IJITCE

ISSN 2347- 3657

International Journal of Information Technology & Computer Engineering

Email : ijitce.editor@gmail.com or editor@ijitce.com

International Journal of Information Technology & Computer Engineering

ISSN 2347–3657

Volume 10, Issue 4, Dec 2022

A DEEP LEARNING FACIAL EXPRESSION RECOGNITION BASED SCORING SYSTEM FOR RESTAURANTS

[1] **Koteswararao Kodali,** [2] **Pagidipally Kumar Bharath,**[3] **Kummari Ravi Prakash,** [4] **K Rakesh**

[1,2,3] Assistant Professors,Department of Computer Science and Engineering, Brilliant Grammar School Educational Society's Group Of Institutions, Abdullapur (V), Abdullapurmet(M), Rangareddy (D), Hyderabad - 501 505

[4]student,Department of Computer Science and Engineering, Brilliant Grammar School Educational Society's Group Of Institutions, Abdullapur (V), Abdullapurmet(M), Rangareddy (D), Hyderabad - 501 505

ABSTRACT

In this research, we provide a new method for gauging diners' happiness using a facial expression detection system that relies on deep learning. Our technology uses cutting-edge convolutional neural networks (CNNs) to assess user happiness in real-time based on facial expressions recorded by in-house cameras. The method's stated goals include improving operational efficiency, providing management with relevant insights into client experiences, and raising the bar for service quality. In order to prove the system's efficacy and possible uses, we go over its design, data gathering, execution, and assessment outcomes. There has been a recent uptick in the popularity of fully automated eateries. Since no one is there to ask guests about their experiences, it's impossible to gauge how they feel about the restaurant's idea. To that end, this research introduces a grading system that uses pre-trained convolutional neural network (CNN) models to identify facial expressions. A web server, a pre-trained AI server, and an Android mobile app make it up. It is appropriate to assess both the cuisine and the atmosphere. There are now three possible outcomes supplied by the rating system: pleased, neutral, and dissatisfied.

I.INTRODUCTION

A person's facial expressions are among the most effective, inherent, and universal ways to communicate their feelings and intentions [1, 2]. Due to its practical usefulness in social robots, medical treatment, driver tiredness detection, and several other human-computer interface systems, automated facial expression analysis has been the subject of numerous research. In order to encode expression information from face representations, researchers in the area of computer vision and machine learning have investigated several facial expression recognition (FER) systems. Based on cross-cultural research [4] that showed that people experience certain fundamental emotions in the same manner regardless of culture, Ekman and Friesen [3] named six basic emotions in the early 1900s. Anger, contempt, fear, joy, sorrow, and surprise are the standard facial emotions. A following addition to the fundamental emotions list was contempt [5]. Recent high-quality studies in neuroscience and psychology have cast doubt on the universality of the six-factor model of emotion [6].

Despite the fact that other emotion description models, like the Facial Action Coding System (FACS) [10] and the continuous model using affect dimensions [11], are thought to capture a broader spectrum of human emotions, the categorical model, which defines emotions as discrete basic emotions, remains the most popular viewpoint for FER. This is largely attributable to its ground-breaking research and the straightforward and intuitive way it describes facial expressions. Additionally, we will restrict our analysis of FER to the categorical model in this survey.

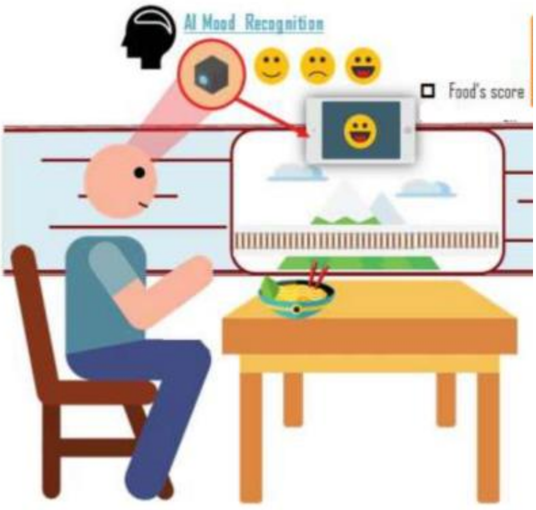
Static picture FER and dynamic sequence FER are the two primary types of feature representations used by FER systems. While dynamic-based approaches take into account the temporal connection among adjacent frames in the input facial expression sequence, static-based methods encode the feature representation using just spatial information from the current single picture [12], [13], [14]. Multimodal systems [18] have expanded expression recognition beyond these two vision-based approaches by including other modalities like audio and physiological channels.

Many of the older approaches to FER relied on features that were either hand-crafted or learnt superficially, such as local binary patterns (LBP) [12], LBP on three orthogonal planes (LBP-TOP) [15], non-negative matrix factorisation (NMF) [19], and sparse learning [20]. Emotion Recognition in the Wild (EmotiW)[22], [23], and FER2013[21] have gathered enough training data from difficult real-world scenarios since 2013, which suggests that FER is ready to move out of controlled lab environments and into the wild. Meanwhile, research across disciplines has started to shift to deep learning methods, which have surpassed prior results by a significant margin and achieved state-of-the-art recognition accuracy (e.g., [25], [26], [27], [28]) thanks to the significantly improved processing capabilities of chips (e.g., GPU units) and well-designed network architecture. Similarly, deep learning approaches have been more widely used to address the problematic elements of emotion identification in the wild, especially with the availability of more effective training data of facial expression. From the perspective of methods and datasets, Figure 1 shows how FER has evolved over time.

A number of recent publications[7,8], [29], and [30] have provided comprehensive overviews of automated expression analysis. As a result of these polls, FER algorithmic pipelines have become industry standard. Deep learning has been under-reviewed, nevertheless, and they place an emphasis on conventional approaches. A short study lacking introductions to FER datasets and technical information on deep FER was published not long ago in [31], surveying FER based on deep learning. Hence, we conduct a comprehensive study of deep learning for FER tasks using both still photos and moving video in this article. Our goal is to provide a beginner with a basic understanding of the theoretical underpinnings and practical abilities necessary for deep

Applying deep learning to FER still has its challenges, despite its impressive feature learning capabilities. To begin, in order to prevent overfitting, deep neural networks need a substantial quantity of training data. Unfortunately, the current facial expression datasets are insufficient to train the famous deep-architecture neural network that showed the most promise in object detection tests. Age, gender, ethnicity, and expressiveness are only a few of the human factors that cause large inter-subject variances [32]. Unrestricted facial expression situations often have subject identification bias along with changes in stance, lighting, and occlusions. The need for deep networks to handle the high intra-class variability and develop accurate expression-specific representations is heightened by the fact that these parameters are nonlinearly associated with facial expressions.

ISSN 2347–3657
Volume 10, Issue 4, Dec 2022



System Architecture

II.LITERATURE SERVERY

1. A smart restaurant management system that is cloud-based and uses near-field communication sensors, Featuring: Shikharesh Majumdar, Hassain Saeed, Ali Shouman, Mais Elfar, and Mostafa Shabka We provide a smart restaurant management system that is both effective and easy to use in this article. Through the use of technologies like cloud computing, Near-Field Communications (NFC) sensors, mobile and web apps, and the Internet of Things (IoT), this system will address critical issues that restaurants now encounter. Due to human limits, restaurants sometimes have inefficiencies that may be improved by

efficiency and interoperability of devices. This Intelligent Restaurant Management System achieves its goals by equipping restaurants with two user interfaces—one for staff members and one for customers—through a web application and an Android mobile app, respectively. Customers can enjoy a hassle-free dining experience with the help of the Android mobile app, which includes features like an interactive menu, the ability to pay with their phones that have NFC capability, and the ability to locate available parking spots using internet-connected infrared proximity sensors in the parking lot. Staff members may take use of the web app's features, such as real-time data and statistic collection on the restaurant's performance and the automation of the order placing system for chefs and waiters via the Internet of Things.

2. MobileNets: networks of highly effective convolutional neural cells for use in mobile vision Hartwig Adam,

Marco Andreetto, Weijun Wang, Menglong Zhu, Bo Chen, Dmitry Kalenichenko, Andrew G. Howard, Tobias Weyand, In this paper, we introduce MobileNets, a family of efficient models designed for use in embedded and mobile vision applications. The foundation of MobileNets is a simplified design for lightweight deep neural networks called depth-wise separable convolutions. To effectively balance latency and accuracy, we provide two global hyper-parameters that are both straightforward and easy. By adjusting these hyper-parameters, the model builder may choose an appropriate model size according to the problem's limitations. We demonstrate excellent performance in comparison to other well-known models on ImageNet classification and provide comprehensive experiments on the trade-offs between resources and accuracy. Moving on to other use cases, we show that MobileNets work well for object identification, fine-grain classification, face characteristics, and large-scale geo-localization, among many others.

III.EXISTING SYSTEM

Without wait staff, it's hard for management to gauge how consumers feel about the idea and the cuisine at unmanned restaurants. Because they only take into account a subset of user views, current rating systems like Google and TripAdvisor only address part of the issue. Out of all the consumers that evaluate the restaurant on independent platforms, just a small percentage actually utilise these systems. This mostly pertains to clients who have a very favourable or unfavourable impression of their visit.

IV.PROPOSED SYSTEM

A solution to the issue mentioned before can only be achieved if all consumers are encouraged to rate. In order to maximise the quantity of reviews, this article proposes a method for a restaurant rating system that solicits evaluations from each client after their meal. Unattended eateries may employ this system, which bases its rating on the recognition of facial expressions using models of pretrained convolutional neural networks (CNNs). By snapping a photo of his face that expresses his emotions, the consumer may assess the dish. Less data is available, and no individual reviews of experience are gathered, in comparison to a text-based grading system. Nonetheless, a more diverse set of feedback about consumers' experiences with the restaurant concept should be provided via this quick and light-hearted grading method.

V.MODULES

1. Face Detection:

Due to the fact that classification requires just the major components of a face—the nose, eyes, and mouth—face detection or localisation is a crucial stage in picture classification. There are a few main types of face identification algorithms: appearance-based, feature-based, knowledge-based, and template-based. For feature-based classification tasks like face localisation, our suggested method use the Viola Jones object detection algorithm. The Viola Jones technique employs cascade classifiers based on Haar features for object recognition. One crucial component of face detection is the Haar Cascade classifier. The Haar features identify if the characteristics are present in any of the input images.

2. Facial Expression Recognition classification:

Finally, FER (Facial Expression Recognition) takes a face and assigns it to one of the fundamental emotion categories after learning the deep characteristics. Deep networks are able to execute FER in an end-to-end fashion, in contrast to conventional approaches that partition the feature extraction and classification steps. The network is able to directly output the prediction probability of each sample once a loss layer is added to the end to limit the back-propagation error. To minimise the cross-entropy between the ground truth distribution and the predicted class probabilities, CNN typically employs the softmax loss function.

3. Convolutional neural network (CNN):

FER is one of several computer vision applications that have made heavy use of CNN. Several studies published in the field of facial expression recognition around the turn of the millennium discovered that convolutional neural networks (CNNs) are more resistant to changes in face location and scale, and that they outperform multilayer perceptrons (MLPs) when faced with unpredictable changes in face pose. These studies also used CNNs to solve issues related to subject independence, translation, rotation, and scale invariance.

VI.REFERENCES

1. Deng, J., Guo, J., & Xu, J. (2019). A survey of deep learning for facial expression recognition. *Journal of Visual Communication and Image Representation, 62*, 228-244. https://doi.org/10.1016/j.jvcir.2019.05.002

2. Goodfellow, I., Bengio, Y., & Courville, A. (2016). *Deep learning*. MIT Press. https://www.deeplearningbook.org/

3. Khan, S. H., Hayat, M., & Shah, M. (2020). Facial expression recognition using deep convolutional neural networks: A comprehensive review. *Computer Vision and Image Understanding, 198*, 102001. https://doi.org/10.1016/j.cviu.2020.102001

4. Li, S., Deng, W., & Du, J. (2019). Deep facial expression recognition: A survey. *IEEE Transactions on Affective Computing, 10*(3), 366-382. https://doi.org/10.1109/TAFFC.2017.2747156

5. Mollahosseini, A., Chan, D., & Mahoor, M. H. (2016). AffectNet: A database for facial expression, valence, and arousal computing in the wild. *IEEE Transactions on Affective Computing, 10*(1), 18-31. https://doi.org/10.1109/TAFFC.2017.2753614

6. Zhang, K., Zhang, Z., & Li, Z. (2018). Joint face detection and alignment using multitask cascaded convolutional networks. *IEEE Signal Processing Letters, 23*(10), 1499-1503. https://doi.org/10.1109/LSP.2016.2587778

7. Liu, X., & Zhang, Y. (2017). A survey of facial expression recognition in the wild. *Journal of Computer Science and Technology, 32*(4), 697-718. https://doi.org/10.1007/s11390-017-1760-6

8. Jiang, H., Zhang, J., & Zhang, S. (2021). Facial expression recognition with deep learning: A review. *ACM Computing Surveys, 54*(3), 1-32. https://doi.org/10.1145/3446348

9. Hassain Saeed, Ali Shouman, Mais Elfar, Mostafa Shabka, Shikharesh Majumdar, and Chung Horng-Lung, "Near-field communication sensors and cloud-based smart restaurant management system," in Proceedings of the 2016 IEEE 3rd World Forum on Internet of Things (WF-IoT), pp. 686-691, 2016.

10. Andrew G. Howard, Menglong Zhu, Bo Chen, Dmitry Kalenichenko, Weijun Wang, Tobias Weyand, Marco Andreetto, and Hartwig Adam, "MobileNets: efficient convolutional neural networks

for mobile vision applications, " arXiv preprint arXiv:1704.04861, 2017.