Mitigating Data Discrepancy in Replay Attack Detection

Presenter: Yongqiang Dou

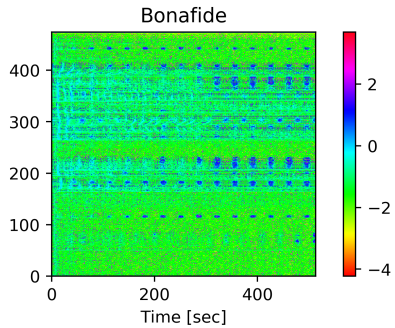
Dynamically Mitigating Data Discrepancy with Balanced Focal Loss for Replay Attack Detection

Yongqiang Dou, Haocheng Yang, Maolin Yang, Yanyan Xu, Dengfeng Ke

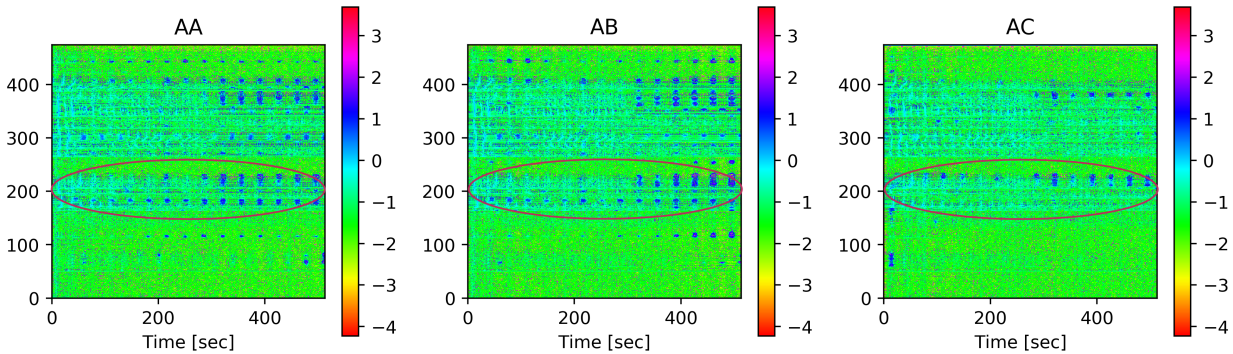
December 2020

Goal: Speaker Verification and Voice Anti-spoofing

Bonafide vs. Spoofed speech utterances feature visualization.

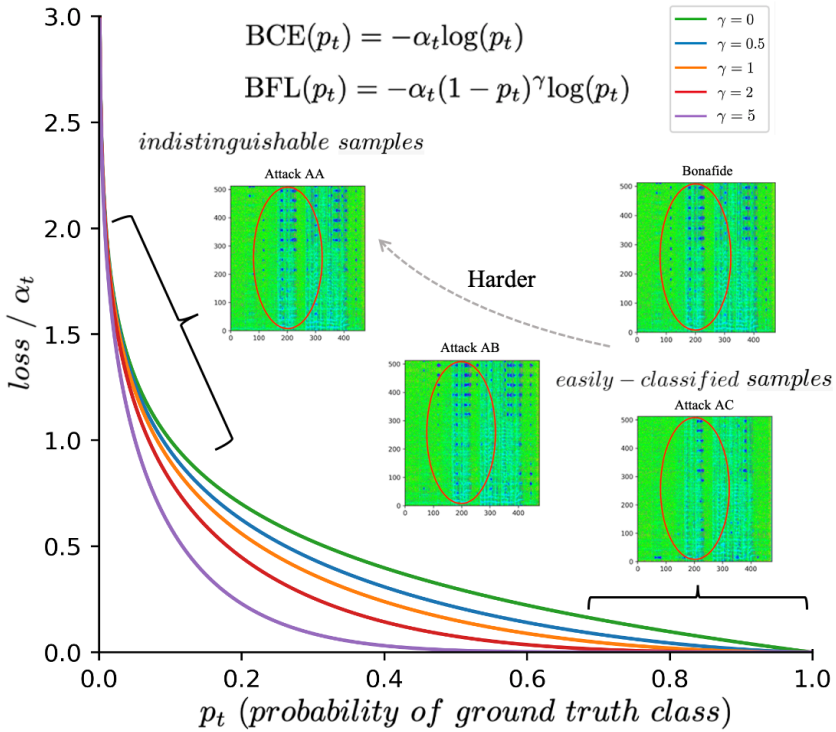


(a) Bonafide



(b) Spoofed utterances with attack type AA, AB, and AC.

Challenges & Solutions



- Data discrepancy between training and inference.
  $\Rightarrow$ Scalable objective for different samples.
- Fusion of numerous hand-designed features.
  $\Rightarrow$ Choose informative ones.
- Unexpected performance for real data.
  $\Rightarrow$ Generalization. Need more effort.

Experiments

Table: Overall Performance of top systems in ASVspoof2019 Challenge.

Method	System	# Models	PA Dev Set		PA Eval Set	
			t-DCF$^{min}_{norm}$	EER(%)	t-DCF$^{min}_{norm}$	EER(%)
Official Baseline	LFCC+GMM	-[a]	0.2554	11.96	0.3017	13.54
	CQCC+GMM	-	0.1953	9.87	0.2454	11.04
DKU(2019)	Fusion System	6	0.0064	0.24	0.0168	0.66
JHU:ASSERT(2019)	Fusion System	5	**0.0030**	**0.13**	0.0160	0.59
This work	BCE + Mean Fusion	3	0.0092	0.40	0.0153	0.62
	BCE + LR Fusion	3	0.0084	0.37	0.0151	0.61
	BFL + Mean Fusion	3	0.0075	0.35	0.0127	0.56
	BFL + LR Fusion	3	0.0077	0.35	**0.0124**	**0.55**

[a] The official baseline adopts conventional methods and therefore does not participate in the comparison of the number of neural networks used for model ensemble.

Saliency Analysis

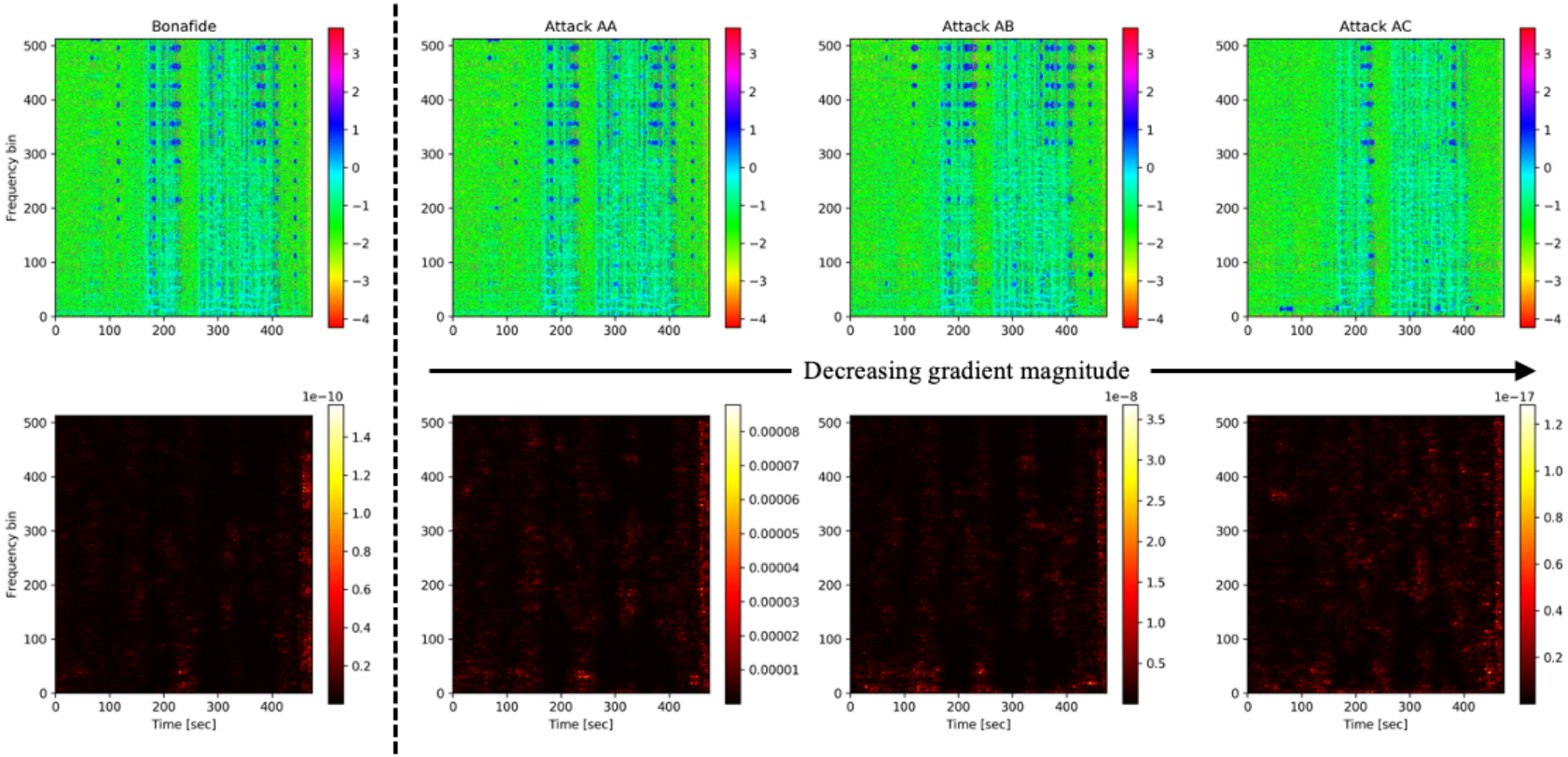


Figure: Visualization of original features (**Top**) vs. saliency maps (**Bottom**)