Smart Glasses Application Based on Face and Facial Expression Recognition for Children with Neurodevelopmental Disorders

Chieh-Ming Yang[1[0000-0002-6614-4635]], Firdaus Golam[1], Jen-Yeu Chen[1], and Wei-Che Chien[1]

[1] Department of Electrical Engineering, National Dong Hwa University, Taiwan
j82887@gmail.com

Abstract. This study presents an invention of a lightweight and portable smart glasses device that incorporates facial expression recognition capabilities. It aims to assist in the behavioral therapy of children with neurodevelopmental disorders by displaying the results of the facial expression recognition model in an augmented reality manner. The model achieved an accuracy of 85.27% on the testing dataset. This invention provides a flexible and customizable approach to therapy with potential benefits for the enhancement of attention and facial expression recognition of children with neurodevelopmental disorders.

Keywords: Neurodevelopmental Disorders, Augmentative and Alternative Communication, Facial Expression Recognition, Convolutional Neural Network, Edge AI, Augmented Reality.

1 Introduction

Children with neurodevelopmental disorders (NDs) often face difficulties in learning, social communication, attention, and facial expression recognition (FER) in inclusive education, which results in their inability to integrate with peers in school or social groups. In the behavioral therapy for children with NDs, therapists use smart assistive technologies (SATs) such as talking pens and electronic communication boards to gradually guide them towards correct oral expression and recognition. Due to the high cost, bulkiness, and limited functional changes of commercially available Augmentative and Alternative Communication (AAC) systems, therapists often use real objects and picture cards instead of AAC devices as assistive tools in therapy.

With the continuous advancement and refinement of artificial intelligence technology, various technological tools have been introduced to assist in the education of children with NDs, helping them integrate into schools and social groups. In the study conducted by Catania et al. [1], three therapists utilized Google Home devices equipped with Google Assistant in the therapy sessions for children with NDs, discussing the benefits and challenges of therapist-child interaction through this technology. Barua et al. [2] provided a comprehensive review of the scope and effectiveness of artificial intelligence-assisted tools developed using machine learning models, which have been applied to address learning challenges faced by students with NDs. Belpaeme et al. [3] focused their research on designing long-term, adaptive social interactions between robots and child users in real-world settings, evaluating the potential of robots in assisting children's social skills training. Therefore, the aforementioned studies confirm that artificial intelligence holds promising prospects in improving social interaction among children with developmental disorders.

According to above reasons, this study has implemented facial expression recognition functions into a portable and lightweight smart glasses assistive device. The application is designed to help children with NDs in their behavioral therapy for emotion recognition, in order to eliminate the inequality caused by their congenital conditions.

2 Materials and Methods

The purpose of this study is to enhance the attention and facial expression recognition of children with NDs. To achieve this goal, a smart glasses application with facial expression recognition was proposed. The application allows the selection of designated individuals for adding stickers and text reminders, thus creating an augmented reality effect. Additionally, users can define their own expression stickers by choosing their preferred and eye-catching design.

For facial expression recognition, the FaceNet model [4] is utilized to extract 128 features of each face image. The Euclidean distance is calculated to measure the similarity between two different faces, with larger distances indicating less similarity and smaller distances indicating more similarity. For facial expression recognition, a lightweight convolutional neural network model, as shown in Fig. 1 is designed and trained using the Facial Expression Recognition 2013 dataset [5]. The dataset contains seven categories of emotions: angry, disgust, fear, happy, sad, surprise, and neutral. Each grayscale image is 48x48 pixels, and the training and testing sets consist of 28,709 and 7,178 images, respectively. The model is trained using the Adam optimizer with a total of 100 epochs in the TensorFlow 2.0 environment under Python 3.6.

The schematic diagram of the smart glasses with facial expression recognition is shown in Fig. 2. The app in the smartphone allows for the addition of facial data to a database, including the name and feature values of specific individuals. The camera on the smart glasses can perform real-time facial recognition and calculate the Euclidean distance of each recognized face from the database. If the distance is less than the set threshold of 0.9, the corresponding name will be displayed; otherwise, if the distance is greater than 0.9, the glasses will not display the person's name, or will label them as an unknown individual. The recognized face will also trigger a sticker based on the output of the facial expression recognition model.

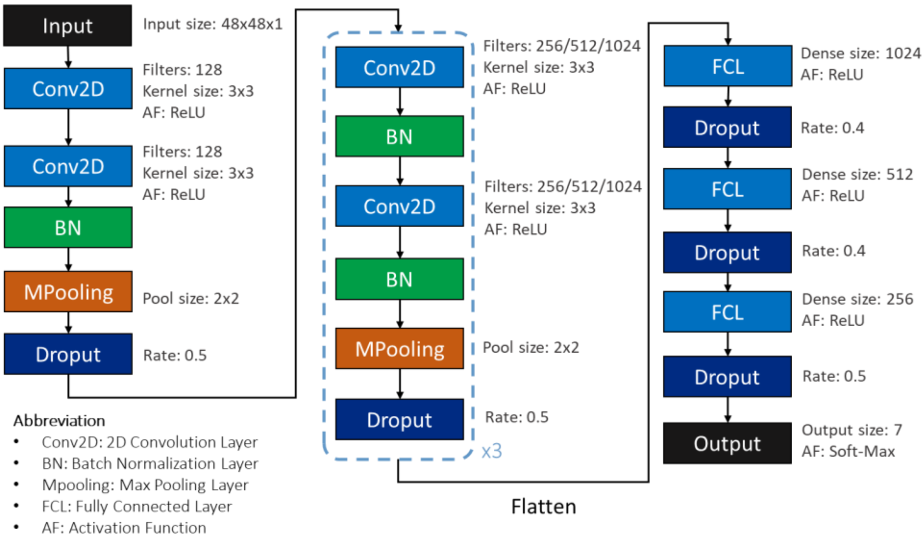

Fig. 1. A convolutional neural network model with lightweight design proposed in this study.

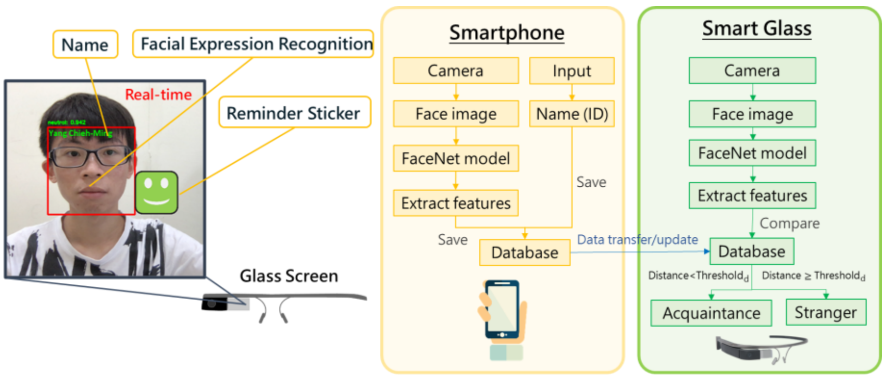

Fig. 2. The flowchart of face and facial expression recognition functions on smartphones and smart glasses.

3 Results

In the training of the facial expression recognition model, the accuracy of the training set and the testing set are 85.23% and 85.27%, respectively. The model use the grad-cam heatmap to visualize the focused areas. Since facial expression recognition relies on capturing small muscle movements on the face, Table 1 shows that the model focuses on distinguishing expression based on features such as mouth opening, eyebrow furrowing, and eye corners. The light yellow-green area indicates the high attention area of the model, while the dark blue-green area indicates the low attention area.

Table 1. Visualization of focused areas in facial expression recognition images using grad-cam heatmap in this study's model.

Label	Original	Heatmap	Label	Original	Heatmap
Angry			Sad		
Disgust			Surprise		
Fear			Neutral		
Happy					

4 Discussion

This study applies edge computing and artificial intelligence techniques on smart glasses using lightweight models and GPU acceleration to achieve real-time facial expression recognition. The stickers and text reminders related to facial expressions are displayed in an augmented reality manner to assist in the behavioral therapy of children with NDs. The results demonstrate that the proposed system is effective in enhancing attention and facial expression recognition of children with NDs. The system's high accuracy and real-time processing provide an efficient and convenient solution for the application of edge AI in the field of behavioral therapy. Additionally, the personalized sticker and text reminder function provides a more flexible and customizable approach to therapy.

This development has been completed and is currently in the testing phase. In the future, extensive testing will be conducted on a large number of children with NDs to evaluate the effectiveness of this device as a therapeutic tool in inclusive education, particularly addressing issues related to emotion recognition and attention.

References

1. Catania. F., Spitale. M., Garzotto, F.: Toward the introduction of google assistant in therapy for children with neurodevelopmental disorders: an exploratory study. In: CHI '21: CHI Conference on Human Factors in Computing Systems, pp. 1–7. Yokohama Japan (2021).
2. Barua, P. D., Vicnesh, J., Gururajan, R., Oh, S. L., Palmer, E., Azizan, M. M., Kadri, N. A., Acharya, U. R.: Artificial intelligence enabled personalised assistive tools to enhance education of children with neurodevelopmental disorders — A review. International Journal of Environmental Research and Public Health 19(3), 1192 (2022).

3. Belpaeme, T., Baxter, P., Read, R., Wood, R., Cuayáhuitl, H., Kiefer, B., Racioppa, S., Kruijff-Korbayová, I., Athanasopoulos, G., Enescu, V., et al.: Multimodal child-robot interaction: building social bonds. Journal of Human-Robot Interaction 1(2), 33–53 (2013).
4. Schroff, F., Kalenichenko, D., Philbin, J.: FaceNet: A unified embedding for face recognition and clustering. In: 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Boston, MA, USA (2015).
5. Goodfellow, I. J., Erhan, D., Carrier, P. L., Courville, A., Mirza, M., Hamner, B., Cukierski, W., Tang, Y., Thaler, D., Lee, D. H., et al.: Challenges in representation learning: A report on three machine learning contests. in International Conference on Neural Information Processing, pp. 117–124 (2013).