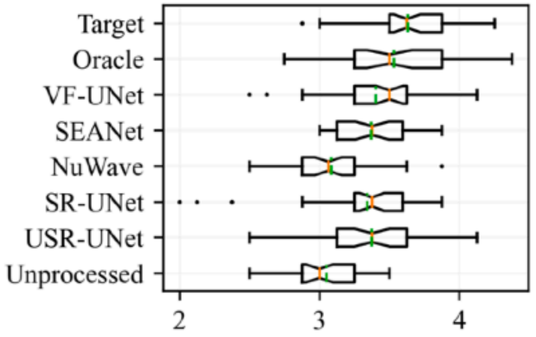
Experiments

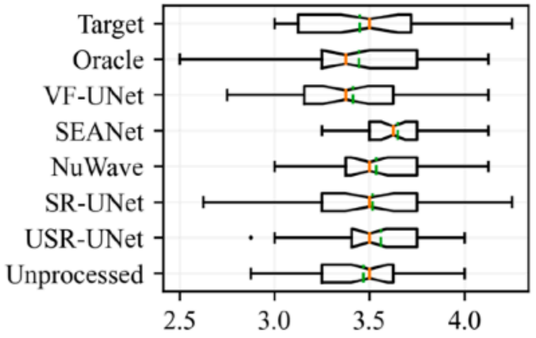
Speech Super Resolution

Table 10. LSD, SiSPNR, SSIM metrics on super resolution evaluation set SR

Training Scheme		Regression based models				VoiceFixer models				Others		
Sample Rate Up Ratio	Metrics	GSR-UNet	SR-UNet	NuWave	SEANet	VF-DNN	VF-BiGRU	VF-UNet-S	VF-UNet	Unprocessed	Oracle	Target
2kHz 22.1	LSD	1.34	1.19	1.41	1.33	1.18	1.08	1.08	**1.05**	3.13	0.89	/
	SiSPNR	11.03	10.89	9.19	9.78	10.67	11.84	11.65	**12.10**	9.18	13.65	/
	SSIM	0.75	0.77	0.73	0.72	0.75	0.77	0.78	**0.78**	0.68	0.85	/
4kHz 11.0	LSD	1.27	1.18	1.35	1.24	1.15	1.03	1.04	**1.02**	2.97	0.89	/
	SiSPNR	11.48	11.10	9.65	10.58	11.07	12.27	11.98	**12.41**	9.52	13.65	/
	SSIM	0.77	0.78	0.76	0.72	0.75	0.79	0.79	**0.79**	0.71	0.85	/
8kHz 5.5	LSD	1.21	1.11	1.24	1.20	1.06	0.99	1.01	**0.99**	2.70	0.89	/
	SiSPNR	12.07	11.82	10.73	11.11	11.94	12.68	12.34	**12.74**	9.93	13.65	/
	SSIM	0.81	**0.82**	0.80	0.74	0.78	0.81	0.81	0.81	0.76	0.85	/
	MOS	3.37	3.34	3.09	3.37	/	/	/	**3.40**	3.05	3.53	3.63
16kHz 2.8	LSD	1.10	0.99	1.18	1.16	1.01	**0.94**	0.96	0.94	2.32	0.89	/
	SiSPNR	13.02	13.01	11.54	11.90	12.37	13.14	12.70	**13.14**	10.08	13.65	/
	SSIM	0.85	**0.88**	0.81	0.75	0.82	0.82	0.82	0.82	0.83	0.85	/
24kHz 1.8	LSD	0.97	**0.91**	1.12	1.15	0.93	0.91	0.94	0.92	1.91	0.89	/
	SiSPNR	**13.96**	13.81	11.63	12.58	13.21	13.38	12.86	13.38	10.40	13.65	/
	SSIM	0.87	**0.91**	0.81	0.75	0.84	0.83	0.83	0.84	0.89	0.85	/
	MOS	3.56	3.52	3.54	**3.65**	/	/	/	3.41	3.47	3.44	3.45
Average Score	LSD	1.18	1.07	1.26	1.21	1.07	0.99	1.01	**0.98**	2.61	0.89	/
	SiSPNR	12.31	12.13	10.55	11.19	11.85	12.66	12.31	**12.75**	9.82	13.65	/
	SSIM	0.81	**0.83**	0.79	0.74	0.79	0.80	0.81	0.81	0.77	0.85	/



(b) Super Resolution (8kHz)



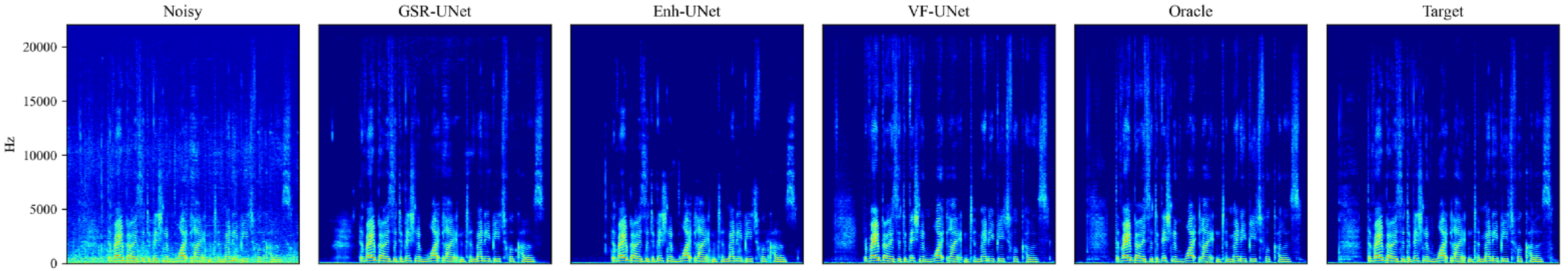
(e) Super Resolution (24kHz)

Hard to distinguish!

Experiments

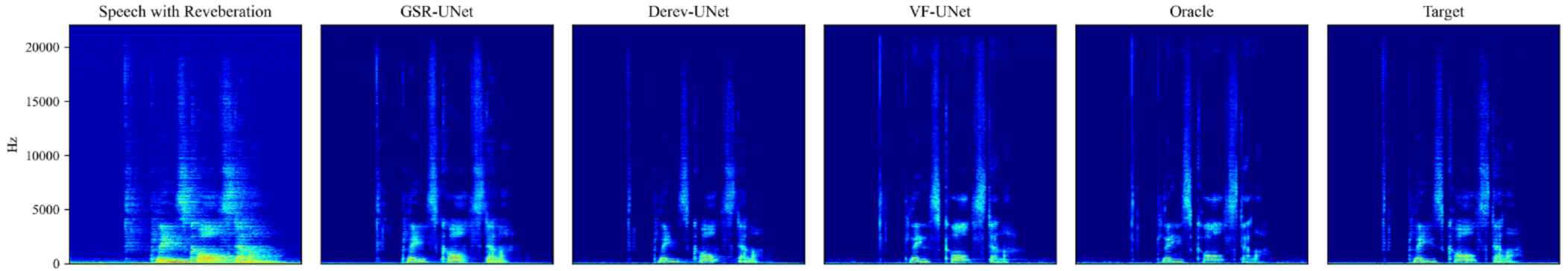
Speech Enhancement and Dereverberation

Enhancement



Comparison between different enhancement methods

Dereverberation



Comparison between different dereverberation methods