

物体検出における形状の類似性が与える影響の調査

原田雅文*¹ 松村遼*² 北風裕教*²

Investigating the Effect of Shape Similarity on Object Detection

Masafumi HARADA, Ryo MATSUMURA and Hironori KITAKAZE

Abstract

In recent years, various damages by crow, a verminous bird has increased. Our research group have been developing an automatic eviction system for verminous birds using drones to reduce the damage caused by crow, cormorants, etc. The input to this system is images of verminous birds taken by a camera mounted on a drone. This image will be used as training data for Convolutional Neural Network. However, the problem with this method is that it cannot identify birds that are similar in color and shape. Therefore, this paper focuses on the shape information, and training has been provided by YOLOv3. As a result, an airplane similar in shape to a crow couldn't be identified as a crow properly. From this, it was confirmed that shape similarity is an impediment to identification of birds. So, training that emphasizes the characteristics of each bird species is necessary.

Keywords: Convolutional Neural Network(CNN), You Only Look Once(YOLO), Data Augmentation(DA)

1 まえがき

野生鳥類（以下、なんらかの被害を与える鳥類を“害鳥”と呼称する）による農業や漁業への被害は、年々増加傾向にあり、日本全国で発生している。農林水産省が発表している全国の農作物被害状況¹⁾によれば、害鳥による被害は2019年度で約31億円となっている。

由井²⁾によれば、旧来から鳥害の防止に、音や光を用いた心理的防止法、網や鉄線を用いた物理的防止法、忌避剤や粘着剤を用いた化学的防止法、人や動物による直接的威嚇法、栽培方法に変化を加える耕種法といった手法が行われている。しかし、これらの手法ではカラスのような知能が高い害鳥はすぐに慣れてしまい、継続的な効果を発揮できない上、コストの高さや二次的影響の発生が懸念されるといった問題点が多く存在する。これらの手法改善を試みたものとして、高山³⁾らは釣り糸を格子状に設置することでカラスの防除効果を得ようとしたが、この手法もカラスによる被害が抑えられたのは設置から7日間のみで、それ以降は効果が示されていない。比較的長期的な効果を得られたのは遠藤⁴⁾の手法である。この手法は、カラスの警戒声等を基にした音声及び近紫外線レーザーを組み合わせる方法であり、カラスの営巣する場所を2年間に渡り追いやることに成功している。しかし、この手法は音量によって忌避効果が左右

されてしまうため、音量次第では近隣住民に対して音による二次的影響の発生が考えられる。

一方、近年ではドローンによる害鳥の防除が検討されている。水元⁵⁾らはドローンを用いることでカラスやムクドリを追い払うことに成功しており、害鳥の防除にドローンが有効であることを報告している。しかし、ドローンの操縦には資格を有した経験豊かな操縦者の協力が必要である。また、長時間にわたって操縦者の拘束も必要となる。さらに、長時間利用可能な大型のドローンとバッテリーを準備する必要もある。このことから、24時間の継続的な害鳥対策が難しいといえる。

ドローンの自動操縦を用いた害鳥の追払いシステムとして、中島⁶⁾らは、果樹園に固定監視カメラを設置し、害鳥を検知するとドローンが自動的に飛来し、害鳥を追い払う手法を提案している。ただし、この手法では果樹園という限定的な領域での適用を前提としているため、害鳥の具体的な種を識別してはいない。基本的に、領域に侵入した鳥類を全て排除すればよい場合を想定している。この方法は、例えば無農薬農法の1つである合鴨農法を実施する場合においては、合鴨にも影響があるため、対応ができない。実際、鳥類には多様な種が混在しているため、多種の野鳥から対象となる害鳥のみを見分けてドローンで対応することは、現在の技術では難しい。

* 1 電子・情報システム工学専攻 * 2 情報工学科

最新のドローンは、目的地を事前に設定することで自動操縦が可能である他、中空からの撮影に利用されている背景がある⁷⁾。著者らの研究室ではこれらの点に着目し、ドローンに搭載したカメラで対象となる害鳥の認識を行い、追跡による追払いを自動で行う害鳥の追払いシステムの開発に着手している。また、通常のドローンには航空法等による規制が発生するため、これらの規制が生じないトイドローン（200g未満のドローン）を想定する。

著者らの求めるシステムは農作物被害の減少に限定しない、多様な害鳥被害に適用できる追払いシステムであり、ドローンの巡回領域を果樹園等に限定しない点で中島らの手法と大きく異なる。それは同時に、一切関係のない鳥類や鳥類以外も存在する領域をドローンで巡回することを意味するため、認識した鳥類の具体的な種を識別する能力が求められる。著者らは、害鳥の識別器として、畳み込みニューラルネットワーク（Convolutional Neural Network、以下 CNN）をベースとした物体検出のアルゴリズムである You Only Look Once（以下 YOLO）⁸⁾のバージョン3（以下 YOLOv3）⁹⁾を利用し、多様な害鳥の認識について研究を行ってきた。このアルゴリズムを本システムに実装する理由として、YOLOv3はリアルタイム処理に優れており、バウンディングボックス（以下 BB）と呼ばれる矩形領域による対象物体の位置の特定が可能のため、ドローン制御に向いていることが挙げられる。著者らのこれまでの研究において、YOLOv3に1種類の害鳥を入力して学習を行い、認識実験を試みた結果、カラスのみでは85%（カワウのみでは91%）の高い認識率で対象物体に対して BB を正確に導くことを確認している¹⁰⁾。しかし、この手法は未学習の対象物体（任意の野鳥）においても、形状が似ていれば、BB でオブジェクトを囲むため、誤認識をする問題があった。また、2種類のカラスとカワウを YOLOv3で同時に学習（2クラス学習）を行い、認識実験も試みているが、カラスは54%、カワウは30%まで認識率が低下した。この原因については明らかになっておらず、CNN における特徴の重要度の鍵を握る。特徴の重要性に関する研究として、Robert¹¹⁾らは、ImageNet をデータセットとして用いた学習において、全体的な輪郭よりも、局所的な特徴（テクスチャ）を重要な特徴として捉えるようになったと報告している。また、テクスチャをスタイル変換によって、テクスチャから



図1 カラスと同時学習を行う物体

特徴が見いだせない状態で学習させると、形状情報から特徴を得ることが可能になることを推察している。一方で、Aditya¹²⁾らは CNN における色情報の重要度はデータセットに依存すると考察している。これらのことから、CNN が見出す特徴はデータセットによって大きく変化することがわかる。すなわち、著者らの既存手法によるデータセットで学習させた場合、形状を重視して特徴を捉えてしまい、それが鳥類識別の障害となっていると予想できる。

そこで本研究では、YOLOv3で鳥類の多クラス認識における認識率の低下の原因を明確にすることを目的とする。対象物体が酷似した場合に認識率が極端に低下することを考慮して、実験では、YOLOv3によってカラスと明らかに形状の異なる物体である木製の『積み木（直方体）』、積み木よりも物体として構造的な意味を持つ形状の『車』、翼を持つことで、車よりも鳥に近い『飛行機』を、『カラス』と同時に学習させた場合、識別結果はカラスとカワウの同時に学習をした時と比較して、どの程度認識結果に違いがあるか検証を行う。例えば、カラスと積み木では明らかに異なる物体であるため、認識率は高く示すはずであるが、カラスと飛行機の場合は、翼が存在することで積み木と学習させた時よりも認識率が低下すると考えられる。これらの形状情報の影響の大きさについて明らかにできれば、カラスとカワウの形状情報の違いに対する認識精度の向上を考察するうえで、今後の方針を検討するための足掛かりになる。

2 学習用データセットの準備

2.1 学習用画像

CNN の学習において、カラス画像と比較するために用いる学習データとして、図1の例で示す木製の『積み木（直方体）』、積み木よりも物体として構造的な『車』、翼を持つことで、車よりも更に鳥に近い『飛行機』の3種類を用いる。これらのデ

クラス番号 中心座標x 中心座標y BBの幅 BBの高さ

```
0 0.324219 0.255208 0.054688 0.072917
0 0.395310 0.277080 0.053125 0.070833
0 0.427344 0.190625 0.029687 0.039583
```

図2 YOLOv3のアノテーションファイルの情報

ータは、以降説明する Data Augmentation によって各クラス7,380枚ほど準備する。

2. 2 従来手法における Data Augmentation

CNNにおける認識率を向上させる方法の1つに、学習用画像を増加させることが挙げられる。著者らの研究グループでは、わずか1枚の2D 画像から学習用画像を効率的かつ大量に生成する独自の Data Augmentation を提案している¹⁰⁾。この手法は、認識対象となる害鳥の2D 画像から知識情報を加えて3D アニメーションを生成し、Unity Technologies 社がリリースしているゲームエンジンである Unity¹³⁾で様々な方向から2D 画像を再取得 (撮影) することで実現している。

一方で、YOLOv3のような物体検出タスクを持つ CNN の正解ラベルとして、画像中の BB を示したアノテーションファイルが必要となる。YOLOv3の学習に対応したアノテーションファイルは、図2の1行目 (枠で囲まれた範囲) で示すフォーマットで記述されたテキストファイルである。クラス番号は BB が属しているクラス番号を示し、BB の中心座標は左上を座標(0,0)、右下を座標(1,1)として0~1の実数値で示される。また、BB の大きさについても同様に0~1の実数値で示される。図2の2行目以降には実際の記述例を示す。画像中に複数の BB が存在する場合は、テキストファイルの行を変える (改行コードを与える) ことで区別を行う。しかし、画像中の BB を座標で指定するのは困難である。そのため、アノテーションファイルは `labelImage`¹⁴⁾ のような補助ツールを使い、画像から視覚的に BB を指定することで作成する。しかし、この作業は学習用画像が増えるほど増加するため、過度な学習用画像の生成は時間及び人的コストの面において Data Augmentation の恩恵を受けにくくなるという問題があった。

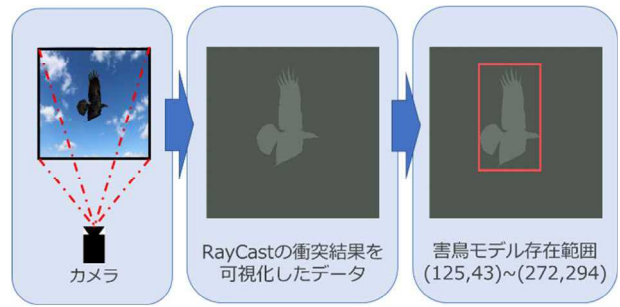


図3 RayCastを用いた自動アノテーション

2. 3 Data Augmentation の改善手法

アノテーションは本来、人が手動で行う。これは学習用画像という単一の情報から認識対象が存在する領域を指定する必要があるからである。ここで、Unity上の害鳥の3Dモデル (以下、害鳥モデル) をキャプチャする本手法の場合、画像情報だけでなく、害鳥モデルの位置等の情報も取得可能である。そこで、この情報を用いて、従来の Data Augmentation を行うプログラムに改良を加え、アノテーション工程も同時に行う処理を追加した。従来手法と大きく異なる点はキャプチャ処理とカメラ回転処理の間にアノテーション関連の処理を記述した点と、カメラの処理と排他的であったアニメーションの再生処理を一連の処理に組み込んだ点である。

YOLOv3の自動アノテーションを実現するためにあたって、Unity 上における物体の衝突判定 (コライダー) と RayCast を用いた。コライダーはゲームにおける“当たり判定”を意味し、Unity 上の衝突はコライダーとの接触を指す。RayCast はカメラの画面上から任意の角度へ照射することができるレーザのようなもので、コライダーと衝突すると、衝突したコライダーの情報を得ることが可能である。そこで、対象物体 (害鳥モデル) にコライダーを持たせ、キャプチャと同時にカメラ画面における全てのピクセルを左上から順番に RayCast をまっすぐに飛ばすことで、物体の位置や形状を数値で得ることとした (図3)。

ここで問題となるのは、通常のアニメーション再生処理ではコライダーが追従しない点である。害鳥モデルの形状に一致したコライダーは Unity の Mesh Collider によって生成可能であるが、3D モデルの細かい形状をコライダーとしてリアルタイムに生成することは計算量の観点で問題があり、現時点のコンピュータの性能では図4のようにズレが

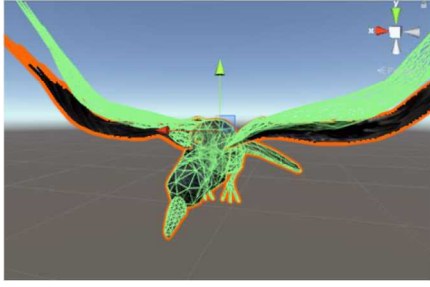


図4 コライダー（緑）生成の遅延によるズレ

生じる。そのため、ゲーム等では、通常、カプセル型のコライダー(Capsule Collider)や直方体のコライダー(Cube Collider)といった簡略化したコライダーをアニメーションの標本となる骨格情報に追従させ、動的なコライダーを表現している¹⁴⁾。しかし、本手法では、アニメーション再生時、害鳥モデルのコライダーと害鳥モデルの形状が完全に一致する必要があるため、上記のような方法は適切ではない。

この問題を解決するために、カメラ処理と排他的にアニメーションを再生するのではなく、害鳥モデルをキャプチャする直前に害鳥モデルのアニメーションをコマ送りする処理を挿入する。この手法による利点は、アニメーションによる害鳥モデルの形状変化に合わせた Mesh Collider の生成までキャプチャ処理を待機させることができる点であり、アニメーションに追従したコライダーを疑似的に再現することを可能にする。

次に、害鳥モデルが存在する矩形領域を求める。これは、上記の手法によって害鳥が存在する領域（画素）から、左端 p_{left} 、右端 p_{right} 、上端 p_{top} 、そして下端 p_{bottom} をそれぞれ求める。これによって得られた値と画像の幅 I_{width} 、高さ I_{height} より、式(1)から式(4)によって、YOLOv3用のアノテーションファイルに必要な値である、 x_{center} （中心座標 x ）、 y_{center} （中心座標 y ）、 x_{scale} （BBの幅）、 y_{scale} （BBの高さ）を求める。

$$x_{center} = \frac{1}{I_{width}} \left(\frac{p_{right} - p_{left}}{2} + p_{left} \right) \quad (1)$$

$$y_{center} = \frac{1}{I_{height}} \left(\frac{p_{bottom} - p_{top}}{2} + p_{top} \right) \quad (2)$$

$$x_{scale} = \frac{1}{I_{width}} (p_{right} - p_{left}) \quad (3)$$

$$y_{scale} = \frac{1}{I_{height}} (p_{bottom} - p_{top}) \quad (4)$$

これらの値と害鳥モデルの所属クラスをテキストファイルとして書き込むことでアノテーションが完了し、次のカメラ移動とコライダー生成、キャプチャ処理へ移行する。

3 実験方法

本実験ではカラスの他に『積み木』、『車』、『飛行機』のいずれかのクラスを1つ追加して、2クラス学習を行う。すなわち、カラスと積み木、カラスと車、カラスと飛行機の3パターンで YOLOv3による学習および認識を行う。

まず、カラスと積み木を同時に学習させる。この時、学習用データセットは前述の改善された Data Augmentation によって1クラスにつき画像7,380枚を生成する。生成したデータセットを YOLOv3 実験用のデータセットと、以降説明するIoUの閾値を指定することで適合率や再現率を算出することができる機能があるため、これを利用する。評価が終了したら、車と飛行機も同様の手順で学習用データセット生成から評価まで行う。

4 評価方法

以上の条件で YOLOv3による学習を行った後、各クラスのテスト用画像35枚による認識率の評価を行う。本研究ではIoU50%以上で正解 BB として判断し、予測 BB が正解なら True Positive（以下TP）、予測 BB が不正解なら False Positive（以下FP）、正解 BB を予測できなかった場合を False Negative（以下FN）として分類する。そして、これらの値から適合率Precision、再現率Recall、F値を求め、評価を行う。IoUは予測 BB と正解 BB が重なっている割合を示し、予測 BB と正解 BB を合わせた領域 union、予測 BB と正解 BB が共通している領域 intersectionから(5)式で求めることができる。適合率は予測 BB が実際に正解 BB である割合のことを指し、間違った物体を認識するとこの値が低下する。適合率は、(6)式で示される。再現率は正解 BB のうち正しく予測できた BB の数の割合を指し、認識漏れが発生するとこの値が低下する。再現率は、(7)式で示される。F値は適合率と再現率の調和平均のことを指し、(8)式で示される。

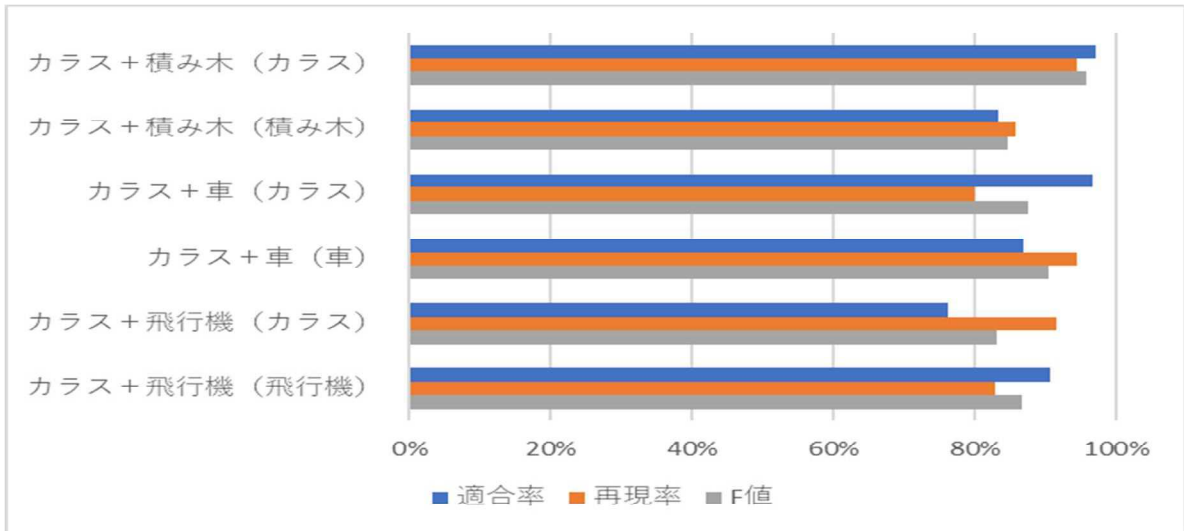


図5 各実験結果

表1 カラスと積み木の認識結果

	適合率	再現率	F値
カラス	97.10%	94.30%	95.7%
積み木	83.30%	85.70%	84.5%

$$Recall = \frac{TP}{TP + FN} \quad (7)$$

$$F = \frac{2(Precision \cdot Recall)}{Precision + Recall} \quad (8)$$

表2 カラスと車の認識結果

	適合率	再現率	F値
カラス	96.60%	80.00%	87.5%
車	86.80%	94.30%	90.4%

表3 カラスと飛行機の認識結果

	適合率	再現率	F値
カラス	76.20%	91.40%	83.1%
飛行機	90.60%	82.90%	86.6%

$$IoU = \frac{intersection \cap union}{intersection \cup union} \quad (5)$$

$$Precision = \frac{TP}{TP + FP} \quad (6)$$

5 実験結果と考察

5.1 2クラス認識実験の結果

カラスと積み木、カラスと車、カラスと飛行機を同時に学習させた時の認識結果を表1から表3に示す。また、各適合率、再現率、F値について比較しやすいようにグラフにした結果を図5に示す。

まず、現在の学習方法でYOLOv3が最低限の識別能力を持っていることを確認するために、『カラス』と『積み木』による2クラス学習を行い、評価実験を行う。これを実験Aとする。これらの物体で実験を行う理由は、お互いの外見の特徴に共通点がないことが明らかであり、識別の難易度は極めて低いと予想されるためである。学習後、『カラス』を入力信号とした際の適合率は97.1%、再現率が94.3%、そしてF値が95.7%となった。この結果、ほぼ全てのカラスを認識することができ、誤検出もほとんどみられなかった。一方、同実験において『積み木』を入力信号とした際の適合率は83.3%、再現率が85.7%、そしてF値が84.5%となり、誤検出や認識漏れが多い結果となった。これを『カラス』の場合と比較した場合、適合率、再現率、F値のいずれにおいても10%以上も結果が悪いといえる。目視で実際の画像を確認した場合、積み木の

誤検出のほとんどは、対象物体に対してひと回りもふた回りも大きく BB が取りすぎたものとなっていた。したがって、カラスを積み木として、または積み木をカラスとして誤検出している例は全くなかった。すなわち、BB の大きさを完璧に推察することはできなかったものの、『カラス』と『積み木』という明らかに異なる2つの物体を識別できることを確認した。

次に、『カラス』と『車』による2クラス学習を行い、構造的特徴の有無による識別能力を実験 A と比較するための検討を行う。これを実験 B とする。これらの物体も実験 A と同様、外見の特徴で共通している部分はない。しかし、『車』は『積み木』と比較してライトやタイヤといった構造的な特徴を保有しており、『カラス』も頭や羽といった構造的な特徴を持っている。これらは異なる構造、外見ではあるが、シンプルな直方体の物体である『積み木』よりも特徴量が複雑になり、実験 A よりも識別難易度が高くなると予想される。学習後、『カラス』を入力信号とした際の適合率は96.6%、再現率が80.0%、そして F 値が87.5%となった。カラスの認識漏れが実験 A と比べて多い結果となった。一方、同実験において『車』を入力信号とした際の適合率は86.8%、再現率が94.3%、F 値が90.4%となった。この結果から、適合率においては、カラスを入力信号とした際よりも劣るが、再現率と F 値においては、カラスよりも良い結果であった。目視で実際の画像を確認すると、カラスを含めた背景部分を車として誤検出する例がいくつかあり、これが車の適合率の低下につながっていることが明らかとなった。ただし、カラスと車の比較においても、カラスを車として、車をカラスとして認識する例はなかった。このことから、2つの学習クラスにそれぞれ構造的特徴が存在した場合においても、外見上の共通点がなければ、お互いのクラスを識別することは可能であることを確認した。実験 B の結果をふまえ、構造的特徴だけでなく、物体の輪郭の特徴がよく似ている構造を持つ場合においても比較実験を試みた。これを実験 C とする。実験 C では『カラス』と同時学習させる対象として、構造的特徴を持った上で、輪郭がカラスと類似している『飛行機』を選択した。本実験において『カラス』と『飛行機』を識別することができれば、輪郭は物体の2つの物体の識別において重要な特徴ではないことが明らかになる。

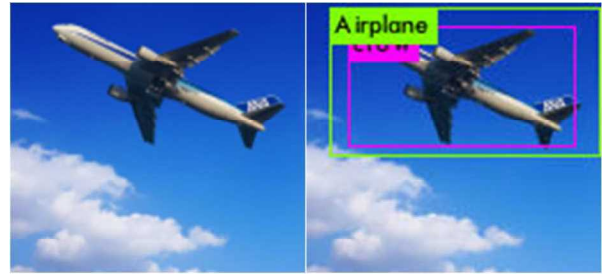


図6 カラスと飛行機を識別できない例

学習後、『カラス』を入力信号とした際の適合率は76.2%、再現率が91.4%、そして F 値が83.1%となった。この結果は、実験 A や実験 B の積み木や車の結果と比べて適合率が明らかに低い傾向がみられた。再現率に関してはそれほど低い傾向はなかった。一方、『飛行機』を入力信号とした際の適合率は90.6%、再現率が82.9%、そして F 値が86.6%という結果となった。こちらは、実験 A や実験 B の積み木や車の結果と比べて適合率が高い結果であったが、再現率は、低い結果であった。カラスと飛行機の適合率、再現率、F 値を比較した場合、適合率は明らかに、飛行機が高い結果であった。目視で実際の画像を確認すると、これらの原因は飛行機をカラスとして認識した例がいくつか存在するためであった。また、カラスを飛行機として認識した例も確認したが、多くの誤認識例としては、飛行機をカラスとして誤検出する、またはカラスと飛行機の両方のクラスに属していると認識したものがほとんどであった(図6)。このことから輪郭の特徴が共通している物体の学習を行う実験 C は、実験 A と実験 B における比較結果とは全く異なる結果となることを確認した。

5. 2 2クラス認識実験の考察

実験 A、実験 B の結果から2クラス学習において、対象物体の構造的特徴の有無に関わらず、外見上の特徴の共通点が見られない場合は、YOLOv3によって十分に識別することができると考えられる。これに対し、実験 C の結果からカラスと飛行機の組み合わせにおいては、適合率において問題を持つことが明らかであった。互いに『胴体と翼を持っている(飛行機の形状が、翼を広げて飛行しているカラスと同じような輪郭を持っている)』という特徴を持つことが原因として推察される。これは、カラスとカワウほどではないものの、認識結

果から2つのクラスが混同して学習されていると解釈できる。

ここで、車と飛行機を考えると、どちらも金属からできて光沢を持つ構造的な特徴を持っている。カラスは金属のような光沢はなく、光源を反射することは少ない。カラスと車は高い適合率である一方、カラスと飛行機は適合率が低いことから、金属のような材質（色やテクスチャ情報）を深層学習で学習しているわけではなく、形状情報の違いを学習している可能性が高いと考える方が自然である。色やテクスチャ情報はカラスと飛行機では類似点はあまり存在しない。飛行している飛行機下部に影ができ、カラスのような黒の単色に見える点ぐらいである。

これらのことから、形状情報が原因で、カラスと飛行機をいくらか混同してしまった可能性が高いため、形状の特徴が似ているものを学習する場合は、色やテクスチャ情報の特徴を別の手段で、学習するための工夫が必要であり、その手法について検討しなければならない可能性がある。この考察を確かなものにするために、以下に、テクスチャ情報の影響を確認するための補足実験を行った。

5. 3 テクスチャ情報に着目した補足実験

前節の結果から、色やテクスチャ情報の重要度に関して影響度を調べるために、図7で示すような積み木のテクスチャをカラスの模様にして、カラスと積み木の2クラス学習における補足実験を行った。認識対象としたのはカラスの画像35枚と積み木の画像35枚である。学習パラメータの変更は行っていない。

補足実験の結果を表4に示す。『カラス』を入力信号とした際の適合率は96.9%、認識率が88.6%、F値が92.5%であった。一方、『積み木 (カラス模様)』を入力信号とした際の適合率は96.6%、再現率が80.0%、F値が90.0%となった。

5. 4 補足実験Dの考察

テクスチャ情報に着目した補足実験の結果において注目する点は、カラスと積み木を全く混同しなかった点と、学習させた積み木と評価実験用の積み木のテクスチャは全く異なるにもかかわらず積み木を認識することができた点である。もし、色やテクスチャが重要な特徴として捉えられている場合、カラスの画像からカラスと積み木の両方

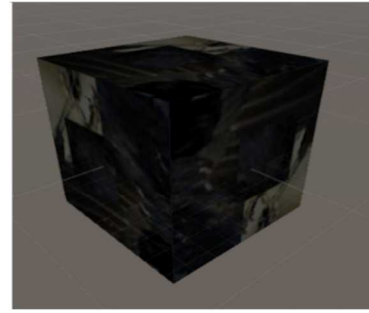


図7 カラスのテクスチャを貼りつけた積み木

表4 カラスと積み木 (カラス模様) の認識結果

	適合率	再現率	F値
カラス	96.90%	88.60%	92.5%
積み木 (カラス模様)	96.60%	80.00%	90.0%



図8 カラスとカワウの外見上の比較

のBBが検出され、積み木画像からは何も検出されないと考えるのが自然である。しかし、補足実験の結果、積み木は立方体形状であるという情報のみでカラスと区別され、積み木を正しく認識することができた。以上のことから、輪郭情報は物体認識における重要な特徴であり、鳥類の識別が困難である原因は輪郭情報が類似している点だけということが明らかになった。この結果から、鳥類の識別に必要なことは、輪郭情報以外の特徴を重点的に学習させる手法の考案である。

前述のとおり、どの特徴が重視されるかはデータセットに大きく依存する。すなわち、本手法で用いているデータセット作成方法では、輪郭情報が重要視されているということである。輪郭情報は、単なる分類問題と異なり位置や大きさの情報も必要となる物体検出において重要な特徴である

ため、それを排除した学習は現実的ではない。よって鳥類識別を行うためには、本手法によるデータセット生成手法を基に、対象となる鳥類の特徴が色濃く出ている部分を強調した学習法を加える必要があると考えている。例えば、カラスの特徴を挙げると、全体的に体色が黒く、首が短い。一方、カワウも体色の大部分が黒色だが、白い模様があり、首が長く、くちばしが黄色いという特徴がある。このように、一見類似しているように見えるが、局所的な特徴はそれぞれ異なる(図8)。これらの特徴を集中的に学習させるための具体的な手法の考案は今後の課題である。

6 あとがき

本研究は、独自開発しているドローンを用いた害鳥追払いシステムにおいて、害鳥の識別がうまく動作しない問題の要因を解明するための調査研究である。識別は CNN を基礎に持つ深層学習 YOLOv3 を用いているため、この YOLOv3 の認識部がどのように対象物体の特徴を捉えているのかを確認することで、その要因について検討を行った。実験では、対象物体を『積み木』、『車』、『飛行機』として、カラスと同時に YOLOv3 でそれぞれ2クラス学習を行い、評価実験を行った。学習効率を上げるために、3D モデルを用いた従来の Data Augmentation 手法を改善し、アノテーション処理を自動化した。認識結果は積み木と車はカラスと完全に識別することができたが、飛行機はいくつかカラスと混同してしまった。この要因から、飛行機の形状はカラスが羽を広げた形状と類似しており、色やテクスチャ情報よりも形状情報が大きく識別に関係していると考察することができた。したがって、現行手法では、飛行機よりもさらに形状が近くなる鳥類同士の識別は難しく、ドローンに搭載するシステムには特別な識別を行うアルゴリズムを導入する必要があると結論づけた。

今後は、害鳥追払いシステムの実システム化を目指し、鳥類の多クラス認識が可能な識別手法の実現に向けてシステム改良に取り組んでいく。

参考文献

- 1) <https://www.maff.go.jp/j/press/nousin/tyozyu/attach/pdf/201223-2.pdf>, (2021.9.06現在)
- 2) 由井正敏, 農作物・果樹の鳥害とその防止対策, 農林水産技術研究ジャーナル, Vol.3, No.4, pp.52-56, (1980)
- 3) 高山耕二, 中村真紀子, 内山雄紀, 福永大吾, 赤井克己, 中西良孝, アイガモ農法におけるカラス害防除に関する研究, 西日本畜産学会報, Vol.51, pp.83-88, (2008)
- 4) 遠藤彰, 音声及びレーザを組み合わせた鳥害防止装置の開発, 電気設備学会誌, Vol.38, No.11, pp.663-666, (2018)
- 5) 水元廣祐, 伊部裕貴, 安彦智史, ドローンを用いた害鳥対策に関する調査研究 —福井の農業への利用性—, 情報システム学会 第12回全国大会・研究発表大会, (2016), https://www.jstage.jst.go.jp/article/proceedingsissj/12/0/12_d23/_pdf/-char/ja (2021.09.07現在)
- 6) 中島幸一, 清尾克彦, 北上眞二, 汐月哲夫, 小泉寿男, 鳥獣被害対策用監視・防除 UAV システムの開発と評価, 研究報告コンシューマ・デバイス&システム(CDS), Vol.14, No.10, pp.1-8, (2015)
- 7) 野波健蔵, 世界のドローン開発動向と農業応用からみた課題と展望, 計測と制御, Vol.55, No.9, pp.780-787, (2016)
- 8) Joseph Redmon, Satosh Divvala, Ross Girshick, Ali Farhadi, You Only Look Once: Unified, Real-Time Object Detection, arXiv preprint arXiv:1506.02640, (2015)
- 9) Joseph Redmon, Ali Farhadi, YOLOv3: An Incremental Improvement, arXiv preprint arXiv:1804.02767, (2018)
- 10) 北風裕教, 吉原蓮人, 岡部蒼太, 松村遼, オブジェクト検出 YOLO を用いた害鳥認識システムの開発, 産業応用工学会論文誌, Vol.8, No.1, pp.q10-16, (2020)
- 11) Robert Geirhos, Patricia Rubisch, Claudio Michaelis, Matthias Bethge, Felix A. Wichmann, Wieland Brendel, ImageNet-trained CNNs are biased towards texture; increasing shape bias improves accuracy and robustness, arXiv preprint arXiv:1811.12231, (2018)
- 12) Aditya Singh, Alessandro Bay, Andrea Mirabile, Assessing The Importance Of Colours For CNNs In Object Recognition, arXiv preprint arXiv:2012.06917, (2020)
- 13) <http://www.unity3d.com/unity>, (2021.09.08現在)
- 14) <https://github.com/tzutalin/labelImg>, git code, (2021.09.08現在)
- 15) <https://github.com/AlexeyAB/darknet>, git code, (2021.09.08現在)