

# Performance Analysis of Face Detection system using HOG and QualHOG Features

Ajay. N. Paithane, Ujwala. G. Patil, Bhagavat. D. Jadhav, Suresh. D. Shirbahadurkar



**Abstract:** Inspired by the expansion of minimal effort advanced cameras in cell phones being conveyed in computerized systems, we think about the connection between perceptual picture quality and an exemplary PC vision errand of face recognition. We measure the corruption in execution of a well known and compelling face detector when human-saw image quality is corrupted by twists usually happening in catch, stockpiling, and transmission of facial pictures, including clamor, obscure, and pressure. It is observed that, inside a certain scope of picture quality, an unobtrusive increment in picture quality can radically enhance face recognition execution. These outcomes can be utilized to guide asset or transfer speed distribution in securing or correspondence/conveyance frameworks that are connected with face location undertakings. In this work a perceptual quality QualHOG feature is used. Face locators prepared on these new components give measurably huge change in resilience to picture bends over a solid gauge. Distortion dependent which is more distorted uninformed variations of the face indicators are proposed and assessed on a huge database of face pictures speaking to an extensive variety of mutilations. A one-sided variation of the preparing calculation is additionally recommended that further improves the power of these face locators. To encourage this exploration, we have developed another dataset in our lab for further study.

**Keywords:** QualHOG, NSS, DFD, NIQ, SVM.

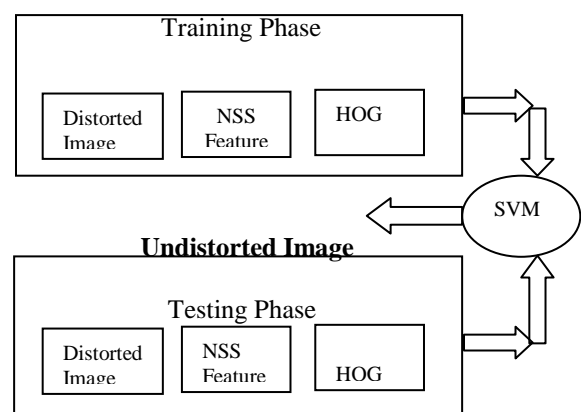
## I. INTRODUCTION

Now day a system with object detection or recognition becomes an essential part of PC based system. The programmed face recognition system is one of the best business ideas in this advanced technical environment. However, the performance of such system degrades when quality of input image is taken into consideration. In recent advances, the gadget likes android mobiles, iphones, tabs and many more inexpensive mobile has face detection system for surveillance is often not restricted and hence there are problem of quality that will surely affect the performance of

image as well as the performance of the system. The essential part of every advanced system like facial expression recognizer or tracker is face detection, best system for face detection improves the quality of the system which is very important in any surveillance system. Currently many researchers are working on No reference- image quality algorithms (NR-IQA) problem, the NR-IQA problem approach is distortion-independent [1], [4], [6], [7]. The models used for solving NR-IQA problem by contaminating modified the distortion in the captured image [2], [3], [17]. The small change in the natural scene statistics (NSS) affects the perceptual quality of the captured image [4]. Hence, to develop distortion free model the natural image quantification is very much helpful.

In this paper, we demonstrate the utilization of effortlessly process able "quality-aware" spatial Natural Scene Statistics (NSS) features that help in the face identification [5]. Similarly the widely-used Histogram of Oriented Gradients (HOG) based recognition calculation is utilized for further study. During this study the Distorted Face Database (DFD) has been developed at our own lab. The developed database is then synthesized which is contaminated with different noise levels.

## II. PROPOSED SYSTEM ARCHITECTURE



**Fig.1 Proposed system block diagram**

The Fig.1 shows a proposed system which consists of two main phases, one is training phase and another testing phase. The processes available in both the phases are same, however the training phase trains the system with predefine features required to recognize. Similarly, the testing phase extracts the input signal so as to obtain the specified features for recognition. The performance accuracy of the system mainly the prime concern while extracting the features those are used for training and testing the query images for recognition.

Revised Manuscript Received on October 30, 2019.

\* Correspondence Author

**Ajay N Paithane\***, Department of Electronics & Telecommunication Engineering, JSPMs Rajarshi Shahu College of Engineering, Pune, India.

**Ujwala G Patil**, Department of Electronics & Telecommunication Engineering, JSPMs Rajarshi Shahu College of Engineering, Pune, India.

**Bhagavat D Jadhav** Department of Electronics & Telecommunication Engineering, JSPMs Rajarshi Shahu College of Engineering, Pune, India.

**Suresh D Shirbahadurkar** Department of Electronics & Telecommunication Engineering, Zeal College of Engineering, Pune, India.

© The Authors. Published by Blue Eyes Intelligence Engineering and Sciences Publication (BEIESP). This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>)

In the proposed system, though the distortion present in the query image is global that query images with face and non-face patches are not used during experimentation. In this proposed system we have developed a dataset prone to distortion from the various images collected from different sources. In database development process, the frontal faces are mainly considered. Hence, the noises at different levels are taken into consideration like, white gaussian noise, gaussianblur, JPEG compressed images [8]. The blind IQA which has natural scene statistics of spatial components are obtained from the spatial-NSS features [9], [10], [11]. The divisive normalization and local mean removal is used to preprocess the input image, the mathematical operation for preprocess is:

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

The preprocessed image is then used to extract HOG features, the complete process to extract HOG features is elaborated here to understand the actually work done. To simplify the process it is observed during experimentation that, a detection window must be divided into different blocks like 2×2, 4×4, 8×8 and 16×16 etc., to ensure better performance. In the feature extraction process a stride of 8 × 8 pixels is used to decompose a dense overlapping block size of 16 × 16. The first feature HOD is used to obtain HOG feature descriptor by concatenating all histograms within a patch. The second important feature used in this study is QualHOG. The supervised classier Linear Support Vector Machines (SVMs) is trained using, this HOG and QualHOG features.

**Database**

The dataset used in this work has been developed in the research lab, also some facial images are used from the available resources. In the proposed work AWGN Additive White Gaussian Noise is adapted for developing dataset. This is a standard noise, where each pixel has been contaminated with Gaussian noise. During the contamination process variance parameter has been added to develop distorted images. The mathematical representation of distorted image is:

$$\tilde{I}(i, j) = I(i, j) + N_{ij}, \quad N_{ij} \approx N(0, \sigma_N^2)$$

$N(0, \sigma_N^2)$  Gaussian distribution,  $\sigma_N^2$  variance

Similarly, the Gblur and Gaussian Blur is one another type of distortion where, each pixel is contaminated using the concept of convolution technique with low pass filter.

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

Where  $-[3\sigma_B] \leq x \leq [3\sigma_B]$  and  $-[3\sigma_B] \leq y \leq [3\sigma_B]$

$$\tilde{I}(i, j) = \sum_{x=-[3\sigma_B]}^{[3\sigma_B]} \sum_{y=-[3\sigma_B]}^{[3\sigma_B]} I(i+x, j+y) * G(x, y)$$

**III. EXPERIMENTATION**

As previously mentioned the database used in this study is readily available on the internet also the author have developed this Distorted Face dataset at the Laboratory.

During experimentation the input images used in this study are from the training dataset. As a basic requirement of this study is only frontal human faces so we have taken all that into consideration, these frontal images are then segmented into face and non face patches after adding appropriate noise levels. In this proposed work 215 images are used, out of which 150 images are used for training and 65 images are used for testing. The positive samples of dataset are obtained manually using 1231 annotated faces. Similarly, negative patches are obtained from 1500 set of images. The proposed algorithm trains the system for 80\*60 patches, hence the detected output has patch size of 80 × 64. from each of the above sets of training images. The Soft-margin linear support vector machine (SVM) is used as a classifier, the SVM is trained in this work using QualHOG features that are extracted from the positive and negative sample of the training datasets[12], [13]. As mentioned previously, the testing is done on the remaining 65 images. The performance if the system is measured on the basis of true positive and true negative samples detected in the system. The performance of the system has been increased due to the hybrid feature technique is used in this research.

**IV. RESULTS AND DISCUSSION**

The performance analysis of human face detection using Perceptual Quality aware HOG and QualHOG features is evaluated in this work. The white Gaussian noise at 10 dB level is used to develop distorted image. The Fig. 2 shows the input image and processed images depicted in Fig. 2(a) to Fig. 2(f)

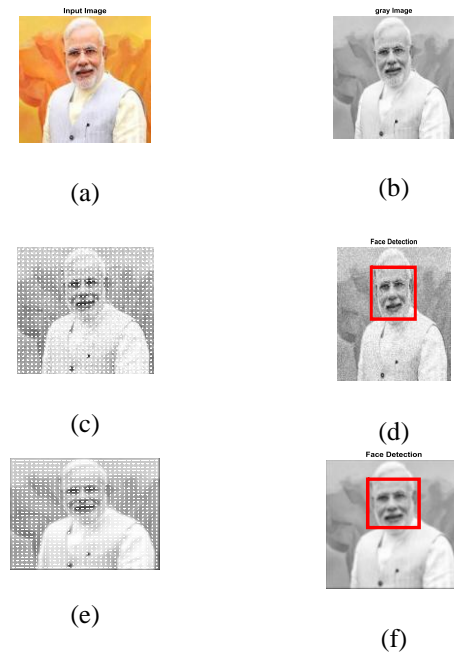
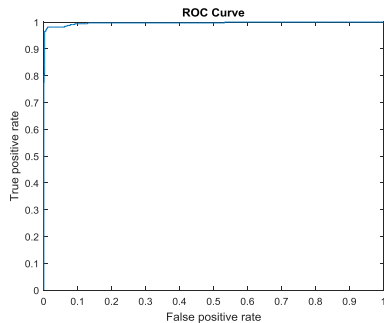


Fig.2 (a)input image(b)gray image(c)feature extraction(d)Face detection(e)feature extraction for blur image(f)face detection for blur image

**QualHOG vs. HOG**

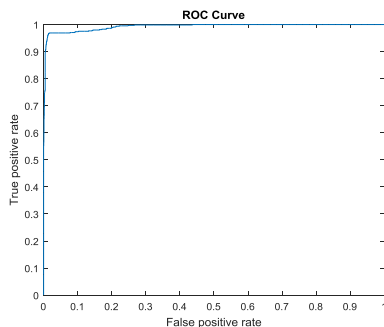
The face locator tests are developed here during experimentation for various distortions. In this experimentation the various identifiers are assessed for different test identifiers.

It took execution of QualHOG and HOG face identifier at two distorted type one added substance White Gaussian Noise and another Gaussian Blur noise (these are universally distributed). After execution a ROC is used for checking the performance of face detector [14], [15]. This collector working bend is graphical plot of the output of LSVM, it relies on upon execution of classifier. Fig.3 shows ROC curve of false positive and true positive rate of the system with QualHOG features. The bend is made by plotting the genuine positive rate against the false positive at different limit settings. A ROC space is characterized by FTR and TPR as x and y axis separately. The mutilated picture is applied to QualHOG identifier and the additives white Gaussian noise is used to contaminate the image during first execution.



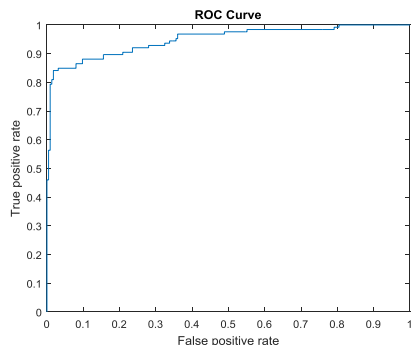
**Fig.3 QualHOG for additive white Gaussian noise**

Similarly, Fig. 4 shows graphical representation of performance analysis of the system for QualHOG feature with Gaussian blur noise.

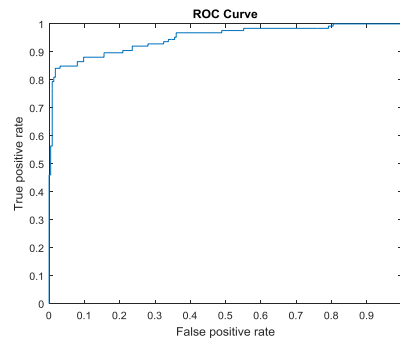


**Fig.4 QualHOG for Gaussian blur**

Fig. 5 shows graphical representation of performance analysis of the face detector system for HOG feature with additive white Gaussian noise.



**Fig.5 HOG output for additive white Gaussian noise**



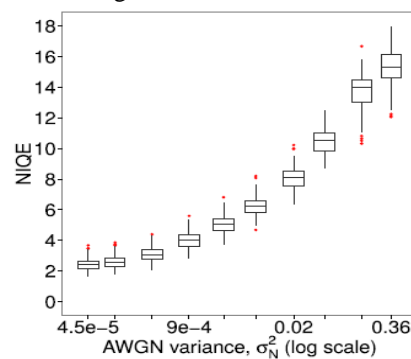
**Fig.6 HOG output for Gaussian blur**

According to Fig. 6, the precision rate of face identifier increases with respect to the expansion of the system

**TABLE I  
Performance of HOG and QUAL HOG  
On distorted types**

V. Distorted	VI. HOG	VII. QualHOG
VIII. AWGN	IX. 74.4467%	X. 99.2623%
XI. GBlur	XII. 74.7373%	XIII. 99.2623%

The experimental results are as shown in table I, where the distorted images are developed using an AWGN and Gaussian Blur. The proposed algorithm has been used to extract features like HOG and QualHOG. After vast experimentation it has been concluded that the performance of the system has been increased due to the QualHOG features set as shown in table 1. In this work, the NIQE scores are computed using the spatial-NSS features those are simple, suitable and not costing a great deal, where as NR quality scores are more expensive compared to the NIQE scores. In this proposed work during experimentation and literature survey the constructive correlation is observed between NIQE and degree of distortion also very less expectation against the distortion and monotonicity in any type of quality model is needed in this work. Fig.7 presented here gives information about the NIQE scores and the levels of AWGN, which gives knowledge about the distortion levels available in the images.



**Fig.7. NIQE vs AWGN**

Similarly, Fig. 8 is the box plot of NIQE and Gblur, which resemble about the level of blur in inputted images from this plot it is infer that the NIQE for Gblur level greater than 15 scales are nearly constant.

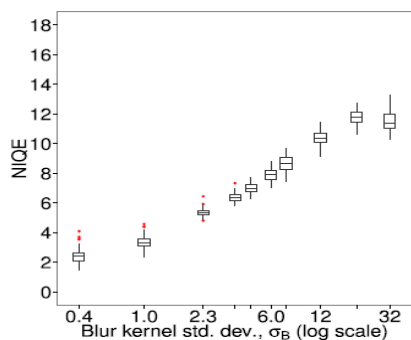


Fig. 8 NIQE Vs GBlur

Fig. 9 gives an idea about the level of JPEG compression against the NIQE scores of distorted images. From the box plot of NIQE and JPEG compression it is observed that the NIQE levels get minimized by changing the levels of JPEG compression.

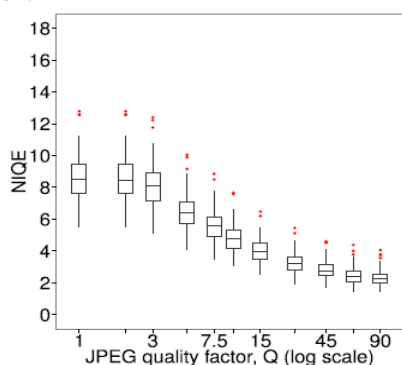


Fig.9. NIQE Vs JPEG

The results shown here are the relation between AUPR and AWGN for HOG-Prist, HOG-AWGN, QualHOG-Prist and QualHOG-AWGN. In this work, we have trained the system with distortion prone QualHOG and HOG with the required distortion levels. Here we have selected a universal face detector for the detection of HOG and QualHOG based system [18], [19]. The results represented in Fig. 10 are the various distortion levels for HOG-Prist, HOG-AWGN, QualHOG-Prist and QualHOG-AWGN.

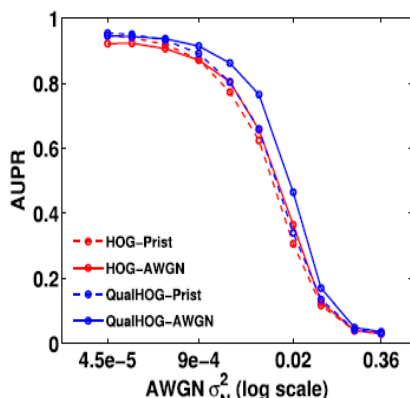


Fig.10. AUPR vs AWGN

The results obtained in this research shown in Fig. 11 are such that, the quality of the system has been drastically improves due to the AWGN QualHOG compared to the other face detectors and the detectors designed using HOG.

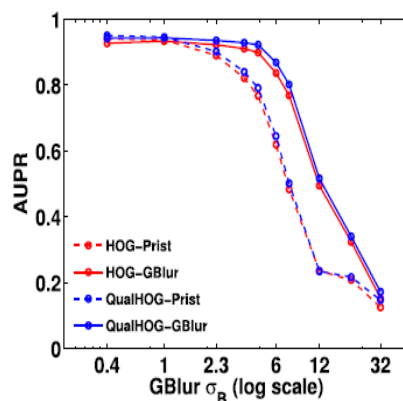


Fig.11AUPR Vs GBlur

The actual improvement of the system measured in terms of quality of distortion is marginal for Gaussian blur. The results compared in the Fig. 6 are based on the proposed features like HOG and QualHOG and the technique adopted for detection is based on the JPEG compression distortions.

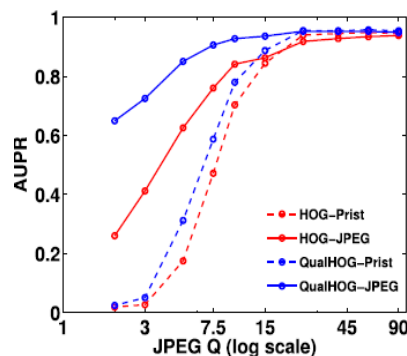


Fig.12 AUPR vs JPEG

The face detector system using newly adopted technique like, JPEG compression QualHOG gives better performance in quality point of view as well as in the robustness of the system as compared to other HOG features based face detector system[20], [21].

#### XIV. CONCLUSION

In this paper we have studied a novel method for trade-off evaluation between face detection performance based on the three common distortion types viz, AWGN, Gaussian blur, JPEG. Similarly we conclude that, the NR image quality score and NIQE are more effective for evaluating actual distortion levels. During experimentation it is observed that, in the HOG-based face detector the NIQE score gets degrade by more than 4 levels. The face detection performance drastically improved due to a modest improvement in image quality measures.

During this study it is observed that, the hybrid of HOG features and spatial NSS features is QualHOG features are more suitable at learning phase of face detector and robust for image distortion.

## REFERENCES

1. P. Ye and D. Doermann, "No-reference image quality assessment using visual codebooks," *IEEE Trans. Image Process.*, vol. 21, no. 7, pp. 3129–3138, Jul. 2012.
2. S. A. Karunasekera and N. G. Kingsbury, "A distortion measure for image artifacts based on human visual sensitivity," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process. (ICASSP)*, Apr. 1994, pp. V/117–V/120.
3. Z. Wang, E. P. Simoncelli, and A. C. Bovik, "Multiscale structural similarity for image quality assessment," in *Proc. Conf. Rec. 37<sup>th</sup> Asilomar Conf. Signals, Syst. Comput.*, Nov. 2003, pp. 1398–1402.
4. A. K. Moorthy and A. C. Bovik, "Blind image quality assessment: From natural scene statistics to perceptual quality," *IEEE Trans. Image Process.*, vol. 20, no. 12, pp. 3350–3364, Dec. 2011.
5. M. A. Saad, A. C. Bovik, and C. Charrier, "Blind image quality assessment: A natural scene statistics approach in the DCT domain," *IEEE Trans. Image Process.*, vol. 21, no. 8, pp. 3339–3352 Aug. 2012.
6. A. Mittal, A. K. Moorthy, and A. C. Bovik, "No-reference image quality assessment in the spatial domain," *IEEE Trans. Image Process.*, vol. 21, no. 12, pp. 4695–4708, Dec. 2012.
7. P. Campisi, M. Carli, G. Giunta, and A. Neri, "Blind quality assessment system for multimedia communications using tracing watermarking," *IEEE Trans. Signal Process.*, vol. 51, no. 4, pp. 996–1002, Apr. 2003.
8. Q. Li and Z. Wang, "Reduced-reference image quality assessment Using divisive normalization-based image representation," *IEEE J. Sel. Topics Signal Process.*, vol. 3 no. 2, pp. 202–211, Apr. 2009.
9. A. Mittal, R. Soundararajan, and A. C. Bovik, "Making a 'completely blind' image quality analyzer," *IEEE Signal Process. Lett.*, vol. 20 no. 3, pp. 209–212, Mar. 2013.
11. M. Abdel-Mottaleb and M. H. Mahoor, "Application notes—Algorithms for assessing the quality of facial images," *IEEE Comput. Intell. Mag.* vol. 2, no. 2, pp. 10–17, May 2007.
12. R.-L. V. Hsu, J. Shah, and B. Martin, "Quality assessment of facial images," in *Proc. Biometrics Symp., Special Session Res. Biometric Consortium Conf.*, Sep./Aug. 2006, pp. 1–6.
13. D. M. Rouse and S. S. Hemami, "Analyzing the role of visual structure in the recognition of natural image content with multi-scale SSIM," *Proc. SPIE*, vol. 6806, p. 680615, Feb. 2008.
14. N. Dalal and B. Triggs, "Histograms of oriented gradients for human detection," in *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2005, pp. 886–893.
15. R. V. Babu, S. Suresh, and A. Perki, "No-reference JPEG-image quality assessment using GAP-RBF," *Signal Process.*, vol. 87, no. 6, pp. 1493–1503, Jun. 2007.
16. M.-H. Yang, D. Kriegman, and N. Ahuja, "Detecting faces in images: A survey," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 24, no. 1, pp. 34–58, Jan. 2002.
17. Paithane, A. N., and D. S. Bormane. "Analysis of nonlinear and non-stationary signal to extract the features using Hilbert Huang transform." *Computational Intelligence and Computing Research (ICCIC), 2014 IEEE International Conference on.* IEEE, 2014.
18. Paithane, A. N., and D. S. Bormane. "Electrocardiogram signal analysis using empirical modedecomposition and Hilbert spectrum." *Pervasive Computing (ICPC), 2015 International Conference on.* IEEE, 2015.

## AUTHORS PROFILE



**Ajay N Paithane** He received the B.E. degree in Electronics engineering from Dr. B. A. M. University of Aurangabad, India, in 1996, and the M.E. from Shivaji University of Kolhapur, India, in 2008, and Ph.D. degrees in Electronics and Telecommunication engineering from S.P.P.U Pune, India, in 2016.

Since July 2015, he has been with the Department of Electronics and Telecommunication engineering, RSCOE, Pune as a Professor. His current research interests include Signal, Image processing and Embedded system design. He is a Life Member

of the Indian Society for Technical Education (ISTE). He has filed six patents and one copyright. He has published more than thirty papers in various IEEE conferences and reputed journals.



**Ms. Ujwala G Patil** She received the B.E. degree in Electronics engineering from Shivaji University of Kolhapur, India, in 1998, and the M.E. from Shivaji University of Kolhapur, India, in 2009, and Pursuing Ph.D. degrees in Electronics and Telecommunication engineering from S.P.P.U Pune, India.

Since July 2006, she has been with the Department of Electronics and Telecommunication engineering, RSCOE, Pune as a Assistant Professor. Her current research interests include Speech, Signal and Image processing. She is a Life Member of the Indian Society for Technical Education (ISTE). She has filed one patent and has published more than ten papers in various IEEE conferences and reputed journals.



**Bhagavat D Jadhav** He received the B.E. degree in Electronics engineering from Pune University, Pune India, in 2002, and the M.Tech. from SPPU, Pune India, in 2008, and Ph.D. degrees in Electronics and Telecommunication engineering from S.P.P.U Pune, India in 2017.

Since July 2015, he has been with the Department of Electronics and Telecommunication engineering, RSCOE, Pune as a Professor. His current research interests include Signal and Image processing. He is a Life Member of the Indian Society for Technical Education (ISTE). He has filed two patent and has published more than twenty papers in various IEEE conferences and reputed journals.



**Suresh D Shirbahadurkar** He received the B.E. degree in Electronics engineering from Dr. B. A. M. University of Aurangabad, India, in 1992, and the M.E. from Dr. B. A. M. University of Aurangabad, India and Ph.D. degrees in Electronics and Telecommunication engineering from Dr. B. A. M. University of Aurangabad, India.

Since July 2015, he has been with the Department of Electronics and Telecommunication engineering, Zeal College of Engineering, Pune as a Professor. His current research interests include Speech, Signal and Image processing. She is a Life Member of the Indian Society for Technical Education (ISTE). She has filed one patent and has published more than 50 papers in various IEEE conferences and reputed journals.