AI-Based Communication Tool for High-Functioning Autistic Children

Farah Ossama, Fatmah Issam, Menna Hesham, Menna Kamel, Yumna Hamdy, Merna Bibars, Ibrahim Sadek

***Abstract*— High-functioning autistic children tend to have poor verbal and nonverbal communication skills, which makes it hard for them to understand people's cues and blend in with their community. This paper presents a solution to help Arabic-speaking children with autism communicate effectively and understand others' feelings. Our AI-based Communication Tool uses a mobile camera to capture the speaker's image and informs the autistic child of their emotions through a mobile app. We used images from the FER+ dataset to train on several TensorFlow models to achieve accurate facial recognition. We achieved 83.97% accuracy. Furthermore, we develop games on the mobile application using Unity to help autistic children practice different social scenarios.**

I. INTRODUCTION

Autistic children, including high-functioning ones, suffer from social interaction difficulties and communication differences. FER (Facial Emotional Recognition) systems help autistic children engage and communicate with society. For example, Premkumar et al. [1] used an FER system to guide autistic children while dealing with people by telling them the age, gender, and emotion of the person they were dealing with. Existing literature, like the study JobTIPS by Strickland et al. [2], was insightful in cultivating autistic individuals' social integration. On the other hand, games are used to simulate real-life situations and evoke emotions to enhance enjoyment and interactivity for children. Our approach is inspired by research, such as Barajas et al. [3]. We chose this topic because it is pivotal to dedicate as much attention as possible to autistic individuals' need for independence and social interactions. Our main goal is to support the independence of Arabic-speaking, high-functioning autistic children with AI-based Communication Tools to help them communicate and understand others' emotions using the FER system and mobile games.

II. MATERIALS AND METHODS

The communication tool is divided into three main subcategories: (A) Facial Emotion Recognition System, (B) Therapeutic Games Development, and (C) Application Development.

A. Facial Emotion Recognition (FER) System

The FER system is a pre-trained deep-learning classification model built using TensorFlow. The model was finetuned to FER+ data and trained on the Google Colab free cloud platform.

- Dataset: FER+ is an improved version of the FER2013 dataset, as FER2013 contains some inaccuracies in the annotations, making it difficult to train models effectively. The eight classes of emotions present in the data are neutral, anger, disgust, fear, happiness, sadness, surprise, and contempt. The dataset exhibits class imbalance.
- Pre-processing: The dataset provided a score for each emotion per image. The emotion with the highest score was taken as the ground truth class for each image. The two classes, "unknown" and "NF," were neglected to focus on classes with valid emotion categories. We applied data augmentation techniques like rotation, zooming, horizontal flipping, and shearing to avoid overfitting. The images were resized to 96x96 pixels. The images were also converted to RGB to accommodate the usage of ImageNet weights in the finetuning step.
- Finetuning Approach: We experimented with several pre-trained models (i.e., MobileNetV2, ResNet50, EfficientB0). All models were pre-trained on the ImageNet dataset. The models were adjusted to accept an input size of 96x96 and to produce predictions for the eight classes. To address the class imbalance, the sci-kit-learn library generated class weights by specifying a dictionary where each class is mapped to its corresponding weight, and then we used these weights to emphasize some classes and deemphasize others during training.

B. Game Development

The four games provided in the App were developed using Unity and were based on therapeutic techniques after consulting with autism specialists in Egypt. The games aimed to aid autistic children in communication and understanding emotions. Fig 1 shows the games selected by therapists which is listed below.

1- "Audio stories": The autistic child will hear a story and then arrange the sequence of actions based on what they heard.

2- "Act out the emotion": They are asked to spin a wheel containing some emotions, and they are supposed to open the mobile camera and mimic the emotion it lands on.

3- "Match emotion with related objects": In this game, they are shown some emotions and are expected to match the corresponding color and photo to this emotion.

4- "What are you feeling?": In this game, a positive or negative scene is presented to the child with a set of feelings from which he chooses the right response he is supposed to feel. This helps the child understand his feelings and relate events to the appropriate feelings that normal people feel.

Farah Ossama, Fatmah Issam, Menna Hesham, Menna Kamel, Yumna Hamdy, and Merna Bibars are with the Systems and Biomedical Engineering Department, Faculty of Engineering, Cairo University.

Ibrahim Sadek is with the Biomedical Engineering Department, Faculty of Engineering, Helwan University (ibrahim_ibrahim@h-eng.helwan.edu.eg).

Figure 1 Games

C. Application Development

The mobile application (Azraq) is an autistic-friendly mobile app that helps autistic children communicate with people. The App can detect people's emotions by taking images and receiving responses from FER and the back end.

- Front-End: It was developed using Unity and consists of the following pages: Start, Signup, Login, and Home (including settings and a game menu). To ensure access to games, the child must complete the pseudo-social script, except for the initial game.
- Back-End: We utilized the Flask framework and SQLite database for development, with deployment on Google Cloud. To make signing up easier for the child, we made the primary key in the user database a composite key between the username and the password to give the user freedom to choose their unique username.
- Pseudo-social script: We used Google speech recognition API to translate from speech to text to get a record from the user and compare it to what was recommended by the application. We could not find an API that can transfer from text to speech in the Egyptian accent. So, we recorded some audio with our voices instead. The final step is to ensure the user said the recommenced sentence, which was done using the cosine similarity metric.

III. Results and Discussion

A. FER

MobileNetV2 had the best results; therefore, it was utilized in the App. After training for 62 epochs, it reached a test accuracy of 83.97%, a precision of 80.40%, a recall of 65.61%, and an F1 score of 72.26%.

During training, a learning rate scheduler was used, which decreased the learning rate from 0.001 to 0.0001; Adam optimizer and categorical cross-entropy loss function were also used. Table 1 shows a comparison between papers that used FER+ dataset.

Table 1: Comparison with Other Methods

Reference	Method	Accuracy (%)
Barsoum et al.[5]	Pre-trained ConvNet	76.90
Miao et al.[6]	SHCNN	86.54
Li et al.[7]	KTN	90.49
Our Pre-trained Approach	MobileNetV2	83.97

B. Game Development

After implementing and testing the four games recommended by specialists, we are currently focused on developing a therapist dashboard. This dashboard will serve as a monitoring tool for therapists, ensuring each autistic child is assigned to a specific doctor who will conduct pre- and post-assessment tests. These assessments are being designed to evaluate the areas targeted by the games. We can determine how much the games contributed to skill development by comparing the assessment scores.

C. Application Development

The mobile app pages were implemented and integrated with the developed games and the implemented APIs.

IV. Conclusion

To sum up, we chose this topic because it is crucial to dedicate more attention to autistic children's need for independence and social interactions. This project sheds light on a new perspective for Arabic-speaking autistic children and fills the lack of mobile applications aiding the Arabic native speakers' region, specifically the Egyptian dialect.

Acknowledgment

This paper owes its existence to the exceptional support of our supervisor, Dr Ibrahim Sadek, and the assistance of Eng. Merna Bibars. We are grateful to Dr Ihab El-Baz, head of the mental health hospital in the medical complex in Maadi, Dr Menan Abd El-Maqsod, Secretary General of Mental Health at the Ministry of Health, and all the respectable specialists who helped us produce this paper.

References

[1] Premkumar*, K., Murugapriya, K., Varsha, M. R., Asmitha, R., & Sureka, S. (2020). Facial Emotion Recognition for Autism Children. In International Journal of Innovative Technology and Exploring Engineering (Vol. 9, Issue 7, pp. 1274–1278). Blue Eyes Intelligence Engineering and Sciences Engineering and Sciences Publication - BEIESP. https://doi.org/10.35940/ijitee.f3772.059720

[2] Strickland DC, Coles CD, Southern LB. JobTIPS: a transition to employment program for individuals with autism spectrum disorders. J Autism Dev Disord. 2013 Oct;43(10):2472-83. doi: 10.1007/s10803-013-1800-4. PMID: 23494559; PMCID: PMC3706489.

[3] A. O. Barajas, H. Al Osman and S. Shirmohammadi, "A Serious Game for children with Autism Spectrum Disorder as a tool for play therapy," 2017 IEEE 5th International Conference on Serious Games and Applications for Health (SeGAH), Perth, WA, Australia, 2017, pp. 1-7, doi: 10.1109/SeGAH.2017.7939266.

[4] FERPlus. (n.d.). FERPlus | Kaggle. https://www.kaggle.com/datasets/ss1033741293/ferplus

[5] Emad Barsoum, Cha Zhang, Cristian Canton Ferrer, and Zhengyou Zhang. 2016. Training deep networks for facial expression recognition with crowd-sourced label distribution. In Proceedings of the 18th ACM International Conference on Multimodal Interaction (ICMI' 16). Association for Computing Machinery, New York, NY, USA, 279–283. https://doi.org/10.1145/2993148.2993165

[6] S. Miao, H. Xu, Z. Han and Y. Zhu, "Recognizing Facial Expressions Using a Shallow Convolutional Neural Network," in IEEE Access, vol. 7, pp. 78000-78011, 2019, doi: 10.1109/ACCESS.2019.2921220.

[7] H. Li, N. Wang, X. Ding, X. Yang and X. Gao, "Adaptively Learning Facial Expression Representation via C-F Labels and Distillation," in IEEE Transactions on Image Processing, vol. 30, pp. 2016-2028, 2021, doi: 10.1109/TIP.2021.3049955.