

I 팀 소개

Where AI & Defense Connect

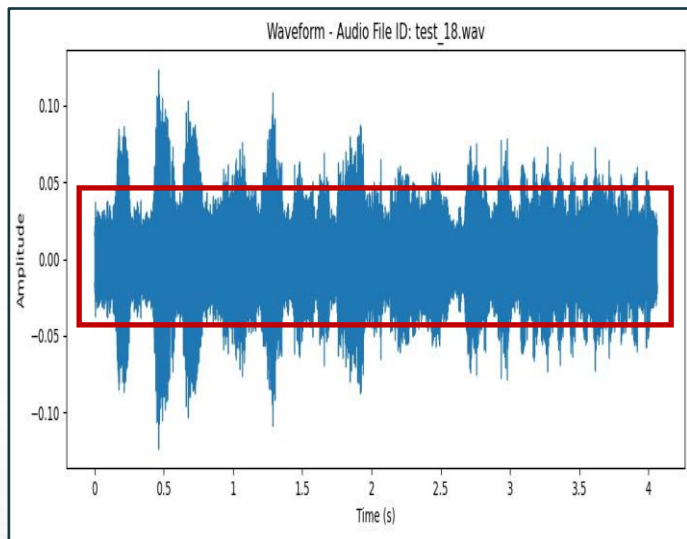
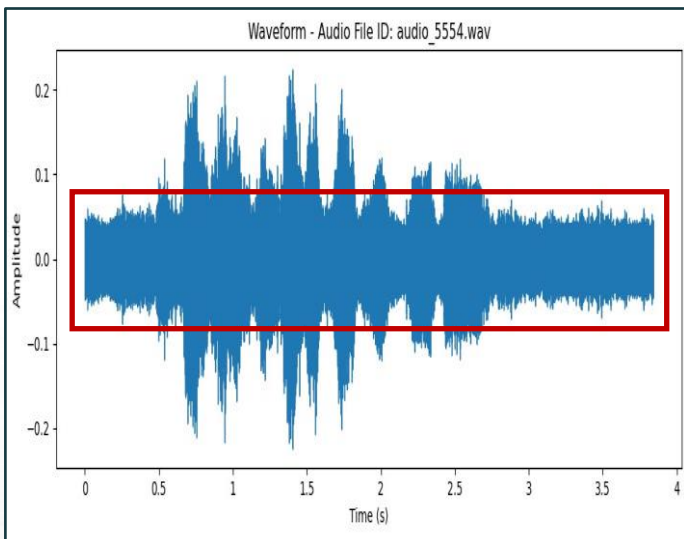


해군 7급 하은택 | 공군 일병 성무열 | 해군 중위 김환범 | 공군 중사 이시현

II 문제 해결 방법 구상

주 제 | 음성인식 기술을 활용한 효과적인 지휘관의 전술 명령 하달

문 제 점 | 노이즈 多 → 인식률 저하 → 한글 문장 생성 → 어색/틀린 문장



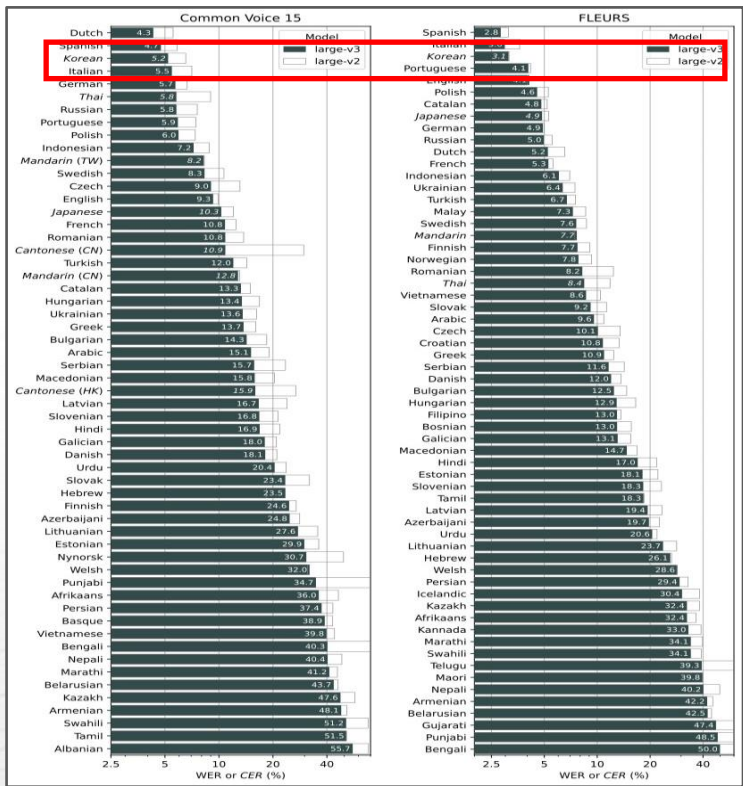
**Whisper
Large-V3** → **CER : 0.579**
(단순 inference)

해 결 방 법 | 노이즈 제거 전처리 → 모델 → 맞춤법, 어순 등 검사 후처리

IV ASR 모델 선택 및 학습

1. 모델 선택 과정

- Transformer 기반 모델 (Whisper / Wav2vec 2.0 / Seamlessm4t 등)
- Robust한 Whisper 선택



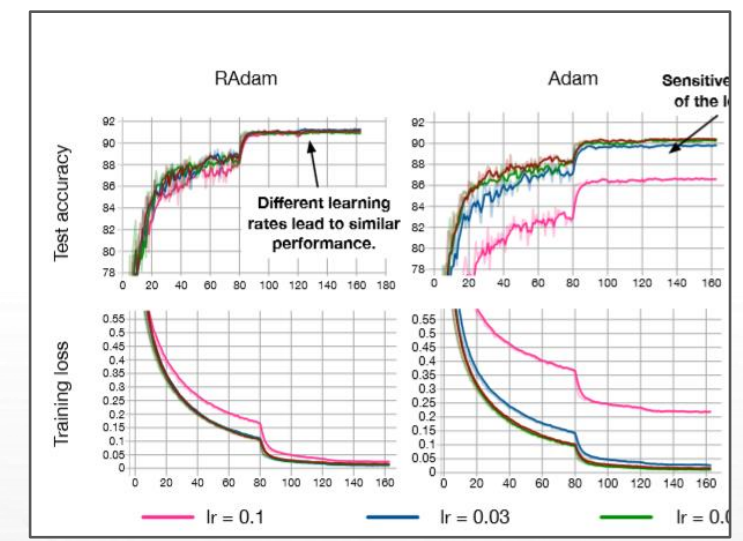
2. Small 선택

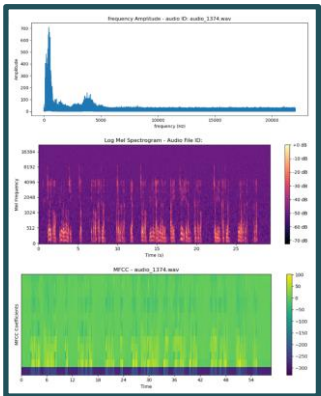
- 할당 VRAM 20GB
- distil/LoRA 등 경량화 모델 고려

Size	Parameters	Required VRAM
tiny	39 M	~1 GB
base	74 M	~1 GB
small	244 M	~2 GB
medium	769 M	~5 GB
large	1550 M	~10 GB

3. 학습 전략

- 6000 steps(약 10 epochs)
- RAdam(lr=1e-4)
- MultiStepLR([500, steps, 500])
- 허깅페이스 기반 베이스라인 구축





EDA

```

# 예시 입력 문장
input_text = "아니 진짜 무어 아니고"

# 입력 문장 인코딩
input_encoding = tokenizer("맞춤법을 고쳐주세요: " + input_text, return_tensors="pt")

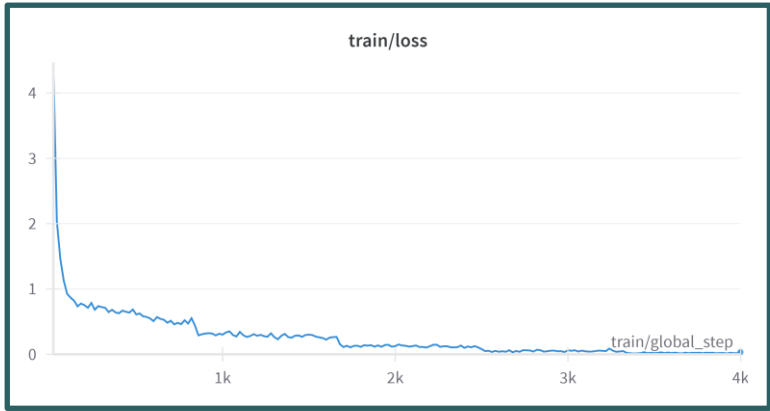
input_ids = input_encoding.input_ids.to(device)
attention_mask = input_encoding.attention_mask.to(device)

# TS 모델 출력 생성
output_encoding = model.generate(
    input_ids=input_ids,
    attention_mask=attention_mask,
    max_length=128,
    num_beams=5,
    early_stopping=True,
)

# 출력 문장 디코딩
output_text = tokenizer.decode(output_encoding[0], skip_special_tokens=True)

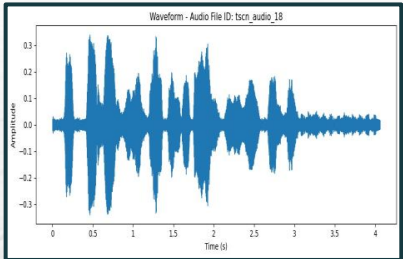
# 결과 출력
print(output_text) # 아니 진짜 뭐 아니고.
    
```

문장 검사기 구현



학습/검증

노이즈 제거 (Sepformer/TSCN)



whisper-small 베이스라인 구현

```

from transformers import Seq2SeqTrainingArguments

training_args = Seq2SeqTrainingArguments(
    output_dir="whisper_small_radam_multistep_3", # 원하는 디렉토리를 이용을 할것이다.
    per_device_train_batch_size=2,
    gradient_accumulation_steps=8, # 배치 크기가 2배 감소할 때마다 2배씩 증가
    learning_rate=1e-5,
    # num_train_epochs=3000,
    max_steps=1000, # epoch 대신 설정
    # gradient_checkpointing=True,
    fp16=True,
    evaluation_strategy="steps",
    # per_device_eval_batch_size=8,
    predict_with_generate=True,
    generation_max_length=325,
    # save_steps=25,
    # eval_steps=25,
    # logging_steps=2,
    save_steps=250,
    eval_steps=250,
    logging_steps=10,
    seeds=777,
    load_best_model_at_end=True,
    metric_for_best_model="dev", # 평가어의 경우 'dev'보다는 'dev'의 더 적절한 것
    greater_is_better=False,
    push_to_hub=False,
)
    
```

앙상블

Name	Status	Score
small_adamw_500steps	Done	0.625
small_radam_multistep_2000steps	Done	0.755
small_radam_multistep_4000steps	Done	0.809
ensemble(3500, 4000)	Done	0.81
ensemble(3500, 4000, tscn4000)	Done	0.8134



2023 국방 AI 경진대회



Where AI & Defense Connect

처음 태스크를 접했을 때 '노이즈 제거'를 핵심이라고 판단 하 모델 학습보다는 데이터 전처리에 중점을 두었습니다. 하지만 이 작업은 상당히 오래 걸렸고, 팀 내 판단으로 전처리 작업을 중단하고 먼저 준비한 모델을 학습시키는 전략을 택했습니다. 모델 학습에 따라 모델 성능은 좋아졌으며 모델 학습 후 전처리, 후처리를 통해 더 좋은 성능을 가진 모델을 도출할 수 있었습니다. 이번 대회를 통해 주어진 문제를 다양한 시각으로 접근해야 정답에 보다 효율적으로 도달할 수 있다는 점을 배울 수 있었습니다.