ISSN: 2278 – 909X
International Journal of Advanced Research in Electronics and Communication Engineering (IJARECE)
Volume 7, Issue 4, April 2018

Multiple Face Detection on Distorted Images using NSS-HOG Features and Neuro-SVM Classifier

[1]Gagan Kumar Pataskar, [2]Abhishek Jain
[1]M.Tech Scholar, [2]Assistant Professor
SATI, Vidisha
[1]gagan.pataskar@gmail.com, [2]jain5662@gmail.com

Abstract- Motivated by the proliferation of low-cost digital Cameras in mobile devices used in automatic surveillance Networks, interaction between the perceptive image quality and the face detection has been identified. There is degradation in performance of a Face detector when quality of natural images degrades, such as noise or blur, while recording, storage and the transmission. It can be observed that in a certain interval of perceived image. In this paper first of all NIQE score is evaluated for each natural images and further NSS_HOG features are extracted that is robust to common and important image distortions such as Gaussian blur and AWGN noise. Secondly Neuro-SVM Classifier is developed which shows the accuracy of classifier greater than some existing classifiers such as LDA, SVM and AdaBoost. To facilitate this study, a new Face Database, containing face and non–face patches from multi-face images impaired by a variety of common distortion types and levels.

Keywords:- **Face detection, NSS, HOG, Neuro-SVM, NIQE, Accuracy.**

I. INTRODUCTION

As computers become faster and faster, new applications for human faces become possible. Examples of such applications are facial recognition in surveillance systems, gestural analysis with user-friendly interfaces or gender recognition in reactive marketing. In this research work a system of face detection and face recognition that has worked well in an uncontrolled environment. In general, working under uncontrolled conditions is one of the most difficult problems of Computer Vision, for example in applications where lighting is radically changing or where we have to handle unpredictable movements. We have tested the system in a real environment without scale and lighting limits to achieve satisfactory performance in real time. According to [1], facial recognition systems can be divided into four different categories, although some methods may belong to more than one category.

Knowledge-based methods, where some rules or relationships between features are encoded.

Feature-invariant approaches, where the idea is to detect the facial features first, such as eyes, mouth, eye brows, and group them into candidate faces [3].

Template-matching methods, where there is a predefined face pattern that is correlated with the image. Point distribution models (PDMs) have also been used for this purpose [4].

Presentation-based methods in which the goal is to form a classifier that learns the characteristics of faces from the learning set with facial and non-facial images.

Numerous classical techniques such as the principal component analysis [5], the Gaussian mixing models [6], the neural networks [7], the hidden Markov models [8], support vector machines [9] and probabilistic models [6]. Many methods of facial recognition have also been proposed. In principle, they can be divided into systems based on the combination of holistic models, on schemes based on local geometric characteristics and on hybrid schemes [10].

Holistic methods use the whole image as raw input for the learning process. Examples of these techniques are the analysis of the main components [5], the independent analysis of the components [11] or the vector support machines [12] applied to the facial recognition.

In the feature-based schemes, some structural features are extracted, such as eyes, mouth, and their local appearance, position, or relative relationship are used for training the classifier. The most successful technique is the elastic bunch graph matching presented in [13] where the authors use Gabor wavelets to extract the basic features for the graph matching scheme.

Hybrid methods try to use the best of the holistic and feature-based approaches combining local features and the whole face to recognize. An example of hybrid methods is the use of eigenfeatures [14], which extends the idea of eigenfaces to specific regions of the face such as mouth, nose, or eyes.

Among the holistic methods, appearance-based methods are the most successful. They are commonly implemented following these steps:

A. Image preprocessing

In Image preprocessing where usually an illumination correction is performed, followed by the localization of some parts of the face for geometrical alignment that makes the feature-based approaches more accurate. Usually the center of the eyes is located, and faces are warped in such a way that distance between eyes remains stable within subjects.

B. Feature extraction

Dimensionality reduction techniques have shown important advantages in some pattern recognition tasks and face processing is not an exception. Usually we achieve a compression of the input data, reducing the storage needs.

In other cases there is also an improvement of the classification results due to reduction of the noise present in the most part of the natural images. Principal component analysis is perhaps one of the most spread dimensionality reduction techniques [5, 16].

C. Feature classification

Once the proper features are extracted, any classifier can be applied. Most of these methods have been successfully used in artificial environments, but do not perform well in many real-world situations as several independent tests have documented.

Today, the most promising approach for face detection is an appearance-based method that is based on the classification, using machine learning methods.

II. RELATED WORKS

This research work combines the ideas of two problems of vision and artificial vision: evaluation of image quality and facial recognition. Image Quality Assessment (IQA) aims to predict the quality of a given image as perceived by human users. The performance of the IQA models is evaluated through correlation measures between the quality values foreseen by the objective on a number of representative test images. Face recognition is a fundamental problem in various image processing applications, including camera focusing, and is a precursor of advanced identification, tracking, etc. Efficient and effective facial recognition algorithms. they have been developed in the last decades. The problem of facial recognition involves the precise identification of the region or regions in any image corresponding to the human face. In the rest of this section, a review of the literature will be made on these two problems.

In [15], a popular regional statistical feature called Histogram of Oriented Gradients (HOG) and SVM classifier to identify dacial portion in an image was introduced. These characteristics are invariant for rotations and 2D illuminations.

In [16], proposed a facial recognition algorithm based on quality aware HOG features. Facial features trained in these functions provide a statistically significant improvement in image distortion tolerance compared to a solid baseline. Distortion-dependent and lens-independent distortion variants are proposed and evaluated on a large database of facial images representing a wide range of distortions. A partial variant of the learning algorithm is also proposed, which further enhances the robustness of these facial sensors. To facilitate this study, a new database of deformed faces will be created, containing face patches and non–face patches of images influenced by a variety of common types and distortion levels.

In [17] a deep convolutional neural network algorithm is proposed for the task of facial recognition. Benchmark experiments available to the public show the success of the method. The face detector can detect faces in a wide range of orientations and expressions. The facial sensor does not require additional modules normally used in in-depth learning methods such as SVM regression or bounding box regression. This work extends the DDFD detector to a lightweight model that improves running and training speed. The facial model combines the results of two different networks formed for facial recognition and local facial parts and shows that the information provided by the latter is crucial to the specific task

In [18] presents a strategic approach for rapid detection and annotation of partially occluded face. Partially Occluded Face Detection (POFD) problem is addressed by using a combination of feature-based and part based face detection methods with the help of face part dictionary. In this approach, the devised algorithm aims to automatically detect face components individually and it starts from mostly un-occluded face component called Nose. Nose is very hard to cover up without drawing suspicion. Keeping nose component as a reference, algorithm search the surrounding area for other main facial features, if any. Once face parts qualify facial geometry, they are normalized (scale and rotational) and tag with annotation about each facial features so that partial face recognition algorithm can be adapted accordingly with the test image.

In [19] performs a face detection based on the use of eyes tracking. By using the harr-like features detect face firstly following by the eyes detection. Locating the position of each eye ball. The series of detecting results can provide to the business to make the further marketing strategies.

Depth information can be used in human face detection. In [20] HOG + LBP were used in feature descriptor of human face in this work and two-level classifier training system was implemented. The test shows the validity and possibility of is as follows: 87 real faces are found out, and 5 non-face objects are error detected out. The detection speed can be maintained at 20 fps if the resolution is set at 320*240.

III. PROPOSED METHODOLOGY

Before describing the proposed methodology, it is necessary to describe the types of image distortions that is considered in this work. A HOG feature uses face–indicative HOG features is then discussed.

A. Image Distortions

In this work two basic types of distortions is considered that commonly occur in digital devices and over communication channels. The image is denoted by a matrix I, such that I(i; j) represents the (i; j)th pixel in the image I.

Additive White Gaussian Noise(AWGN): This is a local distortion, in which a zero mean gaussian noise of variance parameter $\sigma^2{}_N$ is added independently to each pixel.

$I(i; j) = I(i; j) + Nij$; such that $Nij \sim N(0; \sigma^2{}_N)$

Where $N(\mu,\sigma^2)$ is a gaussian distribution with mean μ and variance σ^2.

Gaussian Blur: This is a global distortion in which each pixel is blurred through convolution with a Gaussian low

ISSN: 2278 – 909X
International Journal of Advanced Research in Electronics and Communication Engineering (IJARECE)
Volume 7, Issue 4, April 2018

pass filter of standard deviation σ_B. For computational ease the gaussian kernel is truncated at $6\sigma_B$. The discrete truncated gaussian filter in two dimensions is given as follows:

$$G(x,y) = \frac{1}{2\pi\sigma_B^2} e^{-\frac{x^2+y^2}{2\sigma_B^2}}$$

An image with gaussian blur distortion is given by I= I*G.

B. NSS-HOG face detector

NSS-HOG feature detector is a combined feature detector based on SD-NSS detector and HOG detector.

Spatial Domain NSS is a completely blind Image Quality Assessment (IQA) model is founded on perceptually relevant spatial domain NSS features extracted from local image patches that effectively capture the essential low-order statistics of natural images. The classical spatial NSS model [10] that we use begins by preprocessing the image by processes of local mean removal and divisive normalization:

$$I(i,j) = \frac{I(i,j) - \mu(i,j)}{\sigma(i,j) + 1}$$

where iε{1,2….M} and jε{1,2,….N} are spatial indices, and are the image dimensions, and μ is mean and σ is variance.

The coefficients (1) have been observed to reliably follow a Gaussian distribution when computed from natural images that have suffered little or no apparent distortion [10]. This ideal model, however, is violated when the images do not derive from a natural source (e.g., computer graphics) or when natural images are subjected to unnatural distortions. The degree of modification can be indicative of perceptual distortion severity.

The motivation for these NSS features lies in statistical models of photographs and in low–level models of visual perception. It is well established that the early stages of human vision process images locally. These processes have evolved to encode images using natural statistics for efficient neural transmission and representation in higher–level visual tasks.

HOG Feature is short for Histogram of Oriented Gradient. It is considered that the form feature and color feature of local object, like human face, in an image can be characterized by statistical information of gradient direction density. Since the gradient in an image is mainly lie in boundary region of the local object, the histogram of oriented gradient stands for the edge direction density of detection targets in the image. A detailed discussion on HOG Feature can be viewed in reference [15].

The procedure of HOG Feature extraction is stepped as follow: Firstly, divide the detected image into several connected components that is called as "Cell". Then calculate the gradient pixel by pixel in each Cell and produce the Histogram of Oriented Gradient. Finally, conduct the feature descriptor of detected image by linear combination of all the Cell in the image. Usually the Cells are combined into bigger Block, in which the histogram of Cells is normalized, so that the influence of light and shadow can be reduced in use.

C. Training the Face Detector

To train the face detectors based on NSS-HOG features, implementation of different classifiers such as LDA, SVM, ADABOOST and NeuroSVM was used in the experiments. For each classifier, a preliminary detector was first trained using a small sub–sample of face and non–face patches of the training images.

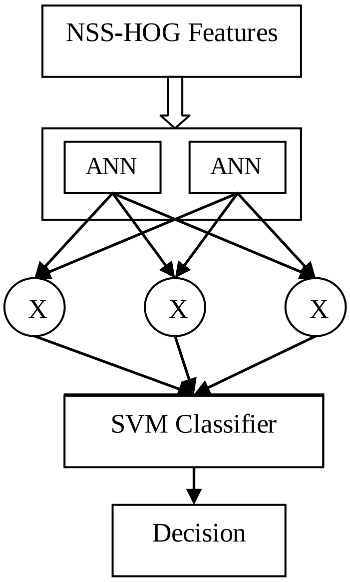

Figure 1: Neuro-SVM Model

D. Testing the Face Detector

Similar to training, test faces are annotated to extract face and non-face region from the corresponding datasets. The precision recall accuracy and F_measure parameters are used as the evaluation metric. Precision is defined as the fraction of detected positives that are faces, i.e., the ratio of true positives to the detected positives. Recall is defined as the fraction of actual positives that is detected, i.e., the ratio of true positives to the total number of positives.

IV. RESULT ANALYSIS

The result analysis is performed on different types of distortion such as AWGN and Gaussian Blur. In practical settings, precise information regarding the distortion types and distortion levels afflicting an image are difficult to estimate. So, NSS_HOG feature are extracted for high performance of distorted image. In training and testing the face detectors, computation involved depends primarily on two tasks: (a) computation of the features (Spatial–NSS and HOG), and (b) learning the Neuro-SVM for classification.

We therefore use NIQE scores as surrogates for perceptual distortion levels. However, as a sanity check, we first assessed the NIQE scores of images against all of the distortion types considered. 10 levels of AWGN were added with the noise variance parameters varying over a log scale, variance= 4.5*10-5; 0.0001; 0.0003, 0.0009; 0.0025; 0.0065; 0.02; 0.05; 0.15; 0.36 as shown in figure 2.

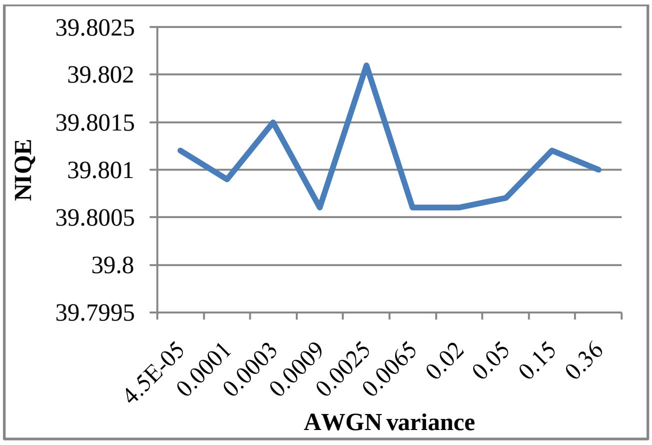

Figure 2: NIQE vs AWGN

The imgaussianfilt() function in MATLAB was used to introduce gaussian blur at 10 levels. The standard deviation of the gaussian filter is equal to {0.4, 1, 2.3, 3.6, 4.5, 6, 7.4, 10, 12, 13} as shown in figure 3.

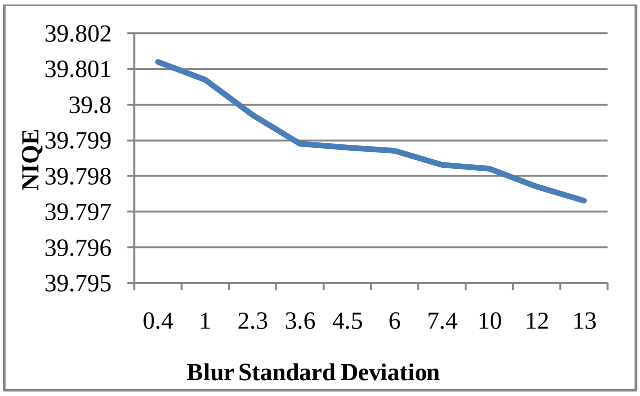

Figure 3: NIQE vs AWGN

To study the performance of proposed face detectors on natural images encountered in real–life, we evaluate the face detectors on a subset of images. One of the image with distortion values are shown below in figure 4-6.

Different classifiers are used to evaluate the performance level of the proposed face detector and result obtained are shown in Table 1-2 and figure 7-8.

Figure 4: Original Image

Figure 5: AWGN Image with variance 0.36

Figure 6: Gaussian Blur Image

Table 1: Performance analysis with AWGN Noise

Classifiers	Recall	Precision	Accuracy	F_measure	Time
LDA	0.8889	1	0.8889	0.9412	0.399
SVM	0.8889	1	0.8889	0.9412	0.364
AdaBoost	0.7222	1	0.7222	0.8387	4.360
Neuro-SVM	0.9444	1	0.9444	0.9714	0.410

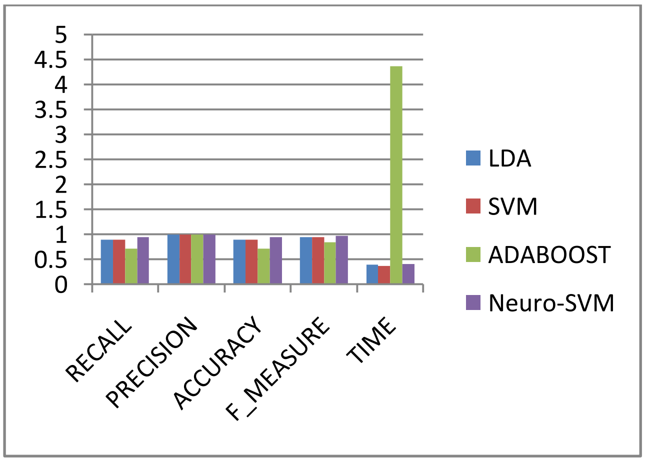

Figure 7: Performance analysis with AWGN Noise

Table 2: Performance analysis with Gaussian Blur

Classifiers	Recall	Precision	Accuracy	F_measure	Time
LDA	0.7143	1	0.7143	0.8333	0.685
SVM	0.625	1	0.625	0.7692	0.675
AdaBoost	0.7143	1	0.7143	0.8333	2.031
Neuro-SVM	1	1	1	1	0.659

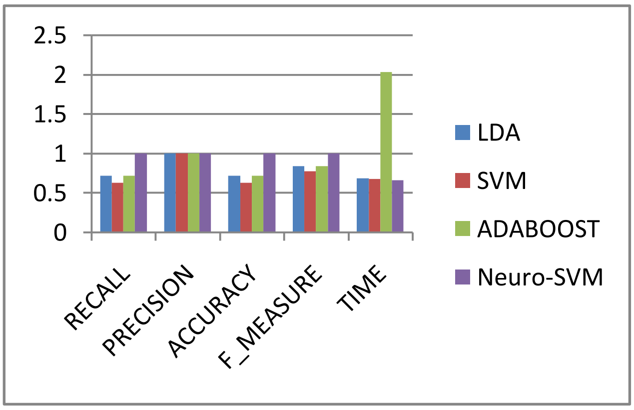

Figure 8: Performance analysis with Gaussian Blur

V. CONCLUSION

This research work examined the degradation of the performance of a popular and effective face detector when the perceived quality of image is degraded by the distortions typically associated with capturing, storing and transmitting images, especially noise or blue. This paper first evaluates the NIQE score for each natural image and extracts additional NSS_HOG features that are robust compared to common and important image distortions such as Gaussian Blur and AWGN noise. Secondly Neuro-SVM Classifier is developed which shows the accuracy of classifier greater than 94%. Interestingly, for AWGN and GBlue, in spite of being distortion–independent, NSS-HOG provides better performance. Thus, the NSS-HOG based face detectors are able to achieve acceptable face detection performance at much higher levels of visual impairments than what is currently possible. As it is known that real life applications have more complex distortion type. In future work, this research work will be extended on such real distorted images. Going forward, it is proceeded towards extensive experimentation on real life dataset of multi-facial images with multiple distortions, and authentic (non–synthetic) distortions drawn from real–life photographic facial imaging applications.

REFERENCES

[1] M. Yang, D. J. Kriegman, and N. Ahuja, "Detecting faces in images: a survey," IEEE Trans. Pattern Anal. Machine Intell., vol. 24, no. 1, pp. 34–58, 2002.

[2] C. Kotropoulos and I. Pitas, "Rule-based face detection in frontal views," in Proc. IEEE Int. Conf. Acoustics, Speech, Signal Processing (ICASSP '97), vol. 4, pp. 2537–2540, Munich, Germany, April 1997.

[3] T. K. Leung, M. C. Burl, and P. Perona, "Finding faces in cluttered scenes using random labeled graph matching," in Proc. 5th International Conference on Computer Vision (ICCV '95), pp. 637–644, Boston, Mass, USA, June 1995.

[4] A. Lanitis, C. Taylor, and T. Cootes, "An Automatic face identification system using flexible appearance models," Image and Vision Computing, vol. 13, no. 5, pp. 393– 401, 1995, online, available: http://citeseer.ist.psu.edu/article/. lanitis95automatic.html

[5] M. Kirby and L. Sirovich, "Application of the Karhunen-Loeve procedure for the characterization of human faces," IEEE Trans. Pattern Anal. Machine Intell., vol. 12, no. 1, pp. 103– 108, 1990.

[6] B. Moghaddam and A. Pentland, "Probabilistic visual learning for object representation," IEEE Trans. Pattern Anal. Machine Intell., vol. 19, no. 7, pp. 696–710, 1997.

[7] K.-K. Sung and T. Poggio, "Example-based learning for viewbased human face detection," IEEE Trans. Pattern Anal. Machine Intell., vol. 20, no. 1, pp. 39–51, 1998.

[8] F. Samaria, Face recognition using hidden Markov models, Ph.D. dissertation, University of Cambridge, Cambridge, UK, 1994.

[9] E. Osuna, R. Freund, and F. Girosi, "Training support vector machines: an application to face detection," in Proc. IEEE Conference on Computer Vision and Pattern Recognition (CVPR '97), pp. 130–136, San Juan, Puerto Rico, USA, June 1997.

[10] W. Zhao, R. Chellappa, P. J. Phillips, and A. Rosenfeld, "Face recognition: a literature survey," ACM Computing Surveys, vol. 35, no. 4, pp. 399–458, 2003.

[11] M. S. Bartlett, M. H. Lades, and T. J. Sejnowski, "Independent component representations for face recognition," in Human Vision and Electronic Imaging III, vol. 3299 of Proceedings of SPIE, pp. 528–539, San Jose, Calif, USA, January 1998.

[12] Z. Li and X. Tang, "Bayesian face recognition using support vector machine and face clustering," in Proc. IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR '04), vol. 2, pp. 374–380, Washington, DC, USA, June– July 2004.

[13] L. Wiskott, J.-M. Fellous, N. Kruger, and C. von der Malsburg, "Face recognition by elastic bunch graph matching," IEEE Trans. Pattern Anal. Machine Intell., vol. 19, no. 7, pp. 775– 779, 1997.

[14] A. Pentland, B. Moghaddam, and T. Starner, "View-based and modular eigenspaces for face recognition," in Proc. IEEE Conference on Computer Vision and Pattern Recognition (CVPR '94), pp. 84–91, Seattle, Wash, USA, June 1994.

[15] N. Dalal and B. Triggs, "Histograms of oriented gradients for human detection," Computer Vision and Pattern Recognition, 2005.

[16] Suriya Gunasekar, Joydeep Ghosh, and Alan C. Bovik, "Face Detection on Distorted Images Augmented by Perceptual Quality–Aware Features", IEEE Transactions on Information Forensics and Security, vol 9, issue 12, pp. 2119 – 2131, 2014.

[17] Danai Triantafyllidou and Anastasios Tefas, "Face detection based on Deep Convolutional Neural Networks exploiting incremental facial part learning", 23rd International Conference on Pattern Recognition, pp. 3560-3565, 2016.

[18] Dr. P.Shanmugavadivu and Ashish Kuma, "Rapid Face Detection and Annotation with Loosely Face Geometry", IEEE, PP.594-597, 2016.

[19] Hsiau Wen Lin, Yi-Hong Lin, "Face detection based on the use of eyes tracking", International Computer Symposium, pp. 402-405, 2016.

[20] Tiemeng Li, Wenjun Hou, Fei Lyu, Yu Lei, Chen Xiao, "Face Detection based-on Depth Information Using HOG-LBP", International Conference on Instrumentation & Measurement, Computer, Communication and Control, pp. 779-784, 2016.