22-27 MAY 2022
ICASSP 2022
SINGAPORE

UNIVERSITY OF CAMBRIDGE
Department of Computer Science and Technology

IEEE Signal Processing Society

IEEE

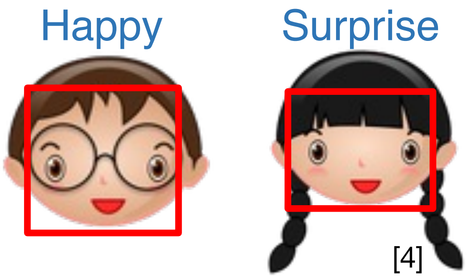
# STATISTICAL, SPECTRAL AND GRAPH REPRESENTATIONS FOR VIDEO-BASED FACIAL EXPRESSION RECOGNITION IN CHILDREN

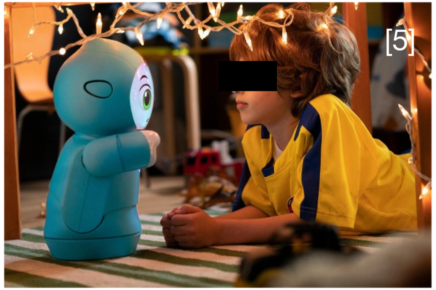
**Nida Itrat Abbasi***, Siyang Song* and Hatice Gunes

*Equal Contribution

# Introduction

- Accurately recognising facial expressions can provide a deeper insight into **human nonverbal behaviours** [1], especially for people that do **not** have advanced **verbal communication skills**, e.g., children.
- Very few machine learning-based **facial expression classification** approaches [2] [3] have specifically investigated facial expression analysis in **children.**
- Understanding child facial expression analysis is important for developing **affect models** for future **child-agent interaction** research.




22-27 MAY 2022
ICASSP 2022
SINGAPORE

# Why construct child-specific facial expression frameworks?

01 Models trained on adult expression datasets do not generalize well on child facial expression recognition tasks [6].

02 Facial expressions of children are often exaggerated, incomplete and unique as compared with their adult counterparts [2].

03 Child facial expressions are very dependent on demographic factors such as gender, age and ethnicity [3].

# Novelty of our Proposed Approach

**1.** First approach that constructs video-level heterogeneous graph representation for facial expression recognition in children.

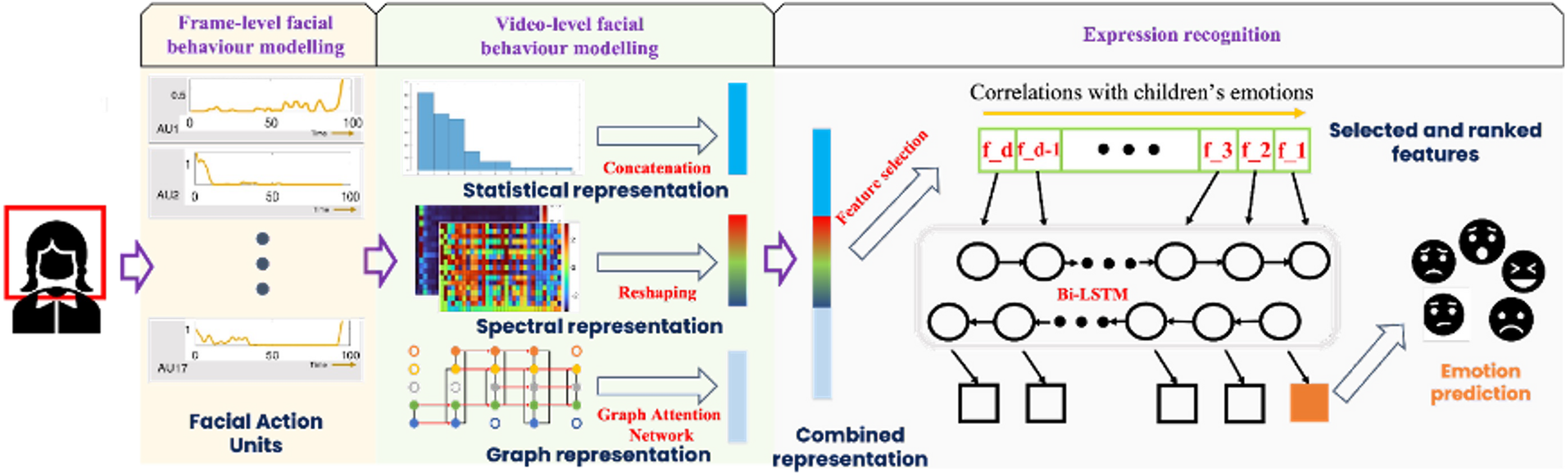
**2.** First approach that predicts children's facial expressions using the automatically detected Action Units (AUs).

## Advantages of using Facial AUs

**01** Objective and descriptive representation for measuring facial behavioral changes.

**02** Ethical advantage over using raw images and videos for vulnerable demographic groups including children and the elderly by discarding video data after AU detection in real-time.

# Proposed Methodology




The proposed pipeline for child facial expression recognition.

# Proposed Methodology

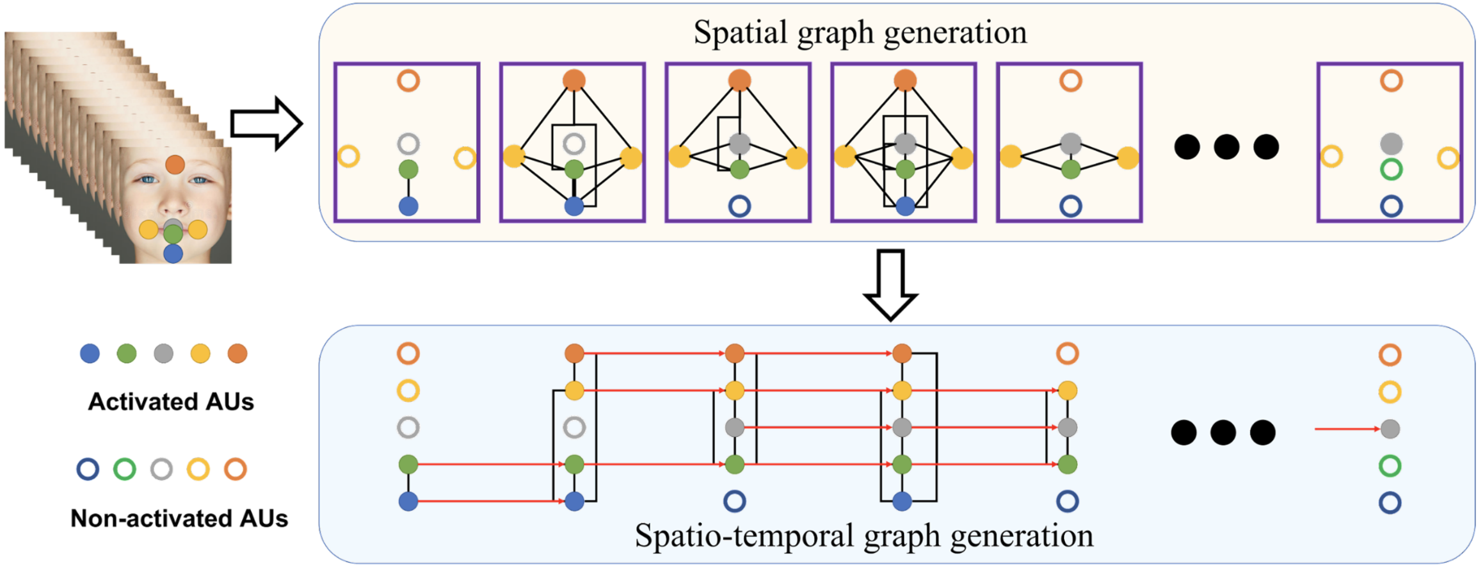



Illustration of the proposed graph representation.

# Experiments

- **Dataset:** LIRIS Children Spontaneous Facial Expression Video Database (12 children of different ages, gender and ethnicities, 208 video clips) [3] [7].

- **Training details:**

1. Excluded video clips belonging to the anger class (not sufficient number of clips), the combined categories, and the clips with very short duration (less than 3% of files).

2. We have conducted **12-fold leave-one-child-out** cross-validation.

3. To train the Bi-LSTM and MLP models, **cross-entropy** was used as the loss function and **Adam** was used as the optimizer with the learning rate of 0.001 and 0.005, respectively.

22-27 MAY 2022
ICASSP 2022
SINGAPORE

# Results

**01. Comparison between video-level representations:**
Combination of all three representations provides the highest accuracy (66 % in MLP and 67.4 % in Bi-LSTM).

**02. Comparison between MLP and the proposed Bi-LSTM model:**
Bi-LSTM outperforms the standard MLP in all representations except in the spectral representation.

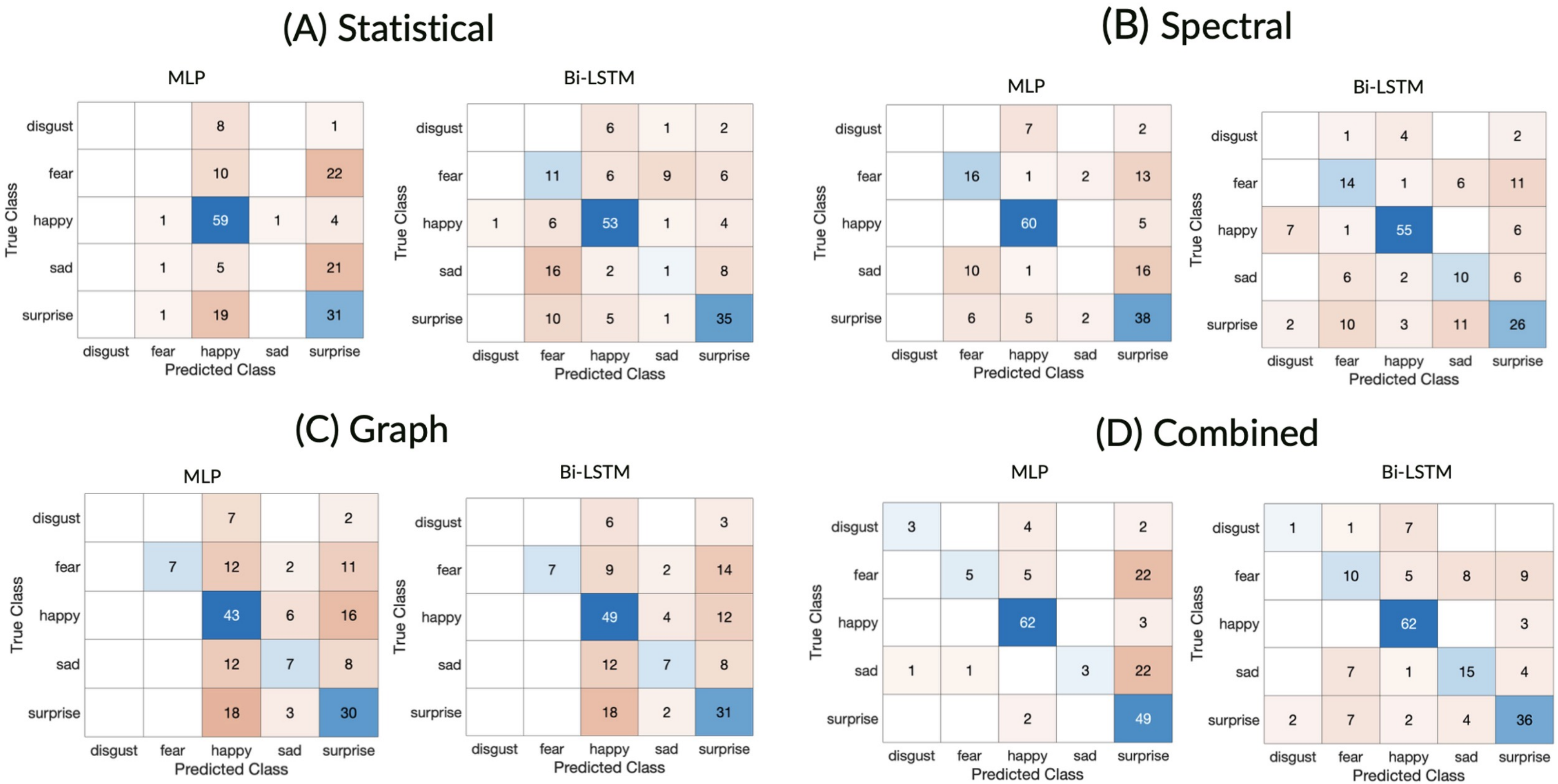
**03. Comparison between expression classes:**
Accurate predictions observed for happy expression class and surprise expression class while negative expressions like disgust and fear are often misclassified.

| Statistical | | Spectral | |
|---|---|---|---|
| MLP | Bi-LSTM | MLP | Bi-LSTM |
| 48.9% | 54.4% | 62% | 57.1% |

| Graph | | Combined | |
|---|---|---|---|
| MLP | Bi-LSTM | MLP | Bi-LSTM |
| 47.3% | 51.1% | 66.3% | 67.4% |

Classification accuracy obtained with different representations proposed

# Results




Confusion matrices for the different facial expression recognition frameworks proposed in this work

# Conclusions

- Combination of all three representations using the Bi-LSTM model provides the highest accuracy for child facial expression recognition.
- Models developed in this work can provide a valuable stepping stone for creating affect recognition frameworks for child-agent interaction research.

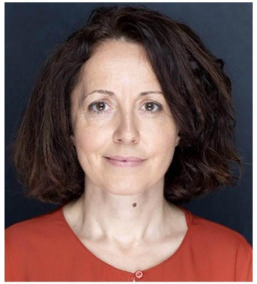
# Future Work

We aim to use more advanced deep-learning frameworks like **gated graph convolutional networks** and also compare other state-of-the-art end-to-end network architectures for improving the accuracy of the models proposed in this work.

# ACKNOWLEDGEMENTS

**Nida Itrat Abbasi**
PhD Student,
Dept. of Computer Science and Technology,
University of Cambridge
nia22@cam.ac.uk

**Siyang Song**
Postdoctoral Researcher,
Dept. of Computer Science and Technology,
University of Cambridge



**Hatice Gunes (Advisor)**
Professor of Affective Intelligence and Robotics,
University of Cambridge

We would also like to thank the following for supporting this work:

- WD Armstrong Trust Fund Studentship, School of Technology, University of Cambridge.
- Cambridge Trust & Computer Science Premium Scholarship, Cambridge Commonwealth, European & International Trust, University of Cambridge.

Nida Itrat Abbasi and Siyang Song contributed equally to this work.

UNIVERSITY OF CAMBRIDGE
Department of Computer Science and Technology

22-27 MAY 2022
ICASSP 2022
SINGAPORE

# References


1. Z. Liu et al., “A facial expression emotion recognition based human-robot interaction system,” IEEE/CAA Journal of Automatica Sinica, vol. 4, no. 4, pp. 668–676,2017.
2. M. A. Witherow et al., “Transfer learning approach to multiclass classification of child facial expressions,” in Applications of Machine Learning. International Society for Optics and Photonics, 2019, vol. 11139, p. 1113911.
3. R. A. Khan et al., “A novel database of children’s spontaneous facial expressions (liris-cse),” Image and Vision Computing, vol. 83, pp. 61–69, 2019.
4. https://www.clipartkey.com/view/iRJboxR_free-girl-face-clipart/
5. https://www.wired.com/story/moxie-is-the-robot-pal-you-dreamed-of-as-a-kid/.
6. A. Howard et al., “Addressing bias in machine learning algorithms: A pilot study on emotion recognition for intelligent systems,” in 2017 IEEE Workshop on Advanced Robotics and its Social Impacts (ARSO). IEEE, 2017, pp. 1–7.
7. R. Ahmed Khan et al., “Automatic affect analysis: from children to adults,” in International Symposium on Visual Computing. Springer, 2015, pp. 304–313.