

Detectron2 を使用した撮影目的の推定

Estimation of Purpose of Shoot using Detectron2

石井 克佳[†] 黒木 啓之[†]

Katsuyoshi ISHII[†] Takashi KUROKI[†]

[†] 東京都立産業技術高等専門学校

[†] Tokyo Metropolitan College of Industrial Technology

1. はじめに

ニューラルネットワークを活用した画像認識は日々発展しており、セグメンテーションはその一つである。

セグメンテーションは、事前に登録されているクラスに対応するオブジェクトを見つけ出し、画像を分割することで認識を可能にしている。こうした技術は、画像内の情報を細かく読み取ることができるが、間接的に関連している情報を読み取ることも可能である。

本研究では、セグメンテーションタスクを使用することで、人物や背景から画像の撮影目的の推定することを目的とする。

2. Detectron2

Detectron2 は Facebook AI が開発した PyTorch ベースの物体検出ライブラリで、オブジェクト検出、セグメンテーション、その他の視覚認識タスクのためのプラットフォームである [1]。本研究ではセグメンテーションを実行する際に使用する。

3. パノプティックセグメンテーション

パノプティックセグメンテーションはいくつか存在するセグメンテーションタスクのうち、セマンティックセグメンテーションとインスタンスセグメンテーションを組み合わせたものである。このセグメンテーションの特徴は画像内すべてのピクセルに対してラベル付けし個別に認識することができることである。

4. COCOdataset

COCO dataset は、物体検出やセグメンテーションなどのコンピュータビジョンタスクのベンチマークとして使用されているデータセットである[2]。数十万枚以上の画像とアノテーションが含まれ、80 のカテゴリに対応している。

5. 研究方法

5.1 セグメンテーション

前述した COCO dataset を使用して、セグメンテーションを実行するためのモデルを作成する。

パノプティックセグメンテーションは画面全体をラベルごとに分けることができるため、同一ラベルの物体も個別に認識できるタスクである。これによって、背景に写っているものや、写っている人の数などを認識することができる。そこから検出された、物体の位置関係や大きさの情報を抽出する。

5.2 学習

次にセグメンテーション処理をした画像を元にそれらをジャンル分けし、再度ニューラルネットワークで学習する。

学習にはセグメンテーションタスクによって得られた情報を活用する。例として物体の数、大きさ、中央からの距離、画像占有率などから画像が何を目的としているものなのかを特定する。

また、セグメンテーション処理を行っていない画像でニューラルネットワークを作成したときとの比較を行う。

6. 研究結果

配信されている COCOdataset を使用したモデルを使用して撮影した画像にセグメンテーション処理を行った。その結果を図 1 に示す。



(a)原画像

(b)処理画像

図 1 パノプティックセグメンテーション処理

図より、パノプティックセグメンテーションとして後ろの物体も個別に認識できていることがわかる。ただし、背景は認識できていない部分もあるため改善が必要である。

7. 今後の予定

今後は推定方法の決定とセグメンテーションから得られた情報の使用方法などを決め実装する。またそれらに使用するサンプル画像を収集する。

その後、学習結果の比較と検討を行う。

参考文献

- [1] github, facebookresearch, Detectron2, <https://github.com/facebookresearch/detectron2>, 参照 Jan. 10, 2024.
- [2] COCO - Common Objects in Context, <https://cocodataset.org>, 参照 Jan. 10, 2024.