

深層学習とロジスティック回帰を組み合わせたシロアリ追跡手法の提案

石川 愛唱¹ 小島 俊輔¹ 木原 久美子¹

概要: シロアリの小集団における個体の行動の解析において、手作業による ID 追跡は多くの時間と労力を要するだけでなく、解析可能なデータ量には限界がある。本研究では、物体検出モデル YOLO とロジスティック回帰およびハンガリアン法を統合した自動追跡手法を提案する。特徴量としてバウンディングボックスの重複度や信頼度変化等を用いたロジスティック回帰による同一個体判定とハンガリアン法による割り当てにより、ID 紐づけを行った。評価の結果、MOTA が 0.952 となり、提案手法が ID 追跡の基盤技術として有用であることを示した。今後の展望として、新たな特徴量の選定と学習データ数の増強が挙げられる。

キーワード: 機械学習・データマイニング - すべて、回帰、動画認識・理解、情報可視化

Proposal of a Termite Tracking Method Combining Deep Learning and Logistic Regression

AISHO ISHIKAWA^{†1} OSHIMA SHUNSUKE^{†1} KIHARA KUMIKO^{†1}

Abstract: Manual tracking in analyzing individual behavior within small termite colonies not only requires significant time and effort but also imposes limitations on the amount of analyzable data. This study proposes an automated tracking method integrating the object detection model YOLO with logistic regression and the Hungarian algorithm. Individual identification was performed using logistic regression with features such as bounding box overlap and confidence change, followed by assignment via the Hungarian algorithm to establish ID binding. Evaluation results yielded a MOTA of 0.952, demonstrating the usefulness of the proposed method as a foundational technology for ID tracking. Future prospects include selecting new features and augmenting the training dataset.

Keywords: Machine Learning/Data Mining - All, Regression, Video Recognition/Understanding, Information Visualization

1. はじめに

ヤマトシロアリは社会性昆虫の一種であり、個体が集団として行動することでその生存戦略が機能するという特徴を持つ。そのため、個体の行動を詳細に観察・解析することは、集団全体の行動理解に向けた重要な手がかりとなる。

先行研究[1]では、ヤマトシロアリを対象とし、シロアリの小集団における個体の行動を詳細な解析を行っていた。先行研究では、シロアリの行動量の解析のために、30頭のシロアリをシャーレの中に入れ、撮影した動画を用いて、手作業でシロアリの行動を1秒間隔で追跡(軌跡を追尾)することによって、軌跡の座標と時刻の情報を取得していた。しかし、手作業による追跡は多くの時間と労力を要し、解析可能なデータ量には限界がある。

一方、近年の機械学習や深層学習の発展により、画像認識や物体検出の分野では大きな成果が得られている。You Only Look Once (YOLO) [2]などのリアルタイム物体検出モデルや、複数物体追跡 (MOT : Multiple Object Tracking)

[3]のアルゴリズムは、監視カメラ映像や自動運転など様々な分野で実用化が進んでおり、生物の行動解析への応用も期待されている。ここで MOT とは、動画内の各フレームにおいて対象となる複数の物体を検出し、それらの同一性をフレーム間で関連付けることで、個々の移動軌跡を推定する技術である。その主な目的は、物体の交差や遮蔽が発生した場合でも、各個体に割り振られた一意の識別番号 (ID) を追跡期間全体に通じて維持することにある。

本研究では、深層学習ベースの物体検出モデルと複数物体追跡アルゴリズムを組み合わせたシロアリの自動追跡システムを構築し、先行研究における追跡作業の効率化を図ることを目的とする。

2. 前提条件と準備

2.1 研究の目的

本研究では、先行研究[1]で作成されたヤマトシロアリの集団行動を記録した動画データを用いて、個体レベルの行動解析を可能とする高精度な自動追跡システムの構築を目指す。具体的には、深層学習モデルによる物体検出と、距離や面積変化などの複数の特徴量を用いたロジスティック回帰による同一個体である確率の推定、ハンガリアン

¹ 熊本高等専門学校
National Institute of Technology, Kumamoto College

法による割り当てを組み合わせた追跡アルゴリズムを提案し、その有用性を検証する。

2.2 使用データセット

研究対象である動画データは、先行研究で作成された、シャーレ内にヤマトシロアリ（兵蟻）が約 30 頭入った環境を、約 30 フレーム/秒 (fps) で計 5 分間撮影したものである。評価では冒頭 30 秒間(約 900 フレーム)のみを対象とする。また、解像度は 1,026×1,022 ピクセルであった。

本システムの構築および評価にあたり、同一条件下で撮影された 2 種類の動画データ（動画 A、動画 B とする）を使用する。役割は以下の通りである。

1) 学習用データ（動画 A）：物体検出モデルである

表 1. 機械学習の実行環境

項目	内容
ホスト OS	Windows 11 Pro 24H2
CPU	AMD Ryzen 9 9950X3D 16-Core
GPU	NVIDIA GeForce RTX5090 32GB
RAM	128GB
GPU ドライバ	NVIDIA GeForceDriver 576.88
CUDA	12.8
CuDNN	9.1.2
環境管理	Conda 24.11.3
言語	Python 3.12.11

表 2. 本研究で使用した主要なライブラリと役割

ライブラリ	主な役割
Ultralytics 8.3.162	YOLO11 モデルの構築、学習、および推論実行
Torch 2.8.0	PyTorch による GPU 加速および深層学習のバックエンド
Statsmodels 0.14.6	ロジスティック回帰モデルの学習および統計的有意性の分析
Scipy 1.15.3	ハンガリアン法（線形和割り当て問題）による最適ペアリング
Opencv 4.11.0.86	動画フレームの読み込み、画像処理、および追跡結果の描画
Pandas 2.3.0	検出データおよび追跡結果の CSV 入出力とデータ管理
Numpy 2.3.0	特徴量計算におけるベクトル演算および行列操作
scikit-learn 1.8.0	教師データの分割および機械学習用前処理の補助
LabelImg 1.8.6	検出モデル学習のためのアノテーションツール

YOLO11 モデルの学習、およびロジスティック回帰モデルの係数の決定に使用する。

2) 評価用データ（動画 B）：構築した追跡システム全体の性能を検証し、既存手法との精度比較を行うためのテストデータとして使用する。

2.3 開発環境

本研究では、解析環境の再現性および独立性を確保するため、パッケージ管理システムである Anaconda Navigator を用いて構築した Conda 仮想環境上で開発を行った。ソースコードエディタは Visual Studio Code を用い、プログラム言語は Python を用いた。機械学習の実行環境を表 1 に示す。

2.4 使用した物体検出アルゴリズムとライブラリ

物体を追跡する際の基盤となる物体検出には、約 30 fps の動画フレームに対し、約 30 頭のシロアリ個体を連続的かつ高速に検出する必要がある。YOLO は、入力画像を一度のネットワーク通過で検出するワンステージ手法であり、高い推論速度と精度を両立させているため、上記の要件を満たす上で最も適切なモデルである。本研究では、安定した性能を誇る YOLO11 (Ultralytics ライブラリ) を使用することとした。

また、本システムの構築において、物体検出からロジスティック回帰による係数の決定、およびハンガリアン法によるデータ結合までの一連の処理を実現するため、複数のライブラリを組み合わせ使用した。本研究で使用した主要なライブラリを表 2 に示す。

3. 物体検出モデルの構築及び真値データ作成

3.1 学習データセットの構築とアノテーション

検出モデルの学習のためにはまず学習データが必要となる。学習用データ（動画 A）から、動画全体をカバーするように 5 分間の動画の中から静止画像 40 枚を抽出した。次に、抽出した 40 枚の画像すべてに対し、アノテーションツールである LabelImg を用いて、ヤマトシロアリを 1 頭ずつ、体全体が入るようにバウンディングボックス (BBBox) をマウス操作で設定した。アノテーションの実行の様子を図 1 に示す。各画像には 30 頭のシロアリが存在するため、合計で 40 枚×30 頭で計 1,200 個のアノテーションデータセットを構築した。

3.2 YOLO の学習設定

学習モデルは YOLO11 のうち、パラメータが多く検出精度が高い YOLO11x.pt (XLarge モデル) を用いた。作成したアノテーションデータセットのうち、訓練データセットに 80%、検証データセットに 20%それぞれ分割し、前者を YOLO11 の学習モデルに用いた。今回の学習はデータセットの総量が大きいこと、確率的勾配降下法の仕組みを考慮した上で、学習の繰り返し回数であるエポック数を



図 1. labelImg を用いたアノテーションの様子

150 に設定した。また、バッチサイズは 16 に、入力画像のサイズは、動画の解像度とほぼ同じである 1,024×1,024 に設定した。

3.3 YOLO の学習結果

モデルの学習における損失 (Loss) および平均適合率 (mAP: mean Average Precision) の変化を図 2 に示す。ここで、IOU とは真値の領域と予測された領域の重なり度合いを 0 から 1 の値で示す指標である。また、mAP とは物体検出モデルの正確性を示す代表的な指標であり、1.0 に近いほど検出性能が高いことを意味する。本研究では、IOU (Intersection over Union) 閾値を 0.5 とした際の指標である mAP50、および IOU 閾値を 0.5 から 0.95 まで変化した際の平均値である mAP50-95、の 2 種類に着目した。

図 2 の学習曲線を見ると、学習が進むに従い訓練データにおける損失 (train loss) および検証データにおける損失 (validation loss) が共に減少しており、モデルの学習が収束していることが分かる。同様に、mAP50 及び mAP50-95 も収束していることが確認できる。全 150 エポックのうち、149 エポック目では mAP50 は 1.000 であり、mAP50-95 は 0.776 となった。いずれも高い値となったことから、構築したモデルがヤマトシロアリの検出において、十分な性能を有していることを確認した。この結果から、以降では、推論データの作成では 149 エポック目の重みファイル (best.pt) を用いることとする。

3.4 推論データの生成

学習を完了した重みファイルの best.pt を用いてそれぞれ 2 種類の動画 (動画 A、動画 B) に対する推論を実行し、検出結果の CSV ファイルを生成した。この CSV ファイルにはフレーム、ラベル、信頼度、および BBox の頂点座標 (x1, x2, y1, y2) の 7 項目が記録されている。

YOLO による物体検出において、信頼度閾値 (Confidence Threshold) を一般的なデフォルト値 0.25 よりも低い 0.1 に設定した。これは、MOT アルゴリズムが持つ履歴情報や特徴量計算の能力を最大限に活用し、低信頼度の検出で

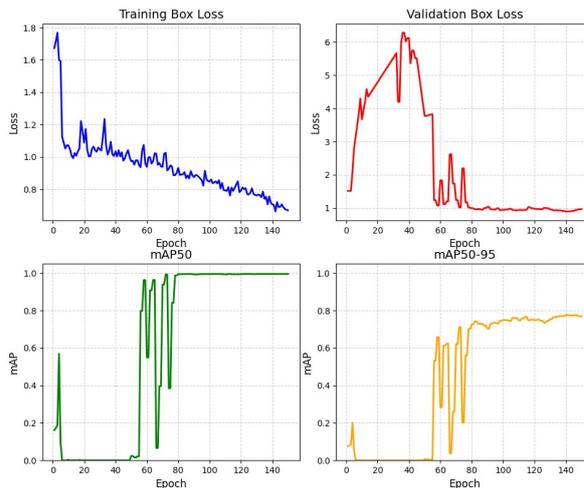


図 2. 学習時の Loss と mAP の変化

あっても軌跡に組み込む可能性を残すことで、追跡における情報の消失を防ぐことを目的としたものである。すなわち、提案する検出器は網羅度 (Recall) を優先し、ノイズの除去 (False Positive の抑制) は追跡時におこなう方針とする。

3.5 手動での ID 割り当て

ロジスティック回帰の式に用いる係数の学習および MOT アルゴリズムの性能を評価するためには、検出対象に一貫した ID を割り当てた真値データ (GT: Ground Truth) を作成しなければならない。

本研究では、システムの性能を客観的に評価するための真値 (Ground Truth) データとして、動画 A および動画 B の 0 フレームから 900 フレーム (計 901 フレーム) に対し、個体識別番号 (ID) を手動で割り当てて作成する。

そのためにはまず、動画を時系列順で横断しながら ID を割り当てていく作業が必要になる。一般的に物体検出で用いられる labelImg 等のツールは、単一フレーム内の物体にクラス名を付与することに特化したツールであり、動画全体を通じて個体の同一性を維持しながら ID を付与す

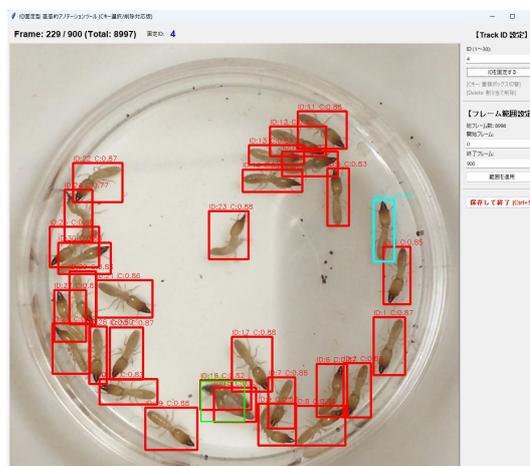


図 3. 手動 ID 割り当ての様子

る作業には適していない。適切なツールが見つからなかったため、Python により専用のツールを作成した。図 3 に手動 ID 割り当ての様子を示す。本ツールは、前後のフレーム情報を参照しながら、目視により ID を割り当てる機能を備えている。

真値データの作成は YOLO の推論で出力した BBox を対象としている。そのため、個体同士の密集等により YOLO が個体を検出できなかった未検出 (FN) 箇所については、真値データの対象から除外している。ヤマトシロアリ 30 頭が 901 フレーム分存在する場合、理論上の最大データ数は約 27,000 個となるが、上記の理由による未検出分を除いた実際の真値データ総数は、動画 A、動画 B それぞれ約 26,800 個となった。

4. 提案するマルチオブジェクト追跡手法

4.1 ID 結合の基本原則

提案する MOT 手法は、フレーム $t-1$ における既存軌跡とフレーム t における新規の検出を対応付けることで、対象ごとに一貫した ID を維持するように動作する。この対応付けには、ロジスティック回帰に基づく「マッチング確率」の推定と、ハンガリアン法 (線形和代入問題の最適解法) による一対一割り当てを組み合わせ合わせたアルゴリズムを採用した。これにより、局所的な判断に左右されず、総合的な結合コストを最小化する ID の割り当てが実現できる。

本手法では、まず、前後のフレーム間における個体の一致度を定量化するため、次に示す 5 つの情報を特徴量としたロジスティック回帰によりマッチング確率を求めることとした。

- 1) 中心座標の距離：前後のフレーム間における移動距離。極端に離れた検出はマッチング確率が低くなる。
- 2) IoU：BBox の重なり合い。数値が高いほどマッチング確率が高くなる。
- 3) BBox の面積変化率の絶対値：個体の急激な外見変化は考えにくいいため、BBox の面積の維持できているとマッチング確率が高くなる。
- 4) 信頼度変化の絶対値：YOLO が算出する信頼度スコアの変化量。個体の急激な外見変化は考えにくいいため変化量が高いほどマッチング確率は低くなる。
- 5) 速度差：過去 5 フレームの移動履歴から算出した平均速度と現在の速度を比較する。慣性の法則に基づき、速度の急激な変化が少ないほどマッチング確率は低くなる。

これらの特徴量は、1 フレームあたりのシロアリの物理的な移動限界や 1 フレーム後のシロアリの向きや大きさの変化が少ないことを考慮し、同一性の判定に寄与すると仮定して選定したものである。ここで、面積および信頼度の変化を絶対値として扱っているのは、同一性の判定にお

いて「変化の方向 (増加・減少)」ではなく「変化の度合い (安定性)」が重要なためである。

これらの特徴量からロジスティック回帰によりマッチング確率を算出し、各フレーム単位で軌跡の更新、一時的に検出が途切れた個体の保持、および新規個体の登録を逐次実行する。

4.1.1 ロジスティック回帰によるマッチング確率の算出

既存軌跡 i の $t-1$ 時点での検出結果 $D_{i,t-1}$ と、現在フレーム t の検出結果 $D_{i,t}$ が同一個体である確率 $P_{i,j}$ をロジスティック回帰モデルで推定する。入力特徴量は、先の 5 種類である。これらの特徴量ベクトルに対して、あらかじめ学習した係数と切片を用いて、以下の式(1)に基づいて、0 から 1 の範囲の確率を得る。

$$P_{i,j} = \sigma \left(\sum_{k=1}^5 w_k x_{k,i,j} + w_0 \right), \quad \sigma(z) = \frac{1}{1 + e^{-z}} \quad (1)$$

ここで、 $x_{k,i,j}$ は $D_{i,t-1}$ と $D_{i,t}$ 間の k 番目の特徴量、 w_k はその係数、 w_0 は切片であり σ はシグモイド関数である。このマッチング確率 $P_{i,j}$ を両個体の同一性を示すコスト行列の要素として用いる。

4.1.2 ハンガリアン法による最適マッチング

ロジスティック回帰により算出された確率 $P_{i,j}$ はマッチング確率の指標であり、これをそのまま最小化アルゴリズムに適用することは難しい。そこで、結合コスト $C_{i,j}$ を「 $1 - P_{i,j}$ 」として定義し、既存軌跡と新規検出のすべての組み合わせを要素とするコスト行列を構築する。このコスト行列に対し、線形和代入問題の最適解法であるハンガリアン法を適用する。個々のペアに対してマッチング確率が最大となる相手を順次割り当てる貪欲法では、部分最適化にとどまる可能性がある。これに対し、ハンガリアン法は全組み合わせを考慮した全体最適解を導出できるため、MOT におけるデータ関連付けの標準的な手法として広く採用されている。本手法では、算出されたマッチング確率が一定値以上の場合に同一の軌跡として結合を行うが、この判定基準をマッチング閾値 (MATCH_THRESHOLD) と定義する。マッチング閾値を高く設定するほど誤結合 (FP) は減少するが、一方で個体の重なり等で一時的に確率が低下した際に軌跡が途切れやすくなる。

4.2 特徴量の係数学習

提案手法のロジスティック回帰モデルは、検出の対応付けに用いる先の 5 種類の特徴量に対する係数 (傾き) と切片を学習することで、既存軌跡と新規検出が同一個体である確率を推定する。学習は、手動で作成した真値データ (GT) から正例と負例のペアを構築し、統計モデルを用いてパラメータを推定する。

4.2.1 教師データの準備とペア構築

ロジスティック回帰モデルの学習には、手動で個体 ID

表 3. 特徴量に対する係数と Z 値と P 値

特徴量	係数	Z 値	P 値
Dist	-0.1252	-2.2613	0.0237
IoU	12.9361	11.8698	0.0000
Area_change_abs	-0.9921	-2.0266	0.0427
Conf_change_abs	-5.4218	-12.7322	0.0000
Delta_V	0.0021	0.0413	0.9671

を付与した真値データ (GT) を用いる。教師データとして、連続する 2 フレーム間において同一の ID を持つ個体のペアを結合すべき対象 (正例) とし、異なる ID を持つ近接個体のペアを結合すべきでない対象 (負例) として抽出した。

負例の選出にあたっては、単にランダムに選ぶのではなく、正例の個体からユークリッド距離に近い順に上位 3 つの異なる ID を持つ BBox を選択する手法を採用した。これにより、正例と負例の比率を 1:3 に固定しつつ、識別が困難な境界付近のデータを重点的に学習させることが可能となる。

作成されたデータセットの内訳は、正例数が 26,768 個、負例数が 80,304 個であり、合計 107,072 個となった。この全データを訓練用データとして 80%、モデルの検証および統計的寄与度の分析用として 20% の割合で分割した。なお、学習の再現性を確保するため、分割時の乱数シード値は固定した。

4.2.2 係数の学習結果と寄与度分析

表 3 に各特徴量に対応する係数、Z 値および P 値を示す。ここで、Dist は距離、IoU は重なり度合い、Area_change_abs は面積変化率の絶対値、Conf_change_abs は信頼度変化の絶対値、Delta_V は速度差を表す。Z 値は、各説明変数の係数をその標準誤差で割った値であり、その絶対値が大きいほど目的変数への影響が大きいことを意味する。P 値は、各説明変数が統計的に有意であるかどうかを示し、一般的には P 値が 0.05 未満の場合、その説明変数は統計的に有意と判断される。本学習結果では、IoU と Conf_change_abs は他の特徴量と比較して Z 値の絶対値が大きく、かつ P 値がいずれも 0.0000 となった。つまりマッチング確率の推定においてこれら 2 指標の寄与度が高い。係数の符号より、IoU の増大はマッチング確率が上昇し、信頼度変化はマッチング確率が低下することがわかる。一方、Dist と Area_change_abs についても P 値が 0.05 未満であり、有意性が認められた。係数が負であることから、距離や面積変化率が增大するほど、マッチング確率は低下する。これら 2 つの指標の P 値は IoU 等と比較して大きい、0.05 未満であり、有向な特徴量である。Delta_V については、P 値が 0.9671 と有意水準 0.05 を大幅に上回っており、マッ

ング確率への寄与は認められなかった。

4.3 マッチング閾値の設定

係数学習済みのロジスティック回帰モデルを真値データに適用し、連続フレーム間の真の関連付けに対するマッチング確率の分布を調査した。その結果、正解ペアにおけるマッチング確率は 0.8 未満のものが極めて少なく、大半が 0.9 以上に集中していることが確認された。

本研究では、ID の保持性の向上と、ID スイッチ (IDSW: 追跡対象の ID が途中で入れ替わる現象) を低く抑えることを優先するため、閾値を、MATCH_THRESHOLD = 0.2 と低めに設定することとした。

この設定値における定量検証の結果、全 26,768 ペアの正解データのうち 26,670 ペアの ID を維持しており、保持率は 99.63% であった

4.4 フレーム保持と ID 割り当ての制御

最後に、今回用いた ID の保持性と誤検出 (FP) の抑制を両立するための工夫について説明する。

4.4.1 フレーム保持

既存 ID が一時的に未検出になった場合でも、即座に ID を消滅させず、次フレームの対応付け候補として一定期間保持する。本研究では、見逃し上限のフレーム数を 30 と設定した。これにより、短時間の未検出による未結合が生じていても、ID は 30 フレーム分の猶予を持って再結合の候補に残るため、ID スイッチ (IDSW) を抑制できる。なお、30 という値は、動画にして約 1 秒である。

4.4.2 新規 ID 割り当ての制御

新規 ID を無制限に発行するほど誤検出 (FP) が増加するため、新規 ID 発行前に連続観測数と信頼度に基づくフィルタを導入する。

本手法を適用する以前は「既存軌跡とマッチしなかった検出はすべて新規 ID」として即時発行していた。このアルゴリズムを用いてテスト動画に対する予測を行い、得られた各軌跡の継続フレーム数 (Frame_Count) と最大信頼度 (Max_Confidence) を抽出した。図 4 は、これらの指標に基づき、真値と一致した軌跡 (TP) と一度も一致しなかった軌跡 (FP) の分布を可視化したものである。な

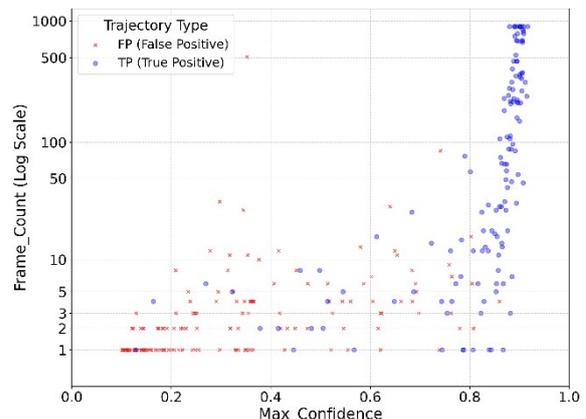


図 4. TP/FP の最大信頼度とフレーム数の分布

お、ここでの TP は評価範囲内で真値データと一回以上照合された ID であり、FP は一度も照合されなかった ID を指す。

図 4 の分布から、FP の多くは出現期間が短く、最大信頼度が低いという特徴を持つことが分かった。そこで、新規 ID を正式な軌跡として昇格させる基準を `Frame_Count > 3` かつ `Max_Confidence > 0.4` とした。この条件を適用することで、初期アルゴリズムで発生していた FP 軌跡の約 78.95% を生成段階で抑制できる。ただし、この基準適用後はハンガリアン法による ID 割り当てが変化するため、実際の運用時（システム全体の動作時）における FP 削減数がこの値とは完全に一致しない。

この新規 ID の発行手法により、軌跡が短時間で消失することを防ぎ、信頼性の高い個体を優先的に追跡することが可能となる。

5. 公式がサポートしているトラッカー

本研究で提案するロジスティック回帰ベースの MOT 手法の性能を客観的に評価するため、Ultralytics YOLO がサポートしている既存の代表的なトラッカーである ByteTrack および BoT-SORT を比較対象として選定した。これらのトラッカーは、高い汎用性と性能から、広く使用されている。また性能評価では Ultralytics YOLO の公式 yaml ファイルに基づき、パラメータはすべてデフォルト値で比較を行った。主要なパラメータを表 4 に示す。

5.1 ByteTrack

ByteTrack[4]は高スコア検出と既存トラックの一次対応付け後、低スコア検出を再度トラックに照合して真の物体復元を図る 2 段階の関連付けを行う。遮蔽などにより一時的に信頼度が低下した物体を低信頼度とし、軌跡の途切れを大幅に抑制できるトラッカーとなっている。人物追跡である MOT17 テストで $MOTA=0.803$ 、 $IDSW=2,196$ を 30 FPS で達成している。

5.2 BoT-SORT

BoT-SORT[5]は、既存の SORT 系トラッカーの課題を克服するために開発された手法である。検出結果のバウンディングボックスの四隅座標 ($x1, y1, x2, y2$) を直接扱う改良されたカルマンフィルタの導入と、事前に学習されたモデルが予測する外観特徴量を組み込むことで予測精度を

表 4. 公式トラッカーで設定できる主要パラメータ

パラメータ名	値	説明
<code>track_high_thresh</code>	0.25	高スコアの信頼度閾値
<code>track_low_thresh</code>	0.1	低スコアの信頼度閾値
<code>new_track_thresh</code>	0.25	新規開始の信頼度閾値
<code>track_buffer</code>	30	見逃し許容フレーム数
<code>match_thresh</code>	0.8	IoU のマッチング閾値

向上させたトラッカーである。MOT17 テストで $MOTA=0.805$ 、 $IDSW=1,257$ を 6.6 FPS で達成している。

6. 性能評価

6.1 MOT アルゴリズムの評価指標

本研究では、提案するロジスティック回帰ベースの性能を定量的に比較するため、MOT Metrics[3]を参考にする。その中でも MOT の評価で広く用いられる Multiple Object Tracking Accuracy (MOTA)を採用した。これは MOT の総合的な性能を示す指標であり、追跡中の誤検出 (FP)、検出漏れ (FN)、および、ID スイッチ数 (IDSW) の 3 つの誤差要因を総合的に考慮したものであり、以下の式 (2) で定義される。MOTA の最大値は 1 であり、この値が高いほど追跡全体における誤差が少ないことを意味する。

$$MOTA = 1 - \frac{FN + FP + IDSW}{GT_Total_Detections} \quad (2)$$

6.2 評価手法

本研究では、構築した追跡システムの性能を定量的に評価するため、作成した真値データ (GT: Ground Truth) と提案手法による追跡結果 (TR: Tracked Results) をフレームごとに照合した。評価の対象は、評価用データセット (動画 B) のフレーム 0 から 900 までの計 901 フレームとした。

フレームにおける GT のバウンディングボックスと TR のバウンディングボックスの重なり度合いを示す IoU を算出し、その値が 0.5 以上である場合に正しく検出されたと判断する。この基準に基づき、GT に存在する個体に対し TR が正しく対応付けできた True Positive (TP)、GT に存在しない位置で TR が誤って対応付けした False Positive (FP)、および GT に個体が存在するにもかかわらず TR がそれを検出できなかった False Negative (FN) をそれぞれ算出する。さらに GT に対応する TR が切り替わった回数である ID スイッチ (IDSW) を計算し、それらを用いて MOTA を評価する。

本評価に用いた真値データの総検出数 (`GT_Total_Detections`) は 26,806 個である。また、実験に用いたシャレ内のシロアリ個体数は 30 頭であり、理想的な環境下での総 ID 数 (ID Count) の理論値は 30 となり、MOTA は 1、IDSW は 0 となる。

6.3 結果の比較

それぞれの TP、FP、IDSW、MOTA 及び ID の総数 (ID Count) の数値を表 5 に示す。

表 5. MOT 手法の性能評価

手法	ByteTrack	BoT-SORT	提案手法
TP (個)	26,763	26,789	26,648
FP (個)	1,002	1,012	986
IDSW (回)	108	94	135
MOTA	0.957	0.958	0.952
ID Count (個)	101	95	127

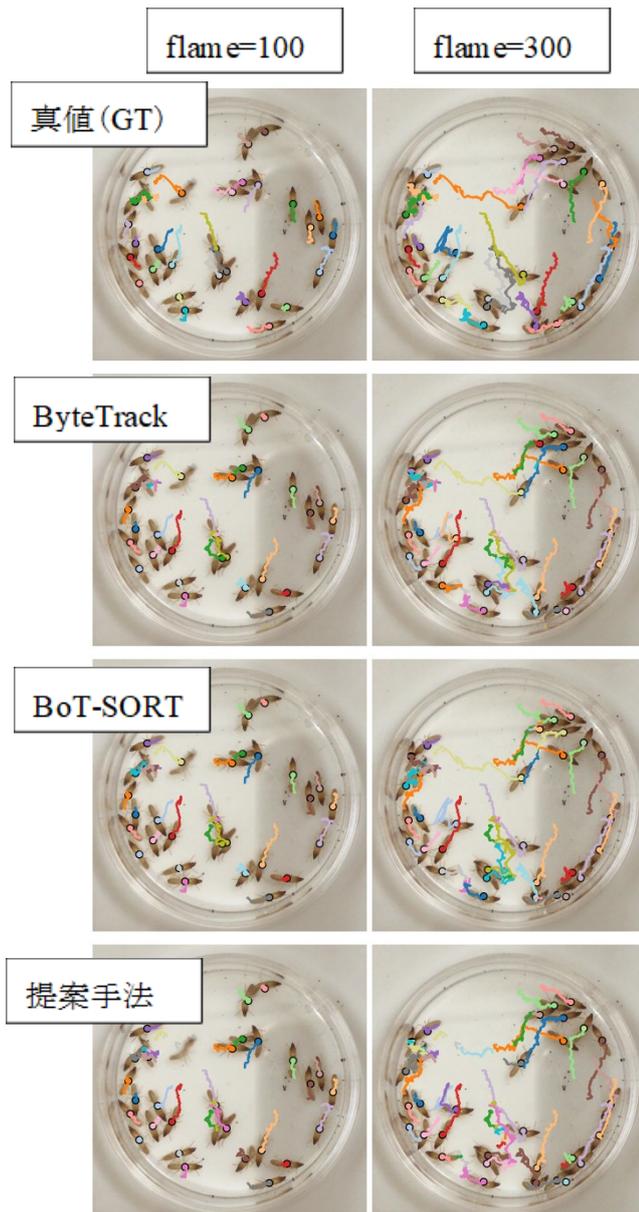


図 5. 真値データと各手法から得た予測検出

また、真値データと各手法によって得た軌跡を図 5 に示す。図は、途中の様子が分かるよう、100 フレーム目と 300 フレーム目のスナップショットである。

MOTA は BoT-SORT が最も高く (0.9581) , ByteTrack (0.9570) , 提案手法 (0.9523) が続いた。提案手法は、

3 手法の中で FP が最小 (986) である一方、TP も最小 (26,648) となっている。これは、4 章で述べた新規 ID の昇格条件 (Frame_Count>3 かつ Max_Confidence>0.4) によって、誤検出だけでなく、短命な正検出も同時に棄却されたためである。このフィルタリングの影響により、FP を抑制できた。

ID が途中で入れ替わらない、つまり ID スイッチ (IDSW) が発生しにくいのは BoT-SORT が最小であり、提案手法は既存手法と比べて 41 回多い結果となった。また、ID Count については、理想値 30 に対して提案手法が 127 と最も高く ByteTrack は 101, BoT-SORT は 95 であった。提案手法のロジスティック回帰を用いたマッチング確率の算出方法はまだ発展途上であり、説明変数の選び方次第で IDSW や IDCount は今後減少すると考える。また、提案手法は未調整のパラメータが多数あり、それらを調整することにより、MOTA の改善が期待できる。

7. 考察

まず、提案アルゴリズムの成果として挙げられるのは、多くの個体で True Positive (TP) が安定して維持され、ヤマトシロアリの主要な移動軌跡を正確に捉えることができた点である。図 5 の観察からも、個体が互いに離れている領域では検出と追跡が安定しており、提案するアルゴリズムは ID を保持する能力が高い。これは、IoU と信頼度の変化を組み合わせたロジスティック回帰モデルによる判定が有効に機能したことを示している。

また、マッチング確率を算出するロジスティック回帰において、IoU と信頼度変化の絶対値が高い寄与率となり、次いで、距離や面積変化率の絶対値が、寄与度としては高いことがわかった。一方、速度差の寄与率は低かった。マッチング確率が IoU と信頼度の変化に依存するモデルであり、個体が離れている場合は有効に機能する一方、個体が密集した場合は、個体間の境界を区別できず、IDSW が発生しやすい。また、新規 ID を割り当てる条件として、「連続 3 フレームの観測」と「最大信頼度 0.4 未満の抑制」を採用した結果、誤検出 (FP) が抑えられたが、検出された正しい ID が短時間の検出で終わった場合、フィルタされやすくなった。

さらに、図 5 の観察から、個体が集まっている領域では近接や重なり、遮蔽が同時に起き、ID が多数発生しやすくなることがわかった。また、個体が離れている領域では検出と追跡が安定していることが分かった。

8. おわりに

本研究では、高精度な物体検出モデル YOLO11 とロジスティック回帰を統合した自動追跡手法を提案し、その有効性を検証した。

本研究における進展として、まず、検出対象であるヤマトシロアリ動きを一貫した ID で記録した追跡用真値データセット (GT) を、手作業で構築したことが挙げられる。この高品質な GT の構築により、提案手法を含む各アルゴリズムについて、MOTA や IDSW といった指標を用いた客観的かつ定量的な性能評価が可能となったことは、今後の行動解析において極めて重要な意義を持つ。また、独自のアノテーションデータを用いた学習によって、YOLO11 による高精度なシロアリ検出を実現した。この検出基盤を確立したことで、提案手法の検証だけでなく、ByteTrack や BoT-SORT といった既存のトラッキングアルゴリズムの適用と比較が可能となった。

評価の結果、提案手法は誤検出 (FP) が最も少なくなることが確認できた。一方で、他の手法と比べ IDSW が多くなる傾向がみられた。提案手法は、ロジスティック回帰の説明変数の選び方に検討の余地があり、また多くの未調整パラメータがある。性能を十分には引き出せておらず、今後の課題としたい。

展望としては、今回の自動検出と手動での修正を組み合わせることで、ヤマトシロアリの高精度かつ高効率な追跡が可能となるツールの開発や、複数のクラス (兵蟻, 職蟻) に対応できるデータセットの構築、またはその他の検出モデルでの検証などが挙げられる。

謝辞

本研究を遂行するに際して、本校技術職員の岩本氏と中村氏、同研究室のメンバーには、進捗報告の時間などに質問や助言をいただき、研究を進める上での大きな助けとなりました。厚く御礼申し上げます。

本研究は JSPS 科研費 JP24K21223 の助成を受けたものです。

参考文献

- [1] 宮本健太郎, 木原久美子, シロアリの局所的な相互作用から引き起こされる系全体の統計量の変化, 情報処理学会九州支部「火の国シンポジウム 2024」, 2024.
- [2] J. Redmon, S. Divvala, R. Girshick, and A. Farhadi, You Only Look Once: Unified Real-Time Object Detection, 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 779-788, 2016.
- [3] K. Bernardin, and R. Stiefelhagen, Evaluating Multiple Object Tracking Performance: The CLEAR MOT Metrics, EURASIP Journal on Image and Video Processing, vol.2008, Article ID 246309, 2008.
- [4] Y. Zhang, P. Sun, Y. Jiang, D. Yu, F. Weng, Z. Yuan, P. Luo, W. Liu, and X. Wang, ByteTrack: Multi-Object Tracking by Associating Every Detection Box, Proceedings

of the European Conference on Computer Vision (ECCV), pp. 1-21, 2022.

- [5] N. Aharon, R. Orfaig, and B. Bobrovsky, BoT-SORT: Robust Associations for Multi-Pedestrian Tracking, arXiv:2206.14651, 2022.