



Developing a Facial Identification System Using Pre-Trained Model and Spoken Dialogue Agent

Chen, Sinan

Nakamura, Masahide

(Citation)

2022 International Balkan Conference on Communications and Networking (BalkanCom):62-67

(Issue Date)

2022-09-29

(Resource Type)

conference paper

(Version)

Accepted Manuscript

(Rights)

© 2022 IEEE. Personal use of this material is permitted. Permission from IEEE must be obtained for all other uses, in any current or future media, including reprinting/republishing this material for advertising or promotional purposes, creating new collective works, for resale or redistribution to servers or lists, or...

(URL)

<https://hdl.handle.net/20.500.14094/0100487656>



Developing a Facial Identification System Using Pre-Trained Model and Spoken Dialogue Agent

Sinan Chen^{1,2}, Masahide Nakamura^{1,3}

¹Kobe University, 1-1 Rokkodai-cho, Nada, Kobe, 657-8501, Japan

²Japan Society for the Promotion of Science, 5-3-1 Kojimachi, Chiyoda-ku, Tokyo 102-0083, Japan

³RIKEN Center for Advanced Intelligence Project, 1-4-1 Nihonbashi, Chuo-ku, Tokyo, 103-0027, Japan

Email: chensinan@ws.cs.kobe-u.ac.jp, masa-n@cs.kobe-u.ac.jp

Abstract—Our research group is currently studying and developing listening services using spoken dialogue agents and IoT technologies to assist the “mind” of the elderly at home. However, the user identification function, an essential part of the service, has not yet been realized. It is difficult to determine the identity of the person who interacts with the spoken dialogue agent. Although with the rapid development of the artificial intelligence field, various smart devices and services using deep learning have appeared in the face recognition technology, problems exist, including costs and computational resources to build and apply a recognition model. The purpose of this paper is to develop a facial identification system using the pre-trained model and spoken dialogue agent. Our key ideas include automatic training data generation by spoken dialogue between the user and the agent and the acquisition and comparison of facial features using a pre-trained model. In this way, our face identification system can be easier introduced and expected with only a general-purpose computer and a Web camera, without needing a conventional Internet connection and manual labeling of training data.

Index Terms—user identification, face recognition, pre-trained model, spoken dialogue, assistive technology

I. INTRODUCTION

As the world’s population ages and the shortage of nursing care personnel and facilities become more serious, the number of elderly people requiring nursing care in Japan who switch from institutions to home care has been increasing yearly. While the elderly are cared for in their familiar homes, they are also at risk of mild dementia and decreased communication skills due to increased loneliness caused by a lack of social interaction. Against this background, our research group has been researching and developing a voice interactive agent listening service. We are currently integrating Internet of Things (IoT) sensors, virtual agents (VA), and cloud technology, aiming to support the “mind” (e.g., How about your feeling today? asked by the VA) monitoring of elderly people at home [1] [2].

Our VA dialogue service links various microservices. Such as a function for reviewing dialogue logs as a personal diary, watching YouTube videos, information retrieval, and calendar event registration have been incorporated into agents, enabling control of various functions through spoken dialogue between an older adult and an agent. Also considered is a personalized health question function that uses a wearable health device to capture changes in an older adult’s features and link the measurement results to a spoken dialogue agent, and an

approach [3] that estimates the cause of health changes from dialogue logs.

However, the function of user identification, a crucial consistency of the service, has not yet been realized. It is difficult to determine the identity of the person with whom the spoken dialogue agent interacts. Various smart devices and services [4] using cloud computing and deep learning have been introduced to face recognition technology in line with the rapid development of artificial intelligence. However, the cost and computational resources required to build and apply recognition models remain an issue.

Recently, TensorFlow [5], an end-to-end open-source platform developed for machine learning, was introduced. Machine learning models are developed in JavaScript and used directly in a Web browser or Node.js. These are commonly referred to as TensorFlow.js pre-trained models, including models for image classification, hand pose detection, human pose detection, object detection, semantic segmentation, and human face recognition. These techniques feature real-time feature extraction from data and can be run offline.

The goal of this study is to develop a face identification system using pre-trained models and a spoken dialogue agent. The key ideas are automatic generation of training data through spoken dialogue between a user and an agent and acquisition and comparison of facial features using pre-trained models. In the proposed method, Step 1 is to load existing data and extract features. In Step 2, new training data and features are extracted. In Step 3, the existing data and the training data are compared. Finally, in Step 4, we perform agent interaction based on the identification results.

In the case studies, we discuss scalable value-added services based on three examples:

- Support for monitoring the daily rhythm of older adults with dementia at home
- Individual logging during agent interaction with multiple people
- Real-time home invasion estimation for older adults living alone

The discussion also shows that the proposed system has the advantages of improving user privacy issues, learning and prehension of data, and the ease of system implementation. On the other hand, the limitations of the proposed system are that it cannot detect face masks from raw images, it takes time

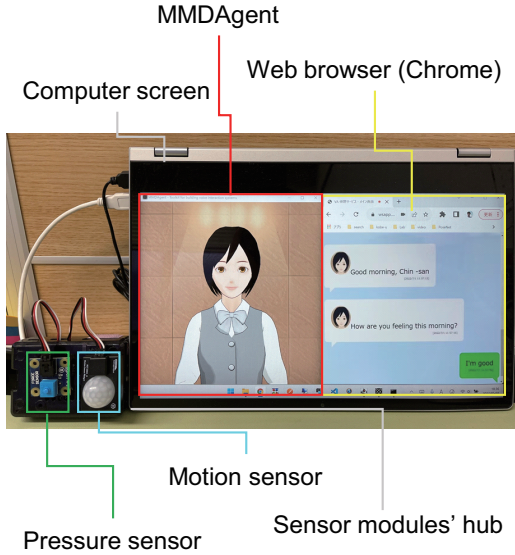


Fig. 1. Virtual agent listening service using purpose-general computer and sensor modules. Copyright 2009-2018 Nagoya Institute of Technology (MMDAgent Model “Mei”)

to load existing data, and it cannot determine the authenticity of the training data. In addition, we describe related work and identify future challenges.

II. RELATED WORK

Many studies have been conducted on face identification using cloud computing. Hu et al. [6] proposed a parallel matching mechanism and a cloud computing-based solution framework to resolve face images, control personal data access efficiently, and obtain personal identity information. They propose a parallel matching mechanism and a cloud computing-based solution framework to efficiently resolve face images, control personal data access, and obtain personal identity information. Another study [7] proposes an affine transformation consisting of permutation, diffusion, and shift transformations to protect facial privacy by using a novel and efficient scheme and randomness method to achieve privacy-preserving face recognition in the cloud. Unlike , these studies focus on the efficiency of face recognition and privacy issues. They are speculated to be used in public rather than private cases such as residences. In contrast, we believe that face identification systems for the elderly and family members at home should focus on personal adaptation (e.g., design of dialogue scenarios based on personal preferences [8], the realization of individualized services through face recognition behavior).

On the other hand, some studies on locally feasible face identification systems use edge computing. Kong et al. [9] proposed an edge computing-based mask (ECMask) identification framework to support public health precautions, consisting of three main stages: video It consists of three main stages: video restoration, face detection, and mask identification. Another study [10] develops a system named IOSTS, an intelligent

online attendance tracking system that can track attendance while using minimal bandwidth and maintaining user privacy. Although these studies differ from our approach, they share the same goals of minimizing local and network communications and improving user privacy issues. In particular, the ability to identify masks from faces can be expected to inform our future research directions.

III. PRELIMINARIES

A. Previous study: monitoring assistance for “mind” of elderly people at home

As the population ages and the shortage of nursing care personnel and facilities become more severe in developed countries, including Japan, Italy, Germany, and France, the number of older adults requiring nursing care who switch from institutions to home care is increasing year by year. While the elderly can live in the comfort of their own homes, they are also at risk of mild dementia and decreased communication skills due to increased loneliness caused by a lack of social interaction.

Our research group is developing a service that integrates IoT sensors, virtual agents (VA), and cloud technology to support the “mind” monitoring of elderly people at home. In this service system, a general-purpose computer is connected to sensor devices (i.e., motion and pressure sensors). A virtual agent (i.e., character material and speech synthesis using MMDAgent [11]) is linked to a Web browser (i.e., speech recognition using Web Speech API). The system is designed to spontaneously initiate spoken dialogue interactions with the agent only while the older adult is near the sensor device. The agent automatically responds with random responses, and the system aims to collect daily nuncupative logs (e.g., what were you doing this morning?) in the form of a “dialogue” regularly. Figure 1 shows the VA listening service using a general-purpose computer and sensors.

Specifically, we utilize Pub/Sub and Web Socket architectural techniques to record and store dialogue logs in real-time in a cloud database (e.g., MongoDB or Next Cloud). A function to look back on these logs as a personal diary implemented to support record-keeping and self-management by the elderly. For the elderly unfamiliar with the operation of smart devices, various microservices, such as video viewing, information retrieval, and calendar event registration, are embedded in the agent to control these functions through voice interaction is also realized. Moreover, we have proposed a personalized health question function to remind and elucidate the causes of health changes that the elderly have not noticed. It captures feature changes (e.g., activity level, heart rate, number of steps) with a wearable health device (e.g., Garmin wearable activity tracker). It links the measurement results with a voice dialogue agent. We also proposed an approach to infer the cause of health changes from dialogue logs. However, the function of user identification, a crucial consistency of the service, has not yet been realized. The problem exists that it is difficult to determine the identity of the person with whom a spoken dialogue agent interacts.

B. Technical challenges

To automate user identity determination, we focus on developing a user identification system using facial features. We consider that there are two technical challenges.

One is the issue of user privacy during face recognition. Many cloud services for face recognition, such as Microsoft Azure Face API, Google Cloud Vision API, and IBM Watson Visual Recognition API, have appeared in the form of Application Programming Interface (API) for face recognition. Although face recognition that fully uses powerful cloud computing resources can naturally provide highly accurate results, there are significant overheads in terms of usage fees and Internet communications when using face recognition in real-time. In addition, there is the issue of user privacy when sending raw time-series image data to a global network for extraction and analysis of user facial feature data.

On the other hand, the other is the time-consuming task of acquiring and training data. The most direct approach is deep learning (e.g., CNN, LSTM) to build a face identification model for each household. However, it requires a massive amount of computational resources and a large amount of pre-labeled training data. However, prior training data is not always available when installing and operating a face identification system provided by the developer at a user's residence. It takes time to obtain and train data individually for each user in each household. Several rounds of cooperation are required from users during the preparation phase of the system, and the effectiveness of training after the training data has been obtained is unknown.

C. TensorFlow.js pre-trained models

TensorFlow, an end-to-end open-source platform developed for machine learning, has recently emerged to make it easier for developers to create and deploy machine learning-powered applications. The TensorFlow library includes TensorFlow.js, a JavaScript library for machine learning. Machine learning models are developed in JavaScript and used directly in the Web browser or Node.js. These are generally called TensorFlow.js pre-trained models and include models involving three aspects: images, text, and general utilities. The image models include models for real-time image classification, hand pose detection, human pose estimation (used to home care assistance [12]), object detection, semantic segmentation, and human face recognition. Figure 2 shows an example of a browser canvas depiction using detection results from a pre-trained model. These techniques are characterized by the ability to extract features from data in real-time and can be performed offline.

D. Face-api.js

Face-api.js [13] is a JavaScript library for one of the TensorFlow.js pre-trained models and provides functions for face detection, face landmark detection, face identification, and age, gender, and emotion estimation for 2D images using Node.js or a Web browser. It provides functions for face detection, face landmark detection, face identification, age,

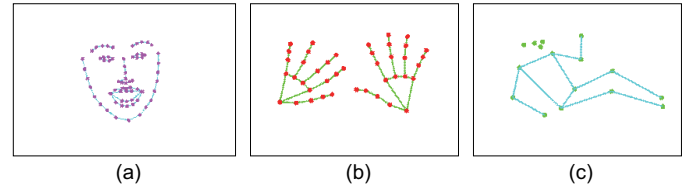


Fig. 2. Examples of drawing results in the Web browser using TensorFlow.js pre-trained models: (a) Face detection. (b) Hand pose recognition. (c) Human pose estimation.

TABLE I
FUNCTIONS, PRE-TRAINED MODELS, AND FEATURES IN THE FACE-API.JS LIBRARY.

Functions	Pre-Trained Models	Features
Face detection	ssdMobilenetv1	Get the facial bounding box with high accuracy
	tinyFaceDetector	Get the facial bounding box with fast speed
Face landmark detection	faceLandmark68Net	Get 68 face feature points
Face identification	faceRecognitionNet	Get the similarity between two faces
Face emotion estimation	faceExpressionNet	Get one of facial expressions including natural, happy, sad, angry, fearful, disgusted, surprised
Face age gender estimation	ageGenderNet	Get age and gender

gender, and emotion estimation. Table I lists the features, pre-trained models, and features of face-api.js. These pre-trained models have single-person and multiple-person modes, each of which is called by a separate function. In addition, these functions can be used on video and real-time video streams and 2D images. The speed of execution of these functions is typically 30 fps.

IV. PROPOSED METHOD

A. Goal and key idea

The goal of this research is to propose a method for developing a feasible face identification system that is local and does not require time for training data preparation in order to improve the technical challenges of Section III-B. Our key idea is to integrate pre-trained models and spoken dialogue agents. Even if the number of training data is zero, we can accumulate training data as the user interacts with the agent, and we expect to improve the recognition system. In this way, it will also allow us to promote the use of face identification systems, which are lighter and easier to implement, as an essential part of VA listening services.

B. Overall architecture

Figure 3 shows the overall architecture of the proposed method in this study. The proposed method uses a general-purpose computer, its built-in camera, and its built-in microphone to accumulate learning data while interacting with a

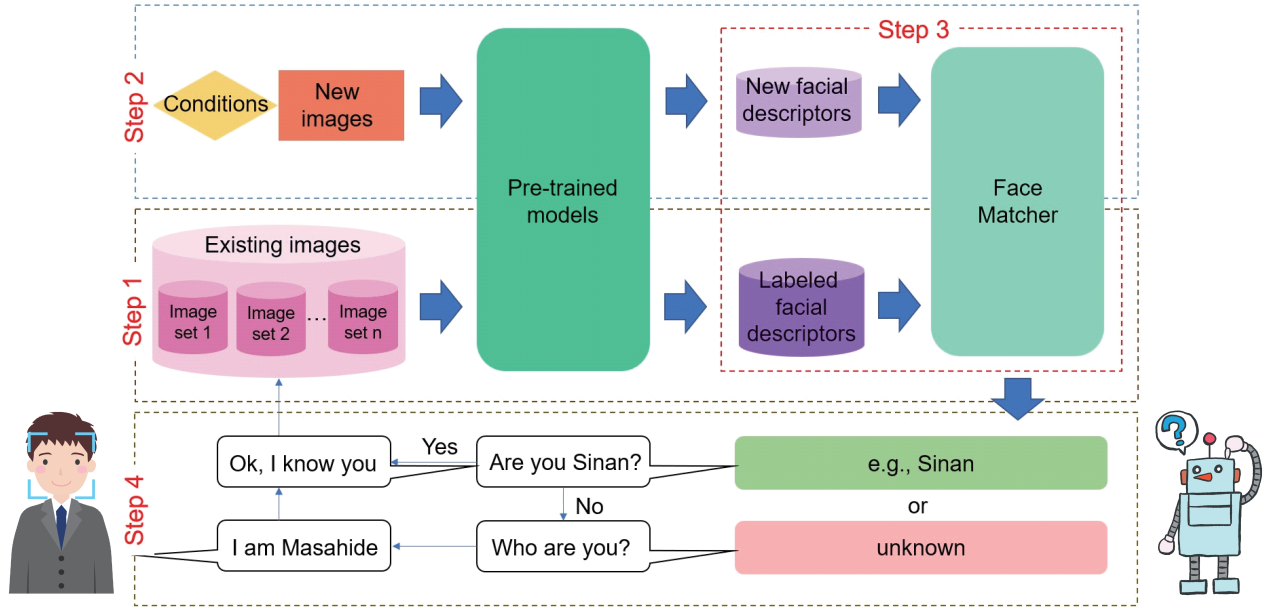


Fig. 3. Overall architecture of the proposed method.

user. It will start from the spontaneous initiation of a spoken dialogue scenario in a Web browser with a VA based on human face detection and identification results. The flow of the proposed method consists of the following four steps. In step 1, the system loads the existing data and extracts their features. In Step 2, the training data are retrieved, and their features are extracted. In Step 3, the system will perform a feature comparison between the existing and training data. Finally, in Step 4, we perform agent interaction based on the identification results. We will describe the details in the following sections.

C. Step 1: Loading existing data and extracting their features

In Step 1, the pre-trained models required for face identification are first loaded from a local storage location using JavaScript. Next, to check for the presence of existing data, we use a Web browser and a Node.js server to create a locally executable API that analyzes a specified local folder path part pattern using the Glob package in the Node Package Manager (NPM). API to analyze local specified folder path part patterns. If the number of existing data is 0, it automatically skips this step and moves to Step 2. On the other hand, if there is existing data, each set of raw image data is input to a pre-trained model of face recognition (i.e., face recognition function, see Section III-D), which outputs a face descriptor for each set. An example of a face descriptor is "label": "Sinan", "descriptors": [Float32Array(128)], which contains a typed array consisting of the label name and 128 32-bit floating-point numbers. Thus, features are extracted from existing raw image data, and labeled face descriptors are obtained.

D. Step 2: Retrieving new data and extracting their features

In Step 2, a video stream captured from the built-in camera of a general-purpose computer is first projected onto a Web

browser, and face detection is performed in real-time as a loop process. Next, the snapshot images are input to the pre-trained model as training data, with the presence of the user's face in the video stream as the trigger condition, and the face descriptors are extracted.

E. Step 3: Comparing features between existing data and training data

In Step 3, the face matcher function in face-API.js is used to compare the new face descriptors output from the pre-trained model with the existing set of face descriptors. Generally, the find best match function of the face matcher is used to calculate the difference distance between multiple face descriptors and output the best value. The probability of a discriminant result is expressed as a number between 0 and 1. On the other hand, if it is close to 1, it does not exist in the labeling results and is more likely to be a "stranger".

F. Step 4: Agent interaction based on the facial identification's results

In Step 4, we first consider a majority voting approach to finalize the identification results. The majority voting approach is common to ensemble learning with weight value [14]. The meaning of this step is to get the final result with the highest number of discriminations out of 10 consecutive discriminations. We then design the following agent interaction scenario using the final result as the trigger condition. As shown in Figure 3, when the labeled result is the final result, the agent spontaneously asks the user, "Are you Sinan (yes/no)?" If the user answers "yes", the system directly stores the training data in the corresponding set of existing data. On the other hand, if the user answers "no", the agent asks again, "What is

your name?”. After obtaining the user’s answer and attaching the correct label to the training data, the agent stores it in the existing data set. When the dialogue ends, real-time face detection starts again. In this way, learning data can be acquired and stored as a “dialogue”. Various value-added services can be expected, such as chatting and watching videos that can be adapted to individual preferences after the user’s identity judgment is correct.

V. CASE STUDY

A. Overview

In this section, as a case study, we consider the personalization of a home care support service using a face identification system based on the pre-trained model and spoken dialogue agents proposed in Section IV. Specifically, we discuss scalable value-added services based on three examples: support for monitoring the daily rhythm of an older adult with dementia at home, individual logging during agent interaction with multiple people, and real-time home invasion estimation for an older adult living alone.

B. Support for monitoring the daily rhythm of an older adult with dementia at home

Older adults with dementia living at home often wander indoors or go out unconsciously and have an irregular life rhythm that reverses day and night. Although many conventional data recording and analysis systems use sensor devices and cameras, they often fail to distinguish between family members and elderly persons with dementia data. They have not yet reached the level of support for monitoring a specific target person. We envision that the proposed face identification system can be linked to smart lock devices (e.g., Switch Bot Lock) and installed at entrances and exits to prevent older adults with dementia from going out without permission. In addition, using an infrared camera, it is promising to extract and identify facial features even during late-night hours. Regardless of whether the older adult with dementia interacts with the agent or not, the system can record the time of activity of the older adult with dementia within the camera’s range as part of his/her daily rhythm. Furthermore, when several older adults care for each other, we are considering using our proposed system to provide individualized monitoring support and automatically contact family members or care facilities in a timely manner when necessary.

C. Individual logging during agent interaction with multiple people

Various spoken dialogue agents, including smart speakers (e.g., Google Assist, Amazon Alexa), are being introduced every year. Among the users of these smart devices, there are usually many situations in which multiple people interact with the agents since the household composition of older adults at home who need support for monitoring is different. Hence, it is difficult to automatically identify “who said ..” with a single device. We propose a face identification system that identifies the user’s identity and records individual logs when

the user interacts with an agent. By automatically classifying the dialogue logs of each individual, it is promising to generate answers and value-added services adapted to the individual automatically. In addition, from the dialogue logs of multiple people, various extensions are expected, such as analysis of family relationships among speakers and estimation of individual requests. Furthermore, our proposed system can be used to register facial images of new members of the elderly person’s household in the form of a “dialogue” with the agent, even if the household composition of the elderly person at home changes.

D. Real-time home invasion estimation for an older adult living alone

In Japan, the number of older adults living alone is rapidly increasing, and the number of break-ins and burglaries targeting older adults’ residences is increasing yearly. Many home monitoring services have emerged, such as caretaker call services and home visitation services using specialized smart equipment. However, the cost of introducing equipment and services and the privacy issues of older adults still exist. Hence, we envision an approach that uses the proposed face identification system to analyze the difference between the frequency of change in the results of continuous identification. They include the frequency of change in the results of labeling within a certain period and estimate the possibility of the “stranger” result. In this way, the proposed face identification system can estimate that a “stranger” has entered a house. At this moment, if it cannot be confirmed by agent interaction with an older adult, it can automatically contact distant family members or the police. It is promising to support the more accessible introduction of a system to watch over the safety of older adults living alone.

VI. DISCUSSION

A. Advantage

We consider the following three advantages of the proposed face identification system with pre-trained models and spoken dialogue agents.

- **Improve user privacy issues:**

Since the system proposed in this research uses pre-trained models to perform face identification, all data communication can be done locally, without needing a global network connection, which improves user privacy issues.

- **Improve the process of training data:**

In this study, the face identification system is linked to a spoken dialogue agent, which can acquire and store data in the form of a “dialogue” between the user and the agent in real-time, even if the number of training data is zero. Based on this, a lightweight process for face identification can be realized using pre-trained models.

- **Improve ease of system implementation:**

The system proposed in this study uses only a general-purpose computer, a built-in camera, and a microphone. It does not require conventional sensor devices (i.e., human

detection and pressure sensors). The computer screen provides an interface that links a VA (i.e., MMDAgent) and a Web browser. It is expected to make it easier to install a face identification system in the homes of the elderly without incurring high costs and computational costs.

B. Limitation

We consider the following three limitations of our proposed face identification system with pre-trained models and spoken dialogue agents.

- **Face mask detection from raw images is not possible:** Along with the impact of the spread of the new coronavirus, older adults often wear face masks even in their residences. Therefore, the face identification system proposed in this study often uses image data, including face masks, as training data. However, masking faces is likely to degrade the accuracy of face descriptors, and the accuracy of face identification may be unstable. In response, we need to check for the presence of human faces in the raw images and the presence of face masks (refer to [15]) in the pre-trained models.
- **Loading existing data takes time:** In the system proposed in this research, training data is stored in the form of “dialogs”, but this increases the number of training data, and there is the problem of the time required to load face descriptors as existing data. Therefore, we are considering storing only the text data of face descriptors with labeling without storing most of the raw image data. In this way, the time required to load existing data can be reduced, and the privacy issue can be improved.
- **Authenticity of training data cannot be determined:** In the agent dialogue scenario proposed in this study, the inability to determine the authenticity of the answers given by the user is a serious problem. That is, the accuracy of the labels in the training data cannot yet be guaranteed. To address this problem, we envision an agent asking multiple users to confirm the authenticity of past image data. Thus, we expect that answers from multiple viewpoints will improve the accuracy of facial image labels.

VII. CONCLUSION

In this paper, we propose a method for developing a face identification system using pre-trained models and spoken dialogue agents to improve the problem that it is difficult to determine the status of a person with whom a spoken dialogue agent interacts. The proposed method automatically generates training data through spoken dialogue between a user and an agent and acquires and compares facial features using pre-trained models. Hence, we expect to realize an easy-to-implement face identification system using only a general-purpose computer and a web camera without requiring a conventional Internet connection or manual labeling of training data. In future work, we will implement and evaluate the

proposed system and consider extending it to a pre-trained model for mask detection from faces. We would also like to conduct clinical experiments using the proposed case study described in Section V in actual residences of older adults.

ACKNOWLEDGMENT

This research was partially supported by JSPS KAKENHI Grant Numbers JP19H01138, JP20H05706, JP20H04014, JP20K11059, JP22H03699, JP19K02973, Grant-in-Aid for JSPS Research Fellow (No.22J13217), and Tateishi Science and Technology Foundation (C) (No.2207004).

REFERENCES

- [1] C. Miura, S. Chen, S. Saiki, M. Nakamura, and K. Yasuda, “Assisting personalized healthcare of elderly people: Developing a rule-based virtual caregiver system using mobile chatbot,” *Sensors*, vol. 22, no. 10, p. 3829, 2022.
- [2] S. Chen, M. Nakamura, and S. Saiki, “Developing a platform of personalized conversation scenarios for in-home care assistance,” in *2021 IEEE International Conference on Industry 4.0, Artificial Intelligence, and Communications Technology (IAICT)*. IEEE, 2021, pp. 148–153.
- [3] S. Chen, H. Ozono, and M. Nakamura, “Integration analysis of heterogeneous data on mind externalization of elderly people at home,” in *International Conference on Human-Computer Interaction*. Springer, 2022, pp. 197–209.
- [4] K. Hirayama, S. Chen, S. Saiki, and M. Nakamura, “Toward capturing scientific evidence in elderly care: Efficient extraction of changing facial feature points,” *Sensors*, vol. 21, no. 20, p. 6726, 2021.
- [5] D. Smilkov, N. Thorat, Y. Assogba, C. Nicholson, N. Kreeger, P. Yu, S. Cai, E. Nielsen, D. Soegel, S. Bileschi *et al.*, “Tensorflow.js: Machine learning for the web and beyond,” *Proceedings of Machine Learning and Systems*, vol. 1, pp. 309–321, 2019.
- [6] P. Hu, H. Ning, T. Qiu, Y. Xu, X. Luo, and A. K. Sangaiah, “A unified face identification and resolution scheme using cloud computing in internet of things,” *Future Generation Computer Systems*, vol. 81, pp. 582–592, 2018.
- [7] S. Guo, T. Xiang, and X. Li, “Towards efficient privacy-preserving face recognition in the cloud,” *Signal Processing*, vol. 164, pp. 320–328, 2019.
- [8] S. Chen and M. Nakamura, “Generating personalized dialogues based on conversation log summarization and sentiment analysis,” in *The 23rd International Conference on Information Integration and Web Intelligence*, 2021, pp. 217–222.
- [9] X. Kong, K. Wang, S. Wang, X. Wang, X. Jiang, Y. Guo, G. Shen, X. Chen, and Q. Ni, “Real-time mask identification for covid-19: An edge-computing-based deep learning framework,” *IEEE Internet of Things Journal*, vol. 8, no. 21, pp. 15 929–15 938, 2021.
- [10] M. Parhi, A. Roul, B. Ghosh, and A. Pati, “Ioats: an intelligent online attendance tracking system based on facial recognition and edge computing,” *International Journal of Intelligent Systems and Applications in Engineering*, vol. 10, no. 2, pp. 252–259, 2022.
- [11] A. Lee, K. Oura, and K. Tokuda, “Mmdagent 脳九 or fully open-source toolkit for voice interaction systems,” in *2013 IEEE International Conference on Acoustics, Speech and Signal Processing*. IEEE, 2013, pp. 8382–8385.
- [12] S. Chen, S. Saiki, and M. Nakamura, “Nonintrusive fine-grained home care monitoring: Characterizing quality of in-home postural changes using bone-based human sensing,” *Sensors*, vol. 20, no. 20, p. 5894, 2020.
- [13] C. Li and C. Li, “Web front-end realtime face recognition based on tfjs,” in *2019 12th International Congress on Image and Signal Processing, BioMedical Engineering and Informatics (CISP-BMEI)*. IEEE, 2019, pp. 1–5.
- [14] A. Dogan and D. Birant, “A weighted majority voting ensemble approach for classification,” in *2019 4th International Conference on Computer Science and Engineering (UBMK)*. IEEE, 2019, pp. 1–6.
- [15] N. Damer, J. H. Grebe, C. Chen, F. Boutros, F. Kirchbuchner, and A. Kuijper, “The effect of wearing a mask on face recognition performance: an exploratory study,” in *2020 International Conference of the Biometrics Special Interest Group (BIOSIG)*. IEEE, 2020, pp. 1–6.