Literature Review of Facial Expression Recognition using Local Binary Patterns, Local Directional Patterns algorithms and Voting Ensemble MetaClassifier

Kennedy Chengeta[1] and Prof Serestine Viriri[2]

University of KwaZulu Natal, Durban, South Africa
kennedychengeta@gmail.com,
WWW home page: http://users/~www.ukzn.ac.za
Phd in Computer Science-Computer Vision, April 2018

Abstract. Face and emotion recognition forms part of image analysis and human computer interfacing, and has been popularized due to applications in law enforcement, authentication and personal identification. This paper proposes a novel approach for recognizing facial expression from facial features using feature descriptors, namely Local Binary Patterns and Local Directional Patterns. This paper presents these two algorithms to detect seven basic facial expressions. The algorithms are compared in terms of accuracy on their own and a proposed algorithm where the 2 algorithms is used with a combined ratio of classifiers called Voting Classifier. The face is divided into several regions where the features are extracted. This research work consists of three parts, namely face representation, feature extraction into a single feature vector(histogram) and classification. Experiments performed on the Cohn-Kanade facial expression and Google-Set databases obtained a good expression classification rate 99% on the local binary pattern over the local directional pattern. The proposed solution included a combination of local binary pattern(LBP) and local directional patterns(LDP+LBP+Voting Classifier) as the feature extractor and voting classifier classification algorithm which is an aggregate classifier composed of k-Nearest Neighbor, decision trees and support vector machines. The results showed improved accuracy results as compared to other local binary pattern and local directional pattern alone.

1 Introduction

Facial behavior automation analysis, benefits fields like medicine, security, accounting, education and computing [3]. Face expression recognition has a wide variety of applications such as in identity authentication, access control and surveillance [1,2]. Faces influence or direct our emotional as well as social behaviors. A human face is not only used for recognition, but also to communicate ones emotions, intended actions and non visual iterations.The human face's emotions guides on a person's gender, attractiveness, age personality and other characteristics as well. For the deaf, sign languages use facial expressions to do grammar encoding in communication. The non-verbal communication signals which are driven by expressions account for more than the spoken words themselves. Indications for sickness, stress, deception and happiness are derived through micro expressions since they leak behavior control. Some of the non-verbal communication signals are voluntary and non voluntary and the latter convey ones personality. Based on Darwin's emotions theory, fear arises when one opens the eyes, raises eyebrows. Anger manifests when one has open eyes and raised nostrils[19]. One will be showing contept by looking the other way. Surprise is shown by opening both the mouth and eyes wide[19].

2 Literature Review-Facial Expressions

The key expressions include, sadness, joy, neutral, surprise, fear and anger[19]. Facial expression recognition algorithms research in personalised gaming has been successfully researched[10]. Chronic pain identification has been successfully identified using facial expression analysers in hospitals and places of work[5, 1]. Automated smile identification was also used in hospitals with focus on chronic patients and young babies using real time embedded solutions[3, 19]. Facial expressions for different emotions have been extracted using various feature extraction algorithms together with various supervised and unsupervised learning classifiers. Local binary patterns and local directional patterns, both local algorithms have been used successfully to identify all the different expressions[14, 16] with holistic algorithms like Principal component analysis achieving great accuracy[4]. Currently two approaches are used to extract facial features of images namely geometric and appearance based. The geometric extraction methods compute the locations and shape data using various facial parts example like nose, mouth, eye, distance between nose and the eye region. These facial parts are then combined into a feature vector which representing the geometry of the face[12]. The feature vector classification accuracy is affected by the dimensional properties and redundancy of some of the features. For appearance based algorithms, the extracted shape image parameters are used as input into the classification for instance using Mahalanobis or Euclidian distance[12]. Appearance based algorithms represent a face as multiple raw intensity images where a given image is made up of multi dimensional vector. Statistical algorithms are used to extract feature space from the images. Some of the features in the given feature vectors are redundant taking for instance non-discriminating features which bring complexity and reduced recognition accuracy rates[1][2]. The global-feature descriptors are also called holistic methods and the local methods indicate breaking down faces into smaller features like nose or the mouth[1][2].

2.1 Holistic and Local Algorithms

Holistic and local features are crucial for facial expression recognition[13][8]. Global features are extracted from the full face images and identify faces using global representations. Descriptions of the features are based on the whole image compared to local facial features only[9]. Major holistic algorithms include eigen faces, fisher faces , principal component analysis(PCA) as well as discrete cosine transform(DCT)[8]. Holistic algorithms are good because they is no information loss even after concatenation since they focus on specific points of interest which are not impacted by dimensional reduction[12]. Local features that include nose, mouse, ears are position as well as face orientation dependent[12]. Local features are extracted by Gabor wavelets, local directional pattern as well local binary patterns. The former's feature vectors are called spartial Local Gabor Feature Vectors (LGFV)[12]. Local Binary Patterns (LBP) are originated from texture analysis for face representation [14]. The LBP image is divided into subregions or blockes and features are extracted based on the LBP operator[14][12]. Parts-based approaches generalise better on previously unseen data and help to work around issues such as occlusion or when the shapes are less rigid (for instance an expression on a face)[12]. LBP and LDP give an added advantage to classification due to their in-variant nature when monotonic gray level images are used[14]. The extracted feature vectors derived from the expression images are concatenated histogram of feature vectors and this is trained and classified by recognised machine learning classifiers like support vector machines, k nearest neighbor and random forest classifiers and voting classifiers[14].

3 Local Binary Patterns

The local binary pattern (LBP) operator was used to describe local features and texture [13][16]. Different variants of it were then enhanced to do face recognition since it gave better accuracy. Grey scale conversion was done and the operator assigned labels to corresponding image pixels by thresholding the n by n neighbourhood of each pixel aligning it to the central pixel, where value of n was 3 for instance[13][16]. The local binary pattern (LBP) operator is defined as a gray-scale invariant texture measure and is derived from a general definition of texture in a local neighborhood [13][14]. Originally, the LBP texture descriptor [15], was computed in a pixel level basis using a 3 x 3 kernel, thresholding the surroundings of each pixel with the central pixel value and taking the result as binary. Local Binary Pattern (LBP) is a 3 x 3 matrix in which eight neighboring pixel intensity is compared with the intensity of the central pixel, resulting negative values are encoded with 0 values, and the positive values are encoded with 1[12]. Local Binary Pattern (LBP) [13][14] and its variants are used in recent times as a feature descriptor for facial expression representation. The most important properties of LBP features are their tolerance against illumination changes and their computational simplicity [14]. LBP feature vectors are represented as histograms. To generate a histogram LBPH or Local Binary Pattern Histogram, the central pixel value of a 3 x 3 neighborhood is considered as a threshold value[14]. Then it labels the result of applying the threshold on the surrounding pixels as a binary number. The binary numbers are converted to a local value by converting it to a decimal number based on the weights. The histogram is computed separately in each of the m regions thus giving a total or m histograms which are then concatenated. Firstly, the image's grey scale value is calculated and the LBP operator will give each pixel a label using n by n neighbourhood where n can be 3 for instance thresholding them against the central pixel.

3.1 LBP variants

Various LBP variants were proposed as LBP extensions and modifications were done to improve facial recognition accuracy[13][16]. Some of the variants used a combination of strong algorithms like SIFT with its interest region descriptors and LBP algorithms in a combination[13][16]. The combination of the LBPs with Gabor features were proposed as well including rotated local binary pattern algorithms [13][16]. The regular local binary pattern (LBP) histogram has eight neighbors with a radius of length one where each pixel is compared to all its 8 neighbours[13][16]. This original version of the LBP operator was based on three by three pixel image blocks with a threshold central value. Eight pixels formed the neighborhood to give a total of 256 separate labels. Prominent among the variants include Multiprocessing split LBP which divides the work of regular LBP over multiple processes and passes only the working range to each process[13][16].

Uniform local binary patterns(ULBP) are patterns that have both circular 1-0 and 0-1 transitions. The patterns 00000000 and 11001111 are uniform, and pattern 10101101 is not uniform. Selecting only uniform patterns contributes to both reducing the feature vector length (LBP histogram) and improving the performance of classifiers using the LBP features[13][16]. Uniform LBPs can also be applied to obtain rotation in-variance. There are several methods for performing LBP histogram comparison. which are histogram intersection, log-likelihood, and Chi-square test.For uniform LBP which also extends the original LBP, only patterns with up to 2 transitions from 0 to 1 and the reverse are included in the calculation[13][16]. Local Binary Pattern which is symmetric (CS-LBP) is a modified LBP algorithm that brings a long feature vector with 256 or more dimensions. Its more robust on flat images than the original LBP algorithm[13][16] and is also enhanced by other algorithms like SIFT descriptors[1][2]. The histograms are made up of the key gradient details of the image and also the gray level variances are in a symmetrically opposing mode[1][2]. Rotated Local Binary Patterns(RLBP) LBP only consider the signs of the differences to compute the final descriptor[13][16]. The information related to the magnitude of the differences is completely ignored. The magnitude provides evidence that has been utilized to increase the discriminative power of the operator[1][2]. The multiprocessing LBP variant works by dividing the input image into x horizontal slices which then spawns y processes. Each process gets as input the entire image and the bounds of the slice that it should work on[1][2].

4 Local Directional Patterns

Local directional pattern algorithms focus on edges in certain directions and based on that the more prominent regions are picked up to generate the Local Directional Pattern[13][16]. The top k-directional bits are given the value of 1 and the remaining bits take the value of zero. The image was then split into small regions where histograms were combined into an LDPv descriptor. LDP encodes the directional information of the face images by involving the face image with the compass mask[13][16]. The compass mask was used to retrieve the edge response values in the neighborhood's 8 different directions. The information was encoded with directional numbers which was used to create the LDN code, corresponding to an LDN facial image. These local featured were then concatenated from various blocks to form feature vector and later used as a face descriptor to distinguish between the face images. Different experiments were done where the descriptors accuracy where measured given the noise, different lighting conditions and varying time conditions and the LDP descriptors exhibited robust resistance to noise.[13][16].

Local directional pattern encodes the directional information of the faces textures producing a more compact discriminative code than current present methods. It uses compass masks that extract directional information, and compute the structure of each micro-pattern and encode such information using the prominent direction indices. This enables one to separate between similar patterns with different intensity transitions[13][16]. The face is formed into key smaller regions from which LDN features are retrieved. The descriptor performs consistently under noise, illumination, expression, and time lapse variations. Encoding is applied using the more prominent edges. It also includes edge detection algorithm called Kirsch algorithm[13][16]. The (Kirsch) compass masks are convoluted using the original image to extract the edge response images. The Kirsch operator or Kirsch compass kernel is a non-linear edge detector that finds the maximum edge strength in a few predetermined directions[13][16]. The Kirsh response values have different importance and weights in different directions as well[16]. Since edge responses are more stable than intensity values, LDP pattern provides the same pattern value even in the presence of noise and non-monotonic illumination changes [16].

5 Methods and Techniques

5.1 Proposed System

The main objectives of this dissertation included an in-depth analysis of facial expression recognition algorithms. The research analysed the impact of such facial expressions analysis on human character. The research developed and implement a software framework for facial feature detection which combines face and feature detection methods developed and identifies human characteristics using

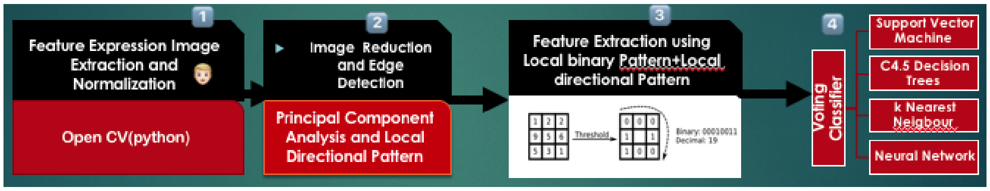

Fig. 1. Facial Expression Recognition Design

local directional patterns and local binary patterns. Our Proposed System is divided into the following modules: The facial expression recognition process was grouped into four interrelated phases or steps namely, face expression detection, normalization, feature extraction and facial expression recognition.

Tools included python, OpenCV libraries as well MongoDB(Storage), AngularJS and java for the user interface tool for the testing and experiments. OpenCV library extensions were used for face detection and normalization in python.

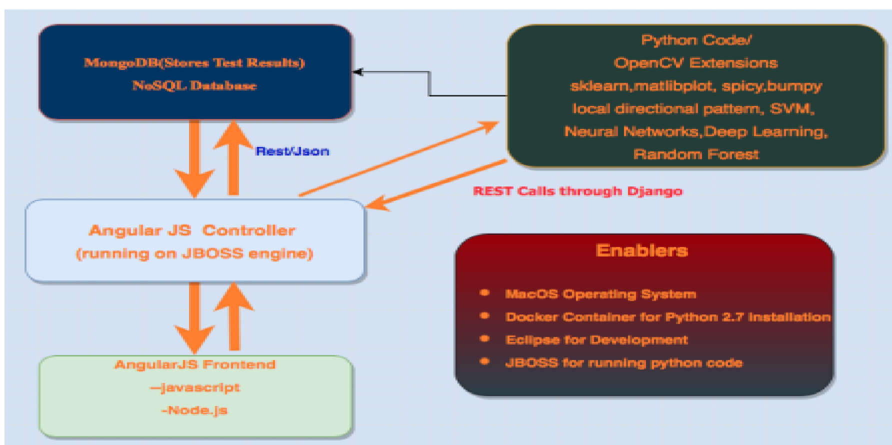

Fig. 2. LBP and LDP Design For Facial Expression Recognition

5.2 Facial Expression Database and Detection

We collected the face images from the Extended CK+ and Google-Set dataset images stores. The data-set images depict different facial expressions reflecting various emotions. The given emotions range from fear, sadness, happiness or being joyful, disgust as well as being neutral.

5.3 Face Alignment, Normalization and Dimensional reduction

Pre-processing the data was done before extraction of the features. This eliminated the noise from the data, and re-sizing of the data was done so that all images had same dimensions. The images were initially converted into grey level images. Kirsch masking at 45 degrees rotation was used to extract edge responses in eight different directions. During normalization, the image went through standardization using size, pose as well as illumination as key parameters in relation to the image. Dimensional reduction was implemented using Principal component analysis (PCA) Agorithm.

5.4 Feature Extraction

Key local features were extracted from the images as feature vectors to form a histogram. The feature vectors were used to measure accuracy and similarities between images using machine learning classifiers. To extract the important features from the face image, the face features were termed as face descriptors. Local directional number pattern(LDN) and local binary pattern(LBP) feature extraction algorithms were used. A given LBP X,Y operator is mathematically defined as

$$LBP_(X,Y)(x_c, y_c) = \sum_{X-1}^{X=0} s(g_x - g_c)2^X. \tag{1}$$

The LBP equation shows that the signs of the differences in given neighborhood is represented as a X -bit binary number which gives dual X unique values for a given LBP code. The gray image texture will be represented by 2 X-bin discrete LBP codes. The proposed local directional number pattern(LDN) represented a 6-bit binary code which was mapped to every given pixel of the input image. Local binary patterns(LBP) were used to determine the local features in the face. Feature vectors were extracted in a matrix originally of size n * n with n being 3 where the values were baselined with the central pixel values to produce a binary pattern. An expression image was divided into small regions from which LDPh histograms were extracted and concatenated into one LDPh descriptor.

$$LDP_h(\sigma) = \sum_{M}^{r=0} \sum_{N}^{r=0} f(LDP_k(r,c),\sigma). \tag{2}$$

5.5 Classification

Several machine learning technique, namely Random Forest, Neural Networks, k-Nearest Neighbor, 4.5 decision tree classifiers and Support Vector Machines were examined. They were also combined in the form of a combined classifier called Voting Classifier. EnsembleVoteClassifier is a meta-classifier that combines various machine learning classifiers for classification via majority, weighted or probability voting[13]. The weighted majority vote involves combining a weight wj with classifier Dz, where Dz represents unique class labels and then assigns the weights for instance 0.35, 0.2, 0.45 which totals a probability of 100%

$$\hat{q} = \arg\max_i \sum_{j=1}^{m} w_j \chi_A \big(C_j(\mathbf{x}) = i\big), \tag{3}$$

$$\hat{q} = \arg\max_i [0.35 \times i_0 + 0.2 \times i_0 + 0.45 \times i_1] = 1 \tag{4}$$

where A is the characteristic function [Cj(x)=iA], and A is the set of unique class labels.

6 Experiment and Results

The research performed around fifty experiments to evaluate the accuracy of the proposed facial expression recognition algorithms using the Google dataset and the CK+ facial data-set. Normalisation and cropping of the images' eyes, mouth and ears was archieved. In the experiments, every image was partitioned into a grid of 10 by 10 and 14 by 14. A combined local binary pattern(LBP) and local directional pattern(LDN) feature extractor together was used. The two algorithms' histograms were then fed into several classifiers namely neural networks, k-Nearest Neighbor, random forest and support vector machines and the voting classifier. The facial image of the one face was compared with the feature extraction image which was stored in the database and a match was recognized between two facial images using a classifier. Two datasets were selected for testing, namely CK+ dataset. The Cohn-Kanade (CK) AU-coded expression dataset encompassed a century of university students ranging between 18 and 30. The facial Google dataset was also used. For the CK+ dataset, almost two thirds were female with about fifteen percent being African American. The algorithms were implemented in python and then tested on the above two face databases. Image sequences were then digitized into 640 * 480 pixel arrays of gray scale frames to generate the feature vectors. In the training phase, the images were extracted in different batches of 100, 1000 and then 10 000. These facial images were feature extracted using the LBP and LDP feature extractors and stored in feature vectors. The mean values of the feature vectors formed histograms which were then fed into the classifiers. The mean vector and the variance vector were used to exclude the features that have wayward variances. In the classification phase, test images' feature vectors were stored from the feature extraction activities of local binary patterns(LBP) and Local Directional Patterns(LDP).

6.1 Ensemble Voting Classifier

For all the classifications a full cross validation training/test of 60% to 40% ratio was implemented in all the image classifiers. The results are explained in Figure 2. The voting classifer combination for kNN, support vector machines(SVM), random forest(RF) and C4.5 decision tree(DT) was done with a 5:2:1:2 ratio, 3:1:3:4 ratio,7:1:1:1 ratio, 6:1:2:1 ratio and 4:1:4:1 ratio. Once the tests were run the accuracy was then averaged on the above 5 classifiers. The modal accuracy however was the one where kNN had high ratio which is 7:1:1:1 and this gave an accuracy for 99.45% for the CK+ database and 99.38% for the Google data-set. The Opponent Color Local Binary Pattern algorithm was not able to be used for comparison in this research to unavailability and time constraints to get the color images of the CK+ database and Google dataset. The classifier algorithms implemented include support vector machines(SVM), random forest(RF), C4.5 decision tree(DT) as well as the voting classifier(VC). Five tests with different ratios for the voting classifier were done for each dataset. This was to ensure an average had to be calculated for the 5 different ratios.

Fig. 3. Voting Meta Classifier

classifier	class a	class b	class c	class d	class e
KNN	x_1 * 0.5	x_1 * 0.3	x_1 * 0.2	x_1 * 0.7	x_1 * 0.4
SVM	x_2 * 0.2	x_2 * 0.1	x_2 * 0.1	x_2 * 0.1	x_2 * 0.1
Random Forest	x_3 * 0.2	x_3 * 0.3	x_3 * 0.1	x_3 * 0.1	x_3 * 0.4
c4.5 Decision Tree	x_4 * 0.1	x_4 * 0.3	x_4 * 0.6	x_4 * 0.1	x_4 * 0.1
Weighted average	0.79	0.67	0.23	0.92	0.86

6.2 Classification Results of the full datasets

The next table shows for Google Dataset as well as the CK+ Dataset facial database for the local facial descriptors in the two tables following below. The next table shows results for Extended CK+ Data-set facial database for the local facial descriptors. The following algorithms were used namely basic LBP algorithm with 16,2 and 8,2 factors respectively. The other variants tested include the symmetric CS-LBP, the Enhanced LBP, local ternary LTP, the rotated RLBP and the combined algorithm with LDP and ELBP algorithm. Various classifiers were used namely k-Nearest neighbor, Adaboost, Random Forest(RF) and a Voting Classifier.

6.3 Facial Expression Results: Small Datasets-Googleset

The facial expressions' results were based on the following emotions namely Anger, Disgust, Fear, Happy or joyness, Sadness,neutral, Surprise plus Contempt. The results are summarised in the following namely figure 6 and 7 based on the Googleset data. The google set data was a small dataset and experiments were carried out from 100 images up to 1500 images though the best results picked were for 75 images as well as for 181 images all achieving an average well beyond 99 percent. Due to the small dataset the 75 images were only able to pick 4 emotions hence the reason the 3 emotions were all showing zero accuracy.

For the Googleset large dataset test with the best accuracy a 1500 images were trained and using a two thirds classification ratio, 1000 images were classified into the 7 emotions.

GoogleSet Data	kNN+	AdaBoost	RF	Voting Classifier		CK+ Data	kNN+	AdaBoost	RF	Voting Classifier
$LBP_{8,2}$	93.09%	97.12%	97.16%	97%		$LBP_{8,2}$	92%	97%	97.6%	97%
$LBP_{16,2}$	94.26%	98.03%	97.28%	95.99%		$LBP_{16,2}$	93%	98%	97%	96%
$CS\text{-}LBP_{8,2}$	93.51%	97.98%	97.33%	97.86%		$CS\text{-}LBP_{8,2}$	93%	96%	96%	98.56%
$CS\text{-}LBP_{16,2}$	94.05%	96.92%	96.08%	98.31%		$CS\text{-}LBP_{16,2}$	92.5%	97.2%	96.08%	97.1
$ELBP_{16,2}$	91.3%	91.02%	96.29%	97.09%		$ELBP_{8,2}$	89%	87.2%	96%	96.9%
$ELBP_{16,2}$	88.21%	87.12%	95.75%	96.91%		$ELBP_{16,2}$	85%	85%	96.5%	96.1%
$LTP_{8,2}$	89.71%	96.31 %	96.77%	96.74%		$LTP_{8,2}$	85.7%	95.1 %	97%	95.54%
$LTP_{16,2}$	88.3%	96.44 %	97.42%	97.24%		$LTP_{16,2}$	86.27%	95 %	97%	96.33%
$RLBP_{8,2}$	86.1%	97.01 %	96.8%	97.48%		$RLBP_{8,2}$	85.1%	95.4 %	96.8%	96.61%
$RLBP_{16,2}$	84.3%	97.21 %	97.09%	97.64%		$RLBP_{16,2}$	85.3%	95.21 %	96%	97.66%
$LDP+ELBP_{8,2}$	94.18%	97.81%	96.92%	99.13%		$LDP+ELBP_{8,2}$	93.1%	98.2%	97.88%	99.23%
$LDP+ELBP_{16,2}$	94.78%	97.99%	97.02%	99.03%		$LDP+ ELBP_{16,2}$	94.3%	97.45%	98.39%	99.27%

Fig. 4. Classifier for CK+ and GoogleSet Dataset

Fig. 5. Google Set Dataset Facial Expression Recognition Small Dataset

	precision	**recall**	**f1-score**	**support**
anger	1	1	1	16
disgust	0.98	1	0.99	44
fear	1	0.92	0.96	12
happy	1	1	1	3
disgust	0	0	0	0
fear	0	0	0	0
happy	0	0	0	0
avg/total	0.99	0.99	0.99	75

Confusion Matrix									
anger	[[**16,**	0,	0,	0,	0,	0,	0]		
disgust	[	0,	**44,**	0,	0,	0,	0,	0]	
fear	[	0,	1,	**11,**	0,	0,	0,	0]	
happy	[	0,	0,	0,	**3,**	0,	0,	0]	
disgust	[	0,	0	0,	0,	0,	0,	0]	
fear	[	0,	0	0	0,	0,	0,	0]	
happy	[	0,	0,	0,	0	0,	0,	0]]	

Fig. 6. Google Set Dataset Facial Expression Recognition

	precision	**recall**	**f1-score**	**support**
anger	1	1	1	16
disgust	1	1	1	44
fear	0.92	1	0.96	12
happy	1	1	1	39
neutral	1	0.94	0.97	18
sadness	1	1	1	23
surprise	0.97	0.97	0.97	29
avg / total	0.99	0.99	0.99	**181**

Confusion Matrix								
anger	[**16,**	0,	0,	0,	0,	0,	0],	
disgust	[	0,	**44,**	0,	0,	0,	0,	0],
fear	[	0,	0,	**12,**	0,	0,	0,	0],
happy	[	0,	0,	0,	**39,**	0,	0,	0],
neutral	[	0,	0,	0,	0,	**17,**	0,	1],
sadness	[	0,	0,	0,	0,	0,	**23,**	0],
surprise	[	0,	0,	1,	0,	0,	0,	**28**]

6.4 Facial Expression Results for Large Datasets-CK+

CK+ database experiments involved tests with a small dataset of 500 up to a range of 4000 images as well. The accuracy in both scenarios was impressive averaging 99%+ in the larger dataset and 98% in the smaller dataset. Only 6 emotions were measured. As indicated in the overral classification table the Voting Classifier showed supreme accuracy compared to the individual classifiers. Sadness was the best accuracy emotion whilst neutral showed few individuals who were incorrectly classified.

LDP+ELBP with a Voting Classifier achieved a high accuracy of 99.23 percent and 99.13 on the Google-Set and CK+ databases respectively which was a high accuracy as opposed to the other Local Binary Pattern Variants. The tests results accuracy percentages were averages of running against small which is around 100 images, medium data-sets of 1 000 images and 10 000 images which

Fig. 7. CK+ Dataset Facial Expression Recognition Large dataset

	precision	**recall**	**f1-score**	**support**
anger	0.98	1	0.99	317
disgust	0.98	0.99	0.99	214
fear	1	0.99	0.99	230
happy	0.99	0.99	0.99	276
neutral	1	0.99	1	214
sadness	1	0.98	0.99	149
avg/total	0.99	0.99	0.99	**1400**

	Confusion Matrix						
anger	[[**317,**	0,	0,	0,	0,	0],	
disgust	[	2,	**212,**	0,	0,	0,	0],
fear	[	2,	0,	**228,**	0,	0,	0],
happy	[	1,	3,	0,	**272,**	0,	0],
neutral	[	0,	1,	0,	1,	**212,**	0],
sadness	[	0,	0,	1,	2,	0,	**146**]]

Fig. 8. CK+ Dataset Facial Expression Recognition small dataset

	precision	**recall**	**f1-score**	**support**
anger	0.99	1	0.99	70
disgust	1	1	1	44
fear	1	1	1	32
happy	1	1	1	40
neutral	1	1	1	36
sadness	1	0.96	0.98	28
avg/total	1	1	1	250

	Confusion Matrix						
anger	[**70,**	0,	0,	0,	0,	0]	
disgust	[	0,	**44,**	0,	0,	0,	0]
fear	[	0,	0,	**32,**	0,	0,	0]
happy	[	0,	0,	0,	**40,**	0,	0]
neutral	[	0,	0,	0,	0,	**36,**	0]
sadness	[	1,	0,	0,	0,	0,	**27**]

is classified large data-sets. The results show the large or small data-sets were giving high accuracy for Local Binary Patterns and Local Directional Patterns with Voting classifier scenario irrespective of size compared to other algorithms. The Local directional pattern was used to smoothen images with Kirsch algorithm and remove the noise from the edges. The Google dataset whilst it had high accuracy also showed very high accuracy results on the combined accuracy results of Local Binary Pattern and Local directional pattern with the voting classifier. The accuracy ith the Voting Classifier also was 99.23 percent which is the highest with all the feature extractors and classifiers as well

7 Conclusion.

The paper proposed using local binary patterns and local directional patterns for feature extraction in facial expression classification scenarios. Based on the results from the several experiments it shows the Local Binary Pattern and Local Directional Pattern with a Voting Classification Algorithm of 70% kNearest Neighbour improves facial emotions accuracy. The classifier with better results was the Voting Classifier and it showed consistent results for low number of images from around 100 and also for medium range of images which is around 1000 as well as for huge number of images ranging around 10 000 facial images as well. Although higher facial expression classification rates were achieved by the proposed methods, there is still some issues which should be furthered addressed such as finding the optimal threshold for each database automatically, or applying the proposed algorithm with other features, in the purpose of improving the recognition rate. Therefore, the hybrid LDP+ELBP Feature Extractor with a Voting Classifier operators presented in this paper show improved accuracy for facial emotion recognition over and above the uniform LBP or LDP on its own. Happiness in all the cases recieved the highest classification accuracy for both google dataset and the CK+ dataset as well ranging from around 99 to almost 100 percent as well.

8 Future Work

Future Improvements include use of Deep Learning in feature extraction and classification in facial expression classification. Other proposed future work activities include using CUDA and other parallel programming interfaces in facial recognition. The research also proposes further work on a combined feature extraction involving the new deep learning algorithms and local binary patterns as well as local directional patterns.

References

1. M. S. Aung, S. Kaltwang, B. Romera-Paredes, B. Martinez, A. Singh, M. Cella, M. Valstar, H. Meng, A. Kemp, M. Shafizadeh, et al.: The auto- matic detection of chronic pain-related expression: requirements, challenges and a multimodal dataset, Transactions on Affective Computing, 2015.
2. P. Pavithra and A. B. Ganesh: Detection of human facial behavioral ex- pression using image processing,
3. K. Nurzynska and B. Smolka, Smiling and neutral facial display recognition with the local binary patterns operator: Journal of Medical Imaging and Health Informatics, vol. 5, no. 6, pp. 13741382, 2015-11-01T00:00:00.
4. Andrew J Calder and A.Mike Burton and Paul Miller and Andrew W Young and Shigeru Akamats: A principal component analysis of facial expressions, Vision Research, "41",number "9",pages:"1179 - 1208", "2001"
5. M. S. H. Aung and S. Kaltwang and B. Romera-Paredes IEEE Transactions on Affective Computing, The Automatic Detection of Chronic Pain-Related Expression: Requirements, Challenges and the Multimodal EmoPain Dataset,2016
6. P. Lemaire, B. Ben Amor, M. Ardabilian, L. Chen, and M. Daoudi, Fully automatic 3d facial expression recognition using a region-based approach, in Proceedings of the 2011 Joint ACM Workshop on Human Gesture and Behavior Understanding, J-HGBU 11, (New York, NY, USA), pp. 5358, ACM, 2011.
7. C. Padgett and G. W. Cottrell, Representing face images for emotion clas- sification, Advances in neural information processing systems, pp. 894900, 1997.
8. P. Viola and M. J. Jones: Robust real-time face detection, Int. J. Comput. Vision, vol. 57, pp. 137154, May 2004.
9. M. S. Bartlett: Face image analysis by unsupervised learning, vol. 612.Springer Science and Business Media, 2012.
10. P. M. Blom, S. Bakkes: C. T. Tan, S. Whiteson, D. Roijers, R. Valenti, and T. Gevers: Towards personalised gaming via facial expression recognition, in Proceedings of Tenth AAAI Conference on Artificial Intelligence and Interactive Digital Entertainment, AAAI Press (Palo, 2014), 2014.

11. S. Biswas, K. Bowyer, and P. Flynn: A study of face recognition of identical twins by humans, in Information Forensics and Security (WIFS), 2011 IEEE International Workshop on, pp. 16, Nov 2011.
12. T. Gernoth, A. Goossen, and R.-R. Grigat: Face recognition under pose variations using shape-adapted texture features, in Image Processing (ICIP), 2010 17th IEEE International Conference on, pp. 45254528, Sept 2010.
13. C. Longmore, C. Liu, and A. Young: The importance of internal features in learning new faces, The Quarterly Journal of Experimental Psychology, vol. 68, no. 2, pp. 249260, 2015.
14. X. Zhao and S. Zhang: Facial expression recognition based on local binary patterns and kernel discriminant isomap, Sensors, vol. 11, no. 10, pp. 9573 9588, 2011.
15. H. Dikkers, M. Spaans, D. Datcu, M. Novak, and L. Rothkrantz: Facial recognition system for driver vigilance monitoring, in Systems, Man and Cybernetics, 2004 IEEE International Conference on, vol. 4, pp. 37873792 vol.4, Oct 2004.
16. A. Ramrez Rivera, R. Castillo, and O. Chae: Local directional number pattern for face analysis: Face and expression recognition, Image Process- ing, IEEE Transactions on, vol. 22, pp. 17401752, May 2013.
17. Nagpal, Shruti and Singh, Maneet and Singh, Richa and Vatsa, Mayank. (2015). Regularized Deep Learning for Face Recognition With Weight Variations. IEEE Access. 3. 1-1. 10.1109/ACCESS.2015.2510865.
18. Anastasia Ioannidou, Elisavet Chatzilari, Spiros Nikolopoulos, and Ioannis Kompatsiaris. 2017. Deep Learning Advances in Computer Vision with 3D Data: A Survey. ACM Comput. Surv. 50, 2, Article 20 (April 2017), 38 pages. DOI: https://doi.org/10.1145/3042064
19. , Turabzadeh, Saeed and Meng, Hongying and Swash, Rafiq M. and Pleva, Matus and Juhar, Jozef,"Facial Expression Emotion Detection for Real-Time Embedded Systems",Technologies,VOLUME = 6,YEAR = 2018,NUMBER = 1,ARTICLE NUMBER = 17,URL = http://www.mdpi.com/2227-7080/6/1/17,ISSN = 2227-7080,DOI = 10.3390/technologies6010017