

XAI（説明可能なAI）を踏まえた畳込みニューラルネットワークによるトラフグ属の種判別モデルの検討

メタデータ	言語: Japanese 出版者: 水産研究・教育機構 公開日: 2024-07-25 キーワード (Ja): キーワード (En): Species Identification; Tigerfish; Convolutional Neural Network; Deep Learning; XAI(eXplainable Artificial Intelligence); Grad-CAM 作成者: 石田, 武志, 芦田, 寛治, 徳永, 憲洋 メールアドレス: 所属:
URL	https://doi.org/10.57348/0002010423

XAI（説明可能なAI）を踏まえた 畳込みニューラルネットワークによる トラフグ属の種判別モデルの検討

石田武志^{1*}・芦田寛治²・徳永憲洋¹

Species Identification Model of the Tiger Pufferfish Genus using eXplainable Artificial Intelligence(XAI) Based Convolutional Neural Network

Takeshi Ishida^{1*}, Kanji Ashida² and Kazuhiro Tokunaga¹

Abstract : The Tiger Pufferfish (*Takifugu rubripes*) is a staple in Japanese cuisine, with over ten species of the *Takifugu* genus found in the surrounding seas. Given that certain parts of the pufferfish are toxic, they are predominantly prepared by trained professionals. Furthermore, species within the *Takifugu* genus are susceptible to hybridization, leading to an increase in hybrid numbers. However, identifying these hybrids is a challenging and time-consuming task, even for experts. To address this, we developed a transfer learning model using the pre-trained VGG16 model to differentiate between pufferfish species. The VGG16 model, commonly used in image recognition, is built on convolutional neural networks. We also implemented Gradient-weighted Class Activation Mapping (Grad-CAM) for visual interpretation of the model. Grad-CAM generates a heat map that highlights the areas focused on by the AI model in the image, allowing us to identify factors contributing to misjudgment and make further improvements. We used seven species from the *Takifugu* genus (excluding hybrids), and approximately 15 colored images of each species were prepared for machine learning. The results showed that our model was able to distinguish between pufferfish species with relatively high accuracy, although some misclassification occurred among species with similar body patterns. The Grad-CAM results revealed that the model was able to distinguish body patterns, but some misclassifications occurred due to gravel and background objects being recognized as patterns.

Key words : Species Identification, Tigerfish, Convolutional Neural Network, Deep Learning, XAI(eXplainable Artificial Intelligence), Grad-CAM

緒 言

日本近海にはトラフグ属 (*Takifugu*) 14 種が生息しており、このうち 11 種が食用とされている¹⁾。これら14種は、種分化して間もない近縁種²⁾で、雑種が生まれやすいという性質があり、近年、日本近海でもトラフグ属の雑種フグが増加傾向にあるとの報告もなされている³⁾。例えば、2013年の茨城県や福島県沖のフグの試験操業漁獲データにおいても、同地域ではショウサイフグ (*Takifugu snyderi*)

とゴマフグ (*Takifugu stictonotus*) の雑種などが全体の40%近くを占めていたとの報告がなされている³⁾。

一般にフグは種により有毒部位が異なり、体内での毒の分布状況が異なるため、可食できる部位が種毎に行政によって定められている。行政上の規定では、フグの雑種については親種共通で食用とされる部位は食べても良いとされている¹⁾が、親の種類を特定することが困難であり⁴⁾、多くの場合において雑種と疑われる全てのフグは、流通過程から排除されている。雑種か純粋種かの判定は専門家(フ

2023年8月22日受付、2023年11月7日受理

1 水産大学校海洋機械工学科(Department of Ocean Mechanical Engineering, National Fisheries University)

2 水産大学校水産学研究科生(Graduate student, National Fisheries University)

グ漁業者、フグ仲卸業者、フグ調理人など）でも困難である場合があり、最終的にはDNA鑑定という時間とコストがかかる方法を行わざるを得ない⁴⁾。

このような中で、漁業現場において利用できる種・雑種判別（純粋種の特定と純粋種に対して雑種を分別する判別）技術が可能となれば、誤判別による危険性をより低下させることができ、安全性の向上につながると考えられる。また、このシステムを現場に導入することで、雑種判別の精度向上だけでなく、同時に交雑発生状況などの情報の収集も可能になる。筆者らは「雑種も含めたフグ種判別システム」の開発を目指している。このシステムは、漁業現場においてスマートフォンなどの携帯端末によりフグ画像を撮影し、それをサーバシステムに伝送し、サーバ側で種・雑種判別を行い携帯端末に結果を送信するものである。

本研究はこの判別システム構築の基礎として、フグ画像からいくつかのフグ種（純粋種）を特定するモデルを、画像認識において近年標準的な手法になっている畳込みニューラルネットワーク（CNN; Convolutional Neural Network）を用いて構築したものである。学習に必要な雑種の画像数を確保することができなかったため、本研究では純粋種のみモデルを構築した。今後このモデルを雑種も含めたモデルに拡張していくためには、雑種の画像を収集していくことが必要であるが、信頼をもって確保できる画像は多くは収集できないことが予想される。一般にCNNモデルの精度向上には大量の画像を学習させることが必要であるが、少ない学習画像でも判別精度の高いモデルの構築が可能である「転移学習」によるモデル構築を検討した。転移学習は既存の学習済みのモデルの一部を利用する方法である。近年は様々なものを学習している大規模モデルを転移学習により用いたほうが、特定のものの判定にも、少ない画像でより精度が高くなる傾向があることが一般的に知られるようになってきているため、今回は少ない画像でより精度向上を図るため転移学習を用いた。

またCNNなどの深層学習においては、「AI（Artificial Intelligence）は何を根拠に判断しているかがわからない」という欠点がある。これに対して、近年は、AIモデルの出力の根拠を検討することを「説明可能なAI（XAI, eXplainable AI）」と呼び、技術的な手法が整ってきている。本研究は、XAIの手法の一つであるGrad CAM（Gradient-weighted Class Activation Mapping）⁵⁾の手法を用いて、フグ判別モデルが画像中の体模様を見て判断しているのかを確認した。

本研究でのモデルは、雑種の判別などを精度良く判定するための前段階として、体模様を確実に見て判別するモデルの構築を検討したものである。本モデルは、今後の「雑種も含めたフグ種判別システム」の基盤として、雑種などの少数の学習用画像しか集まらない状況下でも利用可能であり、体模様に注視して判定をしていることが確認できるモデルの構築を目標としたものである。

魚の画像判別モデルの既往研究としては、Wäldchen and Mäder⁶⁾ や、Qin⁷⁾ などのように、画像から魚の種判別に、画像認識で高い精度を上げているディープラーニング手法を採用した研究報告が増えている。Qin⁷⁾ は、水中のビデオ画像から魚を認識するシステムをディープラーニング手法を用いて開発しており、98%の正答率が得られたと報告されている。

またオンサイトでの画像や映像から直接、魚の計量や種判別を行う研究開発も行われている。Leeら⁸⁾ は水槽における魚を形状の適合性から種の判別をしている。Strachanら⁹⁻¹¹⁾ は、魚の色彩と形状よりベルトコンベヤー上の魚を識別するシステムを開発している。柴田・齊藤¹²⁾ は、オンサイトでの様々な背景の画像に対して利用できるように、魚画像において、手動で特徴点を与えるアプローチを採っている。8種の形状特徴量と6種のテクスチャ特徴量により、81種各20サンプルの魚画像を用いて高い識別精度が得られたと報告している。

また、主な既往研究の識別率をまとめると以下ようになる。Alsmadiら¹³⁾ は魚画像の腹部矩形領域を切り出し、15個の特徴量を提示している。認識処理には3層のニューラルネットワークを適用し、350枚の画像で学習し、150枚の画像でテストを行い、テスト画像に対して86%の精度を得ている。Larsenら¹⁴⁾ は、形状と模様から認識するシステムを線形判別分析の手法により開発し、3種108画像による識別で76%の精度を得ている。Spampinatoら¹⁵⁾ は、形状と模様を組み合わせ10種360画像による識別で、92%の精度を得ている。さらにHuangら^{16,17)} は、水中カメラで撮影された魚を対象とした認識手法を提案している。魚領域を抽出し、色、形状および模様を合わせた66個の特徴量を定義して、階層型の識別機を用いて10種3,179枚の画像より95%の認識精度を得ている。さらにHuangら¹⁷⁾ においては、15種24,150枚の画像を用いて65%の精度を得ている。国内では、仲松ら¹⁸⁾ は、5個の形状特徴量、4個の色特徴量を用いて70種120枚の画像から26%の精度を得ている。平岡ら¹⁹⁾ は、魚の輪郭線とテクスチャの特徴量を用いる手

法を提案している。判別分析の手法を用いて魚形や種判別を行っている事例としては、Larsen ら¹⁴⁾やYakubuら²⁰⁾などがある。Yakubu²⁰⁾は、判別分析による手法を用いて非常に近い2種の識別を行っている。

特にCNNの利用例としてはRauf ら²¹⁾が32層の層を含むモデルを利用して画像の識別を行なっている。Villonら²²⁾は写真データベースでトレーニングされたCNNのパフォーマンスをテストし、20種類の魚種を識別するためのさまざまな後処理決定ルールを考慮している。魚体全体、魚体の一部、環境（サンゴ礁の底や水など）を含む900,000枚の画像でトレーニングされたものは正答率94.9%で、人間の正答率89.3%を上回っていると報告されている。

これらのように魚種判別に深層学習などのニューラルネットワークを用いた事例は多くあるものの、画像のどの部分により判定が行われているのかまでを検討した事例は見られない状況である。

判別モデルの構築

学習・検証用画像の整備

(1) 対象とするフグ種

対象魚種としては、トラフグ属の中でよく漁業現場で水揚げされるもののうち、体模様が似ている4種（コモンフグ *Takifugu poecilonotus*, クサフグ *Takifugu alboplumbeus*, ショウサイフグ *Takifugu snyderi*, ゴマフグ *Takifugu stictonotus*）、および体模様が似ていない4種（アカメフグ *Takifugu chrysops*, トラフグ *Takifugu rubripes*, シマフグ *Takifugu xanthopterus*, クサフグ）、そしてこれらを合わせた7種類とした。本来であれば雑種も判別対象とするのが好ましいが、十分な数の画像を集めることができないため本モデルでは対象外としている。Fig.1に本研究で対象としたトラフグ属画像および、和名と学名の対応表を示す（本報告では以後、フグの種類を和名標記で示す。）

トラフグ属の模様は、体全体に分布する斑点模様や縞模様などの模様に加え、胸鰭横の黒斑の有無により特徴づけることができる⁴⁾。模様については、クサフグなどのような黒地に白い斑点模様もあれば、逆に白地に黒い斑点模様も存在する。コモンフグ、クサフグ、ショウサイフグ、ゴマフグはいずれも黒地に斑点模様であり、模様からの識別が難しい種である。コモンフグには大小の斑紋が混在する模様や、円環状の模様などが観察される。またシマフグの

ようにきれいな縦模様もいる一方で、アカメフグはほとんど模様がみられないフグである。

(2) 学習・検証用画像の収集

フグは水揚げされる漁港や漁期が限られているため、フグ画像を漁業現場から大量に収集することが困難である。このため、本研究ではインターネット上に掲載されているフグ画像を収集して学習・検証用画像を準備した。画像収集はクロールにより行った。これはプログラムを使って自動的にインターネット上の情報を収集するもので、クローラーというプログラムを用いて、ページ間のリンクや検索情報などを基に、効率よく情報を収集することができる。

クロールにより収集した画像から、学習用としては不適当なもの（イラスト、画像が不鮮明なもの、複数の魚体が写っているもの、他の情報が含まれているものなど）を除外し、各フグ種について15枚程度の画像をそろえた。本来、各種ともに少なくとも30~40枚程度の画像を収集することが望ましいと一般に言われているが、不適当なものを除き、特に模様が明確に認識でき、魚体全体が写っている画像を整理すると15枚程度にとどまった。画像の出典は結果としてはほとんどが「魚図鑑」(<https://zukan.com/fish/>)からの画像となった。当該サイトは、ユーザーが相互に魚種を判断したりすることが行われており、比較的魚種に詳しい複数の人が利用するサイトであり、種の特長などもある程度信用できるものと考えられる。今後、フグ雑種の画像などは少数しか収集できない可能性が高く、少ない画像数でのモデル構築の検討は意義があると考えられる。

また、画像の収集時には体模様が確認できる画像に絞ったため、魚の背後の背景による取捨選択を行うと画像数がより少なくなってしまうため、様々な背景の写真を使用している。背景としては、白、灰色などの無模様の背景や、砂場・岩場の背景が多い一方で、一部の画像はタイルや青色の背景もある。様々な背景の画像でも判定できることを目指すという点では、様々な背景の画像を用いることは妥当であると考えられるが、一方で特定の種に特定の背景が集中することは避ける必要がある。本報告の学習用画像では種によって背景に偏りが無いことを確認している。

それぞれの種の画像数をFig.1中に記に示す。判別モデルの学習では、これら画像のうち、各種ともに80%の枚数

の画像をランダムに選択して学習用に利用し、残りの20%を検証用画像として利用した。

フグ種判別モデルの構築

(1) モデルの構築環境

本研究でのフグ種判別プログラムの構築プラットフォームとしてはGoogle Colaboratoryを用いて、Python言語により構築した。Google Colaboratoryは、ブラウザ上で

Image	Scientific name Japanese name	Number of training images
 画像出典 [1]	<i>Takifugu poecilonotus</i> Komonfugu; コモンフグ	15
 画像出典 [2]	<i>Takifugu niphobles</i> Kusafugu; クサフグ	18
 画像出典 [3]	<i>Takifugu snyderi</i> Shousaifugu; ショウサイフグ	18
 画像出典 [4]	<i>Takifugu stictonotus</i> Gomafugu; ゴマフグ	13
 画像出典 [5]	<i>Takifugu chrysops</i> Akamefugu; アカメフグ	14
 画像出典 [6]	<i>Takifugu rubripes</i> Torafugu; トラフグ	15
 画像出典 [8]	<i>Takifugu xanthopterus</i> Shimafugu; シマフグ	15

Fig.1 Seven species of pufferfish subject to classification

Python を記述、実行できるサービスで、パソコン側のプログラミング環境の構築が不要、GPUが利用できるなどの利点がある。深層学習のフレームワークとしてはKerasを用いた。Kerasは、ニューラルネットワークライブラリの1つであり、プログラミング言語Pythonにより比較的短いソースコードで人工知能プログラムを実装することができる。

(2) 畳込みニューラルネットワーク (CNN)

CNNは、ディープラーニングをはじめとするAIが画像分析を行うための学習方法の1つである。CNNは、ニューラルネットワークの多層構造において、畳込み層とプーリング層という、2次元の隠れ層を用いるものである。畳込み層は、フィルター層との畳込み演算により、画像の1部分を取り出して画像の特徴を抽出する層であり、プーリング層は、畳込み層の抽出結果から、画像のカテゴリー分けにおいてあまり重要ではない情報を削ぎ落とし、解像度を下げる役割をしている。通常の深層学習においては、この畳込み層とプーリング層の組をいくつも重ねることで高い認識能力を獲得している。

(3) VGG16モデルの概要

VGG16モデル (Visual Geometry Group16)²³⁾は、オックスフォード大学のVGGチームが構築し、公開されているCNNモデルである。VGG16は、ImageNet データベースの100万枚を超える画像が学習済みであり、出力とし

て入力した画像を1000クラスに分類することができる。1000分類の中には、様々な動植物なども含まれており、「フグ (puffer)」という分類もあるが、フグ種の細かい分類 (トラフグ、ショウサイフグなど) までは無い。このモデルは、Fig. 2に示すように、畳込み層13層+プーリング層5層+全結合層3層=21層の畳込みニューラルネットワークの構成となっている。

(4) 転移学習

転移学習は、学習済みのモデルの一部を別のニューラルネットワークとして利用することである。本研究では、既に学習済みのVGG16モデルの18層の畳込み層及びプーリング層をそのまま利用し、後半の全結合層のみ独自の層 (1024ノードの中間層+出力層) に代替することで、転移学習を行った。VGG16モデルでの出力は1000クラス (1000ノード) に分類されるが、本研究では、フグ種を4種ないしは7種に限定したときの判別モデルを構築したので、最終出力が4ノード、7ノードの構成に変更して実装した。CNNを出た後の全結合層は以下の構成に変更した。

- 最後のプーリング層の出力: $7 \times 7 \times 512$ (=25,088ノード)
→ 1次元に展開
- 全結合層の入力: 25,088ノード
- 全結合層 (中間層): 1,024ノード
- ドロップアウト: 0.5
- 出力層: 4ノード、7ノード

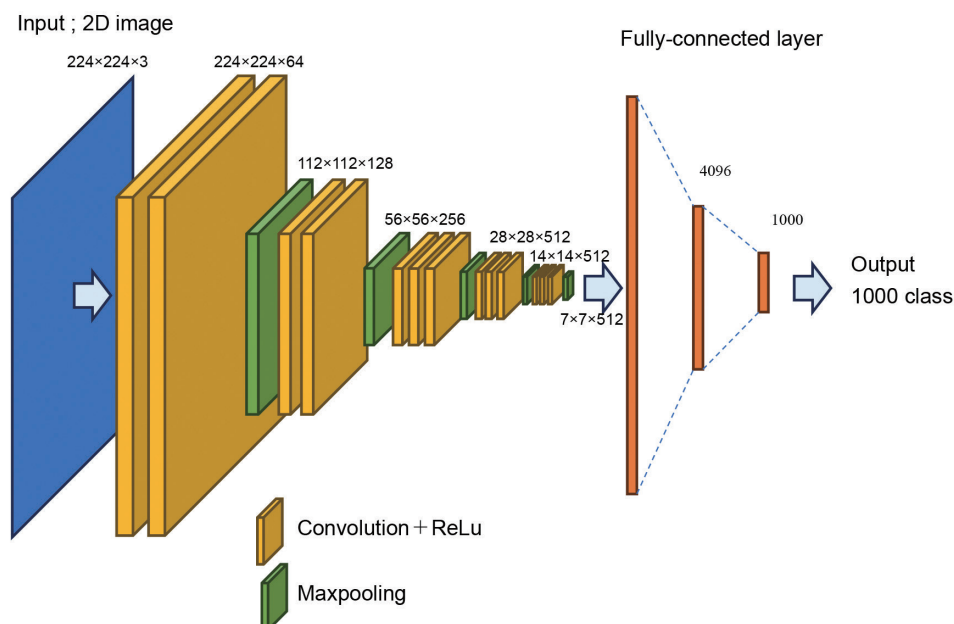


Fig.2 Configuration of learned model VGG16

また学習時の計算条件は以下のように設定した。

- ・学習繰り返し数 (epochs): 50
- ・バッチサイズ (batch_size): 4
- ・誤差関数: 交差エントロピー関数
(categorical_crossentropy)
- ・最適化手法: AdaGrad法

(5) Grad-CAMの概要

Grad-CAM²³⁾の方法は、画像の中の領域で、結果の出力に大きく影響を与える部分を探る方法である。入力された画像の領域の中で、出力結果に影響を与える領域ほど、その部分がもし変化すると出力結果も大きく変化するという考え方に基づいている。具体的には、CNNに画像を入力して計算を実行して出力を求めた後、CNNの最終的な出力から誤差逆伝播法で畳込み層までさかのぼり微分係数(勾配)を計算する。結果を大きく左右する領域ほど勾配の値が大きくなり、この勾配値を色分けして元画像に重ね合わせることで、AIが注目している部分を導出することができる。Fig.3にGrad-CAMを適用した画像と元の画像の例を示す。画像中で赤色の部分ほど、AIモデルが注目していることを示している。

計算ケースの設定

計算ケースとして以下の3つを設定した。

- ケース1: 体模様が似ている4種 (コモンフグ, クサフグ, ショウサイフグ, ゴマフグ)
- ケース2: 体模様が似ていない4種 (アカメフグ, トラフグ, シマフグ, クサフグ)
- ケース3: ケース1およびケース2を合わせた7種

結果

ケース1における判別結果

ケース1における繰り返し学習数が50までの精度と誤差の

推移をFig.4に示す。ここで精度は、判別した画像数に対する正答数の比率であり、誤差は正解となる出力とモデルからの出力を交差エントロピー関数で計算したもので、理想的な出力とモデルの実出力との差を示している。繰り返し数が50回の時点での精度と誤差は以下のとおりであった。

- ・学習用画像に対する誤差(loss): 7.1173e-04, 精度(accuracy): 1.00
- ・検証用画像に対する誤差(val_loss): 0.3322, 精度(val_accuracy): 0.846

これより学習用画像に対しては精度が100%となっているが、検証用画像においては84.6%の精度となっている(検証用画像の13枚中2枚が誤判定)。全ての画像の判定結果をまとめるとTable 1のようになる。

これらの結果よりコモンフグとゴマフグは全ての画像が正しく判定されているが、クサフグの1枚とショウサイフグの1枚が誤判定されている。Fig.5Aに検証用画像のうち、誤判定された画像を示す。図中の[]内数値は、判別モデルからの実際の出力値であり、数値は0から1の範囲となり「コモンフグの確率, クサフグの確率, ショウサイフグの確率, ゴマフグの確率」を示している。

またケース1におけるGrad-CAMの適用結果もFig.5にあわせて示す。Fig.5Aに誤判定された画像2枚を、Fig.5Bに正しく判定された画像のうち2枚についてGrad-CAMに適用した結果を示す。画像中で赤色の部分ほど、フグ判別モデルが注目していることを示している。これらと比較してみると、誤判定された画像は模様の中央部分には注目せず、頭部と尾ひれの部分に注目している。一方で正しく判定された画像のGrad-CAMの結果をみると魚体の胸鰭付近の模様にも着目していることがわかる。今回使用した3種のフグを人間の目で判定する場合も、魚体中央部の模様を比較して判定することから、判別モデルも正しい部分に注目していることがわかる。また誤判定された画像がなぜ中央部に着目しなかったのかについては、今後さらに検証をしていく必要がある。

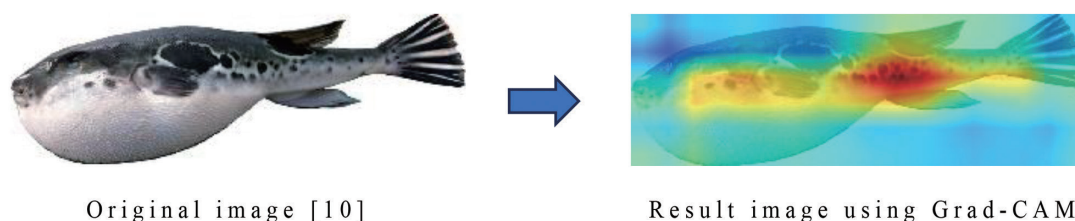


Fig.3 Examples of original image and Grad-CAM applied image

ケース2における判別結果

ケース2における繰り返し数が50回の時点での精度と誤差は以下のとおりであった。

- ・学習用画像に対する誤差(loss): 1.9759e-04, 精度(accuracy): 1.00
- ・検証用画像に対する誤差(val_loss): 0.3322, 精度(val_accuracy): 1.00

これより学習用画像および検証用画像のいずれも精度が100%となっている

ケース3における判別結果

ケース3において、繰り返し数が50回の時点での精度と誤差は以下のとおりであった。

- ・学習用画像に対する誤差(loss): 3.8922e-04, 精度(accuracy): 1.00
- ・検証用画像に対する誤差(val_loss): 0.8599, 精度(val_accuracy): 0.727

これより学習用画像に対しては精度が100%となっているが、検証用画像においては72.7%の精度となっている（検

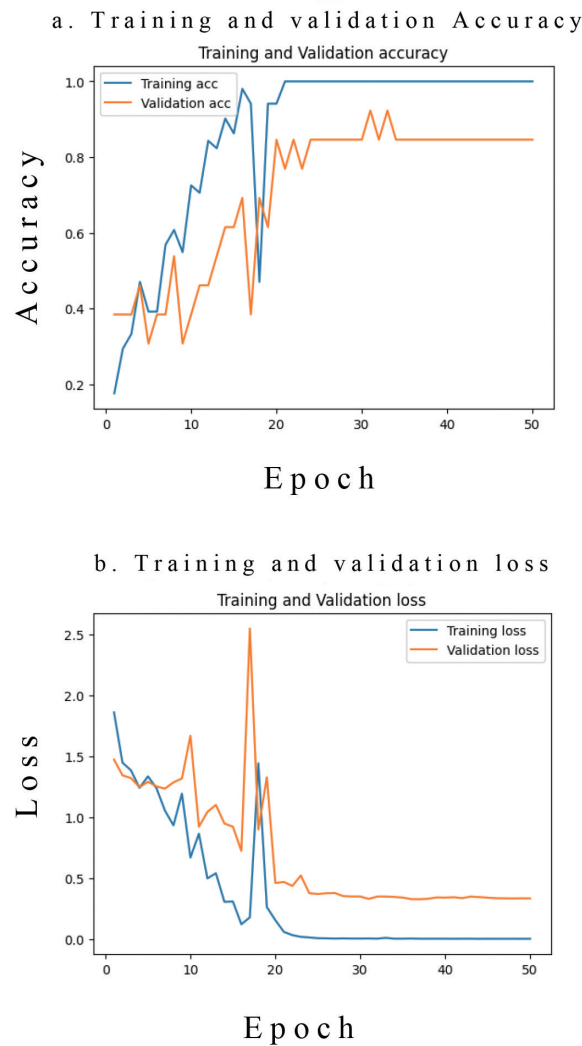


Fig.4 Example of Case 1 Accuracy and Losses in the Learning Process

Table 1 Classification results for all images (Case 1)

		Classified species			
		komonfugu	kusafugu	shousaifugu	gomafugu
Image for classification	komonfugu	15			
	kusafugu		17		1
	shousaifugu		1	17	
	gomafugu				13

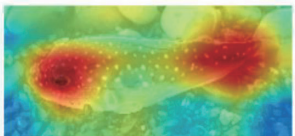
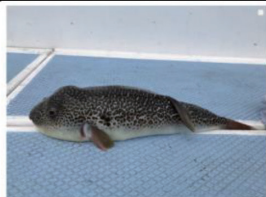
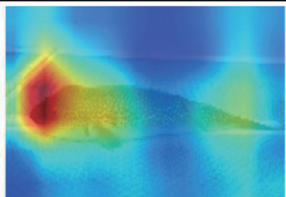
証用画像の22枚中5枚が誤判定)。全ての画像の判定結果をまとめるとTable 2のようになる。これらの結果よりコモンフグ、アカメフグ、シマフグは全ての画像が正しく判定されているが、クサフグの3枚とショウサイフグの2枚、ゴマフグの1枚が誤判定されていることがわかる。

ケース3において誤判定された画像のGrad-CAMの適用結果例をFig.6に、正しく判定された画像の適用例をFig.7に示す。誤判定された画像において判別モデルが注目している傾向をみると、魚体中央の模様よりも、頭部や尾びれの

のほうを注目している傾向があることと、周辺が砂地などの場合は周辺の画像にも注目していることがわかる。このように、誤判定された画像においては、魚体中央部に注目していない、周囲の砂地や砂利なども模様としてみているなどの傾向があることが読み取れる。

学習画像の選択による精度への影響

準備した画像のうち、学習用として利用する画像の選択はモデルにより乱数で自動的に分割される。このため計算

A. Misidentified image			
	Original image	Result image using Grad-CAM	Image source
Kusafugu			[11]
	Output [0.00275458, 0.21266358, 0.6615879 , 0.12299393]* → Kusafugu was judged as Shousaifugu		
Shousai-fugu			[12]
	Output [0.02379026 0.42531833 0.30285785 0.24803355]* → Shousaifugu was judged as Kusafugu		
* Output [probability of Komonfugu, probability of Kusafugu, probability of Shousaifugu, probability of Gomafugu]			

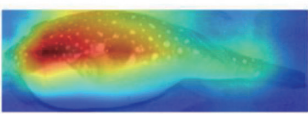
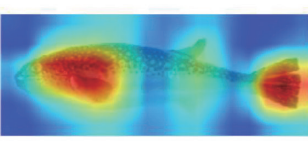
B. Correctly identified image			
	Original image	Result image using Grad-CAM	Image source
Kusafugu			[13]
Shousai-fugu			[14]

Fig.5 Misidentified images of Case 1 and results of Grad-CAM application

Table 2 Classification results for all images (Case 3)

		Classified species						
		komonfugu	kusafugu	shousaifugu	gomafugu	akamefugu	torafugu	shimafugu
Image for classification	komonfugu	15						
	kusafugu	1	15	1	1			
	shousaifugu	2		16				
	gomafugu		1		12			
	akamefugu					14		
	torafugu						15	
	shimafugu							15

を実施するごとに構築されるモデルの精度が変動するが、このような画像の選択の仕方によるモデルへの影響を検討するために各ケースで5回計算を実施して、各種の判別率をまとめたものをFig.8に示す。

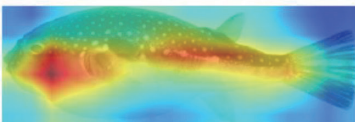
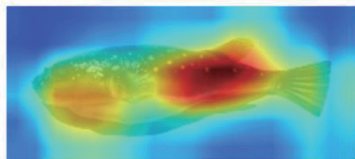
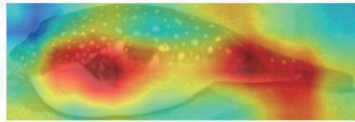
模様が似ているケース1の場合、特にクサフグとショウサイフグ間での誤判定がいずれの計算でも起きていることが分かる。ケース2は模様が大きく異なるケースである。判別精度はいずれの計算でも高い傾向があるが、誤判定される画像がでる場合もある。7種の判定を行ったケース3

においては、模様が似ているクサフグ、ショウサイフグ、ゴマフグ間での誤判定が多いことがわかる。

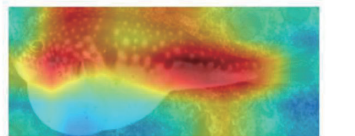
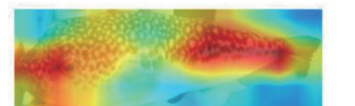
考 察

結果より、フグ種を模様が似ているケース1の4種類の判定と、模様が大きく相違するケース2の4種の判定では、ケース1のほうが判別精度が低い傾向にあることがわかる。ケース3の7種判定においては、4種判別に対して、判

A. Kusafugu

Original image	Result image using Grad-CAM	Image source
		[2]
		[15]
		[16]

B. Shousaifugu

Original image	Result image using Grad-CAM	Image source
		[17]
		[18]

C. Gomafugu

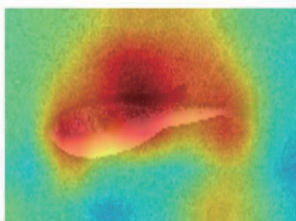
Original image	Result image using Grad-CAM	Image source
		[19]

Fig.6 Results of Grad-CAM application (Misidentified images, Case 3)

定精度が低下することがわかる。模様が似ている種の判別が判別モデルにおいても難しいことがわかる。

また、いずれの種類数でも、学習用画像の精度は100%に達するものの、検証用画像は100%にはならず数枚の誤判定がでる。この原因として考えられることは、学習用画像が少ないことによる過学習を起こしていることが考えられる。今後さらに判別種類数を多くする場合でも判定精度を向上することや過学習の有無の判断のためには、さらに学習画像数を増やすことなどが必要であると考えられる。一方で良質な画像はすぐには増やすことができないことも現実である。

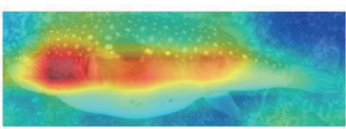
本モデルにおいては各種とも15枚程度の画像を用いているが、画像を増やす効果と同等のData Augmentation（データ拡張）と呼ばれる手法がある。この方法は学習用の画像データに対して、様々な変換を施す事で画像を水増しする手法である。Kerasでは、ImageDataGeneratorクラスが用意されており、これを用いる事で簡単にデータ拡張を実施する事ができる。本研究では、「左右平行移動」、「上下平行移動」、「拡大・縮小」、「画像のせん断」、「明るさの調節」などのパラメータについて予備的な検討を行ったが、精度の大きな向上は得られなかった。これは準備した画像

は、フグが一定の向きで、ほぼ同じ大きさにそろえられていたため、画像の変換の効果が得られなかったと考えられる。

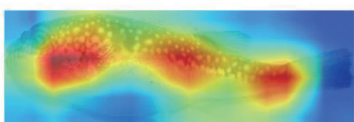
また、準備した画像のうち、学習用として利用する画像の選択により、構築されるモデルの精度が変動する点について予備的に検討したが、このような画像の選択のモデルへの影響を検討するためには、交差検証などの手法を用いることが考えられる。交差検証とは、学習用のデータを複数のセットに分割して、そのうち一つを検証用として用い、学習データと検証データを交互に入れ替えることでモデルの性能を検証するものである。本研究では画像の枚数が少ないため、交差検証を利用するには無理があると考え、今後、画像枚数を増やしたうえで、交差検証の適用を再度検討していく必要がある。

さらに、判別モデルが画像のどの部分に注目しているかを表示することができるGrad-CAM手法を用いて検討を行った。Grad-CAMによる結果をみると、体模様の一部しか注目していないケースや、周囲の砂利模様も注目してしまうケースなどがみられ、CNNの誤判定の要因を探ることができる可能性を示すことができたと考えられる。またFig.5Bの下画像では尾ひれがGrad-CAMによって強調

A. Kusafugu

Original image	Result image using Grad-CAM	Image source
		[20]

B. Shousaifugu

Original image	Result image using Grad-CAM	Image source
		[21]

C. Gomafugu

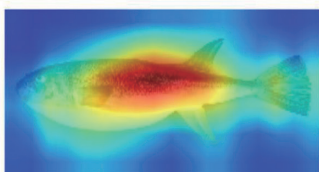
Original image	Result image using Grad-CAM	Image source
		[22]

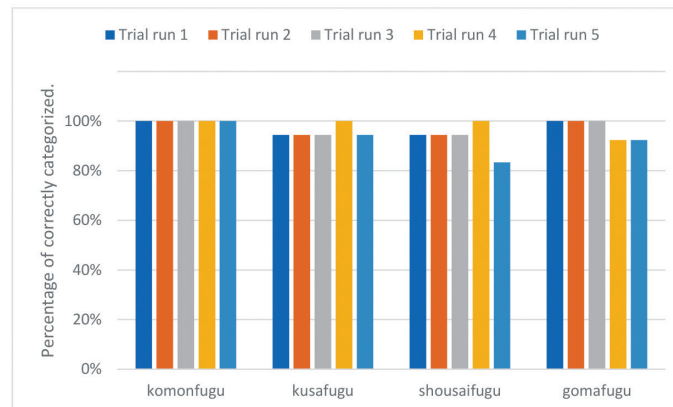
Fig.7 Results of Grad-CAM application (Correctly identified image)

されている。本研究で用いた学習用画像においては尾ひれが開いているものが多いが、各種に一部閉じている画像も含まれている。フグの目利きがフグを判定する場合は、一部の種は尾びれの色なども参照するため、今後は尾ひれも判断要素の一つとして加味されるモデルを検討していくことが必要であると考えられる。

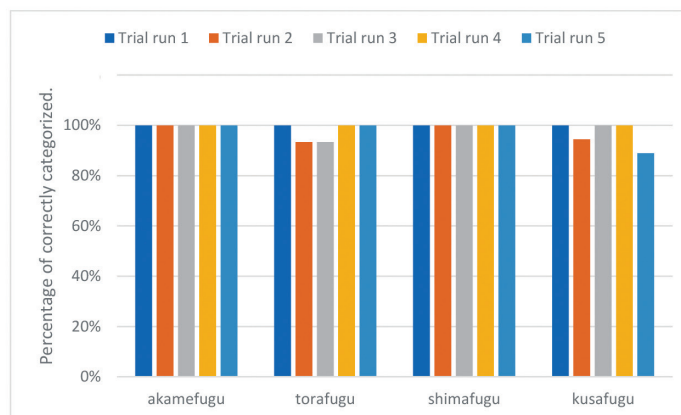
結 論

本研究では既に学習済みのCNNモデルを基礎として、それに転移学習を適用したモデルを構築していくつかのフグ種の判別を行った。今後の「雑種も含めたフグ種判別システム」の構築のための前段階として、少数の画像での学

A. Case 1



B. Case 2



C. Case 3



Fig.8 Variation in calculation accuracy due to image segmentation

習の可能性の検討を「転移学習」モデルで確認し、フグの体模様をみて判定しているかの確認をGrad-CAM手法を用いて検討を行い、判別モデルの構築の可能性が確認できたと考えられる。

今後はフグ漁業者などの協力を求めて、雑種を含めたより多くの正確な画像データを収集していくとともに、モデルの改良を進めて判定率の向上に努めていく必要がある。またGrad-CAMを用いた結果では、AIも人と同じような体模様を見て画像判別を行っていると考えられる。一方で模様に注目しているにもかかわらず誤判定となっている事例もあり、このような場合は、体模様の中でより詳細な注目点をGrad-CAMでは明確にできないため、他のXAI手法と併せて、モデルの判定根拠をより多角的に検討していく必要がある。漁業現場でのフグの「目利き」が注目する点の情報を収集するとともに、XAI手法との知見を合わせていくことで、根拠を明確にしたうえでの判別モデルの構築が可能になり、他の魚種判別にも応用できる基盤ができると考えられる。そして精度を向上させ最終的には実際の現場にこの技術を投入することによる作業効率の向上を図っていきたいと考えている。

また本報告では学習用画像中の魚の背景として様々なものが含まれている状況であり、モデル学習時に画像の背景の影響が完全に排除できているとは言いきれない状況である。様々な背景の画像を大量に学習していれば背景の影響は低減されることが考えられるが、本報告のように学習画像が少ない状況では、背景に影響されている可能性もあり、背景の影響の程度を確認していく必要がある。学習時における背景の影響を除外するには画像から魚の輪郭を判定し、背景を除去するプロセスを検討していくのもひとつの方向であると考えられる。背景を除外できれば、背景の影響を無くすることが可能となり、判定精度の向上につながる可能性もある。一方で、モデルの実用化を考えた場合、魚を抽出するプロセスで様々な画像でも精度高く抽出できるのか、特に尾びれやヒレの部分なども明確に抽出できるのかなどの検討が必要であると考えられる。

謝 辞

本論文の執筆にあたり、フグに関連する知見に関して貴重なコメントをいただいた（国研）水産研究・教育機構水産大学校 酒井治己名誉教授、高橋洋教授、椎木友朗講師に感謝申し上げます。

引用文献

- 1) 厚生労働省編: 三訂日本近海産フグ類の鑑別と毒性. 中央法規出版, 東京, (2013)
- 2) Santini F, Nguyen MTT, Sorenson L, Waltzek TB, Lynch Alfaro JW, Eastman JM, Alfaro ME : Do habitat shifts drive diversification in teleost fishes? An example from the pufferfishes(Tetraodontidae). *J Evol Biol*, **26**:1003–1018 (2013)
- 3) Takahashi H, Toyoda A, Yamazaki T, Narita S, Mashiko T, and Yamazaki Y : Asymmetric hybridization and introgression between sibling species of the pufferfish Takifugu that have undergone explosive speciation. *Marine biology*, **164**(4), 90 (2017).
- 4) 松浦啓一著：日本産フグ類図鑑，東海大学出版部，平塚市，(2017)
- 5) Selvaraju R R, Cogswell M, Das A, Vedantam R, Parikh D, Batra D : Grad-CAM: Visual Explanations from Deep Networks via Gradient-based Localization. <https://arxiv.org/abs/1610.02391>
- 6) Wäldchen J., Mäder P. : Machine learning for image based species identification. *Methods in Ecology and Evolution*, Wiley Online Library (2018)
- 7) Qin H, Li X, Liang J, Peng Y, Zhang C : DeepFish: Accurate under water live fish recognition with a deep architecture. *Neurocomputing, Volume 187*, Pages 49-58, 26 (2016)
- 8) Lee D J, Schoenberger RB, Shiozawa D, Zhan X Xu P : Contour matching for a fish recognition and migration-monitoring system, in: Optics East. *International Society for Optics and Photonics*, pp.37–48 (2004)
- 9) Strachan N, Nesvadba P, Allen AR : Fish species recognition by shape analysis of images. *Pattern Recognit.***23**(5), 539–544 (1990)
- 10) Strachan N : Recognition of fish species by colour and shape. *Image Vis. Comput.* **11**(1)2–10 (1993)
- 11) White D, Svellingen C, Strachan N : Automated measurement of species and length of fish by computer vision. *Fish.Res.***80**(2) 203–210 (2006)
- 12) 柴田利樹: 齊藤剛史, 魚画像認識に有効な特徴量の検討. 情報処理学会研究報告2015-CVIM-197, 26, 1 – 6 (2015)
- 13) Alsmadi MK, Omar KB, Noah SA, and Almarashdeh

- I : Fish Recognition Based on Robust Features Extraction from Size and Shape Measurements Using Neural Network. *Journal of Computer Science* **6** (10): 1059-1065 (2010)
- 14) Larsen R, Olafsdottir H, Ersbøll B K : Shape and texture based classification of fish species. *Image Anal.***745–749** (2009)
- 15) Spampinato C, Giordano D, DiSalvo R, Chen-Burger Y-H J, Fisher R B, Nadarajan G : Automatic fish classification for underwater species behavior understanding. Proceedings of the First ACM International Workshop on Analysis and Retrieval of Tracked Events and Motion in Imagery Streams, ACM, Firenze, Italy, pp.45–50(2010)
- 16) Huang P X, Boom B J, Fisher R B : Under water live fish recognition using a balance-guaranteed optimized tree. Computer Vision-ACCV2012, Springer, Daejeon, Korea, pp.422–433, (2013)
- 17) Huang P X, Boom B J, Fisher R B : Gmm improves the reject option in hierarchical classification for fish recognition. 2014 IEEE W inter Conference on Applications of Computer Vision (WACV) , IEEE, Steamboat Springs CO. USA, pp.371–376(2014)
- 18) 仲松幸信, 姜東植, 宮城隼夫: デジタル画像を用いた魚種の識別. 情報処理学会研究報告132, 49 – 52 (2006-AVM-055)
- 19) 平岡透, 矢野啓司, 瀧山龍三: 輪郭線及びテクスチャ情報に基づく画像認識法: 魚画像認識への応用, 電子情報通信学会技術研究報告. PRMU, パターン認識・メディア理解 **96(492)**, 55-62 (1997-01-24)
- 20) Yakubu A, Okunsebor SA : Morphometric differentiation of two Nigerian fish species (*Oreochromis niloticus* and *Lates niloticus*) using principal components and discriminant analysis. *International Journal of Morphology* **29**:1429-1434. (2011)
- 21) Rauf HT, Lali MIU, Zahoor S, Shah SZH, Rehman AU, Bukhari SAC : Visual features based automated identification of fish species using deep convolutional neural networks. *Computers and electronics in agriculture*. **2019**:167:105075.
- 22) Villon S, Mouillot D, Chaumont M, Darling ES, Subsol G, Claverie T, *et al.* : A deep learning method for accurate and fast identification of coral reef fishes in underwater images. *Ecological informatics*. **48**:238-44(2018)
- 23) Simonyan K, Zisserman A : Very Deep Convolutional Networks for Large-Scale Image Recognition. arXiv:1409.1556, 2015

画像の出典

- [1] <https://zukan.com/fish/internal79>
- [2] <https://zukan.com/fish/leaf19561>
- [3] <https://zukan.com/fish/leaf22002>
- [4] <https://zukan.com/fish/internal201>
- [5] <https://zukan.com/fish/leaf4853>
- [6] <https://zukan.com/fish/leaf11837>
- [8] <https://zukan.com/fish/internal234>
- [9] <http://www.nipic.com/show/1/28/00cc0acd85cd5b50.html>
- [10] <http://www.ko-suisan.co.jp/column/torafugu.html>
- [11] <https://zukan.com/fish/leaf66013>
- [12] <https://fishai.jp/597>
- [13] <https://zukan.com/fish/leaf59994>
- [14] <http://www.miegyoren.or.jp/fisheries-in-mie/uopedia/shosaifugu/>
- [15] <https://zukan.com/fish/leaf81851>
- [16] <https://zukan.com/fish/leaf59994>
- [17] <https://zukan.com/fish/leaf42345>
- [18] <https://zukan.com/fish/leaf51432>
- [19] <https://blog.goo.ne.jp/ryuto50miyu55/e/cf980ea3e82c1c8196181eac66747ace>
- [20] <https://zukan.com/fish/leaf48090>
- [21] <https://zukan.com/fish/leaf3862>
- [22] <https://www.zukan-bouz.com/syu/%E3%82%B4%E3%83%9E%E3%83%95%E3%82%B0>