2024 SW중심대학 디지털 경진대회

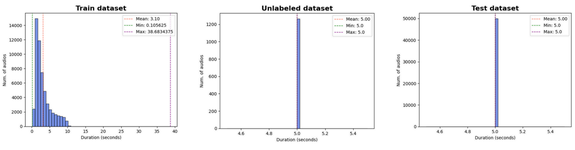
AI부문

가천대학교 Gachon University

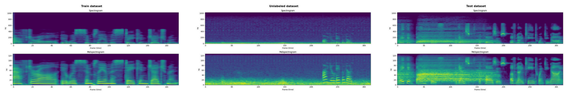
MOTA - 유종문, 김의진, 윤세현, 장희진, 최상현

데이터 분석

학습 및 추론에 사용된 데이터셋은 크게 세 가지로 분류할 수 있다: train, unlabeled, test 데이터셋이다. 이 중 학습에 사용된 것은 train과 unlabeled 데이터셋이다. 각 데이터셋에 포함된 음성의 길이를 확인한 결과, train 데이터셋의 음성은 평균 3.1초이며, 가장 짧은 것은 0.1초, 가장 긴 것은 38초이다. 반면, unlabeled와 test 데이터셋은 모두 5초의 고정된 길이로 구성되어 있다.

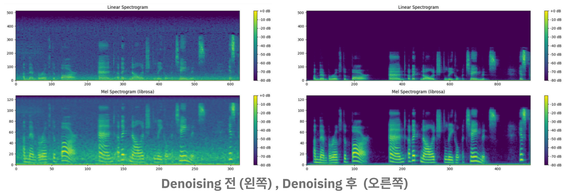


추가로, 음성 데이터셋에 포함된 노이즈도 확인해보았다. 그 결과 train 데이터셋에는 노이즈가 거의 포함되어 있지 않았다. 반면, unlabeled 데이터셋에는 균일한 white 노이즈가 많이 포함되어 있었다. Test 데이터셋의 경우, 매우 복잡하고 소리가 큰 노이즈가 다량 섞여 있었다.



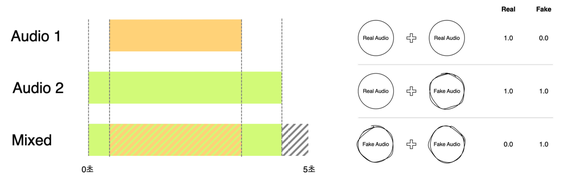
데이터 전처리

우리 팀에서 데이터 전처리 및 증강에 사용한 기법은 총 6가지이다. 먼저, 32,000Hz의 sampling rate를 16,000Hz로 resampling하여 오디오 품질을 해치지 않는 범위 내에서 학습 속도를 개선하였다. 다음으로, 노이즈 제거를 위해 DeepFilterNet v3 모델을 사용하였다.

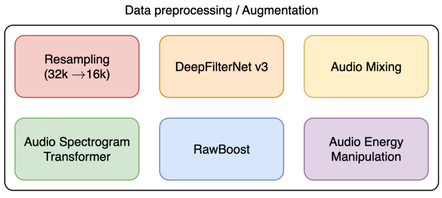


Denoising 전 (왼쪽) , Denoising 후 (오른쪽)

세 번째로, Audio mixing은 임의의 두 음성을 선택하여 겹치는 방식으로, 최대 두 명의 화자로 구성된 test 데이터와 유사한 환경을 만들어 주었다. 여기에 한 쪽 음성의 볼륨을 0 ~ 30 dB까지 낮추는 energy manipulation도 적용하였다.

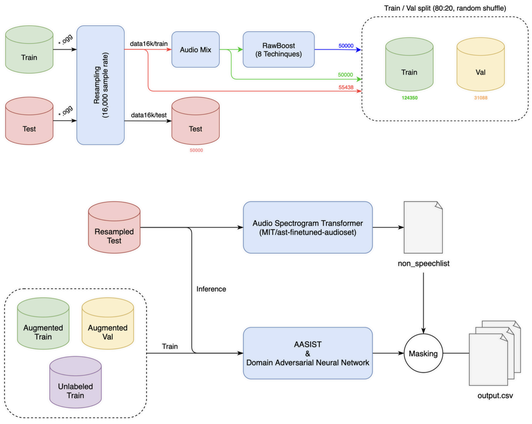


Audio Spectrogram Transformer (AST)는 주어진 음성에 포함된 소리를 logit 값으로 나타낸다. 우리 MOTA 팀은 AST를 활용하여 목소리가 존재하지 않는 데이터를 식별하고, 이를 추론 후 마스킹하는 데 사용하였다. Rawboost는 정교한 노이즈 생성을 위한 알고리즘으로, test 데이터셋에 있는 복잡한 노이즈에 대해 robust한 모델 학습이 가능하도록 하였다.

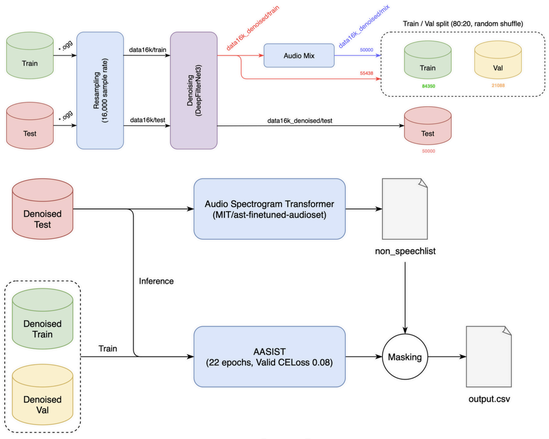


모델 파이프라인

우리가 사용한 모델은 AASIST로, 본 대회와 유사한 ASVSpoof 2019 데이터셋을 사용하여 SOTA를 달성한 모델이다. 첫 번째 파이프라인은 AASIST와 domain adaptation (e.g., DANN)을 결합한 모델로, 데이터 분포가 다른 test 데이터셋에서도 효과적인 추론이 가능하도록 설계하였다.

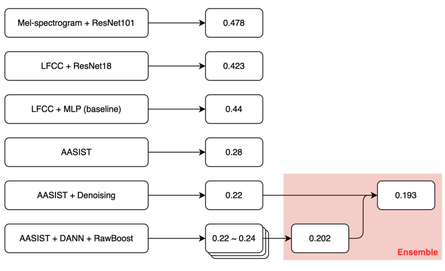


두번째 파이프라인은 AASIST에 denoise 된 데이터를 입력으로 학습하였다. 학습은 22 에폭 진행되었고, CE Loss는 검증과정에서 0.08 이었다.



모델 검증

여러 모델에 대해 실험을 했지만, 위에서 소개한 두가지 방법이 가장 우수한 성능을 보여주어 앙상블을 진행했다. 최종적으로 Dacon 평가산식에서 **0.193**의 점수를 받을 수 있었다.



적용 가능성

최근 Speech Synthesis (Text-To-Speech, TTS) 분야는 실제 음성과 구분할 수 없을 정도로 발전했으며, zero-shot TTS 기법은 숨소리와 감정까지 추론하는 데 큰 성과를 이루었다. 본 모델을 적용하면 생성된 음성과 악의적으로 노이즈가 추가된 음성을 빠르게 판별할 수 있을 것으로 기대된다. 예를 들어, 전화 통화 중 Spoofing attack이 발생하면, 통화 종료 후 상대방의 음성이 가짜인지 신속히 판별하여 사용자에게 주의를 줄 수 있을 것이다.

유종문 | 팀장 | 01062340578 | 0917jong@gachon.ac.kr
김의진 | 팀원 | 01052672383 | whitekuj02@gachon.ac.kr
윤세현 | 팀원 | 01089464638 | wjwjwj2002@gachon.ac.kr
장희진 | 팀원 | 01027657463 | wlsgml628@gachon.ac.kr
최상현 | 팀원 | 01023439034 | shchoi00@gachon.ac.kr

주최 | 과학기술정보통신부 주관 | IITP 정보통신기획평가원 SW 중심대학협의회