RESEARCH ARTICLE

Role of Zoning in Facial Expression Using Deep Learning

TAIMUR SHAHZAD [ID] **[1], KHALID IQBAL** [ID] **[1], MURAD ALI KHAN[2], IMRAN** [ID] **[3], AND NAEEM IQBAL** [ID] **[2,4], (Member, IEEE)**

[1]Department of Computer Science, COMSATS University Islamabad, Attock Campus, Attock 43600, Pakistan
[2]Department of Computer Engineering, Jeju National University, Jeju 63243, Republic of Korea
[3]Department of Biomedical Engineering, Gachon University, Incheon 21936, Republic of Korea
[4]Department of Research and Innovation, SCIEKORE Institute of Scientific Entrepreneurship and Technology, Batkhela 23020, Pakistan

Corresponding author: Imran (imranjofficial@jejunu.ac.kr)

ABSTRACT Facial expression is an unspoken message essential to collaboration and effective discourse. An inner emotional state of a human is expressed using facial expressions and is very effective for communication with actual emotions. Anger, happiness, sadness, contempt, surprise, fear, disgust, and neutral are eight common expressions of humans. Scientific community proposed several face emotion recognition techniques. However, due to fewer face landmarks and their intensity for deep learning models, performance improvement for facial expression recognition still needs to be improved for accurately predicting facial emotion recognition. This study proposes a zoning-based face expression recognition (ZFER) to locate more face landmarks to perceive deep face emotions indemnity through zoning. After face extraction, landmarks from the face, such as the eyes, eyebrows, nose, forehead, and mouth, are extracted. The second step is zoning each landmark into four regions and zone-based face landmarks are passed to the VGG-16 model to generate a feature map. Finally, the feature map is given as input to fully connected neural network (FCNN) to classify facial emotions into multiple classes. Various experiments are performed on facial expression recognition (FER) 2013 and CK+ datasets to evaluate our proposed model with state-of-the-art facial expression recognition approaches using performance assessment metrics like accuracy. The accuracy of the proposed method with face features on CK+ and FER2013 are 98.4% and 65%, respectively. The experimental zoning results improve from 98.47% to 98.74% on the CK+ dataset.

INDEX TERMS Computer vision, deep learning, zoning, FER2013, CK+.

I. INTRODUCTION

Many advances have been made in the field of human-computer interaction, such as vision-based security systems, self-driving cars, and smart applications based on artificial intelligence and predictive analysis [1], [2]. For communication, human beings usually use gesture, speech, and facial expression mediums [3]. Language is a verbal communication medium, and the quality of voice plays an essential role in this medium to sense emotion [4]. However, non-verbal facial expressions and gestures are the medium of communication. According to a study, 7% of communication between humans is communicated through Linguistic language (verbal parts),

The associate editor coordinating the review of this manuscript and approving it for publication was Hossein Rahmani [ID].

38% by paralanguage (vocal parts) and 55% by facial expressions [5]. Therefore, facial expressions are vital in perceiving information in face-to-face communication. The facial expressions define the state of emotion, which is very effective in interpreting human reactions and feelings [6]. Facial expression is an unspoken message. It is an essential part of the collaboration and appropriate information. An inner emotional state of a human is expressed using facial expressions and is very effective for communication with actual emotions. Charles Darwin developed the first conceptual framework for emotional expression. Similarly, a study by Ekman and Friesen enumerated eight basic expressions such as sadness, fear, happy, disgust, anger, and surprise [7]. Fig. 1 illustrates these eight basic facial expressions.
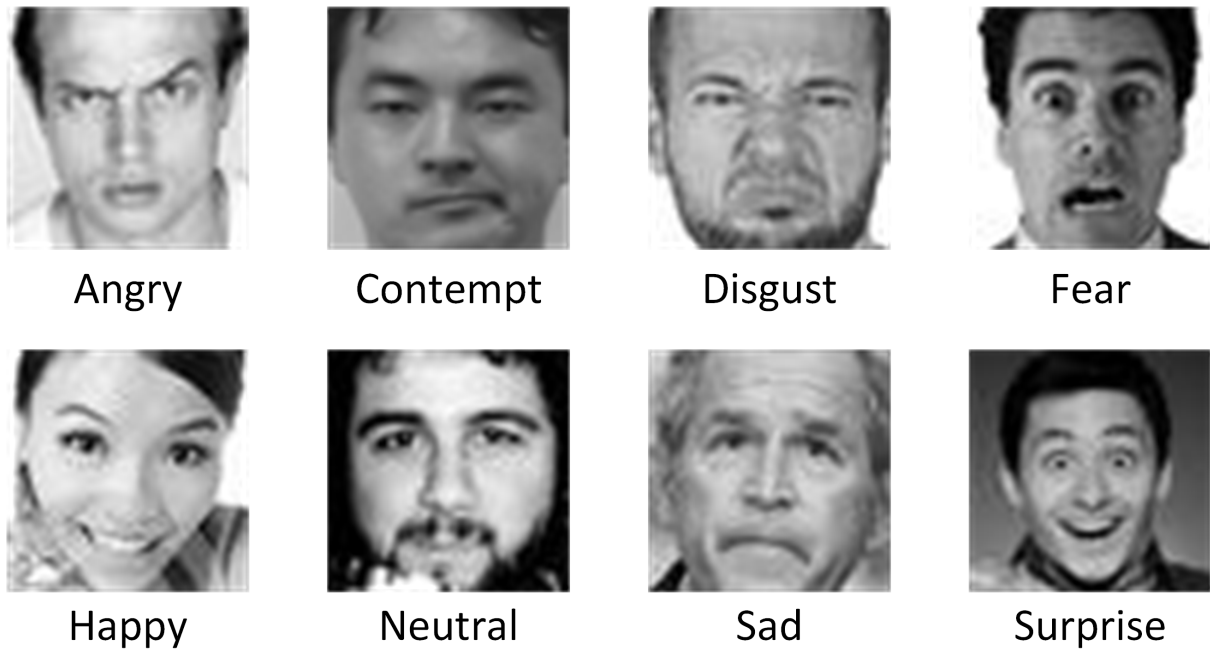
FIGURE 1. Basic facial expressions.

Human perform faces recognition task in their daily lives. People can recognize those faces they have met before and easily differentiate them from strangers. This task is the basis for human cooperation, relationship and collaboration. In the last few decades, due to the advancement in computing power, more powerful computing systems have been available that can have the ability to process a massive amount of data. Image Processing and Computer Vision [8] is the field of Artificial Intelligence, where researchers analyze images and video data to extract patterns and invaluable information. Computer Vision has many application areas, such as human-computer interaction, biometric authentication, Medical Images, multimedia management, and surveillance [9]. Feedback related to user emotions provides valuable information that could positively impact various areas such as e-marketing, robotics, smart products, etc. Facial expression recognition is essential for any image processing application. The accuracy of expression recognition depends upon facial features. Therefore, facial features play a crucial role in accurately predicting the emotion recognition system. A sophisticated technique is required that extract the more relevant facial expression for accurate prediction of emotion recognition.

Face recognition techniques suffers from visual pattern recognition problem when a three-dimensional object is to be identified in the perspective of its two-dimensional image [10]. However, face recognition and facial emotion recognition tasks can be automatically performed using Image Processing and Computer Vision techniques. This area has made significant progress in achieving better face recognition models. The existing face recognition systems can achieve satisfactory results under certain constraints. However, face images suffer from several problems, such as head pose, blur, and illumination [11], that can affect the performance of the facial recognition system. Furthermore, recognizing facial expressions automatically with accuracy is a difficult job. There are two main difficulties with the emotion recognition system. First, the distinct expressions make it difficult to compare different persons. Second, some expressions are difficult to distinguish, which makes it more challenging to differentiate between the emotions of the same person.

Due to fewer facial features and their intensity for deep learning models, improving performance for facial expression detection is still challenging. Therefore, for an accurate prediction of facial emotions prediction. A model called ZFER is proposed to locate more face landmarks to perceive deep face emotions indemnity through zoning. First, face landmarks are extracted after face extraction, such as eyes, eyebrows, Nose, Forehead, and Mouth. In the second step, zoning is applied to get four regions of each landmark [12]. Lastly, zone-based face landmarks are provided to the VG16-16 model as the inputs to generate a features map. Finally, the FCNN is employed to classify facial emotions based on the general features map. To highlight the significance of the proposed research study, two benchmark datasets: FER 2013 and CK+, are used to evaluate our proposed model with state-of-the-art facial expression recognition approaches.

The main contributions of this study are as follows:

- Face landmarks localization and feature extractions based on zoning.
- Extract the forehead as an extra feature for emotion recognition.
- Analyze the impact of zoning on facial features for the classification of emotions.
- Develop a robust hybrid VGG-16-DNN assisted facial emotions classification model to perform extensive

experiments using benchmark FER 2013 and CK+ datasets.

- Finally, the impact of zoning is analyzed to highlight the effectiveness of the proposed VGG-16-DNN model.

The rest of the paper has been organized as follows. Section II presents a comprehensive literature overview of the current research issues and further highlights the motivations for our contributions. Section III presents the proposed methodology. The results and discussion is presented in Section IV. The conclusion is presented in Section V.

II. RELATED WORK

Scientific community proposed various research methods for emotion recognition. This section presents overviews of the existing approaches proposed for the classification of emotions. First, we present facial expressions recognition based on the well-known FER datasets. The later part of related work presents facial expression recognition mechanisms based on deep learning paradigms, ensemble learning models and transfer learning. Lastly, we present an overview of the issues, challenges and research gap to show the significance of this study. Identification and recognition systems based on motion sensing data collected through accelerometer and gyroscope have been proposed using deep learning such as Deep CNN, SVM, and few-shot settings. Rich face representation learning using deep learning helps predict social relation traits using facial images. Introduction to this novel network architecture includes a bridging layer that controls the existing relationship between datasets. Moreover, it enables learning from different attribute sources and coping with missing target attribute labels. One of the extensions to this study includes exploring modelling relationships among more than two people, implemented via a graphical or voting model, where the face is denoted by a node and the relationship between nodes (faces) is an edge [13]. Deep Convolutional Network & Alignment Mapping Network overcome the issue of performance decline for the non-aligned facial features. The method utilizes both aligned and non-aligned features to perform FER on FER 2013 dataset resulting in improved performance [14]. VGG-like standard deep architecture is presented for a more precise definition of categorizing children's emotions related to learning using theoretical and psychological frameworks. It produces effective state prediction accuracy, including neutral and positive states for FER+ Dataset. For meaningful interpretation of information, the work is extended from multiple sources & working with multi-modal datasets [15]. Alenazy and Alqahtani proposed a semi-supervised deep belief network (DBN) and gravitational search algorithm (GSA) based approach [16]. This approach is used to predict seven classes of facial expressions, and each class has its associated 4-bit code for describing it. Training and evaluation of the DBN are performed on JAFFE, Oulu CASIA, CK+, and MMI datasets.

Deep Sparse AutoEncoders (DSAE) are used for the feasible and effective FER using eight facial expressions, including HD features, both geometric and appearance. The research not only provides the learning of robust and discriminative features but also compares the proposed method with three states of the art methods. Furthermore, it can be used for more complex systems catering to various performance requirements [17]. Shallow CNN (SHCNN) with three layers to classify static and micro-expression data is introduced to overcome overfitting on small datasets (FER 2013, FER+, CASME, CASME II, SAMM) and ignore redundancy. This FER is for static and micro-expression image data and provides the best results for FER 2013 and SAMM datasets while comparable results for the remaining datasets [18]. To cope with the challenge of intra-class variation, mini-Xception based on Xception & CNN (mini-Xception Architecture) to FER 2013 dataset is applied, providing a real-time visioning system including face detection & emotion classification simultaneously. The method performs well even with small data with limited features. Adjusting the count of convolutional layers & the filter size can further improve performance. Mini-Xception algorithm, for the original image dimensions, is used [19].

In [20], the author proposed a technique for recognizing facial expressions that predict the expression types on the input image's face region. The photographs acquired in an unrestricted setting were taken into consideration. A facial region is recognized in each image taken during implementation. The colour facial image is then converted into a grey-scaled image to speed up the operations. The goal of using these photographs is to distinguish the many types of expression in them. This few-shot learning mechanism to build a realistic framework improves performance compared to the RNN and handcrafted features. The novel Distribution Matching Approach (FaceBehaviorNet) is a practical framework for heterogeneous multi-task learning. This framework is not limited to large-scale face analysis, but it can also leverage the concept of zero and few-shot in its two case studies. The framework is evaluated using the FER Datasets. The analysis results show that the framework is suitable for continuous effect estimation, facial action unit detection, basic emotion recognition, facial attribute detection, and face identification. Nevertheless, the framework can be extended to improve its generalization abilities [21]. Another IFER-DTFL model is presented by [22], where the author classifies facial expressions using three different stages. For the detection of face ResNet50 and Mask RCNN is used, for the optimization of hyperparameters adam optimizer is implemented, for the features extraction DenseNet121 is used, and for classification WKELM approach is used. Furthermore, [23] proposed densely CNN with the combination of hierarchical spatial attention to classify the facial expressions. By the combination of these two modules, the relevant emotional features are seized to get better classification results. The experiment results show the supremacy of the proposed model.

An Inconsistent Pseudo Annotation to Latent Truth (IPA2LT) with end-end LTNet is proposed to overcome the errors and biases due to inconsistent facial expressions

annotation. Performance is improved for 7 FER Datasets (CK+, MMI, Oulu-CASIA, SFEW, CFEE, RAF, AffectNet), and results are enhanced for noisy, inconsistent, synthetic, and actual data [24]. The basic aim is to cut the framework development cost and create a framework for general affect recognition. One such framework is called: FaceChannel: A fast & furious Deep NN for FER using lightweight NN with few parameters to learn and adapt the learned facial features towards different datasets (AffectNet, FER+, FABD, OMG-E). The system's strengths are improved performance with fewer trainable parameters and cross-data analysis with other affective recognition conditions. Studying and deploying this model on platforms with reduced data processing capabilities, such as social robots, can be a future contribution [17]. Similarly, Deep CNN for noisy labels contributes to developing an enhanced FER+ dataset with ten taggers to label each input and can be implemented via four different approaches: majority voting, multi-label, probabilistic, and cross-entropy loss [25]. Suppressing difficulty while annotating large-scale datasets can be dealt with Self Cure networks (self-attention mechanism and careful labelling). The synthetic and real-world uncertainties of the Synthetic FER, WebEmotion, RAF-DB, AffectNet and FERPlus datasets are suppressed using the proposed approach [26]. As these studies are evaluated using the FER dataset; thus, there is a need for FER system performance enhancement. Therefore, Facial Image Threshing (FIT) machine was developed by following these principles: removing irrelevant facial images, facial image collection, misplaced face data correction, original dataset merging, and data augmentation. The FIT machine recognizes emotions, enhances performance, and can be used by researchers and developers to create an independent FER Dataset [27]. The cognitive model for facial expression recognition is emotion classification using CNN on the FER+ dataset and tested in live validity mode. It can also be implemented in the health sector. It provides faster results with multiple cameras, and efficient results with clear visuals [28]. Making a FER system in real-time video streams using Cycle-GAN's four steps; preprocessing, CycleGAN-based design, generating new synthetic image data, and DNN-based inference system. The transfer learning is modified with a balanced version, including face detection and FER. The method is simple, providing a high recognition rate and speed of facial expressions on the Original FER 2013 dataset. The same can be used in video sequences, and online learning platforms [29]. A hybrid deep learning model to classify human emotions using DNN and transfer learning is proposed by [30]. This study combines multiple deep learning models to make a hybrid classification approach for better prediction results. According to the author proposed hybrid approach achieved 81.42% accuracy on testing data and 95.93% accuracy on training data.

A facial recognition system using a deep learning-based ensemble model can be used as a passive factor in two-factor authentication [31]. Similarly, the ensemble method is presented for real-time facial expression recognition that extracts the facial regions in real time and compresses data using multiple features. It solves issues of insufficient data and expression unrelated to intra-class differences using the Local Pathway network and Attention Module. The datasets used for analysis are FER+, and CK+ [32]. Branching is used instead of a single deep convolution network to reduce the redundancy and cost of the ensemble model training process. Redundancy is thus reduced by varying the branching levels, maintaining diversity and generalization power. Moreover, the generalization error is reduced too. All these experiments are being done on AffectNet and FER+ datasets using an ensemble with shared resources. The study can be further extended for approaches to overcome catastrophic forgetting in ensembles using shared resources [33]. Using datasets such as JAFFE, CK+, PIE, and real-world photographs, this study provides an effective facial expression recognition strategy to discriminate diverse human expressions for the retrieved images. Moreover, this method incorporates traditional preprocessing and deep convolution network-based feature extraction to improve the proposed system's overall accuracy and retrieval time. Improved cat swarm optimization (ICSO) algorithm is used for feature selection and image retrieval using ensemble classifiers such as support vector machine (SVM) and neural network (NN). The recognition rate, precision, recall f-measure, and sensitivity matrices are used to evaluate the performance results of their proposed model [34].

MFP aggregation network based on convolutional neural sub-networks learns deep characteristics from facial patches, which are then gathered into a single model architecture for classifying emotion [35]. The framework used Conditional Generative Adversarial Networks and a set of transformation functions for data augmentation. When evaluated on small datasets, the proposed MFP-CNN performs well, but with just three convolutional layers per sub-network, it cannot outperform contemporary DNN trained on large datasets. On the other hand, the MFP-CNN does not require a lot of processing power or large training datasets, and it trains quickly. Therefore, the proposed design could aid in deciphering the traits of face areas in soft biometric features, emotion recognition, and other attributes such as age, gender, ethnicity, and identification. The Deep Attentive Centre Loss (DACL) method adaptively selects a subset of significant feature elements for enhanced classification [35]. Using the intermediate spatial feature maps in CNN as context, the DACL incorporates an attention method to estimate attention weights connected with feature importance. The estimated weights accommodate the sparse formulation of centre loss to achieve intra-class compactness and inter-class separation for the relevant information in the embedding space. RAF-DB and AffectNet are used wild FER datasets in research. The approach uses sparse re-formulation of centre loss to manage the contribution of the deep feature representations in the Deep Metric Learning's objective function. By provid-

ing attention weights to the sparse centre loss, an attention mechanism fully parameterized by a customizable neural network predicts the probability of contribution along all dimensions.

In summary, FER is one of the most successful methods for providing emotional expressions knowledge, and it is often limited to six fundamental emotional learning plus neutral learning [36]. Several deep learning architectures have recently been proposed and developed. They've offered various services to detect human emotions using databases of random pictures collected from the real world and other laboratories. To verify the neural network, researchers learned and evaluated patterns in various datasets—the viability of architecture. The recognition rates differ from one database to the next while adhering to the same basic principles. The techniques employed in classification, sufficient memory, and model performance will be influenced by the expanding volume of data, processing time, and computer resources required for research activities. Some data augmentation techniques were applied to the training images to improve the learning parameters of the proposed deep-learning models. By finetuning the hyper-parameters, the proposed data augmentation strategies (bilateral faltering, unsharp filter, sharpening filter, image rotation, image scaling, shear mapping, image zooming, image filling, horizontal image flip) not only tackle the problem of overfitting but also improve the performance of the proposed system. The trade-off between used data augmentation and deep learning-based features causes the FER system to accept more difficult test samples to recognize expressions. Moreover, the prevention of vanishing gradient problems using landmarks is done by fusing landmarks (geometric and temporal information) with conventional video-based methods. The methodology is effective for given datasets (CK+, MMI, AFEW), resulting in the fusion of image information and landmarks. The research can be extended to more landmarks and various modalities utilization for multi-modal FER systems [37].

III. PROPOSED METHODOLOGY

This section discussed the steps of the proposed methodology for emotion classification. In the first step, FER2013 and CK+ image is input dataset using the CV2 Library. After pre-preprocessing of datasets, FER2013 and CK+ different experiments are performed. Then, the classification of emotion is performed using a convolutional neural network. Finally, the evaluation of the proposed method is performed using accuracy metrics. Fig.2 shows the proposed methodology diagram for classifying emotions.

First, images of different persons with multiple emotions as input to the model. After this extract, the face features from face images such as mouth, nose, and left and right eyes. Next, zoning is applied to extracted features to extract the more relevant face features. Then, on extracted face features, the classification of emotions is performed. Finally, the evaluation metrics are used to evaluate the performance of the proposed approach.

Furthermore, Fig. 3 shows a step-by-step flow of the proposed facial expression classification model. The proposed emotion classification model aims to classify facial expressions into eight desired classes. Thus, the step-by-step flow of the proposed approach consists of several steps. In the first step, input images are loaded by calling the open CV2 function with the path parameter of the given facial expression dataset. In the second step, loaded images of facial expressions are passed to the preprocessing module to resize all images into some fixed size, as the CNN model only processes images with uniform dimensions. Therefore, resizing is performed to resize all input images into 64 × 64 dimensions. Next, facial features are extracted from the input images, such as mouth, nose, and eyes. The zoning method is applied to the extracted facial features in the next step to extract more relevant and promising features. Furthermore, the prepared features set is divided into three sets: training, validation, and testing. In addition, a robust model is trained using training samples of images and validated using a validation sample set to evaluate the performance of the proposed emotion classification model. Moreover, testing samples are used as unseen samples to evaluate the effectiveness and generalization of the proposed model.

A. DATASETS

For the proposed model, two well-known publicly available datasets of images: FER2013 and CK+ are used. The first dataset, FER2013, is from a paper from Goodfellow et al. [38], written after a workshop/challenge in the Facial Expression Recognition domain. It consists of approximately 30,000 images categorized into seven classes. The seven classes from FER2013 are angry, disgust, fear, happy, sad, surprise, and neutral. However, FER2013 data contain images misplaced and grouped in the wrong class. It also contains irrelevant facial expression images, which cause computational overhead.

The second benchmark dataset, CK+, is acquired from Lucey et al. [39], which consists of 920 trainable facial expression images with eight classes. The CK+ dataset classes include anger, contempt, disgust, fear, happiness, neutral, sadness, and surprise. Each image size of the CK+ dataset is 640 × 490 pixels. It is an extension of an older dataset from the year 2000, namely the CK. Furthermore, the CK+ dataset contains only frontal face pose images and a small number of facial images compared to the FER2013 dataset. Moreover, the small number of training samples of the CK+ dataset produces bias and unreliable results. However, DL models require a large number of training images to build a reliable and generalized model. Therefore, in this research study, both datasets are substituted and merged to form a reliable dataset for classifying facial expressions into eight classes.

B. DATA PREPROCESSING

The final task for the proposed model is to correctly predict the emotion expressed by each face in the FER2013 and CK+
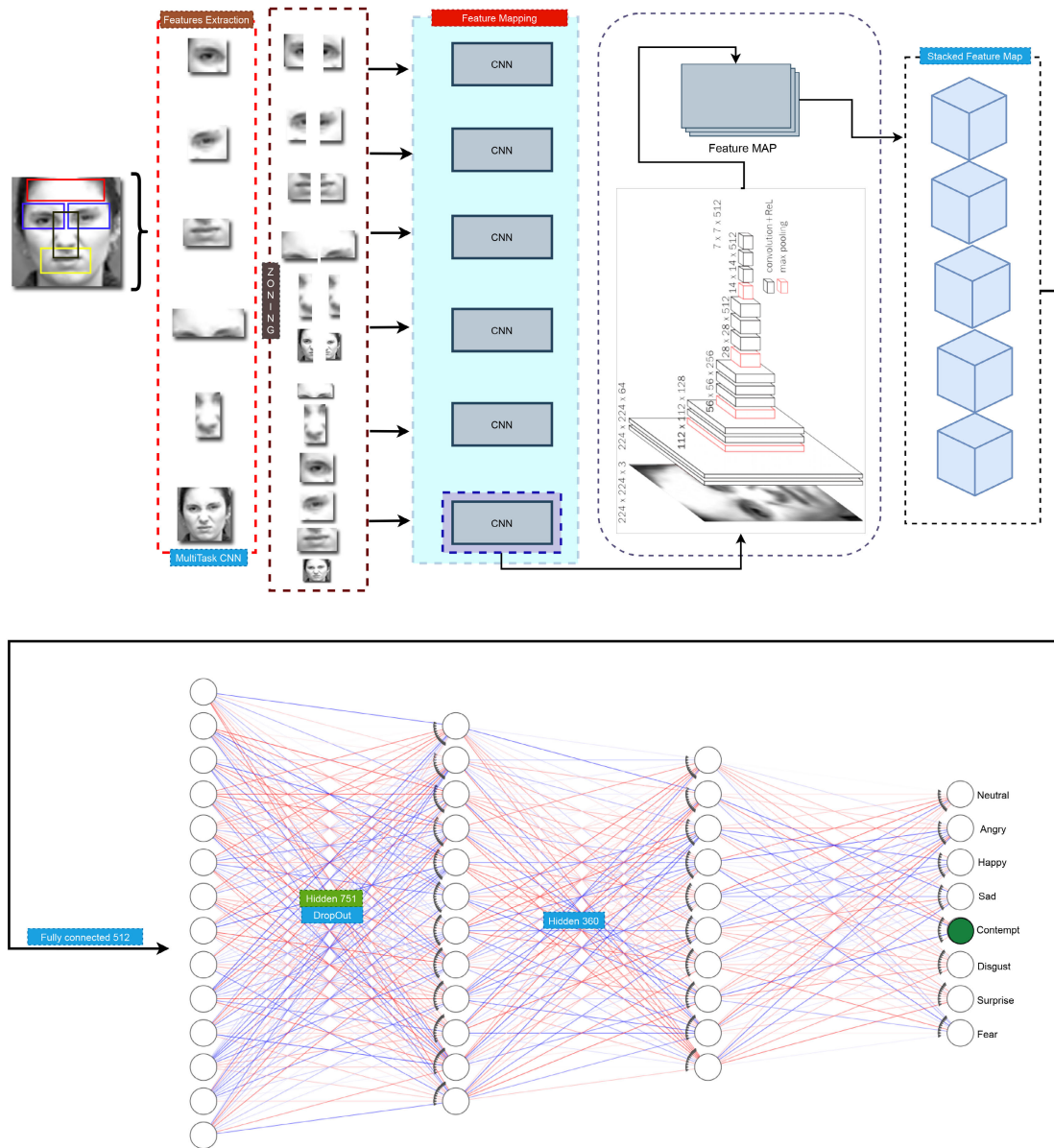
FIGURE 2. Proposed methodology architecture for recognition of emotions.

datasets with high accuracy. To achieve this goal, a technique called zoning is used [40]. The proposed technique utilizes the MTCNN library to calculate the key points needed to locate the four main facial zones in each image as shown in Fig. 4 and enumerated as follows:

1) Left Eye (1 key point)
2) Right Eye (1 key point)
3) Nose (1 key point)
4) Mouth (2 key points)

To make our work more significant, a position estimation algorithm is proposed to calculate the position of both the left and the right eye by calculating the position of an additional key point that is use to locate a new facial zone: the forehead. With this new key point in place, all the facial zones are identified that compose each face (left eye, right eye, nose, mouth, forehead) by calculating the specific areas around each point. Fig. 5 provides an example of localized bounding boxes and key points of features.

1) FEATURE EXTRACTION WITHOUT ZOONING
Selecting the areas belonging to each face feature will be crucial for us, as the face part will contribute to correctly predicting the emotion represented in each image. For this reason, each face part of the original image is cropped and the entire face that is kept in its entirety. Each face part is then considered an ''extracted feature'' for participation in the learning process: for this reason, each face feature is saved as a file into its dataset.
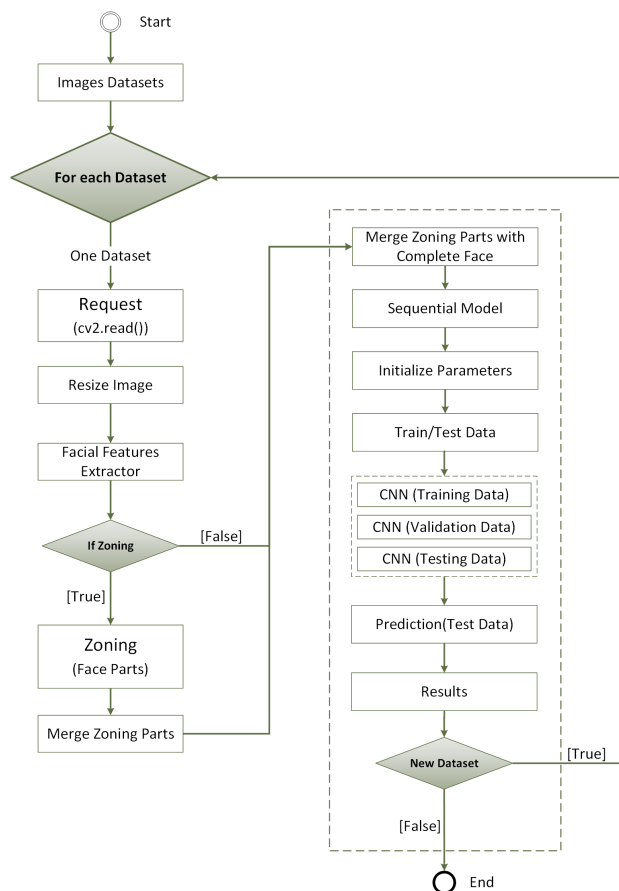
FIGURE 3. Flow model of the proposed facial expression classification approach.

For experimental results six separate datasets are prepared, one for each facial feature and another for the whole face, as shown in Fig. 6.

1) Left eye
2) Right eye
3) Nose
4) Mouth
5) Forehead
6) Cropped face

2) BOTTLENECK FEATURE STACKING

Our bottleneck feature stacking is based on the research study of Wang et al. [32], which improves the proposed model's overall performance. As explained earlier, the data is partitioned into six different datasets based on each facial feature. Then, a separate classifier is trained from each facial feature, eventually extracted from the same facial feature image. The classifier is VGG-16 [41], fine-tuned after hyperparameter optimization. Our needs were the extraction of the feature maps from the last pooling layer [42] - these are the bottleneck features. Eventually, six separate feature maps are obtained with the size of K $\times$ K $\times$ 512, that can be directly stacked to get a K $\times$ K $\times$ 3072 vector. Fig. 7 shows a visual process of the bottleneck feature stacking process to better visualize this process.

3) FEATURE EXTRACTION USING ZONING

Selecting the areas belonging to each facial zone is crucial for the proposed model, as each zone will contribute to correctly predicting the emotion represented in each image [43]. For this reason, we cropped each zone out of the original image and the entire face, which is kept in its entirety. Each zone is then considered an ''extracted feature'' and is used to participate in the learning process: for this reason, each zone (feature) is saved as a file into its dataset. Finally, zones are made for each image after extracting all features by tearing them into two identical Zones. So, now we have 18 features in our dataset. We now have 18 separate datasets, one for each facial feature and another for the whole face, which is shown in Fig. 8.

The extracted features are listed as follows:

1) Left eye
2) Left eye [Left Half Zone]
3) Left eye [Right Half Zone]
4) Right eye
5) Right eye [Left Half Zone]
6) Right eye [Right Half Zone]
7) Nose
8) Nose [Left Half Zone]
9) Nose [Right Half Zone]
10) Mouth
11) Mouth [Left Half Zone]
12) Mouth [Right Half Zone]
13) Forehead
14) Forehead [Left Half Zone]
15) Forehead [Right Half Zone]
16) Cropped face
17) Cropped face [Left Half Zone]
18) Cropped face [Right Half Zone]

Furthermore, the data split is done to report experiment results as follows: 70% for training, 15% for validation and 15% for testing.

C. CLASSIFICATION OF FACIAL EXPRESSION

For the second part of our model, a dataset made of stacked feature maps (bottleneck features) is used. After this process, our proposed classifier is employed to classify facial expressions using such bottleneck features. Fig. 9 shows the flow of the classification part. It consists of a series of fully connected layers [44] that end with a final dense layer of 8 classes designed to produce a vector containing the probabilities for the input to belong to a certain class. As a side note, in the hidden layers part of the network, different dropout layers are also added [45] and tried different activation functions, such as the Leaky-ReLU [46].

D. EVALUATION METRICS

The performance of our proposed model is evaluated using Accuracy, Precision, Recall, and F1 scores. Based on these metrics, our proposed model is evaluated and compared with the previously proposed models.
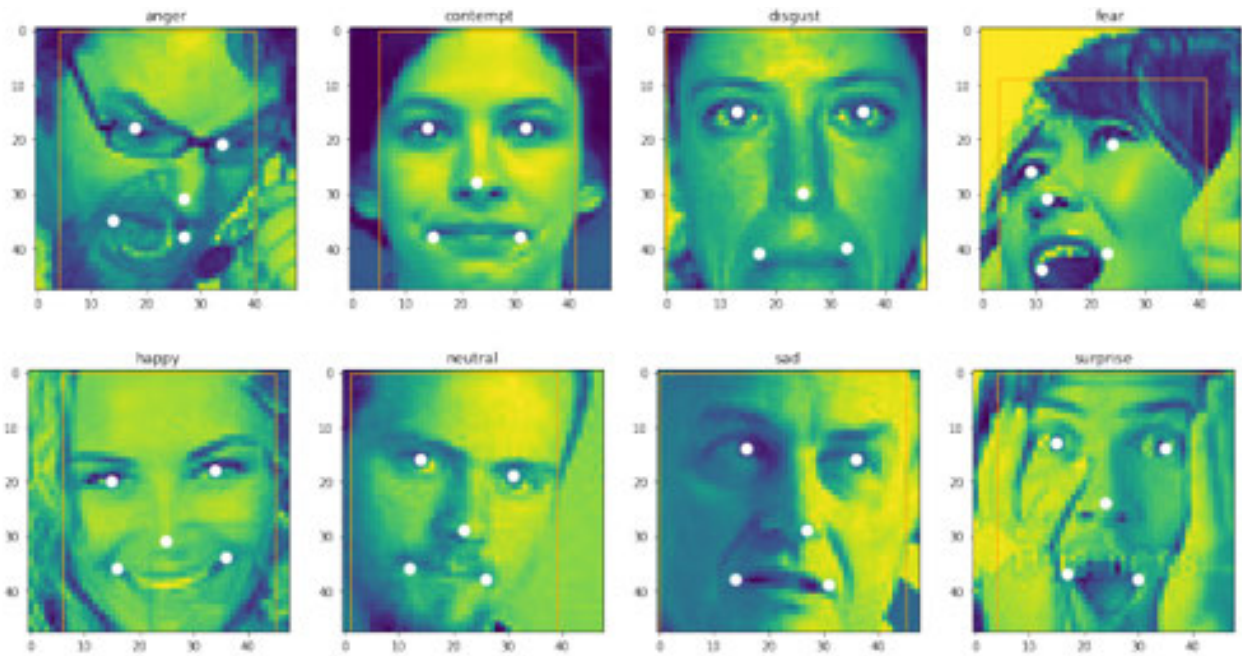
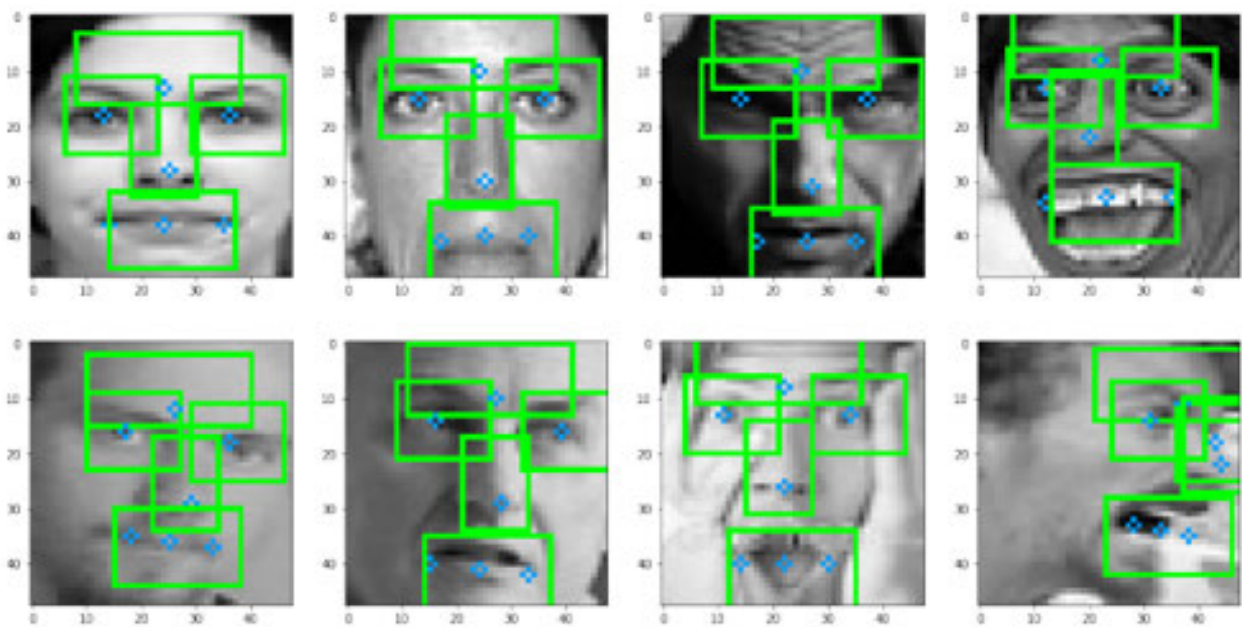**FIGURE 4.** Localized key points for various facial expressions.

FIGURE 5. Localized bounding boxes and key points of features.

1) ACCURACY

The accuracy measure describes how much our classifier correctly predicts facial expressions from all data points. The formula of accuracy is in Eq. 1:

$$Accuracy = (TP + TN)/(P + N) \qquad (1)$$

IV. RESULTS AND DISCUSSION

A. DATA AUGMENTATION

Since the CK+ dataset is quite small, data augmentation would improve the performance of the proposed model. Data augmentation was applied to the training set only, and the validation set was left intact. The augmentations so applied increased the number of samples in the training split based on the ImageDataGenerator augmentation library. More specifically, the detail of the data augmentation mechanisms applied are:

1) Image rotation of ±20 degrees.
2) Re-scaling (required for normalization), factor of 1/255.
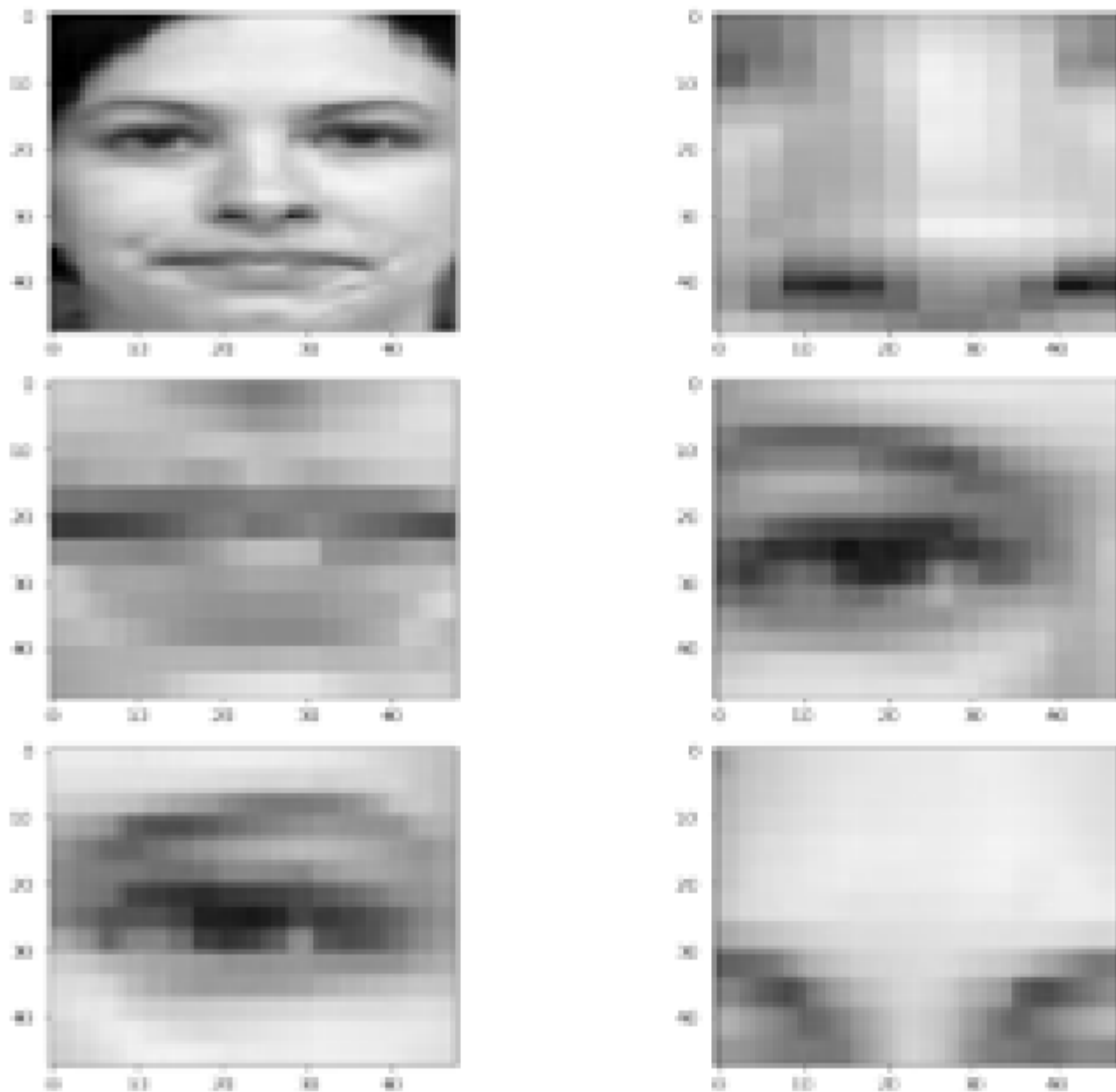3) Shear range (distorting the image along an axis)
4) Cropping

FIGURE 6. Individual features, extracted and separated.

5) Horizontal flipping

B. EXPERIMENTAL RESULTS WITHOUT USING THE ZOONING

1) PROPOSED MODEL RESULTS USING THE FER2013 DATASET

The initial results are a bit disorienting, thus the 6 VGG-16s for our feature maps and the other neural network made of Dense and Dropout layers are used. Fig. 10 shows a comparison of training and validation accuracy and loss analysis using the FER13 dataset. Though in a more comprehensive perspective, the results of FER2013 does not meet our

expectation. The results were not good because the data was highly random. It was because of the unclarity of most images, making it even harder for human experts to classify and tell the exact expression. So the proposed model performance results of FER2013 is reported for other scientific community for any future research insights.

2) PROPOSED MODEL RESULTS USING THE CK+ DATASET

In the given Fig.11, the results in terms of loss and accuracy for the dataset CK+ using all features bottleneck combined like eyes, nose, left eye, right eyes, mouth, forehead and full cropped face, using the above-mentioned model. So, to recap,
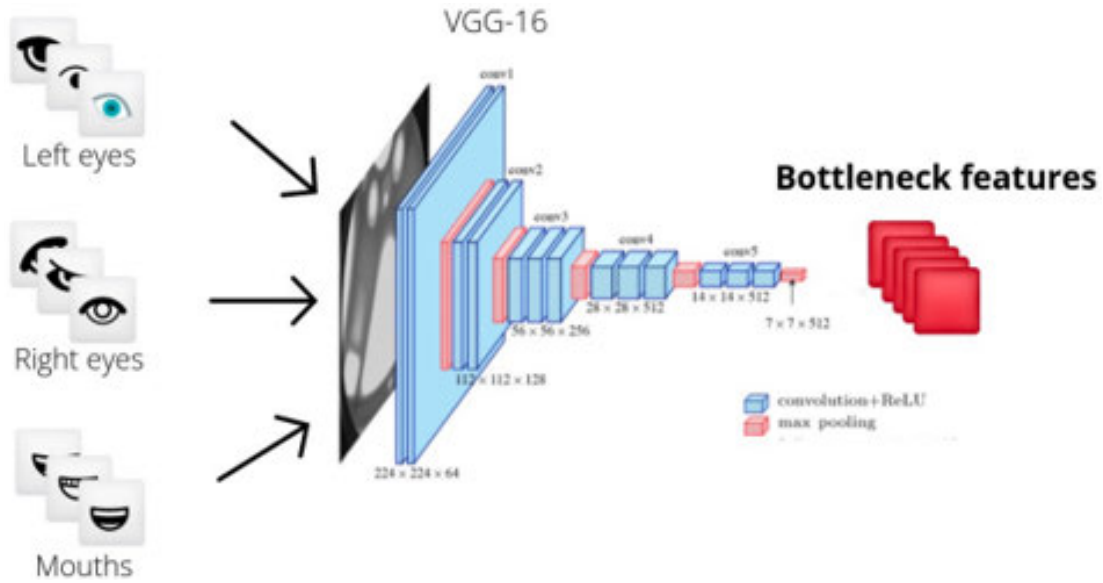
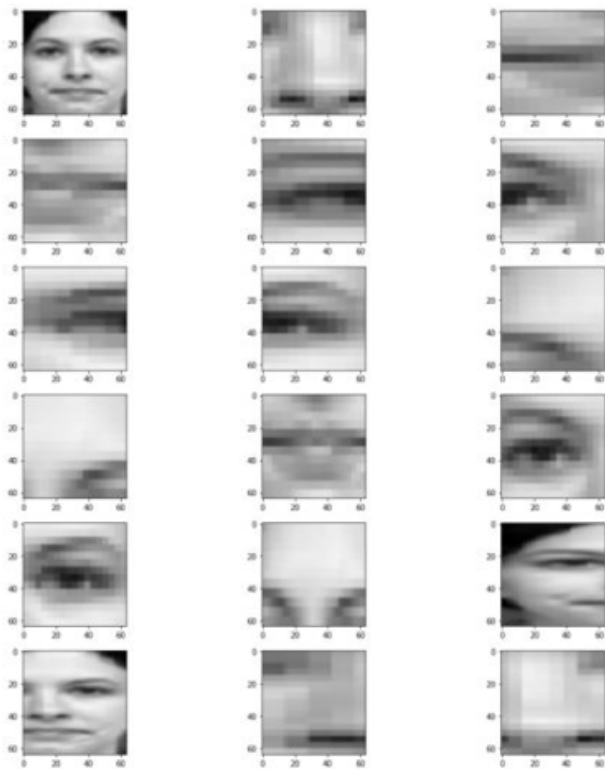**FIGURE 7.** Bottleneck feature stacking process.

FIGURE 8. Individual features, extracted and separated.

hyper-optimization techniques were applied to VGG-16 for fine-tuning the fully connected part from the initial training. After 40 epochs, 0.08 loss, and 98.47% validation accuracy is obtained, which is good, and also, the loss convergence seems very fast, while the validation accuracy is consistent with the training accuracy.

3) EXPERIMENTAL RESULTS OF CK+ WITH CROP FACE IMAGES

Another experiment performed is more comparative in terms of achieving something better than the baseline by using zoning instead of just throwing images of faces to a CNN. For a more in-depth understanding, let us consider a generic VGG-16 as our baseline for a chosen model. The model was realized to make this comparison; The plots for training/validation loss curves and the loss curves are shown without any preprocessing and just take images of faces as the input. As indicated by the caption of Fig. 12, VGG-16 (fine-tuned) is used as our starting point. The results seem quite acceptable, getting us a 91.8% on the validation set. For better performance, the model with zoning is considered, which achieved 98.4% accuracy on the validation set.

C. EXPERIMENTAL RESULTS USING THE ZONING

1) MODEL USING THE CK+ DATASET WITH ZONING

In the given Fig. 13, the results in terms of loss and accuracy for the dataset CK+ using all features bottleneck combined like eyes, nose, left eye, right eyes, mouth, forehead, full cropped face and mentioned zones, using the above-mentioned model. To recap, each VGG-16 has been fine-tuned, and the '' fully connected part'' is trained from scratch.

After 40 epochs, 0.025 loss and 98.74% validation accuracy is obtained, which is great, and a few more decimals than the previous results without zoning.

2) TEST SET ACCURACY AND CONFUSION MATRIX

Fig. 14 shows a confusion matrix to analyze true positive and false negative rates. In the given confusion matrix, the results of test samples are evaluated for each class label.
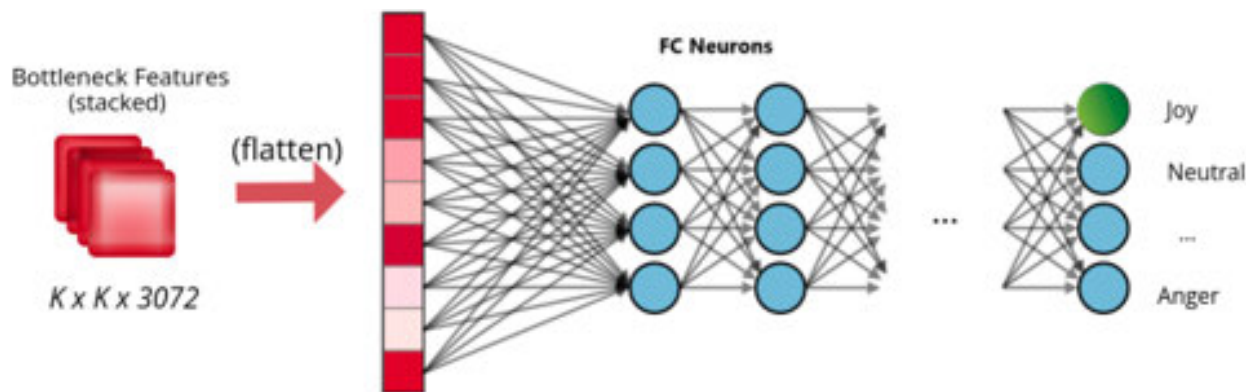
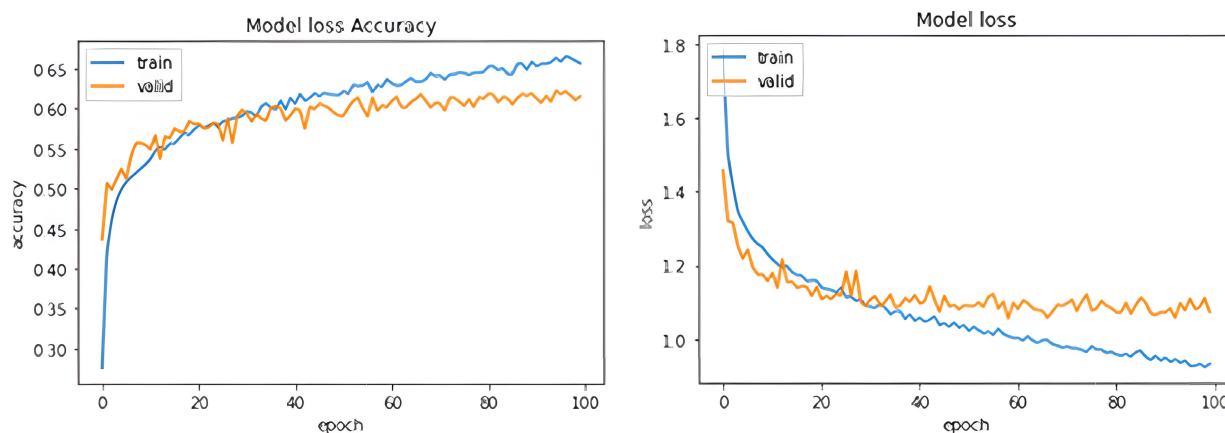**FIGURE 9. Fully connected neural network for the emotions.**

**FIGURE 10. Accuracy-Loss curves for training and validation, FER13.**


FIGURE 11. Accuracy-Loss curves for training and validation, CK+.

Let's assume binary classification with $Pr = TP/(TP + FP)$; $Re = (TP + FN)$; $F1 = 2TP/(2TP + FP + FN)$, the trivial solution to $Pr = Re = F1$ is $TP = 0$. The precision, recall, and $F1$ can have the same value in general. If we solve the system of equations, we find another solution: $FP = FN$. So, if the number of false positives is the same as the number of false negatives, all three metrics have identical values. Even if the accuracy is manually calculated, the confusion

matrices predicts five samples wrong out of 170, so when a formula for accuracy is applied, the same 0.9705% accuracy for the test set is achieved.

D. COMPARISON WITH BASE LINE
This section performs the comparison of the proposed method with the existing method. Table 1 shows that the proposed method improved the results from 89.4% to 97.24 on the
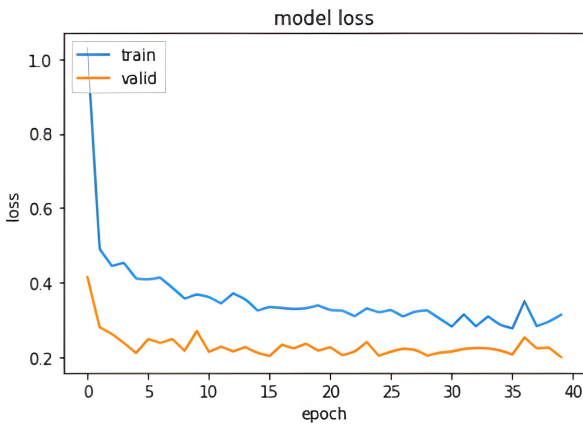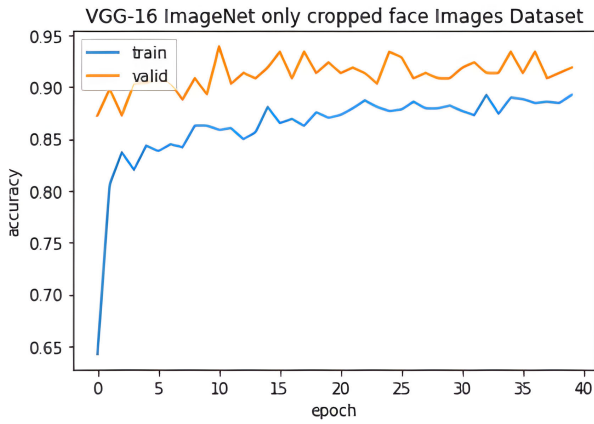
[INFO] accuracy: 91.84%
[INFO] Loss: 0.19902536273002625

FIGURE 12. Metrics of the VGG-16 model using cropped face images.

TABLE 1. Comparison of the proposed emotion classification model and baseline approach results using CK+ dataset.

Method	Accuracy %
ESRs [25]	89.4
VGG-16 [37]	89.90
AUDN [47]	92.84
EFEM [32]	92.84
Proposed Method	**97.24**

CK+ dataset. Furthermore, based on comparative analysis, it is found that the proposed method improved accuracy by 7.84% as compared to the state-of-art method presented in [25]. Similarly, it is also found that our proposed emotion classification method using the CK+ dataset improved an accuracy of 4.4% compared to the latest state-of-the-art study presented in [32].

Similarly, a detailed comparison of the proposed model is summarised in Table 2 using CK+ data set. The proposed model results are compared with three different models:

TABLE 2. A detailed comparison of the proposed emotion classification model with baseline models results using CK+ data.

Model	ACC (%)	PR %	RE %	F1 Score %
AlexNet	95.88	83.0	83.25	82.5
ResNet	93. 87	76.0	73.62	74.25
CNN	94.7	78.5	78.5	78.0
Proposed Model	**97.24**	**96.1**	**97.5**	**96.6**

TABLE 3. Comparison of the proposed emotion classification and baseline approaches results using FER dataset.

Method	Accuracy %
Transfer Learning [48]	48.50
Multiple deep network [49]	52.29
FN2EN [37]	55.15
Proposed Method	**65**

TABLE 4. A detailed comparison of the proposed emotion classification model with baseline model results using the FER dataset.

Model	ACC (%)	PR %	RE %	F1 Score %
AlexNet	63.64	62.6	60.1	60.4
ResNet	61.82	61.2	58.5	58.9
CNN	63.03	62.6	59.9	60.0
Proposed Model	**65.0**	**64.6**	**61.5**	**61.8**

AlexNet, ResNet, and CNN. The comparative analysis shows that the proposed model gives significantly better results than other baseline models. The proposed model achieved slightly higher results than the other three models in terms of accuracy. But our model achieved substantially higher performance when compared the other three matrices: accuracy (ACC), precision (PR), recall (RE), and f1 score.

Furthermore, Table 3 compares the proposed method with the existing method. It is also shown that the proposed method improved the FER dataset results from 55.15% to 65%. Based on comparative analysis, it is found that our proposed method improved by 9.75% as compared to the baseline approach presented in [37].

Similarly, Table 4 illustrates a detailed comparative analysis of the proposed model using the FER data set. This comparative analysis uses AlexNet, ResNet, and CNN as competitor models. A comparative analysis is presented in terms of accuracy, precision, recall, and f1 score to highlight the proposed model's significance compared to baseline models. The proposed model achieved 65% accuracy, which is 1.36% higher than AlexNet, 3.18% higher than ResNet, and 1.97% higher than CNN. Also proposed model gives significantly better precision, recall, and f1 score results than other baseline models.

Moreover, a comparative analysis is visualized in Fig. 15 to evaluate and compare the performance of the proposed and baseline state-of-the-art models using the CK+ dataset. The visualization analysis indicates that the proposed model achieved an accuracy of 97.24%, a precision of 96.1%, a recall of 97.5%, and an f1 score of 96.6%, demonstrating that the proposed model outperformed the state-of-the-art DL techniques, such as AlexNet, ResNet, and baseline CNN.
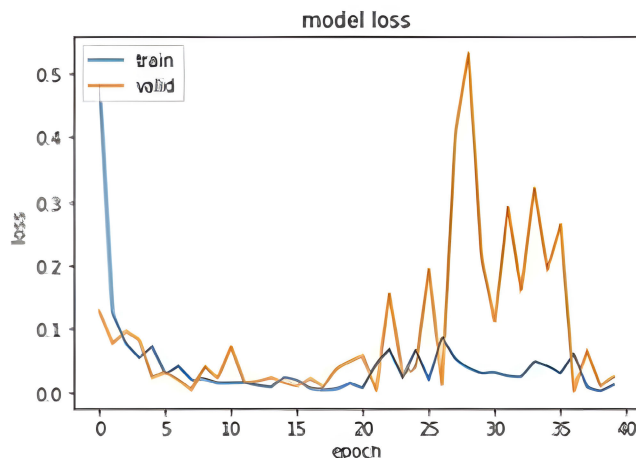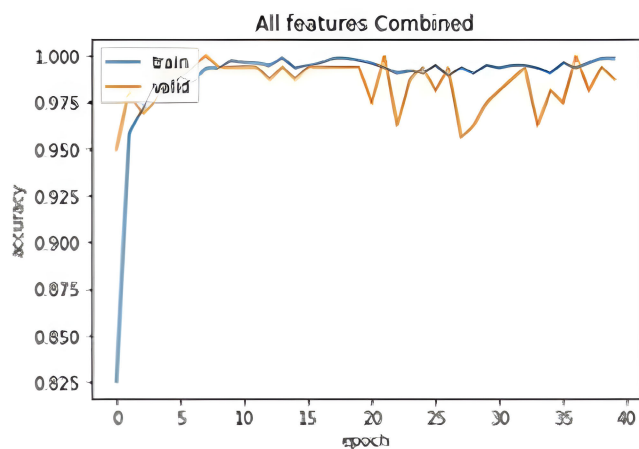
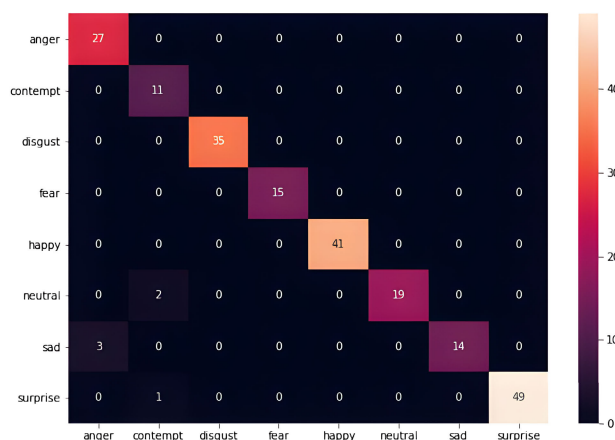**FIGURE 13.** Accuracy-Loss curves for training and validation, CK+.

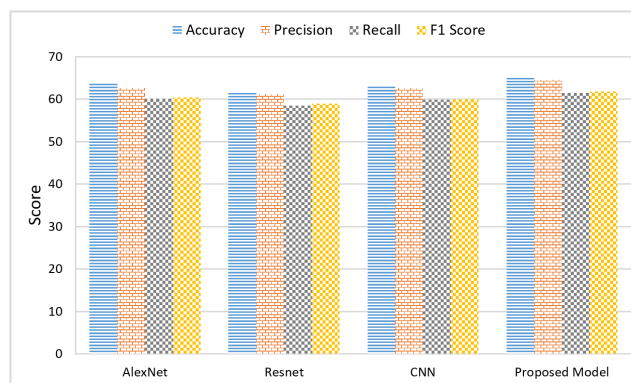**FIGURE 14.** Confusion metrics, CK+.

FIGURE 15. Comparative analysis of CK+ data with proposed and baseline models.
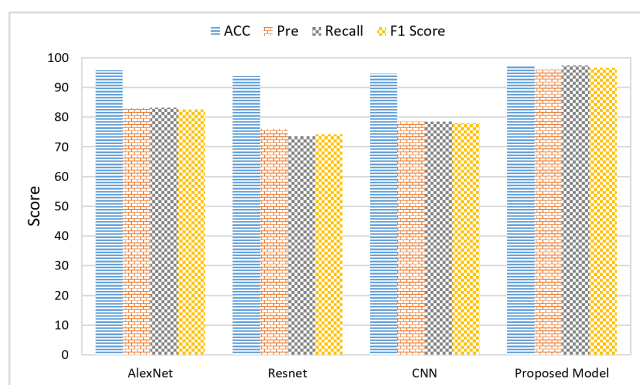

FIGURE 16. Comparative analysis of FER data with proposed and baseline models.

Furthermore, it demonstrates that the proposed model improves the accuracy of 2.01%, 3.37%, and 2.54% compared to the AlexNet, ResNet, and baseline CNN, respectively. Similarly, in terms of precision, the proposed model produced better results than the AlexNet, ResNet, and

baseline CNN models, with 13.1%, 20.1%, and 17.6%, respectively. Moreover, the proposed model also gives better results than AlexNet, ResNet, and baseline CNN models in terms of recall with 14.25%, 23.88%, and 19.0%, respectively. The performance of the proposed model in terms of f1 is improved by 14.1%, 22.35%, and 18.6%. Hence, our proposed model is reliable and effective for facial expressions classification compared to state-of-the-art techniques.

Similarly, Fig. 16 visualizes the performance of the baseline and proposed models given in Table 4 using FER+ data. The visualization analysis indicates that the proposed model achieved an accuracy of 65.0%, a precision of 64.6%, a recall of 61.5%, and an f1 score of 61.8%, which shows that the AlexNet, ResNet, and baseline CNN are all outperformed by our proposed model. As a result, the proposed model performs better than AlexNet, ResNet, and baseline CNN in terms of accuracy by 1.36%, 3.18%, and 1.97%, respectively. Similarly, in terms of precision, the proposed model produced better results than the AlexNet, ResNet, and baseline CNN models with 2.0%, 3.4%, and 2.0% respectively. Moreover, the proposed model also gives better results than AlexNet,

ResNet, and baseline CNN models in terms of recall with 1.4%, 3.0%, and 1.6%, respectively. In the case of the f1 score proposed model improved by 1.4%, 2.9%, and 1.8%. Based on these performance analysis, we conclude that our model is more reliable and effective than state-of-the-art facial expression classification techniques.

V. CONCLUSION

Facial expressions play an important role in perceiving information in communication and define the emotional state to effectively interpret human reactions and feelings. In this study, a robust deep learning-based technique is proposed for analyzing and recognizing facial expressions based on the role of zoning. Zoning of facial features is performed to divide and localize the subject face images. The role of zoning is utilized to locate more face landmarks for the identification of deep face emotions. The proposed approach is categorized into four main steps. First, face landmarks are extracted, and the second role of zoning is employed to get four regions of each localized landmark. Third, the VGG-16 learning model is used to generate the feature map and finally fully connected neural network model is used to classify facial emotions. One of the motivations of this study is to build a robust FER model and realize it on open-source libraries and coding frameworks. Moreover, this study used two benchmark datasets of facial images called CK+ and FER2013 for the evaluation of the proposed approach. The accuracy of the proposed technique with face features on CK+ and FER2013 are 98.4% and 65%, respectively. The experimental zoning results improve from 98.47% to 98.74% on the CK+ dataset. The performance analysis results provide promising directions and motivation for applying facial features in the feature extraction process to improve the FER model's performance. The zoning of images indicates that the application of zoning in the feature extraction process s improves the FER model's performance. However, the full datasets should be explored more in-depth to enhance and solve the problem of occlusion in facial images.

In future work, we will realize the proposed model in smart environments based on online face recognition systems in the Internet of Things environment. Furthermore, the performance of the proposed approach can be further improved by using ensemble learning and hybrid optimization mechanisms. Moreover, the architecture of the fully connected network will be improved to handle complex data.

REFERENCES

[1] L. Kessous, G. Castellano, and G. Caridakis, "Multimodal emotion recognition in speech-based interaction using facial expression, body gesture and acoustic analysis," *J. Multimodal User Interfaces*, vol. 3, nos. 1–2, pp. 33–48, Mar. 2010.

[2] S. Ahmad and D. H. Kim, "A task orchestration approach for efficient mountain fire detection based on microservice and predictive analysis in IoT environment," *J. Intell. Fuzzy Syst.*, vol. 40, no. 3, pp. 5681–5696, Mar. 2021.

[3] S. Wang, H. Hu, and Z. Wang, "The visualization of head-nodding synchronization in online sessions," HIT2021, Vancouver, BC, Canada, Tech. Rep., Apr. 2019.

[4] C. Frith, "Role of facial expressions in social interactions," *Phil. Trans. Roy. Soc. B, Biol. Sci.*, vol. 364, no. 1535, pp. 3453–3458, Dec. 2009.

[5] A. Pennycook, "Actions speak louder than words: Paralanguage, communication, and education," *Tesol Quart.*, vol. 19, no. 2, pp. 259–282, 1985.

[6] U. Hess and P. Thibault, "Darwin and emotion expression," *Amer. Psychologist*, vol. 64, no. 2, p. 120, 2009.

[7] K. Zhang, Z. Zhang, Z. Li, and Y. Qiao, "Joint face detection and alignment using multitask cascaded convolutional networks," *IEEE Signal Process. Lett.*, vol. 23, no. 10, pp. 1499–1503, Oct. 2016.

[8] H. Chen, Y. Wang, T. Guo, C. Xu, Y. Deng, Z. Liu, S. Ma, C. Xu, C. Xu, and W. Gao, "Pre-trained image processing transformer," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2021, pp. 12299–12310.

[9] J. L. Lukinac, M. J. Jukić, K. M. Mastanjević, and M. L. Lučan, "Application of computer vision and image analysis method in cheese-quality evaluation: A review," *Ukrainian Food J.*, vol. 7, no. 2, pp. 192–214, Jun. 2018.

[10] Y. Kortli, M. Jridi, A. Al Falou, and M. Atri, "Face recognition systems: A survey," *Sensors*, vol. 20, no. 2, p. 342, 2020.

[11] L. Boussaad and A. Boucetta, "Deep-learning based descriptors in application to aging problem in face recognition," *J. King Saud Univ.-Comput. Inf. Sci.*, vol. 34, no. 6, pp. 2975–2981, Jun. 2022.

[12] N. Ahmad, K. Iqbal, L. Han, N. Iqbal, and M. A. Abid, "Zone based lossy image compression using discrete wavelet and discrete cosine transformations," in *Proc. Int. Conf. Secur. Privacy New Comput. Environ.* Cham, Switzerland: Springer, 2019, pp. 731–742.

[13] Z. Zhang, P. Luo, C.-C. Loy, and X. Tang, "Learning social relation traits from face images," in *Proc. IEEE Int. Conf. Comput. Vis. (ICCV)*, Dec. 2015, pp. 3631–3639.

[14] B.-K. Kim, S.-Y. Dong, J. Roh, G. Kim, and S.-Y. Lee, "Fusing aligned and non-aligned face information for automatic affect recognition in the wild: A deep learning approach," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. Workshops (CVPRW)*, Jun. 2016, pp. 48–57.

[15] E. Barsoum, C. Zhang, C. C. Ferrer, and Z. Zhang, "Training deep networks for facial expression recognition with crowd-sourced label distribution," in *Proc. 18th ACM Int. Conf. Multimodal Interact.*, Oct. 2016, pp. 279–283.

[16] D. G. R. Kola and S. K. Samayamantula, "A novel approach for facial expression recognition using local binary pattern with adaptive window," *Multimedia Tools Appl.*, vol. 80, no. 2, pp. 2243–2262, Jan. 2021.

[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, Jan. 2018.

[18] P. Barros, N. Churamani, and A. Sciutti, "The FaceChannel: A fast and furious deep neural network for facial expression recognition," *Social Netw. Comput. Sci.*, vol. 1, no. 6, pp. 1–10, Nov. 2020.

[19] S. A. Fatima, A. Kumar, and S. S. Raoof, "Real time emotion detection of humans using mini-Xception algorithm," *IOP Conf., Mater. Sci. Eng.*, vol. 1042, no. 1, 2021, Art. no. 012027.

[20] S. Umer, R. K. Rout, C. Pero, and M. Nappi, "Facial expression recognition with trade-offs between data augmentation and deep learning features," *J. Ambient Intell. Humanized Comput.*, vol. 13, no. 2, pp. 721–735, Feb. 2022.

[21] Q. T. Ngoc, S. Lee, and B. C. Song, "Facial landmark-based emotion recognition via directed graph neural network," *Electronics*, vol. 9, no. 5, p. 764, May 2020.

[22] A. A. Albraikan, J. S. Alzahrani, R. Alshahrani, A. Yafoz, R. Alsini, A. M. Hilal, A. Alkhayyat, and D. Gupta, "Intelligent facial expression recognition and classification using optimal deep transfer learning model," *Image Vis. Comput.*, vol. 128, Dec. 2022, Art. no. 104583. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0262885622002128

[23] C. Gan, J. Xiao, Z. Wang, Z. Zhang, and Q. Zhu, "Facial expression recognition using densely connected convolutional neural network and hierarchical spatial attention," *Image Vis. Comput.*, vol. 117, Jan. 2022, Art. no. 104342. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S026288562100247X

[24] S. Li and W. Deng, "Deep facial expression recognition: A survey," *IEEE Trans. Affect. Comput.*, vol. 13, no. 3, pp. 1195–1215, Jul. 2022.

[25] H. Siqueira, S. Magg, and S. Wermter, "Efficient facial feature learning with wide ensemble-based convolutional neural networks," in *Proc. AAAI Conf. Artif. Intell.*, 2020, vol. 34, no. 4, pp. 5800–5809.

[26] K. Wang, X. Peng, J. Yang, S. Lu, and Y. Qiao, "Suppressing uncertainties for large-scale facial expression recognition," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2020, pp. 6897–6906.

[27] J. H. Kim, A. Poulose, and D. S. Han, "The extensive usage of the facial image threshing machine for facial emotion recognition performance," *Sensors*, vol. 21, no. 6, p. 2026, Mar. 2021.

[28] T. TanujaPatgar, "Convolution neural network based emotion classification cognitive modelforfacial expression," *Turkish J. Comput. Math. Educ.*, vol. 12, no. 13, pp. 6718–6739, 2021.

[29] F. M. A. Mazen, A. A. Nashat, and R. A. A. A. A. Seoud, "Real time face expression recognition along with balanced FER2013 dataset using CycleGAN," *Int. J. Adv. Comput. Sci. Appl.*, vol. 12, no. 6, pp. 1–8, 2021.

[30] R. K. Mishra, S. Urolagin, J. A. A. Jothi, and P. Gaur, "Deep hybrid learning for facial expression binary classifications and predictions," *Image Vis. Comput.*, vol. 128, Dec. 2022, Art. no. 104573. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0262885622002025

[31] C. Benegui and R. T. Ionescu, "Convolutional neural networks for user identification based on motion sensors represented as images," *IEEE Access*, vol. 8, pp. 61255–61266, 2020.

[32] L. Wang, Z. He, B. Meng, K. Liu, Q. Dou, and X. Yang, "Two-pathway attention network for real-time facial expression recognition," *J. Real-Time Image Process.*, vol. 18, no. 4, pp. 1173–1182, Aug. 2021.

[33] D. Kollias, V. Sharmanska, and S. Zafeiriou, "Distribution matching for heterogeneous multi-task learning: A large-scale face study," 2021, *arXiv:2105.03790*.

[34] H. Sikkandar and R. Thiyagarajan, "Deep learning based facial expression recognition using improved cat swarm optimization," *J. Ambient Intell. Humanized Comput.*, vol. 12, no. 2, pp. 3037–3053, Feb. 2021.

[35] A. H. Farzaneh and X. Qi, "Facial expression recognition in the wild via deep attentive center loss," in *Proc. IEEE Winter Conf. Appl. Comput. Vis. (WACV)*, Jan. 2021, pp. 2402–2411.

[36] S. M. S. Abdullah and A. M. Abdulazeez, "Facial expression recognition based on deep learning convolution neural network: A review," *J. Soft Comput. Data Mining*, vol. 2, no. 1, pp. 53–65, 2021.

[37] H. Ding, S. K. Zhou, and R. Chellappa, "FaceNet2ExpNet: Regularizing a deep face recognition net for expression recognition," in *Proc. 12th IEEE Int. Conf. Autom. Face Gesture Recognit.*, May 2017, pp. 118–126.

[38] I. J. Goodfellow et al., "Challenges in representation learning: A report on three machine learning contests," in *Proc. Int. Conf. Neural Inf. Process.* Cham, Switzerland: Springer, 2013, pp. 117–124.

[39] P. Lucey, J. F. Cohn, T. Kanade, J. Saragih, Z. Ambadar, and I. Matthews, "The extended Cohn-Kanade dataset (CK+): A complete dataset for action unit and emotion-specified expression," in *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit.-Workshops*, Jun. 2010, pp. 94–101.

[40] S. Battiato, G. Gallo, and F. Stanco, "A locally adaptive zooming algorithm for digital images," *Image Vis. Comput.*, vol. 20, no. 11, pp. 805–812, Sep. 2002.

[41] C. Alippi, S. Disabato, and M. Roveri, "Moving convolutional neural networks to embedded systems: The AlexNet and VGG-16 case," in *Proc. 17th ACM/IEEE Int. Conf. Inf. Process. Sensor Netw. (IPSN)*, Apr. 2018, pp. 212–223.

[42] J. Ryu, M.-H. Yang, and J. Lim, "DFT-based transformation invariant pooling layer for visual classification," in *Proc. Eur. Conf. Comput. Vis. (ECCV)*, 2018, pp. 84–99.

[43] G. Pirlo and D. Impedovo, "Adaptive membership functions for handwritten character recognition by Voronoi-based image zoning," *IEEE Trans. Image Process.*, vol. 21, no. 9, pp. 3827–3837, Sep. 2012.

[44] W. Ma and J. Lu, "An equivalence of fully connected layer and convolutional layer," 2017, *arXiv:1712.01252*.

[45] P. Baldi and P. J. Sadowski, "Understanding dropout," in *Proc. Adv. Neural Inf. Process. Syst.*, vol. 26, 2013, pp. 1–15.

[46] Y. Liu, X. Wang, L. Wang, and D. Liu, "A modified leaky ReLU scheme (MLRS) for topology optimization with multiple materials," *Appl. Math. Comput.*, vol. 352, pp. 188–204, Jul. 2019.

[47] M. Liu, S. Li, S. Shan, and X. Chen, "AU-aware deep networks for facial expression recognition," in *Proc. 10th IEEE Int. Conf. Workshops Autom. Face Gesture Recognit. (FG)*, Apr. 2013, pp. 1–6.

[48] H.-W. Ng, V. D. Nguyen, V. Vonikakis, and S. Winkler, "Deep learning for emotion recognition on small datasets using transfer learning," in *Proc. ACM Int. Conf. Multimodal Interact.*, Nov. 2015, pp. 443–449.

[49] Z. Yu and C. Zhang, "Image based static facial expression recognition with multiple deep network learning," in *Proc. ACM Int. Conf. Multimodal Interact.*, Nov. 2015, pp. 435–442.

TAIMUR SHAHZAD is a highly skilled computer science professional with a Master of Science degree in Computer Science from COMSATS University Islamabad Attock Campus. He has a deep understanding of Neural Network and Machine Learning and has successfully applied these skills to his work in the software deployment industry. His main focus is to provide fully customized solutions to organizations, designed to meet their specific needs and requirements. With his dedication to delivering innovative and effective technology solutions, he is well-positioned to continue making valuable contributions to the industry.

KHALID IQBAL received the B.Sc. degree from the University of the Punjab, Lahore, the M.S. (CS) degree from SZABIST, Karachi, and the Ph.D. degree in applied computer technology from the University of Science and Technology Beijing, in 2014. He is currently an Associate Professor with the Department of Computer Science, COMSATS University Islamabad, Attock Campus. He is also the Director of the Pattern Recognition, Images and Data Engineering (PRIDE) Research Group, where graduate students are working under his supervision in different fields, such as data mining, social networks, image processing, and simulation-based models. He has worked on Bayesian network application for privacy preserving of XML association rules and text localization in scene images. His research work has been published in several international conference proceedings and journals. His research interests include pattern recognition, machine leaning, data mining, and social networks. He was a recipient of the CSC Scholarship and QCRI/Boeing Travel Grant. He was awarded a fully funded scholarship by the China Scholarship Council for the entire duration of the Ph.D. studies. He also won the Excellent Researcher (or Excellent International Research Student Award) from the University of Science and Technology Beijing, from 2012 to 2013.

MURAD ALI KHAN received the B.S. degree in computer science from COMSATS University Islamabad, Attock Campus, Punjab, Pakistan, in 2020. He is currently pursuing the integrated Ph.D. degree with the Department of Computer Engineering, Jeju National University, Republic of Korea. He has professional experience in the software development industry and academics as well. His research interests include ML and data mining related applications.

IMRAN received the Ph.D. degree from the Computer Engineering Department, Jeju National University, Republic of Korea. He worked as a Researcher with MCL and JNU Big Data Laboratory. He is currently an Assistant Professor with the Department of Biomedical Engineering, IT Convergence College, Gachon University, Republic of Korea. His research interests include software development, IT convergence-based solutions, and entrepreneurship. His research mainly focuses on interdisciplinary scientific applications based on the Internet of Things, machine learning, data science, and blockchain.

NAEEM IQBAL (Member, IEEE) received the B.S. degree in computer science from COMSATS University Islamabad, Attock Campus, Punjab, Pakistan, the M.S. degree in computer science from COMSATS University Islamabad, Attock Campus, in 2019. He is currently pursuing the Ph.D. degree with the Department of Computer Engineering, Jeju National University, Republic of Korea. He has professional experience in the software development industry and in academic as well. He has published more than 30 papers in peer-reviewed international journals and conferences. His research interests include AI-based intelligent systems, data science, big data analytics, machine learning, deep learning, analysis of optimization algorithms, the IoT, and blockchain-based secured applications. He is serving as a professional reviewer for various well-reputed journals and conferences.

● ● ●