

2023

군장병 부문

국방 AI 경진대회

결과물 발표평가



팀명 | MEDIT



대한민국 국방부
Ministry of National Defense



과학기술정보통신부
Ministry of Science and ICT



MEDIT (Medic + IT)



김민영

의무병
5군단 27여단
분자생명공학과



권재영

의무병
육군훈련소 28교육연대
화학교육과(컴퓨터공학)



배기민

군사과학기술병(AI)
육군미래혁신연구센터
컴퓨터소프트웨어학



이학진

소프트웨어개발병(AI)
육군지능정보기술단
컴퓨터소프트웨어학

오디오 분야 분석 모델 개발

음성인식 기술을 활용한 효과적인 지휘관의 전술 명령 전달

(Speech-to-Text: Automatic Speech Recognition)

01

02

03

04

Audio Analysis

Basic Analysis
Spectrogram Analysis

Pre-processing

Noise Reduction
Audio Restructure

Pre-Training Fine-Tuning

Whisper, wav2vec,
DeepSpeech2, etc

Submission

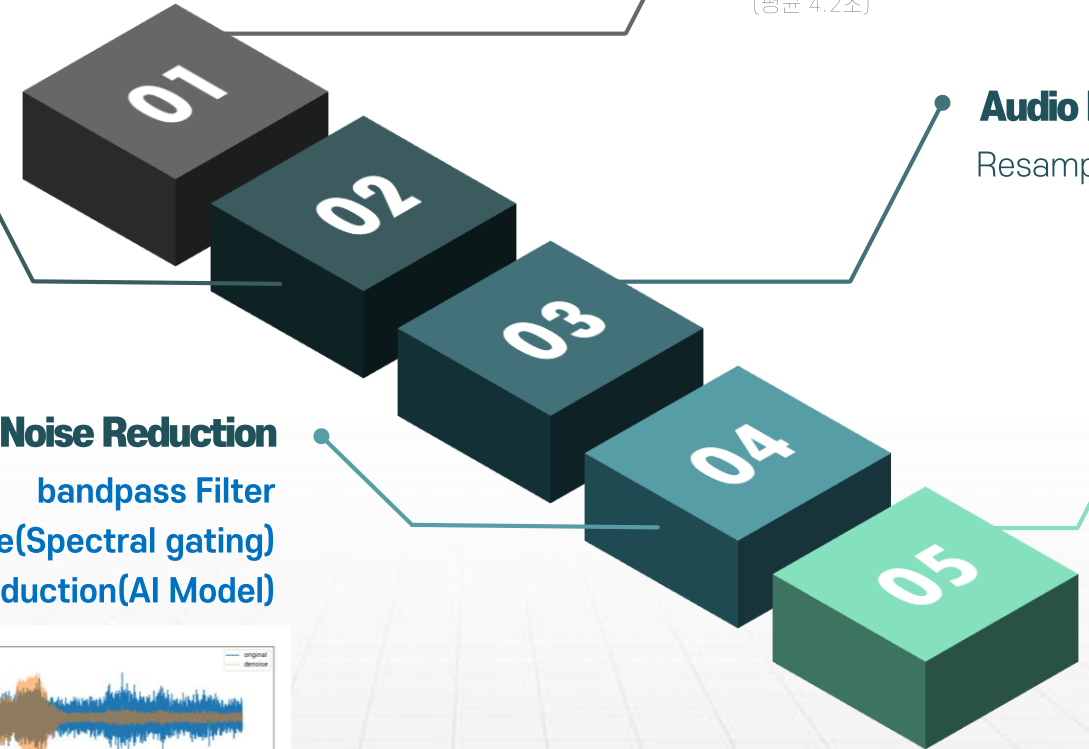
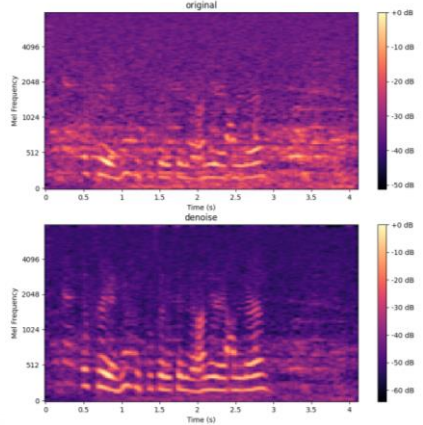
Whisper-base
Whisper-small-ko
Whisper-large-v3-lora-8bit

II 발표 내용_ [평가 항목] 데이터 활용



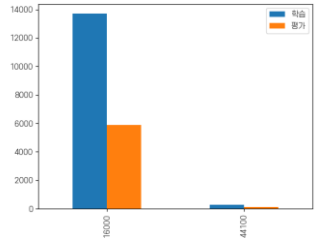
Spectrogram Analysis

Mel-Spectrogram, MFCC



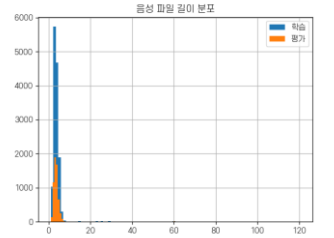
Basic Analysis

(16000, 44100)
샘플링레이트, 채널수, 비트레이트,
오디오 길이, dBFS 등 오디오 특성 확인
 (평균 4.2초)



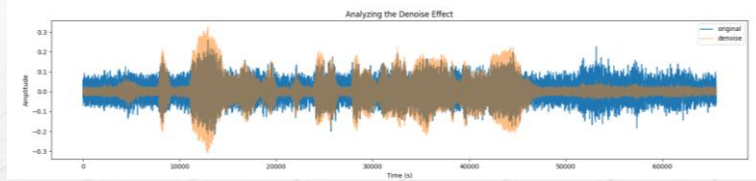
Audio Restructure

Resampling, Audio Cut



Noise Reduction

bandpass Filter
 *noisereducer(Spectral gating)
 **noise-reduction(AI Model)



*원본 음성이 작은 경우 노이즈 제거로 인해 안좋은 상황도 있었음

GOAL

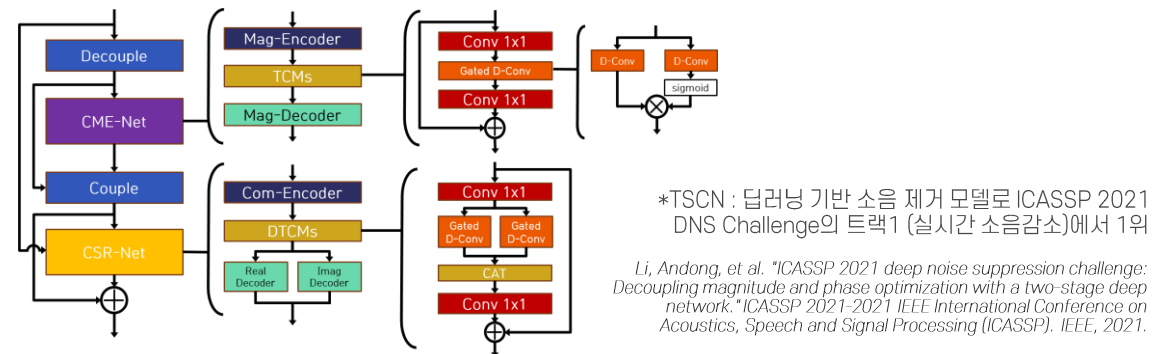
1. 오디오 샘플링레이트 통일
2. 30초 이상 길이 오디오 분할 (test)
3. 오디오 노이즈 억제
4. End(음성)-to-End(텍스트) 모델 사용
5. 전사 규칙 기반 예측 값 변경
6. 중복 문자열 제거
7. 예측 결과 양상블 (투표 기반)

* <https://github.com/timsainb/noisereducer>
 ** <https://github.com/kotechnia/noise-reduction>

II 발표 내용_ [평가 항목] 모델 완성도

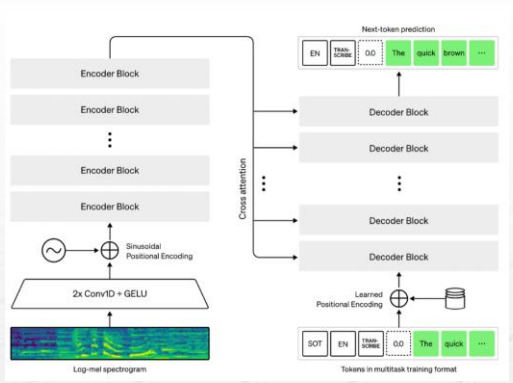


TSCN (CME-Net + CSR-Net)*



CME-Net은 스펙트럼의 크기를 최적화하고, CSR-Net은 위상을 수정 및 크기를 세분화

Whisper-Base & Whisper-Large-V3



Size	Parameters
tiny	39 M
base	74 M
small	244 M
medium	769 M
large	1550 M
large-v2	1550 M
large-v3	1550 M

웹에서 수집한 68만 시간의 다국어 및 멀티태스크 감독 데이터로 학습
 방대하고 다양한 데이터 세트를 사용하여 약센트, 배경 소음, 전문 용어에 대한 견고성이 향상

Radford, Alec, et al. "Robust speech recognition via large-scale weak supervision." International Conference on Machine Learning. PMLR, 2023.

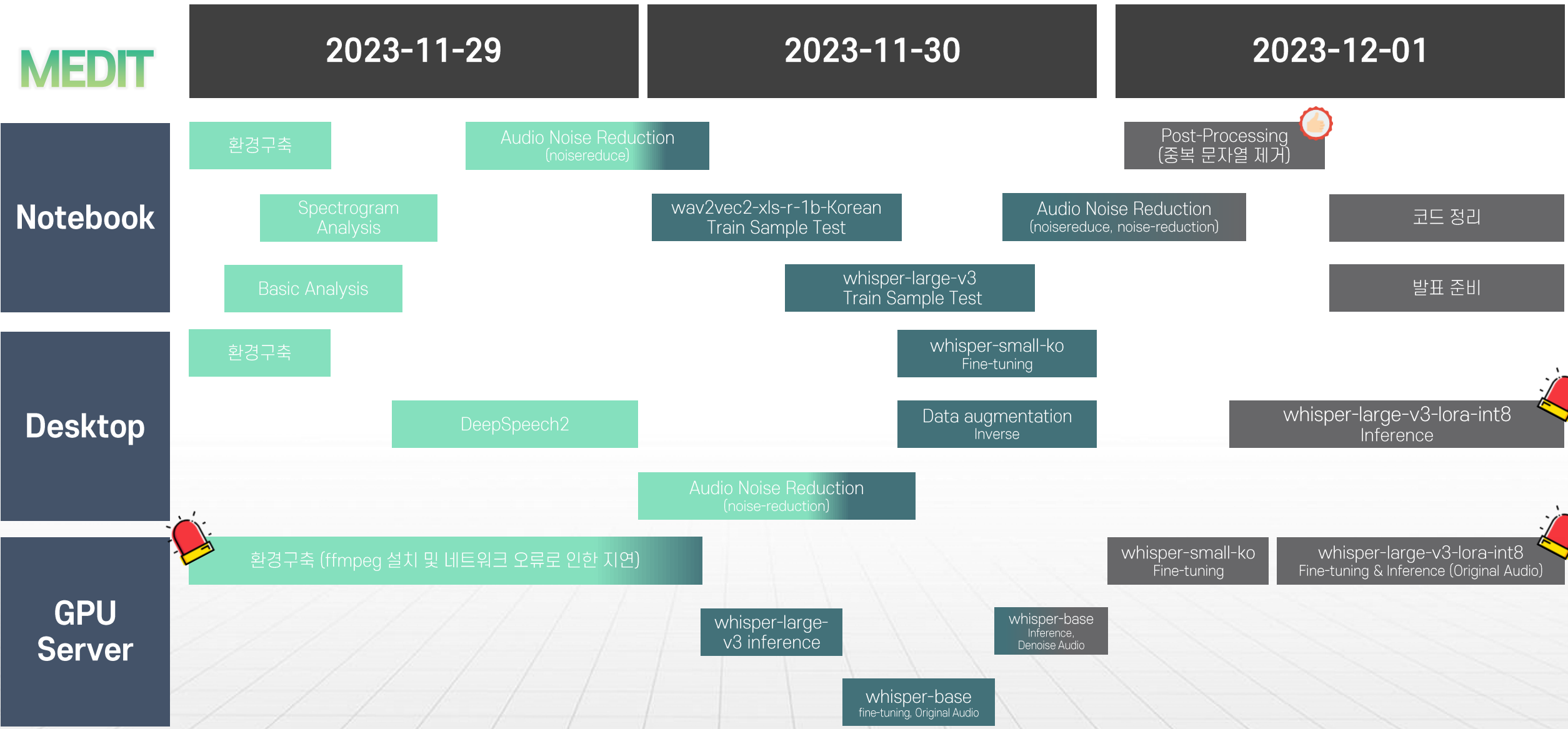
다양한 오디오로 구성된 대규모 데이터 세트를 학습하며
 다국어 음성 인식, 음성 번역, 언어 식별을 수행할 수 있는 멀티태스킹 모델

제출 결과

모델	활용 데이터	1 - CER
wav2vec2-xls-r-1b-korean	Train Sample (100)	0.062
whisper-large-v3	Train Sample (100)	0.596
whisper-large-v3 (Inference, Original Audio)	Train / Test	0.590
whisper-base (Fine-Tuning, Original Audio)	Train / Test	0.691
whisper-base (Inference, Denoise Audio)	Train / Test	0.650
whisper-base (Fine-Tuning, Denoise Audio)	Train / Test	0.692
whisper-base (Post-Processing, Denonise Audio)	Train / Test	0.693
SungBeom/whisper-small-ko (fine-tuning, Original Audio)	Train / Test	0.590
whisper-large-v3-lora-int8 (fine-tuning, Original Audio)	Train / Test	0.725657

- whisper-large-v3 fine tuning 과정에서 CUDA OOM Error로 whisper-base 모델 먼저 학습 진행 (test 데이터 추론에 1시간 소모)
- whisper-large-v3 모델을 학습에 사용하기 위해 LoRA 및 양자화를 적용 완료 (완료 시점 : 12.01 01:00)
- 실제 모델 학습에 사용된 시간이 미비함에 따라 기대한 목표 성능 획득 한계

II 발표 내용_ [평가 항목] 수행 프로세스



MEDIT

2023-11-29

2023-11-30

2023-12-01

Notebook

Desktop

GPU Server

환경구축

Audio Noise Reduction (noisereducer)

Post-Processing (중복 문자열 제거)

Spectrogram Analysis

wav2vec2-xls-r-1b-Korean Train Sample Test

Audio Noise Reduction (noisereducer, noise-reduction)

코드 정리

Basic Analysis

whisper-large-v3 Train Sample Test

발표 준비

환경구축

whisper-small-ko Fine-tuning

DeepSpeech2

Data augmentation Inverse

whisper-large-v3-lora-int8 Inference

Audio Noise Reduction (noise-reduction)

환경구축 (ffmpeg 설치 및 네트워크 오류로 인한 지연)

whisper-small-ko Fine-tuning

whisper-large-v3-lora-int8 Fine-tuning & Inference (Original Audio)

whisper-large-v3 inference

whisper-base Inference, Denoise Audio

whisper-base fine-tuning, Original Audio



2023 국방 AI 경진대회



Where AI & Defense Connect

목표 성능까지 도달하기 위해 적용을 검토했던 방법들이 줄줄이 성공적이지 못해 아쉬웠던 대회였습니다. 대회 네트워크 상황이나 시간 효율성을 고려하여 처음부터 whisper-large 모델을 믿고 나가지 못하고 whisper-base 모델을 학습했던 선택이 결과적으로 아쉽게 다가왔습니다. 그러나 이러한 시행착오들을 버텨내고 팀원들과 함께 최종 결과물을 만들어 낸 경험은 훗날 저에게 큰 성장의 밑거름이 될 것이라 확신합니다. 부대에서도 대한민국의 국방시기술에 많은 관심을 가지고 배워 나가겠습니다.

김민영

아쉽게 소비한 시간으로 인해 목표한 성능까지는 도달하지 못해 아쉽지만, 우수한 역량을 가진 팀원과 함께 문제를 풀어나간 과정은 재미있었습니다. 좋은 경험을 할 수 있게 대회를 준비하신 모든 분께 감사의 말씀을 표합니다. 또한, 국방혁신 4.0의 추동력을 유지하고 첨단과학기술 강군으로의 변혁을 주도하는 육군미래혁신연구센터에서 이번 경진 대회의 경험을 바탕으로 육군이 첨단과학기술군으로 나아가는 데 최선을 다하겠습니다.

배기민

소감 한마디

MEDIT

이번 대회를 기점으로 처음으로 AI 분야에 입문하게 되었는데, 짧은 기간 동안 머신러닝부터 딥러닝까지 공부하고 대회에 참가하며 딥러닝의 절차에 대해 알아갈 수 있는 시간이었습니다. 대회장에서는 처음으로 직접 해보는 AI 대회라 어색하고 뜻처럼 되지 않는 시행착오를 많이 겪었습니다. 그러나 좋은 팀원들을 만난 덕분에 어려움을 해결하고 결과를 도출해 낼 수 있었습니다. 대회를 계기로 육군에 더불어 대한민국의 AI 기술을 발전시켜 나갈 수 있는 일원이 되도록 더욱 노력하겠습니다.

권재영

인공지능 모델을 중점적으로 개발하는 대회는 처음이라 걱정도 많이 하였지만, 좋은 팀원 분들과 함께 할 수 있어서 즐겁게 참여할 수 있었습니다. 또한 지금까지 학습해 온 이론들을 직접 구현해 볼 수 있는 좋은 기회가 되었던 것 같습니다. 이번 대회에서 배운 모델링 기법들을 활용하여 육군지능정보기술단에서 맡은 임무에 성과와 효율성을 최대한 높일 수 있도록 최선을 다하겠습니다.

이학진