ICIC Express Letters
Part B: Applications
Volume **11**, Number **2**, February **2020**

pp. 173–179

CLOUD-BASED FACIAL EXPRESSION RECOGNITION SYSTEM FOR CUSTOMER SATISFACTION IN DISTRIBUTION SECTORS

Jiyoon Lee[1] and Wonil Hwang[1,2,*]

[1]Department of IT Logistics and Distribution
[2]Department of Industrial and Information Systems Engineering
Soongsil University
369 Sangdo-ro, Dongjak-gu, Seoul 06978, Korea
[*]Corresponding author: wonil@ssu.ac.kr

Received July 2019; accepted October 2019

Abstract. *Recently, various studies have been conducted to improve customer satisfaction in the distribution process. In this matter, it is essential how to measure the customer satisfaction, but the traditional methods, like the survey method, are still widely used in the distribution process. The survey method can, however, be expensive, in terms of man power and time, and it is difficult to grasp customer complaints in real time. In this paper, we propose a cloud-based system architecture to investigate customer satisfaction using face expression recognition system based on artificial intelligence. The proposed system architecture was implemented based on the cloud so that heterogeneous clients could access easily and we validated the proposed system with internal data for cross-validation and with external data for showing how the proposed system worked in real-world situations. In conclusion, this study showed that we could successfully apply the method of facial expression recognition for evaluating customer's satisfaction, and the proposed system architecture worked well in the simulated situations like a restaurant.*
Keywords: Customer satisfaction, Facial expression recognition, Facial recognition module, Deep-learning module, Facial expression recognition cloud server

1. **Introduction.** Customer satisfaction is becoming more important as consumers' consumption patterns change and high quality services are required [1,2]. Companies pursue consumer-oriented management to increase re-purchasing customers, reduce advertising and marketing costs, and gain advertising effectiveness through satisfying customers [3,4]. These changes have made customer feedback one of the important factors that determine the durability of the company. In order to improve customer's satisfaction, companies are implementing various satisfaction survey methods. However, traditional methods such as paper, mobile and web-based passive surveys, are time-consuming, costly, and confusing because of the false answers of some customers [5]. For this reason, recently, a customer satisfaction survey method based on an automated facial recognition system has attracted attention in order to investigate customer satisfaction [6].

The facial recognition system is a field that has recently attracted attention due to the development of big data analysis and deep learning technology, and there are methods of analyzing images using algorithms such as CNN (Convolutional Neural Network) [7,8] and YOLO (You Only Look Once) [9,10]. Among existing facial expression recognition systems, there is an emotional expression system which analyzes human facial expressions, extracts their features, and recognizes four basic emotions, namely joy, sadness, anger and surprise [7]. The Principal Component Analysis (PCA) method is used to convert high-dimensional image feature data into low-dimensional feature data and then, through the linear discriminant analysis method, the efficient feature vectors are extracted to recognize the emotions. Thus, the method of expressing the emotions by applying this

DOI: 10.24507/icicelb.11.02.173

procedure to facial data is one of the facial expression systems. In the meanwhile, there is a system, which recognizes facial expressions in images using multiple deep network learning. An example is a classification module with an ensemble of multi-symbol Convolutional Neural Networks (CNN), including a face detection module based on an ensemble of three advanced face detectors [11]. Yu and Zhang [11] fine-tuned the pre-trained models with the SFEW 2.0 training set and studied the methods that minimized log-likelihood loss and hinge loss with the ensemble weights of network responses to combine multiple CNN models. In addition, there is a real-time face recognition and emotion recognition system. In the study of Ju et al. [12], the system automatically detected the front of the video stream and interacted in real-time by reading the seven dimensions, namely neutrality, anger, disgust, fear, joy, sadness and surprise. In the prior research, however, there were few studies, which focused on recognizing customer's satisfaction based on the real-time analysis of facial expressions.

The existing facial expression recognition systems that have been discussed above utilize various artificial intelligence methodologies, such as machine learning and deep learning, to process facial expression recognition and show good performance as data that can be used for learning increase. However, it is necessary to implement a system capable of receiving image data from heterogeneous platform devices located in the circulation phenomenon as well as a facial recognition system in order to recognize the customer's expression in the distribution process and to predict the satisfaction level. In this study, we propose a cloud-based system architecture that can automatically investigate customer's satisfaction using an artificial intelligence-based facial recognition system. This study shows how to construct the proposed system architecture, and validates it with a test data set for cross-validation and with an experimental data set for external validation. The composition of this paper is as follows. In Section 2, we describe the structure, components, and data of the facial recognition system architecture proposed in this study. Section 3 presents the results of the experiment, and we conclude this paper with discussion in the final section.

2. Methods.

In this section, the architecture of the proposed facial recognition system, the functions of each of components and the implementation process are explained first. Second, the image data used for learning, testing and validation in the proposed system is described.

2.1. Architecture of proposed system.

The proposed system architecture is designed to recognize facial expressions from customer face images and provide information about customer satisfaction in real time. The names and roles of each component are as follows. Face Recognition Module (FRM) is a module that can recognize a customer's face and extract it as an image, extracting the customer's face image from the webcam during the distribution process. Deep-Learning Module (DLM) is a module that analyzes facial images extracted from FRM, recognizes facial expressions, and returns results. The final component, Facial Expression Recognition Cloud Server (FERCS), acts as a relay for each component, delivering the customer face images extracted from the FRM to the DLM, and providing the FRM with the satisfaction information recognized by the DLM.

The detailed execution flow of the proposed system architecture is as follows. Figure 1 shows how each component works together. Located on the left-hand side of Figure 1, FRM is a module that recognizes the customer's face through the devices, such as POS and CCTV, during distribution process, and extracts face images. Images extracted from FRM are preprocessed and finally processed as 100×100 grayscale images. Preprocessed images are transferred from FRM to FERCS. FERCS is a module that receives images from FRM and finally provides facial expression recognition results to FRM. In doing so, after storing the received customer image in the cloud, FERCS calls DLM which can return facial expression recognition results. DLM is a module that implements a learning

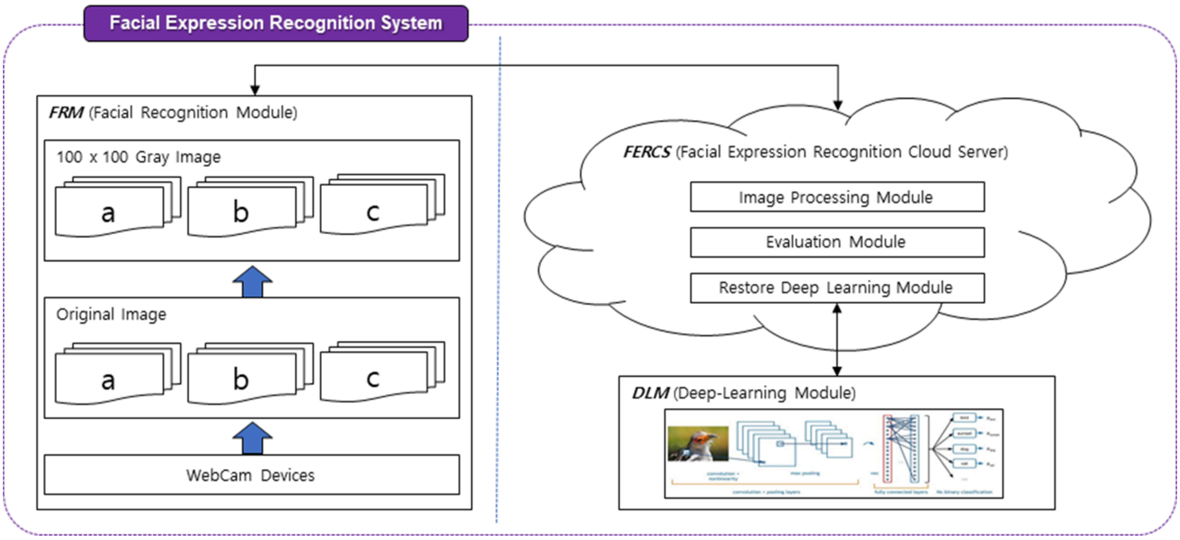

Figure 1. Architecture of facial expression recognition system

process using a deep learning library and returns facial expression recognition results from images. DLM analyzes the images stored in FERCS and provides FERCS with facial expression recognition results. FERCS returns the recognition results of the DLM to the FRM, and finally FRM outputs the recognized results to the screen.

2.2. **System architecture configuration.** Unlike previous studies, the proposed system architecture is designed to contribute to improving customer satisfaction ultimately by recognizing automatically the satisfaction of customers in the distribution process, and in a cloud-based manner, allowing heterogeneous clients to request image upload and facial recognition results using a web-based communication interface. In addition, the artificial intelligence model can be trained with the latest image data uploaded by automated scheduling, which can improve the accuracy of satisfaction analysis. The proposed system architecture consists of FRM (Facial Recognition Module), DLM (Deep-Learning Module) and FERCS (Facial Expression Recognition Cloud Server) introduced in Section 2.1. Details of each component are as follows.

2.2.1. *FRM (Facial Recognition Module).* FRM is a module that operates on devices such as POS and CCTV, installed in a place where customers can directly face the devices during distribution process. FRM recognizes the customer's face image based on the OpenCV library. The recognized face image is transmitted to the FERCS implemented on the cloud, and the image preprocessing is performed for high-speed transmission of data. The image preprocessing includes image size reduction and grayscale processing. The preprocessed image is transferred to the FERCS, which provides facial expression recognition results. FRM displays the recognition results received from the FERCS on the screen so that the customer's satisfaction can be confirmed.

2.2.2. *DLM (Deep-Learning Module).* DLM is a module that recognizes the current facial expression from the face image of the customer and is implemented based on the tensor flow which is a deep learning library. DLM is trained to recognize three facial expressions such as expressionless, smiley and frowned expressions in the face image of the customer. DLM is called by FERCS and implemented based on cloud, and it can continuously learn existing data and new data according to schedule and setting. All clients connected to the FERCS through the DLM can receive the latest facial expression recognition results.

2.2.3. *FERCS (Facial Expression Recognition Cloud Server).* FERCS is a cloud-based web server that relays FRM and DLM. FERCS provides a web service-based interface that can execute the deep learning module. When a request is received, FERCS executes the deep learning module and returns the execution results. In the FERCS, there are three types of modules: ‘Image Processing Module’ for processing the images transmitted for facial expression recognition, ‘Restore Deep Learning Module’ for performing facial expression recognition using the deep learning model, and ‘Evaluation Module’ for evaluating facial expression recognition results. In the rest of this paper, we show how to train and test the data based on ‘Evaluation Module’ to validate the proposed system.

2.3. **Learning and tests.** In the proposed facial expression recognition system, learning and testing involve two types of data, namely internal data (i.e., training and test image data for cross-validation) and external data (i.e., experimental video clip data for external validation). In this section, data and experimental method are described.

2.3.1. *Image data for learning and testing.* For learning and testing of the deep learning model, we used the K-FACE data from NIA [13]. The data provided by the NIA included three facial expressions, namely expressionlessness, laughter and frown, from 200 people. The image pre-processing module in the proposed system processed the data in 100 $\times$ 100 grayscale. Examples of the preprocessed image data are shown in Figure 2. The expressionless face is a state in which the emotions are not revealed on the face, and the laughing and frowning expressions are generally found around the mouth and eyes. For cross-validation, 70% of the total image data was used as learning data, and 30% was used as test data. Data in the database for cross-validation is called as ‘internal data’, whereas data collected from simulated experiments on real-world situations is called ‘external data’ in this study. The external data is image data from video clips of about 3 seconds, assuming the customer’s situation at a restaurant and containing a satisfactory or an unsatisfactory appearance of 15 participants. The external data has a satisfaction score as a label of the image data, which was evaluated by the participants in the experiments on a

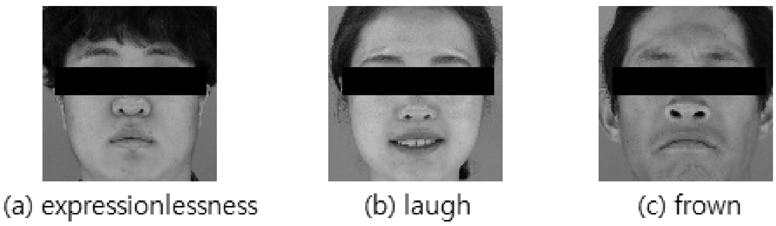

Figure 2. Examples of data for learning

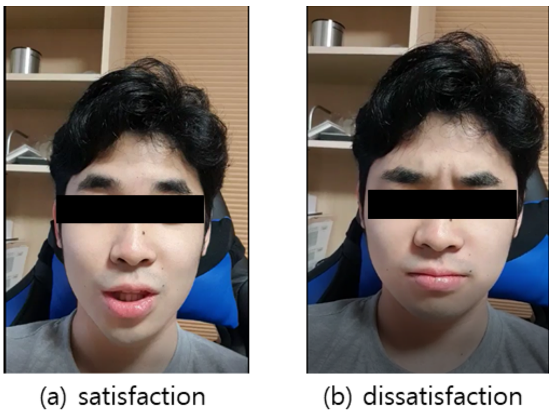

Figure 3. Examples of external data

seven-point scale. Examples of the external image data for satisfaction and dissatisfaction are shown in Figure 3.

2.3.2. *Method of tests.* The tests were conducted with two types of data, namely internal data and external data. The test with internal data shows how well DLM is learning, and the test with external data shows how well the proposed system works on real-world situations by extracting the images from the video clips, analyzing the satisfaction of each image and providing the results of satisfaction or dissatisfaction. The results were displayed as a percentage of the accuracy compared to the actual satisfaction scores.

2.4. **Technical contribution of proposed system architecture.** The proposed system architecture uses a cloud-based structure to provide information about customer satisfaction through facial expression recognition. The components of the system FRM, DLM, and FERCS, use a loosely coupled structure with each other, so this feature allows users to flexibly cope with changes in the technical parts used for image analysis. FERCS also makes it easy to add additional methods for image analysis.

3. **Results.** From the results of internal validation (i.e., tests with internal data), it is shown how well DLM in the facial expression recognition system learned, and the results of external validation (i.e., tests with external data) show how well the proposed system recognized the customer's facial expression and confirmed the satisfaction in this section.

3.1. **Internal validation.** Figure 4 shows the DLM learning based on the tensor flow based on the image data provided by the NIA identified in Section 2.3.1. The DLM extracts only the part of the image data that corresponds to the face, performs grayscale transformation and feature extraction, and then learns using the CNN algorithm. Figure 4 shows the output of the DLM accuracy in scalar format provided by the tensor board. DLM learned until it showed 100% of accuracy with training data and it showed more than 80% of accuracy with test data.

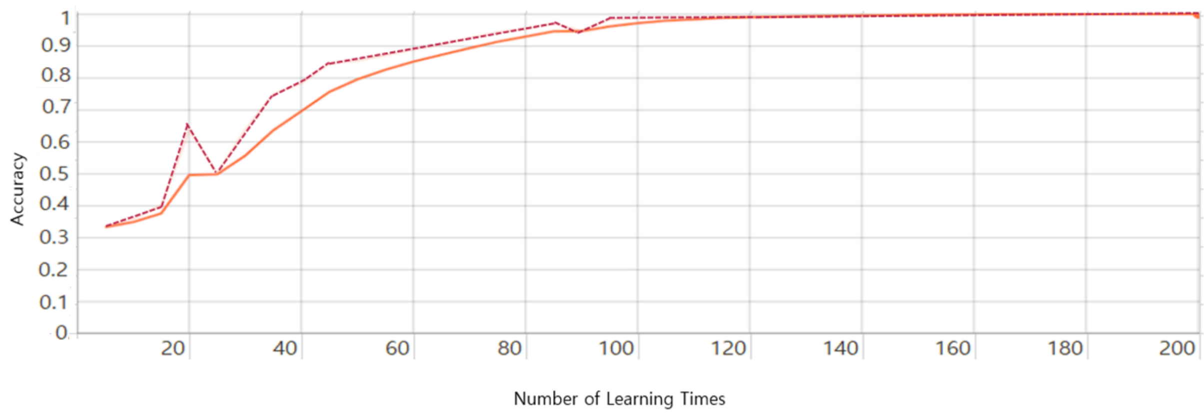

Figure 4. Result of DLM learning (dashed line: values in the raw data, solid line: values adjusted by smoothing)

3.2. **FRM (Facial Recognition Module) implementation.** Figure 5 shows the recognition result of the current facial expression after recognizing the face of the customer through FRM. FRM automatically recognizes customer's faces in the distribution sector (e.g., the customers' faces can be captured as video clips for 2-3 seconds at a restaurant), and delivers it to FERCS. FERCS delivers the requested image to the DLM to analyze the facial expression and predict whether the customer is satisfied or dissatisfied. The analyzed result is transmitted to FRM through FERCS. As shown in Figure 5, the expression recognition result is displayed on the screen with satisfaction or dissatisfaction.

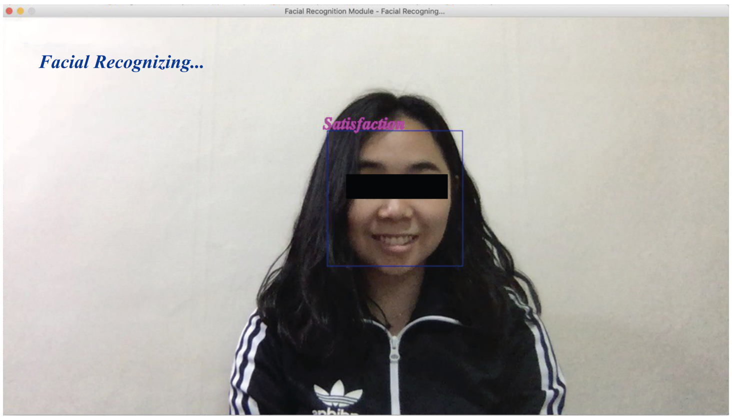

FIGURE 5. Screen shot of facial recognition module

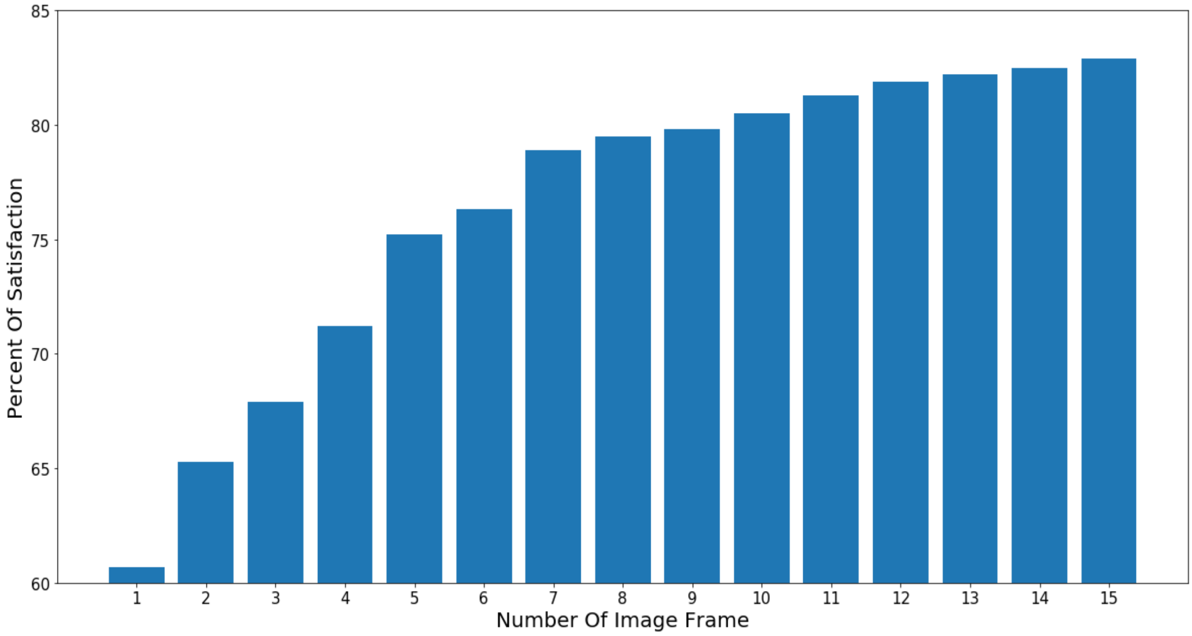

FIGURE 6. Satisfaction evaluation with image data used

3.3. **External validation.** For external validation, meaning how well the proposed system works on real-world situations, tests were conducted with external data as a form of video clip collected from the experiments. During the testing process, the evaluation module was implemented to improve the accuracy of evaluating whether a customer was satisfactory or not through face expression recognition. The evaluation module collected the recognition results from multiple images according to the attributes specified in the system when the customer was recognized as satisfied by receiving the facial images from FRM. Figure 6 shows the accuracy results of evaluating the customer's satisfaction when using the video clips collected in the experiments. The accuracy of evaluating the customer's satisfaction increases as the number of image frames used for facial expression recognition increases, as seen in Figure 6. Specifically the accuracy was lowest when only one frame of image was used for confirming the recognition results of the customer, and the accuracy gradually increased as the number of frames used for the recognition increased. Thus, it can be concluded that the proposed system can improve the recognition accuracy by controlling the properties of evaluation module, such as the number of frames used for facial expression recognition.

4. **Conclusions and Discussion.** In recent years, various studies have been conducted through artificial intelligence, and image analysis is one of them. In this paper, we propose a system architecture that can automate the existing passive satisfaction survey method by using image analysis method based on deep learning. Since the proposed system architecture is based on cloud, heterogeneous clients can easily obtain customer image recognition result using web service-based interface. From two types of tests, such as internal and external validation, the proposed system showed the acceptable level of accuracy (about 80% from cross-validation with internal test data) for recognizing the images of customers and it was concluded that the recognition accuracy could be improved as the number of frames of image used for facial expression recognition increased. Thus, it would be necessary to use multiple images, like video clips, for customer's facial expression recognition. In sum, this study showed that we could successfully apply the method of facial expression recognition for evaluating customer's satisfaction, and the proposed system architecture worked well in the simulated situations like a restaurant. For further study, we need to find optimal conditions for implementing the proposed system with sufficient data for verification.

Acknowledgment. This research was supported by the Ministry of Trade, Industry and Energy (MOTIE), Korea, through the Education Program for Creative and Industrial Convergence (Grant Number N0000717).

REFERENCES

[1] E. W. Anderson, C. Fornell and D. R. Lehmann, Customer satisfaction, market share, and profitability: Findings from Sweden, *Journal of Marketing*, vol.58, no.3, pp.53-66, 1994.

[2] K. Hubacek, D. Guan and A. Barua, Changing lifestyles and consumption patterns in developing countries: A scenario analysis for China and India, *Futures*, vol.39, no.9, pp.1084-1096, 2007.

[3] L. L. Bove and L. W. Johnson, A customer-service worker relationship model, *International Journal of Service Industry Management*, vol.11, no.5, pp.491-511, 2000.

[4] R. Hallowell, The relationships of customer satisfaction, customer loyalty, and profitability: An empirical study, *International Journal of Service Industry Management*, vol.7, no.4, pp.27-42, 1996.

[5] A. Gustafsson, M. D. Johnson and I. Roos, The effects of customer satisfaction, relationship commitment dimensions, and triggers on customer retention, *Journal of Marketing*, vol.69, no.4, pp.210-218, 2005.

[6] J.-H. Geng, Q.-S. Gao and S.-Y. Zhang, Community satisfaction evaluation system's research based on AHP and factor analysis, *Journal of System & Management*, vol.16, no.6, pp.673-677, 2007.

[7] M. S. Bartlett, G. Littlewort, I. Fasel and J. R. Movellan, Real time face detection and facial expression recognition: Development and applications to human computer interaction, *Proc. of the 2003 Conference on Computer Vision and Pattern Recognition Workshop*, Madison, Wisconsin, USA, 2003.

[8] Y. Huang, R. Wu, Y. Sun, W. Wang and X. Ding, Vehicle logo recognition system based on convolutional neural networks with a pretraining strategy, *IEEE Trans. Intelligent Transportation Systems*, vol.16, no.4, pp.1951-1960, 2015.

[9] H.-J. Jeong, K.-S. Park and Y.-G. Ha, Image preprocessing for efficient training of YOLO deep learning networks, *Proc. of 2018 IEEE International Conference on Big Data and Smart Computing*, Shanghai, China, pp.635-637, 2018.

[10] J. Tao, H. Wang, X. Zhang, X. Li and H. Yang, An object detection system based on YOLO in traffic scene, *Proc. of the 6th International Conference on Computer Science and Network Technology*, Dalian, China, pp.315-319, 2017.

[11] Z. Yu and C. Zhang, Image based static facial expression recognition with multiple deep network learning, *Proc. of the 2015 ACM on International Conference on Multimodal Interaction*, Seattle Washington, USA, pp.435-442, 2015.

[12] J. T. Ju, G. J. Park, G.-E. Go, H. C. Yang and G.-B. Sim, Emotion recognition and expression using facial expression, *Proc. of the Korean Institute of Intelligent Systems Conference*, pp.295-298, 2007.

[13] *National Institute on Aging (NIA)*, https://www.nia.or.kr/, Accessed on June 3rd, 2019.