2022年度電気・電子関係学会北陸支部連合大会

深層学習を用いた音楽感情認識における データ構造の最適化

[xx - x - xx]

◎松波旭（福井大学工学部機械・システム工学科）
黒岩丈介　小高知宏（福井大学大学院工学研究科）
諏訪いずみ（仁愛女子短期大学）白井治彦（福井大学工学部）

はじめに

ユーザー行動履歴による推薦

・感情による推薦

これまでの研究では、精度不十分

∴多様な特徴量入力により精度向上

・各々の特徴量は小サイズのほうが、学習時間が短い

代表的な音楽ストリーミングサービスとその提供曲数

5000万

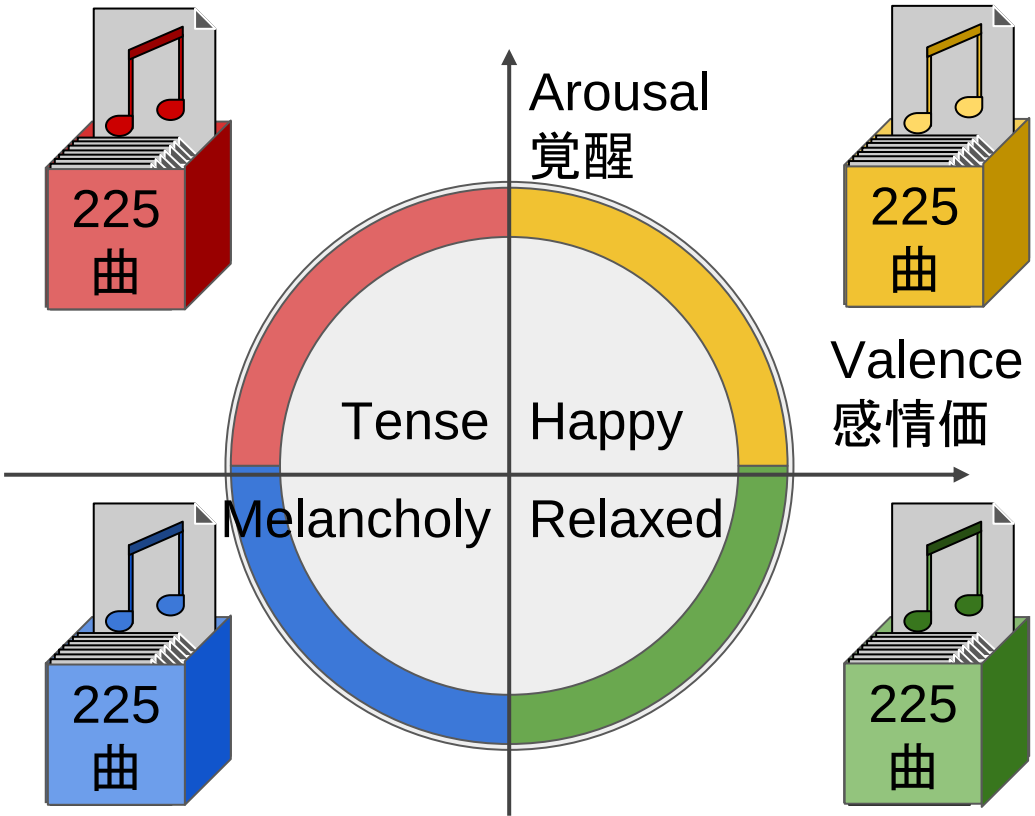
9000万

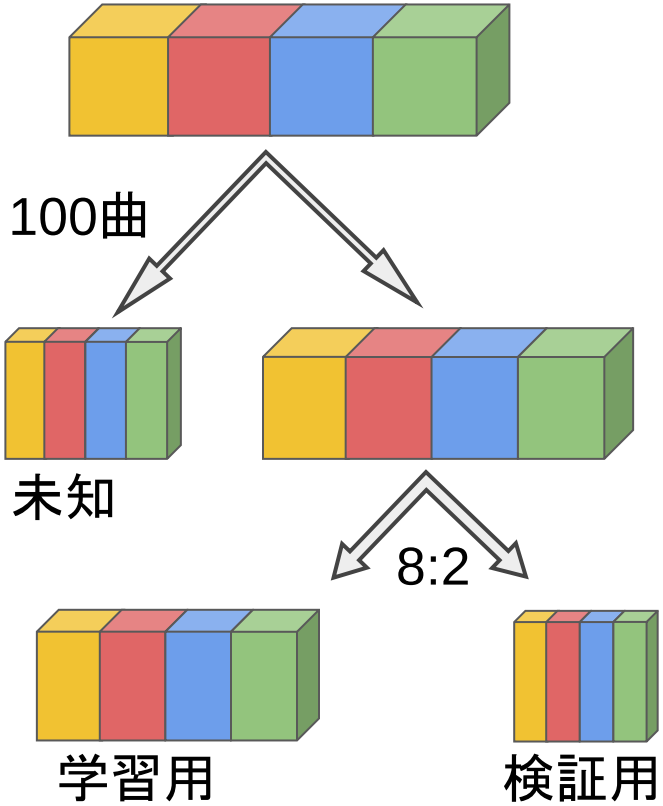
1億

研究目的

スペクトログラムのサイズ削減について考える

データセットの詳細

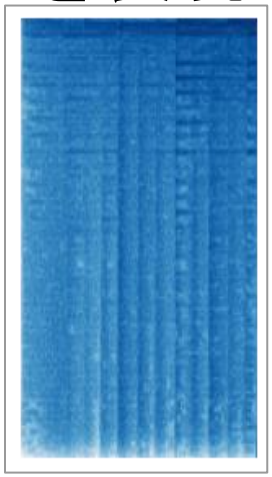

225
曲
Arousal
覚醒
225
曲
Valence
感情価
Tense
Happy
Melancholy
Relaxed
225
曲
225
曲


100曲
未知
8:2
学習用
検証用

特徴量作成

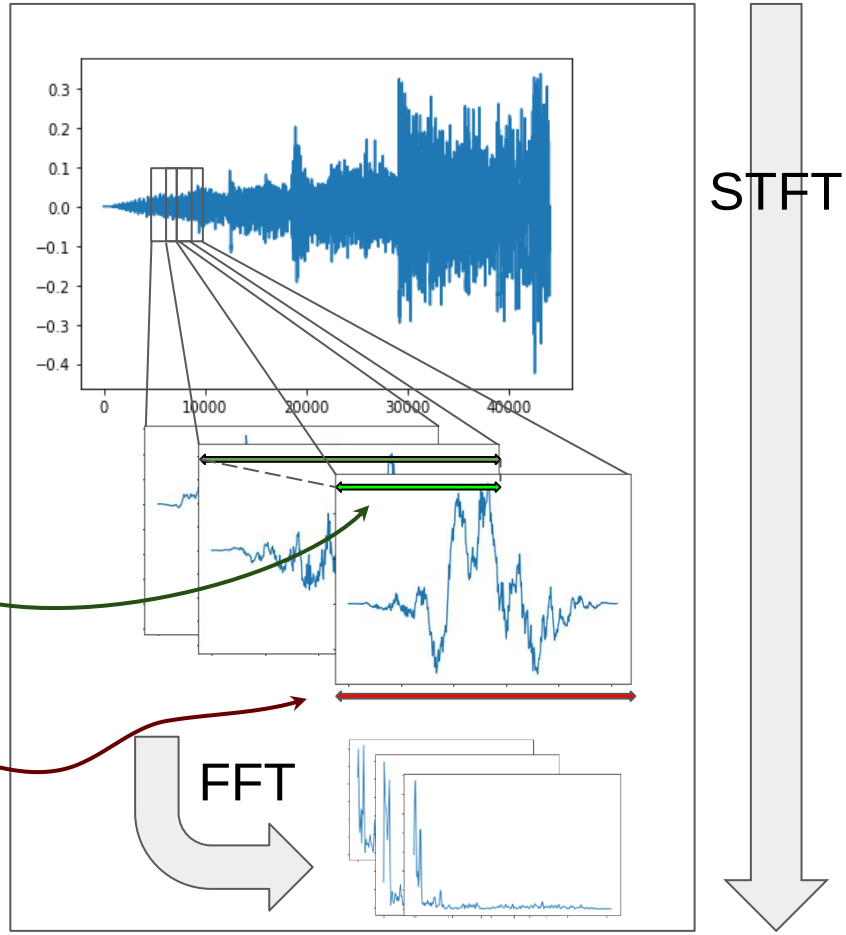
特徴量：スペクトログラム

周波数と音圧の時間変化を表現



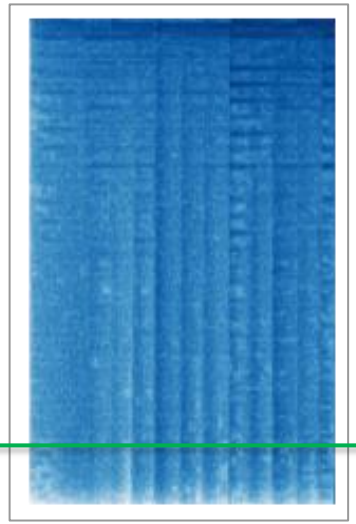
オーバーラップ率	50%
窓サイズ	512

STFT(短時間フーリエ変換)後の二次元配列



データ削減方法

スペクトログラム



閾値

これより下は、
閾値を超えない。

他の時系
列データ

これより上の
データを用いる

モデルについて (Efficient Net V2)

最適化関数	SGD	バッチサイズ	32
学習率	0.0005	epoch数	200

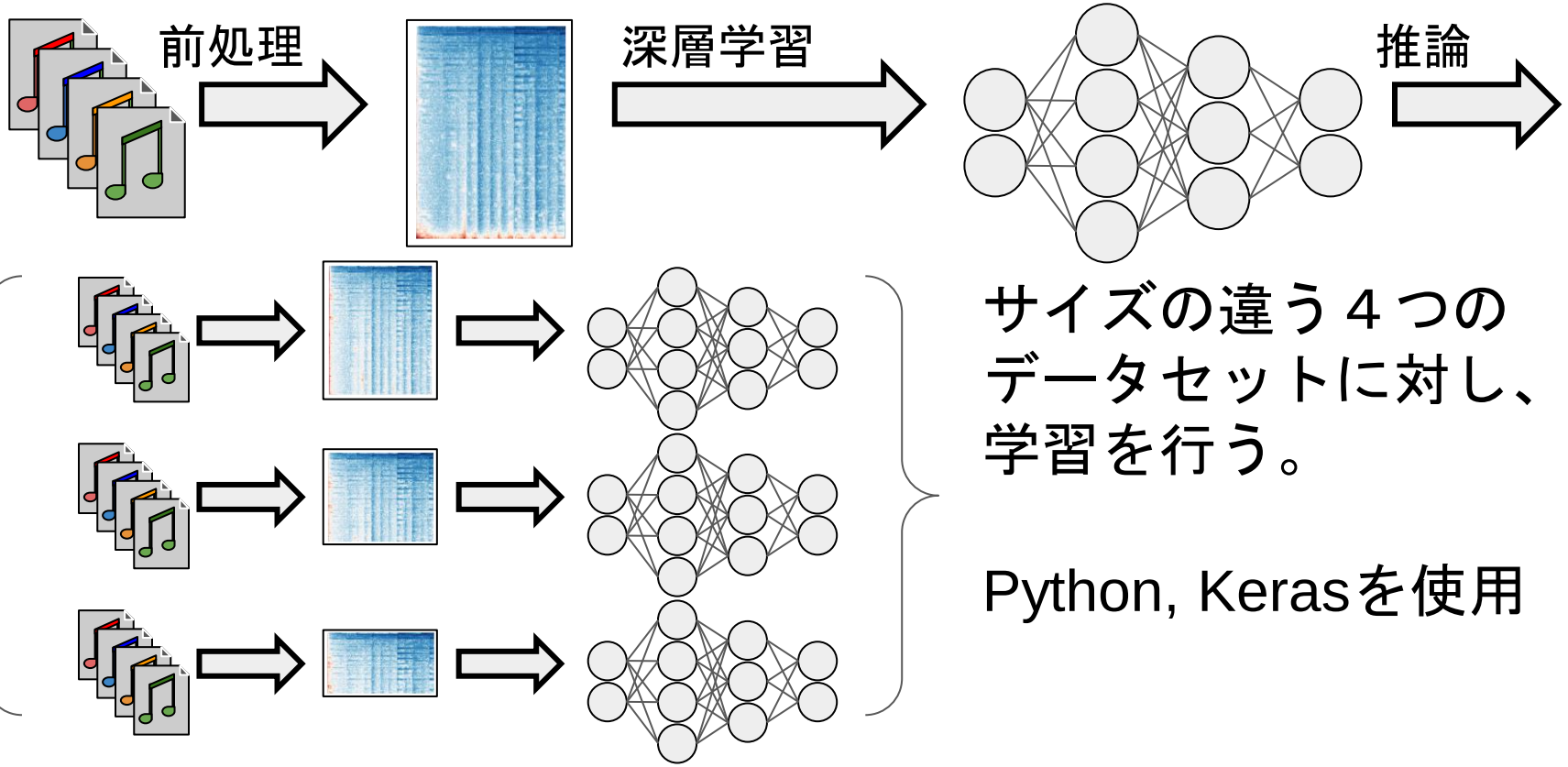
Table 4. EfficientNetV2-S architecture – MBConv and Fused-MBConv blocks are described in Figure 2.

Stage	Operator	Stride	#Channels	#Layers
0	Conv3x3	2	24	1
1	Fused-MBConv1, k3x3	1	24	2
2	Fused-MBConv4, k3x3	2	48	4
3	Fused-MBConv4, k3x3	2	64	4
4	MBConv4, k3x3, SE0.25	2	128	6
5	MBConv6, k3x3, SE0.25	1	160	9
6	MBConv6, k3x3, SE0.25	2	272	15
7	Conv1x1 & Pooling & FC	-	1792	1

・画像認識タスクで
高精度なモデル
・Efficient Net より、
小パラメータ量
小トレーニング時間

引用元：EfficientNetV2: Smaller Models and Faster Training Mingxing Tan Quoc V. Le

感情推定精度とデータ量比較実験の手法



サイズの違う４つのデータセットに対し、学習を行う。

Python, Kerasを使用

分類結果の評価指標

Accuracyの算出方法

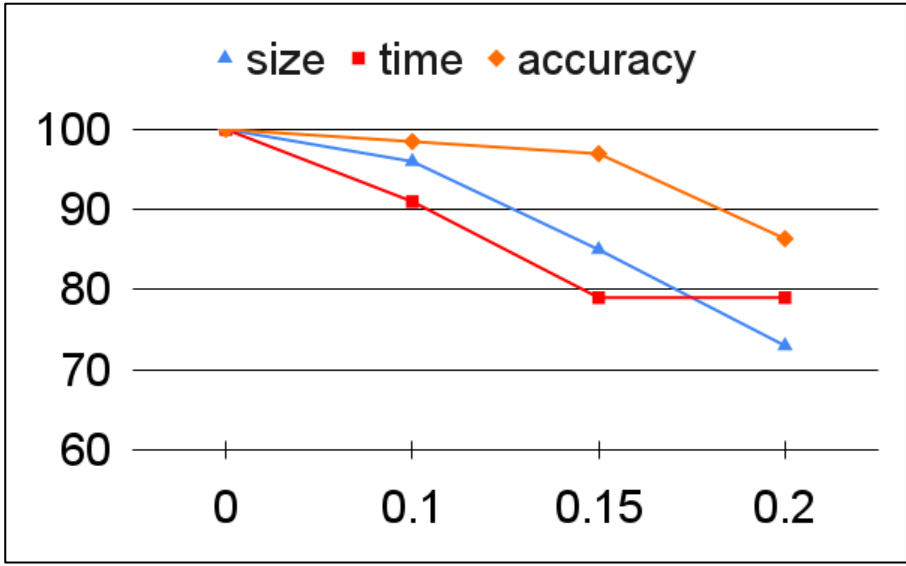
教師 / 予測	Q1	Q2	Q3	Q4
Q1	$\boldsymbol{a_{11}}$	a_{12}	a_{13}	a_{14}
Q2	a_{21}	$\boldsymbol{a_{22}}$	a_{23}	a_{24}
Q3	a_{31}	a_{32}	$\boldsymbol{a_{33}}$	a_{34}
Q4	a_{41}	a_{42}	a_{43}	$\boldsymbol{a_{44}}$

各セル (a_{ij})には、
データ数を表す自然数

$$\text{Accuracy} = \frac{a_{11} + a_{22} + a_{33} + a_{44}}{\text{全データ数}}$$

結果

threshold	size [KiB (%)]	time [min (%)]	accuracy
0	42489 (100%)	126.9 (100%)	0.66
0.1	41134 (96%)	116.0 (91%)	0.65
0.15	36351 (85%)	100.6 (79%)	0.64
0.2	31428 (73%)	100.0 (79%)	0.57

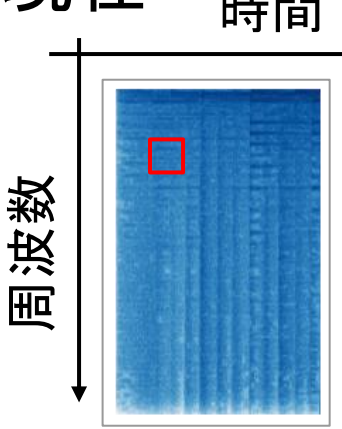


考察

・データ量を削減すると効率的に学習できる。
・データ削減による効率化は閾値0.15が限界

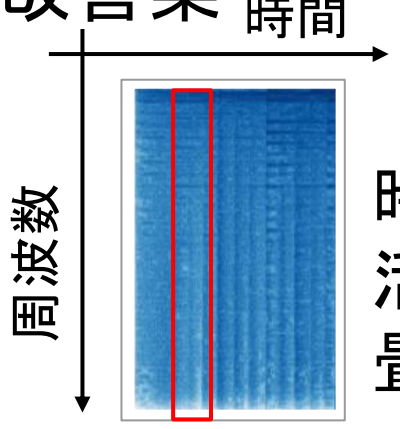
モデル構造を変更するのが良いのではないか。

現在



カーネル
3×3 の
畳み込み

改善案



時系列を
活かした
畳み込み

おわりに

まとめ

・データ削減を行って学習した際の効率を調べた。

・閾値0.2以外、データ削減によって効率化できた。

今後

・データを複数入力できるようなモデルの実現

・モデルの構造の変更