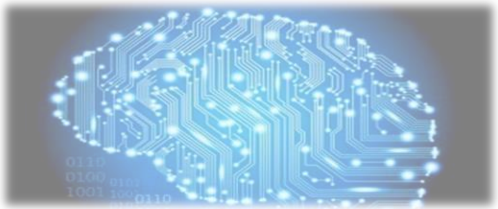
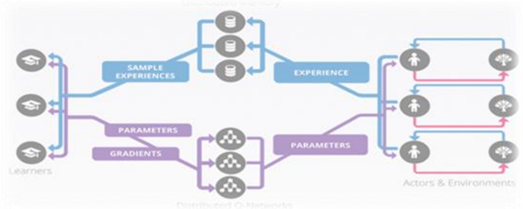


Recommendation System

AutoML 연구실
학부연구생 조성현
21.01.15

Contents

Genetic Algorithm for Multi-label Feature Selection

1. GA problem to be solved

2. Solutions

3. The reason why the problem should be improved in the domain within the paper

4. The reason why the solution worked effectively in the domain of the paper

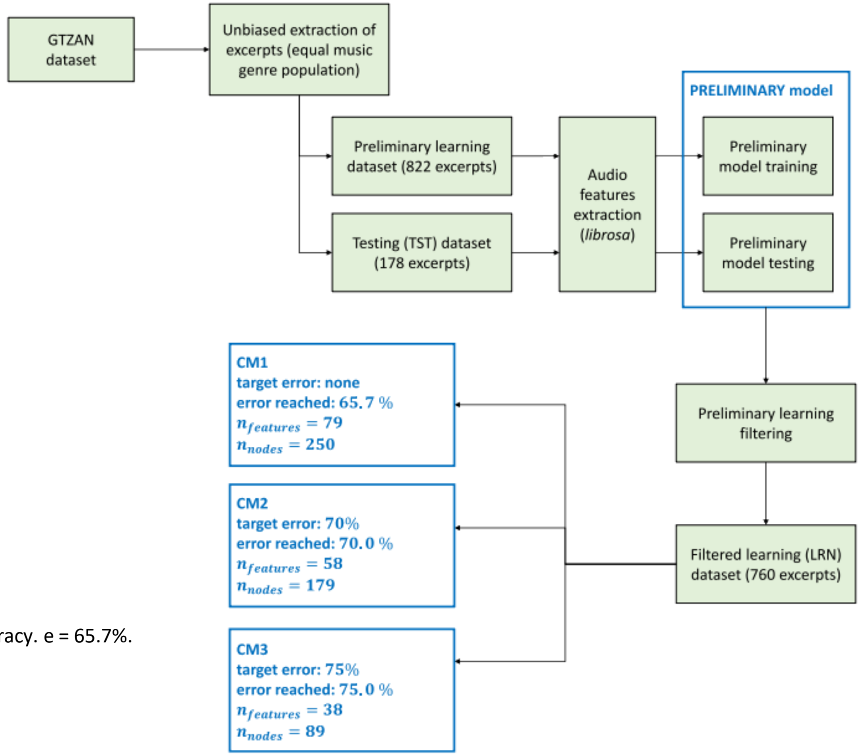
Neuro-genetic programming for multi-genre classification of music content

1. GA problems to be solved

- Most of the music genre classifiers focus on obtaining a single genre to raise accuracy.

2. Solutions

- Low computational complexity and reduced number of audio features yield similar results.



The Classification Model 1 (**CM1**) : largest complexity and greatest accuracy. e = 65.7%.

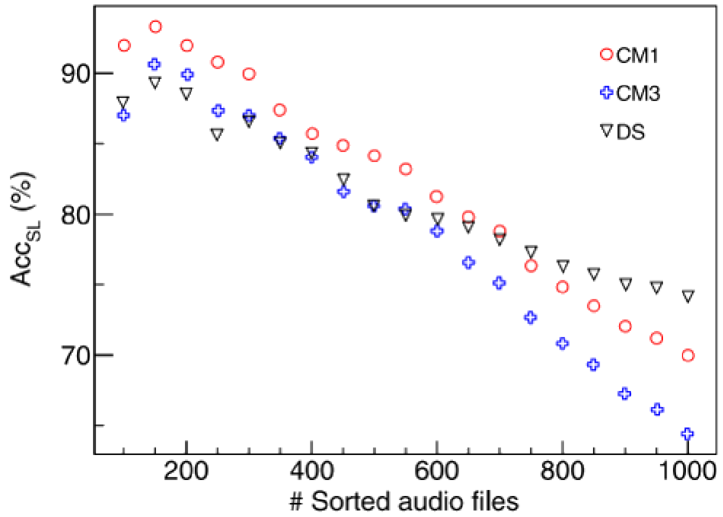
The Classification Model 2 (**CM2**) : e = 70%

The Classification Model 3 (**CM3**) : the simplest of the models. e = 75%

Neuro-genetic programming for multi-genre classification of music content

3. The reason why the problem should be improved in the domain within the paper

- Human agreement regarding the belonging of music content to a particular genre is less than 80%.
- There are many situations in which the classification into more than one single genre is more appropriate for a given song.



Deep Sound(DS) = Computational Neural Networks(CNN) + Long Short-Term Memory(LSTM) Networks

# of features CM1:79 CM3:38 DS: 518

4. The reason why the solution worked effectively in the domain of the paper

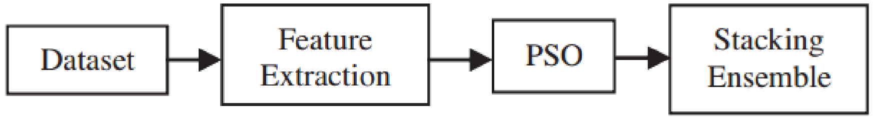
- Reduce the number of features and parameters used for classification(7 statistical values of parameters).
- The classifiers are uniquely based on fully analytical formulas.
- Capable to simultaneously evaluate the degree of belonging of a given music content to several music genres with a fuzzy-like logic, accounting for the existence of fuzzy boundaries between music genres.

Music Genre Classification of audio signals Using Particle Swarm Optimization and Stacking Ensemble

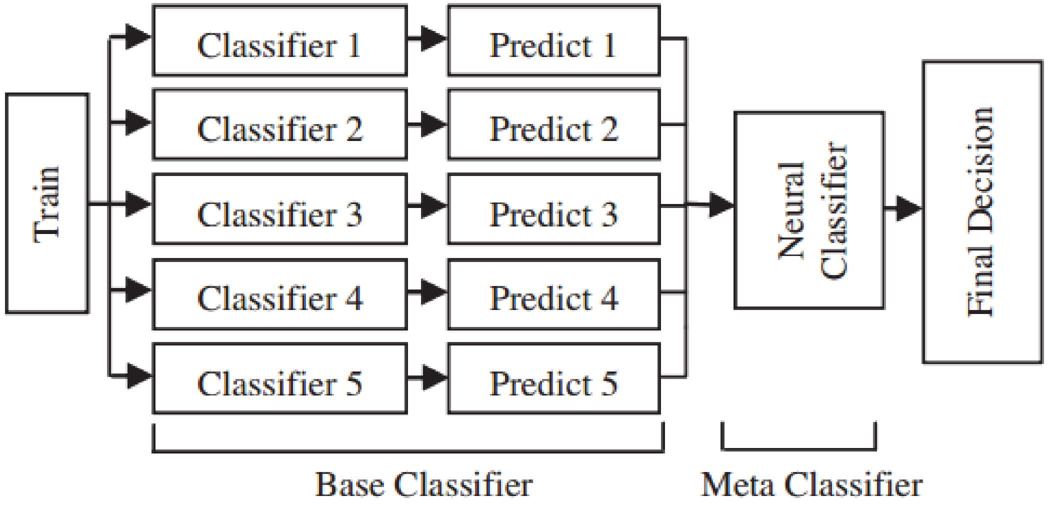
1. GA problems to be solved

- PSO will converge to premature convergence.
- Using all of feature might have duplication, it makes the data confused(less accurate than selecting some features).

2. Solutions



Block diagram of music genre classification



Block diagram of stacking ensemble

Music Genre Classification of audio signals Using Particle Swarm Optimization and Stacking Ensemble

3. The reason why the problem should be improved in the domain within the paper

- Irrelevant or redundant features may even reduce the classification performance.

4. The reason why the solution worked effectively in the domain of the paper

- The population and the number of selections are used primarily via PSO to increase accuracy.
- Results gained from PSO are used in stacking ensemble to improve classification accuracy.
- In stacking ensemble, training model to train the data of each classifier to select base classifier by using k-fold cross validation and prediction from base classifier fed into meta-classifier for creation the new model to predict the final decision.

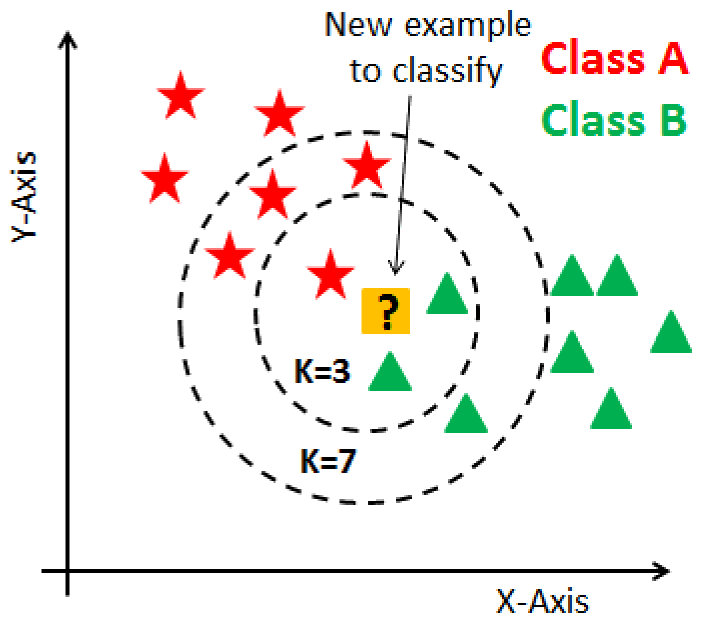
Analysis of Structural Complexity Features for Music Genre Recognition

1. GA problems to be solved

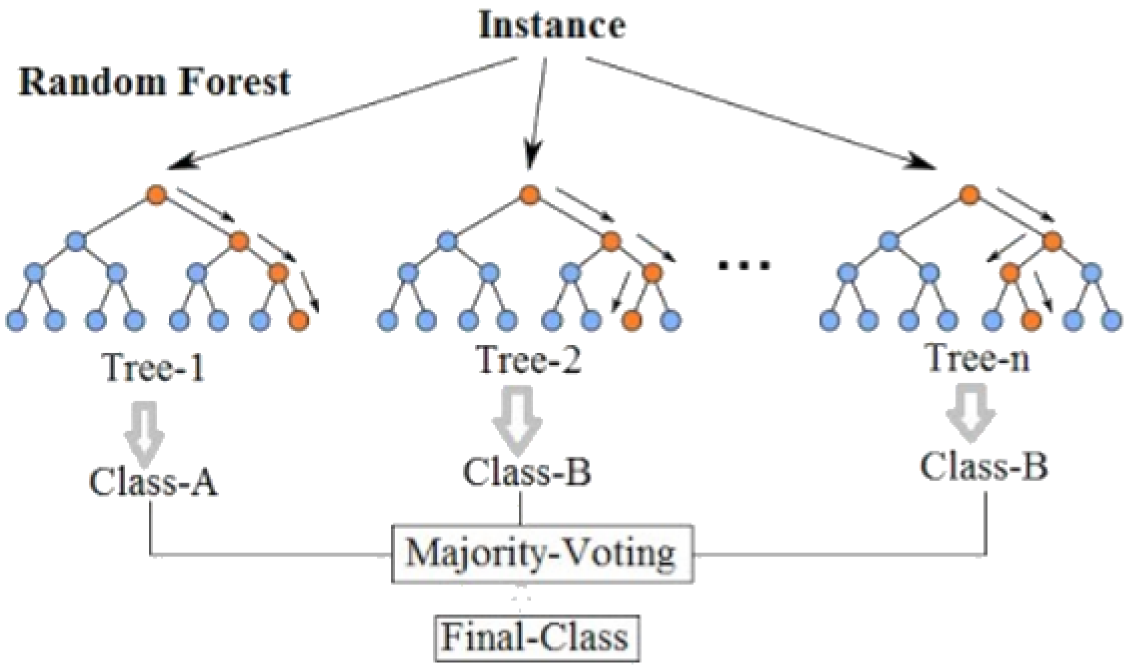
- Genres are hard to be defined into a single label.
- The number of features are too many.

2. Solutions

K-Nearest Neighbor



Random Forest



Analysis of Structural Complexity Features for Music Genre Recognition

3. The reason why the problem should be improved in the domain within the paper

- Classification models built with such descriptors are less interpretable and cannot yield explainable rules why a certain music piece is assigned to a particular genre(e.g. manually engineered audio signal features or deep neural networks).

4. The reason why the solution worked effectively in the domain of the paper

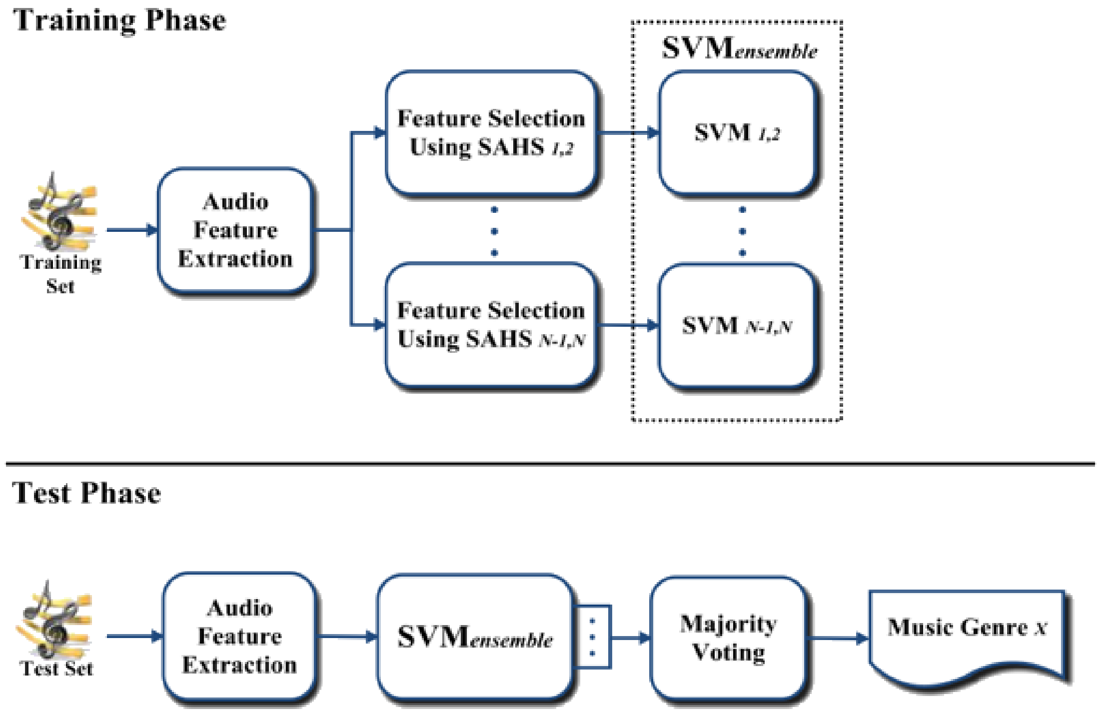
- Implemented a novel feature for chord statistics.
- Using structural complexity, improving performance by using an appropriate time scale according to the feature group.

Music genre classification based on local feature selection using

1. GA problems to be solved

- In general, systems manually classify music using several tags and it is time-consuming.
- Most people have different cognitions of music.

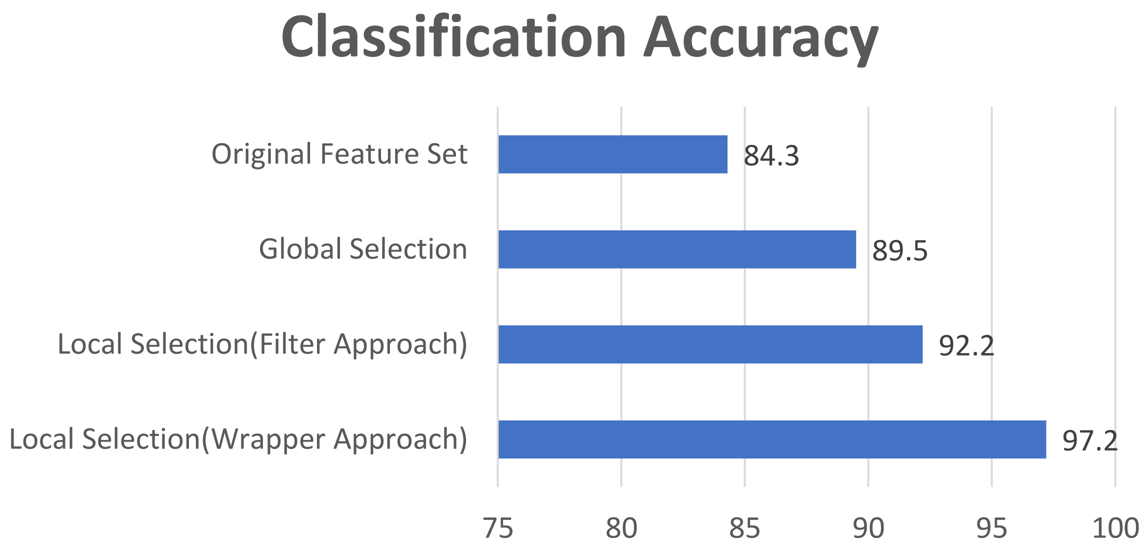
2. Solutions



Music genre classification based on local feature selection using

3. The reason why the problem should be improved in the domain within the paper

- Therefore, developing an automatic music genre-classification system is necessary for the tagging process to be more effective and become standardized.



4. The reason why the solution worked effectively in the domain of the paper

- The parameters of a running multidimensional Gaussian distribution are estimated.
- The mean and standard deviation of all frames in a texture window of 30s are calculated to form two statistical features.
- The mean and standard deviation of all differences between adjacent frames are calculated to form two new features.
- Adopted based on genre pairs instead of a single genre that can be better derived.
- The SVM classifier was adopted because, in general, it demonstrates higher performance than other classifiers do when kernel functions and parameters are appropriately chosen.

Compact feature subset-based multi-label music

1. GA problems to be solved

- The maximum number of features is limited in mobile devices due to low computational capacity and it can neglect important features during the genetic search process.

2. Solutions

- Compact feature subset-based multi-label music categorization

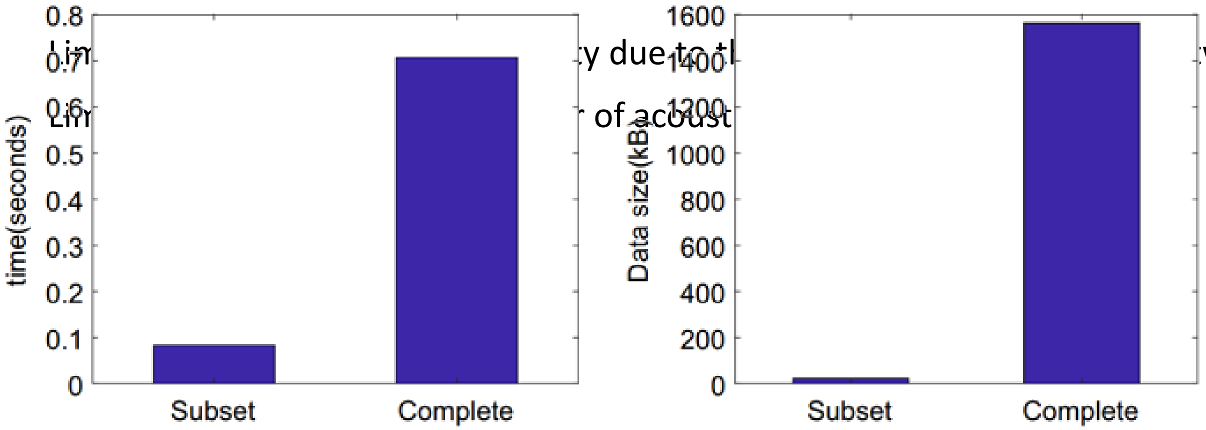
1. $x = \sum_{k=1}^{n} \frac{1}{n} \cdot \frac{\binom{1}{1} \cdot \binom{N-1}{k-1}}{\binom{N}{k}}$, each chromosome randomly selects features less than n

2. Using feature selection based on multi-objective evolutionary algorithm to reduce the number of features.

Compact feature subset-based multi-label music

3. The reason why the problem should be improved in the domain within the paper

- Lim … y due to th … weight mobile devices(usage behavior and battery)
- Lim … of acoust



(a) Average prediction time of 100 music clips in MusicEmo-B

(b) Data size of MusicEmo-B

Comparison results of the average prediction time and data size between the proposed approach and the conventional complete feature set-based approach (Average multi-label accuracy: 0.565 and 0.514, respectively)

4. The reason why the solution worked effectively in the domain of the paper

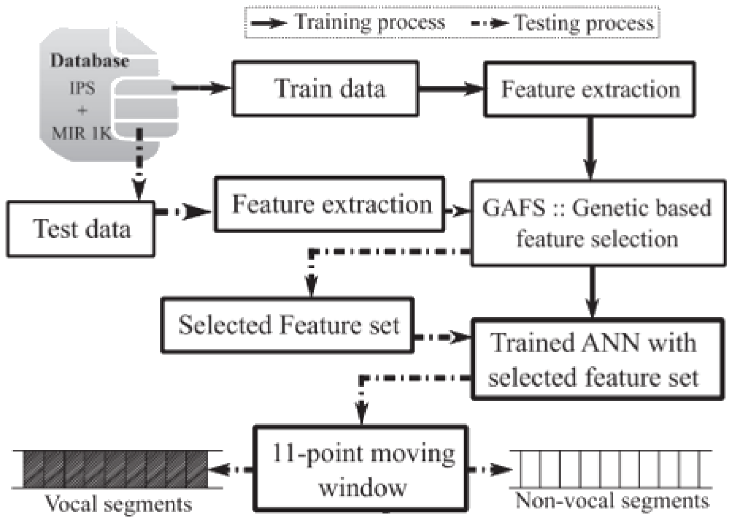
- The cardinality of six datasets is commonly one(allowing the classifier to select only one label for each unseen pattern).
- Music data sets can be associated with multiple concurrent labels, music emotion categorization is modeled as multi-label classification.

Classification of vocal and non-vocal segments in audio clips using

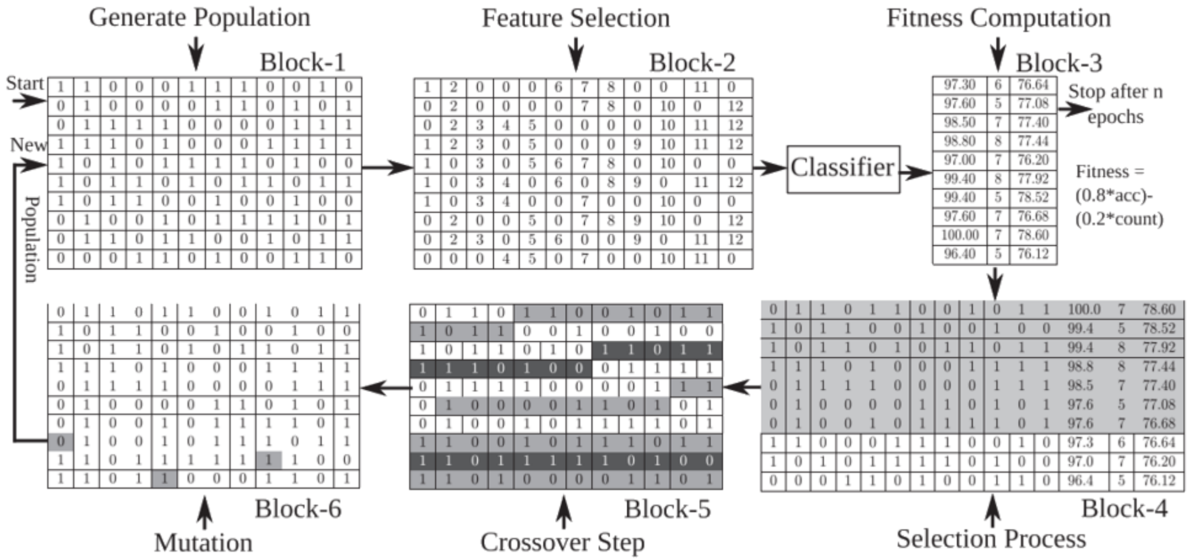
1. GA problems to be solved

- Increased complexity of the task due to the difficulty of vocal and non-vocal segments.

2. Solutions



The proposed flow diagram for vocal and non-vocal segmentation



An example to illustrate the complete process involved in GAFS-ANN

Classification of vocal and non-vocal segments in audio clips using

3. The reason why the problem should be improved in the domain within the paper

- Hard to categorize the digital cloud adequately if proper meta-information is not available.
- Inefficient to manually label a million number of tracks when extracting meta information.
- Music Information Retrieval(MIR) is useful for tagging, but needs an application that can segment vocal and non-vocal.

4. The reason why the solution worked effectively in the domain of the paper

- The standard timbre and temporal features are considered as baseline features to improve the efficiency of the system.
- Several efforts have been made to select the relevant features and reduce the dimensionality of the feature set.(Considering high dimensional feature vectors for real-time applications is not possible)
- It is observed that the evolutionary algorithms are highly efficient when compared to the correlation-based feature selection and wrapper methods.

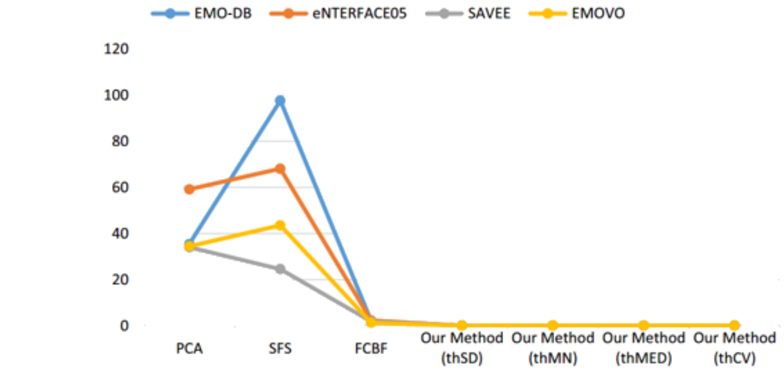
A novel feature selection method for speech emotion recognition

1. GA problems to be solved

- Not all of these features are effective for emotion recognition(different emotions may effect different vocal features).
- Some of existing feature selection methods increase the total workload.

2. Solutions

- Feature selection methods are used to increase the emotional recognition success and reduce workload with fewer features.



	PCA	SFS	FCBF	Our Method (thSD)	Our Method (thMN)	Our Method (thMED)	Our Method (thCV)
EMO-DB	35.42	97.68	1.98	0.014	0.011	0.004	0.014
eNTERFACE05	59.15	68.15	2.23	0.019	0.014	0.015	0.016
SAVEE	33.98	24.52	1.75	0.006	0.005	0.005	0.071
EMOVO	34.43	43.55	1.24	0.007	0.006	0.006	0.008

The workloads for feature selection (second)

Classification successes after feature selection.

Data Set	Classifier	Classification accuracy (%)							
		All	PCA	SFS	FCBF	OM (th_{SD})	OM (th_{MN})	OM (th_{MED})	OM (th_{CV})
EMO-DB	SVM	84.62	↓ 81.71	↓ 82.93	↓ 81.32	↓ 84.07	↓ 83.00	↓ 81.87	↓ 78.57
	MLP	81.32	↓ 60.99	↑ 85.71	↓ 80.77	↑ 82.42	↑ 85.71	↑ 82.97	↑ 82.97
	k-NN	63.74	↓ 30.77	↑ 71.43	↑ 69.78	↑ 68.13	↑ 73.08	↑ 72.53	↑ 71.98
eNTERFACE05	SVM	59.74	↑ 62.31	↓ 49.49	↓ 54.62	↑ 60.51	↑ 60.77	↑ 60.00	↓ 48.72
	MLP	69.23	↓ 49.74	↓ 57.69	↓ 57.69	↓ 67.18	↓ 68.46	↓ 67.95	↓ 48.97
	k-NN	39.74	↓ 23.59	↑ 43.85	↓ 38.46	↑ 41.03	↑ 41.03	↑ 41.79	↓ 33.85
SAVEE	SVM	72.39	↔72.39	↓ 66.26	↓ 55.83	↑ 73.62	↑ 77.92	↑ 74.85	↓ 57.67
	MLP	71.17	↓ 49.69	↓ 65.03	↓ 60.12	↑ 73.62	↔71.17	↔71.17	↓ 54.60
	k-NN	53.37	↓ 22.70	↑ 58.28	↓ 49.69	↑ 57.06	↑ 55.83	↓ 52.76	↓ 47.24
EMOVO	SVM	60.40	↑ 63.91	↓ 51.48	↓ 50.89	↑ 61.54	↓ 59.17	↑ 60.95	↓ 43.20
	MLP	58.58	↓ 42.60	↓ 48.52	↓ 48.52	↔58.58	↓ 56.21	↑ 59.17	↓ 43.79
	k-NN	39.05	↓ 23.67	↑ 59.17	↑ 41.42	↑ 42.60	↑ 46.15	↑ 43.20	↑ 43.79

SD : standard deviation of fc-mean

MN : the mean of fc-mean

MED : the median of fc-mean

CV : the coefficient of variation of fc-mean.

A novel feature selection method for speech emotion recognition

3. The reason why the problem should be improved in the domain within the paper

- The communicative influence of words pronounced in a sentence may differ according to the speakers’ emotions. Therefore, the effects of the emotion on the speech are investigated rather than the words spoken in the Speech Emotion Recognition(SER) studies.

4. The reason why the solution worked effectively in the domain of the paper

- Small feature set of SER reduces the overall workload in real-time applications.
- The success rate can be slightly decreased to reduce the work load.