

自動収穫ロボットの開発③

～DeepLearning を用いた野菜の認識～

福島県立テクノアカデミー浜 職業能力開発短期大学校 計測制御工学科

○発表者名 石川 拓実 指導教官 牛坂 慶太

1. はじめに

野菜の自動収穫を実現するためには、野菜の認識が必要となる。従来の画像処理（パターンマッチング等）による物体の判別は、事前に用意したデータと比べて、新たに取得した画像の背景や明るさが異なった場合、その認識率は大きく低下する。

そこで今回は DeepLearning を用いることで、より高い精度での野菜の位置、良・不良、収穫時期かどうかの判別等を可能にすることを目的とする。

2. DeepLearning による画像認識

DeepLearning による画像認識には大きく 3 つのレベルに分けられる。

(1) Categorization (カテゴリ識別)

画像内に写っている物体の名称のみを知らせる。その際に物体の位置を知らせる必要はない。

(2) Detection (物体検出)

画像内に写っている物体を矩形に切り取り、位置を知らせる。また、切り取った物体の名称も同時に知らせる。

(3) Semantic Segmentation

画像内に写っている物体の領域をピクセルレベルで認識する。

これらは(1)～(3)に行くにしたがって難易度が上がる。今回は野菜の収穫を目的とする上で実際の野菜の位置も認識する必要があるためこの中の(2)を用いる。

3. システム概要

USB カメラに判別させたい野菜を映し、結果をモニターに出して判別するシステム。判別種類については野菜の位置、良・不良、収穫時期かどうかの判別を行うことを目的とする。

使用したソフトウェア等は以下の通りである。

(1) Python

Python とは少ない分量で簡単にプログラムが書けるシンプルでオーソドックスな文法であり、複雑な数学計算（微分演算や行列計算）を 1 行で済ませることも可能。

(2) OpenCV

OpenCV は、画像処理・画像解析および機械学習等の機能を持つオープンソースのライブラリである。オ

ープンソースだから、基本的に無料で使用することができる。Python を始め、C/C++、Java、MATLAB 用として公開されており、もともとはインテルが開発・公開している。

(3) CUDA

CUDA (Compute unified device architecture) とは、NVIDIA が提供する GPU 向けの C/C++ 言語の統合開発環境であり、コンパイラやライブラリなどから構成されている。

(4) YOLOv3

YOLOv3 は物体検出手法シリーズ YOLO (You Only Look Once) の一つである。YOLO はリアルタイムオブジェクト検出アルゴリズム。このアルゴリズムでは検出窓をスライドさせるような仕組みを用いず、画像を一度 CNN に通すことで、オブジェクトを検出することができる。v2 と比較して精度が良くなり、特に小さいものの検出が効果的。

(5) Anaconda

Anaconda(アナコンダ)とは、「データサイエンス」や「機械学習関連アプリケーション」のための Python および R 言語用のディストリビューション。

4. システムの流れ

(1) PC のブラウザで指定した物体（今回の場合はトマト）の学習用画像を元に YOLOv3 で学習を実施し、モデルを作成する。

(2) 作成したモデルを使用し、映像から物体を判別できるのかを確認。（図はデモ用判別画像）

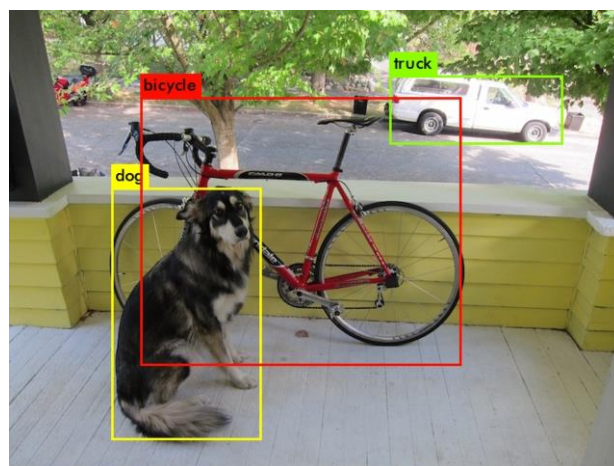


図 1 デモ用判別画像

5. 実験

(1)画像の水増し

野菜（トマト）の画像を水増しし、学習用画像の量を増やす。水増しの方法としては水増しのプログラムを利用し、元画像に対して様々な加工（コントラストを強調・低減させる。画像を反転させる等）を施し、元画像とは別に17枚の画像を水増しした。水増しした画像を含めたフォルダの配置は研究内容に該当したサイトを参考にし、以下のようにした。

```
darknet
|- ...
|- cfg
  |- task
    |- datasets.data
    |- class.txt
    |- train.txt
    |- test.txt
    |- yolov3-voc.cfg
    |- datasets
      |- img1.png
      |- img1.txt
      |- img2.png
      |- img2.txt
      |- ...
```

(2)トマトの座標を取得

画像の中の野菜の部分に矩形に切り取り、切り取った矩形の座標を取得する。方法としては専用のサイトが存在したため、そのサイトを使用し切り取り及び座標の取得を行った。



図2 座標の取得

(3)学習

YOLOv3 で(1)で水増しした画像を用いて学習を実施し、モデルを作成

(4)判別

その後、作成したモデルを用いて映像から野菜を正しく判別できるかを確認する。

6. 結果

今回はトマトの画像データ数百枚程度を使って学習を行った。結果としては高い精度で認識させることに成功した。以下の画像は認識結果である。

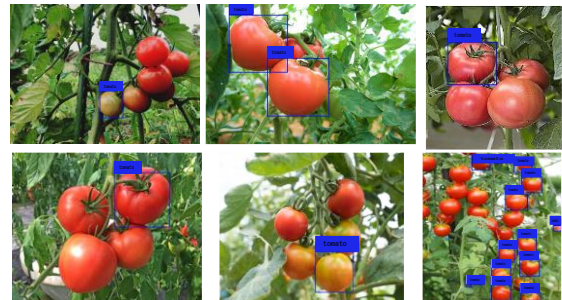


図3 認識結果

7. 考察

今回の結果として、トマトの位置を正確に認識するには成功した。しかし、トマト自体の良・不良、収穫時期かどうかの判別は満足に行うことが出来なかった。その原因としては画像データ量の不足が考えられる。

8. おわりに

今回の卒業研究を通して、DeepLearningを学ぶ機会を得ることができた。初めてDeepLearningを行うにあたり分からない点も多数存在したが、ネットを活用する等して少しずつ理解を深めていった。しかしまだ理解が及ばない点も多いため、今後も機会があれば学びたいと思う。

参考文献

- (1) <https://qiita.com/harmegiddo/items/c3db5fd567fa4c6cc9fb>
Yolo v3 を用いて自前のデータを学習させる + Yolo v3 & opencv のインストール方法付き (Ubuntu 16.04, Opencv 3.3, Conda)
- (2) <https://leapmind.io/blog/2017/06/16/ディープラーニングとは?【入門編】/>
ディープラーニング (DeepLearning) とは?【入門編】
LeapMind BLOG Posted by 坂口真里奈