DÉTECTION D'USURPATION D'IDENTITÉ DANS L'AUDIO

Noé Audemard

PLAN

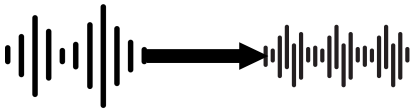
- **Introduction de l’usurpation dans l’audio et de l’axe de recherche**
- Dataset créé
- Entraînement et résultats
- Conclusion

Les méthodes d’usurpation d'identité

- Replay

Attaque physique

- Conversion



- Synthèse

Attaques logiques

Détection de spoofing pour la vérification du locuteur

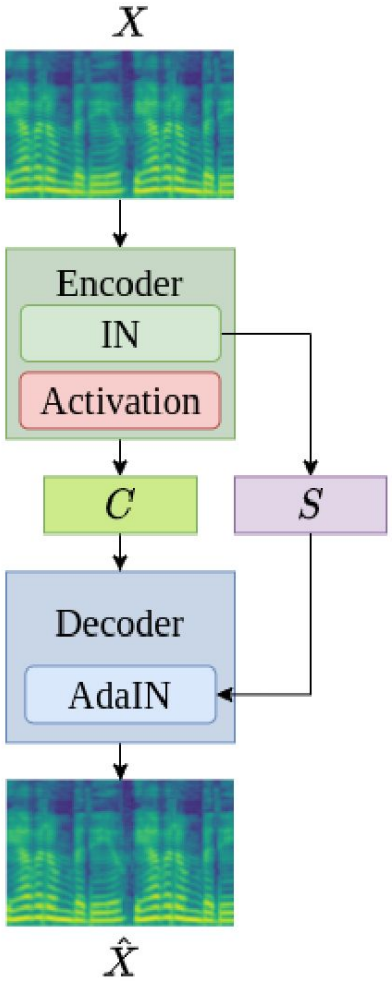
- Les systèmes de vérification du locuteur sont vulnérables aux spoofing

- Il est possible de tester si l'audio est usurpé avec un modèle dédié

- Peut on adapter les modèles de vérification pour refuser les audios usurpés?

PLAN

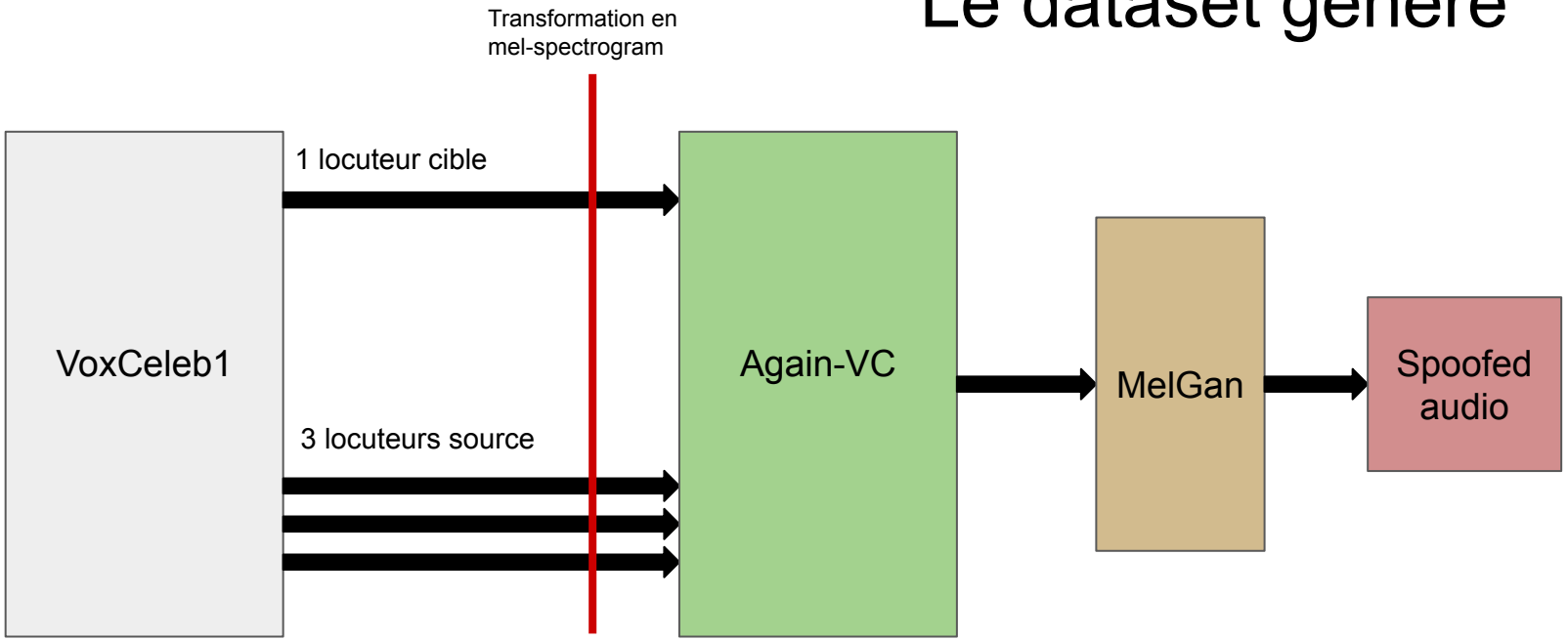
- Introduction de l’usurpation dans l’audio et de l’axe de recherche
- **Dataset créé**
- Entraînement et résultats
- Conclusion



Again-VC

Datasets d’origine: VoxCeleb1

Le dataset généré



PLAN

- Introduction de l’usurpation dans l’audio et de l’axe de recherche
- Dataset créé
- **Entraînement et résultats**
- Conclusion

ECAPA-TDNN modifié

- ECAPA-TDNN d’origine
- ECAPA-TDNN avec modification naïve
- ECAPA-TDNN avec modification informé

Entraînement et résultats (ECAPA naïf)

Voxceleb1: 148 642 utterances

Spoofed audios: 3753 audios * 10

	Genuine audios	Spoofed audios
Trained on genuine audios	3.39	26.4
Trained on both	4.81	20.5

Results in EER%

CYCLEGAN-VC2:
IMPROVED CYCLEGAN-BASED NON-PARALLEL VOICE CONVERSION

Takuhiro Kaneko, Hirokazu Kameoka, Kou Tanaka, Nobukatsu Hojo

Entraînement et résultats (ECAPA naïf)

	Genuine audios	AGAIN-VC audios	CycleGAN audios
Trained on genuine audios	3.32	26.4	35.1
Trained on both (Genuine and Again-VC)	4.89	20.5	33.2

Results in EER%

Résultats de RawGatST sur ASVSpoof 2019: Unseen Logical Attacks

1.06% EER

ECAPA-TDNN modifié

- ECAPA-TDNN d’origine
- ECAPA-TDNN avec modification naïve
- ECAPA-TDNN avec modification informé

Entraînement et résultats (ECAPA “informé”)

	Genuine audios	AGAIN-VC audios	CycleGAN audios
Trained on genuine audios	3.32	25.9	34.4
Trained on both (Genuine and Again-VC)	4.89	21.2	33.0

Results in EER%

Conclusion

- L’entrainement d’un système de reconnaissance du locuteur sur des audios usurpés permettent d’améliorer ses performances en détection d’usurpation mais dégradent ses performances sur la tâche d’origine.

- Cet amélioration semble généralisable mais les résultats ne sont pas conclusif

- La combinaison d’un modèle de vérification et d’un modèle de détection d’usurpation dédié obtient de biens meilleures performancecs

Bibliography

[1] Arsha Nagrani Joon Son Chung Andrew Zisserman. "VoxCeleb: a large-scale speaker identification dataset". In: (2017).

[2] Takuhiro Kaneko Hirokazu Kameoka Kou Tanaka Nobukatsu Hojo. "CYCLEGAN-VC2: IMPROVED CYCLEGAN-BASED NON-PARALLEL VOICE CONVERSION". In: (2019).

[3] Brecht Desplanques Jenthe Thienpondt Kris Demuynck. "ECAPA-TDNN: Emphasized Channel Attention, Propagation and Aggregation in TDNN Based Speaker Verification". In: (2020).

[4] Alexei Baevski Henry Zhou Abdelrahman Mohamed Michael Auli. "wav2vec 2.0: A Framework for Self-Supervised Learning of Speech Representations". In: (2021).

[5] Xuechen Liu Xin Wang Md Sahidullah Jose Patino Hector Delgado Tomi Kinnunen Massimiliano Todisco Junichi Yamagishi Nicholas Evans Andreas Nautsch Kong Aik Lee. "ASVspoof 2021: Towards Spoofed and Deepfake Speech Detection in the Wild". In: (2021).

[6] Yen-Hao Chen Da-Yi Wu Tsung-Han Wu Hung-yi Lee. "AGAIN-VC: A ONE-SHOT VOICE CONVERSION USING ACTIVATION GUIDANCE AND ADAPTIVE INSTANCE NORMALIZATION". In: (2021).

[7] Hemlata Tak Jee-weon Jung Jose Patino Madhu Kamble Massimiliano Todisco and Nicholas Evans. "End-to-End Spectro-Temporal Graph Attention Networks for Speaker Verification Anti-Spoofing and Speech Deepfake Detection". In: (2021).

[8] Hemlata Tak Madhu Kamble Jose Patino Massimiliano Todisco and Nicholas Evans. "RAWBOOST: A RAW DATA BOOSTING AND AUGMENTATION METHOD APPLIED TO AUTOMATIC SPEAKER VERIFICATION ANTI-SPOOFING". In: (2021).

[9] Hsin-Te Hwangh Yu Tsaoh Hsin-Min Wangh Sebastien Le Magueri Markus Beckerj Fergus Hendersonj Rob Clarkj Yu Zhangj Quan Wangj Ye Jiaj Kai Onumak Koji Mushikak Takashi Kanedak Yuan Jiangl Li-Juan Liul Yi-Chiao Wum Wen-Chin Huangm Tomoki Todam Kou Tanakan Hirokazu Kameokan Ingmar Steinero Driss Matroufp Jean-Francois Bonastrep Avashna Govenderb Srikanth Ronankiq Jing-Xuan Zhangr Zhen-Hua Ling Xin Wanga Junichi Yamagishia Massimiliano Todiscoc Hector Delgadoc Andreas Nautschc Nicholas Evansc Md Sahidullahd Ville Vestmane Tomi Kinnunene Kong Aik Leef Lauri Juvelag Paavo Alkug Yu-Huai Pengh. "ASVspoof 2019: A large-scale public database of synthesized, converted and replayed". In: (2021).

[10] Hemlata Tak Massimiliano Todisco Xin Wang Jee-weon Jung Junichi Yamagishi and Nicholas Evans. "Automatic speaker verification spoofing and deepfake detection using wav2vec 2.0 and data augmentation". In: (2021).