

深層学習モデルも間違える：

不完全な個体識別を伴う動物の密度推定

矢島豪太（日本大学） 東出大志（石川県立大学） 中島啓裕（日本大学）

1. はじめに

野生動物の個体数密度推定は、生態学および保全生物学における重要な課題の一つである。特に空間明示型捕獲再捕獲（Spatially Explicit Capture–Recapture, SECR）モデルは、個体の空間情報と検出プロセスを統合することで信頼性の高い密度推定を可能にする手法として広く用いられている（Borchers and Efford 2008）。この手法の主要な要件は、観測されたすべての個体の識別が可能なことである。この点で、カメラトラップ（赤外線センサーが反応すると自動で撮影するシステム）は比較的データ取得コストが低いこと、また外見的特徴から個体の識別が可能であることから有効なデータ取得手段である。しかし、画像から個体を識別する作業は多大な労力を要し、SECR の大規模・長期適用を妨げている。

近年、個体識別の自動化に畳み込みニューラルネットワーク（CNN）による深層距離学習が活用されている。この手法は、画像から個体固有の特徴を多次元ベクトル（画像特徴量）として抽出し、同一個体間の類似度を最大化するように学習する（Vidal et al. 2021）。海域から陸域まで様々な動物種へ適用が進む一方で、この深層距離学習による個体識別を SECR に統合する際には方法論上の課題がある。従来は、モデルの出力を確定的な個体 ID として扱ってきたが、深層学習による判定には常に不確実性が伴う。これを無視すると SECR の推定値に大きなバイアスが生じる。従って、その不確実性を統計的に取り込む枠組みが求められる。

この課題に取り組むうえで参考になるのが、Unmarked SECR model（Chandler and Royle 2013）である。このモデルは、不確実性を伴う識別情報から個体 ID を確率的に推論する枠組みを提供してきた。具体的には、個体の空間的な行動制約（行動圏）を仮定して、観測されたパターンを最もよく説明する個体数と各個体の活動中心を同時に推定する。言い換えれば、空間分布を個体識別のシグナルとして活用しているといえる。この発想は、画像特徴量にも応用可能であると考えられる。すなわち、特徴量の分布を個体識別のシグナルと見做し確率分布として表現することで類似の統計的推論が実現できる。さらに、従来の空間情報も引き続き利用できるため、より頑健な密度推定が見込

める。

そこで本研究では、このアイデアを実現する新たな SECR クラスのモデルを開発した。本モデルではまず、特徴量の不確実性をパラメータとして明示的に組み込み、同一個体から得られる特徴量の変動を統計的に表現した。さらに Unmarked SECR の枠組みを拡張し、トラップ位置と画像特徴量の両方から個体帰属確率を推定できるようにした。モデルの検証には、実際のツキノワグマ (*Ursus thibetanus*) データに深層距離学習を適用して得られた画像特徴量を用いたモンテカルロシミュレーションを実施した。

2. 統計モデル

本研究で提案する統計モデルは、1) トラップと個体の空間情報を考慮した撮影プロセスと、2) 深層学習による画像特徴量の観測プロセスの2つをモデリングの対象とした。そして、これらを階層ベイズモデルとして統合した。

2.1. 個体の撮影プロセス

調査領域 $A \in \mathbb{R}^2$ において J 箇所のカメラトラップを設置し、個体 $i = 1, \dots, N$ の活動中心 $\mathbf{s}_i$ が A 内に一様分布すると仮定した。一般的な SECR モデルに基づき、各トラップ位置 $\mathbf{x}_j, j = 1, \dots, J$ における個体 i の検出回数 $y_{i,j}$ (以下、捕獲履歴) の期待値を、活動中心とカメラの距離に応じて減衰する関数として設計した。具体的には半正規関数として以下のように定義した:

$$\lambda(\mathbf{s}_i, \mathbf{x}_j) = \lambda_0 \exp\left(-\frac{\|\mathbf{s}_i - \mathbf{x}_j\|_2^2}{2\sigma_\lambda^2}\right)$$

ここで λ_0 は、 $\|\mathbf{s}_i - \mathbf{x}_j\|_2^2$ が 0 の時の平均検出回数、 σ_λ はスケールパラメータである。

2.2. 特徴量の観測プロセス

個体の撮影プロセスと同様に、特徴量観測プロセスでは、 L 次元特徴空間 E 上に一様分布するクラスター中心 $\mathbf{g}_i^{true} \in \mathbb{R}^L$ を導入した。そして、 $\mathbf{g}_i^{true}$ から観測特徴量 $\mathbf{g}_{m,l}^{obs}, m = 1, \dots, \sum_{i,j} y_{i,j}, l = 1, \dots, L$ が不確実性を伴い生成されるものとして表現した。捕獲履歴 $y_{i,j}$ が既知の場合、サンプル m に対応する個体の ID は一意に定まる。我々はこれを変数 $ID_m \in \{1, \dots, N\}$ として明示的に扱った。この ID_m のもとで、観測される特徴量 $\mathbf{g}_{m,l}^{obs}$ の生成過程を $\mathbf{g}_{i,l}^{true}$ とパラメータ σ_g をもつ正規分布として記述した:

$$\mathbf{g}_{m,l}^{obs} | ID_m = i \sim \text{Normal}(\mathbf{g}_{i,l}^{true}, \sigma_g^2)$$

ここで σ_g は同一個体の異なる画像から得られる特徴量の変動を表し、深層学習モデル

の識別性能と撮影条件の変動を統合した不確実性パラメータである。この値が大きいほど、同一個体の画像から得られる特徴量のばらつきが大きく、個体識別の精度が低いことを意味する。

3. 推論アルゴリズム

我々は確定的な個体識別情報が得られないことを想定するため、 $y_{i,j}$ を潜在変数として扱った。そこで、マルコフ連鎖モンテカルロ (MCMC) 法によるデータ拡大法 (Tanner and Wong 1987) によって、 $y_{i,j}$ のサンプリングと、 $y_{i,j}$ が既知のもとでの他のパラメータの更新を逐次的に行った。以下ではまず空間情報にのみ基づく $y_{i,j}$ の構成について説明し、その後画像特徴空間を考慮した構成について説明する。

3.1. 空間情報のみに基づく捕獲履歴の構成

捕獲履歴 $y_{i,j}$ は、個体数 (i 方向の次元数 N) と各セルの値の2つが欠測するため、これらを補完する必要がある。1つ目については、 N より十分大きい M を既知の定数として設定し、潜在変数 $z_i \sim \text{Bernoulli}(\psi), i = 1, \dots, M$ を導入した。この二項指示変数は、ある個体が調査エリア A 内に存在するかどうかを表し、次元数を M で固定する役割を持つ。2つ目については、ポアソン分布を和で条件付けると多項分布になる性質を利用した。具体的には、 $y_{i,j}$ が未知であるが、トラップ j における総観測回数 $n_j = \sum_i y_{i,j}$ は既知であるため、以下の関係性が成り立つ:

$$\{y_{1,j}, \dots, y_{M,j}\} \sim \text{Multinomial}(n_j, \{\pi_{1,j}, \dots, \pi_{M,j}\})$$

$$\pi_{i,j} \propto \lambda(\mathbf{s}_i, \mathbf{x}_j) z_i$$

この多項分布によって $y_{i,j}$ の生成が可能となる。

3.2. 画像特徴空間への拡張

上記のサンプリング式に画像特徴量を組み込むために、捕獲履歴 $y_{i,j}$ をイベント単位に分解し、多項分布をカテゴリカル分布に置き換えた。同様に、構成対象も $y_{i,j}$ から ID_m に置き換えた。イベント単位で定義される $\pi_{i,m}$ は、サンプル m と個体 i の空間的な近さと特徴量中心への近さを考慮して以下のように定義した:

$$\pi_{i,m} \propto \lambda(\mathbf{s}_i, \mathbf{x}_{j(m)}) z_i d(\mathbf{g}_i^{\text{true}}, \mathbf{g}_m^{\text{obs}})$$

ここで、 $j(m)$ はサンプル m の観測トラップを示した。また、 $d(\mathbf{g}_i^{\text{true}}, \mathbf{g}_m^{\text{obs}})$ は、正規分布の確率密度を用いて、 $\prod_{l=1}^L \text{Normal}(\mathbf{g}_{m,l}^{\text{obs}}; \mathbf{g}_{i,l}^{\text{true}}, \sigma_g^2)$ とした。この確率密度が高い場合は、 $\mathbf{g}_{m,l}^{\text{obs}}$ に対する $\mathbf{g}_{i,l}^{\text{true}}$ の割り当てが適切であることを示し、それは $\mathbf{g}_{m,l}^{\text{obs}}$ と $\mathbf{g}_{i,l}^{\text{true}}$ の距離が近いことを示す。MCMC の各イテレーションで、観測データを条件と

して、潜在変数 ID_m をサンプリングし、これに基づいて捕獲履歴 $y_{i,j}$ を構築した。
発表では、モデルの詳細とシミュレーション結果を報告する。

4. 引用文献

- Borchers, D. L., and M. G. Efford. 2008. Spatially Explicit Maximum Likelihood Methods for Capture–Recapture Studies. *Biometrics* 64:377–385.
- Chandler, R. B., and J. A. Royle. 2013. Spatially Explicit Models for Inference about Density in Unmarked or Partially Marked Populations. *The Annals of Applied Statistics* 7: 936–954.
- Tanner, M. A., and W. H. Wong. 1987. The Calculation of Posterior Distributions by Data Augmentation. *Journal of the American Statistical Association* 82:528–540.
- Vidal, M., N. Wolf, B. Rosenberg, B. P. Harris, and A. Mathis. 2021. Perspectives on Individual Animal Identification from Biology and Computer Vision. *Integrative and Comparative Biology* 61:900–916.