IJCSE International Journal of Computer Sciences and Engineering Open Access
Research Paper Vol.-6, Issue-9, Sept. 2018 E-ISSN: 2347-2693

Image Training and LBPH Based Algorithm for Face Tracking in Different Background Video Sequence

Ranganatha S[1*], Y P Gowramma[2]

[1]Dept. of Computer Science and Engineering, Government Engineering College, Hassan, Karnataka, India
[2]Dept. of Computer Science and Engineering, Kalpataru Institute of Technology, Tiptur, Karnataka, India

[*]*Corresponding Author: ranganath38@yahoo.co.in*

Available online at: www.ijcseonline.org

Accepted: 21/Sept/2018, Published: 30/Sept/2018

***Abstract*—** Video and video types are changing day by day; due to which, video processing is becoming complex time to time. There is a lack of particular algorithm for automatic detection and tracking of human faces in video, to overcome the challenges that are being faced nowadays. This paper describes a model for detection and tracking of human faces in different background video sequence using OpenCV platform. Both positive and negative image samples are trained and saved as xml file. With the help of trained samples, LBPH algorithm clarifies whether the video frame contain faces or not. Further, HOG descriptor is fed to SVM detector to compute the coefficients that are stored in the xml file. Based on this, face regions are tracked until the last frame is reached. We have tested our proposed algorithm on the videos of a technically challenging dataset. Standard metrics helped to judge the success of the proposed algorithm. Test results indicate the superiority of our proposed model, compared to other similar algorithms.

***Keywords*—** Detection, Tracking of human faces, Different background, Video sequence, OpenCV, LBPH, HOG, SVM.

I. INTRODUCTION

Video processing is an interesting research zone in image processing. Face detection and tracking are part of video processing, where the face regions need to be detected and tracked. Due to the tremendous growth in video technology, the algorithms which worked well few years back, are now lacking behind to overcome the current challenges.

There are three phases in any face tracking model. At first, face regions are detected in the video. This phase is challenging [1], because face regions are detected under various interferences like occlusion, illumination, different background conditions, variations in pose and expression etc. to name a few. Several approaches are available for detection. We can either adopt existing face detection techniques or develop our own methods to serve this purpose. Viola-Jones [2, 3] approach is an existing face detection technique. It detects only those face regions that are posed towards the camera. For detecting the faces with other poses, it requires additional training. In our proposed algorithm, we have trained the image samples; using which, LBPH (Local Binary Patterns Histograms) [4] algorithm detect the face regions successfully.

During the second phase, features are selected and extracted from the region(s) of interest (ROI). Since the features selected directly trouble the computational complexities (i.e. time and space), only few features have to be selected cleverly. Various features are available to select from the ROI. Color, texture, and intensity are grouped as primary; while secondary features e.g. HOG (Histograms of Oriented Gradients) [5], and Haar features [3] are formed using primary features. Our proposed algorithm feeds the HOG descriptor to the SVM (Support Vector Machine) [6] detector to compute the coefficients as a model.

At the end, using the features mined during the second phase, third phase track the face regions in the video until the last frame. In order to accelerate the course of tracking, we have implemented our proposed model in OpenCV [7, 8] platform. It houses open source library software; using which, computer vision and machine learning applications can be implemented. Our proposed algorithm can track the faces, which belong to the following three different background video sequence categories: 1) Moving faces, but unmoving camera, 2) Unmoving faces, but moving camera, and 3) Moving faces, with moving camera.

Rest of the paper is structured as stated next. Section II cover the related work. Section III describes the methodology of our proposed model. Results and discussion are part of section IV, and section V concludes the proposed research work with future directions.

II. RELATED WORK

The works of Fukunaga et al. [9], and Cheng [10] influenced Bradski [11] to develop an algorithm called CAMSHIFT. It is one of the robust chromatic tracking approach developed till date. But, since it is based on color, non-face color regions which appear like skin regions, are also tracked. KLT [12-14] is a feature tracker, which is based on the works of Lucas and Kanade [12]; and later extended by kanade and Tomasi [13], and Tomasi and Shi [14]. Here, the human faces are tracked, based on the displacement of points in the frames. It is more suitable for tracking the faces in outdoor videos. But, KLT misses the small search windows while tracking.

Ranganatha S et al. [15] have implemented a fused method for face tracking by merging centroid of corner points [16] along with KLT. Their merged algorithm worked better, compared to KLT solely. They have implemented another algorithm [17] by combining CAMSHIFT with Kalman filter [18]. The drawbacks of CAMSHIFT were overcome after integrating it with Kalman filter. The two algorithms discussed here, used the videos of Celebrities [19] dataset, for experimentation. But, both the algorithms are capable of tracking only the single face in the video sequences which belong to moving faces, but unmoving camera category.

Ranganatha S et al. [20, 22, 23] have developed three algorithms for tracking the faces in all the three categories of videos being discussed in the introduction section of this paper. The first algorithm [20] was based on BRISK [21] features, the second algorithm [22] adopted Eigen [14] and HOG features; and the third algorithm [23] used trained images, corner and FAST [24] features for tracking the faces in videos. But, may be because of the platform was not OpenCV, the algorithms took more time for some of the video sequences. The three algorithms discussed here, used the videos of Celebrities [19], Choke Point [25], HOHA [26], and VidTIMIT [27] datasets, for experimentation.

Ranganatha S et al. [28] have proposed a method for selected single face tracking in technically challenging videos, using the combined features. But, multiple faces cannot be selected for tracking. To overcome this drawback, the authors then developed a color based selective face tracking approach [29], in which the users are permitted to specify the number (i.e. one, two, three etc.) of faces to be tracked. The two algorithms discussed here, used the videos of Celebrities [19], and VidTIMIT [27] datasets, for experimentation.

The literature being discussed in this section helped in finalizing the topic of this paper, its design, and implementation. The dataset which we have used, is part of the work in paper [19]. To estimate the performance of the proposed model, we have chosen the metrics, which are part of the works in literature [30-32].

III. METHODOLOGY

In this section, we discuss the design and implementation of our proposed model in three sub-sections as follows.

A. Image Training

Training helps to detect the faces in the given input file. It is performed in three steps: 1) Data collection, 2) Preparation, and 3) Model training. OpenCV_Traincascade tool was used to train the face (positive) and non-face (negative) image samples. The default Haar feature assisted to train the boosted cascade classifier. The positive samples were trained manually, while the negative samples were trained by specifying the path of their residence. We have trained 3457 positive images of two datasets, and 1584 negative images of three datasets as mentioned in Table 1 and Table 2 respectively.

Table 1. Statistics of face image samples trained.

Serial No.	Dataset	Number of images
01.	Caltech [33]	2931
02.	Taiwan [34]	526

Table 2. Statistics of non-face image samples trained.

Serial No.	Dataset	Number of images
01.	Caltech [33]	35
02.	101_ObjectCategories [35, 36]	112
03.	256_ObjectCategories [37]	1437

We come across several options while training the cascade; they are: 1) Pressing `c' will confirm the annotation, 2) Pressing `d' will delete the annotation, and 3) Pressing `n' will move to the next image in the folder. After training both positive and negative samples, an xml file is generated, which will be used for face detection purpose later.

B. LBPH Algorithm

LBPH algorithm is used for detecting the faces in the video sequence. Local Binary Patterns (LBP) is a kind of visual descriptor that is used for classification in computer vision. It includes three stages, they are: 1) Face representation, 2) Feature extraction, and 3) Classification.

LBP works with the 8 neighbours of a pixel, by means of the rate of the middle pixel as a threshold. If a neighbour pixel has an upper grey rate than the middle pixel (or the identical grey rate), then a 1 is given to that pixel, else it catches a 0. The LBP watchword for the center element is then formed by joining the 8 ones/zeros to a binary code. Figure 1 illustrates the operation of LBP clearly.

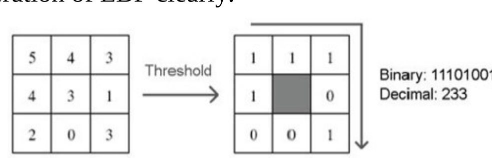

Figure 1. LBP operation.

C. Working of Proposed Algorithm

An assumption is made that, in most of the cases, the person is facing the camera, so that his/her face is successfully detected by the algorithm. Then with the help of trained datasets, LBPH detect the faces in the video sequence; based on which, tracking takes place further. Figure 2 shows the architecture of the proposed system.

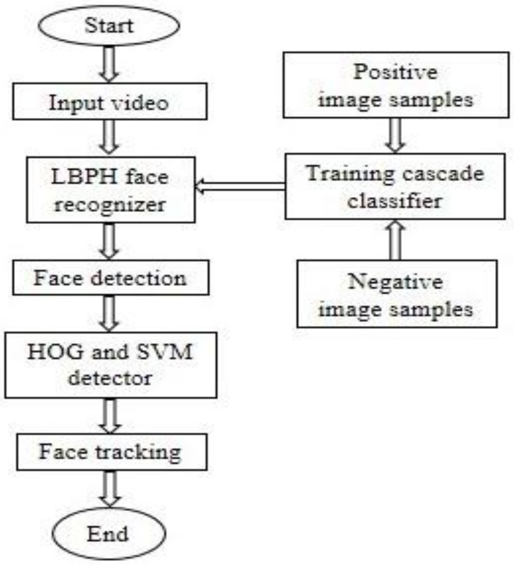

Figure 2. Proposed system architecture.

Face detection procedure is illustrated in the following algorithm in detail.

Algorithm: Face Detection

Input: Video sequence.

Output: Bounding box around the detected face regions.

Steps:

1. Positive and negative image samples are trained using OpenCV_Traincascade tool with the default Haar feature.
2. Load the video file.
3. Check whether the video file is read successfully or not.
4. Allow LBPH to recognize the face regions in the video file loaded.
5. If the face regions are present in the frame, then LBPH perform its work with the help of trained positive image samples.
6. Calculate the height and width of the face regions that are detected.
7. Finally, LBPH draws the bounding boxes around the face regions.

Face tracking procedure is illustrated in the following algorithm in detail.

Algorithm: Face Tracking

Input: Video sequence.

Output: Tracking of the face regions present in the video sequence.

Steps:

1. Load the video file.
2. Convert the colored images into greyscale.
3. Feed the HOG descriptor to the SVM detector, in order to compute the coefficients that are stored in the xml file as a model or classifier.
4. Load all the frames of the video sequence as input to the algorithm.
5. For each frame, starting from the current frame to the last frame, repeat step 6.
6. Call the LBPH algorithm to detect the face regions in the current frame.
7. Move on to the next frame, and go to step 6.

IV. Results and Discussion

We have tested our proposed algorithm on 13 video sequences. The proposed algorithm is compared with 4 algorithms, they are: 1) Algorithm1 [15], 2) Algorithm2 [17], 3) Algorithm3 [20], and 4) Algorithm4 [23]. The comparison of different algorithms is based on the metrics that are part of the works as mentioned in the related work. The important metrics are Recall (R), Precision (P), Accuracy (A), F1 Score (F1S), MOTA (multiple object tracking accuracy), and MOTP (multiple object tracking precision). The tracking results of some of the video sequences are as displayed below.

A. Moving Faces, but Unmoving Camera – Video Sequences

Here, we have included some of the frames of three video sequences as shown in Figure 3, Figure 4, and Figure 5.

Figure 3. Face tracking result of 48 frames low resolution single face video sequence 0044_03_001_adam_sandler.avi with frames #02, #15, #39, and #47 displayed (from left to right).

Figure 4. Face tracking result of 205 frames single face video sequence 0784_01_015_hillary_clinton.avi with frames #23, #03, #36, and #140 displayed (from left to right).

Figure 5. Face tracking result of 217 frames two faces video sequence 0240_02_006_anderson_cooper.avi with frames #05, #40, #80, and #178 displayed (from left to right).

In all the three above video sequences, camera is stationary; but, faces are in motion. We can observe that the videos pose challenges like variations in expression (‘laugh’ in Figure 3), various poses (e.g. frames of Figure 4), low resolution, poor illumination, face skin color like regions etc. But, the proposed algorithm perform its intended work in all the situations robustly.

B. Unmoving Faces, but Moving Camera – Video Sequence

Here, we have included some of the frames of a video sequence in which an actress is singing a song, and she is in the position of rest (i.e. static). But, camera moves around the actress. Hence, the video comes under the above stated video sequence category. The frames of Figure 6 clearly describes the theme of the discussion made till now. We can observe that, only half of the face is visible; but, even then the proposed algorithm perform its intended work perfectly.

Figure 6. Face tracking result of 69 frames single face video sequence 0673_01_004_gloria_estefan.avi with frames #04, #37, #49, and #58 displayed (from left to right).

C. Moving Faces, with Moving Camera – Video Sequence

Here, we have included a frame of multiple face video sequence in which both the faces and the camera are in motion, and some of the faces are merged. Figure 7 shows a frame in which five faces are being detected and tracked.

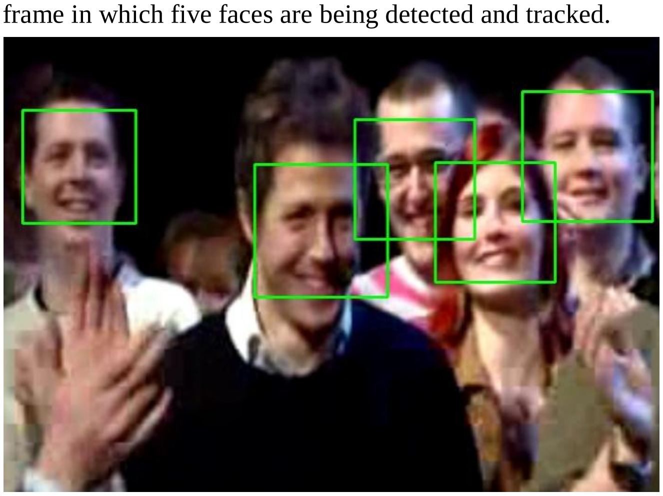

Figure 7. Multiple face tracking result of 25 frames low resolution video sequence 0792_01_001_hugh_grant.avi with frame #09 displayed.

Till now, the results of only five video sequences are discussed; but, the assessable set of results are tabulated by considering all the 13 video sequences as shown in Table 3.

Table 3. Assessable results of five algorithms.

Algorithm	**R** (%)	**P** (%)	**A** (%)	**F1S** (%)	**MOTA** (%)	**MOTP** (%)
Proposed	100	100	100	100	100	100
Algorithm1	69.23	69.23	69.23	69.23	69.23	69.23
Algorithm2	69.23	69.23	69.23	69.23	69.23	69.23
Algorithm3	84.61	84.61	84.61	84.61	84.61	84.61
Algorithm4	100	99.84	99.84	99.84	99.29	100

Figure 8 illustrates assessable results analysis of five algorithms graphically.

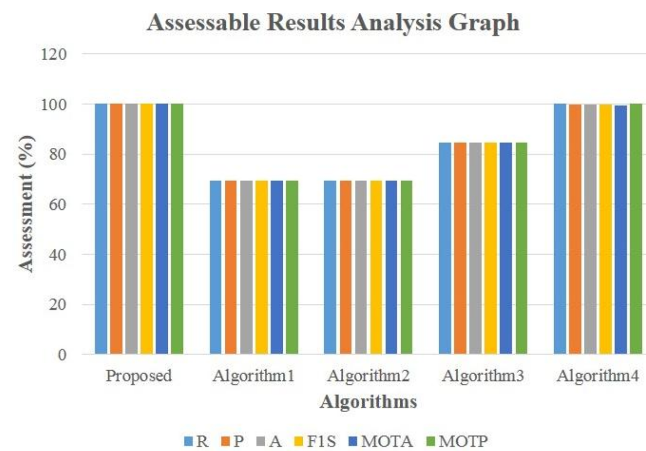

Figure 8. Assessable results analysis of five algorithms.

V. CONCLUSION AND FUTURE SCOPE

This paper fulfils our objective i.e. to detect and trace the faces in video sequence. We have used Haar like features to train the region of interest. LBPH algorithm helped to detect the face section in the specified input frame. The proposed algorithm showed better results for different performance metrics, when compared with other similar algorithms.

1. Since the described face detection and tracking approach delivers adequate results, upcoming work will involve into a more difficult biometric structure that uses improved detection techniques.
2. To further improve the abilities of the detection phase, more and more data-sets can be trained.
3. The occluded face regions can also be considered while training the data-sets for face detection and tracking.

REFERENCES

[1] Ranganatha S, Dr. Y P Gowramma, “Face Recognition Techniques: A Survey”, International Journal for Research in Applied Science and Engineering Technology, Vol.3, Issue.4, pp.630-635, 2015.

[2] P. Viola, M. Jones, "Robust Real-Time Face Detection", International Journal of Computer Vision, Vol.57, pp.137-154, 2004.

[3] P. Viola, M. Jones, "Rapid Object Detection Using a Boosted Cascade of Simple Features", in the Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR), Kauai, USA, pp.511-518, 2001.

[4] Md. Abdur Rahim, Md. Najmul Hossain, Tanzillah Wahid, Md. Shafiul Azam, "Face Recognition using Local Binary Patterns (LBP)", Global Journal of Computer Science and Technology, Vol.13, Issue.4, pp.1-8, 2013.

[5] N. Dalal, B. Triggs, "Histograms of Oriented Gradients for Human Detection", in the Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR), pp.886-893, 2005.

[6] Shantala Giraddi, Jagadeesh Pujari, Shivanand Seeri, "Identifying Abnormalities in the Retinal Images using SVM Classifiers", International Journal of Computer Applications, Vol.111, Issue.6, pp.5-8, 2015.

[7] G. Bradski, "The OpenCV Library", Dr. Dobbs Journal of Software Tools, 2000.

[8] G. Bradski, A.Kaebler, "Learning OpenCV", China: Southeast Univ. Press, 2009.

[9] K. Fukunaga, L. D. Hostetler, "The Estimation of the Gradient of a Density Function, with Applications in Pattern Recognition", IEEE Transactions on Information Theory, Vol.21, Issue.1, pp.32-40, 1975.

[10] Yizong Cheng, "Mean Shift, Mode Seeking, and Clustering", IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol.17, Issue.8, pp.790-799, 1995.

[11] G. Bradski, "Computer Vision Face Tracking for Use in a Perceptual User Interface", Intel Technology Journal, pp.12-21, 1998.

[12] Bruce D. Lucas, Takeo Kanade, "An Iterative Image Registration Technique with an Application to Stereo Vision", in the Proceedings of the International Joint Conference on Artificial Intelligence, pp. 674-679, 1981.

[13] Carlo Tomasi, Takeo Kanade, "Detection and Tracking of Point Features", Carnegie Mellon University Technical Report CMU-CS-91-132, 1991.

[14] Jianbo Shi, Carlo Tomasi, "Good Features to Track", in the Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp.593-600, 1994.

[15] Ranganatha S, Y P Gowramma, "A Novel Fused Algorithm for Human Face Tracking in Video Sequences", in the Proceedings of the IEEE International Conference on Computation System and Information Technology for Sustainable Solutions (CSITSS), Bengaluru, India, pp.1-6, 2016.

[16] C. Harris, M. Stephens, "A Combined Corner and Edge Detector", in the Proceedings of the 4th Alvey Vision Conference, Manchester, UK, pp.147-151, 1988.

[17] Ranganatha S, Y P Gowramma, "An Integrated Robust Approach for Fast Face Tracking in Noisy Real-World Videos with Visual Constraints", in the Proceedings of the IEEE International Conference on Advances in Computing, Communications and Informatics (ICACCI), Udupi, India, pp.772-776, 2017.

[18] Douglas Decarlo, Dimitris N. Metaxas, "Optical Flow Constraints on Deformable Models with Applications to Face Tracking", International Journal of Computer Vision, Vol.38, Issue.2, pp.99-127, 2000.

[19] M. Kim, S. Kumar, V. Pavlovic, H. Rowley, "Face Tracking and Recognition with Visual Constraints in Real-World Videos", in the Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp.1-8, 2008.

[20] Ranganatha S, Y P Gowramma, "Development of Robust Multiple Face Tracking Algorithm and Novel Performance Evaluation Metrics for Different Background Video Sequences", International Journal of Intelligent Systems and Applications, Vol.10, Issue.8, pp.19-35, 2018.

[21] Stefan Leutenegger, Margarita Chli, Roland Y. Siegwart, "BRISK: Binary Robust Invariant Scalable Keypoints", in the Proceedings of the IEEE International Conference on Computer Vision, pp.2548-2555, 2011.

[22] Ranganatha S, Y P Gowramma, "Eigen and HOG Features Algorithm for Face Tracking in Different Background Challenging Video Sequences", International Journal of Computing, in press.

[23] Ranganatha S, Y P Gowramma, "Image Training, Corner and FAST Features based Algorithm for Face Tracking in Low Resolution Different Background Challenging Video Sequences", International Journal of Image, Graphics and Signal Processing Vol.10, Issue.8, pp.39-53, 2018.

[24] E. Rosten, T. Drummond, "Fusing Points and Lines for High Performance Tracking", in the Proceedings of the IEEE International Conference on Computer Vision (ICCV), pp.1508-1515, 2005.

[25] Yongkang Wong, Shaokang Chen, Sandra Mau, Conrad Sanderson, Brian C. Lovell, "Patch-Based Probabilistic Image Quality Assessment for Face Selection and Improved Video-Based Face Recognition", in the proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), pp.74-81, 2011.

[26] Ivan Laptev, Marcin Marszalek, Cordelia Schmid, Benjamin Rozenfeld, "Learning Realistic Human Actions from Movies", in the Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp.1-8, 2008.

[27] C. Sanderson, B.C. Lovell, "Multi-Region Probabilistic Histograms for Robust and Scalable Identity Inference," in the Proceedings of the International Conference of Biometrics, Alghero, Italy, pp.199-208, 2009.

[28] Ranganatha S, Y P Gowramma, "Selected Single Face Tracking in Technically Challenging Different Background Video Sequences Using Combined Features", ICTACT Journal on Image and Video Processing, in press.

[29] Ranganatha S, Y P Gowramma, "Color Based New Algorithm for Detection and Single/Multiple Person Face Tracking in Different Background Video Sequence", International Journal of Information Technology and Computer Science, in press.

[30] D. M. W. Powers, "Evaluation: From Precision, Recall and F-Measure to ROC, Informedness, Markedness & Correlation", Journal of Machine Learning Technologies, Vol.2, Issue.1, pp.37-63, 2011.

[31] T. Fawcett, "An Introduction to ROC Analysis", Pattern Recognition Letters, Vol.27, Issue.8, pp.861-874, 2006.

[32] Keni Bernardin, Rainer Stiefelhagen, "Evaluating Multiple Object Tracking Performance: The CLEAR MOT Metrics", EURASIP Journal on Image and Video Processing, pp.1-10, 2008.

[33] Anelia Angelova, Yaser Abu-Mostafa, Pietro Perona, "Pruning Training Sets for Learning of Object Categories", in the Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR), pp.494-501, 2005.

[34] "Face Database from Robotics Lab of National Cheng Kung University, Taiwan", URL: http://robotics.csie.ncku.edu.tw/Databases/FaceDetect_PoseEstimate.htm.

[35] Li Fei-Fei, R. Fergus, P. Perona, "Learning Generative Visual Models from Few Training Examples: An Incremental Bayesian Approach Tested on 101 Object Categories", in the Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshop (CVPRW), 2004.

[36] Li Fei-Fei, R. Fergus, P. Perona, "One-Shot Learning of Object Categories", IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol.28, Issue.4, pp.594-611, 2006.

[37] G. Griffin, A. Holub, P.Perona, "Caltech-256 Object Category Dataset", Technical Report 7694, California Institute of Technology, pp.1-20, 2007.

Authors Profile

Mr. Ranganatha S received B.E degree (Branch: CS & E, College: KIT, Tiptur, Karnataka, India) and M.Tech degree (Specialization: CS & E, College: RVCE, Bangalore, Karnataka, India) in the years 1998-2002 and 2004-2006 respectively from Visvesvaraya Technological University (VTU), Belagavi, Karnataka, India and currently pursuing his PhD in image processing at VTU Research Resource Centre. He has more than 16 years of teaching experience in reputed organizations at graduate and post graduate level. He is currently working as Assistant Professor, Department of Computer Science & Engineering, Government Engineering College, Hassan, Karnataka, India. He has published papers in various national and international journals and conferences. His current research interests include image processing and algorithms.

Dr. Y P Gowramma received B.E degree (Branch: CS &E, College: AIT, Chikmagalore, University: Mysore, Karnataka, India), M.Tech degree (Specialization: SACA, College: NITK, Karnataka, India) and PhD degree (University: VTU, Belagavi, Karnataka, India) in the years 1994, 1998 and 2014 respectively. She has more than 24 years of teaching experience in reputed organizations at graduate and post graduate level. She is currently working as Professor, Department of Computer Science & Engineering, Kalpataru Institute of Technology, Tiptur, Karnataka, India. She has published papers in various national and international journals and conferences. Her current research interests include image processing and algorithms.