Adversarial training gives an illusion of **privacy-preserving** representation learning.

Representation Learning for Privacy-Preserving Speech Recognition

Brij Mohan Lal Srivastava, Aurélien Bellet, Emmanuel Vincent, Marc Tommasi

CONTEXT

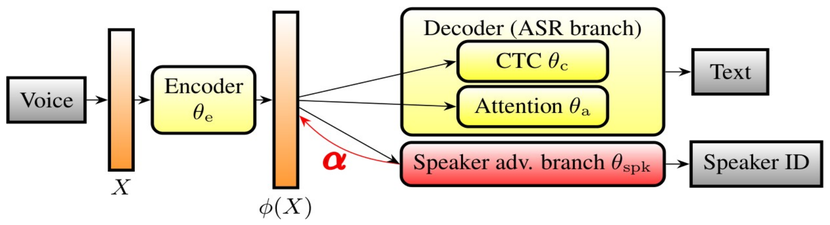
- Speech contains sensitive information, such as identity, gender, emotions, intentions, personality, etc.

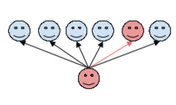
APPROACH

- Private representation shall be built on device/locally
- Sent to cloud for various processing

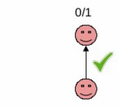
METHOD

- A combination of adversarial branch within ASR might induce speaker-invariance





Closed-set identification



Open-set verification

RESULTS

	fbank	α = 0	α = 0.5	α = 2.0
WER		10.9	12.5	12.5
ACC	93.1	46.3	6.4	2.5
EER Pooled	5.72	23.07	21.97	19.56
EER Male	3.34	19.38	18.26	16.26
EER Female	7.48	26.46	24.45	22.45

OBSERVATION

We observe that the WER of the ASR increases slightly on increasing the privacy tradeoff parameter from 0 to 2.0

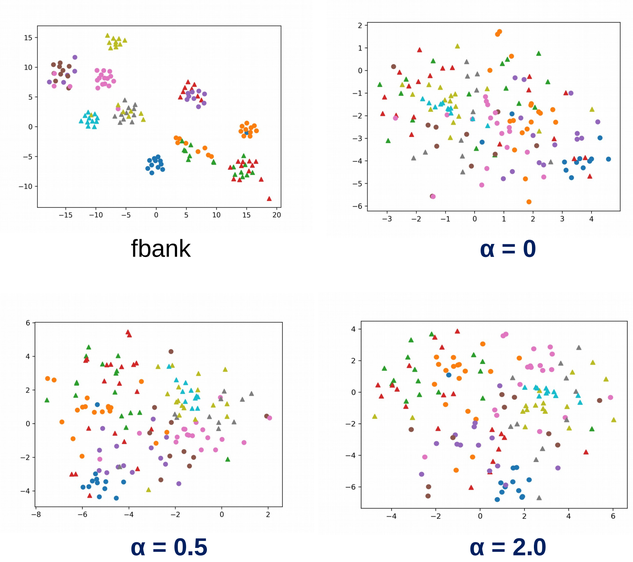
As expected, the speaker classification accuracy decreases. But counterintuitively, the EER decreases instead of increasing.

Table 1: *Splits of Librispeech used in our experiments.*

dataset	data split	# utts	duration (h)
data-full	train-960	281,231	960.98
	test-clean	2,620	5.40
	dev-clean	2,703	5.39
	test-other	2,939	5.34
	dev-other	2,864	5.12
data-adv	train-adv	27,535	97.05
	dev-adv	502	1.77
	test-adv	502	1.77
data-spkv	train-spkv	373,985	1,388.79
	train-plda	422,491	1,443.96
	test-clean-enroll	438	0.75
	test-clean-trial	21,650	51.98

Table 2: *Detailed description of the trial set (test-clean-trial) for speaker verification experiments.*

	Male	Female
# Speakers	13	16
# Genuine trials	449	548
# Impostor trials	9,457	11,196



DISCUSSION

The dramatic anonymization achieved over closed set does not match with open set verification results. Hence we conclude that adversarial training does not immediately generalize to produce anonymous representations

As future work, we plan to investigate several parameters, such as, design choices for adversarial branch, stable range of α, number of speakers, etc.

INVENTEURS DU MONDE NUMÉRIQUE

DATA SCIENCE SUMMER SCHOOL

Take a picture to **download** the **full paper**