多値ラベル分類 ABN を用いたチョウ斑紋の認識と地域変異に関する研究

EP19004 荒井嵩貴　　指導教授：山下隆義

1. はじめに

日本の本州に生息する蝶類であるギフチョウは，翅の斑紋に地域変異があることが知られている．この地域変異は地形，環境，食草などの要因が関与すると考えられているが，その解明には至っていない．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の分析から食草と斑紋の関係性を明らかにすることを目的とし，Attention Branch Network (ABN) を用いた食草分類とその判断根拠の可視化を行う．

2. Attention Branch Network

ABN [1] は，視覚的説明と画像分類タスクの精度向上を目的とし，推論時における判断根拠の可視化を行う手法である．ABN は入力から特徴マップを獲得する Feature extractor，特徴マップを受け取り，クラス識別とアテンションマップの生成を行う Attention branch (AB)，最終的な各クラス確率を出力する Perception branch (PB) で構成される．

3. 提案手法

ギフチョウサンプルの数は種類や採集場所に依存するため，生息地によりサンプル数に偏りが存在する．そのため，少数クラスの分類が困難になり，認識精度の低下に繋がる．また，複数の生息地域にまたがることもある．そこで本研究では，図 1 に示すクラス不均衡を考慮した損失関数の導入及び，多値ラベルに対応した ABN による分類手法を提案し，これらの問題に対応する．

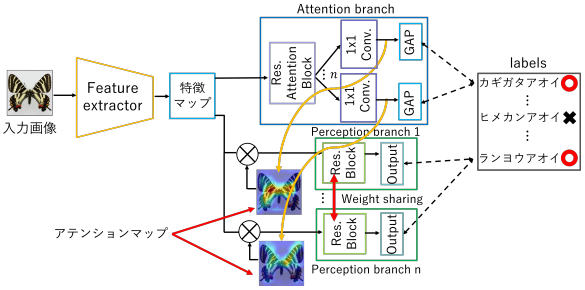

図 1: n ラベル分類に拡張した ABN の構造

提案手法では，AB にて Global Average Polling (GAP) により求めた各クラスの予測結果と正解ラベルから学習し，アテンションマップを出力する．その後，特徴マップとの乗算をクラスごとに行って PB を拡張し，多値ラベル分類を導入する．学習には，クラス不均衡の軽減を目的とした重み付きフォーカルロス [2] を式 (1) の AB の損失関数 L_{att} に導入する．重み付きフォーカルロスを式 (2) に示す．

$$L = L_{att} + \sum_{k=1}^{n} L_{per_k} \tag{1}$$

$$\begin{aligned} L_{att} = -\sum_{k=1}^{n} w_k \big(&(1-\sigma(\hat{y}_p^k))^{\gamma} \log(\sigma(\hat{y}_p^k)) y^k \\ &+ \sigma(\hat{y}_p^k)^{\gamma} \log(1-\sigma(\hat{y}_p^k))(1-y^k) \big) \end{aligned} \tag{2}$$

ここで，w_k はクラス k の重み，$\hat{y}_p^k$ は logits，y^k は正解ラベル，$\sigma(\cdot)$ はシグモイド関数，γ はハイパーパラメータである．PB の損失関数は L_{per_1} から L_{per_n} までの n クラス分の和となる．多値ラベルによる食草のクラス予測後，PB から予測結果を出力する．学習した多値ラベル分類 ABN を用いて，アテンションマップを可視化し，地域ごとに平均した結果を地域変異の調査に用いる．

4. 評価実験

提案手法とサンプル数の調整による評価実験を行う．

4.1. データセット

データセットは藤岡知夫の蝶類コレクションのオスのギフチョウ標本を使用する．ギフチョウが食べる食草である 7 種類のカンアオイを正解ラベルとし，0 から 6 までの ID を割り振る．ラベルごとの内訳を表 1 に示す．本実験では，クラス不均衡を軽減するために，中部地方や近畿地方に分布するヒメカンアオイと新潟県や福島県に分布するユキグニカンアオイのサンプル数を調整した場合と比較する．

表 1: サンプル数調整前後の内訳

ID	正解ラベル	サンプル数 [枚]	
		調整前	調整後
0	カギガタアオイ	235	235
1	カントウカンアオイ	124	124
2	コシノカンアオイ	341	341
3	ヒメカンアオイ	2,075	**400**
4	ミヤコアオイ	340	340
5	ユキグニカンアオイ	145	**245**
6	ランヨウアオイ	359	359

4.2. 実験概要

ギフチョウサンプルを調整せず，重み付きフォーカルロスを用いない場合をベースラインとし，重み付きフォーカルロスのみ導入した場合と，さらにサンプル調整した場合の F 値のスコアとアテンションマップを比較する．モデルは ResNet18 を backbone とした多値ラベル分類 ABN とする．学習回数は 100 エポック，学習率は 0.01，バッチサイズは 8，最適化手法は SGD，ImageNet による事前学習を行う．

4.3. 実験結果

ID ごとの比較結果を表 2 に示す．表 2 より，平均スコアに着目すると，重み付きフォーカルロスのみを導入した場合は 0.4pt 低下したが，さらにサンプル調整をした場合は 5.69pt 向上した．よって，損失関数とサンプル数の両方を考慮することで，クラス不均衡が改善されると考える．

表 2: ID ごとの F 値 [%]

ID	多値ラベル分類 ABN	重み付きフォーカルロス	重み付きフォーカルロス +サンプル調整
0	**100.00**	**100.00**	**100.00**
1	**100.00**	98.30	**100.00**
2	80.32	75.47	**88.33**
3	93.83	**94.51**	94.05
4	**92.61**	87.50	82.79
5	53.33	65.21	**94.73**
6	**99.23**	95.52	**99.23**
平均	88.47	88.07	**94.16**

次に，アテンションマップのラベルごとの平均を図 2 に示す．図 2(a) では羽の左側及び右側中央部分をわずかに注視し，図 2(b) では羽の左上部分を注視する傾向が見られた．

(a) ヒメカンアオイ:3 を食草とするギフチョウ

(b) ユキグニカンアオイ:5 を食草とするギフチョウ

図 2: アテンションマップの平均

5. おわりに

本研究では，重み付きフォーカルロスの導入とサンプル数を考慮した多値ラベル分類 ABN を用いて，食草分類とその判断根拠の可視化を行った．結果から，クラス不均衡の改善とラベルごとで注視領域が変化することを確認した．今後は，食草および採集地の環境情報を考慮した詳細な斑紋の解析を行う予定である．

参考文献

[1] H. Fukui, *et al.*, “Attention Branch Network: Learning of Attention Mechanism for Visual Explanation”, CVPR, 2019.

[2] N. Sarafianos, *et al.*, “Deep Imbalanced Attribute Classification using Visual Attention Aggregation”, ECCV, 2018.