

Article

Not peer-reviewed version

AlexiLearn: Facial Expression Recognition in a Mobile Educational Tool for Alexithymia and Autism

[Iacopo Schianchi](#) *

Posted Date: 2 April 2025

doi: 10.20944/preprints202504.0165.v1

Keywords: Alexithymia; Autism Spectrum Disorder (ASD); Facial Expression Recognition (FER); Mobile Applications; Education; Mental Health; Gamification



Preprints.org is a free multidisciplinary platform providing preprint service that is dedicated to making early versions of research outputs permanently available and citable. Preprints posted at Preprints.org appear in Web of Science, Crossref, Google Scholar, Scilit, Europe PMC.

Copyright: This open access article is published under a Creative Commons CC BY 4.0 license, which permit the free download, distribution, and reuse, provided that the author and preprint are cited in any reuse.

Article

AlexiLearn: Facial Expression Recognition in a Mobile Educational Tool for Alexithymia and Autism

Iacopo Schianchi

Independent Research, USA; iacopo.schianchi@gmail.com

Abstract: Individuals with alexithymia suffer from a reduced ability to feel, express, and recognize emotions; however, the treatments and therapies that are currently in use are often expensive and inaccessible to some. Moreover, autism spectrum disorder is associated with reduced facial expression recognition ability. In this project, the aim was to develop AlexiLearn, a free application that teaches various aspects of emotions to people with these conditions. This application will include a real-time facial expression recognition (FER) system and other educational features like lessons, practices, and emotional reflections. The result was an 80.714% accuracy model performing real-time FER on a mobile device with the use of Google's ML Kit face detection model. Gamified elements were added to several other features to increase engagement, like points and upgrades. This paper also considers possible further research, improvements, and features to address common struggles of individuals with these mental conditions, including expressing and feeling emotions.

Keywords: Alexithymia; autism spectrum disorder (ASD); facial expression recognition (FER); mobile applications; education; mental health; gamification

1. Introduction

A vital part of interpersonal interaction is the use of body language and facial expressions. Alexithymia is a condition of disrupted emotional awareness that influences an average of 7-10% of individuals [1,2]. It is also common for individuals with alexithymia to suffer from a reduced ability to recognize facial expressions and nonverbal emotions [3]. For individuals with alexithymia, emotional unawareness often leads to difficulties in personal relationships and mental health challenges, including depression [4]. In addition, studies suggest a reduced facial expression recognition ability in people with autism spectrum disorder [5]. Although certain therapies, such as group therapy [6], are available, they are often expensive and inaccessible. Furthermore, there are few existing technologies that attempt to solve these problems in a holistic manner. Facial expression recognition systems have proven to be valuable tools in various fields, including software development and human-computer interactions, education, medicine, security, marketing, robotics, and games [7]. Emotion recognition and expression programs (EREP) that were conducted on individuals with alexithymia and schizophrenia have been found to have positive effects on patients [8]. By combining these, AlexiLearn explores the use of a real-time facial expression recognition system jointly with an emotional educational application to create an accessible, effective, free, and interactive tool that may be added to certain alexithymia and autism spectrum disorder therapies.

2. Literature Review

Existing research on facial expression recognition models shows the effectiveness of both deep and transfer learning [9]. Lightweight and efficient facial expression recognition is essential for this application, and the MobileNet-V2 model has given positive results for fast, real-time transfer learning FER systems in previous research [10,11]. Of the many available datasets for facial expression recognition, the open-source FER-2013 dataset is one of the most commonly used in FER studies [12]. However, the 28,709 images are unequally distributed in the seven emotions, and there

are only 436 images of disgust emotions. Therefore, it may be difficult to get a well-performing model in all emotions only using the raw FER-2013 dataset. There have also been efforts to enhance FER datasets with the use of preprocessing techniques, including data augmentation [13]. This preprocessing technique could be used to increase the amount of data that can be provided to the transfer learning model as well as balance each category. Overall, a combination of these techniques could yield a model that can both achieve high accuracy and fast recognition rates, making it ideal for its intended use in AlexiLearn.

3. Methods

3.1. Model Training



Figure 1. Sample data from FER-2013 [14].

The first step to create this app and the FER system was to develop the model using the FER-2013 dataset. The first step in preparing the data was loading the images and resizing them from their original 48 x 48 pixel dimensions to 224 x 224 pixels, which matches the MobileNet-V2 model input format. These images were then randomly augmented with a combination of contrast and horizontal flip. This was performed on each image to achieve an equal number in each category. At the end of this process, each category had 9286 images, and the total became 65002.

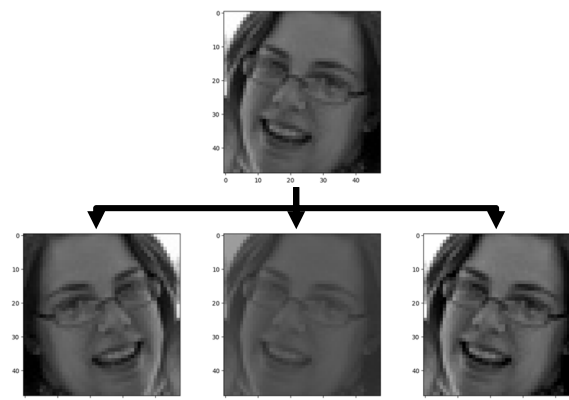


Figure 2. Example of augmented image [14].

Table 1. Dataset size before and after data augmentation [14].

Category	Original	Augmented
Happiness	7215	9286
Neutrality	4965	9286
Sadness	4830	9286
Fear	4097	9286
Anger	3995	9286
Surprise	3171	9286
Disgust	436	9286
Total	28709	65002

The final classification model incorporated custom dense layers (256, 128) and a softmax output layer with 7 neurons, optimized using the Adam algorithm and trained with sparse categorical cross entropy loss over 30 epochs. This model was then exported for use in the Android application. The .tflite file had a size of 9.82 MB, making it suitable for use in the mobile application.

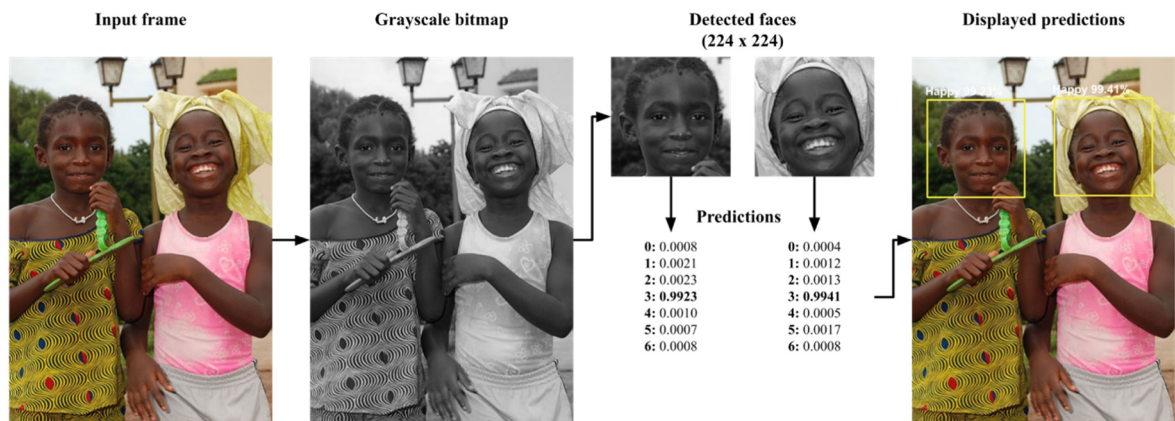


Figure 3: Facial detection and expression recognition process

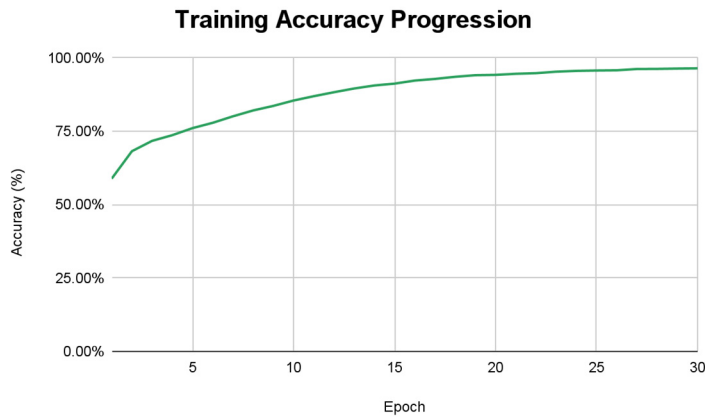


Figure 4. Training accuracy progression.

3.2. Camera Implementation

The next step was creating the AlexiLearn Android application and the functioning FER system using the model that was trained. The proposed app was made in Android Studio using Java, using the CameraX API for real-time camera processing. The camera API's image analysis use case was used to access every frame from either the front or back camera. Then, the image was preprocessed to match the required formats for both the face detection and FER models. The colored frame was converted into a grayscale bitmap of lower resolution to ensure fast face detection using Google's ML Kit face detection. The face detection model was set to the fast performance mode, and landmark, contour, and classification modes were disabled. This model returned the predicted bounding boxes for each face in the frame, which were used to crop the original image around the faces. The bounding box sizes were adjusted slightly to match the FER-2013 dataset's images. Finally, these predictions were preprocessed one more time by setting the dimensions to 224 x 224 pixels and normalized to match the previously trained FER model's requirements. Next, the model was used to predict the facial expression in each cropped bitmap. These predictions and the bounding boxes were drawn on a canvas that was overlaid on top of CameraX's camera preview; therefore, the camera's preview had no delay, while the predictions are asynchronously shown on top of it. This ensures a smooth experience for the user.

3.3. Additional Educational Features

While this feature may be helpful for the improvement of facial expression recognition in individuals with alexithymia and autism, there are various other aspects of emotional intelligence that can be addressed in AlexiLearn. In addition, it is important for AlexiLearn to have gamified aspects in order to enhance the learning experience for users. Studies show the benefits of gamification on learning, engagement, motivation, etc., including its use for learning social and emotional skills [15–17]. One method of gamifying AlexiLearn is to let the user earn points for various tasks in order to buy upgrades, which in turn allow them to earn more points. This cycle can motivate the user to engage in the educational activities provided in the app.

Another way to gamify the FER system is to let the user practice expressing emotions to earn points. In the identification screen, the user may choose to play the "minigame," which prompts them to express a random emotion of those that the FER model can recognize in 8 seconds. If the model recognizes that emotion in any frame, the user is awarded 100 points; otherwise, they lose 50 points. The option to upload an image and receive a prediction was also added to the FER system.

The next educational feature is the Practice section, in which the user may choose a quiz length, and they are given a set of questions where they are asked to either identify the emotion in an image, choose an image with a specified emotion, or choose an image that matches another image's facial expression. These images, from the AffectNet dataset, are stored in Imagekit and accessed randomly by their URL. As a user gets more consecutive questions correct, the points they earn increment by their streak amount. The user's accuracy is tracked and stored, including the emotion-specific accuracy. Finally, the user's accuracy is stored every day, and a streak of the number of successive days the user completes a practice is stored.

Another important feature is the daily emotional reflection, which prompts the user to optionally state which of the seven basic emotions they think they feel and why. This information is also stored. This is intended to assist the user in identifying and expressing their emotions as well as recognizing trends in the events they are generally caused by.

The statistics screen displays the user's average accuracy, accuracy per emotion, a graph of daily accuracy progression, their practice streak, and a calendar containing the user's daily emotional reflections. A user may spend the points they gather in the upgrade store, where they can increase points per question, streak bonus, and insurance for incorrect answers.

The Learn section includes two important features: the emotion descriptions and the lessons. The descriptions allow the user to view thousands of examples of each facial expression and read more about them, helping users gain an in-depth understanding of each emotion. This includes a

description of its facial expression, the sensations felt, behavioral cues in others, real-life examples, coping strategies, and triggers. One user commented, “Nice feature of describing the situations and reasons behind the facial expressions,” highlighting the value of these detailed descriptions in helping users connect with the material on a deeper level. Another user stated that the examples in the Learn section were “really useful,” emphasizing how the information and content within the Learn section support effective learning.

The lessons are another useful way to learn about emotions in detail. Firebase was used to store the lesson data due to its minimal backend requirements. Each emotion has several lessons, each with multiple pages containing text, images, videos, and questions, which allow the user to explore the information in an interactive and engaging way.

4. Results

It was found that the models trained with augmented data resulted in much better generalization and overall accuracy, with especially large improvements in the emotions with less data.

Rather than only testing the accuracy of the model, the entire system was tested in order to get a more accurate representation of each part of the system functioning together. It was tested on a diverse set of 700 non-copyrighted images. The mean accuracy was 80.714%. Happiness was 99%, surprise was 89%, anger was 83%, fear was 80%, sadness was 77%, neutrality was 70%, and disgust was 67%.

Table 2. Final tested accuracy of the FER system.

Emotion	Accuracy
Happiness	99%
Neutrality	89%
Sadness	83%
Fear	80%
Anger	77%
Surprise	70%
Disgust	67%
Mean	80.714%

The face detection functions well in both well-lit and dark environments, with no significant delay between frames due to the fast facial expression recognition and face detection models.

5. Conclusion

This project consisted of developing AlexiLearn, an educational Android application tailored for individuals with alexithymia and autism spectrum disorder, as well as its main feature: a real-time facial expression recognition system. The transfer learning model was trained using an augmented FER-2013 dataset and MobileNet-V2 as the base model. It achieved a mean accuracy of 80.714%. Google’s ML Kit face detection was implemented and allowed the FER model to run efficiently in real-time on mobile devices in bright and dark lighting conditions. The use of other features allowed

for a more diverse, in-depth, and entertaining learning experience, addressing various aspects of emotional intelligence that individuals with these conditions struggle with.

6. Future Work

While AlexiLearn focuses on teaching facial expressions and other key aspects of emotions, it currently offers limited support for users to explore their emotions in depth. Although the daily emotional reflection feature attempts to address this, there is significant potential for improvement through the use of large language models (LLMs). Recent advancements in LLMs have been substantial, and many high-performing models are now open-source. These models have demonstrated considerable emotional intelligence, excelling in emotion recognition, interpretation, and understanding [18]. A potential enhancement to AlexiLearn's effectiveness and interactivity could involve integrating an LLM tailored specifically for emotional exploration. More specifically, users could engage in discussions about significant daily events or sensations, and the LLM could guide them toward identifying the emotions they experienced and understanding their underlying causes. This feature could be combined with the user's daily emotional reflections, enabling the LLM to gain a detailed understanding of their mood, allowing it to provide more personalized, context-aware responses.

Another possible improvement to AlexiLearn is in the facial expression recognition system's accuracy. While the current 80.714% accuracy is sufficient for most use cases, there are various techniques, including the use of landmark detection, that could be used to further improve the system. The accuracy of the model could also be improved by using more data in addition to the augmented FER-2013 dataset. The combination of datasets would result in a significant increase in the size of the data and hence a better-performing model.

Lastly, the current model is trained on seven basic emotions; however, with the use of more datasets, this could be increased, allowing the model to recognize more nuanced emotions and help users understand the true depth and diversity of emotions.

References

1. J. Hogeveen and J. Grafman, "Alexithymia," *Handbook of Clinical Neurology*, vol. 183, pp. 47–62, 2021, doi: <https://doi.org/10.1016/b978-0-12-822290-4.00004-9>
2. M. Joukamaa, A. Taanila, J. Miettunen, J. T. Karvonen, M. Koskinen, and J. Veijola, "Epidemiology of alexithymia among adolescents," *Journal of Psychosomatic Research*, vol. 63, no. 4, pp. 373–376, Oct. 2007, doi: <https://doi.org/10.1016/j.jpsychores.2007.01.018>
3. J. D. A. Parker, G. J. Taylor, and M. Bagby, "Alexithymia and the Recognition of Facial Expressions of Emotion," *Psychotherapy and Psychosomatics*, vol. 59, no. 3–4, pp. 197–202, 1993, doi: <https://doi.org/10.1159/000288664>
4. J. Hogeveen and J. Grafman, "Alexithymia," *Handbook of Clinical Neurology*, vol. 183, pp. 47–62, 2021, doi: <https://doi.org/10.1016/b978-0-12-822290-4.00004-9>
5. D. Tamas, Nina Brkic Jovanovic, Stanka Stojkov, Danijela Cvijanović, and Bozana Meinhardt-Injac, "Emotion recognition and social functioning in individuals with autism spectrum condition and intellectual disability," *PloS one*, vol. 19, no. 3, pp. e0300973–e0300973, Mar. 2024, doi: <https://doi.org/10.1371/journal.pone.0300973>
6. J. S. Ogrodniczuk, W. E. Piper, and A. S. Joyce, "Effect of alexithymia on the process and outcome of psychotherapy: A programmatic review," *Psychiatry Research*, vol. 190, no. 1, pp. 43–48, Nov. 2011, doi: <https://doi.org/10.1016/j.psychres.2010.04.026>
7. O. Ekundayo and S. Viriri, "Facial Expression Recognition: A Review of Trends and Techniques," *IEEE Access*, pp. 1–1, 2021, doi: <https://doi.org/10.1109/access.2021.3113464>
8. Adeviye Aydın, RN, B. Ersoy, and Y. Kaya, "The Effect of an Emotion Recognition and Expression Program on the Alexithymia, Emotion Expression Skills and Positive and Negative Symptoms of Patients with

- Schizophrenia in a Community Mental Health Center," *Issues in Mental Health Nursing*, vol. 45, no. 5, pp. 528–536, Apr. 2024, doi: <https://doi.org/10.1080/01612840.2024.2326951>
9. M. Xu, W. Cheng, Q. Zhao, L. Ma, and F. Xu, "Facial expression recognition based on transfer learning from deep convolutional networks," Aug. 2015, doi: <https://doi.org/10.1109/icnc.2015.7378076>
 10. L. Hu and Q. Ge, "Automatic facial expression recognition based on MobileNetV2 in Real-time," *Journal of Physics: Conference Series*, vol. 1549, p. 022136, Jun. 2020, doi: <https://doi.org/10.1088/1742-6596/1549/2/022136>
 11. S. Kaur and N. Kulkarni, "FERFM: An Enhanced Facial Emotion Recognition System Using Fine-tuned MobileNetV2 Architecture," *IETE journal of research*, pp. 1–15, Apr. 2023, doi: <https://doi.org/10.1080/03772063.2023.2202158>
 12. J. X.-Y. Lek and J. Teo, "Academic Emotion Classification Using FER: A Systematic Review," *Human Behavior and Emerging Technologies*, vol. 2023, p. e9790005, May 2023, doi: <https://doi.org/10.1155/2023/9790005>
 13. N. Yalçın and M. Alisawi, "Introducing a novel dataset for facial emotion recognition and demonstrating significant enhancements in deep learning performance through pre-processing techniques," *Heliyon*, vol. 10, no. 20, p. e38913, Oct. 2024, doi: <https://doi.org/10.1016/j.heliyon.2024.e38913>
 14. I. J. Goodfellow et al., "Challenges in representation learning: A report on three machine learning contests," *Neural Networks*, vol. 64, pp. 59–63, Apr. 2015, doi: <https://doi.org/10.1016/j.neunet.2014.09.005>
 15. Mehrnoosh Khoshnoodifar, Asieh Ashouri, and M. Taheri, "Effectiveness of Gamification in Enhancing Learning and Attitudes: a Study of Statistics Education for Health School Students.," *PubMed*, vol. 11, no. 4, pp. 230–239, Oct. 2023, doi: <https://doi.org/10.30476/jamp.2023.98953.1817>
 16. K. Krishnamurthy et al., "Benefits of gamification in medical education," *Clinical Anatomy*, vol. 35, no. 6, Jun. 2022, doi: <https://doi.org/10.1002/ca.23916>
 17. A. Närä, "How games and gamification can affirm emotional and social skills in children: case: MeikäMeikä-game's functionality," *Urn.fi*, 2022, doi: <http://www.theseus.fi/handle/10024/787882>
 18. X. Wang, X. Li, Z. Yin, Y. Wu, and J. Liu, "Emotional intelligence of Large Language Models," *Journal of Pacific Rim Psychology*, vol. 17, Jan. 2023, doi: <https://doi.org/10.1177/18344909231213958>

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.