Towards Unsupervised Ultrasound Video Clinical Quality Assessment with Multi-modality Data

He Zhao[1], Qingqing Zheng[2], Clare Teng[1], Robail Yasrab[1], Lior Drukker[3,4], Aris T. Papageorghiou[3], and J. Alison Noble[1]

[1] Institute of Biomedical Engineering, University of Oxford, UK   [2] Shenzhen Institute of Advanced Technology, Chinese Academy of Sciences, China
[3] Nuffield Department of Women's and Reproductive Health, University of Oxford, UK   [4] Department of Obstetrics and Gynecology, Tel-Aviv University, Israel

Overview

Background:
- Hard to capture high-quality scans for ultrasound scanning
- Existing methods are highly dependent on annotations

Problem:
- Clinical quality assessment without annotations and protocols
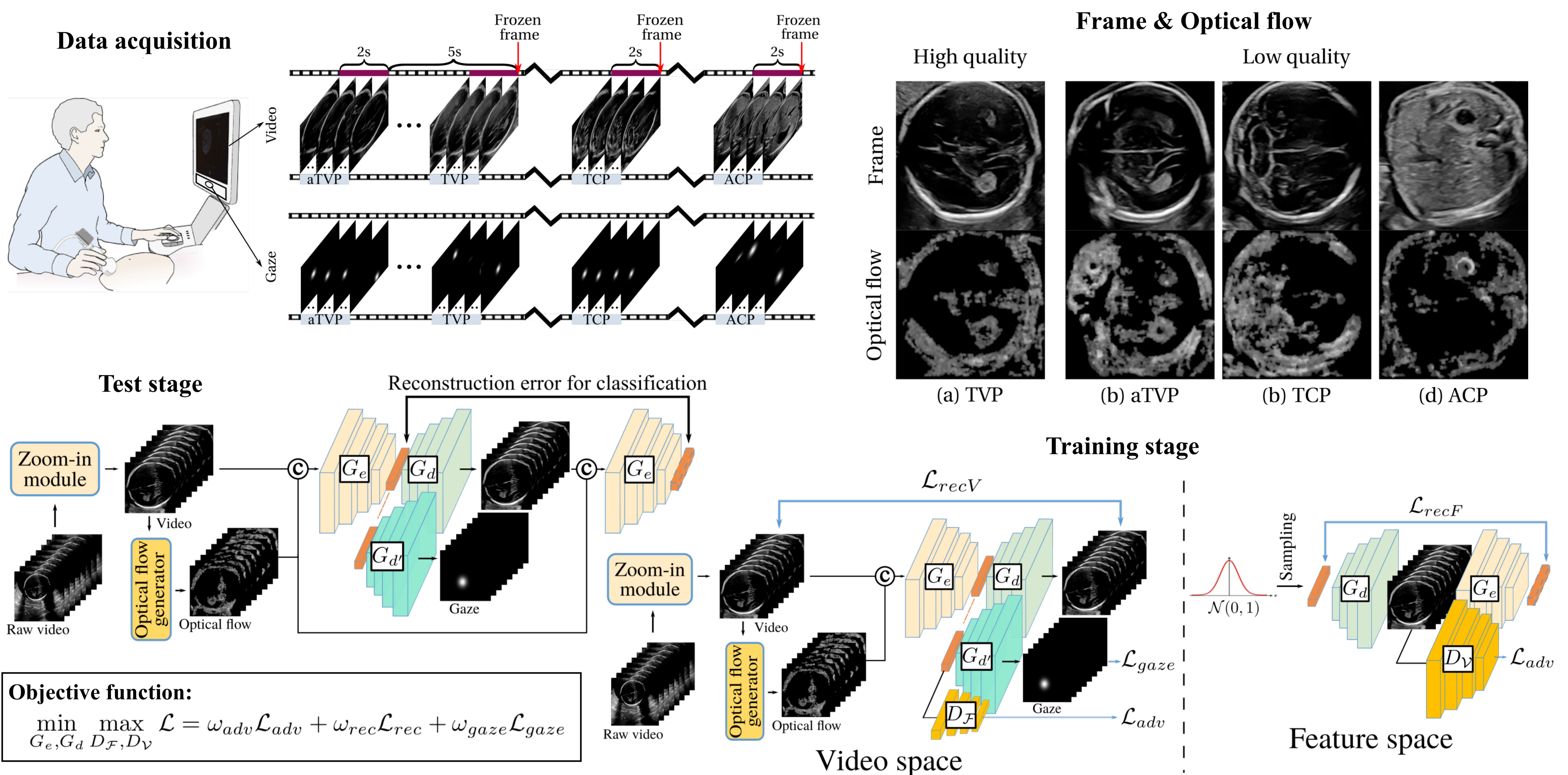
Applications:
- Providing re-scanning guidance & assisting trainee sonographers

Our Contributions

A data-driven approach for ultrasound clinical quality assessment, and its key features can be summarized as:

- Unsupervised video-based framework with only high-quality data
- Bi-directional reconstruction for high-quality representation learning
- Multi-modality for highlighting informative anatomical structures

Pipeline of Our Approach

Frame & Optical flow

(a) TVP   (b) aTVP   (b) TCP   (d) ACP

Objective function:

$$\min_{G_e, G_d} \max_{D_\mathcal{F}, D_\mathcal{V}} \mathcal{L} = \omega_{adv}\mathcal{L}_{adv} + \omega_{rec}\mathcal{L}_{rec} + \omega_{gaze}\mathcal{L}_{gaze}$$

Feature Reconstruction Error Comparison

(a) Single directional reconstruction & single-modality

(b) Bi-directional reconstruction & single-modality

(c) Bi-directional reconstruction & multi-modality
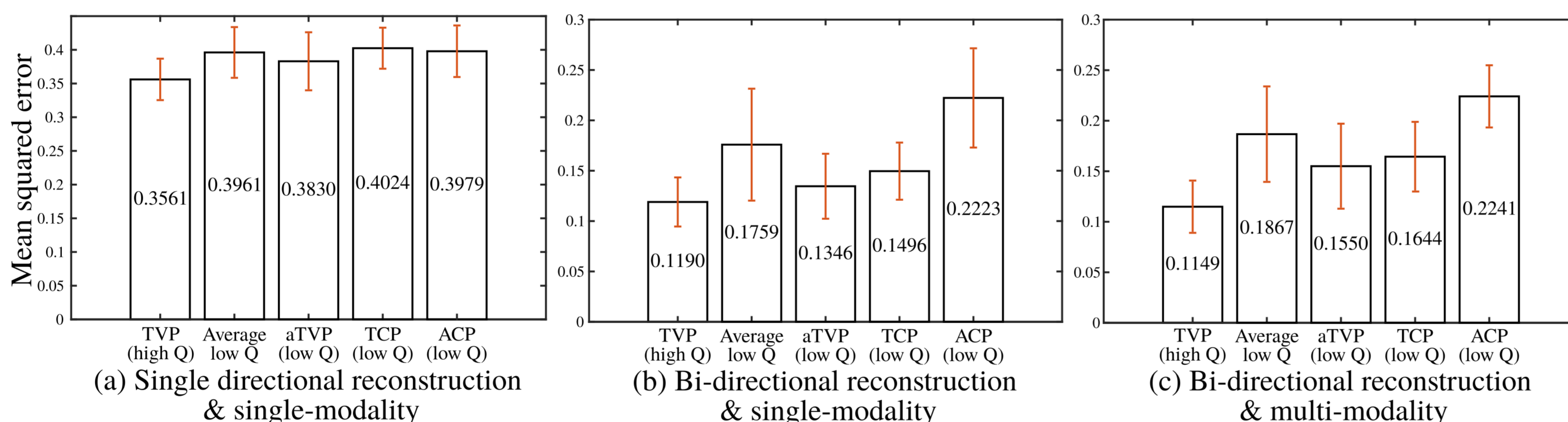
Gaze Prediction
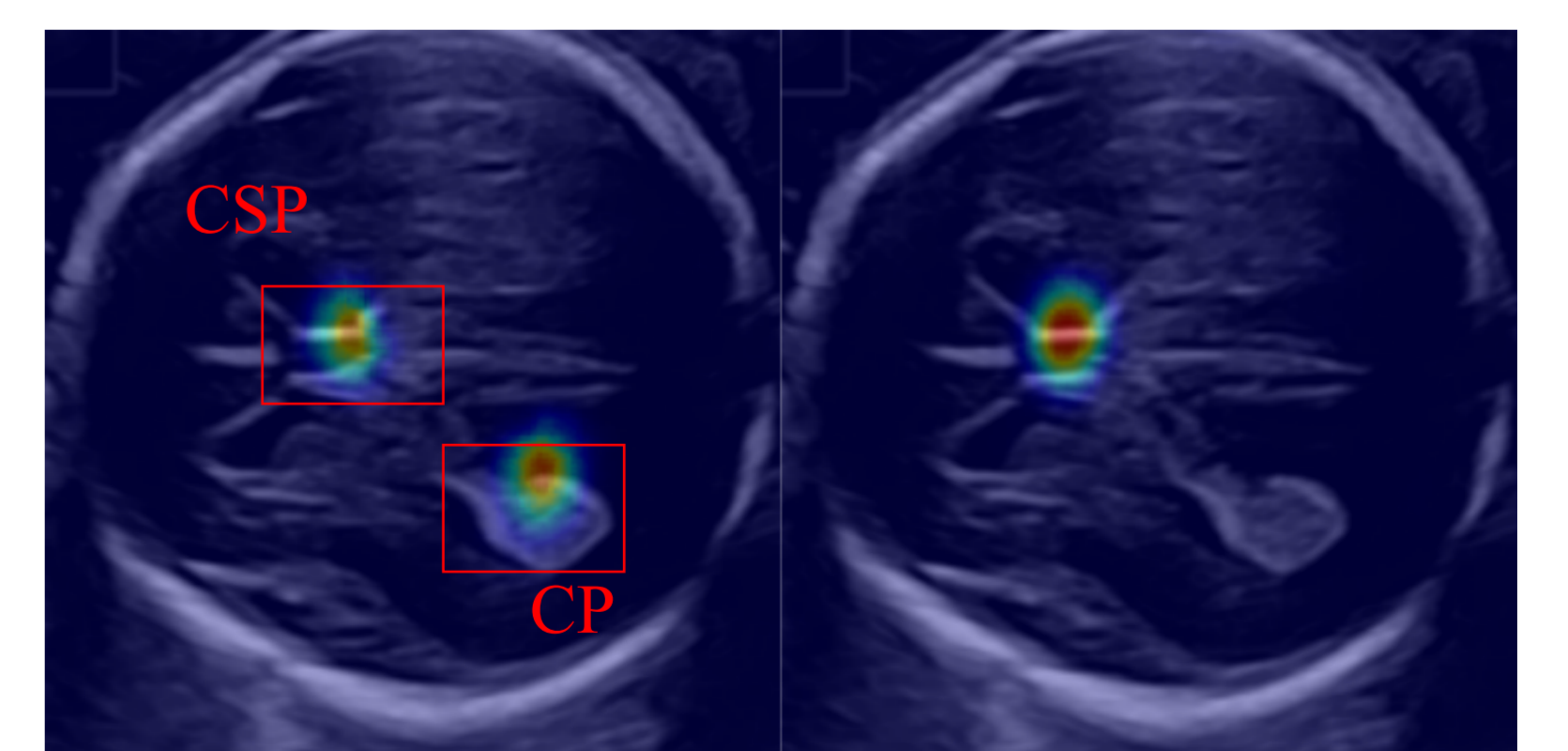
CSP and CP are essential for TVP determination during scanning

Quantitative Results

	Methods	AUC	F1	ACC	SEN	SPE
Single modality	Image-based	0.790 ± 0.006	72.29	71.06	80.11	62.05
	MNAD [1]	0.308 ± 0.009	73.32	57.88	**99.45**	1.54
	STAE [2]	0.824 ± 0.009	80.46	76.07	84.61	64.18
Multiple modalities	Our approach Video only	0.863 ± 0.005	82.66	78.78	86.90	67.47
	with Optical flow	0.889 ± 0.006	85.40	82.54	87.69	75.39
	with Gaze	0.886 ± 0.004	84.88	81.67	88.40	72.31
	All modalities	**0.911 ± 0.003**	**86.99**	**84.56**	88.62	**78.92**

[1] Park, H. et al. CVPR 2020; [2] Zhao, Y. et al. ACMM 2017

Conclusions and Outlook

- A novel unsupervised framework to assess ultrasound clinical video quality
- Directly from high-quality data without clinical protocols or annotations
- Bi-directional reconstruction & multi-modality
- Potential to be applied to different clinical quality assessment problems
- **Future work:** Giving guidance with the learned high-quality representation