BLUR THE BABY!

Harold Kim & Mary Zhuo Ke

What is the Problem?

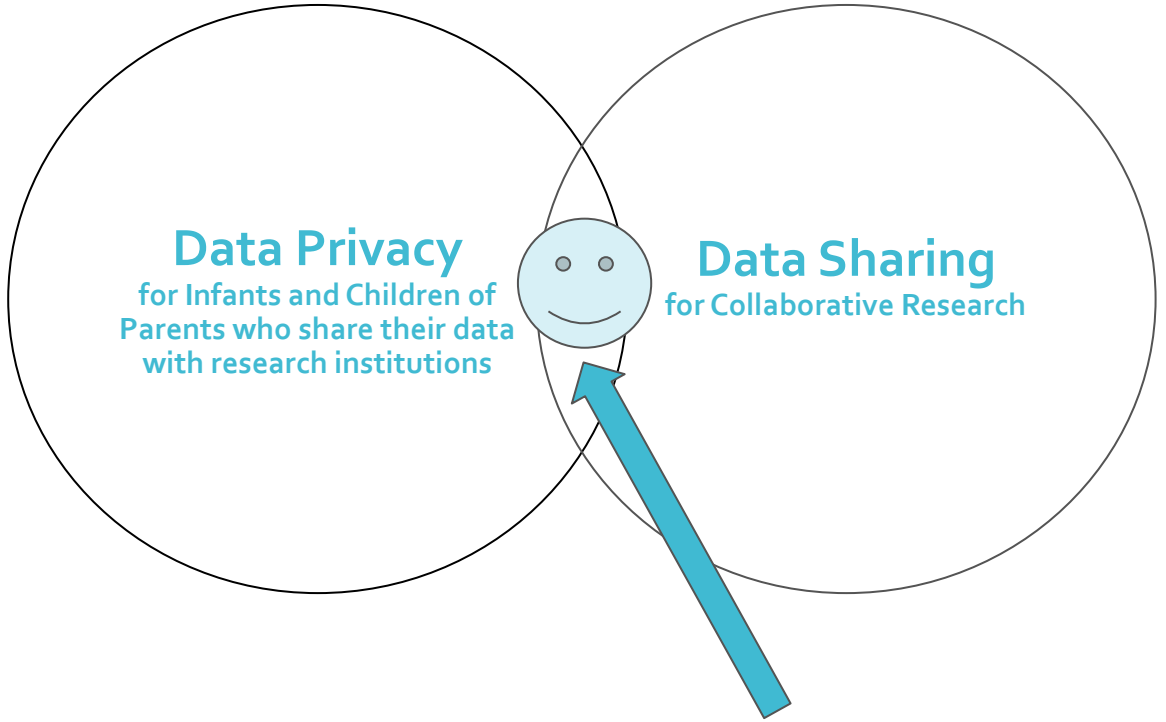
Data Privacy
for Infants and Children of Parents who share their data with research institutions

- In 2016, Feinstein Institute for Medical Research was **fined $3.9 Million** for disclosing identifying data and violating HIPAA.
- https://www.hhs.gov/hipaa/for-professionals/compliance-enforcement/agreements/feinstein/index.html

Data Sharing
for Collaborative Research

- **Identifiable data cannot be shared easily,** limiting further research on data types that can be better informed by video recordings
- **Consent forms** given to participants/patients **are not intuitive,** leading to sharing without full awareness of the risks.

How will we solve it?



Blur the Baby!

- Takes video recordings of infants/children
- Detects identifiable faces with assessment of overall accuracy
- Blurs faces frame by frame
- Compiles a de-identified video that is safer to share in the research community.

Why videos?

- They are a **large data type** that otherwise is just archived, despite the richness in the features possible with video data.
- They take up a lot of space, so extracting features in tabular form would **save a lot of cost and storage** (secondary goal to de-identification).
- They have **more complexities** that pose a challenge to evaluate compared to photos (e.g., animations, transitions from frame to frame, blurred movements), which **makes it a great deep learning problem.**

What will Blur the Baby achieve?

- Protect privacy of **minors who are not able to give consent yet, thus reducing their risk** of privacy concerns when they are older.
- **Increase trust** from study participants that their data is handled securely
- Motivate ways to improve privacy consent **without completely discarding valuable data.**
- **Enable privacy-informed data sharing** by removing the identifiable piece of video recording data so that researchers can use other elements of the data.
- **Stress-test** existing tools, find essential areas of improvement

What kind of solution do we need?

Affordably Efficient

There shouldn't be a high cost to data privacy. If we can propose a solution that d**oesn't take too much time and heavy resources**, that would be ideal.

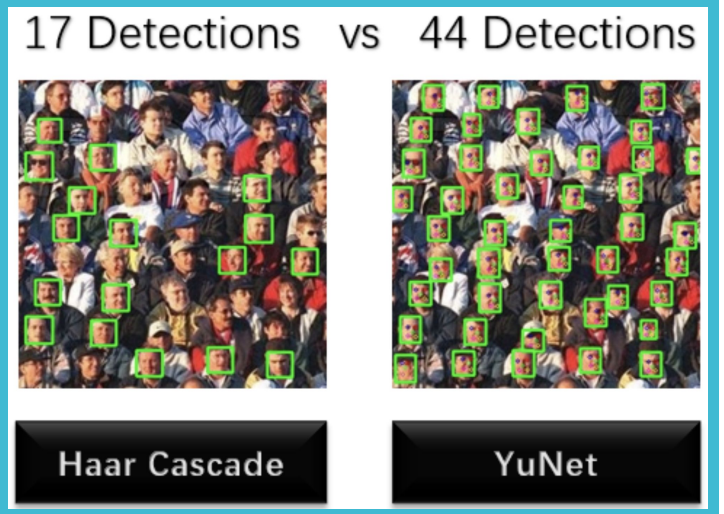
The **solution should scale**, so if you have 1M videos, you don't want it to take forever.

100% Accuracy

We **can't** have partially blurred videos, so we need to aim for 100% accuracy.

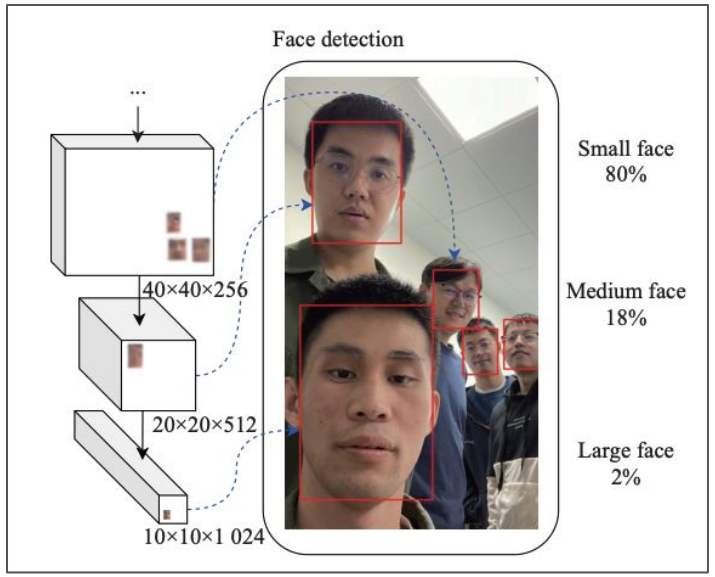
What's the point of de-identification if one of the frames has an exposed face?

Solutions available today



https://opencv.org/blog/opencv-face-detection-cascade-classifier-vs-yunet/

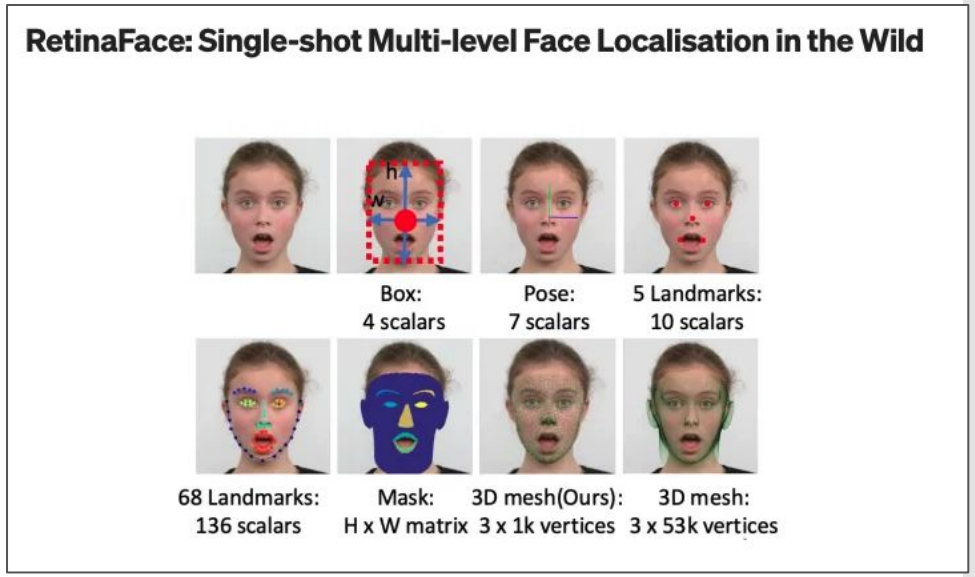
YuNet



Wu, W., Peng, H. & Yu, S. YuNet: A Tiny Millisecond-level Face Detector. Mach. Intell. Res. 20, 656–665 (2023). https://doi.org/10.1007/s11633-023-1423-y

- CNN-based
- trained on WIDER FACE dataset
- lightweight, great for edge devices
- focuses on small faces that are hard to detect
- tiny feature pyramid network

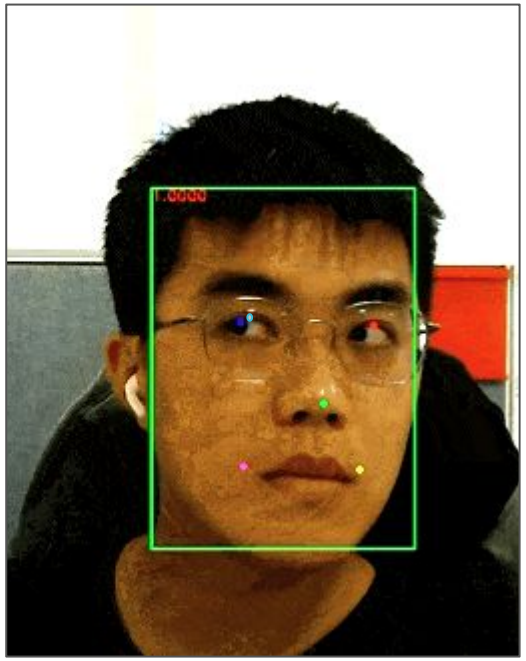
RetinaFace



https://medium.com/analytics-vidhya/exploring-other-face-detection-approaches-part-1-retinaface-9b00f453fd15

- trained on WIDER FACE dataset
- ResNet, MobileNet using pretrained ImageNet
- computationally intense
- uses comprehensive techniques that help estimate 3D attributes too
- focuses on accuracy at multiple scales

Face Detection Output Data

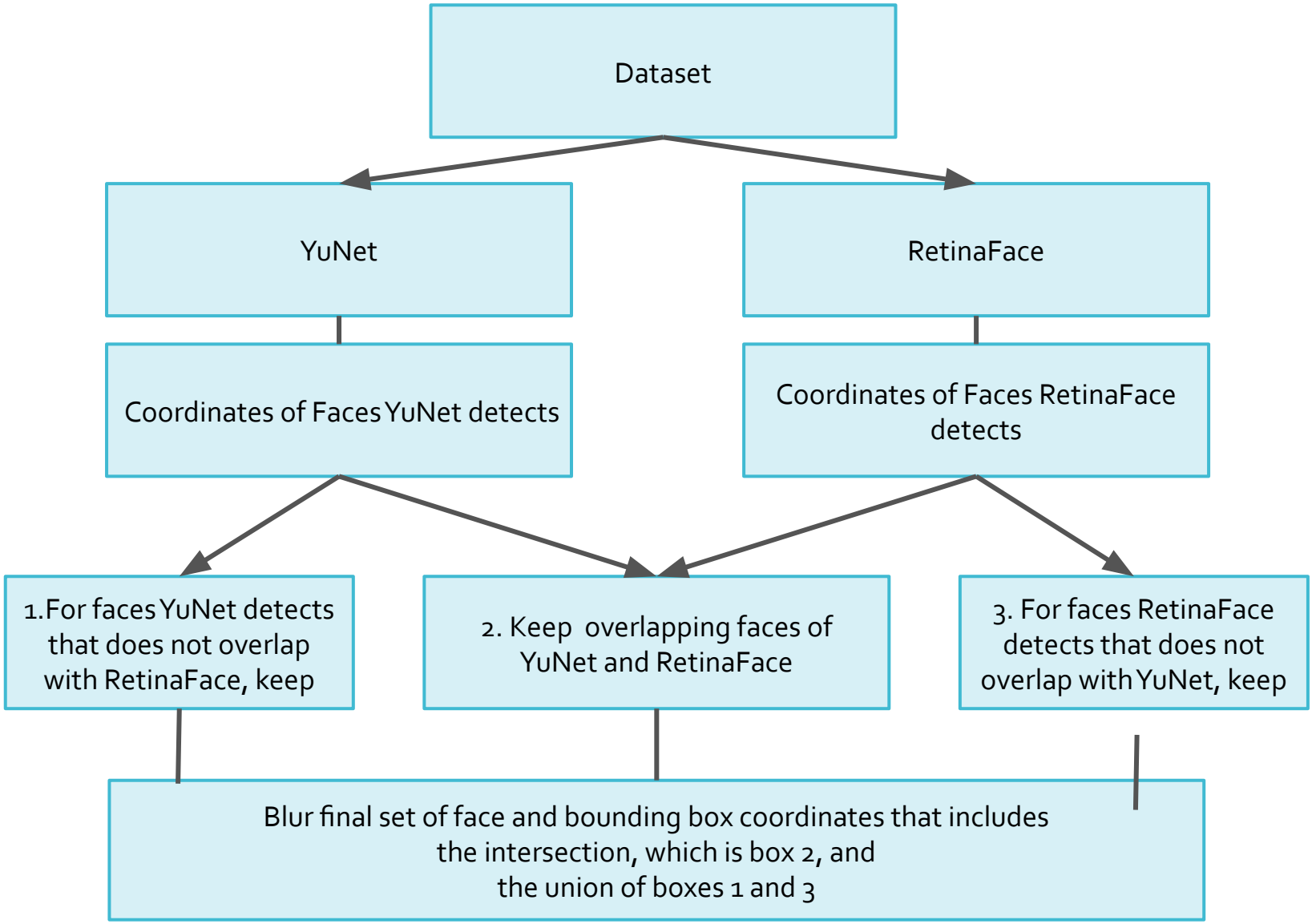


- Face features:
  - left eye
  - right eye
  - nose
  - left edge of mouth
  - right edge of mouth
- bounding box of the face
- confidence score in face detection

https://github.com/opencv/opencv_zoo/tree/main/models/face_detection_yunet

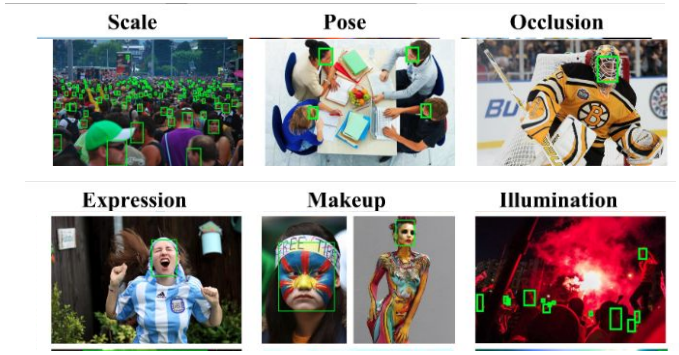
BTB (Blur The Baby) model

Intersection and Union Ensemble Model of YuNet and RetinaFace



Datasets

WIDER FACE validation data



- Focused on validation datasets of the **Family Group** photos
- Ground truth labels of number of faces per photo
- http://shuoyang1213.me/WIDERFACE/

Brainy Baby Video

- Videos of babies and toddlers that are for public use
- https://archive.org/about/
- Used cv2 python library
- Breakdown .mp4 video into individual frames
  - 30 frames per second video
  - 14,500+ frames used
- Over 15 different baby faces, 10 different toddler faces, and 5 different adult faces
- Multiple orientations and occluded faces

Assessment of models using WIDER FACE validation dataset

Quick sanity check: WIDER FACE validation data

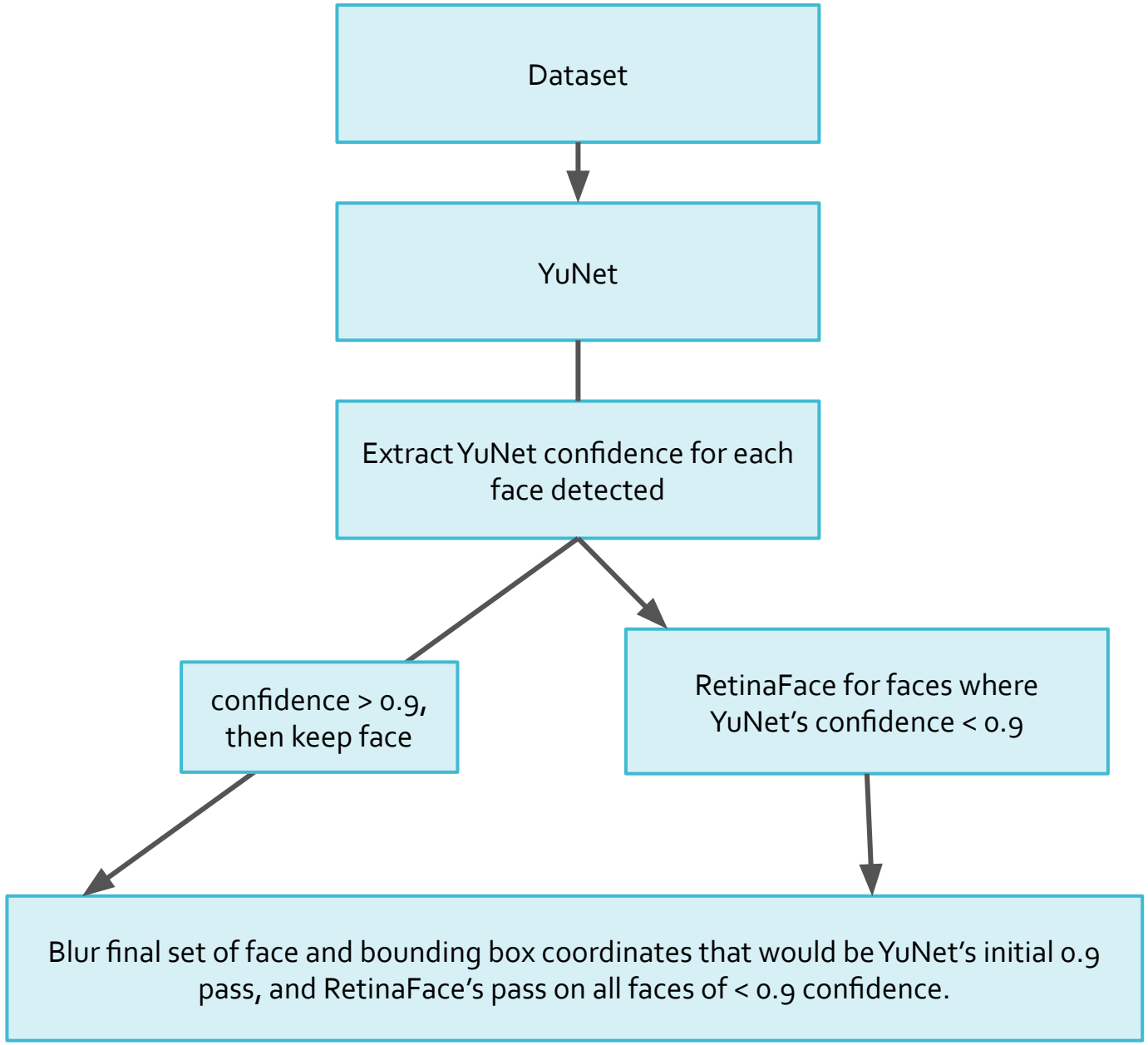
Metric	YuNet	RetinaFace	YuNet + RetinaFace
Number of Images	58	58	58
Time to Blur	1.71 seconds	27.85 seconds	29.49 seconds
Accuracy	36.21%	77.59%	77.59%

We see that **RetinaFace outperformed** for **WIDER FACE validation photos**, as expected, but, this might have something to do with WIDER FACE dataset not being enough of a challenge, like we mentioned.

So let's take it up a notch…

BTB (Blur The Baby) w/ YuNet as first pass and RetinaFace as second pass

that takes ⅓ of the time.

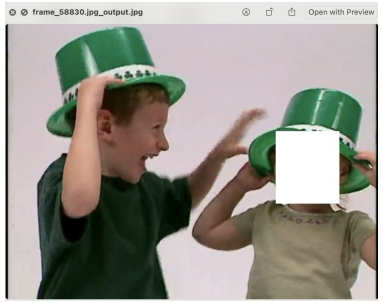


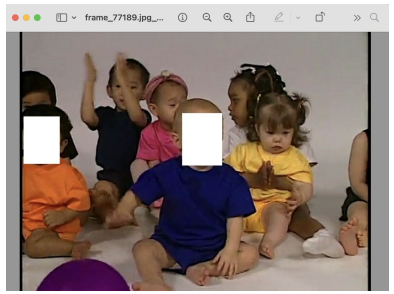
Assessment of BTB using Brainy Baby Video

CHALLENGE CASE: BABY VIDEO

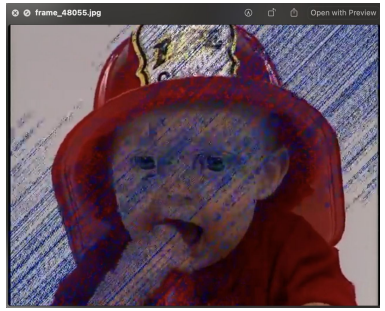
Metric	YuNet	RetinaFace	BTB w/ Intersection and Union	BTB w/ YuNet as first pass and RetinaFace as second pass
Number of Frames (with and without human faces)	14,434	14,434	14,434	14,434
Number of Faces (Human or Not) Detected	11,066	15,456	15,479	~15,400
Time to Blur	5.95 minutes	2.72 hours	2.75 hours	1 hour

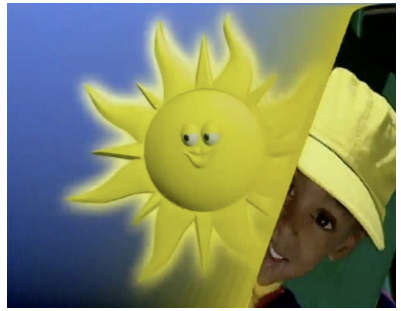
Cases YuNet misses





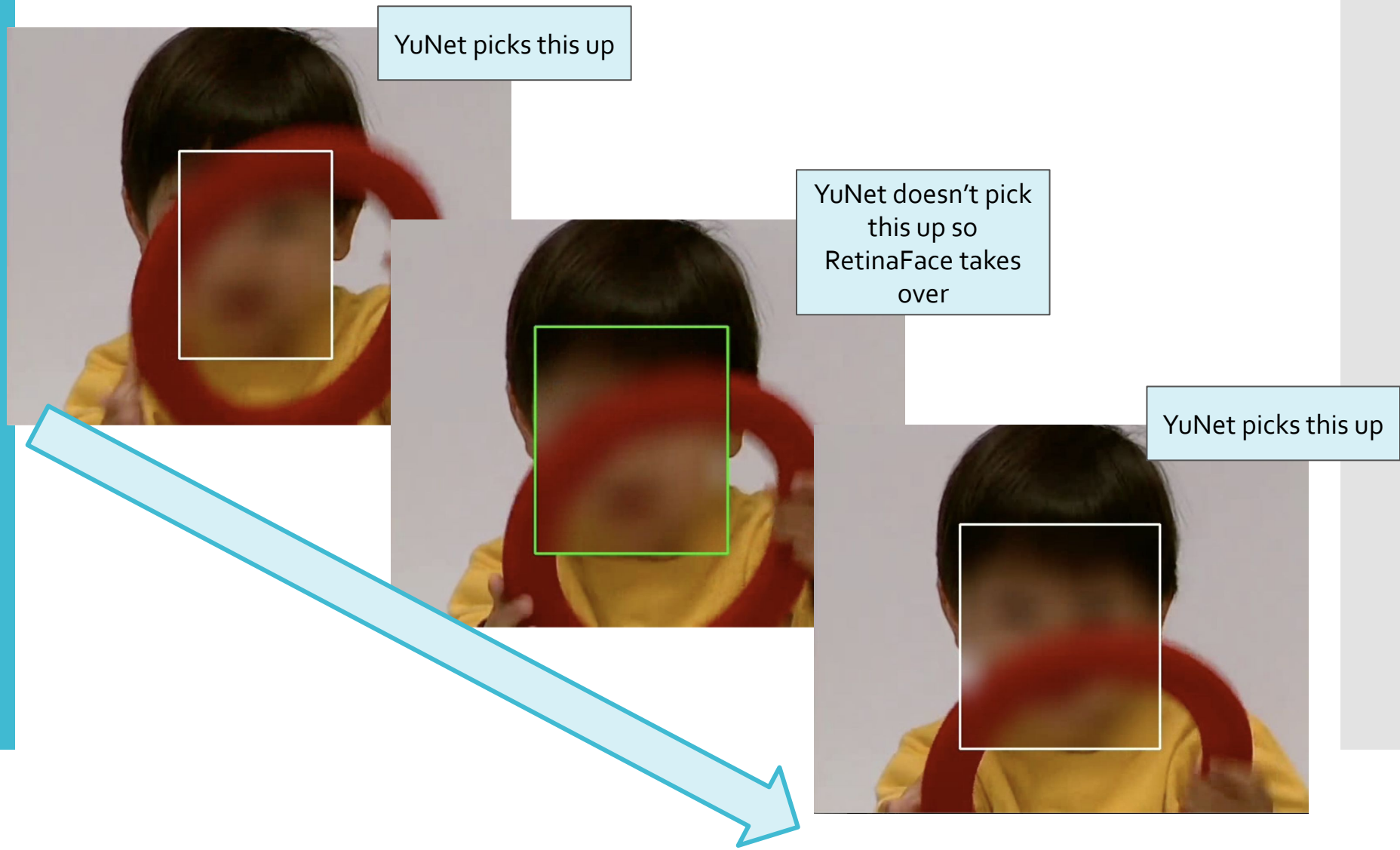
Cases RetinaNet misses





Video Demonstration of Final Output Video (30 seconds)

Final blurred video:
https://drive.google.com/file/d/115nT6ytIaGr_e2PAyakU_4yG_iZp4Jqn/view?usp=sharing



Future work that could enable better solutions

- Having **more face datasets that are labeled** to enable better training and assessment of accuracy
- **Handle transitions/animations more robustly** as videos increase the variations of orientations/scales, especially if it were 60 fps.
- More **efficient ways** to **construct 3D poses** in order to differentiate faces of people from those of animals, objects, portraits

References

1. Yang, S., Luo, P., Loy, C. C., & Tang, X. (2016). WIDER FACE: A face detection benchmark. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 5525–5533. http://shuoyang1213.me/WIDERFACE/

2. Wu, W., Peng, H. & Yu, S. YuNet: A Tiny Millisecond-level Face Detector. Mach. Intell. Res. 20, 656–665 (2023). https://doi.org/10.1007/s11633-023-1423-y
   a. https://github.com/opencv/opencv_zoo/blob/main/models/face_detection_yunet/face_detection_yunet_2023mar.onnx (need to download this model)
   b. Support from chatGPT to help understand YuNet.

3. J. Deng, J. Guo, E. Ververas, I. Kotsia and S. Zafeiriou, "RetinaFace: Single-Shot Multi-Level Face Localisation in the Wild," 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), Seattle, WA, USA, 2020, pp. 5202-5211, doi: 10.1109/CVPR42600.2020.00525.
   a. Support from chatGPT to help understand RetinaFace.

4. Baby video dataset: https://archive.org/details/BrainyBabyShapesandColors