An Auxiliary Notification System for Amnesiac Patients Based on Face Recognition and Voice Synthesis Using Cloud Computing

Qiming Sun¹, Ping Guo² and Changjiang Zhang¹⁺
¹ Department of Computer Science, Wenzhou-Kean University, Wenzhou, Zhejiang, China
² Department of Computer Science, University of Illinois at Springfield, Springfield, Illinois, USA
*czhang@kean.edu

Abstract. Many patients with amnesiac symptoms have difficulties to recognize people, even relatives’ names. This paper aims at building an auxiliary system for patients with memory obstacles, especially for those who have difficulty in recognizing their relatives and close friends. The proposed system can detect the patient’s surroundings in real time and use face recognition and voice synthesis techniques to recognize someone’s approach to the patient. If the face recognition is achieved, it will notify the patient by text and voice messages, which is saved on cloud server during the training. Based on a novel framework, the system has a good usability and does not require complex human operations. It is friendly to patients. All its functionalities are designed to be automatic and efficient.

1. Introduction
Many patients with amnesiac symptoms suffer from recognizing people’s names. For example, people with Alzheimer’s Disease (AD) are very likely to get cognitive dysfunction in few years [1]. According to Laffan and Bier’s research, as AD patients suffer from memory impairment, they may get trouble to make face-name associations. The re-learning of the associations and training of basic cognitive process can be applied to moderate the disease [2, 3] and improve memory for elders [4].

In this paper, we proposed a system to help amnesiac patients address face-name association problems by utilizing the combination of cloud server and smartphone application. The use of cloud computing could guarantee the efficiency and performance of face identification [5]. Backed by cloud server, smartphones with limited power can also accomplish complex computation. In addition, our system adopts the techniques of face recognition and voice synthesis, which can equip our system with automated and intelligent functions for users.

Our system allows users to register faces to the database on the cloud, the system can train a model online in a short time. After the registration, users can open the real-time function for patients. The application is able to automatically monitor the surroundings and detect the person who is approaching the patient. The system will then use voice synthesis technique to notify the patient.

Our proposed system could be a good auxiliary tool for people with memory obstacle. Each module is automated enough and does not require complex human operations. The system can automatically scan the camera’s view and notify the patients in voice if someone else is coming. It runs in a good
performance without interruption unless the application is closed manually, or the smartphone runs out of power.

2. Related Work
Face recognition and voice synthesis, which mean identify specific faces and converting text to speech, could be a helper for patients with amnesiac symptoms. Currently, researchers have proposed some successful examples to setup face recognition service on cloud service [6]. Many models and frameworks have combined mobile processing and cloud computing together with a good performance [7].

Researchers have shown that building an auxiliary system on a mobile phone for health-care has a very good usability [8-10]. Wade’s team suggests that the mobile phone has a potential to enhance independence, reduce stress and could be a good memory aid tool [11]. Migo et al conclude that smartphone applications have the potential to help patients with memory problems [12]. In addition, Evald’s research proved that smartphones are useful compensatory tools in memory recovery process [13].

3. Implementation Details
The proposed system uses offline face detection to check if there is any face information inside the camera view. Once a face is detected, the system can upload the images to the cloud server to do online face registration or recognition. The server maintains a database that organizes different devices and users, and the cloud services, including face recognition and voice synthesis, are under the support of Baidu AI Platform Java SDK. The entire system is built on Android and requires a smartphone with at least one camera and Internet connection. Figure 1 describes the general logic of the system.

![Figure 1. The overall logic of the proposed system.](image)

3.1. Offline Face Detection
The face detection module aims at checking each frame from the camera’s stream and decides whether a face exists or not. The detection process is used to make sure that the current frame is available to do further computation on the cloud, and since its operations are done offline, it significantly reduces the Internet usage and improves the system performance.

OpenCV Library, which is designed for real-time applications and multi-core processing, is used here to do the image process as well as the detection. For each frame captured by the camera, the image will be transformed to grayscale. Then a built-in cascade classifier will be loaded to try to generate the size and features of a face from the grayscale image.

Since the images taken by a cell phone camera are sometimes fuzzy because of slight hands shaking, in order to improve the image quality, an anti-shaking algorithm is introduced. The algorithm selects the sixth image with detected face after a continuous five images with faces, which means the camera should be stable in at least six frames, and it can avoid shakings in most cases.
Once an image in a frame passed the face detection with an available face as well as the anti-shaking check, it will be converted to a byte array and sent to the cloud server to do more computation for face recognition or face registration.

3.2. Online Face Registration and Recognition
Two core cloud-based modules, face registration and face recognition, are demonstrated in this section. Face registration is used to train the system to distinguish a certain face. Users should use this function at first to let the system build a specific model in order to recognize their faces. After a face is registered, the user can be identified by the face recognition module once he/she appears in the camera view.

In face registration, the module accepts three images from the detection module, and the interval between two captures is 5 seconds. In the whole process, instructions will be shown on the screen to suggest users moving their head. In this way, three images of a face in different angles are retrieved, then uploaded and registered in the cloud server. Besides, the system requires users to provide their names, which will be used to synthesize a fragment of voice to notify the patients in recognition module.

Face recognition module can keep monitoring the surrounding environment and notify users in voice once a person appears and the face has a similar matching in the current database. By uploading an image contains face information through the SDK, a JSON string with the similarity score of the best match and its description string will be returned. If the score is greater than a threshold value, such as 80, meaning ‘very similar’, the system will consider the person as a registered user, and the description string will be sent to voice synthesis module to generate and play a speech.

3.3. Voice Synthesis Module
Voice synthesis is an auxiliary function that helps the patients know who is coming or talking to him/her. Based on the Baidu’s platform, the module can get a byte array with the synthetic voice fragment by sending a text to the cloud server.

The module is integrated with a converter and a media player. The converter uses an output file stream to save the voice fragment to a cache directory in MP3 format. The media player then uses an input stream to load and play the MP3 file. The temporary files will be deleted automatically after the user terminates the application.

3.4. Control of Users
The proposed system supports multi-device as well as multi-users. Figure 2 shows the framework of the user repository.

![Figure 2. Organization framework of users.](image)

The system uses a cell phone’s IMEI as a unique key to distinguish different devices. When a user installs and starts the application for the first time, it will automatically register an account on the cloud server using IMEI number of the cell phone. Each device supports multiple users and each user supports multiple face images.

3.5. Multi-thread Framework
To optimize the user experience and reduce the negative influence caused by Internet latency. A novel multi-thread framework is designed to build the system. Figure 3 demonstrates the details of the framework.

![Figure 3](image)

**Figure 3.** The multi-thread framework inside the system.

Once an instance is created, the system will create a UI thread, a camera thread, as well as a handler for cross-thread communications. The UI thread organizes instructions and elements on the screen. The camera thread is used to classify the frames and select suitable images with face information. After a suitable image is selected, a web request thread will be created to communicate with the server and play the synthetic voice. The message returned from the cloud server will then be passed to the handler. The handler can classify the messages, determine the next steps, and update the other threads.

4. Results

A smartphone with Android system is required to run the application. Figure 4 is a flowchart of the detailed usage of the proposed system.

![Figure 4](image)

**Figure 4.** The flow chart to use the system.

Initially, when users open the application, they are required to grant the permissions to allow the program to invoke the cameras and get the phone’s IMEI code. Then people should register their faces before they can be recognized by the system. Users can ask the relatives or friends of the patients with memory problems to register their faces. Before the registration, the system asks the user to enter a
text description, such as a name, as the label to the registered face. The text will be used to generate a voice fragment to notify the patients later. Then the system will open the camera and start to collect the facial information. This process could last for 10-15 seconds. Figure 5 shows a screenshot of the registration module. ‘David’ is registering his face and the system is training a model online.

Figure 5. The user interface of face registration.

After the registration, the system will be able to recognize the specific person. When the user chooses the recognition function, simply place the smartphone somewhere that can capture the faces is enough. For example, fix the smartphone on the patients’ wheelchair. The system can keep monitoring the surroundings. For instance, an amnesiac patient cannot recognize the relatives in a short time, when a relative, who has registered his/her face in the system, comes to talk to the patient, the system can capture the person’s face and tell the patient who is coming. Figure 6 shows the face identification interface; when a registered person comes inside the camera view, the system can recognize him and display his name. At the same time, His name ‘David’, which has been registered as the label text, will be spoken out to notify the patient.

Figure 6. The user interface of face recognition.

5. Conclusions
The proposed system has the ability to recognize registered people’s face in real time and notify the person in voice when the person is approaching him/her. It is easy to learn and operate with a simple user interface and could be a good auxiliary tool for elder people with amnesiac symptoms. Also, it does not need the patients themselves to be able to use the smartphone. Instead, simply install the phone near the patient is enough, the application can keep notifying the patient’s surroundings for someone’s approaching without human intervention until it runs out of power or the application is closed.

In the future, we will ask more patients to try the system. Based on their feedbacks, some new features or functions might be added. We will try to make the application more helpful to the people in need.

We may also consider using an embedded system to build a special face recognition device considering the limited power of a smartphone.
References

[1] Amieva, Hélène, et al. "Prodromal Alzheimer's disease: successive emergence of the clinical symptoms." Annals of neurology 64(5), 492-498 (2008).

[2] Laffan, Amanda J., et al. "Making errorless learning more active: Self-generation in an error free learning context is superior to standard errorless learning of face–name associations in people with Alzheimer's disease." Neuropsychological rehabilitation 20(2), 197-211. (2010).

[3] Bier, Nathalie, et al. "Face–name association learning in early Alzheimer's disease: A comparison of learning methods and their underlying mechanisms." Neuropsychological rehabilitation 18(3), 343-371. (2008).

[4] Hering, Alexandra, et al. "Prospective memory training in older adults and its relevance for successful aging." Psychological research 78(6), 892-904. (2014).

[5] Yuan, Jiawei, and Shucheng Yu. "Efficient privacy-preserving biometric identification in cloud computing." INFOCOM, 2013 Proceedings IEEE, pp. 2652-2660. IEEE (2013).

[6] Soyata, Tolga, et al. "Cloud-vision: Real-time face recognition using a mobile-cloudlet-cloud acceleration architecture." Computers and Communications (ISCC), 2012 IEEE Symposium on, pp. 000059-000066. IEEE (2012).

[7] Shiraz, Muhammad, et al. "A review on distributed application processing frameworks in smart mobile devices for mobile cloud computing." IEEE Communications Surveys & Tutorials 15(3), 1294-1313. (2013).

[8] Yamagata, Christina, et al. "Mobile app development and usability research to help dementia and Alzheimer patients." Systems, Applications and Technology Conference (LISAT), 2013 IEEE Long Island, pp. 1-6. IEEE (2013).

[9] Ventola, C. Lee. "Mobile devices and apps for health care professionals: uses and benefits." Pharmacy and Therapeutics 39(5), 356 (2014).

[10] Zorluoglu, Gokhan, et al. "A mobile application for cognitive screening of dementia." Computer methods and programs in biomedicine 118(2), 252-262. (2015).

[11] Wade, Tracy K., and Juliette C. Troy. "Mobile phones as a new memory aid: a preliminary investigation using case studies." Brain injury 15(4), 305-320. (2001).

[12] Migo, Ellen M., et al. "mHealth and memory aids: levels of smartphone ownership in patients." Journal of Mental Health 24(5), 266-270. (2015).

[13] Evald, Lars. "Prospective memory rehabilitation using smartphones in patients with TBI." Disability and rehabilitation 25(2), 1-10. (2017).