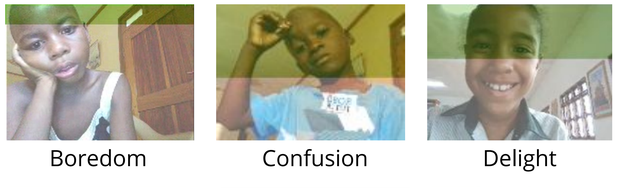
Semi-Supervised Learning to Perceive Children's Affective States in a Tablet Tutor

Mansi Agarwal and Jack Mostow

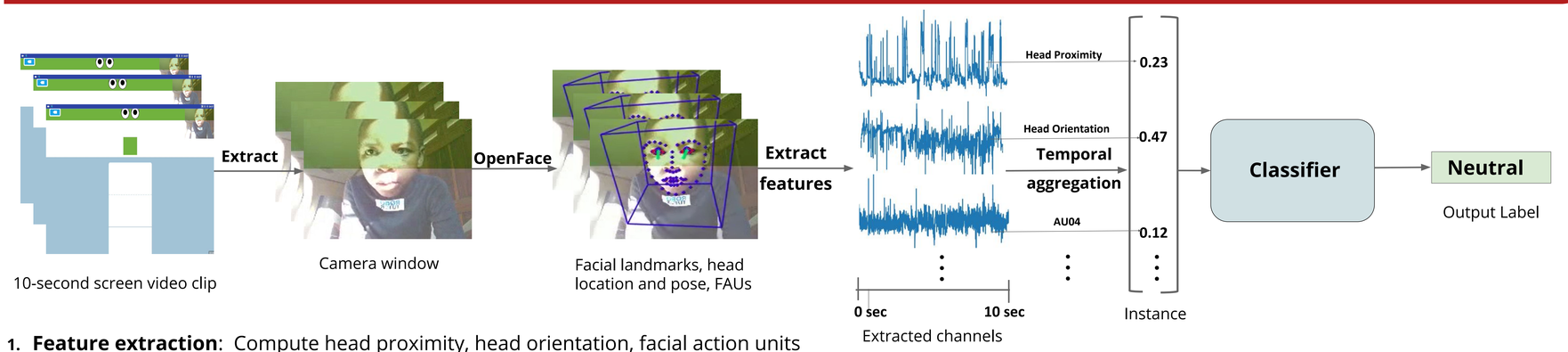
PROBLEM

Detect affective states in videos of children using RoboTutor

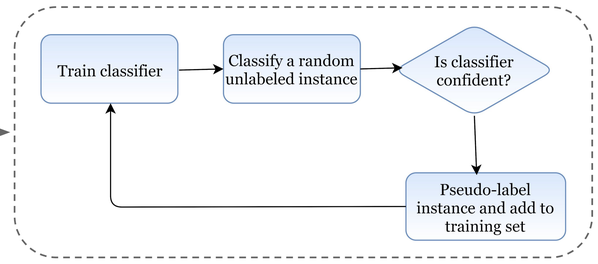


- **Novel population:** children ages 6-12 in Tanzania
- **Camera only:** user-facing tablet camera; no fancy sensors
- **Authentic video:** occlusion, variable indoor and outdoor illumination, limited spatial and temporal resolution

APPROACH

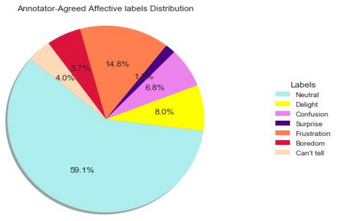


1. **Feature extraction**: Compute head proximity, head orientation, facial action units (FAUs), blink rate, pupil size, eye gaze from OpenFace output.
2. **Temporal aggregation**: Reduce to a single feature value for each 10-second clip.
3. **Semi-supervised learning:** Pseudo-label unlabeled instances to augment sparse hand-labeled training data.



DATASET

- **Dataset:** 229 videos of 30 children using RoboTutor

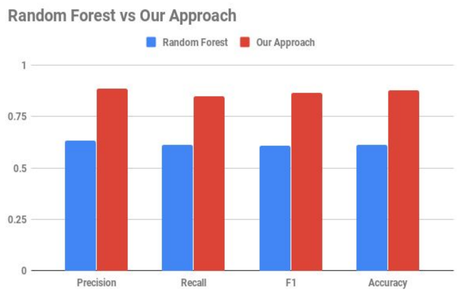


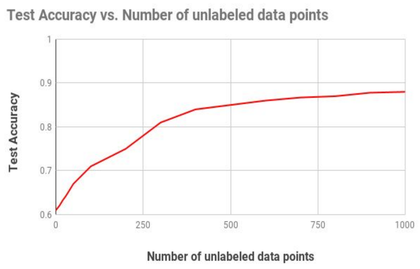
- **Pooled labels:**
  - **Positive:** Delight, Surprise
  - **Negative:** Boredom, Confusion, Frustration
  - **Neutral**
- **Labeled dataset:** 345 10-second clips from 17 videos
  - **Focused sample:** clips with extreme feature values
  - **Random sample:** clips centered at random points
- **Unlabeled dataset**: 1000 clips from 40 other videos

RESULTS COMPARED TO SUPERVISED LEARNING

5-fold cross-validated results

Model	Precision	Recall	F1	Accuracy
Gaussian Naive Bayes	0.52	0.21	0.23	0.21
Decision Tree	0.43	0.44	0.43	0.44
SVM	0.49	0.51	0.46	0.51
Adaboost	0.47	0.47	0.47	0.47
Logistic Regression	0.50	0.53	0.51	0.53
KNN	0.53	0.53	0.53	0.53
Random Forest	**0.63**	**0.61**	**0.60**	**0.61**
Semi-supervised method	**0.88**	**0.84**	**0.86**	**0.88**





CONCLUSION

Unlabelled data helps -- beats supervised learning!

Future work:

- Analyze affective states to guide redesign of RoboTutor.
- Port classifier to run in real-time on tablet.
- Detect and respond to affective states at runtime.
- Use other available inputs -- screen gestures and audio.

ACKNOWLEDGEMENTS

This work would not have been possible without the support of the RoboTutor team, especially Fortunatus Massewe for recording many hours of RoboTutor screen video, and Dr. Leonora Anyango Kivuva, Joash Gambarage, Sheba Naderzad, and Andrew McReynolds for labeling so much data. We would also like to thank Rachel Burcin, Dr. John Dolan and Mikayla Trost for tirelessly supporting us throughout the summer. The first author was a Robotics Institute Summer Scholar supported by an S.N. Bose Scholarship.