International Journal of Scientific and Research Publications, Volume 10, Issue 1, January 2020
ISSN 2250-3153

An Unconventional Framework for Smile Detection using Eye States

C.L.I. Fonseka*, L.S Erandika, S.Sotheeswaran**

Department of Information and System Engineering, UCSC, Sri Lanka
**Department of Information and System Engineering, UCSC, Sri Lanka
** Department of Mathematics, at Eastern University, Sri Lanka

DOI: 10.29322/IJSRP.10.01.2020.p9753
http://dx.doi.org/10.29322/IJSRP.10.01.2020.p9753

Abstract - Facial expression analysis plays a considerable role in under human emotions and behaviours. Analysing facial expressions accurately has board application areas like human behavior analysis, human-human interaction and human-computer interaction. Automatic identifying of smile or non-smile from images has been a challenging and actively studied problem over the past few decades. Since it has many uses like patient observation, camera photo capturing and more. In this research work the smile and non-smile face images classifies through the proposed system which involves the following steps: First, extract the scale-invariant feature transform (SIFT) or speeded-up robust features (SURF) features, then construct the codebook which provides a way to map the descriptors into a fixed-length vector in histogram space. Second, extract the histograms of oriented gradient (HOG) features and Local Binary Pattern (LBP). Third, combine the extracted features and reduce the dimensionality. Finally, the binary-class classify the feature histograms using support vector machines (SVMs). The proposed system focus on detecting smiles from face images that contain either a smile or a non-smile efficiently with highest accuracy by reducing computational needs such as computational time, memory, and disk space.

Index Terms - Smile detection, feature fusion, Support Vector Machine, Local Binary Patterns, classifier, Extract Features,

I. INTRODUCTION

Smile is the most common facial movement that appears on the face. Smile detection can be used to measure the person's mental state. Smile detection has many applications such as intensive care system, resistance training, and interactive systems. Nowadays, research attention has started to toward the more realistic problem of analysing of facial expressions. It seems very difficult to capture the complex decision boundary among facial expressions. When automatic recognition of emotion became feasible, novel challenges has evolved. One of them is the recognition whether a presented emotion is genuine or not. Many face detection methods are already developed and used in many applications. Most of these works focused on object based techniques, in which the facial expressions are identified using facial action coding system (FACS).

In this research, a novel approach will be proposed to identifying the simile and non-smile faces by considering eye state. The novel approach focuses the feature fusion and Bag-of-features (BoF) approach with dimensionality reduction which is to be enhanced the classification accuracy and reduce the time complexity and storage spaces.

The bag-of-features is the most popular approach in recent visual object recognition which makes use of local information extracted at several patches in an image. This BoF approach has proved to yield state-of-the-art performance in large evaluations such as the PASCAL Visual Object Classes and ImageNet Challenges. The general framework of a BoF approach can be summarised in the following four steps: (i) feature extraction from images (ii) cluster analysis on the extracted descriptors from training images (iii) BoF representation of training and testing image sets and (iv) classification of test feature vectors.

The rest of this paper is organised as follows. In Section 2, summarise different techniques that are closely related to smile detection. Section 3 provides the comparison of the previous results. In Section 4, the proposed technique for classifying smile faces is described in detail. Section 5 concludes the paper with a discussion of the findings towards future extensions.

II. LITERATURE REVIEW

A. Smile Detection using Local Binary Patterns and Support Vector Machines

In this paper [1], the authors have proposed an approach for smile detection using Local Binary Pattern (LBP) and Support Vector Machine (SVM). Different LBPs were used as main image descriptors for smile detection. Simplified LBP (SLBP) and Uniform LBP (ULBP) were used to reduce the length of the feature vector.

Rotation invariance is achieved in the LBP based representation considering the local binary pattern as circular. The author has used PCA for reduce the dimension and also k-nearest neighbor's algorithm (k-NN) and SVM were used for classification. The best results came from the following approaches: LBP operator is applied to derive features.

- Two different approaches (SLBP/ULBP) were used to obtain PCA space as a result of the original grayscale

images.

- The concatenation of histograms retrieved from the original image encoded with ULBP and SLBP.
- The concatenation of pixel values from the image encoded with SLBP and ULBP.

For the experiment they have used a dataset of 2421 images of different smiling faces and 3360 images of different non smiling faces with a size OD 59*65 pixels. They have considered two possibilities they are the whole normalized face image and the image parts such as both eyes and the mouth. The classification accuracy was reported 90% using support vector machines (SVM).

This paper said that the distribution of Simplified LBP can be used as a good representation for images with more or less uniform textures. But for the face image to increase the performance it should be restricted to just the mouth.

B. Facial Expression Recognition based on Edge Detection

In this paper [2], the authors have proposed an approach for facial expression recognition using Edge Detection. For the face detection from the images they have used the skin color detection technique which is namely as YCbCr color space. They have used four different edge detection techniques such as Robert, Sobel, Laplace and Canny and they have compared the performance.

The Roberts operator is performed a simple, quick to compute, 2-D spatial gradient measurement on an image. It thus highlights regions of high spatial gradient which often correspond to edges. The Sobel edge detector was made a gradient based method. The first order derivatives are worked with it. Unlike the Sobel edge detector, only one kernel is used by the Laplacian edge detector. It is calculated second order derivatives in a single pass. The Canny edge detector was made an edge detection operator that uses a multi-stage algorithm to detect a wide range of edges in images. And the canny edge detector having this last the best detection performance among the four algorithms.

This paper introduces a study and design of the system of facial expression recognition based edge detection algorithm, first for image preprocessing, image recognition processing allows easy back; and then locate the eyes and lips, individually marked and extract the edge shape feature; finally, the system was trained by using face database, achieve the purpose of identifying other face expression.

Table I - Test Results

Facial Expression Discrimination (%)	Normal	Sad	Smile	Surprise
Canny	100	91.3	99.4	95.7
Laplace	93.2	87.1	89.5	85.6
Sobel	63.4	49.7	66.8	64.9
Robert	49.8	30.1	42.5	48.2

They have chosen the suitable for the color of skin of Japanese Jaffe facial expression database as the material, and carries on the classification according to the requirements of the experiment, selecting four kinds of facial expression such as Normal, Sad, Smile, Surprise and each expression was selected three pictures, including two as the training group, the other as a test group. The obtained results are shown in the Table I.

C. Toward Practical smile detection

In this paper [3], the authors have presented a broad study on automatic smile detection in digital employing a Gabor filter approach.

For the experiment, they were converted all images to grayscale and then normalized by rotating, cropping and scaling the face about the eyes to reach a canonical face of 48x48 pixels, which was based on the manually labeled eyes positions. They have compared five image representations for the feature extraction such as Gabor Energy Filter (GEF), Box Filters (BF), Edge Orientation Histogram (EOH), Combine BF and EOH, Local Binary Pattern (LBP). They have compared two popular classifiers such as GentleBoost and SVMs. GentleBoost is a boosting algorithm that minimizes the x-square error between labels and model predictions. When training with linear SVMs, the entire set of Gabor Energy Filters or Box Filters had used as the feature vector of each image.

Authors have collected two different data sets of facial expression for training set. DFAT data set contains 101 smiles and 848 nonsmiles and GENKI dataset contains 17822 smiles and 7782 nonsmiles face images. The smile detection accuracy on white faces was 97.5% whereas for black faces it was only 90%.

In this paper, they have focused on detecting smiles in poses within approximately ±20 degrees from frontal. As authors have mentioned developing expression recognition systems that are robust to pose variations will be an important challenge for the near future.

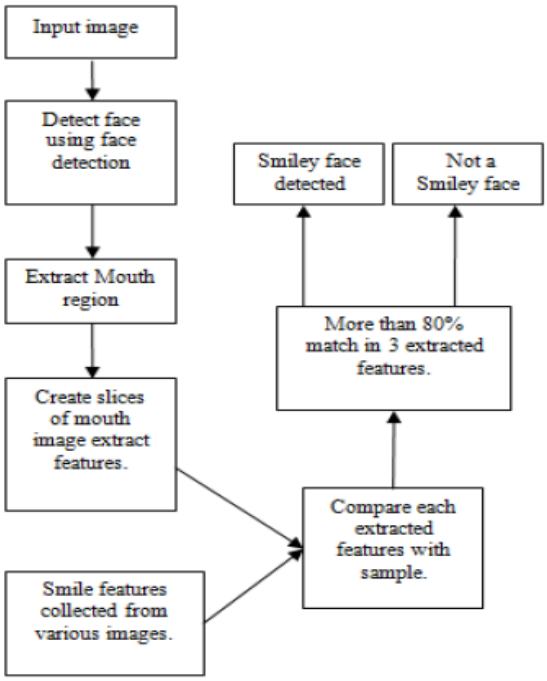

Fig. 1. Block diagram of the proposed system

D. Smile Detection: A Simple Approach

In this paper [4], they have proposed a simple image-based

approach to identify the smile face. To extract the face from the input image they have used three different feature extraction techniques such as feature-based approach, Eigenface-based method and Neural Network. The Principal Component Analysis applied to extracted image slices to reduce the dimensionality of the image data.

Initially a number of different types of human smile pictures represented in gray scale format were collected for a dataset. The faces are extracted from the dataset using above mentioned face extraction techniques. The mouth region of all extracted face images was cropped and created the slices. The features are extracted from the sliced images. The extracted features from the mouth region are compared with smile features which were from various images. If more than 80% match is found in more than those features considered in a smile face. The block diagram of their proposed system is given in the figure 1.

SIMULINK has used to implement this and all the required functions were developed using MATLAB. First two steps were the Face extraction and the localization of the eyes. HSV color model was used to extract the face, while the eyes were detected using LAB color space thus eliminating unwanted areas using Region of Interest and labeling process.

Converted the given image in RGB colour space into YCbCr colour space. The Y, Cb and Cr ranges for skin region were $50<Y<142$, $107<Cb<124$ and $135<Cr<152$. Each pixel of YCbCr was compared against the limits of Y, Cb and Cr to determine if the skin is present and if so by means of thresholding, the face was identified and segmented. In order to eliminate non-skin regions in the image, it is required to erode and dilate the image using a structured element. Subsequently do filling operation to fill the area defined by locations with connectivity. They have compared the dimensions with certain thresholds for each region and percentage of skin in each region, which was helped in removing non -face object. For each region if height and width were within the range then the processed image was a face and otherwise it is not a face.

To find and track eyes, they have applied Lab transform to the extracted face to eliminate unwanted portion. And then applied morphological operations on the output of transformed image to remove noise. Finally, they have determined the region of interest which locates eyes and applied region properties to track the eye pair.

The vector of the pupil and iris area has less gray values than two other white areas. As a result, the mean and standard deviation of the "OPEN" were found and concluded that mean was always >0.2 and standard deviation was always <0.02. If the mean was <0.2 and standard deviation was >0.02, then the state of the eye was concluded as "CLOSE"

A total of 36 different images from GTAV database and 30 images from local database they have been tested in the laboratory. The success rate of the proposed algorithm is 89.5%.

E. Facial Smile detection based on Deep Learning Features

In this paper [6], the authors have proposed an approach for Facial smile detection using deep convolution networks. They have designed a 6-layer deep network (Basic structure of CNN) and then it was modified to a new structure of CNN which is used both recognition and verification signals as supervision to learn expression features. The recognition signal that was responsible for the classification task. The expression verification signal, which was effective to reduce the variation of features which was extracted from the images of the same expression class (smile or non-smile).

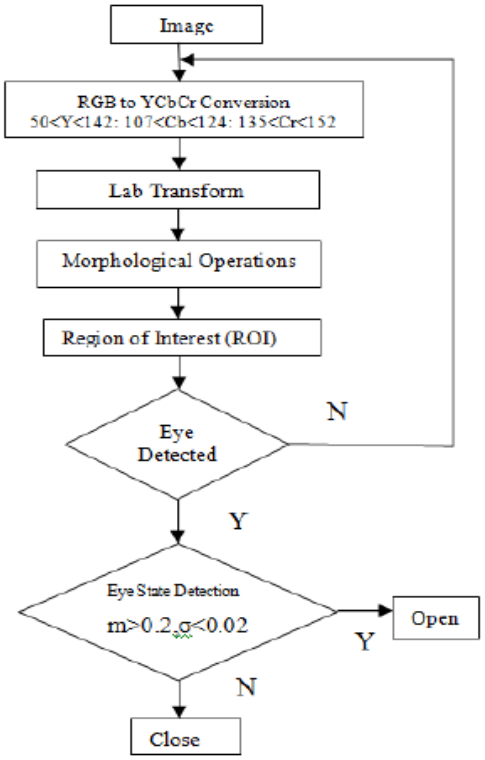

Fig. 2. Flow Chart of proposed algorithm

The network was trained via a two-way soft-max classifier to predict smile or non-smile. Rectified linear unit (ReLU) function was used as the activation function in the convolutional layers and fully-connected layers. To increase the translation invariance and avoid overfitting they have chosen max-pooling with a neighboring region. The two-signal guided structure of CNN is shown in the figure 3.

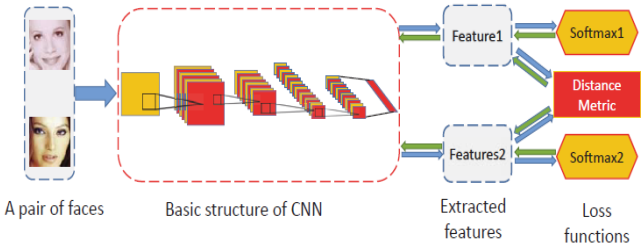

Fig. 3. The two-signal guided structure of CNN.

A pair of images are sent into the structure, with the proposed basic of CNN to generate GENKI-4K database was used for their experiment which contains 2162 smile images and 1828 non-smile images with different ages and races. In each time, they have selected 3000 images for training images and the rest 1000 images for testing.

The authors have achieved a classification accuracy of 94.6%,

which has greater than any accuracy attained by previous methods on the GENKI-4K dataset.

F. Smile Detection by Boosting Pixel Differences

In [07], the author has used the intensity differences between pixels in the grayscale face images as features. They adopt AdaBoost to choose and combine weak classifiers to form strong classifiers for smile detection. In their work, the weak classifier was defined based on the intensity difference of a pair of pixels.

Grayscale face images were normalized to catch a canonical face of 48×48 pixels, which was based on the manually registered eye positions. And they have adopted a fourfold cross- validation with similar number of "smile" and "nonsmile" samples into four groups of 1000 images. Histogram Equalization was used for extract intensity difference features from face images.

GENKI-4K database was used as dataset for the experiment which consists of 2162 different smile images and 1828 different nonsmile images. And they have provided 85% accuracy by examining 20 pairs of pixels and 88% accuracy with 100 pairs of pixels.

G. Efficient Smile Detection by Extreme Learning Machine

In [08], the author has proposed an efficient smile detection approach based on Extreme Learning Machine (ELM).

For the experiment Viola-Jones face detector was used to extract the face from original images. The bicubic interpolation was used to normalize the detected faces to a particular image size. For the face registration they have used a method of holistic flow-based face registration which consist SIFT flow computation and flow-based affine transformation to automatically align the detected faces. They have examined three different feature descriptors such as Local Binary Pattern, Local Phase Quantization (LPQ), and Histogram of Oriented Gradients (HOG) to extract features from registered faces. Those extracted features were input to the ELM classifier to predict the smile status of a given face.

The authors have used two different databases for the experiment. One database called MIX database has generated from four publicly available databases. MIX database has 1534 different smile and 2035 different non-smile face images. The other database was GENKI-4K database which contains 2162 different smile face images and 1828 different non-smile face images.

They have compared ELM with two benchmark classifiers such as Linear Discriminant Analysis (LDA) and Support Vector Machine (SVM). They have achieved 94. % of detection accuracy for the Mix database when LPQ combined with ELM and 88.2% of detection accuracy for the GENKI-4K database when using HOG + ELM.

H. Smile Detection using Multi-Scale Gaussian Derivatives

In [09], authors have used a methodology of the Multi-scale Gaussian Derivatives combined with Support Vector Machines to detect the smile faces as shown in the figure 4.

The OpenCV face detector was used for the face detection from the given face image. The detected face was normalized into 64×64 pixels. That normalized face image was sent into Half -Octave Gaussian Pyramid which is also called Multi-Scale Gaussian derivative (MGD) to extract the features. PCA was used to reduce the dimension. Those extracted features were used by soft Margin Support Vector Machine (SVM) for the classification.

They have used two different datasets for this experiment. GENKI-4k dataset which has 2162 different smile face images and 1828 non-smile face images was used for trained the SVM. To the purpose of validation, they have used Cohn-Kanade dataset. They have achieved a classification accuracy of 92.97%.

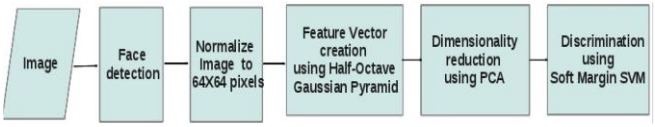

Fig. 4. Schematic of the approach

I. Facial Smile Detection using Convolutional Neural Networks

In this paper [10], they have solved the smile detection problem by proposing more efficient CNN architecture. To make a deeper network they have suggested utilization of small filters, in combination with Bach Normalization and ReLu. For the experimentation they have suggested GENKI – 4K database. By using small filters, they were able to speed up their performance with real time applications. However, they were able to achieve 95.08% accuracy for the smile detection using Softmax classifier for image classification.

III. COMPARISON

The Summary of experimental results based on previous methods as given in the Table II.

.

Table II- Product Comparison

Method used	**Classifier**	**Extract Features**	**Image Size & Pixels**	**Accuracy (%)**
LBP [1]	SVM		OD 59×65	90%
Canny Edge Detection [2]	-		-	99.4%
Gabor Filter [3]	SVM or GentleBoost	GEF, or BF, or EOH, or BF+EOH or LBP	8×8	White face - 97.5% Black face -90%
Simple Image Based Approach [4]	-	Feature based, Eigenface - based and Neural Network	-	-

Image Processing Technique [5]	-	HSV	-	89.5%
Deep Convolution Networks [6]	Recognition signals as supervision	Modified CNN	5×5	94.6%
Boosting pixel differences [7]	AdaBoost	Histogram Equalisation	48×48	20 Pairs of pixels – 85% 100 Pair of pixels – 88%
Extreme Learning Machine [8]	ELM classifier, LDA and SVM	Face - Viola-Jones face detector, Features - LBP, LPQ & HOG	-	Mix database + ELM + LPQ – 94% GENKI – 4K+ ELM + HOG – 88.2%
Multi-scale Gaussian Derivatives [9]	SVM	MGD	64×64	92.97%
BKNet [10]	Softmax	CNN	224×224 RGB	95.08%

IV. PROPOSED METHODOLOGY

We have proposed more efficient approach to smile detection. Face images captured in real-world scenarios are going to use in our study. In our approach, we are trying to achieve highest accuracy in smile detection using eye state by comparing with existing applications.

Evaluation Criteria

$$\text{Accurate Smile Detection Rate} = \frac{\text{No.of detected Smile faces}}{\text{No. of Test images}} \times 100$$

Moreover, it is fully automated smile detection with higher detection accuracy implemented by using GENKI – 4K database which enables the potential for real time application.

V. CONCLUSION

We have summarized ten research papers related to Smile detection with highest accuracy. Based on the comparison highest accuracy achieved by the Canny edge detector [2] which uses a multi-stage algorithm to detect wide range of edges in images. And it achieved 99.4% accuracy for smile detection when compared to other methodologies. Second highest was achieved by the Gabor Filter [3] which has 97.5% accuracy on images with white faces. It uses GEF, or BF, or EOH, or BF+EOH or LBP for the purpose of extract features. And also for classification SVM or GentleBoost plays important role for achieving higher accuracy. More importantly we have focused more on research papers with higher accuracy rate for smile detection other than the parameters like size of the algorithm, time requirement for training, etc. are ignored.

Furthermore, we have proposed a relatively simple smile detection methodology after studying all approaches above and came up with more efficient and accurate solution by using GENKI – 4K database. And to extract features we have proposed to use combination of HOG, SIFT/SURF and LBP. For the dimension reduction we have proposed to use PCA. In the future we would try to implement the above proposed method in MATLAB platform such that it is possible to produce accurate facial expression analysis.

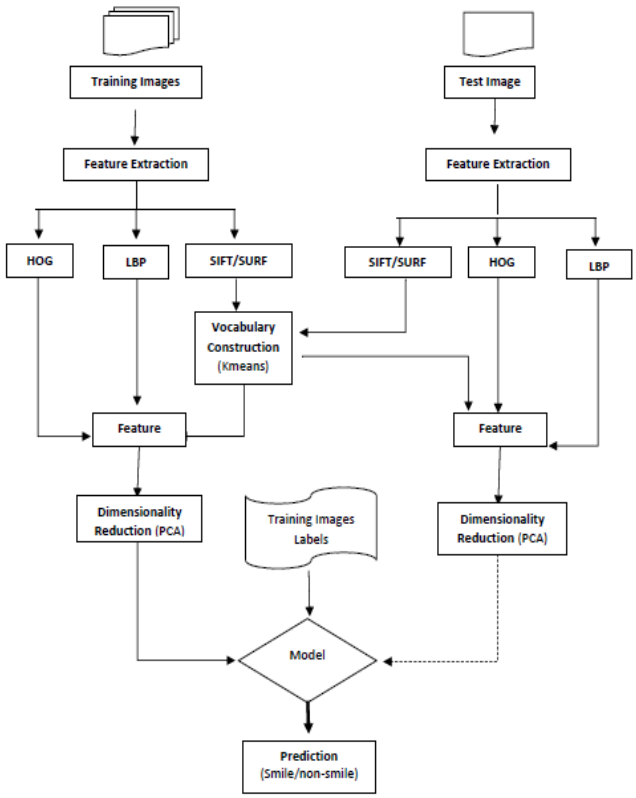

Fig. 5. Framework of Proposed Methodology

REFERENCES

[1] D Freire, Santana,M,C, and Déniz-Suárez,O "Smile Detection using Local Binary Patterns and Support Vector Machines." VISAPP (1). 2009.

[2] X Chen, W Cheng "Facial expression recognition based on edge detection" (IJCSES) Vol.6, No.2, April 2015.

[3] J Whitehill, G Littlewort, I Fasel, "Toward Practical smile detection" Article in IEEE Transactions on Software Engineering · November 2009.

[4] Praveen N, Pavithra K, Archana Yashodhar, Manjula "Smile detection : A simple Approach," International Journal of Advanced Research in Computer Science., May-June 2013.

[5] Vijayalaxmi, D.Elizabeth Rani "Eye State detecting using image processing Technique ," *American Journal of Engineering Research (AJER)*, 2015.

[6] Kaihao Zhang, Yongzhen Huang, Hong Wu, Liang Wang "Facial Smile Detection Based on Deep Learning Features" 3rd IAPR Asian Conference on Pattern Recognition, 2015.

[7] Caifeng Shan, "Smile Detection by Boosting Pixel Differences" *IEEE Transactions on Image Processing, Vol. 21, NO. 1, January 2012*.

[8] Le An, Songfan Yang, Bir Bhanu "Effcient Smile Detection by ExtremeLearningMachine" Center for Research in Intelligent Systems, University of California, Riverside, CA 92521, USA, 2015.

[9] Varun Jain, James L.Crowley "*Smile Detection Using Multi-scale Gaussian Derivatives*" 12th WSEAS International Conference on Signal Processing, Robotics and Automation, Feb 2013, Cambridge, United Kingdom. hal-00807362.

[10] Dinh Viet Sang, Le Tran Bao Cuong, Do Phan Thuan, "Facial smile detection using convolutional neural networks" October 2017.

[11] Vijayalaxmi, D.Elizabeth Rani, "Eye State Detection Using Image Processing Technique " American Journal of Engineering Research(AJER) 2015.

AUTHORS

First Author C.L.I.S. Fonseka is a Lecturer in Computer Science at University of Colombo School of Computing, Sri Lanka. She received her B.Sc. Honors in Computer Science (2017) from the Eastern University, Sri Lanka. Her research interests are in the field of Image Processing, Data Analytics, HCI and Machine Learning.

Second Author L.S Erandika, received her BSc. In Management and Information Technology degree from South Eastern University of Sri Lanka in 2016. Associate Member of Computer Society of Sri Lanka. Her area of interests are Image Processing, Machine Learning, E-Commerce, Business Intelligence and Big Data.

Third Author Sittampalam Sotheeswaran is a Senior Lecturer in Computer Science at the Department of Mathematics at Eastern University, Sri Lanka. He received his B.Sc. Honors in Computer Science (2008) and MPhil in Computer Science (2016) from the University of Jaffna, Sri Lanka. His research interests are in the field of Image Processing, Pattern Recognition, Computer Vision and Machine Learning.