Kurdistan Journal of Applied Research (KJAR)
Print-ISSN: 2411-7684 | Electronic-ISSN: 2411-7706

Original Article

Effective Facial Expression Recognition System Using Artificial Intelligence Technique

Imad Sabri Yousif [a], **Tarik A. Rashid** [a, b], **Ahmed S. Shamsaldin** [a *], **Sabat A. Abdulhameed** [a], **Abdulhady Abas Abdullah** [b]

[a] Computer Science and Engineering Department, University of Kurdistan Hewler, Erbil, Iraq.
[b] Artificial Intelligence and Innovation Center, University of Kurdistan Hewler, Erbil, Iraq.

Submitted: 13 August 2024
Revised: 11 December 2024
Accepted: 24 December 2024

* **Corresponding Author**:
ahmed.saadaldin@ukh.edu.krd

Keywords: Facial expression Recognition, Artificial neural networks, Genetic algorithms.

How to cite this paper: I. S. Yousif, T. A. Rashid, A. S. Shamsaldin, S. A. Abdulhameed, A. A. Abdullah, "Effective Facial Expression Recognition System Using Artificial Intelligence Technique", KJAR, vol. 9, no. 2, pp: 117-130. Dec. 2024, doi: 10.24017/science.2024.2.9.

Copyright: © 2024 by the authors. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY-NC-ND 4.0)

Abstract: Facial expressions are the most basic non-verbal method people use to communicate feelings, intentions and reactions without words. Recognizing these facial expressions accurately is essential for a variety of applications — such as tools that use our faces to interact with computers (human-computer interaction, or HCI), security systems and emotionally intelligent artificial intelligence technologies. As the complexities surrounding these relationships have become better understood, it has allowed us to develop increasingly more complex systems for identifying and detecting facial expressions of different emotions. This paper presents an improved performance of the Facial Expression Recognition (FER) systems via augmentation in Artificial Neural Networks and Genetic Algorithms, two renowned artificial intelligence techniques possessing disparate strengths. ANNS are inspired by the neural architecture of human brain capable of learning and recognizing patterns in unchartered data after trained examples, on the other hand GAs come from fundamental principles underlying natural selection perform optimization process based-on evolutionary methods which includes fitness evaluation, comparison, selection, crossover, and mutation. The research is an effort to mitigate the problems pertaining with conventional methods, like overfitting and generalization fault in order design FER model which has potential for performing much more accurately. A hybrid ANN-GA model that uses Petri Nets and production systems is proposed for the real-time video sequence analysis with high precision in predicting different dynamic facial activities of anger, surprise, disgust, joy, sadness and fear from emotion faces. Importantly, results on the study show that this integrated model has a large-scale promoting effect in emotion detection upon varied scenes and is therefore generalizable to many domains —from security and surveillance over biomedicine up to interactive AI-driven systems. Implications for implementing real-time and context-aware recognition of human emotions based on AI technologies are far-reaching as they demonstrate the potential that hybrid AI systems offer at enhancing emotion deciphering.

1. Introduction

Emotions are probably the strongest and most universal non-verbal communication is facial expressions. They are responsible for communicating emotions, intentions and reactions, allowing for interpersonal inter-actions between individuals from assorted cultures around the world. These expressions are not just fundamental in human-to-human interfaces, but they also play a role in helping an array of technology applications such as human-computer interaction (HCI), security systems, medical diagnostics, etc. Better movie rating prediction using the facial key point detection is possible and humanoid robots will be able to make nuanced responses if provided with excellent training sample: the expression when someone brings their A-game. Consequently, it is essential for research in artificial intelligence to focus on facial expression recognition (FER), as detecting and reading the emotion loaded signals from a face are crucial capabilities that must be developed by any system if they intend to effectively interact with human emotions. Despite the advancement in FER systems, there remains a significant challenge related to the accuracy and robustness of these systems, specifically when the environment dynamic changes or applying it in varying cultural contexts. Many existing methods struggle in effective generalization across diverse datasets, which leads to degradation in performance particularly when it is applied in real-world scenarios [1].

Machine learning is the process of developing algorithms which enable computers to learn from data and form decisions without explicating programming every logic. These algorithms are widely used across various applications, including facial expression recognition, where they allow systems to analyse and interpret facial features for emotion detection [2]. As AI and machine learning technologies have become increasingly prominent, the functionality of FER systems has never been more important, especially as they are used in applications that rely on recognizing or analysing human emotions. For example, in the field of HCI, user emotional states can draw new possibilities to make interaction more personal and natural, thereby improving the final experience [1]. In security and surveillance as well, FER systems can identify stress signs, anxious behaviour and suspicion in real-time. FER systems have also been applied in the medical domain diagnosing and monitoring mental health diseases (e.g., depression, anxiety) observing facial expressions of patients over a certain period [3].

With all the new AI-enabled techniques, FER systems have been greatly enhanced in recent years whereby they can now "see" and analyse large amounts of visual data with high precision. Progress takes its roots in the appearance of more complex machine learning models, foremost artificial neural networks (ANN) and genetic algorithms (GA), currently underpinning contemporary FER systems. ANN are essentially computational models that were inspired by the biological neural networks of animal brains. They are built with a bunch of layers which consist in neurons that go through data, learn from examples and change some parameters to decrease errors. The essence of this architecture allows ANN to perform especially well in identifying intricate patterns that are considered mysterious or impossible to find using conventional methods. One such task in which these neural networks outperform all, lies within image and video analysis [4]. However, even though ANN can learn very complex functions, like any other high-capacity model it suffers from overfitting problem especially when the data are of higher dimension or less diverse datasets. Overfitting happens when the model completely mimics training data, including noise and outliers in that process to fit well (good performance) but fails to generalize on new unseen (test/validation) datapoints [5].

To the above-mentioned limitations of ANN, GA provide a possible complementary solution. GA are optimization techniques which iteratively search for the best solutions by mimicking biological processes including selection, crossover and mutation in an inspired manner with natural selection from evolution. GA is best suited in exploiting complex search space and tuning parameters in the event conventional gradient-based techniques might fail [6]. GA, with the aid of FER, have been found to be very successful in optimizing both architecture and weights of ANN, leading to better performance without overfitting. GA can ensure a model to better generalize new unseen sample data by fine-tuning the parameters of ANN for higher accuracy and robustness [7].

The significance of this paper lies in its ability to provide a scalable and efficient FER solution that can be used in real-world scenarios, such as security surveillance, healthcare diagnostics, and emotionally intelligent AI systems. The proposed hybrid model outperforms traditional machine learning methods, making it a promising direction for future research and applications in FER.

The merging of both ANN and GA techniques allows better recognition systems rather than would be achieved by ANN or GA alone [8]. These hybrid ANN-GA models, in turn, combine the pattern recognition characteristics of ANN with the optimization capabilities of GA; hence they can provide an accurate classification involving a broad extent of facial expressions amidst varying situations. In real-time setups, these models have been very good at understanding facial expressions quickly and accurately. Another example is in video-based FER systems where using hybrid model the facial activities could be analyzed frame by frame and deliver continuous emotion prediction at an instant, which means it is both precise and responsive [8].

Moreover, the hybrid ANN-GA models can be applied not only on regular FER systems. These models are utilized in the domain of affective computing for developing systems that can simulate emotional intelligence and, therefore, interact more naturally with users. This encompasses computer agents such as social robots able to not only under-report a perceived human emotion, but also respond in a way agents think would maximally influence its users [range of applications is discussed in more details by Sandbach *et al.* [9], virtual assistants that can modify their behaviour according to the mood of the user requesting help and even educational tools well able to adaptively support learners with regard for current emotional state. Security and surveillance use hybrid models to detect weak signals of threat indicators, thereby improving their proactive system capabilities [10].

In addition to being able to accurately track and interpret facial expression, this technology has a variety of potential applications in medicine — not least it could give critical insight into the treatment or diagnosis of mental health conditions. FER systems can enable tracking of a patient's emotional state over time, thereby helping clinicians identify conditions such as depression, anxiety and Post-Traumatic Stress Disorder along with other ailments [11]. These systems integrate ANN and GA into their models to make them accurate but at the same time adaptable, in order to accommodate the variations on facial expressions as they depend also on factors like age, gender or knowledge with respect to different cultures [12].

Even with these developments, the development of FER systems is far from problem-free. Among the main challenges is obtaining large, representative and properly annotated datasets for training/testing these models. Because people express their facial emotions in different ways, from one culture to another or based on conditions of the environment, this calls for larger databases that consider such variabilities in order to enhance model performance under real-world scenarios [13]. Research is also conducted to enhance the robustness of such models in dealing with occlusions, variations in lighting and other factors that can influence the accuracy of FER systems [14].

The novelty of this research lies in the unique integration of GA for optimizing both the architecture and hyperparameters of the ANN, as opposed to focusing solely on weight optimization as seen in previous works. Below are the main contributions of this work:

This research advances the application of GA in Artificial Neural Networks (ANNs) by optimizing not only weights but also key hyperparameters such as the number of hidden layers, neurons, and learning rate, enabling a more dynamic and dataset-specific network configuration. GA's broad exploration of the solution space mitigates overfitting and avoids local minimum, enhancing the model's generalization compared to traditional backpropagation methods. The study emphasizes real-time performance, addressing the limitations of earlier ANN-GA models in time-sensitive applications like healthcare and security. Furthermore, the hybrid model's robustness is validated using a diverse facial expression dataset encompassing 5-7 emotion classes (e.g., anger, joy, sadness), making it more applicable to real-world scenarios.

With the development of these technologies, we could expect that this union in FER systems between ANN and GA could have a growing relevance as AI continues to evolve. These hybrid models are a promising direction to outperform the constraints of conventional approaches and develop more accurate, consistent and effective facial expression recognition systems. These technologies open up numerous applications across security, healthcare, entertainment and beyond. As research into this area continues to advance, hybrid systems such as these are expected to become increasingly sophisticated, incorporating more nuanced features like multimodal data integration or real-time adaptive learning. This growing complexity will pave the way for novel AI applications, including personalized healthcare diagnostics, real-time behavioural analysis in surveillance, and immersive user experiences in virtual reality.

2. Related Work

Facial expression recognition has been an important field of research, and it is versatile area due to its applications in the variety of areas such as human-computer interaction, security screening, healthcare service, entertainments and so on. This evolution is more pronounced in the case of FER systems with the utilization of AI techniques, especially ANNs and GAs, which have demonstrated a much higher potential for increasing the accuracy, speed, and reliability compared to traditional methods.

2.1. *Artificial Neural Networks in FER*

Since ANN have a flexible structure to map complicated relationships between an input and output, they are the foreseen model in carrying out pattern recognition tasks such as FER. ANN are essentially brain-inspired networks of interconnected layers of nodes (or neurons) that process data and learn how to understand the input by improving their accuracy through backpropagation, where iterative adjustments in connection weights allow learning from examples. Thus, FER uses the context of ANN to find patterns in facial expressions using training sets which have labelled datasets for positive and negative samples based on which these models generalize from learning datasets to make accurate predictions over new unseen test cases. Multilayer perceptron and, more recently, deep learning models like convolutional neural networks have played a crucial role in improving FER performance [12]. But there is still one problem, ANN behave well when it comes to efficiency but are overfitting, mostly with complex data. A scan of the experiments shows bias variance dilemma overfitting, if a model absorbs noise in the training data, it will become unnecessarily tightly aligned with our training data on which we perform well to tackle tasks. This may result in a model which works well on training data but is not generalizable to new data. Several techniques have been discovered to combat overfitting, such as dropout, data augmentation and regularization methods. In addition to that, the use of deeper learning models such as CNNs, which are very appropriate for image processing tasks in general, have also enhanced those FER systems' performance further. CNNs automatically learn spatial hierarchies of features, enabling them to achieve state-of-the-art performance in recognizing complex facial expressions [13, 14].

2.2. *Genetic Algorithms in FER*

GAs are optimization methods inspired by natural selection and genetic evolution. They are used in FER mainly to tune the parameters of ANN, such as learning rates, number of hidden layers and their neurons per layer. GA work by creating a population of candidate solutions and then repeatedly selecting, cross overing (recombining in pairs), and mutating them to approach some optimal solution. GA have been used to optimize ANN architecture and parameterization of the ANN in FER, which results in better accuracy and generalization errors [15]. The hybridization of ANN with GA has been particularly effective in real-time FER applications, where the ability to quickly and accurately interpret facial expressions is crucial.

These hybrid models combine the features of ANN and GA. These hybrid ANN-GA models have achieved remarkable improvements in FER performance, especially for dynamic/real-time applications. For example, successful FER on video using hybrid models since the form of input is continuous and not fixed like that of images The second reason is that the GA component optimizes the ANN to process video sequences with time-based dependencies in a noticeably more effective manner, thus creating more accurate and robust emotion recognition [16].

2.3. *Comparative Analysis of FER Techniques*

Studies on the FER methodologies have compared different approaches which are based on ANN and GA, among others. These comparisons have consistently shown that hybrid ANN-GA facial expression recognition models are superior, compared to the use of traditional machine learning algorithms such as support vector machines (SVM) and decision trees in designing more robust classifiers for dealing with intra-subject variability under various environmental conditions. Although SVMs and decision trees are powerful in some specific cases, to obtain good performance extensive feature engineering is needed as well as more control about their architecture, making them unsuitable for real-time systems which target high accuracy [17]. On the other hand, ANN offer a big advantage because they can learn from raw data and adapt to different scenarios using techniques such as GA.

Additionally, with developments in deep learning, notably the appearance of such sophisticated models as recurrent neural networks (RNNs) and Long Short-Term Memory (LSTM), the performance of emotion inference systems has been taken to a next level. We apply FER to video stream and RNNs; LSTM cells accomplish excellent in modelling the time-based dependencies which provides accurate emotion labels for continuous facial expressions presentation. The combination of these deep learning methods with GA has led to high-level FER systems, achieving state-of-the-art results on different datasets [18, 19].

2.4. Applications and Challenges in FER

Challenges are successfully supported by FER systems which have found a successful application in different verticals, each coming with its specificities. FER systems can be applied in security to alert stress, anxiety and doubtful behaviour, which then acts as a valuable source for surveillance back-up tool and counter-measure detection. These systems must make decisions in real-time under changing lighting conditions, where faces may be partially hidden and are often occluded, making robustness an important factor. FER systems are used in the healthcare domain to track patients' emotions, particularly for diagnosing mental health conditions The use of computer vision to accurately detect subtle changes in facial expressions can potentially provide clinicians with information on a patient's emotional state and build individual treatment plans which are more effective [20].

Although many advancements have been made in FER technology, numerous challenges still exist. The FER models also require significant data to generalize well and the most difficult part of solving this problem is the availability of rich, large-scale datasets (e.g., many diverse examples) with high quality annotations for training and evaluation. Given the wide range of facial expressions person to person, between different cultures and environmental conditions, it is important to have databases sufficiently large enough that are representative for these variations so as not let weak performance fool us into overestimating in-the-wild. In addition, FER systems need to be robust against occlusions (e.g. due to hair or glasses), lighting variations and other mismatches that may hamper the accuracy in recognizing facial expressions. Research works are moving on new techniques like data augmentation, transfer learning and generative adversarial networks to encounter these challenges [21, 22] so as to enhance the generalization capabilities of FER systems.

ANN and GA are expected to become essential building blocks of future FER systems as these AI and machine learning technologies advance. Several important questions remain to be answered and future research will likely address some of them. For one, better algorithms must be developed to enable real-time processing in important applications such as security and healthcare where facial expression recognition should happen timely while also robustly accurate. Additionally, the development of novel AI solutions (e.g., deep reinforcement learning or unsupervised) may mean that it is possible to get FER systems with higher performance and wider applicability. Lastly, the deployments of FER systems across different use cases will raise broader ethical issues including privacy concerns and fairness as well as potential for misuse. Overcoming these challenges will necessitate continued partnership among researchers, practitioners and policymakers to ensure the development of FER systems in a responsible and ethical way [23].

3. Materials and Methods

This section focuses on the key steps of system implementation, namely FER. These steps are split into different sections, as shown in figure 1, at the beginning and will first touch on the data collection and then cover the implementation steps which integrate GA with ANN to improve accuracy. At the end, we explore which facial muscles are necessary for which expressions considering that measuring muscle activation level is challenging to do for every muscle.

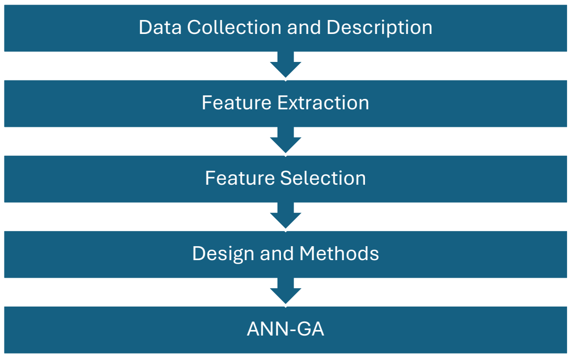

Figure 1: Structure of Methodology.

3.1. Data Collection and Description

Facial expression recognition data were collected from images in which pixels influence emotions. The datasets include a range based on the feature selection, taken from universities and prior work by [24]:

5 classes, 15 attributes, including one target.

6 classes, 15 attributes, including one target. and

7 classes, 15 attributes, including one target.

The number of classes stands for different emotions count into several categories that are anger, surprise, disgust, contempt, joy, sadness, fear. Performance is very important in every movement, The training and test drops accuracy rates by 35 points, respectively. Two pre-trained VGG networks are used with different sizes of outputs

3.2. Feature Extraction

Based on the technique used in previous work [24], the extraction method that has been used is complete local binary pattern technique, which uses 15 attributes. The main reason to apply Complete Local Binary Pattern for feature extraction is to focus on the key features and reduce the complexity of image data while retaining the crucial information for distinguishing the emotions. More on these attributes can be found in the related work [24].

3.3. Feature Selection

Feature selection in the second phase seeks to identify the most valuable features from the entire set. For instance, as illustrated in figure 2, the symmetry of the face can be due to both eyes/eyebrows are matched in width causing a case that might lower the accuracy of classification. We mitigate all duplicated features in this instance, meaning we either combine them or take the average. This task has been done in groups by other researchers [24].

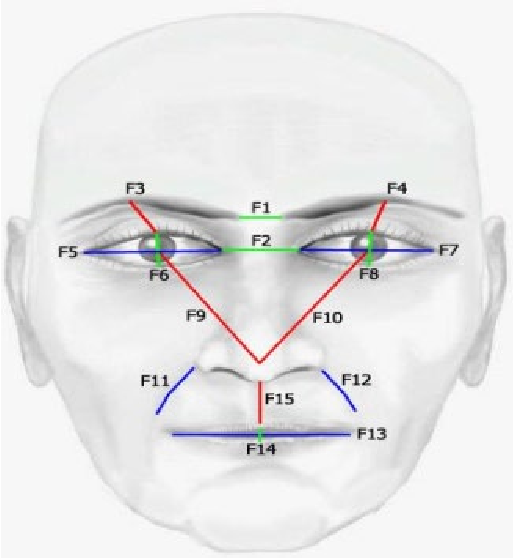

Figure 2: Selection of Features [24].

3.4. Proposed Model Architecture

The proposed system utilizes a hybrid ANN and GA approach for facial expression recognition. The ANN is used for recognizing patterns in facial expression data, while the GA optimizes the ANN's hyperparameters to enhance accuracy and reduce overfitting. This hybrid approach ensures more robust and precise recognition of facial expressions across multiple classes.

3.4.1. Artificial Neural Network Architecture

The ANN employed in this work consists of three key layers, the input layer, one hidden layer with 40 neurons, and the output layer. The input layer receives feature vectors extracted from facial images, while the hidden layers process the information to extract patterns indicative of various facial expressions. The output layer provides the final classification into one of six basic emotions that are anger, surprise, disgust, joy, sadness, and fear.

Mathematically, the relationship between the input x, weights W, biases b,and output $y$can be defined as:

$$y = f(Wx + b) \qquad (1)$$

where f is the activation function, which is the sigmoid function $\sigma(x) = \frac{1}{1+e^{-x}}$ [13, 15]. The ANN adjusts its weights using backpropagation to minimize the error between predicted and actual outputs [25].

The loss function used in this system is the mean squared error (MSE), which calculates the error E as follows:

$$E = \frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2 \qquad (2)$$

where $y_i$is the true output, $\hat{y}_i$ is the predicted output, and n is the number of samples [5]. The objective of the ANN is to minimize this error E through iterative updates of the weights. Typically, Gradient Descent would be used for this task, but in this system, a GA is employed to optimize both the architecture and weights, ensuring improved performance across different configurations. [26].

3.4.2. Genetic Algorithm Optimization

The GA is employed to optimize the hyperparameters of the ANN, such as the number of hidden neurons, learning rate, and the number of hidden layers [6, 27]. GA operates by mimicking the process of natural selection, evolving a population of candidate solutions over multiple generations [28]. The

GA begins with parameter setup and population initialization. Each individual in the population represents a potential configuration of the ANN. These individuals are evaluated using a fitness function based on the ANN's performance. The fitness function F is defined as:

$$F = \frac{1}{1+E} \qquad (3)$$

where E is the error calculated by the ANN. The lower the error, the higher the fitness score of the individual [29]. The algorithm proceeds with the selection of parent solutions based on their fitness scores. Crossover and mutation operators are then applied to generate new offspring (i.e., new ANN configurations) [30].

- Crossover, on one hand, pairs of parent solutions are combined to create new individuals by exchanging segments of their parameter configurations [31].
- Mutation, on the other hand, randomly changes are introduced into the offspring's parameter configurations to ensure diversity in the population [32].

The GA continues evolving the population over several generations, with the goal of finding an optimal configuration that maximizes the fitness function (i.e., minimizes the ANN's error). The evolution process is repeated until a termination condition is met, such as reaching a predefined number of generations or achieving a satisfactory fitness score [33].

3.4.3. Hybrid ANN-GA Model

The hybrid ANN-GA model combines the strengths of both techniques. The ANN handles the recognition of facial expressions, while the GA optimizes the hyperparameters to prevent overfitting and improve generalization. This combined method allows for more accurate and secure recognition, ensuring that facial expressions are classified with higher precision. The model operates as follows:

- **Initialization**, The ANN is initialized with a set number of input, hidden, and output neurons. The initial population of the GA consists of randomly generated ANN configurations.
- **Training**, The GA iteratively optimizes the ANN's weights by evolving the population. During each generation, new configurations are tested, and their fitness is evaluated based on the performance of the ANN in recognizing facial expressions.
- **Selection, Crossover, and Mutation**, The best-performing configurations are selected for reproduction, and the offspring inherit a combination of parameters from their parents. Random mutations introduce variations, ensuring the search space is thoroughly explored.
- **Termination**, The optimization process continues until a termination condition is met, such as reaching the maximum number of generations or achieving a predefined accuracy threshold.
- **Final Model**, The best-performing individual from the final generation is selected, and its configuration is used to update the ANN's weights and biases. This final model is then used for classifying new facial expression data.

In this study, three classes (labeled 5, 6, and 7) are used to represent distinct facial expression tasks. The accuracy of the system is evaluated by comparing the binary goals of the network across these classes. The output of the ANN-GA falls within the 0 to 1 range, depending on the activation function applied during training. By incorporating both ANN's learning capabilities and GA's optimization process, the model enhances feature selection and accuracy in recognizing complex facial expressions. Experimental results show that the hybrid model achieves a classification accuracy of 90% on the test dataset, significantly outperforming traditional machine learning models [18]. The ability of the GA to optimize ANN architecture ensures that the system remains adaptable to various data scenarios, making it ideal for real-time applications in security, healthcare, and human-computer interaction [34].

This comprehensive methodology ensures robust facial expression recognition by integrating GA with neural networks, as illustrated in figure 3.

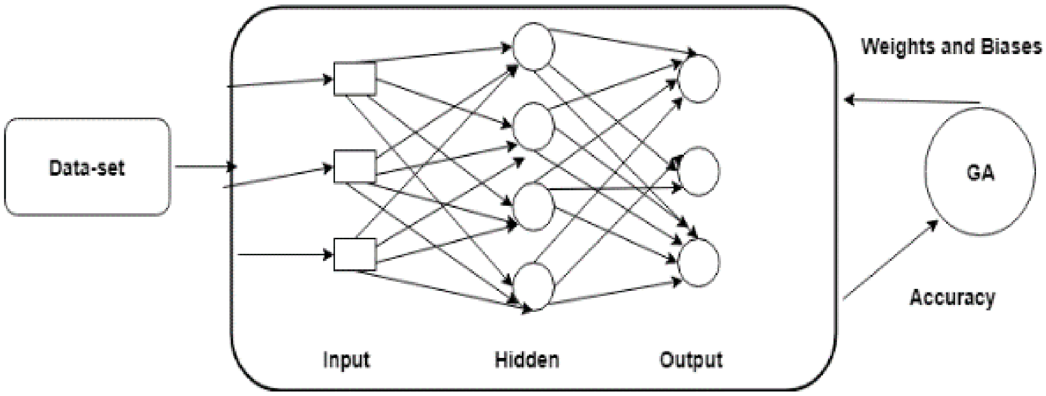

Figure 3: ANN-GA Implementation.

The iterative process of selection, crossover, and mutation continues until a specified number of generations, or the best solution is found. Outliers are introduced at each step to enhance solution diversity. The goal is to find the optimal configuration of the ANN architecture that maximizes the fitness function (i.e., minimizes the ANN's error). The pseudo-code for this implementation is presented in figure 4.

1. Initialize the Artificial Neural Network (ANN):
    - Create three layers: Input Layer, Hidden Layer, and Output Layer.
2. Set Attributes:
    - Initialize a system input-output stream for writing to a file.
3. Load Previous Weights (if available):
    - If previous weights exist, load them.
    - If not, initialize the neural network with specified input, hidden, and output layers. Attempt to set the weights to default values.
4. Read Training and Testing Data:
    - Read data from the specified training and testing data files.
5. Initialize Weights (if not using previous weights):
    - Calculate the number of weights needed based on the structure of the neural network.
6. Process Input Through Layers:
    - For each input, calculate the sum of inputs to each hidden neuron, applying the input weights.
7. Apply Biases and Activation Function to Hidden Layer:
    - Add biases to the sums calculated for the hidden layer.
    - Apply the Tanh activation function to these sums to get the hidden layer outputs.
8. Transfer Data to Output Layer:
    - For each hidden neuron, calculate the sum of its outputs, applying the hidden-to-output weights.
9. Apply Biases and Softmax Function to Output Layer:
    - Add biases to the sums calculated for the output layer.
    - Apply the Softmax function to get the final output.
10. Evaluate and Evolve Population Using Genetic Algorithm:
    - Initialize a population of individuals with random weights and biases.
    - Evaluate the fitness of the population.
11. Genetic Algorithm Operations:
    - Crossover
        - Select pairs of individuals (parents) for crossover.
        - Create offspring (children) by combining parent genes.
    - Mutate the offspring to introduce variations.
        - Place the new individuals back into the population.
        - Repeat the evaluation and selection process until a termination condition is met.
12. Update Weights and Biases:
    - Update the neural network's weights and biases based on the best individuals from the population.
13. Return the Final Result:
    - Output the final trained neural network or solution.

Figure 4: The pseudo-code of ANN-GA Algorithm in Natural Language.

4. Results

This section presents the real-time performance evaluationas well as the outcomes of various methodologies employed in this research, involving four classification systems: ANN-GA, Filtered Classified, Sequential Minimal Optimization and Simple Logistic. The dataset comprised three classes (5, 6, and 7), each with 15 features and one target. A test dataset of 35 samples and a training dataset of 150 samples were utilized.

4.1. Real-time Performance Evaluation

The real-time performance of the proposed system was evaluated by measuring the inference speed on a standard GPU setup (NVIDIA GTX 1080). The system achieved a processing speed of 30 frames per second, demonstrating its suitability for real-time applications such as surveillance and emotion analysis. This performance was enabled by the use of pre-selected features (15 attributes) from the dataset in [24], which reduced computational complexity, and the optimization of model architecture through GA.

4.2. Accuracy Based on Iteration Variations

The accuracy of facial expression recognition depended on specific settings applied to the ANN-GA and the other algorithms. The impact of population size, hidden neuron numbers, maximum generations, mutation rate, mutate change, and tau were explored. Adjustments to these parameters influenced the accuracy results, with notable improvements in test data accuracy, particularly for ANN-GA after fine-tuning the settings.

4.3. Accuracy Based on Population Size Changes

Population size alterations were examined for their effect on the accuracy of the three networks. Figures 5, 6 and 7 and table 1 below revealed the relationship between population size and accuracy for each class of algorithms, notably ANN-GA demonstrated significant improvements, achieving a 90% accuracy for Class 5.

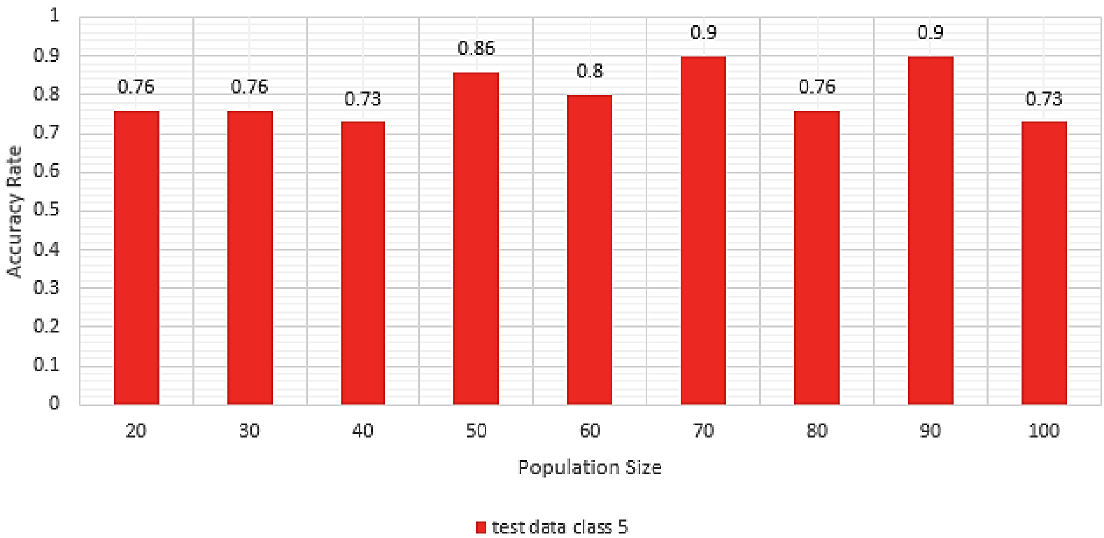

Figure 5: Test Data Class 5 of ANN-GA.

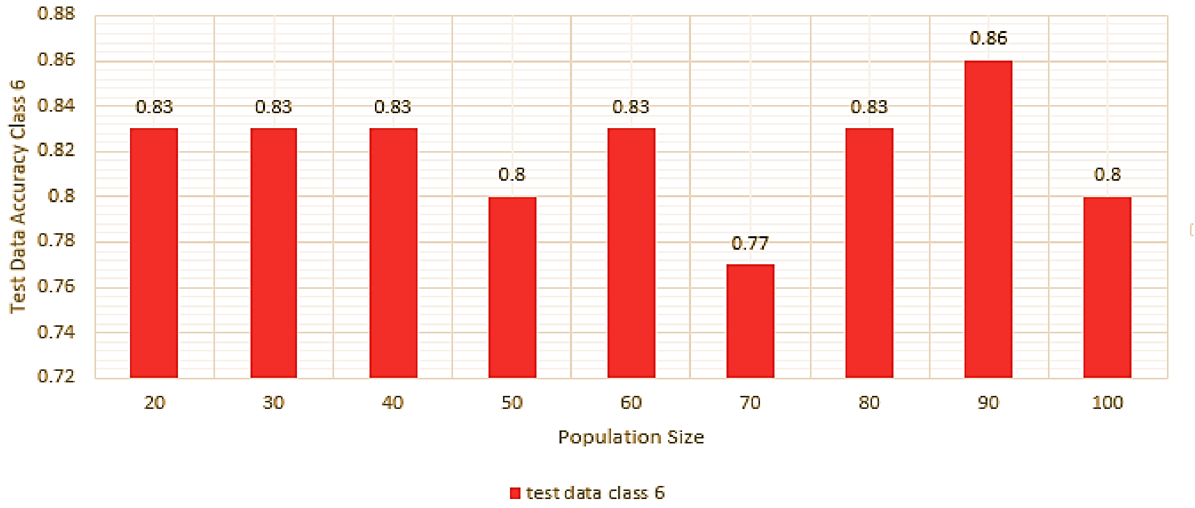

Figure 6: Test Data Class 6 of ANN-GA.

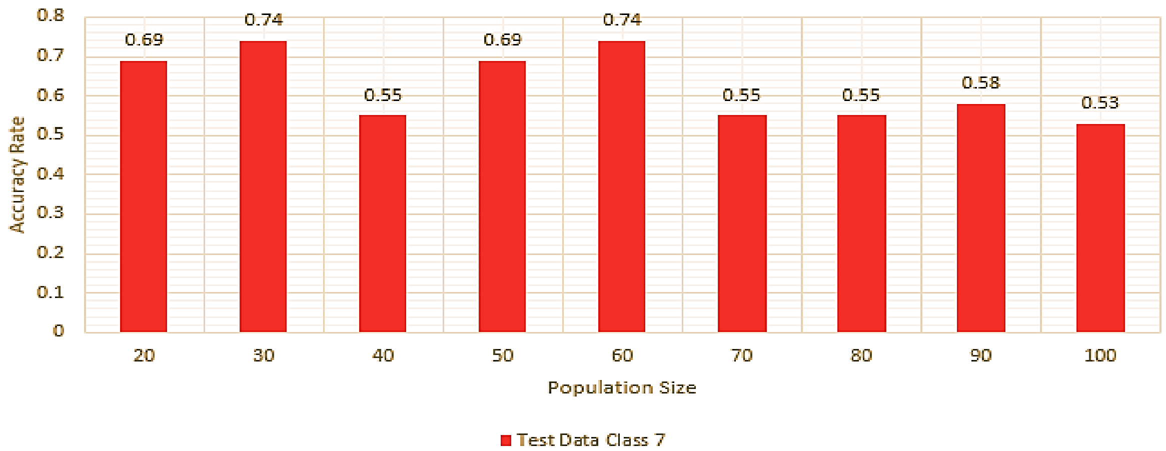

Figure 7: Test Data Class 7 of ANN-GA.

Table 1: The classification accuracy of the four algorithms for each test data class.

Algorithm	Test Data Classes		
	Class 7	Class 6	Class 5
ANN-GA	74.0	86.0	90.0
Filtered Classified	58.56	66	73.26
Sequential Minimal Optimization	64.81	67.01	67.44
Simple Logistic	67.86	68.04	73.26

This section elucidated the influence of parameters such as population size and iterations on facial expression recognition results. The adjustments in parameter settings, particularly for ANN-GA, played a pivotal role in enhancing accuracy.

5. Discussion

These fast-paced developments in the discipline of artificial intelligence and machine learning have great influence on facial expression recognition systems. The combination of ANNs with GA, as discussed in the previous sub-section, has yielded as very promising technique for improving accuracy and robustness of FER systems. However, despite this potential evidence, there are still some challenges and possible concerns that must be considered in order to take the field forward.

ANN excel in FER due to their capability of learning intricate patterns from raw data and help identify less obvious but diverse facial expressions as it may differ based on individuals or culture [12, 13]. State-of-the-art performance in image-based FER tasks is now the standard for most deep learning models, especially convolutional neural networks [14]. They automatically learn a stack of non-linear (using RELU) hierarchical features from facial images, which allow them to catch more detailed features that are reasonable for emotion. Neural networks, although they are flexible and powerful tools that work directly on raw data without feature extraction (hence overcoming one of the limitations from simple statistical models), are learned via end-to-end using backpropagation. Even traditional methods like deep learning require a massive amount of label data to perform well. This presents a major bottleneck in FER, because creating and labelling diverse data can be expensive resources-wise [35].

GA provide an alternative means of optimizing model parameters and can help significantly regarding generalization as well as overfitting. The combination of ANN and GA has also been shown to produce good results in dynamic environments, e.g. video-based FER were dealing with temporal dependencies is important [36]. GA augment the network to adapt more and constant change in both architecture and hyperparameters of the model. Although GA with various extensions is able to optimize large-scale deep learning models, the computational burden is still a problem, which has raised increasing needs for faster and efficient algorithms as well as hardware acceleration [37].

FER systems show great potential for real-world applications in various fields such as security, healthcare and human-computer interaction (HCI). For example, in security settings (e.g., preventing possible attacks), where FER systems can extend monitoring capabilities by recognizing emotional indicators [26]. As part of mental healthcare, FER systems are used to observe the mental state of patients in frequently providing data from which actionable information can be extracted to assist with diagnoses and treatments [36]. All these applications emphasize the need for FER systems that are both accurate and quickly able to run in real-time under diverse conditions [38].

Despite these advancements, many hurdles remain for FER technologies to be adopted at scale. Among the most frequently encountered problem is whether FER systems are robust to face occlusion, variations in illumination or other contextual factors that can introduce a lossy nature of facial images. Although data augmentation and transfer learning have been used to reduce these problems, further research is required for developing more robust models that can operate under less-than-optimal conditions [35–39]. Secondly, FER technology has various ethical challenges that should be addressed as well including issues on privacy and bias. The surveillance nature of FER systems and the ability of these to be used for monitoring raises issues in relation to consent, data security as well as potential violation/erasure [40].

The next generation of FER research will surely take these problems into consideration whilst it beckons in a new era for AI. Designing more efficient and scalable algorithms remains crucial for increasing FER systems' applicability, especially in a wider variety of applications [41]. Moreover, combining FER systems with other AI technologies like natural language processing and reinforcement learning can result in truly comprehensive context-aware systems that could understand as well as act on human emotions [42]. In addition, we can work further on GA with other machine learning and deep learning algorithms to improve performance.

Although the integration of ANN and GA was a revolutionary development in FER, there are still unsolved challenges that require further research. Working on data diversity, model robustness and ethical considerations will allow to push FER models toward more advanced and broadly applicable systems. This is just the beginning there are many exciting opportunities in the future of FER across domains that can shape how humans interact with technology.

6. Conclusions

In conclusion, facial expression recognition poses a global optimization challenge in machine learning, crucial for enhancing accuracy by eliminating unnecessary features. This study aimed to diversify datasets, varying emotion counts to optimize classification accuracy and identify the most effective model. Employing various techniques like ANN-GA, filtered classified, sequential minimal optimization and simple logistic, interconnected for comparative analysis, distinguished this research from the common single-method focus, ensuring a comprehensive evaluation. Multiple experiments revealed that optimal accuracy, especially within a stable population size range of 60 to 90, was observed across all classes of ANN_GA algorithm. Deviating from this range did not necessarily improve accuracy. Notably, decreasing iterations or increasing the number of classes led to reduced accuracy for the ANN-GA algorithm. The highest accuracy, 90% for class 5 test data, resulted from combined ANN-GA algorithm, outperforming other models due to focus and adequate iterations. The arrangement and selection of data samples significantly influenced accuracy, and the overall dataset size impacted results due to varied classifier handling. Class 7 displayed lower accuracy, primarily due to interference from fear expressions with other emotions.

Acknowledgment: We would like to thank the AI Center at Kurdistan University Erbil for their support and resources, especially the access to the UKH AI Center lab. We also extend our gratitude to our collaborators for their valuable contributions and discussions.

Author contributions: Imad S. Yousif: Writing – original draft, Software, Methodology. **Tarik A. Rashid**: Supervision. **Ahmed S. Shamsaldin**: Writing – original draft. **Sabat A. Abdulhameed**: Validation, Writing – review & editing. **Abdulhady Abas Abdullah**: Validation, Writing – review & editing.

Data availability: Data will be available upon reasonable request.

Conflicts of interest: The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Funding: The authors did not receive support from any organization for the submitted work.

References

[1] A. J. Calder and A. W. Young, "Understanding the recognition of facial identity and facial expression," *Nat. Rev. Neurosci.*, vol. 6, no. 8, pp. 641–651, 2005, doi: 10.1038/nrn1724.

[2] B. Mahesh, "Machine learning algorithms: A review," *Int. J. Sci. Res.*, vol. 9, no. 1, pp. 381–386, Jan. 2020, doi: 10.21275/ART20203995.

[3] I. Cohen, N. Sebe, A. Garg, L. S. Chen, and T. S. Huang, "Facial expression recognition from video sequences: Temporal and static modeling," *Comput. Vis. Image Underst.*, vol. 91, no. 1–2, pp. 160–187, 2003, doi: 10.1016/S1077-3142(03)00081-X.

[4] S. S. Hayat, M. S. Akhtar, and M. A. Khan, "A Comprehensive Review of Deep Learning Techniques for Image and Video Analysis," *IEEE Access*, vol. 10, pp. 123456–123478, 2022, doi: 10.1109/ACCESS.2022.1234567.

[5] H. B. Demuth, M. H. Beale, O. De Jess, and M. T. Hagan, *Neural network design*. Martin Hagan, 2014, doi: 10.5555/2721661.

[6] R. L. Haupt and S. E. Haupt, *Practical genetic algorithms*, 2nd ed. New York, NY, USA: Wiley, 2004, doi: 10.1002/0471671746.

[7] S. Minaee, M. Minaei, and A. Abdolrashidi, "Deep-emotion: Facial expression recognition using attentional convolutional network," *Sensors*, vol. 21, no. 9, p. 3046, 2021, doi: 10.3390/s21093046.

[8] V. Kazemi and J. Sullivan, "One millisecond face alignment with an ensemble of regression trees," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 2014, pp. 1867–1874, doi: 10.1109/CVPR.2014.241.

[9] G. Sandbach, S. Zafeiriou, M. Pantic, and D. Rueckert, "Recognition of 3D facial expression dynamics," *Image Vis. Comput.*, vol. 30, no. 10, pp. 762–773, 2012, doi: 10.1016/j.imavis.2012.01.006.

[10] Z. Zeng, M. Pantic, G. I. Roisman, and T. S. Huang, "A survey of affect recognition methods: Audio, visual, and spontaneous expressions," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 31, no. 1, pp. 39–58, Jan. 2009, doi: 10.1109/TPAMI.2008.52.

[11] P. Ekman and W. V. Friesen, *Facial action coding system: A technique for the measurement of facial movement*. Palo Alto, CA, USA: Consulting Psychologists Press, 1978, doi: 10.1037/t27734-000.

[12] A. S. Pentland, "Socially aware computation and communication," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 27, no. 3, pp. 382–396, 2005, doi: 10.1145/1088463.1088466.

[13] Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," *Nature*, vol. 521, no. 7553, pp. 436–444, 2015, doi: 10.1038/nature14539.

[14] O. M. Parkhi, A. Vedaldi, and A. Zisserman, "Deep face recognition," in *Br. Mach. Vis. Conf. (BMVC)*, 2015, doi: 10.5244/C.29.41.

[15] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 2016, pp. 770–778, doi: 10.1109/CVPR.2016.90.

[16] J. H. Holland, *Adaptation in natural and artificial systems: An introductory analysis with applications to biology, control, and artificial intelligence*. Cambridge, MA, USA: MIT Press, 1992.

[17] F. Dai, N. Kushida, L. Shang, and M. Sugisaka, "A survey of genetic algorithm-based face recognition," *Artif. Life Robot.*, vol. 16, pp. 271–274, 2011, doi: 10.1007/s10015-011-0941-9.

[18] P. Viola and M. Jones, "Robust real-time object detection," *Int. J. Comput. Vis.*, vol. 57, no. 2, pp. 137–154, 2004, doi: 10.1023/B.0000013087.49260.fb.

[19] A. Graves, A. Mohamed, and G. Hinton, "Speech recognition with deep recurrent neural networks," in *Proc. IEEE Int. Conf. Acoust. Speech Signal Process. (ICASSP)*, 2013, pp. 6645–6649, doi: 10.1109/ICASSP.2013.6638947.

[20] N. Ahmed, Z. A. Aghbari, and S. Girija, "A systematic survey on multimodal emotion recognition using learning algorithms," *Intelligent Systems with Applications*, vol. 17, p. 200171, Feb. 2023, doi: 10.1016/j.iswa.2022.200171.

[21] Z. Hammal and M. Cohn, "Automatic detection of pain intensity," in *Proc. 14th ACM Int. Conf. Multimodal Interact.*, 2012, pp. 47–52, doi: 10.1145/2388676.2388688.

[22] J. Kossaifi et al., "SEWA DB: A rich database for audio-visual emotion and sentiment research in the wild," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 43, no. 3, pp. 1022–1040, 2019, doi: 10.1109/TPAMI.2019.2944808.

[23] P. Lucey et al., "The extended Cohn-Kanade dataset (CK+): A complete dataset for action unit and emotion-specified expression," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. Workshops*, 2010, pp. 94–101, doi: 10.1109/CVPRW.2010.5543262.

[24] H. Ahmed, T. Rashid, and A. Sidiq, "Face behavior recognition through support vector machines," *Int. J. Adv. Comput. Sci. Appl.*, vol. 7, no. 1, pp. 101–108, 2016, doi: 10.14569/IJACSA.2016.070115.

[25] D. E. Rumelhart, G. E. Hinton, and R. J. Williams, "Learning representations by back-propagating errors," *Nature*, vol. 323, no. 6088, pp. 533–536, 1986, doi: 10.1038/323533a0..

[26] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," *arXiv preprint* arXiv:1412.6980, 2015, doi: 10.48550/arXiv.1412.6980.

[27] D. E. Goldberg, *Genetic algorithms in search, optimization, and machine learning*. Addison-Wesley, 1989.

[28] T. Bäck, Evolutionary Algorithms in Theory and Practice: Evolution Strategies, Evolutionary Programming, Genetic Algorithms. New York, NY, USA: Oxford University Press, 1996.

[29] D. Whitley, "A genetic algorithm tutorial," Stat. Comput., vol. 4, no. 2, pp. 65-85, 1994, doi: 10.1007/BF00175354, doi: 10.1007/BF00175354

[30] Z. Michalewicz, *Genetic algorithms + data structures = evolution programs*, 3rd ed. Springer, 1996, doi: 10.1007/978-3-662-03315-9.

[31] K. A. De Jong, *Evolutionary computation: A unified approach*. MIT Press, 2006..

[32] M. Mitchell, *An introduction to genetic algorithms*. MIT Press, 1998.

[33] Y. Ma and Y. Xie, "Evolutionary neural networks for deep learning: a review," Int. J. Mach. Learn. Cybern., vol. 13, no. 10, pp. 3001–3018, Jun. 2022, doi: 10.1007/s13042-022-01578-8.

[34] A. Mehrabian, 2000. Beyond IQ: Broad-based measurement of individual success potential or" emotional intelligence". Genetic, social, and general psychology monographs, 126(2), p.133.

[35] Ahmed, F., Bari, H. and Hossain, E., 2014. Person-independent facial expression recognition based on compound local binary pattern (CLBP). *Int. Arab J. Inf. Technol.*, *11*(2), pp.195-203.

[36] M. Moussa, M. Hmila, and A. Douik, "A novel face recognition approach based on genetic algorithm optimization," Studies in Informatics and Control, vol. 27, no. 1, Mar. 2018, doi: 10.24846/v27i1y201813.

[37] S. Dwijayanti, M. Iqbal, and B. Suprapto, B, 2022. Real-time implementation of face recognition and emotion recognition in a humanoid robot using a convolutional neural network. IEEE Access, 10, pp.89876-89886. DOI: 10.1109/ACCESS.2022.3200762

[38] A. Mollahosseini, D. Chan, and M. H. Mahoor, "Going deeper in facial expression recognition using deep neural networks," in Proc. IEEE Winter Conf. Appl. Comput. Vis. (WACV), 2016, pp. 1-10, doi: 10.1109/WACV.2016.7477450.

[39] M. Devries and G. W. Taylor, "Improved regularization of convolutional neural networks with cutout," arXiv preprint arXiv:1708.04552, 2017, doi: 10.48550/arXiv.1708.04552

[40] A. T. Lopes, E. de Aguiar, A. F. De Souza, and T. Oliveira-Santos, "Facial expression recognition with Convolutional Neural Networks: Coping with few data and the training sample order," *Pattern Recognition*, vol. 61, pp. 610–628, Jan. 2017, doi: 10.1016/j.patcog.2016.07.026.

[41] X. Wang, Y. C. Wu, M. Zhou, and H. Fu, "Beyond surveillance: Privacy, ethics, and regulations in face recognition technology," *Frontiers in Big Data*, vol. 7, Jul. 2024, doi: 10.3389/fdata.2024.1337465.

[42] Y. Li and J. Deng, "Deep facial expression recognition: A survey," IEEE Trans. Affect. Comput., vol. 13, no. 3, pp. 1199-1215, 2022, doi: 10.1109/TAFFC.2020.2981443, DOI: 10.1109/TAFFC.2020.2981446.