

放牧支援のための空撮画像における CNN に基づく牛検出

CNN-based cattle detection in aerial images for grazing management assistance

邵文[†] 福田 誠一郎[†] 吉橋 亮太[†] 川上 玲[†] 尤 少迪[‡]
Wen Shao[†] Seiichiro Fukuda[†] Ryota Yoshihashi[†] Rei Kawakami[†] Shaodi You[‡]
川瀬 英路[§] 苗村 健[†]

[†] 東京大学 [‡]Data61, オーストラリア国立大学 [§] カミエンス・テクノロジー株式会社
[†]University of Tokyo [‡]Data61, Australian National University [§]Kamiens Technology Inc.

Abstract: 畜産業において労力不足が深刻化しており、負担の少ない放牧が注目されている。放牧牛管理の支援のため、本稿ではドローンにより撮影された空撮画像から牛を検出するシステムを提案する。牛の検出には CNN を用い、その中でも物体検出で高い精度と実行速度を兼備する YOLOv2 を導入する。この際、空撮画像では対象物体の画像上の大きさがある程度既知であるため、これとネットワークのダウンサンプリング数により、対象物にとってより相応しい入力解像度が決定できる。放牧環境での空撮画像を用いた実験から、YOLOv2 が高い精度で牛を検出可能であり、また対象物に適している入力解像度を用いた場合に、より高い性能を得た。

1 はじめに

近年、人口減少及び高齢化による労力不足が畜産業で深刻化している。農家の負担軽減のため、牛舎管理作業の軽減や牛の健康増進などの利点を持つ放牧が勧められている。放牧牛の管理で重要な作業は、牛の個体数や活動状態を把握した上での怪我や病気の早期発見とその対応であり、コンピュータビジョンによる動物監視は労力を減らす有望な手立てである。また、広範囲にわたる動物の監視には、空撮画像が適している。

本稿はドローンによる空撮画像を用いる放牧牛の検出システムを提案する。牛の検出は CNN (Convolutional Neural Network) を用い、その中でも物体検出で高い精度と実行速度を兼備する YOLOv2 (You Only Look Once version 2) [1] を導入し、現場での実用に近い高精度で頑健な検出器を目指す。一般に画像内における対象物体の大きさは様々に変化する。しかし、本研究で対象とするのは、ドローンを一定の高さに維持して撮影した空撮画像であり、牛の解像度がある程度予測することができる。このような場合、これとネットワークのダウンサンプリング数により、より適した入力解像度を決定できることを示す。

2 関連研究

ドローンによる空撮画像を用いる牛の検出として、王らの研究 [2] と Chamoso らの研究 [3] がある。王らは、HOG 特徴量と SVM を検出器として用いたが、道路や木などの背景の誤検出や牛の検出漏れが多く見られ、精度が十分でない。一方、Chamoso らは CNN を用いたが、それは合計 3 層と浅く、また出力は特徴マップであり、各個体の推定には別の画像処理を必要とする。

3 システム概要

検出システムには YOLOv2 [1] を用いる。従来の検出器が各ウィンドウ毎に特徴マップの計算をしていたのに対し、

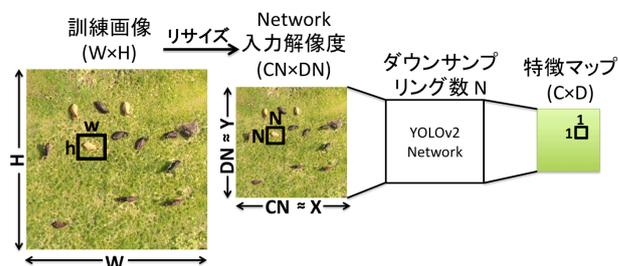


図 1: 入力解像度の調整

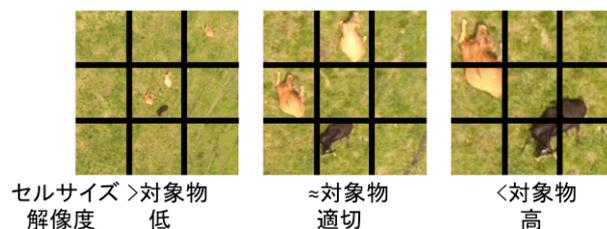


図 2: セルサイズと対象物の大きさの関係

YOLOv2 は、画像全体が畳込層に入力されるため、特徴マップの計算が画像全体に対して一度で済む。YOLOv2 は、一連の畳込層とプーリング層が出力する特徴マップから回帰により対象物の位置とサイズを予測する。

特徴マップの各画素から回帰を行うことを踏まえ、この特徴マップで対象物が一画素に入るような入力解像度が対象物にとって相応しいと考える。特徴マップの一画素の入力画像での領域をセルと呼ぶ。図 1 のように、ネットワークのダウンサンプリング数を N とする場合、セルサイズは $N \times N$ になる。したがって、入力画像の牛の領域がおおよそ $N \times N$ に収まることが望ましい。図 2 に示すように、牛の解像度が低い場合、一つのセルに複数の牛が入り、逆の場合、セルに牛の一部分しか入らず、いずれも適切でない。そこで図 1 のように入力解像度 $X \times Y$ を定めることを提案する。原画像のサイズを $W \times H$ 、その中で牛のサイズの中央値を $w \times h$ とすると、 X と Y は

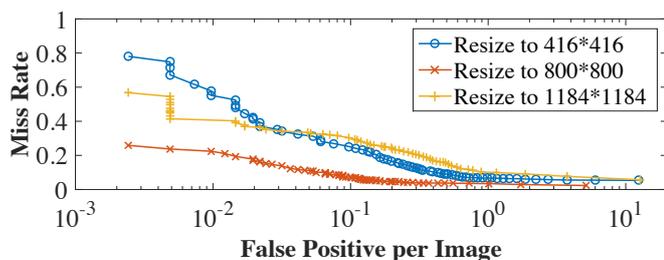
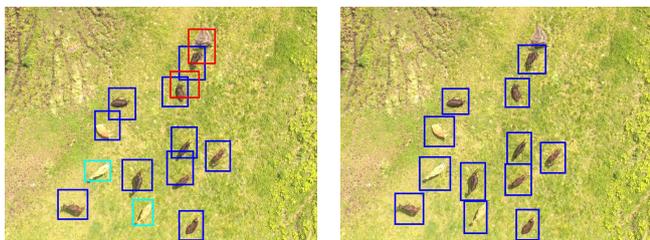


図 3: ROC 曲線による比較



(1) Resize to 416 × 416 (2) Resize to 800 × 800

図 4: 検出結果例 (青: TP, 赤: FP, シアン: TN)

$$X = \frac{W \times N}{w}, \quad Y = \frac{H \times N}{h} \quad (1)$$

で求まる。アンカーボックスを用いる検出精度の向上のため特徴マップのサイズ $C \times D$ は奇数が望ましい [1]。これも考慮して、 $X \times Y$ に最も近い入力解像度 $CN \times DN$ を決める。

4 実験

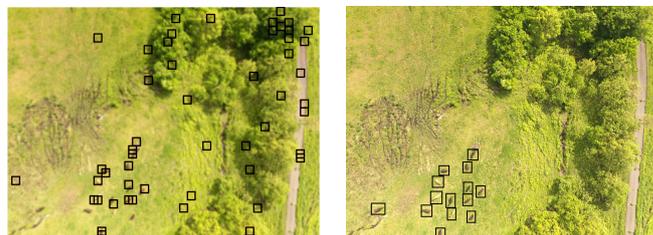
提案する YOLOv2 による検出システム及び入力解像度の改良手法の有効性を示すため、空撮画像から牛を検出する実験で精度を評価した。実験では、Darknet [4] をフレームワークとして使用した。

空撮画像は王ら [2] が作成したものを使用する。これは 2016 年 5 月に撮影された合計 656 枚の $4,000 \times 3,000$ 画素のデータである。全部で 1,887 頭の牛にアノテーションを手動で付与した。牛のサイズの中央値は $w = 87$, $h = 90$ であった。

訓練画像として 245 枚の画像を選び、 $3,000 \times 3,000$ 画素に両端から切り出して 490 枚の正方形画像を得た。この正方形画像の中の $2,122 \times 2,122$ 画素の領域に対して 30 度ずつの回転による 12 倍のデータオーグメンテーションを行った。全部で 5,880 枚の訓練画像が得られた。訓練における入力解像度 $X \times Y$ は式 (1) で計算した。 $W = H = 2,122$ であり、YOLOv2 のダウンサンプリング数は $N = 2^5 = 32$ である。したがって、 $X = 781$, $Y = 754$ が求められた。それに特徴マップのサイズが奇数である条件を加え、入力解像度を 800×800 画素 ($C = D = 25$) に設定した。実験では、YOLOv2 の事前学習で使われた ImageNet 画像のサイズに最も近い設定の 416×416 画素 ($C = D = 13$)、及び更に高解像度の 1184×1184 画素 ($C = D = 37$) の場合と比較した。

表 1: 従来研究と検出性能の比較

	適合率	再現率	F 値
王らの研究 [2] (HOG+SVM)	0.319	0.538	0.401
本稿 (Resize to 416 × 416)	0.853	0.887	0.870
本稿 (Resize to 800 × 800)	0.945	0.942	0.944
本稿 (Resize to 1184 × 1184)	0.901	0.750	0.819



(1) HOG 特徴量と SVM を (2) YOLOv2 及び入力解像度の調整を用いた検出結果

図 5: 従来研究と検出結果の比較

テストについて、訓練画像と同じ日に撮影した 411 枚のテスト画像を訓練画像と同じサイズに分割し、検出を行った。確信度の閾値を変化させ、Miss Rate (牛の総数に対する検出漏れの割合) と FPPI (False Positive per Image) を計算した。図 3 で示す ROC 曲線の比較により、牛のサイズの中央値が出力される特徴マップで一画素になるように入力解像度を 800×800 画素に設定した場合に最も高い検出性能を持つことがわかった。具体的な検出例を図 4 で示す。また、各確信度の閾値における適合率、再現率、F 値を計算した結果、閾値が 0.29 の時 F 値が最大値を取り、最も性能が良い。王らの研究 [2] との検出性能の比較を表 1 に示し、図 5 でそれぞれの一例を示す。本手法によって検出性能を向上させることができた。

5 おわりに

本研究では、YOLOv2 による放牧牛の検出器を構築し、対象物が出力の特徴マップで一画素になるような入力解像度が適することを示した。今後は学習データに含まれない新しい牧場での検出性能を調べる予定である。

参考文献

- [1] J. Redmon *et al.*: “YOLO9000: Better, Faster, Stronger”, CVPR, 2017.
- [2] 王ら: “畜産業支援に向けたドローンによる空撮画像の撮影と牛検出への応用”, SSII, 2017.
- [3] P. Chamoso *et al.*: “UAVs Applied to the Counting and Monitoring of Animals”, Ambient Intelligence, 2014.
- [4] J. Redmon: “Darknet: Open Source Neural Networks in C”, <http://pjreddie.com/darknet/>, 2013–2016.