Turkish Online Journal of Qualitative Inquiry (TOJQI)
Volume 12, Issue 8, July, 2021:7126 – 7141

Research Article

HOG, Wavelet Moments and Cartoon Features based Facial Expression Recognition using Random Forest Classifier

Priyanka Nathawat, Dr. Vivek Chaplot

Bhupal Noble's University, Udaipur, India-313001
nathawat.pre@gmail.com

Abstract – Facial expressions play an important role in nonverbal communication as they provide information about people's feelings and are used in a wide variety of fields, especially in education, health, law and entertainment. Recognizing emotions is one of the most complex areas of science. In recent years, more and more applications have tried to automate it. These innovative applications concern several areas such as helping children with autism, video games, human-machine interaction. This paper presents a novel technique of facial expression recognition using HOG, Wavelet Moments and Cartoon descriptors based feature extraction methods. Random Forest Classifier is utilized to calculate the similarity score using two public datasets of facial expressions; KDEF and JAFFE. The performance evaluation is done on the basis of a confusion matrix plot with accuracy, precision and sensitivity graphs.

Keywords – Cartoon Features, HOG, Random Forest Classifier, Viola Jones Method, Wavelet Moments, etc.

I. Introduction

A face contains information about a person's personality, such as eye color, mouth shape, hair color, ear shape, etc. It can also express communicative and emotional forms of expression that can be recognized by facial expressions.

More importantly, facial expressions are a series of visually visible facial signals resulting from the activation (voluntary or not) of one or more of the 44 muscles that make up the face, which express the sensation of a change in the face [1].

Facial expression plays an important and vital role in interpersonal and social communication and reveals various information such as emotions. Automatic facial expression detection has been an active area of research over the past several decades with a number of important applications such as human-machine interaction and visual monitoring. Although humans can easily recognize and analyze faces and facial expressions, developing automated systems to accomplish this task presents many challenges.

The issue of recognizing facial expressions has been debated for many years, the question is how to recognize emotions, and many researchers have actively participated in recognizing emotions from facial expressions using visual, computer and image processing techniques [2].

Over the past decade, the computer vision research community has shown great interest in the analysis and automatic recognition of facial expressions. The research and computer vision community is envisioning the development of a system that can recognize facial expressions in videos or images. Most of these facial expression analysis systems attempt to categorize expressions into several broad emotional categories such as: B. Joy, sadness; Anger, surprise, fear and disgust [3]. Today, facial recognition technology is used in public places to help law enforcement agencies, identify those who may be dangerous, combat passport fraud, identify missing persons, and much more. [4].

Research into the automatic recognition of facial expressions is very important in our modern technological age. Significant investments have been made in facial recognition in recent years thanks to significant applications in areas such as virtual reality, intelligent course systems, healthcare, and data-driven animation [5]. The main purpose of facial expression recognition is to describe a person's emotional state based on a given facial image. Several approaches are proposed for automatic recognition of facial expressions as they are; It is split into two classes: attribute-based and model-based. Numerous functions can classify machine learning algorithms. Several techniques have been developed to address the problem of reducing redundant and irrelevant variables. Feature selection helps understand data, reduces computational overhead, reduces the impact of dimensional problems, and improves evaluator performance [6].

This research focuses on recognizing facial expressions by features. Seven different emotions (happy, angry, sad, annoyed, scared, surprised, and neutral) were classified based on facial expressions. The median filter is used to adjust the image. To classify facial expressions, a hybrid structure was created based on a random forest classifier.

In the second part of the study, studies on facial expression recognition are summarized in the literature. Section 3 briefly describes HOG, wavelet moments and feature extraction algorithms based on cartoon descriptors and classifiers based on machine learning methods. The simulations and results associated with the methodology proposed in the research paper are described in section 4. The last section finally concludes the research paper.

II. Previous Studies

There are various studies in the literature on the recognition of facial expressions. In the study conducted by [7], the problem of 3-dimensional facial expression recognition was discussed with multi-layered (manifold) data and kernel-based methods. The performance of the developed structure has been tested on BU-3DFE and Bosphorus datasets. In the study carried out by [8], features were extracted by examining face markers on a three-dimensional plane. For feature extraction, the positions of 83 face markers were determined in 3D and the distances between each pair of points were calculated. In total, 3403 distances were calculated for 83 points. It is aimed to find the most suitable subsets of these points. For this, Fisher Criterion-based feature selection method was used. Various studies in the literature were compared with the current study and it was shown that successful results were obtained with 88.5%. Authors of[8] detected facial expressions by generating masks from facial markers over BU-3DFE and BU-4DFE datasets. The resulting masks were classified with Convolutional Neural Networks.

Authors of [9] proposed a structure based on Gabor method-based feature extraction and Adaboost Algorithm-based feature selection, which will increase the accuracy rate and reduce the computation time in solving the facial expression recognition problem. Multilayer Feed Forward Neural Networks were used as classification method. The performance of the proposed structure has been tested on JAFEE and Yale datasets. The algorithm developed by authors of [10] captures the image from the dataset and reduces the size with Principal Component Analysis (PCA). Facial expressions are detected by Artificial Neural Networks (ANN) according to the trained dataset. Images have been converted to grayscale and resized as 256x256 pixels for faster processing. At the end of the study, it was seen that PCA + ANN method successfully handled the facial expression detection problem. Authors of [11] created a facial expression dataset for 7 different emotions with 5 female and 7 male participants. Viola Jones algorithm was used for face detection. The location of 68 face markers has been found with the dlib library. The performances of Support Vector Machines, kNN, Random Forest and Regression Tree methods were compared and the best result was obtained from Support Vector Machines with 97%. A facial expression classification structure using Viola Jones Algorithm and Principal Component Analysis was proposed in the study by authors of [12]. Authors of [13] tested their new feature extraction approach on JAFEE dataset. Support Vector Machines were preferred as the classification method and an average of 87% classification success was achieved. Authors of [14] applied a general optimization algorithm to select a sub-feature set based on minimizing the classification error in the facial expression recognition problem.

Authors of [15] proposed a hybrid information system combining computer vision and machine learning technologies for visual and interactive e-learning systems. Authors of [16] tested the success of Felzenszwalb's directed gradient histogram in facial expression recognition. Authors of [17] considered facial expression recognition as an image pair matching problem. In the study performed by [18], the most efficient subset was determined from the features obtained by using the triangulation points obtained from the face image by applying sequential forward feature selection. Authors of [19] ensured the recognition of facial expressions by using local binary pattern (LBP) features obtained from the vicinity of the landmarks found using active shape models. Authors of [20] proposed a new and effective method for the detection of face action units using variable decision thresholds in the estimation phase of binary learning methods. Authors of [21] carried out a study that will enable automatic recognition of smiles on face images.

III. Proposed Methodology

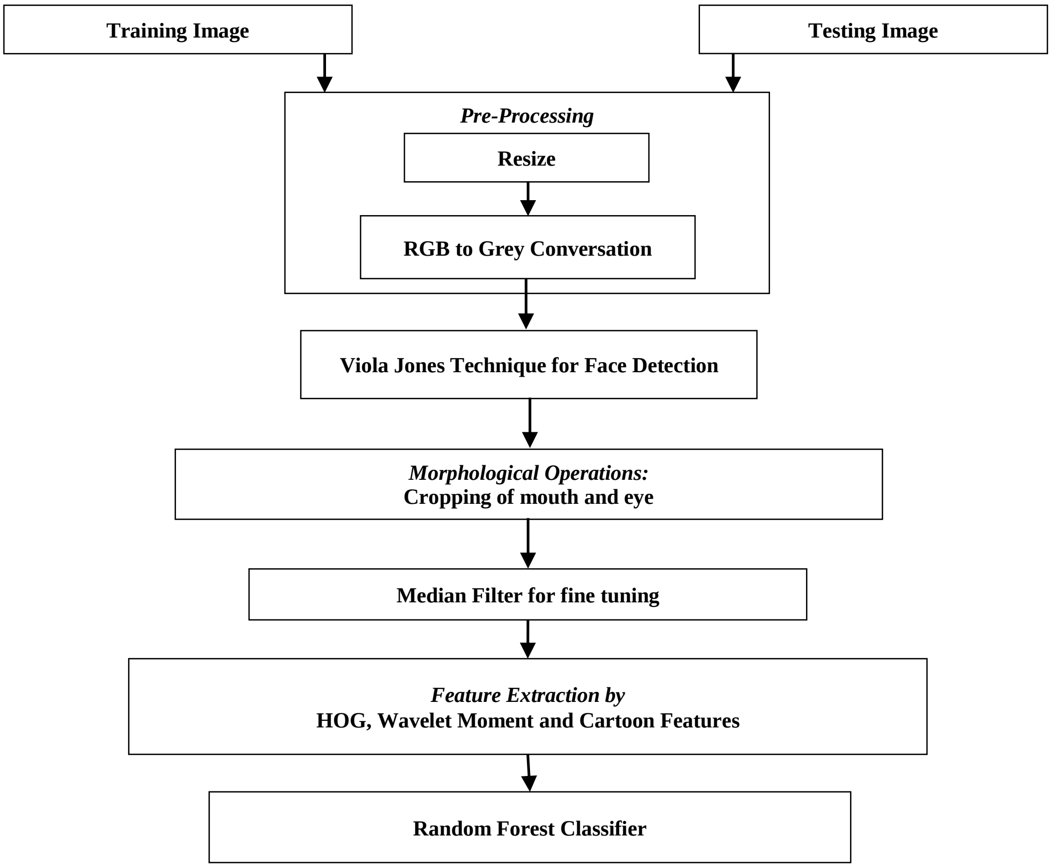


Figure 1: Proposed flow diagram for facial expression recognition

This section describes facial expression recognition using the Histogram of Oriented Gradient, wavelet moments and cartoon features. The flow diagram for proposed work is shown in Figure 1. The facial expression recognition process is a sequential action. Facial recognition techniques are usually studied in three areas, classified according to their approach. Pixel groups are identified in the test image using a pattern matching method similar to the sample image.
Rest of methodology is explained as follows:

A. Pre-Processing of Face

1. Initially, the input image is resized to 224×224 pixels using the built-in resizing function available in MATLAB.
2. After resizing, the face image in RGB format is converted to grayscale using the rgb2gray function if the input image is colored.

B. Viola Jones Method for Face Detection

The automatic detection of faces in a digital image is a problem that presents numerous drawbacks, since there is no a priori information on the size of the face or its location within the image. For these reasons, the detection algorithms have to be prepared to handle faces with multiple sizes and different appearances, which can change drastically depending on the position, lighting and facial expression. In images where many individuals appear, or where costumes cover part of people's faces, a face may be partially hidden and algorithm glitches.
The detection technique proposed by Viola and Jones uses a series of classifiers, grouped in successive stages that respond to a set of features to detect the presence of a face in the image. As stated above, it is an image recognition algorithm, which was the first object detection framework to provide competitive real-time object detection rate proposed in 2001 by its two creators, and it in turn can be trained to detect other types of objects based on the face detection problem [22].
Viola and Jones modified the Haar-like feature rectangles in such a way that 3 different types of Haar rectangles emerged, such as [23]:
Characteristic of Two Rectangles: It is the difference between the sums of the pixels that are within two different rectangles:

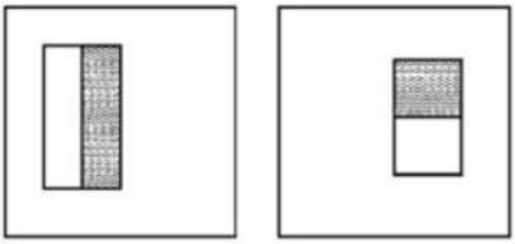


Figure 2: Two rectangles characteristic [24]

Characteristic of Three Rectangles: It is the sum of the pixels that are inside the two outer rectangles, subtracted with the sum of the pixels of the middle rectangle:

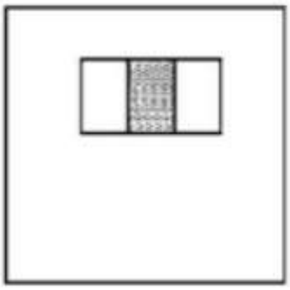


Figure 3: Three rectangles characteristic [24]

Characteristic of Four Rectangles: It is the difference between the diagonals of the pairs of rectangles.

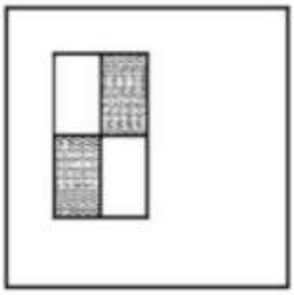


Figure 4: Four rectangles characteristic [24]

These values indicate specific properties of a specific area of the image. Each feature can indicate the presence or absence of certain features in the image, such as edges or texture changes. For example, two rectangular elements can indicate where the border between the dark area and the light area lies.

This algorithm for facial recognition creates an integral image, which makes the computational cost for the sum of the pixels lower than that of the Haar-like feature. This is because an integral image is a data structure and algorithm designed for greater efficiency in summing the values of the pixels in a rectangle subset of the grid.

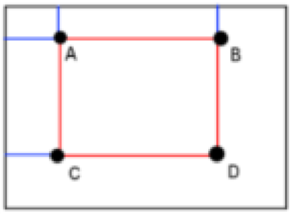
The algorithm can be defined as [22]:

$$I(x,y) = \sum_{\substack{x' \leq x \\ y' \leq y}}^{\infty} i(x', y') \tag{1}$$

Where $i(x,y)$ is the value of the pixel at position (x,y).

$$I(x,y) = i(x,y) + I(x,y-1) + I(x-1,y) - I(x-1,y-1) \tag{2}$$

After calculating the total area, exactly four control matrices are required to estimate the sum of the intensities over the area of the rectangle, regardless of the size of the area. What is it, designations in the figure below:



$Sum = D - B - C + A$

Figure 5: Evaluation of the sum of the pixel values [24]

C. Morphological Operations

Mathematical morphology is a field of digital image processing based on set theory, algebraic and geometric principles [25]. The main operations of mathematical morphology are erosion and dilation [26]. The erosion operation $\varepsilon_B(f)$ is obtained by getting the smallest value in window B, called the structure element.

$$\varepsilon_B(f)(r, c) = \min\{f(r + d, c + l)|(d, l) \in B\} \tag{3}$$

The dilation operation δ_B is obtained with the element with the highest value in the area of the element of style B.

$$\delta_B(f)(r, c) = \max\{f(r - d, c - l)|(d, l) \in B\} \tag{4}$$

Mathematical morphology has other operations, such as opening and closing, which are performed by performing a combination of basic operations. The morphological disclosure of an image f by a structuring element B is called $f \circ B$ and is defined as blurring f by means of B followed by expansion by the structuring element itself [26].

$$f \circ B = \delta\{\varepsilon(f, B), B\} \tag{5}$$

The closure of an image f by a structuring element B is denoted $f \cdot B$ and is defined as an extension of f to B followed by erosion by the structuring element itself.

$$f \cdot B = \varepsilon\{\delta(f, B), B\} \tag{6}$$

D. Median Filter

The processes and filtering algorithms that have been implemented will be discussed below. The median filter is based on ordering in ascending or descending order the pixels that fall within the neighborhood or window that is analyzed. Once ordered, the value that is located in the central position of the window is chosen and that is the value that is located in the pixel that is in that position of the target image. The mathematical modeling of this method is as shown in equation (7):

$$\begin{gathered} K(i) = K(0), K(1), K(2), \dots K(i), \dots, K(M \cdot N - 1) \\ K(0) < K(1) < K(2), \dots < K(i), \dots < K(M \cdot N - 1) \\ G(x, y) = K(v) \end{gathered} \tag{7}$$

Where M and N represent the dimensions of the neighborhood, $K(i)$ are the values read and ordered, $G(x, y)$ is the value in the coordinate pair (x, y) that is obtained as a result of the operation and v represents the central position of the neighborhood where it operates and can be calculated using equation (8):

$$v = \frac{M \cdot N - 1}{2} \tag{8}$$

Generally $M = N$ and both must be odd numbers. This method is particularly effective if you want to preserve contours, borders as well as fine details of objects.

E. Feature Extraction

The goal of extracting feature in image processing field is to express the feature in numerical or symbolic form called encoding. Depending on the case, the values of these features can be real, integer or binary. The vector composed of n features represent a point in the new space of n dimensions. The steps involved in feature extraction are shown in the flow diagram in Figure 1 using three feature extraction methods explained in following subheadings.

1. Histogram of Oriented Gradients (HOG)

HOG is a technique for extracting characteristics that acts by dividing the image into Gradient vectors. Gradient vectors are vectors that, in addition to direction and sense, give us the greatest value regarding some magnitude (in our case the gradient applies to edges), so our gradient vector at each point will have direction, direction and magnitude pointing to where we have the greatest contrast relative to edges.
In the technique, the image is divided into cells of the same size and the gradient vectors are calculated in each cell with its direction and magnitude [27]. After calculating the gradients, all these vectors are organized in a Histogram depending on their angle and magnitude. After the generation of the histograms is complete, blocks are normalized, aiming at the elimination of noise and distortions.
The histogram oriented gradients is based on the assumption that every image will have the same changes in its illumination regardless of whether it is light or dark, and it also works efficiently when evaluating images with change of scale; thanks to the fact that it gives a description of the gradients within the image and their direction [28].
Given a grayscale image divided into cells (in this case contemplated as 8×8), the computational gradient described by equations (9) and (10):

$$G_x = [-1,0,1] \cdot \left[I\big((x-1),y\big), I(x,y), I\big((x+1),y\big)\right]^T \tag{9}$$
$$G_y = [-1,0,1] \cdot \left[I\big(x,(y-1)\big), I(x,y), I\big(x,(y+1)\big)\right]^T \tag{10}$$

The magnitude of each pixel is defined by equation (11):

$$G(x,y) = \sqrt{G_x(x,y)^2 + G_y(x,y)^2}, \tag{11}$$

While its direction is given as follows:

$$\varphi(x,y) = arctan\frac{G_y(x,y)}{G_x(x,y)} \tag{12}$$

Conventionally (12) containers are used that will have all angles between 0 and 180 degrees. When the angle of a pixel falls within the range of the container, its magnitude is added to it and considered as 0 for all the others, as described in equation (13):

$$hog_i = \begin{cases} G(x,y)\ and\ \varphi(x,y) & \in \quad (i) \\ 0 & otherwise \end{cases} \tag{13}$$

We add each of the obtained values to the feature vectors of each of the cells within the entire image. In order to have a correct description of all the lighting and contrast changes, it is necessary to normalize locally, so we must group the cells into larger groups. Given the above, the block descriptor is given by the concatenation of the normalized histograms generated in each cell. These blocks generally overlap, which means that each of the cells may be considered on more than one occasion for the final descriptor [29]:

$$hog_B = \frac{hog_B}{\sqrt{\|hog_B\|^2 + \varepsilon'}} \tag{14}$$

Where hog_B is the descriptor of each block and ε represents a small constant value (theoretically it doesn't matter which one) added to avoid numerical problems in the square root. Then, the final descriptor will be given by the concatenation of each of the blocks within the (12) containers.

2. Wavelet Moments

Directional decomposition for image analysis can be initiated by a Gaussian windowing system in a two-dimensional Fourier interval and a search for four-dimensional filters, as well as Gabor waves directed by J. Daugman [30]. The use of Gabor waves for visualization is often associated with viewing the human visual system.
Wavelet is constructed through an isotropic Gaussian window of a complex plane wave with frequency F in the direction θ[31]:

$$\psi^\theta(x) = \frac{e^{-\|x\|^2/2}}{2\pi} e^{-j(x^T\omega_0)} \tag{15}$$

Where, $\omega_0 = F[\cos(\theta)\,;\sin(\theta)]^T$
The expansion depends on the number of orientations K, which is fixed and can be regularly distributed in $[0,\pi]$.

$$\theta \in \Theta = \left\{\frac{k\pi}{K}; 0 \le k \le K\right\} \tag{16}$$

Therefore, we can decompose any real 2D signal (x) into a dot product with the following expression:

$$\left\{\psi_{j,u}^{\theta}(x) = 2^{-j}\psi^{\theta}\left(2^{-j}(x-u)\right)\right\}_{\theta\in\Theta, j\in\mathbb{Z}, u\in\mathbb{R}^2} \quad (17)$$

3. Feature Extraction using Cartoon Descriptors

The cartoon feature in images is classically obtained by using a pair of low-pass and high-pass filters to the image *I* by the following minimization:

$$\min_u\left\{\sigma^4\int|Du|^2 + \|I-u\|_{H^{-1}}^2\right\} \quad (18)$$

Where u represents the cartoon part of the image I.

F. Classification by Random Forest Classifier

Breiman (2001) presented a random forest with the following very general definition [32]:
"A random forest is a classifier consisting of a set of basic decision tree type classifiers, which are defined":

$$\{h(x,\Theta_k),\ k = 1,\dots L\} \quad (19)$$

Let $\left(\hat{h}(\Theta_1),\dots,\hat{h}(\Theta_q)\right)$be a set of predictive trees with$\Theta_1,\dots,\Theta_q$independent random variables$\mathcal{L}_n$. The $\hat{h}_{RF}$ random forest predictor was obtained by adding this set of random trees as follows:

- $\hat{h}_{RF}(x) = \frac{1}{q}\sum_{l=1}^{q}\hat{h}(x,\Theta_l)$, the mean of the predictions of individual trees in the regression.
- $\hat{h}_{RF}(x) = \arg\max_{1\le k\le K}\sum_{l=1}^{q}1_{\hat{h}(x,\Theta_l)=k}$, the majority of votes of individual forecast trees in the classification.

The term random forest comes from the fact that individual predictors are explicit here for predictors for each tree, and each tree depends on an additional random variable (i.e., different from $\mathcal{L}_n$).
The classification process followed by the random forest method consists of [33]:

1. Assignment to a node if it is terminal, deciding whether a node will be labeled as a sheet or it will carry a test.
2. If the node is not terminal, then we have to select a test to assign it.
3. If the node is terminal, then we must give it a class.
The general algorithm for decision trees is as follows:
Input: sample S
Initialize the current tree to the empty tree;
The root designates the current node
Repeat
See if the current node is terminal
If the node is terminal then
Assign it a class
If not
Select a test and generate as many new child nodes as there are answers to this test
End if
Explore another node if there is one
Until a decision tree A is obtained
Exit: decision tree A.

IV. Simulation and Results

A. Facial Expression Datasets

1. Karolinska Directed Emotional Faces (KDEF) Database

The KDEF database [34] was originally developed for neuroscience research. However, it has since been used in the field of computer vision because of its applicability. The KDEF database contains 4900 images taken from 70 people (35 men and 35 women), ranging in age from 20 to 30 years old.

Figure 6: Examples of images extracted from the KDEF database. From left to right: Neutrality, Joy, Sadness, Surprise, Anger, Fear, and Disgust. From top to bottom: 0°, 45°, -45°, 90°, -90° [34]

Each individual displays 7 expressions (anger, disgust, fear, joy, neutrality, sadness, surprise), which are captured twice from 5 different angles (-90, -45, 0, +45, +90 degrees) and recorded the JPEG format with a resolution of 562×762 pixels (Figure 6). The authors of this database kept an eye on several important issues when developing the database. The environment (lighting, background, and distance from camera) was kept constant throughout the capture process and subjects were asked to remove all accessories (hats, glasses, facial hair and makeup).

2. Japanese Female Facial Expression (JAFFE) Database

The JAFFE database [35] includes 213 facial expression images of ten Japanese women. The latter simulated 3 to 4 examples for each of the six basic emotions, as well as neutral emotion. The resolution of the images is 256×256 pixels. Figure 7 shows some examples of images from the JAFFE database.

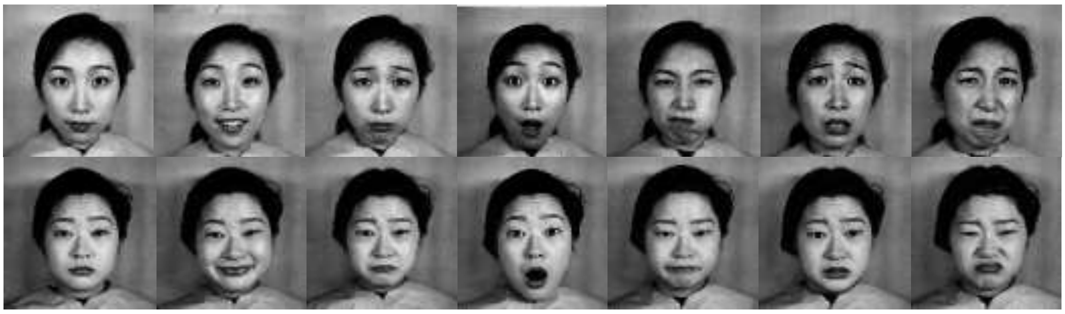


Figure 7: Examples of images extracted from the JAFFE database. From left to right:Neutrality, Joy, Sadness, Surprise, Anger, Fear, and Disgust [35]

B. Simulation Resutls Using KDEF Database

The simulation is carried out by using MATLAB software image processing toolbox. Following steps are involved in series for facial expression recognition. Input image is taken and reduced to 224×224 as follows:

Figure 8: Input image

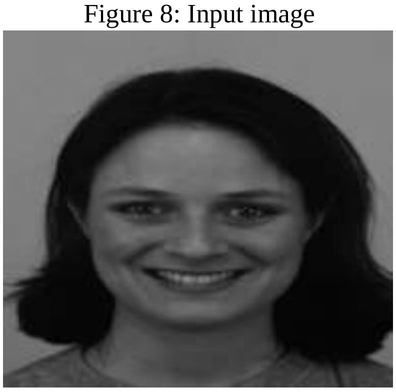


Figure 9: Resized to 224×224

Then using Viola Jones Method face is detected and cropped to small size image shown in figure 10 and 11.


Figure 10: Face detection using Viola Jones method


Figure 11: Cropped image

Morphological operations are performed detecting eyes and mouth, then using viola jones method eyes and mouth are detected and forming a new cropped image. Feature extraction is performed usingHOG, cartoon descriptors and wavelet moments techniques. This cropped image is again resized to 110×110 for further classification as shown in figure from 12 to 14.

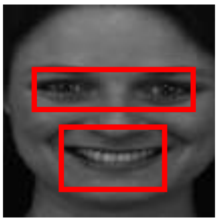


Figure 12: Eye and mouth detection


Figure 13: Cropped image


Figure 14: Resized to 110×110

The above feature extraction method was applied to the image above and the features was extracted for further classification using a random forest classifier. In the proposed approach, there is no threshold to face recognition. The RF itself performs similarity measurements and recognizes test images. Finally, the confusion matrix diagram shows the performance of the RF-based method (see Figure 15).

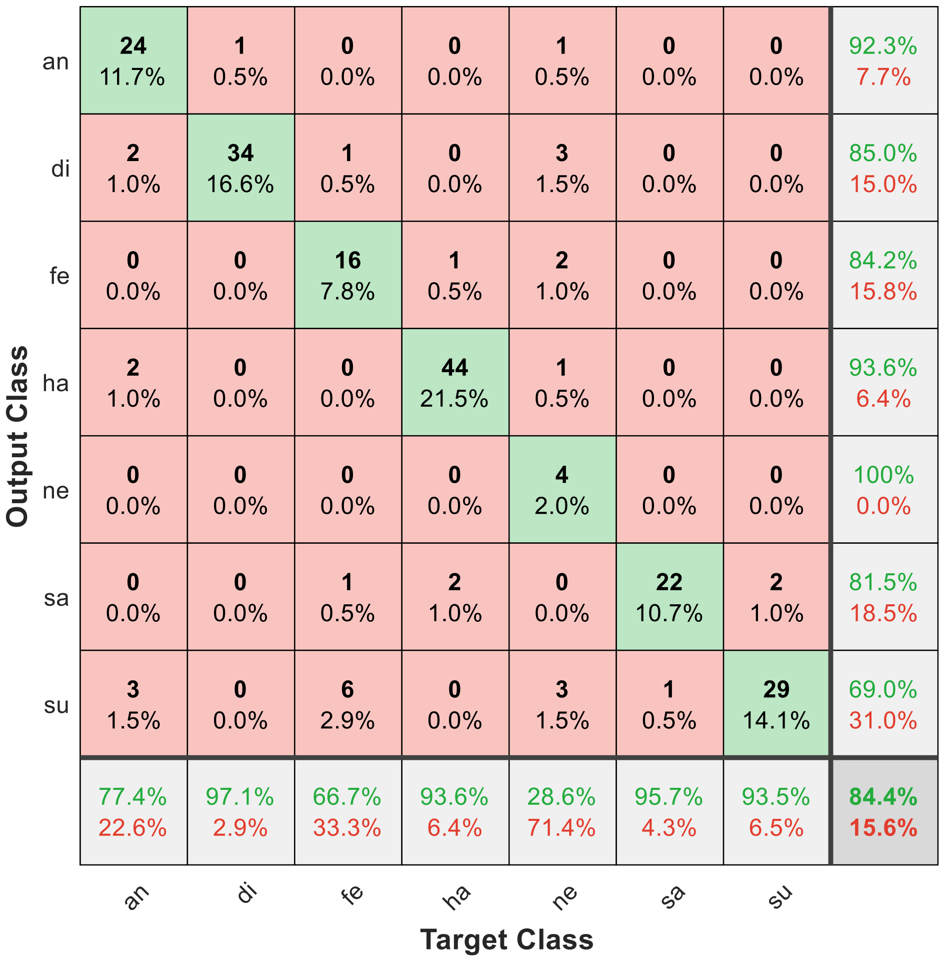


Figure 15: Confusion matrix plot of proposed approach for KDEF database

Rows and columns represent classes in the KDEF facial expressions database. We visited 7 classes and each class represented 7 different facial expressions. The confusion matrix plot shown above concludes the specificity for proposed algorithm i.e. 84.4%.

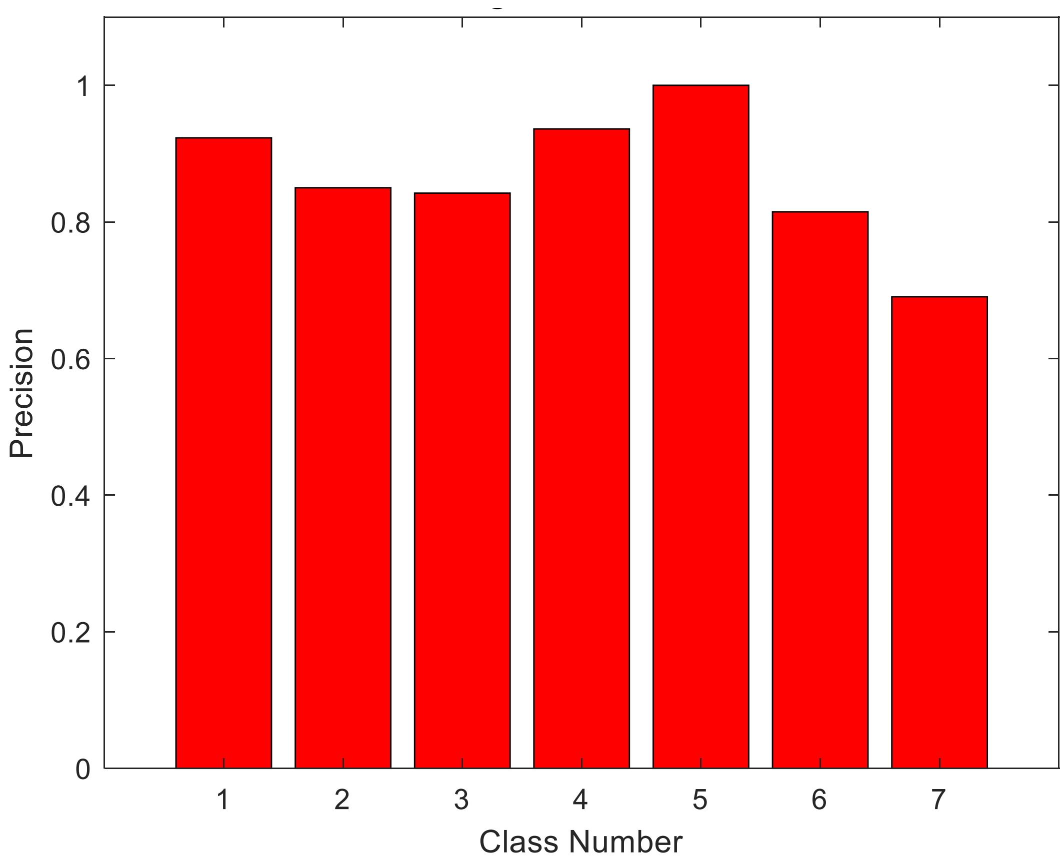


Figure 16: Precision graph of proposed approach for KDEF database

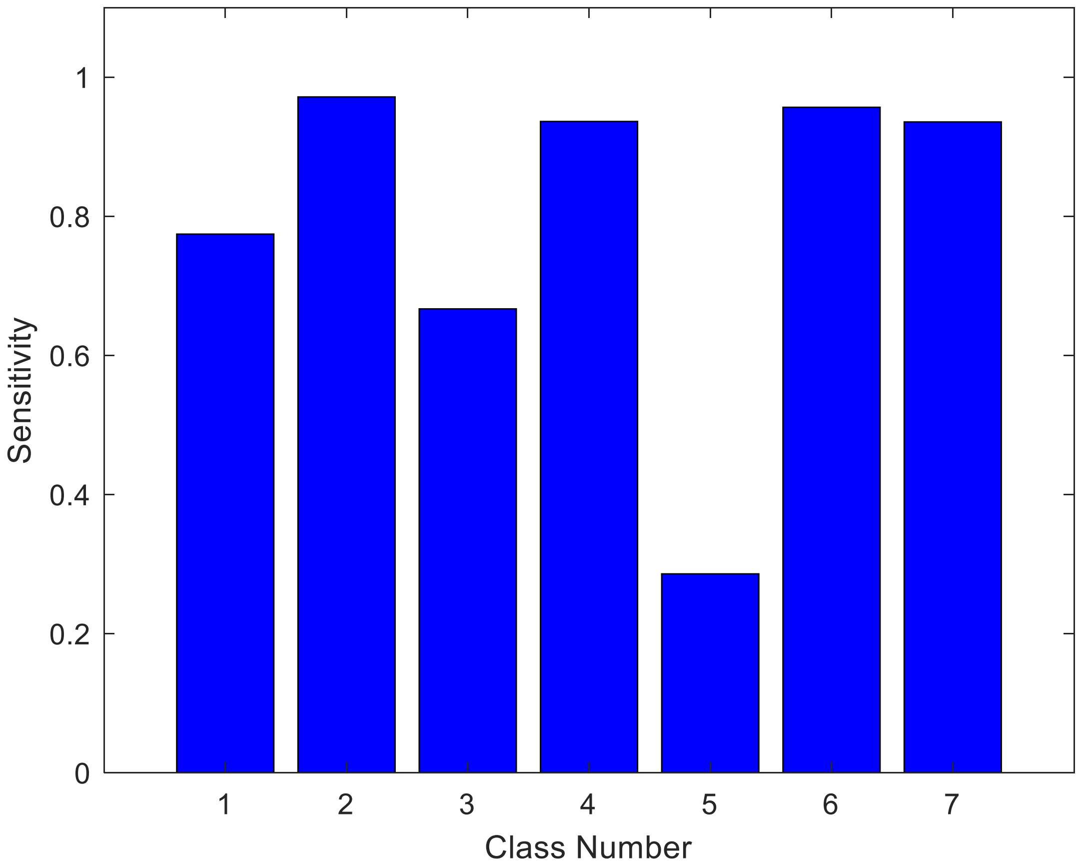


Figure 17: Sensitivity graph of proposed approach for KDEF database

C. Simulation Resutls Using JAFFE Database

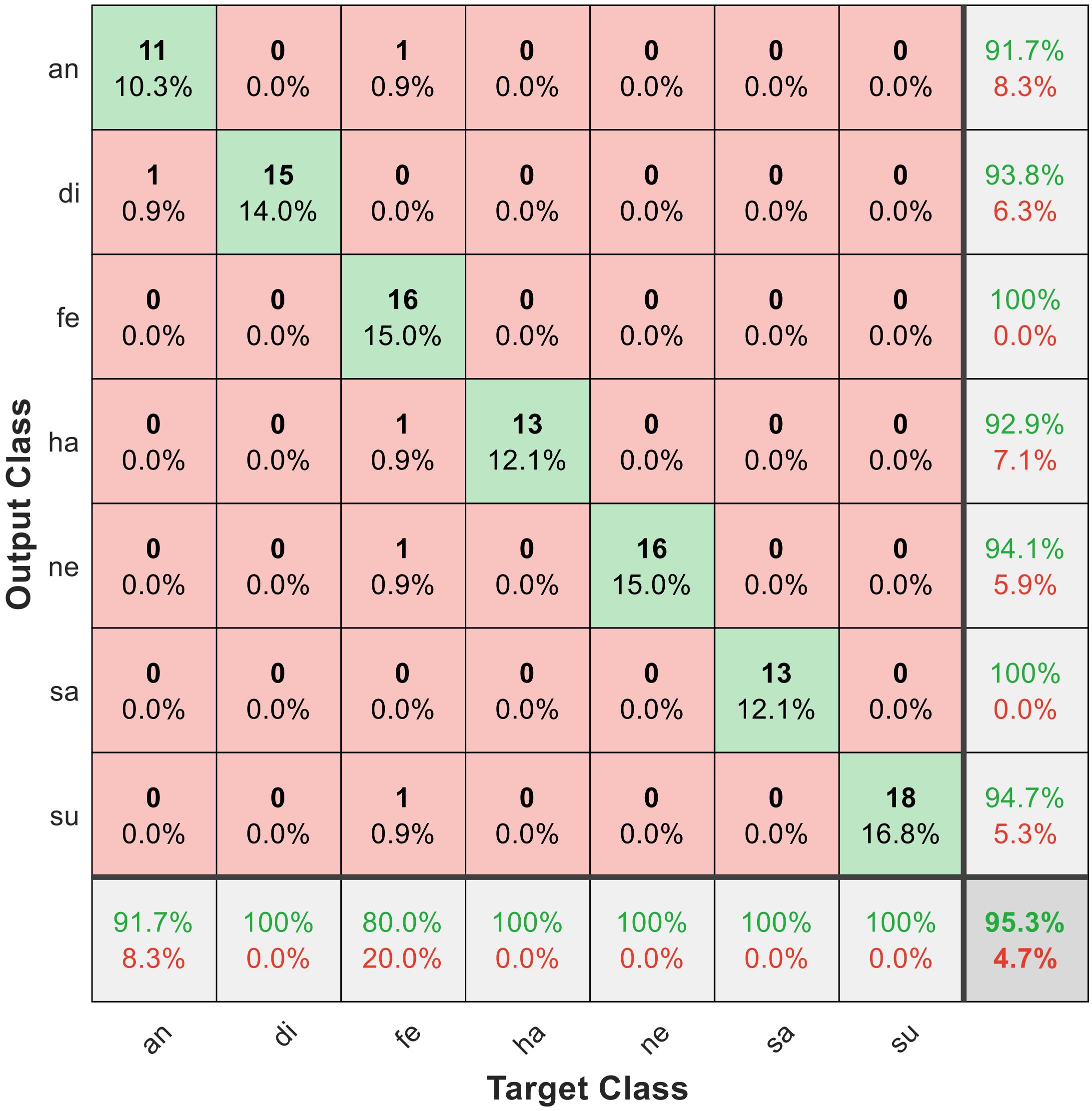


Figure 18: Confusion matrix plot of proposed approach for JAFFE database

Rows and columns represent classes in the JAFFE facial expressions database. We attended 7 classes and each class represented 7 different facial expressions. The confusion matrix plot shown above concludes the specificity for proposed algorithm i.e. 95.33%.

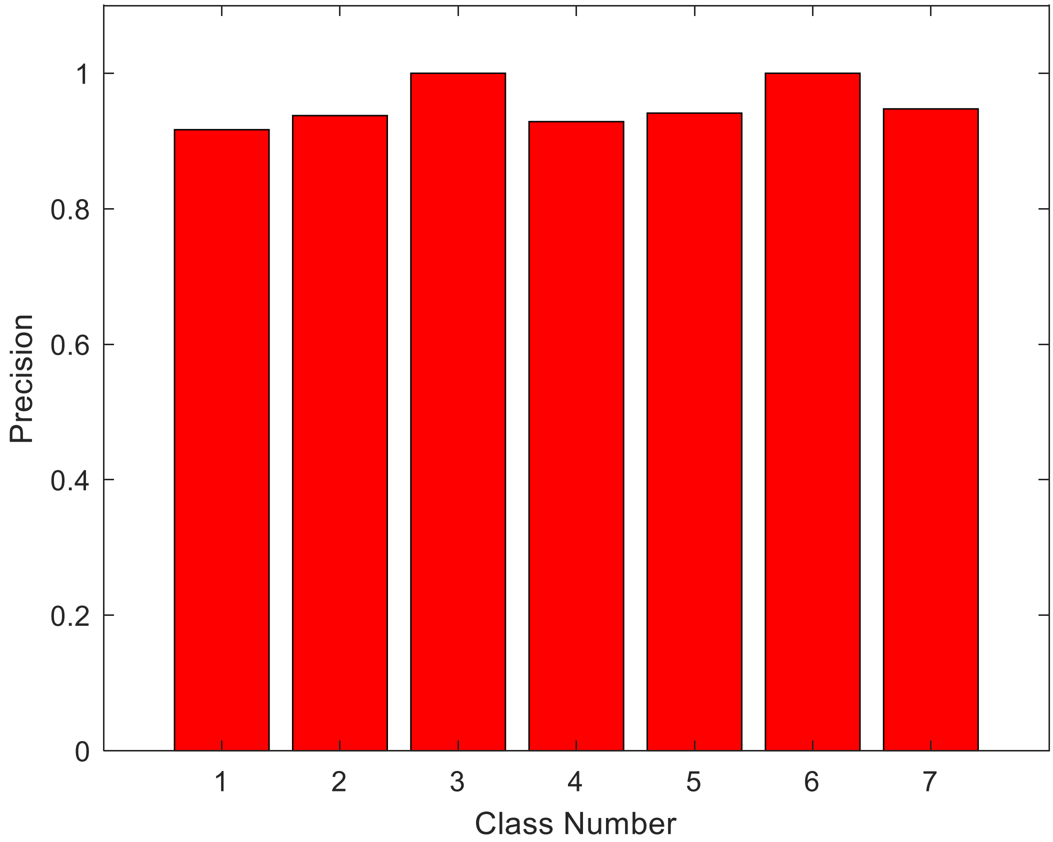


Figure 19: Precision graph of proposed approach for JAFFE database

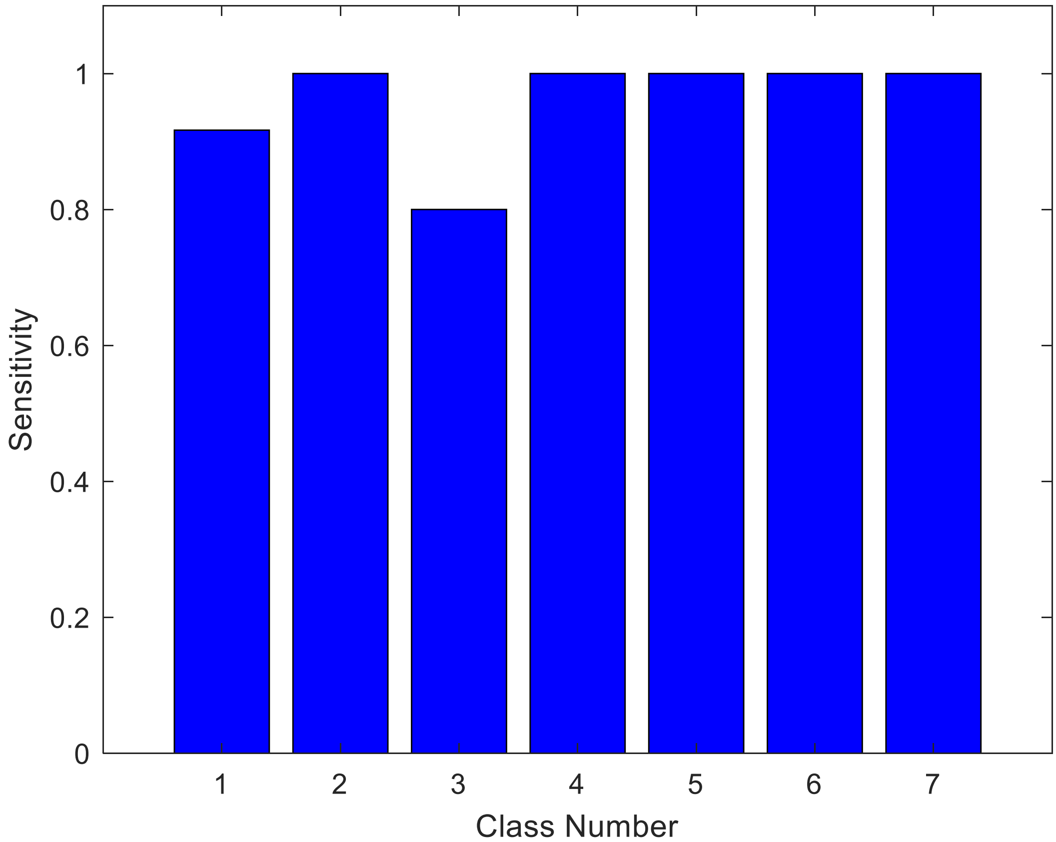


Figure 20: Sensitivity graph of proposed approach for JAFFE database

Table 1: Comparative results for both the databases

Evaluation Parameters	KDEF Database	JAFFE Database
Accuracy	95.54%	98.66%
Error	4.46%	1.34%
Sensitivity	97.40%	99.22%
Specificity	84.39%	95.33%
Precision	97.40%	99.22%
False Positive Rate	15.61%	4.67%
F-Score	97.40%	99.22%
Matthews Correlation Coefficient	0.8179	0.9455
Kappa	0.8179	0.9455

V. Conclusion

In this paper, a facial expression recognition system was developed for seven basic emotions. Features are obtained with three feature extraction approaches based on the HOG, wavelet moments and cartoon descriptors. For the classification stage, Random Forest classifier is used. The important thing is to choose well a specific training dataset for the data environment with which you later want to test. The performance of the proposed method has been tested in the experimental studies carried out. When the three different feature extraction methods are considered, it is seen that successful results are achieved with 95.54% using the KDEF database and 98.66 using the JAFFE database.

References

[1] Rajan, S., Chenniappan, P., Devaraj, S. and Madian, N., 2019. Facial expression recognition techniques: a comprehensive survey. *IET Image Processing*, *13*(7), pp.1031-1040.

[2] Kalsum, T., Anwar, S.M., Majid, M., Khan, B. and Ali, S.M., 2018. Emotion recognition from facial expressions using hybrid feature descriptors. *IET Image Processing*, *12*(6), pp.1004-1012.

[3] Huang, Y., Chen, F., Lv, S. and Wang, X., 2019. Facial expression recognition: A survey. *Symmetry*, *11*(10), p.1189.

[4] Tarnowski, P., Kołodziej, M., Majkowski, A. and Rak, R.J., 2017. Emotion recognition using facial expressions. *Procedia Computer Science*, *108*, pp.1175-1184.

[5] Samadiani, N., Huang, G., Cai, B., Luo, W., Chi, C.H., Xiang, Y. and He, J., 2019. A review on automatic facial expression recognition systems assisted by multimodal sensor data. *Sensors*, *19*(8), p.1863.

[6] Balasundaram, A. and Ashokkumar, S., 2020. Study of Facial Expression Recognition using Machine Learning Techniques. *Journal of Critical Reviews*, *7*(8), pp.2429-2437.

[7] Hariri, W., Tabia, H., Farah, N., Benouareth, A. and Declercq, D., 2017. 3D facial expression recognition using kernel methods on Riemannian manifold. *Engineering Applications of Artificial Intelligence*, *64*, pp.25-32.

[8] Zarbakhsh, P. and Demirel, H., 2017, September. Fuzzy SVM for 3D facial expression classification using sequential forward feature selection. In *2017 9th International Conference on Computational Intelligence and Communication Networks (CICN)* (pp. 131-134). IEEE.

[9] Yang, H. and Yin, L., 2017, October. CNN based 3D facial expression recognition using masking and landmark features. In *2017 Seventh International Conference on Affective Computing and Intelligent Interaction (ACII)* (pp. 556-560). IEEE.

[10] Ruiz, L.Z., Alomia, R.P.V., Dantis, A.D.Q., San Diego, M.J.S., Tindugan, C.F. and Serrano, K.K.D., 2017, December. Human emotion detection through facial expressions for commercial analysis. In *2017IEEE 9th International Conference on Humanoid, Nanotechnology, Information Technology, Communication and Control, Environment and Management (HNICEM)* (pp. 1-6). IEEE.

[11] Ayvaz, U. and Gürüler, H., 2017, May. Real-time detection of students' emotional states in the classroom. In *2017 25th Signal Processing and Communications Applications Conference (SIU)* (pp. 1-4). Ieee.

[12] Agrawal, S. and Khatri, P., 2015, February. Facial expression detection techniques: based on Viola and Jones algorithm and principal component analysis. In *2015 Fifth International Conference on Advanced Computing & Communication Technologies* (pp. 108-112). IEEE.

[13] Tivatansakul, S., Ohkura, M., Puangpontip, S. and Achalakul, T., 2014, September. Emotional healthcare system: Emotion detection by facial expressions using Japanese database. In *2014 6th computer science and electronic engineering conference (CEEC)* (pp. 41-46). IEEE.

[14] Lajevardi, S.M. and Hussain, Z.M., 2009, July. Feature selection for facial expression recognition based on optimization algorithm. In *2009 2nd International Workshop on Nonlinear Dynamics and Synchronization* (pp. 182-185). IEEE.

[15] Ayvaz, U., Gürüler, H. and Devrim, M.O., 2017. Use of facial emotion recognition in e-learning systems. *Інформаційні технології і засоби навчання*, (60, вип. 4), pp.95-104.

[16] Gacav, C., Benligiray, B., Özkan, K. and Topal, C., 2018, May. Facial expression recognition with FHOG features. In *2018 26th Signal Processing and Communications Applications Conference (SIU)* (pp. 1-4). IEEE.

[17] Engin, D. and Ekenel, H.K., 2017, May. Facial expression pair matching. In *2017 25th Signal Processing and Communications Applications Conference (SIU)* (pp. 1-4). IEEE.

[18] Gacav, C., Benligiray, B. and Topal, C., 2016, May. Sequential forward feature selection for facial expression recognition. In *2016 24th Signal Processing and Communication Application Conference (SIU)* (pp. 1481-1484). IEEE.

[19] Özbey, N. and Topal, C., 2018, May. Expression recognition with appearance-based features of facial landmarks. In *2018 26th Signal Processing and Communications Applications Conference (SIU)* (pp. 1-4). IEEE.

[20] Aksoy, N. and Sert, M., 2016, May. Facial action unit detection using variable decision thresholds. In *2016 24th Signal Processing and Communication Application Conference (SIU)* (pp. 2185-2188). IEEE.

[21] Akkoca, B.S. and Gökmen, M., 2015, May. Automatic smile recognition from face images. In *2015 23nd Signal Processing and Communications Applications Conference (SIU)* (pp. 1985-1988). IEEE.

[22] Egorov, A.D., Divitskii, D.U., Dolgih, A.A. and Mazurenko, G.A., 2018, January. Some cases of optimization face detection methodes on image (Using the Viola-Jones method as an example). In *2018 IEEE Conference of Russian Young Researchers in Electrical and Electronic Engineering (EIConRus)* (pp. 1075-1078). IEEE.

[23] Lu, W.Y. and Ming, Y.A.N.G., 2019, June. Face detection based on Viola-Jones algorithm applying composite features. In *2019 International Conference on Robots & Intelligent System (ICRIS)* (pp. 82-85). IEEE.

[24] Dabhi, M.K. and Pancholi, B.K., 2016. Face detection system based on viola-jones algorithm. *International Journal of Science and Research (IJSR)*, *5*(4), pp.62-64.

[25] Mahmud, F. and Mamun, M.A., 2017. Facial expression recognition system using extreme learning machine. *International Journal of Scientific & Engineering Research*, *8*(3), pp.26-30.

[26] Rani, A.M. and Durgadevi, R., 2017. Image Processing Techniques to Recognize Facial Emotions. *International Journal of Engineering and Advanced Technology (IJEAT)*, *6*(6).

[27] Nigam, S., Singh, R. and Misra, A.K., 2018. Efficient facial expression recognition using histogram of oriented gradients in wavelet domain. *Multimedia tools and applications*, *77*(21), pp.28725-28747.

[28] Nazir, M., Jan, Z. and Sajjad, M., 2018. Facial expression recognition using histogram of oriented gradients based transformed features. *Cluster Computing*, *21*(1), pp.539-548.

[29] Eng, S.K., Ali, H., Cheah, A.Y. and Chong, Y.F., 2019, November. Facial expression recognition in JAFFE and KDEF Datasets using histogram of oriented gradients and support vector machine. In *IOP Conference series: materials science and engineering* (Vol. 705, No. 1, p. 012031). IOP Publishing.

[30] Daugman, J.G., 1988. Complete discrete 2-D Gabor transforms by neural networks for image analysis and compression. *IEEE Transactions on acoustics, speech, and signal processing*, *36*(7), pp.1169-1179.

[31] Navastara, D.A., Wiranto, H., Fatichah, C. and Suciati, N., 2021. Facial Expression Recognition Using Wavelet Transform and Convolutional Neural Network. In *Advances in Computer, Communication and Computational Sciences* (pp. 941-952). Springer, Singapore.

[32] Breiman, L., 2001. Random forests. *Machine learning*, *45*(1), pp.5-32.

[33] Tiwari, K. and Patel, M., 2020, February. Facial Expression Recognition Using Random Forest Classifier. In *International Conference on Artificial Intelligence: Advances and Applications 2019* (pp. 121-130). Springer, Singapore.

[34] Goeleven, E., De Raedt, R., Leyman, L. and Verschuere, B., 2008. The Karolinska directed emotional faces: a validation study. *Cognition and emotion*, *22*(6), pp.1094-1118.

[35] Michael, L., Kamachi, M. and Gyoba, J., 1998. The Japanese female facial expression (JAFFE) database. In *Proceedings of the third international conference on automatic face and gesture recognition*.