Multi-Level Fusion for Facial Expression Recognition in Human Behavior Identification

Aqeel Hussein[1], Ibraheem H. M.[2], Sarah Ali Abdulkareem[3], Ryam Ali Zubaid[4], Noor Thamer[5]

[1]Department of Medical device technology Engineering, Alfarahidi University, Baghdad, Iraq
[2] Computer Communications Engineering Department, Alrafidain University College, Baghdad, Iraq
[3] Department of computer engineering techniques, Al-turath University College, Baghdad, Iraq
[4] Department of computer engineering techniques, Mazaya University college, Thi Qar, Iraq
[5] Accounting Department, Al-Mustaqbal University College , 51001 Hillah, Babylon , Iraq
Emails: aqeel.hussien@alfarahidiuc.edu.iq; ibraheem.hatem.elc@ruc.edu.iq; sarah.ali@turath.edu.iq; ryamzubaid34@gmail.com; noorthamer2020@mustaqbal-college.edu.iq;

Abstract

In this study, we present a multi-level fusion of deep learning technique for facial expression identification, with applications spanning the fields of cognitive science, personality development, and the detection and diagnosis of mental health disorders in humans. The suggested approach, named Deep Learning aided Hybridized Face Expression Recognition system (DLFERS), classifies human behavior from a single image frame through the use of feature extraction and a support vector machine. An information classification algorithm is incorporated into the methodology to generate a new fused image consisting of two integrated blocks of eyes and mouth, which are very sensitive to changes in human expression and relevant for interpreting emotional expressions. The Transformation of Invariant Structural Features (TISF) and the Transformation of Invariant Powerful Movement (TIPM) are utilized to extract features in the suggested method's Storage Pack of Features (SPOF). Multiple datasets are used to compare the effectiveness of different neural network algorithms for learning facial expressions. The study's major findings show that the suggested DLFERS approach achieves an overall classification accuracy of 93.96 percent and successfully displays a user's genuine emotions during common computer-based tasks.

Keywords: Artificial Intelligence, Facial Expressions; Deep fusion features; Smart Decision System; Behavior traits.

1. Introduction

Due to advancements in deep learning and machine learning, we are able to automatically recognize facial emotions [1]. Hybrid face expression recognition systems (DLFERS) use a support vector machine for feature extraction and classification to determine human activity [2]. Because the eyes and the mouth are so responsive to human expression changes and so important for interpreting emotional expressions, DLFERS employs an information classification algorithm to create a new fused picture made of two integrated blocks of eyes and mouth [3]. DLFERS' feature extraction is significantly more accurate and efficient when multi-level fusion deep learning techniques are used, such as Transformation of Invariant Structural Features (TISF) and Transformation of Invariant Powerful Movement (TIPM) [4].

Expression of emotion and intent through facial expressions is essential in social interactions [5]. Face recognition and image preprocessing, function retrieval, and expression classification are the three cornerstones of facial expression recognition [6]. Convolutional neural networks (CNNs) and recurrent neural networks (RNNs) are two examples of deep learning approaches that may be fused together at several levels to improve this procedure [7,8]. These methods can improve recognition accuracy by extracting deeper and more reliable elements from facial emotions. Emotions may also be recognized by artificial intelligence, which does this by learning what each facial expression means and then applying that information to new data [9,10]. Multi-level fusion of deep learning algorithms may also be used to create more advanced systems for properly identifying nuanced human emotions and expressions [11-13].

Doi: https://doi.org/10.54216/FPA.100210
Received: November 23, 2022 Accepted: March 17, 2023

Such expressions reflect the viewer's psychological response. It's a non-verbal means of communication. In interpersonal relations, it serves a communicative function [14]. Recognition for Facial Expression is an image processing issue in the broader area of computer vision. Image Classification problems can be found where images from different classes need to be computationally allocated to a name. [15] The representations are of human aspects, and the categories in Face Expression Recognition (FER) structures, particularly, are a collection of emotions. All FER approaches machine learning to involve a set of explanations, each with a single category of excitement. Seven classifications of emotion are also used as normal and are categorized into the expression of happiness, sadness, anger, surprise, fear, disgust, and neutral, which is illustrated in figure 1.

Figure 1: Classification of Facial Expressions

Various human emotions can not be described simply by subtle facial expressions, frequently displayed similarly in emotions such as rage and disgust. Everyone's emotional gestures can be very unusual, with different twists and facial signs. The photos to be listed can include a wide variety of different positions and orientations of people's heads. FER is harder than most other Picture Classification functions for such sorts of reasons. Even then, excellently-designed machines can deliver reliable results if restrictions all through advancement are considered. Facial expressions, in different ways, are used by people to communicate different meanings. The spectrum of significations encompasses theoretically inherent social and psychological constructs like 'surprise' and 'with care' in terms of nuanced, culture-specific concepts.

In this study, proposes a Deep Learning assisted Hybridized Face Expression Recognition system (DLFERS) has been proposed and from a single image frame that utilizes feature extraction and classification with a support vector machine for identifying human behavior. The main contributions of this study are:

- The study proposes a improved approach to facial expression identification using a multi-level fusion of deep learning techniques; the method is named DLFERS. Human social/physiological interaction

Doi: https://doi.org/10.54216/FPA.100210
Received: November 23, 2022 Accepted: March 17, 2023

detection, cognitive science, personality development, and mental disease diagnosis are only few of the many potential applications of this method.

- Using a support vector machine (SVM) for feature extraction and classification, the suggested technique can determine human activity from a single frame of video. Emotional states may be accurately interpreted from face features using this method.
- The method incorporates an information classification technique to generate a new fused image made up of two integrated blocks of eyes and mouth, regions of the face that are especially sensitive to changes in human expression and thus relevant to the task of decoding emotional expressions.
- The suggested technique makes advantage of the Transformation of Invariant Structural Features (TISF) and the Transformation of Invariant Powerful Movement (TIPM) to extract features from data stored in a "Storage Pack of Features" (SPOF).
- The study assesses the efficacy of comprehensive facial-expression learning using a number of different neural network techniques and data sets. The results show that the suggested DLFERS approach is able to obtain an overall classification accuracy of 93.96%, which is quite effective in exposing a user's genuine emotions when dealing with basic computer-based activities.

The remaining work is given as follows: part 2 provides insights about background studies, part 3 discusses Storage Pack of Features (SPOF) that uses Transformation of Invariant Structural Features (TISF) and Transformation of Invariant Powerful Movement (TIPM) for feature extraction of facial expression recognition. Finally, part 5 concludes the research with future scope

2. **Related Works**

This section discusses several works that have been carried out by researchers; Chao Q et al. [16]. A new approach to recognition of the term is proposed in this paper, which is focused on awareness and binary designs. The method to remove facial contours is based on the LBP operator. Second, a pseudo-3-D model is often used to divide the face into 6 facial expressions. In this sense, the sub-sectors and the regional facial expression photos use the marked LBP objective of feature extraction and then use the supporting vector machine in two classrooms and the softmax with two classes of emotional classroom models, the simple emotional model, and the world's leading emotional model.

Xiaoming Zhao et al. [17] proposed a new method of facial expression recognition based on profound learning combined with multi-layer perceptrons (MLP). This includes the benefit of unattended functional training from the DBNs, with the MLP identification advantage. The experimental results on the JAFFE show the encouraging output of the experimental facial expression recognition process, which outperforms the other state-of-the-art classification systems.

Fadi Dornaika et al. [18] studied the output of the classifiers which exploit the independent temporal parameters of face reaction, which a 3D facial tracker gives with systematic 3D head and facial activity. The usage is clear in the identification pose and the texture. The first approach is a static time-spinning technique used to identify emotions where temporary fingerprints of various uniform facial expressions generate learning data. The first one is based on the static period distorting methodology.

Hayet Boughrara et al. [19] proposed a constructive Multi-Layer Perceptron (MLP) training algorithm for face recognition implementations. A single layer of hidden neurons and minimal training patterns composes the developmental algorithm. The proposed building MLP training algorithm gives a synthesis. Parameters, such as the number of behaviors, the number of objects during the training stage, and the MSE predefined value, for each sub-set of the class, are to be initialed throughout the learning process. The proposed algorithm is designed to classify the face. A biological facial description, namely Perceived Facial Pictures PFI, has been used for removing characteristics from human face images for the functional extraction phase.

Khandait SP et al. [20] present an approach to the question of automated retrieval of facial features from an already prefrontal picture and classification and recognizance of facial expression and, therefore, of an individual's expression and condition. Feedback propagation is used to classify the facial emotions into seven fundamental categories, including surprise, neutrality, sorrow, disgust, fear, happiness, and rage. Morphological image processing operations are used to classify and position the face part. JAFFE facial expression database is experimented with and increases efficiency in respect of 100 percent training set accuracy and 95.26 percent test set accuracy.

Hakan Boz et al. [21] proposed that artificial intelligence techniques are sometimes employed in various research topics that conventional computer solutions cannot solve. In this context, it is one of the many common fields of interest in recent years to employ computational modeling to understand complex features. The purpose

Doi: https://doi.org/10.54216/FPA.100210

Received: November 23, 2022 Accepted: March 17, 2023

of this analysis is to provide explanations so far, Introducing work where a method has been built and established which can derive feelings from personal facial expressions. The system here suggested an artificial neural network of Cascade Feedforward developed by a recent optimization algorithm known as the Vortex optimization algorithm.

Kaihao Zhang et al. [22] analyzed facial expressions in time sequences using the Part-Based Hierarchical Bidirectional Recurring Neural Network (PHRNN). Our PHRNN modeling facial morphological variation and dynamic creation of phrases successfully extract "time-related characteristics" from successive marks based on visual marks. Meanwhile, a Multi-Signal Convolutionary Neural Network (MSCNN) is being developed to retrieve information that is still present. These profound Evolutionary Spatial-Temporary Networks (PHRNN and MSCNN) extract all-partial geometry and dynamic information, thus effectively enhancing the efficiency of the recognition of facial expressions.

Isidoros Perikos et al. [23] presented an emotion recognition device to instantly identify the basic emotional conditions of human facial expressions. At first, the device analyzes the face image and detects and calculates characteristic human facial abnormalities, such as eyes, eyebrows, and mouth. A neural multi-layer system can then be used to identify the expression of the face in the right state of emotion. JAFFE database pictures of real humans have tested the device, and the findings obtained are very acceptable.

Tong Zhang et al. [24] introduced a new form of feature learning powered by the deep neural network (DNN) is used for the recognition of facial expressions (FER). In this way, invariant feature transformers (SIFTs) are first extracted from each image of the facial, which corresponds to one set of landmarks of every other facial image. Then, a SIFT-derived feature vector matrix is used to process data for training optimized discriminatory characteristics for speech identification in a well-designed DNN model. The DNN model suggested provides many layers to describe the partnership between functionality matrices of SIFT and their strong-level conceptual data.

Guihua Wen et al. [25] In this paper, an assembly of convolutionary neural grids is presented with a probability-based facial expression hybrid detection that adapts the design of CNN to the first layer and multiple hidden maxout layers using the linear activation layer. It was designed around the optimum values of CNN using arbitrarily specific parameters and design, where each CNN was the base grade Educated to give every class a chance. The probability-based hybrid algorithm then merged these variables. Our process, which had better exactness than comparisons, was evaluated through experiments on benchmarking data sets.

Based on the statistical survey, a Deep Learning assisted Hybridized Face Expression Recognition system (DLFERS) has been proposed from a single image frame that utilizes feature extraction and classification with a support vector machine for identifying human behavior

3. Deep Learning assisted Hybridized Face Expression Recognition system

The Deep Learning assisted Hybridized Face Expression Recognition system (DLFERS) has been proposed from a single image frame that utilizes feature extraction and classification with a support vector machine to identify human behavior. DLFERS uses the Transformation of Invariant Structural Features (TISF) and the Transformation of Invariant Powerful Movement (TIPM) for feature extraction within the segmented section. The Storage Pack of Features (SPOF) is composed of TISF and TIPM, and it is mainly used for reduction comparison in both the training and testing phase. The facial expression recognition and the behavioral task are validated with an SVM classifier to enhance the performance in identifying human facial expressions like happiness, fear, surprise, anger, sad, disgust, and neutral.

The key steps in the proposed system DLFERS are preprocessing, extraction of features, codebook creation, and classification. There are a variety of measures in the proposed emotional recognition system. Feature-based methods in new research are favored in identifying feelings, as it is a difficult problem to adapt a design to different face kinds and forms. The models using universal facial characteristics through segmenting or adding regional location data are easy and quicker, and with a transition of subject position and illumination, the accuracy rate decreases. However, due to its robustness to lighting and changes, algorithms that work with local characteristics are more suited to normal psychological recognition. Specific function descriptors in the chosen areas were shown to be excellent for tasks such as object recognition, image matching, and object processing and classification

Doi: https://doi.org/10.54216/FPA.100210
Received: November 23, 2022 Accepted: March 17, 2023

3.1. Preprocessing

The first preprocessing step involves using the face detection algorithm; the facial component of the input image or photos is detected. A part of the identified face is cut from the left, right, and top. (Q_1, Q_2) and (Q_3, Q_4) are the left and right corners of the identified face image.

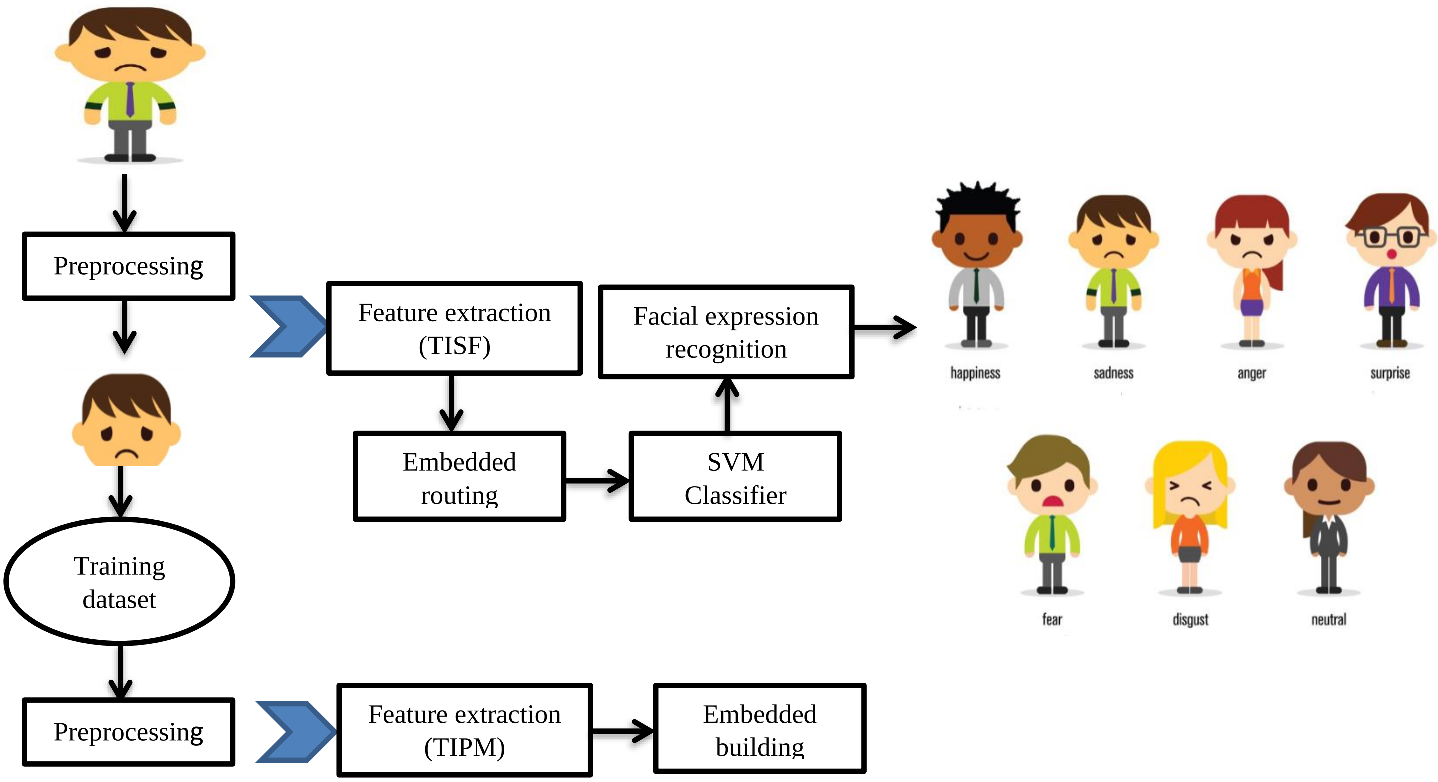


Figure 2 Flow of Deep Learning assisted Hybridized Face Expression Recognition system

The pixel value representation for the left and right of the cropped image is denoted as (Qm_1, Qm_2) and (Qm_3, Qm_4) and the calculated value is illustrated shown in Equations (1) to (4)

$$Q_{m1} = Q_1 + [\gamma K] \tag{1}$$

$$Q_{m2} = Q_2 + [\delta L] \tag{2}$$

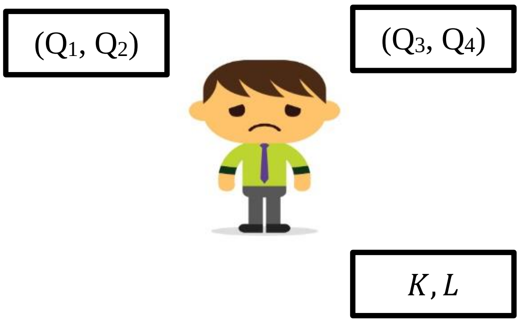
$$Q_{m3} = Q_3 + [\gamma K] \tag{3}$$

$$Q_{m4} = Q_4 + [\delta L] \tag{4}$$

Here K and L are horizontal and vertical measurements of the image. δ and γ are the horizontal and vertical image path factors for cropping, respectively. The cutting system offers a central face that can be seen whenever the feeling is demonstrated in the significant changes. The system output had an adverse impact on the bottom side of face images, particularly when surprise emotion is detected.

DLFERS involves the preprocessing, feature extraction, and classification of facial expressions, as shown in figure 2. The input image after cropping the left and right corners of the image is shown in figure 3(a). Only the facial part is concentrated, and the optimized values for the cropping scaling factor are given as δ and γ, and it is illustrated in figure 3 (b).

Doi: https://doi.org/10.54216/FPA.100210
Received: November 23, 2022 Accepted: March 17, 2023



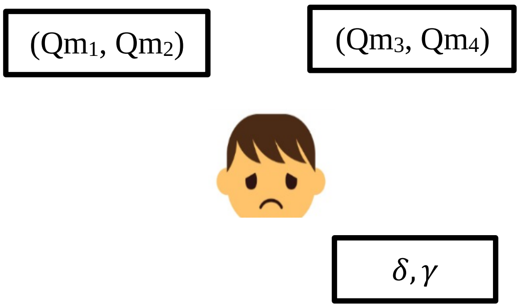


Figure 3 : Left and right corner cropping Figure 4: After cropping the image

Distinct measurements of the pictures cut in and taken from JAFFE data sets. All images are multiplied with vector quantization such that the set of attributes in each picture is removed. Furthermore, these flexible photos and the corresponding wavelet transform are used. Data is contained in the subband of low-low frequency. Wavelet Transform is used, and the low-low frequency substring is retained, reducing the amount of pixels to process. The down-samples also show that certain point face areas are still included, which remain constant if various feelings are explained. This may lead to similarities in the corresponding segmentation process between conflicting faces.

3.2. Feature Extraction

According to the figure, the differentiated face is separated into 29 components of various sizes in the pre-processing stage. 4. The complete face is divided into 2 × 2 and 4 × 4 segments. The total face area is divided into the upper and lower part. The upper part includes the area from head to eyes, and the lower part is the remaining area of the face. The principle of TISF is introduced for the retrieval of the main points from the image when TISF data is retrieved separately within each area.

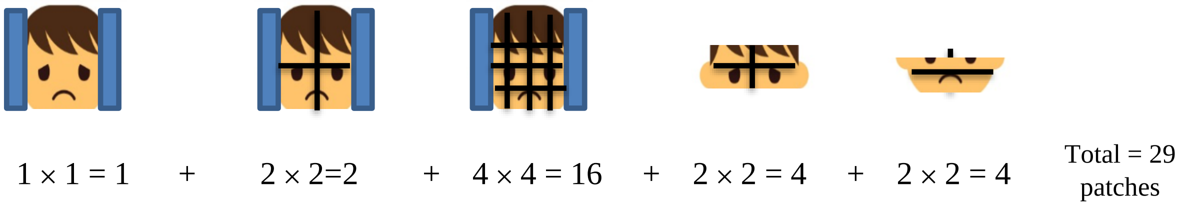


Figure 5: The differentiated faces of various sizes

All patches invariant to the varying image transformations result in powerful functionality. The weight matrix for each picture is a combined function derived globally from the chosen areas, which integrates location data. In the case of TISF selection and extraction, they give a clear view of an object described by the feature A(r, s, μ) determined by a logarithmic function GF (r, s, μ) input feature IP (r, s). The logarithmic function A(r, s, μ) and GF (r, s, μ) is calculated as shown in Equations (5) and (6)

$$A(r, s, \mu) = GF(r, s, \mu) * IP(r,s) \tag{5}$$

$$GF(r,s,\mu) = \frac{1}{2\pi\mu^2} e^{-(a^2+b^2)/2\mu^2} \tag{6}$$

Here μ is the gaussian probability density variance. The gradient magnitude of the n (r, s) and position of the $\emptyset$ (r, s) is determined using the following pixel differences, and it is illustrated given in Equations (7) and (8)

Doi: https://doi.org/10.54216/FPA.100210

Received: November 23, 2022 Accepted: March 17, 2023

$$n(r,s) = \sqrt{((K(r+1,s) - K(r-1,s))^2 + (K(r,s+1) - K(r,s-1))^2} \tag{7}$$

$$\emptyset(r,s) = \tan^{-1}[\frac{K(r,s+1)-K(r,s-1)}{K(r+1,s)-K(r-1,s)}] \tag{8}$$

A function identifier of 1 × 128 and 1 × 64 sizes is quantified for TISF and TIPM separately, mostly with the aid of the differential intensity and direction variables. A kernel function is calculated on the strength of a size 4 × 4 regions per stage identified.

3.3. Embedded building

An embedded building is used to view broad-face images compactly. Related objects are grouped into a given number of observations. Each group has a simple average value that covers the total cluster and contributes to an essential decrease in functionality. In DLFERS, the SPOF-based embedding development information selected for each expression is all quantified, as shown in the figure. 5. Figures extracted from all images of a particular emotion are initially grouped and distributed by clustering methods.

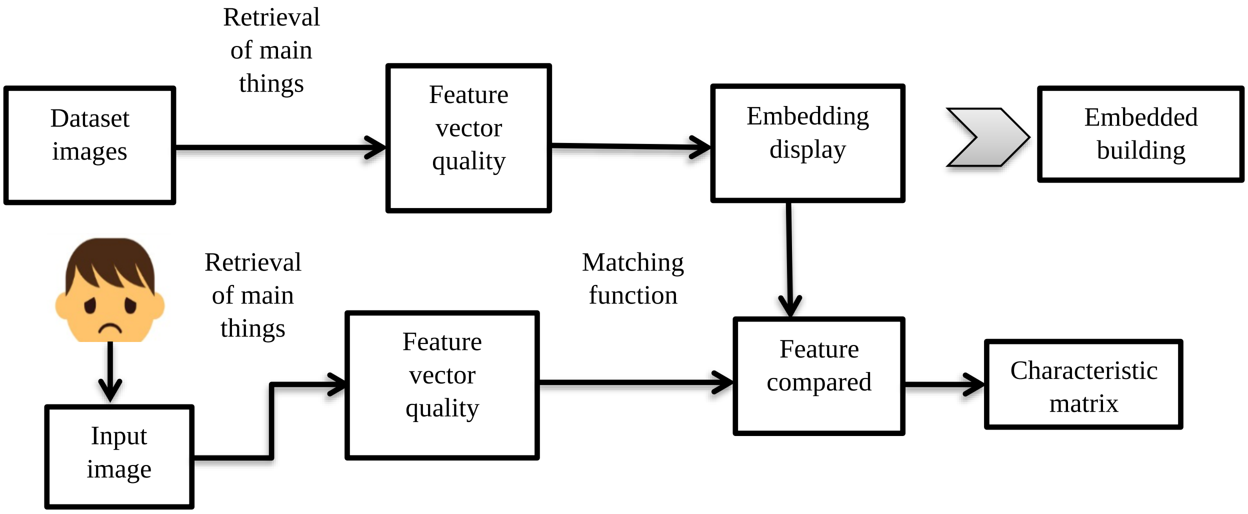


Figure 6: SPOF-based embedded development

SPOF-based embedded development creates the vector image method for classification training and validation. For feature representation, the closest embedding neighbors for all features are obtained on a quick database level for the closest neighbors, which calculates the biggest data obtained from the nearest neighbors easily. Regardless of the drop in the rate of function similarities, the calculation time is decreased using the database. Rather than each particular target within each image in the training sets, each 1 × 128 extracted feature is now being compared with D number of clusters. Finally, a histographic vector between each picture is produced. A correlating expression symbol for the source of the training sample is allocated to the picture input vectors. A function vector called m × D is used for the performance measurement where m is the number of pictures for learning in a set of data and D is the correct symbol.

In the experiment phase, the test picture for the expression tag is allocated for pixel mapping of the full test digital images. The SVM classification for identifying emotional responses is designed to obtain functionalities in both the training and evaluation process. It is a classification method that adapts and divides the data sets into groups. SVMs are controlled training methods and are perfectly suited to classification tasks. Multi-class categorization can also be lengthened to integrate the SVM. SVM classification efficiency depends on the matrix type, for example. Normal, quadratic, etc. In this proposed method, SVM classification systems from multiple classes are educated related to a Storage Pack of Features (SPOF) of Transformation of Invariant Structural Features (TISF) and Transformation of Invariant Powerful Movement (TIPM). All D embedding inputs are compared to each feature vector, and the pixel bin for the maximum deciding set is increased. The descriptor is now being used as an operating frequency by each group core. Picture function variable for training and validation of the classifier. The SVM classifier achieves the informative classifier training process based on derived functionalities. The training set for SPOF-TISF and SPOF-TIPM functionality trains a total of M classifications, in which M represents an overall number of courses. TISF and the TIPM descriptors 1 × 128 and 1 × 64 measurements define distinguishing figures from differentiated areas. Depending on histograms, such

Doi: https://doi.org/10.54216/FPA.100210
Received: November 23, 2022 Accepted: March 17, 2023

identifiers are then plotted to a code word for matrix building. At the finishing point, both TISF and TIPM classifications are generated randomly. The qualified SVM classification is equipped in a series to determine a feeling, and the picture is expected to indicate the category in which a game is located. SVM is a classification algorithm used to describe the images using a decision boundary. It is able to determine both linearly and non-linearly. SVM uses iterative methods to find the optimum vectors. A limited sample of test images called a training set is used for this appropriate algorithm feature space choice. Essentially the attribute values are the models on each unit. Estimating such training examples for the ideal hyper-plane corresponds with the resolution of sequentially restricted binomial optimization problems.

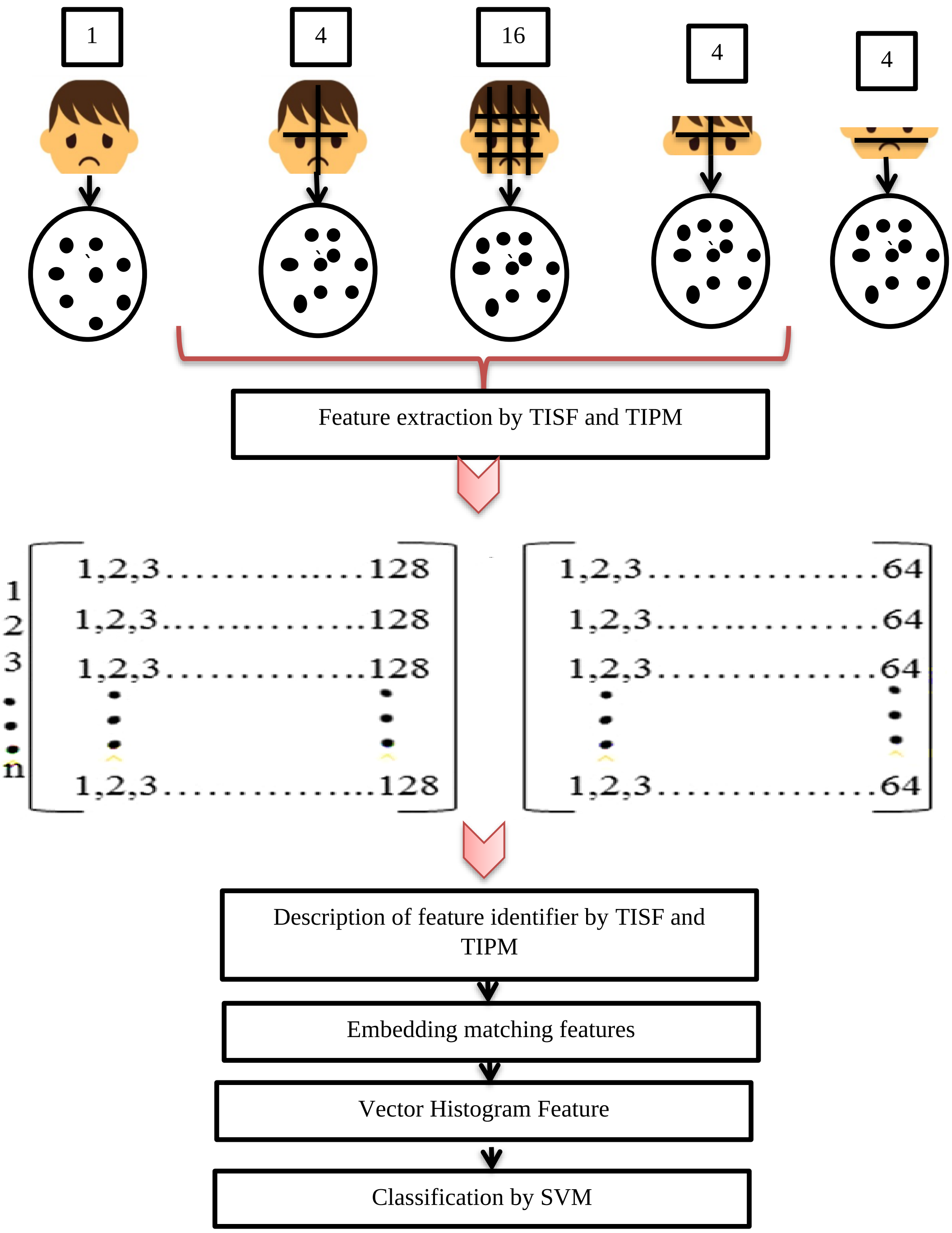


Figure 7: Feature extraction and classification method

The key distinction between all the classifiers and the SVM is that all the other classification algorithms are focused on reducing the risk of the experimental analyses, i.e., reducing testing errors, whereas the SVM is focused on reducing the systemic risks to reduce the anticipated training error in the invisible training set. The normal function of SVM is illustrated shown in Equation (9)

Doi: https://doi.org/10.54216/FPA.100210
Received: November 23, 2022   Accepted: March 17, 2023

$$g(y) = s(\left(\sum_j \beta_j K_j(r_j, r)\right) + a) \tag{9}$$

Here $g(y)$ is the kernel function, r_j is the pixel value, β_j is the scaling factor, $g(y)$ is the Gaussian function.

The SVM is modified by using the kernel function, and it is illustrated shown in Equation (10)

$$g(y) = \sum_{j=1}^{M_r} \beta_j Z_j l(y_j, y)) + a \tag{10}$$

M represents the overall vectors in figure 6. 1, 4, 16, 4, 4 is the number of segments in the input image, and TISF and TIPM carry out the feature extraction. These extracted features are then matched with the embedded codebook. The final step involves the classification by SVM. The findings of comprehensive facial emotional recognition studies employing feature-based methods with conversational speech segmentation are described in DLFERS. Furthermore, the differentiated face region is pre-processed to hold critical details only. For organized spatial knowledge collection, the spatial SPOF method is used. TISF and TIPM descriptors are used in the proposed method to remove feature lights, sizes, and rotations from the differentiated facial patches. For the purpose of the embedded word construction, an ordered spatial SPOF function is used. The entire facial forehead is taken in one method, and the face is divided into the upper and lower parts. The whole facial expression is classified mainly by the movement of the mouth and eyes, and the performance accuracy is obtained by SVM.

Based on the Deep Learning assisted Hybridized Face Expression Recognition system (DLFERS), it is clear that SVM achieves feature extraction and classification performance from a single image frame. The reactions recorded are then compared with the expected reaction of the image. Feature extraction is achieved by a Storage Pack of Features (SPOF) that uses the Transformation of Invariant Structural Features (TISF) and Transformation of Invariant Powerful Movement (TIPM).

4. Results and discussions

In this Section, Deep Learning assisted Hybridized Face Expression Recognition system has been validated by conducting various experiments on the JAFEE (Japanese Female Facial Expression) data set, which has a series of images showing various facial expressions.

4.1. JAFEE Datasets

Including six fundamental images (sadness, surprise, joy, anger, fear, and disgust) and one neutral expression, this data pack contains 213 images. The images were rated as belonging to a specific expression by 60 Japanese subjects. The picture is listed as the highest-voted word. Each disclosure of such information for a given subject begins with an impartial expressive framework in both datasets and finishes mostly with the top phrase slot. Deep Learning assisted Hybridized Face Expression Recognition system trains seven classifiers based on corresponding expression pictures, i.e., sadness, surprise, joy, anger, fear, disgust, and neutral, as illustrated in table 1. After experiments on the justification of the mapped and transformed images, such variables are specifically chosen by using wavelet transforming photos, and the overall result of emotion detection is equated with a capability that allows large time savings.

Table 1: Classification Accuracy for facial expression

Accuracy	Scaled Image	Entire Face	Scaling Factor 1	Scaling Factor 2	Scalinf Factor 3
happy	95.89	94.33	96.89	98.77	97.86
Sad	80.12	80.09	79.98	81.23	88.9
Surprise	86.78	85.67	88.66	96.77	95.88
Disgust	83.45	82.33	91.23	87.66	96.77
Fear	76.65	76.88	61.34	95.17	95
Anger	83.67	83.44	87.56	98.99	92.33
neutral	88.99	81.22	77.89	97.89	90.98
Average	85.07	83.42	83.36	93.78	93.96

The classification accuracy for all facial expressions like happy, sad, surprised, disgusted, fearful, angry, and neutral is illustrated in figure 7. The specificity of emotional understanding of Deep Learning assisted Hybridized Face Expression Recognition system is measured by various criteria (a) Feature selection is achieved by selecting the region of interest from the whole face, (b) The spatial detection data achieve the

Doi: https://doi.org/10.54216/FPA.100210
Received: November 23, 2022 Accepted: March 17, 2023

classification rate used for emotion detection, (c) The cluster center is used to produce the function of the embedded building.

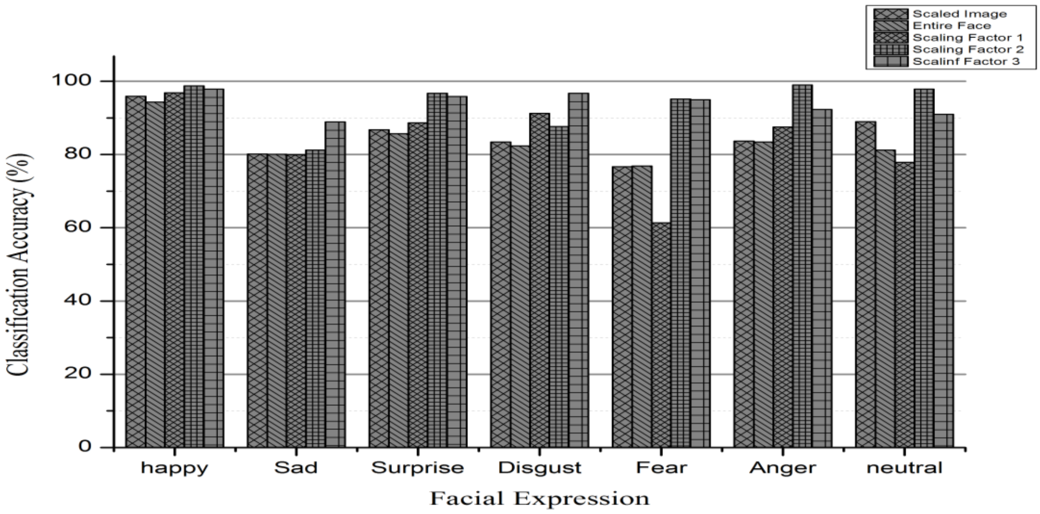


Figure 8: Classification Accuracy for facial expression

The graph is drawn between scaled image, complete face, scaling factor1 $\delta = 0.10, \gamma = 0.7$, scaling factor 2 $\delta = 0.12, \gamma = 0.15$, scaling factor 3 $\delta = 0.15, \gamma = 0.15$. High classification accuracy is obtained for the scaling factor 3. The scaling factor is nothing but cropping the face image; that is, all unwanted places in the images are removed. The highest scaling factor gets the highest classification accuracy. Accuracy with different patches with all segmentation levels is illustrated in table 2

Table 2: Performance Accuracy with different patches with all segmentation levels

Patches	Patches 1	Patches 2	Patches 16	Patches 4 (upper part)	Patches 4 (lower part)
Happy	96.6	93.44	97.01	98.01	97.11
Sad	78.9	81.99	80.05	80.33	87.66
Surprise	80.11	84.33	89.01	97.03	94.21
Disgust	86.33	80.01	90.03	88.67	95.66
Fear	79.01	77.89	62.89	94.55	98.12
Anger	80.12	84.66	88.56	92.12	90.12
neutral	85.67	82.77	78.66	98.89	89.02

The performance accuracy is calculated for all the patches of the segmented image, and it is illustrated in figure 8. Patches1 is nothing, but the image is not segmented; patches2, the image is divided into two groups; patches 16, the image is divided into 16 groups; patches 4 (upper part), the upper part of the face includes the eyes part, patches 4 (lower part) the lower part of the face include the mouth section.

Doi: https://doi.org/10.54216/FPA.100210
Received: November 23, 2022 Accepted: March 17, 2023

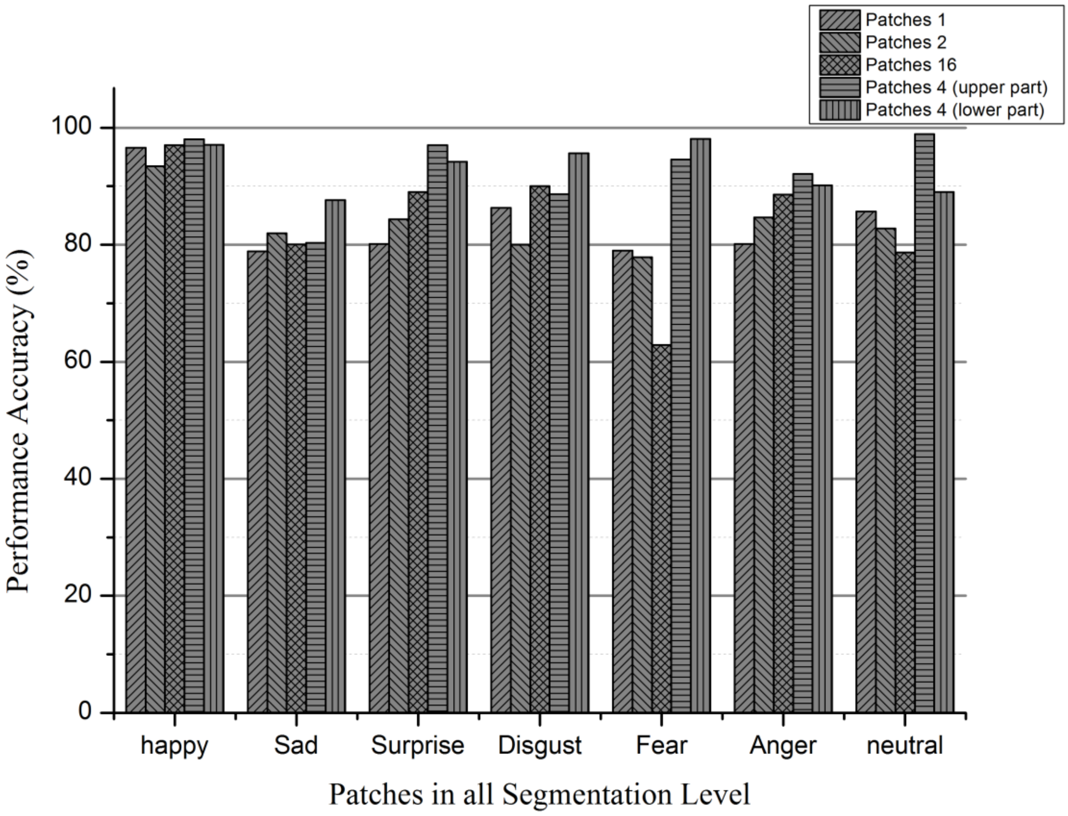


Figure 9: Performance Accuracy with different patches with all segmentation levels

The performance accuracy for clusters 20, 40, 60, 80, and 100 is achieved in embedding construction. Face image is divided into clusters, and the clusters also define its accuracy, and it is illustrated in table 3. Rather than each particular target within each image in the training sets, each 1×128 extracted feature is now being compared with D number of clusters.

Table 3: Performance Accuracy for various clusters in face image

samples	Clusters 20	Clusters 40	Clusters 60	Clusters 80	Clusters 100
happy	85.66	91.12	95.12	96.77	96.12
Sad	72.12	79.03	78.09	79.88	86.11
Surprise	81.23	81.23	88.12	96.12	93.12
Disgust	85.66	79.03	89.04	87.23	94.55
Fear	77.99	76.9	60.88	92.34	96.99
Anger	76.88	81.34	85.34	90.12	91.11
neutral	82.33	80.09	75.66	97.81	88.09

The specificity of emotional understanding of the Deep Learning assisted Hybridized Face Expression Recognition system is measured by the cluster center and is used to produce the function of the embedded building, and its performance accuracy for clusters in the image is illustrated in Figure 9. Then after patches from a differentiated facial are picked, tests are carried out with different clusters to produce embedding language. The detection output has been shown to focus on the amount of embedding creation clusters, the number of characteristics identified, features extraction rates, and the scale of the training sample.

Doi: https://doi.org/10.54216/FPA.100210
Received: November 23, 2022 Accepted: March 17, 2023

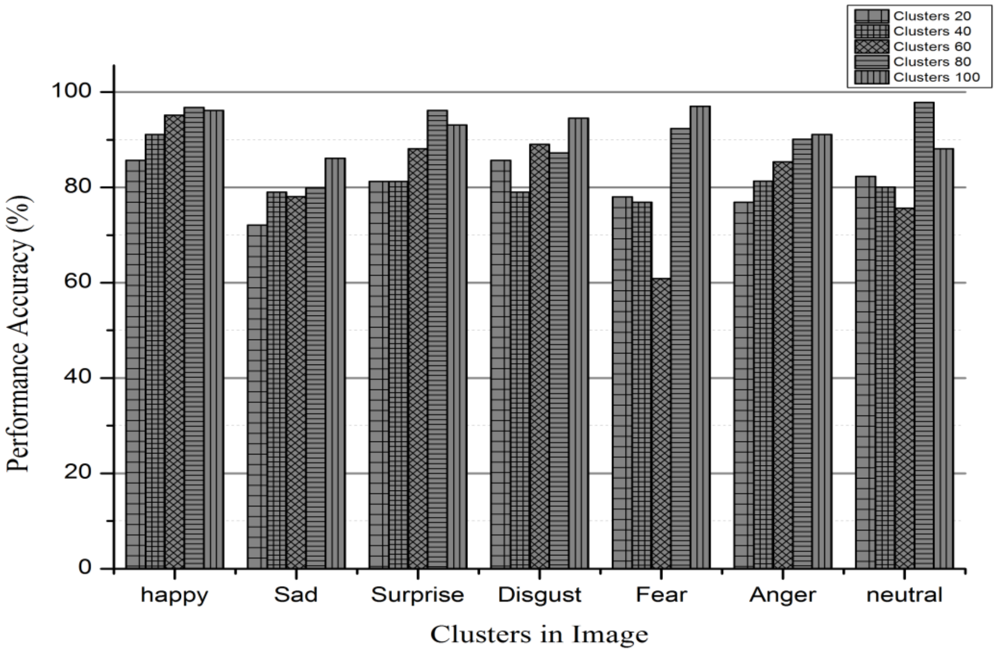


Figure 10: Performance Accuracy for various clusters in face image

The overall recognition accuracy of all the existing classification algorithms is compared and illustrated in figure 10. The comparative analysis illustrated that the suggested approach contributed to increased accuracy with respect to emotions. DLFERS is used to recognize facial emotion, which extracts TISF for each second pixel.

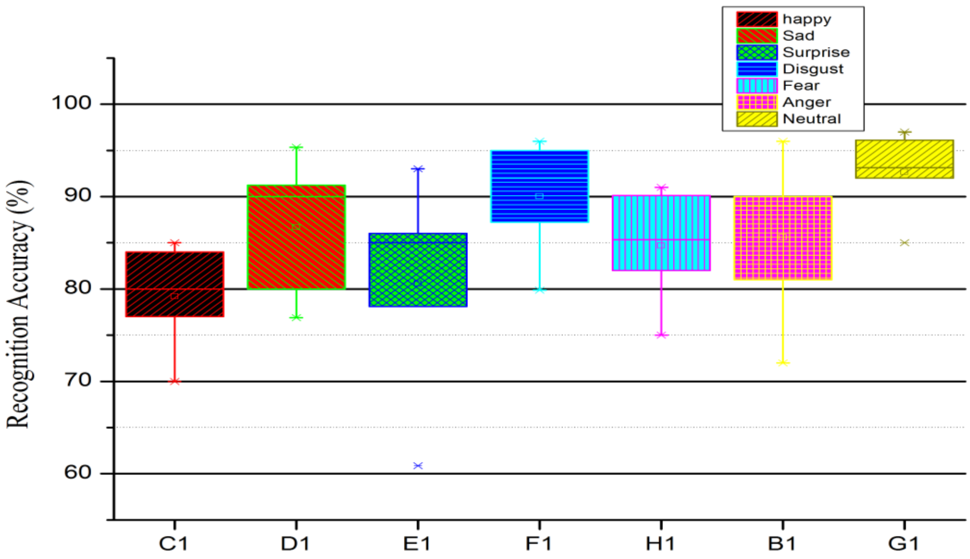


Figure 11: Recognition Accuracy of DLFERS

The overall performance accuracy for DLFERS is illustrated in figure11. It is indicated that with the absence of the aspect of emotions that remain valid during facial expression and the efficiency improves. Unless unwanted areas are removed, feature overlaps lead to the deterioration of efficiency.

Doi: https://doi.org/10.54216/FPA.100210
Received: November 23, 2022 Accepted: March 17, 2023

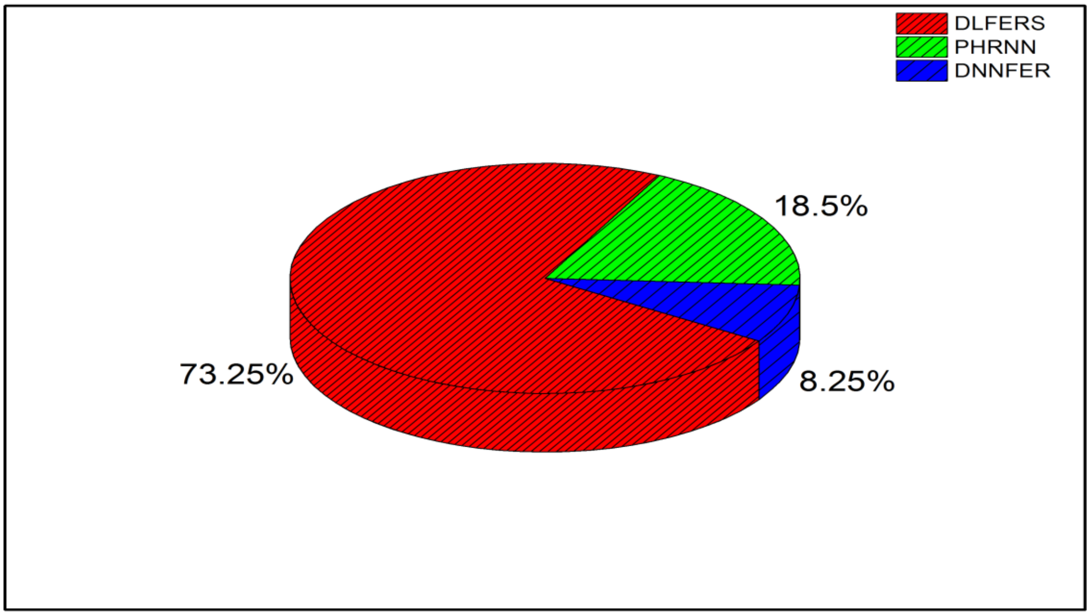


Figure 12: The overall performance accuracy for DLFERS

The detection output has been shown to rely on the amount of embedding creation groups, the range of factors detected, image segmentation rates, and the width of the training sample. The DLFERS can be prohibitively costly and does not consider image transformations in many other practical applications. The severity value restrictions in the proposed DLFERS are not difficult coded; therefore, they can be adjusted with training samples to different situations.

The Deep Learning assisted Hybridized Face Expression Recognition system (DLFERS) has been proposed from a single image frame that utilizes feature extraction and classification with a support vector machine to identify human behavior. The proposed method also involves the Storage Pack of Features (SPOF) that uses Transformation of Invariant Structural Features (TISF) and Transformation of Invariant Powerful Movement (TIPM) for feature extraction of facial expression recognition. The overall classification accuracy obtained for the proposed DLFERS method is 93.96%.

5. **Conclusion**

This research provides information regarding Deep Learning assisted Hybridized Face Expression Recognition system (DLFERS) that has been proposed from a single image frame that utilizes feature extraction and classification with a support vector machine for identifying human behavior. In addition, DLFERS uses an information classification technique into a new fused image composed of two integrated blocks of eyes and mouth. These areas are very sensitive to human expression changes and particularly relevant for decoding emotional expressions. The feature extraction is achieved by a Storage Pack of Features (SPOF) that uses Transformation of Invariant Structural Features (TISF) and Transformation of Invariant Powerful Movement (TIPM) for the feature extraction process. The overall classification accuracy obtained for the proposed DLFERS method is 93.96%.

Funding: “This research received no external funding”
Conflicts of Interest: “The authors declare no conflict of interest.”

Reference

[1] Elfenbein HA. Nonverbal dialects and accents in facial expressions of emotion. Emotion Review. 2013 Jan;5(1):90-6.
[2] Weise T, Bouaziz S, Li H, Pauly M. Realtime performance-based facial animation. ACM transactions on graphics (TOG). 2011 Jul 25;30(4):1-0.
[3] Bourke C, Douglas K, Porter R. Processing of facial emotion expression in major depression: a review. Australian and New Zealand Journal of Psychiatry. 2010 Aug 1;44(8):681-96.

Doi: https://doi.org/10.54216/FPA.100210
Received: November 23, 2022 Accepted: March 17, 2023

[4] Garg A, Bajaj R. Facial expression recognition & classification using hybridization of ica, ga, and neural network for human-computer interaction. Journal of Network Communications and Emerging Technologies (JNCET) www. jncet. org. 2015 May;2(1).

[5] Olney NT, Goodkind MS, Lomen-Hoerth C, Whalen PK, Williamson CA, Holley DE, Verstaen A, Brown LM, Miller BL, Kornak J, Levenson RW. Behaviour, physiology and experience of pathological laughing and crying in amyotrophic lateral sclerosis. Brain. 2011 Dec 1;134(12):3458-69.

[6] Wood A, Rychlowska M, Korb S, Niedenthal P. Fashioning the face: sensorimotor simulation contributes to facial expression recognition. Trends in cognitive sciences. 2016 Mar 1;20(3):227-40.

[7] Shin E. A Neural Network approach to Automated Essay Scoring: A Comparison with the Method of Integrating Deep Language Features using Coh-Metrix.

[8] O'Connor RC, Nock MK. The psychology of suicidal behaviour. The Lancet Psychiatry. 2014 Jun 1;1(1):73-85.

[9] Wong N, Beidel DC, Sarver DE, Sims V. Facial emotion recognition in children with high functioning autism and children with social phobia. Child Psychiatry & Human Development. 2012 Oct 1;43(5):775-94.

[10] Kaplan A, Haenlein M. Siri, Siri, in my hand: Who's the fairest in the land? On the interpretations, illustrations, and implications of artificial intelligence. Business Horizons. 2019 Jan 1;62(1):15-25.

[11] Belhumeur PN, Jacobs DW, Kriegman DJ, Kumar N. Localizing parts of faces using a consensus of exemplars. IEEE transactions on pattern analysis and machine intelligence. 2013 Jan 15;35(12):2930-40.

[12] Galbally J, Marcel S, Fierrez J. Biometric antispoofing methods: A survey in face recognition. IEEE Access. 2014 Dec 18;2:1530-52.

[13] Ranjan R, Castillo CD, Chellappa R. L2-constrained softmax loss for discriminative face verification. arXiv preprint arXiv:1703.09507. 2017 Mar 28.

[14] Stins JF, Roelofs K, Villan J, Kooijman K, Hagenaars MA, Beek PJ. Walk to me when I smile, step back when I'm angry: emotional faces modulate whole-body approach–avoidance behaviors. Experimental brain research. 2011 Aug 1;212(4):603-11.

[15] Giannopoulos P, Perikos I, Hatzilygeroudis I. Deep learning approaches for facial emotion recognition: A case study on fer-2013. InAdvances in Hybridization of Intelligent Methods 2018 (pp. 1-16). Springer, Cham.

[16] Qi C, Li M, Wang Q, Zhang H, Xing J, Gao Z, Zhang H. Facial expressions recognition based on cognition and mapped binary patterns. IEEE Access. 2018 Mar 22;6:18795-803.

[17] Liu P, Han S, Meng Z, Tong Y. Facial expression recognition via a boosted deep belief network. InProceedings of the IEEE conference on computer vision and pattern recognition 2014 (pp. 1805-1812).

[18] Dornaika F, Moujahid A, Raducanu B. Facial expression recognition using tracked facial actions: Classifier performance analysis. Engineering Applications of Artificial Intelligence. 2013 Jan 1;26(1):467-77.

[19] Boughrara H, Chtourou M, Amar CB, Chen L. Facial expression recognition based on a mlp neural network using constructive training algorithm. Multimedia Tools and Applications. 2016 Jan 1;75(2):709-31.

[20] Khandait SP, Thool RC, Khandait PD. Automatic facial feature extraction and expression recognition based on neural network. arXiv preprint arXiv:1204.2073. 2012 Apr 10.

[21] Boz H, Kose U. Emotion extraction from facial expressions by using artificial intelligence techniques. BRAIN. Broad Research in Artificial Intelligence and Neuroscience. 2018 Feb 28;9(1):5-16.

[22] Zhang K, Huang Y, Du Y, Wang L. Facial expression recognition based on deep evolutional spatial-temporal networks. IEEE Transactions on Image Processing. 2017 Mar 30;26(9):4193-203.

[23] Perikos I, Ziakopoulos E, Hatzilygeroudis I. Recognizing emotions from facial expressions using neural network. on Artificial Intelligence Applications and Innovations 2014 Sep 19 (pp. 236-245). Springer, Berlin, Heidelberg.

[24] Zhang T, Zheng W, Cui Z, Zong Y, Yan J, Yan K. A deep neural network-driven feature learning method for multi-view facial expression recognition. IEEE Transactions on Multimedia. 2016 Aug 3;18(12):2528-36.

[25] Wen G, Hou Z, Li H, Li D, Jiang L, Xun E. Ensemble of deep neural networks with probability-based fusion for facial expression recognition. Cognitive Computation. 2017 Oct 1;9(5):597-610.

[26] Ali, M.H., Jaber, M.M., Abd, S.K., Alkhayyat, A. and Jasim, A.D., 2022. Artificial Neural Network-Based Medical Diagnostics and Therapeutics. International Journal of Pattern Recognition and Artificial Intelligence.

Doi: https://doi.org/10.54216/FPA.100210

Received: November 23, 2022 Accepted: March 17, 2023

[27] Alajanbi, M., Mohd Arfian Ismail, Raed Abdulkareem Hasan, & Junaida Sulaiman. (2021). Intrusion Detection: A Review . Mesopotamian Journal of CyberSecurity, 2021, 1–4. https://doi.org/10.58496/MJCS/2021/001

[28] Ahmed Hussein Ali, Mostafa Abdulghfoor Mohammed, & Munef Abdullah Ahmed. (2021). Character Recognition By Implementing FPGA-Based Artificial Neural Network. Mesopotamian Journal of Computer Science, 2021, 14–19. https://doi.org/10.58496/MJCSC/2021/003

Doi: https://doi.org/10.54216/FPA.100210
Received: November 23, 2022 Accepted: March 17, 2023