

ドローンによる牧場空撮画像における乳牛の個体識別

田口航平† 邵文† 川上玲† 川瀬英路‡ 苗村健†

†東京大学 ‡カミエンス・テクノロジー株式会社

E-mail: shao@nae-lab.org

Abstract

放牧における牛の監視作業は、労働の負担が大きい一方で、個体の状態を管理し病気や怪我の牛を早期発見するために重要である。本稿ではドローンを用いて撮影した空撮画像から乳牛の模様によって個体識別を行い、それによる個体ごとの監視を目指す。空撮画像では牛の向きは不揃いであるが、牛の識別を行う際には、体の向きが揃っていた方が識別精度がよいと予測される。このため、畳込みニューラルネットワーク (CNN) である YOLOv3 を用いてまず牛の体と頭部を検出し、画像における牛の体軸整合を行う。その後 Siamese Network を用いて乳牛の模様の類似度による個体識別を行う。従来手法であるテンプレートマッチングと比較し、提案手法の有効性を確認した。

1 はじめに

畜産農家の作業時間の中で、大きな割合を占めるのは監視作業である。その中でも病気や怪我をした牛の早期発見が重要である。近年、一戸あたりの牛飼養頭数は増加する一方、労働人口は高齢化しており、農家の管理負担が増している。放牧の監視作業の負担を軽減するため、長期的なコストが少ない、保守運用が楽などのメリットを持つコンピュータビジョンを用いた監視が注目されている。特に、広い範囲にわたる放牧地をカバーできるドローンによる空撮画像の使用が有望である。

本稿は、ドローンによる空撮画像から乳牛の個体識別を目指す。これにより、遠隔地から牛を個体ごとに観察することが可能となり、農家の支援につながる。空撮画像では対象物が様々な向きで写されるが、個体識別のためには、体の向きが揃っている方が精度が良いことが期待される。そこで、YOLOv3 [1] を用いた牛の体と頭部の検出を行い、これに基づいて個体の向きを揃える体軸整合を行う。また、模様による個体識別については類似度学習を行う Siamese Network [2] を用いる。実験では、Siamese Network を用いる場合とテンプレートマッチングを用いる場合の比較を行い、またそれぞれでは体軸整合の処理の有無による比較も行う。

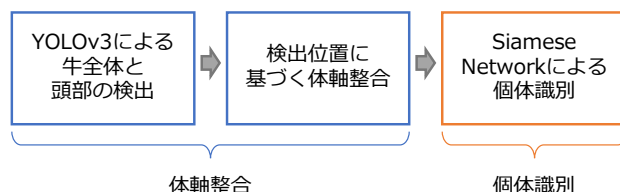


図 1: 手法の概要

2 関連研究

畜産におけるコンピュータビジョンを利用した牛の監視に関する研究はいくつかある。Ardöら [3] は、牛舎で撮影したビデオから、CNN を用いて牛を検出し、写っている牛が2頭以下のフレームを省くシステムを作成した。また、Liら [4] は牛の模様検出において、牛の臀部の模様について QDA (Quadratic Discriminant Analysis) や SVM (Support Vector Machine) を用いることにより、牛の個体識別を実現した。しかし、これらは牛舎における固定カメラによる監視で、空間的に広い放牧地における監視に適用するのは困難である。

空撮画像からの牛検出について、王ら [5] や Longmoreら [6] はドローンにより撮影した空撮画像から HOG (Histogram of Oriented Gradients) と SVM を用いて牛の検出を行った。また、邵ら [7] はドローンの飛行高度が一定の場合、空撮画像に写っている対象物のサイズが概ね同じである特徴を利用し、CNN による牛の検出を実現した。しかし、広範囲にわたる空撮画像において、個体レベルでの監視はまだ実現されていない。

3 提案手法

牧場空撮画像において乳牛の個体を識別するため、以下のようなステップで問題解決を目指す。まず牛とその頭部の検出を行った後、牛の体全体の中心位置と頭部の中心位置から算出した角度だけ牛の画像を回転させ、牛の体軸整合を行う。その後、模様の類似度学習により個体識別を行う。手法の概要を図 1 に示す。

3.1 牛の体軸整合

まず牛の体軸整合を行うことで個体識別における牛の向きによる影響を軽減する。ここで体軸整合とは画

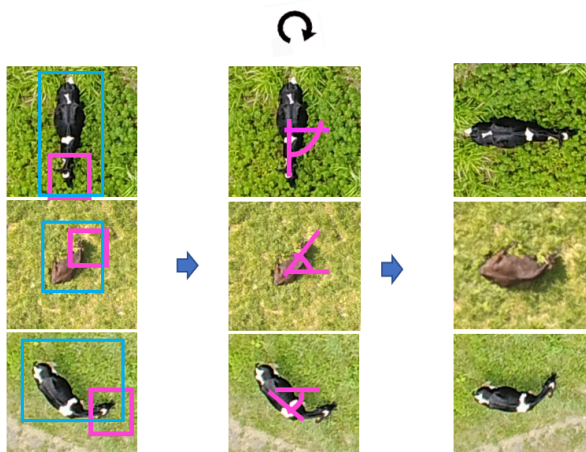


図 2: 体軸整合の流れ

像中の牛の体の向きを揃えることを指す．牛とその頭部の検出を YOLOv3 [1] を用いて行い，牛の体全体の中心位置と頭部の中心位置から算出した角度だけ牛の画像を回転させる．牛の頭部と体の中心位置から牛の向きを以下の式で計算する．

$$\theta = \arctan \frac{y_b - y_h}{x_b - x_h} \quad (1)$$

ここで x_b, y_b を牛の体の検出位置の画像における xy 座標， x_h, y_h を牛の頭の検出位置の xy 座標とした．またここで牛の頭部については牛の体の中心位置からの距離がもっとも近く，かつ体のバウンディングボックス内にあるものをこの牛の頭として採用した．これにより牛同士が近い場合に異なる牛の頭を誤って判断する可能性が軽減できる．また同一の牛に対して複数の検出結果が得られた場合はもっとも IOU が大きいものを結果として採用した．体軸整合の流れを図 2 に示す．

3.2 模様の類似度学習による個体識別

個体識別には通常多くのデータが求められ，人の個体監視の問題に関する先行研究の多くでは膨大な数のデータを必要とした．しかし，今回使用する乳牛の空撮画像についてはドローンのバッテリー容量により飛行時間が限られることから，データ数を人の個体監視で用いられるレベルの数まで増やすことは困難である．そこで少ない画像からでも個体識別が可能なネットワークが必要であると考えられ，二つの画像の類似度学習を行う Siamese Network [2] を導入する．模様による個体識別の流れを図 3 に示す．

4 実験

4.1 データセット

本研究では，ドローンにより撮影した乳牛の空撮画像データセットを作成した．撮影は 2018 年 7 月・8 月・9 月に埼玉県秩父市の牧場にて行われた．7 月の撮影で

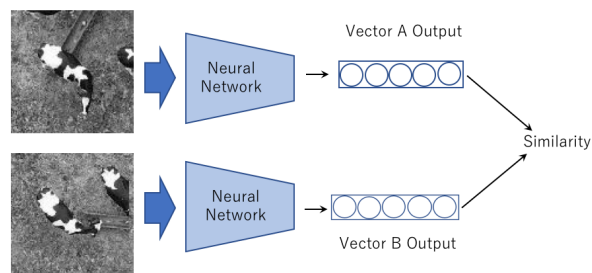


図 3: 模様による個体識別の流れ



図 4: データセットの例

は 4 回の飛行で 5472×3648 画素の 83 枚の画像と 3 つの動画から 2 秒ごとに切り出した 1920×1080 画素の画像 125 枚を取得した．8 月と 9 月の撮影ではそれぞれ 5472×3648 画素の 83 枚の画像を入手した．合計 374 枚の画像について全てに牛の位置と頭部位置を記録し，そして個体ごとに ID を付与するアノテーションを行った．撮影された牛は 84 個体 2177 頭であり，手作業によりラベル付けを行った．今回用いたデータの例を図 4 に示す．

4.2 牛の体軸整合

牛の頭部と体の検出による体軸整合の実験では今回作成したデータセットと王ら [5] の肉牛の空撮画像データセットを使用した．訓練画像として，4.1 のデータセットから 7 月に撮影した画像から 66 枚と王ら [5] のデータセットから 242 枚で使用し，分割及び回転によるデータオーグメンテーションを行った．分割について，今回作

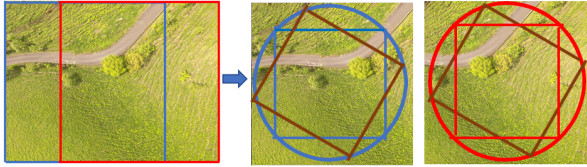


図 5: 回転によるオーグメンテーションの流れ

成したデータセットについてはまず牛の大きさが平均して 100×100 画素前後になるように画像を 6000×4000 画素にリサイズしたあと 2000×2000 画素になるように 396 枚に切り分けた。王らの画像も 4000×3000 画素を 2000×2000 画素になるように 968 枚に切り分けた。分割により 1364 枚のオーグメンテーション画像が得られた。回転について、王ら [5] のデータセットについては 4000×3000 画素の画像 242 枚、今回撮影したデータセットのうち 7 月の 6000×4000 画素にリサイズした 66 枚の合計 308 枚を 1 枚につき回転し 24 枚にオーグメンテーションを行い、計 7392 枚が得られた。回転によるオーグメンテーションの流れを図 5 に示す。分割と回転で得られた合計 8756 枚の画像を用いて訓練を行った。

次にテスト画像として 8 月に撮影したものは撮影牧区に応じて 40 枚を 4000×3000 画素に、 44 枚を 6000×4000 画素にリサイズし、訓練画像と同じ大きさに切り分け 424 枚に切り分けた。9 月に撮影した 53 枚は 3000×2000 画素についても同様にリサイズを行い、 106 枚に切り分けた。これに王ら [5] のデータセット 4000×3000 画素 407 枚についても同様に 1628 枚に切り分け、合計 2158 枚の画像を用いた。

YOLOv3 [1] により、牛の体と牛の頭部の 2 クラスの検出を行った。評価には IOU を用い、検出結果のバウンディングボックス BB_1 が正解データ BB_2 とどれだけ一致しているかを以下の式で計算する。

$$IOU = \frac{area(BB_1 \cap BB_2)}{area(BB_1 \cup BB_2)} \quad (2)$$

YOLOv3 の検出結果における信頼度の閾値を 0.20 に設定し、信頼度が 0.20 より大きければ検出結果として扱った。この検出結果と Ground Truth の IOU を計算し、IOU が 0.2 より大きければ TP (True Positive) とし、 0.2 より小さければ FP (False Positive) とした。二重にカウントを行っているものについては IOU の大きい方を TP とし、小さい方を FP とした。また Ground Truth と検出結果の IOU を計算し、検出結果の中で IOU の値が 0.2 以下のものについては FN (False Negative) としてカウントを行った。そして Miss Rate (誤り率)、Recall (再現率)、Precision (適合率)、F-Measure (F 値) について求めた。

牛の頭部と体の検出結果の精度を表 1 に示す。また、検出結果の例を図 6 に示す。頭も体も検出されてい

表 1: 検出精度の評価

	Precision	Recall	F-Measure
牛全体	0.882	0.831	0.860
牛の頭部	0.814	0.801	0.808



図 6: YOLOv3 による検出結果。青:正検出, 赤:誤検出, シアン:検出漏れ

い牛だけでなく、頭または体の片方だけしか検出されない牛も存在する。

4.3 模様類似度学習による個体識別

模様認識には本研究で作成した乳牛のデータセットを使用した。画像における牛の大きさを統一するために牛の大きさの平均が約 100 画素前後になるように画像のリサイズを行い、アノテーションに基づいて牛の中心位置が中心に来るように 200×200 画素の大きさに画像を切り分け、牛が 1 頭ずつ表示されている画像を作成した。作成した画像を ID ごとに分類することで個体識別を行うためのデータを作成した。また切り分けた画像をさらに 105×105 画素になるようにリサイズし、白黒に画像を変換した。今回訓練に 40 個体、テストに 37 個体の牛の画像を使用した。また、今回訓練に用いる画像は背景画像の影響を考慮し、背景を変化

	Template Matching			Siamese Network		
	Precision	Recall	F-Measure	Precision	Recall	F-Measure
回転なし	0.521	0.281	0.365	0.245	0.468	0.321
YOLOv3 の検出に基づく回転	0.446	0.246	0.317	0.284	0.542	0.372
Ground Truth に基づく回転	0.497	0.306	0.379	0.387	0.431	0.408

表 2: Siamese Network と Template Matching との比較: (1) 回転なし (2) YOLOv3 の結果に基づく回転 (3) Ground Truth に基づく回転

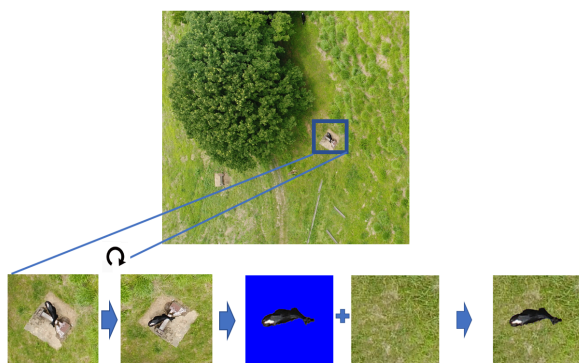


図 7: 背景のオーグメンテーションの流れ

させてオーグメンテーションしたものを訓練に用いた。訓練画像は全てオーグメンテーションしたものを使用し、一個体につき 20 枚となるようにオーグメンテーションを行ったので、合計 800 枚の画像について訓練を行った。今回の訓練では画像のペアについて訓練を行うので 400 枚数の画像のペア 79800 通りについて訓練を行った。背景のオーグメンテーションの流れを図 7 に示す。

評価の際には背景をオーグメンテーションした画像は使用せず、オーグメンテーション前の画像 212 枚を用いて評価を行った。その画像を撮影場所により 4 つのグループに分類し、個体ごとにグループ内からランダムに 100 ペアを抽出し、そのペアについて類似度を計算し、評価を行った。グループ毎に 100 ペアについては同じ個体の牛からペアが選ばれるように調整を行った。今回 37 個体に 3700 ペアについて評価を行った。

Siamese Network の性能を評価するため、テンプレートマッチングとの比較を行った。また、個体識別における体軸整合の有効性について、回転を行わない場合、YOLOv3 の検出結果に基づいて回転を行った場合、Ground Truth に基づいて回転を行った場合の 3 つに分けて評価を行った。類似度が閾値 0.4 以上かつ ID が一致すれば TP とする。

結果を表 2 に示す。Siamese Network と Template Matching においてそれぞれ比較を行ったところ Siamese Network における F 値の方が高い結果となり、Siamese Network を用いて模様識別を行ったことについての有効性が示された。また、Siamese Network を用いた実



図 8: 個体識別の結果例

験では回転を行わない場合より、YOLOv3 を用いて体軸整合を行なった結果の方が高いを得ることができた。これにより今回の実験において牛の画像を回転させ、体軸整合を行った画像で検出を行うことの個体識別に対する有用性を示した。個体識別の結果の例を図 8 に示す。

5 おわりに

本研究では牛の個体ごとの監視のため、ドローンを用いて撮影した空撮画像から牛の個体識別を行った。牛の識別精度向上のため、YOLOv3 を用いて牛の位置とその頭部位置を検出し、牛の体軸整合を行った。その後、牛の模様の情報から Siamese Network による類似度学習を行い、二枚の画像の牛が同一であるかどうかの判定を行った。実験により、Siamese Network による個体識別は従来の Template Matching の手法に比べ、高い性能が得られることを確認した。また、Siamese Network を使用する場合、体軸整合を行う方が、回転を行わずに個体識別を行う場合よりも高い性能が得られた。したがって、体軸整合が個体識別に有用であることが確認できた。

謝辞

本研究は、JSPS 科研費 JP18K11348 の助成を受けたものです。

参考文献

- [1] Redmon Joseph, Farhadi Ali: YOLOv3: An Incremental Improvement, <https://pjreddie.com/media/files/papers/YOLOv3.pdf>, last visit: 2019/03/01.
- [2] Gregory Koch, Richard Zemel, Ruslan Salakhutdinov: Siamese Neural Networks for One-shot Image Recognition, Proceedings of the 32nd International Conference on Machine Learning, JMLR: W&CP volume 37, 2015.
- [3] Håkan Ardö, Oleksiy Guzhva, Mikael Nilsson: A CNN-based Cow Interaction Watchdog, in Proc. of IEEE International Conference on Pattern Recognition, 2016.
- [4] Wenyong Li, Zengtao Ji, Lin Wang, Chuanheng Sun and Xinting Yang: Automatic individual identification of Holstein dairy cows using tail-head images, Computers and Electronics in Agriculture, volume 142, 622–631, 2017.
- [5] 王晋, 福田誠一郎, 吉橋亮太, 川上玲, 川瀬英路, 苗村健: 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, 画像センシングシンポジウム (SSII), 2017.
- [6] S. N. Longmore, R. P. Collins, S. Pfeifer, S. E. Fox, M. Mulero-Pzmy, F. Bezombes, A. Goodwin, M. De Juan Ovelar, J. H. Knapen and S. A. Wich: Adapting astronomical source detection software to help detect animals in thermal images obtained by un-manned aerial systems, International Journal of Remote Sensing, 38 (8-10): 2623-2638, 2017.
- [7] 邵文, 福田誠一郎, 吉橋亮太, 川上玲, 尤少迪, 川瀬英路, 苗村健: 放牧支援のための空撮画像におけるCNNに基づく牛検出, 映像メディア処理シンポジウム (IMPS2017), pp.1–6, 2017.