

Classroom Engagement and Feedback Analysis Using Multimodal Emotion Recognition

Ashwitha Devasya, Dr. Andrew Abel, Dr. Martin Goodfellow, Dr. Laibing Jia

University of Strathclyde, Glasgow

ashwitha.devasya@strath.ac.uk



1

Rethinking Classroom Engagement: From Traditional Methods to AI

- **Teacher's perspective:** Hard to track whole-class engagement; traditional methods like surveys are biased and time-consuming.
- **Our solution:** AI-driven, class-level feedback that's objective, timely, and teacher-focused.

RQ1

What are teachers' views on using MER to find classroom engagement compared to traditional methods?

RQ2

How can deep learning techniques be optimised to find classroom engagement?

2

Our Approach

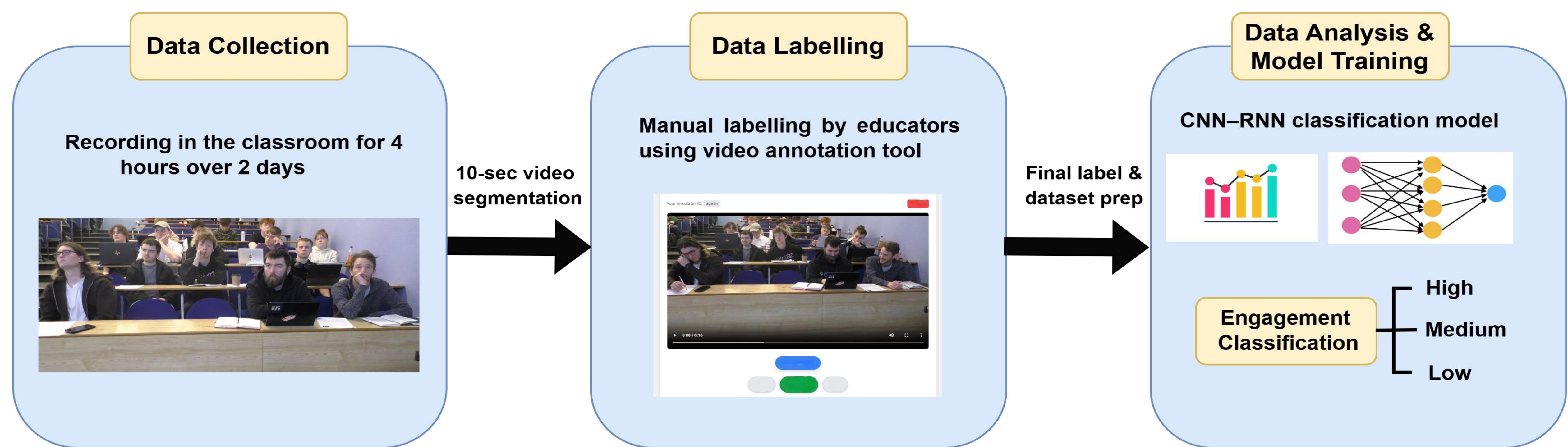


Fig 1. Overview Of Processing And Analysing Classroom Group Engagement From Multimodal Data

- **MERICE** – A Multimodal (audio-visual) classroom engagement dataset
- Focused on classroom-level engagement rather than individual students, labelled as high, low and medium.
- Baseline experiments using deep learning (DL) models.

3

Key Insights

- Engagement annotation showed subjectivity across annotators.
- The temporal analysis showed how engagement changes over time.
- The DL model's performance varied by level; no model was superior.

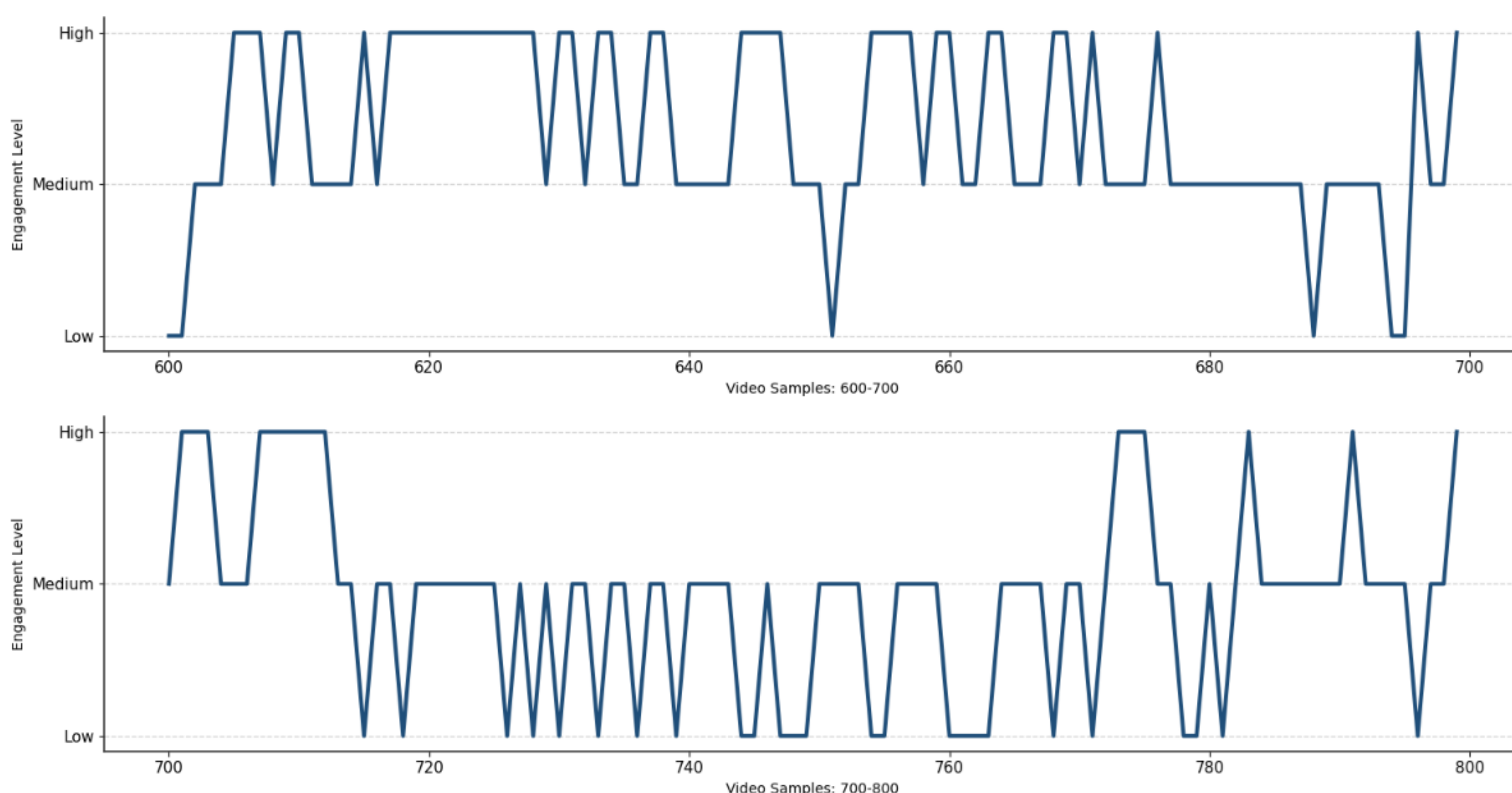


Fig 2. The Engagement Over Time In The Classroom

Model	High (%)	Medium (%)	Low (%)
MobileNetV2-BiLSTM	0.646	0.630	0.641
AttentionNet + MobileNetV2	0.471	0.613	0.627
EfficientNetB0-BiLSTM	0.705	0.482	0.642
VideoMAE	0.716	0.428	0.713

Fig 3. Engagement Level Classification Accuracy

4

Looking Ahead

- Improving DL models by integrating multimodal signals.
- Interviewing teachers before and after developing DL models to understand their views on an AI-enabled classroom.

Please visit our Website Here!



haiti.cis.strath.ac.uk

REFERENCES

- [1] Whitehill, J., Serpell, Z., Lin, Y.C., Foster, A. and Movellan, J.R., 2014. The faces of engagement: Automatic recognition of student engagement from facial expressions. *IEEE Transactions on Affective Computing*, 5(1), pp.86-98.
- [2] Lu, W., Yang, Y., Song, R., Chen, Y., Wang, T. and Bian, C., 2025. A Video Dataset for Classroom Group Engagement Recognition. *Scientific Data*, 12(1), p.644.
- [3] Zaletelj, J. and Košir, A., 2017. Predicting students' attention in the classroom from Kinect facial and body features. *EURASIP journal on image and video processing*, 2017(1), p.80.