

kaggle 음성인식 챌린지 수행계획서

과제명	Kaggle 음성인식 챌린지		
참여기관명	충남대 알고리즘 동아리 ANT	수행책임자	서영학
참여인원	임진수, 서영학, 김휘준, 정윤수, 김동하, 차복, 신하람 ,김남형		
알고리즘	MFCC 와 Recurrent CNN 을 이용한 음성인식		

○ 과제 수행 환경:

충남대학교 인공지능 랩, Ubuntu 14.4 도구 Tensorflow GPU / 그래픽 환경 GTX1080, GTX1070, GTX960

○ 과제 내용:

- 데이터

케글 Tensorflow Speech Recognition 데이터

- 데이터 전처리 과정

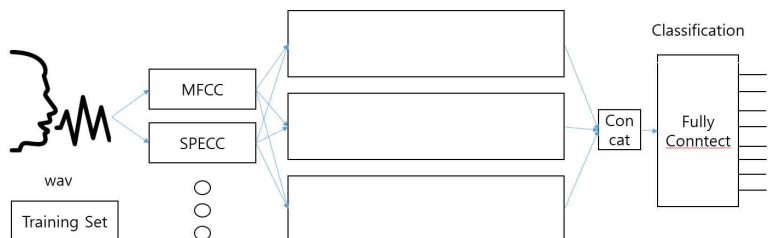
wav 파일을 MFCC, Spectrogram 등 특징을 추출하는 알고리즘을 이용하고 추출된 데이터를 이미지화 시킵니다. 그 후 강인한 모델을 만들기 위해서 이미지에 노이즈를 추가하거나 자르는 등 방법을 사용하여 데이터양을 늘려 학습합니다.

- 모델 수행 과정

기존 MFCC, Spectrogram 등 음원에서 feature을 추출할 수 있는 알고리즘을 사용하여 다양한 feature을 추출 한 후 각 특징 간 유의한 feature들 끼리 묶어 다양한 CNN 네트워크(ResNet, CapsuleNet, DenseNet)를 통하여 각 네트워크 별로 결과를 내고, 각 결과를 앙상블기법을 이용하여 최종적인 분류를 하는 모델로 과제를 수행할 것이다.

모델

Neural Network model



```
graph LR; wav[wav] --> MFCC; wav --> SPECC; MFCC --> L1[ ]; SPECC --> L1; MFCC --> L2[ ]; SPECC --> L2; MFCC --> L3[ ]; SPECC --> L3; L1 --> Concat[Con cat]; L2 --> Concat; L3 --> Concat; Concat --> FullyConnect[Fully Connect]; FullyConnect --> Classification
```