

擬態によって学習・識別困難な対象の GAIN に基づく深層学習

Deep learning based on GAIN for mimesis

友野 海^{*1}

Kai Tomono

延原 肇^{*1}

Hajime Nobuhara

河合 新^{*1}

Shin Kawai

^{*1}筑波大学大学院システム情報工学研究科知能機能システム専攻

Intelligent Interaction Technologies, Graduate School of Systems and Information Engineering, University of Tsukuba

The bias of the training image data is considered to have a great influence on the identification of mimicable organisms. The purpose of this study is to examine the effectiveness of Guided Attention Inference Networks (GAIN) for recognizing the *Anolis carolinensis*. We also confirmed the improvement of classification accuracy from the verification results.

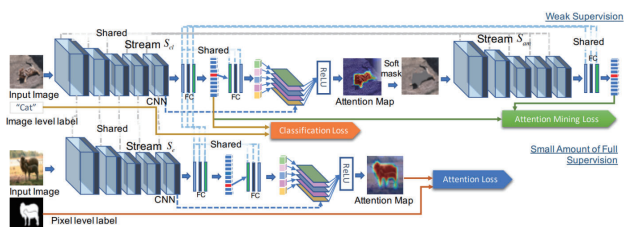


図 1: GAIN



図 2: 保護色の変化

1. はじめに

本研究では、擬態機能を有する生物、例えば、グリーンアノール（学名：Anolis carolinensis）を識別のターゲットとする。従来の深層学習の枠組みでは、これらの擬態機能を有する対象の識別をするための学習が困難であるため、判断根拠の可視化だけでなく Attention Map を訓練に用いることで、さらなる精度の向上や説得性をモデルに持たせることを可能にした Guided Attention Inference Networks (GAIN) (図 1) に着目する [Li 18]. GAIN は学習データに偏りがある場合にも有効であり、必ずしも Attention Map に教師データを必要としない。

対象とする爬虫類の種類はアノールトカゲのみとし、個体の画像データは小笠原諸島父島で撮影したものを扱う。擬態を行う生物は環境（色、構造）に合わせた変化を見せ（図 2）、CNN のフィルターが画像データの偏りを学習しやすく、検出精度の低下が起こりやすいと考えたため、GAIN を適用することで改善できる。

VGG16 を基本モデルとしたネットワークに対し、GAIN を適用することの有効性を検証する。試験データ 378 枚と別に用意した訓練データ及び検証データ約 1,000 枚を用いたホルドアウト検証を行う。

2. GAIN に基づく提案システム

GAIN[Li 18] は、画像領域分割タスクを行うことを可能とする end-to-end な学習手法であると同時に、データセットのバイアスに影響を受けにくいモデルの学習を行うことが可能であ

る（データセットのバイアスに影響を受けていると考えられる学習の例を図 3 に示す）。GAIN はピクセルレベル教師データを用いない学習手法と用いる方法をシームレスに選択することが可能であるため、ピクセルレベル教師データを用意することが困難な大規模データとそうでない小規模データの両方で使い分けが可能な汎用性が特徴である。GAIN では 3 つの stream を扱い、それぞれに対応した処理を行う。classification stream (Scl) は、CNN を利用した一般的な画像のクラスタリングと処理は同じであるが、attention mining stream (Sam) 及び external stream (Se) は異なる。Sam では同じ入力画像にマスキングを施したものを新たに同モデルに入力するが、マスキング画像の生成には畳み込み層の最終層と出力層の関係から生成した attention map を用いる。Se ではクラスラベルデータとは別に用意したピクセルレベル教師データを用いて学習を行う。これら全ての stream は共通して同じモデル、重みを使用する。また、Scl, Sam 及び Se に対応する損失関数を classification loss (L_{cl}), attention mining loss (L_{am}) 及び attention loss (L_{ext}) とし、これらの最小化を行うが、本研究では Se は用いない。attention map を生成するために、まず Scl の最終出力に対する CNN の最終層で勾配を計算し、その後 Global Average Pooling (GAP) [Zhou 16] を行うことで出力に対する各重みの寄与度合いを知ることができる式 (1)。GAP は各特徴マップの数値を平均する方法である。ただし、1 層 k 番目の特徴マップを $f_{l,k}$ とし、c クラスの出力を s^c とおく。

$$w_{l,k}^c = \text{GAP} \left(\frac{\partial s^c}{\partial f_{l,k}} \right) \quad (1)$$

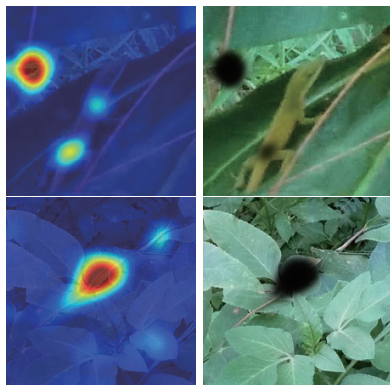


図 3: Attention Map の可視化と比較 (左) 誤った学習 (右) 正しい学習

表 1: 拡張前のデータセット

	Train	Valid	Test	Sum
Positive	839	226	378	1443
Negative	839	226	378	1443
Sum	1678	452	756	2886

また式 (2) より, attention map を生成できる.

$$A^c = \text{ReLU}(\text{conv}(f_l, w^c)) \quad (2)$$

これらの手順は Gradient-weighted Class Activation Mapping (Grad-CAM) と呼ばれる手法であり [Selvaraju 17], Class Activation Mapping (CAM) と比較して出力クラスへの視覚的な説明性を向上させている. ピクセルレベル教師データを用いない場合は式 (3) より, シグモイド関数 T を用いてマスキング画像を生成する.

$$I^{*c} = I - T(A^c) \circ I \quad (3)$$

n をデータ数, α を 1 とし各誤差関数を式 (4), (5) に示す.

$$L_{am} = \frac{1}{n} \sum_c S^c(I^c) \quad (4)$$

$$L_{self} = L_{cl} + \alpha L_{am} \quad (5)$$

本研究では L_{self} の最小化を行う.

3. 実験条件

3.1 データセット

学習に用いるデータセットを表 1 に示す. 表 1 のデータは拡張前のデータセットであり, 画像の収集にはスマートフォン (Nexus 5) を用いた. また, 学習時には学習データを 18,971 枚, 評価データを 5,006 枚に拡張したものを使用した. また, 試験データは拡張を行わずに用いた. 拡張方法は画像の位置ずらし及び切り取りを行った. アノールトカゲが確認された周囲の植物の画像を負例データとし, 同様に拡張を行った.

3.2 モデル構造及び学習条件

アノールトカゲとそれ以外のクラスを扱う 2 値分類を行う. VGG16 の畳み込み層を入力層側から 7 層固定し, 残りの層を

表 2: 拡張前のデータセット

	GAIN なし	GAIN あり
分類精度	0.9762	0.9788

学習させる. VGG16 の特徴抽出層の出力は Global Average Pooling により処理され, 全結合層は入力層をユニット数 512, 中間層をユニット数 256, 出力層をユニット数 1 として構築したものを扱う. よって, VGG16 本来の全結合層は用いずに新たに構築・接続したものを学習モデルとして扱う. また, 学習時の最適化手法は Adam (learning rate = $1.0e-7$) を用い, バッチサイズを 32 とした. 最終層に sigmoid 関数を用い, それ以外は ReLU を活性化関数とする. Scl 実行時の損失関数はバイナリークロスエントロピーを用いる.

3.3 実験手順

縦横が 224 ピクセルの RGB 画像データを入力とするモデルを用いて順伝播させ, 誤差逆伝播法による重みの更新は行わずに Sam に関する順伝播も同時に行う. ただし, Scl 時に入力として与えた画像を対象に式 (3) によるマスキングを行い, Sam 時の入力とする. Scl 及び Sam 時のモデルの重みは共有されており, 同じモデルを扱う. Scl, Sam の処理後に出力から式 (5) を計算し, 誤差逆伝播法から重みの更新を行う.

4. 実験結果と考察

GAIN を使用しないモデルでは 25 エポック以降から検証データの分類精度が低下したため, 25 エポック目の状態で学習を打ち切った. 同様に GAIN を使用したモデルでは 45 エポック目で学習を打ち切った. これらの学習済みモデルを使用し, 試験データによる検証を実施した結果, 表 2 に示すような結果となり GAIN を使用した場合の分類精度の向上が確認できた. GAIN による分類精度の向上は学習データのバイアスを学習してしまう問題の改善によるものであると考えられるが, 実験回数が不十分のため, 本研究結果は可能性の確認にとどまる.

5. 総括

本研究では外来生物のアノールトカゲを対象とする GAIN を用いた深層学習を行い, バイアスの学習への効果を確認した. 今後は正しい学習が行われているかの定量的な評価を行うとともに, GAIN ext の導入による効果の検証を行う方針である.

参考文献

- [Li 18] Li, K., Wu, Z., Peng, K., Ernst, J., and Fu, Y.: Tell Me Where to Look: Guided Attention Inference Network, *CoRR*, Vol. abs/1802.10171, (2018)
- [Selvaraju 17] Selvaraju, R. R., Cogswell, M., Das, A., Vedantam, R., Parikh, D., and Batra, D.: Grad-CAM: Visual Explanations from Deep Networks via Gradient-Based Localization, in *2017 IEEE International Conference on Computer Vision (ICCV)*, pp. 618–626 (2017)
- [Zhou 16] Zhou, B., Khosla, A., Lapedriza, A., Oliva, A., and Torralba, A.: Learning Deep Features for Discriminative Localization, in *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 2921–2929 (2016)