big data and cognitive computing

MDPI

Article

Eliciting Emotions: Investigating the Use of Generative AI and Facial Muscle Activation in Children's Emotional Recognition

Manuel A. Solis-Arrazola [1], Raul E. Sanchez-Yanez [1], Ana M. S. Gonzalez-Acosta [2], Carlos H. Garcia-Capulin [1] and Horacio Rostro-Gonzalez [1,3,*]

1 Department of Electronics Engineering, DICIS-University of Guanajuato, Carretera Salamanca-Valle de Santiago km 3.5 + 1.8 kms, Salamanca 36885, Mexico; ma.solisarrazola@ugto.mx (M.A.S.-A.); sanchezy@ugto.mx (R.E.S.-Y.); carlosg@ugto.mx (C.H.G.-C.)

2 Laboratory of Artificial Intelligence, Robotics and Control, Faculty of Biological Systems and Technological Innovations, Benito Juárez Autonomous University of Oaxaca, Av. Universidad S/N, Ex-Hacienda 5 Señores, Oaxaca 68120, Mexico; goaa800904.fmc@uabjo.mx

3 GEPI Research Group, IQS-School of Engineering, Ramon Llull University, Via Augusta 390, 08017 Barcelona, Spain

* Correspondence: horacio.rostro@iqs.url.edu

Abstract: This study explores children's emotions through a novel approach of Generative Artificial Intelligence (GenAI) and Facial Muscle Activation (FMA). It examines GenAI's effectiveness in creating facial images that produce genuine emotional responses in children, alongside FMA's analysis of muscular activation during these expressions. The aim is to determine if AI can realistically generate and recognize emotions similar to human experiences. The study involves generating a database of 280 images (40 per emotion) of children expressing various emotions. For real children's faces from public databases (DEFSS and NIMH-CHEFS), five emotions were considered: happiness, angry, fear, sadness, and neutral. In contrast, for AI-generated images, seven emotions were analyzed, including the previous five plus surprise and disgust. A feature vector is extracted from these images, indicating lengths between reference points on the face that contract or expand based on the expressed emotion. This vector is then input into an artificial neural network for emotion recognition and classification, achieving accuracies of up to 99% in certain cases. This approach offers new avenues for training and validating AI algorithms, enabling models to be trained with artificial and real-world data interchangeably. The integration of both datasets during training and validation phases enhances model performance and adaptability.

Keywords: generative artificial intelligence; facial emotion recognition; facial muscle activation; artificial neural networks

check for updates

Academic Editors: Alberto Abelló and Giuseppe Maria Luigi Sarnè

Received: 6 November 2024
Revised: 6 January 2025
Accepted: 14 January 2025
Published: 20 January 2025

Citation: Solis-Arrazola, M.A.; Sanchez-Yanez, R.E.; Gonzalez-Acosta, A.M.S.; Garcia-Capulin, C.H.; Rostro-Gonzalez, H. Eliciting Emotions: Investigating the Use of Generative AI and Facial Muscle Activation in Children's Emotional Recognition. *Big Data Cogn. Comput.* **2025**, *9*, 15. https://doi.org/10.3390/bdcc9010015

1. Introduction

In recent years, Artificial Intelligence (AI) has made significant progress in understanding and replicating complex human behaviors such as emotions [1–3]. This research delves into the intersection of emotion recognition in children [4], Facial Muscle Activation (FMA) [5], and the innovative use of generative AI [6–9]. The central inquiry guiding this study reveals the capability of AI to effectively generate and discern emotions in children's faces, thereby contributing to a deeper understanding of human behavioral analysis.

Children, inherently expressive beings, manifest a diverse spectrum of emotions crucial for their social and cognitive development [10–16]. Traditional methods of emotion recognition often rely on facial expressions, but the dynamic and nuanced nature of children's emotional expressions necessitates a more precise approach [17–20], particularly those facing challenges such as autism that may impact their communication skills [21]. Leveraging Facial Muscle Activation, which focuses on the subtle movements of facial muscles, provides a nuanced understanding of the intricate interplay of emotions on young faces [22–24].

This study builds upon the existing body of work by integrating Generative AI and Facial Muscle Activation to explore emotion recognition in children. By generating a dataset of AI-generated facial expressions and combining it with real-world data, this research not only expands the scope of emotions used for training but also enhances the performance and adaptability of emotion recognition systems. The novel integration of synthetic and real-world data sets presents a promising direction for advancing emotion recognition technology, with applications ranging from behavioral science to emotion-aware AI systems.

To complement this analysis, generative AI, renowned for its ability to synthesize realistic content, is introduced as a novel tool [7,25]. The central question guiding this research is whether AI can generate facial images that authentically evoke emotions in a manner indistinguishable from genuine human expressions. This inquiry ignites a captivating exploration into the realm of emotional authenticity in AI-generated content.

This research comprises a multifaceted exploration with the following primary objectives:

1. Enhance the understanding of children's and adolescents' emotional expressions through precise Facial Muscle Activation [22], innovatively determined solely from images, eliminating the need for sensor-based methods prone to discomfort and recording errors [26].
2. Utilize generative AI with Midjourney [27] to create images of children expressing seven emotions (anger, disgust, fear, happiness, sadness, surprise, and neutrality) based on Ekman's framework [28], extending current databases that are limited to five emotions. For AI, generating these emotions is easier, as it does not require the actual experience of the emotion, unlike humans who often need to feel the emotion to express it effectively.
3. Investigate the effectiveness of emotion recognition methods based solely on muscular activation using real and artificially generated images. This is achieved by implementing and evaluating two neural network training approaches: the first involves training a multilayer perceptron neural network using real databases (NIMH-CHEFS [29] and DEFSS [30]) of children and adolescents expressing emotions, followed by performance evaluation with both real and artificially generated images. The second approach trains the neural network with a database of artificially generated images (Midjourney) and validates its performance using both real and artificial images, exploring the model's adaptability across different data types.

2. Related Work

A significant portion of the literature has focused on using traditional datasets, such as NIMH-CHEFS and DEFSS, which include photographs of children and adolescents expressing a limited variety of emotions like happiness, sadness, anger, and fear. These datasets have provided foundational resources for training machine learning models aimed at identifying emotions based on facial expressions [31–34]. However, a limitation of these methods is the use of databases that are limited in size, diversity, and coverage of all emotional states, which has prompted the exploration of alternative methods, such as Generative AI (GenAI), to address these gaps. In this context, the use of Generative AI for creating realistic images to evoke emotions in people has gained traction in recent years [35–37]. Tools like Midjourney have made it possible to generate images of people (including children) displaying a wide array of emotions based on textual descriptions. This method of emotion elicitation offers significant advantages over traditional datasets. It eliminates the challenges of obtaining ethically sourced, varied, and high-quality data from human participants, particularly when it comes to sensitive populations such as children. Previous research has shown that synthetic facial expressions, when used for emotion recognition tasks, can achieve comparable or even superior results to real-world data [38–40]. By leveraging GenAI, it becomes possible to expand the range of emotions available for training AI models, as we will demonstrate below, a critical factor in enhancing the robustness and accuracy of emotion recognition systems. Furthermore, a key concept in emotion recognition is Facial Muscle Activation (FMA), which refers to the movements of facial muscles that correspond to specific emotional expressions [41,42]. Early work by Ekman and Friesen [43] introduced the Facial Action Coding System (FACS), which provides a detailed analysis of facial movements associated with emotional states. More recent developments in FMA have incorporated machine learning algorithms to track facial muscle activation patterns, offering a more precise understanding of how emotions manifest in facial expressions [22,44]. For instance, the activation of muscles like the zygomaticus major (for happiness) and corrugator supercilia (for anger) are commonly associated with specific emotional states, and analyzing these movements can improve the accuracy of emotion recognition systems [45,46]. Although much of the research has focused on either real or synthetic data, the combination of both in several applications of Artificial Intelligence has received increasing attention. Recent studies have demonstrated that integrating real and synthetic data can enhance the generalization ability of emotion recognition models, making them more robust and adaptable [47,48]. This approach is particularly valuable when training models for complex tasks, such as recognizing subtle emotional differences, or when working with limited datasets. By combining real-world data with AI-generated images, models can be trained to recognize a wider range of emotional expressions, thus improving performance across diverse scenarios. Khan et al. (2024) highlighted the benefits of combining real and synthetic data in training machine learning models, showing that this technique can significantly enhance model accuracy and applicability in various technologies [49].

Despite these advancements, the application of Generative AI and FMA in emotion recognition remains largely unexplored. Furthermore, there are no existing databases or algorithms that utilize AI-generated images for training or validating their models. While generated images have been used to evoke emotional responses in individuals, there is a lack of images specifically designed to enhance emotion recognition algorithms, as proposed in this study. Moreover, this work uniquely investigates whether muscle activation features, commonly extracted from real human images, can also be derived from AI-generated images.

3. Materials and Methods

Our methodology consists of three stages. The first involves obtaining and generating images of frontal faces of children depicting some of the seven aforementioned emotions. The subsequent stage aims to determine a feature vector based on facial muscle activation generated when children in the images express an emotion. This feature vector, in turn, serves the third stage, which entails an artificial neural network learning the relationships between the muscles activated during the expression of emotions. These stages are detailed below and visually represented in Figure 1.

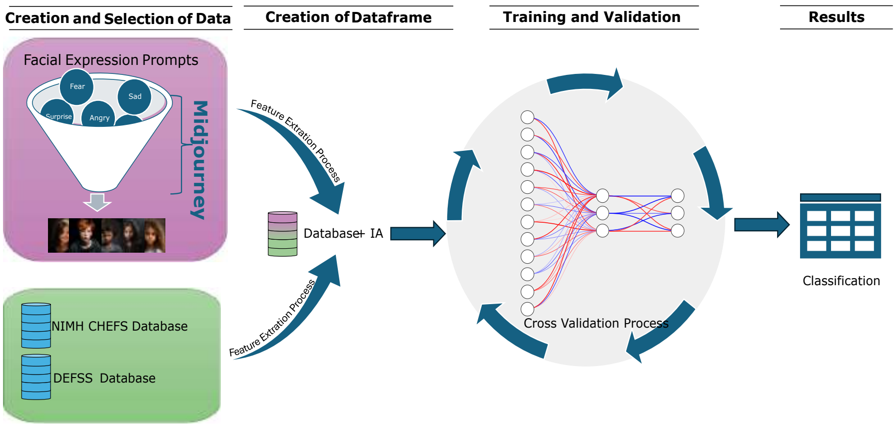

Figure 1. General diagram of the proposed methodology.

3.1. Data

Public databases for research involving people, especially children, are severely limited, primarily due to concerns related to their protection and ethical considerations. This challenge has served as the impetus for the current study. In this section, we elucidate the process of generating images featuring children expressing various emotions for research purposes. Despite these constraints, we have managed to gain access to two public databases (DEFSS and NIMH-CHEFS) featuring images of real children, enabling the validation of our study.

3.1.1. Database of Children's Faces Generated Using Generative AI

One crucial aspect under investigation in this study is the utilization of artificially generated images for validating AI algorithms. This exploration aims to assess algorithm performance with such synthetic images or data, recognizing the challenges of obtaining real data, especially when dealing with large datasets or requiring sophisticated equipment. Additionally, this approach aims to streamline the information collection process, reducing the time devoted to gathering data and enabling a more rapid validation of our algorithms. In this context, a database comprising 240 images of children expressing the aforementioned emotions has been generated using Midjourney [50]. Midjourney is a generative artificial intelligence program and service developed and hosted by the San Francisco-based independent research lab, Midjourney, Inc. Similar to OpenAI's DALL-E [51] and Stability AI's Stable Diffusion [52], Midjourney generates images based on natural language descriptions, commonly referred to as prompts [53].

The images shown in Figures 2–8 were generated on the previously described Midjourney platform using prompts. Each set of images features children's faces expressing a specific emotion, with the accompanying prompt also provided. These prompts can be entirely adaptable for different ages, races, and very specific features that one may want to assess on faces, such as nose, eyes, mouth, etc. (see Table A1). While our prompts are centered on generating images of faces portraying 8-year-old children, the algorithm for feature extraction and subsequent emotion recognition and classification exhibits a high level of robustness, allowing it to effectively handle diverse age groups. Hence, alongside the generation of these images, we incorporated other databases featuring images captured from real children. The details of these databases will be expounded upon in the next subsection. It is crucial to note at this juncture that access to databases of this nature is significantly limited. Consequently, the experiment primarily targets children, with validation extending to adolescents and/or young adults due to the constrained availability of appropriate datasets.

Figure 2. "Fear" expression captured in artificially generated children's faces (Table A1).

Figure 3. "Sad" expression captured in artificially generated children's faces (Table A1).

Figure 4. "Happy" expression captured in artificially generated children's faces (Table A1).

Figure 5. "Surprise" expression captured in artificially generated children's faces (Table A1).

Figure 6. "Angry" expression captured in artificially generated children's faces (Table A1).

Figure 7. "Disgust" expression captured in artificially generated children's faces (Table A1).

Figure 8. "Neutral" expression captured in artificially generated children's faces (Table A1).

3.1.2. Databases of Real Children's Faces

In this study, we explored two publicly available databases that showcase children expressing a spectrum of emotions. However, a limitation of these databases is twofold. Firstly, they are not exclusively dedicated to children, encompassing a broader age range. Secondly, these databases are restricted to portraying only five distinct emotions. In this regard, the NIMH-CHEFS database [29] stores 482 images of children aged between 10 and 17 years, exhibiting expressions of fear/being afraid, anger, happiness, sadness, and neutrality (refer to Figure 9). Additionally, we utilized the DEFSS database [30], comprising 404 validated facial photographs of individuals aged between 8 and 30 years. For the purposes of our study, we selectively employed 74% of the dataset, focusing on the age range relevant to our research (see Figure 10). The subjects in this dataset express five distinct emotional expressions: happiness, anger, fear, sadness, and neutrality, similar to NIMH-CHEFS.

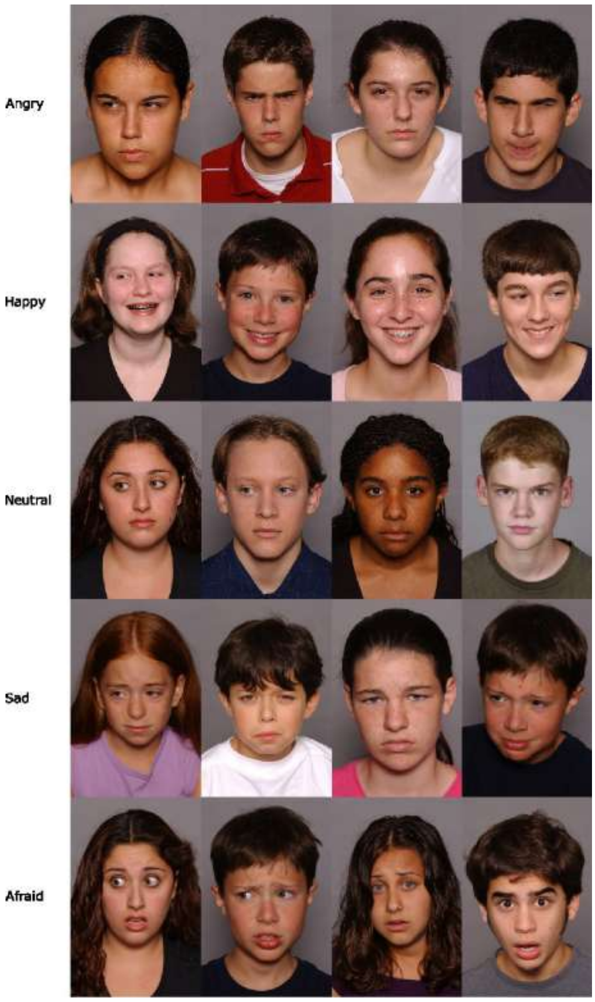

Figure 9. Sample images depicting various emotions found in the NIMH-CHEFS database.

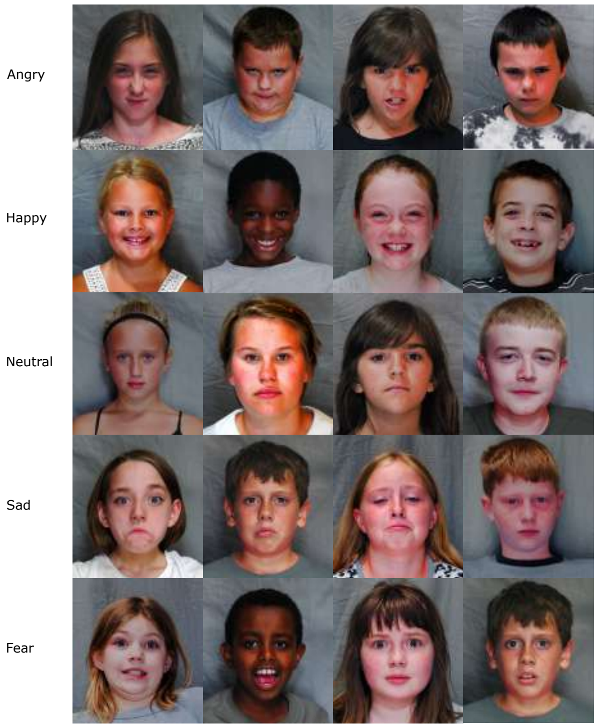

Figure 10. Sample images depicting various emotions found in the DEFSS database.

3.2. Feature Extraction

From the images presented in the previous section, we construct our feature vector, which will, in turn, be used to train an artificial neural network, as discussed in the next section. This concept was extended from the work presented in [22]. In this regard, the steps for the generation of the feature vector are described below:

1. Landmark detection: This process involves identifying 68 specific points on the face, including the eyes, nose, mouth, and other distinctive features. This step is performed by using the Dlib standard library [54], which provides a comprehensive set of landmarks, allowing for a detailed representation of facial expressions. These landmarks serve as crucial reference points for subsequent analyses, enabling us to map the nuanced muscular activities associated with different emotions. This meticulous landmark detection process lays the foundation for our in-depth investigation into children's emotional expressions through Facial Muscle Activation (FMA). Results for this step are illustrated in Figure 11 for both types of children's faces, generated and real.

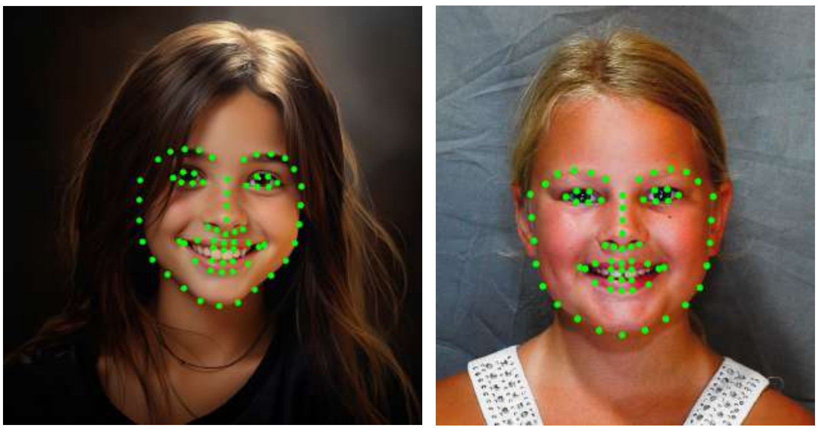

Figure 11. Locations of the standard 68-point facial landmarks in generated and real children's faces.

2. Centroids estimation: In this step, we calculate the centroids of specific triangles conformed with triads from the landmark set. The centroids, representing the geometric centers of specific facial regions, provide a condensed yet informative representation of the facial features. By determining the centroids, we gain valuable insights into the spatial relationships among the facial landmarks. This step is integral to our subsequent analysis, particularly in quantifying the variations in facial muscle activation during emotional expressions in children. The centroid estimation process contributes to a more focused and efficient representation of facial dynamics, enhancing the accuracy of our investigation into the intricacies of emotional recognition. This step is depicted in Figure 12.

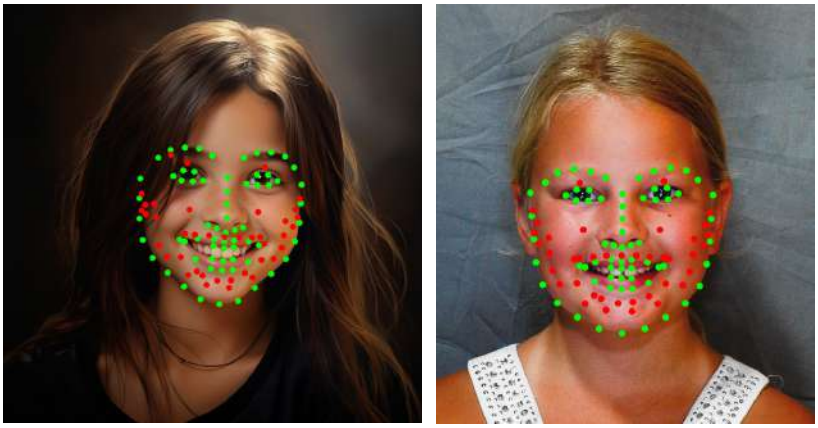

Figure 12. Locations of the estimated centroids (red dots) in generated and real children's faces.

3. FMA calculation: In the calculation of Facial Muscle Activation (FMA), we leverage the previously identified centroids to quantify the dynamic changes in facial muscle movements during emotional expressions. FMA involves measuring the variations in distances between the centroids associated with specific facial regions. By monitoring these spatial changes, we gain insights into the nuanced muscular activities responsible for conveying various emotions. This meticulous calculation process allows us to discern patterns of muscle activation, providing a comprehensive understanding of how different emotions manifest in the facial expressions of children. The FMA analysis, depicted in Figure 13, serves as a critical link in our exploration of emotion recognition, bridging the gap between facial landmarks and the underlying muscular dynamics. A comprehensive set of 56 features was computed and categorized across three distinct facial regions (upper, middle, and lower) encompassing measurements such as facial width and angle.

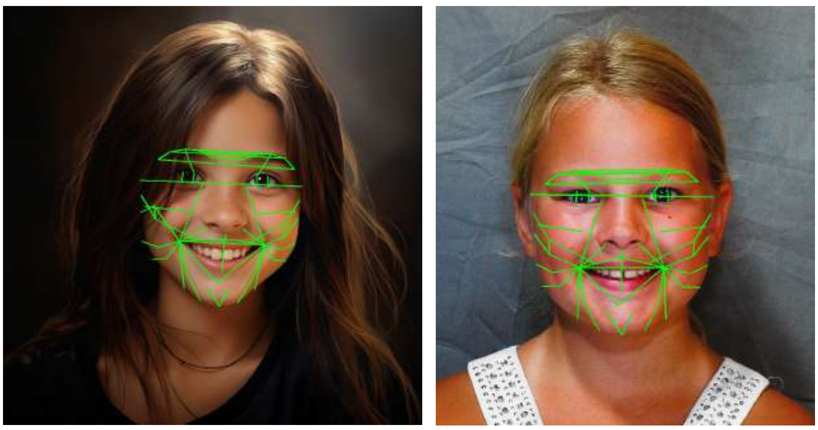

Figure 13. Segments connecting landmarks and centroids in generated and real children's faces.

3.3. Emotion Recognition and Classification

An Artificial Neural Network (ANN) was employed to recognize and categorize emotions based on the extracted feature vector. The ANN (Figure 14) was trained in WEKA [55] using the parameters shown in Table 1 to discern patterns within the intricate dataset, enabling it to associate specific combinations of facial features with distinct emotional expressions in children. This trained model serves as a robust classifier, contributing to the accurate identification of emotions from facial data. The utilization of an ANN in this phase signifies a pivotal aspect of our methodology, as it encapsulates the mixture of intricate facial muscle activations into a coherent framework for emotion classification. The results of this classification process are elaborated upon and presented in subsequent section, shedding light on the efficacy of our approach in accurately recognizing and classifying emotions in children. Refer to Figure 15 for a screenshot displaying the algorithm's execution.

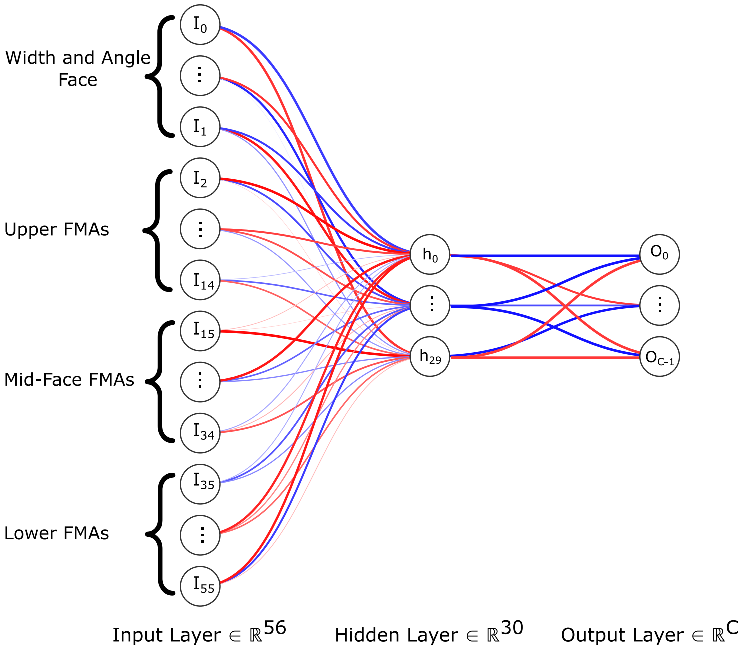

Figure 14. Artificial Neural Network architecture: The network consists of 56 input neurons, 30 neurons in the hidden layer, and C neurons in the output layer. The value of C varies depending on the number of emotions, with 5 for real children's faces and 7 for artificially generated faces, for instance.

Figure 15. Screenshot displaying the algorithm's execution.

Different configurations of the data were implemented during the training and validation of the neural network. For this purpose, real children's data (see Figures 9 and 10) were utilized to train the network, which was subsequently validated using artificially generated children's faces (see Figures 2–8). Additionally, the inverse scenario was explored, training the network with generated images and validating it with real children's faces. Finally, a cross-validation approach was adopted by merging real and generated children's data. The outcomes of these experiments are presented in the following section. In Table 2, the total number of images used and how they were distributed across the training and testing datasets are shown.

Table 1. Parameters used by the classification algorithms.

Method	Parameters	Weka Function
MLP	Learning rate (L = 0.3), momentum (M = 0.2), training time (N = 500), neurons in the hidden layers (H = 31), and seed (S)	MLP

Table 2. Distribution of real and artificially generated images used for training and testing the neural network, including the total number of images for each dataset configuration.

Source	Instances	Training	Validation
NIMH-CHEFS	533	426	107
DEFSS	270	216	54
IA	294	235	59
NIMH-CHEFS + IA	827	661	166
DEFSS + IA	564	451	113

4. Results

Our method, as previously mentioned, was validated using both real and generated images. In this section, we specifically highlight the performance of our algorithm in recognizing and classifying emotions based on the associated muscle activation patterns. These patterns are learned by an artificial neural network, enabling it to discern the expressed emotion solely from an image of a child. The method was quantitatively evaluated, and the ensuing results are delineated below for various combinations of training and validation. To assess the efficacy of our proposal, four statistical measures, Accuracy, Precision, Recall, and F1 score, were utilized in the validation process.

4.1. Results with Real Children Faces

The initial experiment focused on training and validating our method using two publicly available databases featuring real children expressing five emotions (Happy, Angry, Fear, Sad, and Neutral). The results of this experiment are shown in Table 3. Notably, in the case of the NIMH-CHEFS database, although apparently centered on children, the age range spans from 10 to 17 years old. Observing the results, it became apparent that the "Happy" and "Fear" classes demonstrated the most favorable outcomes, while the "Sad" class exhibited the least success. Subsequently, a decision was made to exclude the "Sad" class from both the training and validation sets. This strategic omission led to a significant improvement in results for the other two classes that had initially shown suboptimal performance, specifically, "Neutral" and "Angry" (see Table 4).

The performance across the NIMH-CHEFS and DEFSS databases was also assessed by mapping their classes to arousal and valence categories (see Table 5). Emphasizing why it is crucial to consider these classes during the analysis of muscle activation. This approach is essential for comprehending and contextualizing muscle activation in emotion analysis. The relationship between classes and arousal and valence categories provides a deeper insight into how emotions are expressed in children's faces, allowing for a more comprehensive interpretation of emotional dynamics [56]. By mapping these classes to arousal and valence dimensions, our analysis is enriched, considering not only specific facial expressions but also the intensity and emotional polarity. This connection between classes and categories contributes to a nuanced understanding of muscle activation associated with different emotions, thus enhancing the interpretation of the obtained results.

Table 3. Performance across the NIMH-CHEFS and DEFSS databases for all classes.

Source	Metric	Happy	Angry	Fear/Afraid	Sad	Neutral
NIMH-CHEFS	Accuracy (%)	99.81	88.15	97.13	85.85	89.29
	Precision	0.99	0.75	0.92	0.58	0.75
	Recall	1.00	0.68	0.93	0.66	0.74
	F1 Score	1.00	0.71	0.93	0.61	0.75
DEFSS	Accuracy (%)	96.57	92.65	87.75	82.84	81.37
	Precision	0.96	0.73	0.64	0.14	0.77
	Recall	0.91	0.80	0.66	0.27	0.61
	F1 Score	0.94	0.76	0.65	0.19	0.68

Table 4. Performance across the NIMH-CHEFS and DEFSS databases without *Sad* class.

Source	Metric	Happy	Angry	Fear/Afraid	Neutral
NIMH-CHEFS	Accuracy (%)	99.76	91.92	96.67	90.74
	Precision	0.99	0.83	0.94	0.82
	Recall	1.0	0.83	0.92	0.82
	F1 Score	1.0	0.93	0.93	0.82
DEFSS	Accuracy (%)	94.32	91.48	83.52	87.50
	Precision	0.93	0.73	0.53	0.85
	Recall	0.89	0.80	0.61	0.76
	F1 Score	0.91	0.76	0.57	0.80

Table 5. Performance across the NIMH-CHEFS and DEFSS databases by mapping their classes with the arousal and valence categories.

Source	Metric	Activation	Deactivation	Pleasant	Unpleasant
NIMH-CHEFS	Accuracy (%)	96.18	89.87	100	90.63
	Precision	0.90	0.73	1.0	0.89
	Recall	0.90	0.77	1.0	0.87
	F1 Score	0.90	0.75	1.0	0.88
DEFSS	Accuracy	87.75	82.35	98.04	83.82
	Precision	0.53	0.77	1.0	0.67
	Recall	0.70	0.63	0.93	0.76
	F1 Score	0.60	0.69	0.96	0.71

Validation with Generated Children's Faces

The subsequent phase in validating our results involved the utilization of artificially generated images, although the training itself was conducted exclusively with real images. For this validation, the neural network was trained with the NIMH-CHEFS database since it was the one that best covered the age defined for the generated images. The results are presented in Table 6. The findings reveal that the use of both real and artificial data can significantly enhance the robustness and generalization capabilities of our model. This strategic integration not only broadens the diversity of the dataset but also fortifies the model's ability to extrapolate and accurately classify emotions in novel instances. The synergy between real and artificial data serves as a key component in fortifying the model's adaptability, contributing to a more reliable and comprehensive emotion recognition framework. Similarly to the results obtained with images featuring real children, optimal outcomes were observed in the "Happy" and "Fear" classes, while the "Sad" class exhibited the least favorable performance.

Table 6. Performance of the validation phase using generated images.

Source	Metric	Happy	Angry	Fear/Afraid	Sad	Neutral
NIMH-CHEFS	Accuracy (%)	95.45	90.91	95.45	81.82	90.91
	Precision	1.0	0.60	1.00	0.75	0.50
	Recall	0.80	1.00	0.83	0.50	1.00
	F1 Score	0.89	0.75	0.91	0.60	0.67

4.2. Results with Generated Children's Faces

One of the most significant findings of this investigation is the utilization of generative Artificial Intelligence. Through this method, we have constructed a database of artificially generated children's faces, each portraying seven distinct emotions. This advancement not only broadens the emotional spectrum under study, surpassing the typical focus on five emotions in publicly available databases, but also confronts the sensitive issue of utilizing children's imagery. The outcomes stemming from this aspect of the study are striking, yielding classification results surpassing 90% for all emotional categories. Results of this experiment are shown in Table 7.

Table 7. Performance during training and validation across all classes for the AI-generated images.

Source	Metric	Happy	Angry	Fear	Sad	Neutral	Disgust	Surprise
IA	Accuracy (%)	98.94	98.23	92.23	94.35	96.11	97.17	94.7
	Precision	1.0	0.95	0.78	0.79	0.75	0.88	0.82
	Recall	0.94	0.93	0.74	0.86	0.71	0.92	0.84
	F1 Score	0.97	0.94	0.76	0.83	0.73	0.90	0.83

4.3. Results Combining Both Real and Generated Children's Faces

Based on the findings from both real and artificially generated children's faces, we opted to conduct an experiment aimed at combining both real and artificial data to enrich the datasets. This simulated scenarios where datasets contain limited data, resulting in artificial intelligence algorithms struggling to learn or perform poorly. The results are highly promising, with the majority of cases showing improvements over previous findings using solely real data. In other instances, the results remain unchanged, while a smaller proportion saw a decrease in performance. It is important to note that the incorporation of artificial data also increased the size of the dataset, leading to overall improved performance of the network. The results of this step are shown in Table 8. Similarly, we conducted validation with the same number of classes (or emotions) found in the databases with real faces, excluding the emotions of "surprise" and "disgust". In this regard, we also observe that the results perform quite well. In many cases, the outcomes are superior to using only real faces, as can be seen in Table 9. It is important to mention that in both tables, those results that showed lower performance compared to the use of images with real faces still have very good performance.

Table 8. Performance combining real children and generated facial expressions during training and validation, using all generated classes (including Disgust and Surprise). ▼ indicates that the obtained results perform worse than the data with only real faces. ▲ denotes better performance, while ■ indicates the same performance. The (*) indicates that this emotion was not present in the images with real children.

Source	Metric	Happy		Angry		Disgust		Fear/Afraid		Sad		Neutral		Surprise	
NIMH-CHEFS + IA	Accuracy (%)	99.81	99.64▼	88.15	93.23▲	*	98.55▲	97.13	95.04▼	85.85	88.27▲	89.29	90.93▲	*	98.53▲
	Precision	0.99	0.99■	0.75	0.81▲	*	0.78▲	0.92	0.87▼	0.58	0.63▲	0.75	0.79▲	*	0.91▲
	Recall	1.00	0.99▼	0.68	0.81▲	*	0.91▲	0.93	0.86▼	0.66	0.68▲	0.74	0.73▼	*	0.84▲
	F1 Score	1.00	0.99▼	0.71	0.81▲	*	0.84▲	0.93	0.86▼	0.61	0.65▲	0.75	0.76▲	*	0.87▲
DEFSS + IA	Accuracy (%)	96.57	97.87▲	92.65	92.38▼	*	96.28▲	87.75	90.43▲	82.84	89.54▲	81.37	89.89▲	*	97.16▲
	Precision	0.96	0.97▲	0.73	0.74▲	*	0.70▲	0.64	0.68▲	0.14	0.57▲	0.77	0.79▲	*	0.80▲
	Recall	0.91	0.93▲	0.80	0.75▼	*	0.76▲	0.66	0.72▲	0.27	0.63▲	0.61	0.71▲	*	0.84▲
	F1 Score	0.94	0.95▲	0.76	0.74▼	*	0.73▲	0.65	0.70▲	0.19	0.60▲	0.68	0.75▲	*	0.82▲

Table 9. Performance combining real children and generated facial expressions during training and validation (without Disgust and Surprise). ▼ indicates that the obtained results perform worse than the data with only real faces. ▲ denotes better performance, while ■ indicates the same performance.

Source	Metric	Happy		Angry		Fear/Afraid		Sad		Neutral	
NIMH-CHEFS + IA	Accuracy (%)	99.81	99.60▼	88.15	91.78▲	97.13	96.36▼	85.85	87.74▲	89.29	91.11▲
	Precision	0.99	0.99■	0.75	0.80▲	0.92	0.91▼	0.58	0.67▲	0.75	0.78▲
	Recall	1.00	0.99▼	0.68	0.78▲	0.93	0.91▼	0.66	0.69▲	0.74	0.78▲
	F1 Score	1.00	0.99▼	0.71	0.79▲	0.93	0.91▼	0.61	0.68▲	0.75	0.78▲
DEFSS + IA	Accuracy (%)	96.57	97.29▲	92.65	94.78▲	87.75	90.19▲	82.84	89.77▲	81.37	87.47▲
	Precision	0.96	0.98▲	0.73	0.85▲	0.64	0.75▲	0.14	0.60▲	0.77	0.74▼
	Recall	0.91	0.91■	0.80	0.76▲	0.66	0.74▲	0.27	0.72▲	0.61	0.71▲
	F1 Score	0.94	0.95▲	0.76	0.85▲	0.65	0.75▲	0.19	0.65▲	0.68	0.72▲

5. Strengths and Limitations

Our paper presents several strengths that make it a significant contribution to the field of emotion recognition, especially in children. One of the key strengths lies in the efficiency of our proposed method, which achieves excellent performance in recognizing and classifying emotions across both real and synthetic datasets. The use of generative artificial intelligence (GenAI) for creating synthetic images is particularly advantageous, as it helps to overcome the limitations of manually curated datasets where actors are asked to simulate emotions. In these datasets, emotions such as “Sad”, “Angry”, and “Neutral” often appear ambiguous, leading to challenges in both human and machine recognition. In contrast, GenAI-generated images provide more realistic and diverse emotional expressions, enhancing the system’s accuracy and reliability.

A significant practical strength of our system is its ability to perform real-time emotion recognition once the neural network is trained. This feature makes the system highly suitable for applications that require immediate feedback, such as in educational or therapeutic settings.

The architecture of our neural network also facilitates reproducibility. We have intentionally designed it to be straightforward and not overly complex, ensuring that researchers and practitioners can easily replicate our results. Additionally, the system was developed using WEKA, a user-friendly programming environment known for its ease of use, which further enhances the accessibility and reproducibility of our work.

Another important aspect is the democratization of synthetic image generation. Tools like MidJourney, DALL-E, or other similar platforms make it possible for virtually anyone with access to these technologies to generate synthetic images. This accessibility broadens the potential for widespread use and further development of emotion recognition systems based on synthetic data.

Regarding the limitations of our current system, it is important to note that it is designed to work primarily with static images. The next logical step would be to evaluate its performance with video data, which would simulate more dynamic and realistic scenarios. Additionally, our system currently performs best with front-facing images of faces.

Misaligned or angled images can reduce the algorithm's performance. However, this issue can be mitigated by using synthetic images to create a more robust training dataset that accounts for various angles and orientations.

In summary, our paper's strengths include the system's real-time analysis capabilities, its reproducible architecture, and the practical application of GenAI for synthetic image generation. These elements collectively advance the field of emotion recognition and offer a strong foundation for future research and application development.

6. Conclusions

In conclusion, our study investigated the complex domain of emotion recognition in children by integrating generative artificial intelligence and facial muscle activation analysis. The use of the Midjourney platform for generating artificial images of children expressing seven emotions enabled us to evaluate the adaptability and effectiveness of the proposed methodology.

Initial validation with real images of children revealed high accuracy in recognizing emotions such as "happy" and "fear", while the "sad" class presented challenges. These findings prompted refinements in the algorithm to enhance its performance. Incorporating AI-generated images into the training and validation process significantly improved the model's generalization capabilities, as demonstrated by increased accuracy, precision, recall, and F1 scores across various emotion categories.

By mapping results to arousal and valence dimensions, the study offered deeper insights into the interplay between facial muscle activation and emotional expression.

However, the study also acknowledged limitations, including the constrained availability of comprehensive databases of children's facial expressions. These constraints highlight the need for consideration of factors such as age, race, and specific facial features in future research. Moreover, challenges in accessing real databases and ethical concerns surrounding sensitive data, such as images of children, underscore the value of using AI-generated images, which mitigate privacy issues and streamline data collection.

7. Future Work

Building upon the findings of this study, future work will aim to expand the scope and applications of the proposed methodology. One key direction is to extend the approach to other populations, such as older adults. Prior research with real images has revealed limitations in the algorithm's ability to recognize certain emotional expressions in this demographic. Leveraging Generative AI may help address these challenges, offering the potential to enhance emotion recognition in populations with unique facial features or age-related changes.

Additionally, the methodology could be applied to the healthcare domain, particularly in areas such as autism detection, facial paralysis rehabilitation, and the identification of microexpressions associated with conditions like strokes. These applications would allow the proposed approach to contribute to critical areas where understanding and interpreting emotional expressions are vital for diagnosis, treatment, and patient care.

Another important avenue is the expansion of the existing dataset. Developing a more diverse and extensive database featuring a wider range of emotions, demographics, and facial features will improve the algorithm's robustness and generalization capabilities. This step is essential for ensuring that the system performs well across diverse real-world scenarios.

Finally, the methodology can be extended to dynamic emotion recognition by analyzing expressions in video sequences. This shift from static images to video would provide richer temporal data, enabling the system to capture and interpret emotional transitions more effectively. Incorporating video-based analysis could further refine the model and open up new possibilities for applications in areas like behavioral analysis and human–computer interaction.

Author Contributions: H.R.-G., C.H.G.-C. and R.E.S.-Y. contributed to Conceptualization, Formal analysis, Investigation, Supervision, and Writing—original draft. M.A.S.-A., A.M.S.G.-A. and H.R.-G. contributed to Data curation, Resources, Methodology, and Validation. M.A.S.-A., A.M.S.G.-A., H.R.-G. and R.E.S.-Y. contributed to Software, Visualization, and Writing—review and editing. H.R.-G. contributed to Funding acquisition, Project administration, and Supervision. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Informed Consent Statement: Not applicable.

Data Availability Statement: The NIMH-CHEFS (https://psychiatry.duke.edu/research/ (accessed on 5 November 2024)) and DEFSS (http://reflectionsciences.com/resources/researchers/ (accessed on 5 November 2024)) databases used in this research are publicly available for research purposes. The GenAI images are available in the following LINK (https://github.com/BioInspiredLab-UGTO/GenAI_images (accessed on 5 November 2024)).

Acknowledgments: This research has been supported by the National Council of Humanities, Sciences, and Technology of Mexico (CONAHCYT) through the scholarship 916172. During the preparation of this study, the authors used Midjourney for the purposes of image generation. The authors have reviewed and edited the output and take full responsibility for the content of this publication.

Conflicts of Interest: The authors declare no conflicts of interest.

Appendix A

Table A1. Midjourney prompts.

Emotion (Class)	Prompt
Fear	An ultra realistic 8K headshot of a 8-year-old child, capturing the intense emotion of fear. The child's eyes are wide open, and the pupils are dilated, fear vividly expressed in their gaze. The eyebrows are arched upward, creating tension lines on the forehead. The corners of the lips are slightly trembling, revealing the child's vulnerability. Focus on the facial muscles that convey fear, such as wrinkles on the forehead and around the eyes. Capture the tension in the muscles of the jaw and cheeks. Ensure the expression is realistic and compelling, effectively conveying the child's distress in that moment.
Sad	An ultra-realistic 8K photographic portrait of an 8-year-old child displaying an intense emotion of sadness. The photograph should authentically capture the essence of sadness, with the child's facial muscles and features convincingly reflecting the sorrowful expression. Pay special attention to aspects such as drooping eyelids, a lowered gaze, and a melancholic expression in the eyes. Observe the eyebrows, which should be slanted downward, creating lines of sorrow on the forehead. The muscles around the mouth should exhibit a slight downturn, reflecting sadness in the corners of the lips. Capture tension in the jaw and cheek muscles. Illuminate the image carefully to emphasize details and colors that convey the authentic emotion of sadness. The resulting photograph should be highly realistic, poignantly capturing the true essence of the child's sadness.
Happy	An ultra realistic 8K headshot of a 8 year old child displaying an intense emotion of joy. The photograph should capture the essence of happiness, with facial muscles convincingly reflecting the child's joyful expression. The eyes should gleam with joy, exhibiting wrinkles around the eye corners. Focus on the muscles of the upper lip and the corners of the lips, creating an authentic and visible smile. The cheek muscles should lift, adding a natural touch to the joyful expression. The forehead and eyebrows should display relaxation, without tension lines, to convey genuine happiness. Highlight the sparkle in the eyes and ensure that the overall expression of the child is open and relaxed. Illuminate the image to complement the joyful atmosphere, emphasizing details and colors that reflect the child's happiness. Capture the innocence and spontaneity of childhood joy, and ensure that the resulting image is an authentic and heartwarming representation of the child's positive emotion.

Table A1. *Cont.*

Emotion (Class)	Prompt
Surprise	An ultra-realistic 8K headshot a 8-year-old child displaying an intense emotion of surprise or astonishment. The photograph should authentically capture the essence of surprise, with the child's facial muscles convincingly reflecting the surprised expression. Focus on aspects such as wide- open eyes with dilated pupils and raised upper eyelids, along with raised eyebrows creating horizontal lines on the forehead to signify surprise. The child's mouth should be open in an "O" shape or a similar manner, clearly indicating surprise. If appropriate, ensure that the corners of the lips are slightly lifted, contributing to the overall expression of surprise. Make sure that the overall expression of the child genuinely and visibly reflects surprise. Illuminate the image thoughtfully to complement the surprised atmosphere, emphasizing details and colors that convey the authentic emotion of surprise. The resulting photograph should appear highly realistic, capturing the true essence of the child's surprise.
Angry	Generate an ultra-realistic 8K photographic portrait of an 8-year-old child displaying an intense emotion of anger, resembling a passport-style photograph. The background should be light gray, and the complete face of the subject must be visible. Focus on capturing specific changes in facial muscles associated with the expression of anger, such as raised eyebrows, contraction of the muscles between the eyebrows, narrowed eyes, an intense gaze, raised upper eyelids, nasal muscle contraction, narrowed lips with a possible thin line formation, and a tense jaw. Ensure that the photograph authentically conveys the child's anger intensity. Illuminate the image carefully to highlight details and colors that reflect the authenticity of the emotion.
Disgust	An ultra-realistic 8K photograph of an 8-year-old child making a disgusted pout because something displeased them. The expression should include wrinkling of the nose, furrowing of the brow, lifting of one side of the upper lip, and narrowing of the eyes in a manner that conveys repugnance. These elements should be pivotal and highly expressive in the photograph. Illuminate the image carefully to highlight details and colors that convey the authenticity of the child's emotion. The photograph should resemble a passport-style portrait, with a light gray background, showcasing the complete face of the subject in a close-up shot.
Neutral	An ultra-realistic 8K photographic portrait of an 8-year-old child with a neutral facial expression, devoid of any evident emotion. The image should capture serenity and the absence of emotional expression, highlighting the natural features of the child's face. Illuminate the image carefully to emphasize details and colors in a balanced manner. The photograph should feature a neutral background, and the complete face of the subject should be clearly visible in a close-up shot, akin to a passport-style portrait

References

1. Ali, O.; Abdelbaki, W.; Shrestha, A.; Elbasi, E.; Alryalat, M.A.A.; Dwivedi, Y.K. A systematic literature review of artificial intelligence in the healthcare sector: Benefits, challenges, methodologies, and functionalities. *J. Innov. Knowl.* **2023**, *8*, 100333. [CrossRef]
2. Zhao, J.; Wu, M.; Zhou, L.; Wang, X.; Jia, J. Cognitive psychology-based artificial intelligence review. *Front. Neurosci.* **2022**, *16*, 1024316. [CrossRef] [PubMed]
3. Bhattacharya, S. Artificial intelligence, human intelligence, and the future of public health. *AIMS Public Health* **2022**, *9*, 644–650. [CrossRef] [PubMed]
4. Negrao, J.G.; Osorio, A.A.C.; Siciliano, R.F.; Lederman, V.R.G.; Kozasa, E.H.; D'Antino, M.E.F.; Tamborim, A.; Santos, V.; de Leucas, D.L.B.; Camargo, P.S.; et al. The Child Emotion Facial Expression Set: A Database for Emotion Recognition in Children. *Front. Psychol.* **2021**, *12*, 666245. [CrossRef] [PubMed]
5. Schumann, N.P.; Bongers, K.; Scholle, H.C.; Guntinas-Lichius, O. Atlas of voluntary facial muscle activation: Visualization of surface electromyographic activities of facial muscles during mimic exercises. *PLoS ONE* **2021**, *16*, e0254932. [CrossRef] [PubMed]
6. Gozalo-Brizuela, R.; Garrido-Merchán, E.C. A survey of Generative AI Applications. *arXiv* **2023**, arXiv:2306.02781. [CrossRef]
7. Lv, Z. Generative artificial intelligence in the metaverse era. *Cogn. Robot.* **2023**, *3*, 208–217. [CrossRef]
8. Bandi, A.; Adapa, P.V.S.R.; Kuchi, Y.E.V.P.K. The Power of Generative AI: A Review of Requirements, Models, Input and Output Formats, Evaluation Metrics, and Challenges. *Future Internet* **2023**, *15*, 260. [CrossRef]
9. Elasri, M.; Elharrouss, O.; Al-Maadeed, S.; Tairi, H. Image Generation: A Review. *Neural Process. Lett.* **2022**, *54*, 4609–4646. [CrossRef]
10. van Kleef, G.A.; Côté, S. The Social Effects of Emotions. *Annu. Rev. Psychol.* **2022**, *73*, 629–658. [CrossRef] [PubMed]
11. Decety, J.; Holvoet, C. The emergence of empathy: A developmental neuroscience perspective. *Dev. Rev.* **2021**, *62*, 100999. [CrossRef]
12. Denervaud, S.; Mumenthaler, C.; Gentaz, E.; Sander, D. Emotion recognition development: Preliminary evidence for an effect of school pedagogical practices. *Learn. Instr.* **2020**, *69*, 101353. [CrossRef]

13. Walle, E.A.; Reschke, P.J.; Knothe, J.M. Social Referencing: Defining and Delineating a Basic Process of Emotion. *Emot. Rev.* **2017**, *9*, 245–252. [CrossRef]
14. Housman, D.K. The importance of emotional competence and self-regulation from birth: A case for the evidence-based emotional cognitive social early learning approach. *Int. J. Child Care Educ. Policy* **2017**, *11*, 13. [CrossRef]
15. Oleszkiewicz, A.; Frackowiak, T.; Sorokowska, A.; Sorokowski, P. Children can accurately recognize facial emotions from emoticons. *Comput. Hum. Behav.* **2017**, *76*, 372–377. [CrossRef]
16. Campbell, S.B.; Denham, S.A.; Howarth, G.Z.; Jones, S.M.; Whittaker, J.V.; Williford, A.P.; Willoughby, M.T.; Yudron, M.; Darling-Churchill, K. Commentary on the review of measures of early childhood social and emotional development: Conceptualization, critique, and recommendations. *J. Appl. Dev. Psychol.* **2016**, *45*, 19–41. [CrossRef]
17. Ruba, A.L.; Repacholi, B.M. Beyond Language in Infant Emotion Concept Development. *Emot. Rev.* **2020**, *12*, 255–258. [CrossRef]
18. Ruba, A.L.; Pollak, S.D. The Development of Emotion Reasoning in Infancy and Early Childhood. *Annu. Rev. Dev. Psychol.* **2020**, *2*, 503–531. [CrossRef]
19. Aviezer, H.; Ensenberg, N.; Hassin, R.R. The inherently contextualized nature of facial emotion perception. *Curr. Opin. Psychol.* **2017**, *17*, 47–54. [CrossRef]
20. Batty, M.; Taylor, M.J. The development of emotional face processing during childhood. *Dev. Sci.* **2006**, *9*, 207–220. [CrossRef] [PubMed]
21. Rump, K.M.; Giovannelli, J.L.; Minshew, N.J.; Strauss, M.S. The development of emotion recognition in individuals with autism. *Child Dev.* **2009**, *80*, 1434–1447. [CrossRef]
22. Solis-Arrazola, M.A.; Sanchez-Yañez, R.E.; Garcia-Capulin, C.H.; Rostro-Gonzalez, H. Enhancing image-based facial expression recognition through muscle activation-based facial feature extraction. *Comput. Vis. Image Underst.* **2024**, *240*, 103927. [CrossRef]
23. Grabell, A.S.; Huppert, T.J.; Fishburn, F.A.; Li, Y.; Jones, H.M.; Wilett, A.E.; Bemis, L.M.; Perlman, S.B. Using facial muscular movements to understand young children's emotion regulation and concurrent neural activation. *Dev. Sci.* **2018**, *21*, e12628. [CrossRef] [PubMed]
24. Mehu, M.; Mortillaro, M.; Bänziger, T.; Scherer, K.R. Reliable facial muscle activation enhances recognizability and credibility of emotional expression. *Emotion* **2012**, *12*, 701–715. [CrossRef]
25. Demmer, T.R.; Kühnapfel, C.; Fingerhut, J.; Pelowski, M. Does an emotional connection to art really require a human artist? Emotion and intentionality responses to AI- versus human-created art and impact on aesthetic experience. *Comput. Hum. Behav.* **2023**, *148*, 107875. [CrossRef]
26. Jarque-Bou, N.J.; Sancho-Bru, J.L.; Vergara, M. A Systematic Review of EMG Applications for the Characterization of Forearm and Hand Muscle Activity during Activities of Daily Living: Results, Challenges, and Open Issues. *Sensors* **2021**, *21*, 3035. [CrossRef]
27. Oppenlaender, J. The Creativity of Text-to-Image Generation. In Proceedings of the 25th International Academic Mindtrek Conference, Academic Mindtrek'22, New York, NY, USA, 16–18 November 2022; pp. 192–202. [CrossRef]
28. Ekman, P.; Friesen, W.V. Constants across cultures in the face and emotion. *J. Personal. Soc. Psychol.* **1971**, *17*, 124–129. [CrossRef] [PubMed]
29. Egger, H.L.; Pine, D.S.; Nelson, E.; Leibenluft, E.; Ernst, M.; Towbin, K.E.; Angold, A. The NIMH Child Emotional Faces Picture Set (NIMH-ChEFS): A new set of children's facial emotion stimuli. *Int. J. Methods Psychiatr. Res.* **2011**, *20*, 145–156. [CrossRef] [PubMed]
30. Meuwissen, A.S.; Anderson, J.E.; Zelazo, P.D. The creation and validation of the Developmental Emotional Faces Stimulus Set. *Behav. Res. Methods* **2017**, *49*, 960–966. [CrossRef] [PubMed]
31. Naveen, D.; Rachana, P.; Swetha, S.; Sarvashni, S. Mental Health Monitor using Facial Recognition. In Proceedings of the 2023 2nd International Conference for Innovation in Technology (INOCON), Bangalore, India, 3–5 March 2023; pp. 1–3.
32. Shehu, H.A.; Browne, W.; Eisenbarth, H. Particle Swarm Optimization for Feature Selection in Emotion Categorization. In Proceedings of the 2021 IEEE Congress on Evolutionary Computation (CEC), Krakow, Poland, 28 June–1 July 2021; pp. 752–759. [CrossRef]
33. Bryant, D.; Howard, A. A Comparative Analysis of Emotion-Detecting AI Systems with Respect to Algorithm Performance and Dataset Diversity. In Proceedings of the 2019 AAAI/ACM Conference on AI, Ethics, and Society, AIES'19, New York, NY, USA, 27–28 January 2019; pp. 377–382. [CrossRef]
34. Khan, R.A.; Crenn, A.; Meyer, A.; Bouakaz, S. A novel database of children's spontaneous facial expressions (LIRIS-CSE). *Image Vis. Comput.* **2019**, *83–84*, 61–69. [CrossRef]
35. Rapp, A.; Di Lodovico, C.; Torrielli, F.; Di Caro, L. How do people experience the images created by generative artificial intelligence? An exploration of people's perceptions, appraisals, and emotions related to a Gen-AI text-to-image model and its creations. *Int. J. Hum.-Comput. Stud.* **2025**, *193*, 103375. [CrossRef]
36. Carrasco, M.; González-Martín, C.; Navajas-Torrente, S.; Dastres, R. Level of Agreement between Emotions Generated by Artificial Intelligence and Human Evaluation: A Methodological Proposal. *Electronics* **2024**, *13*, 4014. [CrossRef]

37. Sezgin, E.; McKay, I. Behavioral health and generative AI: A perspective on future of therapies and patient care. *Npj Ment. Health Res.* **2024**, *3*, 25. [CrossRef] [PubMed]
38. Yalçin, N.; Alisawi, M. Introducing a novel dataset for facial emotion recognition and demonstrating significant enhancements in deep learning performance through pre-processing techniques. *Heliyon* **2024**, *10*, e38913. [CrossRef] [PubMed]
39. Manalu, H.V.; Rifai, A.P. Detection of human emotions through facial expressions using hybrid convolutional neural network-recurrent neural network algorithm. *Intell. Syst. Appl.* **2024**, *21*, 200339. [CrossRef]
40. Mishra, S.; Shalu, P.; Singh, R. Enhancing Face Emotion Recognition with FACS-Based Synthetic Dataset Using Deep Learning Models. In *Proceedings of the Computer Vision and Image Processing*; Kaur, H., Jakhetiya, V., Goyal, P., Khanna, P., Raman, B., Kumar, S., Eds.; Springer: Cham, Switzerland, 2024; pp. 523–531.
41. Krumhuber, E.G.; Skora, L.I.; Hill, H.C.H.; Lander, K. The role of facial movements in emotion recognition. *Nat. Rev. Psychol.* **2023**, *2*, 283–296. [CrossRef]
42. Wingenbach, T.S.H. Facial EMG—Investigating the Interplay of Facial Muscles and Emotions. In *Social and Affective Neuroscience of Everyday Human Interaction: From Theory to Methodology*; Boggio, P.S., Wingenbach, T.S.H., da Silveira Coêlho, M.L., Comfort, W.E., Murrins Marques, L., Alves, M.V.C., Eds.; Springer International Publishing: Cham, Switzerland, 2023; pp. 283–300. [CrossRef]
43. Ekman, P.; Friesen, W.V. *Facial Action Coding System: A Technique for the Measurement of Facial Movement*; Consulting Psychologists Press: Palo Alto, CA, USA, 1978.
44. Cai, Y.; Guo, Y.; Jiang, H.; Huang, M.C. Machine-learning approaches for recognizing muscle activities involved in facial expressions captured by multi-channels surface electromyogram. *Smart Health* **2018**, *5–6*, 15–25. [CrossRef]
45. Littlewort, G.; Bartlett, M.S.; Fasel, I.; Susskind, J.; Movellan, J. Dynamics of facial expression extracted automatically from video. *Image Vis. Comput.* **2006**, *24*, 615–625. [CrossRef]
46. Littlewort, G.; Whitehill, J.; Wu, T.; Fasel, I.; Frank, M.; Movellan, J.; Bartlett, M. The computer expression recognition toolbox (CERT). In Proceedings of the 2011 IEEE International Conference on Automatic Face & Gesture Recognition (FG), Santa Barbara, CA, USA, 21–23 March 2011; pp. 298–305. [CrossRef]
47. Khare, S.K.; Blanes-Vidal, V.; Nadimi, E.S.; Acharya, U.R. Emotion recognition and artificial intelligence: A systematic review (2014–2023) and research recommendations. *Inf. Fusion* **2024**, *102*, 102019. [CrossRef]
48. Siddhad, G.; Iwamura, M.; Roy, P.P. Enhancing EEG Signal-Based Emotion Recognition with Synthetic Data: Diffusion Model Approach. *arXiv* **2024**, arXiv:2401.16878.
49. Khan, A.A.; Chaudhari, O.; Chandra, R. A review of ensemble learning and data augmentation models for class imbalanced problems: Combination, implementation and evaluation. *Expert Syst. Appl.* **2024**, *244*, 122778. [CrossRef]
50. Available online: https://www.midjourney.com/ (accessed on 5 November 2024).
51. Available online: https://openai.com/dall-e-2 (accessed on 5 November 2024).
52. Available online: https://stability.ai/ (accessed on 5 November 2024).
53. Wong, M.; Ong, Y.S.; Gupta, A.; Bali, K.K.; Chen, C. Prompt Evolution for Generative AI: A Classifier-Guided Approach. In Proceedings of the 2023 IEEE Conference on Artificial Intelligence (CAI), Santa Clara, CA, USA, 5–6 June 2023. [CrossRef]
54. Available online: http://dlib.net/ (accessed on 5 November 2024).
55. Frank, E.; Hall, M.A.; Witten, I.H. *Data Mining: Practical Machine Learning Tools and Techniques*, 4th ed.; Morgan Kaufmann: Burlington, MA, USA, 2016.
56. Zsidó, A.N. The effect of emotional arousal on visual attentional performance: A systematic review. *Psychol. Res.* **2024**, *88*, 1–24. [CrossRef] [PubMed]