

平成 30 年度修士論文

一般環境下でのマウスの識別と追跡

宮崎大学大学院 工学研究科 工学専攻
機械・情報系コース 情報システム工学分野

学籍番号 T1703008

岡崎浩佑

指導教員 椋木雅之教授

平成 31 年 1 月 28 日

概要

本研究では、人物追跡において有効である物体検出、個体識別、未知個体判定の手法を、一般環境下のマウスに適用するために改良を行う。一般環境下でのマウスの識別追跡を困難にする要因として「マウスの体毛と同色の背景であること」「ギャラリーデータが少ないこと」「個体同士の外見が類似していること」が挙げられる。従来の手法を適用すると、物体検出では、マウスと背景が類似しているために両者を混合してしまい、誤検出数が増加する。個体識別では、物体追跡の性質上、個体識別のための見本となるギャラリーデータは映像の最初の数フレームから取り出すことになる。データの少なさは偏りの原因となり、識別誤り率が増加する。未知個体判定では、個体別の外見差がある人物は1つのパラメーターで未知と既知の条件分けが可能であったが、外見差の少ないマウスで条件分けを行うにはパラメーター数が不十分である。

この問題に対して、物体検出では、画像分類手法を取り入れて、背景のみを削減することでマウス検出の誤り率を改善する。個体識別では、データ拡張手法を取り入れて、画像変換処理によるデータの増強を行い、ギャラリーの偏りを削減する。未知個体識別では、パラメーターを増やし、より細かい条件分けをすることで未知個体識別精度を向上させる。

従来の手法と改良した手法とで比較実験を行い、それぞれの手法で改善できたことを確認した。

目次

| | | |
|--------|--------------------------|----|
| 1 | 序論 | 3 |
| 2. | 一般環境下でのマウスの識別・追跡問題 | 5 |
| 2.1. | マウスの識別・追跡問題の定義 | 5 |
| | 想定する一般環境 | 5 |
| 2.2. | 一般環境下での識別追跡の困難性 | 6 |
| 2.2.1. | 検出対象と背景の誤検出 | 6 |
| 2.2.2. | マウス識別の困難性 | 7 |
| 2.2.3. | 未知個体への対処の困難性 | 8 |
| 2.3. | 物体識別・追跡の要素技術 | 9 |
| 2.3.1. | 物体検出 | 9 |
| 2.3.2. | 物体識別 | 10 |
| 2.3.3. | 未知個体識別 | 13 |
| 3. | 一般環境に対応するための識別追跡の要素技術の改良 | 14 |
| 3.1 | 深層学習と画像分類を用いた物体検出 | 14 |
| 3.2 | データ拡張による物体識別の改良 | 17 |
| 3.3 | 画像分類パラメーターを用いた未知個体識別 | 18 |
| 4. | 評価実験 | 19 |
| | 実験の設定 | 19 |
| 4.1. | マウス検出の評価実験 | 20 |
| 4.1.1. | 実験準備 | 20 |
| 4.1.2. | 比較実験 | 20 |
| 4.1.3. | 考察 | 22 |
| 4.1.4. | 追加実験 | 22 |
| 4.2. | マウス識別の評価実験 | 23 |
| 4.2.1. | 比較実験 | 23 |
| 4.2.2. | 実験結果・考察 | 24 |
| 4.3. | 未知個体識別の評価実験 | 25 |
| 4.3.1. | 比較実験 | 25 |
| 4.3.2. | 結果・考察 | 26 |
| 5 | 結論 | 27 |
| | 謝辞 | 28 |
| | 参考文献 | 29 |

1 序論

物体追跡とは、動画や一連の画像群の中から対象物体を検出し、識別を行い、個々の軌跡を求める課題である。近年の深層学習の発展により、画像内の物体を認識する画像認識技術が急激に進展した。現在では、この技術を活用した顔認証システムや犯罪者検知など、多くのソリューションが実用化されている。物体追跡においても、画像認識技術は人物照合などに利用されている。これにより、複数監視カメラ間での特定人物の経路の把握が可能になり、人物行動解析や防犯システムの実用化に向けて試験導入が進められている。このように人物追跡に商業的可能性があり、また非剛体である人物は物体追跡の研究対象として理想的とされていることもあり、現在の物体追跡に関する研究の 7 割は歩行者追跡に向けられている [1]。

本研究では、マウスを追跡対象としている。本学農学部の坂本講師らの研究グループは、動物の生態把握の研究を行っている。生態把握には監視カメラ映像の解析が有効であるが、人手で行うには時間と労力がかかってしまう。本研究の目的である映像中のマウスの識別・追跡が実現できれば、この人的コストを削減できる。

一般的な環境下でのマウスの追跡を考える。背景に土や落ち葉があれば、マウスの体毛と似た色の背景はマウスの検出を困難にする。マウスが集団で生活していれば、類似個体の存在により個体識別が困難になる。また、カメラに映る個体数が増減することを考慮すると、映像に現れる個体が既知の個体か新しい個体か識別する必要があり、外見に大きな差がないネズミで行うにはとても困難な処理である。服や顔など様々な特徴を抽出できて識別が容易な人間と異なり、マウスの場合は検出と個体識別、未知個体判定が困難となる。そのため人物追跡において有効である手法そのままでは、必ずしも良い結果は得られない。

そこで本研究では、人物追跡において有効である物体検出、個体識別、未知個体判定の手法をマウス追跡のために改良する。具体的には、マウスの検出では検出結果に対して画像分類を行い、背景であると分類されたものを結果から除外することで誤検出数を少なくしてマウスの検出率を向上させる。個体識別では、ギャラリーの画像に変換処理を施すことでデータ拡張を行い、識別率の向上をはかる。未知個体識別では判別のためのパラメータを追加してより細かい状況設定を行うことで識別誤り数を削減する。

本稿の構成は以下の通りである。2章ではマウスの識別追跡問題の定義、一般環境の定義、一般環境下での識別追跡の困難さ、従来の人物追跡に用いられた手法の説明を行う。3章では提案手法の詳細を述べる。4章では提案手法の有効性を示すために行った元の手法との比較実験について述べる。5章では本研究の結論と今後の課題について述べる。

2. 一般環境下でのマウスの識別・追跡問題

2.1. マウスの識別・追跡問題の定義

マウスの識別・追跡で達成したいことは、入力された映像に対して出力として個体ごとの正確な経路を得ることである。本研究では、1つの固定カメラで飼育ケース内部を撮影した映像を使用する。個体識別のために必要なギャラリーデータには、追跡したい映像の数フレーム分の画像から個体別に切り出した画像を与える。

想定する一般環境

本研究では、一般環境として従来研究 [2] よりも制約を緩めた状況を想定する(図 1)。背景に落ち葉を敷き詰め、遮蔽物を置いた飼育箱を用意した。その中に追跡対象として4匹のアカネズミ(以下マウスと呼ぶ)を投入する。マウスの体毛と落ち葉の色は類似しており、検出が困難である。また、4匹のマウスは外見が類似しており、少量のギャラリーデータでは識別が困難である。遮蔽物の存在により、映像に映る個体数は増減する。正しい追跡を行うには個体が現れた際に、過去に現れたことがある個体か新しく現れた個体かを識別する必要がある。しかし、外見が類似した個体間で既知か未知かを判断することは難しい。

このように一般環境下でのマウスの追跡では「物体検出」「個体識別」「未知個体識別」の3つが課題となる。

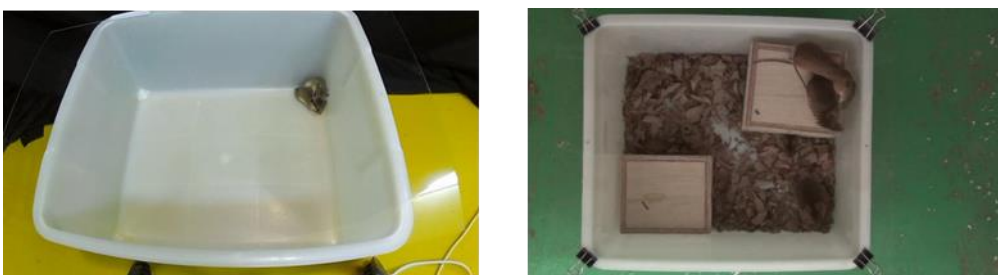


図 1 従来研究の撮影環境(左)と本研究で想定する一般環境 (右)

2.2. 一般環境下での識別追跡の困難性

2.2.1. 検出対象と背景の誤検出

物体検出とは、入力された画像に対して、画像上の物体の位置座標を返す問題のことである。本研究での物体検出の対象はマウスであり、それ以外を背景と呼ぶ。マウスの持つ茶色の体毛は、飼育ケースに敷き詰められた落ち葉の色と似ており、人の目で見てマウスと背景を区別することが難しい。

物体のクラスと位置を同時に識別する手法に **Single Shot MultiBox Detector(SSD)** [3]がある。この手法では迷彩服で背景に擬態した人物の静止画から人物を検出すること [4]も可能であったため、背景に紛れたマウスの検出にも効果があると考えた。

しかし、今回検出する映像は一連の動きをするマウスの映像であるので、マウスの移動で画像にボケが生じる場合や、マウスが遮蔽物に隠れて体の大部分が見えなくなる場合は識別が困難になることを考慮しなければならない。

SSD の検出結果には、検出位置を表すボックスとクラスの信頼度がセットで出力される(図 2)。このクラスの信頼度に閾値を設けることで、SSD によって予測されたクラスと正解との一致率を向上させることができる。しかし、ボケや遮蔽が生じると SSD で予測されるクラスの信頼度も低くなり、このときに閾値処理を行うと正解であった検出も取り除いてしまう(図 3)。SSD が出力するクラスの信頼度によらずに、背景の検出のみを取り除く仕組みが必要である。

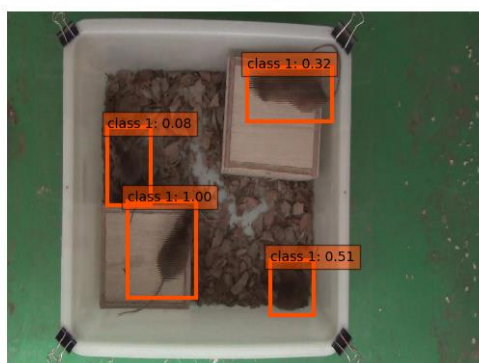


図 2 SSD によるマウスの検出結果

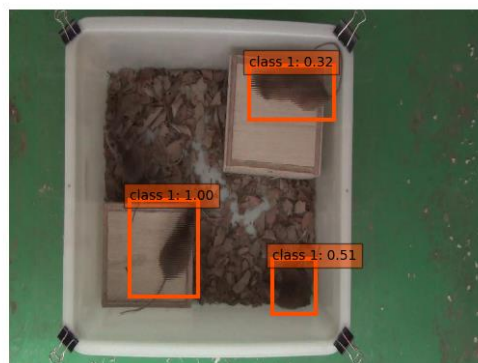


図 3 閾値処理後のマウスの検出結果

2.2.2. マウス識別の困難性

個体識別とは、入力として与えられたマウス画像が、事前に個体情報として保持しているギャラリーのどの個体に対応するかを識別する問題である。このときの、与えられた画像がどのクラスに属するかを分類する問題を画像分類という。本論文では個体識別のために画像分類手法を用いる。

画像分類手法には深層学習を用いた手法や類似度アルゴリズムによる手法などがある。深層学習を用いた手法は分類のために大量のデータが必要であるのに対し、類似度アルゴリズムによる手法はクラスのお手本であるギャラリーのデータとの類似度で分類を行うためクラスごとに画像を数枚用意するだけでよい。物体追跡では数フレーム分のデータでその後の個体識別を行う必要があるため、本研究では少量の画像で分類が可能な類似度アルゴリズムによる手法を採用する。

画像分類ではギャラリーの偏りを無くすために、膨大な数の画像群の中から無作為に取り出したものをギャラリーとして使用するなどの工夫がなされている。しかし、物体追跡の場合は大量のデータを用意することができず、連続するフレーム画像数枚という偏ったデータ(図 4)を与えざるを得ない。このギャラリーの偏りは、識別結果の性能に直結する問題である。正しい個体識別のために、ギャラリーの偏りを緩和する仕組みが必要である。

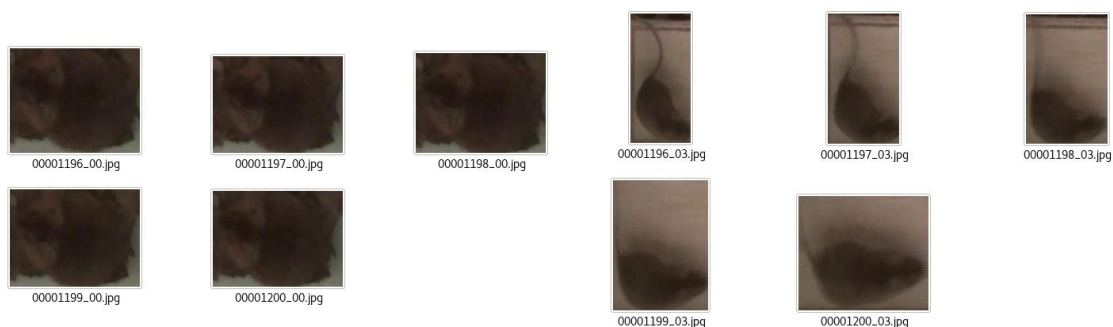


図 4 同フレーム画像から切り出した個体 A(左 5 枚)と個体 B(右 5 枚)の画像姿勢のバリエーション差によってギャラリーの偏りが生じている

2.2.3.未知個体への対処の困難性

未知個体識別とは、追跡対象が映像に現れた際に、その個体が今までの映像に登場していてギャラリーに登録されている既知の個体か、初めて映像に登場した未知の個体かを識別する問題である。

多くの物体追跡の研究では未知個体は存在しないことを前提にしている。そのため、与えられたテストデータはギャラリーに存在するクラスのどれかに分類される。本研究で使用するマウスの映像では遮蔽物が存在する。未知個体がないものとして個体識別を行うと、遮蔽物から出てきた既知の個体か、初めて映像に現れた未知の個体かを区別することができない。区別できず、未知個体が既知の個体として分類されると、画面上に同一個体が2体以上存在するという矛盾が生じる場合がある。

従来の人物追跡の研究で、未知個体識別の分類を含んだものの1つとして Lu らの手法 [5]がある。この手法では、識別したい人物をギャラリーで近似し、最も類似度が高いギャラリーの人物に識別する。この際、近似に使用したあるパラメータ β が複数クラスに分散している場合、識別された人物が未知の人物であると判定する。

人物の識別では顔や姿勢、服の色など様々な特徴が利用でき、ギャラリーのクラス間に特徴差を設けられる。マウスの場合は個体同士の外見が似ており、人物のように抽出できる特徴のバリエーションもないため、ギャラリーのクラス間の特徴差が少ない。そのため、マウスの分類では、1つのクラスのみ類似することがない。このことは、 β の複数クラスへの分散につながる。 β が分散する原因が「未知個体」によるものと「個体の類似性」によるものの二つとなり、この二つを条件分けしなければ未知個体の識別誤りを引き起こす。この二つの状況を識別するための新たな状況設定が必要となる。

2.3. 物体識別・追跡の要素技術

2.3.1. 物体検出

物体検出 [6]とは、物体認識 [7]タスクの一つで、画像中のどの位置に物体があるかを求める問題である。

物体検出の従来手法には、テンプレートマッチングや知識ベースの物体検出手法がある。これらの手法は、識別器をラスタスキャンして物体領域を検出する。

近年、深層学習の発展により CNN が画像分類分野において成果を上げた。物体検出においても、画像上の物体の候補領域を生成すれば、候補領域に対して画像分類問題に持ち込むことが可能である。そのため、深層学習ベースの様々な物体検出手法が提案された。本研究では、深層学習ベースの物体検出手法である **Single Shot Multi Box Detector** [3]を用いてマウスの検出を行う。

Single Shot Multi Box Detector

● 検出

CNN から画像の特徴マップを抽出する。用意されたアスペクト比のデフォルトボックスを特徴マップに当てはめることで、物体のクラスと位置、信頼度を予測して物体検出を行う(図 5)。

CNN から得られる特徴マップは、畳み込み層とプーリング層により縮小していく。得られた特徴マップそれぞれに、デフォルトボックスによる検出を行うことで、様々なスケールの物体を検出することが可能になる。

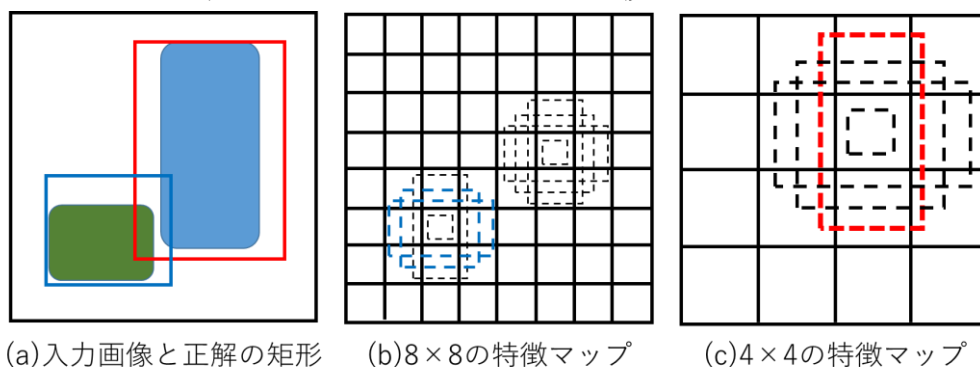


図 5 特徴マップによる検出図

2.3.2. 物体識別

画像分類 [6]とは、物体認識タスクの一つで、物体が事前に定義されているカテゴリの内、どのカテゴリに属しているかを求める問題である。

物体追跡での個体識別処理において、ギャラリーとして与えられる画像は、数としては少数である。そのため、大量の画像を必要とする深層学習を用いた手法は利用できない。本研究では、少量のギャラリー画像で物体識別を行える CMA 法 [8]を利用する。

CMA 法

CMA 法は、未知のテストデータを複数の既知のカテゴリのいずれかに分類する。既知のカテゴリは、そのカテゴリに属する 1 枚以上の学習用画像の特徴量集合として与えられる。テストデータも、テスト画像の特徴量として与えられる。CMA 法は、テストデータを既知のカテゴリ全ての学習用画像で近似する最適化段階と、この近似に最も貢献したカテゴリを選択する分類段階の 2 段階からなる(図 6)。

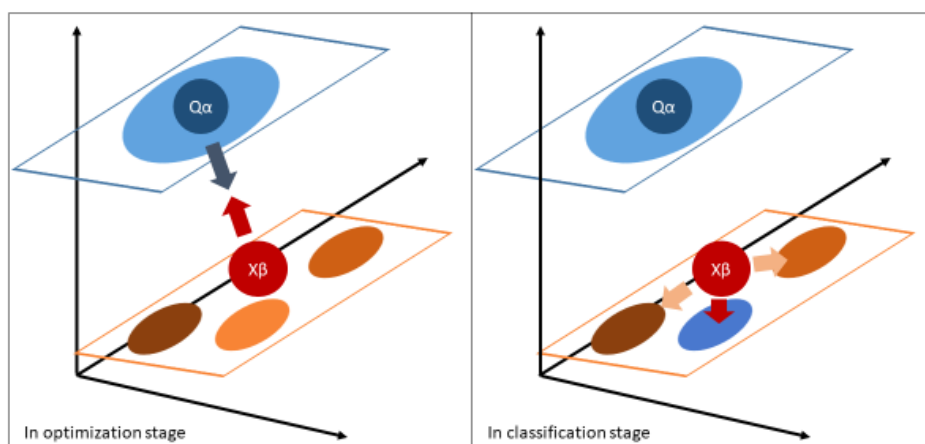


図 6 CMA 法の最適化段階と分類段階 [9]

- 最適化段階

テストデータの近似点を、全ての学習用画像の特徴量の線形和により表現する。最適化段階では、この近似点を構成する線形和の係数を求める。この際、テストデータと近似点との距離が小さくなるように係数を定める。一方で、近似点がデータの性質を保持するよう、学習用画像の平均点 (Mean Point) に近づけるという制約を課す。テスト画像の特徴量ベクトルを \mathbf{y} 、既知のカテゴリ $i \in \{1, \dots, n\}$ (n はカテゴリの種類数) の学習用画像の特徴量ベクトルを並べた行列を $\mathbf{X}_i \in \mathbf{R}^{m \times N_x^i}$ (N_x^i はカテゴリ i の学習用画像数)、全学習用画像の特徴量ベクトルを並べた行列を $\mathbf{X} \in \mathbf{R}^{m \times N_x}$ ($N_x = \sum_i N_x^i$ は全学習用画像数) とする。この時、CMA 法の最適化段階では、以下の式 $f(\boldsymbol{\beta})$ を最小化する係数ベクトル $\boldsymbol{\beta} \in \mathbf{R}^m$ を求める。

$$f(\boldsymbol{\beta}) = \|\mathbf{y} - \mathbf{X}\boldsymbol{\beta}\|^2 + \lambda \left\| \boldsymbol{\beta} - \frac{\mathbf{1}_{N_x}}{N_x} \right\|^2 \quad (1)$$

ここで、 $\mathbf{1}_k$ は各要素が1 の k 次元ベクトル、 $\|\cdot\|$ はベクトルの L_2 ノルムである。式 (1) の $\mathbf{X}\boldsymbol{\beta}$ が近似点を表している。第1項は、テストデータと近似点を近くするための項である。第2項は、近似点を平均点に近くするための制約項 (正則化項) である。 λ は正則化項の重みパラメータである。式 (1) は2次形式であり閉じた解を持つ。具体的には、下記の連立方程式を解けば良い。

$$(\mathbf{X}^T \mathbf{X} + \lambda \mathbf{1}) \boldsymbol{\beta} = \mathbf{X}^T \mathbf{y} + \lambda \frac{\mathbf{1}_{N_x}}{N_x} \quad (2)$$

- 分類段階

最適化段階で求めた係数 $\boldsymbol{\beta}$ を利用して、近似点を構成する際に最も貢献したカテゴリを求め、そのカテゴリにテストデータを分類する。近似点の係数 $\boldsymbol{\beta}$ の各次元が、学習用画像に対応している。学習用画像は、既知のカテゴリのいずれかに含まれている。即ち、カテゴリ i の学習用画像の係数を $\boldsymbol{\beta}_i$ とし、 $\boldsymbol{\beta}^T = (\boldsymbol{\beta}_1^T, \dots, \boldsymbol{\beta}_n^T)$ と分解できる。この時、各カテゴリ i について下式を求め、最小となるカテゴリに分類する。

$$d^i = \|X_i\|_* \cdot \frac{\|\mathbf{y} - X_i \boldsymbol{\beta}_i\|^2 \|\boldsymbol{\beta}\|}{\|\boldsymbol{\beta}_i\|} \quad (3)$$

$\|\mathbf{y} - X_i \boldsymbol{\beta}_i\|^2$ が小さい程、カテゴリ i のみで代表点の近似がより良く行えていることになる。また、 $\|\boldsymbol{\beta}_i\|/\|\boldsymbol{\beta}\|$ が大きい程（その逆数が小さい程）カテゴリ i の係数が既知カテゴリ全体の中で大きな役割を果たしていると言える。 $\|\cdot\|_*$ は、行列の核ノルム（行列の特異値の和）である。データのばらつきが大きい程、この値は大きくなる。カテゴリ間のばらつきの違いを考慮するための重み付けに利用している。

2.3.3.未知個体識別

未知個体識別とは、テストデータがギャラリーに存在するクラスかどうかを判別する処理である。個体識別では通常、テストデータは必ずギャラリーに存在するクラスのどれかに分類される。

従来的人物追跡の研究で、Lu らの手法 [5]では、個体識別で利用するパラメーターの値によって未知個体を識別する方法を提示した。この項では、その手法の個体識別法と未知個体識別法について説明する。

個体識別法

ギャラリーの集合 \mathbf{X} とテストデータ \mathbf{y} が与えられたとき、テストデータ \mathbf{y} はギャラリーの集合 \mathbf{X} と線形和係数 $\boldsymbol{\beta}$ 、誤差 $\boldsymbol{\epsilon}$ によって近似できると仮定する。

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\epsilon} \quad (4)$$

近似するために必要な係数 $\boldsymbol{\beta}$ を以下の式で求める。

$$\hat{\boldsymbol{\beta}} = \operatorname{argmin}\{\|\mathbf{y} - \mathbf{X}\boldsymbol{\beta}\|_2 + \lambda\|\boldsymbol{\beta}\|_1\} \quad (5)$$

ギャラリー \mathbf{X} をクラス単位で分割した \mathbf{X}_i と $\boldsymbol{\beta}$ で近似し、テストデータ \mathbf{y} に最も近く近似できたクラス i を \mathbf{y} のクラスとする。

$$\min_{\mathbf{X}_i} \|\mathbf{y} - \mathbf{X}_i \delta_i(\boldsymbol{\beta})\|_2, 1 \leq i \leq N \quad (6)$$

$\boldsymbol{\beta}$ の特性と未知個体識別法

テストデータ \mathbf{y} がギャラリー \mathbf{X} に存在する場合は、 $\boldsymbol{\beta}$ は最も近似できた1クラスを指し示す。ギャラリー \mathbf{X} に存在しない場合、 $\boldsymbol{\beta}$ は複数のクラス間で分散する傾向がある。この傾向から、分類に最も貢献した $\max\{\beta_i\}$ が $\boldsymbol{\beta}$ 全体に比べて小さい場合、そのテストデータ \mathbf{y} は未知個体であることが言える。

$$\frac{\max\{\beta_i\}}{\sum_j \beta_j} < \gamma, 0 < \gamma < 1 \quad (7)$$

3. 一般環境に対応するための識別追跡の要素技術の改良

3.1 深層学習と画像分類を用いた物体検出

SSD が検出した領域に対して CMA による画像分類を行い、誤検出である背景画像を削除する。処理手順を以下に示す(図 7)。

(1) SSD の学習

SSD によるマウス検出は、VGG16 という 1000 カテゴリ・120 万枚の画像を学習した畳み込みネットワークモデルを利用して行われる。このモデルにはマウスのカテゴリが存在しないため、マウスの検出を行うには追加でモデルの学習を行う必要がある。学習済みのモデルを転用して、新たなモデルを生成する方法をファインチューニングという。

ファインチューニングによって学習されたモデルがある程度の検出性能を得るために、学習用画像は 100 枚程度用意する必要がある。学習用画像とそれに対応する対象の位置情報を入力として与え、モデルを再学習することで新しいカテゴリの検出が可能となる。

(2) SSD によるマウス検出

ファインチューニングにより「マウス」を学習した SSD をマウス動画に適用し、動画中からマウスの候補領域をすべて切り出す。切り出した候補領域に対して特徴抽出し、特徴量を得る。

(3) 画像分類に必要なギャラリーデータの作成

マウスの候補領域のうち、一部の結果に対して目視で「マウス」「背景」を判別する。判別したそれぞれのクラスの特徴量を、ギャラリーとして CMA に与える。

(4) 背景の削除

残りのすべての候補領域について、CMA による「マウス」「背景」の 2 クラス分類を行う。このうち、「マウス」と識別された候補領域のみをマウスの検出結果とする。

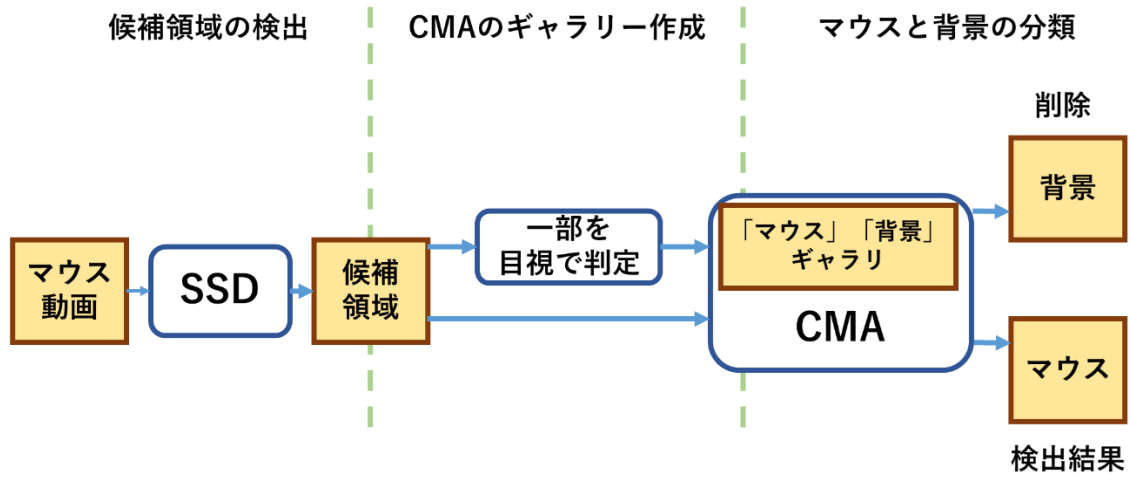


図 7 提案手法の流れ図

SSD の出力には、画像上のどの位置に物体があるかを示す矩形の位置座標と、何の物体であるかを表すクラス ID、クラスらしさを表す信頼度がある。この内、クラス ID と信頼度を求める処理は、物体認識のような性質を持つ処理である。この二つの処理をまとめて「認識」と呼び、物体の位置を求める処理を「検出」と呼ぶこととする。

事前に行う学習によって、検出と認識の性能は向上する。しかし、検出と認識それぞれの性能を向上させるために必要なデータは異なる。検出は、対象らしいものの位置を提示する処理であり、学習に与えるデータとしては検出対象のデータを与えればよい。一方、認識では、矩形内の画像が何を表すかを認識する処理である。類似物体が存在する中で正しい認識を行うためには、認識したい対象と対象に類似する物体の違いを示すようなデータを与える必要がある。

提案手法では、検出は SSD から流用して、認識は画像分類手法の CMA 法を用いて行う。CMA 法でマウスと背景を分類する際に、『認識したい対象と対象に類似する物体の違いを示すようなデータ』をギャラリーに与えている。これは、前述した『正しい認識を行うため』の条件を満たしている。よって、SSD の閾値処理による背景の削除よりも、提案手法の画像分類による背景の削除の方が正しく背景の削除ができ、マウスの誤検出数を削減できると考えた。

画像分類の性能は、与えるギャラリーに依存する。ギャラリーをマウス動画の連続するフレームから選ぶと同じように撮影されたマウスが得られ、ギャラリーに偏りが生じる。そのため、識別率が低下する。これに対処するために、広い範囲のデータから無作為に選んだマウスと背景画像を CMA で用いるギャラリーデータとした。

物体検出の改善方法に SSD 再学習があるが、検出したい動画ごとに背景を学習させるのは現実的ではない。一方、提案手法は、10 枚程度のギャラリーで分類することができる。優位性としては、少量のデータで改善できることであり、背景の異なる動画にも対応しやすいことが挙げられる。

3.2 データ拡張による物体識別の改良

データ拡張 [10]とは、画像にアフィン変換やコントラスト調整、平滑化、ノイズ処理等の画像処理を加えて、データを水増しする処理のことである。関連研究 [11]では、学習用データセットにデータ拡張処理を行うことで、一般物体認識精度を改善している。本論文でも、データ拡張による個体識別率の改善を試みる。

物体追跡の性質上、追跡前にそれぞれの個体データを大量に用意することは難しい。そのため、少量の画像をギャラリーに追加して識別・追跡処理を行うことになる。その場合の個体識別結果は、ギャラリーのデータに強く依存し、類似する画像しか正しく識別することができない。例えば、個体 A のギャラリーデータが上向きのものしかない場合に、右向きの個体 A のデータを識別させると、識別結果は A 以外の個体を示す可能性が高い。正しい個体識別のためは、ギャラリーに含まれていない向きの画像も正しく識別できる必要がある。

提案手法では回転処理によって向きのバリエーションを増やした画像をギャラリーデータに追加する。具体的な処理としては、ギャラリーに与える数フレーム分のサンプル画像全てに 0 度、90 度、180 度、270 度の回転処理を加えて、4 倍に増やしたサンプル画像をギャラリーとする(図 8)。向きのバリエーションを増やすことで、一方向のデータしかない個体は単純に 4 倍のバリエーションを得られ、複数方向のデータを持っていた個体はあまり回転処理による影響を受けない。この手法によって、個体の向きによるギャラリーの偏りを解消して、個体の識別率を向上させることができると考えた。

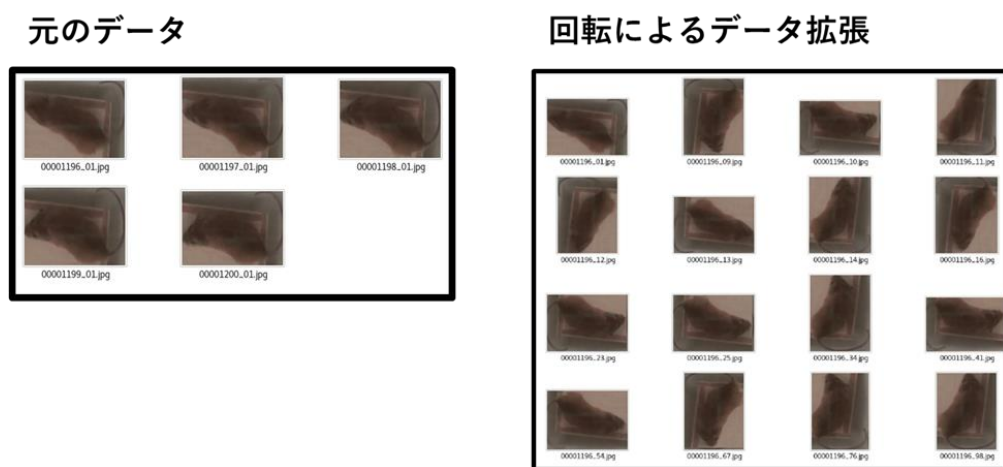


図 8 元のサンプル画像(左図)とデータ拡張を行ったサンプル画像(右図)

3.3 画像分類パラメーターを用いた未知個体識別

CMA 法でテストデータ \mathbf{y} をギャラリー X_i で近似したときの差 d^i を、評価のパラメーターとして使用する。以下、 d^i を類似度と呼称する。

$$d^i = \|X_i\|_* \cdot \frac{\|\mathbf{y} - X_i \boldsymbol{\beta}_i\|^2 \|\boldsymbol{\beta}_i\|}{\|\boldsymbol{\beta}_i\|} \quad (8)$$

$\frac{\min\{d^i\}}{\sum_j d^j}$ の値が小さい場合、そのテストデータはある一つのクラスのみに関連していることを示し、分類の信頼度が高いと言える。一方、 $\frac{\min\{d^i\}}{\sum_j d^j}$ の値が大きい場合、そのテストデータは他のクラスにも分類される可能性があったことを示し、分類の信頼度が低いと言える。このように $\frac{\min\{d^i\}}{\sum_j d^j}$ は分類の信頼度を表す指標となる。

ギャラリーに存在しない未知個体を分類した場合、1つのクラスのみに関連することはないので、基本的に未知個体の分類の信頼度は低くなる。この傾向を利用して、従来手法より細かく条件分けをして未知個体を識別する。具体的には以下の条件

$$\frac{\min\{d^i\}}{\sum_j d^j} > \delta \quad \text{かつ} \quad \frac{\max\{\beta_i\}}{\sum_j \beta_j} < \gamma \quad (9)$$

を満たすとき、テストデータは未知個体であると判定する。分類の信頼度が低く、 $\boldsymbol{\beta}$ が分類に貢献していないと判断したとき、テストデータは未知個体だと言える。

4. 評価実験

実験の設定

飼育ケースの背景に落ち葉、遮蔽物を2つ設置する(図9)。カメラを固定して映像を撮影する。映像はアカネズミが2匹の場合と4匹の場合の2種類を撮影した。2匹の映像は、29分49秒の映像で、フレームレートは30fps、フレーム数は53654フレームである。4匹の映像は、20分14秒の映像で、フレームレートは30fps、フレーム数は36389フレームである。どちらの映像も画像サイズは1440×1080ピクセルである。



図9 映像の撮影環境

4.1. マウス検出の評価実験

画像分類による検出結果の絞り込みで検出精度を改善できるか比較実験で検証する。

4.1.1. 実験準備

- アカネズミ検出器作成

SSD をアカネズミ検出器に調整する。ImageNet と呼ばれる大規模なデータセットで訓練された CNN モデルの VGG16 に、アカネズミ画像をファインチューニングさせてアカネズミ検出器として利用する。

学習用データとしてアカネズミ 2 匹の映像のフレーム画像 364 枚と画像上のアカネズミの位置情報を示すアノテーションファイル 364 個を用意して、5 万回の学習を行った。

- CMA ギャラリーデータの作成

4 匹の映像の 3000~3999 フレームを入力として SSD による検出を行う。SSD の検出結果の中からマウスを囲う矩形と背景を囲う矩形を無作為に 12 枚取り出し、CMA のギャラリーに加えた。

4.1.2. 比較実験

実験手順を以下に示す。

1. SSD にマウス 4 匹の映像 720~1200 フレームを入力として与え、マウスの検出矩形を得る。
2. CMA 法で 1 の検出結果をマウスと背景に分類する。分類の結果、マウスは 1470 枚、背景は 482 枚に分けられた。背景を取り除いた 1470 個の矩形を提案手法による検出結果とする。
3. それぞれのマウス検出の結果を正解データと比較して評価する。正解の矩形と検出結果の矩形の一致度を表す指標として **Intersection over Union(IoU)**(図 10)を用いる。IoU は矩形全体の領域に対する矩形の重複領域の割合であり、矩形の一致率を示す。この実験では、IoU の閾値を 0.4 とし、閾値以上なら、二つの矩形は一致しているとみなす。この指標を用いて、**True Positive(TP)**:「マウスを検出できた数」、**False Negative(FN)**:「背景と予測したものがマウスだった数」、**False Positive(FP)**:「マウスと予測したものが背景だった数」を求める。その後、TP、FN、FP を用いて、適合率:「マウスと予測したデータのうち、実際にマウスであったものの割合」、再現率:「実際にマウスで

あったもののうち、マウスであると予測されたものの割合」、F 値(式 (10)) : 「適合率と再現率の二つの指標を 1 つにまとめた値」によって従来手法と提案手法の性能評価を行う。

$$F\text{値} = \frac{2 \times \text{適合率} \times \text{再現率}}{\text{適合率} + \text{再現率}} \quad (10)$$

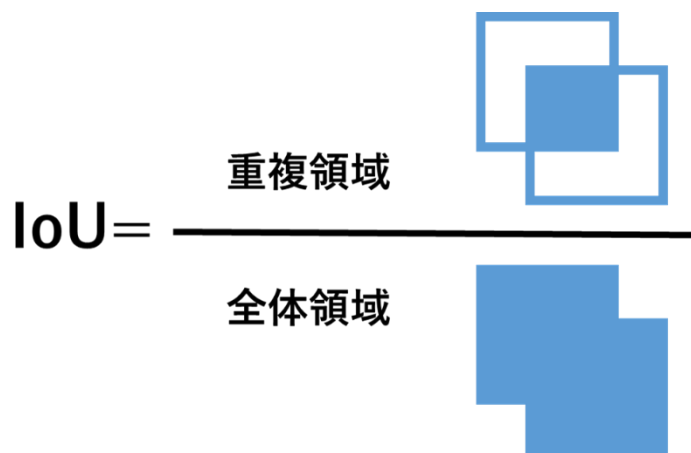


図 10 Intersection over Union(IoU) : 領域の一致率

実験結果

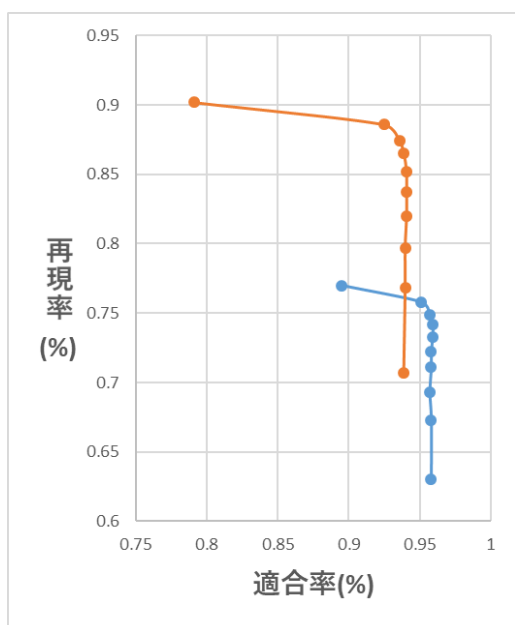


図 11 従来手法(赤)と提案手法(青)の適合

率と再現率の変化

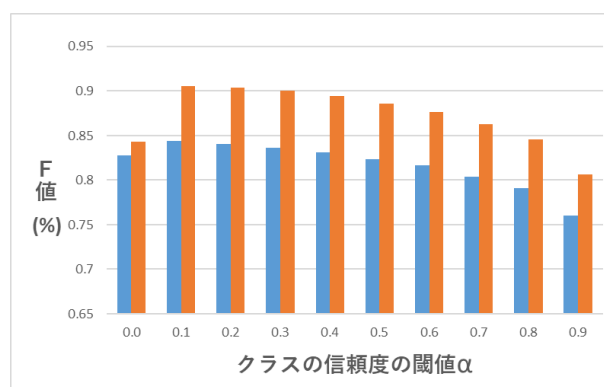


図 12 従来手法(赤)と提案手法(青)の F 値

4.1.3. 考察

SSD によるマウス検出処理によって検出された矩形の、クラスの信頼度に閾値 α を設ける。閾値以下の矩形を背景であるとみなし、取り除く。 α を 0 から 1 まで 0.1 刻みに変更して、背景削減処理を行ったときの、適合率と再現率の変化を図 11 に示す。また、図 12 に、従来手法と提案手法の、SSD 閾値の変化における F 値を示す。グラフを見ると、全体的な性能は従来手法のものが良く、提案手法によって改善することはできなかった。しかし、図 11 を見ると、検出の正しさを示す適合率において、提案手法が従来手法を上回っている。物体追跡において、対象の誤検出は追跡結果に影響を及ぼす。物体追跡を目的とした検出方法としては、提案手法は有効である場合がある。

4.1.4. 追加実験

事前に用意されているデータは利用できるものとして、SSD の学習で用いた画像を CMA のギャラリーに追加して再度実験を行う。SSD の学習で使用したアカネズミ 2 匹の映像のフレーム画像から、マウスの画像 565 枚と、ランダムに切り出した背景 458 枚を CMA のギャラリーに加えて、再度検出実験を行った。結果を図 13 図 14 に示す。ギャラリーデータ数を増やしても、性能の改善がみられなかった。

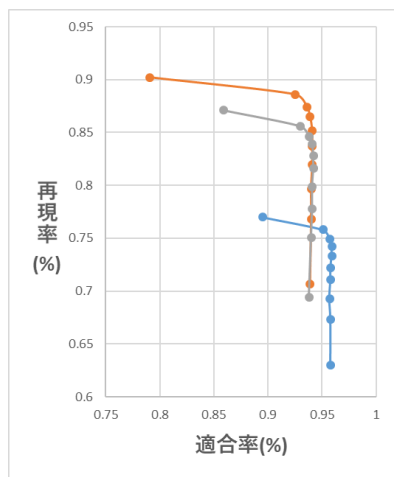


図 13 提案手法(グレー)の適合率と再現率の変化

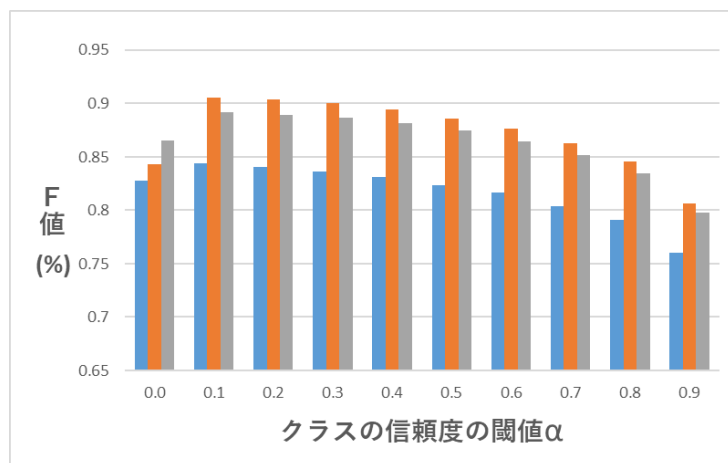


図 14 提案手法(グレー)の F 値

4.2. マウス識別の評価実験

ギャラリー内のデータに回転処理を加えてデータを拡張し、ギャラリーの偏りを減らすことで識別率の低下を防ぐ。拡張処理を行う場合と行わない場合で比較実験を行い、提案手法によって識別率の改善ができるか比較実験で検証する。

4.2.1. 比較実験

実験手順を以下に示す。

1. 4匹の映像の1196~1200フレームの画像をサンプルとしてギャラリーに追加する。4個体×5フレームの20個のデータをギャラリー1とする(図15)。
2. サンプルを回転処理によって拡張した、4個体×4方向×5フレームの計80枚のサンプルをギャラリー2とする(図16)。
3. 1201~1400フレームの識別を行い、ギャラリー1と2の結果を比較する。

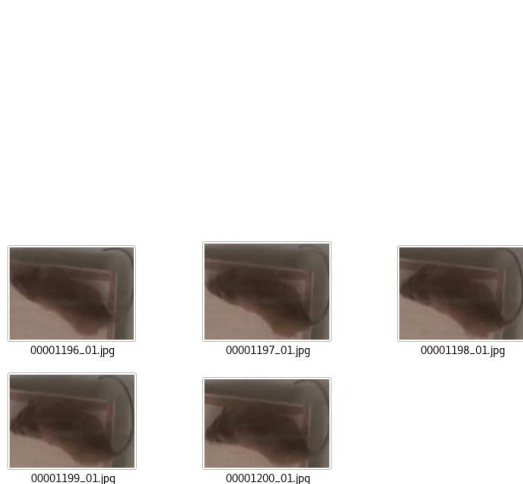


図 15 ギャラリー1のマウス B

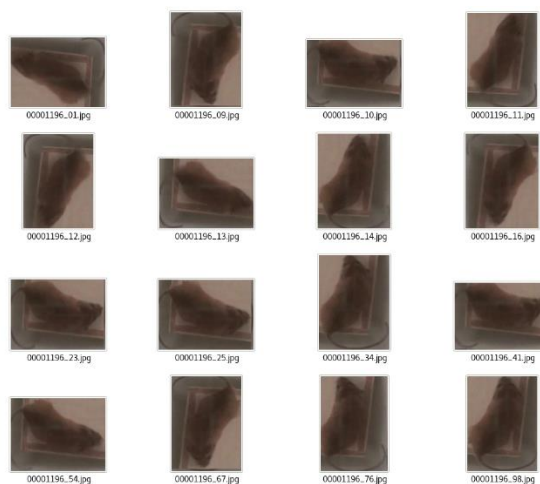


図 16 ギャラリー2のマウス B

4.2.2. 実験結果・考察

下の図 17 に、従来手法と提案手法の個体識別結果を示している。個体別の識別率を見ると、識別率の悪かった個体 A の識別率が約 30% 上がり、識別率の高かった個体 D の識別率は約 20% 低下した。これは、ギャラリーデータの偏りのために、多くの画像が個体 D へ分類されていたものが、データ拡張による偏りの緩和により、どの個体にも満遍なく分類されるようになったことを示す。

全体の識別率を見ると 2% 改善しており、提案手法によってギャラリーの偏りを改善できたといえる。

データの少なさによるギャラリーの偏りに対してデータ拡張は有効である。

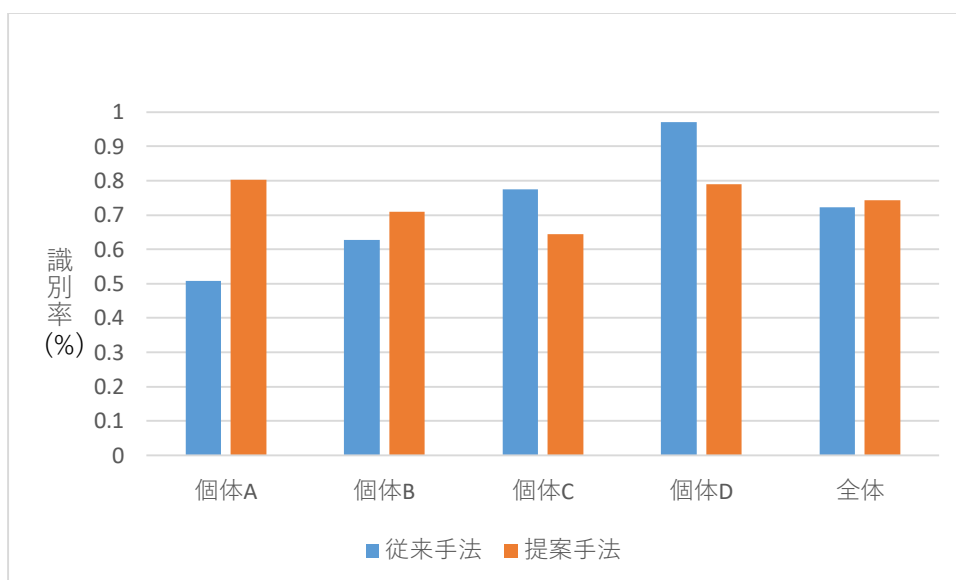


図 17 従来手法(青)と提案手法(赤)による個体識別結果

4.3. 未知個体識別の評価実験

テストデータ \mathbf{y} をギャラリー \mathbf{X}_i で近似したときの差である d^i と、テストデータ \mathbf{y} をギャラリー \mathbf{X} で近似する係数 $\boldsymbol{\beta}$ とを組み合わせることで、未知個体検出の条件を細かく分ける。個体分類の信頼度が低く、かつ、 $\boldsymbol{\beta}$ が分類に貢献していない場合に、未知個体であると判定する(式(9))。

この方法によって従来手法から誤検出率を改善できるか比較実験によって確認する。

4.3.1. 比較実験

1196~1200 フレームの画像をギャラリーとし、後続の 1201~1400 フレームの画像をテストデータとして未知個体識別を行う。

まず、従来手法において未知個体識別に最適な γ の値を求める。 γ の値を変更しながら未知個体識別を行った結果を図 18 に示す。F 値が最も高いときの γ 、つまり $\gamma = 0.5$ を未知個体識別の最適な値とする。

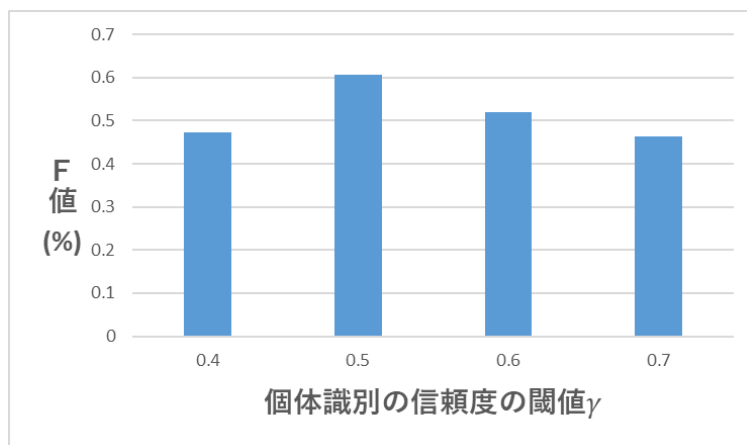


図 18 従来手法による未知個体識別結果の F 値

$\gamma = 0.5$ のときの従来手法に対して、分類の信頼度を加えた提案手法が、未知個体識別率を改善できるか比較実験によって検証する。信頼度の閾値 δ の値を変えながら提案手法による未知個体分類を行った。

4.3.2. 結果・考察

結果を図 19 示す。青の横線は、 $\gamma = 0.5$ のときの従来手法による未知個体識別結果の F 値を表す。赤の棒グラフは、 $\gamma = 0.5$ で、分類の信頼度の閾値 δ を 0.05 ずつ変更したときの、提案手法による未知個体識別結果の F 値である。図を見ると δ が 0.25~0.3 の範囲で、従来手法の結果を上回っている。 $\delta = 0.35$ 以降では、未知個体判定の条件が厳しくなりすぎたため、未知個体の識別数が 0 となり、F 値が従来手法を下回った。

未知個体判定のパラメーターを追加したことで、より細かい状況設定が可能になり、適切な値を与えれば識別誤りを改善できることが分かった。

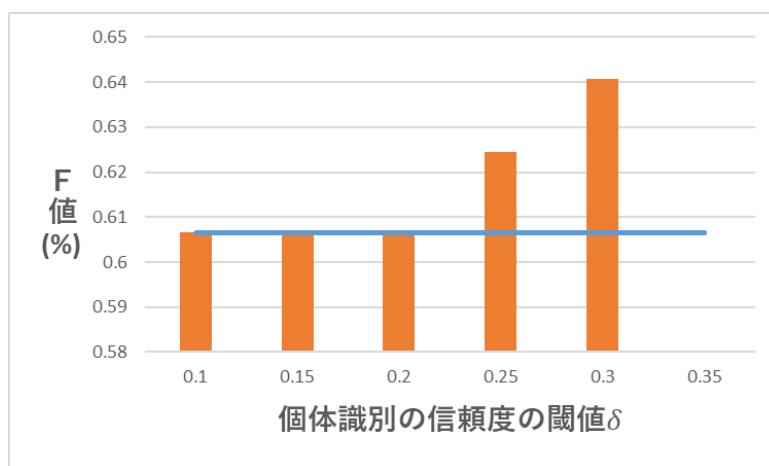


図 19 提案手法未知個体識別結果の F 値

5 結論

本論文では、一般環境下でのマウスの識別と追跡に必要な要素技術の課題を検討した。

「物体検出」処理では、一般背景下において背景とマウスの色が似ている問題がある。従来の物体検出手法である SSD では、背景とマウスを混合させてしまい、クラスの信頼度をもとにした閾値処理での背景の除去処理でマウスを除去する場合があった。これに対して、提案手法では画像分類によるマウスと背景の分離を行い、比較実験によって従来手法の結果より、適合率の点で高い結果を表した。

「個体識別」処理では、与えるギャラリーが偏ることで識別結果に偏りが生じる問題があった。これに対して、提案手法で画像変換によるデータ拡張を行うことでギャラリーの偏りをなくし、比較実験によって改善できたことを示した。

「未知個体識別」処理では、マウスは個体別の特徴差が少ない問題があり、人物の場合に比べて未知と既知の場合分けが困難であった。これに対して、提案手法では識別のためのパラメータを追加することで、細かい条件分けが可能となり、比較実験によって未知個体検出率の改善できたことを示した。

今後の課題として、まず、「物体検出」処理の検証が挙げられる。本研究の実験で、SSD によるマウス検出は高い検出精度を示した。しかし、それは学習用データの元映像と実験に用いた映像の類似性によるものの可能性がある。汎用性の高い検出処理を得るためにも、異なる環境で撮影した映像を用いて実験し、検証する必要がある。その後、改めて提案手法との比較実験を行い、提案手法の評価を行う。

他の課題としては、ギャラリーに加えるデータの評価が挙げられる。マウス識別の評価実験で、ギャラリーに偏りがあると識別性能が低くなる傾向が見て取れた。識別結果は追跡結果に直結するので、ギャラリーが偏らないためにも追加データの評価方法が必要である。また、正しい識別とギャラリーの追加を行うために、個体識別結果の誤りを修正する仕組みも必要である。これらの課題を解決した後で、要素技術を総括して追跡システムを構築することが挙げられる。

謝辞

本研究を行うにあたり、多くの方々に協力していただきました。最後まで熱心に指導をしていただいた指導教員の椋木雅之教授に感謝いたします。また、今回の実験で使用したアカネズミの映像の撮影に協力してくださった坂本信介講師、三原さん、右京さんに感謝いたします。研究の相談や助言を快く受けていただいた岡さん、荻原さん、並びに椋木研究室の皆様感謝いたします。

大変お世話になりましたことをこの場をお借りしてお礼申し上げます。

参考文献

- [1] Wenhan Luo, Junliang Xing, Anton Milan, Xiaoqin Zhang, Wei Liu, Xiaowei Zhao, Tae-Kyun Kim, “Multiple Object Tracking: A Literature Review,” arXiv:1409.7618v4, 2017.
- [2] 岡崎浩佑, 阿山駿希, 椋木雅之, “尤度計算に AKAZE 特徴量を利用したパーティクルフィルタによる小動物追跡,” 第 70 回電気・情報関係学会九州支部連合大会, 2017.
- [3] Wei Liu, Dragomir Anguelov, Dumitru Erhan, Christian Szegedy, Scott Reed, Cheng-Yang Fu, Alexander C. Berg, “SSD: Single Shot MultiBox Detector,” arXiv:1512.02325v5, 2016.
- [4] 大岐建次郎, “深層学習と画像変換を用いた複雑背景における物体検出,” H29 年度 宮崎大学 情報システム工学科卒業論文, 2017.
- [5] Weizhi Lu, Cong Bai, Kidiyo Kpalma, Joseph Ronsin, “Multi-object tracking using sparse representation,” 2013 IEEE International Conference on Acoustics, Speech and Signal Processing, 2013.
- [6] 藤吉弘亘, 山下隆義, “深層学習による画像認識,” 日本ロボット学会誌 35 巻 3 号, 2017.
- [7] 柳井啓司, “一般物体認識の現状と今後,” vol.48, no.SIG16(CVIM19), pp.1-24, 2007.
- [8] Yang Wu, Masayuki Mukunoki, Michihiko Minoh, “Collaborative Mean Attraction for Set Based Recognition,” MIRU, 2014.
- [9] Hiroki ogihara, Masayuki Mukunoki, “Image Classification using Collaborative Mean Attraction with Sparse Optimization,” ICPRAI, 2018.
- [10] 今西勁峰, “深層学習におけるデータ前処理ノウハウ,” 日本化学会 情報化学部会誌 36 巻 2 号 pp. 35-38, 2018.
- [11] 外山託海, 黄瀬 浩一, “一般物体認識における学習データセットのサイズ拡張による影響の解析,” 電子情報通信学会技術研究報告. MI 医用画像 pp. 145-150, 2010.

