
進捗報告

1 今週やったこと

- データセット準備

2 データセット準備

Video ViT を使用しようとしていたが、現在整備している画像データだとデータ長不足（1つのデータが3コマしかないもの）やデータの長さが不均一という問題があったため、このデータを ViViT に与えるのは難しいと判断した。

そこで Youtube 上の動画から動画データセットを作成する方針と一旦 ViT に一枚ずつデータを入力する方針に分けて進める。

2.1 ViT

事前学習した ViT に自作の画像フォルダを読み込ませるところまでは動作を確認した。MLP に通す前の特徴ベクトルを獲得すれば、絵描き歌の1フレームの情報は取得できる状態になった。

2.2 動画データセット

動画をデータセットにする場合を調査した。Youtube からダウンロードはできたので、open cv で画像に書き出せばデータセットの読み込みは出来ると思われる。

1データあたりの情報量は増えるが、グラデーションがかかっていたり歌詞の字幕ついていたりアスペクト比がばらばらだったりするので、ノイズは画像よりも多い。

3 予定

- 実験方針を決める