A Deep Learning Facial Expression Recognition based Scoring System for Restaurants

Wan-Jung Chang[1,4*], Miriam Schmelzer[2,4], Florian Kopp[3,4], Chia-Hao Hsu[1,4],
Jian-Ping Su[1,4], Liang-Bi Chen[1,4], and Ming-Che Chen[4]

[1]Department of Electronic Engineering, Southern Taiwan University of Science and Technology, Tainan, Taiwan
[2]Department of Computer Science, University of Applied Sciences Constance (HTWG), Constance (Konstanz), Germany
[3]Department of Electrical Engineering, University of Applied Sciences Constance (HTWG), Constance (Konstanz), Germany
[4]Artificial Intelligence over Internet of Things Applied Research Center (AIoT Center),
Southern Taiwan University of Science and Technology, Tainan, Taiwan
{*allenchang, liangbichen, jerryhata}@stust.edu.tw; {miriam.schmelzer, florian.kopp}@htwg-konstanz.de

***Abstract*—Recently, the popularity of automated and unmanned restaurants has increased. Due to the absence of staff, there is no direct perception of the customers' impressions in order to find out what their experiences with the restaurant concept are like. For this purpose, this paper presents a rating system based on facial expression recognition with pre-trained convolutional neural network (CNN) models. It is composed of an Android mobile application, a web server, and a pre-trained AI-server. Both the food and the environment are supposed to be rated. Currently, three expressions (satisfied, neutral and disappointed) are provided by the scoring system.**

Keywords—automated restaurants; deep learning; facial recognition; restaurant scoring; unmanned restaurants; Internet of things (IoT)

I. INTRODUCTION

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 trend, the popularity of automated and unmanned restaurants continues to grow. Particularly in Japan and Taiwan, the number of automated or unmanned restaurants that run successfully is growing. In Fig. 1 examples of two restaurants of this type are shown.

Saeed *et al.* [1] proposed a cloud-based smart restaurant management system, which is composed of an Android mobile application for customers and a web application for restaurant staff members. The mobile application provides the customer among other things an interactive menu for ordering dishes and paying their bill.

As there is no staff available in unmanned restaurants, it is difficult for the restaurant management to estimate how the concept and the food is experienced by the customers. In addition, it is not possible to estimate the age and gender composition of customers to successfully perform targeted marketing. Existing rating systems, such as Google and TripAdvisor, only partially solve this problem, as they only cover a part of the customers' opinions. These rating systems are only used by a subset of the customers who rate the

(a)

(b)

Fig. 1. Two automated/unmanned restaurants examples. (a) An automated restaurant opearted in Tokyo, Japan. (b) A unmanned restaurant opearted in Kaohsiung, Taiwan. (Photographs courtesy of Internet)

This work is supported in part by the Ministry of Science and Technology (MoST), Taiwan, under grant, MOST 107-2637-E-218-003. **Corresponding author: Prof. Wan-Jung Chang.*

restaurant on independent rating platforms on their own initiative. This applies mainly to customers who experience their visit as very positive or negative.

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. It runs within the framework of an Android mobile application running on a mobile device that is mostly used in unmanned restaurants anyway as seen for example in [1]. The scoring system is based on facial expression detection using pre-trained convolutional neural network (CNN) 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.

Compared to text-based rating systems, there is much less information and no individual experience reports collected. However, this simple, fast and playful rating system should give a wider range of opinions about the experiences of the customers with the restaurant concept. In addition, statistics about the estimated age and gender composition of the customers can be generated.

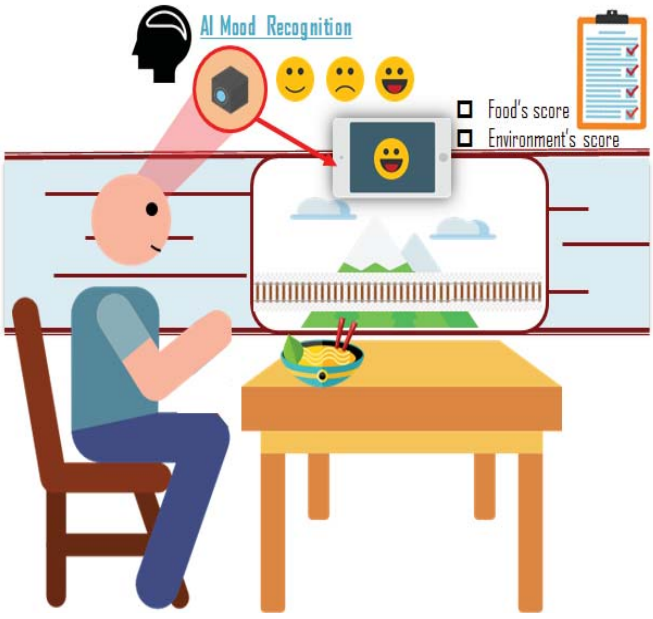


Fig. 2. Application scenario of the restaurant scoring system.

II. Architecture and Design

Fig. 2 shows an application scenario of the proposed scoring system. After the customer has finished eating and paid, he could be asked to rate the food as well as the environment on the same tablet that has already been used for ordering the food and payment.

The proposed restaurant scoring system is composed of an Android mobile application, a web server and a pre-trained AI server. The mobile application forms the interface through which the customer interacts with the scoring system. The facial expression recognition is performed on the AI server. The mobile application and the AI server do not communicate directly with each other. Instead, the entire data transfer between these two runs via a web server on which the data is stored in a database.

In Fig. 4, a flowchart is shown, which represents the execution steps of the scoring system. It is divided into three activities, each representing the flow of one of the three components (mobile application, web server and AI server) of the scoring system. Fig. 5 depicts the corresponding sequence diagram. It illustrates the communication sequences between the system components including the customer. The design of all three components as well as their execution steps are described below.

A. Android mobile application

Fig. 3 shows screenshots of the mobile application. On the first in Figure 3a, the customer is asked to take a picture of his face, in which the expression should rate the food. When the picture is accepted by the customer, it is uploaded to the web server. The app directly displays the next page, which is shown in figure 5b. On this page you can see the customer's rating image, under which the result of the AI server's facial expression recognition is displayed. This result is downloaded from the web server as soon as there is one available. This is done through polling the web server. As long as no result of the facial expression detection is available, a hint is displayed that the recognition is being waited for.

In order to intercept potential incorrect detections of the facial expression detection algorithm and thus to get statistics about its accuracy, the customer additionally is asked to rate the food manually in this experimental version. Presently, there are three different facial expressions (satisfied ☺, neutral 😐 and disappointed ☹) provided for the rating. By clicking the "Next"-button, the manual rating is uploaded to the web server and the customer is redirected to the next page, where he is asked to rate the environment. The procedure is the same as for the food rating. The corresponding screenshots are shown in Fig. 3(c) and Fig. 3(d).

B. Web Server

A web server was chosen over a direct communication between the mobile application and the AI server, as this makes it easier to replace and adapt individual components of the system. PHP and MySQL are used on the web server side. An API was written in PHP scripting language, which serves as an interface to both, the mobile application and the AI server. It manages the reading and writing of data from and to the MySQL database. In the database table the image URL and name, the customer and the AI server rating, the customer's estimated age and gender as well as the time and date of the ratings are stored.

C. AI Server

The customers rating image is downloaded from the web server as soon as it is available. This is done by polling the web server. In order to achieve the facial expression detection, we combine the FaceNet [2] with the mobileNet [3]. The FaceNet

is a 3-layer CNN-based approach, which can detect face locations on an image. After that, the mobileNet is used in the same way as in [4] for the facial expression recognition.

It is a light weight deep neural network for mobile and embedded vision applications. It was chosen because it can also achieve fast and precise results on low-cost hardware systems such as the used Nvidia Jetson TX2. It is able to detect seven different facial expressions.

For this use case it is modified so that these expressions are mapped to the three expressions used in our scoring system. The negative expressions sad, fear, disgust and anger are mapped to disappointed, happy and surprised to satisfied and neutral remains neutral. As soon as a result of the facial expression recognition is available, it is uploaded to the web server.

The AI server runs on an Nvidia Jetson TX2. This is an embedded AI-computing device with small form factor and low energy consumption. It is optimized for AI applications and power efficient.

III. Conclusion

In this paper, a restaurant scoring system, which is based on facial expression recognition, is proposed. It is possible to get a wider range of customer opinions compared to independent 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.

In a next step, the system could be combined with existing text-based rating platforms like Google rating to bring together the advantages of both systems. A further development could lead to a system where the customer can rate touchless in the restaurant. For this, it must be ensured that the accuracy of the facial expression recognition is high enough. It is also an idea to extend the image-based rating system with a speech recognition feature. The customer could express his opinion and impressions verbally or make suggestions for improvement like it is already done with Google ratings.

It is planned to extend the system with a web application that will enable the restaurant management to get a quick graphical overview and easy insight into the statistics. The aim is to integrate the proposed restaurant rating system into existing unmanned restaurants.

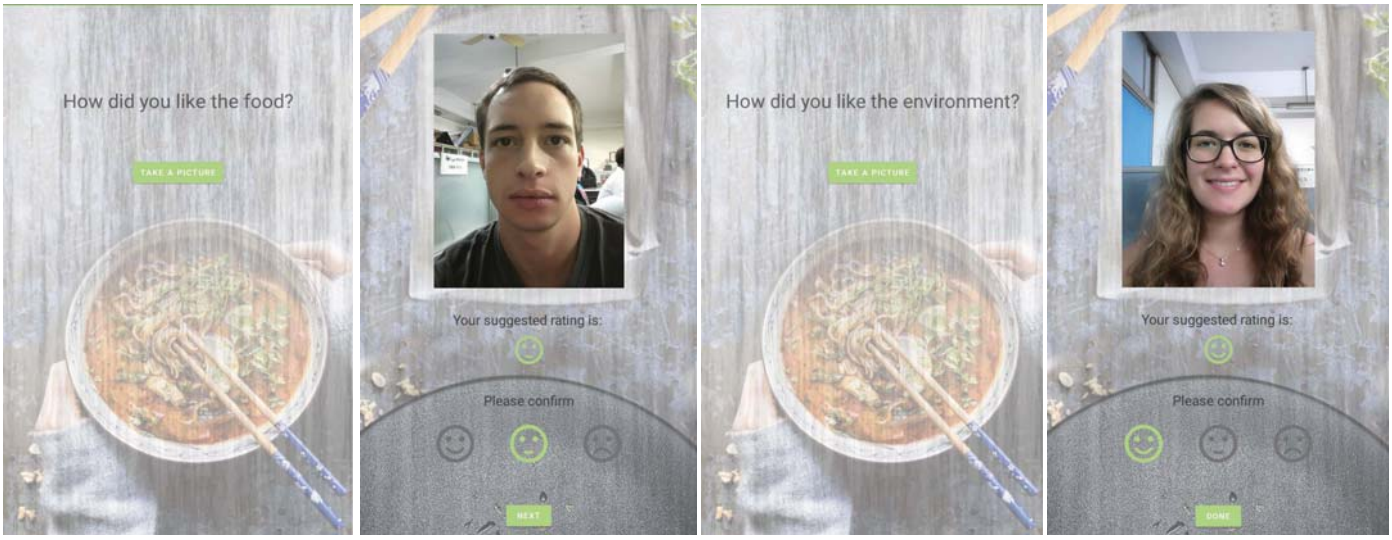
References

[1] 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.

[2] Florian Schroff, Dmitry Kalenichenko, and James Philbin, "FaceNet: a unified embedding for face recognition and clustering," n *Proceedings of the 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 815-823, 2015.

[3] 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

[4] Janne Tommola, Pedram Ghazi, Bishwo Adhikara, and Heikki Huttunen, "Real time system for facial analysis," in *Proceedings of the 7th European Workshop on Visual Information Processing (EUVIP'18)*, 2018.



(a) (b) (c) (d)

Fig. 3. Screenshots of the Android mobile application. (a) Request to rate the food. (b) Mood recognition and rating of the food. (c) Request to rate the environment. (d) Mood recognition and rating of the environment.

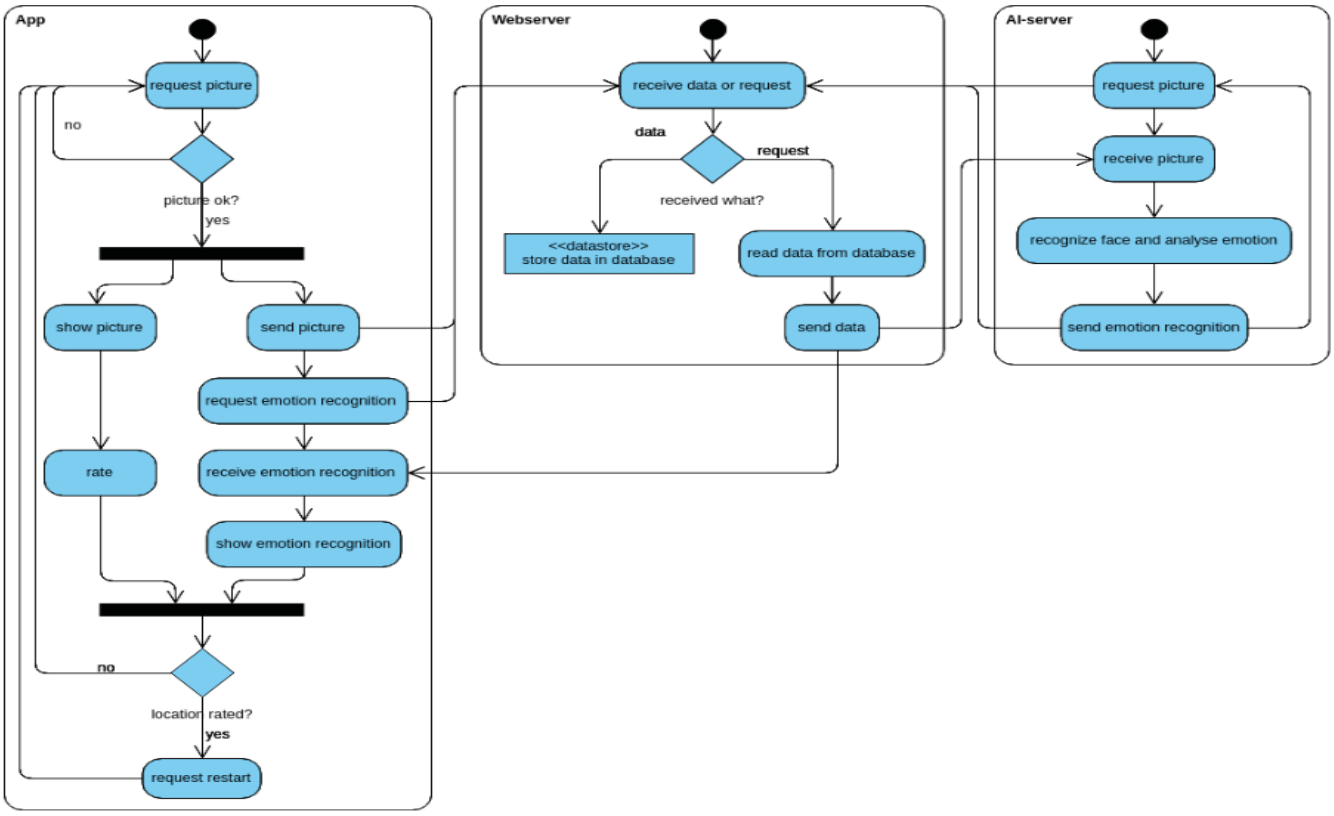


Fig. 4. Flowchart of the restaurant scoring system

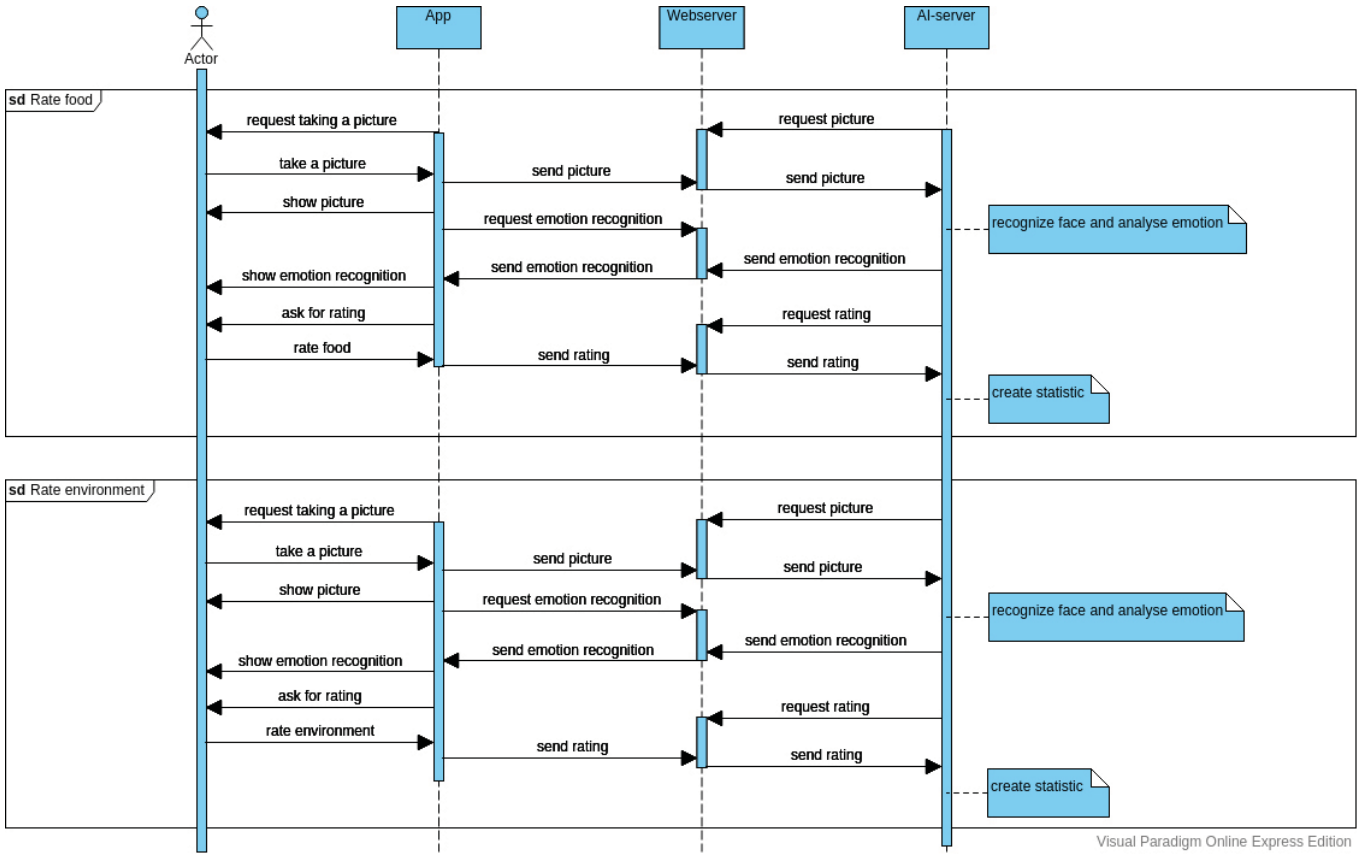


Fig. 5. Sequence diagram for the restaurant scoring system