

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

永田研究室 EMM2509 西田 尚生

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(OD)用のYOLOXモデルと、インスタンスセグメンテーション(IS)用のSOLOv2を設計できるように機能を拡張したので報告する。

ODとISについて

オブジェクト検出とは図1(b)のように画像や動画内の中から特定のオブジェクトをバウンディングボックスとラベル名で特定するコンピュータービジョン技術である。また、インスタンスセグメンテーションとは図2(b)のように画像や動画内のオブジェクトを個別のインスタンス(個体)としてピクセル単位で識別し、それぞれの正確な輪郭を抽出する技術である。例えば、対象物が同じpersonであってもそれを「person 1」「person 2」のように個別のオブジェクトとして認識することができる。

アノテーション

アノテーションとは、AIの学習に不可欠な教師データを作成することを目的とした、画像内に存在する物体や領域、属性などにラベルを付け、コンピュータに意味を理解させるための作業のことである。また、特定のオブジェクトにラベルを割り当てる作業はラベリングと呼ばれる。

企業より依頼のあったperson, chair, desk, door, wheel_chair, 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モデルの訓練用には図1(a)のようにバウンディングボックスで、また、ISモデルの訓練用には図2(a)のようにポリゴン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。

学習

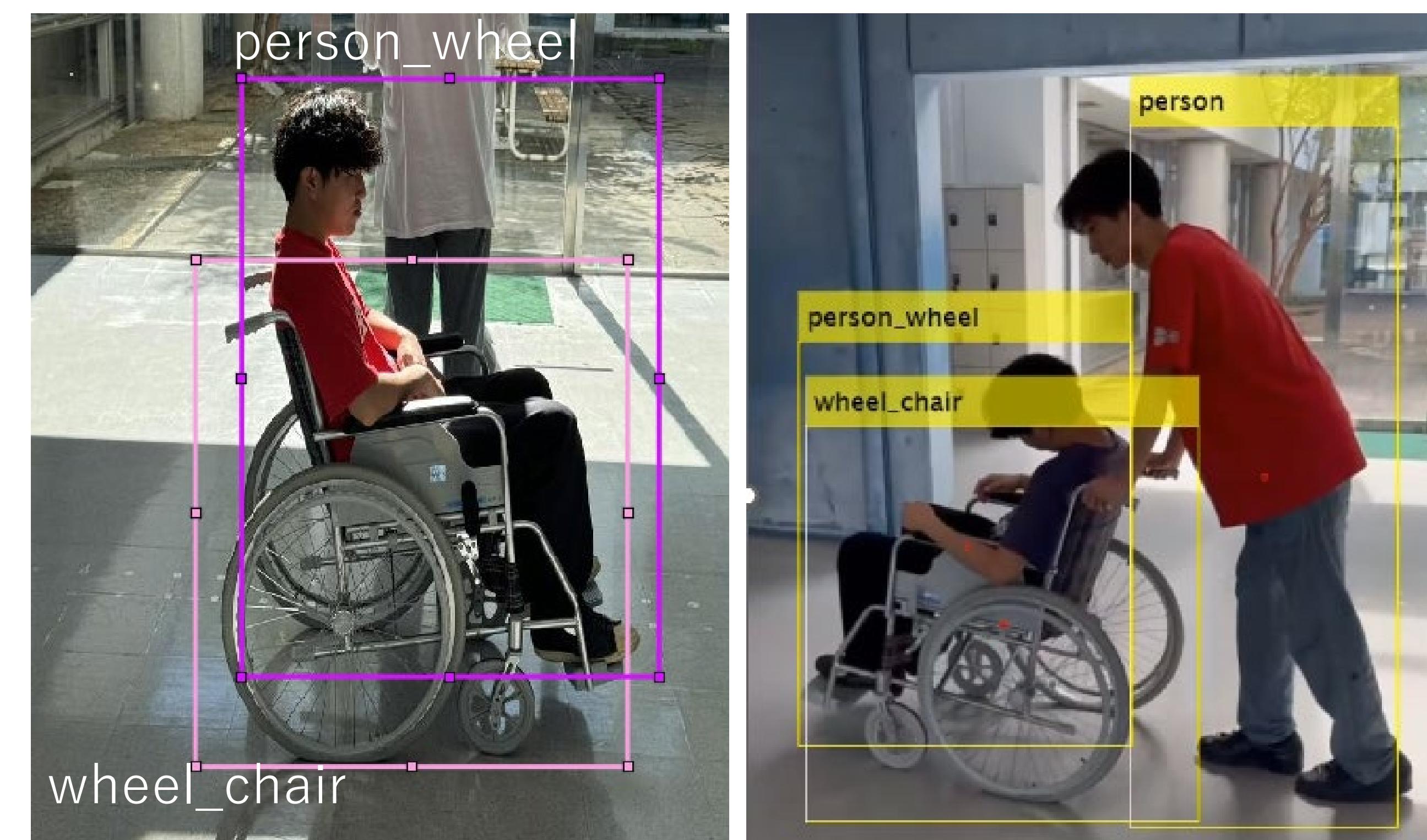
アノテーション結果をもとに、MATLAB上で開発したYOLOモデルとSOLOモデル設計用ツールを使って学習を行うことができる。図3(左)には研究室で開発したツールのメインダイアログを示す。ここで保存したgTruthファイルの選択や、学習条件などの設定を簡単に行うことができる。

YOLO, SOLO共に様々なバージョンがあるものの、本研究ではMATLAB環境で推奨されているYOLOXとSOLOv2の各モデルの構築を主テーマにしている。このアプリを用いることで、YOLOXモデルでは既に公開されているnano-coco, tiny-coco, small-coco, medium-coco, large-cocoをバックボーンに選択後に、また、SOLOv2モデルではresnet50-coco, light-resnet8-cocoを選択後に、事前に準備したラベリング画像を用いて転移学習を実施し、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。また、図3(右)サブダイアログでは各モデルの汎化性能の評価を簡易操作で行うことができる。

学生は学習に伴う計算負荷を抑えるためにラベリングを行う画像の解像度を低くしたり、適切な学習条件の摸索により訓練時間の短縮や検出精度の向上を試みることができる。

結果とまとめ

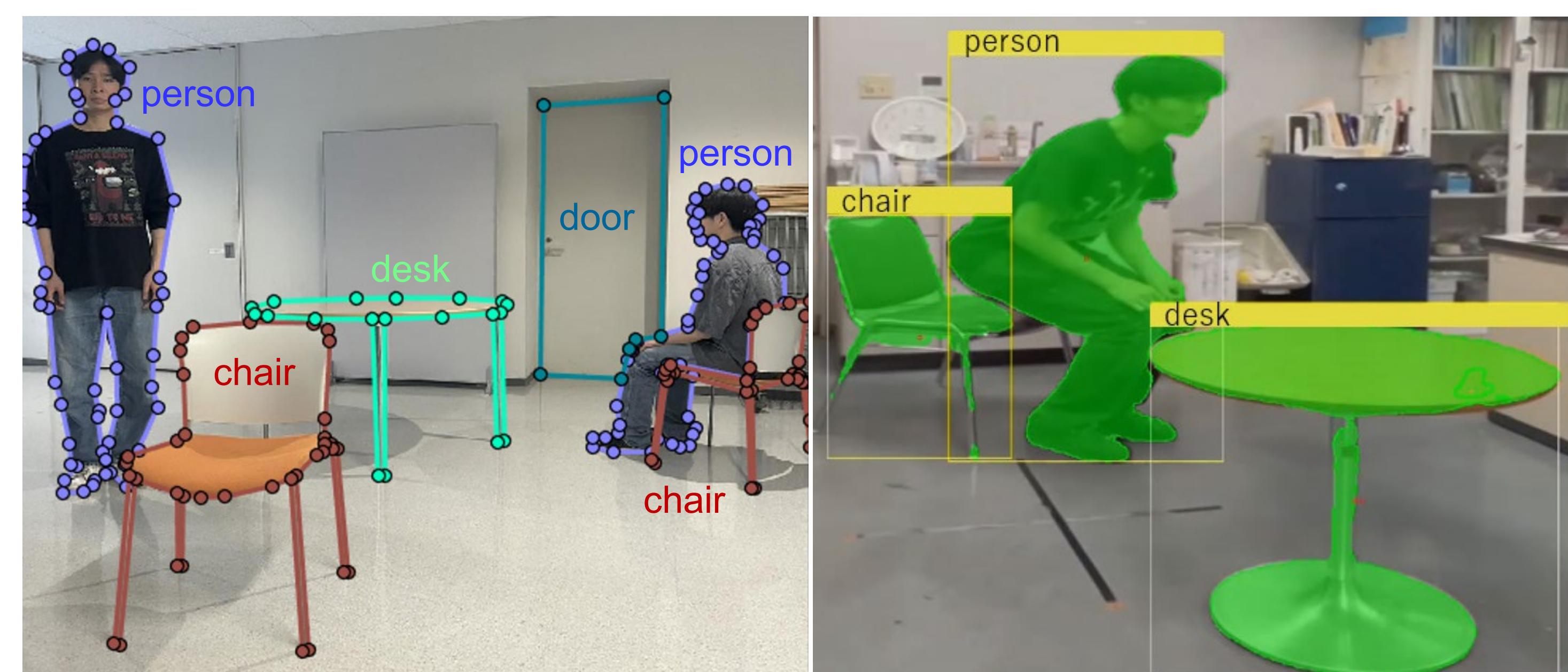
OD用のYOLOXについては、訓練用のラベリング画像を増やしていくことで汎化性能を良好に向上させていくことができた。一方、当初は1モデルのSOLOv2により6カテゴリの検出を試みたが訓練パラメータの調整、学習時間、検出精度などの点で実用的な性能が得られなかった。そこで各カテゴリに特化させた1クラス認識用のSOLOv2モデルを独立に構築し、それらを統合して実行できるようにシステムを拡張した。



(a) Labeled by rectangle

(b) Predicted objects

Fig. 1 Labeling and detection results by YOLOX.



(a) Labeled by polygon

(b) Segmented objects

Fig. 2 Labeling and segmented results by SOLOv2.

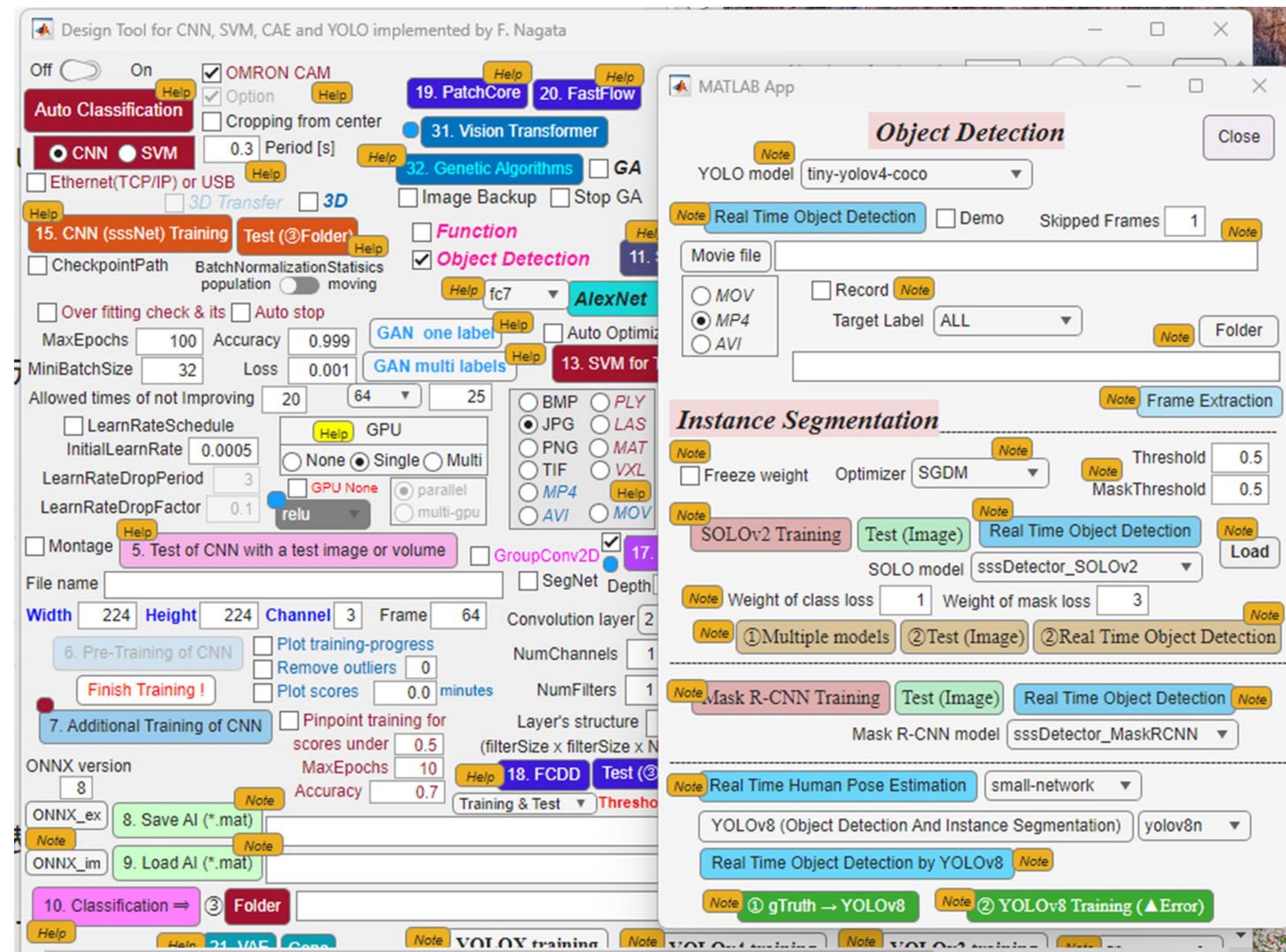


Fig. 3 Developed application for modeling YOLO and SOLO.

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。

- [1] 西田, 今村, 永田, 渡辺, リアルタイムオブジェクト検出用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, 産業応用工学会全国大会2025講演論文集, pp. 97-98, 長崎大学文教キャンパス, 2025年9月11日.
- [2] N. Nishida, F. Nagata, A. Kurozumi, K. Otsu, K. Watanabe, M.K. Habib, Design Support Tool of Deep Learning Models for Real-Time Object Detection and Instance Segmentation, Procs. of The 31st International Symposium on Artificial Life and Robotics (AROB 31st 2026), 2026 (to be presented).