Postural developments modulate children's visual access to social information

Alessandro Sanchez*, Bria Long*, Allison Kraus, & Michael C. Frank
Department of Psychology, Stanford University

Background

In the first few years of life, children's view of the world changes radically (e.g., Fausey et al., 2016) due to improvements in locomotive capacity (e.g., Kretch et al., 2014).

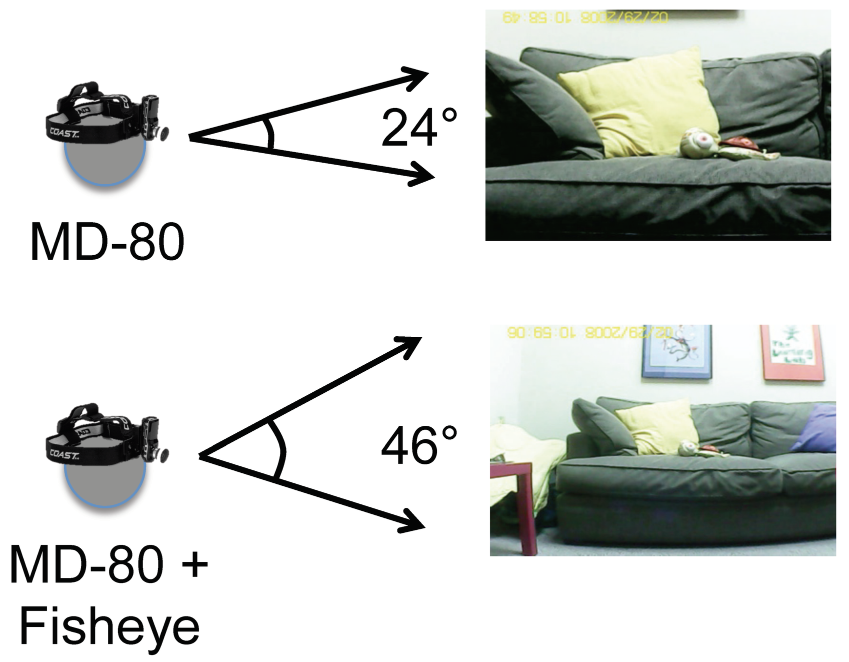
For example, 12-month-olds saw more faces when they were sitting/standing vs. prone/crawling (Franchack, Kretch, & Adolph, 2017).

Postural developments may shape infant's access to social information relevant for linguistic and cognitive development (Frank et al., 2013; Franchack, Kretch, & Adolph, 2017).

Can we replicate these results using a broader age range and modern tools from computer vision?

Methods

Head-mounted camera



Posture: carry, prone, sit, stand
Orientation: behind, close, far
Hand-annotated from 3rd-person video

Participants

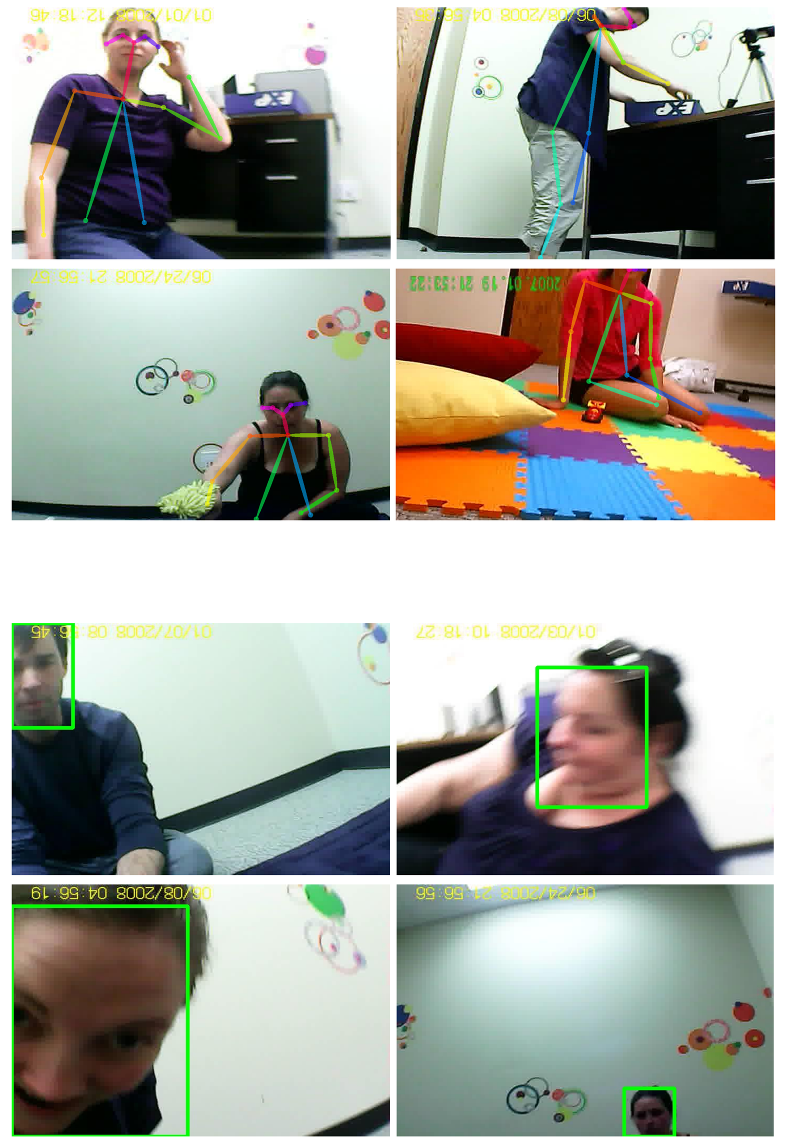
Group	N	% Incl	Avg. age	Avg. Video Length
8 mo.	12	0.46	8.71	14.40
12 mo.	12	0.40	12.62	12.71
16 mo.	12	0.31	16.29	15.10

Computer Vision Algorithm Evaluation

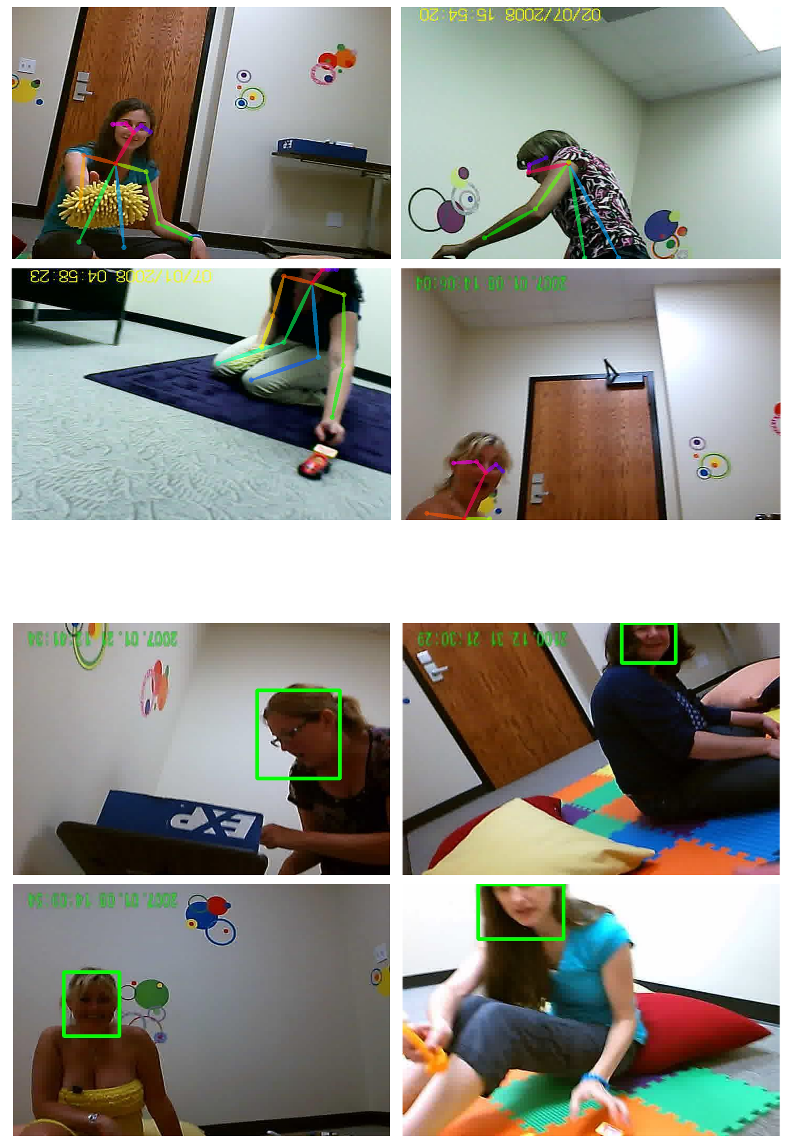
MTCNN: Joint face/key-point detector (Zhang et al., 2016)
OpenPose: Pose/face detector (Cao et al., 2017)
ViolaJones: Widely used face detector pre-CNNs

Algorithm	*Sample*	*P*	*R*	*F*
MTCNN-Faces	High	0.89	0.92	0.90
MTCNN-Faces	Random	0.94	0.62	0.75
OpenPose-Faces	High	0.78	0.93	0.84
OpenPose-Faces	Random	0.72	0.80	0.76
ViolaJones-Faces	High	0.96	0.44	0.60
ViolaJones-Faces	Random	0.44	0.38	0.41
OpenPose-Wrists	High	0.66	1.00	0.79
OpenPose-Wrists	Random	0.88	0.29	0.43

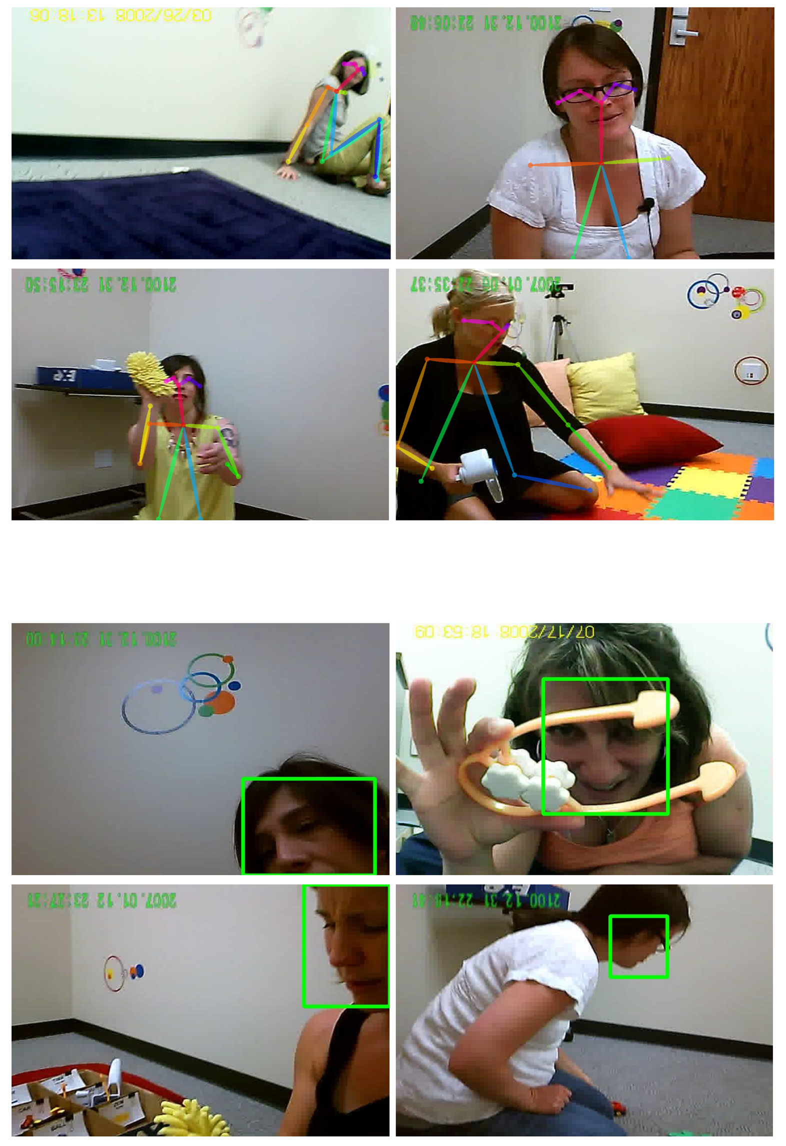
8-month-olds



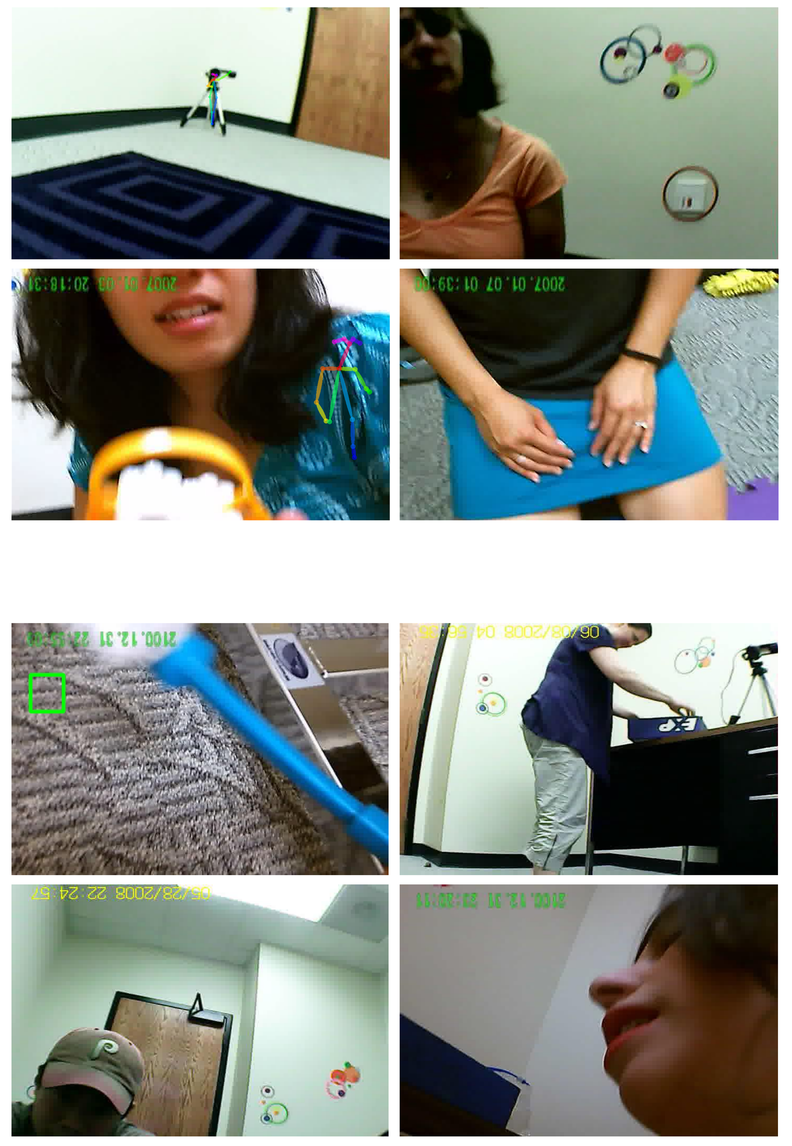
12-month-olds



16-month-olds

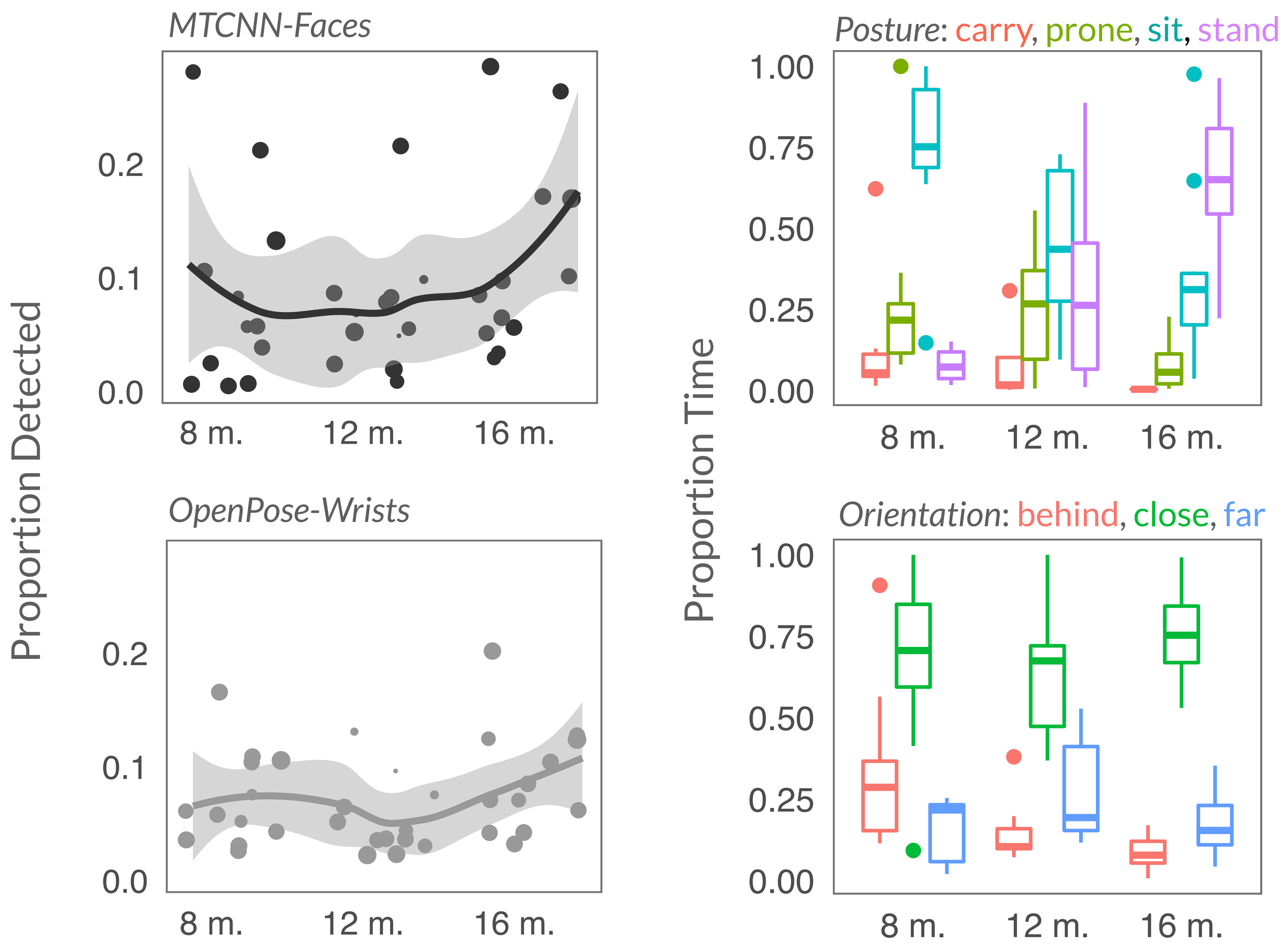


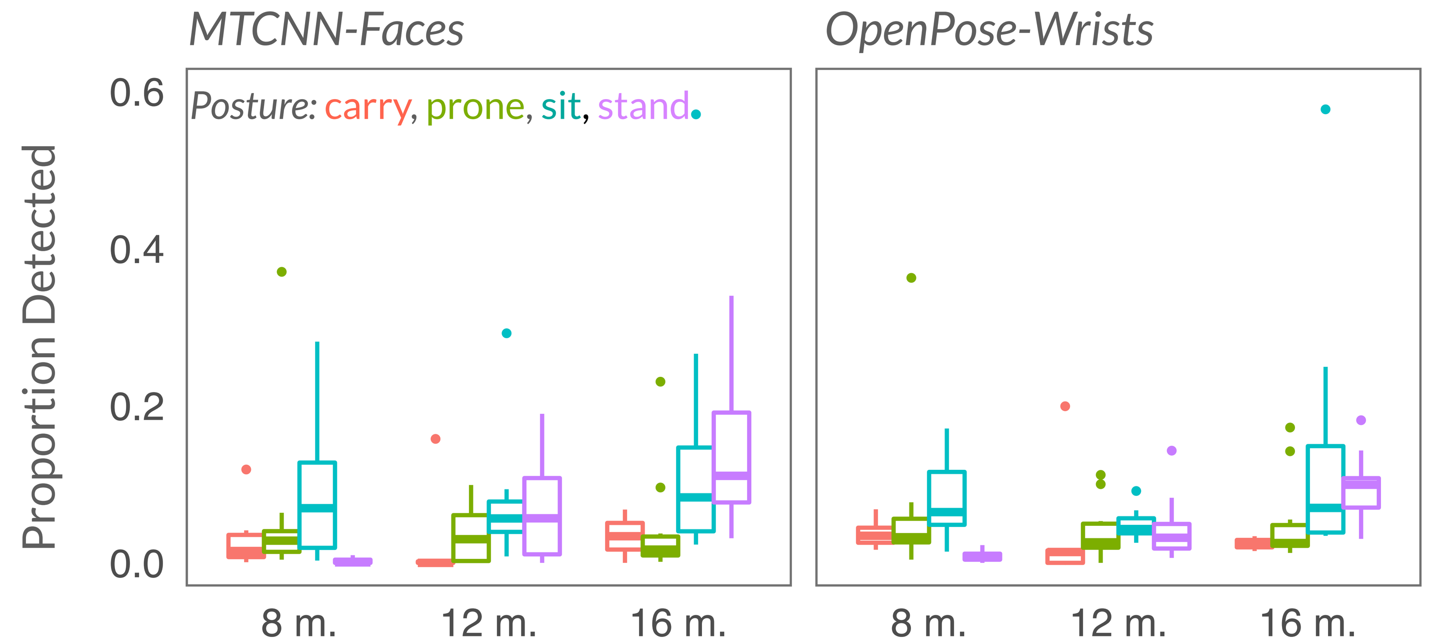
Detection failures

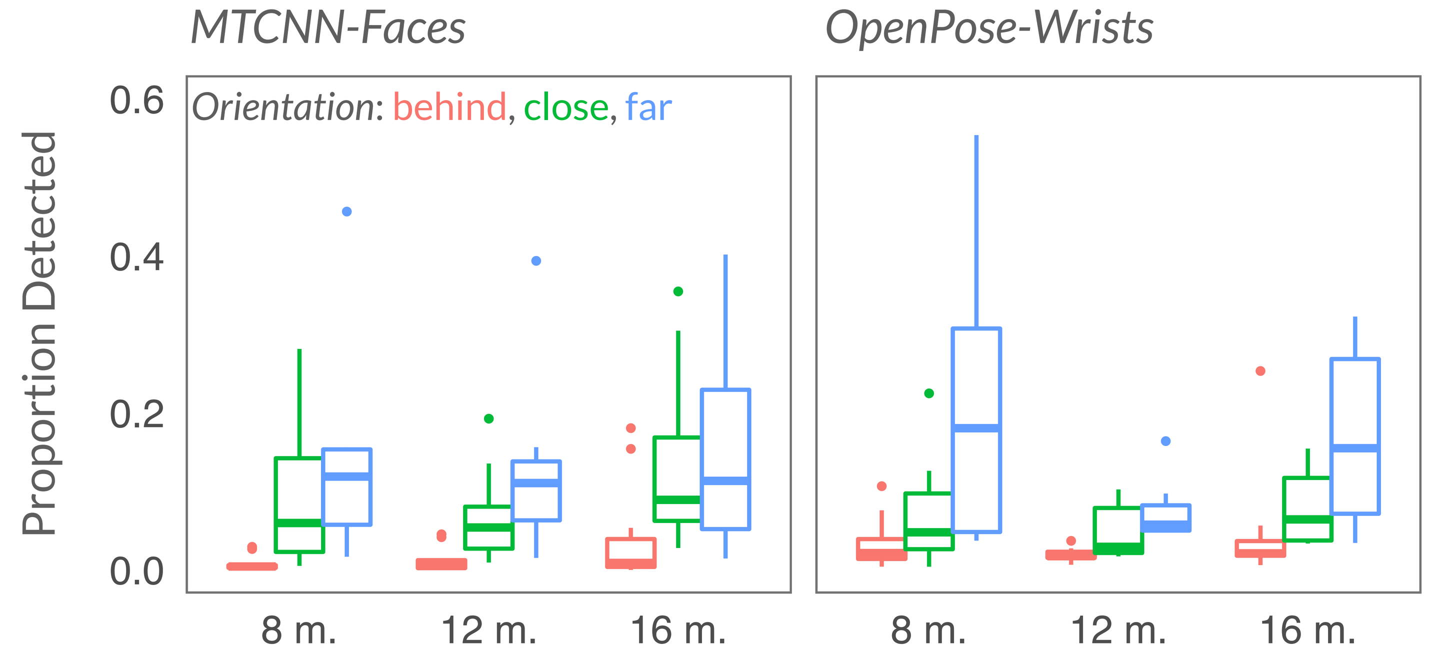


Results

Infant's posture and orientation relative to their caregiver modulated the proportion of faces/wrists detected







Discussion

Children's posture and orientation towards their caregiver changed systematically across age, and both of these factors influenced the proportion of faces/hands in the child's visual field.

Hands/wrists were not as well detected as faces, yet still showed the same overall trends

Future work will explore how object detection models can be applied to egocentric videos

Goal: create open source tools for analyzing natural egocentric visual experience at scale

References

Cao, Z., Simon, T., Wei, S.-E., & Sheikh, Y. (2017). Realtime multi-person 2D pose estimation using part affinity fields. *CVPR* .

Franchak, J. M., Kretch, K. S., & Adolph, K. E. (2017). See and be seen: Infant–caregiver social looking during locomotor free play. *Developmental Science.*

Frank, M. C., Simmons, K., Yurovsky, D., & Pusiol, G. (2013). Developmental and postural changes in childrens visual access to faces. In *Proceedings of the 35th Annual Meeting of the Cognitive Science Society*.

Fausey, C. M., Jayaraman, S., & Smith, L. B. (2016). From faces to hands: Changing visual input in the first two years. *Cognition.*

Kretch, K. S., Franchak, J. M., & Adolph, K. E. (2014). Crawling and walking infants see the world differently. C*hild Development.*

Zhang, K., Zhang, Z., Li, Z., and Qiao, Y. (2016). Joint face detection and alignment using multitask cascaded convolutional networks. I*EEE Signal Processing Letters.*

All data and code for these analyses are available at https://github.com/amsan7/xs-face