International Journal of
INTELLIGENT SYSTEMS AND APPLICATIONS IN ENGINEERING
ISSN:2147-6799 www.ijisae.org Original Research Paper

Effects of Different Datasets, Models, Face-parts on Accuracy and Performance of Intelligent Facial Expression Recognition Systems

Sharmeen M. Salim Abdullah[1], Subhi R. M. Zeebaree[2], Maiwan B. Abdulrazzaq[3]

Submitted: 11/12/2023 **Revised**: 28/01/2024 **Accepted**: 03/02/2024

Abstract: Facial expression recognition is a crucial area of study in the field of computer vision. Research on nonverbal communication has shown that a significant amount of deliberate information is sent via facial expressions. Facial expression recognition is a crucial field in computer vision that deals with the significant impact of nonverbal communication. Expression recognition has lately been extensively used in the medical and advertising sectors. Difficulties in Facial Emotion Recognition. Facial emotion recognition is a technique that examines facial expressions in static photos and videos to uncover information about an individual's emotional state. The intricacy of facial expressions, the versatile use of the technology in any setting, and the incorporation of emerging technologies like artificial intelligence pose substantial privacy hazards. Facial expressions serve as non-verbal cues, offering indications of human emotions. Deciphering emotional expressions has been a focal point of study in psychology for many years. This study will examine several prior studies that have undertaken comprehensive facial analysis, including both total and partial face recognition, to identify expressions and emotions. The datasets and models used in previous studies, as well as the findings gained, show that employing the whole face yields more accuracy compared to using specific face-parts, which result in lower accuracy ratios. However, emotional identification often does not rely only on the whole face, since it is not always feasible to have the full face available. Contemporary research is now prioritising the identification of facial expressions based on certain facial features. Efficient deep learning algorithms, particularly the CNN algorithm, can do this task.

Keywords: *Facial behavior analysis, Facial expression recognition, Datasets, Tools, Models, Complete facial recognition, incomplete facial recognition.*

1. Introduction

It is possible to effectively communicate sentiments via the expressions that individuals have on their faces since these expressions convey the substance of people's feelings in an instant. Research conducted in the fields of computer vision and artificial intelligence demonstrates that methods specifically designed to differentiate between facial expressions and emotions have a major impact. Identifying emotions based on facial expressions was the motivation for the development of these approaches [1],[2]. In order to recognize emotions, it is preferable to rely on visual inputs rather than making a physical connection. This is the case despite the fact that wearable sensors might be used to do this task. Because of the enhanced relevance and variety that these modalities give, the preference for visual modalities of emotion detection is driven by the fact that these modalities offer[3]. A vast range of facial expressions are shown by people, and the expressions that are displayed are influenced by the core emotions that are being experienced. There is a distinct collection of personality traits and a distinctive distribution scale that distinguishes each face expression from the others. There are references to sources 4 and 5 included in the text that was submitted by the user[4],[5].

of the face Expression recognition, which is often referred to as FER, is an important part of human-computer interaction (HCI). By examining photographs of people's faces, it improves the ability of computers to detect and understand the emotions that people are emotionally experiencing [6]. There is a possibility that the rise in popularity of artificial intelligence systems may be attributed to the progress that has been made in computer technology over the course of the many years that have passed. When it comes to real-time applications that include human-computer interaction, the detection of emotions via the study of facial expressions is of the highest relevance. apps that make use of artificial intelligence systems are one example of this kind of application [7],[8]. Humanoid robot apps are another possibility. The development of the validity and widespread implementation of FER algorithms in sectors such as healthcare, security, safe driving, video games, and other disciplines has been considerably supported by the applications that recognize emotions. It has been shown that the use of these approaches is crucial in the

[1] *Information Technology Dept., Technical College of Informatics-Akre, Akre University for Applied Sciences, Duhok, Iraq, sharmeenbarwary2020@gmail.com*
[2] *Energy Eng. Dept., Technical College of Engineering, Duhok Polytechnic University, Duhok, Iraq, subhi.rafeeq@dpu.edu.krd*
[3] *Computer Science Dept., Faculty of Science, University of Zakho, Duhok, Iraq, maiwan.abdulrazzaq@uoz.edu.krd*

interaction between humans and computers, as it results in intelligent outcomes in a range of disciplines [9],[10],[11],[12].

There are a variety of technologies that have been presented with the intention of automating the process of recognizing facial expressions in order to make the implementation of these applications easier. These tactics yield remarkable results, especially in circumstances where the surroundings are regulated in a controlled way [13]. The conventional methods of FER created a number of difficulties when it comes to dealing with the complex and changeable characteristics of facial expressions of emotion. Deep learning, and more especially Convolutional Neural Networks (CNNs), has brought about a change in the area by making it feasible for automated algorithms to swiftly acquire precise patterns from images of faces [14]. This has resulted in a revolution in the field. CNNs have recently shown great success in face identification tasks, especially when they are provided with photos that are facing ahead and do not contain any obstacles [15],[16],[17]. CNNs have also demonstrated exceptional performance in other face recognition tasks. On the other side, the work gets more difficult when faces are hidden, either by angles that mask facial characteristics or by other aspects. This may be done in combination with other factors. Having the ability to detect and interpret faces that present a vast range of postures, angles, sizes, and facial emotions is an important talent to possess. Nevertheless, the topic of establishing precise automated FER continues to provide a considerable barrier in conditions that are not regulated [18],[19]. However, in usual circumstances, the face is often obscured by hands or other accessories, such as masks, scarves, hats, glasses, and other similar items. This makes the process of identification more challenging. Masks, scarves, hats, glasses, and related items are examples of such accessories. When it comes to dealing with these occlusions, the research literature has proposed a few distinct solutions [20],[21]. These strategies are described in more detail below.

Human emotions are often expressed not just through words, but also through subtle changes in our facial expressions. A raised eyebrow, a tightened jaw, or a fleeting smile – these tiny movements can speak volumes about how we feel. FER steps into this fascinating realm, using the power of computer vision and artificial intelligence to decipher these emotional cues from images and videos. Imagine a world where: Computers can understand our emotions from our faces, enabling more natural and nuanced human-computer interaction[22],[23]. Educational platforms can adapt to students' emotional states, tailoring learning experiences to boost engagement and understanding. Autonomous vehicles can react to pedestrians' expressions, enhancing safety and preventing accidents. Market research tools can analyze customer reactions to products and advertisements, providing valuable insights for better marketing strategies[19],[22],[24].

These are just a glimpse of the vast potential of FER. This technology is rapidly evolving, driven by advancements in computer vision, deep learning, and the availability of increasingly sophisticated datasets and tools. The foundation of FER rests on two main aspects: datasets and tools. Datasets, such as FER2013, AffectNet, and JAFET, play a crucial role in teaching our systems about facial expressions. For instance, FER2013 provides grayscale images labeled with seven basic emotions, while AffectNet offers over a million images covering a broader range of emotions and includes details like gender, age, and facial muscle movements. JAFET focuses on high-quality images with subtle expressions, aiding in training models to detect nuanced emotions. Each dataset has its unique strengths and challenges, influencing the capabilities of the models trained on them. These datasets form the basis for developing effective FER tools, enhancing our understanding of facial expressions[25].

To analyze facial expressions effectively, powerful tools and models are essential. Tools like OpenCV provide built-in facial feature detection, while Py-Feat offers a Python toolbox for facial expression analysis. Dlib is crucial for detecting facial landmarks and predicting shape changes. These tools empower developers and researchers in the field of FER. Deep learning models, such as CNNs, excel with large datasets like AffectNet, while Support Vector Machines (SVMs) offer an efficient alternative for smaller datasets. Ensemble methods, combining different models, can enhance performance. However, challenges like intra-class variability and inter-class similarity require continuous research. Context, including non-facial cues like body language, is vital for understanding emotions. Despite challenges, FER holds promise for improving human-computer interaction and understanding human emotions as tools, datasets, and models continue to advance [25],[17].

2. Background Theory

The human face, a canvas of intricate movements and subtle changes, speaks volumes beyond spoken words. Deciphering this unspoken language through facial expression recognition (FER) has captivated researchers for decades, leading to the development of intelligent systems that analyze facial features and infer emotional states. But the accuracy and performance of these systems are not monolithic; they are intricately interwoven with the interplay of three key factors: datasets, models, and face-parts [10]. Imagine a child attempting to identify animals based solely on pictures of puppies and kittens. Their understanding, while initially skewed, will broaden and refine with exposure to a wider range of creatures. Similarly, the size and diversity of the dataset used to train an FER system directly impact its performance. Size matters: Larger datasets expose the model to a greater variety of facial expressions, increasing its ability to generalize and adapt to previously unseen situations. An abundance of data allows the model to learn finer nuances within and between expressions, leading to more accurate recognition. However, simply throwing data at the problem isn't enough [4].

Diversity is key: A truly robust FER system should be trained on a dataset that reflects the rich tapestry of human faces. This goes beyond ensuring gender and ethnic balance; it necessitates accounting for age, cultural variations in expression, and even environmental factors like lighting and occlusion. Biases can lurk within homogenous datasets, leading to misinterpretations and inaccuracies, particularly for marginalized groups. Facing the bias monster: Facial expressions are not universal. Cultural expressions, subtle micro-expressions, and even individual differences in muscle movement can lead to misinterpretations if the training data is biased towards a specific group or demographic [19]. Addressing bias requires conscious data curation, incorporating diverse datasets, and employing techniques like fairness-aware algorithms to mitigate imbalances. With the rise of deep learning, Convolutional Neural Networks (CNNs) have become the dominant force in FER. Their ability to automatically extract features from large datasets has led to significant performance improvements. However, CNNs are data hungry and computationally expensive, limiting their applicability in resource-constrained settings [15]. Beyond the convolutional crown: Traditional machine learning models like Support Vector Machines (SVMs) and Random Forests offer advantages in terms of interpretability and efficiency. They excel in smaller datasets and provide insights into the model's decision-making process. Hybrid approaches that combine deep learning models with traditional methods are also gaining traction, leveraging the strengths of each to achieve optimal performance [21].

Emerging possibilities: The future of FER holds exciting possibilities beyond established models. Recurrent Neural Networks (RNNs) show promise in understanding the temporal dynamics of facial expressions, while generative models could create synthetic datasets to address data scarcity issues. Exploration of these frontiers can further enhance the accuracy and robustness of FER systems. Not all parts of the face are created equal when it comes to expressing emotions. Eyebrows, eyes, and the mouth form a potent trio, often conveying the majority of emotional information. However, neglecting other subtle cues like forehead wrinkles, dimple formation, and even neck tension can lead to inaccurate interpretations [9]. The selective gaze: Focusing on key facial features can improve performance, particularly in resource-constrained situations. Feature selection techniques can identify the most informative regions of the face, reducing computational load and improving efficiency. However, over-reliance on specific features can make the system vulnerable to occlusions or variations in individual facial anatomy. Holistic harmony: The ideal approach balances feature selection with holistic analysis. By extracting information from key regions while maintaining awareness of the entire facial landscape, FER systems can achieve optimal accuracy and robustness. The interplay of individual features and their relationship to the whole face provides a richer understanding of the emotional state being conveyed [16].

Optimizing the accuracy and performance of intelligent FER systems demands a nuanced understanding of the complex interplay between datasets, models, and facial features. A diverse, unbiased dataset fuels the learning process, while carefully chosen models ensure efficient and accurate analysis. Focusing on key facial features, while maintaining holistic awareness, unlocks the full potential of the facial landscape. By meticulously orchestrating these factors, we can develop FER systems that truly hear the whispers of the human face, paving the way for richer interactions and deeper understanding between humans and machines [1]. Facial expressions are intricate tapestries woven from a multitude of threads – the arch of an eyebrow, the crinkling of an eye, the subtle quiver of a lip. Capturing these nuances requires technologies that go beyond simply perceiving pixels. Enter the realm of deep learning, where algorithms akin to the human brain learn to extract and interpret these minute details. Convolutional Neural Networks (CNNs) reign supreme in this domain, excelling at uncovering spatial patterns and extracting features from images. Imagine them as meticulous detectives, sifting through pixelated landscapes to discern the raised ridge of a furrowed brow or the telltale drop of a sad mouth. By meticulously analyzing these features, CNNs learn to associate them with specific emotions, paving the way for accurate recognition [8].

But the story doesn't end there. Facial expressions rarely exist in isolation; they unfold in dynamic sequences, conveying emotions with a nuanced interplay of movement and context. This is where Recurrent Neural Networks (RNNs), particularly Long Short-Term Memory (LSTM) networks, step onto the stage. Imagine them as skilled storytellers, piecing together the temporal threads of an expression, capturing the subtle shifts and transitions that reveal the true emotional narrative. By analyzing sequences of images or even video streams, LSTMs can decode the temporal dynamics of facial expressions, adding a layer of depth and understanding to FER systems. While these technologies hold immense promise, their practical application hinges on efficiency [3]. Deploying complex algorithms on resource-constrained devices like smartphones or embedded systems demands a delicate balance between accuracy and computational cost. Model pruning emerges as a powerful tool in this quest. Imagine meticulously trimming away unnecessary branches from the neural network, streamlining its architecture without compromising its ability to recognize emotions. By removing redundant connections and neurons, model pruning reduces the computational footprint of the system, enabling its deployment on devices with limited processing power [7].

Knowledge distillation offers another intriguing pathway. Picture a seasoned teacher imparting their wisdom to a eager student. In this analogy, the teacher is a large, pre-trained model, and the student is a smaller, more efficient model. Knowledge distillation involves transferring the accumulated knowledge of the larger model to the smaller one, enabling the latter to achieve comparable accuracy with a significantly lower computational burden. True efficiency often lies in the harmonious blending of technologies [11]. Ensemble learning leverages the complementary strengths of different models, creating a robust and accurate system that surpasses the capabilities of its individual components. Imagine a team of detectives, each with their own expertise, pooling their insights to crack a complex case. Similarly, combining diverse models, like CNNs and LSTMs, can lead to superior FER performance, capturing both the spatial and temporal nuances of facial expressions. Facial expressions rarely exist in a vacuum. They are woven into the fabric of our interactions, colored by the context of the situation, the tone of the conversation, and even the cultural background of the individual. Integrating contextual cues into FER systems adds another layer of sophistication, leading to more accurate and nuanced interpretations[14].

Imagine a system that not only recognizes a surprised expression but also understands that it might be caused by an unexpected loud noise rather than genuine shock. By incorporating information like speech content, gestures, and environmental factors, FER systems can move beyond simply labeling emotions to truly comprehending the underlying intentions and motivations behind them. As FER technology advances, its ethical implications demand careful consideration. Issues of bias, privacy, and misinterpretation loom large, requiring thoughtful solutions. Bias in training data can lead to discriminatory outcomes, while privacy concerns arise when deploying FER systems in public spaces [18]. Additionally, misinterpreting subtle expressions can have real-world consequences, particularly in sensitive applications like healthcare or education. Addressing these ethical concerns requires a multi-pronged approach. Utilizing diverse and representative datasets, implementing robust privacy-preserving mechanisms, and developing transparent and accountable systems are crucial steps towards ensuring the responsible and ethical development and deployment of FER technologies [20].

2.1 Artificial Intelligence

The term "artificial intelligence" refers to the process of reproduction of human brain processes, such as learning and discerning, as well as the imitation of mental abilities and conduct. This is what is meant when people talk about "artificial intelligence." In specifically, it involves the development of robots that are able to acquire knowledge, comprehend information, and draw conclusions, so imitating the patterns of thinking and behavior that are associated with human beings. [26],[27]. Applications that make use of artificial intelligence may be found in a wide variety of fields. There are a variety of applications that fall under this category, including medical applications, robotics, simulations (similar to video games), and the resolution of challenges that are associated with discriminating[28]. One subfield of artificial intelligence is known as machine learning, and its primary objective is to develop intelligent systems that are capable of acquiring knowledge and improving their performance depending on the information that they are given. The creation of intelligent systems is the primary focus of the field of study known as machine learning. The use of machine learning is now being utilized in a broad range of domains, including but not limited to the following: cybersecurity, discrimination, classification, detection, identification, commerce, and social networking [29],[30]. It is recommended that one start with the core notions that Professor Tom Mitchell articulated in 1998 in order to produce an explanation of machine learning that is both accurate and succinct. The development of a comprehensive and precise definition of machine learning will be made possible as a result of this. The computer code is intended to acquire knowledge via experience and to enhance its performance on a job as it acquires more experience [31]. The purpose of the computer code is to accomplish this.

Artificial intelligence (AI) is a vast and complex field, encompassing a wide range of technologies and approaches designed to mimic or surpass human intelligence. It's an area of rapid advancement, constantly pushing the boundaries of what's possible[31],[28]. To give you sufficient details about AI, let's break it down into key points:

What is AI?

At its core, AI is the ability of a machine to learn and perform tasks that typically require human intelligence, such as problem-solving, decision-making, and pattern recognition. This can be achieved through various techniques, including[32]:

- Machine learning: Algorithms that learn from data without being explicitly programmed.
- Deep learning: A subset of machine learning inspired by the structure and function of the brain, using artificial neural networks.
- Natural language processing (NLP): Understanding and generating human language.
- Computer vision: Extracting information from images and videos.
- Robotics: Designing and building intelligent machines that can interact with the physical world.

Types of AI:

There are different ways to categorize AI, but some common types include [33]:

- Artificial narrow intelligence (ANI): Also known as weak AI, this refers to systems that are good at one specific task, like playing chess or recognizing faces. Most AI applications today fall under this category.
- Artificial narrow intelligence (ANI)
- Artificial general intelligence (AGI): Sometimes called strong AI, this is a hypothetical type of AI that would be able to perform any intellectual task that a human can, and potentially even surpass human intelligence. We are still far from achieving AGI.
- Artificial general intelligence (AGI)
- Artificial super intelligence (ASI): An even more speculative concept, ASI refers to an AI that would be significantly more intelligent than any human. Some experts believe that ASI could pose an existential threat to humanity, while others believe it's unlikely to be achieved or could be beneficial.

Artificial super intelligence (ASI)

Applications of AI:

AI is already having a major impact on many aspects of our lives, from healthcare and finance to transportation and entertainment. Here are just a few examples [34]:

- Healthcare: AI is being used to develop new drugs, diagnose diseases, and personalize treatment plans.
- Finance: AI is used to detect fraud, manage risk, and make investment decisions.
- Transportation: AI is powering self-driving cars and drones, and optimizing traffic flow.
- Entertainment: AI is creating personalized recommendations for music, movies, and TV shows, and even writing its own creative content.

Challenges of AI:

Despite its immense potential, AI also raises a number of challenges that need to be addressed. These include [33]:

- Bias and fairness: AI algorithms can be biased based on the data they are trained on, which can lead to discriminatory outcomes.
- Job displacement: As AI becomes more sophisticated, it could automate many jobs currently performed by humans.
- Safety and security: AI systems can be vulnerable to hacking and manipulation, which could have serious consequences.
- Ethical considerations: The development and use of AI raise a number of ethical questions, such as who is responsible for the actions of AI systems, and how do we ensure that AI is used for good.

The future of AI [23]:

The future of AI is full of both possibilities and challenges. As the field continues to advance, it's important to have a thoughtful and informed discussion about how we want to develop and use this powerful technology.

2.2 Machin learning

Machine learning (ML) is a subfield of AI that empowers computers to learn from data without being explicitly programmed. Imagine sifting through mountains of information and automatically detecting patterns, drawing connections, and predicting outcomes – that's the essence of machine learning. Let's delve deeper into its key aspects[35], [36]:

1. The Learning Process:

- Algorithms: ML thrives on algorithms, mathematical recipes that learn from data, identify patterns, and make predictions. Popular algorithms include:
  - Linear Regression: Models linear relationships between variables.
  - Decision Trees: Classify data based on a series of branching rules.
  - K-Nearest Neighbors: Predicts an instance's class based on its closest neighbors in the data.
  - Support Vector Machines: Find optimal hyperplanes to separate data points into different classes.
  - Artificial Neural Networks: Mimic the human brain with interconnected layers to learn complex relationships.
- Data: The fuel for ML algorithms. The quantity, quality, and relevance of data significantly impact the learning process and model performance.
- Training: The phase where algorithms analyze data, internalize patterns, and refine their predictive abilities.
- Validation and Testing: Evaluating the trained model's performance on unseen data to assess itsgeneralizability and avoid overfitting, memorizing the training data without true understanding.

2. Types of Machine Learning [37]:

- Supervised Learning: Labeled data, with each data point associated with a known class or value, guides the algorithm to learn the mapping between features and the target variable. Used for tasks like classification (spam detection) and regression (predicting house prices).
- Unsupervised Learning: Deals with unlabeled data, where the algorithm must discover hidden patterns and structures on its own. Used for tasks like clustering (grouping similar data points) and dimensionality reduction (compressing data without losing essential information).
- Reinforcement Learning: An agent interacts with an environment, receiving rewards for desirable actions and penalties for undesirable ones, iteratively improving its behavior through trial and error. Widely used in robotics and game playing.

3. Applications of Machine Learning[38], [39]:

- Healthcare: Predicting disease outbreaks, analyzing medical images, and personalizing patient care.
- Finance: Detecting fraud, optimizing investment strategies, and assessing credit risk.
- Retail: Recommending products, analyzing customer behavior, and optimizing pricing strategies.
- Transportation: Predicting traffic flow, optimizing delivery routes, and developing self-driving cars.
- Entertainment: Personalizing content recommendations, creating realistic special effects, and generating creative content.

4. Challenges and Ethics [39]:

- Bias and Fairness: Algorithms can inherit biases from the data they are trained on, leading to discriminatory outcomes. Careful data selection and mitigation techniques are crucial.
- Explainability and Transparency: Complex models can be opaque, making it difficult to understand their decision-making process. This raises concerns about accountability and fairness.
- Security and Privacy: ML systems can be vulnerable to attacks, and the use of personal data raises privacy concerns. Robust security measures and ethical data handling are essential.

5. The Future of Machine Learning [40]:

As research advances, it can be expected that:

- More powerful and sophisticated algorithms: Deep learning and other techniques will continue to evolve, enabling machines to tackle even more complex tasks.
- Increased automation and personalization: ML will automate routine tasks and personalize experiences across various domains.
- Human-machine collaboration: Humans and machines will work together, leveraging each other's strengths to solve complex problems.

2.3 Machin learning Methods

a. Supervised learning

Supervised machine learning algorithms are like having a supervisor from whom the algorithm learns, the inputs to the algorithm are examples of the required inputs and outputs. In this process, it is essentially training the model on some data (training data) which is named (data labeled) correctly. Next, some new Datasets are given model (test data), expecting the correct result to be generated based on its previous analysis for the data entered [28],[41]. Once the model has undergone this training regimen, it is put to the test using new, unseen datasets known as the test data. The expectation is that the model, having learned from the labeled training data, will now be able to generate accurate predictions or classifications for the new inputs. The performance of the model is then evaluated based on how well it aligns with the correct outcomes for the test data. Supervised learning, therefore, hinges on the availability of labeled training data to guide the algorithm's learning process. The success of the model in making accurate predictions on new, unseen data is a testament to its ability to generalize patterns learned during training. This paradigm is widely employed in various applications, ranging from image and speech recognition to natural language processing, where the algorithm's ability to understand and replicate patterns is crucial for its effectiveness[42],[43].

b. Unsupervised Learning

Unsupervised learning is a kind of learning that does not depend on a supervised model. This type of learning is also known as "learning without supervision" as it does not need an output matrix in the same manner as supervised learning does is implied by the fact that this is the case [44]. When dealing with data for which there is no previous knowledge concerning the result that is desired to be reached, this approach is used as opposed to other methods that are utilized. An examination of the data is performed by the network, and a function is constructed in order to determine the level of accuracy that is present. After then, it makes an effort to reduce the amount of error as much as it can in order to obtain the best degree of accuracy that is achievable. Data that has not been labelled in any manner is the major focus of unsupervised learning, which is a kind of machine learning. Additionally, it allows algorithms to categories, label, and cluster data inside datasets on their own, without the need for any external supervision or influence in the process of carrying out this operation. This is because it eliminates the requirement for any external supervision or influence. This is made possible by the fact that it enables algorithms to do all of these tasks on their own, which makes it possible for this to be successful [45].

c. Semi-supervised learning

An innovative approach to machine learning, known as semi-supervised machine learning, is a technique that blends supervised and unsupervised learning into a single learning process. Using a little amount of data that has been annotated in conjunction with a significant amount of data that has not been annotated is the method that is used in this methodology. The technique is able to make use of the benefits that are associated with both supervised and unsupervised learning as a result of this. As a result of this, it is very probable that we will be able to train the model to classify data without the need of a significant quantity of data that has been categorized for the purpose of training [38].

d. Reinforcement Learning

It is a type of machine learning technique where we have an agent that we want to train over a period of time to interact with a particular environment and improve its performance in relation to the type of actions it performs on the environment within this period, that is, enabling the agent to learn in an interactive environment by trial and error using observations from its own procedures and experiments [35].

2.4 Deep Learning

Deep learning is a major field within the field of machine learning that has garnered a significant amount of interest over the course of the last few years [46]. Deep learning's major goal is to bridge the gap that exists between the area of machine learning research and the ultimate goal of building intelligent machines. This gap has been identified as a significant obstacle in the field. This goal is one of the most essential motivating factors for the development of deep learning software[43],[30]. The deep learning-based approach tries to develop an artificial intelligence model by representing data as a hierarchical structure of ideas. This is necessary in order to accomplish the goal. In its most basic form, this is the objective of the technique. Deep learning is the method that should be used in order to accomplish this. In the process of constructing this system, each layer is constructed by building upon thoughts from layers that came before it that were of a lower level of sophistication. This procedure is carried out several times till the construction is finished[47],[30]. The class structure is an essential component that must be included into each and every deep learning strategy that need to be taken into account. Note Figure (1) that illustrates the relationship between intelligence Artificial, machine learning, and deep learning.

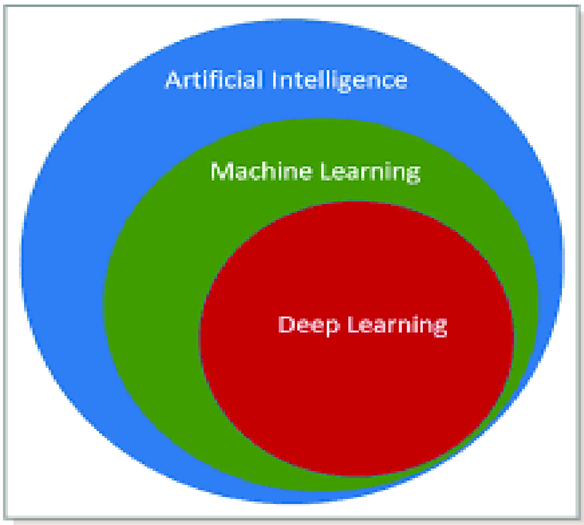

Fig (1): The relationship between artificial intelligence, machine learning and deep learning [46].

3. Literature Review

3.1 Trends toward the Complete FER

In 2023, Kenan et al.[48] suggested a tripartite framework for discerning emotions from face photos. The first phase involves training the CNN-based network using the FER+ dataset. The Binary Particle Swarm Optimization (BPSO) technique is used to choose features from the feature vector in the fully connected layer of the CNN network that has been trained in the second stage. The Support Vector Machines (SVM) algorithm is used to classify the selected characteristics. The suggested system comprises CNN, BPSO, and SVM. Using their CNN model, they retrieved 128 features that reflect the pictures from the FER+ dataset. SVM was used to do out classification utilizing these characteristics. The SVM model outperforms the CNN model in terms of accuracy. In order to enhance the precision, the BPSO heuristic optimization technique is used. The BPSO algorithm was used to choose the features that had the highest classification accuracy from a pool of 128 features that represented the photos. By the conclusion of the BPSO process, the number of features had been effectively halved. The SVM approach was used to reclassify it using the characteristics gathered from BPSO. The combination of BPSO + SVM resulted in an accuracy of 85.74% and improved speed values compared to SVM alone for classification.

In 2023, Lee and Yoo [49] Proposed a novel learning approach using the divide-and-conquer strategy to improve the efficiency of Facial Expression Recognition (FER). This approach aims to minimise the distance between samples belonging to the same class while increasing the distance between samples from other classes. The MobileNet technique was used to detect the facial area in an image, while a ResNet-18 model was utilised as the underlying deep neural network for precise identification of face emotions. Subsequently, groups exhibiting similar facial expressions were categorised by analysing the confusion matrix, which presents the results of the trained ResNet-18 model's inferences. These cohorts exhibiting comparable facial expressions were later used to educate the deep learning model. The evaluation included the application of the proposed methodology to two thermal datasets (Tufts and RWTH) and two RGB datasets (RAF and FER2013). The results demonstrate improved facial expression recognition (FER) ability, with an accuracy rate of 97.75% for Tufts, 86.11% for RWTH, 90.81% for RAF, and 77.83% for FER2013. The proposed method has the capacity to accurately classify large amounts of facial expression data.

In 2023, Sudheer Babu et al.[50] Presented EfficientNet-XGBoost, a novel methodology that employs the Transfer Learning (TL) mechanism. The system's originality is shown by the use of a distinct combination of pre-trained EfficientNet architecture, fully connected layers, an XGBoost classifier, and tailored parameter fine-tuning. To accelerate the learning process of the network and tackle the problem of the vanishing gradient, their model included fully connected layers with global average pooling, dropout, and dense operations. EfficientNet is enhanced by replacing the denser layer(s) and using the XGBoost classifier, hence customising it for FER (Facial Expression Recognition). Feature map visualisation is performed to illustrate the reduction in the dimensionality of feature vectors. The proposed methodology has undergone thorough validation on established benchmark datasets such as Cohn-Kanade (CK+), Karolinska Directed Emotional Faces (KDEF), Japanese Female Facial Expression (JAFFE), and FER2013. To tackle the issue of data imbalance, the CK+ and FER2013 datasets used artificial data augmentation techniques, including geometric alteration approaches.

The proposed methodology is implemented individually on each of these datasets, and the corresponding outcomes are recorded for performance assessment. The performance is assessed based on many factors, including accuracy, recall, and F1 measure. Performing comparison analysis involves independently applying competent strategies to the same sample datasets. The recommended strategy outperforms the others in terms of accuracy, achieving rates of 100%, 98%, and 98% for the first three datasets, regardless of their features. Nevertheless, the efficacy of the proposed research is less encouraging when applied to the FER2013 datasets.

In 2023, Winyangkun et al. [51] suggested a method to enhance the working environment, primarily influenced by the advancing science of deep learning, specifically including CNN and FER. The FER system is used to categorize seven distinct emotions seen on human faces. In order to enhance efficiency, they also used other crucial approaches such as histogram equalization and background removal in the classification process. Their suggested model achieved an average accuracy of 97 percent in recognizing seven different classes. Table (1) illustrates a comparison of prior trends in comprehensive FER research.

3.2 Trends toward the Incomplete IFER

In 2020, Bita and Naimul Mefraz [52] addressing the problem of facial expression recognition when using a head-mounted Virtual Reality (VR) headset is a difficult undertaking since the top half of the face is totally hidden. Their attention was directed towards FER while dealing with a significant blockage, while the user was using a head-mounted display in a VR environment. The researchers suggested a geometric framework to replicate the blocking of visual perception caused by a Samsung Gear VR headset. This framework may be used with current facial expression recognition datasets. Subsequently, a transfer learning methodology is used, commencing with the utilization of two pre-trained networks, namely VGG and ResNet. They refine the networks even further by optimizing them using the FER+ and RAF-DB datasets. Their methodology yielded similar outcomes to current approaches when trained on three modified benchmark datasets that simulate realistic occlusion caused by using a common VR headset.

In 2020, Palaniswamy and Soumya [53] the objective of their study is to categorize five fundamental human emotions based on partly obstructed facial photographs. The absence of fundamental datasets including obstructed pictures poses a challenge in constructing a model that can effectively generalize. Meta-learning algorithms provide a resolution to this issue. In their study, they conducted research on using meta-learning principles in conjunction with prototypical networks to categorize emotions in a few-shot scenario. For training purposes, a set of 50 occluded picture samples is used, with 10 examples allocated to each emotion category. We used the Carnegie Mellon University Multiple Pose and Illumination Face (CMU Multi-PIE) database, the AffectNet database, and internet-sourced photos for the purposes of training and testing. The approach is called MERO, which stands for meta-learning for emotion recognition under occlusion.

In 2020, Ding et al. [54] suggested the use of a landmark-guided attention branch to identify and eliminate distorted elements from obscured areas, ensuring that they are not utilized for recognition purposes. A preliminary attention map is first created to identify if a certain facial feature is obstructed and direct our model to focus on unobstructed areas. In order to enhance resilience, they suggested including a facial area branch that would divide the feature maps into distinct face blocks, assigning each block the responsibility of independently predicting the expression. This leads to a wider range of distinct and selective characteristics, allowing the expression recognition system to regain functionality even when the face is partly obstructed. The occlusion adaptive deep network demonstrates superior performance compared to state-of-the-art approaches on two tough benchmark datasets in natural settings and three real-world occluded expression datasets, due to the synergistic effects of the two branches.

In 2021, Yang et al. [55] suggested a dual-phase attention model to enhance the precision of face-mask-aware FER. During stage 1, the training process focuses on the masked/unmasked binary deep classifier. This classifier is capable of producing attention heatmaps that can roughly differentiate between the face regions that are covered by a mask and the parts that are not occluded. During stage 2, the FER classifier undergoes training. This training process involves guiding the classifier to focus more on the area that is crucial for facial expression classification. Both occluded and non-occluded regions are considered, but their importance is adjusted. The proposed technique surpasses previous cutting-edge occlusion-aware FER algorithms on face-mask-aware FER datasets, regardless of whether the data is collected in natural environments or controlled laboratory settings. This work only focused on FER categorization, specifically limited to three emotion categories: positive, negative, and natural.

In 2021, El Barachi et al. [56] suggested a hybrid machine learning model that integrates CNNs and SVMs to accurately analyze face sentiment and emotions in facial pictures that are missing or partly occluded. The suggested model underwent successful testing utilizing a dataset of 4,690 photos, which included a total of 25,400 faces. These photographs were gathered at a large-scale public event. The model successfully classified the test dataset, which consisted of faces captured from various perspectives, camera distances, occlusion regions, and picture resolutions. The findings demonstrate a classification accuracy of 89.9% for facial sentiment analysis, and an accuracy of 87.4% for discerning between seven emotions in partly obscured faces. This renders their concept well-suited for real-world practical applications.

In 2022, Delphine, Benjamin et al. [57] suggested a methodology for identifying facial expressions even when there are obstructions on parts of the face. The suggested methodology involves the direct reconstruction of optical flows that have been affected by partial occlusion. These rebuilt optical flows are then used in the subsequent recognition phase. To address this, they suggested using an auto-encoder that is taught to restore distorted optical flow. This is achieved by utilizing a loss function that compares the reconstructed optical flow with the actual optical flow estimated from pictures that do not include any occlusions. The researchers conducted a comparative analysis of several loss functions and different methods for calculating the training optical fluxes. Various studies have shown that the endpoint loss and the Midflows technique, which take into account a broad range of optical flow intensities, seem to be the most appropriate options for this specific job. They demonstrated the high competitiveness of the suggested approach by comparing its findings with the most advanced ones available. They verified the effectiveness of their method using the controlled CK+ datasets, which included various occlusions.

In 2023, Khoeun, Ratanak et al. [58] Proposed a technique for extracting distinctive features of the top face mask to identify emotions in incomplete facial photographs, using Star-Like Particle Polygon Estimation (SLPPE). The proposed Sequential Linear Programming and Project Execution (SLPPE) has three primary stages. At first, a synthetic mask is used to cover the lower part of the actual facial input image. Consequently, the method of representing regions is used for the image of the top face. Moreover, SLPPE serves as a technique for extracting characteristics. This technique is used to generate probability-based feature vectors that accurately capture the attributes of a partly obscured facial image. Unlike previous CNN-based approaches that use several Convolutional modules to analyse the input picture in raster form and extract features, this approach use the SLPPE feature extraction methodology to directly extract features from the original input image. This is done to mitigate the issue of extracting unwanted features resulting from the face mask area supplied by the Convolutional modules. Their proposed SLPPE feature extraction module produces potential transformed data in the form of feature vectors without using Convolutional modules. The proposed method is assessed by applying the SLPPE to the CK+, FER2013, and RAF-DB datasets using Long Short-Term Memory (LSTM) and Artificial Neural Network (ANN) architectures. The CK+, FER2013, and RAF-DB datasets yielded accuracy rates of 99.01%, 98.7%, and 94.62%, respectively.

In 2023, Chen et al. [59] Proposed is a new neural network structure known as the Feature Fusion Residual Attention Network (FFRA-Net). The FFRA-Net consists of three modules: a multi-scale module, a local attention module, and a feature fusion module. The multi-scale module divides the intermediate feature map into many sub-feature maps, distributing them equally over the channel dimension. Subsequently, a convolution operation is executed on each of these feature maps in order to get a diverse set of global features. The local attention module divides the intermediate feature map into many sub-feature maps according on the spatial dimension. Subsequently, a convolution operation is applied to each of these feature maps, resulting in the extraction of distinct local features using the attention mechanism. The feature fusion module is crucial for integrating global and local expression information and establishing residual links between inputs and outputs to offset the loss of intricate traits. Two datasets, namely the Face Masked Real World Affective Face Database (FM_RAF-DB) and the Sun Glasses Real World Affective Face Database (SG_RAF-DB), were constructed specifically for studying occlusion expression. The construction of these databases was based on the RAFDB dataset. The comprehensive experiments undertaken validate that the FFRA-Net, as originally planned, achieves exceptional results on four datasets: FM_RAF-DB, SG_RAF-DB, RAF-DB, and FERPLUS, with accuracies of 77.87%, 79.50%, 88.66%, and 88.97%, respectively. Hence, the approach described in this work demonstrates strong appropriateness for blocked facial expression recognition (FER). Table (2) presents a comparison of previous patterns in extensive IFER research.

4. Comparison Among Reviewed Works

In this section we will present detailed comparisons among the trends toward the complete face expression recognition in one side which is illustrated in Table (1), and among the trends toward the incomplete face expression recognition in other side as presented in Table (2).

Table (1): Comparison among the previous trends of complete FER works.

Ref No.	CNN Model	Dataset used	accuracy
[48] 2023	CNN	FER+	85.74%
[49] 2023	MobileNet, ResNet-18	Tufts	97.75%
		RWTH	86.11%
		RAF	90.81%
		FER2013	77.83%
[50] 2023	EfficientNet	CK+	100%
		KDEF	98%
		JAFFE	98%
		FER2013	98%
[51] 2023	VGG	FER2013	97%

Table (2): Comparison among the previous trends of incomplete FER works.

Ref. No.	Dataset	Model	Depended Face-Parts	Accuracy
[52] 2020	FER+, RAF-DB AffectNet	VGG, ResNet	Lower half face part	For VGG16: FER=79.98%, AffectNet=50.13%, RAF=73.37%
				For ResNet50: FER=79.90%, AffectNet=47.35%, RAF=74.76%
[53] 2020	(CMU Multi-PIE) AffectNet	CNN+ ResNet	Complete face with little covering.	(CMU Multi-PIE)= 84%, AffectNet=72%
[54] 2020	RAF AffectNet	ResNet50	Landmark points of Non-occluded face-areas	RAF=87.16%, AffectNet=64.06%
[55] 2021	Masked-LFW-FER Masked-KDDI-FER	VGG19 MobileNet CNN	Upper unmasked face-part	For Masked-LFW-FER: VGG19=47.12%, MobileNet=52.08%, CNN=87.92%
				For Masked-KDDI-FER: VGG19=46.21%, MobileNet=48.71%, CNN=90.01%%
[56] 2021	Manually collected from AffectNet and Google FERC	CNN	Vertical half of the face.	89.9% (As best result)
[57] 2022	CK+	CNN+ Autoencoder	CFER Eyes Nose+Mouth	CFER=92.8%, Eyes=83.1%, Nose+Mouth=89.6%

[58] 2023	CK+ RAF	LSTM+ANN	Eyes	CK+=99.01%, RAF=94.62%
[59] 2023	FM_RAF SG_RAF RAF FERPLUS	ResNet18	Eyes (Nose+Mouth)	FM_RAF=77.87%, SG_RAF=79.50%, RAF=88.66%, FERPLUS=88.97%

5. Extracted Statistics from Reviewed Works

In this section of the review paper, we will extract data from nearly all of the research listed in Tables (1 and 2) in order to analyze and compare the findings, we will also need to collect data from other relevant studies.

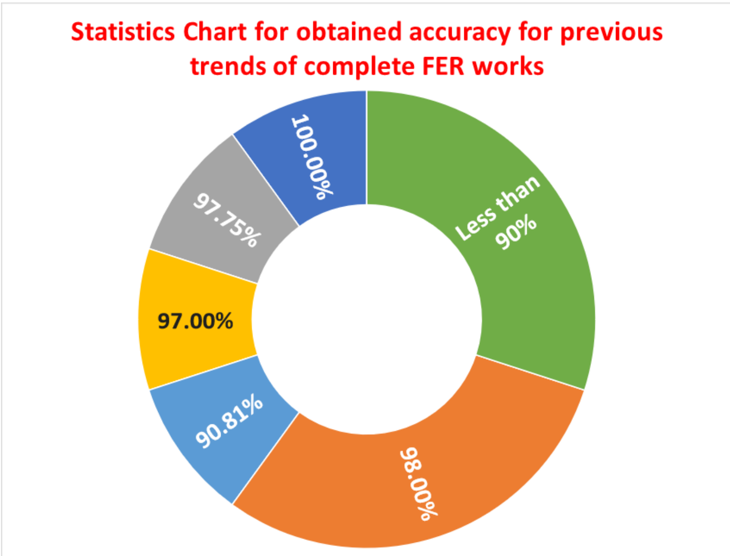

Fig (2): The percentage representation of Statistics Chart for obtained accuracy for previous trends of complete FER works.

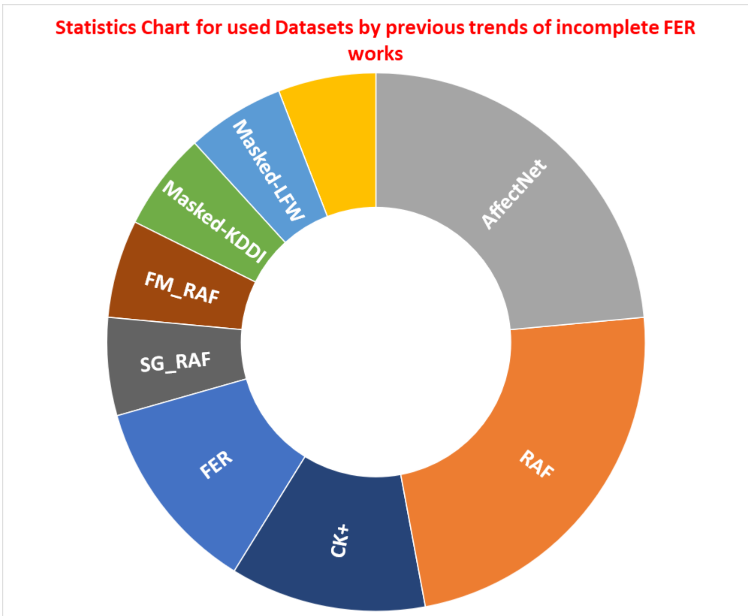

Fig (3): The percentage representation of Statistics Chart for used Datasets by previous trends of incomplete FER works.

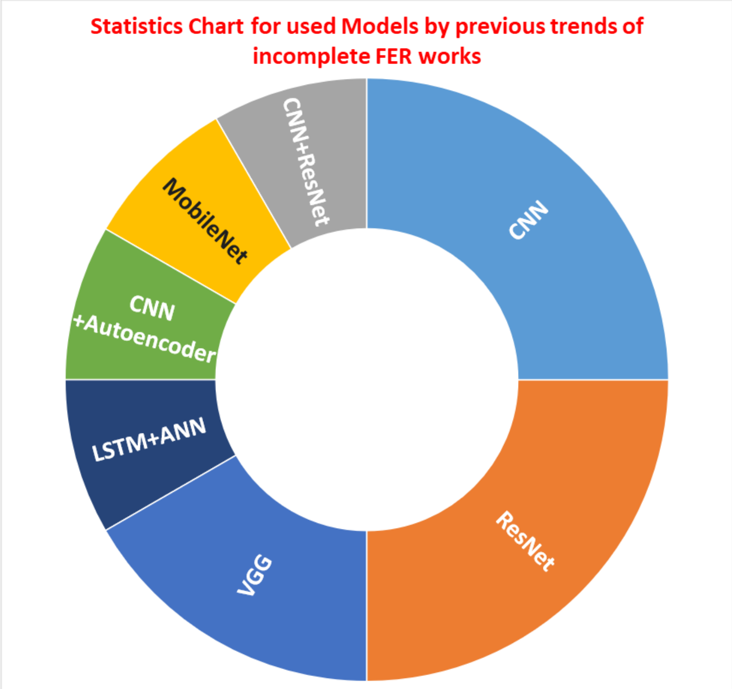

Fig (4): The percentage representation of Statistics Chart for used Models by previous trends of incomplete FER works.

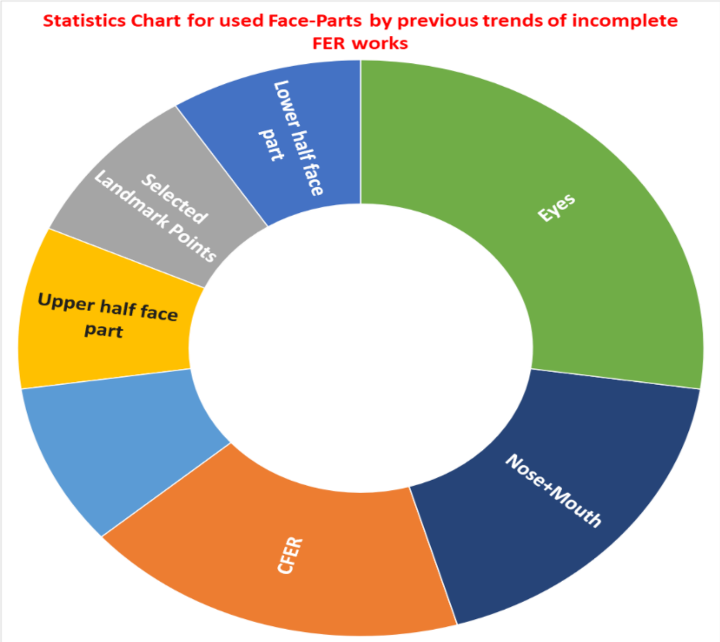

Fig (5): The percentage representation of Statistics Chart for used Face-Parts by previous trends of incomplete FER Works.

6. Conclusion

From the reviewed different previous works related to facial expression recognition including face emotions, it can be concluded that this field is an important one among the new technologies in our life. There two main famous trends in this field which are (FER and IFER). It is unable always to have all face-parts which leads to simple and quick emotions recognition. But, the usual and wide range occurred in our life, is the availability of parts or even (only one part) of the face. However, this paper addressed eight famous datasets and five models for the FER, the obtained accuracy ratios of them were between (77.83% to 100%). While elven datasets, ten models, and seven face-parts selections have been addressed for the IFER, and the obtained ratios were between (47.35% to 99.01%). We extracted important statistics ratios as pie chart plots representing the obtained (accuracies, datasets, and models). Consequently, the researchers who want to work in this field can depend on these comparison tables and extracted statistics to avoid the expected drawbacks and get as maximum as optimum results.

References

[1] S. LokeshNaik, A. Punitha, P. Vijayakarthik, A. Kiran, A. N. Dhangar, B. J. Reddy, *et al.*, "Real Time Facial Emotion Recognition using Deep Learning and CNN," in *2023 International Conference on Computer Communication and Informatics (ICCCI)*, 2023, pp. 1-5.

[2] D. Duncan, G. Shine, and C. English, "Facial emotion recognition in real time," *Computer Science*, pp. 1-7, 2016.

[3] A. Patwal, M. Diwakar, A. Joshi, and P. Singh, "Facial expression recognition using DenseNet," in *2022 OITS International Conference on Information Technology (OCIT)*, 2022, pp. 548-552.

[4] H. Dino, M. B. Abdulrazzaq, S. Zeebaree, A. B. Sallow, R. R. Zebari, H. M. Shukur, *et al.*, "Facial expression recognition based on hybrid feature extraction techniques with different classifiers," *TEST Engineering & Management*, vol. 83, pp. 22319-22329, 2020.

[5] S. M. Saleem, S. R. Zeebaree, and M. B. Abdulrazzaq, "Real-life dynamic facial expression recognition: a review," in *Journal of Physics: Conference Series*, 2021, p. 012010.

[6] S. Meriem, A. Moussaoui, and A. Hadid, "Automated facial expression recognition using deep learning techniques: an overview," *International Journal of Informatics and Applied Mathematics*, vol. 3, pp. 39-53, 2020.

[7] I. Talegaonkar, K. Joshi, S. Valunj, R. Kohok, and A. Kulkarni, "Real time facial expression recognition using deep learning," in *Proceedings of international conference on communication and information processing (ICCIP)*, 2019.

[8] D. Dukić and A. Sovic Krzic, "Real-time facial expression recognition using deep learning with application in the active classroom environment," *Electronics*, vol. 11, p. 1240, 2022.

[9] M. Sajjad, F. U. M. Ullah, M. Ullah, G. Christodoulou, F. A. Cheikh, M. Hijji, *et al.*, "A comprehensive survey on deep facial expression recognition: challenges, applications, and future guidelines," *Alexandria Engineering Journal*, vol. 68, pp. 817-840, 2023.

[10] A. Y. Maghari, "Recognition of partially occluded faces using regularized ICA," *Inverse Problems in Science and Engineering*, vol. 29, pp. 1158-1177, 2021.

[11] S. Ziccardi, F. Crescenzo, and M. Calabrese, "“What is hidden behind the mask?” facial emotion recognition at the time of COVID-19 pandemic in cognitively normal multiple sclerosis patients," *Diagnostics*, vol. 12, p. 47, 2021.

[12] Y. Nan, J. Ju, Q. Hua, H. Zhang, and B. Wang, "A-MobileNet: An approach of facial expression recognition," *Alexandria Engineering Journal*, vol. 61, pp. 4435-4444, 2022.

[13] [C. Bisogni, A. Castiglione, S. Hossain, F. Narducci, and S. Umer, "Impact of deep learning approaches on facial expression recognition in healthcare industries," *IEEE Transactions on Industrial Informatics*, vol. 18, pp. 5619-5627, 2022.

[14] G. Zhao, H. Yang, and M. Yu, "Expression recognition method based on a lightweight convolutional neural network," *IEEE Access*, vol. 8, pp. 38528-38537, 2020.

[15] J. D. Bodapati, U. Srilakshmi, and N. Veeranjaneyulu, "FERNet: a deep CNN architecture for facial expression recognition in the wild," *Journal of The institution of engineers (India): series B*, vol. 103, pp. 439-448, 2022.

[16] M. Akhand, S. Roy, N. Siddique, M. A. S. Kamal, and T. Shimamura, "Facial emotion recognition using transfer learning in the deep CNN," *Electronics*, vol. 10, p. 1036, 2021.

[17] K. Mohan, A. Seal, O. Krejcar, and A. Yazidi, "FER-net: facial expression recognition using deep neural net," *Neural Computing and Applications*, vol. 33, pp. 9125-9136, 2021.

[18] S. Bellamkonda, N. Gopalan, C. Mala, and L. Settipalli, "Facial expression recognition on partially occluded faces using component based ensemble stacked cnn," *Cognitive Neurodynamics*, vol. 17, pp. 985-1008, 2023.

[19] A. S. F. Rodrigues, J. C. Lopes, R. P. Lopes, and L. F. Teixeira, "Classification of facial expressions under partial occlusion for VR games," in *International Conference on Optimization, Learning Algorithms and Applications*, 2022, pp. 804-819.

[20] R. Magherini, E. Mussi, M. Servi, and Y. Volpe, "Emotion recognition in the times of COVID19:

Coping with face masks," *Intelligent Systems with Applications*, vol. 15, p. 200094, 2022.

[21] D. Pamod, C. Joseph, V. Palanisamy, and S. Lekamge, "Emotion Analysis of Occluded Facial Expressions-A Review of Literature," in *2022 ASU International Conference in Emerging Technologies for Sustainability and Intelligent Systems (ICETSIS)*, 2022, pp. 423-429.

[22] P. Bambharolia, "Overview of Convolutional Neural Networks," in *Proceedings of the International Conference on Academic Research in Engineering and Management, Monastir, Tunisia*, 2017, pp. 8-10.

[23] N. Omar, S. R. Zeebaree, M. A. Sadeeq, R. R. Zebari, H. M. Shukur, A. Alkhayyat, *et al.*, "License plate detection and recognition: A study of review," in *AIP Conference Proceedings*, 2023.

[24] A. Khan, A. Sohail, U. Zahoora, and A. S. Qureshi, "A survey of the recent architectures of deep convolutional neural networks," *Artificial intelligence review*, vol. 53, pp. 5455-5516, 2020.

[25] A. Ghosh, A. Sufian, F. Sultana, A. Chakrabarti, and D. De, "Fundamental concepts of convolutional neural network," *Recent trends and advances in artificial intelligence and Internet of Things*, pp. 519-567, 2020.

[26] C. Manresa-Yee, S. Ramis, and J. M. Buades, "Analysis of Gender Differences in Facial Expression Recognition Based on Deep Learning Using Explainable Artificial Intelligence," 2023.

[27] N. Michael, "Artificial intelligence a guide to intelligent systems," ed: Addison Wesley, 2005.

[28] A. Pannu, "Artificial intelligence and its application in different areas," *Artificial Intelligence*, vol. 4, pp. 79-84, 2015.

[29] J. X.-Y. Lek and J. Teo, "Academic Emotion Classification Using FER: A Systematic Review," *Human Behavior and Emerging Technologies*, vol. 2023, 2023.

[30] B. S. Tahir, Z. S. Ageed, S. S. Hasan, and S. R. Zeebaree, "Modified Wild Horse Optimization with Deep Learning Enabled Symmetric Human Activity Recognition Model," *Computers, Materials & Continua*, vol. 75, 2023.

[31] D. Sarkar, R. Bali, and T. Sharma, "Practical machine learning with Python," *Book" Practical Machine Learning with Python*, pp. 25-30, 2018.

[32] C. Dalvi, M. Rathod, S. Patil, S. Gite, and K. Kotecha, "A survey of ai-based facial emotion recognition: Features, ml & dl techniques, age-wise datasets and future directions," *Ieee Access*, vol. 9, pp. 165806-165840, 2021.

[33] D. S. Shakya, "Analysis of artificial intelligence based image classification techniques," *Journal of Innovative Image Processing*, vol. 2, pp. 44-54, 2020.

[34] B. Jena, S. Saxena, G. K. Nayak, L. Saba, N. Sharma, and J. S. Suri, "Artificial intelligence-based hybrid deep learning models for image classification: The first narrative review," *Computers in Biology and Medicine*, vol. 137, p. 104803, 2021.

[35] J. Saeed and S. Zeebaree, "Skin lesion classification based on deep convolutional neural networks architectures," *Journal of Applied Science and Technology Trends*, vol. 2, pp. 41-51, 2021.

[36] P. Wang, E. Fan, and P. Wang, "Comparative analysis of image classification algorithms based on traditional machine learning and deep learning," *Pattern Recognition Letters*, vol. 141, pp. 61-67, 2021.

[37] J. Luo, Z. Xie, F. Zhu, and X. Zhu, "Facial expression recognition using machine learning models in fer2013," in *2021 IEEE 3rd International Conference on Frontiers Technology of Information and Computer (ICFTIC)*, 2021, pp. 231-235.

[38] D. A. Zebari, H. Haron, S. R. Zeebaree, and D. Q. Zeebaree, "Enhance the mammogram images for both segmentation and feature extraction using wavelet transform," in *2019 International Conference on Advanced Science and Engineering (ICOASE)*, 2019, pp. 100-105.

[39] S. Kim, G. H. An, and S.-J. Kang, "Facial expression recognition system using machine learning," in *2017 international SoC design conference (ISOCC)*, 2017, pp. 266-267.

[40] A. N. Dixit and T. Kasbe, "A survey on facial expression recognition using machine learning techniques," in *2nd international conference on data, engineering and applications (IDEA)*, 2020, pp. 1-6.

[41] B. Mahesh, "Machine learning algorithms-a review," *International Journal of Science and Research (IJSR).[Internet]*, vol. 9, pp. 381-386, 2020.

[42] S. M. S. A. Abdullah, S. Y. A. Ameen, M. A. Sadeeq, and S. Zeebaree, "Multimodal emotion recognition using deep learning," *Journal of*

Applied Science and Technology Trends, vol. 2, pp. 52-58, 2021.

[43] M. J. Mohammed Jasim, B. K. Hussan, S. R. Zeebaree, and Z. S. Ageed, "Automated Colonic Polyp Detection and Classification Enabled Northern Goshawk Optimization with Deep Learning," *Computers, Materials & Continua,* vol. 75, 2023.

[44] K. B. Obaid, S. Zeebaree, and O. M. Ahmed, "Deep learning models based on image classification: a review," *International Journal of Science and Business,* vol. 4, pp. 75-81, 2020.

[45] M. Usama, J. Qadir, A. Raza, H. Arif, K.-L. A. Yau, Y. Elkhatib*, et al.*, "Unsupervised machine learning for networking: Techniques, applications and research challenges," *IEEE access,* vol. 7, pp. 65579-65615, 2019.

[46] M. Karnati, A. Seal, D. Bhattacharjee, A. Yazidi, and O. Krejcar, "Understanding deep learning techniques for recognition of human emotions using facial expressions: a comprehensive survey," *IEEE Transactions on Instrumentation and Measurement,* 2023.

[47] S. R. Zeebaree, O. Ahmed, and K. Obid, "Csaernet: An efficient deep learning architecture for image classification," in *2020 3rd International Conference on Engineering Technology and its Applications (IICETA)*, 2020, pp. 122-127.

[48] K. DONUK, A. Ali, M. F. ÖZDEMİR, and D. HANBAY, "Deep feature selection for facial emotion recognition based on BPSO and SVM," *Politeknik Dergisi,* vol. 26, pp. 131-142, 2023.

[49] D.-H. Lee and J.-H. Yoo, "CNN Learning Strategy for Recognizing Facial Expressions," *IEEE Access,* 2023.

[50] S. B. Punuri, S. K. Kuanar, M. Kolhar, T. K. Mishra, A. Alameen, H. Mohapatra*, et al.*, "Efficient net-XGBoost: an implementation for facial emotion recognition using transfer learning," *Mathematics,* vol. 11, p. 776, 2023.

[51] T. Winyangkun, N. Vanitchanant, V. Chouvatut, and B. Panyangam, "Real-Time Detection and Classification of Facial Emotions," in *2023 15th International Conference on Knowledge and Smart Technology (KST)*, 2023, pp. 1-6.

[52] B. Houshmand and N. M. Khan, "Facial expression recognition under partial occlusion from virtual reality headsets based on transfer learning," in *2020 IEEE Sixth International Conference on Multimedia Big Data (BigMM)*, 2020, pp. 70-75.

[53] K. Soumya and S. Palaniswamy, "Emotion recognition from partially occluded facial images using prototypical networks," in *2020 2nd International Conference on Innovative Mechanisms for Industry Applications (ICIMIA)*, 2020, pp. 491-497.

[54] H. Ding, P. Zhou, and R. Chellappa, "Occlusion-adaptive deep network for robust facial expression recognition," in *2020 IEEE International Joint Conference on Biometrics (IJCB)*, 2020, pp. 1-9.

[55] B. Yang, W. Jianming, and G. Hattori, "Face mask aware robust facial expression recognition during the COVID-19 pandemic," in *2021 IEEE International conference on image processing (ICIP)*, 2021, pp. 240-244.

[56] M. El Barachi, M. AlKhatib, and S. S. Mathew, "A Hybrid Machine Learning Approach for Sentiment Analysis of Partially Occluded Faces," in *2021 IEEE International Conference on Artificial Intelligence and Computer Applications (ICAICA)*, 2021, pp. 389-392.

[57] D. Poux, B. Allaert, N. Ihaddadene, I. M. Bilasco, C. Djeraba, and M. Bennamoun, "Dynamic facial expression recognition under partial occlusion with optical flow reconstruction," *IEEE Transactions on Image Processing,* vol. 31, pp. 446-457, 2021.

[58] R. Khoeun, W. Yookwan, P. Chophuk, A. Rodtook, and K. Chinnasarn, "Emotion Recognition of Partial Face using Star-Like Particle Polygon Estimation," *IEEE Access,* 2023.

[59] Y. Chen, S. Liu, D. Zhao, and W. Ji, "Occlusion facial expression recognition based on feature fusion residual attention network," *Frontiers in Neurorobotics,* vol. 17, 2023.