Spectrum of Engineering Sciences
ISSN (e) 3007-3138 (p) 3007-312X Volume 3, Issue 4, 2025

A ROBUST CONVOLUTIONAL NEURAL NETWORK-BASED APPROACH FOR HUMAN EMOTION CLASSIFICATION: CROSS-DATASET VALIDATION AND GENERALIZATION

Ghulam Gilanie[*1], Syed Naseem Abbas[2], Sana Cheema[3], Akkasha Latif[4], **Muhammad Ahsan**[5], Hina Shafique[6], **Anum Saher**[7], Syeda Naila Batool[8], Fatima Bibi[9], Hunaina Arshad Chohan[10], Iqra Mubeen[11]

[*1,3,4,6,7,8,9,10,11]*Department of Artificial Intelligence, Faculty of Computing, the Islamia University of Bahawalpur, Pakistan*
[2,5]*Department of Computer Science, Faculty of Computing, the Islamia University of Bahawalpur*

[*1]ghulam.gilanie@iub.edu.pk, [2]nasim.naqvi@iub.edu.pk, [3]sanacheema887@gmail.com, [4]akashacheema70@gmail.com, [5]chahsan146@gmail.com, [6]hinach1912@gmail.com, [7]saheranum1@gmail.com, [8]nailashah313@gmail.com, [9]fatimaibrahim.edu.pk@gmail.com, [10]hunaina.edu.pk@gmail.com, [11]iqramubeen085@gmail.com

DOI: https://doi.org/10.5281/zenodo.15267426

Keywords
Facial Expressions Recognition, Human Social Interaction, Convolutional Neural Networks

Article History
Received on 14 March 2025
Accepted on 14 April 2025
Published on 23 April 2025

Corresponding Author: *
Ghulam Gilanie

Abstract
This paper presents advancements in the field of facial expression recognition, emphasizing the development and application of machine learning algorithms to enhance the automated understanding of human emotions. We detail our methodological approach, which integrates deep learning techniques with real-time image processing to identify and classify a wide range of facial expressions across diverse human populations. The core of our research lies in the creation and validation of a robust model trained on an extensive dataset that includes varied emotional states captured under different environmental conditions. Our findings demonstrate significant improvements in accuracy and speed over existing systems, highlighting the model's adaptability to both static images and dynamic video streams. We also address the challenges related to the interpretability of machine learning models in this context and propose solutions to increase transparency and reliability. The implications of this technology for applications in sectors such as security, healthcare, and customer service are discussed, along with ethical considerations concerning privacy and the potential for bias. The paper concludes with suggestions for further research to refine these technologies and expand their practical applications. The proposed model of facial expression recognition can deal with subjects of any ethnicity.

INTRODUCTION

Emotions are feelings that correspond with a person's mood and internal state. They are mental states influenced by individual circumstances and behaviors. In humans, the face plays a crucial role in conveying a person's current mood, as facial expressions can vividly demonstrate emotions [1]. This is why psychologists often interpret a person's feelings by observing their facial expressions, a skill that requires professional expertise. However, accurately interpreting these expressions can be challenging due to individual differences. In the 1970s, psychologist Paul Ekman identified six basic emotions happiness, sadness, disgust, fear, anger, and surprise—that are universally experienced across different ethnicities, ages, genders, races, and cultures [2]. Robert Plutchik, another psychologist, proposed that combinations of these basic emotions could lead to new, complex emotions [3]. For example, combining joy and trust can give rise to love [4]. Facial expressions are key to identifying emotions, as each emotion is associated with distinct expressions, which can be analyzed by skilled psychologists [5]. Yet, for the average person, discerning genuine from feigned expressions can often be difficult [6]. Emotions are instinctive responses that guide our reactions [7]. They are evident even in babies who cannot name or fully express these emotions. As we mature, we develop the ability to recognize and articulate our emotions, a skill enhanced by emotional awareness [8]. This skill aids in understanding our desires, resolving personal issues, managing relationships, and overcoming past difficulties [9].

We are living in the age of technology, socially active now, and daily communication with many people without principal meetings [10]. A small phone call does not explain a person's mood, and we do not have any information about the emotions of another person, so a small harsh word can spoil a relationship. Similarly, most of the people around us are good at hiding their feelings and emotions [11]. Most of them are suffering from some issues but they do not want to tell anybody, or they have a lack of confidence or have some communication issues. Understanding another person's emotional state can greatly enhance communication [12], serving as a powerful tool for building trust [13]. This understanding prompts several crucial questions: How can we discern a person's mood? How do we ascertain someone's mental state? What is the best way to respond? How should we handle the situation? These questions are fundamental for effective interpersonal interactions as they guide us in choosing the right words and actions that align with the emotional context of the conversation [14]. This ability not only improves the quality of communication but also strengthens relationships by fostering empathy and understanding [15].

The answers to these questions often lie in facial expressions. Indeed, the face provides a wealth of information that can reveal much about a person's emotional state and feelings [16]. Facial expressions, driven by the movement of small muscles beneath the skin, allow observers to discern someone's current emotions [17]. These expressions serve as a form of nonverbal communication, offering cues that, when accurately interpreted, can enhance understanding and interaction without needing words [18]. The human face can configure itself in almost 16,384 unique ways by the movement of these micro motors to impose different facial expressions. These expressions are the primary way to convey social information not only in humans but also in other mammals. A human can adopt these expressions voluntarily or involuntarily, and the brain is involved in controlling these expressions [19]. Mainly, these expressions are categorized into micro and macro expressions. Micro expressions are involuntary expressions and difficult to observe as the duration of these expressions is between 1/25th to 1/3rd of a second [20]. These expressions are also called true expressions, as they are difficult to self-create [21]. These expressions are also used to identify whether a person is lying or not. Detecting subtle facial expressions with the naked eye can be challenging, but technology such as slow-motion cameras and other specialized devices can significantly aid in distinguishing these nuances [22]. Macro expressions, which last from three-quarters of a second up to two seconds, are commonly observed during daily interactions. These are also referred to as voluntary expressions and can be intentionally manipulated by individuals to present fake emotions, potentially misleading observers [22]. There are six universally recognized facial expressions corresponding to the basic emotions of happiness, sadness, fear, surprise, anger, and disgust. These expressions are consistent across different cultures and are fundamental to human nonverbal communication [23].

Happiness is the positive state of mind that comes when a person feels good, safe, relaxed, excited or joyful. Happiness exists in many shapes and sizes; it comes from a small or slight pleasure to an extremely exciting situation. According to the research, there are about 17 expressions for happiness. We can easily detect that a person is happy or not by these micro-expressions including cheeks are raised, wide smile with open mouth, teeth are exposed, the eyes squint, lower eyelid may show wrinkles or be tense, "crow's feet" around or near the outside the eyes, some possible wrinkles around the nose, the body is relaxed, an upbeat and pleasant tone of voice or corners of the lips are stretched back and upward may help to detect [24]. In fake happiness, we just smile, while real happiness involves all or most of the above points especially the movement of muscles around the eyes and nose. These expressions help us to find the emotion of happiness and to distinguish between real and fake happiness.

Sadness is a negative emotion that explains pain, loss, helplessness, grief, and disappointment. It also comes in combination with some other expressions like fear, anger, etc [25]. Loss of valued things, including people, makes us sad. Long-lasting sadness is a profound emotional state that can potentially evolve into depression if not addressed. This emotion is reflected in subtle micro-expressions, such as the inner corners of the eyebrows being drawn inward and upward, creating a triangulated shape of the skin below the eyebrows. Additionally, the corners of the lips may droop, the jaw may tighten, and the lower lip might pout [26]. These physical manifestations often include a distinct frown, which involves slanting and raising the inner eyebrows—a movement that is difficult to simulate artificially. Other signs might include tension in the neck and chin areas, often as an effort to restrain tears, along with a lowered or hunched posture and a tendency to look away or downward [27]. Fear, on the other hand, is a potent emotion triggered by perceived danger, activating the body's fight-or-flight response, which is critical for survival. This state of alertness results in physical changes such as tensed muscles, an increased heart and respiration rate, and a heightened state of mental alertness [28]. The facial expressions associated with fear are quite distinct: the eyebrows are raised and drawn together, creating wrinkles on the forehead and between the eyebrows, but not across. The eyes widen, often showing the upper white portion but not the lower, and the eyelids become tense. The mouth may open with the lips tensed or stretched back, adding to an overall expression of intense alertness and readiness to respond to the threat [29].

Surprise is another type of emotion that triggers the body into fight or flight mode. It can be categorized into three forms: positive, negative, and neutral. A positive surprise is when you reach home and find your loved ones who arranged a surprise birthday for you. A negative surprise is when someone jumps out behind you and scares you at night [30]. The emotion of surprise is marked by distinct facial expressions that are often instantaneous and involuntary. These expressions include eyebrows that are raised and curved, which contribute to the stretching of the skin below the brow and the formation of horizontal wrinkles across the forehead. The eyelids open wide, revealing a greater amount of the white of the eyes both above and below, while the jaw often drops open, and the teeth part without the tension or stretching typically seen in other expressions. When the mouth opens in surprise, the initial reaction might evolve depending on the context of the surprise. If the surprise is pleasant, the open mouth may quickly transition into an open smile [31]. Conversely, if the surprise is negative or shocking, the facial expression may shift towards those associated with anger or fear, such as tensed eyelids or additional wrinkling of the forehead [32]. These reactions underscore the rapid shift in emotional state that surprise can trigger, reflecting the immediate and often profound impact of unexpected events [33].

It is simple to create a fake surprise; widely open eyes and raised eyebrows are enough to create a fake expression. The actual difference between real and fake expressions is the timing and duration of expression [34]. Anger is a powerful emotion characterized by the feeling of frustration and antagonism. It also activates the fight-or-flight response [11]. When a stimulus generates feelings of anger, the body is ready to respond against the danger to protect yourself. It is also a negative emotion [35]. Uncontrolled anger can turn into aggression and violence. Facial expressions associated with anger are characterized by several visual clues that can indicate a person's emotional state. These signs include eyebrows that are lowered and drawn together, often forming a closed 'V' shape, and the appearance of vertical lines between the eyebrows. The eyes may take on a hard stare or appear bulging, reflecting intensity or aggression [36]. The lower eyelids are typically tensed, adding to a stern facial appearance. Lips might be pressed firmly together, with the corners turned downward, or they might take on a squared shape as if the person is shouting. Additionally, the nostrils may dilate as part of the physiological response to anger, increasing air intake [3].The lower jaw might jut out, and the mouth can be either open and squared or closed with a visibly tense chin and jawline [37]. In more dramatic expressions of anger, the eyes might squint, accompanied by a fixed icy stare, and the nose might "flare," further emphasizing the intensity of the emotion.These visual cues are crucial for recognizing anger and can help in understanding the severity and immediacy of the emotional response in various interpersonal situations [38].

In an anger situation, heartbeat and respiration rate also become high and fake anger can be distinguished by only focusing on the eyebrows [13]. In real anger, there is high tension between the eyebrows and shaped like "V". If a person is smiling, with a "V" shape between his eyes, we can guess there is something wrong. During disgust, the feeling of strong disapproval or revulsion can originate due to several things like unpleasant smell, scent, taste, or something immoral that is not acceptable by body or person [39]. It is also negative feelings or feelings of disapproval. Expression including upper lip is raised, upper teeth may be exposed, nose is wrinkled, cheeks are raised, or lowered eyebrows may help in determining the disgust condition [10]. Emotion recognition through automated systems has remained an active research area during recent years. Several Artificial Intelligence (AI) based agents required this application to recognize the emotions/mood of the subjects during interactive sessions in an automated manner [40]. Convolutional Neural Networks (CNNs) are highly effective for extracting detailed micro-expressions from facial images, making them extremely valuable for emotion recognition tasks. This capability forms the basis of Facial Expression Recognition (FER) systems, which are instrumental in a variety of applications across different fields [41, 42]. FER systems contribute significantly to understanding and analyzing human behavior and interactions. In human-machine interaction, for instance, these systems enable machines to respond to human emotions in a way that is nuanced and appropriate, enhancing the user experience. Similarly, in the field of mental health, FER can aid in the detection of mental disorders by analyzing the subtleties in facial expressions that might indicate psychological distress or other issues [43]. Additionally, FER is used in studying human psychology, helping researchers and practitioners gain insights into emotional responses and mood fluctuations. This understanding can lead to improved interpersonal relationships and more effective communication by providing a clearer comprehension of the emotional cues conveyed during interactions.

In more technical applications, such as video surveillance, FER enhances security systems by enabling the detection of suspicious behaviors or emotional states that require attention. It also plays a crucial role in face detection and tracking systems used in various security and commercial settings. Moreover, real-time emotion detection systems powered by FER technology are becoming increasingly popular in customer service and other interactive applications, where understanding and responding to user emotions can significantly impact service quality and customer satisfaction. Overall, the integration of CNNs in FER systems offers profound advantages for both technological applications and deeper human-centric studies, bridging the gap between human emotions and digital technology. The following research questions have been addressed through this conducted study: is a model trained on one dataset to recognize facial expression work on another dataset? Which one among these expressions is more likely to be determined from grey-scale images or colored images? To address these questions, an automated approach for facial expression recognition from digital images using Convolution Neural Network (CNN) has been proposed in this research work.

In Section 1, Introduction, we discuss the significance of emotions and their manifestations through facial expressions. Section 2, Related Work, reviews the existing literature on facial expression recognition technologies. In Section 3, Methods, we detail the methodological approach used in this study. Section 4, Results, presents the findings from the application of the proposed CNN model. Finally, Section 5, Conclusions and Future work, summarizes the contributions of the study and discusses the potential applications of this technology. It outlines future research directions, including the expansion of the model's capabilities to video streams and more complex emotional recognition tasks. The conclusion also addresses ethical considerations and the impact of these technologies on privacy and interpersonal communications.

1. Related work

In the research conducted by [44], A novel approach to facial expression recognition has been introduced, utilizing Convolutional Neural Networks enhanced with an attention mechanism (ACNN). This method effectively captures the facial regions of interest through an end-to-end learning framework [45], offering a more nuanced understanding of facial expressions by combining multiple representations. The approach includes two variations of ACNN: the patch-based ACNN (pACNN) and the global-local based ACNN (gACNN) [46]. The pACNN focuses on local facial patches, honing in on specific areas of the face to gather detailed emotional cues. Conversely, the gACNN combines these local representations from individual patches with global representations that consider the entire facial image [47]. This dual approach allows for a comprehensive analysis of both detailed and overall facial expressions, enhancing the accuracy and reliability of the recognition process. The effectiveness of this system has been tested on both real and synthetically generated datasets, ensuring robustness across different scenarios. For research and validation purposes, two widely recognized datasets, RAF-DB and AffectNet, were employed [48]. These datasets are globally available and provide a diverse range of facial expressions, making them ideal for rigorous testing and refinement of the facial expression recognition technology [15]. This innovative method promises significant advancements in the field, potentially improving applications in various domains such as human-computer interaction, psychological research, and security systems [49]. An input image is decomposed by ACNN for diverse local patch extraction [14], the weight vector is used for concatenating local and global gate unit images for prediction and then fed to two fully connected layers to minimize the softmax loss [43]. pACNN focused on representative and criminative patches by using region decomposition and occlusion perception, while gACNN uses integration with full face region for ensemble learning to promote diversity of learned features [12]. The experiment is performed on both types of datasets in-the-wild datasets (RAF-DB, AffectNet, SFEW) and in-the-lab datasets (CK+, MMI, Oulu-CASIA). Achieved accuracy ranges from 47.76% to 91.64%. The model utilizes multiple techniques that need further modification to enhance performance in environments with occlusion and dynamic settings [50]. It relies on facial landmarks to predict expressions, which does not always ensure accurate outcomes due to the variability in facial movements [51]. Another research proposes a new method involving a Part-Based Hierarchical Bidirectional Recurrent Neural Network (PHRNN) to capture dynamic variations of facial expressions [52]. This technique focuses on extracting temporal features from consecutive frames [53]. It combines PHRNN with a Multi-Scale Convolutional Neural Network (MSCNN) to create the Deep Evolutional Spatial-Temporal Network (DESTN), designed to extract comprehensive details including both partial and whole information, dynamic and static elements, as well as geometry and appearance, thus enhancing the Facial Expression Recognition (FER) system's performance [54]. The PHRNN model maps input sequences to sequences of hidden states to capture dynamic features [55], while the MSCNN extracts static features using four convolutional layers [56], a fully connected layer, and a softmax layer with 10 filters of size 5x5x1 [53]. By integrating these models, the system aims to improve performance and reduce error rates. The effectiveness of this integrated model was tested on three databases, each providing different results [57]. The CK+ Database, with 123 subjects and 593 sequences, showed that the spatial network model using MSCNN achieved 95.54% accuracy, and the temporal network using PHRNN achieved 96.36% accuracy. Combined, the spatial-temporal network (PHRNN-MSCNN) reached an accuracy of 98.50%. The Oulu-CASIA Database, including expressions from 80 subjects, resulted in 77.67% accuracy for the spatial model and 78.96% for the temporal model, with the combined model achieving 86.25% accuracy [58]. Lastly, the MMI Database, featuring 30 subjects with 205 sequences, yielded 77.05% accuracy for the spatial model, 76.17% for the temporal model [59], and 81.18% for the combined network model [60]. However, this method struggles with subjects in motion, highlighting a need for further enhancements to increase its viability in dynamic or real-world scenarios [61].

The researchers [62], proposed an algorithm for a facial expression recognition system that consists of a hierarchical deep neural network structure and local binary patterns (for feature extraction). The dataset is divided into 10 parts, 09 of which are used for training while the remaining is used for testing purposes. The image is passed through an appearance-based network after a geometric feature-based network using CNN captured the movement of landmarks [63]. The experiments are performed on two globally available datasets, i.e., CK+ and JAFFE. CK+ achieved 96.46% results, while the JAFFE database came up with 91.27% results when tested using the 10-cross cross-validation method [64]. Similarly, the researchers [65] introduced a model for humanoid robot facial expression detection using CNN and Long Short-Term Memory (LSTM). CNN models are used to crop and convert them into greyscale images [66]. This model is tested on multiple databases, including the AFFECTNET database (450,000 manually & 500,000 automatically annotated), and CK+ database (593 images). Experiments obtained 94.41% results. Since, the multi-modal technique is used, which requires a large amount of, labeled data for training. Further, this model is limited to a humanoid robot [67]. It is established from the literature that multiple techniques have been used to perform experiments on different datasets [68]. Luminance, head poses, and some other obstacles have a major impact on accuracy. Keeping this motivation, in the research light weighted, with a lower number of layers, size and parameters have been proposed to recognize human facial expressions in robust manners [69].A study introduced a shallow convolutional neural network (SHCNN) model to address the limitations of traditional CNN models, particularly their inefficacy with small datasets and neglect of temporal information [70], which are crucial for accurate expression detection [71]. The SHCNN was tested across multiple datasets, including FER2013, FERPlus, CASME, CASME-II, and SAMM, achieving a 75% accuracy rate.

Another research initiative reported in [72] tackled the challenge of occlusion in facial expression recognition. This study introduced a Lossless Attention Model (LLAM) specifically designed for use with CNNs. Utilizing the ResNet-18 architecture for image classification and object detection [73], the experiments were conducted on the Aff-Wild2 dataset, which includes 564 videos comprising approximately 2.8 million frames. The LLAM achieved an accuracy of 0.49 and an unweighted mean (F1) score of 0.38 [74]. Further research [75] proposed a robust vectorized convolutional neural network that incorporates an attention mechanism to focus on regions of interest (ROIs) in static images. This methodology was tested using globally available datasets such as AffectNet, CK+, Fer2013, and JAFFE, which collectively contain over 1,000,000 facial images. This approach achieved a maximum classification rate of 90%.

The literature indicates that a variety of techniques have been employed to enhance facial expression recognition across different datasets. Factors such as luminance and head poses significantly impact accuracy. Motivated by these challenges, the referenced research proposes a lighter, less complex model with fewer layers and parameters. This approach aims to robustly recognize human facial expressions even in the presence of various obstacles.

2. Methods

For facial expression recognition, traditional methods such as feature selection, extraction, reduction, and classification have been utilized by some researchers. These engineering-based approaches, while systematic, often result in time-consuming processes with relatively low performance outcomes. In contrast, deep learning-based architectures, particularly those utilizing convolutional neural networks (CNNs), offer more promising results due to their ability to automatically extract and process features directly from raw images.

This significant advantage has inspired further investigation into the potential of CNN architectures for recognizing emotions from facial images. The ability of CNNs to efficiently handle large volumes of data and their effectiveness in feature learning directly from the spatial hierarchies of images makes them particularly suitable for the complex task of emotion recognition, pushing the boundaries of what can be achieved compared to traditional techniques.

3.1 Dataset

Publicly available benchmarked datasets such as Jaffe, Senthil, KDEF, UTKFace, and FERG have been utilized for research and experimentation in the study. The data distribution for the proposed system involved allocating 60% for training, 20% for validation, and 20% for testing. The specifics of these datasets, including their size, composition, and other relevant characteristics, are detailed in Table 1 for easy reference and comparison. This structured approach to data management ensures a balanced evaluation of the system's performance across different phases of model development and assessment.

Table 1: Detailed dataset and other relevant characteristics

Sr No.	Dataset	Expressions	Real or animated	Total images
1	Jaffe	07 (06 basic expressions, 01 neutral)	Real	213
2	FERG	07 (06 basic expressions, 01 neutral)	Animated	55767
3	Senthil	06 basic expressions	Real	80
4	KDEF	06 basic expressions	Real	4900
5	UTKFace	06 basic expressions	Real	20,000

3.2. Preprocessing

All the images from each dataset were standardized by resizing them to a resolution of 512x512 pixels and converting them to PNG format. Additionally, each image was decomposed into its constituent color channels—red, green, and blue—as well as an intensity channel. This process of unfolding the images into separate channels is crucial for detailed analysis and processing in image recognition tasks, allowing for more nuanced feature extraction and improving the accuracy of the subsequent analysis shown in Figure 1.

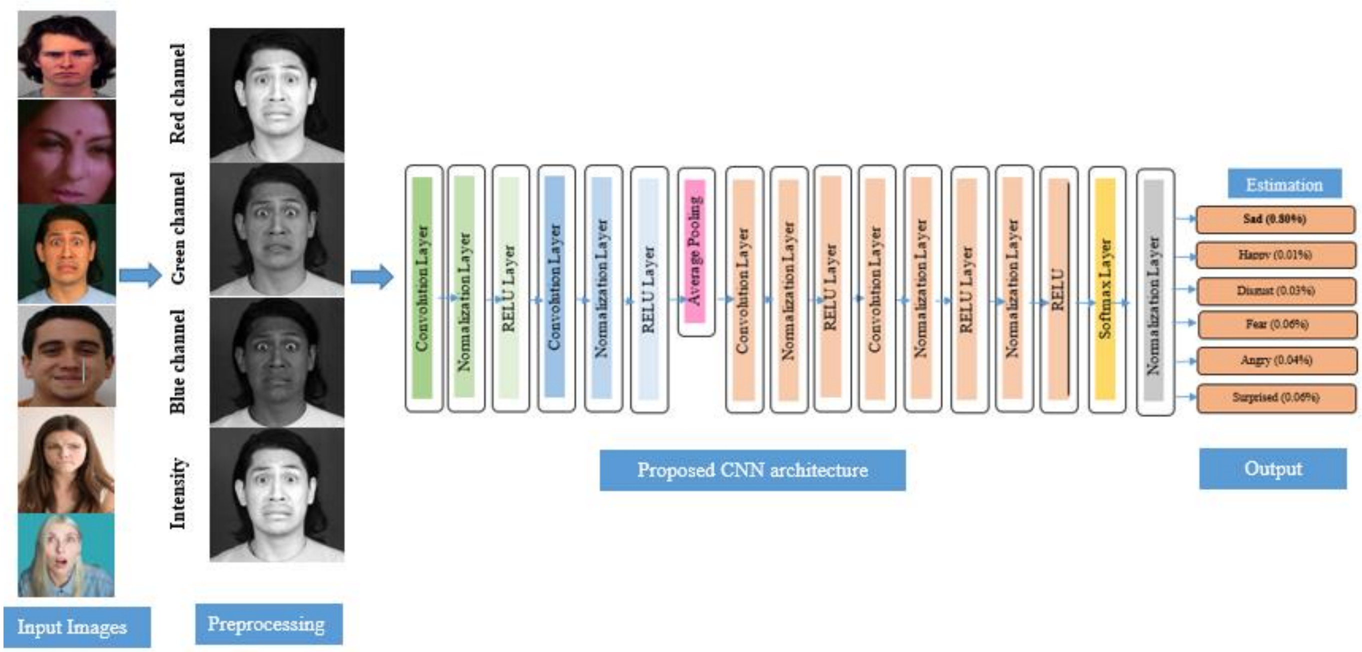

Figure 1. The proposed framework for facial expression recognition

2.3. The proposed CNN architecture

In this study focused on recognizing facial expressions, Convolutional Neural Networks (CNNs) have been thoroughly investigated. The proposed framework for this research consists of three main stages: inputting images, preprocessing these images, and classifying them using a specially designed CNN architecture.

CNNs represent a cutting-edge field in machine learning, drawing inspiration from the human brain's visual processing capabilities. Neuroscientists have identified that the neocortex, part of the cerebral cortex involved in processing visual and auditory information in mammals, manages sensory signals through a complex hierarchical system. Similarly, CNNs mimic the human visual system and are specifically tailored for analyzing two-dimensional images. This design premise allows CNNs to encode specific properties inherent in visual data effectively.

The core operation within a CNN involves convolving images with kernels (also known as filters), which helps in extracting feature maps. These feature maps represent various aspects of the input images, such as edges, textures, or specific shapes, crucial for the subsequent classification tasks. This method enables the CNN to learn and identify intricate patterns in facial expressions, thereby enhancing the accuracy and effectiveness of facial expression recognition.

Units in a feature map are connected to the previous layer through weights of kernels. Weights of kernels are tweaked during the training through backpropagation. Because the same kernels have been used by all units, convolutional layers have fewer weights to train. Information about neighborhood has been considered using kernels. The following are key components implemented within the Convolutional Neural Networks (CNNs) framework to effectively achieve facial expression recognition:

1) **Initialization**:

Initialization plays a crucial role in ensuring the convergence of the CNN model during training. It is used to maintain a proper scale of activations and gradients throughout the layers of the network, preventing issues like vanishing or exploding gradients that can hinder learning.

2) **Activation Function**:

The activation function introduces non-linearity into the network, enabling it to learn and model complex patterns in data. In this study, Rectified Linear Units (ReLU) are used as the activation function. ReLU is widely adopted due to its simplicity and effectiveness, and it is mathematically defined in Equation No. 1.

$F(x)=\max(0,x)$ 1)

3) **Convolutional Layer**: The convolutional layer is a fundamental component of CNNs that detects and summarizes the presence of specific features within the input image. It is responsible for identifying local patterns by analyzing small regions of the image, enabling the network to understand spatial hierarchies of features.

In this process, the image is divided into smaller regions, often referred to as perceptions or receptive fields. Each region is examined individually to identify patterns such as edges, textures, or other distinguishing features.

Given an input image Img with dimensions (M, N) and a filter F of size (p, q), the convolution operation is mathematically expressed in Equation No. 2. This operation involves sliding the filter across the image, starting from the top-left corner and moving to the bottom-right, applying the filter at each position. The result is a feature map, which highlights the presence and location of specific features within the input image. These feature maps serve as the building blocks for deeper layers of the network, allowing CNN to progressively learn more complex features.

$conv=(Img*F)(x,y)=\sum u\sum v Img(x-u,y-v)\cdot F(u,v)$
(2).

4) **Pooling:**

The pooling layer is utilized to combine spatially adjacent features into feature maps, helping the network retain the most significant information while reducing computational complexity. In most CNN architectures, max-pooling or average-pooling are the standard techniques used to achieve this.

In the proposed model, average pooling is employed. This method performs down-sampling by dividing the feature map into rectangular regions and computing the average of the values within each region. The result is a smaller activation map that retains the essential spatial structure of the original features while significantly reducing data dimensionality.

This process not only minimizes the size of the activation maps—making the network more efficient—but also helps prevent overfitting by simplifying the feature representations. Although max pooling is another common approach that selects the maximum value within each region and discards the rest, it is not applied in this architecture. Instead, the average-pooling technique has been preferred for its smoother feature representation, contributing to the reduction of computational load while preserving meaningful patterns in the data.

5) **Loss Function:**

The loss function is a critical component in training neural networks, as it measures the difference between the predicted output and the actual target values. The objective during training is to minimize this loss to improve the model's accuracy and performance.

In this study, Categorical Cross-Entropy is used as the loss function, which is particularly well-suited for multi-class classification problems like facial expression recognition. It evaluates how well the predicted probability distribution matches the actual distribution of the classes. The mathematical representation of this loss function is provided in Equation No. 3.

By minimizing the categorical cross-entropy during training, the model learns to make more accurate predictions, assigning higher probabilities to the correct classes and refining its understanding of the underlying features in the data.

$$H=-\sum_j p_j \log(p_j) \quad (3)$$

6) **Fully Connected Layer:** T he fully connected (FC) layer is a vital component of any CNN-based model, primarily responsible for classification and decision-making tasks. In this layer, each neuron from the previous layer is fully connected to every neuron in the hidden and output layers. This dense connectivity allows the network to combine the extracted features from earlier convolutional and pooling layers to make final predictions.

In the context of facial expression recognition, the fully connected layer plays a crucial role in analyzing the learned features and generating output probabilities for different expression categories. It essentially translates the spatial feature maps into class scores, enabling the model to identify and classify human emotions accurately based on the input facial images.

The proposed 17 layered architectures for facial expressions recognition are presented in Tables 1.

Table 1: Architecture of the CNN designed for facial expression recognition

Layer #	Type of Layer	Filter size	Stride	No. of filters
Layer-1	Convolution Layer	5x5	2x2	64
Layer-2	Normalization Layer			
Layer-3	RELU Layer			
Layer-4	Convolution Layer	5x5	2x2	128
Layer-5	Normalization Layer			
Layer-6	RELU Layer			
Layer-7	Average Pooling	5x5	2x2	128
Layer-8	Convolution Layer	5x5	2x2	128
Layer-9	Normalization Layer			
Layer-10	RELU Layer			
Layer-11	Convolution Layer	5x5	2x2	64
Layer-12	Normalization Layer	-	-	-
Layer-13	RELU Layer	-	-	-
Layer-14	Normalization Layer	-	-	-
Layer-15	RELU Layer			
Layer-16	Softmax Layer			
Layer-17	Normalization Layer			

3. Results

The classification of facial images containing various expressions involves a systematic process, beginning with preprocessing and leading up to the final recognition stage. In this study, the dataset was divided into 60% for training, 20% for validation, and 20% for testing. To prepare the data, images were preprocessed by separating them into individual channels—red, green, blue, and intensity—which helped generate uniform and descriptive input images. Multiple experiments were conducted to evaluate the effectiveness of different channels and input combinations in classifying facial expressions.

In Experiment 01, only the intensity channel was used to classify two basic expressions across 1,323 images, resulting in an accuracy of 64.85%. Experiment 02 and Experiment 03 also used the intensity channel but with 1,727 and 5,150 images respectively, both achieving an accuracy of 57.88%. In Experiment 04, the RGB channels were combined with the intensity channel to classify six expressions, yielding a 70% classification rate. However, Experiment 05, which focused on classifying two expressions using 7,000 images, resulted in a lower accuracy of 50%. Experiment 06 showed promising results by using only the green channel to classify six expressions from 1,500 images, achieving an impressive accuracy of 93.22%. Building upon this, Experiment 07 used the green channel on a larger dataset of 25,000 images and achieved the highest classification accuracy of 95.0% among all experiments. To evaluate the effectiveness of the red and blue channels, Experiment 08 and Experiment 09 were conducted, achieving accuracy of 59.10% on 1,400 images and 78.43% on 1,500 images respectively.

These results indicate that the green channel significantly contributes to improved classification performance in facial expression recognition, outperforming other channels across different dataset sizes. The results obtained using the proposed model were also compared with several state-of-the-art facial expression recognition studies, as summarized in Table 3. One notable study by [18] employed a Variational Autoencoder–Generative Adversarial Network (VAE + GAN) on a local dataset, achieving an accuracy of 91.10%. While the results are promising, the study did not use any benchmarked datasets for evaluation, which limits its generalizability. In another study [44], the researchers proposed a convolutional neural network (CNN) model enhanced with an attention mechanism. This model was evaluated on both in-the-wild datasets (RAF-DB, AffectNet, SFEW) and controlled lab datasets (CK+, MMI, Oulu-CASIA). The model achieved an impressive 91.64% accuracy for clear images and 88.17% for images with occlusion. The Part-Based Hierarchical Bidirectional Recurrent Neural Network (PHRNN), introduced by [50] , yielded an accuracy of 86.25% on the Oulu-CASIA database and 81.18% on the MMI dataset. Although this model demonstrates effective temporal analysis, its classification rates are comparatively lower than other models. Another study by [76] proposed a deep feature fusion convolutional neural network, which was tested on the Bosphorus dataset and achieved 79.17% accuracy. This, again, reflects relatively lower performance in facial expression detection. Similarly, the work presented in [77] utilized a CNN architecture in combination with a VGG pre-trained model. This model reported classification accuracies of 85.19% on the CK+ dataset, 65.17% on the JAFFE dataset, and 84.38% on the FACES dataset. While the results are considered reasonable, performance varied significantly across datasets. Another approach, De-Expression Residual Learning (DeRL), proposed in [78] , was evaluated on multiple benchmarked datasets. It achieved 88.00% on Oulu-CASIA, 73.23% on MMI, 84.17% on BU-3DFE, and 81.39% on the BP4D+ dataset. Although DeRL was tested on widely recognized datasets, its classification rates remained moderate.

Overall, these comparisons highlight that while several existing models demonstrate good accuracy, many struggle with consistency across various datasets or under challenging conditions like occlusion. The proposed model, in contrast, shows strong performance and higher classification accuracy, especially when optimized with selected channels and refined preprocessing techniques.

The proposed model, comprising a 17-layered architecture shown in Table 2, is lightweight and designed with fewer parameters compared to other CNN-based models. Despite its simplicity and efficiency, it outperforms existing state-of-the-art facial expression recognition methods by achieving the highest accuracy of 95.0%. This demonstrates the effectiveness of the proposed approach in recognizing facial expressions accurately while maintaining low computational complexity, making it suitable for real-time applications and resource-constrained environments shown in Table 3.

Table 2: Results obtained through the proposed model during different experiments

Sr. No.	Experiment No	Number of Expressions	Datasets / Channel	Number of images used	Accuracy
1	Experiment No.01	2	Local Dataset / Intensity	1323	64.85%
2	Experiment No.02	2	Local Dataset / Intensity	1727	57.88%
3	Experiment No.03	2	Local Dataset / Intensity	5150	57.88%
4	Experiment No.04	6	FERG database / IRGB	25000	70.0%
5	Experiment No.05	2	Local Dataset / IRGB	7000	50.0%
6	Experiment No.06	6	FERG database /IRGB	1500	93.22%
7	Experiment No.07	6	FERG database /IRGB	25000	95.0%
8	Experiment No.08	7	FERG database /IRGB	1400	59.10%
9	Experiment No.09	7	FERG database /IRGB	1500	78.43%

Table 3: Comparison with state-of-the-art methods

Sr. No.	Study	The proposed methodology	Dataset used	Evaluation measures achieved
1	[79]	Variational Autoencoder-Generative Adversarial Network (VAE + GAN)	Local Dataset having 237 participants	81.10%
2	[44]	Convolutional Neural Network with attention mechanism (ACNN)	In-the-wild datasets (RAF-DB, AffectNet, and SFEW) In-the-lab datasets (CK+, MMI, and Oulu-CASIA)	91.64% on clear images 88.17% for occlusion
3	[50]	Part-Based Hierarchical Bidirectional Recurrent Neural Network (PHRNN)	Global datasets (Oulu-CASIA and MMI)	86.25% on Oulu-CASIA 81.18% on MMI
4	[76]	Deep feature fusion convolutional neural networks	Bosphorus dataset	79.17%
5	[77]	Convolutional neural networks and VGG pre-trained model	CK+, JAFFE, and FACES	85.19% on CK+ 65.17% on JAFFE 84.38% on FACES
6	[78]	De-Expression Residual Learning (DeRL)	Oulu-CASIA, MMI, BU-3DFE, and BP4D+	88.00% on Oulu-CASIA 73.23% on MMI 84.17% on BU-3DFE 81.39% on BP4D+
7	**The Proposed Model**		Jaffe, FERG, Senthil, KDEF, and UTKFace	**95.0%**

4. Conclusions

In this research, CNN-based architecture has been proposed to classify the seven basic facial expressions—happy, sad, angry, neutral, disgusted, surprised, and fearful—to effectively predict a person's emotional state. The designed model is lightweight, featuring a lower number of layers, reduced size, and fewer parameters compared to other state-of-the-art CNN models, making it efficient and suitable for real-time applications. Recognizing human emotions in daily life is often complex and subjective; therefore, this study contributes significantly to improving the accuracy and reliability of mood detection through facial expressions.

To the best of our knowledge, the proposed model has achieved the highest reported accuracy of 95% among comparable methods. This performance can support a wide range of applications, including human behavior analysis, human-machine interaction, mental health assessment, psychological research, mood and response evaluation, interpersonal communication, video surveillance, face detection and tracking systems, and real-time emotion detection systems.

Looking ahead, the future scope of this work includes extending the model to detect facial expressions from video sequences for real-time tracking of individuals. Since head and body postures can impact facial recognition accuracy, efforts will also be directed toward recognizing expressions in crowded environments. Additionally, the study aims to further explore and analyze facial micro-features, which may help in identifying subtle emotional cues and predicting internal physiological or psychological changes with greater precision.

Compliance with Ethical Standards

Funding: This study has no funding.

Conflict of Interest: The authors declare that they have no conflict of interest.

Ethical approval: This article does not contain any studies with human participants performed by any of the authors.

Informed Consent: There is no individual participant included in the study.

REFERENCES

1. Saher A, Gilanie G, Cheema S, Latif A, Batool SN, Ullah H. A Deep Learning-Based Automated Approach of Schizophrenia Detection from Facial Micro-Expressions. Intelligent Automation & Soft Computing. 2024;39(6).
2. Ekman P. An argument for basic emotions. Cognition & emotion. 1992;6(3-4):169-200.
3. Janjua HU, Andleeb F, Aftab S, Hussain F, Gilanie G. Classification of liver cirrhosis with statistical analysis of texture parameters. International Journal of Optical Sciences. 2017;3(2):18-25.
4. Plutchik R. The emotions: University Press of America; 1991.
5. Bajwa UI, Shah AA, Anwar MW, Gilanie G, Ejaz Bajwa A. Computer-aided detection (CADe) system for detection of malignant lung nodules in CT slices-a key for early lung cancer detection. Current Medical Imaging. 2018;14(3):422-9.
6. Plutchik R. The nature of emotions: Human emotions have deep evolutionary roots, a fact that may explain their complexity and provide tools for clinical practice. American scientist. 2001;89(4):344-50.
7. Nazir A, Ullah H, Gilanie G, Ahmad S, Batool Z, Gadhi A. Exploring Breast Cancer Texture Analysis through Multilayer Neural Networks. Scientific Inquiry and Review. 2023;7(3):32-47.
8. Batool SN, Yang J, Gilanie G, Latif A, Yasin S, Ikram A, et al. Forensic Radiology: A robust approach to biological profile estimation from bone image analysis using deep learning. Biomedical Signal Processing and Control. 2025;105:107661.
9. Lane RD, Quinlan DM, Schwartz GE, Walker PA, Zeitlin SB. The Levels of Emotional Awareness Scale: A cognitive-developmental measure of emotion. Journal of personality assessment. 1990;55(1-2):124-34.
10. Gilanie G, Bajwa UI, Waraich MM, Asghar M, Kousar R, Kashif A, et al. Coronavirus (COVID-19) detection from chest radiology images using convolutional neural networks. Biomedical Signal Processing and Control. 2021;66:102490.
11. Gilanie G, Nasir N, Bajwa UI, Ullah H. RiceNet: convolutional neural networks-based model to classify Pakistani grown rice seed types. Multimedia Systems. 2021:1-9.
12. Janjua HU, Jahangir A, Gilanie G. Classification of chronic kidney diseases with statistical analysis of textural parameters: a data mining technique. International Journal of Optical Sciences. 2018;4(1):1-7.

13. Attique M, Gilanie G, Mehmood MS, Naweed MS, Ikram M, Kamran JA, et al. Colorization and automated segmentation of human T2 MR brain images for characterization of soft tissues. PloS one. 2012;7(3):e33616.
14. Gilanie G, Cheema S, Latif A, Saher A, Ahsan M, Ullah H, et al. A Robust Method of Bipolar Mental Illness Detection from Facial Micro Expressions Using Machine Learning Methods. Intelligent Automation & Soft Computing. 2024;39(1).
15. Naveed S, Husnain M, Alsubaie N, Samad A, Ikram A, Afreen H, et al. Drug efficacy recommendation system of glioblastoma (GBM) using deep learning. IEEE Access. 2024.
16. Russell JA, Dols JMF. The psychology of facial expression: Cambridge university press Cambridge; 1997.
17. Rafiq M, Bajwa UI, Gilanie G, Anwar W. Reconstruction of scene using corneal reflection. Multimedia Tools and Applications. 2021;80(14):21363-79.
18. Ahmed M, Gilanie G, Ahsan M, Ullah H, Sheikh FA. Review of Artificial Intelligence-based COVID-19 Detection and A CNN-based Model to Detect Covid-19 from X-Rays and CT images. VFAST Transactions on Software Engineering. 2023;11(2):100-12.
19. Shafiq H, Gilanie G, Sajid M, Ahsan M. Dental radiology: a convolutional neural network-based approach to detect dental disorders from dental images in a real-time environment. Multimedia Systems. 2023;29(6):3179-91.
20. Khera EA, Ullah H, Hussain F, Abubakar S, Majeed A, Tabssum I, et al. Characterizing nickel oxide thin films for smart window energy conversion applications: Combined experimental and theoretical analyses. ChemistrySelect. 2023;8(37):e202302320.
21. Amjad M, Ullah H, Andleeb F, Batool Z, Nazir A, Gilanie G. Fourier-Transform Infrared Spectroscopy (FTIR) for Investigation of Human Carcinoma and Leukaemia. Lasers in Engineering (Old City Publishing). 2021;51.
22. Ullah H, Batool A, Gilanie G. Classification of Brain Tumor with Statistical Analysis of Texture Parameter Using a Data Mining Technique. International Journal of Industrial Biotechnology and Biomaterials. 2018;4(2):22-36.
23. Shreve M, Godavarthy S, Goldgof D, Sarkar S, editors. Macro-and micro-expression spotting in long videos using spatio-temporal strain. Face and Gesture 2011; 2011: IEEE.
24. Gilanie G, Bajwa UI, Waraich MM, Anwar MW. Risk-free WHO grading of astrocytoma using convolutional neural networks from MRI images. Multimedia Tools and Applications. 2021;80(3):4295-306.
25. Gilanie G, Bajwa UI, Waraich MM, Anwar MW, Ullah H. An automated and risk free WHO grading of glioma from MRI images using CNN. Multimedia tools and applications. 2023;82(2):2857-69.
26. Ghani M, Gilanie G. The IOMT-Based Risk-Free Approach to Lung Disorders Detection from Exhaled Breath Examination. INTELLIGENT AUTOMATION AND SOFT COMPUTING. 2023;36(3):2835-47.
27. Asghar K, Gilanie G, Saddique M, Habib Z. Automatic Enhancement Of Digital Images Using Cubic BÃ© zier Curve And Fourier Transformation. Malaysian Journal of Computer Science. 2017;30(4):300-10.
28. Afzal F, Ullah H, Amjad M, Akhtar M, Shah MI, Batool Z, et al. Detection of Uric Acid in UV-VIS wavelength Regime. JOURNAL OF NANOSCOPE (JN). 2023;4(1):75-81.
29. Rashid MS, Gilanie G, Naveed S, Cheema S, Sajid M. Automated detection and classification of psoriasis types using deep neural networks from dermatology images. Signal, Image and Video Processing. 2024;18(1):163-72.

30. Gilanie G, Rehman N, Bajwa UI, Sharif S, Ullah H, Mushtaq MF, editors. FERNet: A Convolutional Neural Networks Based Robust Model to Recognize Human Facial Expressions. International Conference on Soft Computing and Data Mining; 2022: Springer.
31. Wazir E, Gilanie G, Rehman N, Ullah H, Mushtaq MF, editors. Early Stage Detection of Cardiac Related Diseases by Using Artificial Neural Network. International Conference on Soft Computing and Data Mining; 2022: Springer.
32. Batool SN, Gilanie G. CVIP-Net: A Convolutional Neural Network-Based Model for Forensic Radiology Image Classification. Computers, Materials & Continua. 2023;74(1).
33. Ullah H, Gilanie G, Hussain F, Ahmad E. Autocorrelation optical coherence tomography for glucose quantification in blood. Laser Physics Letters. 2015;12(12):125602.
34. Ullah H, Gilanie G, Attique M, Hamza M, Ikram M. M-mode swept source optical coherence tomography for quantification of salt concentration in blood: an in vitro study. Laser Physics. 2012;22:1002-10.
35. Gilanie G, Saher A, Batool SN, Khursheed A, Shafique H, Perveen S, et al. Digital Image Processing for Ultrasound Images: A Comprehensive. Digital Image Processing. 2021;15(3).
36. Gilanie G. Automated Detection and Classification of Brain Tumor from MRI Images using Machine Learning Methods: Department of Computer Science, COMSATS University Islamabad, Lahore campus; 2019.
37. Yaseen M, Khurshed A, Ullah H, Batool Z, Nazir A, Gilanie G, et al. In-vitro Evaluation of Anticancer Activity of Rhodamine-640 perchlorate on Rhabdomyosarcoma cell line. 2022.
38. Khera EA, Ullah H, Hussain F, Abubakar S, Majeed A, Tabssum I, et al. Characterization of Nickel Oxide Thin Films for Smart Window Energy Conversion Applications: Comprehensive Experimental and Computational Study. Available at SSRN 4235112.
39. Gilanie G, Bajwa UI, Waraich MM, Habib Z, Ullah H, Nasir M. Classification of normal and abnormal brain MRI slices using Gabor texture and support vector machines. Signal, Image and Video Processing. 2018;12:479-87.
40. Asghar S, Gilanie G, Saddique M, Ullah H, Mohamed HG, Abbasi IA, et al. Water classification using convolutional neural network. IEEE Access. 2023;11:78601-12.
41. Gilanie G, Asghar M, Qamar AM, Ullah H, Khan RU, Aslam N, et al. An Automated and Real-time Approach of Depression Detection from Facial Micro-expressions. Computers, Materials & Continua. 2022;73(2).
42. Gilanie G, Bajwa UI, Waraich MM, Habib Z. Automated and reliable brain radiology with texture analysis of magnetic resonance imaging and cross datasets validation. International Journal of Imaging Systems and Technology. 2019;29(4):531-8.
43. Gilanie G, Attique M, Naweed S, Ahmed E, Ikram M. Object extraction from T2 weighted brain MR image using histogram based gradient calculation. Pattern Recognition Letters. 2013;34(12):1356-63.
44. Li Y, Zeng J, Shan S, Chen X. Occlusion aware facial expression recognition using cnn with attention mechanism. IEEE Transactions on Image Processing. 2018;28(5):2439-50.
45. ULLAH H, BATOOL Z, NAZIR A, AHMED K, RASHEED M, AL-KHATTAF FS, et al. Potential application of CeO2/Au nanoparticles as contrast agents in optical coherence tomography. Journal of Optoelectronics and Advanced Materials. 2024;26(July-August 2024):307-15.

46. Siddique MA, Akhtar M, Majid MA, Khera EA, Ahmad M, Gilanie G, et al. A Multi-Modal Approach for Exploring Sarcoma and Carcinoma Using FTIR and Polarimetric Analysis. Microscopy Research and Technique. 2025.
47. Gilanie G, Ullah H, Mahmood M, Bajwa UI, Habib Z. Colored Representation of Brain Gray Scale MRI Images to potentially underscore the variability and sensitivity of images. Current Medical Imaging Reviews. 2018;14(4):555-60.
48. Anwaar M, Gilanie G, Ahmad F, Sharif W, Shaheen M, Ashraf M, et al. Enhanced Relative Discrimination Criterion and RSS-ELM Techniques to Improve the Document Classification. 2024.
49. Gilanie G, Bajwa UI, Waraich MM, Habib Z. Computer aided diagnosis of brain abnormalities using texture analysis of MRI images. International Journal of Imaging Systems and Technology. 2019;29(3):260-71.
50. Zhang K, Huang Y, Du Y, Wang L. Facial expression recognition based on deep evolutional spatial-temporal networks. IEEE Transactions on Image Processing. 2017;26(9):4193-203.
51. Hafeez HA, Elmagzoub MA, Abdullah NAB, Al Reshan MS, Gilanie G, Alyami S, et al. A CNN-model to classify low-grade and high-grade glioma from mri images. IEEE Access. 2023;11:46283-96.
52. Gilanie G, Shafiq H, Batool SN, Abbas SN, Shafique H, Cheema S, et al. PARAMETER OPTIMIZATION OF AUTOENCODER FOR IMAGE CLASSIFICATION USING GENETIC ALGORITHM. Spectrum of Engineering Sciences. 2025;3(4):201-13.
53. Malik S, Ullah H, Andleeb F, Batool Z, Nazir A, Gilanie G, et al. Enhancing Contrast in Optical Imaging of Cancer tissues and Study the Spectral Properties of Methylene Blue. Acta Microscópica. 2021;30(2):49-57.
54. Ullah H, Zafar M, Batool Z, Nazir A, Gilanie G, Rehman J. Early Detection of Liver, Ovary, Breast and Stomach Tumours in the Visible (λ= 630 nm) and Infrared (IR)(λ= 10.5 to 5.5 μm) Wavelength Regimes. Lasers in Engineering (Old City Publishing). 2023;54.
55. Gilanie G, Batool SN, Khursheed A, Shafique H, Mahmood N, Cheema S, et al. Bit Pattern Selection Based Novel Method of Steganography in RGB Encoding Scheme Based Digital Images.
56. Sajid M, Sharif W, Gilanie G, Mazher M, Iqbal K, Akhtar MA, et al. IoMT-Enabled Noninvasive Lungs Disease Detection and Classification Using Deep Learning-Based Analysis of Lungs Sounds. International Journal of Advanced Computer Science & Applications. 2025;16(2).
57. Gilanie G, Batool SN, Shafique H, Khursheed A, Mahmood N, Cheema S, et al. An Overview on X-Rays ImagesProcessing: Methods, Challenges& Issues, and Future Work.
58. Gilanie G, Batool SN, Abbas SN, Cheema S, Latif A, Shafique H, et al. DEEP LEARNING-BASED APPROACH FOR ESTIMATING THE AGE OF PAKISTANI-GROWN RICE SEEDS. Spectrum of Engineering Sciences. 2025;3(1):557-72.
59. Gilanie G, Javedb M, Rauf B, Cheemaa S, Latif A, Perveena S, et al. RiceAgeNet: Age Estimation of Pakistani Grown Rice Seeds using Convolutional Neural Networks.
60. Gilanie G, Batool SN, Abbas SN, Shafique H, Iqbal M, Cheema S, et al. READABLE TEXT RETRIEVAL FROM NOISE-INFLUENCED DOCUMENTS USING IMAGE RESTORATION METHODS. Spectrum of Engineering Sciences. 2025;3(3):337-60.

61. Gilanie G, Batool SN, Abbas SN, Latif A, Iqbal M, Shafique H, et al. STEGANOGRAPHIC SECRET COMMUNICATION USING RGB PIXEL ENCODING AND CRYPTOGRAPHIC SECURITY. Spectrum of Engineering Sciences. 2025;3(3):323-36.
62. Kim J-H, Kim B-G, Roy PP, Jeong D-M. Efficient facial expression recognition algorithm based on hierarchical deep neural network structure. IEEE Access. 2019;7:41273-85.
63. Rubab SF, Mushtaq MF, Tahir MH, Amna, Samad A, Gilanie G, et al., editors. The Comparative Performance of Machine Learning Models for COVID-19 Sentiment Analysis. International Conference on Soft Computing and Data Mining; 2022: Springer.
64. Ghaffar AA, Mushtaq MF, Amna, Akram U, Samad A, Gilanie G, et al., editors. Refined Sentiment Analysis by Ensembling Technique of Stacking Classifier. International Conference on Soft Computing and Data Mining; 2022: Springer.
65. Li T-HS, Kuo P-H, Tsai T-N, Luan P-C. CNN and LSTM Based Facial Expression Analysis Model for a Humanoid Robot. IEEE Access. 2019;7:93998-4011.
66. Iqbal MJ, Bajwa UI, Gilanie G, Iftikhar MA, Anwar MW. Automatic brain tumor segmentation from magnetic resonance images using superpixel-based approach. Multimedia Tools And Applications. 2022;81(27):38409-27.
67. Ullah H, Batool Z, Nazir A, Akhtar U, Gilanie G, Rehman J, et al. Proteins and Triglycerides Measurement in Blood Under Ultraviolet (UV)/Visible (Vis) Spectroscopy at & lambda;= 190 to 1100 nm with an Additional He-Ne Laser Source. LASERS IN ENGINEERING. 2023;55(3-6):157-67.
68. ULLAH H, GHALIB A, AKHTAR M, MEHMOOD R, AKRAM A, BATOOL Z, et al. Measurements of Hyperproteinaemia in Human Blood Using Laser Interferometry: In Vitro Study. Lasers in Engineering (Old City Publishing). 2024;57.
69. Ullah H, Batool Z, Nazir A, Mohsin S, Raouf A, Rasheed M, et al. Assessing Graphene Oxide (GO) and CuO Nanocomposites for Effective Antibacterial Properties Using Laser Interferometry. Lasers in Engineering (Old City Publishing). 2023;55.
70. Miao S, Xu H, Han Z, Zhu Y. Recognizing Facial Expressions Using a Shallow Convolutional Neural Network. IEEE Access. 2019;7:78000-11.
71. Gilanie G. Spectroscopy of T2 weighted brain MR image for object extraction using prior anatomical knowledge based spectroscopic histogram analysis 2013.
72. Wang C, Hu R, Hu M, Liu J, Ren T, He S, et al. Lossless Attention in Convolutional Networks for Facial Expression Recognition in the Wild. arXiv preprint arXiv:200111869. 2020.
73. Ullah H, Faran M, Batool Z, Nazir A, Gilanie G, Amin N. Diagnosis of Ocular Diseases Using Optical Coherence Tomography (OCT) at λ= 840 nm. Lasers in Engineering (Old City Publishing). 2022;53.
74. Gilanie G, Abbas SN, Cheema S, Latif A, Ahsan M, Shafique H, et al. A ROBUST ARTIFICIAL NEURAL NETWORK APPROACH FOR EARLY DETECTION OF CARDIAC DISEASES. Spectrum of Engineering Sciences. 2025;3(4):499-511.
75. Sun X, Zheng S, Fu H. ROI-Attention Vectorized CNN Model for Static Facial Expression Recognition. IEEE Access. 2020;8:7183-94.
76. Tian K, Zeng L, McGrath S, Yin Q, Wang W, editors. 3D Facial Expression Recognition Using Deep Feature Fusion CNN. 2019 30th Irish Signals and Systems Conference (ISSC); 2019: IEEE.

77. Sajjanhar A, Wu Z, Wen Q, editors. Deep learning models for facial expression recognition. 2018 Digital Image Computing: Techniques and Applications (DICTA); 2018: IEEE.
78. Yang H, Ciftci U, Yin L, editors. Facial expression recognition by de-expression residue learning. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition; 2018.
79. Saha S, Navarathna R, Helminger L, Weber RM, editors. Unsupervised deep representations for learning audience facial behaviors. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops; 2018.