

Extreme Amodal Face Detection

Changlin Song¹, Yunzhong Hou¹, Michael Randall Barnes², Rahul Shome¹,
Dylan Campbell¹

¹The Australian National University, ²University of Oslo



Australian
National
University

WACV
TUCSON, AZ

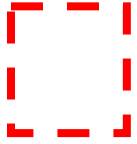


2026
3/6 - 3/10



UNIVERSITY
OF OSLO

Motivation



- A standard positive detection: Object visible and in-frame
- In reality: Objects don't just disappear if out-of-frame

Motivation

Privacy
(Active Avoidance)



REC
Recording Blocked:
Face Predicted
Off-Screen

IMU

Predicts approaching faces & prevents capture.

Safety
(Autonomous Driving)



RISK:
Pedestrian entering

Anticipates risks before they enter frame.

Accessibility
(Visually Impaired)



Audio Cue:
"Someone standing approx. 60° to your right."

Provides out-of-view situational awareness.

Extreme Amodal Face Detection

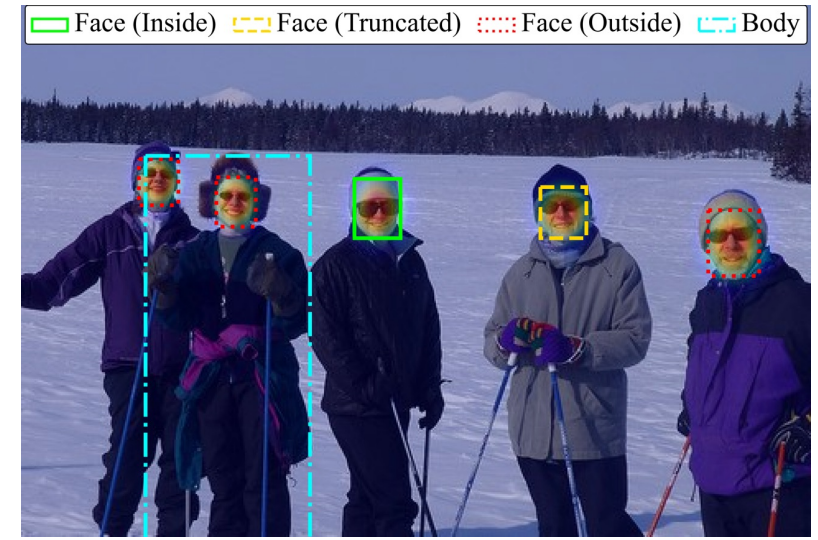
$$\underbrace{\mathbb{R}^{H \times W \times 3}}_{\text{Input Image } (x)} \rightarrow \underbrace{\{(c_i, b_i)\}_{i=1}^N}_{\text{Objects } (o)} \times \underbrace{[0,1]^{KH \times KW}}_{\text{Heatmap } (h)}$$



(a) Input Image x



(b) Prediction (o, h)

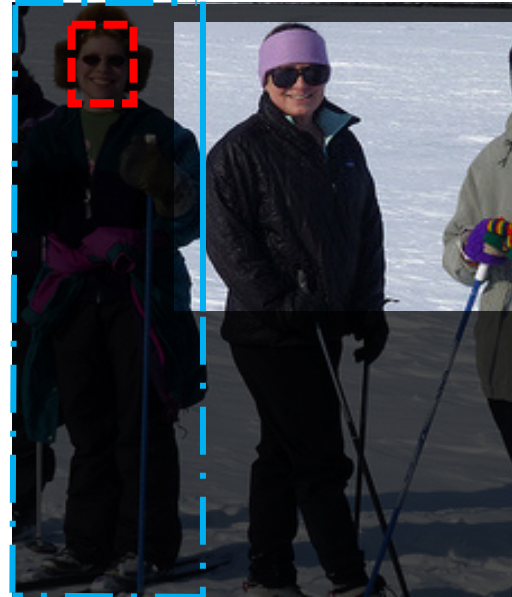


(c) GT Realization with Annotations

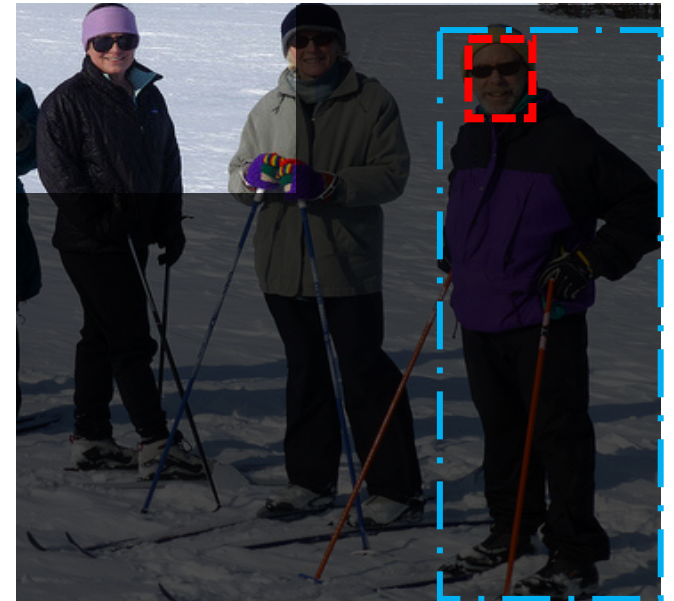
Taxonomy of Faces



Truncated



Outside +

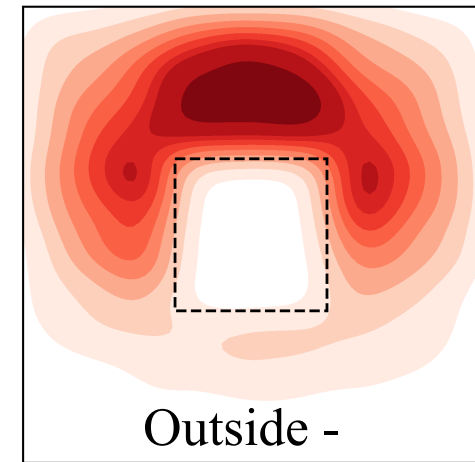
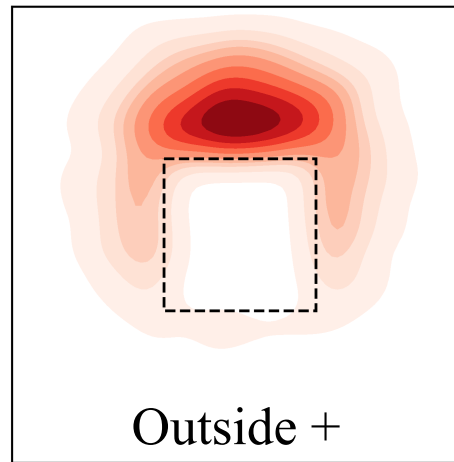
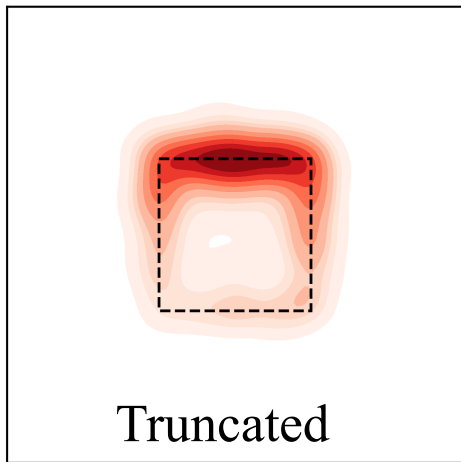


Outside -

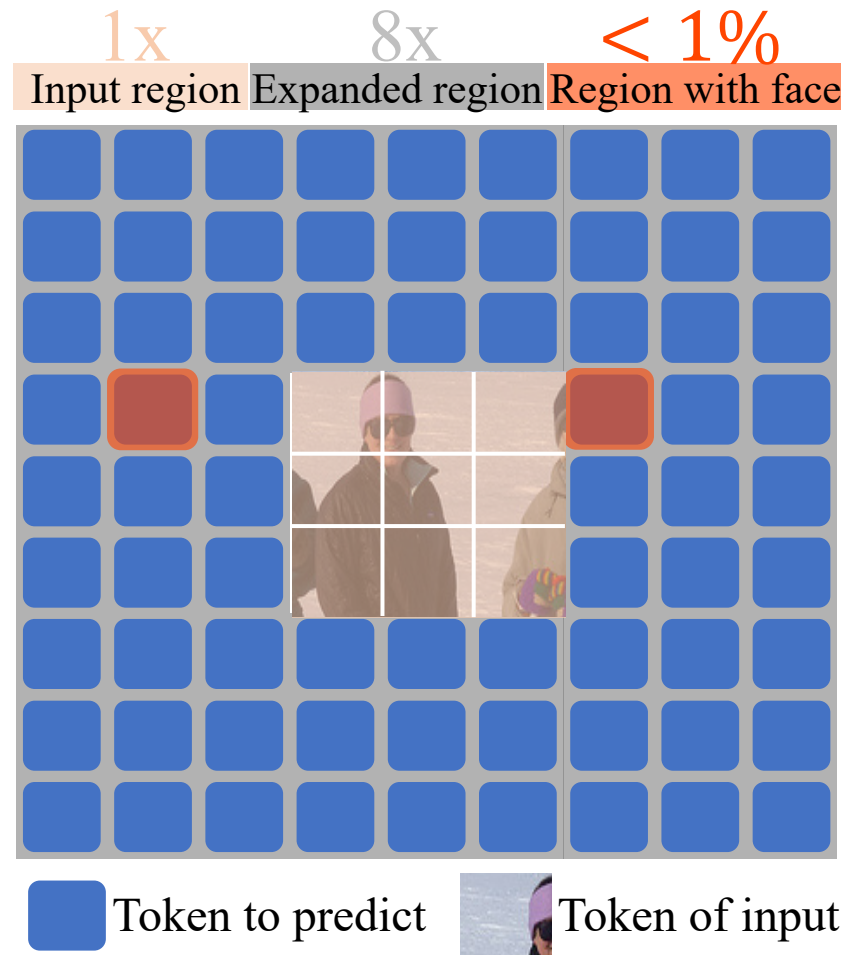
Extreme Amodal Face Dataset (EXAFace)

~182k images for train

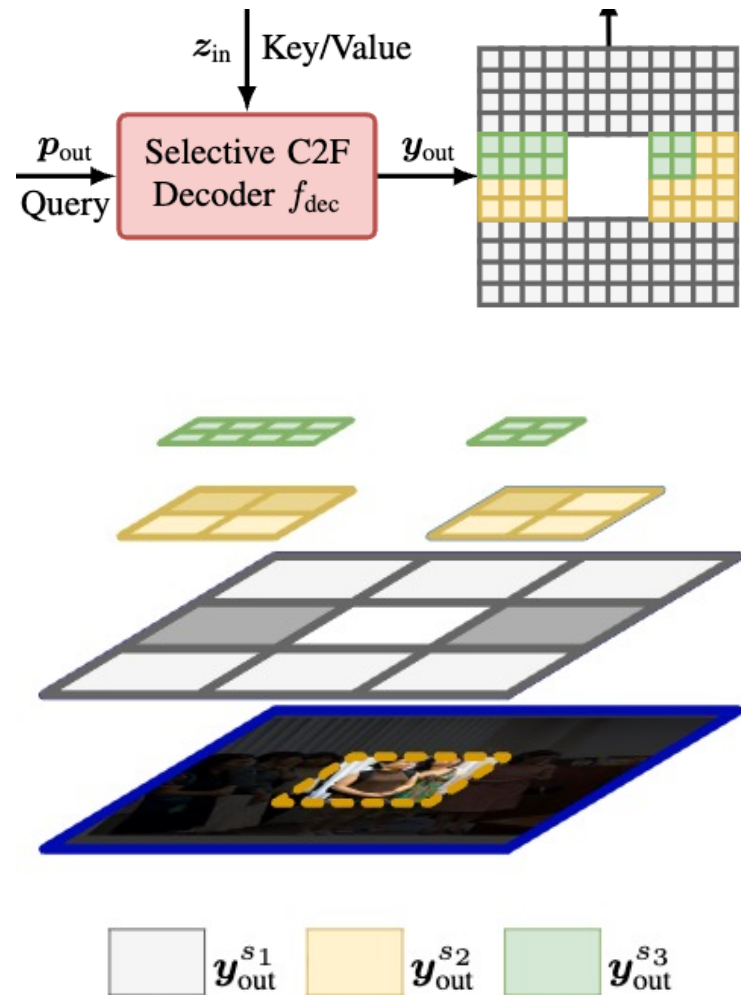
~7k images for test



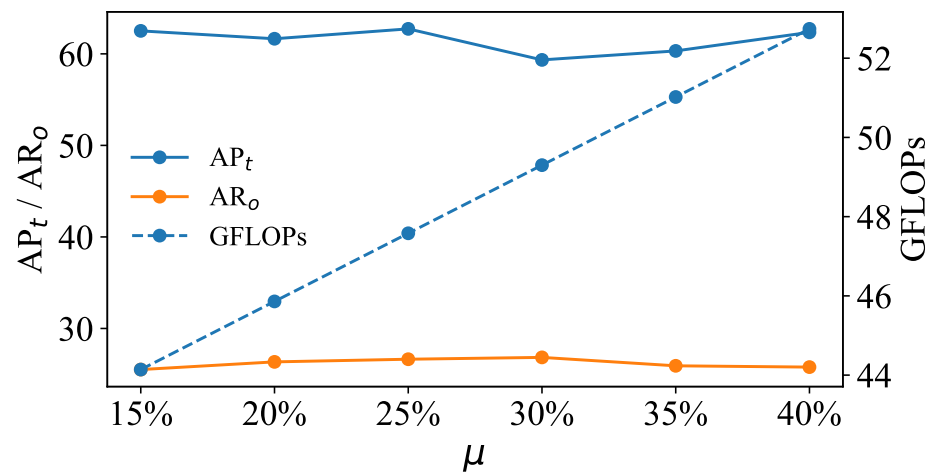
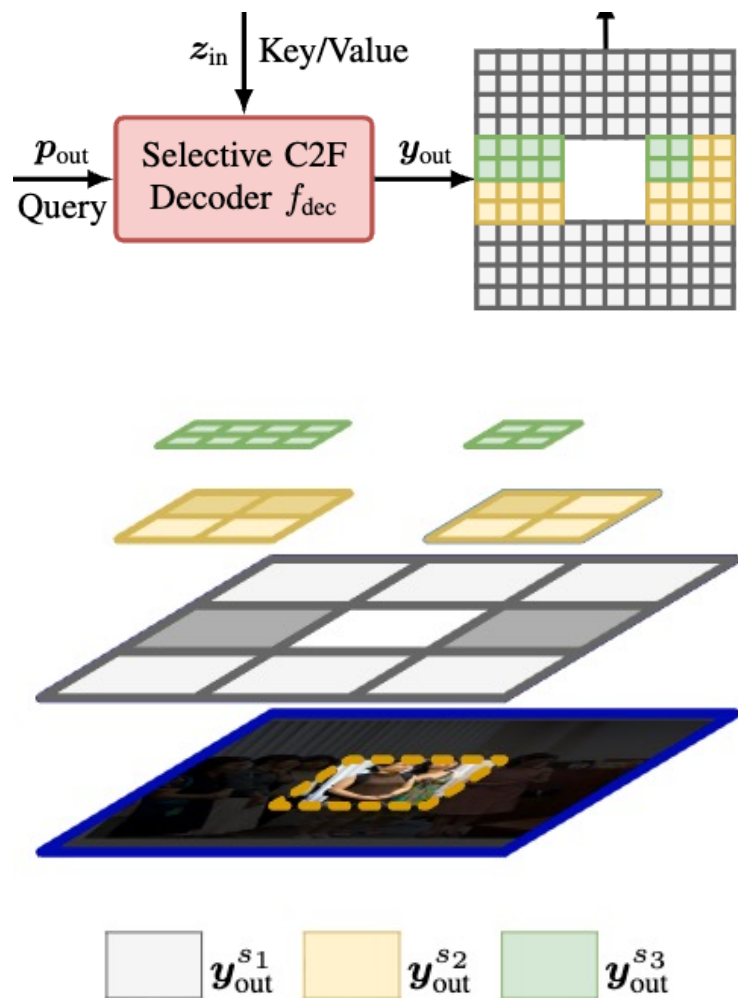
Challenges



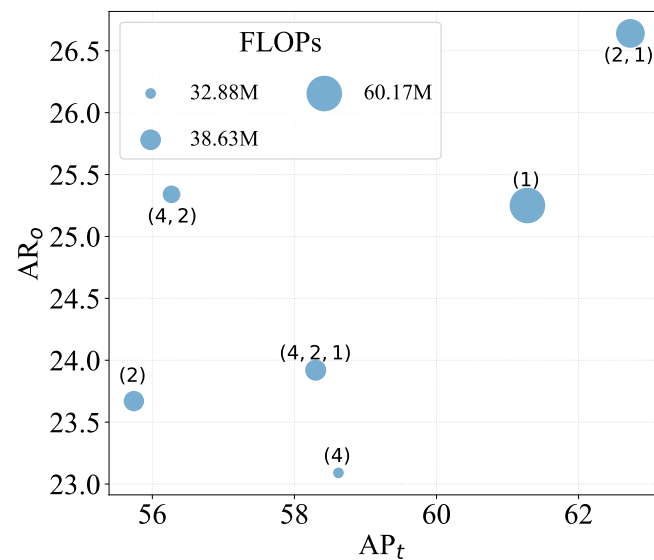
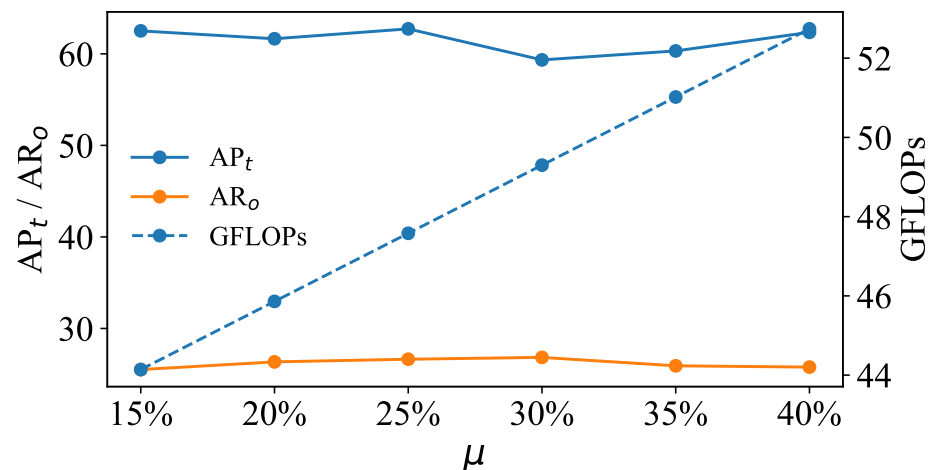
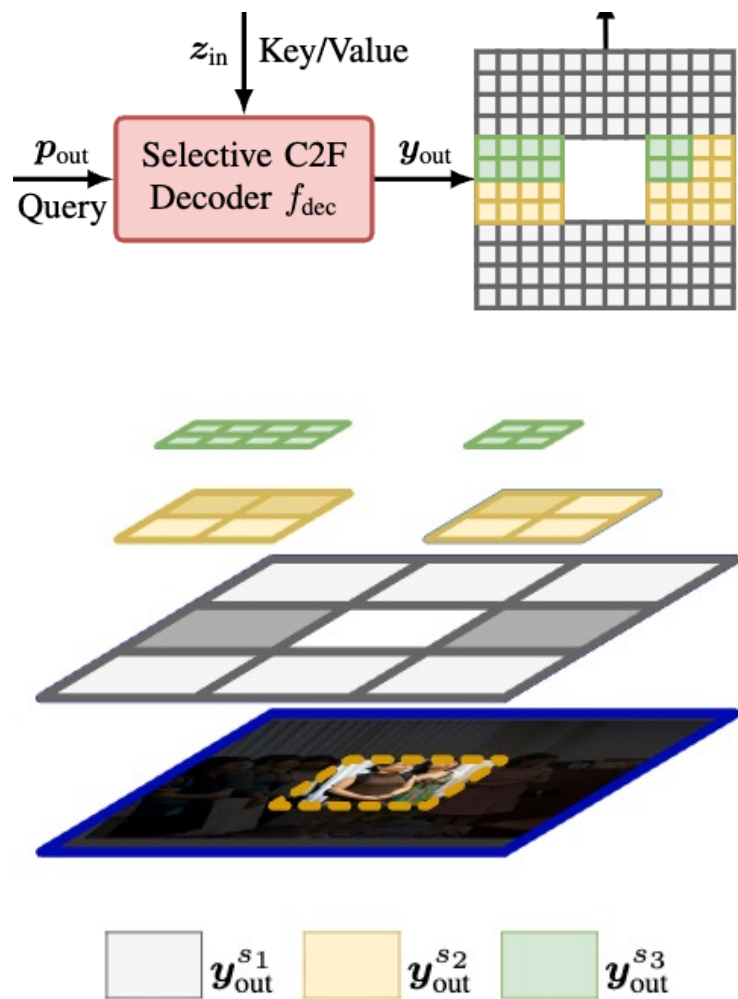
Method: Selective Coarse-to-Fine Decoder



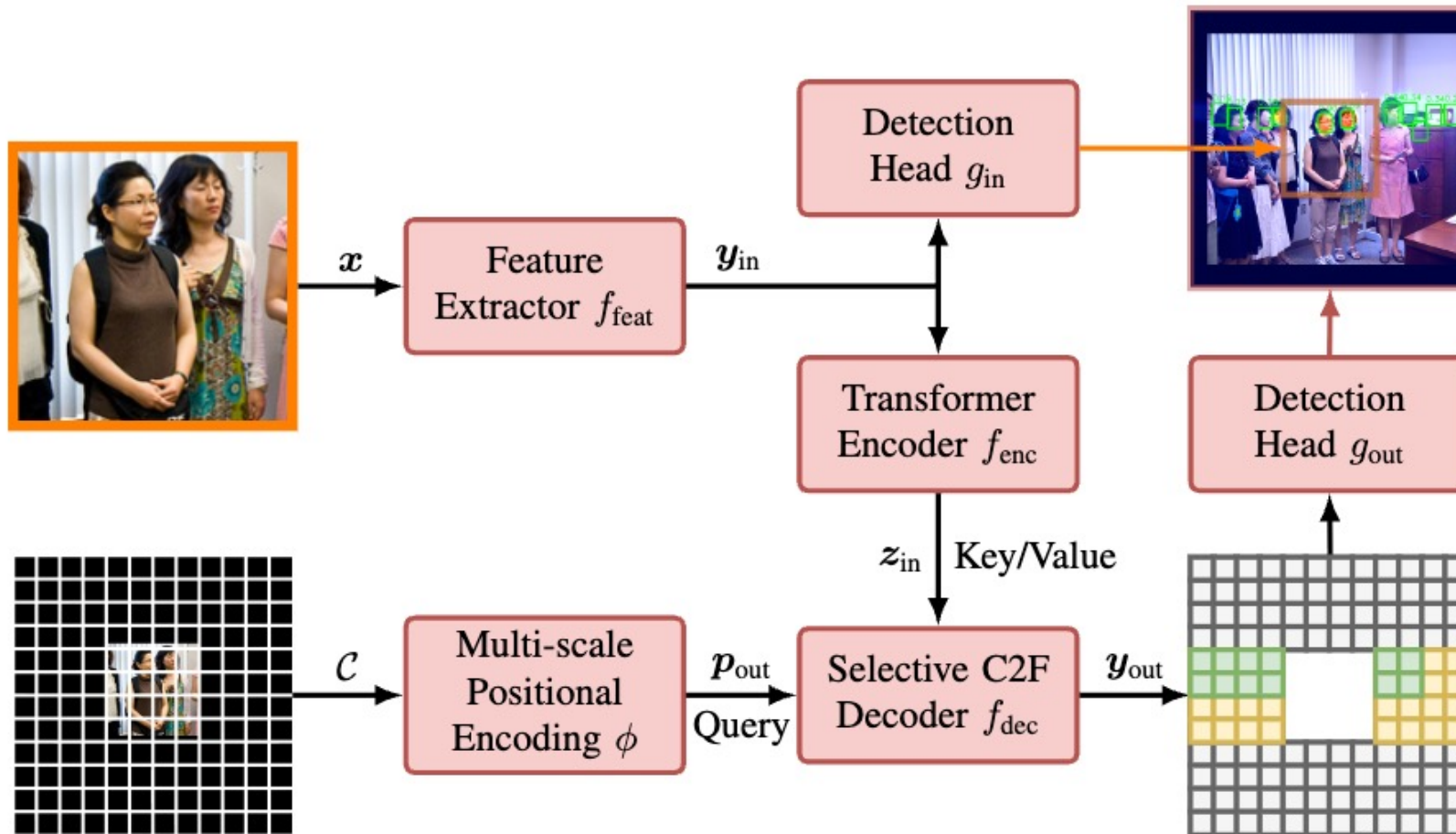
Method: Selective Coarse-to-Fine Decoder



Method: Selective Coarse-to-Fine Decoder



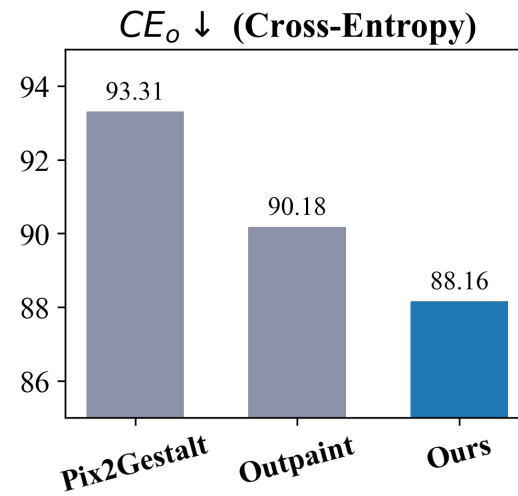
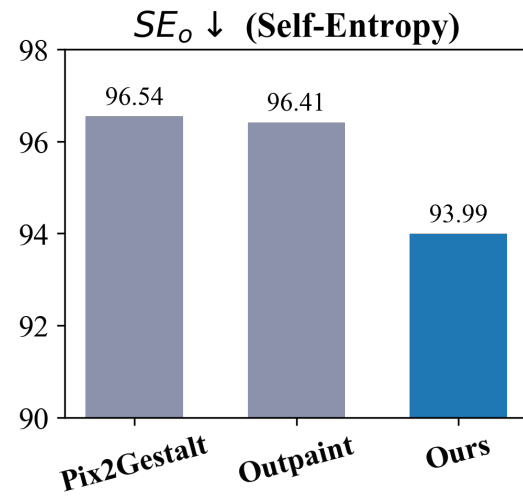
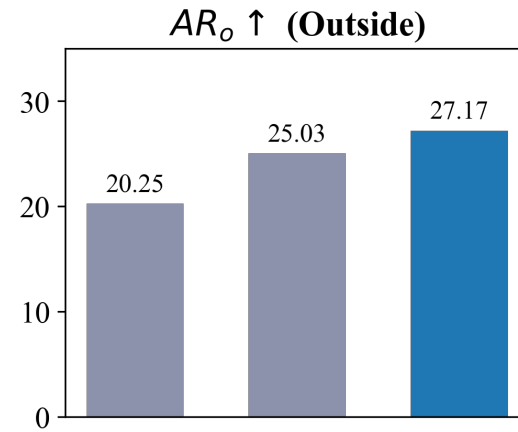
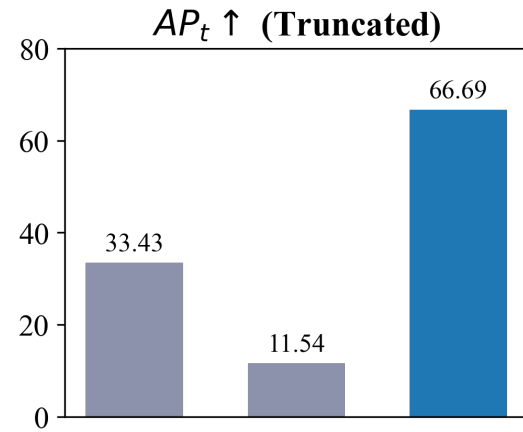
Method: Architecture



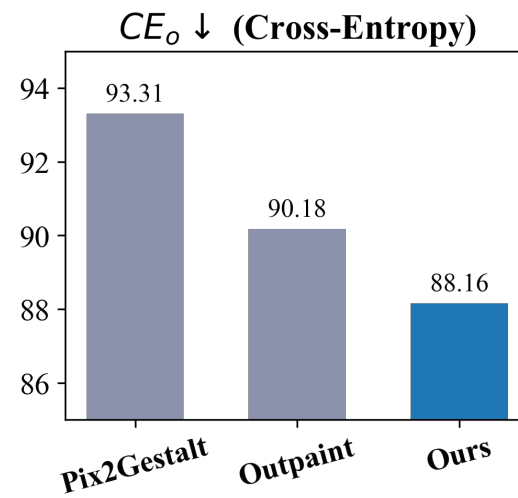
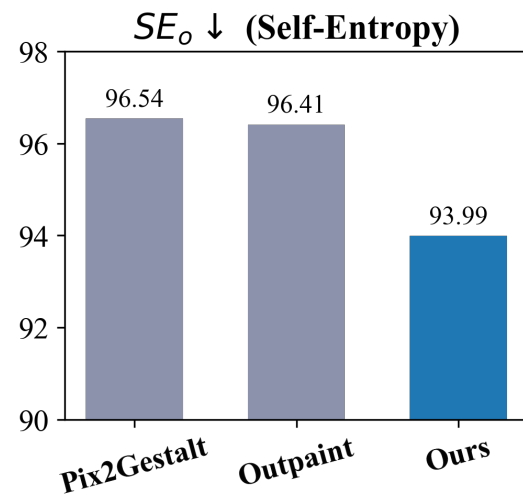
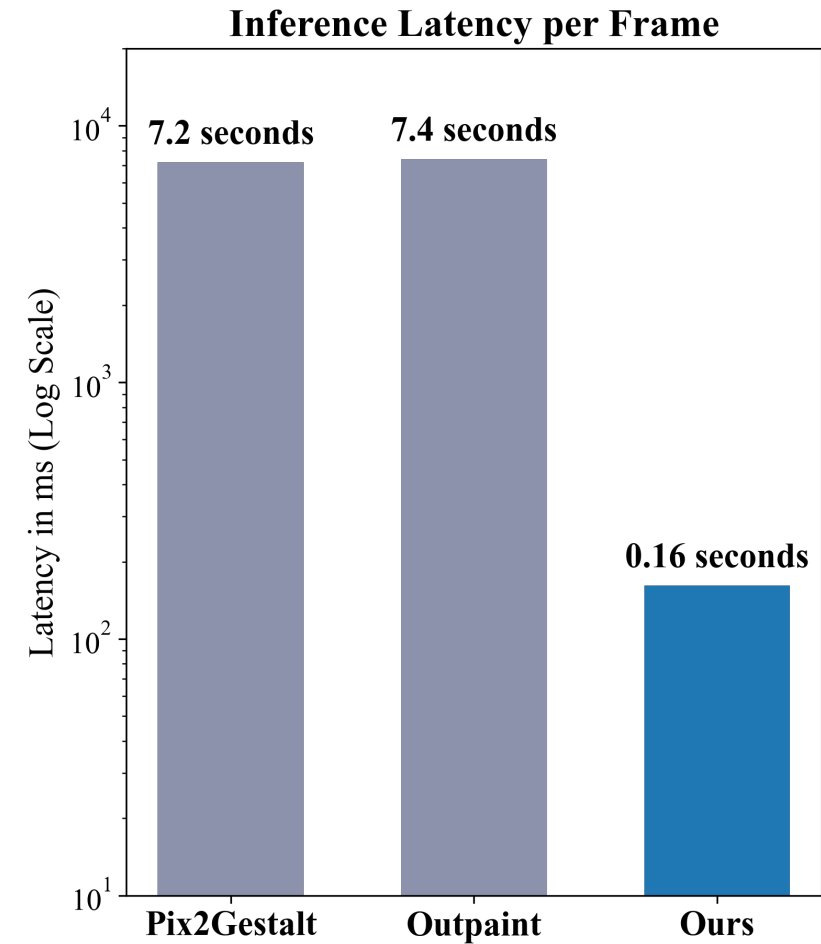
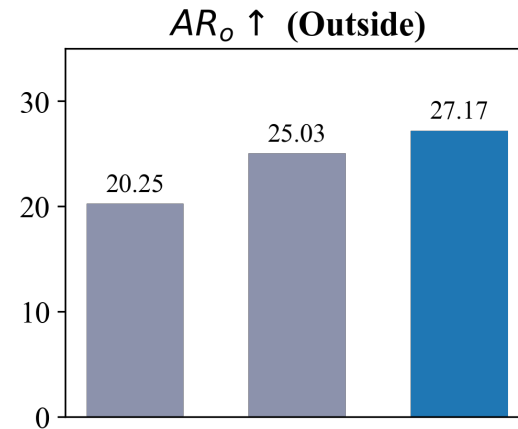
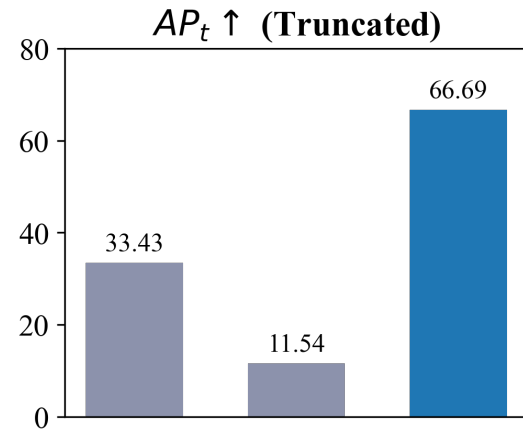
Results



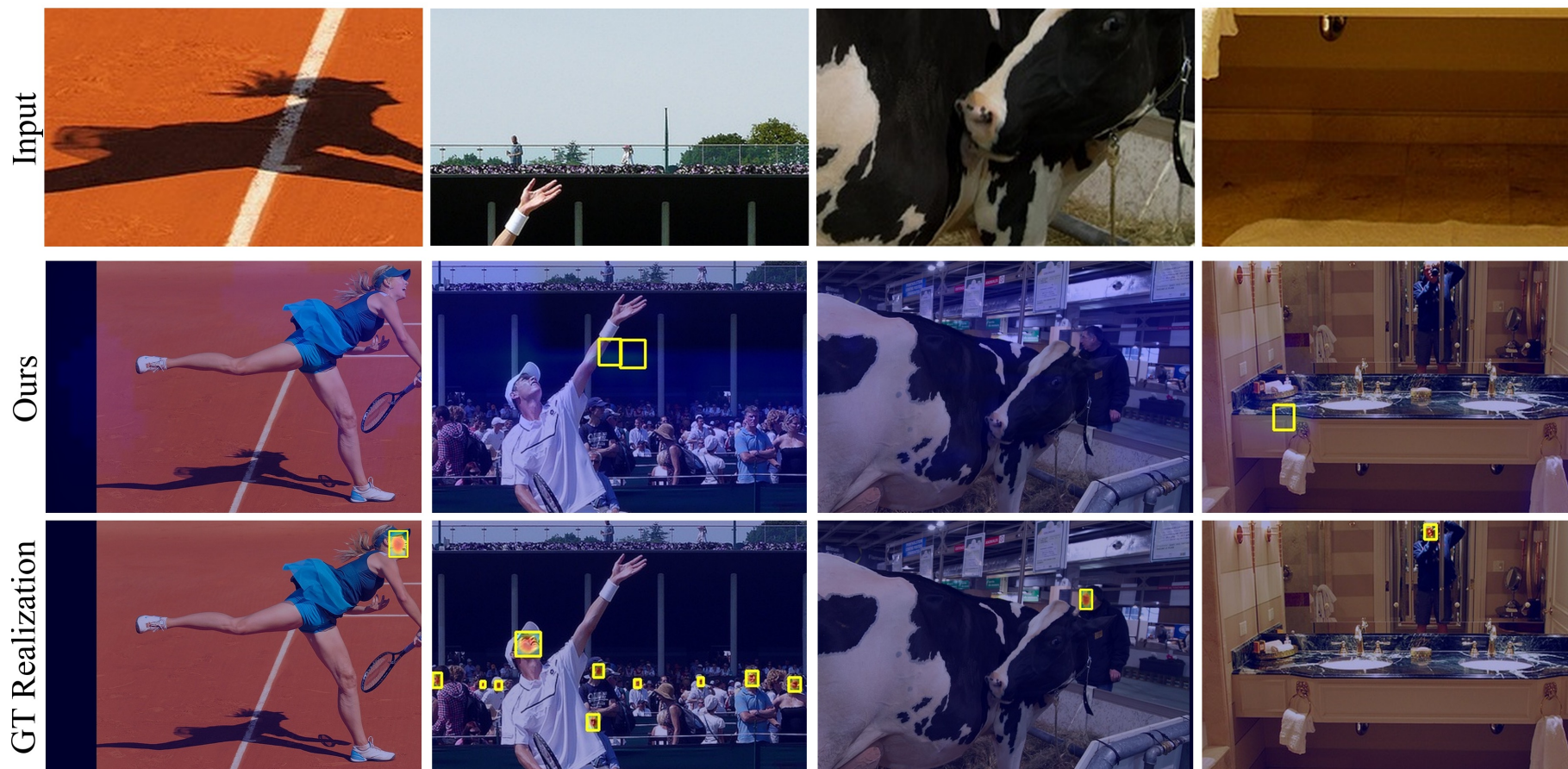
Results



Results



Failure Cases and Limitations



Limitation 1: Weak contextual cues (e.g., shadows).

Limitation 2: Predicts a single conditional distribution.

Conclusion

- **New Task:** Extreme Amodal Face Detection
- **New Dataset & Benchmark:** EXAFace
- **Efficient Solution:** Selective Coarse-to-Fine Decoder
- **State-of-the-Art:** Outperforms generative baselines while running 46x faster (0.16s/frame).

Thank you!

Welcome discussion 🙋 changlin.song@anu.edu.au



Project



Code



Data