1}^{T} \alpha_t h_t(x) \geq \frac{1}{2} \sum_{t=1}^{T} \alpha_t \\ 0 & \text{otherwise} \end{cases}$$

where $\alpha_t = \log \frac{1}{\beta_t}$

V. SIMULATION RESULTS

We perform comprehensive experiments on Dell Precision Tower 7810, Dual Intel Xeon Processor E5-2699 V3machine with on board 192 GB of RAM. Aforementioned is a high performance machine that can handle the real-time execution of complex object detection and recognition algorithms. All the simulations have been executed in OpenCV.

It is important to state here again that the accuracy of the AdaBoost algorithm has been evaluated under bad lighting conditions.

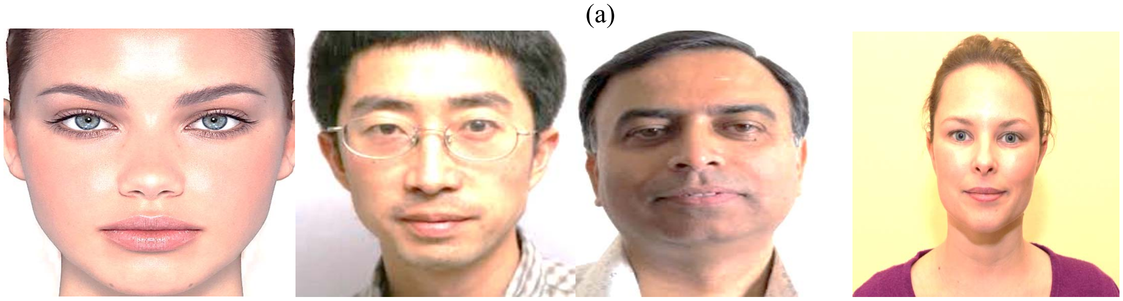
Definition: Bad lighting condition: A natural lighting condition that results in an image that is not easy for most of the algorithms to detect or recognize objects. For example, a very bright image that is obtained due to strong sunlight, a very dark image because of low light, or a blurred image due to object/camera motion. This paper interchangeably uses non-uniform illuminations, strong sun light, or dark image to mean the bad lighting conditions. Moreover, the original images shown in this paper are naturally degraded without manual noise addition. Fig. 1(a), 2(a), and 3(a) are effected by severe dark contrast when captured in strong sunlight, which we refer bad lighting conditions.

Before discussing the accuracy boosting comparison, we briefly state the object classifiers details below.

Face Classifier: We use the face classifier[1] as provided by [7]. The face detector was obtained by manually labeling and

[1] www.opencv.com

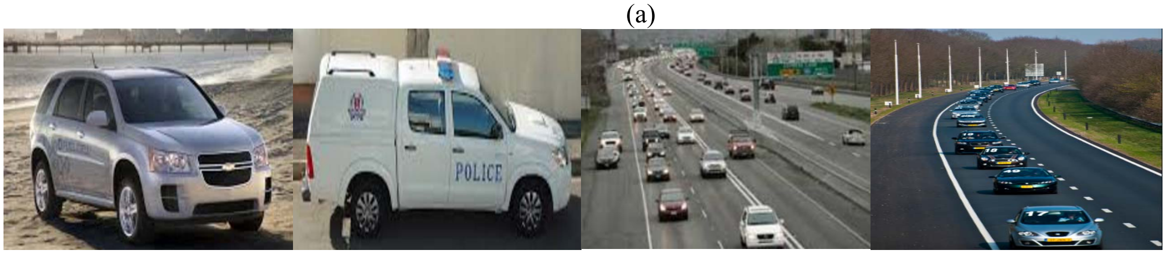
(a)



(b)

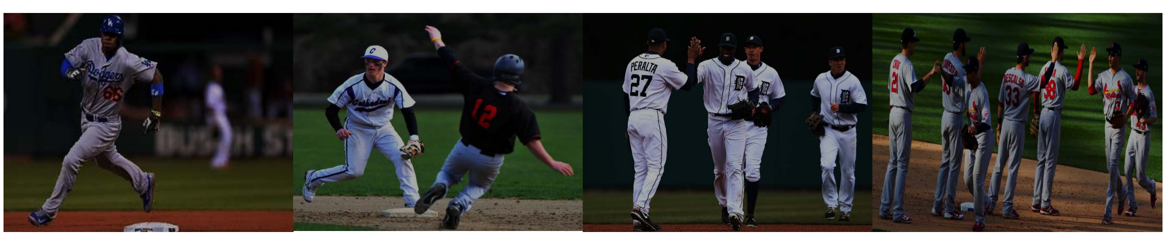
Fig. 1: (a) Original degraded facial images (b) Enhanced facial images using the MSR

(a)



(b)

Fig. 2: (a) Original degraded vehicle images (b) Enhanced vehicle images using the MSR



(a)

(b)

Fig. 3: (a) Original degraded player images (b) Enhanced player images using the MSR

training 5000 significantly cropped facial images and 10,000 negative or non-facial images of size 24×24 pixels. While face recognizer was obtained from [10]. Since face detection precedes the face recognition. Therefore in our experiments we only evaluate face recognition accuracy before and after facial image enhancement. Moreover, it should be noted that there can be variations in human faces, such as change in expression (smile, laugh, yawn, angry, sad, occlusion, or neutral) and pose (from frontal to profile view). This paper only evaluates the frontal facial images of human in a neutral position.

Vehicle Classifier: We use the vehicle classifier developed in [9] that contains 17,200 images out of which 6200 images are vehicle images, while 11,000 are non-vehicle images. Base resolution of the detector is 20×20 pixels.

Player Classifier: We use the player classifier developed in [1] that contains 1730 images, out of which 630 are cropped and labeled player images of resolution 20×40 pixels. While, 1100 are non-player images of resolution 640×480 pixels.

Discussion on Results:

In sections below, we discuss our observations and describe some novel facts for object detection. It should be noted that results presented in Table 2 have been obtained on 3512 different images. Precisely, there were 812 facial images that contained up to 34 frontal facial images, 2500 vehicle images that included up to 45 vehicles in an image, and 2000 player images that contained up to 19 players' in an image.

A. Image Enhancement

The employed image enhancement scheme has significant impact on degraded images. Some of the observations regarding the object image enhancement are highlighted below.

- As can be seen in enhanced image in Fig. 1(b) that background of the face is much better. Similarly, important facial details, such as eye brows, nose, cheeks, lips, and forehead are nicely enhanced that ultimately improves the recognition capability of the face recognizer.
- Vehicle detection has always remained a challenging task. Vehicle detection becomes further a challenging task when there is a bad lighting condition. The highways colors further get dim that result in a blackish appearance of the input images as shown in Fig. 2(a). Due to this fact important details and features of the object change. Therefore, most of the detectors struggle to perform well in such situations. The image enhancement effectively eradicates the blackish regions as shown in Fig. 2(b). Especially, the bottom right image in Fig. 2(b) illustrates the usefulness of the image enhancement. Most of the vehicles travelling on the road are black/silver in color. After the image is enhanced, the vehicle details are much clear than the original image. It will be explained in the next section that enhancing such image helps to achieve the high detection rate.
- The player images shown in Fig. 3(a) have poor aesthetic quality. The enhancement on such images results in a nice quality output image. The players' dress, caps, and facial regions are very clear after enhancement. Moreover, the grass color in stadium is also restored to natural shape as indicated by Fig. 3(b).

B. Accuracy Comparison

This section discusses in detail the accuracy comparison before and after the image enhancement is applied. Table 2 summarizes the comparison results. From Table 2, some of the important observations are highlighted below.

- It can be seen in column 2 and column 3 of Table 2 that the facial image enhancement has a significant effect on the recognition accuracy. Face recognition accuracy abruptly boosts to 94% from 49%. For up to 5 facial images appearing in an image produces 100% recognition accuracy after facial image enhancement. Moreover, we observe a decreasing trend in face recognition accuracy even after the application of image enhancement as the number of faces increase more than 17 in an image.
- For face recognition, the image enhancement scheme also introduced some illuminations for the case when there were eyeglasses on the face. Due to the aforementioned fact, we conclude the average face recognition accuracy to be least among vehicle and player detection after the application of image enhancement scheme. These facts are also shown in bottom line of column 3, column 5, and column 7.
- We observe almost similar results for vehicle and player detection after the image enhancement scheme is applied. Up to 8 vehicles or players appearing in the input image, the image enhancement scheme helps detector to achieve 100% detection.
- Like the face recognition, we observe a decreasing trend in detection accuracy when the numbers of objects increase more than 17 for vehicles and players even after the application of image enhancement scheme. This fact is also shown in column 5 and column 7 in Table 2.
- Applied image enhancement scheme significantly boosts the detection accuracy of the AdaBoost to 94.9352% from 41.5141% and 97.2538% from 42.9453% for vehicles and players', respectively.
- Generally speaking, the MSR boosts the detection accuracy of objects from 41.5141% (vehicles) to 97.2538% (players) under non-uniform or bad lighting conditions.
- It is important to state here that the MSR barely interrupts the real-time detection capability of the detector. We observe that the MSR image enhancement algorithm consumes about 0.6991 Seconds to enhance a 640×480 pixels image.
- Table 2 also contains the words "NOT TESTED" for faces and players columns. We believe that it is hard to obtain more than 34 neutral frontal faces in the input image. Although there are cases where we can

TABLE 2: ACCURACY COMPARISON OF FACES, VEHICLES, AND PLAYERS BEFORE AND AFTER IMAGE ENHANCEMENT

Objects per Image	Type of Object					
	Face(s) Recognition Accuracy %		Vehicle(s) Detection Accuracy %		Players Detection Accuracy %	
	Before Enhancement	After Enhancement	Before Enhancement	After Enhancement	Before Enhancement	After Enhancement
1	40	100	0	100	0	100
2	45	100	0	100	0	100
3	45	100	45	100	50	100
4	45	100	45	100	50	100
5	50	100	40	100	55	100
6	50	98	40	100	55	100
7	52.65	97.90	42	100	42	100
8	54	94.95	45	100	45	100
10	57.75	94.90	50	100	50	98
12	58.61	94.90	51.75	95	51.75	96
15	58.67	89.65	54	95	54	95
17	58.95	88.40	54.54	88	54.54	87.80
19	54.45	85.80	51	88	51	87.50
21	48	85	49.80	88	NOT TESTED	NOT TESTED
34	31	85	49.75	87.50	NOT TESTED	NOT TESTED
41	NOT TESTED	NOT TESTED	44	87.50	NOT TESTED	NOT TESTED
45	NOT TESTED	NOT TESTED	43.90	84.90	NOT TESTED	NOT TESTED
Average	**49.9386**	**94.3000**	**41.5141**	**94.9352**	**42.9453**	**97.2538**

obtain even more than 50 faces in an image, such as audience in a stadium. But it is hard to get the neutral faces in this case as we will come across with expression and pose challenge.

- Similarly, we believe that there are maximum 22 players' in sports, such as soccer or cricket. It barely happens that all 22 players are focused in the input image. For our work, we were able to capture 19 players' in an input image. We sincerely hope that a suitable justification has been given for this argument.

VI. CONCLUSIONS AND FUTURE WORK

In this paper, we apply our own developed Multi-Scale Retinex (MSR) algorithm as a pre-processing module to boost the accuracy of AdaBoost algorithm, which is considered to be state-of-the-art algorithm for robust object detection and recognition. Simulations performed on 5312 images revealed that the MSR can be reliably and effectively used under non-uniform illuminations to boost the accuracy of the AdaBoost.

With the rapid advancements in object detection and recognition algorithms, completely automated and reliable object detectors and recognizers seem to operate effectively in non-uniform illumination in the future for several entities, such as face, vehicles, players, pedestrian, license plates, fingerprints, or speech. We observed that the detection and recognition accuracies of even state-of-the-art object detectors and recognizer rapidly degrade under non-uniform illuminations. For an algorithm to be more robust to detect and recognize objects, a state-of-the-art image enhancement algorithm is needed to rectify the images degraded due to non-uniform illuminations under strong sun light to boost the accuracy. This was the focus of our work that considered general and practical aspects of image enhancement.

Based on the simulation results and analysis presented in Section 5, we conclude that in future, modifications in the MSR can be proposed to further boost the accuracies of state-of-the-art object detection and recognition algorithms. In the future, we also anticipate to develop the parallel version of the MSR to drastically reduce the execution time. This will indeed help to achieve the higher detection and recognition accuracies in near real-time.

REFERENCES

[1] Z. Mahmood, T. Ali, S. Khattak, L. Hassan, and S. U. Khan, "Automatic player detection and identification for sports entertainment applications," *Pattern Analysis and Applications*, 2015, Vol. 18, No. 4, pp. 971-982.

[2] A. Broggi, E. Cardarelli, S. Cattani, P. Medici, M. Sabbateli, "Vehicle detection for autonomous parking using a soft-cascade AdaBoost classifier," *Intelligent Vehicles Symposium Proceedings* (*IV*)*, 2014 IEEE*, USA, pp. 912-917.

[3] M. Kafai, B. Bhanu, "Dynamic bayesian networks for vehicle classification in video," *IEEE Transactions on Industrial Informatics*, Vol. 8, No. 1, Feb 2012, pp. 100-109.

[4] Z. Mahmood, T. Ali, and S. Khattak, "Automatic player detection and recognition in images using AdaBoost," *Applied Sciences and Technology* (*IBCAST*)*, 9th International Bhurban Conference*, 2012, pp. 64-69.

[5] S. Zafeiriou, G. Atkinson, M. F. Hansen, W. Smith, "Face recognition and verification using photometric stereo: The Photo face database and a comprehensive evaluation," *IEEE Transactions on Information Forensics and Security*, Vol. 8, No. 1, Jan 2013, pp. 121-135.

[6] Z. Mahmood, T. Ali, and S. Khattak, "A Color Image Enhancement Technique using Multi-Scale Retinex," *11th International Conference on Frontiers of Information Technology (FIT)*, Islamabad, Pakistan, 2013, pp. 119-124.

[7] P. Viola and M. J. Jones, "Robust real-time face detection," *Int. J. Comput. Vis.*, Vol. 57, May 2004, pp. 137–154.

[8] Z. Mahmood, M. U. S. Khan, M. Jawad, S. U. Khan, and L. T. Yang "A parallel framework for object detection and recognition for secure vehicle parking," *17th International Conference on High Performance*

Computing and Communications (*HPCC*), New York, Aug. 2015, USA, pp. 892-895.

[9] Z. Mahmood, T. Ali, and S. Khattak, S. U. Khan, and L. T. Yang, "Automatic Vehicle Detection and Driver Identification Framework for Secure Vehicle Parking," *13th International Conference on Frontiers of Information Technology (FIT)*, Islamabad, Pakistan, 2015, pp. 6-11.

[10] J. Lu, K. Plataniotis, A. Venetsanopoulos, and S. Li, "Ensemble-based discriminant learning with boosting for face recognition," *IEEE Trans. Neural Networks*, Vol. 17, no. 1, pp. 166- 178, Jan. 2006.

[11] Z. Mahmood, T. Ali, and S. U. Khan, "The effects of pose and image resolution on automatic face recognition," *IET Biometrics*, 2016, DOI. 10.1049/iet-bmt.2015.0008.