PiscoMed Publishing

Automatic Detection of Face in Video Sequences by using Extended Semi Local Binary Patterns

Ithaya Rani Panneerselvam[1]*, Hari Prasath T.[2]

[1] Associate Professor Dept. of Computer Science and Engineering Sethu Institute of Technology, Pulloor, Kariapatti, Virudhunagar- 626005, India

[2] Assistant Professor, Dept. of Electrical & Electronics Engineering, Kalasalingam University, Anand Nagar, Krishnan-koil-626126, India

Abstract: Machine analysis of detection of the face is an active research topic in Human-Computer Interaction today. Most of the existing studies show that discovering the portion and scale of the face region is difficult due to significant illumination variation, noise and appearance variation in unconstrained scenarios. To overcome these problems, we present a method based on Extended Semi-Local Binary Patterns. For each frame, an aggregation of the pixel values over a neighborhood is considered and a local binary pattern is obtained. From these a binary code is obtained for each pixel and then histogram features is computed. Adaboost algorithm is used to learn and classify these discriminative features with the help of exemplar face and non-face signature of the images for detecting the location of face region in the frame. This Extended Semi Local Binary Pattern is sturdy to variations in illumination and noisy images. The developed methods are deployed on the real time YouTube video face databases and found to exhibit significant performance improvement owing to the novel features when compared to the existing techniques.

Keywords: Extended Semi Local Binary Pattern; Ensemble Classifier; Human-Computer Interaction

1. Introduction

One of the most interesting fields of image analysis is the automatic detection of the human face. The major applications of finding the face areas are face recognition, facial expression recognition, gender identification, face registration, human-machine interaction, surveillance, etc. Face discovery methods identify the human faces in the video sequences and give a rough estimation of the portion and scale of all faces in real time. But finding the region of the human face is a challenging task as the human faces are non-rigid and they appear in different background (simple, clutter and dynamic) and have a high degree of variability in terms of location, poses, expressions and illumination (good and bad)[1-2].

To overcome these problems, we propose a novel approach for detecting the face region by using Extended Semi-Local Binary Pattern (ESLBP). Our work considers the local appearance descriptors by extending the Semi Local Binary Pattern (SLBP) proposed in[14]. Since the proposed work encodes the information in a block (small overlap- ping block) of pixels by comparing the central block with locally neighboring blocks into a binary code. Each block is an aggregation of 3x3 pixel values, due to which, it is robust to illumination variations and noise. The remaining part of the paper is organized as follows: Section 2 briefly reviews the several related works. Section 3 proposes the extraction of ESLBP features. Section 4 shows the experimental setup and results. Section 5, offers conclusion and directions for

doi: 10.18282/hrms.v3i1.656

future work.

2. Related work

Detection of face techniques has been examined immensely in the previous decades. The methodology for discovering the face area utilizing skin color and the maximum morphological gradient combination image was exhibited[3-4]. The system failed when it manages with skin color areas including similar color background, region of dress, and the overlapping of faces. H. Sagha et.al proposed a methodology for discovering sparse features using a genetic algorithm for multi view face detection. Discovering these features was time intensive[5]. The Gabor Wavelet (GW) that catches the properties of orientation selectivity, spatial localization and is optimally localized in the space and frequency domains was used in face detection[6-8]. Despite that, it is both time and memory consuming to convolve the images with a bank of Gabor filters to extract multi-scale and multi-orientation co-efficient. The detection of the facial components utilizing speeded up robust features could achieve only moderate performance[12]. The techniques using the Local Binary Pattern (LBP), Modified Census transforms (MCT), and Local Gradient Pattern (LGP) based features for detecting the faces are presented in the existing work[9-11]. These techniques pixel based features are created at each location compare a central pixel with neighboring pixels.

While pixel based features are good for solving the computer vision problem, they are limited due to their sensitivity to noise and illumination variations. This is because pixel comparison is more sensitive to changed pixels than to block comparison. The block features such as Harr-Like features[13] used for face detection. This method compared the sum of intensity of two rectangular blocks which are neighboring to each other. Hence, comparing the intensity of the blocks reduce the effect of the noise and illumination variations. We concentrate on comparing neighboring blocks to make a new feature. We propose the use of a novel system known as ESLBP; a set of binary patterns with embedded local blocks. The proposed methodology is insensitive to head poses and robust to variations in lighting condition and noisy images by ESLBP features.

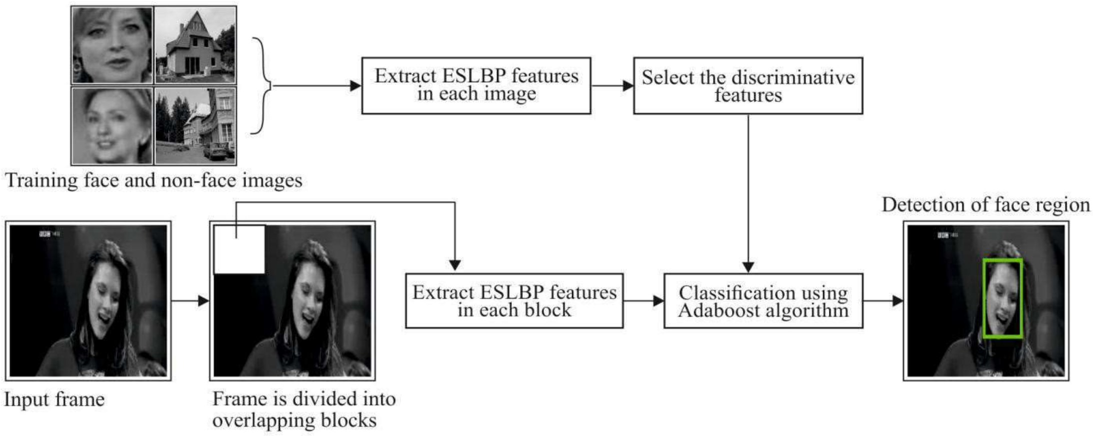


Figure 1; Overview of the system diagram for detecting the location of face area**.**

3. Proposed work

The video with single subject contains multiple frames depicting the temporal variations in different poses, expressions and varying lighting conditions of the individual. The overall framework of the proposed approach for the detection of face area based on ESLBP features is illustrated in Fig.1. The following steps describe the proposed $ESLBP_{5x5}$ approaches:

Each frame is extracted from video. The frame is divided into overlapping blocks. The local features are extracted by using $ESLBP_{5x5}$ in each block.

In each block contains $ESLBP_{5x5}$ features are passed to AdaBoost classifier for detecting the face region.

The performance of the proposed work using $ESLBP_{5x5}$ features are also compared with the LBP, SLBP, $ESLBP_{6x6}$ and $ESLBP_{7x7}$ by deploying the Ensemble classifiers. For conducting and evaluating the work, YouTube (YT) video face databases [15] are taken. The following sub-sections describe the procedure in detail.

3.1 ESLBP features extraction

Initially, each frame is extracted and represented as set of frames $\{f_1, f_2, ..., f_k\}$ from video V, where k is the number of frames. We propose the ESLBP procedure by dividing the k[th] frame f_k into overlapping blocks represented as $B_1, B_2, ...B_{m*n}$. The number of blocks is (m)*(n) and m and n are the number of rows and columns in each frame respectively. Typically, each block is the size of 30x30 as follows.

$$f_{k,B}(i,j) = f_{k,B}(i:i+29, j:j+29) \qquad 1 \le i \le m; 1 \le j \le n \tag{1}$$

The reason for the selection of block size of 30*30 is to match against the training images of 30*30. Furthermore, each block 'B' is divided into overlapping sub-blocks represented as $SB_1, SB_2, ..., SB_{m_1*n_1}$ where each sub- block is of size 5x5. The number of sub-blocks is $(m_1)*(n_1)$ and m_1 and n_1 are the number of rows and columns in each block respectively (30 in this case). Following this, $ESLBP_{5x5}$ is defined by a binary coding function [9] to find the sum of intensity in a window of size 3x3 within a sub-block as follows:

$$P^{a,b}_{k,B,(x,y)\varepsilon SB}(i,j) = \sum_{i_2=a}^{a+2}\sum_{j_2=b}^{b+2} f_{k,B,SB}(i+i_1+i_2, j+j_1+j_2) \qquad 1 \le a \le 3; 1 \le b \le 3 \tag{2}$$

$$where\ 1 \le x \le m_1;\ 1 \le y \le n_1 \quad i_1 = x : x+4; j_1 = y : y+4$$

Where $P^{a,b}_{k,B,(x,y)\varepsilon SB}(i,j)$ the sum of intensities is over a window of 3*3 centered on each pixel for the sub-block 'SB' whose size is 5x5 at each pixel position (i,j). The LBP [9] is obtained on these processed values over these sub blocks. The representations of these quantities are as follows:

Let M to represent the matrix as:

$$M = P^{a,b}_{k,B,(x,y)\varepsilon SB}(i,j) \tag{3}$$

The value by using the proposed method $ESLBP_{5x5}$ is obtained as:

$$f_{ESLBP}(i,j) = \sum_{i=0}^{2}\sum_{j=0}^{2} T(i,j)\, 2^8$$

$$where\ \ i \ne 1 and\ \ j \ne 1 \tag{4}$$

and

$$T(i,j) = \begin{cases} 1 & M(i,j) \ge M(1,1) \\ 0 & else \end{cases} \tag{5}$$

Fig. 2 shows a worked example of the sample block with the proposed method.

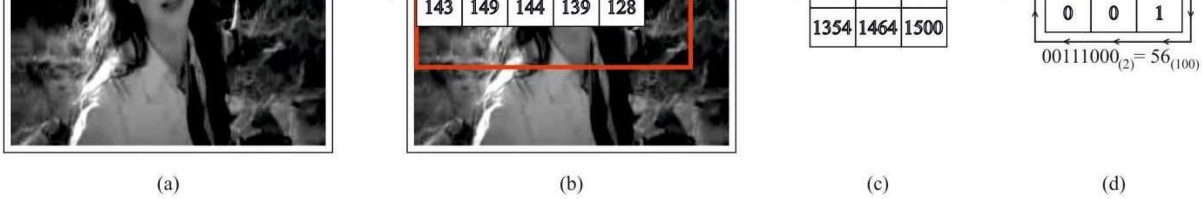


Fig.2 ESLBP procedure a) original image b) red color box denoted as block 'B' size 30x30 and Sub-block size 5x5 within a block c) gray color represented as sum of window of each size 3x3 within a sub- block image. d) The LBP operation transforms the sum of sub-block into a decimal number.

sidered, which leads to better features extraction. After obtaining the value using $ESLBP_{5x5}$ method for each pixel associated with a block, a 59-bin histogram is computed to extract the features for each block. A histogram (H) of the block

f_{ESLBP} can be defined as:

$$h_L = \sum_{i,j} I\,(lower_L < f_{ESLBP}(i,j) \geq higher_L)\ 1 \leq i \leq m, 1 \leq j \leq n\ ,\quad 1 \leq L \leq 59 \tag{6}$$

where L is the number of bins for the values produced by the $ESLBP_{5x5}$ operator for each block. The interval of each bin is represented by the range $lower_L$ and $higher_L$.

$$I(A) = \begin{cases} 1 & A \text{ is true} \\ 0 & A \text{ is false} \end{cases} \tag{7}$$

The ESLBP histogram contains information about the description of the local micro-patterns such as edges, spots and flat areas, over the whole image, so it is used to statistically describe image characteristics. We obtained 59-histogram bins for each block.

3.2 Learning ESLBP-Histogram bins

We propose to learn discriminative ESLBP-Histogram bins for better face detection. Adaboost provides a simple yet effective approach for stage wise learning of a nonlinear classification function [13]. So, we adopt the AdaBoostalgorithm to learn the discriminative ESLBP histogram bins. It is not only useful to select the discriminating features, but also to train the classifier. AdaBoost learns a small number of weak classifiers whose performance is just better than random guessing, and boosts them iteratively into a strong classifier of higher accuracy/minimum error. The process of AdaBoost maintains a distribution of the training samples. At each iteration, a weak classifier which minimizes the weighted error rate is selected, and the distribution is updated to increase the weights of the misclassified samples to reduce the importance of the others. Similar to the proposal in [16], the weak classifier $p_i(x)$ consists of a feature h_i which corresponds to a single $ESLBP_{5x5}$ histogram bin, a threshold θ_i and polarity g_i indicating the direction of the inequality sign:

$$p_i(x) = \begin{cases} 1 & \text{if } g_i h_i(x) \leq g_i \theta_i \\ 0 & \text{otherwise} \end{cases} \tag{8}$$

The proposed work for detecting the location of face region is summarized in Algorithm 1 below.

Algorithm 1

Input: Video, face and non-face training images

Output: detection of face region

Step-1: For a given video, frames are extracted from video. Each frame is resized into 240x240 pixels.

Step-2: Each frame is divided into overlapping blocks. Each block size is 30x30 pixels. The reason for the selection of block size of 30*30 is to match against the training images size of 30*30 pixels.

Step-3: The $ESLBP_{5x5}$ procedure as follows:

Each block is divided into overlapping sub-blocks. The sub-blocks size is 5*5 pixels.

From the sub-block, aggregation of 3x3 pixels over a neighborhood pixels. The resultant value is reduced to 3x3 pixels from 5*5 pixels sub-blocks.

LBP is obtained from the resultant value. It returns one scalar value.

Repeat the steps from a to c until, reach the end of each block (in the case 30*30 pixels)

Step-4: Adaboost algorithm is used to classify the face or non- face region in each block with the help of face and non-face signature of training images.

Step 5: Repeat the steps from Step-2 to Step- 4 until, reach the end of each frame.

4. Experimental data, results and discussion

To evaluate the performance of our proposed method, YT video datasets were used for the experiment. YT video clips contain 47 celebrities, mostly actors/actresses and politicians. Most of the videos are low resolution and recorded at high compression rates. This leads to noisy, low-quality image frames. The dataset consist of about 1910 video clips, each containing hundreds of frames. The frame rate of 30 fps and resolution sizes varies from (180 × 240) to (240 × 320). Out of the 1910 video sequence studies, 1870 of them consists of only one person and the remaining have more than one person. Detection of the face and its component has been a challenging task using this database since the video exhibit large variations in face pose, illumination, expression, and other conditions. A total number of 388 videos pertaining 47 celebrities were used for gallery and the remaining were used for testing. Table 1 shows the detailed contents of the gallery and testing data set. Fig. 3 shows some samples of the gallery images.

Fig.3. Sample of gallery images a) face, non-face images

Table 1 Description of the Experimental data selected

Training Image Data set			Testing Data	Features dimensions (training and testing)
388 faces (different poses, view and expressions) from YT database	450 Non-face images (some collected from background images and some manually created)	838 Total gallery images	1522 videos from YT database	59 (corresponding to 59 bins)

We collected the gallery images, which comprise of face images and non-face images. Initially features are extracted such as LBP[9], SLBP[14], $ESLBP_{5x5}$, $ESLBP_{6x6}$ and $ESLBP_{7x7}$ from each gallery image. We evaluate the $ESLBP_{6x6}$ scheme which represents a sub-block of size 6x6 and sum of the window of each size 4x4 within a sub-block and the $ESLBP_{7x7}$ which represents a sub-block of size 7x7 and sum of the window each size 5x5 within a sub-block which is compared with our proposed ESLBP5x5 scheme. To begin with, we make LBP, SLBP, $ESLBP_{5x5}$, $ESLBP_{6x6}$ and $ESLBP_{7x7}$ histograms of each gallery images described by a 59- histogram bin. In all cases we adopted Adaboost algorithm to learn discriminative histogram bins and boost a strong classifier. The minimum error performance of the boosted strong classifier achieve as a function of the number of features selected. With the selected 27 bins, the boosted strong classifier achieves the minimum error rate in Good Illumination (GI), Bad Illumination (BI), noisy image (N) and Multiple Subjects (MS) videos as plotted in Fig.4. From Fig. 4 the following observations can be made:

GI videos – The SLBP, $ESLBP_{5x5}$ and $ESLBP_{6x6}$ features have error rates of .1, .06 and .09 respectively which is less than the error rates of LBP and $ESLBP_{7x7}$ features. The $ESLBP_{7x7}$ features have an error rate of .12 which is somewhat less than LBP feature's error rate of.13.

BI videos – The $ESLBP_{5x5}$ features have an error rate of .09 which is less than SLBP and LBP features error rate of .15 and .12 respectively. The $ESLBP_{7x7}$ and $ESLBP_{6x6}$ features have error rates of .083 and .087 respectively which is slightly less than error rate of $ESLBP_{5x5}$ features.

N videos – The $ESLBP_{6x6}$ and $ESLBP_{7x7}$ features have an error rate of .124 and .13 respectively which is less than SLBP features error rate. The SLBP features have an error rate of .15 which is less than LBP features error rate of .17. The $ESLBP_{5x5}$ features have an error rate of .11 which is less than $ESLBP_{6x6}$ and $ESLBP_{7x7}$ features.

MS videos – The SLBP features has an error rate of .163 which is less than the LBP features error rate of .178. The $ESLBP_{5x5}$ features have an error rate of .114 which is less than $ESLBP_{6x6}$ and $ESLBP_{7x7}$ feature's error rate of .137.

As a result, we conclude that the proposed method $ESLBP_{5x5}$ features reduces the error compared with other methods such as LBP, SLBP, $ESLBP_{6x6}$ and $ESLBP_{7x7}$ on YT database.

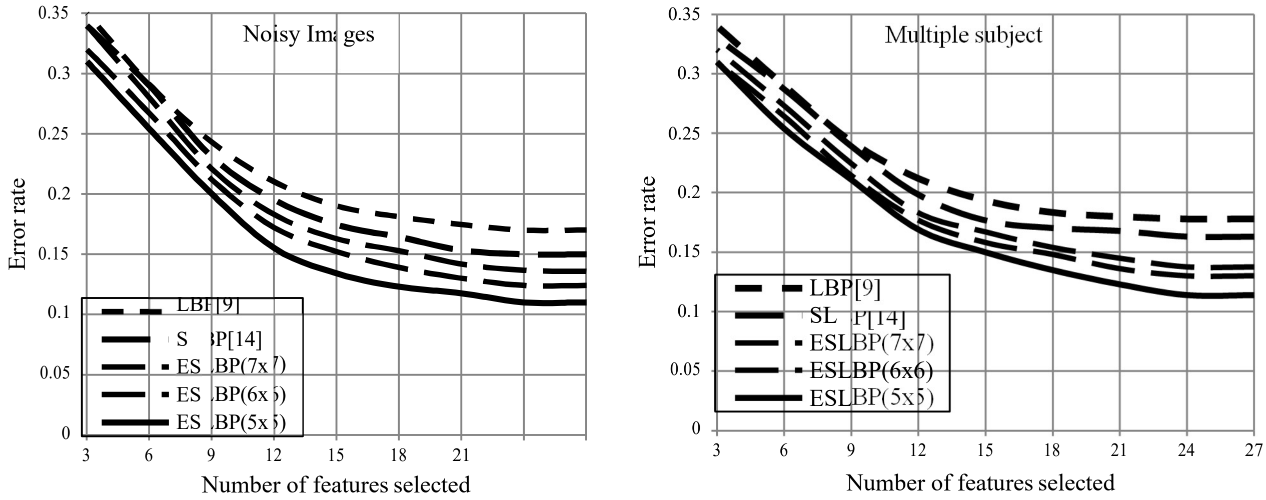


Fig.4 Error rate of LBP, SLBP, ESLBP $_{7x7}$, ESLBP$_{6x6}$ and ESLBP$_{5x5}$ features in different types of videos.

Fig.5 shows the receiver operating characteristic curves. The curve is plotted from the results obtained using the YT databases for the various features such as: LBP, SLBP, ESLBP$_{7x7}$, ESLBP$_{6x6}$ and ESLBP$_{5x5}$ tested with GI, BI, N and MS in the video set. Fig.4 depicts the relationship between a number of false positives and the detection rate. It can be observed the detection rate as follows:

GI videos – The LBP and ESLBP$_{7x7}$ schemes have a features detection rate of 87% and 88% respectively which is lower than SLBP, ESLBP$_{5x5}$ and ESLBP$_{6x6}$ features detection rate. Because, the ESLBP$_{7x7}$ features follows aggregation of pixel size 5x5 over a neighborhood this leads to over brightness and does not detect the face region correctly. The LBP features slightly tolerate the illumination variation in the images. Both LBP and ESLBP$_{7x7}$ features do not achieve the high detection rate in GI videos. The ESLBP$_{5x5}$ features detection rate of 94% is higher than SLBP and ESLBP$_{6x6}$ features. ESLBP$_{6x6}$ features detection rate of 91% is little greater than SLBP features detection rate of90%.

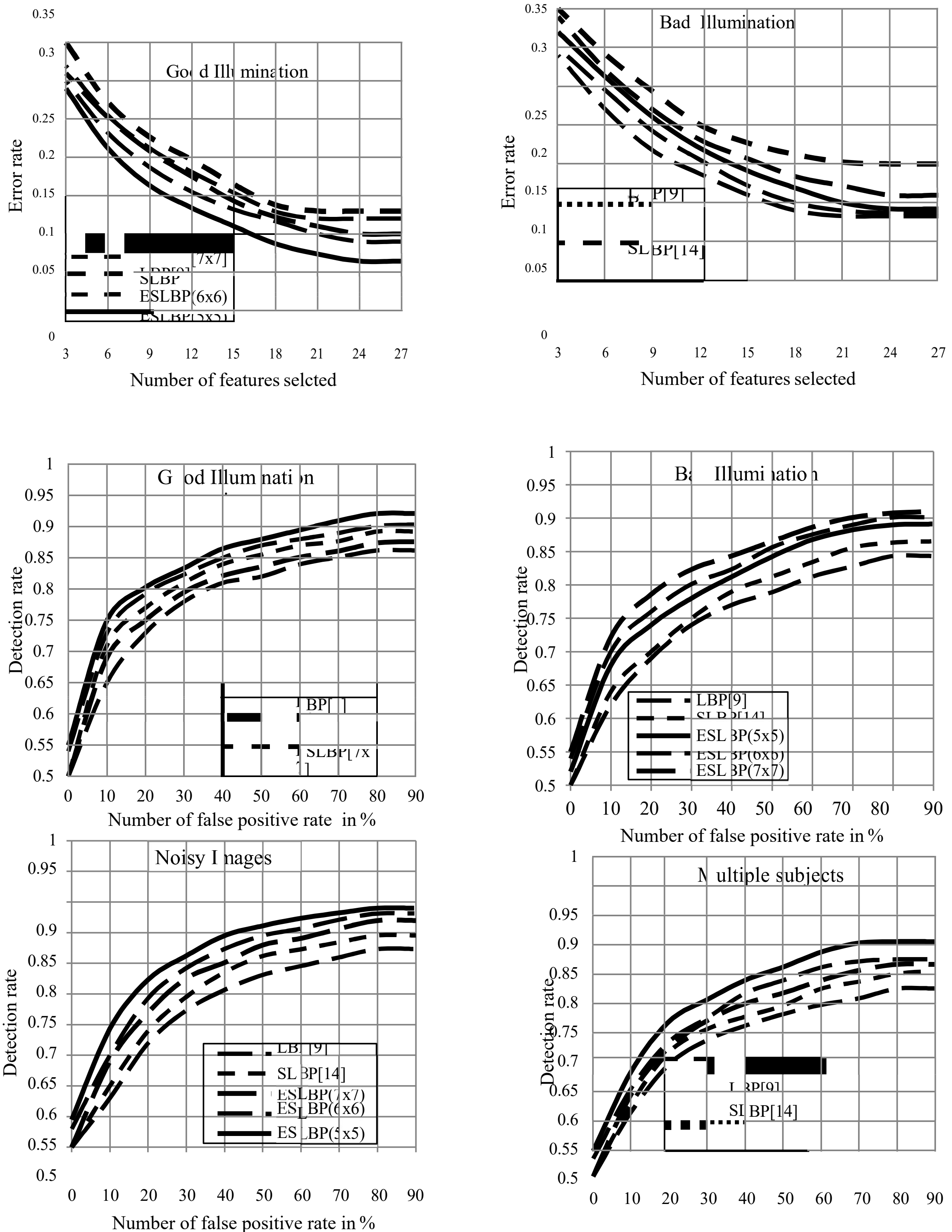
BI videos – The LBP and SLBP schemes have a features detection rate of 85% and 88% respectively which is lower than ESLBP$_{5x5}$, ESLBP$_{6x6}$ and ESLBP$_{7x7}$ features detection rate. The ESLBP$_{6x6}$ features detection rate of 91.3% is somewhat higher than ESLBP$_{5x5}$ features detection rate of 91%. However ESLBP$_{7x7}$ features detection rate of 91.67% is slightly higher than ESLBP$_{5x5}$ and ESLBP$_{6x6}$ features with a detection rate of 91%.

N videos – The SLBP features has a detection rate of 85% which is greater than LBP features detection rate of 83% since the LBP features does not aggregate the pixel values. It is sensitivity to noisy images. Hence the face region is not detected in the noisy images. The SLBP features detect the face region in the noisy images somewhat better than LBP features. The ESLBP$_{6x6}$ and ESLBP$_{7x7}$ features have detection rates of 87.6% and 87% respectively which is higher than SLBP and LBP features. However ESLBP$_{5x5}$ features detection rate of 89% which is little bit higher than ESLBP$_{6x6}$ and ESLBP$_{7x7}$ features.

MS videos – The ESLBP$_{6x6}$ features detection rate of 87% is higher than SLBP features detection rate of 85%. The LBP features detection rate is 83% which is lower than SLBP features detection rate of 85%. The ESLBP$_{5x5}$ features detection rate of 90% is higher than ESLBP$_{6x6}$ and ESLBP$_{7x7}$ features. The SLBP and ESLBP$_{7x7}$ features have the same detection rate. Since the ESLBP$_{7x7}$ features are over summed, it leads to non detection of the face region with multiple poses and different expressions.

From Fig. 5, it is observed that the proposed ESLBP$_{5x5}$ features deliver the best performance for the face detection rate on YT databases. Mainly, they have good characteristics to represent the face with high detection rate in a little false positive number case.

The performance indicators in the detection of the face includes calculating the Sensitivity, Precision and F measure and the results are shown in Table 2. Fig.6 shows the sample result. ESLBP$_{7x7}$ features would fail to detect the face in the GI due to aggregation of 5x5 sub-block size. But it is insensitivity in BI and N images. The LBP and SLBP features would fail to detect the face in the noise image. The LBP features slightly insensitivity to illumination variations and more sensitivity to noisy images. The SLBP features slightly insensitivity to noisy images. The SLBP and $ESLBP_{7x7}$ features robust against the appearance variations. The $ESLBP_{6x6}$ features are better than $ESLBP_{7x7}$ features to detect the face in appearance variations, illumination variations and noisy images. However $ESLBP_{5x5}$ features is better than $ESLBP_{6x6}$ in all types' videos. From Fig.6, we can show that the proposed $ESLBP_{5x5}$ features are robust against noise and illumination and differences pose variations and expressions. The size of the bounding box is determined using the scale on the detected face on the video sequences.



N umber of false poitive rate in %

Fig. 5 The performance evaluation with the five features. The proposed ESLBP features present the best performance for the face detection rate on YT database

Table 2 - Performance Comparison of face detection procedure

Features Extraction	Sum of intensity window size	Sensitivity in (%)				Precision in (%)				F measure in (%)			
		No. of videos				No. of videos				No. of videos			
		575	390	522	35	575	390	522	35	575	390	522	35
		GI	BI	N	MS	GI	BI	N	MS	GI	BI	N	MS
LBP [9]	-	87	85	83	83	88	84	85	83	87	84	84	83
SLBP $_{4x4}$ [14]	2x2	90	88	85	85	90	87	87	84	90	87	86	84
ESLBP $_{7x7}$	5x5	88	91.67	87	87	88	89	89.4	86	88	90	88	86
ESLBP $_{6x6}$	4x4	91	91.3	87.5	87	89	88	90	86	89	89	89	87
ESLBP $_{5x5}$ Proposed	3x3	94	91	88.56	90	92	88	91	90	93	89	90	90

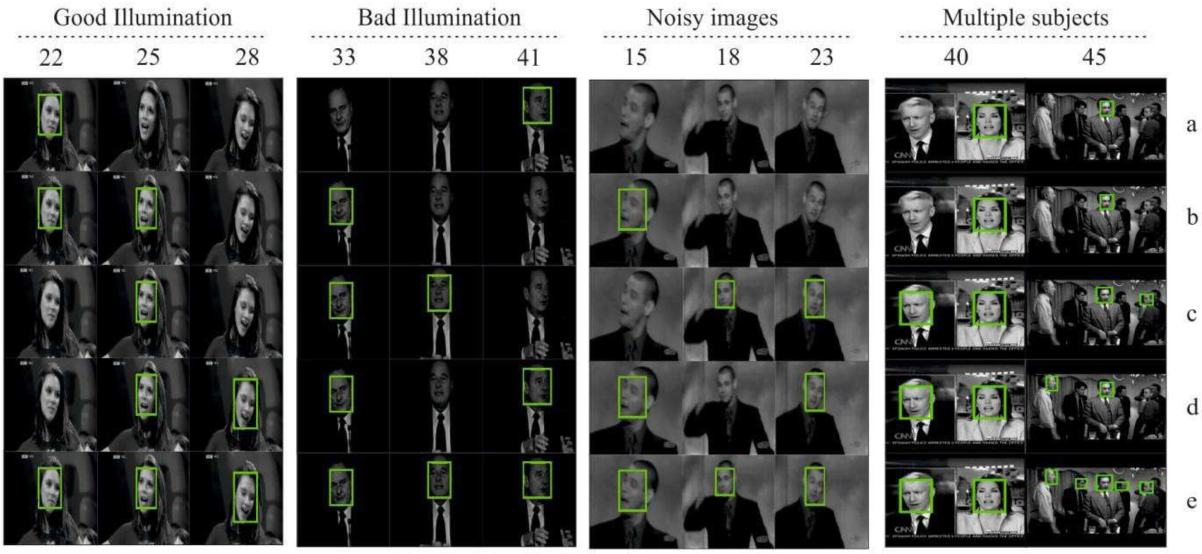


Fig. 6 Sample detection of face result in YT databases a) LBP b) SLBP c) ESLBP$_{7x7}$ d) ESLBP$_{6x6}$ e) ESLBP$_{5x5}$

ble 3 shows that the empirical time for detecting the location of face region.

Table 3 Comparison Execution Time of the proposed approach with existing approach

Features Extraction	Sum of intensity window size	Execution time of features extraction from training images	Average Execution time for the detection of face in a Videos			
			GI (575 videos)	BI (390 videos)	N (522 videos)	MS (35 videos)
LBP [9]	-	0.0025	1.6939	2.0238	2.3246	18.0632
SLBP $_{4x4}$ [14]	2x2	0.0028	1.7080	2.0547	2.3730	18.2179
ESLBP $_{5x5}$ Proposed	3x3	0.00285	1.7103	2.0831	2.4067	18.2521
ESLBP $_{6x6}$	4x4	0.00295	1.7141	2.0935	2.4119	18.2571
ESLBP $_{7x7}$	5x5	0.002989	1.7200	2.1026	2.4163	18.2727

5. Conclusion and Future work

This paper investigated the benefits of ESLBP features for strong discovery of face area in video sequences under uncontrolled scenario. The experimental results showed that the proposed method is showing promising results when compared with the existing methods. The optimal selection of the size of the neighborhood area over which the aggregation operation is done helps the elimination of noise and illumination variations. The ESLBP$_{5x5}$ are encoded to discriminate face and non-faces. Also they are robust to the varieties of image condition. The test results exhibit that

ESLBP features finds its appropriate role in face detection applications carried out on the YT databases. The extension of the work is in progress in the recognition of facial expression.

References

[1]M.Yang, D. Kriegman, N. Ahuja , Detecting faces in images: a survey, IEEE Transactions on Pattern Analysis and Machine Intelligence, 24(1), 2002, 34–58.

[2]Hielm, E., Low, B.: `Face detection: a survey`, Computer Vision and Image Understanding, 2001, 83, (3), 236–274.

[3] Borah, S. Konwar, S. ; Tuithung, T. ; Rathi, R. A human face detection method based on connected component analysis , Communications and Signal Processing (ICCSP), 2014 International Conference on 1205 – 1208,2014.

[4] Byung-Hun Oh , Kwang-Seok Hong , A study on facial components detection method for face-based emotion recognition , International Conference on Audio, Language and Image Processing (ICALIP), 2014 , 256 – 259

[5]H. Sagha, S. Kasaei, E. Enayati, M. Dehghani , Finding Sparse Features in Face Detection Using Genetic Algorithms , IEEE International Conference on Computational Cybernetics, ICCC 2008 ,179 - 182

[6] Li Xiaohua , Kin-Man Lam , Shen Lansun , Zhou Jiliu Face detection using simplified Gabor features and hierarchical regions in a cascade of classifiers , Pattern Recognition Letters , (2009) 717–728

[7] Lin-Lin Huang *, Akinobu Shimizu, Hidefumi Kobatake, Robust face detection using Gabor filter features , Pattern Recognition Letters 26 (2005) 1641–1649

[8] T. Yun and L. Guan, "Automatic face detection in video sequences using local normalization and optimal adaptive correlation techniques," Pattern Recognit., vol. 42, no. 9, pp. 1859–1868, Sep. 2009.

[9] T. Ojala, M. Pietikainen, and D. Harwood, "A comparative study of texture measures with classification based on features distributions," Patt. R ecognit., vol. 29, no. 1, pp. 51 - 59, 1996.

[10] B. Froba and A. Ernst, "Face detection with the modified census transform," in Proc. IEEE Int. Conf. Automatic Face and Gesture Recognition, 2004, pp. 91 - 96.

[11] B. Jun and D. Kim, "Robust face detection using local gradient patterns and evidence accumulation," Patt. Recognit., vol. 45, no. 9, pp. 3304 - 3316, 2012.

[12] Donghoon Kim, Rozenn Dahyot, Face Components Detection using SURF Descriptors and SVMs , Machine Vision and Image Processing International Conference, 2008. 51 – 56 .

[13] Paul Viola and Michael Jones, "Robust real-time object detection," IJCV, vol. 57, no. 2, pp. 137–154, 2004.

[14] Kyungjoong Jeong, Jaesik Choi and Gil-Jin Jang , Semi-Local Structure Patterns for Robust Face Detection, IEEE Signal Processing Letters, VOL. 22, NO. 9, 2015 pp 1400-1403

[15] Kim M, Kumar S, Pavlovic V, Rowley H (2008) Face Tracking and Recognition with Visual Constraints in Real-World Videos. IEEE Conf. Computer Vision and Pattern Recognition

[16] P. Viola, M. Jones, Rapid Object Detection using a Boosted Cascade of Simple Features, Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2001,Vol (1), pp.511–518.