**COSE361(01)\_hw2**

2021150449 통계학과 신진섭

**1. Definition of your translation task**

자신의 영어 목소리를 이성의 영어 목소리로 음성 변환

**2. Description of the source code that you implemented or downloaded**

#코드 베이스 다운

!git clone <https://github.com/hpjang/VAE_experiment.git>

#가상환경 생성

!pip install virtualenv

#패키지 설치

!pip3 install torch pyworld pysptk h5py numpy kaldi-io librosa scipy scikit-learn fastdtw

#경로 설정

cd "/content/drive/My Drive/cyclevae2/VAE\_experiment"

#데이터 전처리

!python3 preprocess/preprocess-vcc2018.py

#모델 학습 & 변환

!bash run\_all.sh > vae\_exp\_0520.txt

**3. Description of the training/validation/test data that you**

Train data

<https://datashare.ed.ac.uk/bitstream/handle/10283/3061/vcc2018_database_training.zip>

다운로드 후 사용

Test data

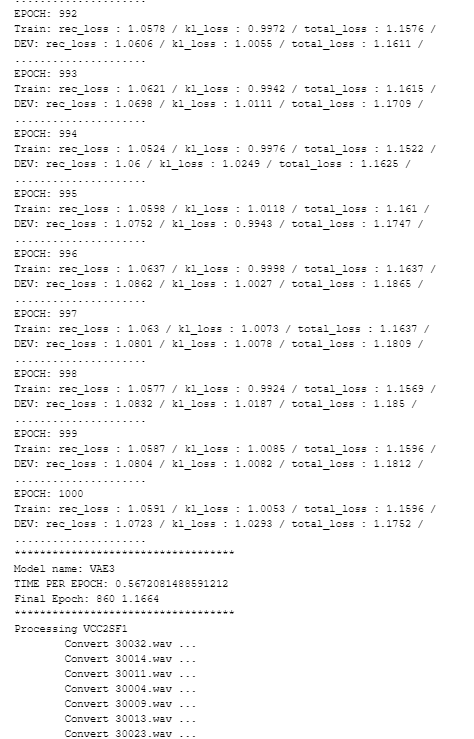
<https://datashare.ed.ac.uk/bitstream/handle/10283/3061/vcc2018_database_evaluation.zip>

다운로드 후 사용

에딘버러 대학의 DataShare 플랫폼에 호스팅된 ZIP 파일로, 2018년 음성 변환 챌린지(VCC 2018)를 위한 훈련 데이터. VCC 2018 데이터베이스는 음성 변환 시스템은 한 사람의 목소리를 다른 사람의 목소리처럼 들리도록 변환하여 음성 변환 시스템을 개발하고 평가하는 데 사용.

**4. Description of the compiling, training, validation, and testing procedures**

**vae\_exp\_0520.txt**



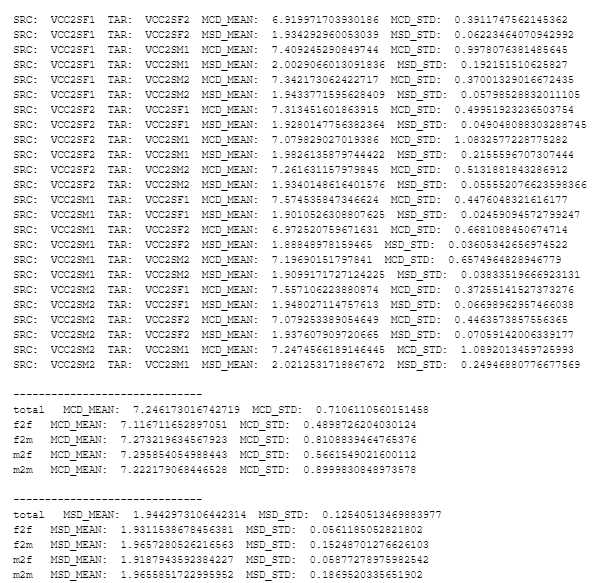
Epoch (학습 반복 횟수)

Train (훈련 데이터)

DEV (검증 데이터)

최종 epoch는 860에서 total\_loss 1.1664로 종료

**VAE3\_0.txt**



**MCD\_MEAN**: 변환 음성의 스펙트럼 차이를 측정하는 지표의 평균

**MCD\_STD**: 변환 음성의 스펙트럼 차이의 표준 편차

**MSD\_MEAN**: 주파수 편차의 평균

**MSD\_STD**: 주파수 편차의 표준 편차

f2f : 여성 -> 여성

f2m : 여성 -> 남성

m2f : 남성 -> 여성

m2m : 남성 -> 남성

**5. Resulting graphs of the hyperparameter tuning experiments**

!bash run\_graph\ \(1\).sh

import pandas as pd

df=pd.read\_csv('/content/drive/My Drive/cyclevae2/VAE\_experiment/graph/VAE3.csv')

df.reset\_index(inplace=True)

df.rename(columns={'index': 'epoch', 'epoch': 'Train', 'Train': 'Dev', 'Dev': 'nan'}, inplace=True)

df = df.drop(columns=['nan'])

import matplotlib.pyplot as plt

plt.figure(figsize=(10, 6))

plt.plot(df['epoch'], df['Train'], label='Train Loss', color='orange')

plt.plot(df['epoch'], df['Dev'], label='Dev Loss', color='red')

plt.xlabel('Epoch')

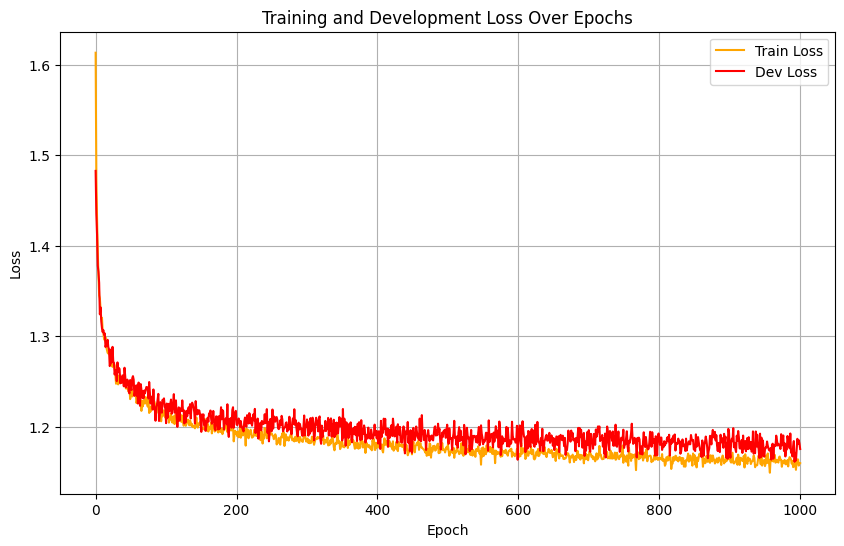
plt.ylabel('Loss')

plt.title('Training and Development Loss Over Epochs')

plt.legend()

plt.grid(True)

plt.show()



**6. A sample set of source (input), target, and converted (output) test data**

함께 제출하는 wav 파일 참고 부탁드립니다.

Input data : ‘input\_voice.wav’

Output data : ‘output\_voice.wav’