

楽曲生成 AI による日本語曲雰囲気推定のための データセット構築とその分析

渡邊 大輝^{1,2,a)} 塩見 昌裕¹ 滝口 哲也²

概要: 近年、音楽情報検索 (Music Information Retrieval: MIR) 分野では、楽曲の雰囲気・ジャンル・感情を推定する研究が注目されている。特に、Valence-Arousal (V/A) 平面を用いた楽曲雰囲気推定は音楽感情認識 (Music Emotion Recognition: MER) の代表的な枠組みとして広く用いられている。一方、日本語楽曲を対象とした雰囲気推定用データセットは依然として限られており、十分な検討が行われていない。本研究では、日本語楽曲の雰囲気推定を目的として、楽曲生成 AI を用いた Valence-Arousal アノテーション付き楽曲データセットを構築した。構築したデータセットに対し、音楽基盤モデルである MERT を用いた雰囲気推定実験を行い、既存データセットである EMO および DEAM との比較を行った。その結果、本データセットでは Arousal 推定において比較的良好な性能が得られた一方で、Valence 推定においては性能が低下する傾向にあった。また、評価値の分布を可視化した結果、Valence および Arousal がともに低い第三象限のデータが不足していること、および評価者ごとの評価傾向の違いが存在することが確認された。本研究は生成 AI 楽曲を用いた日本語楽曲雰囲気データセット構築の可能性を示すとともに、データ分布や主観評価に起因する課題を明らかにするものである。

1. はじめに

Peeters ら [1] および Casey ら [2] により、音楽情報検索の分野では音楽信号から高次の意味情報を推定する研究が進展していることが整理されている。その一つとして楽曲の雰囲気や感情を推定する研究が注目されている。楽曲の雰囲気や感情を高精度に推定することは、楽曲推薦やプレイリスト自動生成、さらには楽曲生成 AI における条件制御など、音楽情報処理技術の高度化に寄与すると期待されている。

楽曲雰囲気推定において、雰囲気を定量化する代表的な手法として、Russell ら [3] によって提案された Valence-Arousal (V/A) 平面が挙げられる。Han ら [4] によると、近年の音楽感情認識研究では、従来の音響特徴量と機械学習手法に基づくアプローチに加え、深層学習を用いた手法が提案された。さらに、大規模音楽データを用いて事前学習された音楽基盤モデルを下流タスクに応用する研究が進展している。このような事前学習モデルは、音楽感情認識研究の分野で高い表現能力を有する。

一方で、楽曲の雰囲気推定に用いられてきた既存データセットの多くは英語圏の楽曲を対象としており、日本語楽

曲を対象とした Valence-Arousal アノテーション付きデータセットは依然として限られている。日本語楽曲を対象とした例として、大野 ら [5] によって VOCALOID 楽曲を含む音楽動画への印象評価データセットが報告されているが、同データセットでは著作権保護の観点から音源データのオープンな配布が行われておらず、研究者各自による個別取得を前提としている。このため、機械学習の学習データとしての再配布や実験の再現性を担保した利用には制約がある。

また、Hu ら [6] によると言語や文化的背景の違いが楽曲の感情知覚に影響を与えることは先行研究において指摘されており、特に Valence 推定においてはデータセット間や文化間での汎化が難しいことが報告されている。さらには、文化の違いを超えて利用可能な楽曲生成 AI の発展も進んでおり、大量の楽曲を自動生成することが可能となっているが、生成楽曲を対象とした雰囲気分析や、そのデータ特性に関する検証は十分に行われていない。このような背景から、日本語楽曲および生成 AI 楽曲を対象とした雰囲気データセットの構築と評価は、音楽感情認識分野において重要な課題の一つであると考えられる。

本研究では、日本語楽曲の雰囲気分析を目的として、楽曲生成 AI を用いた Valence-Arousal アノテーション付き楽曲データセットを構築する。具体的には、楽曲生成 AI に

¹ 国際電気通信基礎技術研究所: ATR

² 神戸大学

^{a)} 2215016t@atr.jp

よって生成した日本語楽曲に対し、専用のアノテーション用アプリケーションを用いて主観評価データを収集した。さらに、構築したデータセットを用いて、Li ら [7] によって提案された音楽基盤モデルである Music Representation Transformer(MERT) による雰囲気推定実験を行い、既存データセットとの比較およびデータ分布や評価者間の差、データ数依存性の分析を通して、本データセットの特性を明らかにする。

2. 関連研究

2.1 楽曲の雰囲気の定量化

楽曲の雰囲気を定量化する方法はとして、Xiao ら [8] が提案した、楽曲の雰囲気を多次元意味空間として整理し、PCA / クラスタリングによる定量化、Plewa ら [9] が提案したクラスタ・トポロジーベースの Self-Organizing Map(SOM) など、様々なアプローチがある。本研究では、Russell ら [3] が提案した Valence-Arousal(V/A) 平面を採用する。Valence は感情の快・不快を、Arousal は覚醒・沈静の度合いを表す指標である。この提案手法はもともと「人間の感情モデル」であったが、近年では、楽曲の雰囲気推定にも広く用いられている。

2.2 楽曲の雰囲気推定モデル

楽曲の雰囲気推定モデルは音楽感情認識の分野で、近年特に研究されている。前章で述べた通り、雰囲気推定モデルは、音響特徴・機械学習から始まり、Yang ら [10] は楽曲の MFCC・スペクトル・リズムなどの音響特徴を SVM などの機械学習モデルで推定した。続いて、深層学習モデルが提案され、Xin ら [11] のスペクトログラム入力の CNN、Choi ら [12] の時間構造を扱う CNN + RNN/LSTM、Chaki ら [13] をはじめとする Attention が主流となった。そして近年、雰囲気推定モデルの転換点として、事前学習モデルである MERT が提案された。MERT 本来は雰囲気推定モデルではなく、音楽理解のための「基盤表現モデル」であり、V/A 推定の下流タスク用の回帰器として Multi-Layer Perceptron(MLP) を学習させ、推定する。

Arousal 推定は比較的、高性能を示す一方、Valence 推定は依然として性能が低いという課題が挙げられる。

2.3 既存データセット

本研究では、MER の分野で広く用いられている既存データセットとして、Soleymani ら [14] による EMO、および Aljanaki ら [15] による DEAM を採用する。これらのデータセットは各 45 秒の楽曲データに対して、Valence-Arousal のラベルがアノテーションされている。また、Zhang ら [16] が提案した PMemo は、複数評価者による Valence-Arousal アノテーションを特徴とする代表的な MER データセットであるが、各楽曲クリップはコーラス部分を中心とした可

表 1: Suno AI による楽曲生成条件

項目	設定
使用楽曲数	282 曲
生成日時	2025 年 11 月上旬
モデル	Suno AI v5
利用プラン	有料プラン (Premier)
生成方法	テキストプロンプト入力 (Simple)
プロンプト形式	固定テンプレート
生成楽曲長	約 2:00
利用目的	研究目的

変長区間で構成されている。

これらのデータセットはデータ量、質、ともに十分であるが、やはり英語圏を対象としており、文化差の汎化が課題として挙げられる。

3. データセット構築

今回構築するデータセットは、Valence-Arousal 回帰のための Valence-Arousal アノテーション付き楽曲群である。EMO と DEAM と作成手順を模倣し、各 45 秒の楽曲に対して Valence, Arousal をアノテーションした。EMO と DEAM との差分は、EMO と DEAM は 1 つの楽曲に対してランダムに 1 部分だけ切り抜いているのに対し、我々が構築したデータセットは、1 つの楽曲を 45 秒で切り抜けるだけ楽曲前半から順に複数切り抜いている点である。

さらに、EMO および DEAM と同一条件での比較を行うため、45 秒固定長クリップを前提とするデータセットのみを使用し、PMemo は比較対象から除外した。

以降、楽曲生成に使用した生成 AI、楽曲生成に関わる雰囲気語・軸、アノテーション用アプリケーション、雰囲気評価の手続きを述べる。

3.1 楽曲生成 AI

データセット構築において、楽曲生成 AI 「Suno AI」を用いた。「Suno AI」は商用の楽曲生成 AI であり、本研究ではその内部構造には立ち入らず、ブラックボックスとして利用するとともに、研究目的に限り利用した。

本研究で用いた「Suno AI」による楽曲生成条件を表 1 に示し、楽曲生成に用いたテキストプロンプトを表 2 に示す。

生成させた楽曲に対する選別基準として、生成条件と音楽的特性が一致しない楽曲を除外した。また、「Suno AI」は 2 曲ずつ生成する仕組みのため、生成条件と音楽的特性が一致する場合は、2 曲のうち 1 曲をランダムに選別した。

3.2 雰囲気軸と雰囲気語

本研究では、楽曲生成においても Russell らの円環モデルを参考にした。図 1 に示す。図 1 より、V/A 平面には各象限ごとに雰囲気語が分布していることが分かる。楽曲生成の際、この雰囲気語をプロンプトに組み込み楽曲生成

表 2: Suno AI によるプロンプト条件

項目	指定内容
言語	日本語
歌詞生成	自動生成
曲調	和風要素を含まない
ボーカル	男/女
雰囲気	指定された雰囲気に準拠

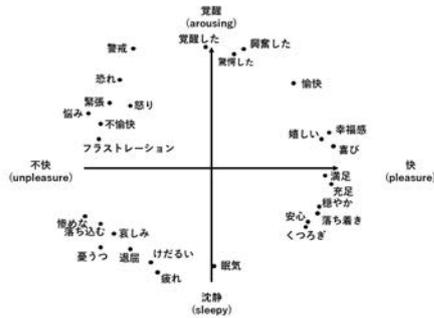


図 1: Russell の円環モデルを基に作成した Valence-Arousal 平面

表 3: 雰囲気語とデータ数の関係

象限	データ数
第一象限	54
第二象限	70
第三象限	70
第四象限	88

した. 各象限の雰囲気語とデータ数の関係は, 雰囲気語一つに対して, 2曲採用し, さらに各象限内の雰囲気語の組み合わせで楽曲生成した. また, 男/女ボーカルを均等にするために, 同じ条件で楽曲を生成した. 雰囲気語の数とデータ数の関係を表3に示す.

3.3 アノテーション用アプリケーション

本研究では, 作成したデータセットに Valence-Arousal のアノテーションを行うため, Betella らによるアフェクティブスライダー [17] を基にアプリケーションを作成した. アノテーション用アプリケーションの画面を図2に示す. EMO および DEAM の評価手続きに合わせるため, 各楽曲に対して Valence-Arousal を 1.00~9.00 の範囲で評価可能な設計とした.

3.4 評価データの収集

2025年11月14日から, 作成した楽曲の評価データを収集した. 2025年12月22日現在も引き続き, データ収集中である. アノテーション協力者は株式会社 国際電気通信基礎技術研究所 (ATR) の2人の職員であった. 作成した282曲を無音区間を含まず, 重複しないように45秒ずつ曲の冒頭から切り取り, 計801曲を作成. その中の500曲

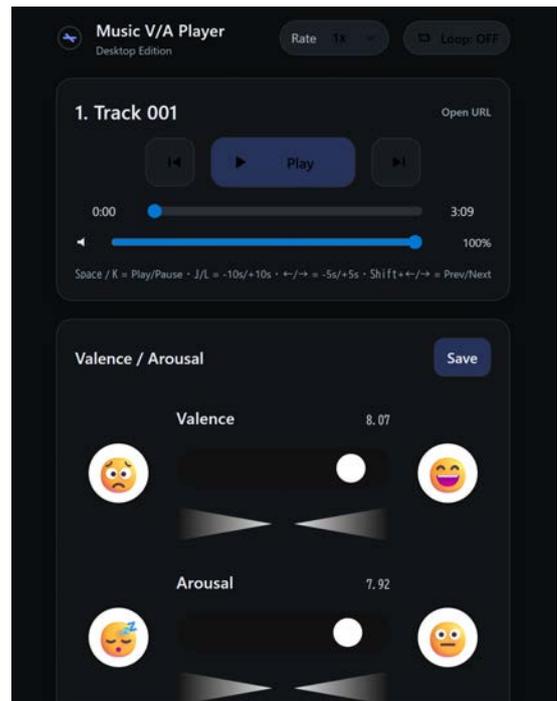


図 2: Valence-Arousal アノテーション用アプリケーション画面

をランダムに選び, 評価データの対象とした. この500曲に対し, 3.3節で述べたアプリケーションで各曲に対して, Valence-Arousal を 1.00~9.00 の範囲でアノテーションを行い, データセットを構築した.

また, EMO および DEAM では, 複数の評価者による主観評価を標準化し, その平均値を Valence-Arousal のラベルとして用いている. 本研究で構築したデータセットにおいても, 既存データセットとの比較を可能にするため, 同様の手続きに基づきアノテーション値を標準化した.

4. 実験

4.1 実験目的

本実験の目的は, 本研究で構築した日本語楽曲雰囲気データセットの有効性を検証することである. 具体的には, 2.3節で述べた既存データセットである EMO および DEAM と比較し, 同一の雰囲気推定モデルを用いた場合の推定性能を評価する.

4.2 使用データセット

実験には, EMO, DEAM, および本研究で構築したデータセットの3種類を用いた. 各データセットはいずれも45秒の楽曲クリップに対して Valence-Arousal のアノテーションが付与されている.

各データセット間で楽曲数は異なるが, 本実験ではデータ数の差を補正する処理は行わず, 各データセットが本来有する規模と分布を保持したまま評価を行った. これは, データセット固有の特性が雰囲気推定性能に与える影響を

分析することを目的としているためである。

また、EMO, DEAM, および本研究で構築したデータセットは、いずれも評価者ごとの主観評価を標準化した値を Valence-Arousal の正解ラベルとして用いている。このため、異なるデータセット間においても同一スケールでの性能比較が可能である。

4.3 実験設定

雰囲気推定モデルとして、2.2 節で述べた事前学習済み音楽基盤モデル MERT を用いた。MERT は音楽信号から汎用的な音楽表現を抽出するモデルであり、本実験では特徴抽出器として固定し、その出力を入力とする MLP を回帰器として学習した。MERT のパラメータは更新せず、すべてのデータセットに対して同一のネットワーク構造および学習条件を適用した。

MLP は Valence および Arousal を同時に推定する回帰モデルとし、学習および評価の手順は、全データセットに対して同一とした。

4.4 評価指標

推定性能の評価には、決定係数 R^2 を用いた。 R^2 は予測値と正解値の相関を表す指標であり、データ数の違いによる影響を受けにくいことから、データセット間比較に適した指標である。

Valence および Arousal それぞれに対する R^2 に加え、両者をまとめた全体的な推定性能を評価するため、total R^2 を算出した。

total R^2 は、Valence および Arousal の予測値と正解値を結合し、2次元回帰問題として算出した決定係数である。

5. 結果

5.1 Valence-Arousal 推定性能

各データセットに対する Valence および Arousal の推定結果を決定係数 R^2 により評価した。結果を表 4 に示す。表 4 より、データセット間で推定性能に差が確認された。全体として EMO が最も高い性能を示した。また、DEAM は Valence $R^2 = 0.4329$, Arousal $R^2 = 0.6119$, total $R^2 = 0.5224$ であった。一方、本データセットは Valence $R^2 = 0.2078$, Arousal $R^2 = 0.5695$, total $R^2 = 0.3886$ であった。

DEAM と本データセットを比較すると、Arousal の R^2 は近い値を示した一方で、Valence および total の R^2 は本データセットの方が低い結果となった。

5.2 追加実験: 評価者間差およびデータ数依存性

本節では、本研究で構築したデータセットに対して、評価者ごとの推定結果の違い、および学習に用いるデータ数の違いが雰囲気推定性能に与える影響を分析する。

表 4: 各データセットにおける Valence / Arousal 推定性能 (R^2)

データセット	Valence R^2	Arousal R^2	total R^2
EMO	0.5490	0.7612	0.6551
DEAM	0.4329	0.6119	0.5224
本データセット	0.2078	0.5695	0.3886

表 5: 本データセットにおける評価者別推定性能 (R^2)

評価者	Valence R^2	Arousal R^2	total R^2
評価者 A	0.0294	0.0868	0.0580
評価者 B	0.2023	0.7163	0.4592

表 6: データ数の違いによる Valence-Arousal 推定性能の変化 (R^2)

データ数	Valence R^2	Arousal R^2	total R^2
200	-0.1178	0.2086	0.0484
500	0.2078	0.5695	0.3886

5.2.1 評価者間差

評価者間の評価性能の差異を表 5 に示す。ここでも、各評価者の正解ラベルは z-score 標準化しており、使用データセットのデータ数は 500 とした。

表 5 より、評価者間で Arousal は 0.6295, Valence は 0.1729 の性能差が見られる。評価者 A の推定性能は 0 近傍であるのに対し、評価者 B の Arousal は EMO との差が 0.0449 であった。

5.2.2 データ数依存性

データ数の違いによる Valence-Arousal 推定性能の変化を表 6 に示す。

データ数を 200 曲から 500 曲に増加させることで、Valence, Arousal, および total R^2 のいずれにおいても、推定性能が 0.3 ほど向上する傾向が確認された。

5.3 評価値の分布

本研究で構築したデータセットの Valence-Arousal 評価値について、正規化後の散布図 (raw plot) を図 3 に示す。図 3(a) が評価者 A, 図 3(b) が評価者 B による結果である。

両評価者の raw plot より、Valence と Arousal がともに低い第三象限に対応する点はほとんど見られず、第三象限のデータ不足が視覚的に確認できる。

6. 考察と分析

本章では、第 5 章で得られた実験結果および追加実験の結果を踏まえ、本研究で構築したデータセットの特性とその限界について考察する。特に、Valence-Arousal 平面上のデータ分布、評価者間のアノテーション特性、およびデータ数が推定性能に与える影響に着目する。

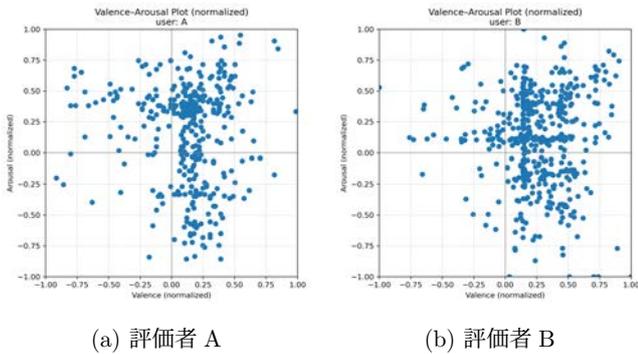


図 3: 評価者別 Valence-Arousal 評価値の分布 (正規化後)

6.1 Valence-Arousal 平面におけるデータ分布の偏り

評価値の分布を可視化した結果 (図 3) より, 本研究で構築したデータセットでは, Valence および Arousal がともに低い第三象限に該当するデータがほとんど存在しないことが確認された。

この分布の偏りは, 楽曲生成に用いた生成 AI の特性や, プロンプト設計に起因している可能性がある。特に, 第三象限に対応する「沈静」「抑鬱」といった雰囲気は, 音楽としての成立性や聴取体験の観点から, 生成されにくい傾向があると考えられる。実際 EMO の第三象限のデータを確認すると, 不協和音, バックグラウンドノイズなどが曲全体で含まれ, 全体として, 不快感を演出する音源であった。

このような分布の偏りは, Valence 推定性能が相対的に低下した要因の一つである可能性がある。

6.2 評価者間のアノテーション特性

評価者別の評価値分布 (図 3) より, 評価者 A では特定の評価値にデータが集中し, 評価値が離散的に付与される傾向が確認された。一方, 評価者 B では, より連続的な分布が観測された。

この差異は, 評価者ごとの主観的判断基準や評価スタイルの違いを反映していると考えられる。主観評価に基づくデータセット構築において, このような評価者間差は避けがたい課題である。

本研究では, 評価者ごとに z-score 標準化を適用することで, 評価スケールの違いによる影響を低減した。しかしながら, 評価値の分解能そのものに起因する差異は完全には解消されておらず, これが推定性能に影響を与えている可能性がある。

6.3 データ数と推定性能の関係

データ数依存性の実験より, 学習に用いるデータ数を 200 曲から 500 曲に増加させることで, Valence, Arousal, および total R^2 のいずれにおいても推定性能が向上する傾向が確認された。

この結果は, 本研究で構築したデータセットが, データ数

の増加に伴い汎化性能が向上し, 雰囲気推定に有用な情報をより多く提供できることを示唆している。一方で, 現時点ではデータ数が限定的であり, 特に Valence 推定においては, 十分な汎化性能を確保できていない可能性がある。

6.4 本研究の限界と今後の課題

本研究の限界として, Valence-Arousal 平面上の分布の偏り, 評価者数の少なさ, およびデータ数の制約が挙げられる。

今後は, 第三象限に対応する雰囲気を意図的に生成・収集することで, データ分布の均衡を図る必要がある。また評価者を増やし, 評価者間一致度の分析を行うことで, アノテーションの信頼性向上を目指す。

7. まとめ

本研究では, 日本語楽曲の雰囲気推定を目的として, Valence-Arousal アノテーション付き楽曲データセットの構築と, その評価を行った。

具体的には, 楽曲生成 AI を用いて日本語楽曲を生成し, アノテーション用アプリケーションを設計することで, Valence-Arousal 評価データを収集した。また, 構築したデータセットに対して, 音楽基盤モデルである MERT を用いた雰囲気推定実験を行い, 既存データセットである EMO および DEAM との比較を行った。実験の結果, EMO が最も高い推定性能を示した一方で, 本研究で構築したデータセットでは, Arousal 推定において比較的良好な性能が得られることが確認された。一方, Valence 推定は相対的に低い傾向が示され, データ分布の偏りや主観評価の影響を受けやすいことが示唆された。さらに, 評価値の分布を可視化した結果, Valence および Arousal がともに低い第三象限に該当するデータがほとんど存在しないことが明らかとなった。また, 評価者ごとの評価傾向に差が存在し, 主観評価に基づくデータセット構築における課題が確認された。これらの結果より, 生成 AI 楽曲を用いた雰囲気データセット構築の可能性とともに, データ分布や評価者差といった課題を定量的・視覚的に示すことができた。今後の課題として, 第三象限に対応する楽曲の意図的な生成・収集, 評価者数およびデータ数の拡張の検討が挙げられる。

謝辞

本研究は JSPS 科研費 (JP25K00838, JP25H01175, JP24K21327) および JST ムーンショット型研究開発事業 JPMJMS2011 の助成を受けたものです。本研究の実施にあたり, 楽曲アノテーション (評価実験) にご協力いただいた皆様は心より感謝申し上げます。また, 研究遂行に関して有益な助言をいただいた関係者の皆様に感謝いたします。

参考文献

- [1] Peeters, G., Raffi, Z., Fuentes, M., Duan, Z., Benetos, E., Nam, J. and Mitsufuji, Y.: Twenty-Five Years of MIR Research: Achievements, Practices, Evaluations, and Future Challenges, *ICASSP 2025 - 2025 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, IEEE, p. 1–5 (online), DOI: 10.1109/icassp49660.2025.10888947 (2025).
- [2] Schedl, M., Gómez, E. and Urbano, J.: Music Information Retrieval: Recent Developments and Applications, *Foundations and Trends® in Information Retrieval*, Vol. 8, No. 2-3, pp. 127–261 (online), DOI: 10.1561/15000000042 (2014).
- [3] Russell, J. A.: A Circumplex Model of Affect, *Journal of Personality and Social Psychology*, Vol. 39, No. 6, pp. 1161–1178 (1980).
- [4] Han, D., Kong, Y., Han, J. and Wang, G.: A Survey of Music Emotion Recognition, *Frontiers of Computer Science*, Vol. 16, No. 6, p. 166335 (online), DOI: 10.1007/s11704-021-0569-4 (2022).
- [5] Ohno, N., Nakamura, S., Yamamoto, T. and Goto, M.: Construction of an Impression Evaluation Dataset for Music Videos and Analysis of Its Characteristics, Technical Report 7, Information Processing Society of Japan (2015).
- [6] Hu, X. and Yang, Y.-H.: Cross-Dataset and Cross-Cultural Music Mood Prediction: A Case on Western and Chinese Pop Songs, *IEEE Transactions on Affective Computing*, Vol. 8, No. 2, pp. 228–240 (online), DOI: 10.1109/TAFFC.2016.2523503 (2017).
- [7] Li, Y., Yuan, R., Zhang, G., Ma, Y., Chen, X., Yin, H., Xiao, C., Lin, C., Ragni, A., Benetos, E., Gyenge, N., Dannenberg, R., Liu, R., Chen, W., Xia, G., Shi, Y., Huang, W., Wang, Z., Guo, Y. and Fu, J.: MERT: Acoustic Music Understanding Model with Large-Scale Self-Supervised Training, *International Conference on Learning Representations (ICLR)* (2024).
- [8] Hu, X. and Downie, J.: Exploring mood metadata: Relationships with genre, artist and usage metadata, *Proceedings of the 8th International Conference on Music Information Retrieval, ISMIR 2007*, Proceedings of the 8th International Conference on Music Information Retrieval, ISMIR 2007, pp. 67–72 (2007). 8th International Conference on Music Information Retrieval, ISMIR 2007 ; Conference date: 23-09-2007 Through 27-09-2007.
- [9] Plewa, M. and Kostek, B.: Music Mood Visualization Using Self-Organizing Maps, *Archives of Acoustics*, Vol. 40, No. 4, pp. 513–525 (online), DOI: 10.1515/aoa-2015-0051 (2015).
- [10] Yang, Y.-H., Lin, Y.-C., Su, Y.-F. and Chen, H. H.: A Regression Approach to Music Emotion Recognition, *IEEE Transactions on Audio, Speech, and Language Processing*, Vol. 16, No. 2, pp. 448–457 (online), DOI: 10.1109/TASL.2007.911513 (2008).
- [11] Liu, X., Chen, Q., Wu, X., Liu, Y. and Liu, Y.: CNN based music emotion classification (2017).
- [12] Choi, K., Fazekas, G., Sandler, M. and Cho, K.: Convolutional Recurrent Neural Networks for Music Classification (2016).
- [13] Chaki, S., Doshi, P., Patnaik, P. and Bhattacharya, S.: Attentive RNNs for Continuous-time Emotion Prediction in Music Clips, *AffCon@AAAI*, (online), available from <https://api.semanticscholar.org/CorpusID:219602013> (2020).
- [14] Soleymani, M., Caro, M. N., Schmidt, E. M., Sha, C.-Y. and Yang, Y.-H.: 1000 Songs for Emotional Analysis of Music, *Proceedings of the 1st ACM International Workshop on Crowdsourcing for Multimedia (CrowdMM)*, ACM, pp. 1–6 (online), DOI: 10.1145/2506364.2506365 (2013).
- [15] Aljanaki, A., Yang, Y.-H. and Soleymani, M.: Developing a Benchmark for Emotional Analysis of Music, *PLOS ONE*, Vol. 12, No. 3, p. e0173392 (online), DOI: 10.1371/journal.pone.0173392 (2017).
- [16] Zhang, K., Zhang, H., Li, S., Yang, C. and Sun, L.: The PMemo Dataset for Music Emotion Recognition, *Proceedings of the 2018 ACM International Conference on Multimedia Retrieval (ICMR '18)*, Yokohama, Japan, Association for Computing Machinery, pp. 135–144 (online), DOI: 10.1145/3206025.3206037 (2018).
- [17] Betella, A. and Verschure, P. F. M. J.: The Affective Slider: A Digital Self-Assessment Scale for the Measurement of Human Emotions, *PLOS ONE*, Vol. 11, No. 2, p. e0148037 (online), DOI: 10.1371/journal.pone.0148037 (2016).