Face Quality Analysis Using Deep Machine Learning

Dr. Shikha Nema, Shweta Mishra, Deepa Jilla, Aditi Bansal
HOD (Electronics and Communication Department), Student, Student, Student
Electronics and Communication Department
Usha Mittal Institute of Technology, SNDT University, Mumbai, India

Abstract: Initially, three different biometric characteristics (finger print, iris and face) were identified for determining an individuals identity but finally face came to be regarded as the most suited to the practicalities of the real world. The only requirement is to make a facial image in compliance with a standard which is universal. The aim of this project is to design an automatic system which verifies compliance of images with ICAO standards (ISO/IEC 19794-5). This paper compares three classification models having feature extracted from Local Binary Pattern, Histogram of Oriented Gradients and a Deep Neural network model. An image analysis methodology using Support Vector Machine (SVM), Convolutional Neural Network (CNN) and K-Nearest Neighbors (KNN) was used to classify images.

IndexTerms - **Face image quality, face recognition, ICAO, ISO/IEC 19794-5, HOG, LBP, CNN, SVM, KNN.**

I. INTRODUCTION

Face identification has proved to be one of the most promising and reliable biometric trait involving applications in video surveillance, forensics, competitive exams, jobs, access control and various other commercial areas. It offers measurability, universality, high performance, uniqueness, permanence, and acceptability. It can provide the capability for real-time verification of an individual's identity. Countries can enforce face identification for protecting their interiors, achieving effective control of borders and enforce immigration laws [3]. Furthermore, it does not require any complex and costly enrolment procedures.

More diverse the data is, more it becomes difficult to achieve interoperability. Therefore, the International Civil Aviation Organization (ICAO) endorsed in 2002, the use of face recognition as the globally interoperable biometric characteristic for machine-assisted identity confirmation with machine-readable travel documents. In sequence, according to ICAO directives, the ISO proposed the ISO/IEC 19794-5 standard which specifies some quality requirements for facial images. For instance, a face image should not present an opened mouth to be included in e-passport [2].

The objective of this work is to propose an automated system approach for evaluating the conformance of facial images to some ISO/ICAO requirements. This might offer numerous benefits such as faster production of a document and improve quality sufficiently to ensure accurate recognition by both humans and computer system.

Machine Learning, a very popular domain has been increasingly used for image-based quality analysis. It is a sub-area of Artificial Intelligence, wherein a set of different models learn from past data, to generalize to new data with an additional feature of handling noisy input and complex data environment and form new concepts from prior knowledge. Supervised learning model was used to classify images as acceptable or not acceptable. Here, we are given a dataset and already know what our correct output should be, and the relationship between the input and output has already been derived. In classification problem, we try to predict results in the discrete output (i.e. map input variables into discrete categories), say a good or bad image.

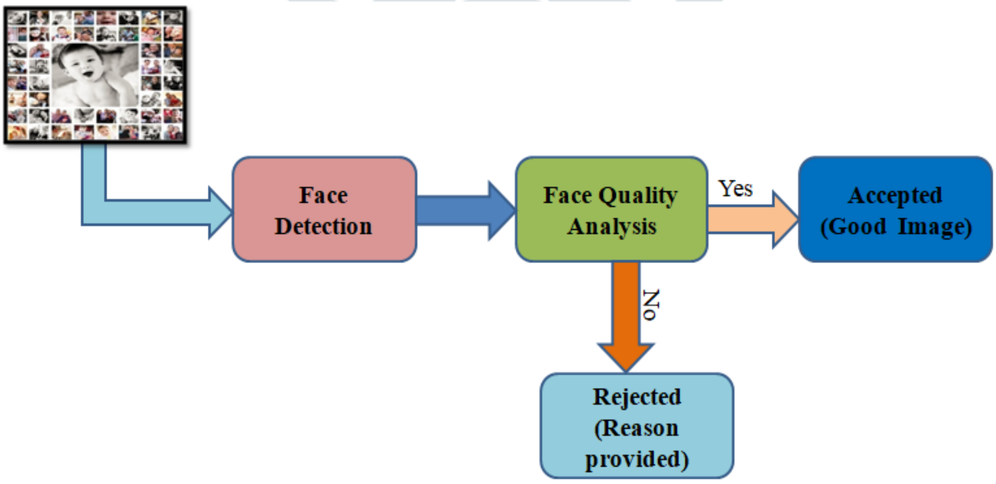

Fig.1 Block diagram of the System

Multiple colored images are given to the face detector system which extracts the facial region for each image. Face detection algorithm finds the location and size of the face in an image. The image is matched with images stored in a database. Any facial feature changes will invalidate the matching process. After detection, the face image is tested for its validity according to ICAO standards. If the image quality complies with the specified standard, then it is considered good and accepted otherwise rejected. In this paper, three image analysis methodologies KNN, SVM and CNN classifier, are used to differentiate good and bad images. The image dataset was formed by manual segregation of publicly available images and databases into different categories, as specified by ICAO. For the two classification problem, we consider combination of bad quality parameters as rejected, the absence of which makes the image acceptable. These are then applied to three different classifiers and comparison results are obtained.

The study is divided into five sections with a feature descriptor and classification algorithm discussed in section III. The image database is described in section IV. Experiments and results have been provided in section V. Conclusions are drawn in section VI.

II. Related Work

In the past, there has been numerous research on facial image quality analysis and bringing them in compliance with ICAO standards. Some of them are discussed here.

Nasrollahi, Kamal, and Thomas B. Moeslund [1] proposed system consisting of face detection followed by face quality assessment using four simple features including out-of-plan rotation, sharpness, brightness and resolution to assess the face quality in a video sequence.

Castillo, Oriana Yuridia Gonzalez [5] presented "Two Dimensional Facial Image Quality" (2DFIQ) project for measuring the quality of digital passport photos based on the specifications of the international standards such as ICAO-MRTD and ICAO/IEC 19794-5. In the analysis and design phase of the project, there were some problems in assigning precedence and relevance of various attributes of a passport photo occurred.

Ferrara, Matteo, Annalisa Franco, and Davide Maltoni [6] proposed a system which focuses on the requirements for face images which was used in Machine Readable Travel Documents, defined in the ISO/IEC 19794-5 standard. In particular, an evaluation framework is proposed for testing software which will automatically verify the compliance of an image to ICAO standard, as mentioned above. The results obtained for the commercial software are reported and compared.

Lawrence, Steve, et al.[8] described a hybrid neural network solution which combines local image sampling, a Self-Organizing Map Neural Network and Convolutional Neural Networks. The CNN extracts successively larger features in a hierarchical set of layers. The system is designed for rapid Classification, fast approximate Normalization, and Preprocessing, Consistency.

Adrian Rosebrock [11] build an image classifier, which uses few training examples. The dataset consisting of dogs and cats images was downloaded from kaggle. Images were then classified as dogs and cats using few hundred or thousand images.

Tiwary, K. [13] proposed article which briefs us about Gradients for Human image gradients and histograms, what role Detection do they play in face or human recognition and how are they used in it. We will get to know how computer differentiate human body and recognizes it based on the feature.

Tiago F. and Bottino, Andrea and Laurentini, Aldo and De Simone, Matteo [10], presented a paper which aims at understanding the most image pairs relevant facial features, how effective can be computer algorithms for detecting siblings pairs, and if they can beat human evaluation. A high quality database containing pictures of sibling pairs includes frontal, profile, expressionless, and smiling faces.

Mahmoud Afifi and Abdelrahman Abdelhamed [9] proposed the face image standard for making face image data suitable for human visual inspection, verification and authentication. It adopts ISO /IEC 19794-5:2005(E). The ISO standard offers all possible applications of automated as well as human visual inspection, which is more restrictive.

The study by US-VISIT consist of manual quality assessment of 20,000 images and application of graphical image categorization tool to label images presenting certain defects such as cropped images, Intensity Saturation, Head Pose, etc. Various quality metrics along with their notable experimental conclusion was presented.

A collection of images by Peter Hancock [12] is useful for conducting experiments in psychology, primarily faces, though other submissions are welcome. These images are free for research use. Three databases Utrecht ECVP, Iranian face, and Pain Expression matched our project concern. Images for parameters such as good, non-frontal, mouth open etc.were collected using these 3 dataset.

The experiments conducted on dataset using LPB, HOG and deep networks features shows that LPB feature extractor has outperformed the other methods. As the results obtained by HOG feature descriptor where not upto the mark and the accuracy was low so a combination of LBP and HOG was used to increase the accuracy. This outperformed the SVM classifier.

III. Background

3.1 Feature Descriptor

3.1.1 Histogram of Oriented Gradients (HOG): It is a feature descriptor which captures features by counting the occurrence of gradient orientation. Traditional HOG divides the image into different cells and computes a histogram of gradient orientations over them. HOG is being applied extensively in object recognition areas as facial recognition [14] [15]. The process for computing HOG is explained in Fig. 2. HOG prediction accuracy, was high when orientation was found to be 10, pixel per cell (PPC) as (18,18) and cell per pixel (CPC) as (2,2).

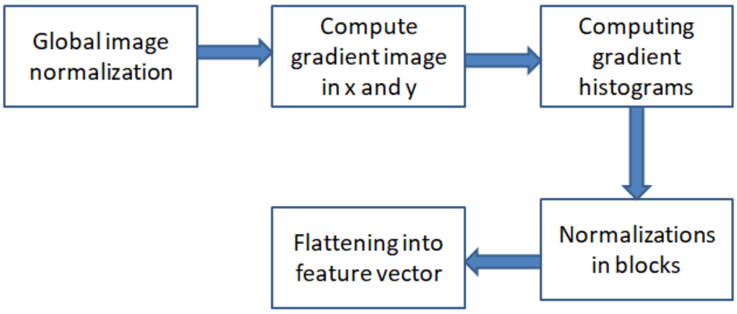

Fig.2 Block diagram of Histogram of Oriented Gradients (HOG)

3.1.2 Local Binary Patterns (LBP): The texture descriptor computes a local representation of texture constructed by considering each pixel with its surrounding neighborhood of pixels. The steps for constructing a LBP texture descriptor are shown in Fig. 3. LBP provides us the benefit of capturing extremely fine-grained details in the image, however, this also serves as a drawback, as the details cannot be captured at varying scales, only the fixed 3x3 scale[14] [16]. Hence two parameters were introduced, by Ojala et al. to handle varying neighborhood sizes:

1. The number of points p in a circularly symmetric neighborhood to consider (thus removing relying on a square neighborhood).
2. The radius of the circle r, which allows to account for different scales.

LBP with numpoints and radius as 48 and 44 respectively was used with SVM classifier to achieve a good accuracy.

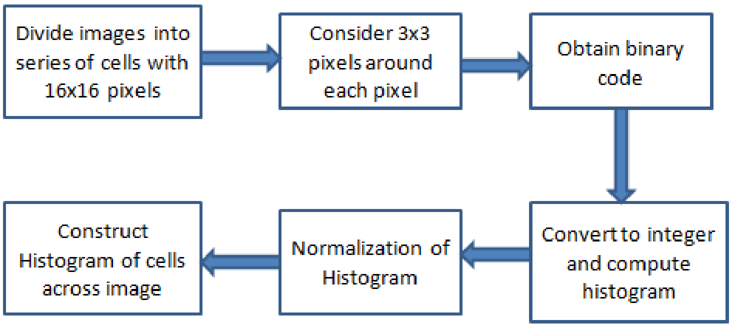

Fig.3 Block diagram of Local Binary Patterns (LBP)

3.2 Classifiers

3.2.1 K-Nearest Neighbor (KNN) Classifiers: This is an important classification and regression predictive algorithm in supervised Machine Learning. KNN is basically used because of its low calculation time and easy interpretation, hence has applications such as pattern recognition, data mining and intrusion detection. In KNN the object is always classified on the basis of majority votes of its neighbors, where 'K' refers to the nearest neighbor we wish to take vote from. The two parameters accessed on different values of 'K' are: training error rate and validation error rate. For training sample, the error rate at k = 1 is always zero because the nearest point to any training data point is itself.

3.2.2 Support Vector Machine (SVM): SVM is a classifier designed for binary problems with extension to multi-class problems. SVM was introduced by Cortes and Vapnik in 1995. The main idea was ensure the networks high generalization ability by mapping inputs non-linearly to high dimensional feature spaces, where linear decision surfaces were constructed with special properties. Some advantages of the SVM are the generalization of binary and regression forms and notation simplification. SVM uses several kernels such as the polynomial, linear, and the Gaussian radial basis function (RBF) kernel [14]. Linear SVC in multi classification is handled according to one-vs-rest scheme. It is similar to SVC with a kernel equal to linear and offers more flexibility in the choice of penalties and loss functions.

3.2.3 Convolution Neural Networks (CNN): Convolution Neural Network also called as Shift Invariant or Space Invariant Artificial Neural Network. It consists of inputs and output layer as well as multiple hidden layers which has their own functions. Hidden layers consist of: Convolution layer, RELU layer i.e. activation function, polling layer, fully connected layer and normalization layer. For this we have used Keras, a high level neural network API, written in python and runs on top of Tensor flow, CNTK or Theano. It is known for its user friendliness, modularity, easy extensibility, runs seamlessly on CPU or GPU, supports CNN, RNN as well as their combination and allows for easy prototyping. Label binarizer converts multi-class labels to binary labels.

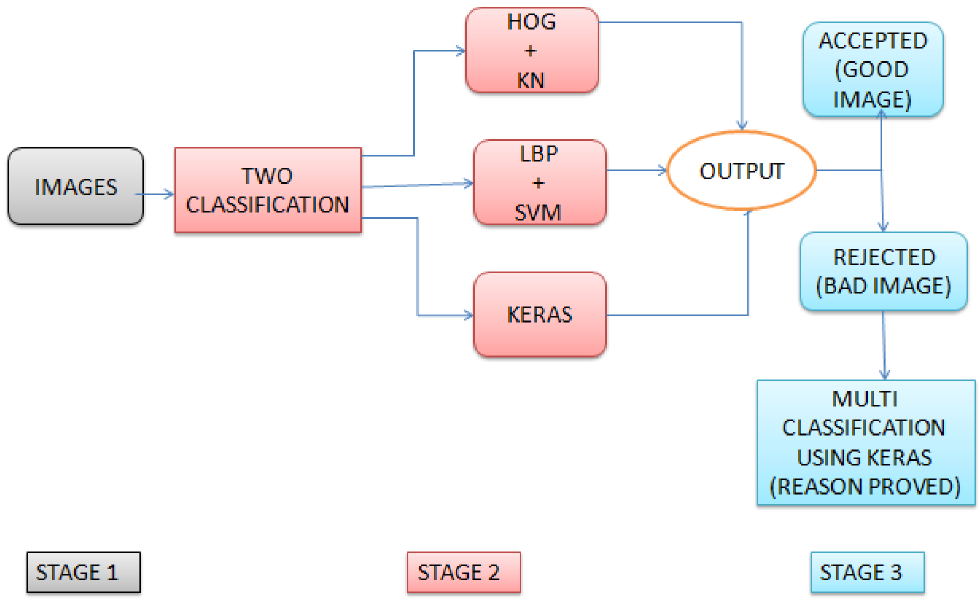

Fig.4 Block diagram of the system presenting three stages

IV. IMAGE DATABASE

The system dataset has been created by manual segregation of publically available images and databases [9] [10] [12]. These images are classified into different parameters listed in Table 1, to bring facial image under the scrutiny of ICAO Standard. Images that are defective or cannot be classified in any category and are simply discarded.

4.1 Methodology

For classification of dataset, three different models based on the aforementioned feature extractors have been considered which will accept or reject a given test image. Python Programming language is used for conducting the feature extraction experiments with its supporting libraries and packages provided by Anaconda Distribution. Scikit-Image library is used for three classification model i.e. LBP, HOG and CNN.

TABLE I: SELECTED PARAMETERS IN COMPLIANCE WITH ICAO STANDARDS

Parameter	Description
Blur	The image shall be in sharp focus also with no smoothening effect
Pixelation	Images should not be grainy or pixilated
Low resolution	Images should not be hazy, foggy or dull
Spectacles	The portrait shall show the eyes with no light reflection off the glasses and no tinted lenses
Non frontal	Looking away, tilt, out of plan rotation is not acceptable
Shadows on face	The lightening shall be uniform, any remarkable shadow is not acceptable
Mouth opened	The applicant shall be looking at the camera with mouth closed
Red eyes	Appearance of red eye due to flash are not accepted
Closed eyes	The applicant shall be looking at the camera with eyes open
Head and face covering	Head covering shall not be accepted in any circumstances (except religious, medical or cultural) and no hair obscuring the face
Too bright	No over exposure to light
Lines on face	Appearance of some lines on images due to low quality are not accepted
Damaged/Ink mark	The photo should be of high quality with no creases or ink marks
Good image	Absence of aforementioned parameters

V. EXPERIMENTS AND RESULTS

5.1 Local Binary Pattern and SVM

LBP is based on texture description. We construct a directory/package 'kimtaehyung', from which class 'LocalBinaryPatterns' is extracted. This class is used for computing the LBP representation of image and building histogram. Sad.py, a python file, placed in the same directory is used to call LocalBinaryPatterns, training the model and classifying images.

Linear SVC, with its SVM C parameter equal to 50 and random state in range of 42-45 was used. LBP with its aforementioned feature i.e. (numpoint, radius = (48, 44)), when used with SVM gave exactly 75 percent accuracy. It was found to be slower when compared with other two models. Images in the training as well as testing folder were preprocessed (for conversion of RGB image to grayscale) and also normalize to an optimal size.

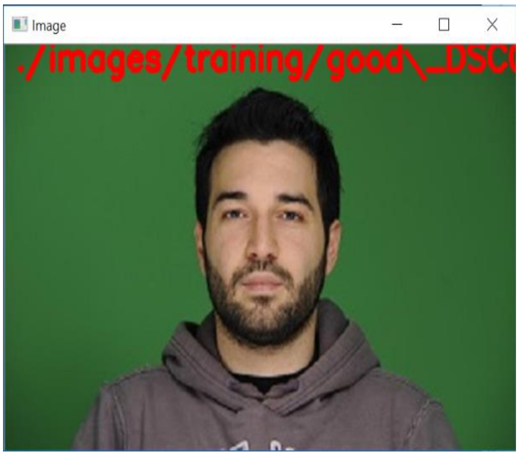

Fig.5 LBP output: good image

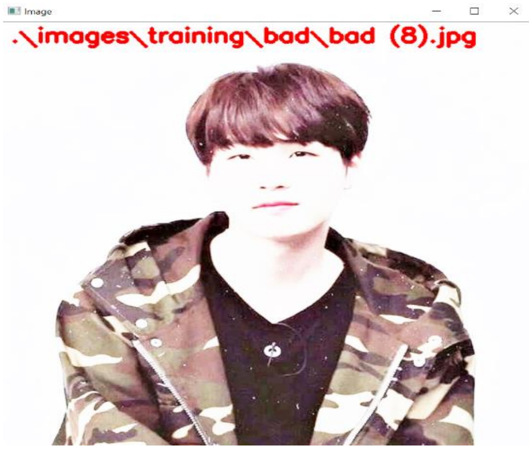

Fig.6 LBP output: bad image

5.2 Histogram of Oriented Gradients and KNN

Here, the testing and training images were same as that of LBP. HOG computes the histogram in the manner, different from LBP and is based on edge description. Fig.7 shows the computation of HOG for the given input image. HOG with its feature, described in section 3, was used for the experiment. HOG predicted 80 percent accuracy when used with KNN. It was also faster than LBP. For proper results, images are to be preprocessed.

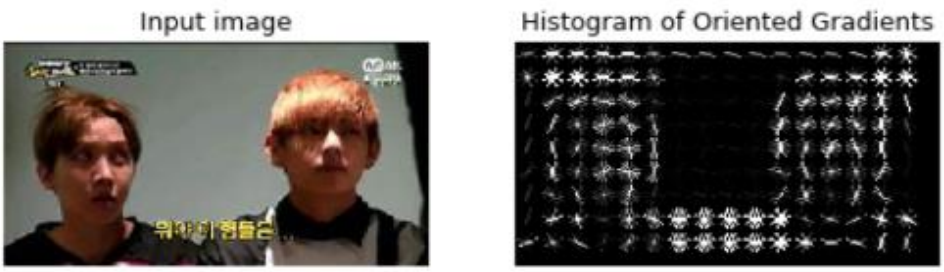

Fig.7 HOG Output

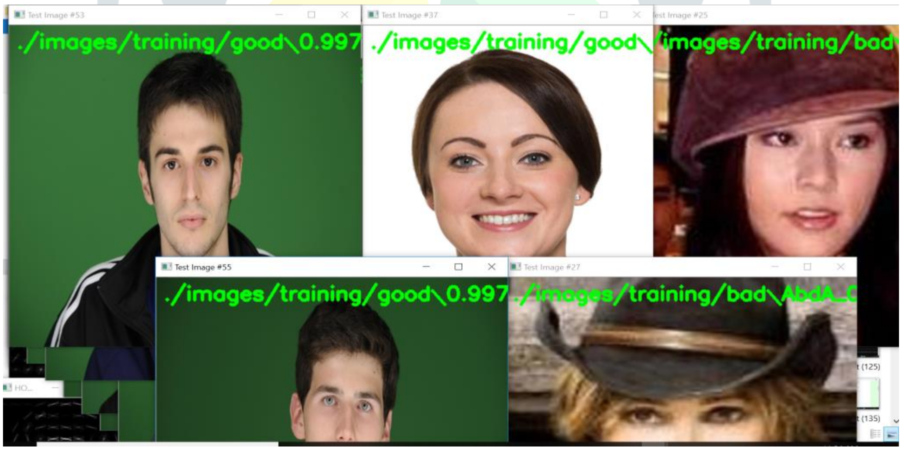

Fig.8 HOG Output: Good and Bad images

5.3 Neural Network

In this model, firstly, the path of training and validation data was specified. Keras, on the other hand facilitated easy prototyping and it's simplest model, Sequential was used. Then, three layers Convolutional/CNN model was set up followed by a fully connected layer. After this, RMSprop optimizer was instantiated by passing it to model.compile and trained the model and validated its output (i.e good or bad image) which is shown in Fig.9 and Fig.10.

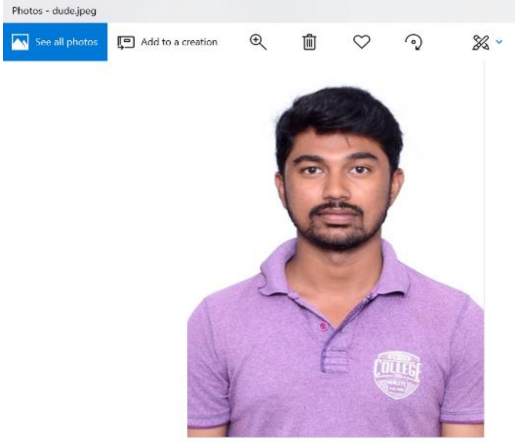

Fig.9 CNN output: good image

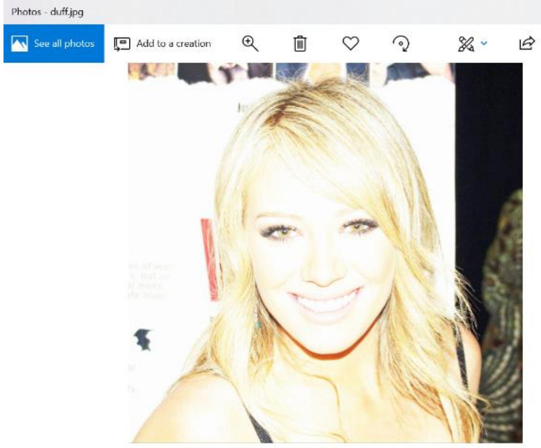

Fig.10 CNN output: bad image

5.4 Multi-classification

Multi-classification was performed by building 14 directories, with set of parameters given in Table 1. When a user uploads an image and checks for quality requirement of his/her facial image, the system generates a result informing whether the image is acceptable or rejected according to the ICAO standards. Furthermore, if the image is rejected, the system will provide valid reasons for the same. The user can then immediately rectify their errors and upload good quality images. For eg: If the user uploads an image which is blur, the system will instantly tell the user that these images are not acceptable and does not match the quality requirement. Multi-classification was performed using Keras, as it outperformed other models. The result of multi-classification is shown in Fig.11-14.

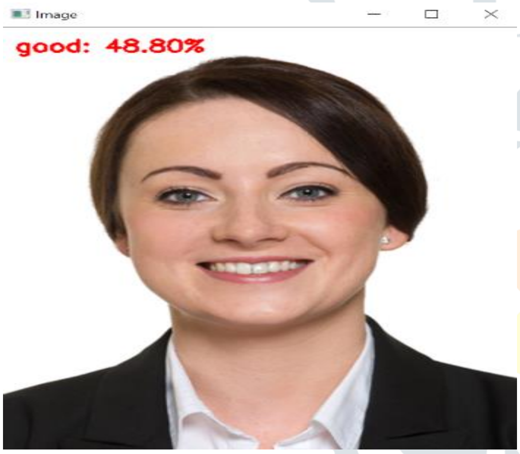

Fig.11 Good image

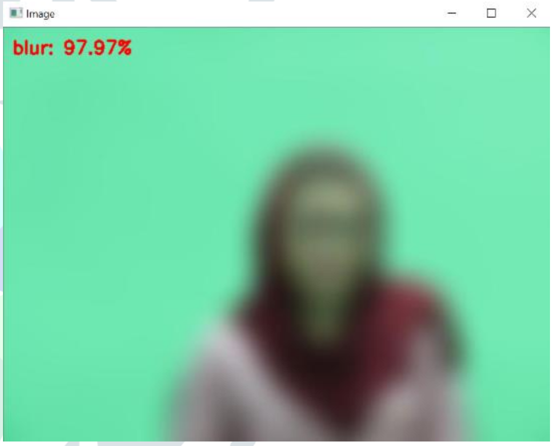

Fig.10 Bad image: Blur

Fig.13 Bad image: Non-frontal

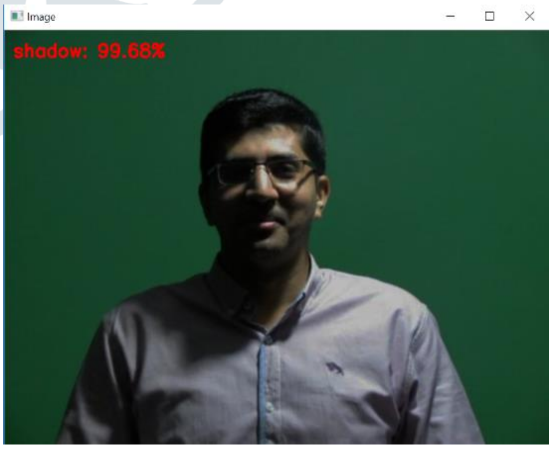

Fig.14 Bad image: Shadow

VI. CONCLUSION

The dataset classified into 14 parameters was used for multi-classification problem. This can further be used for the development of an application or a website which will predict whether a given image is suitable for a particular application, such as migration from one country to another, competitive exams etc. Experiments based on two classification problem were conducted using three different models (i.e. HOG and KNN, LBP and SVM, CNN). This suggests that HOG as a feature descriptor, is better than LBP, since it gives us an accuracy of 80 percent, as against 75 percent of LBP. Keras, however performed the best amongst all three, with an accuracy of 87 percent. Table II summarizes accuracies obtained by all three feature extraction models and classifiers.

Unlike LBP, CNN did not require normalization in two classification problem. Since Keras outperformed the other two models, therefore we proceeded with Keras to perform multi-classification.

TABLE II: SHOWS THE BEST ACCURACY VALUE FOR ALL THREE FEATURE EXTRACTION MODELS

Feature Extraction Models	**Accuracy Achieved (percent)**
LBP and SVM	75
HOG and KNN	80
CNN	85

VII. ACKNOWLEDGMENT

It gives us immense pleasure to express our deep gratitude and sincere thanks to Dr.Shikha Nema, H.O.D., and all the faculty members from Department of Electronics and Communication, Usha Mittal Institute of Technology, Mumbai for their valuable and useful support and comments for making this project successful. We'll not forget to mention that their approach kept our working environment alive and their encouragement promoted us to do our task rigorously.

REFERENCES

[1] Nasrollahi, Kamal, and Thomas B. Moeslund. ”Face quality assessment system in video sequences.”European Workshop on Biometrics and Identity Management. Springer, Berlin, Heidelberg, 2008.

[2] R. L. Parente, L. V. Batista, I. L. P. Andrezza, E. V. C. L. Borges, and R. A. T. Mota, ”Assessing Facial Image Accordance to ISO/ICAO Requirements,” 2016 29th SIBGRAPI Conference on Graphics, Patterns and Images (SIBGRAPI), 2016.

[3] Epic.org. (2009).”Facial Image Quality Improvement and Face Recognition Study Final Report” [online] Available at: https://epic.org/privacy/biometrics/10-0795

[4] Egovstandards.gov.in. (2010). ”Face Image Data Standard for e-Governance Applications in India”. [online] Available at: http://egovstandards.gov.in/sites/default/files/FaceImageDataStandardVer 1.0.pdf

[5] Castillo, Oriana Yuridia Gonzalez. ”Survey about facial image quality.”Fraunhofer Institute for Computer Graphics Research(2005):10-15.

[6] Ferrara, Matteo, Annalisa Franco, and Davide Maltoni. ”Evaluating systems assessing face-image compliance with ICAO/ISO standards.”European Workshop on Biometrics and Identity Management. Springer, Berlin, Heidelberg, 2008.

[7] Best-Rowden, Lacey, and Anil K. Jain. ”Automatic Face Image Quality Prediction.”arXiv preprint arXiv:1706.09887(2017).

[8] Lawrence, Steve, et al. ”Face recognition: A convolutional neural network approach.”IEEE transactions on neural networks8.1 (1997): 98-113.

[9] Sof dataset Mahmoud Afifi and Abdelrahman Abdelhamed, ”AFIF4: Deep gender classification based on an AdaBoost-based fusion of isolated facial features and foggy faces”. arXiv:1706.04277, arXiv 2017

[10] Tiago F. and Bottino, Andrea and Laurentini, Aldo and De Simone, Matteo, ” Detecting siblings in image pairs ” ,Journal= ”The Visual Computer”, 1st Dec 2014, vol.30, pg=1333-1345

[11] The Keras Blog “Building powerful image classification models using very little data” , available at https://blog.keras.io/building-powerfulimage- classification-models-using-very-little-data.html

[12] Psychological Image Collection at Stirling (PICS).2D face sets. [ONLINE] Available at: http://pics.psych.stir.ac.uk/ .

[13] Tiwary, K. (2019).Histograms Of Oriented Gradients for Human Detection, N. Dalal and B. Triggs. [online] Medium. Available at: https://medium.com/@ktiwary2/scattered-thoughts-on-ml-68d30f44da19

[14] Alhindi, Taha J., et al. ”Comparing LBP, HOG and Deep Features for Classification of Histopathology Images”. 2018 International Joint Conference on Neural Networks (IJCNN), 2018, doi:10.1109/ijcnn.2018.8489329.

[15] “Histogram of Oriented Gradients,” Histogram of Oriented Gradients - skimage v0.15.dev0 docs. [Online]. Available: http://scikitimage. org/docs/dev/autoexamples/featuresdetection/plothog.html.

[16] “Local Binary Patterns with Python and OpenCV,” PyImageSearch, 21-Jun-2018. [Online]. Available: https://www.pyimagesearch.com/2015/12/07/local-binary-patternswith- python-opencv/ .

[17] G. B. Huang, M. Ramesh, T. Berg, and E. Learned-Miller, ”Labeled faces in the wild: A database for studying face recognition in unconstrained environments,” Univ. of Mass., Amherst, Tech. Report 07-49, Oct. 2007.