

マルチラベル分類 ABN を用いた ギフチョウ斑紋画像からの食草分類に関する研究

平川 翼[†] 荒井 嵩貴[†] 山下 隆義[†] 藤吉 弘亘[†] 大場 裕一[†]

福井 弘道[†] 矢後 勝也^{††}

[†] 中部大学 〒487-8501 愛知県春日井市松本町 1200

^{††} 東京大学 総合研究博物館 〒113-0033 東京都文京区本郷 7 丁目 3-1

E-mail: [†]{hirakawa,arataka}@mprg.cs.chubu.ac.jp,{takayoshi,fujiyoshi,yoba,fukui}@isc.chubu.ac.jp,

^{††}myago@um.u-tokyo.ac.jp

あらまし チョウ類はその派手な色彩から人目に付きやすく、身近な生物であり、愛好家も多い分類群である。また、標本の収集量が豊富で標本間の比較検討が容易なことから、斑紋の地域変異が認められ、これらは地史や地形、気候、食草などの様々な要因が関係している。そこで本研究では、チョウ斑紋の地域変異と食草の分布に着目し、チョウ斑紋画像からの食草分類を行うことで、その関係性を明らかにすることを目的とする。具体的には、斑紋の地理的変異が知られており、分布解明度が非常に高いチョウ類の一種であるギフチョウを対象とする。デジタル標本および採集地に関するメタデータから、ギフチョウ斑紋画像分類データセットを作成する。本データセットを用いてマルチラベル分類 Attention Branch Network を学習することで、斑紋画像から高精度な食草分類が可能となると同時に、アテンションマップの解析により、専門家の知見と同様の判断根拠が獲得できることを示す。

キーワード Attention Branch Network, マルチラベル分類, 視覚的説明, ギフチョウ, 地域変異

1. 緒 言

チョウ類は人目に付きやすく、最も身近な生物の一つである。また、標本の収集量が豊富なことから、生息分布の解明度が極めて高い生物群であり [1], 標本間の比較検討が容易なことから、斑紋の地域変異（地理的変異）が知られている種が多い。この地域変異には、地史や地形、環境、食草との関連など、様々な要因が関係していると考えられているため、チョウ類標本データと様々なデータを関連づけて解析することで、チョウ類に纏わる多様な問題・要因の解明が期待できる。本研究では、特に愛好家からの人気が高いことで分布解明度が高いギフチョウ (*Luehdorfia japonica*) に着目する。ギフチョウは翅の斑紋の地域変異がよく知られている。また、図 1 に示すように、ギフチョウは幼虫の食草で種分化が著しいカンアオイ属 [1], [2] の分布にギフチョウの地域変異が対応すると考えられている。

そこで本研究では、ギフチョウの標本画像と各生息地域の食草分布の情報をもとに、ギフチョウ斑紋の地域変異と食草との関係性を明らかにすることを目的とする。具体的には、中部大学の保有する「藤岡知夫日本産チョウ類コレクション」と食草分布の情報からギフチョウ斑紋画像データセットを作成し、チョウ斑紋画像からの食草分類とその判断根拠の可視化を通じた斑紋の地域変異の解析を行う¹。データセットの作成では、デジ

タル標本画像を前処理し、画像サンプルを収集する。それと同時に、各画像サンプルに対して、採集地情報から幼虫期に利用される食草ラベルを付与する。作成したデータセットに対して、視覚的説明と画像分類タスクの精度向上を目的とする Attention Branch Network (ABN) [5] を用いて学習及び分類を行う。

作成したデータセットのサンプル数は現存する標本の種類や採集場所に大きく依存する。各採集地に生息するギフチョウの数や収集された個体数は大きく異なるため、潜在的にクラス不均衡を抱えたデータセットとなっている。クラス不均衡を抱えたデータセットでは、サンプル数の少ないクラスにおいて認識精度が低下することが知られている。そこで本研究では、クラス不均衡の軽減のために、クラス不均衡を考慮した誤差関数やデータサンプリングを導入する。これにより、分類精度の改善だけでなく、獲得される視覚的説明の変化を確認し、斑紋と食草の関係性を裏付ける適切な注視領域の獲得を目指す。

2. 関連研究

2.1 コンピュータビジョン技術の生物学への応用

コンピュータビジョン技術を応用した生物学的解析はいくつか行われている [6]~[11]。Cuthill ら [6] は、Triplet network [12] と Categorical cross-entropy loss により獲得した埋め込み空間でのサンプル間距離を活用し、チョウ類のミュラー型擬態 (Müllerian mimicry) を定量的に解析している。

Fan ら [11] は、中国雲南省に生息するアゲハチョウ種 80 種のデータセットの構築と分類を行っており、Lin ら [9], [10] は、

(注 1)：本稿では、文献 [3], [4] にて実施した画像分類の予備実験の結果をもとに、データセットの再整理や新たに実施した実験結果について報告する。

チョウの亜科 (subfamily), 属 (genus), 亜種 (subspecies) を同定するための深層学習に基づいた画像分類モデルを提案している. いずれの研究も詳細画像識別によるチョウ類の生物学的分類の同定である. 一方, 本研究ではギフチョウという単一の種における斑紋の地域変異を扱うことから, 非常に詳細な画像識別タスクを扱う. また, 単なる分類にとどまらず, 識別モデルの推論過程で獲得される視覚的説明 (アテンションマップ) を用いて, 食草と斑紋の関係性を明らかにする.

2.2 マルチラベル分類

本研究ではマルチラベル分類タスクを扱う. マルチラベル分類では, クラス不均衡 (クラスインバランス) を考慮した学習方法が提案されており [13]~[15], 主なアプローチとして, 誤差関数の工夫 [16]~[18] や学習中のデータのサンプリング方法の工夫 [19]~[22] が存在する. 本研究で扱うデータセットは, 潜在的にクラス不均衡を抱えているため, クラス不均衡を考慮した学習を行い, 分類精度の改善と解析を行う.

2.3 視覚的説明

視覚的説明は機械学習モデルの推論結果に対する判断根拠を提示する手法である. 主なアプローチとして, GradCAM [23] をはじめとする Post-hoc な方法や, モデル内部にアテンションマップを生成する機構を埋め込んだモデル [5], [18], [24] が存在する. 本研究では視覚的説明をアテンション機構により推論結果に反映させる Attention Branch Network (ABN) [5] を活用し, 各クラスラベル (食草) を分類する際に注視した領域を示しつつ分類と解析を行う.

3. データセット

3.1 藤岡知夫日本産チョウ類コレクション

藤岡知夫日本産チョウ類コレクション (藤岡コレクション) とは, 日本産全種及び関連する近隣諸国標本の約 29 万頭, 標本箱にして約 1,750 箱を有する世界有数の貴重なコレクションであり, これを有効利用するために, データベース整備やヴァーチャルミュージアムの作成が行われている [26], [27]. 本コレクションデータベースおよび植生や植物相等の GIS 関連データを活用することにより, チョウ斑紋の地域変容の解明が可能になると考えられる. 本研究では, 藤岡コレクションを活用し, ギフチョウ斑紋画像からの定量的な食草分類を行う. 以下では, 作成したギフチョウのデータセットの詳細について述べる.

3.2 ギフチョウデータセット

3.2.1 デジタル標本の撮影と前処理

図 2(a) に示すように, 標本箱から 1 頭ずつ取り出し, 標本番号や採集値ラベル, 大きさを示す目盛, 補正のためのカラーチャートとともに撮影を行う. この際, 複数枚の被写界深度で撮影を行ったのちに, 深度合成を行うことで, 全ての箇所焦点が合った画像を作成する. 撮影したギフチョウのデジタル標本の一例を図 2(b) に示す. デジタル標本の画像サイズは約 6,500×4,000 [pixels] であり, 非常に高精細な画像である.

これらのデジタル標本は, 前述の通り, データベース整備やヴァーチャルミュージアムの作成を目的として撮影されており, カラーチャートなどの画像分類に不要な箇所が存在する.

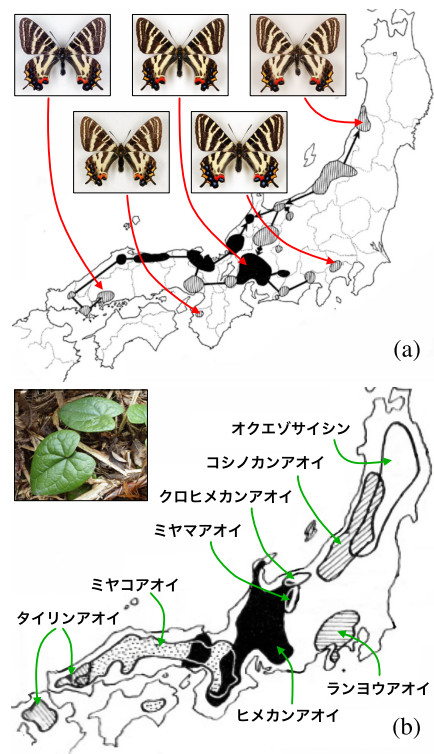


図 1 (a) ギフチョウ翅型・斑紋の地域変異から見た個体群の概念図. 個体群間の線は類似性を示しており, 黒で示した個体群ほど斑紋の黒帯が発達している. (b) ギフチョウが利用する主要なカンアオイ属の分布概念図. それぞれ文献 [25] より改変.

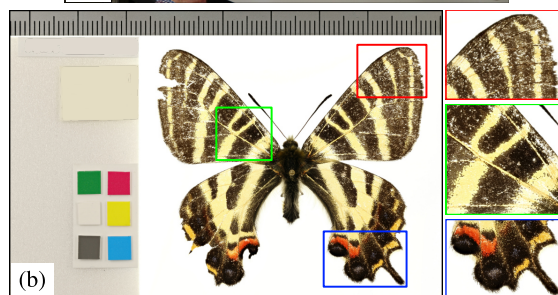


図 2 藤岡コレクションデジタル標本の撮影. (a) 撮影の様子および (b) 撮影したデジタル画像データの一部.



図 3 前処理後のギフチョウ画像の例.

そのため, これらを除去することで, 画像サンプルを収集した. サンプル画像を図 3 に示す. 切り取った画像パッチは全て 2,048×2,048 [pixels] にリサイズし, 画像サイズを統一する. 作成したギフチョウ画像の総枚数は 4,434 枚であり, そのうちオ

表 1 ギフチョウデータセットのラベル分布

ラベル (カンアオイ)	実験に使用	サンプル数	
		オス	メス
カギガタアオイ	✓	235	45
カントウカンアオイ	✓	124	29
クロヒメカンアオイ	✓	55	10
コシノカンアオイ	✓	341	99
サンインカンアオイ	✓	54	13
スズカカンアオイ	✓	91	21
タイリンアオイ	✓	92	24
トウゴクサイシン	✓	42	10
ヒメカンアオイ	✓	2,075	517
フタバアオイ		1	0
ミチノクサイシン		0	1
ミヤコアオイ	✓	340	148
ミヤマアオイ		3	1
ユキグニカンアオイ	✓	145	30
ランヨウアオイ	✓	359	74

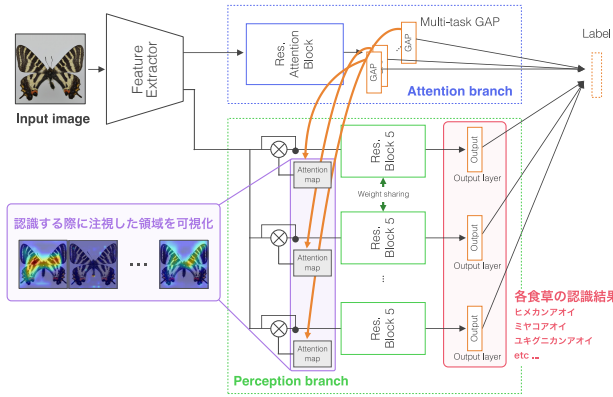


図 4 マルチラベル分類 ABN のネットワーク構造

スが 3,510 枚、メスが 924 枚である。

3.2.2 食草ラベルデータの作成

次に各ギフチョウサンプルに対する食草ラベルを付与する。各ギフチョウ標本には、「採集地、採集年月日、採集者」などのメタデータが付与されている。そこで、採集地の情報と日本の植生情報から食草ラベルを割り当てた [1], [2]。

作成したギフチョウデータセットのラベル内訳を表 1 に示す。1 つのサンプルに複数の食草ラベルが付与される場合があるため、ラベルの合計数がサンプル枚数と一致しないことに注意されたい。表 1 より、現時点では十分数の存在しないラベルが複数ある。そのため、使用するラベルを一部に限定し実験を行う。

4. 手 法

4.1 モ デ ル

図 4 にマルチラベル分類 ABN のネットワーク構造を示す。入力された画像は Feature extractor によりラベル共通の特徴を抽出し、抽出した特徴を Attention branch へと入力する。Attention branch では、各ラベルに応じた Global average pooling (GAP) [28] を行い、それぞれのラベル分類の結果を獲得する。これと同時に、GAP 直前の畳み込み層から取得される特徴マップであ

るアテンションマップを Perception branch へと入力し、Feature extractor から出力される特徴マップに対する重み付け（アテンション機構）を適用した後に、最終的な分類結果を出力する。これにより、分類する際に注視した領域を可視化しつつ、その領域の特徴を強調した特徴を用いて分類を行うことが可能となる。

4.2 クラス不均衡を考慮した学習

本研究で使用するデータセットは、クラス不均衡を抱えている。そのため、ナイーブな学習では、サンプル数の多いラベルを中心に学習が進み、分類精度に偏りが生じることが考えられる。さらに、本研究では、高い分類精度を達成することだけが目的ではなく、モデルが分類時に注視した領域の可視化を通じた斑紋の解析を目的とするため、適切な分類と妥当な注視領域の獲得が必要となる。本研究では、クラス不均衡を考慮する学習方法を導入することで、精度と獲得されるアテンションマップの変化について実験的に考察を行う。具体的には、クラス不均衡を考慮した (i) 誤差関数および (ii) データサンプリングを学習に導入する。

4.2.1 誤 差 関 数

ABN の誤差関数 L_{abn} は次のように定義される。

$$L_{abn} = L_{att} + L_{per} \quad (1)$$

マルチラベル分類 ABN では、Attention branch 及び Perception branch それぞれの出力と正解ラベルの誤差 (L_{att}, L_{per}) を計算し、学習を行う。

従来のマルチラベル分類 ABN では、(L_{att}, L_{per}) それぞれに Binary cross-entropy loss (BCE) を用いているが、サンプル数の不十分なクラスラベルについて十分な学習が行われず、認識精度の低下を招く恐れがある。そこで本研究では、クラス不均衡を考慮した誤差関数として以下の誤差関数を使用する。

Weighted focal loss (WFL) [18] WFL は Focal Loss [29] にラベルの事前分布をもとに重みづけした誤差関数である。ここで、あるクラス c に対するロジットを x 、正解ラベルを y^c とする。また、シグモイド関数 $\sigma(\cdot)$ と x を用いて、確率 $p = \sigma(x)$ を定義するとき、WFL は次のように定義される。

$$L_{wfl} = - \sum_c w_c (y^c (1-p)^\gamma \log(p) + (1-y^c) p^\gamma \log(1-p)) \quad (2)$$

ここで、 $w_c = e^{-a_c}$ はクラス c に対する重みであり、クラス c の事前分布 a_c から算出される。

Asymmetric loss (ASL) [16] ASL はポジティブクラスとネガティブクラスのバランスを考慮した誤差関数であり、Focal loss をもとに次のように定義される。

$$L_{asl} = - \sum_c (y^c (1-p)^{\gamma_+} \log(p) + (1-y^c) p_m^{\gamma_-} \log(1-p_m)) \quad (3)$$

ここで、 $p_m = \max(p-m, 0)$ は分類の簡単なネガティブクラスの誤差を無視するための Shifted probability であり、 m は誤差を

無視する確率のマージンを表すパラメータである。パラメータ (γ_+ , γ_-) のバランスを調整することで、ポジティブとネガティブクラスの間のバランスを考慮した誤差計算を行う。

4.2.2 データサンプリング

不均衡なデータを用いて学習と分類を行う際のバランスを考慮するアプローチとして、希少なクラスのサンプルを多くサンプリングする「オーバーサンプリング」やメジャークラスのサンプルを少なくサンプリングする「アンダーサンプリング」が存在する。本研究では、前述の誤差関数に加えて、データのサンプリング方法を調整することで識別精度及び獲得されるアテンションマップの違いを考慮する。具体的には、表 1 に示したオスのラベル数を参考に、著しくサンプル数の多いヒメカンアオイのサンプルをアンダーサンプリングし、ユキグニカンアオイなどのサンプル数の少ないデータをオーバーサンプリングすることで、クラスバランスを考慮した学習を行う。

5. 実験

5.1 実験条件

データセットとして、3.2 節にて述べたギフチョウデータセットを活用する。画像サンプルは、448×448 [pixels] にリサイズしてネットワークへ入力する。このデータセットの内、オスの画像データ 3,510 枚のみを使用する。その中でも学習及び評価を行うことが可能なサンプル数が確保されている食草ラベル 12 種類を持つ画像データ 3,503 枚を実験に使用する（表 1 の「実験に使用」の列を参照）。3,503 枚をランダムに分割し、2,799 枚を学習、704 枚を評価に用いる。さらに、学習サンプルについては、次の 2 種類をそれぞれ用いて実験を行う。

Unbalanced 上記のランダムに分割した 2,799 枚をそのまま用いる学習用セット。クラス不均衡を含んだデータセットとして実験を行う。

Balanced 4.2.2 節にて述べたように、属性ラベル数に応じて学習サンプルの属性ラベル数の割合が可能な限り等しくなるようにデータサンプリングを行なった学習用セット。

各学習サンプルの内訳は、実験結果の表 3 に示す。

ネットワークモデルとして、ImageNet [30] で事前学習した ResNet18 [31] をバックボーンとするマルチラベル分類 ABN を使用する。ミニバッチサイズは 32、学習回数は 1,000 [epochs] とする。学習時には、左右反転や Random crop、コントラスト変換などの弱いデータ拡張を適用する。全ての学習において、最適化手法に Momentum SGD（学習率 = 0.01, Momentum = 0.9）を使用し、500 および 750 [epochs] 目で学習率を 10 分の 1 に変更する。

評価指標として、F1-score および mean Average Precision (mAP) を採用する。また、獲得したアテンションを可視化することで、定性的な評価を行い、チョウ類の専門家の視点から注視領域と斑紋との関連性についても議論を行う。

5.2 分類精度の評価

表 2 に属性平均の F1-score と mAP を示す。クラスバランスを未調整のデータセット (Unbalanced) と調整したデータセット (Balanced) を比較すると、Balanced を用いた場合に F1-score が

表 2 分類性能

	Training set: unbalanced		Training set: balanced	
	F1-score [%]	mAP [%]	F1-score [%]	mAP [%]
BCE	74.49	92.26	88.68	92.81
WFL [18]	74.61	93.78	88.79	92.74
ASL [16]	81.74	89.23	80.14	92.91

向上している。特に、BCE と WFL を用いた場合に F1-score が大きく向上しており、ポジティブとネガティブクラスのバランスを考慮しない誤差関数では、サンプルバランス調整が精度向上に大きく寄与している。

また、誤差関数の違いに着目すると、Unbalanced を用いた場合、ASL は他の誤差関数と比較して、高い F1-score を獲得している。一方、Balanced なデータセットでは、BCE や WFL よりも F1-score が低下している。このことから、サンプルバランスが著しく悪いデータセットで学習を行う場合には、ASL が有効であることが確認できる。

次に、表 3 に、食草ラベル毎の F1-score を示す。また、表には各食草ラベルの学習用サンプル数を示している。Unbalanced を用いた場合の結果より、学習サンプル数の少ないクロヒメカンアオイやサンインアオイ、トウゴクサイシンにおいて、ASL を用いることで F1-score が改善されている。ASL はサンプル数の少ない、すなわちネガティブサンプルの多い、属性クラスの精度改善に大きく寄与している。

一方、Balanced を用いた結果では、ASL の精度は大きな変化がないものの、BCE と WFL において大きな精度改善が見られた。この精度向上の理由として、食草ラベルの強い共起性が考えられる。具体的には、図 1(b) より、ランヨウアオイとタイリンアオイが共にポジティブになる可能性はなく、ミヤマアオイがポジティブであれば、ヒメカンアオイもポジティブであるなどのラベルの共起性が考えられる。この食草分類のポジティブラベルの共起性は、一般的なコンピュータビジョンタスクである歩行者や顔画像の属性推定と比較して、非常に強いと考えられる。これらのことから、ラベルの分布は比較的シンプルなため、単純なサンプル数の調整により精度が大きく改善したものと考えられる。より複雑なポジティブラベルの組み合わせが存在する場合は、ASL などの誤差関数による改善が期待できるが、食草ラベルの共起性や分布の詳細な解析については今後の予定の一つである。

5.3 アテンションマップの評価

図 5 に各食草ラベルの平均のアテンションマップを示す。

図 5(a-c) に示す、Unbalanced を用いた場合のアテンションマップでは、翅全体を注視している場合や分類に不要な背景に強いアテンションが発生する場合などがある。一方、サンプルバランス調整後の結果では、背景への注視領域が抑制されており、ギフチョウの領域に注視領域が存在している。このことから、サンプルバランスの調整が安定したアテンションの獲得に貢献しており、データサンプリングを行うことで安定したアテンションマップの獲得が可能である。

表3 属性ごとの F1-score [%]. Training set の行は、各食草ラベルのサンプル数を示している。

	カキガタアオイ	カントウカンアオイ	クロヒメカンアオイ	コシノカンアオイ	サンインアオイ	スズカンアオイ	ダイリニアオイ	トウゴクサイシン	ヒメカンアオイ	ミヤコアオイ	ユキグニカンアオイ	ランヨウアオイ	mean
Training set: unbalanced	188	99	44	272	43	72	73	33	1659	272	115	287	—
BCE	100.00	91.30	53.33	66.66	30.76	41.66	97.29	87.50	85.30	88.37	52.38	99.30	74.49
WFL [18]	100.00	95.83	30.76	75.96	30.76	53.84	94.44	87.50	80.00	89.70	57.14	99.30	74.60
ASL [16]	98.94	93.61	77.77	67.82	77.77	57.14	97.29	94.11	74.28	85.00	58.53	98.59	81.74
Training set: Balanced	282	294	418	408	408	223	365	330	610	408	436	576	—
BCE	100.00	97.95	73.68	80.85	95.23	63.41	97.43	100.00	96.17	84.76	74.57	100.00	88.67
WFL [18]	100.00	100.00	77.77	76.11	73.68	80.00	100.00	100.00	96.60	88.23	73.07	100.00	88.79
ASL [16]	100.00	100.00	72.72	71.42	80.00	68.57	88.37	56.25	93.57	71.35	67.69	91.71	80.14

次に、図 5(e-f) に示す、Balanced を用いたアテンションマップにおいて、誤差関数を変更した場合の違いを比較する。F1-score の高い BCE 及び WFL では、背景部分への強いアテンションがほとんど見られず、胴体や翅を主に注視している、一方、ASL は Unbalanced な場合と比較するとノイズは低減されているものの、BCE や WFL と比較すると認識すべきでない領域を注視している。以上より、説明性の高いアテンションマップの獲得にはサンプルバランスの調整が大きく寄与しており、誤差関数の変更では精度は改善するものの、アテンションマップの説明性はあまり改善しない傾向が明らかとなった。

5.3.1 専門家知見を踏まえた考察

次に、上記の結果より最も精度が高く、ギフチョウの翅に安定して注視しているサンプル調整後の WFL を用いた場合のアテンションマップ (図 5(e)) をチョウ類の専門家に提示し、獲得した注視領域と斑紋との関係性について議論を行う。

全体的な傾向に対する考察として、前側の翅の上部の縁付近 (前翅前縁中央～中室) にアテンションが当たることが多い。これは専門家が採集地を推定する際に確認する、代表的な部分であることが知られている [32]。そのため、マルチラベル分類 ABN が獲得した注視領域と専門家の一般的な知見が一致しており、アテンションマップが提示する説明性が妥当であると言える。

また、図 5(e) のユキグニカンアオイに対するアテンションマップでは、前側の翅の下部の縁付近 (前翅内縁付近) にアテンションが集中している、この箇所は専門家の知見とは異なる箇所であり、今後の詳細な解析を行うことで、従来は認識されていなかった新たな斑紋の特徴を発見できる可能性があり、今後の課題の一つである。

6. 結 言

本稿では、マルチラベル分類 ABN を用いて、ギフチョウ斑紋画像からの食草分類およびその分類結果に対する判断根拠を可視化を行った。クラス不均衡を考慮した誤差関数およびデータサンプリングを適用し学習した結果、高い分類精度を達成した。また、アテンションマップを用いた食草分類とチョウ斑紋の解

析により、学習により獲得したアテンションマップが専門家の知見と一致する箇所を注視していることが明らかとなった。

今後の予定としては、より詳細な検証および専門家の知見以外の新たな地域変異の傾向を明らかにすることなどが挙げられる。また、本データセットへのギフチョウサンプルの追加およびデータベースの公開なども順次行う。

謝辞 本研究は中部大学問題複合体を対象とするデジタルアース共同利用・共同研究 IDEAS202006・IDEAS202105 および JSPS 科研費 21H02215 の助成を受けたものである。

文 献

- [1] T. Inomata, ed., Atlas of the Japanese Butterflies, Takeshobo, 1986.
- [2] Y. Okuyama, *et al.*, “Radiation history of Asian Asarum (sect. Heterotropa, Aristolochiaceae) resolved using a phylogenomic approach based on double-digested RAD-seq data,” *Annals of Botany*, vol.126, no.2, pp.245–260, 2020.
- [3] 矢後勝也, 他, “画像認識技術を用いたチョウ斑紋地域変異の定量的考察,” IDEAS Joint Usage/Joint Research Report, 2021, pp.49–53, 2022.
- [4] 矢後勝也, 他, “深層学習による画像認識技術を用いたチョウ斑紋の地理的変異に関する定量的考察,” IDEAS Joint Usage/Joint Research Report, 2022, pp.47–52, 2023.
- [5] H. Fukui, *et al.*, “Attention branch network: Learning of attention mechanism for visual explanation,” *CVPR*, pp.10697–10706, 2019.
- [6] J.F.H. Cuthill, *et al.*, “Deep learning on butterfly phenotypes tests evolution’s oldest mathematical model,” *Science Advances*, vol.5, no.8, p.eaaw4967, 2019.
- [7] D.C. Wham, *et al.*, “Measuring perceptual distance of organismal color pattern using the features of deep neural networks,” *bioRxiv*, 736306, 2019.
- [8] R. Zhao, *et al.*, “Butterfly recognition based on faster r-cnn,” *Journal of Physics: Conference Series*, vol.1176, no.3, p.032048, 2019.
- [9] Z. Lin, *et al.*, “Increasingly specialized perception network for fine-grained visual categorization of butterfly specimens,” *IEEE Access*, vol.7, pp.123367–123392, 2019.
- [10] Z. Lin, *et al.*, “Fine-grained visual categorization of butterfly specimens at sub-species level via a convolutional neural network with skip-connections,” *Neurocomputing*, vol.384, pp.295–313, 2020.
- [11] M. Fan, *et al.*, “Identification of papilionidae species in yunnan province based on deep learning,” *International Conference on Image, Vision and Computing (ICIVC)*, pp.611–614, 2022.
- [12] F. Schroff, *et al.*, “Facenet: A unified embedding for face recognition and clustering,” *CVPR*, pp.815–823, 2015.

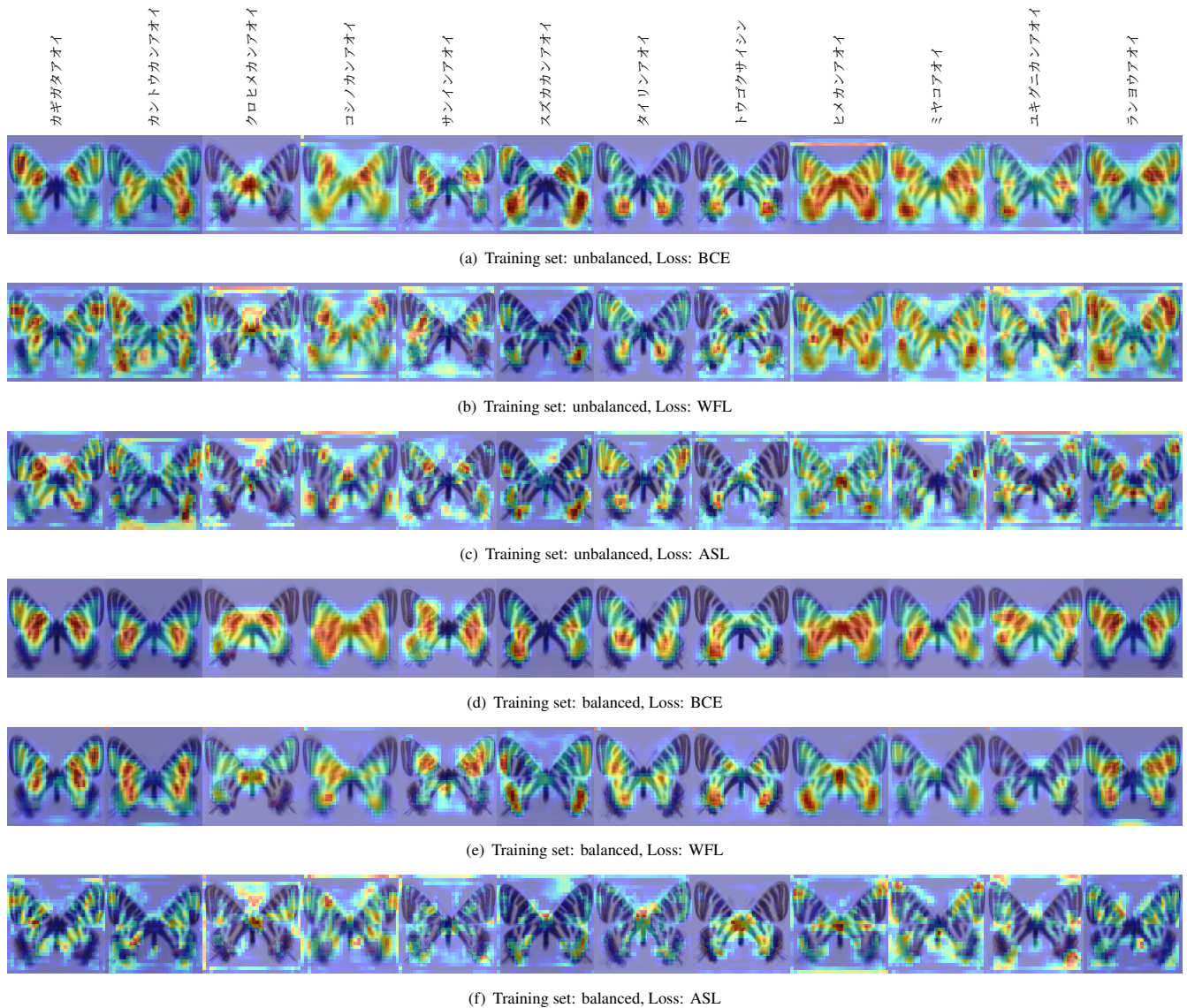


図5 食草ラベルごとのアテンションマップの平均画像。図上部に各列のアテンションマップが対応する食草ラベル名を示している。

- [13] C. Huang, *et al.*, “Learning deep representation for imbalanced classification,” CVPR, pp.5375–5384, 2016.
- [14] C. Huang, *et al.*, “Deep imbalanced learning for face recognition and attribute prediction,” IEEE Transactions on Pattern Analysis and Machine Intelligence, vol.42, no.11, pp.2781–2794, 2020.
- [15] Y. Wang, *et al.*, “Dynamic curriculum learning for imbalanced data classification,” ICCV, pp.5017–5026, 2019.
- [16] T. Ridnik, *et al.*, “Asymmetric loss for multi-label classification,” ICCV, pp.82–91, 2021.
- [17] T. Kobayashi, “Two-way multi-label loss,” CVPR, pp.7476–7485, 2023.
- [18] N. Sarafianos, *et al.*, “Deep imbalanced attribute classification using visual attention aggregation,” ECCV, pp.680–697, 2018.
- [19] F. Charte, *et al.*, “Resampling multilabel datasets by decoupling highly imbalanced labels,” Hybrid Artificial Intelligent Systems, pp.489–501, 2015.
- [20] F. Charte, *et al.*, “Dealing with difficult minority labels in imbalanced multilabel data sets,” Neurocomputing, vol.326–327, pp.39–53, 2019.
- [21] B. Liu and G. Tsoumakas, “Synthetic oversampling of multi-label data based on local label distribution,” ECML PKDD, pp.180–193, 2019.
- [22] B. Liu, *et al.*, “Multi-label sampling based on local label imbalance,” Pattern Recognition, vol.122, p.108294, 2022.
- [23] R.R. Selvaraju, *et al.*, “Grad-cam: Visual explanations from deep networks via gradient-based localization,” ICCV, pp.618–626, 2017.
- [24] B. Zhou, *et al.*, “Learning deep features for discriminative localization,” CVPR, pp.2921–2929, 2016.
- [25] 渡辺康之, 日本の昆虫 1 ギフチョウ, 文一総合出版, 1985.
- [26] 小檜山賢二, 福井弘道, “生物の地域分布可視化,” IDEAS Joint Usage/Joint Research Report, 2017, pp.45–73, 2019.
- [27] 小檜山賢二, 他, “生物の地域分布可視化によるサイエンス・コミュニケーション支援システムの開発,” IDEAS Joint Usage/Joint Research Report, 2018, pp.95–104, 2000.
- [28] M. Lin, *et al.*, “Network in network,” ICLR, pp.1–10, 2014.
- [29] T.-Y. Lin, *et al.*, “Focal loss for dense object detection,” ICCV, pp.2999–3007, 2017.
- [30] J. Deng, *et al.*, “Imagenet: A large-scale hierarchical image database,” CVPR, pp.248–255, 2009.
- [31] K. He, *et al.*, “Deep residual learning for image recognition,” CVPR, pp.770–778, 2016.
- [32] 蝶研出版編集部 (編), ギフチョウ 88 か所めぐり, 蝶研出版, 大阪, 1986.