

リアルタイムオブジェクト検出モデルと インスタンスセグメンテーションモデルの設計支援ツール

Design Support System for Realtime Object Detection Model and Instance Segmentation Model

永田研究室 EMM2509 西田 尚生
Nagata Laboratory EMM2509 Naoki Nishida

1. 研究の背景と目的

山口東京理科大学では令和6年度に数理・データサイエンス・AIに関する科目を履修することを可能とした学部横断型プログラムとして、「数理・データサイエンス・AI 専門応用プログラム」を開設しており、これは文部科学省の「数理・データサイエンス・AI 教育プログラム（応用基礎レベル）」にも認定されている。このため、研究室では現在、データサイエンス教育における学生のニューラルネットワーク関連の応用的実務能力の向上と、企業との共同研究にも活用できるシステムの構築を行っている。本研究では自動運転システム、製造工程の自動化ライン、医療診断システムに加えて、社会生活の場面でも移動ロボットなどへの応用が期待されているリアルタイムオブジェクト検出用(OD用)のYOLOXモデルと、図1のようなインスタンスセグメンテーション用(IS用)のSOLOv2, Mask R-CNNモデルも設計できるように機能を拡張したので報告する。

2. 研究内容

MATLAB アプリケーション(mlapp)上にYOLOX, SOLOv2, Mask R-CNN各モデルの設計、学習及び評価の機能を追加し、mlappのメインダイアログでCOCOデータセットを使って事前学習されたバックボーンモデルを基に、新しいドメインに対応したモデルを構築することができるようにした。まず、企業より依頼のあったperson, chair, desk, door, wheelchair, alcoholの6カテゴリ(クラス)が含まれたラベリング用の画像を学内で撮影し、400枚程度収集した。これらの画像に対してMATLABから提供されているラベリングツールを用いて、OD用モデルではバウンディングボックスで、IS用モデルでは多角形でそれぞれ検出させたい領域を指定し、ラベル(クラス名)を付与した。ラベリング結果はgTruthファイルとして保存し、各モデルの訓練に用いた。その後、テスト用の動画内のオブジェクト検出実験により、訓練後のYOLOX, SOLOv2, Mask R-CNN各モデルの汎化性能を評価した。

3. 結果

OD用のYOLOXについては、訓練用のラベリング画像を増やしていくことで汎化性能を良好に向上させていくことができた。一方、IS用のMask R-CNNは、SOLOv2に比べて訓練時間や検出精度の点で十分な性能が得られなかったため、SOLOv2のみで評価を行った。当初は1モデルのSOLOv2により、6カテゴリの検出を試みたが、訓練パラメータの調整、学習時間、検出精度などの点で実用的な性能が得られなかった。そこで各カテゴリに特化させた1クラス認識用のSOLOv2モデルを独立に構築し、それらを統合して実行できるようにシステムを拡張した。その結果、各クラスの検出性能を改善できただけでなく、クラス毎のラベリング用画像の更新、訓練時間の短縮、新たな検出対象クラスの追加や不要となったクラスの削除といった管理の容易さなど多くの機能改善を図ることができた。現在は、オブジェクト検出時の確からしさ(スコア)を高めるためにラベリング画像を蓄積している。例えば、doorのテスト画像では、ガラス製の透明のもの、正面ではなく斜め方向から撮影されたもの、あるいは一部のみが映った画像などの検出精度に課題があったため、より多くの種類のドア、様々なアングルと距離から撮影したドアの写真を収集しながら、それらの検出精度と信頼度を高める取り組みを継続的に行っている。



Fig. 1 Instance segmentation by SOLOv2.