UIBAIFED – a dataset of facial expressions generated by AI: development and validation

Pablo Núñez-Pérez[1,†], Esperança Amengual-Alcover[1,*,†], Maria Francesca Roig-Maimó[1,†], Ramon Mas-Sansó[1,†] and Miquel Mascaró-Oliver[1,†]

[1]*University of the Balearic Islands, Carretera de Valldemossa, km 7.5, Palma de Mallorca, Spain*

Abstract

This paper introduces UIBAIFED, a novel facial expression dataset designed to enhance Facial Expression Recognition (FER) by providing high-quality, realistic images labeled with detailed demographic attributes, including age group, gender, and ethnicity. Unlike existing datasets, UIBAIFED incorporates a fine-grained classification of 22 micro-expressions, based on the universal facial expressions defined by Ekman and the micro-expression taxonomy proposed by Gary Faigin. The dataset was generated using advanced diffusion models and validated through a convolutional neural network (CNN), achieving an accuracy of 82% in expression classification. The results highlight the dataset's reliability and potential to improve FER systems. UIBAIFED fills a critical gap in the field by offering a more comprehensive labeling system, enabling future research on expression recognition across different demographic groups and advancing the robustness of FER models in diverse applications.

Keywords

HCI, machine learning, facial expression dataset, FER

1. Introduction

Facial Expression Recognition (FER) has experienced significant advances in recent years, largely driven by improvements in deep learning techniques [1] and the increasing availability of high-quality datasets [2]. These datasets play a crucial role in training models that can accurately interpret facial expressions in various contexts. However, existing datasets still present challenges related to demographic diversity, class imbalances, and ethical concerns such as bias in representation [3]. Addressing these issues is essential for developing more robust and generalizable FER models.

Despite the increasing availability of FER datasets, widely used collections such as Fer2013 [4], CK+ [5], RAF-DB [6], and AffectNet [7] have limitations. These include class imbalances where some emotions, like happiness, are overrepresented, while others, such as fear or disgust, remain underrepresented [8]. Additionally, many datasets primarily feature young and Western populations, limiting the generalization of FER models to underrepresented groups [9]. To mitigate these shortcomings, previous research has explored alternative approaches such as data augmentation techniques and synthetic facial expressions, to improve data diversity and model performance [10, 11]. However, to the best of our knowledge, no publicly available FER dataset has been entirely generated using AI.

In this work, we introduce UIBAIFED (UIB Artificial Intelligence Facial Expression Dataset), the first AI-generated dataset designed to improve FER model training and evaluation. Unlike traditional datasets, UIBAIFED uses generative AI techniques to create a diverse and balanced dataset of facial expressions. This approach ensures a more representative training corpus for modern FER systems, reducing demographic biases and improving overall model robustness.

2. Related work

2.1. Traditional FER datasets

Several datasets have been widely used in FER research, including FER2013, CK+, RAF-DB, and AffectNet. These datasets have significantly contributed to the advancement of deep learning models for emotion classification. However, they often suffer from limitations such as:

- Demographic Imbalances: Many datasets focus on younger and Western populations, resulting in models that generalize poorly to underrepresented groups [9].
- Class Imbalances: Some emotions, such as happiness and neutrality, are more frequently represented than others, such as fear or disgust, which can lead to biased model performance [12].
- Labelling Inconsistencies: Differences in how emotions are annotated across datasets can introduce noise and hinder model generalization [13].

These limitations have motivated researchers to develop new datasets that offer more balanced and diverse samples, ensuring better generalizability of FER models.

2.2. Micro-expression recognition

Micro-expressions are brief, involuntary facial expressions that reveal suppressed emotions. Their fleeting nature makes them difficult to capture and classify, yet they are crucial in fields such as psychology, security, and Human-Computer Interaction (HCI) [14].

One of the major gaps in current FER datasets is the absence of systematic labelling for micro-expressions. Unlike standard datasets that focus on broader emotional categories, micro-expressions require finer granularity and precise annotation. This limitation hinders the development of models capable of detecting subtle emotional cues in real-time applications [15].

Faigin's categorization of facial expressions provides a comprehensive framework for understanding facial dynamics beyond the traditional seven emotional categories [16]. This taxonomy emphasizes the complexity of expressions, capturing subtle variations that are often overlooked in conventional FER studies. However, existing datasets rarely incorporate this level of detail, limiting the ability of current models to recognize nuanced emotional states. Bridging this gap requires datasets explicitly designed to align with Faigin's categorization.

3. Methods

To address the aforementioned challenges, in this work we introduce UIBAIFED, an AI-generated dataset designed to provide a more balanced and diverse representation of facial expressions. By leveraging generative models, we ensure controlled variations in age, gender, and ethnicity while maintaining realistic differences in pose, lighting, and expression intensity. This approach aims to mitigate biases in traditional datasets and enhance the robustness of FER models.

3.1. Facial models

To ensure the quality of the dataset, the generated images adhere to the following criteria: the face must be centred and occupy between 40% and 70% of the image area; lighting should be sufficient to clearly highlight facial expression details, while the background remains uniform and neutral to prevent potential classification interference. Additionally, facial expressions must accurately replicate the descriptions proposed by Gary Faigin [16]. Furthermore, visual artefacts should be minimal and should not compromise the expressiveness of the face.

The UIBAIFED dataset ensures a balanced distribution across sex, five distinct age groups (see Figure 1) and three body composition categories (see Figure 2).

(a) 15   (b) 25   (c) 45

(d) 65   (e) 85

Figure 1: Age group distribution in the UIBAIFED dataset.

(a) Underweight   (b) Normal weight   (c) Overweight

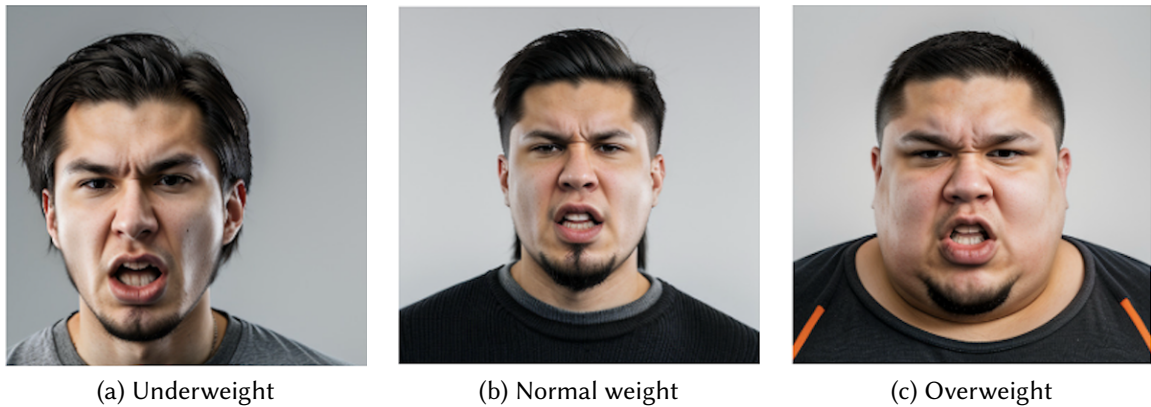**Figure 2:** Body composition categories in the UIBAIFED dataset.

Ethnic diversity is considered based on the classification provided by the Office of Management and Budget (OMB) [17], which includes groups such as Native Americans, Asians, Black individuals, Hispanics, Native Hawaiians or other Pacific Islanders, and White individuals of European, North African, or Middle Eastern descent (see Figure 3).

| (a) White | (b) Black | (c) Asian |

| (a) Alaskan-Native | (b) Hispanic | (c) Hawaiian |

Figure 3: Ethnic diversity distribution in the UIBAIFED dataset.

3.2. Image creation and filtering process

For the generation of facial expression images in the UIBAIFED dataset, the Stable Diffusion model [18] was utilized. This open-source technology can be run locally, offering the advantage of generating an unlimited number of images. Its flexible nature and the continuous contributions from the community have enabled the development of improved versions, enhancing the variety and quality of the results, ensuring that the images meet the criteria established for facial expression analysis.

The Stable Diffusion checkpoints are pre-trained models designed to generate images from textual descriptions. Large datasets are used to learn the correlations between words and visual elements. The selection of a checkpoint requires considering the ability to generate a variety of images with all the required characteristics, while also minimizing the generation time. Based on empirical findings, it has been found that the Realistic Vision checkpoint [19] best meets the needs of the dataset.

To optimize the model for facial expression generation, Low-Rank Adaptation (LoRA) [20], [21] was employed. LoRA allows the adaptation of machine learning models to new contexts quickly by adding lightweight components to the original model rather than modifying the entire structure. In the case of Stable Diffusion, LoRAs specifically tailored for facial expression generation were sourced from CivitAI [22]. Table 1 depicts the LoRAs used for the generation of the UIBAIFED dataset.

Table 1

LoRAs Used for Generating the UIBAIFED Datasets

LoRa	Emotional categories
Sad - Facial Expression	Sadness
Look alive! Excited Facial Expression/Emotion	Excitement
Ashamed - Embarrassed Facial Expression	Shame
Look alive! Smirking Facial Expression/Emotion	Joy
Disgusted - Disapproving Facial Expression	Disgust
Angry - Facial Expression	Anger
Shocked - Facial Expression	Surprise
Pleading Eyes - Facial Expression	Fear
Scared - Facial Expression	Fear
Expressions Helper Realistic	Multipurpose use

Additionally, the necessary prompts for generating the micro-expressions that make up the dataset were developed. Out of the 33 micro-expressions described by Gary Faigin, a subset of only 22 was successfully reproduced due to the difficulty in describing certain subtleties for generative models.

An example of the generated (positive and negative) prompts is as follows:

```
--prompt "White Man, 15y.o, (AngryShouting:0),(angry!!), (((shouting!!!))),
<lora:l\_ang\_ae\_sd\_64\_32:0.9>, Underweight, ((looking at the camera)),
hyperrealistic, professional photo, studio lighting, sharp focus,
centered on the image, vertical alignment, face, plain grey background"

--negative\_prompt "((Deformed)), disfigured, hat,(artifacts in eyes, bad iris),
((artifacts in face)), hawaiian clothes, worse quality, low quality, jpeg,
pixelated, anime, ((poorly illuminated face)), red eyes, ((bad teeth)),
((body, arms, hands, legs, naked))"
```

The prompt described above generates the image shown in Figure 4, which represents a 15-year-old male of lean build with the *Anger* expression, specifically the micro-expression *AngryShouting*, according to Gary Faigin's taxonomy.

Figure 4: Example of a generated image with the *Anger* expression (15-year-old White male, underweight).

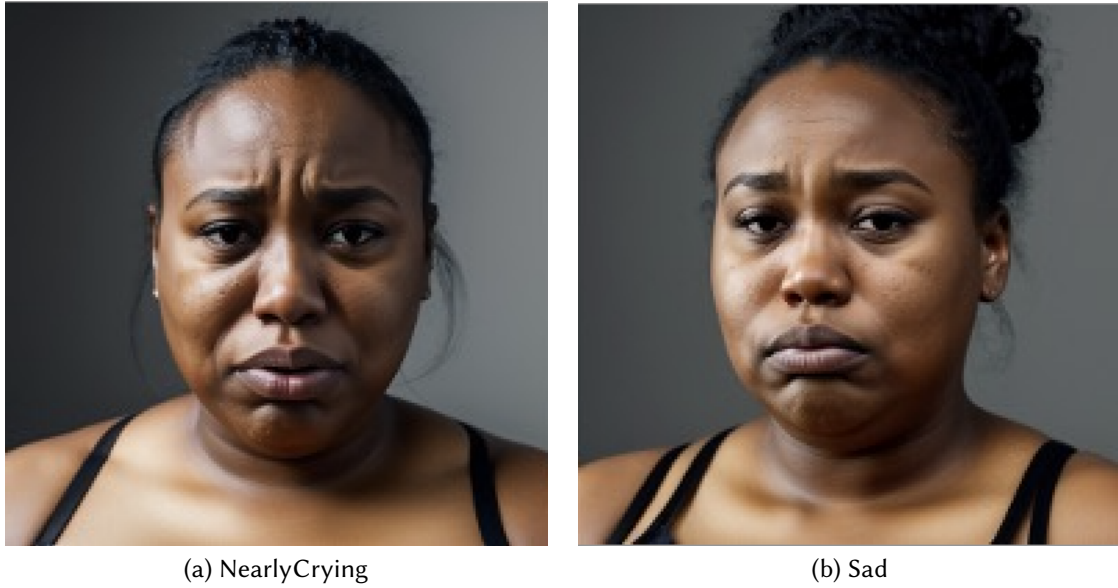|	
(a) NearlyCrying	(b) Sad

Figure 5: Expressions *NearlyCrying* and *Sad* generated with the same LoRA.

The structure of the different prompts is consistently maintained, following this format:

```
"Ethnicity, gender, age,<description of the expression>"
```

Within the description of the expression, the reference to the LoRA is included using the following nomenclature:

```
<lora: (LoRA name):(weight)>
```

In this structure, "weight" refers to the intensity of the expression. Certain micro-expressions are generated using the same LoRA but with different descriptors. For example, the micro-expressions *NearlyCrying* and *Sad*, both representing sadness, are generated with the following two prompts, producing the images shown in Figure 5, while utilizing the same LoRA.

```
--prompt "Black Woman, 25y.o, (NearlyCrying:0), ((sad mouth)), miserable face,
(sad:1.2), <lora:l\_sad\_se\_sd\_64\_32:1>, Overweight, ((looking at the camera)),
hyperrealistic, professional photo, studio lightning, sharp focus,
centered on the image, vertical alignment, face, plain grey background"

--prompt "Black Woman, 25y.o, (Sad:0), (sad), (melancholic face), closed lips,
small mouth, <lora:l\_sad\_se\_sd\_64\_32:1>, Overweight, ((looking at the camera)),
hyperrealistic, professional photo, studio lightning, sharp focus,
centered on the image, vertical alignment, face, plain grey background"
```

Parentheses and numerical values are used to emphasize specific words or phrases.

Figure 6 displays the 22 expressions generated for a 15-year-old White male. An automated script was developed to generate 3960 prompts, resulting from the combination of 2 genders, 6 ethnicities, 5 age groups, and 3 body types, all organized according to the six universal expressions according to Ekman's classification [23].

The images corresponding to the generated prompts were produced using the Automatic111 application [24]. Due to the random nature of the image generation process, not all images are expected to be accurate on the first attempt. Therefore, for each micro-expression, between 15 and 30 images were generated to ensure the desired quality and consistency.
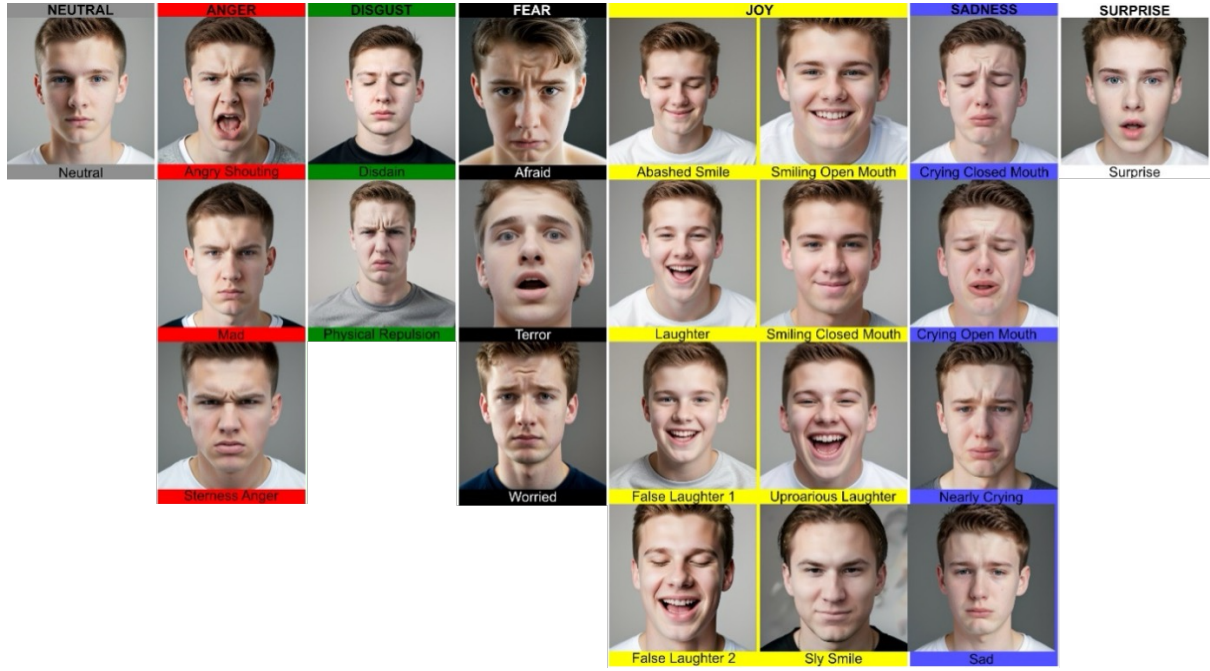
Figure 6: Examples of the 22 expressions generated for a 15-year-old white male.

The images generated using the specified prompts were manually selected based on their alignment with the descriptions and graphical representations provided by Gary Faigin [16]. Figure 7 illustrates the manual matching process for the micro-expression *SlySmile*. The left image represents the generated expression from the UIBAIFED dataset, while the right image corresponds to the reference illustration from Gary Faingin's work. The selection process ensured that each image accurately represented the intended facial expression and adhered to the established criteria.
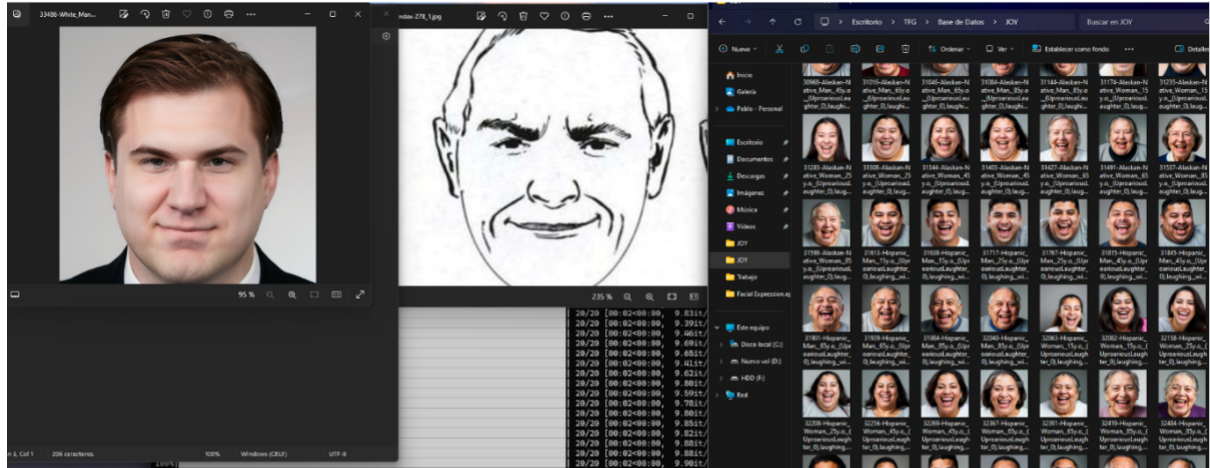
Figure 7: Manual matching process for the *SlySmile* expression.

4. The UIBAIFED dataset

The total number of images in the dataset is 2948. Images generated for prompts with different body types were removed due to the minimal differences observed between those labelled as Normal Weight and Underweight. A greater number of representations were retained for more complex expressions, leading to the distribution of images per micro-expression shown in Table 2.

Table 2
Distribution of images per micro-expressions

Universal Expression	Micro-expression	Number of images
Neutral	Neutral	123
Anger	AngryShouting	121
	Mad	165
	SternnessAnger	126
Disgust	Disdain	161
	PhysicalRepulsion	151
Fear	Afraid	121
	Terror	124
	Worried	135
Joy	AbashedSmile	129
	Laughter	121
	FalseLaughter1	120
	FalseLaughter2	122
	SmilingOpenMouth	120
	SmilingClosedMouth	121
	UproariousLaughter	119
	SlySmile	120
Sadness	CryingClosedMouth	132
	CryingClosedMouth	132
	CryingOpenMouth	158
	NearlyCrying	158
	Sad	155
Surprise	Surprise	146

The database is organized into folders, each containing images with a resolution of 512×512 pixels. There is one folder for each of the six universal expressions according to Ekman's classification [23]. It is important to note that the seventh expression, *Contempt*, which Ekman later added to his original classification, is labelled in our dataset as the micro-expression *Disdain*. This expression is included in the *Disgust* folder, following Gary Faigin's classification approach.

Within each folder, the images are named according to the following format:

```
Num\_ethnicity\_gender\_age\_microexpression.png
```

"Num" represents the generation number assigned by Stable Diffusion and indicates the order of the images within each folder. The images are organized first by micro-expression, followed by ethnicity, gender, and age.

5. UIBAIFED validation

To initially validate the UIBAIFED dataset, a simple Convolutional Neural Network (CNN) model was employed for facial expression classification. The model takes grayscale images of size 128×128 pixels as input, which are processed through three convolutional layers. These layers are followed by a Rectified Linear Unit (ReLU) layer and a max-pooling layer to extract key features. The architecture also includes four Fully Connected (FC) layers, which are used to classify the facial expressions into one of the 22 target micro-expressions described in the dataset. The overall network structure is shown in Figure 8.

To enhance generalization and prevent overfitting, a Dropout layer is applied between the fully connected layers. This dropout technique helps the network learn more robust features by randomly dropping units during training, which improves the model's ability to generalize to unseen data.

The dataset is split into training and test sets, with 67% of the data used for training and 33% for testing. The data distribution is balanced based solely on micro-expression types, and other factors
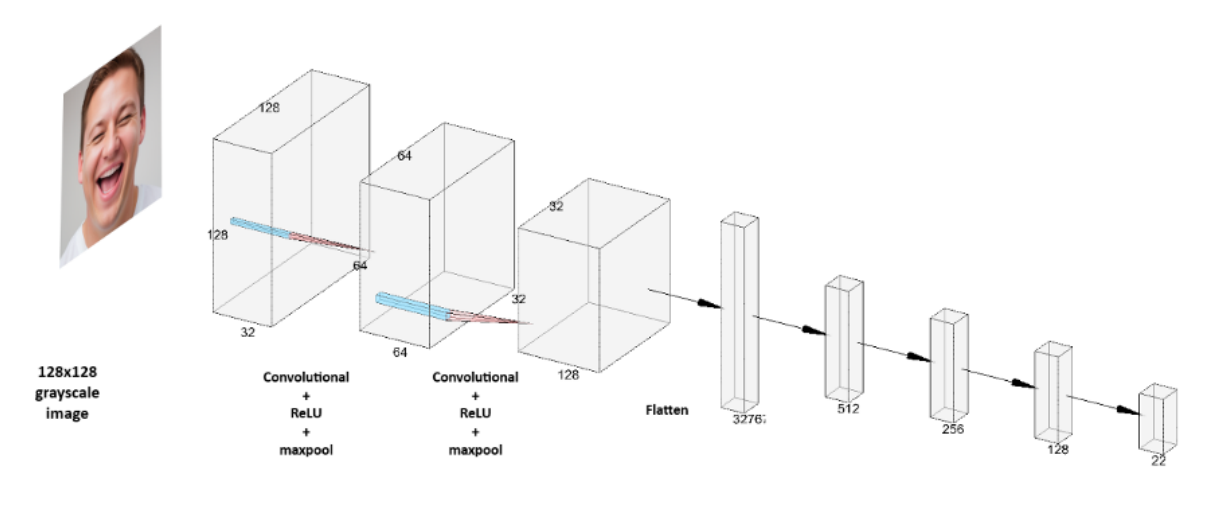
Figure 8: Structure of the Neural Network Model for UIBAIFED Dataset Validation.

such as gender, body type, ethnicity, and age are not considered in this validation step. These factors will be explored in future studies.

6. Results

After completing the training process, a Loss value close to 0.5 and an overall Accuracy of 82% were achieved. These results were obtained using 67% of the images for training, 1975 images in total. Figure 9 shows the evolution of these values as a function of the training epochs.
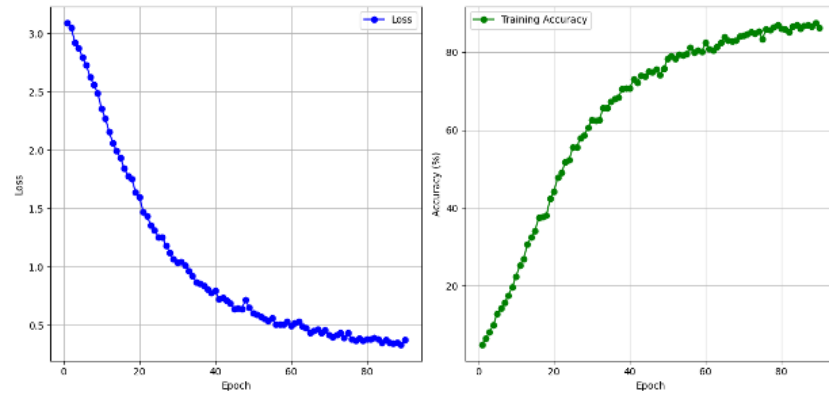
Figure 9: Loss and accuracy evolution across epochs.

The trained CNN model was tested with the test dataset (formed by 1975 images), achieving an overall accuracy of 85,71%. The resulting confusion matrix is presented in Figure 10, while Table 3 details the performance metrics for each of the 22 micro-expressions. Additionally, Table 4 provides a summary of the overall classification metrics.

The test results indicate that the CNN model has successfully learned and generalized most facial expressions in a validation set of over 900 images that were not used during training.

Most facial expressions achieve an accuracy above 75%. However, expressions related to *Joy* present greater classification challenges. Specifically, *AbashedSmile* is sometimes misclassified as *Sad* or *Worried*, while *FalseLaughter1* is frequently confused with *CryingOpenMouth*. This misclassification likely occurs because some samples of *FalseLaughter1* include eyes that are sufficiently closed, making them visually similar to *CryingOpenMouth*.

Confusion Matrix

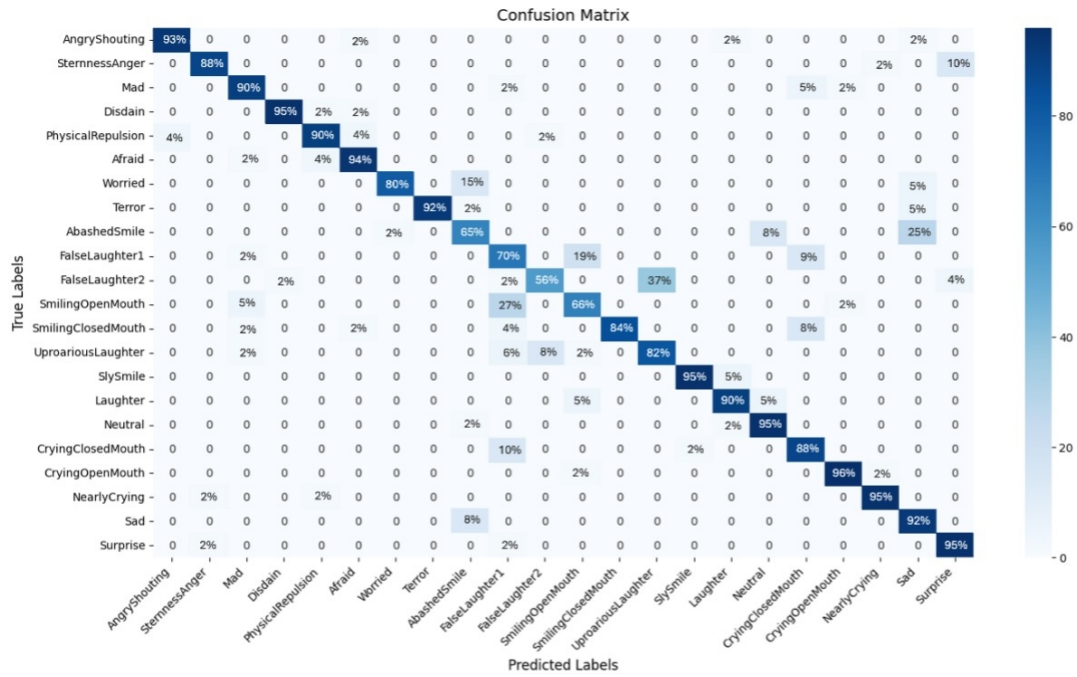| True \ Pred	AngryShouting	SternnessAnger	Mad	Disdain	PhysicalRepulsion	Afraid	Worried	Terror	AbashedSmile	FalseLaughter1	FalseLaughter2	SmilingOpenMouth	SmilingClosedMouth	UproariousLaughter	SlySmile	Laughter	Neutral	CryingClosedMouth	CryingOpenMouth	NearlyCrying	Sad	Surprise
AngryShouting	93%	0	0	0	0	2%	0	0	0	0	0	0	0	0	0	2%	0	0	0	0	2%	0
SternnessAnger	0	88%	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2%	0	10%
Mad	0	0	90%	0	0	0	0	0	0	2%	0	0	0	0	0	0	0	5%	2%	0	0	0
Disdain	0	0	0	95%	2%	2%	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
PhysicalRepulsion	4%	0	0	0	90%	4%	0	0	0	0	2%	0	0	0	0	0	0	0	0	0	0	0
Afraid	0	0	2%	0	4%	94%	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Worried	0	0	0	0	0	0	80%	0	15%	0	0	0	0	0	0	0	0	0	0	0	5%	0
Terror	0	0	0	0	0	0	0	92%	2%	0	0	0	0	0	0	0	0	0	0	0	5%	0
AbashedSmile	0	0	0	0	0	0	2%	0	65%	0	0	0	0	0	0	8%	0	0	0	0	25%	0
FalseLaughter1	0	0	2%	0	0	0	0	0	0	70%	0	19%	0	0	0	0	0	9%	0	0	0	0
FalseLaughter2	0	0	0	2%	0	0	0	0	0	2%	56%	0	0	37%	0	0	0	0	0	0	0	4%
SmilingOpenMouth	0	0	5%	0	0	0	0	0	0	27%	0	66%	0	0	0	0	0	0	2%	0	0	0
SmilingClosedMouth	0	0	2%	0	0	2%	0	0	0	4%	0	0	84%	0	0	0	0	8%	0	0	0	0
UproariousLaughter	0	0	2%	0	0	0	0	0	0	6%	8%	2%	0	82%	0	0	0	0	0	0	0	0
SlySmile	0	0	0	0	0	0	0	0	0	0	0	0	0	0	95%	5%	0	0	0	0	0	0
Laughter	0	0	0	0	0	0	0	0	0	0	0	5%	0	0	0	90%	5%	0	0	0	0	0
Neutral	0	0	0	0	0	0	0	0	2%	0	0	0	0	0	2%	0	95%	0	0	0	0	0
CryingClosedMouth	0	0	0	0	0	0	0	0	0	10%	0	0	0	0	2%	0	0	88%	0	0	0	0
CryingOpenMouth	0	0	0	0	0	0	0	0	0	0	0	2%	0	0	0	0	0	0	96%	2%	0	0
NearlyCrying	0	2%	0	0	2%	0	0	0	0	0	0	0	0	0	0	0	0	0	0	95%	0	0
Sad	0	0	0	0	0	0	0	0	8%	0	0	0	0	0	0	0	0	0	0	0	92%	0
Surprise	0	2%	0	0	0	0	0	0	2%	0	0	0	0	0	0	0	0	0	0	0	0	95%

Figure 10: Confusion matrix of the CNN model classification results.

Table 3

Performance Metrics for Each Micro-expression

Micro-expression	Precision	Recall	F1-Score	Support
Neutral	0.92	0.82	0.87	40
AngryShouting	0.84	0.95	0.89	43
Mad	0.77	0.82	0.80	40
SternnessAnger	0.86	0.90	0.88	40
Disdain	0.97	0.86	0.91	43
PhysicalRepulsion	0.88	0.87	0.87	52
Afraid	0.94	0.85	0.89	53
Terror	0.93	0.97	0.95	40
Worried	0.73	0.80	0.76	40
AbashedSmile	0.68	0.65	0.67	40
Laughter	0.93	0.97	0.95	40
FalseLaughter1	0.57	0.72	0.63	54
FalseLaughter2	0.67	0.27	0.38	52
SmilingOpenMouth	0.61	0.61	0.61	41
SmilingClosedMouth	1.00	0.82	0.90	50
UproariousLaughter	0.51	0.76	0.61	51
SlySmile	0.97	0.97	0.97	40
CryingClosedMouth	0.72	0.69	0.71	42
CryingOpenMouth	0.87	0.98	0.92	48
NearlyCrying	0.97	0.93	0.95	41
Sad	0.82	0.79	0.81	39
Surprise	0.85	0.80	0.82	44

A recurring pattern observed across all training-test cycles is the confusion between *FalseLaughter2* and *UproariousLaughter*. The primary difficulty in distinguishing these expressions lies in their strong resemblance. Both feature a wide, open mouth and eyes that are either closed or nearly closed. This issue was already anticipated during the image filtering process, where it was noted that the visual differences between these expressions were minimal (see Figure 11).

Table 4
Summary of Overall Classification Metrics

Metric	Precision	Recall	F1-Score
accuracy		0.80	
macro avg	0.82	0.81	0.81
weighted avg	0.81	0.80	0.80

(a) FalseLaughter2

(b) UproariousLaughter

Figure 11: Comparison of *FalseLaughter2* and *UproariousLaughter*.

7. Conclusion and future work

In this study, we have introduced and tested UIBAIFED, a novel facial expression dataset that features high-quality, realistic color images labeled according to age group, gender, ethnicity, and facial expression. The labeling follows the universal expressions, encompassing a total of 22 micro-expressions based on the terminology proposed by Gary Faigin.

To validate the dataset, a Convolutional Neural Network (CNN) was employed, achieving an accuracy of 80%, with strong performance across most expressions.

Compared to existing facial expression datasets, UIBAIFED introduces a key innovation by providing a more detailed level of labeling. To the best of our knowledge, no other database currently offers this granularity in annotation.

Moving forward, the dataset enables new research opportunities, particularly in analyzing FER performance across different age and ethnic groups. Addressing these challenges will contribute to further advancements in the field of facial expression recognition.

Acknowledgments

Declaration on generative AI

During the preparation of this work, the authors used Stable Diffusion 1.5 and Realistic Vision V6.0 B1 for the generation of the images that comprise the UIBAIFED dataset. Additionally, ChatGPT (GPT-4, March 2025 version) was employed to assist with grammar, spelling, and language refinement. After using these tool(s)/service(s), the author(s) reviewed and edited the content as needed and take(s) full responsibility for the publication's content.

References

[1] S. Li, W. Deng, Deep facial expression recognition: A survey, IEEE Transactions on Affective Computing 13 (2022) 1195–1215. doi:10.1109/TAFFC.2020.2981446.

[2] D. Kollias, A. Schulc, E. Hajiyev, S. Zafeiriou, Analysing affective behavior in the first abaw 2020 competition, in: 2020 15th IEEE International Conference on Automatic Face and Gesture Recognition (FG 2020), IEEE, 2020, pp. 637–643. doi:10.1109/FG47880.2020.00126.

[3] J. Buolamwini, T. Gebru, Gender shades: intersectional accuracy disparities in commercial gender classification, in: S. A. Friedler, C. Wilson (Eds.), Proceedings of the 1st Conference on Fairness, Accountability and Transparency, volume 81 of *Proceedings of Machine Learning Research*, PMLR, 2018, pp. 77–91. URL: https://proceedings.mlr.press/v81/buolamwini18a.html.

[4] I. J. Goodfellow, D. Erhan, P. L. Carrier, A. Courville, M. Mirza, B. Hamner, W. Cukierski, Y. Tang, D. Thaler, D.-H. Lee, et al., Challenges in representation learning: a report on three machine learning contests, in: Neural information processing: 20th international conference, ICONIP 2013, daegu, korea, november 3-7, 2013. Proceedings, Part III 20, Springer, 2013, pp. 117–124.

[5] P. Lucey, J. F. Cohn, T. Kanade, J. Saragih, Z. Ambadar, I. Matthews, The extended cohn-kanade dataset (CK+): a complete dataset for action unit and emotion-specified expression, in: 2010 IEEE computer society conference on computer vision and pattern recognition-workshops, IEEE, 2010, pp. 94–101. doi:10.1109/CVPRW.2010.5543262.

[6] S. Li, W. Deng, J. Du, Reliable crowdsourcing and deep locality-preserving learning for expression recognition in the wild, in: Proceedings of the IEEE conference on computer vision and pattern recognition, 2017, pp. 2852–2861.

[7] A. Mollahosseini, B. Hasani, M. H. Mahoor, AffectNet: A database for facial expression, valence, and arousal computing in the wild, IEEE Transactions on Affective Computing 10 (2017) 18–31.

[8] J. Fan, J. Zhou, X. Deng, H. Wang, L. Tao, H. K. Kwan, Combating uncertainty and class imbalance in facial expression recognition, in: TENCON 2022 - 2022 IEEE Region 10 Conference (TENCON), IEEE, 2022, pp. 1–4. doi:10.1109/TENCON55691.2022.9977693.

[9] I. Dominguez-Catena, D. Paternain, M. Galar, Metrics for dataset demographic bias: a case study on facial expression recognition, IEEE Transactions on Pattern Analysis and Machine Intelligence 46 (2024) 5209–5226. doi:10.1109/TPAMI.2024.3361979.

[10] A. Psaroudakis, D. Kollias, MixAugment & Mixup: augmentation methods for facial expression recognition, in: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2022, pp. 2367–2375.

[11] J. Yu, Y. Liu, R. Fan, G. Sun, Mixcut:a data augmentation method for facial expression recognition, 2024. URL: https://arxiv.org/abs/2405.10489. arXiv:2405.10489.

[12] Y. Zhang, Y. Li, l. Qin, X. Liu, W. Deng, Leave no stone unturned: mine extra knowledge for imbalanced facial expression recognition, in: A. Oh, T. Naumann, A. Globerson, K. Saenko, M. Hardt, S. Levine (Eds.), Advances in Neural Information Processing Systems, volume 36, Curran Associates, Inc., 2023, pp. 14414–14426. URL: https://proceedings.neurips.cc/paper_files/paper/2023/file/2e6744370a8616c90d1e3b7a41993b7c-Paper-Conference.pdf.

[13] Y. Chen, J. Joo, Understanding and mitigating annotation bias in facial expression recognition, in: Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV), 2021, pp. 14980–14991.

[14] I. P. Adegun, H. B. Vadapalli, Facial micro-expression recognition: a machine learning approach, Scientific African 8 (2020) e00465. URL: https://www.sciencedirect.com/science/article/pii/S2468227620302039. doi:https://doi.org/10.1016/j.sciaf.2020.e00465.

[15] H. Guerdelli, C. Ferrari, W. Barhoumi, H. Ghazouani, S. Berretti, Macro- and micro-expressions facial datasets: a survey, Sensors 22 (2022). doi:10.3390/s22041524.

[16] G. Faigin, The artist's complete guide to facial expression, Watson-Guptill, New York, 2012.

[17] Office of Institutional Research, Race/ethnicity FAQs, 2025. URL: https://provost.tufts.edu/institutionalresearch/race-ethnicity-faq/.

[18] Stable Difussion, Stable Diffusion online, 2025. URL: https://stabledifffusion.com.

[19] Stable Difussion, Realistic vision V6.0, 2025. URL: https://civitai.com/models/4201/realistic-vision-v60-b1.

[20] E. J. Hu, Y. Shen, P. Wallis, Z. Allen-Zhu, Y. Li, S. Wang, L. Wang, W. Chen, et al., LoRA: low-rank adaptation of large language models, ICLR 1 (2022) 3.

[21] Y. Zeng, K. Lee, The expressive power of low-rank adaptation, arXiv preprint arXiv:2310.17513 (2023).

[22] Civitai, LoRA Stable Diffusion & Flux AI models, 2025. URL: https://civitai.com/tag/lora.

[23] P. Ekman, T. Dalgleish, M. Power, Handbook of cognition and emotion, Wiley Online Library, 1999.

[24] AUTOMATIC1111, Stable Diffusion web UI, 2022. URL: https://github.com/AUTOMATIC1111/stable-diffusion-webui.

A. Online Resources

The data underlying the results presented in the study are available at https: https://zenodo.org/records/15181380