

1. 序論

ChatGPT1)などに代表される生成系AIは、文章・画像・楽曲など様々なコンテンツを生成できるAIの総称である。ここ数年間で生成系AIが生成する文章・画像・楽曲などの精度は大きく向上しており、テキストで指示するだけで、人間が制作した作品と判別しづらい文章や画像を作り出すことができるAIもある。近年では生成系AIを用いることにより、容易に文章、画像等を作成できるため、教育的な問題、倫理的な問題、法的な問題など様々な問題が発生している2)3)。学習に使用されるデータは原則、著作権者の許諾なしに利用可能なため、任意のクリエイターの作品を無断使用し、その特徴に似せてAIに出力させることもできてしまう。その結果、学習元への経済的還元が無かったり、模倣物と競合してしまったりすることに繋がる。楽曲の面では現在、いくつかの生成AIモデルが、任意のアーティストの声や歌詞のテンポを学習し、本物と酷似したトラックを作成する機能を提供している。それらの問題に対処するために、生成系AIが作成したものと、人間が作成したものとを区別できる手法の必要性が上がっており、「AI or Not」5)をはじめとした画像の判別するツールが開発されている。しかし、楽曲を判別するツールは現在確認できない。そこで本研究では、AI生成楽曲を判別などの生成系AIをめざす。本研究では、AI生成楽曲の特徴や人間が制作した楽曲との差などを調べ、その差を元にAI生成楽曲であるかどうかを判別するAIを作成する。また、AIおよびAI作成楽曲の利害やあり方についての考察を行う。

2. 研究内容

本研究では、楽曲生成AIのひとつAIVA6)を用いて楽曲を生成し、それを音声認識AIの学習データとすることでAI生成楽曲が持つ特徴を検証する。機械学習等で使用するプログラムは、C++言語で作成する。また、AIが生成したものではない楽曲は、「フリーBGM(音楽素材)無料ダウンロード | DOVA-SYNDROME」7)や、「フリーBGM・音楽素材 MusMus」8)からダウンロードしたものを使用する。

2.1 楽曲を生成

AIVA内で、Chord progressionからランダムに楽器を選択し、1分30秒~2分30秒程度の楽曲を50曲作成する。また、DOVA-SYNDROMEとMusMusから人間が作曲し

た同じような長さの楽曲を50曲ダウンロードする。

2.2 楽曲データの読み取り

本研究では音声ファイルを読み取るためのプログラムを作成し、AIVAで作成したものと、DOVA-SYNDROMEやMusMusでダウンロードしたものを、読み取り、データ化した。

2.3 教師あり学習

楽曲の特徴を得て、教師あり学習でサポートベクトルマシンを用いて、AIと人間の2つのグループに分類するプログラムにより機械学習していく。

2.4 楽曲を判別

学習に使用していない、新たな楽曲の音声データから特徴ベクトルを生成し、SVMモデルを使用して特徴ベクトルを分類することで、AIが生成したものかどうかを判別させる。

3. 結果・考察

それぞれボーカルありの楽曲で判別できるようにしたかったが、DOVA-SYNDROMEやMusMusにアップロードされているボーカル有りの楽曲が少なく、学習データ不足になると考え、ボーカルの無い楽曲を用いての研究になった。

4. 結論

参考文献

- 1) ChatGPT, Open AI, <https://chat.openai.com/>
- 2) 生成AIをめぐる議論, 情報通信白書令和5年度版, 総務省(2023), <https://www.soumu.go.jp/>
- 3) 武田俊之: 大学は生成系AIの影響をいかに認識しているか?, 日本教育工学会 研究報告 Vol.2023, No.2, pp.88094 (2023), https://doi.org/10.15077/jsetstudy.2023.2_88
- 4) A.Sablayrolles, M.Douze, C.Schmid, H.Jégou: Radioactive data: tracing through training, arXiv Machine Learning, Cornell University (2020), <https://arxiv.org/abs/2002.00937>
- 5) AI or Not, <https://www.aiornot.com/>
- 6) AIVA, <https://www.aiva.ai/>
- 7) DOVA-SYNDROME, <https://dova-s.jp/>
- 8) MusMus, <https://musmus.main.jp/>