

INFORMATICS INSTITUTE OF TECHNOLOGY

In Collaboration with UNIVERSITY OF WESTMINSTER



University of Westminster, Coat of Arms

**SafeView: Predicting Sensory Overstimulation in Children's Videos with
Multimodal Ordinal Regression Modeling**

A Dissertation by
Mr Taalib Asaf
W1866997–2021075

Supervised by Mr Achala Aponso

Submitted partially fulfilling the requirements for the BEng in Software Engineering degree at the University of Westminster.

ABSTRACT

Digital content exposure continues to grow for children adding stress to their sensory abilities because it produces cognitive overload effects and emotional struggles especially if they possess sensory processing sensitivities. Manual annotation of overstimulating content requires human effort and does not scale well to large video datasets thus the method also comes with inherent subjectivity. The SafeView project establishes an automated detection system for sensory overstimulation in videos based on audio-visual characteristics that generate high cognitive demand because of scene transitions, loud effects, and multiple auditory elements to ensure child-friendly media usage.

The SafeView system employs a machine learning pipeline to detect sensory overstimulation, utilizing a Random Forest model trained on a diverse set of audio-visual features extracted from 6-minute video segments. Audio features were derived using advanced signal processing techniques, including sound classification with YAMNet to identify elements like speech, music, and sound effects, and spectral analysis with Librosa. Visual features, including scene frequency and visual complexity, were extracted through computer vision methods using Opencv for scene detection and frame analysis..A few models, such as the Random Forest, were optimised with GroupKFold cross-validation to ensure robustness across diverse video content, enabling a scalable and automated approach to overstimulation detection.

Random Forest achieved an accuracy with an RQWK of 0.86, an RMSE of 0.62, and a Spearman's rank correlation coefficient of 0.82 for identifying overstimulation scores in SafeView. The performance surpassed both SVR measurements (QWK: 0.65, RMSE: 0.85) as well as XGBoost models (QWK: 0.84, RMSE: 0.65).

Subject Descriptors:

- Computing methodologies → Machine learning → Machine learning approaches → Ensemble methods → Random forests
- Computing methodologies → Artificial intelligence → Computer vision → Computer vision tasks → Scene understanding