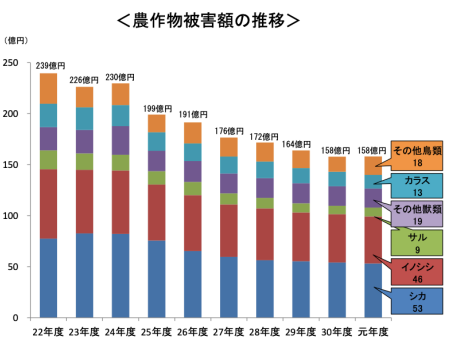
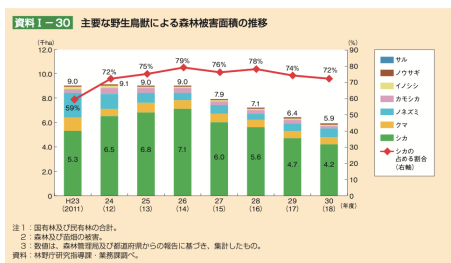


機械学習を用いた野生生物の行動認識

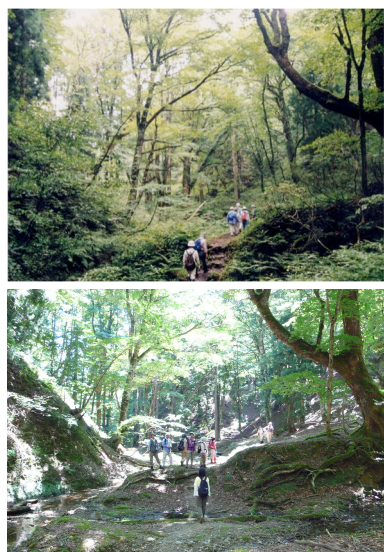
京都大学大学院農学研究科地域環境科学専攻
フィールドロボティクス分野
木俣 稜



野生鳥獣による農作物被害額の推移
<https://www.maff.go.jp/j/seisan/tyozyu/higai/attach/pdf/index-384.pdf>より引用



野生鳥獣による森林被害面積の推移
<https://www.rinya.maff.go.jp/j/kikaku/hakuso/r1hakuso/attach/pdf/zenbun-6.pdf>より引用



京都大学芦生研究林内の同一地点の画像
(上：被害前 下：被害後)

定点カメラ



カメラトラップのイメージ

日本における獣害

国内ではイノシシ・シカ等の野生生物による農作物や森林への被害、「獣害」が深刻な問題です。獣害は単なる一時的な経済的被害だけではなく、離農などによる食糧生産の低減や生態系破壊、土壌流出を招く可能性があるなど将来的にも影響を及ぼす問題であり、対策が必要です。

このような状況の下、近年では自治体レベルで捕獲や狩猟が推進された結果、イノシシやシカなどの個体数は現在減少傾向にあるとみられています。しかし依然として被害は大きく、更なる対策が必要です。

獣害対策

効果的、効率的な獣害防止対策のためには対象生物の行動を分析することが重要です。対象の行動の分析によって効果の高い予防策を検討することや、猟をする際に対象の行動パターンから最も高い狩猟成績があげられると予測されるポイントで待ち構えることが可能になるためです。

この野生生物の行動分析手法の一つに「カメラトラップ」があります。これは対象地点に設置した定点カメラの前を通過した生物を撮影したのち、この画像または映像を人の目で確認して行動を記録するという手法です。個体の詳細な動きを調査できる、対象生物にストレスを与えないといった利点がありますが、労力が高いという問題があります。

先行研究

既往の研究ではM. S. Norouzzadehら^a(2018)、安藤ら^b(2019)など、機械学習を用いて生物種の同定と頭数のカウント、行動分類等を行なった事例があります。しかし、これらのモデルは画像もしくは動画のフレーム1枚に対して予測を行うものであり、動画全体の分析を行なっているものではありません。

研究目的

本研究の目的はカメラトラップの動画から野生生物の行動を自動で分析することです。機械学習を用いて動画から特徴量を抽出し、対象生物の数や位置、行動について出力するモデルを検討します。

また、対象生物は最も大きな被害をもたらすシカとし、カメラトラップ動画のデータセットには昼間に可視光で撮影したカラー動画と夜間に近赤外領域で撮影したモノクロ動画が存在します。

研究アプローチ

動画を直接入力として分析する機械学習手法は3D Convolution (S. Jiら^a, 2013)などが挙げられます。しかし、3D Convolutionは学習パラメータが多いため必要データ数も多くなるという特徴があり、大量のデータを集めることが難しいカメラトラップとは相性が良くありません。

そこでタスクを①部位検出：動画から対象生物の部位の軌跡情報を抽出するパート、②行動分類：軌跡情報から行動を出力するパートの2つに切り分けました。このように分割することで難しい問題をより単純なタスクに落とし込み課題解決を図ります。

①部位検出

部位検出パートでは動画の各フレームで対象の頭部と下半身の位置を検出したのち、その位置が時間経過とともにどのように変化したかを示す軌跡情報を画像として出力します。

部位の検出器はYOLOv3(J. Redmon^d, 2018)を転移学習させたモデルを使用します。画像の収集が比較的容易な奈良公園で撮影したシカの画像とカメラトラップで収集した画像を合わせたデータセットで学習を行うことで、カメラトラップ画像のテストにおいて2クラス平均で正答率95%、発見率85%を達成しました。

この部位の検出器を動画の一定間隔のフレームで作動させ、部位ごとに検出した領域の軌跡を描きます。軌跡画像では寒色が頭部、暖色が下半身、色のグラデーションで時間経過をそれぞれ表しています。

②行動分類

行動分類パートでは部位検出パートで出力した部位の軌跡情報を示す画像(以下Box Flow)をNeural Network (NN)の入力として行動の分類を行います。

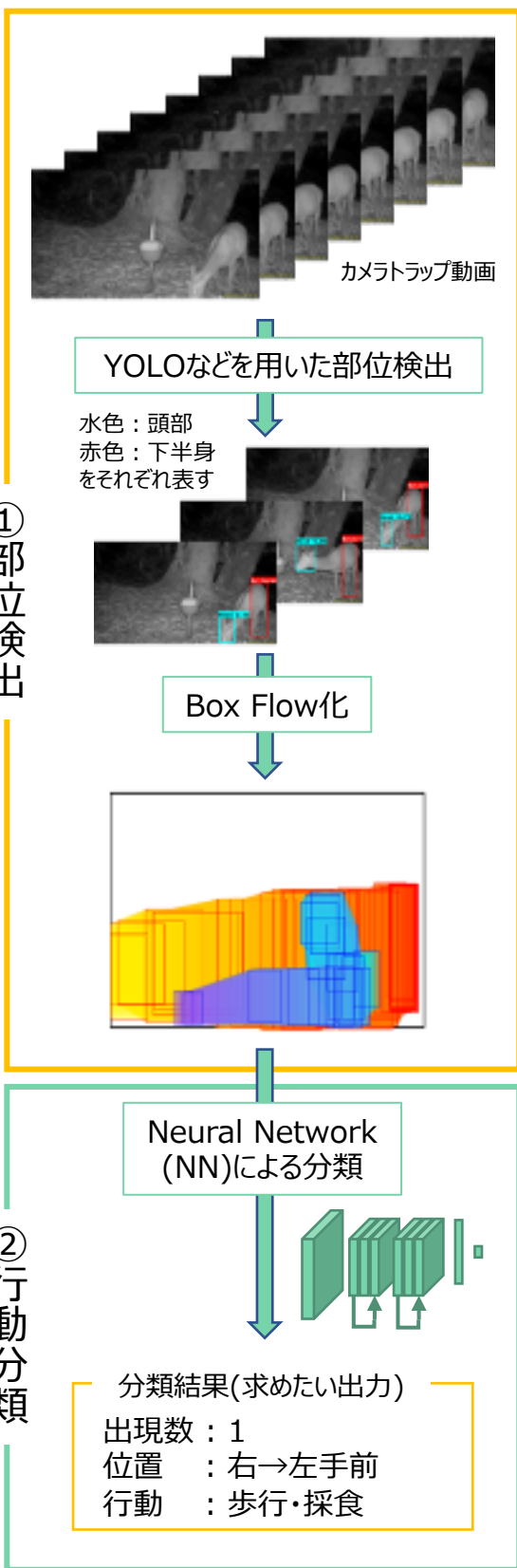
出力は位置、頭数、行動などですが位置と頭数は部位検出の情報から判定が可能であると考えられるためNNは行動についてのみ出力を行います。

本手法では動画全体の情報がBox Flow 1枚に圧縮されているため計算量が少なくなる、動画の長さによらず同じモデルで分析が可能といった利点があります。

今後の展望

①について、現在シカが複数写っている場合は軌跡を個体ごとに独立に描くことができないため、②の出力に影響することが考えられます。今後は同一個体の部位同士をマッチングする機能が必要です。

また、②についてはBox Flowを用いた手法の精度を評価しつつ、①の途中で生成される特徴ベクトルをモデルの入力とする手法など他の手法も検討し、精度の高い分類手法を探索します。



参考文献

- ^a Norouzzadeh, M. S., et. al. (2018) Automatically identifying, counting, and describing wild animals in camera-trap images with deep learning, *Proceedings of the National Academy of Science of the United States of America* 115:E5716-E5725.
- ^b 安藤 正規, 他. (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(1):49-60.
- ^c Ji, S., et. al. (2013) 3D Convolutional Neural Networks for Human Action Recognition *IEEE Transactions on Pattern Analysis and Machine Intelligence* 35(1):221-231.
- ^d Redmon, J., et. al. (2018) YOLOv3: An Incremental Improvement, <https://arxiv.org/abs/1804.02767>. (2021年2月10日アクセス)
- 鳥獣被害の現状と対策 (令和3年1月), <https://www.maff.go.jp/j/seisan/tyozyu/higai/attach/pdf/index-384.pdf>. (2021年2月10日アクセス)
- 令和元年度 森林・林業白書 第3節 森林保全の動向, <https://www.rinya.maff.go.jp/j/kikaku/hakusyo/r1hakusyo/attach/pdf/zenbun-6.pdf>. (2021年2月10日アクセス)