Pairwise comparisons of WER using other noise generation techniques

	Librispeech			Commonvoice		
	Deepspeech	Sphinx	Google	Deepspeech	Sphinx	Google
All vs Random	0.001	0.001	0.02	0.011	0.001	0.041
All vs Important	0.1543	0.001	0.03	0.40	0.001	0.19
Important vs Random	0.001	0.001	0.001	0.035	0.001	0.9

Table 1: P-values using One way Anova and Tukey's HSD for pairwise comparison of WER achieved by frame selection methods (using GL noise generation).

	Librispeech			Commonvoice		
	Deepspeech	Sphinx	Google	Deepspeech	Sphinx	Google
All vs Random	0.001	0.001	0.06	0.038	0.001	0.9
All vs Important	0.036	0.001	0.09	0.6	0.001	0.9
Important vs Random	0.001	0.032	0.9	0.07	0.035	0.9

Table 2: P-values using One way Anova and Tukey's HSD for pairwise comparison of WER achieved by frame selection methods (using DE noise generation).

	Librispeech			Commonvoice		
	Deepspeech	Sphinx	Google	Deepspeech	Sphinx	Google
GL vs OP	0.001	0.001	0.001	0.009	0.001	0.81
GL vs DE	0.001	0.001	0.001	0.001	0.001	0.79
OP vs DE	0.55	0.66	0.9	0.75	0.63	0.81

Table 3: P-values using One way Anova and Tukey's HSD for pairwise comparison of WER achieved by noise generation methods (using Important frames).

	Librispeech			Commonvoice		
	Deepspeech	Sphinx	Google	Deepspeech	Sphinx	Google
GL vs OP	0.001	0.001	0.007	0.009	0.001	0.9
GL vs DE	0.001	0.001	0.001	0.006	0.001	0.9
OP vs DE	0.9	0.66	0.9	0.9	0.21	0.9

Table 4: P-values using One way Anova and Tukey's HSD for pairwise comparison of `WER` achieved by noise generation methods (using `Random` frames).

Pairwise comparisons of `Similarity` using other noise generation techniques

	Librispeech	Commonvoice
Random VS All	0.001	0.001
Important VS All	0.001	0.001
Random VS Important	0.041	0.32

Table 5: P-values using One way Anova and Tukey's HSD for pairwise comparison of `Similarity` achieved by frame selection methods (using `GL` noise generation).

	Librispeech	Commonvoice
Random VS All	0.001	0.001
Important VS All	0.001	0.001
Random VS Important	0.034	0.9

Table 6: P-values using One way Anova and Tukey's HSD for pairwise comparison of `Similarity` achieved by frame selection methods (using `DE` noise generation).

	Librispeech	Commonvoice
OP VS GL	0.001	0.001
DE VS GL	0.001	0.001
DE VS OP	0.01	0.001

Table 7: P-values using One way Anova and Tukey's HSD for pairwise comparison of `Similarity` achieved by noise generation methods (using `Important` frames).

	Librispeech	Commonvoice
OP VS GL	0.001	0.001
DE VS GL	0.001	0.001
DE VS OP	0.38	0.001

Table 8: P-values using One way Anova and Tukey's HSD for pairwise comparison of `Similarity` achieved by noise generation methods (using `Random` frames).

Pareto front with `Similarity` and `WER` using other attack generation techniques and other frame selection techniques on different ASRs and datasets

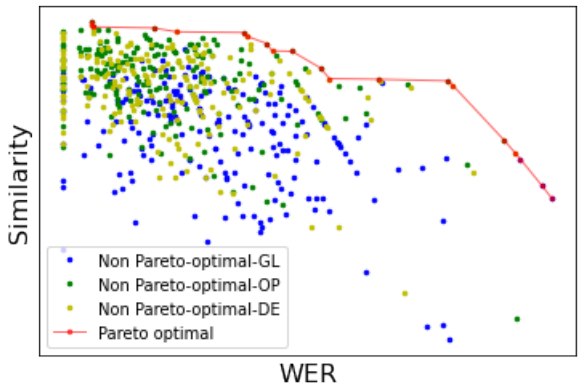

Figure 1: Pareto front in samples generated by `GL`, `OP` and `DE` on Commonvoice dataset and Deepspeech ASR using `All` frames.

In Figure 1,`GL` holds 3 points on the pareto front, `OP` holds 7 points and `DE` holds 10 points.

In Figure 2,`GL` holds 1 point on the pareto front, `OP` holds 9 points and `DE` holds 25 points.

In Figure 3,`GL` holds 4 points on the pareto front, `OP` holds 7 points and `DE` holds 8 points.

In Figure 4,`GL` holds 4 points on the pareto front, `OP` holds 6 points and `DE` holds 17 points.

In Figure 5,`GL` holds 1 points on the pareto front, `OP` holds 12 points and `DE` holds 10 points.

In Figure 6,`GL` holds 0 points on the pareto front, `OP` holds 11 points and `DE` holds 26 points.

In Figure 7,`All` holds 2 points on the pareto front, `Random` holds 5 points and `Important` holds 8 points.

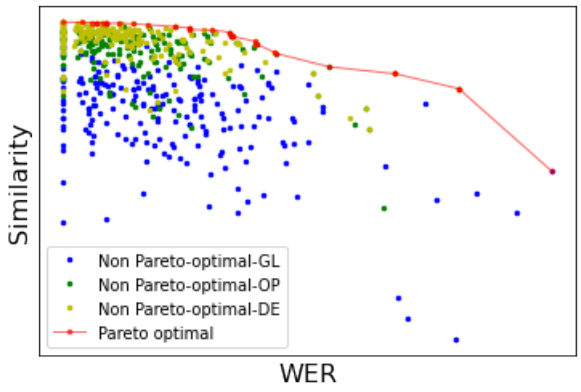

Figure 2: Pareto front in samples generated by `GL`, `OP` and `DE` on Commonvoice dataset and Deepspeech ASR using `Random` frames.

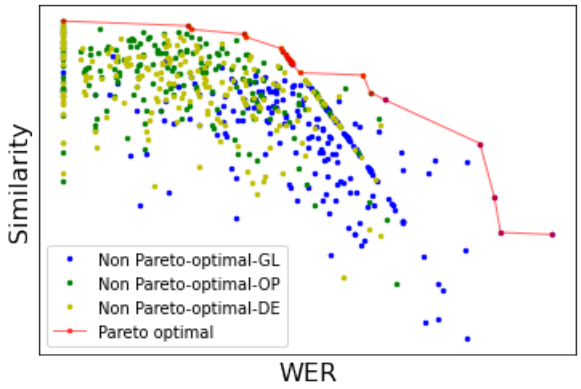

Figure 3: Pareto front in samples generated by `GL`, `OP` and `DE` on Commonvoice dataset and Sphinx ASR using `All` frames.

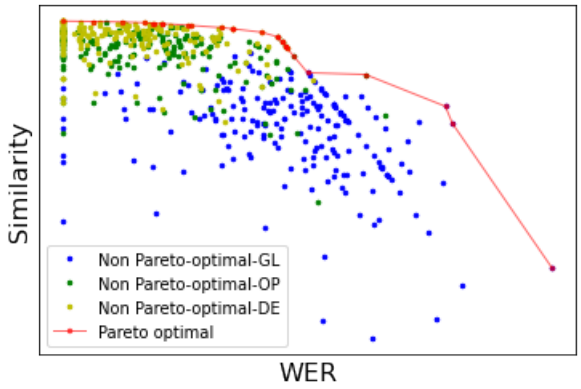

Figure 4: Pareto front in samples generated by `GL`, `OP` and `DE` on Commonvoice dataset and Sphinx ASR using `Random` frames.

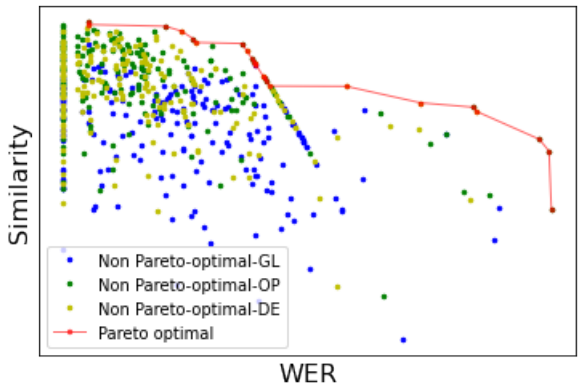

Figure 5: Pareto front in samples generated by `GL`, `OP` and `DE` on Commonvoice dataset and Google ASR using `All` frames.

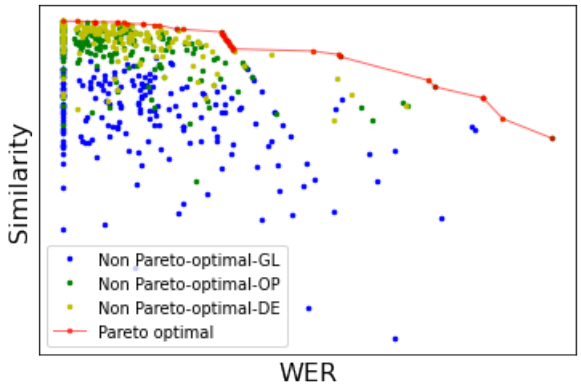

Figure 6: Pareto front in samples generated by `GL`, `OP` and `DE` on Commonvoice dataset and Google ASR using `Random` frames.

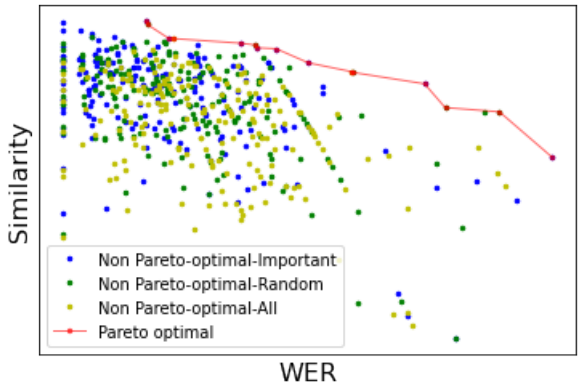

Figure 7: Pareto front in samples generated by `Random`, `Important` and `All` frame selection techniques on Commonvoice dataset and Deepspeech ASR using `GL`.

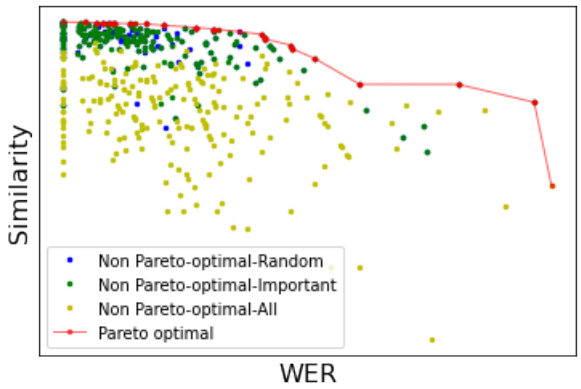

Figure 8: Pareto front in samples generated by `Random`, `Important` and `All` frame selection techniques on Commonvoice dataset and Deepspeech ASR using `DE`.

In Figure 8,`All` holds 1 points on the pareto front, `Random` holds 24 points and `Important` holds 27 points.

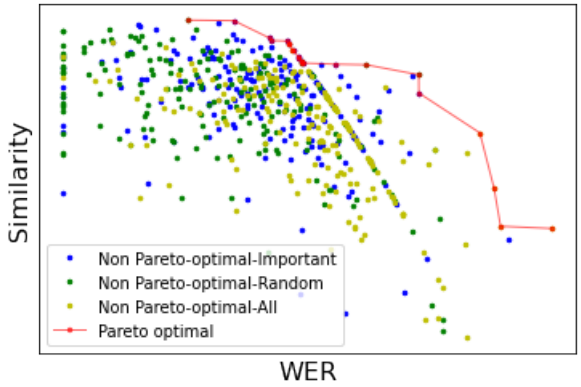

Figure 9: Pareto front in samples generated by `Random`, `Important` and `All` frame selection techniques on Commonvoice dataset and Sphinx ASR using `GL`.

In Figure 9,`All` holds 5 points on the pareto front, `Random` holds 6 points and `Important` holds 12 points.

In Figure 10,`All` holds 9 points on the pareto front, `Random` holds 15 points and `Important` holds 17 points.

In Figure 11,`All` holds 3 points on the pareto front, `Random` holds 11 points and `Important` holds 11 points.

In Figure 12,`All` holds 0 points on the pareto front, `Random` holds 25 points and `Important` holds 27 points.

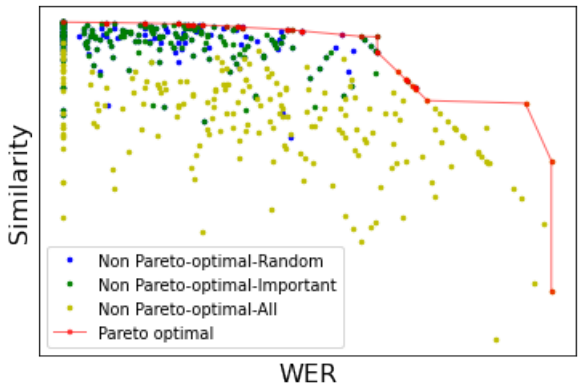

Figure 10: Pareto front in samples generated by `Random`, `Important` and `All` frame selection techniques on Commonvoice dataset and Deepspeech ASR using `DE`.

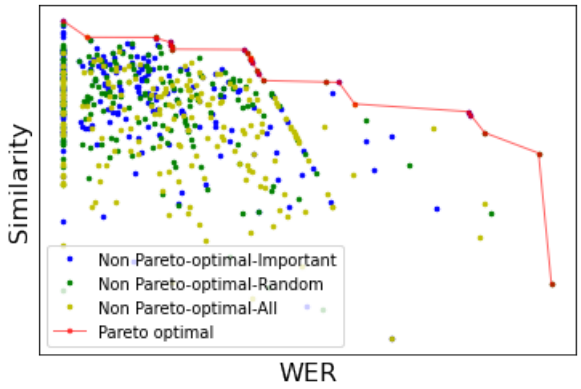

Figure 11: Pareto front in samples generated by `Random`, `Important` and `All` frame selection techniques on Commonvoice dataset and Google ASR using `GL`.

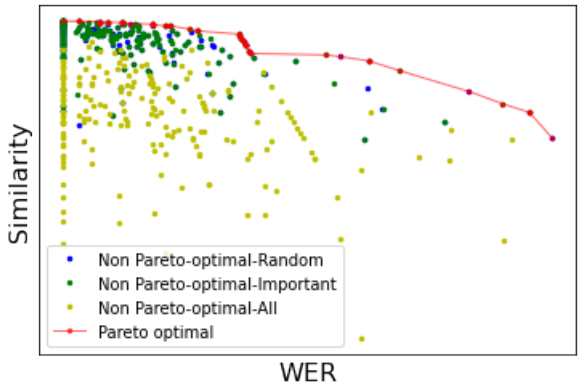

Figure 12: Pareto front in samples generated by `Random`, `Important` and `All` frame selection techniques on Commonvoice dataset and Google ASR using `DE`.