NUML International Journal of Engineering and Computing
Volume: 4 Issue 1
https://numl.edu.pk/journals/nijec
Print ISSN: 2788-9629
E-ISSN:2791-3465
DOI:https://doi.org/10.52015/nijec.v4i1.92

A Novel Framework for the Accuracy Enhancement of Facial Expression Recognition System

[a]Naila Batool, [b]Muhammad Tahir, [a]Hamid Nawaz

[a] Department of Software Engineering, Govt. College University, Faisalabad, Pakistan
[b] School of Electronics and Information Engineering, Changchun University of Science and Technology, China

Submitted	Revised	Published
04-Aug-2025	26-Oct-2025	30-Oct-2025

Abstract

Facial Expression Recognition has become a promising field for more natural interactivity with computing devices and machines and has become the focus of attention for many research scholars over the past decade. Newly developed facial expressions recognition methods focus on neutral expression or six expressions used in most state-of-the-art methods. Accuracy is the main problem in the face recognition results. The problem that must be tackled is the optimization of the expression recognition algorithm i.e. to detect, isolate and correctly translate one of the major expressions of the human face with accuracy targeted towards 100%. This work will try to improve the accuracy of recognizing facial expression by using Histogram of Oriented Gradients (HOG) and Local Ternary Pattern (LTP).

Keywords: facial expressions, histogram of oriented gradients, local ternary pattern

Corresponding Author: Naila Batool (nailabatool93@gmail.com)

1. Introduction

Human correspondence has two fundamental viewpoints: verbal (sound-related, for example, talking, singing and once in a while manner of speaking and non-verbal (visual, for example, non-verbal communication, gesture-based communication, paralanguage, contact, eye to eye connection, or the utilization of content. The nuclear units of verbal correspondence are words and for non-verbal marvels like outward appearances, body developments, and physiological responses [1].

Facial Expression Recognition has become a promising field for more natural interactivity with computing devices and machines and has become the focus of attention for many research scholars over the past decade. FER is a non-intrusive way of communication and is a potential way of data input to a computing device without the need of uttering any words or typing any commands.

The nature and speed of development in the technological field especially in computing and interactivity has grown tremendously in the past decade. The requirements of potential consumers a decade ago have been met beyond their expectations and more explorations and innovations have generated quite more needs and wants for the end user. Among these innovations [3], the speciality of interactivity between human and machine have been of keen interest among technologists and it has given programmers and developers a good share of tough time to deal with. From touch interactions to voice control algorithms, from gestures translation to face recognition [2], all of these software solutions are in practice in gadgets, smartphones and daily life smart home and automation solutions.

Concentrating on nonverbal correspondence uncovers that 55% of an individual's enthusiastic or purposeful data is passed on through outward appearances. As of late, looks into passionate examinations have made incredible accomplishments. On one hand, improvement of neuroscience [4] and intellectual science will push the advancement of passionate investigation. Then again, specialized development in PC vision and AI makes applications identified with enthusiastic investigation accessible to the general population. As a significant subfield of enthusiastic investigation, looks into Facial Expression Recognition (FER) grow rapidly too.

The most generally utilized set is maybe human widespread outward appearances of feeling which comprises six essential articulation classifications that have been demonstrated to be unmistakable crosswise over societies Figure 1. Basic facial expressions are six, including happiness, anger, sadness, disgust, surprise and fear [5]. These appearances, or facial setups have been perceived in individuals from broadly disparate social and social foundations, and they have been watched even in the essences of people brought into the world hard of hearing and visually impaired.

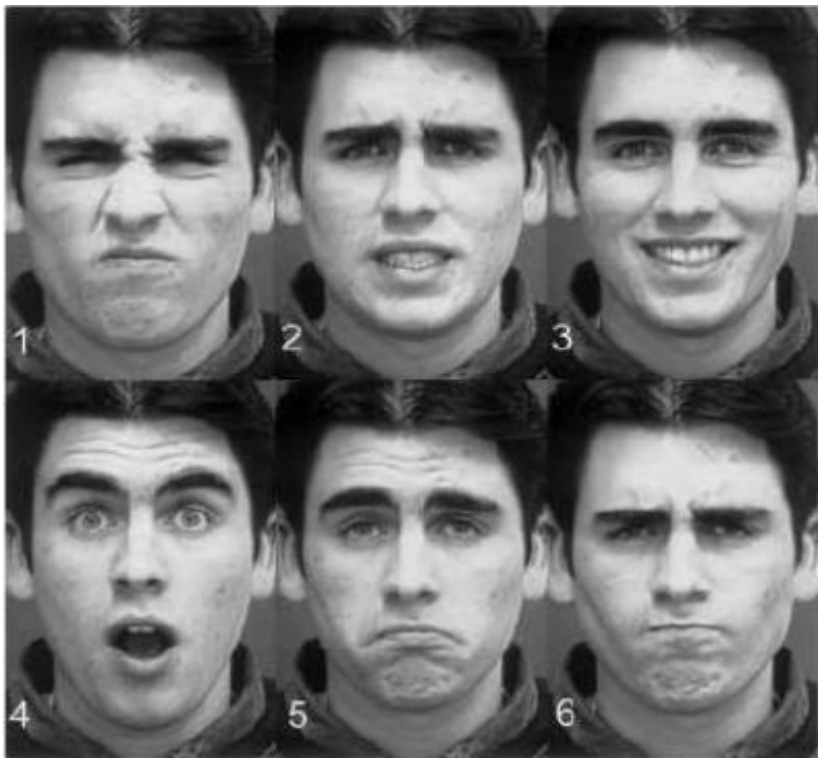


Figure 1: Basic facial appearance phenotypes

The aim of our research is to develop a novel method for the recognition of these expressions. Accomplishing the automatic recognition of facial expressions will encourage growth of various fields (e.g. computer science technology, security system technology, and medical disease diagnosis technology, human and computer interaction systems, driver fatigue predication, discomfort and sadness recognition, and mental tension monitoring system). Due to the importance and significance of medical and cognitive scientists, facial expressions recognition research is a multi-disciplinary research area of interest. Surely, this area of research made huge progress because of improvements in related research sub- areas, such as image/video-based face detection, recognizing and tracking of human face, and the ongoing research and developments in the field of machine learning, such as classification and regression (supervised learning) and feature detection and extraction. Though, correct recognition of facial expressions is still an interesting and challenging problem because of its complexity, refined facial expression changes and different scale, translation and rotation changes of the human head. Newly developed facial expressions recognition methods focus on neutral expression or six expressions used in most state-of-the-art methods.

Some state-of-the-art methods use landmarks selection on faces and only extract features from the land-marks location; this is a common method for facial expression recognition. The problem of facial expressions recognition is like most of the other object recognition techniques, first phase of facial expression recognition or face recognition system is face detection, the face is detection using the cascading object detector method, and the coordinates of the face region on the image is obtained, the coordinates are used to crop the face region from images automatically. In the second phase facial features are extracted and labeled, in the last phase a classifier is trained by the obtained features and labels from the training images; this trained classifier is used to predict a class (expression) for the new testing image/frame.

While it may seem quite difficult to achieve at first glance, FER is a subset of pattern recognition problems in mathematics and is currently under doing research and development all over the world. The problem that has to be tackled is the optimization of the expression recognition algorithm i.e. to detect, isolate and correctly translate one of the major expressions of the human face with accuracy targeted towards 100%. This in itself is a difficult task to achieve as the facial structure, size, features, color, illumination, distance and clarity of the face all affects the algorithm's performance and accuracy.

To assess and ease these issues, in the past multiple proposed algorithms have been coined, all targeted towards achieving the perfect classification and analysis of expressions on the human face. The literature review highlights those issues.

The objectives of our research are to propose enhanced solutions to improve the quality of facial expressions recognition and induction of classifiers in random forest for more precise results. The aim of our research is to develop a novel structure for acknowledgment of the outward appearances with the high precision by consolidating two strategies i.e., Histogram of Gradient (HOG) and Local Ternary Pattern (LTP).

2. Background Study

Extract features using a two-dimensional mesh, known as Potential Net. Every one of the strategies which are said above, are considered as holistic features, since they are connected to the entire image structure. Local feature

is a type of feature, which concentrates on a small area. The simplest idea is to use the sub-windows of the image as a local feature: for example in figure 1.2. [6].

Analyzed that PCA has been a well-known technique in the past and was based on the eigen decomposition of images. A detection system based on eigen decomposition, has been described which is called principal component analysis (PCA) [7]

[8] The Eigenface coefficients can be used as features and recently, their extension as defined here. Tensor face that showed a promising selection of features. [9] combined the Active Shape Model (ASM) with the Gabor filter; and: [10] used Facial Animation Parameters which are based on Active Shape Model.

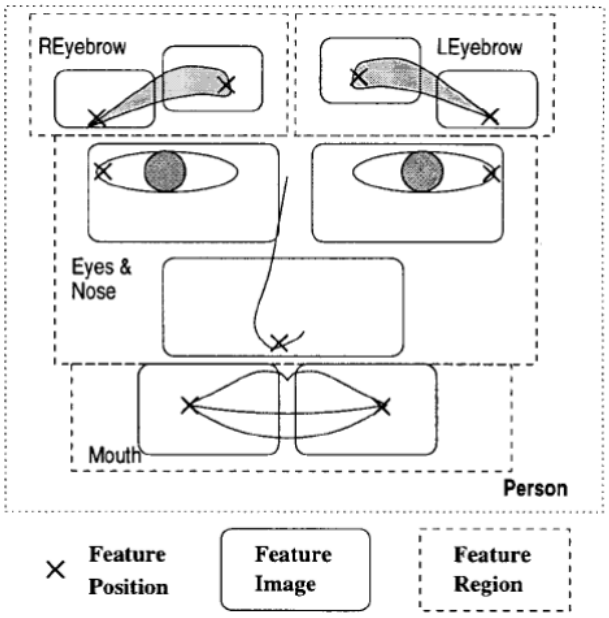


Figure 2: Scheme of Facial Recognition and Features

Facial Expression Recognition is basically a subset of Facial Recognition and has a major effect on the overall speed and performance of the architecture of the algorithm. For face detection, an algorithm can be trained to see skin color, head shape or the appearance of a human face in general as visual cues programmatically but among these the appearance cue is the most well-performing. At the detection level of the face, many tricks are used, for example, the motion (video), and skin color, shape of the head / face and facial appearance. The most effective face detection algorithms based on their appearance [10].

Proposed approach: a new method is presented for facial expressions recognition through the shape and texture features of facial landmarks, it is based on Active Appearance Model (AAM). First of all the facial landmarks are detected and then selected for features extractions. These two kinds of features are fused to classify expressions which rely on machine learning supervised learning algorithms. These features are tested with different state-of-the-art classifiers and achieved a high accuracy [11]. [12] worked on Deep Convolutional Neural Networks (CNNs) which is used for feature extraction. The work mechanism is different from other state-of-the-art machine learning techniques. The features extracted using CNN are very robust in terms of recognition accuracy, this method is also very fast in extracting features from video frames and images. Furthermore, the classifier used is ensemble CNN that has achieved an average accuracy of 75.2% on FERin 2013.

This work suggests a simple key for the identification of facial expressions that uses a mixture of CNN and stages of pre-processing of specific images. Convolutional neural networks achieve greater precision with big data. Though, there are no freely accessible data sets with adequate data for identification facial expressions with deep architectures. Then, to address the problem, authors implement roughly pre-processing methods to extract first the specific expression features of a face image and explore the instruction of performance of samples during training. The experiments used to evaluate our technique were performed using three widely used public databases (CK +, JAFFE and BU-3DFE). A study of the impact of each image preprocessing operation on the accuracy speed is presented. The projected method: attains good results compared to other approaches of identification of facial expression: accuracy of 96.76% in the CK + database. It is quick to train and allows real-time facial expression recognition with standard computers [13].

Authors propose a bidirectional recurrent hierarchical part based neural network (PHRNN) to analyze facial expression timing information. Morphological differences of the face and the dynamic progress of the expressions of their PHRNN model, which is effective for mining "temporal features" created on facial innovations from consecutive frames. Meanwhile, to make the information look immobile, a convoluted neural network of multiple signals (MSCNN) proposes to eliminate "spatial features" of still frames. We use recognition and verification signals as supervision to calculate various loss functions, which are useful for increasing variations of different expressions and reducing differences between identical expressions. This deep space-time Network Evolutional (composed of PHRNN and MSCNN) extracts the complete partial info, the presence of geometry and dynamic-stationary info, efficiently cumulative the identification performance of facial expressions. Experimental consequences show that this technique far surpasses the most contemporary. Three databases of widely used facial expressions (CK +, Oulu-Cassia and MMI), our method reduces the error rates of the best previously in 45.5%, 25.8% and 24.4%, respectively [14].

In this research, the stationary wavelet transform is used to abstract landscapes for facial expression identification for its good localization features, both in the spectral and spatial domain. More specifically, a combination of horizontal and vertical substrates of stationary wavelet transform is used, since these sub bands contain information on muscle measure for most facial expressions. The dimensionality of the representative is more compact by implementing a discrete alteration of the cosine in these sub-bands. The designated features are then approved to the forward moving NN that is skilled through the backward propagation procedure. An average recognition rate of 98.83% and 96.61% was achieved for the JAFFE and CK + dataset, respectively. The achieved accuracy is 94.28% for the locally registered MS-Kinect dataset. It has been observed that the proposed technique is very promising for the recognition of facial expression compared to other latest generation techniques [15].

Present a new facial descriptor, ternary local directional model (LDTP), for the identification of facial expression. LDTP competently encodes info related to expressions by means of directional info and the ternary pattern to exploit the strength of edge designs in the border area, overcoming the edges of the edges. Edge approaches in the soft area Our proposal. They practice a thick grid for constant codes (highly related to non-expression) and a thinner one for active codes (strictly related to expression). This multilevel approach allows us to make a more detailed description of facial movements, while continuing to characterize the gross characteristics of the expression. They have tested their technique using independent individual authentication systems to assess performance [16].

In this article, they intend to present a new framework for recognizing facial expressions to automatically distinguish expressions with great precision. Above all, a large feature consisting of a combination of facial and geometric features is introduced in the recognition of facial expression because it contains accurate and complete information about expressions. In addition, Deep Dispersion Autoencoders (DSAEs) are set up to recognize facial expressions with great precision when learning solid and discriminating data characteristics. The outcomes of the testing specify that the obtainable structure can attain a high identification accuracy of 95.79% in the extended Cohn Kanade database (CK +) for seven facial expressions, which exceeds the other three advanced methods up to 3.17%, 4.09% and 7.41%, respectively. The method offered is also functional to identify eight facial expressions and offers suitable accuracy of identification, which effectively validates the possibility and effectiveness of the method in this document [17].

The philosophies based on the motor of the identification of the facial expression suggest that the visual insight of the facial expression is maintained by sensorimotor processes that are also used to produce the same expression. As an outcome, sensorimotor and visual procedures must deliver congruent expressions info about a facial expression. Here, we report evidence that questions this view. The repeated execution of facial expressions has the opposite effect of recognizing a subsequent facial expression with respect to the repeated display of facial expressions. Furthermore, the results of the motor condition, but not the visual condition, were correlated with a non-sensory condition in which the participants imagined an expressions situation. These outcomes can be well described by the idea that the identification of facial expression is not always refereed by motor procedures but can also be identified only in visual information [18].

In this study, authors present an effective and fast facial expression identification system. They present a new characteristic called W_HOG where W designates the distinct wavelet alteration and HOG specifies the histogram of the oriented gradients characteristic. The proposed structure includes four phases:

- facial processing,
- domain transformation,
- extraction of characteristics and
- recognition of expressions.

Facial processing consists of phases of face detection, cropping and normalization. In the conversion of the area, the features of the spatial area are transformed in the frequency area by the application of discrete wavelet conversion (DWT). Feature extraction is performed by retrieving the Histogram of Directed Gradients (HOG) function in the DWT domain that is called the W_HOG function. For expression recognition, the W_HOG function is provided to a multi-cellified support vector (SVM) classifier based on a well-structured tree with one-to-all architecture. The projected system is skilled and verified with reference facial expression data sets CK +, JAFFE and Yale. The investigational outcomes of the planned technique are efficient for the identification of facial expression and overcome current approaches [19].

In this study [20] propose another plan for the FER framework dependent on intensive various leveled instruction. Capacities obtained from presentation based rasters are joined with geometric highlights in a various leveled structure. This appearance-based system removes facial highlights utilizing LBP pictures that are prepared during geometry-based preparation. This adjusts the directions of the AU, which is essentially a muscle that moves in outward appearances. The proposed strategy joins the aftereffects of the SoftMax work with two highlights,

considering the mistakes related with the second most noteworthy consequence of enthusiastic forecast. What's more, we offer methods for making pictures of individuals with nonpartisan feelings utilizing programmed coding. With this procedure, we can separate unique facial highlights among impartial and passionate pictures without succession information. We contrast the proposed calculation and other more current CK + and JAFFE dataset calculations, which are normally viewed as an approved dataset when outward appearances are identified. The 10-overlap cross-approval results demonstrate a precision of 96.46% in the CK + informational collection and an exactness of 91.27% in JAFFE. Contrasted with different techniques, the consequences of the proposed progressive system structure show an expansion in exactness of up to 3% and a normal increment of 1.3% in the CK + database. In the JAFFE dataset, exactness expanded to around 7% and the normal increment was changed to around 1.5%.

Table 1: literature work review

Sr.	Author	Year	Technique	Result
1	Liu and Wang [8].	2006	Multiple Gabor based Features facial extraction	Fusion methods works better than original
2	Kotsia and Pitas [9].	2007	Active Shape Model (ASM) with the Gabor filter	Accuracy with multiclass SVMs: 99.7% FAU: 95.1%
3	Li and Jain [10].	2011	Facial Animation Parameters using Active Shape Model	Not meet accuracy requirements
4	Zheng and Liu [11].	2016	Facial emotion recognition using AAM	Achieve accurate classification
5	Pramerdorfer and Kampel [12].	2016	Feature extraction using deep CNN	Average accuracy75.2% on FER2013
6	Lopes *et al* [13].	2017	Facial recognition with mixture of CNN and Pre-processing steps	Accuracy of 96.76% in the CK + database
7	Zhang *et al* [14].	2017	bidirectional recurrent hierarchical part based neural network (PHRNN)	Reduces error rates: CK+ = 45.5% Oulu-Cassia = 25.8% MMI = 24.4%
8	Qayyum *et al* [15].	2017	stationary wavelet transforms	achieved accuracy is 94.28%
9	Ryu *et al* [16].	2017	ternary local directional model (LDTP)	Better ability for recognition
10	Zeng *et al* [17].	2018	Deep Dispersion Autoencoders (DSAEs)	Accuracy: 95.79%

11	Rose *et al* [18].	2018	Motor and visual facial expression detection	Motor gives higher results than visual
12	Nigam *et al* [19].	2018	Wavelet histogram-oriented gradients	Shows better performance than other
13	Kim et al.	2019	LBP	Average Accuracy 96.46%
13	Proposed	2019	HOG and LTP	100% accuracy

3. Methodology and Materials

Facial Expression Recognition is a class of Facial Recognition System and can be represented by the flowchart in figure 3. The input image is used for the detection of Facial features first. If a video is provided according to a supported format to the algorithm, the algorithm first finds a face in a key frame. The interval to the next key frame is used for tracking the position of the face w.r.t. the image dimensions. The next step in the algorithm is to align the face in the center of the tracker. For that purpose, the image in the video or the image file is taken, if needed it is rotated, then the extra useless image information is discarded, if the face is distorted e.g. skewed and transposed, the algorithm warps it to proper shape and proportions, and then the illumination of the image is calculated and normalized. This procedure is called "geometrical normalization."

When the face is according to a certain pre-defined proportion and size and is well aligned, the feature extraction algorithm is applied to it. Facial landmarks are already assigned to the face such as eyebrow and mouth and jawline position, so then the expression detection algorithm is carried out. The results of the algorithm are matched to a database in the feature matching algorithm and the weights of the expressions in the image are saved in memory as a vector entry. This vector entry can then be used to compare with a predefined vector table with names i.e. a completely sad face has certain weights that are well defined if the calculated vector matches a vector for sadness, the classifier gives us the result "sad". For other expressions, the algorithm is intended to deliver accurate results with the competitive weight of each expression comparable with relevant recent research.

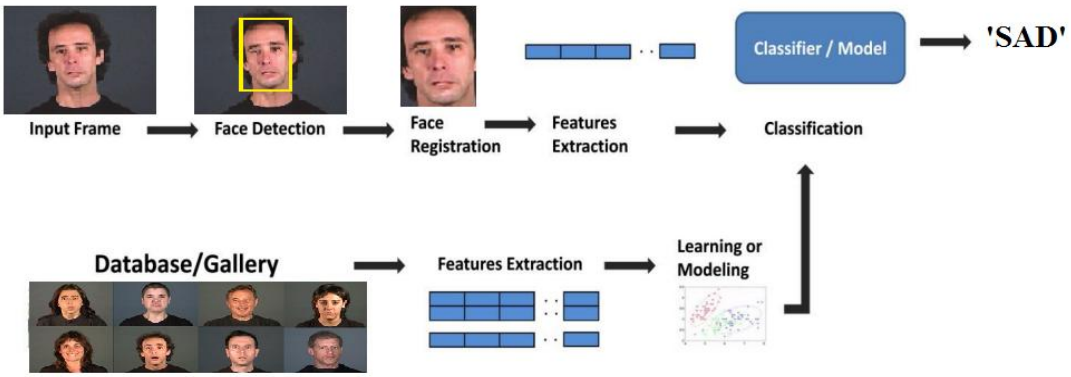


Figure 3: Steps in Facial expression Recognition

3.1 *Algorithm*

TRAINING:

Step 1: Feature extraction - Extracting LTP and HOG features for all training images in the database. The idea behind choosing LTP as the classifier is, it is easier to compute in quick time, and efficient in capturing the texture pattern of the images. Moreover, the reference used LBP and obtained better results than Geometric and Wavelet features. Since LTP is more detailed than LBP it can further increase the classification accuracy. And also, the HOG features can capture the edge information, which is crucial in determining eyebrows and mouth position in the facial image, which in turn is very important in recognizing the expressions.

Step 2: Classifier training - 3 classifiers are trained. 1) NN, 2) Template matching 3) SVM.

Each classifier is trained in two ways:

- Only using LTP features
- Combined LTP and HOG features

TESTING:

Step 1: Feature extraction

Step 2: Classification- All the test images are classified using all the (3x2 =) 6 classifier models created in the training stage.

ANALYSIS:

The classification accuracy of all 8 models is compared. And also, the results are compared with the reference. (This uses LBP for feature extraction, Template matching, and SVM for classification)

3.2 *Facial Component Extraction using LTP*

In request to shape LBP extra vigorous against small alterations in picture component estimations of level picture districts any place dim estimations of pixels don't change significantly, the thresholding subject of the descriptor is changed and a three-level administrator known as neighborhood ternary example (LTP) is anticipated [21]. The LTP descriptor is suitable to wear out the issues in close consistent dim-worth territories. In parallel encryption, the differentiation between the center picture component and a neighboring picture component is encoded by two qualities zero (0) and one (1) while in ternary encryption the qualification between the center picture component and a neighboring picture component two paired examples mulling over its positive and negative components. This is encoded by 3 qualities (1, 0 and −1). The ternary example has been separated into histograms from these positive and negative components figured over an area of intrigue are then linked.

LTP might be a three esteemed coding created on the possibility of LBP (Facial, Recognition, In, & Situations, 2014). In LTP, the area pixel dark qualities are contrasted and the focal pixel dim qualities by utilizing a limit. upheld this correlation, the area worth will be allotted one among the 3 qualities, +1 or 0 or - 1, as appeared in equation 1:

$$LTP\ (T_i) = \{1\ p_i \geq (i_c + l)\ 0\ |p_i - i_c| < l\ -1\ p_i \leq (i_c - l) \qquad (1)$$

Pi indicates the field of choice, ic means the dim-worth focus point, l signifies the edge, and afterward encoded the qualities that got from the upper equation.

In spite of the fact that LTP has a specific capacity to stifle clamor, its limit is physically set, typically a fixed worth. At the point when the light changes all of a sudden become rough when the edge is getting greater or the commotion sifting execution will decrease. To take care of the above issues, an Adaptive LTP Threshold calculation (ALTP) is proposed in this paper. Explicit computations are as per the following: Calculate the mean worth g of dark qualities in (P, R) area in equation 2.

$$\underline{g} = \frac{\sum_{i=0}^{p} g_i}{n} \quad i = 0, 1, \dots, p \qquad (2)$$

Locate the discrete qualities delta g of the qualities in the field with respect to the g bar using equation 3.

$$\Delta g_i = g_i - \underline{g} \quad i = 0,1, \dots, p \qquad (3)$$

At that point the standard deviation (SD) of particular numbers is utilized to determine the level of scattering using equation 4.

$$\delta = \sqrt{\frac{\sum_{i=0}^{p} (\Delta g_i - \Delta \underline{g})^2}{p+1}} \qquad (4)$$

Among them $\underline{\Delta g} = \frac{\sum_{i=0}^{p} \Delta g_i}{p+1}$ speaks to the mean value of discrete qualities. Simultaneously, the weight coefficient S appears in equation 5.

$$S_{(LTP)} = \{1 \left(g_i + \underline{g}\right) > (\underline{\Delta g} + \delta)\ 0\ else\ -1 \left(g_i + \underline{g}\right) < (\underline{\Delta g} + \delta) \qquad (5)$$

3.3 *Feature Extraction using HOG*

Presently, inclination has been thought of in light of the fact that the ground-breaking component for highlight extraction when contrasted with others. Inclination holds extra information rather than the force/surface. The preeminent normal is that the HOG. Slope local Ternary Pattern (GLTP), local Gradient Pattern, local Gradient Increasing Pattern are the contrary inclination based generally systems for FER.

HOG-descriptor was created by [22] and has been with progress utilized by a few scientists working inside the area of pc vision. mainly it's been utilized for human recognition, object ID and person on foot distinguishing proof. Hoard is registered to misuse greatness and direction. Flat and vertical the two angles are of the info picture that figured misuse the ensuing Equations 6 and 7.

$$G_x = I_f * [-1,0,1] \qquad (6)$$

$$G_y = I_f * [-1,0,1]^T \qquad (7)$$

3.4 *Classification using SVM*

Order is the remainder of the FER framework inside which the classifier classifies the articulation like grin, tragic, shock, outrage, dread, disturb and impartial. SVM is one among the characterization methods inside which

two sorts of methodologies are concerned. they're one against one and one against all approaches. One against all characterization proposes that it develops one example for each class [23]. One against one characterization recommends that it builds one class for each pair of classifications [24].

SVM is one among the most grounded characterization procedures for cutting edge spatiality inconveniences [25]. SVM is the administered AI procedure and it utilizes four types of pieces for its higher presentation [26]. They're straight, polynomial, Radial Basis work (RBW) and sigmoid. The direct portion maps the high dimensional information and it's straightly detachable. The RBF portion utilizes the play out that maps the one component into the high dimensional information. The polynomial piece learns the nonlinear models and moreover settles their likeness.

SVM[27]could be an entrenched AI approach, that has been with progress received for different data characterization issues. The possibility of SVM depends on the cutting-edge factual learning hypothesis. For data arrangement, SVM starting verifiably maps the data into a superior dimensional component zone at that point develops a hyperplane in such a way that the isolating edge between the examples of two classes is immaculate. This isolating hyperplane acts then as the choice surface. Given a training set of labeled models T= {(xi, yi), I =1… . l}, where xi ∈ Rn and yi ∈, the new learning is classed by the resulting capacity using equation 8:

$$f(x) = sign\,(\sum_{i=1}^{I} \quad a_i y_i k(x_i, x) + b \qquad (8)$$

where k (xi, x) could be a bit work, ai are Lagrange multipliers of the twin improvement issue, and b is that the parameter of the best hyperplane. The instructing tests xi with ai > 0 are alluded to as the help vectors, and furthermore the isolating hyper-plane amplifies the edge with pertinence these help vectors. Among the varying pieces found inside the writing, straight, polynomial, and spiral. The SVM finds a straight isolating hyperplane with the top edge to isolate the data in the highlight house.

3.5 *Testing Methods*

For the purpose of testing following methods are used:

- Cross Validation Method
- Percentage Method

3.6 *Materials*

We used mostly commonly knowing datasets for our research i.e. JAFFEE, YALE, CK+. We applied all the steps on these datasets. These datasets contain the basic facial expressions.

3.7 *Tools*

For the purpose of testing following methods are used:

- MATLAB- (Matrix laboratory)

- WEKA- (Waikato Environment for Knowledge Analysis)

4. Results and Discussions

We have executed and tried our proposed outward appearance acknowledgment Method more than three dataset's authentic articulation acknowledgment. Here, we present outcomes for three most prominently utilized datasets that are the Japanese female outward appearance (JAFFE) acknowledgment dataset, the Extended Cohn-Kanade dataset (CK+), and the Yale outward appearance acknowledgment dataset. The investigations have been performed in MATLAB 2019 condition on an Intel® Core™ i3 machine.

For assessment, we have isolated the subjects arbitrarily into n gatherings of generally equivalent size and connected a forget about one gathering cross approval test conspire. These gatherings are subject-autonomous. Similar subjects did not show up in both preparing and testing. In this manner, the testing is finished with novel appearances and individual free.

4.1 *Visual and Numerical Results of JAFFEE*

The JAFFE [28] dataset contains 213 pictures of 10 female Japanese subjects. Each subject was solicited to play out different postures from seven essential prototypic articulations. The articulation mark for each picture speaks to the articulation of the subject. was approached to present. The pictures are given in a .tiff picture group with a goal of 256×256 pixels. We split the JAFFE dataset into training and testing images.

Experimental Results for JAFFE

The experimental results of the JAFFE dataset are shown in table 2. For JAFFE dataset K-Fold and %age split validation method used and we get the accuracy 100% for both in 3.5 sec time.

Table 2: JAFFE results with 70% training and 30% testing images

Validation Method	Training	Testing	TP Rate	FP Rate	Precision	F-Measure	Time (sec)
K-Fold 10	70%	30%	100%	0%	100%	100%	3.5
%age split	70%	30%	100%	0%	100%	100%	3.95

Confusion Matrices JAFFE

The confusion matrix for JAFFE dataset is created by using k-folds 10 cross validations method as shown in table 3 and for percentage split validation in table 4.

Table 3: Confusion Matrix for JAFFE k-folds 10 cross-validation

a	b	c	d	e	F	G	Classified As
70	0	0	0	0	0	0	**a = angry**
0	77	0	0	0	0	0	**b = disgust**
0	0	134	0	0	0	0	**c = fear**
0	0	0	77	0	0	0	**d = happy**
0	0	0	0	63	0	0	**e = neutral**
0	0	0	0	0	63	0	**f = sad**
0	0	0	0	0	0	84	**g = surprise**

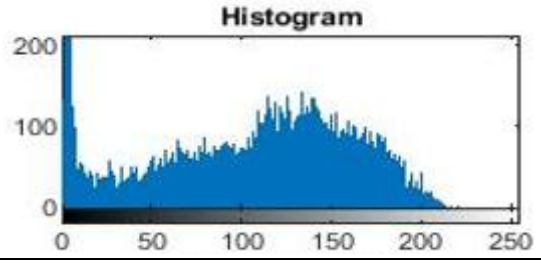
Table 4: JAFEE percentage split validation

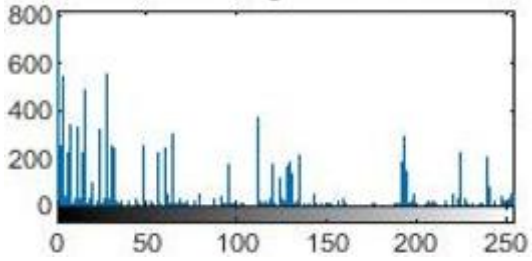
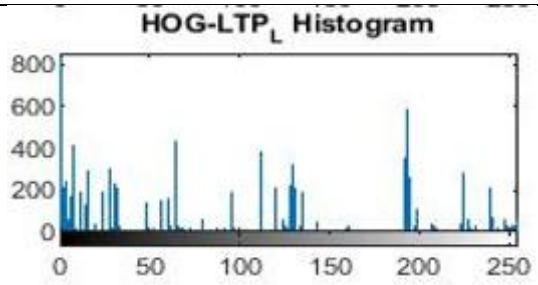
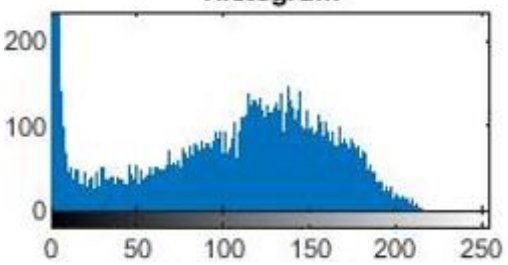
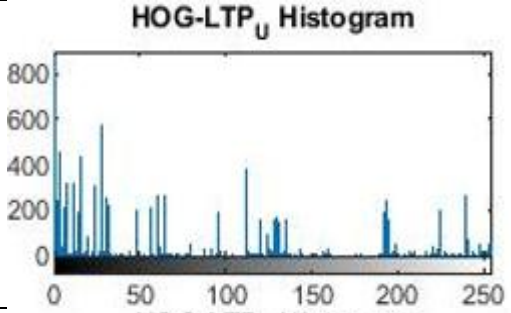
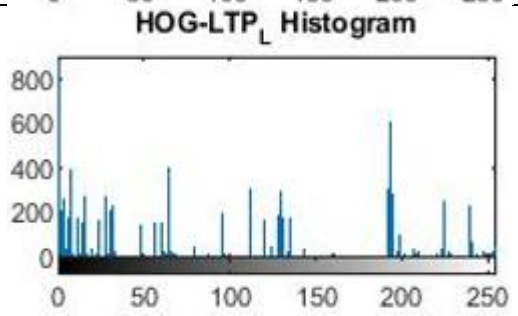
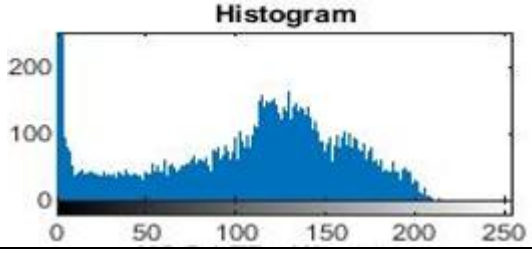
a	b	c	d	e	f	G	Classified As
125	0	0	0	0	0	0	**a = angry**
0	105	0	0	0	0	0	**b = disgust**
0	0	131	0	0	0	0	**c = fear**
0	0	0	130	0	0	0	**d = happy**
0	0	0	0	99	0	0	**e = neutral**
0	0	0	0	0	102	0	**f = sad**
0	0	0	0	0	0	121	**g = surprise**

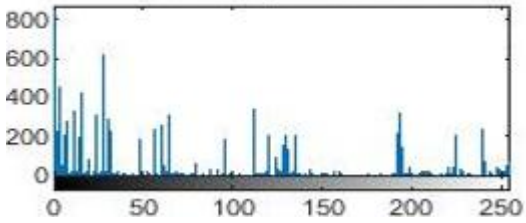
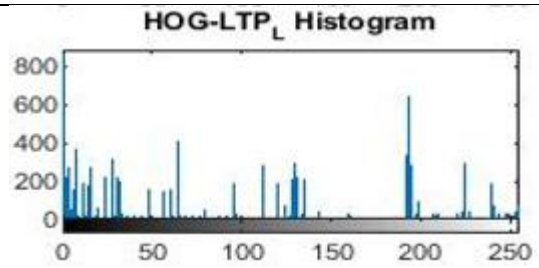
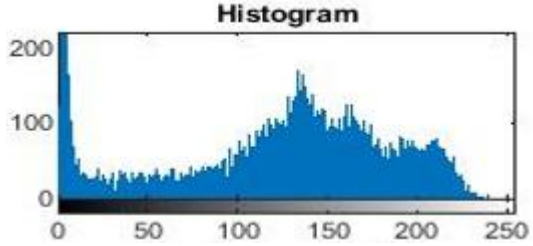
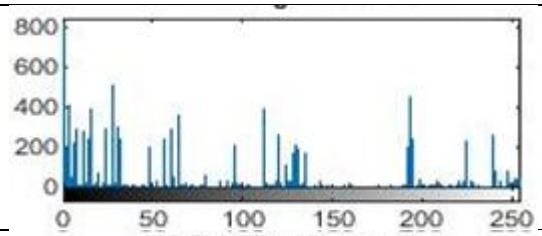
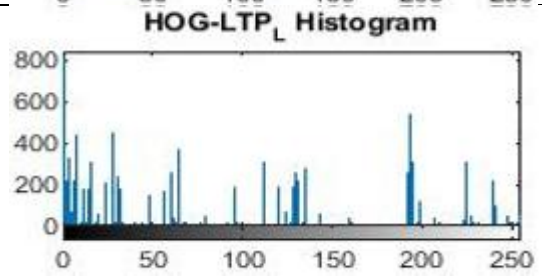
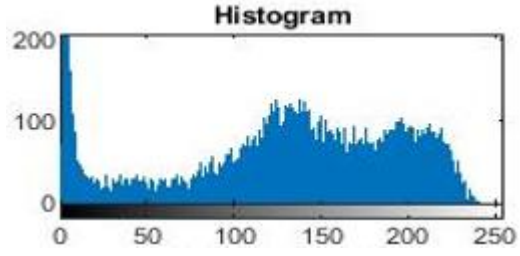
Visual Results for JAFFE with Histogram

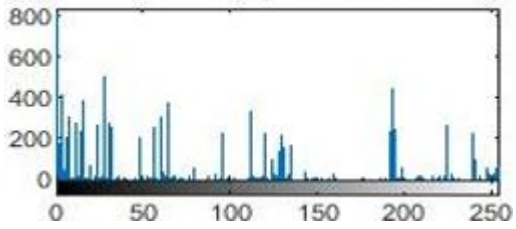
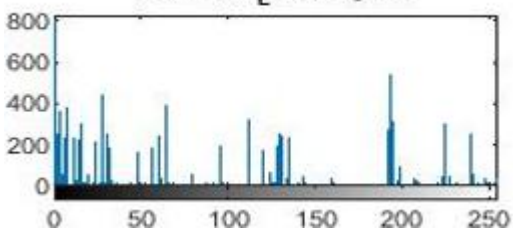
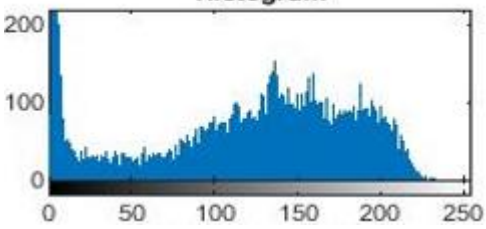
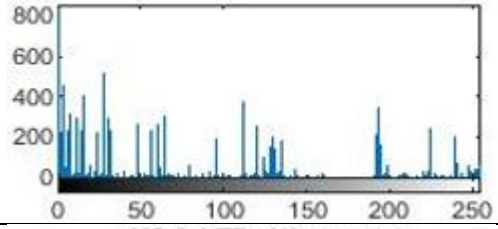
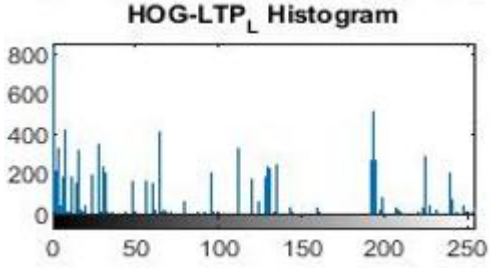
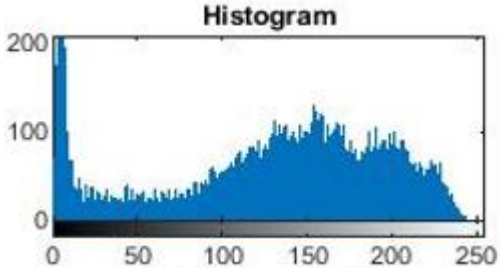
Visual results of JAFFE data for basic expressions using the Histogram are shown in table 5.

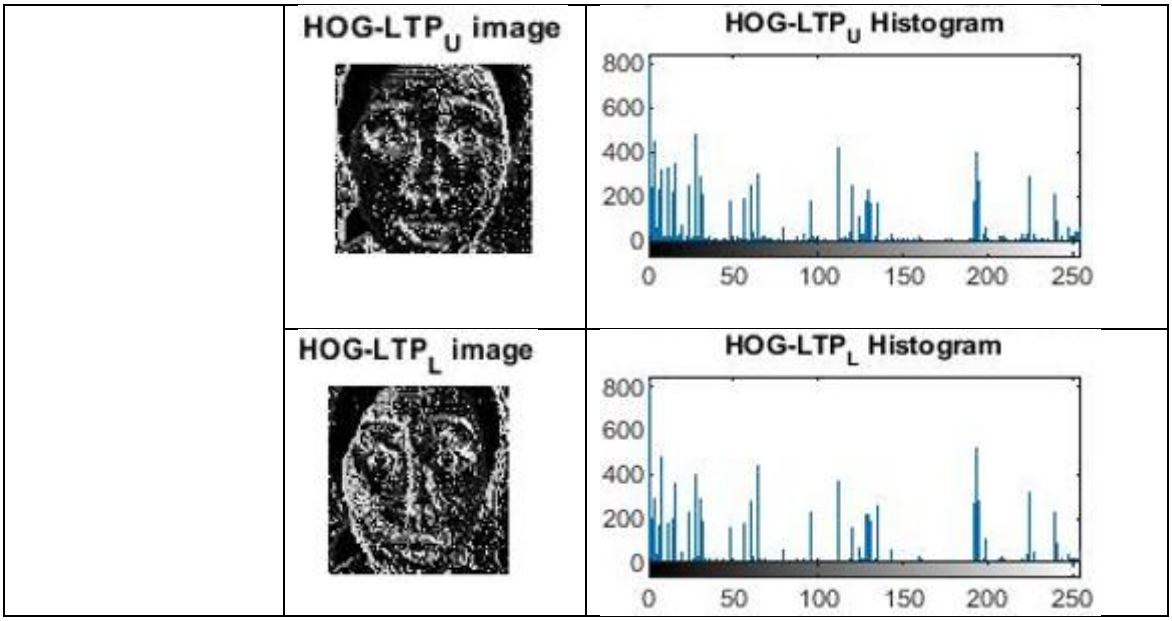
Table 5: JAFEE visual results with Histogram

Expression	Image	Histogram
1. Angry		

	HOG-LTP$_U$ image 	HOG-LTP$_U$ Histogram 
	HOG-LTP$_L$ image 	HOG-LTP$_L$ Histogram 
2. Disgust	original image 	Histogram 
	HOG-LTP$_U$ image 	HOG-LTP$_U$ Histogram 
	HOG-LTP$_L$ image 	HOG-LTP$_L$ Histogram 
3. Fear	original image 	Histogram 

	HOG-LTP$_U$ image 	HOG-LTP$_U$ Histogram 
	HOG-LTP$_L$ image 	HOG-LTP$_L$ Histogram 
4. Happy	original image 	Histogram 
	HOG-LTP$_U$ image 	
	HOG-LTP$_L$ image 	HOG-LTP$_L$ Histogram 
5. Neutral	original image 	Histogram 

	HOG-LTP$_U$ image 	HOG-LTP$_U$ Histogram 
	HOG-LTP$_L$ image 	HOG-LTP$_L$ Histogram 
6. Sad	original image 	Histogram 
	HOG-LTP$_U$ image 	HOG-LTP$_U$ Histogram 
	HOG-LTP$_L$ image 	HOG-LTP$_L$ Histogram 
7. Surprise	original image 	Histogram 



4.2 *Visual and Numerical Results of CK+*

The CK AU-coded articulation dataset comprises 97 college understudies between 18 to 30 years old. Sixty-five percent were female, 15% were African-American and 3% were Asian or Latino. Subjects were approached to perform up to six distinctive prototypic feelings (for example euphoria, shock, outrage, dread, disturb and pity) just as an impartial articulation. Picture groupings from the nonpartisan articulation to target articulation were caught utilizing a frontal confronting camera and digitized to 640 × 480 or 490 pixels in .png picture design. Discharged in 2010, The CK+ dataset [29] builds the quantity of subjects from CK by 27% to 123 subjects and the quantity of picture arrangements by 22% to 593 successions.

Experimental Results for CK+

The experimental results of the CK+ dataset are shown in table 2. For CK+ dataset Cross validation method is used and we get the accuracy 99.99% in 162.34 sec time.

Table 6: CK+ results with 70% training and 30% testing images

Validation Method	Training	Testing	TP Rate	FP Rate	Precision	F-Measure	Time (sec)
Cross Validation	70%	30%	99.99%	0%	99.99%	99.99%	162.34

Confusion Matrices CK+

The confusion matrix for CK+ dataset is created by using 10 cross validations methods as shown in table 7.

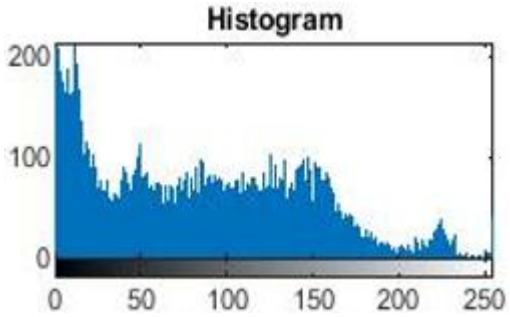
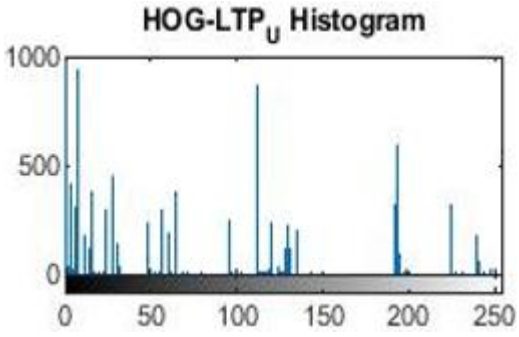
Table 7: Confusion Matrix for CK+ using cross validation

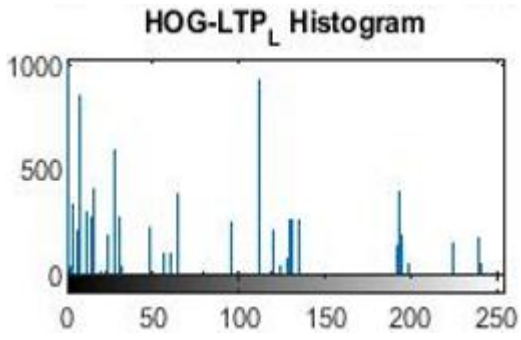
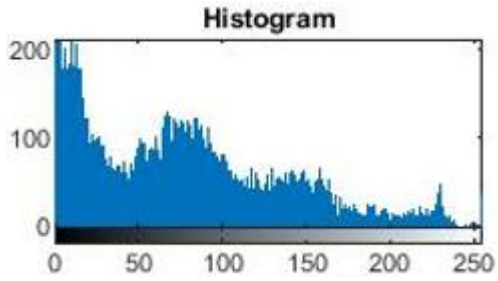
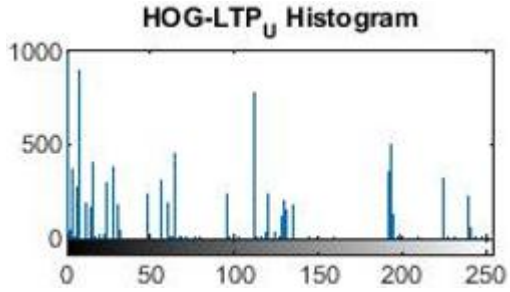
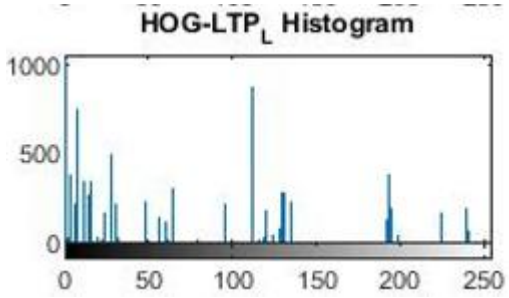
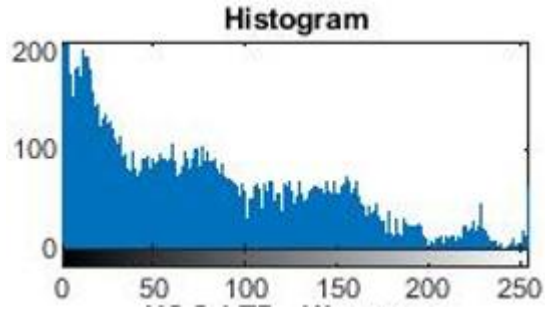
a	b	c	d	e	f	g	Classified As
396	0	0	0	0	0	0	**a = angry**
0	378	0	0	0	0	0	**b = disgust**
0	0	413	0	0	0	0	**c = fear**
0	0	0	431	0	0	0	**d = happy**
0	0	0	0	357	0	0	**e = neutral**
0	0	0	0	0	368	0	**f = sad**
0	0	0	0	0	0	366	**g = surprise**

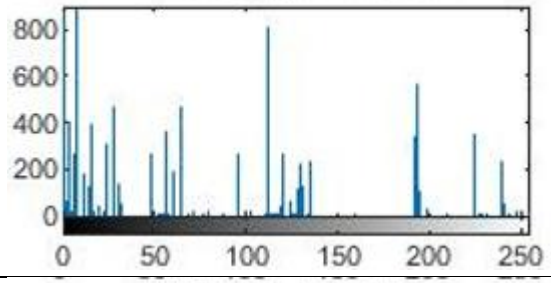
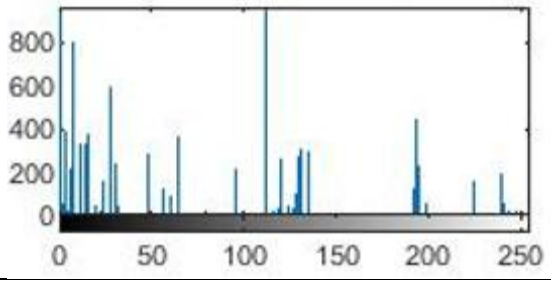
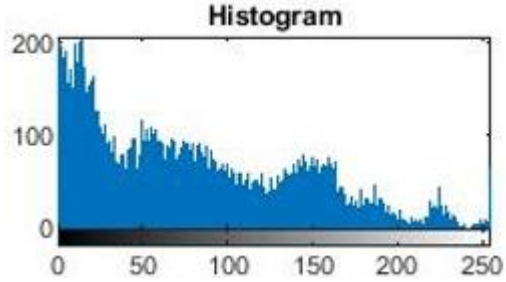
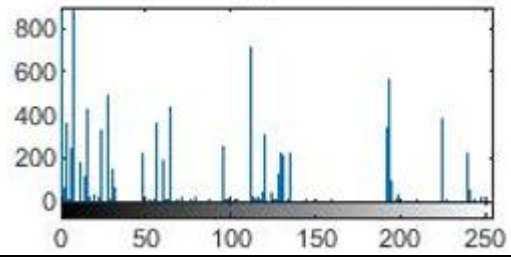
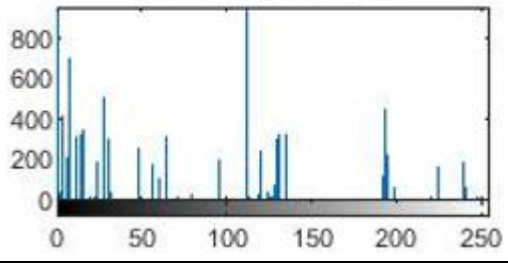
Visual Results for CK+ with Histogram

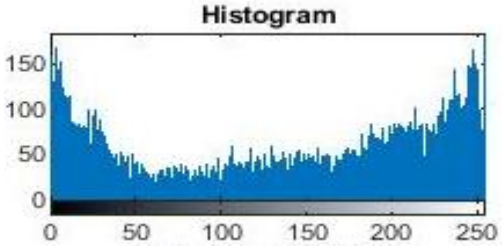
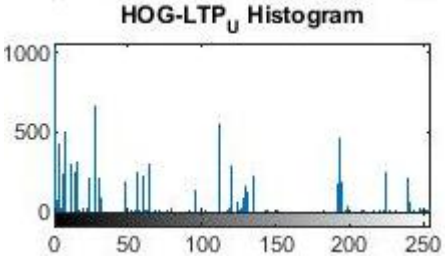
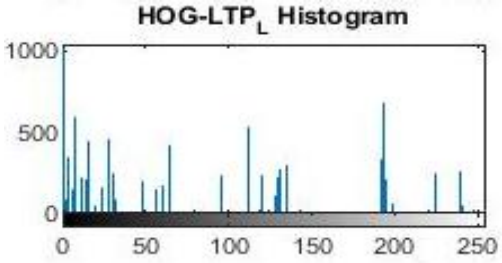
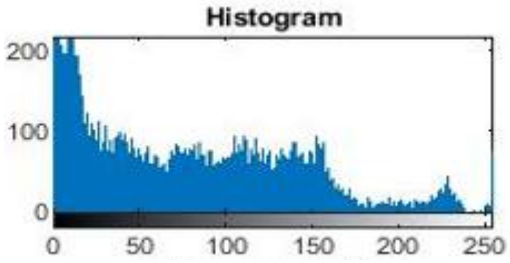
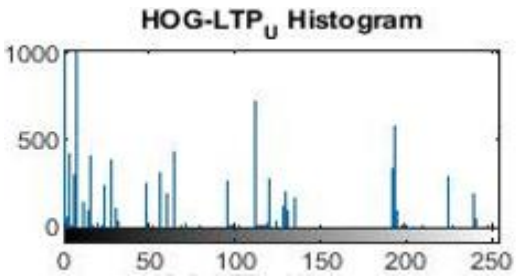
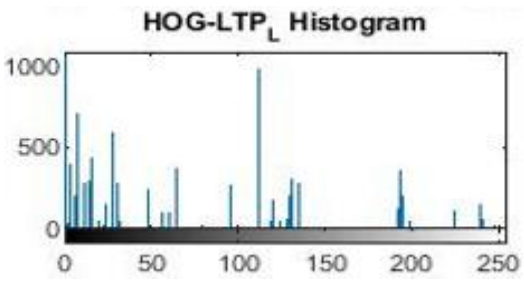
Visual results of CK+ data for basic expressions using the Histogram are shown in table 8.

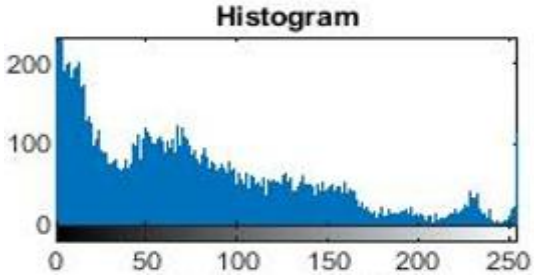
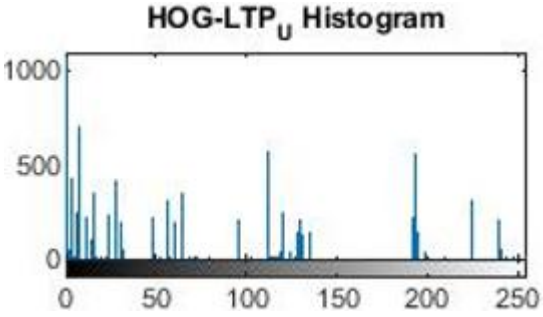
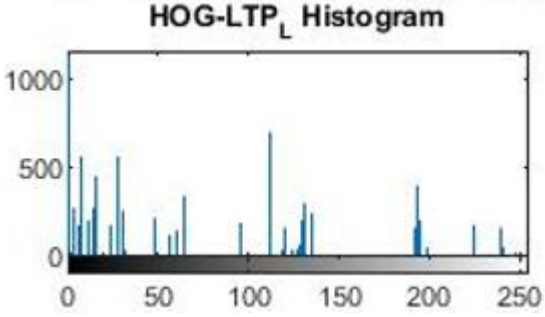
Table 8: CK+ visual results with Histogram

Expression	Image	Histogram
1. Angry		
		

	HOG-LTP$_L$ image 	
2. Disgust	original image 	
	HOG-LTP$_U$ image 	
	HOG-LTP$_L$ image 	
3. Fear	original image 	

	HOG-LTP$_U$ image 	HOG-LTP$_U$ Histogram 
	HOG-LTP$_L$ image 	HOG-LTP$_L$ Histogram 
4. Happy	original image 	Histogram 
	HOG-LTP$_U$ image 	HOG-LTP$_U$ Histogram 
	HOG-LTP$_L$ image 	HOG-LTP$_L$ Histogram 

5. Neutral	original image 	
	HOG-LTP$_U$ image 	
	HOG-LTP$_L$ image 	
6. Sad	original image 	
	HOG-LTP$_U$ image 	
	HOG-LTP$_L$ image 	

7. Surprise	original image 	Histogram 
	HOG-LTP$_U$ image 	HOG-LTP$_U$ Histogram 
	HOG-LTP$_L$ image 	HOG-LTP$_L$ Histogram 

4.3 *Visual and Numerical Results of YALE*

The Yale facial expression dataset contains 165 grayscale images of 15 individuals. There are 11 pictures for every subject.

Experimental Results for YALE

The experimental results of the YALE dataset are shown in table 9. For YALE dataset K-Fold-10, LOOCV, and %age split validation method used and we get the accuracy 97.6%,98.8%, and 100% respectively.

Table 9: YALE results with 70% training and 30% testing images

Validation Method	Training	Testing	TP Rate	FP Rate	Precision	F-Measure	Time (sec)
K-Fold 10	70%	30%	97.6%	0.08%	97.6%	97.6%	0.69
LOOCV	70%	30%	98.8%	0.4%	98.9%	98.8%	0.85
%age	70%	30%	100%	0%	100%	100%	1.03

Confusion Matrices YALE

The confusion matrices for YALE dataset are created by using 10-cross validation, LOOCV, and %age methods as shown in table 10, table 11, table 12 respectively.

Table 10: Confusion Matrix for YALE k-folds 10 cross validation

a	b	c	d	Classified As
42	0	0	0	a = happy
0	40	2	0	b = neutral
0	1	41	0	c = sad
0	1	0	41	d = surprise

Table 11: Confusion Matrix for YALE LOOCV

a	b	c	d	Classified As
42	0	0	0	a = happy
0	42	0	0	b = neutral
0	1	41	0	c = sad
0	1	0	41	d = surprise

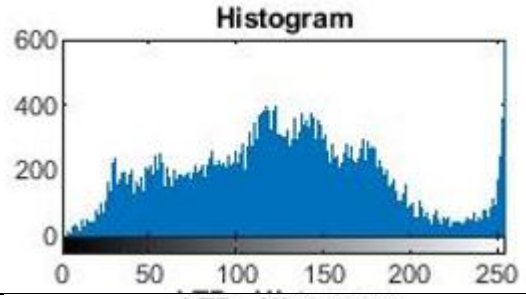
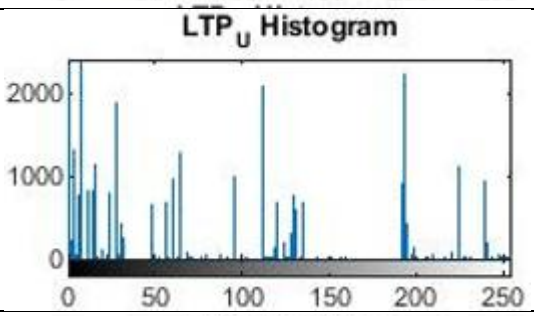
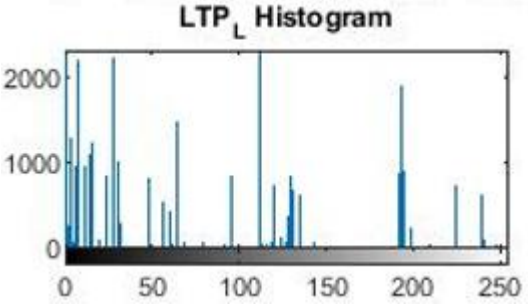
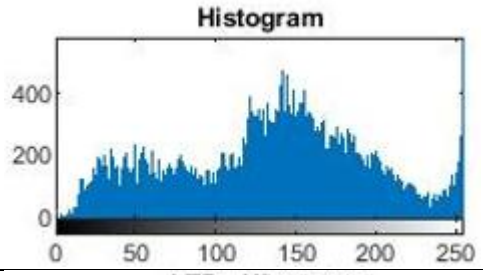
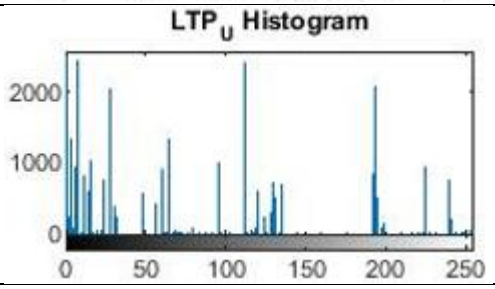
Table 12: Confusion Matrix for YALE Percentage

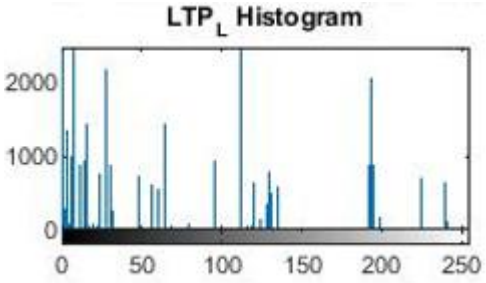
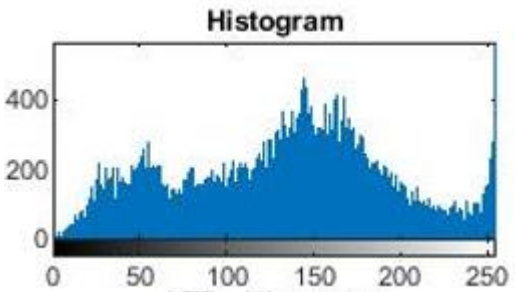
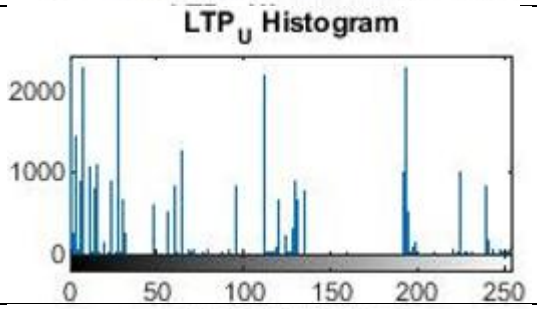
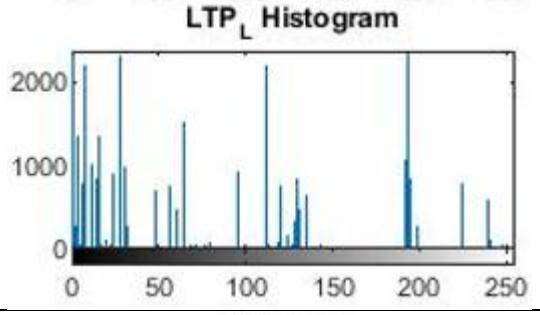
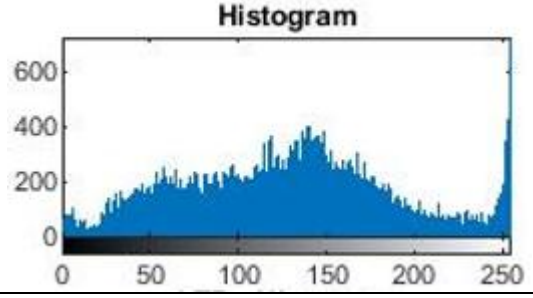
a	b	c	d	Classified As
8	0	0	0	a = happy
0	10	0	0	b = neutral
0	0	8	0	c = sad
0	0	0	8	d = surprise

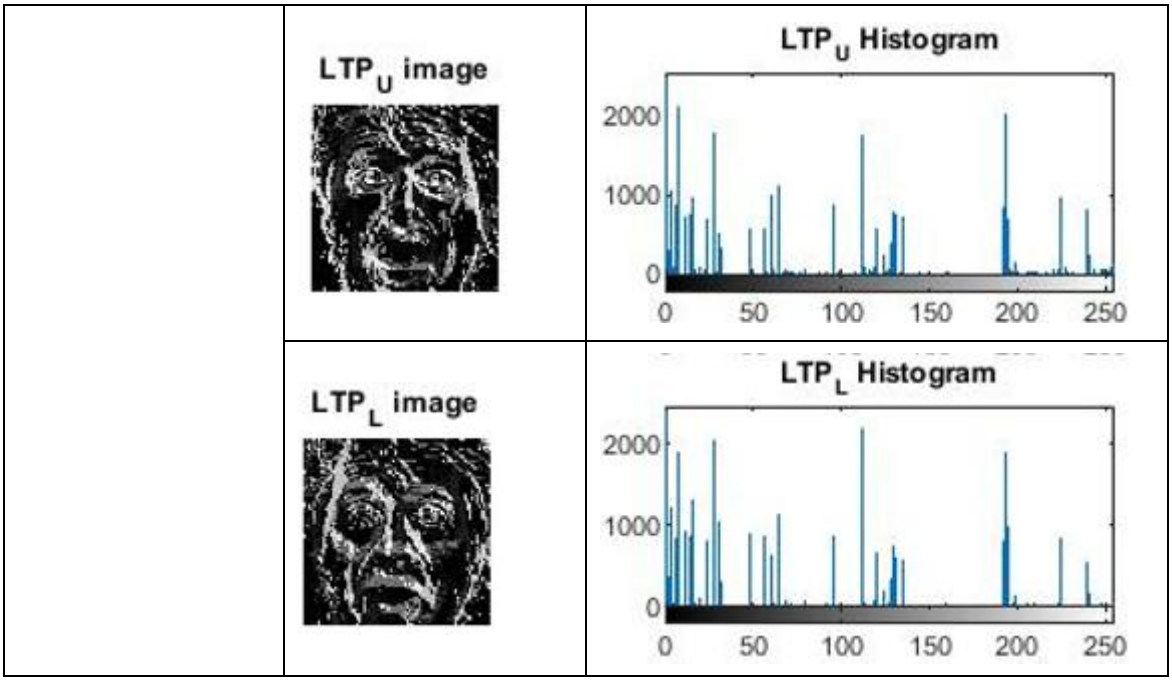
Visual Results for YALE with Histogram

Visual results of YALE data for basic expressions using the Histogram are shown in table 13.

Table 13 YALE visual results with Histogram

Expression	Image	Histogram
1. Happy		
		
		
2. Natural		
		

	LTP_L image 	LTP_L Histogram 
3. Sad	original image 	Histogram 
	LTP_U image 	LTP_U Histogram 
	LTP_L image 	LTP_L Histogram 
4. Surprise	original image 	Histogram 



5. Conclusion and Future Work

Feature extraction is a method for facial recognition. It involves several steps like dimensionality reduction, feature extraction, and feature selection. Dimensionality reduction is an important task in the pattern recognition system. In this work three steps are followed first is face detection second is feature extraction and the last step is expression recognition if matched with the arranged data sets then show the positive or matched output.

Three large datasets CK+, JAFEE and YALE are used with six and more than an expression of facial images. This work provides good results as compared to the previous state of artwork related to accuracy and clarity in poor images as well. In the pre-processing stage focus on face image detection using MATLAB code then extracted the valid features using Histogram of Oriented Gradients (HOG) and Local Ternary Pattern (LTP).In post-processing more binary filter, the principle of rules used for clear the gestures and after extraction and separation of features three classifiers neural network (NN), Support vector machine (SVM) and template matching using for further classification.

In the results and discussions, part three validation method is used for the analysis of results listed as Leave one out-cross validation (LOO-CV), k-fold and percentage method using the software WEKA. After classified images apply the best filter principle of rules and binary orientation for pure clarity of expression images. Compared results with the previous art of works gain good accuracy for JAFEE is 100%, CK+ is 99.99% and YALE is 98.9% with large datasets of images.

For future work and recommendation will be taken a shot at the biggest informational indexes utilizing application run time quick face finder. Give more than great precision to non-untestable feelings utilizing the application with video identifier and picture indicator moreover.

References

[1] Cai, Y.; Li, X.; Li, J. Emotion Recognition Using Different Sensors, Emotion Models, Methods and Datasets: A Comprehensive Review. Sensors 2023, 23, 2455.

[2] Stankovi´c, M.; Neši´c, M.; Obrenovi´c, J.; Stojanovi´c, D.; Miloševi´c, V. Recognition of facial expressions of emotions in criminal and non-criminal psychopaths: Valence-specific hypothesis. Personal. Individ. Differ. 2015, 82, 242–247.

[3] LeCun,Y.; Bengio, Y.; Hinton, G. Deep learning. Nature 2015, 521, 436–444.

[4] Dulguerov, P.; Marchal, F.; Wang, D.; Gysin, C. Review of objective topographic facial nerve evaluation methods. Am. J. Otol. 1999, 20, 672–678.

[5] Ekman,P.; Friesen, W.V. Unmasking the Face: A Guide to Recognizing Emotions from Facial Clues; Ishk: Los Altos, CA, USA, 2003; Volume 10.

[6] Ekman, P., & Rosenberg, E. (2005). What the face reveals. What the Face Reveals.

[7] M. Vasilescu and D. Terzopoulos, “Multilinear independent components analysis,” 2005 IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit., vol. 1, no. June, pp. 1–7, 2005.

[8] WeiFeng Liu and ZengFu Wang, “Facial Expression Recognition Based on Fusion of Multiple Gabor Features,” 18th Int. Conf. Pattern Recognit., no. 2, pp. 536–539, 2006.

[9] I. Kotsia and I. Pitas, “Facial expression recognition in image sequences using geometric deformation features and support vector machines,” IEEE Trans. Image Process., vol. 16, no. 1, pp. 172–187, 2007.

[10] S. Z. Li, A. K. Jain, and F. Recognition, “Handbook of Face Recognition,” pp. 1–15, 2011.

[11] W. Zheng and L. Cuicui, “Facial Expression Recognition Based On Texture And Shape,” 25th Wirel. Opt. Commun. Conf. Facial, pp. 1–5, 2016.

[12] C. Pramerdorfer and M. Kampel, “Facial Expression Recognition using Convolutional Neural Networks : State of the Art,” 2016.

[13] A. T. Lopes, E. de Aguiar, A. F. De Souza, and T. Oliveira-Santos, “Facial expression recognition with Convolutional Neural Networks: Coping with few data and the training sample order,” Pattern Recognit., vol. 61, pp. 610–628, 2017.

[14] K. Zhang, Y. Huang, Y. Du, and L. Wang, “Facial Expression Recognition Based on Deep Evolutional Spatial-Temporal Networks,” IEEE Trans. Image Process., vol. 26, no. 9, pp. 4193–4203, 2017

[15] H. Qayyum, M. Majid, S. M. Anwar, and B. Khan, “Facial Expression Recognition Using Stationary Wavelet Transform Features,” Hindawi, vol. 2017, no. 1, pp. 1–9, 2017.

[16] B. Ryu, A. R. Rivera, J. Kim, and O. Chae, “Local Directional Ternary Pattern for Facial Expression Recognition,” IEEE Trans. Image Process., vol. 26, no. 12, pp. 6006–6018, 2017.

[17] N. Zeng, H. Zhang, B. Song, W. Liu, Y. Li, and A. M. Dobaie, “Facial expression recognition via learning deep sparse autoencoders,” Neurocomputing, vol. 273, pp. 643–649, 2018.

[18] S. de la Rosa, L. Fademrecht, H. H. Bülthoff, M. A. Giese, and C. Curio, “Two Ways to Facial Expression Recognition? Motor and Visual Information Have Different Effects on Facial Expression Recognition,” Psychol. Sci., p. 095679761876547, 2018.

[19] S. Nigam, R. Singh, and A. K. Misra, “Efficient facial expression recognition using histogram of oriented gradients in wavelet domain,” 2018.

[20] Kim, J., Kim, B., Roy, P. P., & Jeong, D. (2019). Ji-hae kim 1 , byung-gyu kim 1 , partha pratim roy 2 , da-mi jeong 1, 4(c), 1–14. https://doi.org/10.1109/ACCESS.2019.2907327.

[21] X. Tan and B. Triggs, "Enhanced Local Texture Feature Sets for Face Recognition Under Difficult Lighting Conditions," in *IEEE Transactions on Image Processing*, vol. 19, no. 6, pp. 1635-1650, June 2010, doi: 10.1109/TIP.2010.2042645.

[22] L. Dang, B. Bui, P. D. Vo, T. N. Tran and B. H. Le, "Improved HOG Descriptors," 2011 Third International Conference on Knowledge and Systems Engineering, Hanoi, Vietnam, 2011, pp. 186-189, doi: 10.1109/KSE.2011.36.

[23] Valstar MF, Patras I, Pantic M. Facial action unit detection using probabilistic actively learned support vector machines on tracked facial point data. In2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05)-Workshops 2005 Sep 21 (pp. 76-76).

[24] Lei G, Li XH, Zhou JL, Gong XG. Geometric feature based facial expression recognition using multiclass support vector machines. In2009 IEEE International Conference on Granular Computing 2009 Aug 17 (pp. 318-321).

[25] Burges CJ. A tutorial on support vector machines for pattern recognition. Data mining and knowledge discovery. 1998 Jun;2(2):121-67.

[26] Osuna E, Freund R, Girosit F. Training support vector machines: an application to face detection. InProceedings of IEEE computer society conference on computer vision and pattern recognition 1997 Jun 17 (pp. 130-136)

[27] Chih-Chung Chang and Chih-Jen Lin. 2011. LIBSVM: A library for support vector machines. ACM Trans. Intell. Syst. Technol. 2, 3, Article 27 (April 2011), 27 pages. https://doi.org/10.1145/1961189.1961199

[28] Minaee, S., & Abdolrashidi, A. (2019). Deep-Emotion : Facial Expression Recognition Using Attentional Convolutional Network.

[29] Lucey, P., Cohn, J. F., Kanade, T., Saragih, J., Ambadar, Z., Matthews, I., & Ave, F. (2010). The Extended Cohn-Kanade Dataset (CK +): A complete dataset for action unit and emotion-specified expression, (July), 94–101.