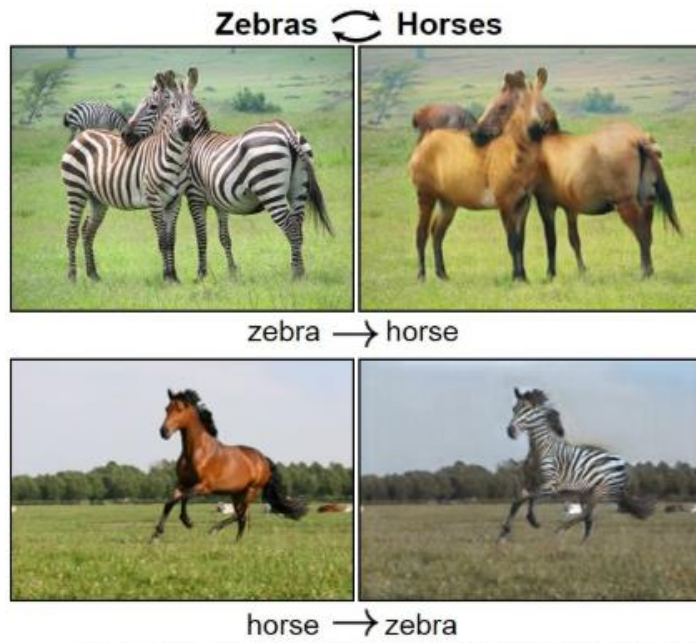




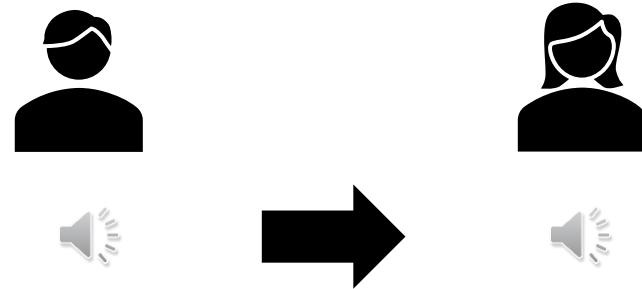
인공지능 과제 설명

과제 내용

- CycleGAN 기반 데이터 변환 코드 실습
 - 원본 정보를 유지하면서 다른 도메인으로 데이터를 바꾸는 모델



Image



Voice

과제 내용

□ 과제 수행 방법

1. GPU 서버 신청 or 개인 PC에서 작업

- E-mail: sgleem@ai.korea.ac.kr

2. CycleGAN 코드 받아오기

- Github에서 CycleGAN 모델을 사용하여 데이터를 변환하는 코드 다운로드
 - ex) https://github.com/leimao/Voice_Converter_CycleGAN
- CycleGAN 논문을 보고 직접 코드를 작성
 - Unpaired Image-to-Image Translation using Cycle-Consistent Adversarial Networks [J. Y. Zhu, 2017]

3. 데이터 수집 후 모델 학습

- 공개 데이터 사용
 - 다운로드 받은 github 코드에서 이용하는 데이터 직접 사용
- 직접 데이터 수집
 - 직접 사진을 찍거나 목소리 녹음

□ 코드는 공유해도 상관 없으나 데이터는 직접 돌릴 것

과제 내용

□ 제출 결과물

- 실험 보고서
 - 데이터 수집 방법
 - CycleGAN 모델 학습에 사용한 데이터에 대한 설명
 - Source train set, validation set
 - Target train set, validation set
 - 파일을 직접 첨부할 필요는 없음 (링크 or 간단한 설명)
 - 실험 목적
 - 변환할 특징
 - 유지시킬 특징
 - 사용한 프로그램에 대한 간단한 설명
 - 코드를 직접 첨부할 필요는 없음
- 학습한 CycleGAN 모델을 이용해 데이터를 변환한 결과물
 - Source test set
 - Target test set
 - 파일을 압축하여 첨부할 것

과제 내용

□ 평가 기준

- 과제 목적에 맞게 제출하였는지 (20)
- 사용한 데이터가 실험 목적에 부합하는지 (40)
 - ex1)
실험 목적: 실제 사진을 스케치한 그림으로 바꾸는 모델 학습
[O] 사용한 데이터: 실제 사진 (Source) -> 스케치한 그림 (Target)
[X] 사용한 데이터: 실제 사진 (Source) -> 실제 사진 (Target)
 - ex2)
실험 목적: 내 목소리를 A라는 여성의 목소리로 바꾸는 모델 학습
[O] 사용한 데이터: 내 목소리(Source) -> A 목소리(Target)
[X] 사용한 데이터: 내 목소리 (Source) -> 다른 여성인 B의 목소리(Target)
- 데이터가 실험 목적에 따라 잘 변환되었는지 (40)
 - 변환된 데이터가 의미 있게 변환되었는가



GPU 서버 설명

서버 환경

- ❑ 리눅스 환경에서 GPU를 활용하는 **Python** 기반 딥러닝 라이브러리 제공
 - Ubuntu 18.04.2
 - Python 2.7.15 && Python 3.6.7
- ❑ GTX 1080 Ti 4장 보유
 - CUDA 10.0
 - CUDNN 7.4.1

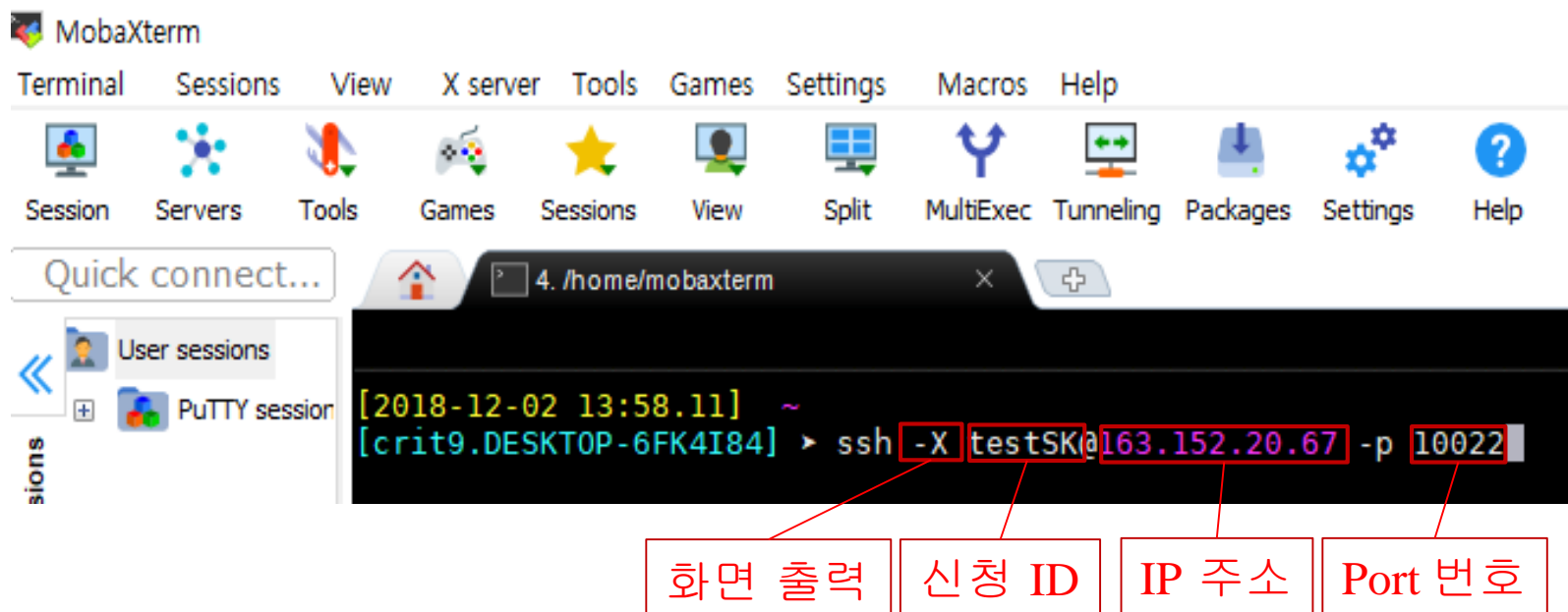
서버 접속 방법

- ❑ 서버 정보
 - DNS : **csgpu01.korea.ac.kr** or **csgpu02.korea.ac.kr**
 - Port 번호 : **10022** or **10023**
- ❑ 외부에서 접속이 불가능
 - 교내 **WiFi**에서만 접속이 가능함
- ❑ MobaXterm
 - GPU 서버 원격 접속을 위한 프로그램
 - <https://mobaxterm.mobatek.net/download.html>
 - Home Edition으로 다운로드

서버 접속 방법

❑ MobaXterm

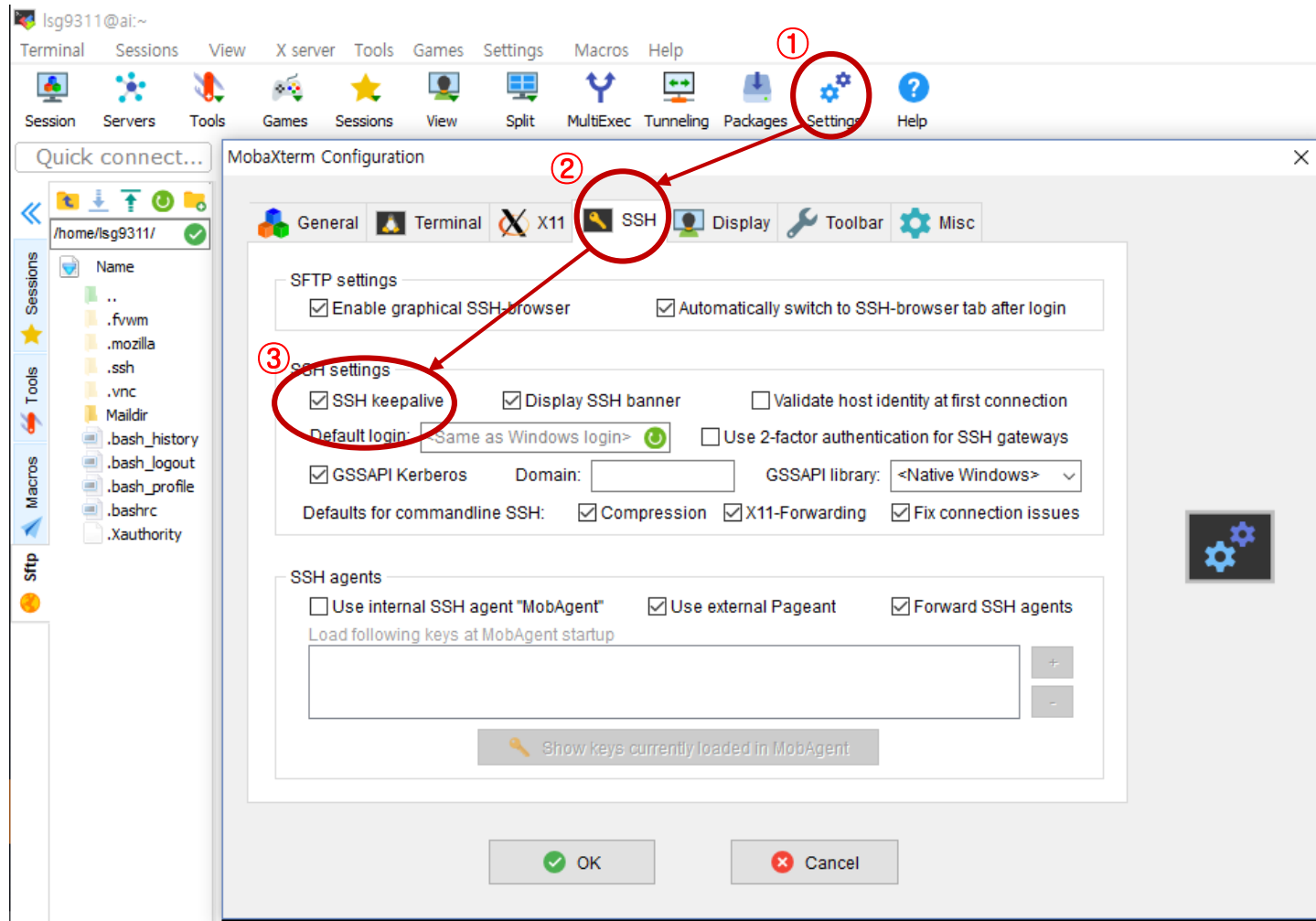
- 실행 후 아래 명령어를 입력
- `ssh -X <신청 ID>@<할당 받은 DNS 주소> -p <할당 받은 port 번호>`
- 이후 신청한 비밀번호를 입력



서버 접속 방법

❑ MobaXterm

- 서버 연결이 끊길 수 있으니 아래와 같이 설정해주세요.





과제 진행 예시

과제 진행 예시

□ CycleGAN voice conversion 코드 다운로드

- https://github.com/leimao/Voice_Converter_CycleGAN 접속
- Github 접속 후 주소 복사

Voice Converter Using CycleGAN and Non-Parallel Data


voice-conversion cyclegan speech

32 commits 1 branch 0 releases 1 contributor MIT

Branch: master New pull request Find File **Clone or download**

leimao update

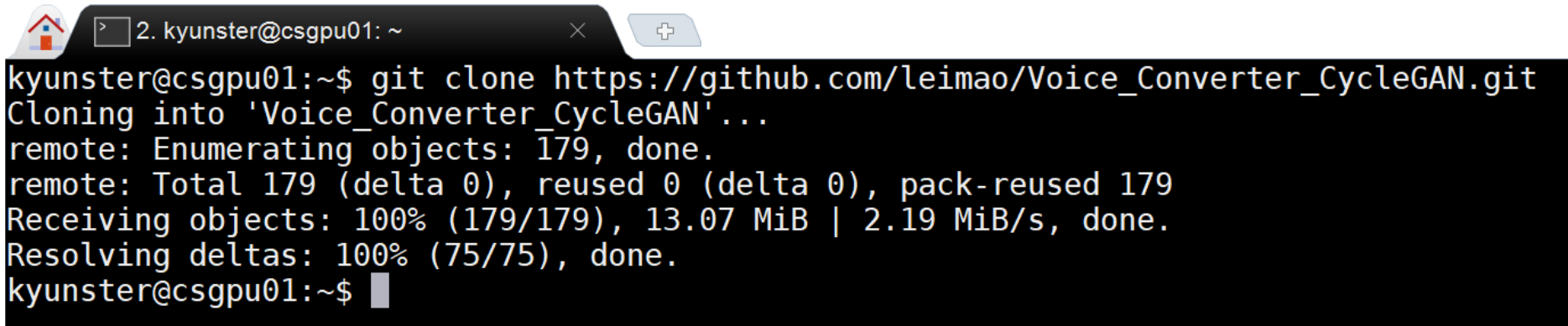
demo	update demo and training information	
figures	figure	
train_log	update demo and training information	
LICENSE.md	Initial commit	
README.md	update	9 months ago
convert.py	fix the bug that the two inputs are the same	10 months ago
download.py	update	11 months ago
model.py	update	11 months ago
module.py	fix GLU implementations	10 months ago
preprocess.py	fix pitch conversion bug	11 months ago
train.py	fix the bug that the two inputs are the same	10 months ago
utils.py	update	11 months ago

Clone with HTTPS ?
Use Git or checkout with SVN using the web URL.
https://github.com/leimao/Voice_Converter_CycleGAN 

Open in Desktop Download ZIP

과제 진행 예시

- ❑ CycleGAN voice conversion 코드 다운로드
 - git clone <주소 붙여넣기>

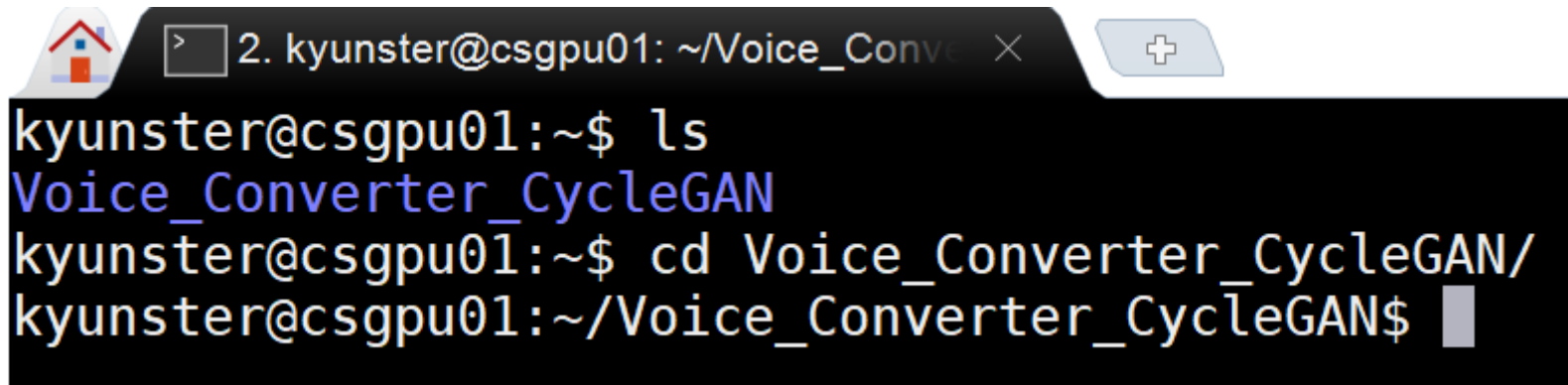
A terminal window with a dark background and light-colored text. The window title bar shows a home icon, a tab labeled '2. kyunster@csgpu01: ~', and a close button. The terminal content shows the execution of a git clone command to download a repository from GitHub. The output indicates that 179 objects were enumerated and cloned successfully, with a total size of 13.07 MiB.

```
kyunster@csgpu01:~$ git clone https://github.com/leimao/Voice_Converter_CycleGAN.git
Cloning into 'Voice_Converter_CycleGAN'...
remote: Enumerating objects: 179, done.
remote: Total 179 (delta 0), reused 0 (delta 0), pack-reused 179
Receiving objects: 100% (179/179), 13.07 MiB | 2.19 MiB/s, done.
Resolving deltas: 100% (75/75), done.
kyunster@csgpu01:~$
```

과제 진행 예시

❑ CycleGAN voice conversion 코드 다운로드

- `ls`
 - 현재 디렉토리 내 파일 및 폴더 확인
- `cd <이동할 디렉토리>`
 - 디렉토리 이동




```
2. kyunster@csgpu01: ~/Voice_Conve ×  
kyunster@csgpu01:~$ ls  
Voice_Converter_CycleGAN  
kyunster@csgpu01:~$ cd Voice_Converter_CycleGAN/  
kyunster@csgpu01:~/Voice_Converter_CycleGAN$
```

과제 진행 예시


- VCC2016 데이터 다운로드
 - 현재 자동으로 다운로드 받는 코드에 오류가 있음
 - <https://datashare.is.ed.ac.uk/handle/10283/2211> 접속


Download all files 

 Readme.txt (3.284Kb)

 Listening test results: VCC questionnaire responses - spreadsheet with all descriptions of systems provided by participants (503.5Kb)


 Listening test results: the MOS and similarity scores (35.24Kb)


 Participant submissions: sound files, source, target, baseline and participants' submissions (1.940Gb)

 VCC training data: readme file released with the training data (1.666Kb)

 VCC training data: source and target evaluation data (43.42Mb)

Test data

 VCC training data: evaluation data released to participants during the challenge (21.52Mb)

 VCC training data: training data released to participants during the challenge (142.7Mb)

Training data

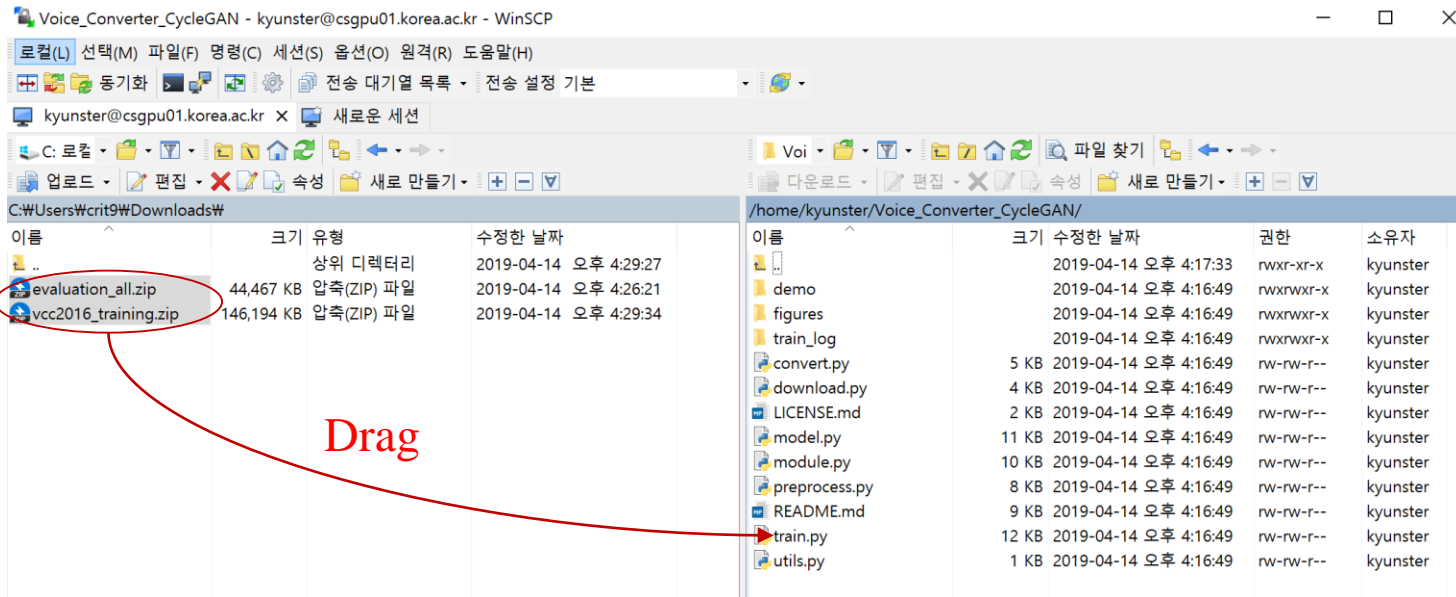
 Tomoki Toda, Ling-Hui Chen, Daisuke Saito, Fernando Villavicencio, Mirjam Wester, Zhizheng Wu, Junichi Yamagishi "The Voice Conversion Challenge 2016" in Proc. of Interspeech, San Francisco. (195.8Kb)

 Mirjam Wester, Zhizheng Wu, Junichi Yamagishi "Analysis of the Voice Conversion Challenge 2016 Evaluation Results" in Proc. of Interspeech 2016. (227.3Kb)

 license_text (17.00Kb)

과제 진행 예시

- 데이터 서버로 업로드
 - WinSCP 이용
 - 수집한 데이터를 서버로 업로드



과제 진행 예시

- 데이터 압축 해제
 - unzip <압축파일 이름>

```
kyunster@csgpu01:~/Voice_Converter_CycleGAN$ ls
convert.py  download.py  figures      model.py    preprocess.py  train_log  utils.py
demo        evaluation_all.zip  LICENSE.md  module.py   README.md     train.py   vcc2016_training.zip
kyunster@csgpu01:~/Voice_Converter_CycleGAN$ unzip evaluation_all.zip
Archive:  evaluation_all.zip
  creating: evaluation_all/
  inflating: evaluation_all/.DS_Store
   creating: __MACOSX/
   creating: __MACOSX/evaluation_all/
  inflating: __MACOSX/evaluation_all/._.DS_Store
```

과제 진행 예시

❑ 직접 녹음을 진행하고 싶은 경우

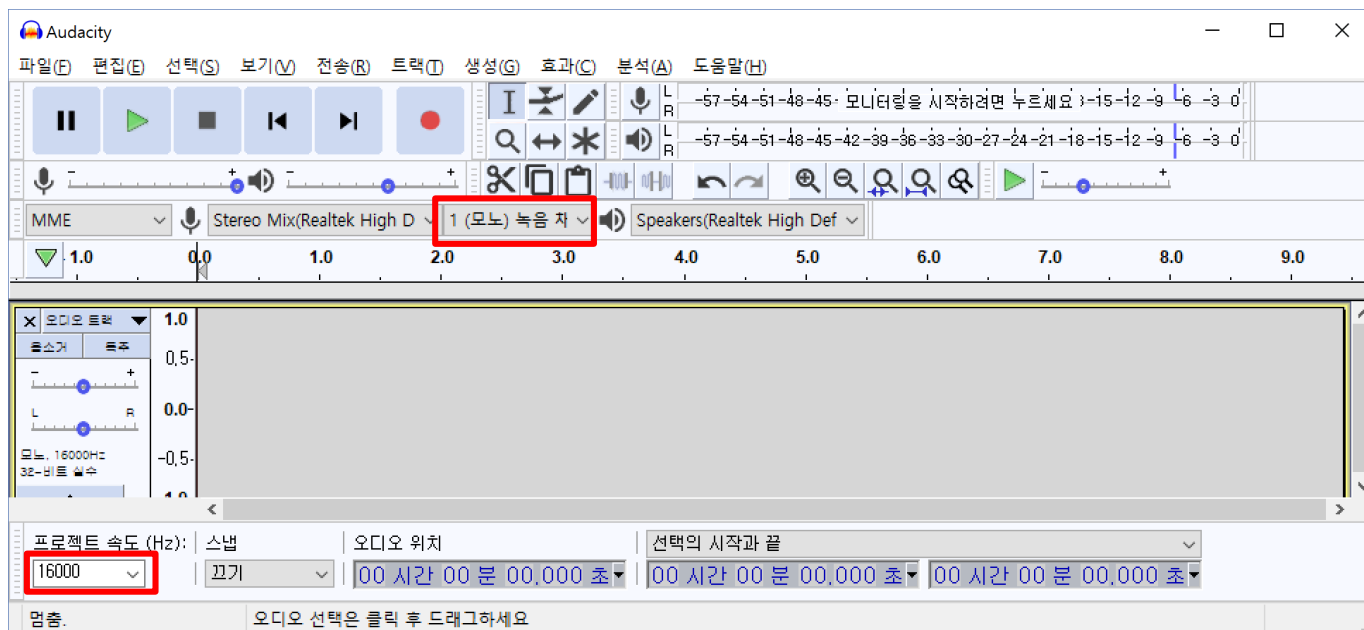
■ Audacity 다운로드

- <https://www.audacityteam.org/download/windows/>

❑ 녹음기 설정

■ 아래와 같이 설정

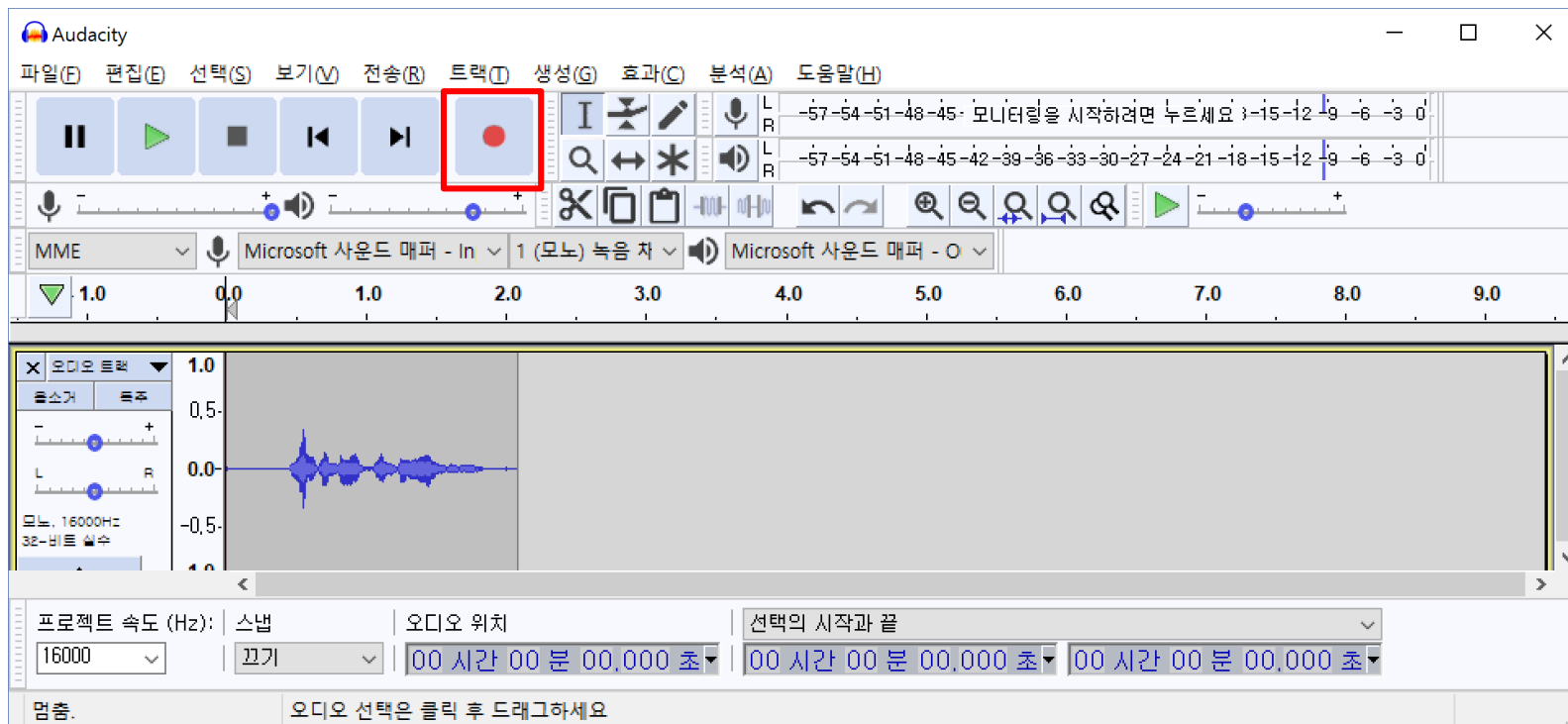
- 아래 설정과 같이 하지 않을 경우 코드가 실행되지 않을 수 있음



과제 진행 예시

□ 녹음 문장

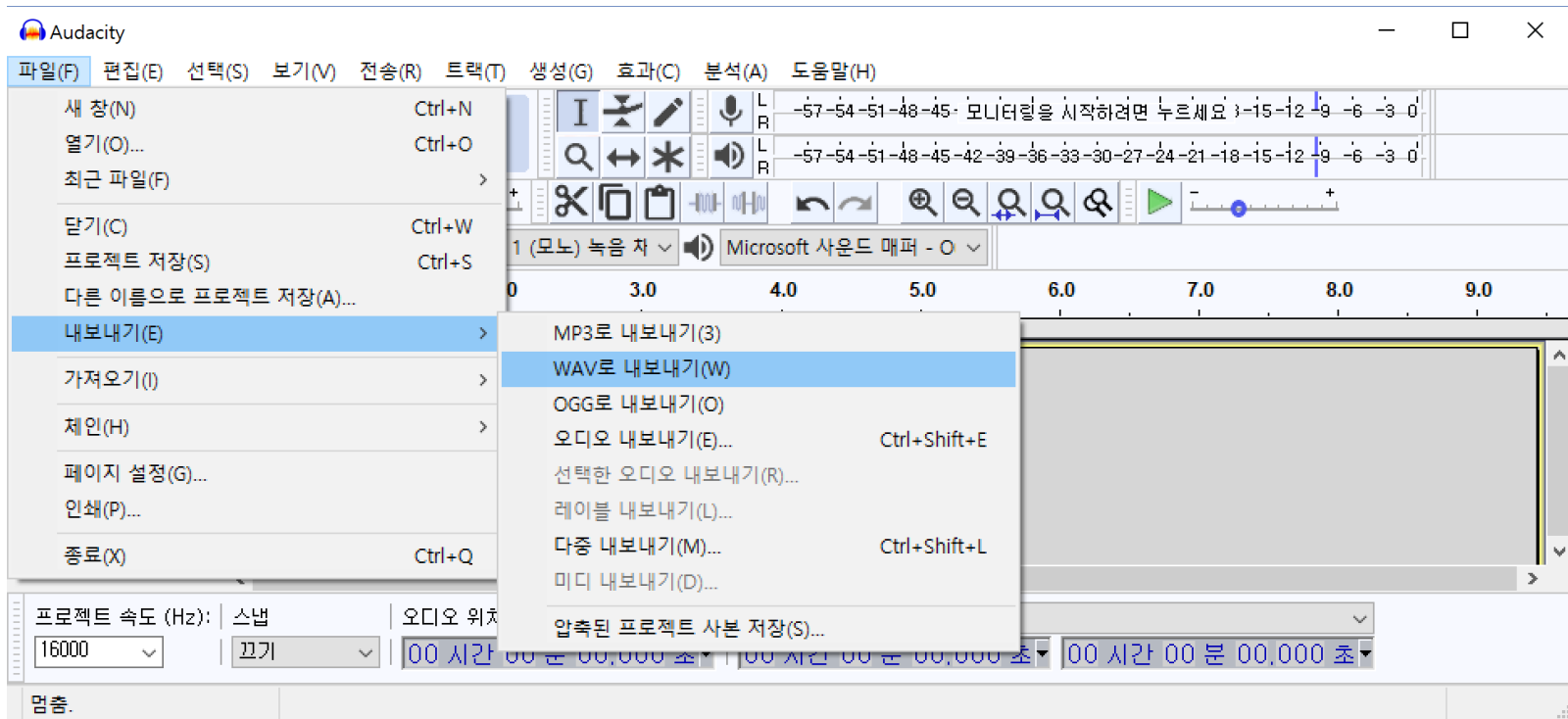
- VCC2016 데이터셋은 대본이 제공되지 않음
 1. 문장 듣고 직접 따라하기
 2. 아무 영어 문장이나 162 문장 녹음



과제 진행 예시

□ 녹음 파일 저장

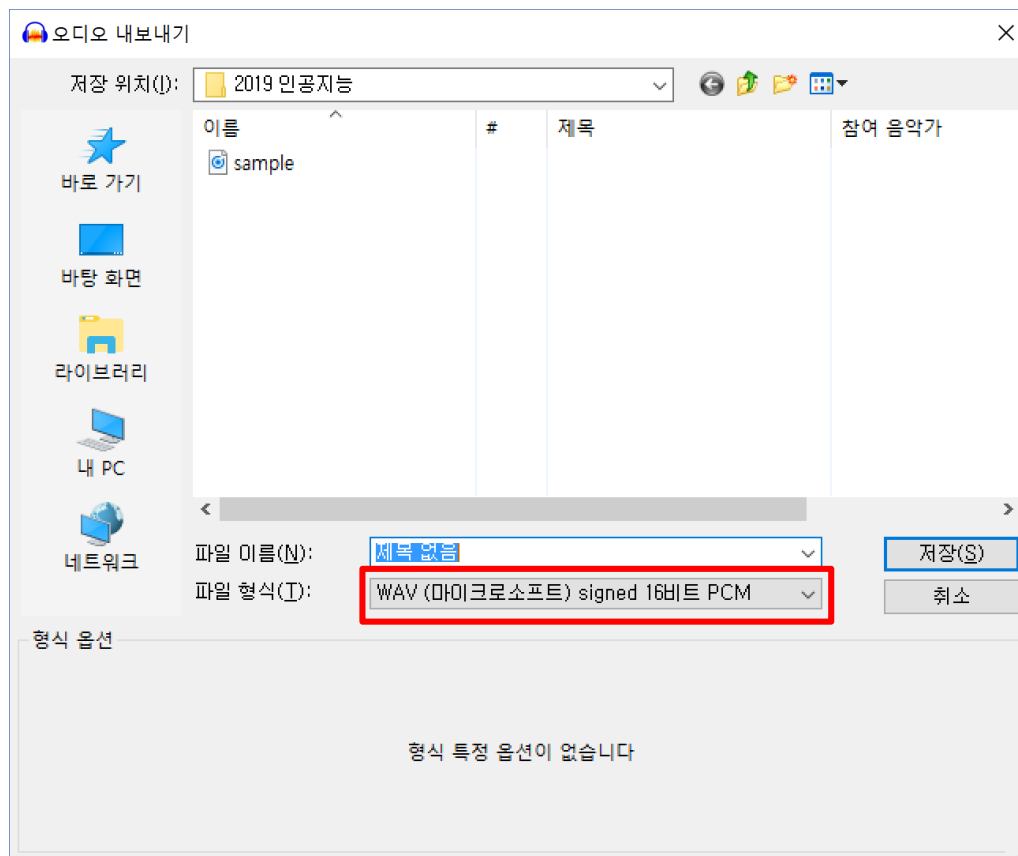
- 한 문장 녹음 후 => 파일 => 내보내기 => WAV로 내보내기
- WAV(마이크로소프트) signed 16비트 PCM으로 저장
- 저장 이후 새 창 만들어서 다시 녹음할 것



과제 진행 예시

□ 녹음 파일 저장

- 한 문장 녹음 후 => 파일 => 내보내기 => WAV로 내보내기
- WAV(마이크로소프트) signed 16비트 PCM으로 저장
- 저장 이후 새 창 만들어서 다시 녹음할 것

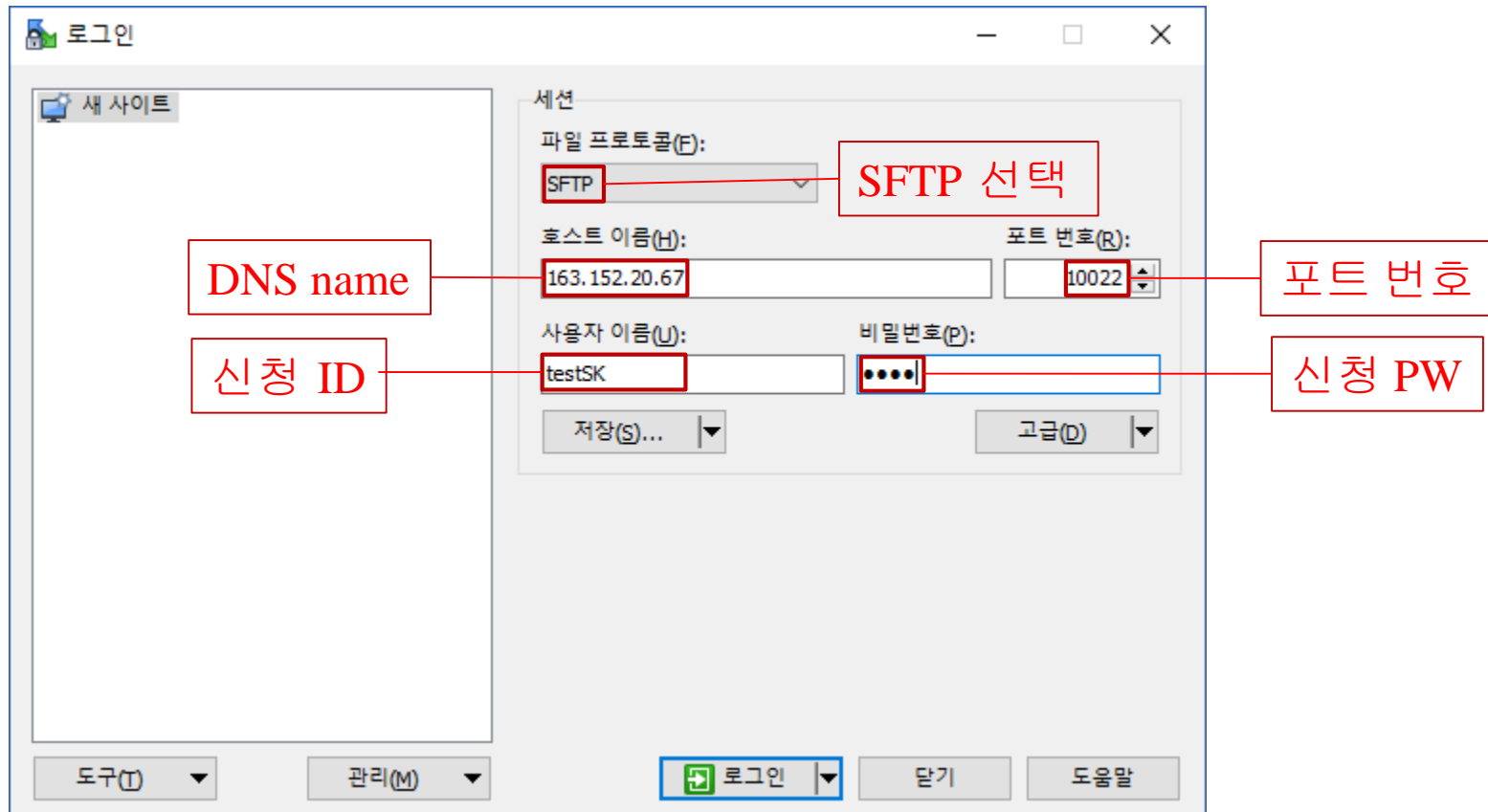


과제 진행 예시

❑ 데이터 서버로 업로드

▪ WinSCP 이용

- <https://winscp.net/download/WinSCP-5.13.5-Setup.exe>
- 아래와 같이 입력한 뒤 서버 접속



과제 진행 예시

□ 모델 학습

- 해당 폴더로 이동: `cd Voice...`
- `python3 train.py`
 - `--train_A_dir` <source 화자 train 폴더>
 - `--train_B_dir` <target 화자 train 폴더>
 - `--model_dir` <학습한 모델 저장할 폴더>
 - `--model_name` <모델 파일 이름>
 - `--random_seed` 0
 - `--validation_A_dir` <source 화자 dev(test) 폴더>
 - `--validation_B_dir` <target 화자 dev(test) 폴더>
 - `--output_dir` <성능 측정 결과 저장 폴더>
 - `--tensorboard_log_dir` <텐서보드 결과 저장 폴더>

```
kyunster@csgpu01:~/Voice_Converter_CycleGAN$ python3 train.py --train_A_dir ./vcc2016_training/SF1 --train_B_dir ./vcc2016_training/TM1 --model_dir ./model/sf1_tm1 --model_name sf1_tm1.ckpt --random_seed 0 --validation_A_dir ./evaluation_all/SF1 --validation_B_dir ./evaluation_all/TM1 --output_dir ./validation_output --tensorboard_log_dir ./log
```

과제 진행 예시

□ 설정 변경

- vim train.py [vim 에디터 진입]
- 아래 부분의 코드 수정
- **num_epochs = 5000 => num_epochs = 1000**

```
import os
import numpy as np
import argparse
import time
import librosa

from preprocess import *
from model import CycleGAN

def train(train_A_dir, train_B_dir, model_dir, model_name, random_seed, validation_A_dir, validation_B_dir, output_dir, tensorboard_log_dir):
    np.random.seed(random_seed)

    num_epochs = 5000
    mini_batch_size = 1 # mini_batch_size = 1 is better
    generator_learning_rate = 0.0002
    generator_learning_rate_decay = generator_learning_rate / 200000
    discriminator_learning_rate = 0.0001
    discriminator_learning_rate_decay = discriminator_learning_rate / 200000
    sampling_rate = 16000
    num_mcep = 24
    frame_period = 5.0
    n_frames = 128
    lambda_cycle = 10
    lambda_identity = 5
```


과제 진행 예시

□ 데이터 변환

- `python3 convert.py`
 - `--model_dir` <학습한 모델 저장한 폴더>
 - `--model_name` <모델 파일 이름>
 - `--data_dir` <변환 시킬 데이터 폴더>
 - `--conversion_direction` A2B
 - `--output_dir` <변환된 음성 저장 폴더>

```
kyunster@csgpu01:~/Voice_Converter_CycleGAN$ python3 convert.py --model_dir model/sf1_tm1 --model_name sf1_tm1.ckpt --data_dir data/test --conversion_direction A2B --output_dir result_2
```



과제 진행 시 유의사항

유의사항

- ❑ 외부에서는 서버 사용이 불가능합니다.
- ❑ 과제와 관련 없는 일을 서버에서 수행할 경우 서버 사용 권한 박탈
- ❑ 24시간 이상 GPU 점유 시 강제종료
 - PPT에 있는 예제 사용하시는 경우 꼭 설정 변경해주세요.
- ❑ 미리미리 사용해주세요.
 - 여러 사람이 동시에 사용할 경우 GPU 서버를 이용하지 못할 수 있음.
 - GPU 서버 점유로 인해 과제 진행을 못하셔도 책임지지 않습니다.
- ❑ 모델이나 변환된 데이터는 꼭 로컬로 옮겨 저장해 둘 것
 - 서버 문제로 인해 데이터가 날아가도 책임을 지지 않습니다.
- ❑ 코드 실행 후 그대로 창을 닫아버리면 프로그램이 꺼짐
 - Solution) nohup [명령어] > [log 출력파일] &
 - https://zetawiki.com/wiki/%EB%A6%AC%EB%88%85%EC%8A%A4_nohup_%EC%82%AC%EC%9A%A9%EB%B2%95

유의사항

- ❑ 궁금한 점이 있으실 경우
 - E-mail 1(권장): sgleem@ai.korea.ac.kr
 - E-mail 2(답장이 느릴 수 있음): crit9311@gmail.com
- ❑ PPT 예제 관련 및 서버 사용 관련 질문 받습니다.
- ❑ Tensorflow, Pytorch와 같은 라이브러리 사용법에 대한 질문도 받습니다.
- ❑ 단, 다른 코드 사용법에 대한 질문은 답변 드릴 수 없습니다.