

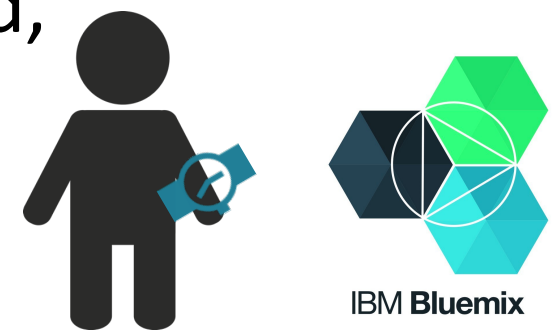
Motivation

- Verbal protest is a commonly reported challenging behavior observed in children with autism
- Definition: sensory overload-induced crying, screaming, shouting, and yelling
- Detection will help us
 - Understand frequency and context of occurrence
 - Predict and prevent
 - Reduce burden on caregiver
 - Develop metric to evaluate the efficacy of therapy
- Can we detect whether a child with autism is exhibiting verbal protest from audio recordings?

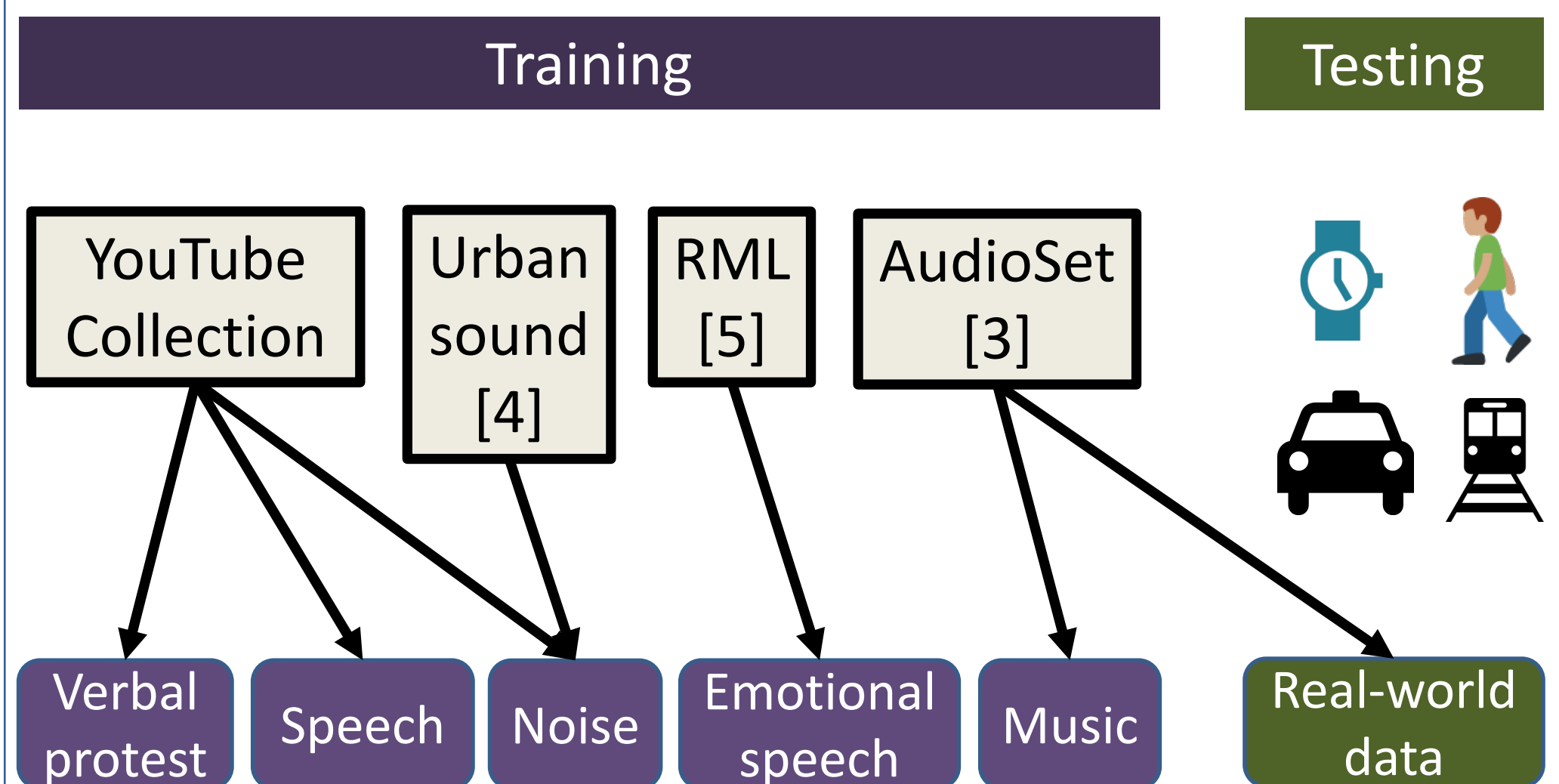


Contribution

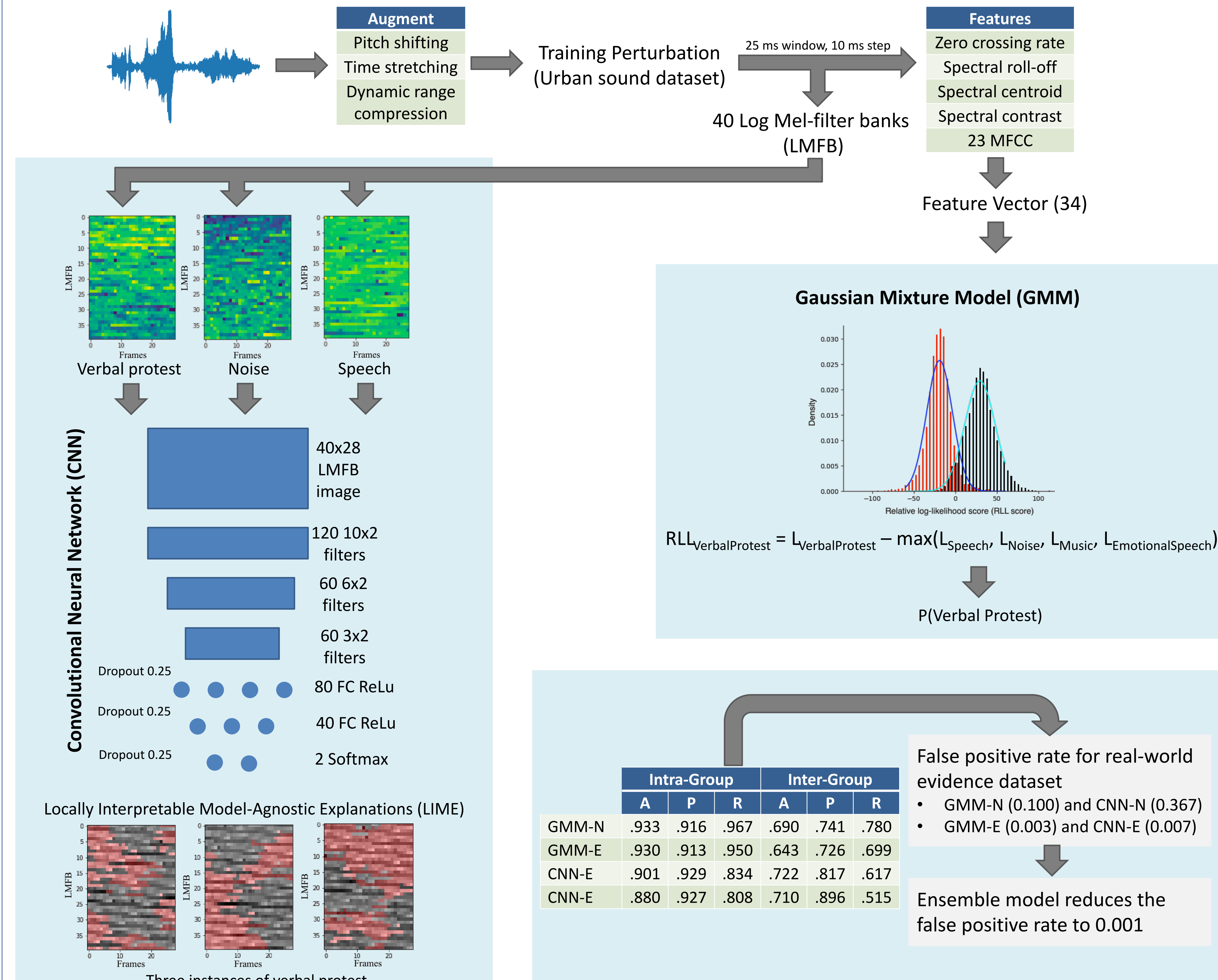
- Curation of a verbal protest audio dataset
- Development of verbal protest detection models
 - Can work in a resource-constrained, real-time setting
 - Noise tolerant
 - Limited false positives



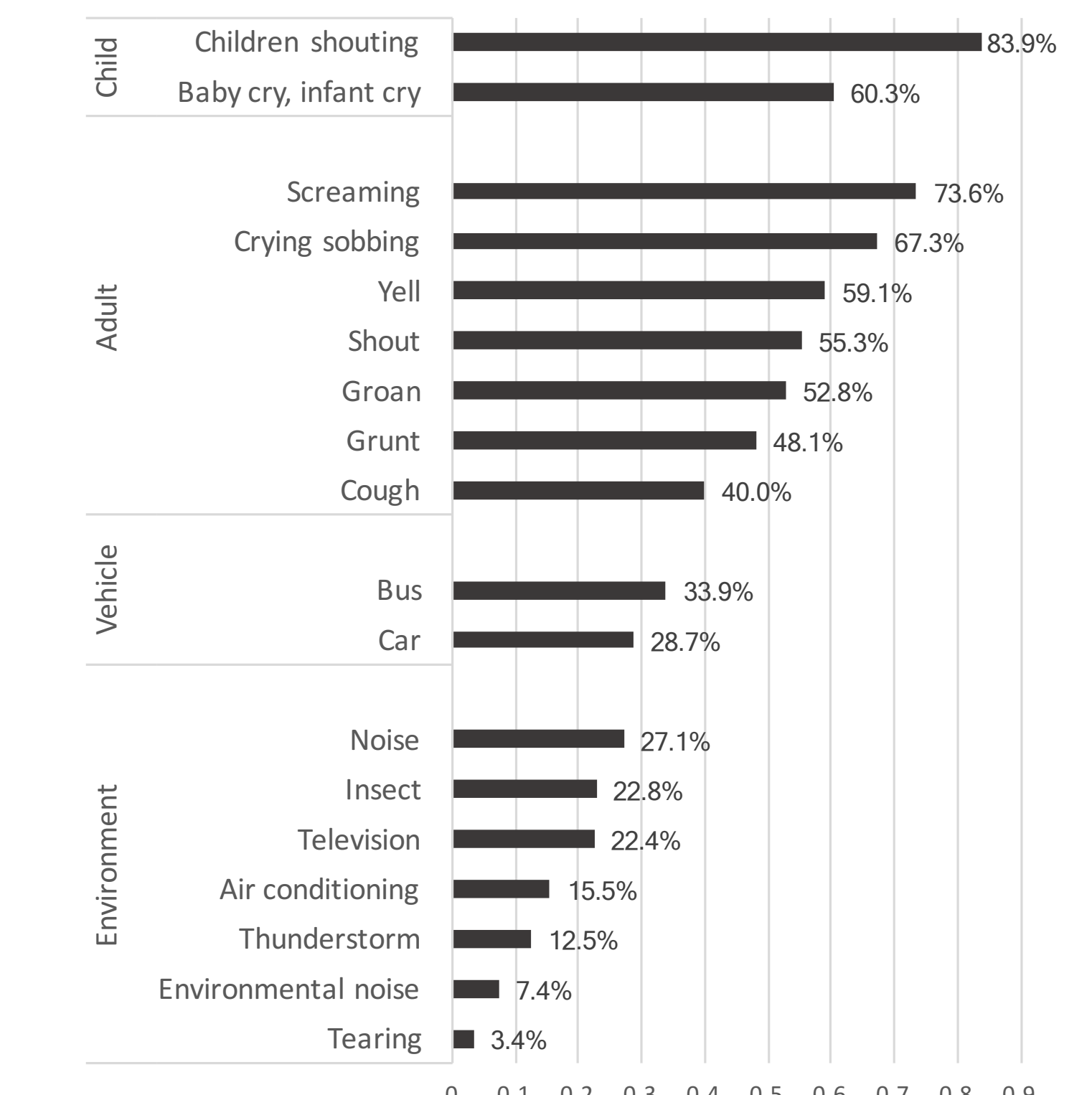
Dataset



Modeling Approach



Validation



Conclusion

- GMM model is lightweight but produces more false positives
- CNN model reduces false positive rate but requires more resources
- Ensemble model can be a superior choice
- Future work is to explore reducing verbal protest episodes by utilizing biomarkers and characteristics of the child and their environment.

Acknowledgements

We would like to thank Soumya Ghosh from IBM Research for the useful comments and discussions on the methodology.

References

- [1] Yizhar Lavner, et al., "Baby cry detection in domestic environment using deep learning," in ICSEE 2016.
- [2] Jouni Pohjalainen, et al., "Detection of shouted speech in the presence of ambient noise," in Interspeech 2011.
- [3] Jort F Gemmeke, et al., "Audio set: An ontology and human-labeled dataset for audio events," in IEEE ICASSP 2017.
- [4] Justin Salamon, et al., "A dataset and taxonomy for urban sound research," in MM 2014.
- [5] Yongjin Wang and Ling Guan, "Recognizing human emotional state from audiovisual signals," IEEE transactions on multimedia, 2008.