

ユズ果実計測のための物体検出モデルの性能向上

1250289 上杉颯大 【 知能情報学研究室 】

1 はじめに

近年, 人工知能技術の進展によって, 農業分野においても自動収穫ロボットなどの研究が活発化している. 深層学習を用いた画像認識も広く用いられるが [1], 深層学習では大量の教師データが必要なため, YOLO のような学習済みモデルを用いることが一般的である. 学習済みモデルの入力サイズは 1280 画素程度であることが多く, 一方カメラの画素は 4K 画素以上が多い. そこで本研究では, ドローンからユズ果実を撮影した画像を対象に, カメラの画像数を縮小化せず分割して YOLO モデルに入力することで, 果実計測のために認識精度の向上をはかる. また, 同時に分割を行うと, 入力サイズを小さくすることもできることから, メモリ使用の抑制効果もある.

2 提案手法

本研究では, UAV(Unmanned Aerial Vehicle, ドローン) によって撮影された 4000×2250 画素の画像を, 学習済みモデルの入力サイズ 640×640 , 1280×1280 ごとに分割して順に入力し, 果実を検出することで, 認識精度の向上, GPU メモリ使用量の抑制をはかる. 画像の分割は元の画像左上から順に, モデルの入力サイズに合わせて定められた正方形領域を取り出し, 右あるいは下の残った部分を, 分割サイズ未満のデータとして作成した.

3 実験

データセットは, ドローンにより撮影された着果しているユズの木の航空写真 174 枚にユズの実と判断した場合には果実として視認できる部分にのみラベルを付与し, 学習用データ 139 枚, 検証用データ 35 枚に分ける. その後, 4000×2250 画素の元画像を 320×320 画素, 640×640 画素, 1000×1000 画素, 1280×1280 画素, 2000×2000 画素の各サイズに分割を行い, それらを学習用データおよび検証用データとした. YOLOv8 の nano モデルを用いて分割をしていない画像と各サイズに分割した画像で学習を行う. YOLOv8 の内部前処理機構を利用してモデル入力サイズを 640, 1280 の固定解像度にリサイズし, 分割しない場合と比較を行う. モデル



図 1 分割画像例

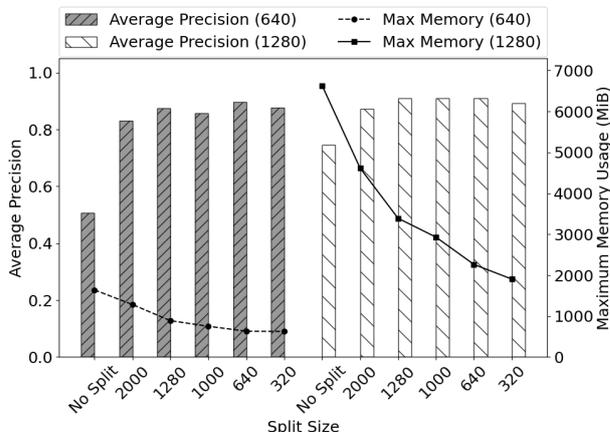


図 2 学習条件別の平均適合率 (AP, 棒), 最大メモリ使用量 (折れ線)

の性能評価には, Average Precision(平均適合率, AP) と GPU の最大メモリ使用量を用いる.

4 結果・考察

図 2 にモデル入力サイズを 640, 1280 とした時の分割画像の大きさ別の平均適合率 (AP) と最大メモリ使用量を示す. 棒と左目盛りが AP, 折れ線と右目盛りがメモリ使用量を示している. 画像を分割して学習を行った全てのモデルで分割をしていないモデルより高い AP の値を示した. モデル入力サイズ 640 では 640×640 のサイズで学習を行ったモデルで AP が 0.897 と最も高く, 最大メモリ使用量も 631MiB と, 分割を行わないモデル 1639MiB に比べて大幅に低減していることが確認できた. さらに, 分割サイズが大きくなると AP の値が低下することが確認された. これは学習時のリサイズ処理において, 対象物であるユズが相対的に小さく入力されるようになるため精度が低下したと考えられる. モデル入力サイズ 1280 ではモデル入力サイズ 640 の結果よりも高い AP の値のモデルが複数存在した.

5 おわりに

本研究では, YOLOv8 の nano モデルを用いて分割有無や分割サイズ, リサイズ解像度の違いによるモデルの性能評価を行なった. 結果より, AP, GPU メモリの使用量を考慮するとモデル入力サイズ 640 の条件下で 640×640 サイズで学習したモデル的していることが確認できた.

参考文献

[1] 三本他, “少データ状況下における敵対的生成ネットワークを用いた果実識別 AI の性能向上法,” 知能システムシンポジウム, 2023.