

自動撮影カメラ等の動画を用いたニホンジカ自動判別技術の開発

石田 朗・釜田淳志¹・鈴木万里子（愛知県森林セ）江口則和²（北大院農）

安達貴広・柴田幸俊（MTGフォレスト）

シカのモニタリング調査の一環として自動撮影カメラで収集された静止画や動画から省力的にシカ情報を抽出・データ化するために、深層学習で画像・映像から自動的にシカの有無や頭数をデータ化する手法を開発した。自動撮影カメラ等で撮影された映像からシカが映っている3,607の動画ファイルを抽出し、これらを分解してシカがマーキングされたフレーム573,149を得た。同様にイノシシで84,475、カモシカで33,526、それ以外の動物等で278,122のフレームが得られた。これらを合わせたフレームのうち、80%でTensorflow Object Detection APIのSSD inception v2 モデルを使い深層学習を行った。残り20%を用いてシカの検出精度を検証したところ、平均適合率(AP)は85.4%となり、動画サンプル活用することで効率的に自動撮影カメラ等の画像を自動解析でデータ化できるモデルを構築できる可能性が示された。

キーワード：ニホンジカ、自動撮影カメラ、動画、自動判別、深層学習

I はじめに

ニホンジカ (*Cervus nippon*, 以下、シカ) による様々な被害が各地で深刻な状況となっている(8)。愛知県では、HPを活用したアンケート調査システム「シカ情報マップ」(<https://shikadoko.jp/>)を(株)マップクエストや(国研)森林総合研究所と共同開発し、省力的に県内のシカとその被害の分布やその変化を把握する取り組みを進めている(5, 7)。また、県内の地域ごとにより詳細なシカの生息状況や動向を把握するために、自動撮影カメラを活用している(6)。しかしながら、自動撮影カメラで収集した静止画・動画サンプルは膨大になり、個々に確認して、シカの頭数を数えるのに多大な労力が必要となる。そこで、本研究では深層学習を用いて画像情報から自動的にシカの有無や頭数をデータ化する手法の開発を行った。

II 材料と方法

1. 調査地およびデータ収集

愛知県の北設楽群設楽町田峯、北設楽群豊根村坂宇場、豊田市稻武、岡崎市下山、新城市黄楊野、新城市上吉田、新城市作手徳定、および奈良県の奈良公園の8地域で動画映像を収集した。愛知県では平成28~29年度にかけて、各地域に4~15台の自動撮影カメラ(BMC SG560K-8mHDとBushnellトロフィーカムXLT HD MAX)を設置した。撮影は24時間稼働で行い、撮影時間は10秒/回、撮影間隔は1分とした。また、奈良県では、平成31年3月にデジタルビデオカメラ(Sony HDR-PJ790V)で撮影を行い、撮影時間1分

以内とした。収集した映像はカラー及び3チャンネルのグレースケールであり、映像ごとに映っている動物の種類と個体数を確認した。

2. 画像処理

アノテーションツールVATIC(Video Annotation Tool from Irvine, California)(3)を用いて、動画内に映っているシカやその他の動物などの部分に矩形のマーキングを施すとともに(図-1)、動画を構成する静止画(以下、フレーム)に分解した。



図-1. アノテーションツールVATICのシカ等動物のマーキング作業画面

見やすくするため、ボックスを強調し、矢印で注釈を加えた。

3. シカ情報の自動判別学習および検出精度評価

軽量なモデルであるSingle Shot MultiBox Detector(SSD)(10)をベースとしたTensorflow Object Detection APIで公開されているSSD inception v2 モデルのうち、

ISHIDA Akira*, KAMATA Atsushi, SUZUKI Mariko, EGUCHI Norikazu, ADACHI Takahiro, SHIBATA Yukitoshi

Development of technique for automatic discrimination of animal species using video footage by automatic photographic camera and video camera

akira_ishida@pref.aichi.lg.jp

¹現所属：愛知県新城設楽農林水産事務所、²現所属：人間環境大学人間環境学部



図-2. 得られたフレームの例

シカやその他動物、人がバウンディングボックスでマーキングされる（見やすくするため、ボックスを強調し、矢印で注釈を加えた）。動画を構成する静止画に分解することで、動物の様々な角度でのサンプルが得られる。

転移学習を行うための学習済みモデルである SSD inception v2 atrous coco (9) を用いて、愛知県内で収集した映像からシカの含まれるフレームについて深層学習を進めた。

学習には、アノテーションで得られたフレームのうち愛知県内のシカが映っている 80%をランダムに選定して用い、バッチサイズを 24 枚、エポック数を 3 とした。1 エポック当たりの学習回数（バッチ数）は、得られたシカのフレーム数×0.8/24 である。

構築されたモデルについては、残り 20% の全フレームを用いて、Tensorflow Object Detection API の検証プログラムで算出される平均適合率 (Average Precision, 以下 AP) により評価を行った。具体的には、バウンディングボックスの一致度 (Intersection over Union, 以下 IoU) を出した。IoU は、真のバウンディングボックス R_p と予測されたバウンディングボックス R_g の一致度合を共通部分/和集合 (①式) として定義され (4), シカがいない箇所をシカとして判別した場合, $\text{IoU} = 0$ となる。

$$IoU = \frac{area(R_p \cap R_g)}{area(R_p \cup R_g)} \dots \dots \dots \quad (1)$$

つぎに、IoU が閾値 0.5 を超えたものを正答とし、再現率 r の時の精度 $p(r)$ の平均として AP を求めた。

以上の学習で使用した PC の CPU は intel i7-6700K, メモリは 32G, グラフィックスは GTX 1080 (NVIDIA 社) であった。

4. シカおよびそれ以外の動物情報の自動判別学習および検出精度評価、アプリ化

2.と同様の方法および同じPCを用いて、愛知県および奈良県で収集された映像のうち、シカ、イノシシ、カモシカおよびそれ以外（3種以外の動物、鳥、人など）の4つの分類について、それらが含まれるフレームについて1つのモデルで物体検出とその判別を行う

深層学習を進めた。学習には、アノテーションで得られたフレームのうち動物が映っている 80%をランダムに選定して用い、バッチサイズを 24 枚、エポック数を 3 とした。

構築されたモデルについては、残り 20% の全フレームを精度の評価に用いて、動物の 4 つの分類についてそれぞれで 3. と同様の方法で AP を求めた。動物がない箇所を動物として判別した場合や動物がいる箇所でも別の動物として判別した場合、IoU=0 となり、AP を低下させる。

III 結果

収集した動画映像について動物の映っているものを選別し、動物の種類を判別したところ、愛知県内で収集した自動撮影カメラ動画で 7,270（うちシカは 3,324）、奈良県で収集したビデオカメラ動画で 283（うちシカは 283）の映像が得られた。アノテーション作業の結果、愛知県のシカが映っていた動画映像からは、573,149 のフレームが得られた。また、奈良県のシカが映っていた動画映像を合わせて、634,452 のフレームが得られた。さらに、イノシシでは 84,475、カモシカでは 33,526、それ以外ではノウサギ、タヌキ、キツネ、アナグマ、アライグマ、ハクビシン、ネコ、テン、ホンドリス、ヤマドリ等の鳥類、人などで 278,122 のフレームが得られた（図-2）。

愛知県のシカについて得られたフレームを用いて構築されたモデルでは、APは90.07%であった。また同様の方法で愛知県および奈良県のシカ以外の動物について得られたフレームを用いて構築されたモデルでは、APはシカで85.4%，イノシシで51.0%，カモシカで68.50%，それ以外の動物で69.90%であった。

IV 考察

本研究では動画からのフレームサンプルを用いることで、量的・質的なデータを増やすことができ、APはシカでは約85%，カモシカやそれ以外の動物で約70%と効率的に判別学習を進めることができたと考えられる。一方で、イノシシではAPが約50%と低かった。フレームサンプル数がより少ないカモシカよりも形状や色彩での判別の難易度が高かったためと推察される。

構築したモデルの実用化に向けての課題としては、動画をフレームに分解してデータ数を増やす手法を用いたが故に、モデル構築データと検証データに類似のものが多く含まれており、今回構築されたモデルを他の場所でのデータに適用した場合に精度が落ちる可能性が考えられる。安藤ら(1)は約12万枚の静止画サンプルを用いて、すでに深層学習を活用してシカだけの画像では最大96.4%，シカ以外を含めた画像ではシカで79.6%，イノシシで84.4%，カモシカで91.6%の正答率の動物判別モデルを構築し公開している。本研究で用いたフレームは63万を超えていたが、元となる動画数は約8千であり、そこを種類ごとのサンプル数のバランスを取りながら量的・質的に増強することで各種の判別精度が上がることが期待される。また、Beeryら(2)は、開発に用いた画像を得た場所以外のサンプルでの正答率については同様のレベルを得られない可能性を指摘している。本学習モデルについても今後実用化に向けて様々な場所で撮影されたサンプルを用いて正答率を検証し、改良を積み重ねることも必要と考えられる。

本研究で構築したシカ、イノシシ、カモシカ、その他の動物についての学習モデルを搭載した自動動物判別アプリ(図-3)を作成した。アプリはWindows環境のみで使用が可能で、以下のURLで公開している。

<http://www.mtg-forest.com/index.php/works/adammmm>

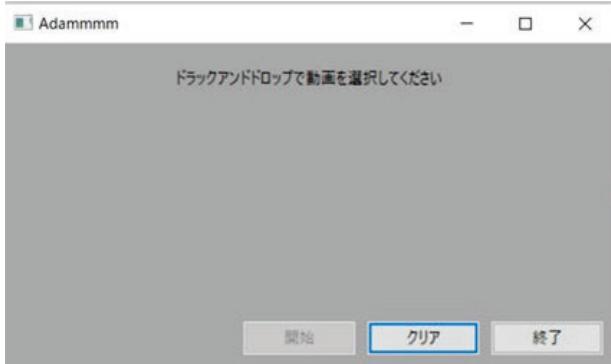


図-3. 動物自動判別アプリの画面

画面内に動画や静止画のファイルをドラッグ&ドロップで各ファイルの判別結果(動物種と頭数)がCSV形式で得られる。

謝辞

本研究は、戦略的情報通信研究開発事業(No.172306001)の支援を受けて実施した。

引用文献

- (1) 安藤正規・中塚俊介・相澤宏旭・中森さつき・池田敦・森部絢嗣・寺田和憲・加藤邦人 (2019) 深層学習(Deep Learning)によるカメラトラップ画像の判別. 哺乳類科学 59: 49-60
- (2) Beery S, Van Horn G, Perona P (2018) Recognition in Terra Incognita arXiv preprint arXiv:1807.04975
- (3) Vondrick C, Patterson D, Ramanan D (2013) Efficiently Scaling Up Crowdsourced Video Annotation. A set of best practices for high quality, economical video labeling. Int. J. Comput. Vis. 101: 184-204
- (4) 原田達也 (2017) 画像認識(機械学習プロフェッショナルシリーズ). 講談社
- (5) 石田朗・江口則和 (2018) シカの森林被害の軽減化に向けて -実態調査とICTの活用-. 現代林業 625: 46-50
- (6) 石田朗・釜田淳志・栗田悟・日置順昭 (2019) 愛知県における自動撮影カメラを用いたシカの生息状況調査. 中森研 67: 39-40
- (7) 釜田淳志・石田朗・栗田悟・江口則和・寺田行一・早川雅人・佐藤亮介・高橋啓・立脇隆文・安達貴広・岡輝樹 (2019) 「シカ情報マップ」の利用状況と効果的な情報収集手法の検討. 中森研 67: 41-42
- (8) 小泉透 (2011) 拡大するシカの影響. 森林科学 61: 2-3
- (9) Tensorflow detection model zoo.
https://github.com/tensorflow/models/blob/master/research/object_detection/g3doc/detection_model_zoo.md
- (10) Wei L, Anguelov D, Erhan D, Szegedy C, Reed S, Fu C, Berg A C (2016) SSD: Single Shot Multibox Detector. Europ. Conf. Compt. Vis. 9905: 21-37