

氏名	OBADA AL AAMA (シリア)		
学位の種類	博士 (工学)		
学位記番号	生工博甲第483号		
学位授与の日付	令和6年3月25日		
学位授与の条件	学位規則第4条第1項該当		
学位論文題目	Generating Synthetic Datasets to Enhance Robotic Perception in Food Pick-and-Place Tasks (フードピックアンドブレースタスクにおけるロボティック認識向上のための合成データセットの生成)		
論文審査委員会	委員長	教授	我妻 広明
		〃	石井 和男
		〃	田向 権
		〃	張 力峰

学位論文内容の要旨

本論文では、食品を対象とした物体認識システムの構築と、産業用ロボットを用いたお弁当詰め自動化への応用について提案している。物体認識システムの主体となるニューラルネットワークの訓練には大規模データが必要であるが、人手により数多ある食材の大規模データを準備することは非常に困難である。そこで本論文では、Generative Adversarial Network (GAN) を用いて疑似画像を生成することで食品画像データベースを構築することを提案している。また、生成したデータベースと物理シミュレータを組合せることで、大量の画像をシミュレータ内で撮影し、大規模データを自動構築することを提案している。自動生成した大規模訓練データを用いてニューラルネットワークを学習させることで、実用に耐えうる物体認識システムの構築が可能なことを数値的に検証し、その有効性を示している。さらに、提案システムを産業用ロボットへ統合し、ミニトマトやブロッコリーなどの実際の食品のPick-and-Placeタスクにて有効性を評価している。

本論文の構成は、以下の通りである。第1章は序論である。本論文の背景として、人口減少が続く先進国におけるロボットの重要性、ロボットの自動化における人工知能の役割、深層学習に必要な訓練データの構築の困難さ、Robot Operating System (ROS) などのAIやロボットに関する解説から始まり、本論文のモチベーションとなる深層学習の訓練のための大規模データ構築に関して現状の課題を述べている。その上で、本論文が取り組む問題として、1. 現状存在しない日本食に用いられる食材の大規模データベースの構築、2. 大規模訓練データ構築のための人手による教師情報 (アノテーション) 付与作業にかかる膨大な時間の削減、3. 産業用ロボット応用のためのリアルタイム物体認識システムの構築、以上の3点を設定することが述べられている。

第2章では、本論文の関連研究が述べられている。GANの拡張であるCycle-GANや各種Convolutional Neural Networks (CNN), You Look Only Once (YOLO), You Only Look At CoefficientTs (YOLACT)といった、本研究と関連深いニューラルネットワークの各種アルゴリズムが概説されている。また、トマト収穫ロボットなど、ニューラルネットワークのロボット応用に関する文献も紹介されている。

第3章では、本論文での提案手法が述べられている。まず、Cycle-GANによるシミュレータ画像の精緻化による大規模データセット構築を提案している。3Dシミュレータに精緻化後のオブジェクトを取込み様々な角度から対象物を撮影することで、大量の画像を準備可能なことを示している。準備した大規模データセットを用いてVGG-16というCNNを学習することで、食品の認識システムが構築可能なことを示している。次に、PyBulletという物理シミュレータを用いて、大量の食品オブジェクトをランダムに物理シミュレータ内に展開、同様に様々な角度で自動撮影することで、アノテーションの自動化と大規模データセットの生成が可能なことを示している。得られた大規模データセットを用いてYOLACTを学習することで、ロボット応用が容易な、物体検出と認識を同時に実行できるシステム構築が可能なことを示している。さらに、ロボット応用で重要となる把持点推定手法について述べている。最終的にロボットのハンド部にカメラを装着し、提案した物体検出・認識システムと組み合わせることで、産業用ロボットを用いたPick-and-Placeタスクを実現できることを示している。

第4章では、実験結果と考察が述べられている。まず、Cycle-GANによるシミュレータ画像の精緻化について述べられている。提案手法を用いて22,000枚(11クラス、各クラス2000枚)の訓練用画像を生成しVGG-16を学習させたところ、テスト精度52%になることが示されたが、認識精度は十分でないことが分かった。また、VGG-16は認識のみしか実行できないため、ロボット応用では物体検出を別途用意する必要がある。これらの課題を解決するために、PyBullet物理シミュレータを導入したアノテーション自動化と大量データセット生成手法を提案している。提案手法に各種データ水増し(Augmentation)手法を組み合わせることで120,000枚(11クラス)の画像を生成し、物体検出、認識の両方を同時実行できるYOLACTを学習させたところ、94.56%という高いテスト精度が得られることが述べられている。さらに、本手法を産業用ロボットへと実装し、Pick-and-Placeに適用、エビフライ等、6種のお弁当具材の自動把持へと応用したところ、物体毎に差があるが、最低でも90%以上の物体認識率と88%以上の物体把持成功率を達成することが出来たことが述べている。さらに、人手によるデータ準備と比べて、提案手法によるデータ生成は約80倍高速であることが示されている。

第5章は本研究のまとめである。本論文が達成した項目、議論、今後の展望が述べられている。

学位論文審査の結果の要旨

本論文に関し、審査委員から、問題設定の根拠、本研究が達成した内容と特にロボット応用で未達成な内容の整理と今後の展望、画像生成と画像アノテーションの違い、関連研究との比較の妥当性など、多数の質問がなされたが、いずれも著者より満足な結果が得られた。

また、公聴会は現地と遠隔のハイブリッド形式にて開催した。ハイブリッド形式であるため、公聴会の予告は、各種WEB、メール、Facebook、Slackなどを通して行い、発表は現地参加に加えて、ZOOMおよびYouTube Liveで公開した。その結果、学内外より合計42名の出席があった（YouTubeのリアルタイム視聴者を含む）。質疑では、データ生成に必要な時間と認識精度のトレードオフの考え方、ロボット応用における把持点推定手法や把持力の設計法の詳細などに関する質問が多数あったが、どれも著者の明快な説明により質問者の理解が得られた。

以上により、論文調査及び最終試験の結果に基づき、審査委員会において慎重に審査した結果、本論文が、博士（工学）の学位に十分値するものであると判断した。