IJCRT.ORG ISSN : 2320-2882

INTERNATIONAL JOURNAL OF CREATIVE RESEARCH THOUGHTS (IJCRT)

An International Open Access, Peer-reviewed, Refereed Journal

A Deep Learning Facial Expression Recognition Based Scoring System For Restaurants

Ravindra Changala, Akula Tarun Sai, Govreddy Jaya Prakash Reddy, Vighnesh Cilamkoti, Gaddam Bharath Reddy.

Associate Professor, Department of Information Technology, Guru Nanak Institutions Technical Campus, Hyderabad, India.
Department of Information Technology, Guru Nanak Institutions Technical Campus, Hyderabad, India.

Abstract: Automated and unmanned restaurant appeal has surged in recent times. Since there is no staff present, it is impossible to directly gauge what patrons think of the restaurant idea or learn about their experiences. This research provides a grading system based on pre-trained convolutional neural network (CNN) models for face expression identification. It consists of a web server, an AI server that has already been taught, and an Android mobile application. It is intended that both the cuisine and the setting be appraised. Restaurants that are both automated and staff-free are growing in popularity. It is unable to directly evaluate the opinions of the customers due to a staffing shortage, thus it is impossible to find out about their experiences with the restaurant concept. This study uses convolutional neural network (CNN) models that have already undergone training to produce a facial emotion recognition grading system. It consists of an AI server that has previously been trained, a web server, and an Android mobile application. It is intended that reviews would be left for both the food and the environment. Three expressions are now available in the grading system: neutral, dissatisfied, and satisfied.

Keywords: CNN, AI, DCNN, deep neural networks.

I.INTRODUCTION

Emotion and purpose may be communicated through facial expressions, which are strong, natural, and universal [1, 2]. Due of its potential applications in medical diagnostics, social robotics, and driver fatigue monitoring. An extensive body of research has been conducted on automated face image analysis.

In computer vision and machine learning, a number of approaches for facial expression recognition (FER) have been investigated to encode expression data from facial expressions. Six fundamental emotions were discovered by Ekman and Friesen [3] at the start of the 20th century. Their study was inspired by cross-cultural studies [4], which revealed that some basic emotions are felt in comparable ways by people in different cultures.

Common facial emotions include surprise, amusement, contempt, joy, sadness, and rage. Eventually, contempt was included in the list of essential emotions [5]. The hypothesis that there are six basic emotions that all people experience has been called into question by recent research in psychology and neuroscience [6]. Continuous models with affect parameters and the Facial Action Coding System (FACS) [10] are two examples. Two further emotion description models that are frequently thought to adequately represent human emotion.

In recent years, the rapid development of information and communication technology (ICT), the Internet of Things (IoT) and artificial intelligence (AI) has resulted in an increasing number of applications based on these technologies. Following this trends, the popularity of automated and unmanned restaurants continues to grow. Particularly in Japan and Taiwan, successfully if growing. The field of computer vision known as facial expression classification (FEC) has greatly profited from the quick developments in machine learning. These advancements have made it possible to use data sets that include a variety of facial expressions taken of various people in order to train classifiers that can distinguish between various types of facial expressions. Specifically, the use of deep learning to FEC has resulted in notable enhancements in accuracy while dealing with intricate circumstances like fluctuating illumination, angles, or occlusion.

The tremendous computational and architectural complexity of the deep neural networks that power these systems is one of the main obstacles to the general deployment of deep learning-based FEC systems, despite the fact that their performance keeps improving. This obstacle is especially restrictive in real-time embedded settings where low latency operation on the low-cost embedded devices is necessary. When it comes to the use of assistive technologies to enhance quality of life, for instance, most people who use them are reluctant to carry around big, expensive, and bulky devices because doing so would make it more difficult for them to make the most of the technologies seamlessly. Because of this, the assistive gadgets need to make use of tiny, inexpensive embedded computers that nonetheless have minimal latency in order to give the user real-time input.

II.RELATED WORKS

Recently, a number of comprehensive studies on automated expression analysis were published [7]. FER's standard algorithm pipeline has been developed utilising these surveys. They underline the use of conventional approaches and highlight the lack of attention that deep learning has received. You can find a succinct deep learning-based description of FER that skips over the introduction of the FER dataset and the technical specifics of deep FER here. Thus, utilising both still photographs and films In this study, we comprehensively look at deep learning for photo sequence (FER) issues. Our objective is to familiarise a novice with the foundational techniques and skill sets required for in-depth FE.

Over the past thirty years, it has been clear that automatic facial expression identification and classification for multi-pose and multilayer face photos is an appealing and difficult topic. According to a study of the literature, early research has concentrated on a number of statistical and structurally based techniques. On the other hand, various feature- and template-based methods have also been studied. Many researchers have adopted the classical methods to obtain texture features in statistical ways, such as the Histogram-of-Orientation Gradient (HOG), the Scale Invariant Feature Transform (SIFT), the LBP (Local Binary Pattern) features, and some spatio-temporal features (STM-ExpLet). However, these methods are labour intensive and require high performance.
Convolutional neural networks, or CNNs, have been applied by researchers recently [22, 23], with remarkable results for large-scale static image and video sequence recognition. The CNN has been extensively used for the FER system and has greatly enhanced cutting-edge procedures in addition to examining ImageNet classification problems' performance [22]. Character recognition challenges were formerly solved by CNN models [24], however now days, CNN is frequently utilised to handle a variety of object identification issues. Here, the availability of huge amounts of training data—that is, the use of picture augmentation techniques—is the key component for CNN's success [15]. Furthermore, rather of using traditional handmade features, the CNN learns strong high-level features by fusing global appearances with local geometric features. This allows the CNN to function at a high level. Nevertheless, motion blur, low resolution, lighting, posture and expression change, occlusion by hair artefacts, and absence of intensity sounds are all present in the training image examples.

CNN seeks to address the difficulties that emerge while taking pictures in an unrestricted imaging environment, as well as the use of people-sentiment analysis, multimodal human-machine or computer interactions, and intelligent systems.

III.LITERATURE WORK

Hassain Saeed ; Ali Shouman ; Mais Elfar ; Mostafa Shabka in the title Near-field communication sensors and cloud-based smart restaurant management system, introduced an efficient and user-friendly Smart Restaurant Management System. This system will solve key problems faced by restaurants today through the use of technologies such as Mobile and Web applications, Internet of Things (IoT), Near-Field Communications (NFC) sensors, and cloud computing. Restaurants have many inefficiencies due to human limitations that can be resolved through automation and device-to-device communication. This Smart Restaurant Management System accomplishes this by providing two interfaces for the two types of users in restaurants; an Android mobile application for customers and a web application for restaurant staff members. The Android mobile application allows customers to have a seamless dining experience with features such as finding available parking spaces easier through internet-connected infrared proximity sensors in the parking lot, finding available tables at the restaurant easier through NFC sensors, ordering dishes through an interactive menu, and being able to pay the bill from their NFC equipped phones. The web application provides staff members benefits such as collecting data and statistics on the restaurant's performance in real time and automating the order placement system for waiters and cooks via IoT technology.

Florian Schroff, Dmitry Kalenichenko, James Philbin in the title, FaceNet: A Unified Embedding for Face Recognition and Clustering, implemented face verification and recognition efficiently at scale presents serious challenges to current approaches. In this paper we present a system, called FaceNet, that directly learns a mapping from face images to a compact Euclidean space where distances directly correspond to a measure of face similarity. Once this space has been produced, tasks such as face recognition, verification and clustering can be easily implemented using standard techniques with FaceNet embeddings as feature vectors. Their method uses a deep convolutional network trained to directly optimize the embedding itself, rather than an intermediate bottleneck layer as in previous deep learning approaches. To train, we use triplets of roughly aligned matching / non-matching face patches generated using a novel online triplet mining method. The benefit of our approach is much greater representational efficiency: we achieve state-of-the-art face recognition performance using only 128-bytes per face.

On the widely used Labeled Faces in the Wild (LFW) dataset, our system achieves a new record accuracy of 99.63%. On YouTube Faces DB it achieves 95.12%. Our system cuts the error rate in comparison to the best published result by 30% on both datasets. Also introduce the concept of harmonic embeddings, and a harmonic triplet loss, which describe different versions of face embeddings (produced by different networks) that are compatible to each other and allow for direct comparison between each other.

Andrew G. Howard, Menglong Zhu, Bo Chen , in the title, MobileNets: Efficient Convolutional Neural Networks for Mobile Vision Applications, presented a class of efficient models called MobileNets for mobile and embedded vision applications. MobileNets are based on a streamlined architecture that uses depth-wise separable convolutions to build light weight deep neural networks. We introduce two simple global hyper-parameters that efficiently trade off between latency and accuracy. These hyper-parameters allow the model builder to choose the right sized model for their application based on the constraints of the problem. We present extensive experiments on resource and accuracy tradeoffs and show strong performance compared to other popular models on ImageNet classification. We then demonstrate the effectiveness of MobileNets across a wide range of applications and use cases including object detection, finegrain classification, face attributes and large scale geo-localization.

Janne Tommola, Pedram Ghazi, Bishwo Adhikari , in the title Real Time System for Facial Analysis, described the anatomy of a real-time facial analysis system. The system recognizes the age, gender and facial expression from users in appearing in front of the camera. All components are based on convolutional neural networks, whose accuracy we study on commonly used training and evaluation sets. A key contribution of the work is the description of the interplay between processing threads for frame grabbing, face detection and the three types of recognition. The python code for executing the system uses common libraries--keras/tensorflow, opencv and dlib--and is available for download.

Sergio Escalera ; Mercedes Torres Torres , in the title ,ChaLearn Looking at People and Faces of the World: Face AnalysisWorkshop and Challenge presented the 2016 ChaLearn Looking at People and Faces of the World Challenge and Workshop, which ran three competitions on the common theme of face analysis from still images. The first one, Looking at People, addressed age estimation, while the second and third competitions, Faces of the World, addressed accessory classification and smile and gender classification, respectively. We present two crowd-sourcing methodologies used to collect manual annotations. A custom-build application was used to collect and label data about the apparent age of people (as opposed to the real age). For the Faces of the World data, the citizen-science Zooniverse platform was used. This paper summarizes the three challenges and the data used, as well as the results achieved by the participants of the competitions.

IV.PROPOSED APPROACH

In order to solve this problem, all customers must be motivated to give a rating. This paper introduces an approach for a restaurant rating system that asks every customer for a rating after their visit to increase the number of ratings as much as possible. The scoring system is based on facial expression detection using pre-trained Convolutional neural network models. It allows the customer to rate the food as well as the environment by taking a picture of his face that reflects the corresponding feelings.

Convolutional neural network (CNN)

In order to solve this problem, all customers must be motivated to give a rating. This paper introduces an approach for a restaurant rating system that asks every customer for a rating after their visit to increase the number of ratings as much as possible. The scoring system is based on facial expression detection using pre-trained Convolutional neural network models. It allows the customer to rate the food as well as the environment by taking a picture of his face that reflects the corresponding feelings.

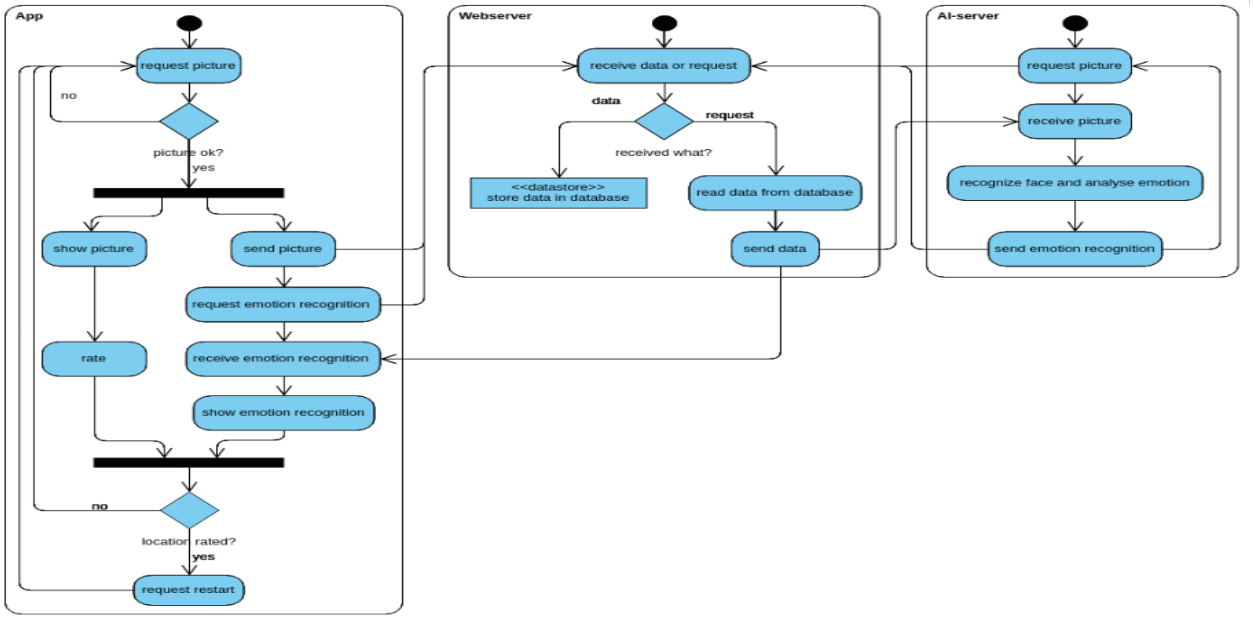


Fig.1 Data Flow Diagram of Restaurant rating system.

In order to solve this problem, all customers must be motivated to give a rating. This paper introduces an approach for a restaurant rating system that asks every customer for a rating after their visit to increase the number of ratings as much as possible. The scoring system is based on facial expression detection using pre-trained Convolutional neural network models. It allows the customer to rate the food as well as the environment by taking a picture of his face that reflects the corresponding feelings.

FER applications still have issues despite deep learning's strong feature learning capabilities. To prevent overfitting, deep neural networks first need a lot of training data. Popular deep learning neural networks that have shown the most promise in identification tasks are currently unable to be trained with the current facial expression databases. High inter-subject variances are also a result of individual differences in age, gender, ethnic background, and expressiveness . These variables are nonlinearly correlated with face expressions, which makes deep networks even more necessary to deal with the high variability and to acquire efficient expression-specific representations.

The face expression is just one of several angles from which human expressive behaviours in realistic applications must be encoded. Combining many models of a max-level framework, which can include supplementary information, can further boost resilience. Pure expression recognition
using only visible facial images can lead to promising outcomes. For instance, participants in the (AVEC) and EmotiW challenges used a range of fusion algorithms of multimodal identification and ranked the audio model as the second-most important component. This is also an intriguing
research field because of the great complementarity between facial expression other methods.

V.FACE DETECTION

Face expression detection is the most important phase for image classification since, the principal component analysis of nose, mouth, eyes are needed for the classification. Facial detection algorithms are classified based on knowledge feature template models. This paper uses object detection algorithm for facial expression recognition. In this algorithm, Haar Cascade Classifier is used based on is this input image is detected with the help of Haar features. 3.2 Facial Expression Recognition classification: The last step of facial expression recognition is to classify the given input image into any one of the basic emotions such as happy,sad,angry etc.Instead of feature extraction step, the feature classification step are independent. FER is performed in end to end way. Especially, the last layer is is added to the networks and to monitor the errors.

Fig 2 Detection of facial Expression

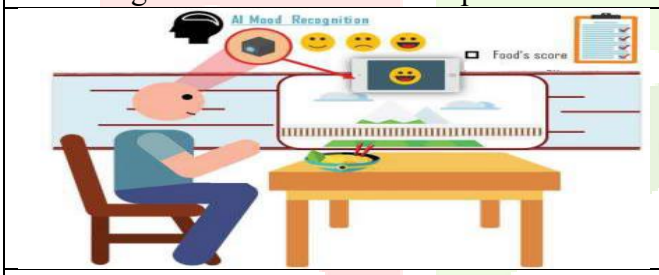


Fig 3 Uploading the image of the person

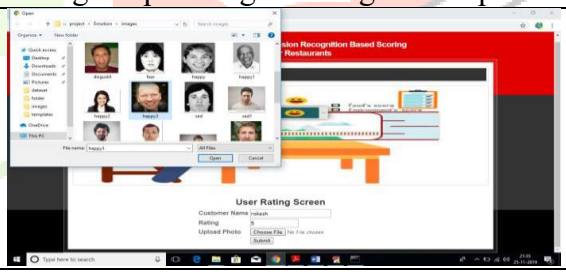


Fig 4 ratings given by various users

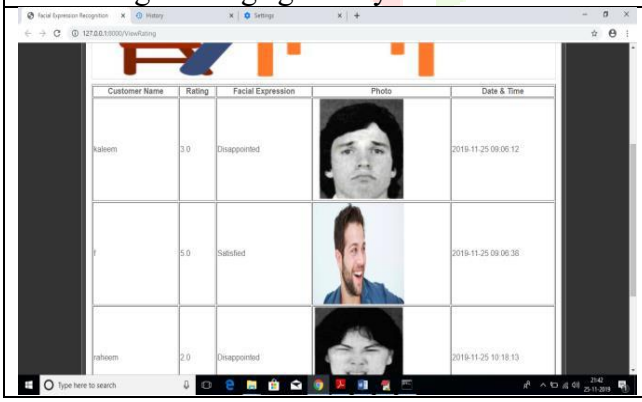
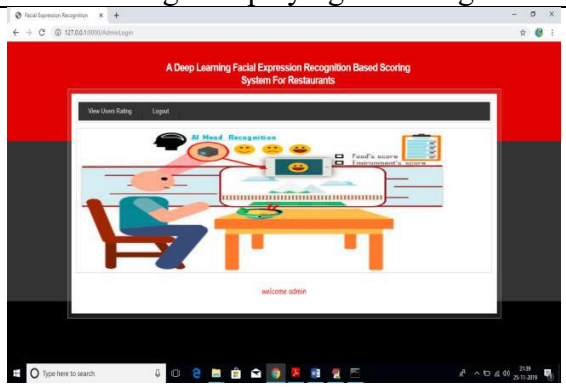


Fig 5 displaying the rating



We have design this as a web application using python DJANGO web server. This application can run on user browser where he can upload his photo with rating, uploaded photo will be sent to webserver where machine learning algorithm will be used to extract expression from photo and then saved result to MYSQL database. Another user called ‘admin’ can login to application and see all users visited to restaurant and can view all customer feedback with facial expression and photo. By seeing this result admin can understand whether customers are happy with their services

and foods or not. To run this project, install MYSQL and then create database by copying content from 'DB.txt' file and paste in MYSQL. Install python and then install DJANGO web server and deploy code on DJANGO. After deployment start server and run the code from browser.

VI.CONCLUSION

In this paper, a restaurant scoring system, which is based on facial expression recognition, is proposed. It is possible to get scoring platforms by making a direct request at the end of the visit to the customer. But there is only a rough insight, since only two ratings are requested. Because facial expression recognition is a modern technology that is used in a playful setting for the scoring system, the interest of the customer to give a rating is aroused.

REFERENCES

[1] Seyed-Farzad Mohammadi, Ghasem Saeedi-Anari, Cyrus Alinia, Elham Ashrafi, Ramin Daneshvar, and Alfred Sommer. Is screening for glaucoma necessary? a policy guide and analysis. Journal of ophthalmic & vision research, 9(1):3, 2014.

[2] Maya M Jeyaraman, Maya Jeyaraman, William G Hodge, Cindy Hutnik, John Costella, and Monali S Malvankar-Mehta. The effectiveness of teleglaucoma versus in-patient examination for glaucoma screening: A systematic review and meta-analysis. 2014.

[3] Ravindra Changala, "Evaluation and Analysis of Discovered Patterns Using Pattern Classification Methods in Text Mining" in ARPN Journal of Engineering and Applied Sciences, Volume 13, Issue 11, Pages 3706-3717 with ISSN:1819-6608 in June 2018.

[4] Rohit Varma, Paul P Lee, Ivan Goldberg, and Sameer Kotak. An assessment of the health and economic burdens of glaucoma. American journal of ophthalmology, 152(4):515–522, 2011.

[5] A. T. Broman H. A. Quigley. The number of people with glaucoma worldwide in 2010 and 2020. British Journal of Ophthalmology, 90:262—-267, 2006.

[6] J.C. Rueda, D.P. Lesmes, J. C. Parra, R. Urrea, J.J. Rey, L.A. Rodriguez, C.A. Wong, and V. Galvis. Valores de paquimetr´ıa en personas sanas y con glaucoma en una poblaci´on Colombiana. MedUNAB, 10(2):81–85, 2007.

[7] Ministerio de Salud y Protecci´on Social. Analysis de situacion de salud visual en Colombia, 2016.

[8] Ravindra Changala, "Secured Activity Based Authentication System" in " in Journal of innovations in computer science and engineering (JICSE), Volume 6, Issue 1,Pages 1-4, September 2016.ISSN: 2455-3506.

[9] S. S. Kanse and D. M. Yadav. Retinal fundus image for glaucoma detection: A review and study. Journal of Intelligent Systems, 28(1):43– 56, 2017.

[10] S. Nawaldgi. Review of automated glaucoma detection techniques. In 2016 International Conference on Wireless Communications, Signal Processing and Networking (WiSPNET), pages 1435–1438, 2016.

[11] Ravindra Changala, "Online Attendance using PCA Based Image Face Recognition" to publish in "International Journal for Research in Applied Science and Engineering Technology (*IJRASET*)" with Impact Factor 1.241, ISSN: 2321-9653,Volume 2, Issue XII, December 2014.

[12] C. Anusorn, W. Kongprawechnon, T. Kondo, S. Sintuwong, and K. Tungpimolrut. Image processing techniques for glaucoma detection using the cup-to-disc ratio. Science and Technology Asia, 18(1):22–34, 2013.

[13] C. Dhumane and S.B. Patil. Automated glaucoma detection using cup to disc ratio. International Journal of Innovative Research in Science, Engineering and Technology, 4(7):5209–5216, 2015.

[14] Ravindra Changala, Framework for Virtualized Network Functions (VNFs) in Cloud of Things Based on Network Traffic Services, International Journal on Recent and Innovation Trends in Computing and Communication, ISSN: 2321-8169 Volume 11, Issue 11s, August 2023.

[15] Ravindra Changala, "Optimization of Irrigation and Herbicides Using Artificial Intelligence in Agriculture", International Journal of Intelligent Systems and Applications in Engineering, Volume 11, Issue 3), July 2023.

[16] Ravindra Changala ,"Development of Predictive Model for Medical Domains to Predict Chronic Diseases (Diabetes) Using Machine Learning Algorithms And Classification Techniques", ARPN Journal of Engineering and Applied Sciences, Volume 14, Issue 6, 2019.

[17] J. Ayub, J. Ahmad, J. Muhammad, L. Aziz, S. Ayub, U. Akram, and I. Basit. Glaucoma detection through optic disc and cup segmentation using K-mean clustering. In 2016 International Conference on Computing, Electronic and Electrical Engineering (ICE Cube), pages 143–147, 2016.

[18]Ravindra Changala, "A Generalized Association Rule Mining Framework for Pattern Discovery" published in Ravindra Changala et al, / (IJCSIT) International Journal of Computer Science and Information Technologies, Impact Factor 2.93 in Vol. 5 (4) , 2014, pp:5659-5662, ISSN: 0975-9646,August 2014.

[19] S.M. Nikam and C.Y. Patil. Glaucoma detection from fundus images using matlab gui. In 2017 3rd International Conference on Advances in Computing, Communication & Automation (ICACCA)(Fall), pages 1–4, 2017.

[20] Jyotiprava Dash and Nilamani Bhoi. A thresholding based technique to extract retinal blood vessels from fundus images. Future Computing and Informatics Journal, 2(2):103–109, 2017.

[21] Jose Abel De La Fuente-Arriaga, Edgardo M Felipe-River´on, and Eduardo Gardu˜no-Calder´on. Application of vascular bundle displacement in the optic disc for glaucoma detection using fundus images. Computers in biology and medicine, 47:27–35, 2014.