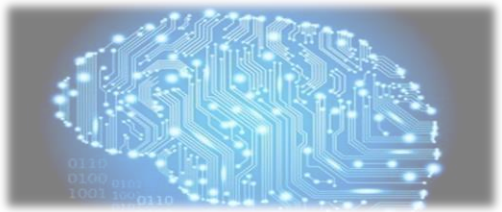
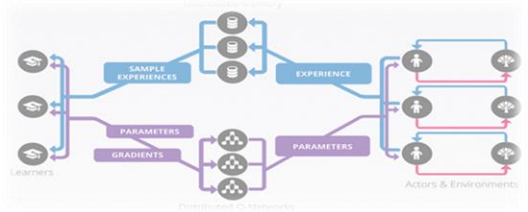



Musical Genre Classification via CNN

AutoML 연구실
석사 과정 조성현
21.10.12

The standard statistics of the datasets

Dataset	|W|	Avg. Length	Suggested Domain	Used Domain	|C|	Avg. Size of Class
Ballroom	698	30	Genre	Genre	8	87.25 ± 17.56
Ballroom-Extended	4180	30	Genre	Genre	13	321.54 ± 195.70
emoMusic-G	1000	45	Emotion	Genre	8	125.00 ± 0.0
Emotify-G	400	43	Emotion	Genre	4	100.00 ± 0.0
FMA-MEDIUM	24984	30	Genre	Genre	16	1561.50 ± 2069.36
FMA-SMALL	7996	30	Genre	Genre	8	999.50 ± 0.50
GiantStepsKey-G	604	80	Key	Genre	24	24.9130 ± 28.0274
GMD-G	1042	300	Groove	Genre	18	57.89 ± 70.07
GTZAN	1000	30	Genre	Genre	10	100.000 ± 0.000
HOMBURG	1886	10	Genre	Genre	9	209.56 ± 134.87
ISMIR04	2187	120	Genre	Genre	6	121.17 ± 95.60
MICM	1137	360	Genre	Genre	7	162.43 ± 119.72
Tropical Genres	1500	30	Genre	Genre	5	300.000 ± 0.000

|W| : the number of patterns
Avg. Length : the average length of music excerpts in the collection
Suggested Domain : the domain suggested originally by the creator of the collection
Used Domain : the domain used for creating the corresponding dataset
|C| : the number of classes or tags
Avg. Size of Class : the average number of patterns per class and its standard deviation

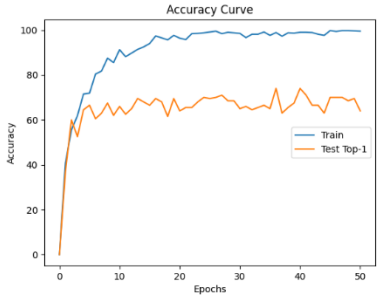
Experiment Settings

Dataset	Avg. Length
Ballroom	30
Ballroom-Extended	30
emoMusic-G	45
Emotify-G	43
FMA-MEDIUM	30
FMA-SMALL	30
GiantStepsKey-G	80
GMD-G	300
GTZAN	30
HOMBURG	10
ISMIR04	30
MICM	360
Tropical Genres	30

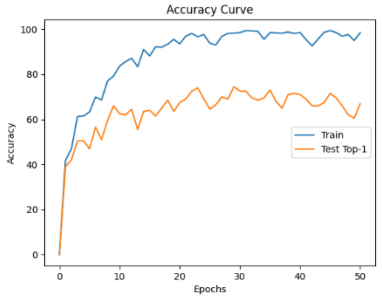
- Overall preprocessings and settings
  1. Convert to .wav files
  2. Compute mel spectrograms and save as .npy files
  3. Contvert to mel spectrogram (0~10s, 5~15s, 10~20s, 15~25s, 20~30s)
  4. Train : Test = 8 : 2

Comparison with other CNN models

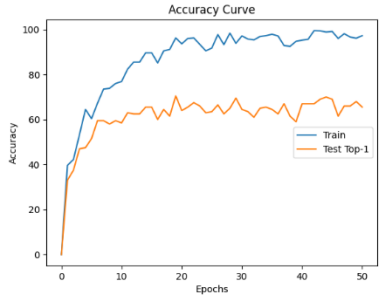
Dataset	Proposed	ResNet18	ResNet50	GoogleNet	DenseNet	ShuffleNet v2
Ballroom	-	-	-	-	-	-
Ballroom-Extended	-	-	-	-	-	-
emoMusic-G	-	-	-	-	-	-
Emotify-G	-	-	-	-	-	-
FMA-MEDIUM	-	-	-	-	-	-
FMA-SMALL	-	-	-	-	-	-
GiantStepsKey-G	-	-	-	-	-	-
GMD-G	-	-	-	-	-	-
GTZAN	64.0	67.0	65.5	57.5	70.5	62.5
HOMBURG	-	-	-	-	-	-
ISMIR04	-	-	-	-	-	-
MICM	-	-	-	-	-	-
Tropical Genres	-	-	-	-	-	-



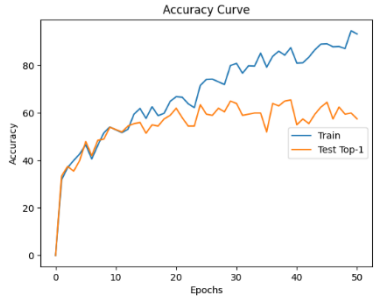
Proposed



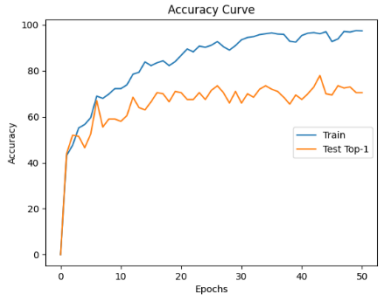
ResNet18



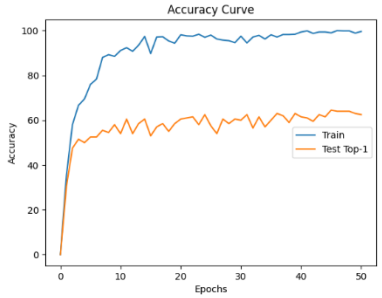
ResNet50



GoogleNet



DenseNet



ShuffleNet v2

Further Discussion Points

Use only datasets that are about 30 seconds long for genre domain classification.
(Included datasets : Ballroom, Ballroom-Extended, FMA-MEDIUM, FMA-SMALL, GTZAN, Tropical Genres)

1. Numpy values are not fixed.
    - Reduced or expanded data may differ from the original value