Towards Real-world Video Face Restoration: A New Benchmark

Ziyan Chen*, Jingwen He*, Xinqi Lin, Yu Qiao, Chao Dong†

上海人工智能实验室 Shanghai Artificial Intelligence Laboratory

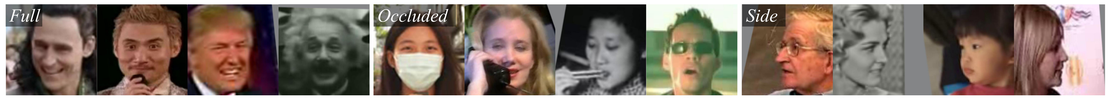
codes & more results in paper available at https://github.com/ziyannchen/VFRxBenchmark

CVPR SEATTLE, WA JUNE 17-21, 2024 Workshops (NTIRE 2024)

Motivation

Current SOTA face restoration (FR) methods achieve perfect results on (a) widely used test sets, but fails on cases from (b) our proposed FOS datasets with complex motions from videos.

Thus we propose the FOS benchmark to re-think current FR methods, including image subsets of FOS-real, FOS-syn with a toxonomy of ***full***, ***occluded*** and ***side*** faces sampled from videos.



FOS-real subset includes lq face mages sampled from real-world video clips.

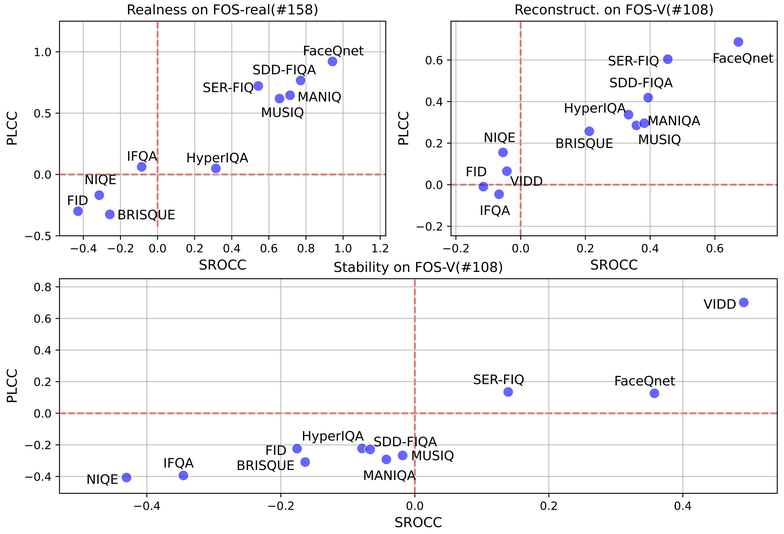
FOS-syn subset is lq-hq image pairs in which lq is synthesized from CelebA-HQ.

We further propose FOS-V to fill the gap that no video FR test set from real-world is available.

FOS-V subset includes real-word lq video clips with various face motion.

Benchmark Results

Validation in Effectiveness of popular IQA and FIQA metrics reffering to their correlation with subjective scores.



Subjective scores achieved by 4 image FR medthos on FOS-real(#158).

	F.		O.		S.	
	Real.↑	Fidel.↑	Real.↑	Fidel.↑	Real.↑	Fidel.↑
CodeFormer [57]	3.64	3.64	3.43	3.47	3.29	3.31
RestoreFormer [48]	2.72	2.90	2.57	2.83	2.36	2.63
VQFR [13]	3.42	3.32	2.92	2.89	2.91	2.85
GFP-GAN [46]	3.05	3.22	2.83	3.04	3.00	3.23

Quantitative comparison of image and video FR methods on FOS-V dataset.

	SER-FIQ↑	MANIQA↑	FID↓	VIDD↓
GPEN [54]	0.596	0.639	79.21	0.51
GCFSR [14]	0.572	0.473	98.53	0.48
GFP-GAN [46]	0.601	0.515	96.50	0.48
VQFR [13]	0.596	0.514	85.60	0.62
RestoreFormer [48]	0.556	0.524	86.82	0.50
CodeFormer [57]	0.616	0.520	98.85	0.50
BasicVSR [3]	0.567	0.300	123.26	0.36
EDVR [45]	0.567	0.302	124.20	0.38
EDVR-GAN [45]	0.562	0.499	84.84	0.44
BasicVSR-GAN [3]	0.520	0.505	83.19	0.52

Quantitative comparison of image and video FR methods on FOS-syn dataset.

	PSNR↑				SSIM↑				LPIPS↓				SER-FIQ↑				MANIQA↑				FID↓			
	F.	O.	S.	Total	F.	O.	S.	Total	F.	O.	S.	Total	F.	O.	S.	Total	F.	O.	S.	Total	F.	O.	S.	Total
PULSE [33]	22.01	20.56	21.24	21.25	0.623	0.584	0.643	0.616	0.466	0.518	0.513	0.500	0.734	0.700	0.661	0.698	0.590	0.573	0.554	0.572	88.29	67.86	70.66	75.40
PSFR-GAN [4]	24.30	23.70	25.15	24.37	0.620	0.596	0.669	0.628	0.409	0.444	0.440	0.431	0.515	0.413	0.365	0.430	0.626	0.612	0.582	0.606	68.77	58.86	63.38	63.56
HiFaceGAN [52]	24.75	23.64	24.71	24.35	0.622	0.592	0.644	0.618	0.418	0.456	0.460	0.445	0.766	0.616	0.531	0.637	0.595	0.553	0.530	0.559	71.90	62.59	72.12	68.73
GPEN [54]	24.60	23.41	24.75	24.24	0.664	0.635	0.707	0.668	0.394	0.402	0.388	0.395	0.768	0.629	0.556	0.650	0.695	0.695	0.674	0.688	69.39	63.67	66.95	66.60
GCFSR [14]	26.31	25.11	26.43	25.93	0.699	0.677	0.734	0.703	0.311	0.353	0.354	0.340	0.772	0.654	0.579	0.668	0.656	0.640	0.622	0.639	79.71	64.36	65.31	69.65
DMDNet [30]	25.67	24.48	25.79	25.30	0.681	0.656	0.719	0.685	0.355	0.398	0.389	0.381	0.769	0.637	0.555	0.653	0.622	0.586	0.558	0.588	67.61	60.91	61.91	63.41
GFP-GAN [46]	25.27	23.95	25.13	24.76	0.673	0.642	0.702	0.672	0.331	0.381	0.385	0.366	0.771	0.662	0.589	0.673	0.664	0.660	0.641	0.655	78.06	62.75	63.39	67.92
VQFR [13]	23.95	22.49	23.69	23.36	0.650	0.618	0.685	0.651	0.328	0.373	0.375	0.359	0.765	0.664	0.587	0.672	0.654	0.644	0.624	0.641	75.90	61.68	66.26	67.79
RestoreFormer [48]	24.74	23.41	24.68	24.26	0.636	0.609	0.671	0.638	0.328	0.373	0.381	0.361	0.781	0.656	0.576	0.670	0.648	0.634	0.613	0.631	72.09	56.92	56.81	61.80
CodeFormer [57]	25.29	25.17	25.17	25.21	0.661	0.692	0.692	0.682	0.317	0.358	0.358	0.345	0.783	0.680	0.603	0.688	0.666	0.666	0.645	0.659	80.90	65.02	68.78	71.40