The 32nd Annual Conference of the Japanese Society for Artificial Intelligence, 2018

牛の分娩予兆として映像から観測可能な状態の検知

Detection of Cattle States Relevant to Calving from Video

沖本祐典 *1 Yusuke OKIMOTO
菅原一真 *1 Kazuma SUGAWARA
斎藤奨 *1 Susumu SAITO
中野鐵兵 *1*2 Teppei NAKANO
赤羽誠 *1*2 Makoto AKABANE
小林哲則 *1 Tetsunori KOBAYASHI
小川哲司 *1 Tetsuji OGAWA

*1早稲田大学 Waseda University
*2知能フレームワーク研究所 Intelligent Framework Lab

We investigated the capability of general neural-network-based object detection and recognition techniques to automatically detecting in video characteristic states of cows which are relevant signs of calving. To prevent fatal accidents during calving, it is desirable to detect calving signs in video and image information from camera. End-to-end estimation of calving signs from video is not realistic without large-scale datasets. This study therefore attempts to detect from videos characteristic states that are observable as calving signs (e.g., standing/lying, tail rising, and protrusion of the allanto and amnion). For that purpose, we built a cow monitoring dataset by crowdsourcing, and trained and evaluated state detection systems based on convolutional neural networks on the developed dataset. Experimental comparisons demonstrated that cow standing and allanto and amnion detections perform well and that these features contribute to calving sign detection from video.

1. はじめに

分娩時における牛の死亡事故は，畜産農家にとって最も回避したい事故である．実際，子牛の死亡原因の 35%は難産に起因しており [Mee 13]，事故を防止するためにも畜産農家が分娩の介助を行うことが望ましい．そのため，分娩の予兆を通知する装置が多く製品化されている [Saint-Dizier 15]．これらの装置では，接触型のセンサを用いて分娩時の牛の特徴的な変化を検出している．例えば，膣内の温度低下を温度センサで検出するもの，子宮と腹部の収縮を腹部に装着したハーネスで検出するもの，尾の挙上回数の増加を加速度センサで検出するものなどがある．しかし，このような接触型センサは一般的に高価であり，接触型であるが故に牛に負担を与え，取り付けの際は農家にとって危険が伴う．そのため，畜産農家と牛の双方に負担を掛けない分娩予兆検知の手段として，非接触型センサであるカメラから取得した映像情報を用いることが望まれている．

牛の状態や行動パターンは牛の個体差や周辺環境の影響により，分娩予兆，通常時に依らず極めて多様であるため，牛の観測映像から分娩予兆かどうかを直接推定するシステムを構築するには，多様な変動を含む，整備された大規模なデータセットが必要となる．しかし，一般に利用可能な牛映像のデータセットは存在しない．さらに注意すべきは，分娩前に特徴的な変化が見られるような牛の行動 (例えば，起立と臥床など) の多くは，分娩前でなくても観測されることである．これは，分娩前数時間の映像データを用いて分娩予兆を予測するモデルを構築しても，その信頼性は低いことを示唆している．そこで，本研究では，映像情報から分娩予兆を直接推定するのではなく，映像情報から観測可能な分娩予兆に寄与する情報を抽出し，その上で分娩予兆のモデル化を行うアプローチについて検討を行う．

分娩予兆に関する臨床的，行動学的な特徴は，畜産学において詳しく調査されてきた [Saint-Dizier 15, Jensen 12, Burfeind 11]．例えば，臨床的な分娩予兆として，肉牛と乳牛とともに，分娩の 48 時間以内に膣内の温度が 0.6 度から 0.7 度降下することが知られている [Saint-Dizier 15, Burfeind 11]．乳牛において報告されている行動学的な分娩予兆としては，分娩前 2 時間における臥床時間の増加 [Jensen 12]，分娩前 2〜6 時間前からの起立・臥床の変化の回数の増加 [Saint-Dizier 15, Jensen 12]，分娩前 4〜6 時間前からの尾の挙上回数の増加 [Jensen 12] などが知られている．また，牛の分娩の過程では，尿膜・羊膜と呼ばれる，水風船上の組織が産道から露出する [Saint-Dizier 15]．このような変化の多くはカメラで捉えることが可能である．

そこで本研究では，映像から分娩予兆に寄与する情報の抽出を目指し，画像情報から観測可能な分娩の予兆として起立/臥床，羊膜・尿膜の露出，および尾の挙上という 3 状態に着目し，それらを畳み込みニューラルネットワーク (convolutional neural network, CNN) を用いて検出可能かどうかを調査した．CNN は物体認識・物体検出のタスクで高い性能を達成するが，その学習には整備された大規模データが必要となる．大規模な画像データセットとしては，ImageNet [Deng 09] が物体認識に Pascal VOC2007 [Everingham] および COCO [Lin 14] が物体検出に用いられているが，これらはクラウドソーシングを用いてラベル付与がなされている．本研究の対象である牛の行動に関するデータベースの構築に際しては，データの収集に加えて，クラウドソーシングに基づきラベルの付与を行うシステムの構築も必要となる．本研究を通じ，分娩予兆検知において画像情報を活用するための要件およびその効果に関する知見が得られることが期待できる．

本稿の構成は以下の通りである．2. では，分娩予兆に関係する牛の状態を説明し，それらを映像情報から検知するシステムについて述べる．3. では，クラウドソーシングを用いて構築した牛のモニタリングデータセットについて述べる．4. では，画像データを用いた実験により，開発した牛の状態検知システムの性能について述べる．最後に，5. にて本稿の結論を述べる．

連絡先: 早稲田大学基幹理工学部，東京都新宿区早稲田町 27，okimoto@pcl.cs.waseda.ac.jp

2. 分娩予兆に関わる牛の状態の検出

本章では，牛の分娩予兆とそれに関わる牛の状態について述べた後，それらの状態を動画像から検出するシステムについて述べる．

2.1 分娩予兆とそれに関わる牛の状態

本研究では次の 3 つの分娩予兆と，それぞれに関わる牛の状態に注目する．

- **臥床時間の増加**: 牛の臥床時間は分娩前 2 時間に増加する [Jensen 12]．臥床時間は牛の起立状態を検出することで算出可能となる．
- **尿膜・羊膜の露出**: 水風船状の組織である尿膜と羊膜は，分娩の過程に産道から露出する [Saint-Dizier 15]．一般的に，尿膜は分娩 3 時間から 7 時間前に露出し，また尿膜は分娩 2 時間半から 5 時間前に露出する．尿膜と羊膜は，前者が黒っぽく，後者が白っぽいという色の違いを除くと，外見は似ている．本研究では，尿膜の露出と羊膜の露出を一つの状態として扱う．
- **尾の拳上時間の増加**: 尾の拳上時間は分娩 4 時間から 6 時間前以降に増加する [Jensen 12]．尾の拳上時間は尾の拳上状態を検出することで算出可能である．

2.2 ニューラルネットワークによる牛の状態検出

図 2 に先述した分娩予兆に関わる牛の状態を検出する流れを示す．牛の領域が画像から切り抜かれた後，それらの領域がニューラルネットワークを用いた牛の状態検知器に入力される．

2.2.1 牛領域検出

牛領域の検出には，ニューラルネットワークを用いた高精度な物体検出アルゴリズム YOLOv2 [Redmon 17] を用いる．YOLOv2 の重みには，COCO Train/Val 2014 [Lin 14] で学習した既存のモデル YOLOv2 608x608** を用いる．このモデルは一般的な物体検出タスクを想定しているため，牧場という特殊な環境下において，牛の検出は再現率が低くなる．そこで今回，牛の可能性がある領域を多めに検出した後，ヒューリスティックに候補領域を削減する方法を取った．まず，10 種類の動物のクラス（cow, bear, dog, bird, elephant, horse, cat, sheep, mouse, giraffe）を牛の可能性があるとして検出し，領域の確信度 (confidence) が閾値 ($v_{\mathrm{th}}^{\mathrm{conf}}$) を超えたものを牛領域の候補とする．その次に，各画像において得られた候補領域を，以下のヒューリスティックな方法で特定の数まで削減する．2 つの候補領域 A, B が重なっていたとき ($A \cap B \neq \emptyset$)，次式で定義される intersection over union(IoU)

$$IoU(A, B) = area(A \cap B)/area(A \cup B) \qquad (1)$$

が閾値 ($v_{\mathrm{th}}^{\mathrm{IoU}}$) が超えていた場合，A，B のうち面積が小さい領域を破棄する（ここで $area(A)$ は領域 A の面積）．この操作を，候補領域の数が閾値 ($v_{\mathrm{th}}^{\mathrm{reg}}$) 以下となるまで繰り返す．もし候補領域数が閾値以下まで削減できなかった場合，その画像は訓練及び評価データとしては用いない．以上の操作により得られた領域を牛の領域として，CNN を用いた牛の状態検知システムの入力とする．

2.2.2 牛の状態検知

画像から検出した牛の候補領域ごとに，CNN を用いて，分娩予兆に関わる牛の状態を検知する．先述した分娩予兆に関わる牛の 3 つの状態ごとに，それぞれ状態検知器を構築する．状態検知器のネットワーク構造は状態ごとに同じであり，畳込み層は VGG16 [Simonyan 14] あるいは ResNet50 [He 16] と同様の構造，それに続く全結合層は 1 層 1024 ユニット出力層は，牛の分娩予兆に関わる状態の事後確率を出力する．ImageNet で学習済みの畳込み層とランダムに初期化された全結合層を，集めた牛のモニタリングデータを用いてファインチューニングした．ここで，YOLOv2 によって切り出された牛領域は 224x224 にリサイズされたのち，z-正規化を施し，CNN を用いた状態検知器に入力される．

3. データ収集

本章では，先述した 3 つの状態をクラウドソーシングを用いてアノテーションした，牛のモニタリング画像データセットについて述べる．

3.1 データ収録

分娩が近い黒毛和種雌牛を，分娩房の側面に設置したカメラで継続的に収録した．収録例を図 3 に示す．収録は 2016 年 11 月から鹿児島県の牧場で開始した．本研究では，収録された動画から 1 秒につき 1 枚ずつ画像として切り出したものを生の収録データとして扱う．

3.2 クラウドソーシングによるアノテーション

切りだされた牛領域の画像に対し，3 つの分娩予兆に関わる牛の状態のアノテーションを行った．アノテーションは，評価データを筆者らが，大規模な訓練データを Amazon Mechanical Turk のクラウドワーカーがそれぞれ行った．クラウドワーカーに対する質問は次のとおりである．

- 牛は起立状態か．
- 牛から尿膜か羊膜が露出しているのが見えるか．
- 牛は尾を拳上しているか．

また，上記の質問それぞれに対し，次の 4 つの選択肢を用意した．

- **はい**: 質問は正しい (例：牛は起立している)
- **いいえ**: 質問は正しくない (例: 牛は臥床している)
- **わからない**: 判断を下すのが困難
- **切り抜きエラー**: そもそも判断を下せない

回答の品質保証のため，一枚の画像に対して一つの状態のみをアノテーションする単純なタスク設計とした．タスク画面には，牛領域と，それを含む分娩房全体をワーカーに提示されている．わからない以外のラベルについての画像例も表示されており，必要な場合には別ページでさらに多くの画像例を確認できる．それぞれのタスクにつき 2 人分の回答を集め，両方のラベルが一致した画像のみを学習に用いた．

**https://pjreddie.com/darknet/yolo


(a) 起立状態


(b) 臥床状態


(c) 尿膜の露出状態


(d) 羊膜の露出状態


(e) 尾の挙上状態

図 1: 分娩予兆に関係する牛の状態の例

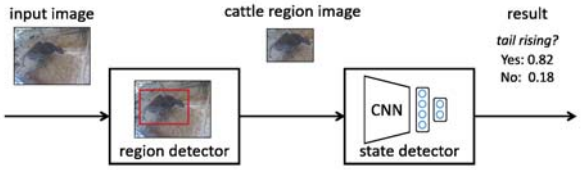


図 2: 分娩に寄与する状態検出の流れ．画像より牛領域を切り出し，状態検知システムの入力とする．

図 3: 収録された画像の例 (2 頭の牛が分娩房の斜め上方から収録される)．

4. 牛の状態の検知実験

4.1 使用データ

本実験では，独自に収集した牛のモニタリング画像データセットのうち 2017 年の 3–8 月収録分を用いた．ここには計 18 回の分娩シーンが含まれている．表 1 に，第 2.1 節で述べた分娩に関わる牛の状態ごとの，牛領域の画像の数を示す．これは，分娩 5 時間前から分娩時までのすべての画像データから 6 秒毎に 1 枚ずつ選び，それらの画像から牛領域を切り出し，アノテーションした結果である．分娩予兆に関わる状態についての訓練・検証データは 8 回の出産シーンを含み，評価データは訓練・検証データと重複しない 4 回の出産シーンを含むように選択した．さらに，データ数が多いクラスのデータの一部を無作為に削除することで，訓練データと評価データそれぞれの正例・負例の数が等しくなるようにした．

4.2 実験設定

牛領域の検出に関して，先述した検出におけるパラメータは経験的に決定し，$v_{\mathrm{th}}^{\mathrm{conf}}=0.1$，$v_{\mathrm{th}}^{\mathrm{IoU}}=0.7$，$v_{\mathrm{th}}^{\mathrm{reg}}=2$ とした．また，YOLOv2 の tensorflow による実装である darkflow[*†] を用いた．状態推定器のニューラルネットの学習には，モーメンタムを 0.9，学習率を $1e-4$ とした確率的勾配降下法を用いた．学習は validation loss が二回続けて上昇した場合に早期終了した．また，学習データに対しては，画像反転，回転，並進移動，拡大のデータオーグメンテーションを行った．学習の各エポックごとにデータオーグメンテーションを施し，エポック毎の画像枚数を 20000 まで拡張した．またミニバッチサイズは 50 とした．

[*†] https://github.com/thtrieu/darkflow

表 1: 牛領域画像の枚数.

cattle state	train	validation	test
standing	10170	1130	800
allanto & amnion	7612	844	800
tail rising	5714	634	800

表 2: 牛の状態検知システムの性能．表中の数値は AUC.

network	standing	allanto & amnion	tail rising
VGG16	0.95	0.90	0.56
ResNet50	0.98	0.92	0.48

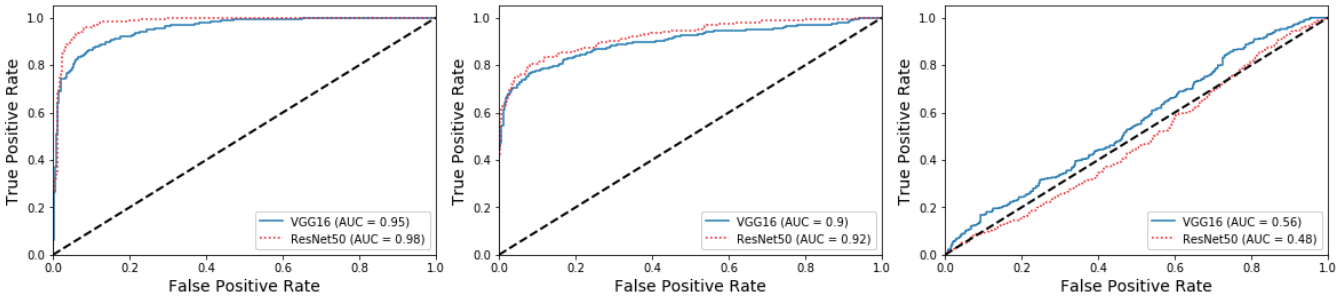
4.3 実験結果

図 4 に 3 つの分娩予兆に関わる状態をニューラルネットで検知した，それぞれの受信者操作特性 (Reciever Operation Characteristics, ROC) 曲線を示す．表 2 に ROC 曲線の下部の面積である Area Under the Curve(AUC) を示す．また図 5 は，VGG16 と同じ畳込み層の構造を持つニューラルネットワークにおいて，Grad-CAM [Selvaraju 17] を用いて得た結果を出力する上で注目された入力画像の領域を示す．

牛の起立状態と尿膜・羊膜の露出状態の AUC は 2 つの状態において 0.8 を超えている．Grad-CAM によって得られた入力画像における注目されている部位もこの結果を支持する．すなわち，起立状態の検知においては足の部分が注目されており，尿膜・羊膜の露出状態の検知においては尻の部分が注目されている．これらの結果は，起立状態および尿膜・羊膜の露出状態は動画からの分娩予兆の早期検知に有効であることを示唆する．対して，尾の挙上の検知性能は良くない．尾の挙上を検知では，尾の部分が注目されることが期待されるが，際立って注目されている部分はない．尾が挙上しているか，いないかの判断はクラウドワーカーにとって極めて難しく，尾の挙上のアノテーションが他の 2 状態に比べ信頼性が低いと考えられる事実，起立状態と尿膜・羊膜の露出状態のアノテーションの，2 人のワーカーの回答の一致率はそれぞれ 81.5%と 73.3%であった一方で，尾の挙上状態における一致率は 48.1%にとどまった．この結果は，クラウドソーシングを用いたアノテーションの設計を改良することが必要であることを示唆する．

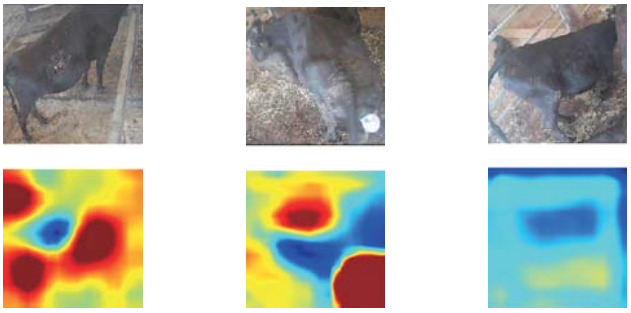
5. 結論

本研究では，画像情報から分娩予兆に関わる牛の状態を検知するシステムを開発した．CNN を用い，分娩時に観測される状態である牛の起立・尿膜及び羊膜の露出・尾の挙上の 3 状態を入力画像から検出した．ここで，それぞれの画像から切りだされた牛領域に対して，クラウドソーシングによるアノテーションシステムを開発しラベル付与を行った．これらの画像を用いた実験の結果，起立と尿膜・羊膜の露出に関しては良い性



(a) 起立状態　(b) 尿膜・羊膜の露出状態　(c) 尾の拳上状態

図 4: 3 つの牛の状態検知システムで得られた ROC 曲線.



(a) 起立状態　(b) 尿膜・羊膜の露出状態　(c) 尾の拳上状態

図 5: CNN による状態検知において顕著に利用される入力画像の領域．起立状態の検出 (a)，尿膜・羊膜の検出 (b) では，想定通り足と尻に注目があたっている．一方，尾の拳上状態の検出 (c) では，尾の部分に注目があたっていることが望ましいが，そうなっていない．

能で検出できたが，尾の拳上に関しては改善の余地があることが示された．

6. 謝辞

本研究を行うにあたり，牛のモニタリング動画を提供してくださった東村牧場様，また春日良一様 **, 高橋美博様 *, 山下博美様，田中知明様，そして有益な議論を頂いた春日久志様に感謝申し上げます．

参考文献

[Burfeind 11] Burfeind, O., Suthar, V., Voigtsberger, R., Bonk, S., and Heuwiser, W.: Validity of prepartum changes in vaginal and rectal temperature to predict calving in dairy cows, *Journal of Dairy Science*, pp. 5053–5061 (2011)

[Deng 09] Deng, J., Dong, W., Socher, R., Li, L., Li, K., and Fei-Fei, L.: ImageNet: A large-scale hierarchical image database, in *2009 IEEE Conference on Computer Vision and Pattern Recognition*, pp. 248–255 (2009)

[Everingham] Everingham, M., Gool, L. V., Williams, C., Win, J., and Zisserman, A.: The PASCAL Visual Object Classes Challenge 2007 (VOC2007) Results, http://www.pascal-network.org/challenges/VOC/voc2007/workshop/index.html

[He 16] He, K., Zhang, X., Ren, S., and Sun, J.: Deep Residual Learning for Image Recognition, in *The IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 770–778 (2016)

[Jensen 12] Jensen, M.: Behaviour around the time of calving in dairy cows, *Applied Animal Behaviour Science*, Vol. 139, pp. 195–202 (2012)

[Lin 14] Lin, T., Maire, M., Belongie, S., Bourdev, L., Girshick, R., Hays, J., Perona, P., Ramanan, D., Dollár, P., and Zitnick, C.: Microsoft COCO: Common Objects in Context, in *Proceedings of the European Conference on Computer Vision (ECCV)*, pp. 740–755 (2014)

[Mee 13] Mee, J.: Why Do So Many Calves Die on Modern Dairy Farms and What Can We Do about Calf Welfare in the Future?, *Animals*, pp. 1036–1057 (2013)

[Redmon 17] Redmon, J. and Frhadi, A.: YOLO9000: Better, Faster, Stronger, in *The IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 7263–7271 (2017)

[Saint-Dizier 15] Saint-Dizier, M. and Chastant-Maillart, S.: Methods and on-farm devices to predict calving time in cattle, *The Veterinary Journal*, Vol. 205, pp. 349–356 (2015)

[Selvaraju 17] Selvaraju, R., Das, A., Vedantam, R., Cogswell, M., Parikh, D., and Batra, D.: Grad-CAM: Why did you say that? Visual Explanations from Deep Networks via Gradient-based Localization, in *The IEEE International Conference on Computer Vision (ICCV)*, pp. 618–626 (2017)

[Simonyan 14] Simonyan, K. and Zisserman, A.: Very Deep Convolutional Networks for Large-Scale Image Recognition, *arXiv preprint arXiv:1409.1556* (2014)

** Farmer's Support, LLC.