Eyeglasses detection based on learning and non-learning based classification schemes

Conference Paper · April 2017

DOI: 10.1109/THS.2017.7943484

3 authors:

Ahmad Saeed Mohammad
Al-Mustansiriya University

Ajita Rattani
Wichita State University

Reza Derakhshani
University of Missouri - Kansas City

Eyeglasses Detection based on Learning and Non-learning based Classification Schemes

Ahmad Saeed Mohammad
School of Computing and Engineering
University of Missouri-Kansas City
Missouri, Kansas City 64110
Email: asm2x9@mail.umkc.edu

Ajita Rattani
School of Computing and Engineering
University of Missouri-Kansas City
Missouri, Kansas City 64110
Email: rattania@umkc.edu

Reza Derahkshani
School of Computing and Engineering
University of Missouri-Kansas City
Missouri, Kansas City 64110
Email: derakhshanir@umkc.edu

***Abstract*—In this paper, we propose two schemes for prescription eyeglasses detection. The first proposed scheme is non-learning which uses Viola-Jones to detect Region of Interest (ROI) followed by glass detection yielding an overall accuracy of 99.0% for FERET and 97.9% for VISOB datasets. The second scheme is the learning-based scheme consisting of three main steps (a) ROI detection, (b) Local Binary Pattern (LBP) and Histogram of Gradients (HOG) feature extraction and fusion them together, and (c) applying classifiers such as Support Vector Machine (SVM), Multi-Layer Perceptron (MLP), and Linear Discriminant Analysis (LDA), and fusing the output of these classifiers. The latter obtained a best overall accuracy of about 99.3% on FERET and 100% on VISOB dataset.**

I. Introduction

Face biometrics is one of the most popular and widely adopted biometric traits due to being natural and non-intrusive. General face recognition pipeline consists of image acquisition, face detection, image pre-processing, feature extraction and matching. Face detection consists of detecting the facial region of interest (ROI) from the image. Then the detected face is preprocessed to account for intra-class variations such as those caused by lighting and pose, followed by feature extraction and matching [1].

The challenges associated with face recognition can be attributed to factors such as pose, facial expression, illumination variations and the presence of intervening structural components such as eyeglasses among other things [1], [2]. These challenges have attracted researchers from various backgrounds such as psychology, pattern recognition, computer vision, and computer graphics.

Eyeglasses are considered as a confounding factor of face recognition systems. This is because eyeglasses due to frame occlusion and reflections may cover a significant part of the face over the ocular region. Furthermore, the presence of eyeglasses may cause inaccurate eye pair detection, a detrimental outcome given that most of the face recognition systems depend on precise eye pair detection for accurate face detection and registration, Occlusion of the ocular and periocular regions, due to the presence of eyeglasses may result in degraded biometric performance given the importance of these regions [2], [3], [4].

To mitigate the effect of eyeglasses, researchers have been working on eyeglasses detection [5], [6], [7], [8], [9], localization, and removal [10]. Accurate eyeglass detection which is the focus of this work is an important first step towards its localization and removal.

The challenges associated with accurate eyeglasses detection are due to reflections, shadows, and dark images due to low lighting or skin tone (Figure 1).

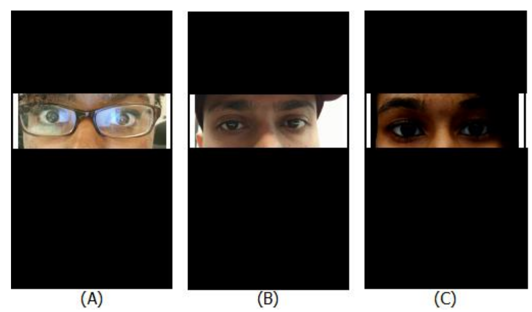

Fig. 1. (A) reflective glasses, (B) shadow, and (C) dark images causing challenges in accurate eyeglasses detection. The upper and lower facial regions have been masked to preserve volunteers' privacy (VISOB database).

In this paper we propose and compare robust learning and non-learning based schemes for accurate eyeglasses detection. The proposed schemes minimize the effects of confounding factors such as glass reflections, shadows, adverse illumination, and under-exposed images, which usually lead to inaccurate eye and eyeglasses detection, and have not been satisfactorily addressed in related literature to date.

In summary, the two-fold contributions of this work are as follows:

1) Introducing two schemes for eyeglasses detection that are robust under varying illumination conditions and skin tones. Challenges such as glass reflections and shadows are also mitigated.
2) Large scale evaluation of the proposed schemes on the publicly available VISOB database. Comparative assessment of the obtained results were also done with the existing work on FERET database.

978-1-5090-6356-7/17/$31.00 ©2017 IEEE

Experimental results suggest that the proposed schemes obtain better eyeglasses detection accuracy in comparison to existing methods.

The rest of this paper is organized as follows: section 2 discusses the existing work on eyeglasses detection. Section 3 discusses our proposed approach to eyeglasses detection. Experimental evaluations and results are discussed in section 4. Conclusions are drawn in section 5.

II. PREVIOUS WORK

The existing literature on eyeglasses detection can be broadly classified into (a) learning-based, and (b) non-learning based.

Fernndez et al., [5] proposed a learning based eyeglasses detection. they used the local binary pattern (LBP) and Support Vector Machines (SVM) achieving a detection accuracy of 98.65%.

Wu et al., [6] proposed detecting eyeglasses based on AdaBoost and Support Vector Machine (SVM) using Haar and Gabor features. Their experimental investigation on FERET database showed an overall correct detection accuracy in the range of 95.5% to 98.0%.

Jing et al., [7] proposed a combination of edge information (such as strength and orientation) and geometric features (such as convexity, symmetry, smoothness, and continuity with a deformable contour) for eyeglasses detection. The proposed method was applied to a subset of frontal face images, reporting on eyeglasses detection rate of 99.5%.

Wu et al. [8] used 3D Hough transform to obtain the 3D plane belonging to the frame of the eyeglasses in stereo face images. The reported accuracy of eyeglasses detection only for the frame was 80%.

Jang et al., [9] calculated the likelihood of eyeglasses using edge information by computing Fishers criterion value on different region inside inter-ocular area. This value give a cue of eyeglasses existence. The proposed scheme was evaluated on Olivetti Research Ldt (ORL) database containing 40 subjects with 10 images per subject. The reported fisher's values ranged between 4.9 to 9.6.

III. THE PROPOSED APPROACH

In this paper, we propose two schemes for prescription eyeglasses detection following the earlier mentioned dichotomy of methods:

A. Non-learning based

The first proposed scheme uses Viola-Jones to detect ocular ROI using facial geometric information, followed by glass detection (Figure 2). Figure 3 shows the Viola-Jones [11] based ROI detection using geometric information. In this scheme, we find the maximum width $(Wmax)$ which is the maximum of $(w1)$ and $(w2)$ (see Figure 3) as detected by Viola-Jones. The minimum width (Wmin) is estimated in the same fashion. Further, we divide the minimum width (Wmin) by the maximum width (Wmax) to estimate the ratio (Wr). Other facial parts, such as the nose and the eyebrows were used as alternative landmarks to detect the ocular ROI when the earlier ROI finder fails due to reflection on glasses, as shown in Figure 4.

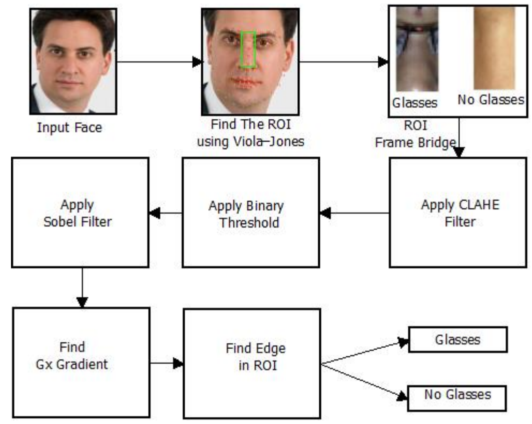

Fig. 2. The block diagram for our non-learning based scheme.

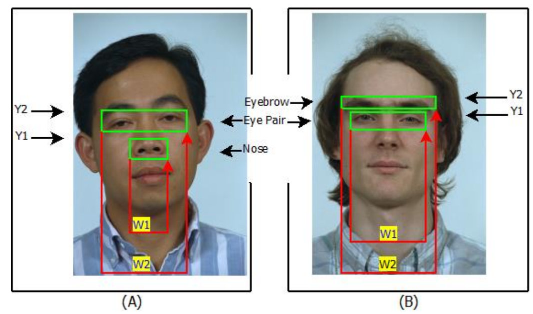

Fig. 3. Viola-Jones based nose detector applied to eye pair region estimation; (A) original face, (B) detected nose region, and (C) detected eye pair region using facial geometric information.

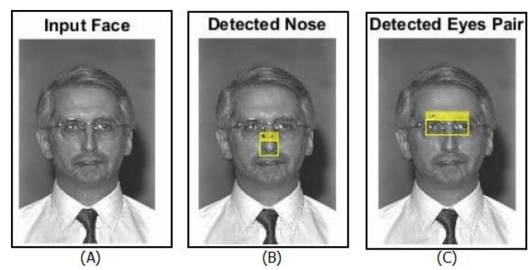

Fig. 4. Viola-Jones based nose detector applied for eye pair region estimation; (A) original face, (B) detected nose region, and (C) detected eye pair region using facial geometric information.

Under-exposed images are mitigated by applying CLAHE [3] to grayscale image. After that, Sobel filter is applied [12], and the gradients (Gx) and (Gy) are estimated. Further, the threshold for glass detection is selected in an adaptive manner based on the illumination and skin tone, as shown in Figure 5.

B. Learning-based

Our learning-based scheme consists of three main steps (a) ROI detection, (b) Local Binary Pattern (LBP) and Histogram

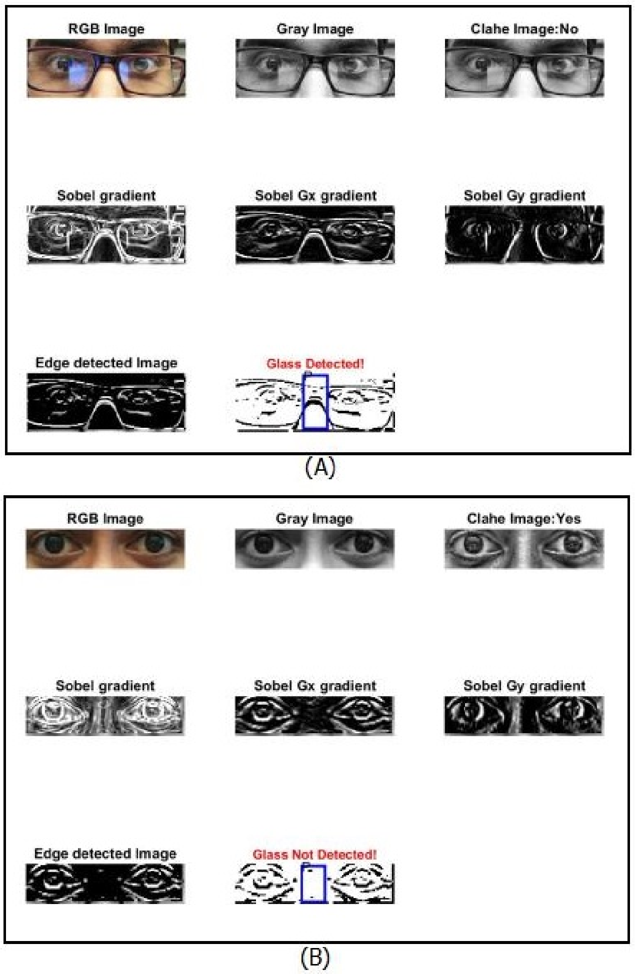

Fig. 5. The proposed eyeglasses detection approach on sample images (A) with and (B) without eyeglasses (VISOB database).

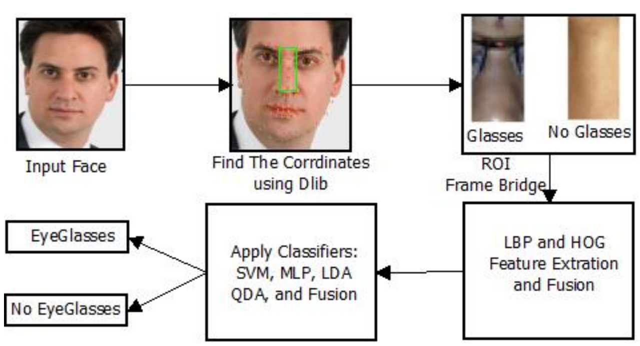

Fig. 6. The block diagram of learning based eyeglasses detection scheme.

of Gradient (HOG) based feature extraction from the ROI, and (c) applying suppvised learning using Support Vector Machine (SVM), Multi-Layer Perceptron (MLP), and Linear Discriminant Analysis (LDA) [13], [14], [15], [16], then fusing the output of these classifiers, as shown in Figure 6.

For ROI detection, we used facial landmark localization using dlib-ml library [17] instead of Viola-Jones and generated 68 landmarks for localizing eyeglasses region and the that of

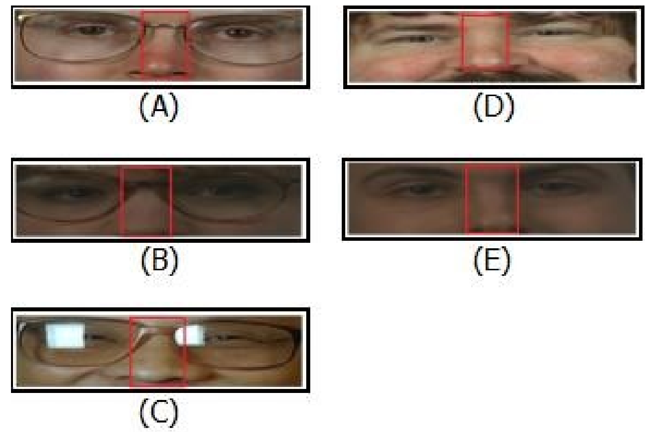

Fig. 7. ROI detection using landmarks from dlib (A) normal condition, (B) dark condition, (C) reflection, (D) no eyeglasses in office light, and (E) no eyeglasses in low light condition.

eyeglasses bridge using location of the nose, ear and eyebrows (see Figure 7). The dlib's 68 landmark seem to be more robust than Viola-Jones in presence of reflection or over and under exposure.

IV. EXPERIMENTS

In this section, we describe our databases and results.

A. Databases

1) Visible Ocular Biometric Database (VISOB): VISOB database [18] consist of eye images from 550 healthy adult volunteers acquired using three different smartphones namely, iPhone 5, Samsung Note 4 and, Oppo N1. Volunteers were asked to take "selfie" like captures using front facing cameras of these mobile devices in two different sessions that were about 10 to 15 minutes apart. The volunteers used the mobile phone naturally within the distance of 8 to 12 inches from the face.

For each session, burst of images were captured in three indoor settings for each volunteer. Under each lighting condition, each shot was repeated with and without glasses, if the volunteer used prescription glasses. Each shot was also repeated while asking the volunteer to look straight at the cell phone screen or glance away to the left and right. The indoor conditions differed in lighting as follows: room lights on, room lights off (but dim ambient lighting still present), and finally in front of big windows in bright daylight setting.

2) Face Recognition Technology Database (FERET): This database contains both gray scale and color images [19]. The number of frontal face images from FERET database is 1978 from 989 subjects, where 233 wore eyeglasses.

B. Results

The above mentioned VISOB and FERET databases were used for performance evaluation of the proposed glass detection schemes. For VISOB, we randomly selected 20491 samples with and without glasses and 2274 samples with and without glasses, for training and validation, respectively,

978-1-5090-6356-7/17/$31.00 ©2017 IEEE

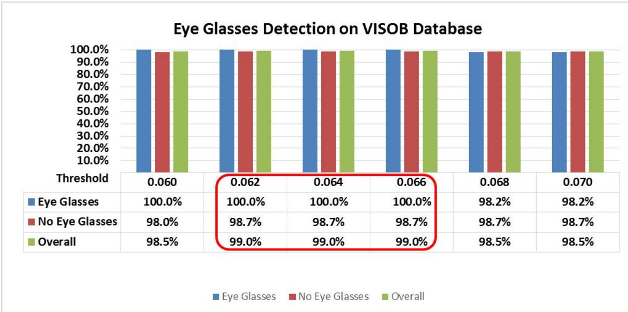

Fig. 8. The bar graph showing the performance of the proposed eyeglasses detection approach on VISOB database. The maximum overall obtained accuracy is 99.0%.

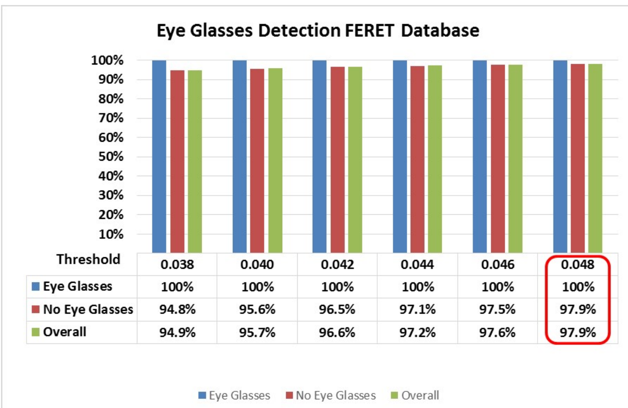

Fig. 9. The bar graph showing the performance of the proposed eyeglasses detection approach on FERET database. The maximum overall obtained accuracy is 97.9%.

of the proposed scheme. The performance evaluation was done on unseen 1456 samples with and without glasses. For FERET, we randomly selected 1467 samples with and without glasses and 620 samples with and without glasses, for training and validation, respectively, of the proposed scheme. The performance evaluation was done on unseen 616 samples with and without glasses.

Figure 8 shows the performance of the proposed non-learning based eyeglasses detection approach on VISOB database. The eyeglasses detection accuracy is ascertained as percentage of correctly classified ROI, correct non-glasses detection rate, correct glasses detection rate, and the overall accuracy; representing the True Positive Rate (TPR), True Negative Rate (TNR), and overall performance, respectively. It can be seen (Figure 14) that the best performances are obtained when the adaptive threshold used to detect eyeglasses varies in the range $[0.062, 0.066]$ with an overall accuracy of about 99.0%.

Figure 9 shows the performance of the proposed approach on FERET database using our proposed method. The best performance of 97.9% accuracy for non-eyeglasses, 100% for eyeglasses, and 97.9% of overall performance with adaptive threshold was obtained.

Table I, shows a comparative evaluation of our proposed

TABLE I
COMPARATIVE ASSESSMENT OF THE PROPOSED APPROACH WITH EXISTING STUDY IN [6]

Method Name	eyeglasses [%]	No eye-glasses [%]	Overall [%]
Proposed Method on VISOB	100	98.7	99.0
Proposed Method on FERET	100	97.9	97.9
Bo el al., [6] on FERET	97.1	97.7	97.2

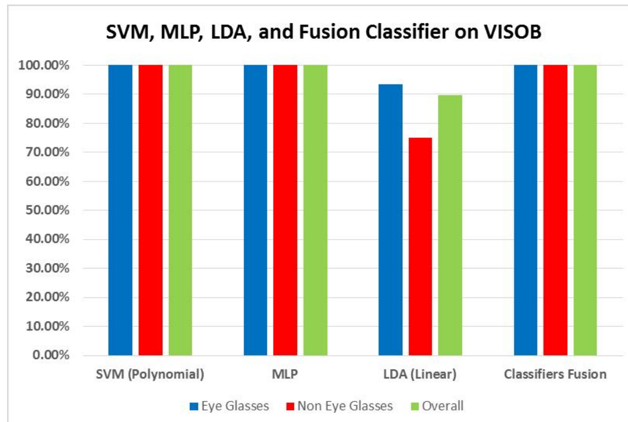

Fig. 10. The bar graph showing the performance of the Learning based eyeglasses detection approach on VISOB database. The maximum overall accuracy is 100%.

eyeglasses detection approach with an existing study [6] reporting their *best* results on FERET database. As can be seen, our method performs very well on challenging mobile VISOB database (Figure 14), and on FERET database (Figure 15) higher eyeglasses detection accuracy and no eyeglasses detection accuracy over [6].

Figure 10 shows the performance of the proposed learning based eyeglasses detection approach on VISOB database. Again, the eyeglasses detection accuracy is ascertained as percentage of correctly classified ROI, correct non-glasses detection rate, correct glasses detection rate and the overall accuracy, representing True Positive Rate (TPR), True Negative Rate (TNR), and overall performance, respectively (Figure 10).

Figure 11 shows the performance of the proposed approach on FERET database using adaptive threshold for eyeglasses detection. The best performance obtained is 99.2% accuracy for non-eyeglasses detection, 100% for eyeglasses detection, and 99.3% for overall performance using SVM and fused classifiers.

Table II shows the performance of learning-based scheme for SVM, MLP, LDA and their fusion. SVM outperformed all other classifiers with an overall accuracy of 100%. Two level of fusion have been used in this scheme. At first level, feature fusion has been applied to concatenate LPB features with HOG

978-1-5090-6356-7/17/$31.00 ©2017 IEEE

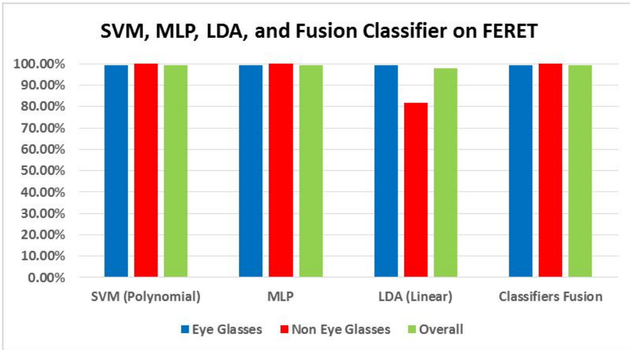

Fig. 11. The bar graph showing the performance of the learning based eyeglasses detection approach on FERET database. The maximum overall accuracy is 99.3%.

TABLE II
OVERALL ACCURACY OF THE PROPOSED LEARNING BASED SCHEMES FOR EYEGLASSES DETECTION USING THREE DIFFERENT CLASSIFIERS I.E., SVM, MLP, LDA AND THEIR FUSION.

Classifier	Database	EyeGlasses [%]	No Eye-Glasses [%]	Overall [%]
SVM	VISOB	100	100	100
SVM	FERET	99.2	100	99.3
MLP	VISOB	100	100	100
MLP	FERET	99.2	100	99.3
LDA	VISOB	93.5	75.0	89.7
LDA	FERET	99.2	100	97.9
Fused Classifier	VISOB	100	100	100
Fused Classifier	FERET	99.2	100	99.3

features. At the second level, output of SVM, MLP and LDA have been fused at the decision level for majority voting. The proposed fusion classifier outperforms existing learning-based schemes based on individual features such as LBP [5].

ACKNOWLEDGMENT

This work has been supported by a research grant from EyeVerity[1].

REFERENCES

[1] J. Anil and L. P. Suresh, “Literature survey on face and face expression recognition.” in *IEEE International Conference on Circuit, Power and Computing Technologies (ICCPCT)*, Mar 2016, p. 16.
[2] A. Ramchandra and R. Kumar, “Overview of face recognition system challenges,” *International Journal of Scientific and Technology Research*, vol. 2, pp. ISSN 2277–8616, August 2013.
[3] U. Park, A. Ross, and A. K. Jain, “Periocular biometrics in the visible spectrum: A feasibility study in biometrics: Theory, applications, and systems,” in *BTAS IEEE 3rd International Conference Biometrics: Theory, Applications, and Systems*, Sept 2009, p. 16.
[4] A. Rattani and R. Derakhshani, “Ocular biometrics in the visible spectrum: A survey,” *Image and Vision Computing*, vol. 59, pp. 1 – 16, 2017.
[5] A. Fernndez, R. Garca, R. Usamentiaga, and R. Casado, “Glasses detection on real images based on robust alignment,” in *IEEE 3rd International Conference on Future Internet of Things and Cloud (FiCloud)*, May 2015, pp. 519–531.

[1]www.eyeverify.com.

[6] B. Wu, H. Ai, and R. Liu, Eds., *Glasses Detection by Boosting Simple Wavelet Features*. ICPR , Cambridge , UK, Aug. 2004.
[7] Z. Jing and R. Mariani, “Glasses detection and extraction by deformable contour,” in *ICPR*, 2000.
[8] H. Wu, G. Yoshikawa, T. Shioyama, S. Lao, and M. Kawade, “Glasses frame detection with 3d hough transform,” in *ICPR*, 2002.
[9] X. Jiang, M. Binkert, B. Achermann, and H. Bunke, “Towards detection of glasses in facial images,” in *Proc. 14th ICPR*. Brisbane, 1998.
[10] C. Wu, C. Liu, H. Shum, Y. Xy, and Z. Zhang, “Automatic eyeglasses removal from face images,” *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 26, no. 3, pp. 322–336, Mar. 2004. [Online]. Available: http://dx.doi.org/10.1109/TPAMI.2004.1262319
[11] P. Viola and D. S. MJ. Jones, “Detecting pedestrians using patterns of motion and appearance,” *International Journal of Computer Vision*, vol. 63, no. 2, pp. 153–161, 2005. [Online]. Available: http://dx.doi.org/10.1007/s11263-005-6644-8
[12] Zuiderveld and Karel, “Contrast limited adaptive histograph equalization,” *Graphic Gems IV. San Diego: Academic Press Professional*, pp. 474 – 485., 1994.
[13] K. Kuman, T. Babu, N. vaishy, and K. Lavanya, “Stock market prediction by non-linear combination based on support vector machine regression model.” *International Journal of Advanced Research in Computer Science*, vol. 7, no. 7, 2017.
[14] A. Mohammad, “Weather temperature forecasting using artificial neural network.” *Journal of Engineering and Development and Al-Mustansyriah University*, vol. 15, no. 2, 2011.
[15] Y. Geng, J. Chen, R. Fu, G. Bao, and K. Pahlavan, “Enlighten wearable physiological monitoring systems: On-body rf characteristics based human motion classification using a support vector machine.” in *IEEE transactions on mobile computing*, vol. 15(3), 2016, pp. 656–671.
[16] A. Mohammad, A. Hamoudi, and Y. Kareem, “Hand writing numbers detection using artificial neural networks,” *Iraqi Journal of Computer and Communication and Control and System Engineering and University of Technology*, vol. 13, no. 2, 2013.
[17] E. D. King, “Dlib-ml: A machine learning toolkit,” *Journal of Machine Learning Research*, vol. 10, pp. 1755–1758, July 2009.
[18] A. Rattani, R. Derakhshani, S. K. Saripalle, and V. Gottemukkula, “Icip 2016 competition on mobile ocular biometric recognition.” in *2016 IEEE International Conference on Image Processing (ICIP).*, 2016.
[19] http://www.nist.gov/itl/iad/ig/feret.cfm., 2011, [Date created: January 25, 2011 and Last updated: May 13, 2015].