

研究テーマについて

1 漫画の研究を続けるのか

- 感情推定はメンタル的に続けたくない.
- モチベーション的には、白抜きされた連番コマ画像を入力とし、セリフを生成する研究を目標として、まずはセリフの選択問題タスクに取り掛かる?

2 その他興味ある研究テーマ

2.1 歌詞・メロディの相互変換

Zhonghao Sheng ら [1] は Transformer を用いて、マスク部分を予測するという教師なしの事前学習を行うことで、新しく高精度な歌詞と旋律の変換を達成した.

2.2 The NES Music Database[2]

The NES Music Database とは、Chris Donahue らによって作成されたファミコンのゲーム楽曲 (BGM・SE) を収録したデータセット. 397 タイトルの計 5278 曲が含まれている. MIDI データを始めとした扱いやすい形式に変換されているので扱いやすい. 生音を用いた複雑な音楽とは異なり、制限された音階と楽器 (2 つのパルス波, 1 つの三角波, 1 つのノイズ) をうまく組み合わせているので、機会学習に向いている. 先行研究では、筆者らによって LakhNES モデルが提供されており、使うことができる. (制作物例)

以上を踏まえて、例えばゲームのタイトルやジャンル、Wikipedia に載っている概要などをまとめた自然言語情報を入力として、そのゲームにふさわしい 8bit BGM の自動生成モデルの作成とかは面白そう (それ以上に難しそう) と思っている.

3 次やること

やること決める. 腹くくる.

参考文献

- [1] Zhonghao Sheng, Kaitao Song, Xu Tan, Yi Ren, Wei Ye, Shikun Zhang, and Tao Qin. Song-mass: Automatic song writing with pre-training and alignment constraint, 2020.
- [2] Chris Donahue, Huanru Henry Mao, and Julian McAuley. The nes music database: A multi-instrumental dataset with expressive performance attributes, 2018.