

第5回ミーティング

MCNの再現実験のための準備

ER20038 小林 亮太

2023 年 5 月 27 日

1 はじめに

卒業研究を始めるにあたり、マルチモーダルデータを用いた自己教師あり学習のなかでも 3 モーダルの場合の知識を得るためにラベルなしビデオデータからの自己教師あり学習のための Multimodal Clustering Networks (MCN)[1] について論文調査を行った。それに引き続き、調査した論文内で紹介されていた手法で論文と同じ精度が達成できるかを確認するために再現実験を行う。

2 実験準備

再現実験を行うにあたり、その準備を可能な限り行った。以下に完了したものとそうでないものを示す。

2.1 完了した準備

2.1.1 プログラムのダウンロード

論文で紹介されている手法のソースコードが github に存在しているため、clone コマンドでダウンロードした。また、ダウンロードしたファイルを解凍した。

2.1.2 Dockerfile の作成

実行予定のプログラムから必要なモジュールなどの情報を読み取り、適切な dockerfile を作成を行う。しかしながら、一部のモジュールはビルド段階ではエラーとなり実装できなかったため、ビルド後にコンテナ内で手動にてインストールをした。

2.2 未完了の準備

2.2.1 データセットのダウンロード

今回の手法で利用する HowTo100M データセットのダウンロードが必要である。そのため、公式サイトから All-in-One とされている zip ファイルをダウンロードした。しかしながら、ダウンロードしたファイルはデータセットそのものではなく、別途フォームから申請を行い、ユーザー ID とパスワードを受け取る必要がある。その他のデータセットについての情報を以下にまとめる。

- ナレーション付きビデオの大規模データセット
- 120 万本の Youtube ビデオのビデオにキャプションを付けた 1 億 3600 万本のビデオクリップ
- 23000 のカテゴリが存在
- 全てをダウンロードする場合、12TB ほどの容量が必要

以下がデータセットのダウンロードに関する現在までの進捗である。

- 4/13：フォームを送信
- 4/27：著者にメールを送信
- 現在は返答待ち
- youtube の動画をコマンドラインからダウンロードする方法を試験的に実行

データセットが youtube 上の動画を利用しているため、外部の youtube 動画をダウンロードするようなサービスを利用することでデータセットを用意することが可能である。以下に使用を試みたものをまとめる。

- youtube-dl：長期間、更新が途絶している。自身の環境では動作を確認できなかった。
- yt-dlp：youtube-dl のフォーク。より高速なダウンロードが可能である。自身の環境で動作を確認したため、youtube-dl の代替として利用した。

HowTo100M の公式サイトで配布されている動画 ID のリストを参照して yt-dlp で自動ダウンロードするようなプログラムを作成し、正常に動作することを確認した。

3 おわりに

今回はマルチモーダルデータの自己教師あり学習の再現実験の準備を行った。進捗としては、プログラム実行時にプログレスバーが表示されるところまでは進んでいるので、データセットを入手すればすぐにも動かすことができると考えられる。今後としては、必要な動画の解像度などダウンロードに関する詳細な情報を集め、早急にダウンロードを開始したいと考えている。

参考文献

- [1] Brian Chen, et al., “Multimodal Clustering Networks for Self-supervised Learning from Unlabeled Videos”, ICCV2021.