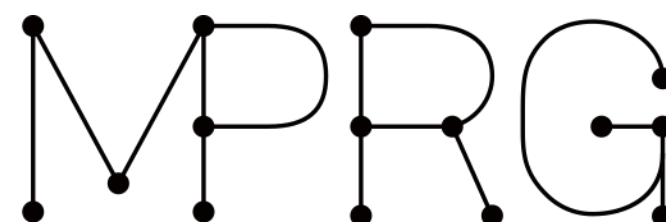


第5回ディスカッション

MCNの再現実験のための準備

ER20038 小林亮太

担当：岡本（主），張，岩垣



MACHINE PERCEPTION AND ROBOTICS GROUP

- 再現実験
 - Multimodal Clustering Network (MCN)
 - 準備
 - データセットの準備

- 完了済
 - プログラムのダウンロード
 - 一部データセットのダウンロード
 - 必要なモジュールのインストール
 - Numpy, Panda, librosa, torch, fast_pytorch_kmeans, apex, Tqdm, gensim
- 未完了
 - データセットのダウンロード
 - HowTo100MのYouTube上の動画データのダウンロード
 - フォームからの利用申請を行ってユーザー名とパスワードの入手が必要
 - 進捗
 - フォームを送信 (4/13)
 - 著者にメールを送信 (4/27)
 - 現在は返答待ち

- HowTo100Mデータセット
 - 様々な手法のトレーニングで使用
 - Youtube上の動画を利用
 - 自分でダウンロードすることが可能
- Youtube-dlを用いた動画のダウンロード
 - 現在, 自分の環境ではYoutube-dlは使用不可
 - 長期間, 更新が途絶
- Yt-dlpを用いた動画のダウンロード
 - Youtube-dlの派生
 - より高速なダウンロードが可能
 - Youtube-dlの代替として利用

- HowTo100Mデータセットのダウンロード
 - 公式サイトで配布されている動画IDのリストを利用
 - リストを参照して自動でダウンロードするプログラムを作成
 - プログラムが正常に動作することを確認

```
video_id,category_1,category_2,rank,task_id
nVbIUDjzWY4,Cars & Other Vehicles,Motorcycles,27,52907
CTPAZ2euJ2Q,Cars & Other Vehicles,Motorcycles,35,109057
rwmt7Cbuvs,Cars & Other Vehicles,Motorcycles,99,52907
HnTLh99gcxY,Cars & Other Vehicles,Motorcycles,35,52907
EyP3HVhg1u0,Cars & Other Vehicles,Motorcycles,95,52906
w6zbbdK1ewY,Cars & Other Vehicles,Motorcycles,17,52906
RAidUDTPZ-k,Cars & Other Vehicles,Motorcycles,10,52907
nssig0FNZVU,Cars & Other Vehicles,Motorcycles,99,52906
m2YyhYyleII,Cars & Other Vehicles,Motorcycles,153,52906
u5B1RAufWVM,Cars & Other Vehicles,Motorcycles,150,52906
AEytW9ScgCw,Cars & Other Vehicles,Motorcycles,108,52906
tYQoPHwNkho,Cars & Other Vehicles,Motorcycles,18,52907
DUxVMAebfrM,Cars & Other Vehicles,Motorcycles,131,52907
M1Ta54rkeQQ,Cars & Other Vehicles,Motorcycles,54,52907
```

動画IDのリストの一部

- 再現実験の準備：データセットのダウンロード
 - フォーム, メールからの利用申請：返答待ち
 - Youtube-dlを用いたダウンロード：自身の環境では利用不可
 - Yt-dlpを用いたダウンロード：動画IDリストを元にダウンロード可能なことを確認
- 今後の予定：データセット準備, Everything at Onceの再現実験準備