

Facial emotion analysis through image processing using AI

Student: Turki Arif

Institution Name: Midocean University

Advisor: Dr. Ashraf Elsayed.

Abstract

Recognizing emotions from facial expressions is challenging due to the variability of human emotion patterns, which are influenced by factors such as facial orientation, lighting conditions, and subtle differences in complex emotions. To address these complexities and enhance real-world applications, I focus on the use of a deep convolutional neural network (VGG19) trained with precision. We will use the CK+ and FER2013 datasets. We aim to enhance interactions and expand the scope of its application.

1-Interdiction

Within the scope of human-computer interaction, a highly important technology has emerged: facial emotion recognition, which is used in many fields. Accurate interpretation of human facial emotions is essential in our era and is the foundation for many applications, including psychological assessments, digital marketing, virtual reality experiences, and the improved use of AI. However, despite recent advances, reliable interpretation of facial emotions remains challenging due to the subtleties of expression variations influenced by environmental factors such as facial orientation, lighting conditions, and subtle differences in individual emotions. We can overcome these challenges by using AI to improve accuracy, which will be satisfactory to both users and systems.

Regarding current research, much of it is focused on approaches to dealing with unpredictable real-world facial conditions. These systems often face difficulties in consistency and generalization, especially when facing a changing environment. Therefore, in this project, we will try to address the challenges by using the powerful VGG19 convolutional neural network model based on the ready-made facial expression data CK+ & FER2013.

2-Problem Definition:

the problem addressed by this project is the accurate classification of human emotions from facial expression images. This is formally known as a multi-class classification problem, encompassing emotions such as happiness, sadness, anger, disgust, surprise, neutrality, and fear. This approach leverages the deep convolutional neural network known as VGG19, which is pre-trained for image-based emotions using publicly available and pre-curated datasets such as CK+ and FER2013. These datasets are subject to constraints on consistent image quality and lighting conditions. These constraints include the availability of computational resources and the reliance on annotated dataset, which is problematic due to the large size and diversity of the data. The primary goal is to understand facial emotions and the accuracy of their classification. Through VGG19, we will explore the performance and efficiency of the network, which may facilitate practical applications such as psychological assessment, automation, and user experience design for interested users.

3-Motivations

The ability to accurately recognize human emotions through facial expressions has implications that benefit various fields. In practical applications, emotion recognition systems improve user experiences in customer service, marketing strategies, health diagnosis, and ultimately human-computer interaction. Academically, research is pivotal in the field of computer vision and AI. This provides an opportunity to address limitations and develop new methodologies for accurately detecting emotions. This interest is growing due to its potential to revolutionize gaming, virtual reality, and psychographics. However, significant gaps remain, particularly in generalization across diverse datasets and dealing with a changing environment. Personal motivations stem from the ability of technologies to significantly enhance daily interactions between humans and intelligent systems, thus improving overall quality of life.

4-Objectives

In this project, we aim to implement systems and examine their effectiveness in recognizing facial emotion outputs using the VGG19 deep convolutional neural network model through transfer learning. We aim to reliably classify emotions across seven different categories using the VGG19 model. The objectives also include evaluating the model's performance across multiple reference datasets (CK+ & FER2013), applying appropriate preprocessing and amplification techniques, and analyzing the system's ability to adapt to noise and variance in image data. We also aim to compare the performance of the proposed models with current state-of-the-art technologies to highlight strengths and potential areas for improvement.

In the long term, the project aims to understand the mechanism of facial emotion recognition using artificial intelligence. This will have implications for healthcare, education, security, and interactive AI platforms, enhancing user experience

.

5-Literature Review

Facial expressions play a crucial role in nonverbal communication and have been at the heart of emotion recognition research since Ekman and Friesen laid the groundwork in 1978. Thanks to recent strides in deep learning, especially with convolutional neural networks (CNNs), we've seen a remarkable boost in the accuracy of facial emotion recognition. Models like ResNet-50 and VGG-16 are particularly adept at picking up on subtle differences in images, but they often focus more on pixel details than on the emotional context, which makes them more susceptible to adversarial attacks [5]. In the past, traditional machine learning methods like KNN and SVM managed to hit recognition rates of about 84%, but CNNs, especially when paired with advanced optimizers like Adam, have pushed that accuracy up to 89% [4].

However, there are still some significant hurdles to overcome. Detecting emotions in real-time is quite challenging due to issues like computational inefficiency, reliance on specific datasets, and the necessity for extensive training data [1, 4]. Factors in the environment, such as lighting and variations in posture, can also greatly affect how reliable and accurate these models are [1]. While popular datasets like AffecNet, RAF-DB, and CK+ have helped improve model generalization, they still carry biases that can limit their effectiveness in real-world situations [2, 5].

On a brighter note, transfer learning techniques, particularly with the VGG-19 model, have made a significant impact on sentiment analysis, achieving accuracy levels as high as 99% on datasets like CK+ [2]. There are suggestions to optimize the VGG-19 architecture and its hyperparameters to further enhance classification performance and simplify the model, which would help computational challenges and dataset constraints [2]. Moreover, fine-tuning deep learning models like AlexNet-DCNN has shown recognition accuracies above 92%, reinforcing the value of transfer learning [3]. To truly assess and enhance the effectiveness of transfer learning models in real-time emotion detection, it's essential to conduct comparative analyzes and optimize hyperparameters across various datasets.

Hybrid methods that blend CNNs with Facial Action Coding Systems (FACO) have been introduced to boost robustness, accuracy, and resilience against adversarial attacks. This approach effectively merges pixel-level data with emotional context [5]. On a similar note, traditional feature extraction techniques, like Gabor filters paired with genetic algorithms, have achieved impressive accuracies over 90%, indicating that classical methods can play a supportive role in enhancing CNN performance [3].

To sum it up, the progress in CNNs and transfer learning has greatly improved facial emotion recognition. However, we still face challenges such as sensitivity to environmental factors, computational limitations, biases in datasets, and vulnerabilities to adversarial attacks. Moving forward, research should prioritize computational efficiency, hybrid approaches, strong transfer learning techniques, and hyperparameter optimization to ensure reliable, real-time emotion recognition.

In conclusion, the VGG19 model remains one of the most effective and widely used models for facial emotion recognition. This literature review highlights the necessary improvements for practical application.

6- Proposed Methodology

The proposed approach relies on deep learning techniques, specifically the VGG19 convolutional neural network model, known for its powerful precursor in image classification, and its deep

architecture based on transfer learning. This model is trained on the ImageNet dataset, using facial expression datasets to effectively classify emotions. This model also has the ability to extract deep spatial features in images, which is important for capturing subtle facial cues associated with different emotions.

Preprocessing includes face detection, cropping, grayscale transformation, normalization, and data augmentation techniques such as scaling, rotation, and flipping. These steps are essential for ensuring consistent input quality and enhancing the model's generalization ability. Features are extracted internally by the model's convolutional layers, eliminating the need for manual feature engineering.

The project will be implemented using the PyTorch framework, thanks to its flexibility, dynamic drawing, and robust GPU acceleration support. Supporting libraries include NumPy and Pandas for data processing, OpenCV for preprocessing, and Matplotlib for visualization. Training parameters, such as learning rate, batch size, and number of stages, will be optimized through experimentation, along with performance, precision, recall, and the F1 score of the confusion matrix.

This methodology is designed to produce an efficient, scalable, and effective emotion recognition system that performs well under a variety of conditions and provides a foundation for future improvements or deployment in real-time applications.

7- Dataset Description

The study uses two widely known open-source facial emotion recognition datasets: CK+ and FER2013. These datasets were chosen due to their availability and importance in analyzing facial expressions and representing diverse emotional states.

FER2013 is a large dataset containing 35,887 grayscale images (approximately 15 GB) with dimensions of 48 x 48 pixels. It classifies the seven emotions: disgust, anger, fear, happiness, sadness, surprise, and neutral. It is derived from the Kaggle platform and is used repeatedly to train deep learning models and measure their performance.



CK+ consists of 920 images (approximately 18 GB) that are displayed in a sequential frame and labeled with one of the seven standard emotions. All images undergo preprocessing steps such as resizing, grayscale normalization, histogram equalization, and augmentation techniques (flip, rotate, and scale) to improve model generalization and reduce overfitting.



Challenges facing these two groups include class imbalances (e.g., a lack of samples expressing fear and disgust), discrepancies in resolution between images, and inconsistent facial orientations. These challenges are mitigated through data augmentation, balanced sampling strategies, and processing pipelines.

8- Project implementation timeline

Week	Description
1-2	Download CK+ database preparation (Apex peak image extraction)
3-4	Setting up the working environment (PyTorch) and image processing
5-7	Building and training the model using VGG19 Transfer Learning
8	Applying Augmentation and Performance Improvement Techniques
9	Model evaluation using performance metrics and confusion matrix analysis
10	Writing the final report and presenting the presentation

9- Data source:

<https://www.kaggle.com/datasets/msambare/fer2013>

<https://zenodo.org/records/11221351>

10- References

1-Olfa Askria,b, Ghaith Manitaç, Mohamed Ali Hajjajib. (2024) Efficient Facial Emotion Recognition Using An Optimized Deep Learning Model Based On Quantum Gazelle Optimization Algorithm , ScienceDirect .

<https://www.sciencedirect.com/science/article/pii/S1877050924024311>

2-Gaurav Meena, Krishna Kumar Mohbey, Ajay Indian, Sunil Kumar.(2022) Sentiment Analysis from Images using VGG19 based Transfer Learning Approach, ScienceDirect.

<https://www.sciencedirect.com/science/article/pii/S1877050922007888>

3-Gaurav Meena, Krishna Kumar Mohbey. (2023) Sentiment analysis on images using different transfer learning models. ScienceDirect

<https://www.sciencedirect.com/science/article/pii/S1877050923001424>

4-Sahana M, Praneetha Umesh, Ashwini Kodipalli, Trupthi Rao. (2024) EmoCNN Unleashing human emotions with customized CNN using different optimizers. ScienceDirect

<https://www.sciencedirect.com/science/article/pii/S1877050924008007>

5- Harisu Abdullahi Shehu, Will N. Browne, Hedwig Eisenbarth. 2024 Emotion categorization from facial expressions: A review of datasets, methods, and research directions

ScienceDirect <https://www.sciencedirect.com/science/article/pii/S0925231225000396>