DOI: 10.15849/IJASCA.240730.19

Int. J. Advance Soft Compu. Appl, Vol. 16, No. 2, July 2024
Print ISSN: 2710-1274, Online ISSN: 2074-8523

Emotion Unveiled: A Deep Learning Odyssey in Facial Expression Analysis for Intelligent HCI

Shadi AlZu'bi[1*], Mohammed Elbes[1], and Ala Mughaid[2]

[1] Department of Computer Science. Faculty of Sciences and IT.
Al Zaytoonah University of Jordan. P.O.Box 130 Amman (11733), Jordan

[2] Department of Information Technology.
Faculty of prince Al-Hussien bin Abdullah for IT.
The Hashemite University. P.O. Box 330127, Zarqa (13133), Jordan

[*] smalzubi@zuj.edu.jo

Received 10 January 2024; Accepted 1 April 2024

Abstract

Examining facial expressions is a crucial aspect that garners attention due to its significance in divulging emotional states. This study delves into employing a robust deep learning method for automatically analyzing and identifying facial emotions in images. The chosen technique revolves around the convolutional neural network (CNN) algorithm. A dataset containing images of individuals, each exhibiting distinct facial expressions, was curated. The emotions in these images were categorized into seven groups (angry, disgust, fear, happy, neutral, sad, surprise) based on the depicted emotional states. The approach comprises four primary steps: preprocessing the input facial images, utilizing image adjustments and data augmentation, employing the Viola and Jones technique for face detection and landmark localization, creating a numerical feature vector from the registered image for feature extraction, and inputting the extracted features into the CNN for classification. The proposed CNN model underwent application to classify facial emotions within the image dataset. Additionally, a pretrained VGG-16 model was incorporated into the classification process for facial images. A comparative analysis between the proposed CNN approach and the pretrained VGG-16 model revealed that the latter outperformed the former in terms of accuracy rate and loss function values when determining individuals' facial expressions.

Keywords: *facial emotion detection, computer vision, convolutional neural network, deep learning, facial expression recognition*

Received 2 January 2024; Accepted 1 March 2024

1 Introduction

In real life, people express their emotions on their face to show their psychological activities and attitudes in the interaction with other people. Nowadays, facial emotions detection of human becomes an important issue in social communications and many other applications such as computer vision, pattern recognition, human computer interaction, security, cognitive science, and image understanding. Mainly, it helps in identifying the state of human emotion (e. g. neutral, happy, sad, surprise, fear, anger, disgust...) based on specific features which appear on the face of a human. The recent researches of the scientific community pay attention to facial expressions detection and its techniques Alzu'bi et al. [2019], AlZu'bi et al. [2022a], Yin et al. [2007]. In general, most systems of facial expressions recognition have three main steps: face detection in an image, facial features extraction, and facial expressions classification.

Human face is complex to interpret, so emotions recognition can be divided into the classification of basic emotions and the classification of compound emotions. The main challenge in facial emotion detection is to automatically detect the facial emotion state in an accurate way. Moreover, it is difficult to find the similarity of the same emotion state between different people since they may express the same emotion state in different ways. Due to all of these reasons, a lot of effort has been made to develop reliable and automatic facial emotion recognition systems based on the use of classification machine learning methods and deep learning techniques Zhao and Pietikäinen [2009], Zhang et al. [2017], Nigam et al. [2018], Clawson et al. [2018], Bendjillali et al. [2019], Li et al. [2020], Minaee et al. [2021], Umer et al. [2022], Boughida et al. [2022], Kavitha and RajivKannan [2023]. Facial emotion recognition commonly employs the support vector machine (SVM) as a machine learning technique Cortes and Vapnik [1995]. Conversely, convolutional neural networks (CNN) Gu et al. [2018] emerge as a prevalent deep learning method extensively applied in computer vision and the recognition of facial expressions.

This paper concentrates on the automated classification of seven fundamental human facial emotions: happy, sad, surprised, angry, disgust, neutral, and fear. The study introduces a classification model employing CNN deep learning method, capable of accurately and automatically categorizing the seven basic emotions. Notably, the system can determine the emotional state depicted in a grayscale input image, assigning it to one of the specified facial expressions. Figure 1 visually depicts the seven basic facial emotions used in the model's classification within this research. The subsequent sections are organized as follows: Section 2 reviews related work, Section 3 outlines the methodology of the proposed system and the utilized CNN model, Section 4 details experimental evaluations and analysis, and finally, Section 5 presents

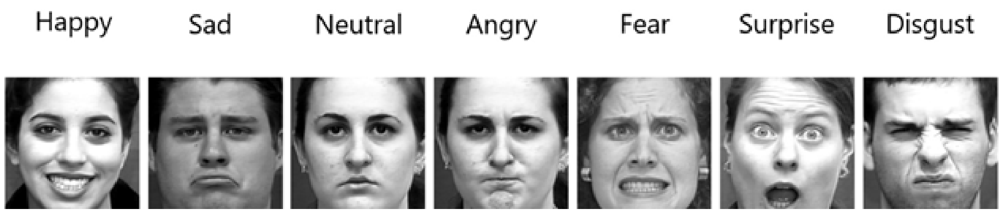

Figure 1: The basic facial expression emotions

the paper's conclusion and outlines potential future work.

2 Related Works

Previous researchers have extensively explored diverse realms, delving into areas such as facial emotion detection, computer vision, and human-computer interaction Adaileh [2020], Alflaieh [2022], Alqudah [2023]. In the pursuit of understanding psychological activities and attitudes through facial expressions, their work has laid crucial foundations for applications in pattern recognition, cognitive science, and image understanding AlZu'bi et al. [2022b], AlZu'bi et al. [2022], Mughaid et al. [2023]. The scientific community's attention has particularly centered on facial expressions detection techniques, with a focus on face detection, feature extraction, and expressions classification in several fields including legal environments Barqawi [2023], Alrai [2023] and educational fields Abdalla et al. [2022], Al-Shafei [2022]. Noteworthy efforts have been dedicated to developing reliable automatic recognition systems, employing a spectrum of classification machine learning methods and deep learning techniques.

The employment of deep learning in facial expression analysis explores the intricate interplay between emotions and technology, paving the way for advanced human-computer interactions through sophisticated facial expression analysis. many previous research have been conducted in this field AlZoubi et al. [2024].

Recently, the scientific research has shown an increasing attention in the domain of facial expressions. Many studies applied different machine learning and deep learning techniques for recognizing facial emotions automatically. In this section, the recent research on facial expression recognition will be described.

As an illustration, the investigation referenced in Zhao and Pietikäinen [2009] utilized the spatiotemporal local binary pattern operator from three orthogonal planes (LBP-TOP) to extract features. The AdaBoost technique was employed for the selection of crucial expression-related features, and the final classification of facial expressions was carried out using the support vector machine (SVM) algorithm.

A different investigation outlined in Zhang et al. [2017] introduced a method relying on convolutional neural networks (CNN) to identify human facial expressions in facial images sourced from the JAFFE and extended Cohn-Kanade (CK+) datasets. The technique involved isolating the face foreground from the background in the images. Subsequently, certain patches of facial components were produced, containing both local and global personal identity information. These generated image patches were later input into a CNN for the purpose of classifying facial expressions.

The approach demonstrated by Nigam et al. [2018]) utilized the SVM machine learning algorithm on a set of facial images, taken from CK+, JAFFE, and Yale facial expression datasets, for facial expression recognition. The first phase of this approach was face processing, which is consisted of these three steps: face detection, cropping, and normalization. Then, discrete wavelet transform (DWT) was applied on the images for transforming the spatial domain features into the frequency domain. In the feature extraction phase, the Histogram of Oriented Gradients (HOG) feature in the DWT domain was retrieved. Then, this feature was fed into the SVM classifier for the classification of facial expressions. The experimental results proved that the applied approach was efficient and outperformed the existing methods used for facial expressions recognition.

In Clawson et al. [2018], the authors introduced a CNN deep learning approach centered around human characteristics for facial emotion recognition. Their method underwent evaluation on the CK+ dataset, achieving a notable classification accuracy of 93.3% in recognizing human facial expressions.

The investigation highlighted in Bendjillali et al. [2019] presented a method for facial expression identification, combining the discrete wavelet transform (DWT) for feature extraction and the convolutional neural network (CNN) deep learning algorithm. The approach involved four crucial steps: utilizing the Viola–Jones face detection algorithm Viola and Jones [2001] for facial localization, enhancing facial images with the contrast limited adaptive histogram equalization (CLAHE) algorithm, employing DWT for facial feature extraction, and training the CNN with these features for automatic facial expression classification. The Viola–Jones algorithm detected facial features, the CLAHE algorithm improved facial image quality, and DWT extracted facial features. The CNN model achieved a classification accuracy of 96.46% on the CK+ dataset and 98.43% on the JAFFE dataset.

In contrast, the research detailed in Li et al. [2020] introduced a CNN-based approach for recognizing diverse facial expressions in images. This method incorporated local binary patterns (LBP) features and convolutional features to enhance the CNN model's performance. The authors created the Nanchang University Facial Expression (NCUFE) dataset, annotated with seven facial expressions (anger, disgust, fear, happy, sad, surprise, and neutral). Evaluation on the NCUFE dataset and four primary facial expression datasets (Oulu-CASIA, JAFFE, CK+, and FER2013) demonstrated the superior efficacy of this approach compared to various existing methods.

The authors in Minaee et al. [2021] described a deep learning approach for facial expression recognition. This approach applied an attentional convolutional neural network on the FER-2013, CK+, FERG, and JAFFE datasets. A visualization technique, that finds the most salient regions of a facial image, was also used to discover different emotions based on the output of the classifier. The experimental results of their approach showed that the utilized approach outperformed the existing approaches for facial expressions recognition in terms of accuracy. It also showed that different emotions were sensitive to different chunks of the face.

The research highlighted in Umer et al. [2022] introduced a system designed for identifying various expressions within the human face region. Utilizing the CNN method based on deep learning and incorporating data augmentation techniques, the system aimed to classify different facial expressions. When implemented on the KDEF dataset (seven expression classes), GENKI-4k dataset (two expression classes), and CK+ dataset (seven expression classes), the system exhibited promising results in accurately classifying human facial expressions.

The work outlined in Boughida et al. [2022] presented a facial expression recognition technique incorporating Gabor filters and a genetic algorithm. Gabor features were extracted from designated regions of interest on the human face, pinpointed through facial landmarks. The genetic algorithm played a role in refining Support Vector Machine (SVM) hyperparameters and selecting the most optimal features. Experimental results of this method demonstrated classification accuracy rates of 96.30%, 94.20%, and 94.26% for the JAFFE, CK, and CK+ datasets, respectively.

In contrast, the authors Kavitha and RajivKannan [2023] combined the hybrid CNN model and the Long Short-Term Memories (LSTMs) approach for facial expressions recognition. The Viola-Jones (VJ) algorithm was also used for recognizing faces from the preprocessed facial images. Their combined approach was evaluated on the Facial Expression Recog Image Ver (FERC) dataset and shown effective results.

3 Methodology

Facial emotion recognition typically involves a four-step process. Initially, a face is detected in an image, and a rectangle is outlined around it. Subsequently, features within this facial region are identified as the second step. The third stage involves extracting spatio-temporal features from the facial components. In the last step, a feature extraction classifier is utilized to pro-

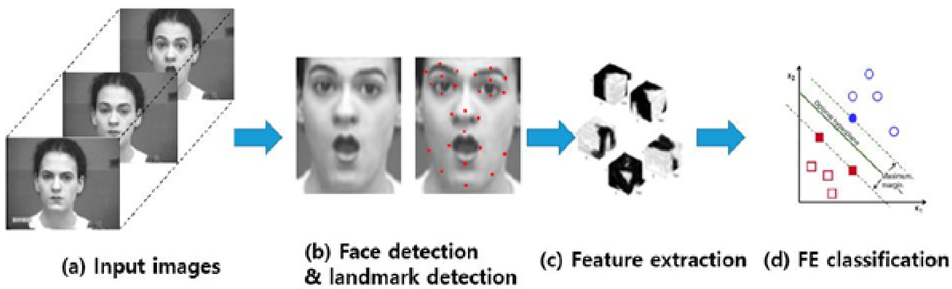

Figure 2: The facial emotion recognition procedure

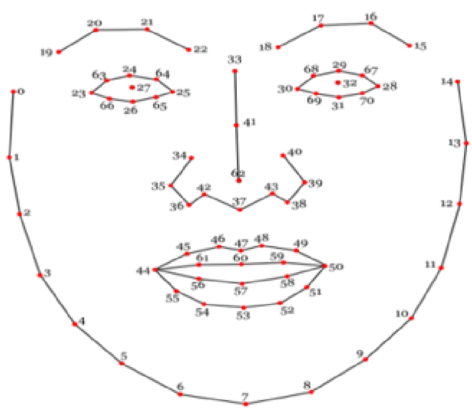

Figure 3: Visually prominent points of the contours of the face

duce recognition results based on the extracted features. Figure 2 visualizes the facial emotion recognition process in an input image, covering the detection of the facial area and facial features.

The facial contours consist of visually distinctive points, including the tip of the nose, the endpoints of the eyebrows, and the mouth, as depicted in Figure 3.

In this research, we have fetched a dataset of facial images of human feelings, where this dataset has been divided into two groups. The first group contains a large number of images that the deep learning method will be trained on, and these images have facial expressions of anger, disgust, fear, happiness, assertiveness... The second group will be used to verify whether the prediction of feelings made by the applied deep learning method is true. Each image in the first group was labeled with a number and a type of the facial expression or feeling. We counted the number of pictures to know the number of pictures in each one of the 7 groups of emotions (happy, sad, surprised, angry, disgust, neutral, and fear). We noticed that there is a group that contains few pictures, so we generated more images from the training and verification data for this group. This was done by using the ImageDataGenerator technique in order to generate more images artificially by giving them some properties. After that, the model was built using the Adam optimizer Kingma and Ba [2014] and we fitted the model with the training and validation data, and then we saved the model. Lastly, we displayed the results of the facial expression classification in the form of diagrams and the accuracy measure was estimated for the model as well as the accuracy of drawing, loss, and some other information. The achieved results were good and the constructed model was able to identify accurately the facial expressions captured by the camera.

3.1 The Proposed System

The proposed system in this research is based on the following six steps:

1. Input image
2. Processing
3. Face registration
4. Facial features extraction
5. Emotion rating
6. Output / Identified Expression

A description for each one of the above steps for the proposed system is given below:

1. preprocessing: Image preprocessing is a crucial stage in computer vision, encompassing various processes in this study. These processes involve enhancing image clarity and scaling, adjusting contrast, and employing additional enhancements to improve expression frames. The size of all images is standardized to 48x48 pixels. To enhance the model's resilience to noise and subtle transformations, data augmentation is incorporated. Each image undergoes ten amplifications through distinct linear transformations, including horizontal flipping, rotation with a random angle, skewing the center area, and zooming at the four corners of the image.

2. Face detection: The second component of the proposed system involves creating a module for face detection and landmark localization in the image. The face detection algorithm employed is based on the work of Viola and Jones Viola and Jones [2001]. This method represents an image using a set of Haar-like features, including two-, three-, and four-rectangular features (as illustrated in Figure 4). Computed in a single

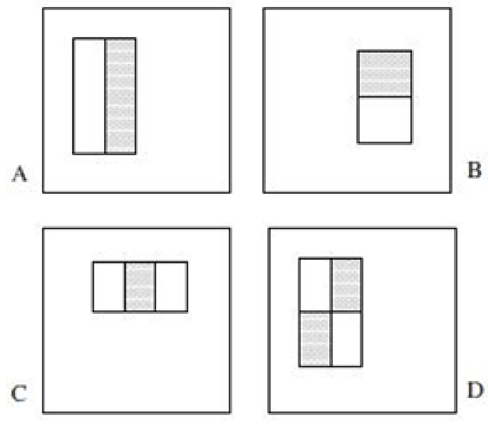

Figure 4: Face detection and landmark localization-A

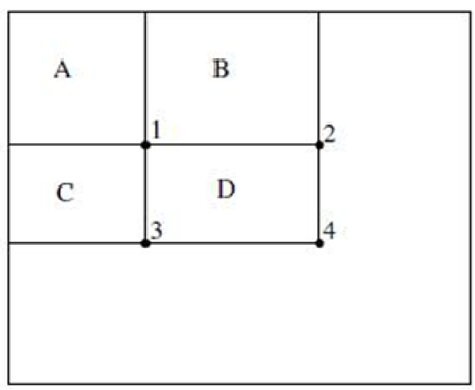

Figure 5: Face detection and landmark localization=B

pass, these features are efficient, as each rectangle only requires 4-pixel references (depicted in Figure 5). This allows for rapid computation, aiding in determining whether the image contains the sought object (face) or not.

This algorithm is widely utilized for face detection and has the versatility to be trained for detecting various objects. Its speed and efficiency make it suitable for real-time applications. In the proposed system, this algorithm is implemented for localizing faces, eyes, and mouths, utilizing pre-trained classifiers from the OpenCV library.

The face detection procedure involves sequential steps applied to the input image. The flow of the procedure is depicted through the output of each function called on the input image. Initially, the classifier trained for face detection scans the image (as shown in Figure 6). If no face is

Figure 6: Face detection

Figure 7: Eyes and mouth detection

detected, further processing is carried out, and the system generates an appropriate error message.

Upon successful face detection, the classifiers for eye detection are selectively applied to the upper part of the face. The left and right eyes are individually detected in the corresponding left and right upper face regions (illustrated in Figure 7). Subsequently, the fourth classifier is utilized to locate the mouth region in the lower part of the face. To enhance the time efficiency of the algorithm, the search area for the detectors of facial elements is appropriately narrowed.

With the locations of the face and facial landmarks determined, the face representation can be constructed. In cases where multiple faces are detected in the image, the algorithm selects the largest one for subsequent processing.

3. Feature Extraction: Feature extraction plays a vital role in face recognition, entailing the identification of distinct regions, points, landmarks, or curves/contours in a given 2-D image or a 3D range image. In this phase, a numerical feature vector is created from the registered image. Frequently extracted features from images encompass lips, eyes, eyebrows, and the nose tip. Figure 8 provides visual examples of the feature extraction process.

4. Emotion Classification: after extracting and preparing the most distinguished features from the face images, the images are classified to the appropriate class label of facial expressions by utilizing a particular clas-

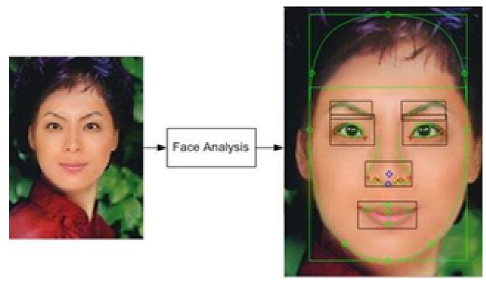

Figure 8: Facial features extraction

Figure 9: The seven basic emotions

sification technique. The classification algorithm employed in this study endeavors to categorize presented faces into one of the seven fundamental emotions. The effectiveness and precision of the facial expression recognition system hinge on the employed techniques for feature extraction and classification. In this investigation, the Convolutional Neural Network (CNN) is utilized to classify facial expressions into one of the seven basic emotions. Figure 9 provides visual examples of images displaying the seven basic emotions (neutral, happy, sad, surprise, fear, angry, disgust).

3.2 The Dataset

In this research, the dataset employed consists of 35,887 facial images, each standardized to dimensions of 48x48 pixels. To ensure uniformity, the faces have undergone automatic registration, ensuring approximate centering and consistent spatial occupancy in every image. These images represent facial expressions categorized into seven distinct classes: (0=Angry, 1=Disgust, 2=Fear, 3=Happy, 4=Sad, 5=Surprise, 6=Neutral). The dataset is divided into approximately 28,000 images for the training set and 7,000 images for the validation set. The original data is structured as arrays, with each pixel assigned a grayscale value. Conversion of this data into raw images was performed, and these images were systematically organized into distinct folders, allocating 80% to the training folder and the remaining 20% to the validation folder. For a detailed breakdown of the count of training images corresponding to each expression, refer to Table 1.

Table 1: Number of training images for different expressions

Expression	**Number of Images**
Angry	3993
Disgust	436
Fear	4103
Happy	7164
Neutral	4982
Sad	4938
Surprise	3205

Based on the conducted counting process and analysis of the images of the used dataset, there are 28,821 images which belong to the training class and 7,066 images that belong to the validation class. The maximum number of images was for the happy emotion, while the minimum number of images was for the disgust emotion.

3.3 The Proposed Model

The proposed work is an automatic facial emotion analysis and recognition system using a robust deep learning technique. The study involves the creation of a dataset containing facial images categorized into seven emotion groups. The research focuses on classifying seven basic facial emotions (happy, sad, surprised, angry, disgust, neutral, and fear). The work contributes to the field of computer vision, pattern recognition, and human-computer interaction, addressing the importance of automatic facial emotion recognition in various applications. Figure 10 illustrates an overview of the proposed Facial Expression Recognition system, while figure 11 illustrates an explanation of each emotion compared to the emoji stile. In the related work given in Section 2 of this research, the recent studies have shown the efficiency and the effectiveness of using the CNN deep learning algorithm for facial emotions recognition comparing with different machine learning techniques. For this reason, the convolutional neural network (CNN) algorithm is used in this research for the process of classifying the facial emotions.

The convolutional neural network (CNN) deep learning algorithm is a type of artificial neural network (ANN), used widely in the image processing and computer vision fields. It is specifically designed for processing pixel data and structured arrays of data such as images. This deep learning algorithm selects the best properties (features) in an image without any human supervision. Generally, a CNN is a feedforward neural network with up to 20 or 30 layers. It consists of an input layer, hidden layers, and an output layer.

Each layer in the network comprises a set of nodes, where each node in

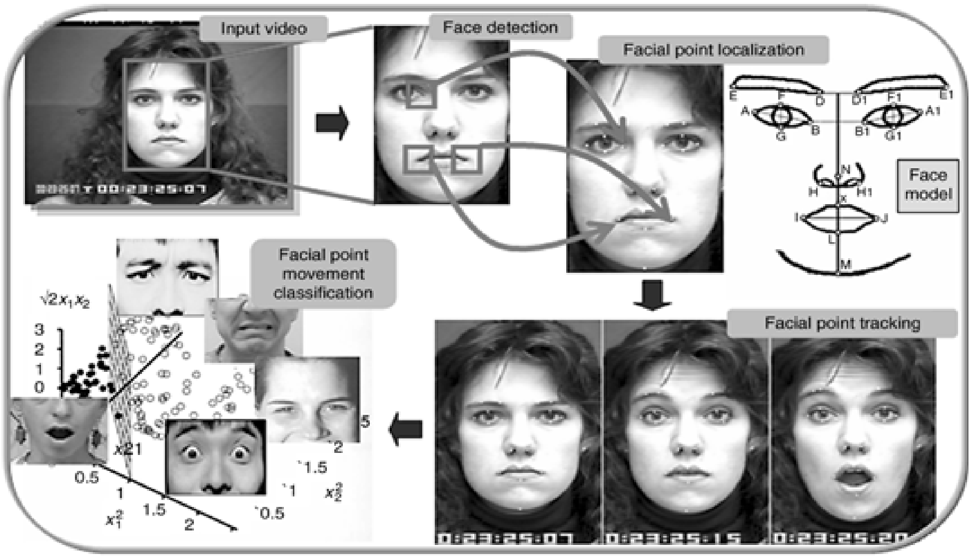

Figure 10: The Proposed system

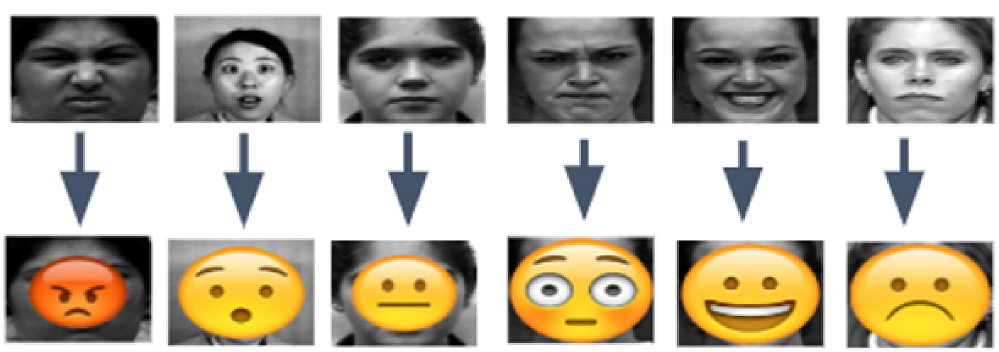

Figure 11: Emoji Description for basic emotions

one layer is connected to every node in the next layer. Every node receives a weighted input, processes it through an activation function, and outputs the network's result. Specifically, a Convolutional Neural Network (CNN) consists of one or more convolutional layers (hidden layers), a pooling layer, and is followed by one or more fully connected layers, similar to a multilayer neural network. Convolutional layers employ sliding functions on adjacent groups of pixels in an image, enhancing their ability to comprehend patterns observed in images Gu et al. [2018].

In this study, we establish our CNN with the following global architecture:

1. Four convolutional layers

2. Two fully connected layer

We have chosen the layered desks that we will be using in our project. From the Keras Python library, we have imported the following layers: Dense, Dropout, Flatten, Conv2D, BatchNormalization, Activation, MaxPooling2D, and Sequential. Then, the process has been optimized using Adam optimizer. The following paragraphs gives a brief description for each one of the above-mentioned layers.

1. **Dense Layer:** This layer represents a standard deeply connected neural network layer where each node receives input from all nodes of its previous layer.

2. **Dropout Layer:** Dropout is employed to mitigate model overfitting by randomly setting outgoing edges of hidden nodes to 0 during each update of the training phase. The last activation function chosen is softmax, commonly used for multi-label classification.

3. **Flatten Layer:** This layer transforms a matrix into a feature vector.

4. **Conv2D Layer:** The Conv2D layer in Keras is a 2D convolution layer that creates a convolution kernel convolved with the layer input, producing a tensor of outputs.

5. **Batch Normalization Layer:** Batch Normalization enhances performance and stability of Artificial Neural Network (ANN) models by normalizing inputs to have zero mean and unit variance, thus stabilizing the learning process, especially in training very deep neural networks.

6. **Activation Function:** The rectified linear activation function (ReLU) is commonly used in neural networks due to its ease of training and often better performance.

7. **MaxPooling Layer:** MaxPooling calculates the maximum value in each patch of each feature map, aiding in downsampling feature maps and highlighting the presence of features.

8. **Sequential Model:** Sequential model is suitable for a stack of layers with one input tensor and one output tensor per layer, forming the complete network.

9. **Adam Optimizer:** Adam optimizer is commonly used for optimizing neural networks, adapting learning rates during training by maintaining an exponentially decaying average of past gradients and squared gradients.

10. **Padding Parameter:** Padding parameter in the Conv2D class ensures uniform length of input sequences, with options including 'valid' for no zero-padding and 'same' for preserving spatial dimensions.

In our approach, the categorical cross- entropy function was used as a loss function, where it is quite relevant for classification tasks. The formula of the categorical cross- entropy function is given below:

$\text{Loss} = -\sum_{c=1}^{m} \left(y_{o,c} \log \left(p_{o,c} \right) \right)$

Where:

y: binary indicator (0 or 1)
p: predicted probability
m: number of classes (angry / disgust / fear / happy / neutral / sad / surprise)

In this investigation, we employed an additional model, namely the VGG-16 algorithm Russakovsky et al. [2015], to conduct facial image classification, demonstrating the efficacy of our CNN approach.

The VGG-16 algorithm Russakovsky et al. [2015] is a convolutional neural network comprising 16 layers with associated weights. Its input is an image sized at 224x224 pixels, with three channels: red (R), green (G), and blue (B). The architecture includes 13 convolutional layers, 3 fully connected layers, max-pooling layers for volume size reduction, and a SoftMax activation function, culminating in the final fully connected layer. Notably, this network is substantial, boasting approximately 138 million parameters.

4 Experimental Results

In this section, the classification rate of facial emotions recognition is estimated for our CNN approach on the applied dataset of facial images. Moreover, another model called a pretrained VGG-16 model was used in the classification process to enhance the effectiveness of the approach. Different experiments have been carried out for the CNN and VGG-16 models. We carried out experiments to compare between the two models for 50 epochs and measured the accuracy and loss functions for each model. Additionally, more 3 experiments have been conducted for both models by changing the learning rate values in epoch 3 only to be 0.01 in the first experiment, 0.001 in the second experiment, and 0.0001 in the third experiment. The accuracy rates and the loss function values have been also estimated for both models for the different learning rate values in epoch 3. The number of trainable and non-trainable parameters used for each model was also estimated in the performed experiments. The following table 2 shows the number of parameters used in each one of the two models:

Model	**CNN**	**VGG-16**
Total parameters	4,478,727	6,981,895
Trainable parameters	4,474,759	6,976,903
Non-trainable parameters	3,968	4,992

Table 2: Comparison of total and trainable parameters between the CNN and VGG-16 models.

Table 3 shows the accuracy and loss functions for both the CNN and VGG-16 models for 50 epochs on each model. Figure 12 is a graphical illustration of the achieved results.

Model	**loss**	**accuracy**	**val_loss**	**val_accuracy**
CNN	1.0678	0.8927	1.0497	0.9048
VGG-16	1.0265	0.9095	1.0812	0.9074

Table 3: Comparison of loss and accuracy metrics between the CNN and VGG-16 models.

5 Conclusions and Future Work

In this study, a comprehensive approach for automatic facial emotion analysis and recognition is presented, leveraging a robust deep learning technique, specifically the convolutional neural network (CNN) algorithm. The methodology includes dataset creation, image processing, Viola and Jones face detection, landmark localization, and feature extraction. The CNN model, along with a pretrained VGG-16 model, classifies seven basic facial emotions.
The achieved results demonstrate competitive performance, with the CNN model exhibiting a training accuracy of 89.27% and validation accuracy of 90.48%, while the VGG-16 model achieves a training accuracy of 90.95% and validation accuracy of 90.74%.

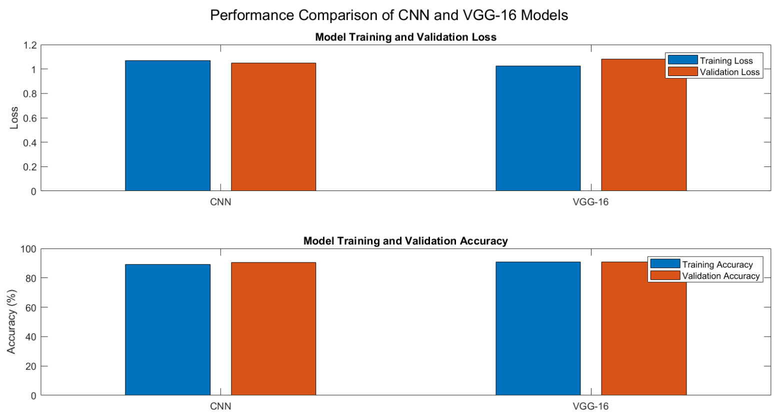

Figure 12: Performance evalutation comparison between the CNN and VGG-16 models

The research significantly contributes to computer vision, pattern recognition, and human-computer interaction. Future work entails refining model accuracy, exploring real-time applications, and further optimizing the classification process. An intelligent application could involve integrating emotion-aware interfaces for more nuanced and responsive human-machine interactions.

Acnowledgement

We gratefully acknowledge the support of the Deanship of Scientific Research at Al-Zaytoonah University of Jordan for supporting this work via Grant # 08/10/2022-2023.

References

Shadi Alzu'bi, Omar Badarneh, Bilal Hawashin, Mahmoud Al-Ayyoub, Nouh Alhindawi, and Yaser Jararweh. Multi-label emotion classification for arabic tweets. In *2019 Sixth International Conference on Social Networks Analysis, Management and Security (SNAMS)*, pages 499–504. IEEE, 2019.

Shadi AlZu'bi, Darah Aqel, and Mohammad Lafi. An intelligent system for blood donation process optimization-smart techniques for minimizing blood wastages. *Cluster Computing*, 25(5):3617–3627, 2022a.

J. Yin, Y. Li, and J. Li. Face feature extraction based on principle discriminant information analysis. In *2007 IEEE International Conference on Automation and Logistics*, pages 1580–1584. IEEE, 2007.

G. Zhao and M. Pietikäinen. Boosted multi-resolution spatiotemporal descriptors for facial expression recognition. *Pattern Recognition Letters*, 30(12): 1117–1127, 2009.

C. Zhang, P. Wang, K. Chen, and J. K. Kämäräinen. Identity-aware convolutional neural networks for facial expression recognition. *Journal of Systems Engineering and Electronics*, 28(4):784–792, 2017.

S. Nigam, R. Singh, and A. K. Misra. Efficient facial expression recognition using histogram of oriented gradients in wavelet domain. *Multimedia Tools and Applications*, 77:28725–28747, 2018.

K. Clawson, L. Delicato, and C. Bowerman. Human centric facial expression recognition. *Proc. Br. HCI*, pages 1–12, 2018.

R. I. Bendjillali, M. Beladgham, K. Merit, and A. Taleb-Ahmed. Improved facial expression recognition based on dwt feature for deep cn. *Electronics*, 8(3):324, 2019.

J. Li, K. Jin, D. Zhou, N. Kubota, and Z. Ju. Attention mechanism-based cnn for facial expression recognition. *Neurocomputing*, 411:340–350, 2020.

S. Minaee, M. Minaei, and A. Abdolrashidi. Deep-emotion: Facial expression recognition using attentional convolutional network. *Sensors*, 21(9):3046, 2021.

S. Umer, R. K. Rout, C. Pero, and M. Nappi. Facial expression recognition with trade-offs between data augmentation and deep learning features. *Journal of Ambient Intelligence and Humanized Computing*, pages 1–15, 2022.

A. Boughida, M. N. Kouahla, and Y. Lafifi. A novel approach for facial expression recognition based on gabor filters and genetic algorithm. *Evolving Systems*, 13(2):331–345, 2022.

M. N. Kavitha and A. RajivKannan. Hybrid convolutional neural network and long short-term memory approach for facial expression recognition. *Intelligent Automation & Soft Computing*, 35(1), 2023.

C. Cortes and V. Vapnik. Support vector machine. *Machine Learning*, 20(3): 273–297, 1995.

J. Gu, Z. Wang, J. Kuen, L. Ma, A. Shahroudy, B. Shuai, and T. Chen. Recent advances in convolutional neural networks. *Pattern Recognition*, 77:354–377, 2018.

A. Adaileh. Virtual digital currency as a method for funding terrorism. *Al-Zaytoonah University of Jordan Journal for Legal Studies*, 1(1):32–56, 2020.

M. T. Alflaieh. Electronic fraud in the context of e-commerce under jordanian legislation. *Al-Zaytoonah University of Jordan Journal for Legal Studies*, 3 (3):67–82, 2022. 10.15849/ZUJJLS.221130.04.

F. Alqudah. The scope of electronic arbitration in the settlement of electronic commerce disputes. *Al-Zaytoonah University of Jordan Journal for Legal Studies*, 4(2):214–224, 2023. 10.15849/ZUJJLS.230730.11.

Shadi AlZu'bi, Raed Abu Zitar, Bilal Hawashin, Samia Abu Shanab, Amjed Zraiqat, Ala Mughaid, Khaled H Almotairi, and Laith Abualigah. A novel deep learning technique for detecting emotional impact in online education. *Electronics*, 11(18):2964, 2022b.

Shadi AlZu'bi, Amjed Zraiqat, and Samar Hendawi. Sustainable development: A semantics-aware trends for movies recommendation system using modern nlp. *International Journal of Advances in Soft Computing & Its Applications*, 14(3), 2022.

Ala Mughaid, Ibrahim Obeidat, Shadi AlZu'bi, Esraa Abu Elsoud, Asma Alnajjar, Anas Ratib Alsoud, and Laith Abualigah. A novel machine learning and face recognition technique for fake accounts detection system on cyber social networks. *Multimedia Tools and Applications*, 82(17):26353–26378, 2023.

L. Barqawi. The impact of using artificial intelligence in pharmaceutical companies. *Al-Zaytoonah University of Jordan Journal for Legal Studies*, 4(1): 217–236, 2023. 10.15849/ZUJJLS.230330.11.

A. Alrai. Investigating and inferring crimes through artificial intelligence systems. *Al-Zaytoonah University of Jordan Journal for Legal Studies*, 4(1): 156–175, 2023. 10.15849/ZUJJLS.230330.08.

A. Abdalla, E. Aljheme, and F. Abdulhadi. The impact of using visual materials in enhancing learning english vocabulary at libyan preparatory schools. *Al-Zaytoonah University of Jordan Journal for Human and Social Studies*, 3(1):217–229, 2022. 10.15849/ZJJHSS.220330.12.

H. Al-Shafei. Computerization of programs for teaching arabic to non-native speakers: Android applications as a model. *Al-Zaytoonah University of Jordan Journal for Human and Social Studies*, 3:301–323, 2022. 10.15849/ZJJHSS.220508.15. (special issue).

Omar AlZoubi, Mohammad Abu Awad, Ayman M Abdalla, and Laaly Samrraie. Varicocele detection in ultrasound images using deep learning. *Multimedia Tools and Applications*, pages 1–18, 2024.

P. Viola and M. Jones. Robust real-time object detection. In *Proceedings of the Second International Workshop on Statistical and Computational Theories of Vision*, Vancouver, CA, Canada, 2001.

D. P. Kingma and J. Ba. Adam: A method for stochastic optimization. *arXiv preprint arXiv:1412.6980*, 2014.

O. Russakovsky, J. Deng, H. Su, J. Krause, S. Satheesh, S. Ma, and L. Fei-Fei. Imagenet large scale visual recognition challenge. *International Journal of Computer Vision*, 115:211–252, 2015.