

# ラベルマップを用いて学習したCNNによる 地理情報と環境情報からの生物種推定

竹ノ内達哉<sup>†</sup> 青野 雅樹<sup>††</sup>

<sup>†</sup> 豊橋技術科学大学 情報・知能工学専攻 〒441-8580 愛知県豊橋市天伯町雲雀ヶ丘 1-1

<sup>††</sup> 豊橋技術科学大学 情報・知能工学系 〒441-8580 愛知県豊橋市天伯町雲雀ヶ丘 1-1

E-mail: <sup>†</sup>takenouchi@kde.cs.tut.ac.jp, <sup>††</sup>aono@tut.jp

あらまし 地球上に生息する生物の種は非常に多い。故に、種を同定する作業は非常に困難であり、そのプロセスやツールの改善が強く求められている。また、生物はそれぞれの種に適した環境に生息しており、その分布は地理的な位置に加え、その周囲における気温や湿度等の環境に依存する。それら位置情報や環境情報を用いることで、特定の場所に生息する種を予測することができる。しかしながら、特定の場所に生息する種を漏れなく観測し網羅することは困難であり、作成されるデータセットには少なからずノイズが混入する。そのため、種の観測漏れ等を考慮して、近傍で観測された種をもとにそれらを補う必要がある。本研究では、ラベルマップを用いて学習したCNNによる地理情報と環境情報からの生物種推定手法を提案し、実験結果を報告する。

キーワード 機械学習, ニューラルネットワーク, CNN, ResNet, 生物多様性, 地理情報, 環境情報, 生物種推定

## 1 はじめに

地球上には多様な生物が生息しており、生物多様性 (Biodiversity) という言葉があるようにその性質は種ごとに異なる。Mora ら [1] によると、地球上の生物の種の総数は真核生物のみで 870 万 (標準誤差  $\pm 130$  万) 種であると予測されている。通常、それだけの種を一人の人間が把握するのは非常に困難であり、生物多様性科学 (Biodiversity Science) の研究には計算機を活用する情報学 (Informatics) が大きく役立つ。それら生物多様性科学と情報学を合わせた分野は生物多様性情報学 (Biodiversity Informatics) と呼ばれる。

また、現在盛んに取り組まれているニューラルネットワークは、情報学のみならず様々な分野においても研究されている。生物多様性情報学における生物種の推定タスクもその一つであり、2019 年に開催されたコンペティションである GeoLife-CLEF2019 [2] では、地理情報と環境情報から種の推定を行うタスクに取り組んでいる。

本研究では、地理情報等からそこで最も観察される可能性が高い種を予測することにより、生物多様性情報学における種の同定プロセスやそれに用いるツールを改善することが可能となる。それにより、インベントリと呼ばれる地域の生物種の調査を促進し、位置ベースの予測サービスの開発、非専門家の補助、アプリケーションの作成による教育への応用等に役立つことが期待される。しかし、地球全体を網羅することは困難であることや、サンプリングバイアスなどの要因により、地理情報から直接種の分布予測を行うモデルを学習することは非常に困難である。

また、生物は種によって様々な性質を持ち、生存に適した気候や土壌などを始めとする環境は種の分布に強く関係している

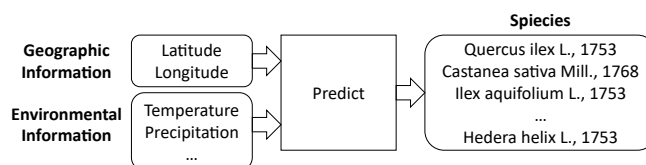


図 1 生物種予測システム概略図

と考えられている [3, 4]。そのため、地理情報から分布モデルを直接学習する代わりに、地理情報に加えて、その周囲における平均気温や降水量、土壌の種類や表土、水までの距離などを始めとした変数からなる環境情報に基づき種の予測を行うことにより、その予測精度を改善することが期待される。

本稿では、図 1 に示すように地理情報と環境情報による特定の場所で最も観察される可能性が高い種の推定を行う、ラベルマップを用いた予測システムを提案する。実験には GeoLife-CLEF2019 で使用されるデータセットを用いて、生物種推定手法の比較と検証を行う。

2 節では、画像処理を行うニューラルネットワークや、生物種の推定に関する研究について述べる。3 節では提案する手法の詳細について説明する。4 節では各種法の比較実験を行い、5 節ではその結果について考察する。6 節では結論を述べ、今後の課題について検討する。

## 2 関連研究

本節では、過去の GeoLifeCLEF コンペティションにおけるテクニカルレポートについて調査し、過去の参加者が得た知見や高い精度を達成した手法について述べる。また、一般的な CNN やデータ拡張などの論文についても調査し、その概要を述べる。

表 1 GeoLifeCLEF2018 データセット

データセット	標本数
訓練データ	218,543
テストデータ	72,849
合計	291,392

## 2.1 GeoLifeCLEF2018

GeoLifeCLEF2018 [5] は、2018 年に開催されたコンペティションであり、その課題では地理情報と環境情報から 3,336 種の生物の予測を行う。データセットは、合計 291,392 個の観察による地理情報と、33 種類の環境情報で構成される。その内訳は表 1 に示すとおりである。

Deneu ら [6] は、本コンペティションにおいて最高の精度を達成しており、その提案は CNN と Random Forest [7] の予測結果を合成する手法であった。Random Forest においては、位置情報のみを使用しており、CNN では、序数変数 (Ordinal Variables) とカテゴリカル変数 (Categorical Variables) の違いに着目し、それぞれを個別の入力データとして CNN で畳み込み、ネットワークの途中で結合している。本手法により、入力画像のチャンネル数の差異による予測結果の依存を軽減し、実際に交差検証法により精度の改善が確認された。これより、カテゴリカル変数をネットワークへの入力として分離することにより精度の向上に寄与すると考えられる。

Taubert ら [8] は画像の回転や反転によるデータ拡張を用いた CNN に加え、地理情報と環境情報を結合したベクトルを用いた eXtreme Gradient Boosting (XGBoost [9,10]) による予測モデルを提案した。77 チャンネルの画像として与えられる環境情報を、XGBoost で用いるために縦横方向に平均してベクトルに変換するが、広い範囲よりも狭い範囲で平均した場合が高精度であった。結果として、中央の 4 画素のみを平均化したほうがより良い結果が得られることを示した。これより、画像全体の特徴と中央の局所的な特徴はそれぞれ別の特徴を持っており、どちらも予測に使用することができるのではないかと考えられる。

Aravindan ら [11] は LSTM [12] で拡張された ResNext [13] を用いて、CNN と LSTM を組み合わせたモデルを生物の 5 段階の分類体系 (Class, Order, Family, Genus, Species) に基づき分類する手法を提案した。

## 2.2 GeoLifeCLEF2019

GeoLifeCLEF2018 に続き、同様に GeoLifeCLEF2020 [2] が 2019 年に開催されている。Negri ら [14] は、本コンペティションにおいて最高の精度を達成しており、その提案は環境情報を単純にチャンネルとして結合して予測する手法であった。ネットワークとして Inception V3 [15] を用いており、通常の RGB 画像を入力とする代わりに 77 チャンネルの環境情報を用いるよう変更を加えている。また、出力には softmax 関数を適用し、Cross Entropy Loss により学習を行っている。ネットワーク自体は既に存在するものを用いたことは、GeoLifeCLEF2018 で複雑なネットワーク構造を用いて高い精度を達成した Deneu ら

の手法と異なる。

Si-Moussi ら [16] は Negri らと同様に CNN を用いた手法に加え、種の共起ベクトルのみを用いた DNN を使用する手法を提案した。CNN では、変数の種類 (Continuous, Ordinal, Categorical) に応じて異なる畳み込みを適用した後に連結される。これは、GeoLifeCLEF2018 における Deneu らの手法と同じ戦略である。また、Cross Entropy Loss にクラスに応じた重みを適用することにより、クラス間の不均衡に対処した。そして、共起ベクトルのみを使用する手法は、専門家により作成された種の共起集合を用いてニューラルネットワークの学習を行う。結果として CNN を用いた手法が精度が高いことが確認されている。

Krishna ら [17] は、主に古典的な機械学習手法を用いた手法を提案している。地理情報のみを用いた手法や地理情報と環境情報を用いた手法を XGBoost や Random Forest 等を用いて予測している。この手法により Negri らや Si-Moussi らに次ぐ精度を達成している。

## 2.3 CNN

CNN に関する研究としては、He ら [18] の ResNet が存在する。深い層に及ぶ CNN ではその深さにより学習が阻害されるという勾配消失問題や勾配爆発問題 [19–21] が存在していたが、残差 (Residual) を用いて学習させることによりそれを改善した。さらに、Bottleneck 構造により、VGG [22] と比較してネットワークの深さに対してパラメータ数を抑えることができる。また、Hang ら [23] の研究によると、ResNet におけるダウンサンプル時に MaxPooling を適用することや、さらにその適用を畳み込みの後に遅らせることで精度の向上が期待できると報告されている。

一方、画像のデータ拡張や正規化に関する研究として、Sangdoo ら [24] の提案した CutMix がある。これは、データセットの異なるサンプルを合成してデータ拡張を行う手法の一つであり、ノイズのあるデータセットに対して特に有効であることが示されている。また、類似する手法として Mixup [25] や Cutout [26] が存在する。

## 3 提案手法

ベースラインを図 2 に、提案手法を図 3 に示す。本研究では ResNet のバリエーションの一つである ResNet18 を使用する。ResNet18 は全 18 層の畳み込み層を持ち、残差接続により多数の畳み込みを実現している。

提案手法はベースラインとなる ResNet18 を元に、データセットのノイズに対処した学習を行う。先行研究では、同じ環境に複数の種が生息しているということに対して根本的な対策がされていなかった。そこで、本研究では訓練データを用いて擬似的なマルチラベルのデータとしてラベルマップを作成することで、学習後の予測精度の改善を図る。また、その学習に適したデータ拡張や Loss の改善を行う。

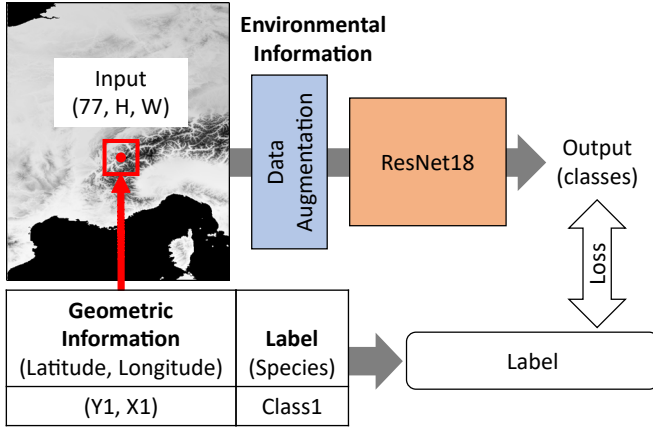


図 2 ベースライン

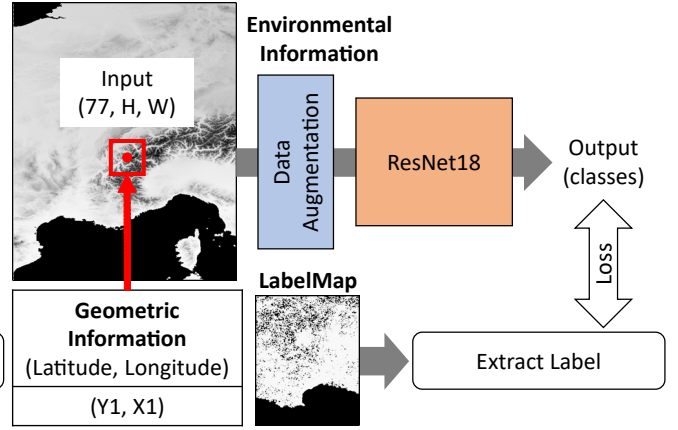


図 3 提案手法

### 3.1 ベースライン

ベースラインでは、最初に入力として地理情報と環境情報が与えられると、与えられた地点を中心とする周囲の環境情報を抽出する．ネットワークへの入力として、抽出された環境情報が合計 77 チャンネルを持つ画像として使用される．なお、ResNet では入力画像のチャンネルは RGB の 3 次元を用いるが、今回用いる入力画像に合わせ最初の畳み込み層の入力を 77 チャンネルに拡張した．また、最後の全結合層も同様に正解ラベルの総数に合わせて調整する．

学習は Deneu ら [6] と同様に SoftMax 関数と CrossEntropy-Loss を用いる．

### 3.2 ラベルマップ

GeoLifeCLEF2019 のデータセットにおいて、地理情報と生物種は 1 対 1 となっている．すなわち、データセットを未加工で用いると、ある環境において生息する生物種は 1 種類のみであると学習することとなる．しかしながら、多くの環境では複数の生物種が同時に生息していると考えられる．これは、一般的な画像分類タスクにおいて、犬と猫が同時に写った画像をどちらか片方のクラスのみで分類するような学習に近いと考えられる．そのような入力はノイズとなり、正常な学習を阻害する可能性が存在する．

そこで、地理的に十分に近い複数のデータを統合し、擬似的なマルチラベルのデータを作成する．1 対 1 で対応しているデータセットに前処理を施し、同じ環境情報に対して複数の生物種を正解ラベルとして扱い学習を行う．本稿では、ネットワークに学習させる正解クラスのラベルを二次元平面上で表すことから、これをラベルマップと表現する．GeoLifeCLEF2019 のデータセットを用いて作成したラベルマップの例を図 4 に示す．ここで、図中の白点は任意のラベルの地理情報を画像中の座標としてプロットしたものである．実際には、縦横方向に加えてクラス数に等しいチャンネル方向の次元が存在するが、可視化のためグレースケールで表している．

また、一般的に種はそれぞれが生存に適した環境に生息すると考えられている [3]．環境の中でも、気温や標高といったものは空間に対して連続的に変化することが多く、それらの環境

がある地点において急激に変化することは稀であると考えられる．そのため、ある環境で種が観測された場合、その付近においても同様の種が生息している可能性は非常に高いと考えられる．これにより、訓練データに対する過学習を防ぎ精度の向上が期待される．

### 3.3 CutMix の拡張

本研究で使用する CutMix [24] について説明する．本手法では、2 組の訓練画像とラベルの組  $(x_A, y_A)$  及び  $(x_B, y_B)$  を合成して、1 組の新しい画像とラベルの組  $(\tilde{x}, \tilde{y})$  を生成する．その合成は式 1 により定義される．

$$\begin{aligned}\tilde{x} &= M \odot x_A + (1 - M) \odot y_B, \\ \tilde{y} &= \lambda y_A + (1 - \lambda) y_B\end{aligned}\quad (1)$$

ここで、 $M$  は各要素が 0 もしくは 1 の値を持つ入力画像と同じ大きさの  $W \times H$  のマスク行列であり、 $\lambda$  は 0 もしくは 1 の間を取る実数である．まず、 $\lambda$  は式 2 のベータ分布に従う．

$$\lambda \sim \text{Beta}(\alpha, \alpha) \quad (2)$$

ここで、 $\alpha$  はハイパーパラメータである．次に、訓練画像の合成を行うマスク行列  $M$  は 4 つのパラメータから生成され、式 3 で定義される．生成される行列は訓練画像  $x_B$  から  $r_x, r_y$  を中心として、 $r_w, r_h$  の大きさの画像を切り取るような行列となる．

$$\begin{aligned}r_x &\sim \text{unif}(0, W), & r_w &= W\sqrt{1 - \lambda}, \\ r_y &\sim \text{unif}(0, H), & r_h &= H\sqrt{1 - \lambda}\end{aligned}\quad (3)$$

$r_x, r_y$  は一様分布に従い画像上の任意の画素を選択した座標であり、 $r_w, r_h$  は切り取る画像の幅と高さとなる．このようにして行われる訓練画像の合成を図 5 に示す．

そして、生成された訓練画像  $\tilde{x}$  には  $x_A$  が元の画像に対して  $\lambda$  の割合で含まれており、 $x_B$  も同様に  $(1 - \lambda)$  の割合で含まれている．ラベルについても同様となる．ここまでは、CutMix の主な処理である．

ここで、パラメータ  $r_x, r_y$  に着目したい． $r_x, r_y$  は訓練画像から切り取る画像の中心座標を表すが、その定義域は切り取る領域  $r_w, r_h$  が考慮されていない．そのため、例えば  $r_x$  が端点

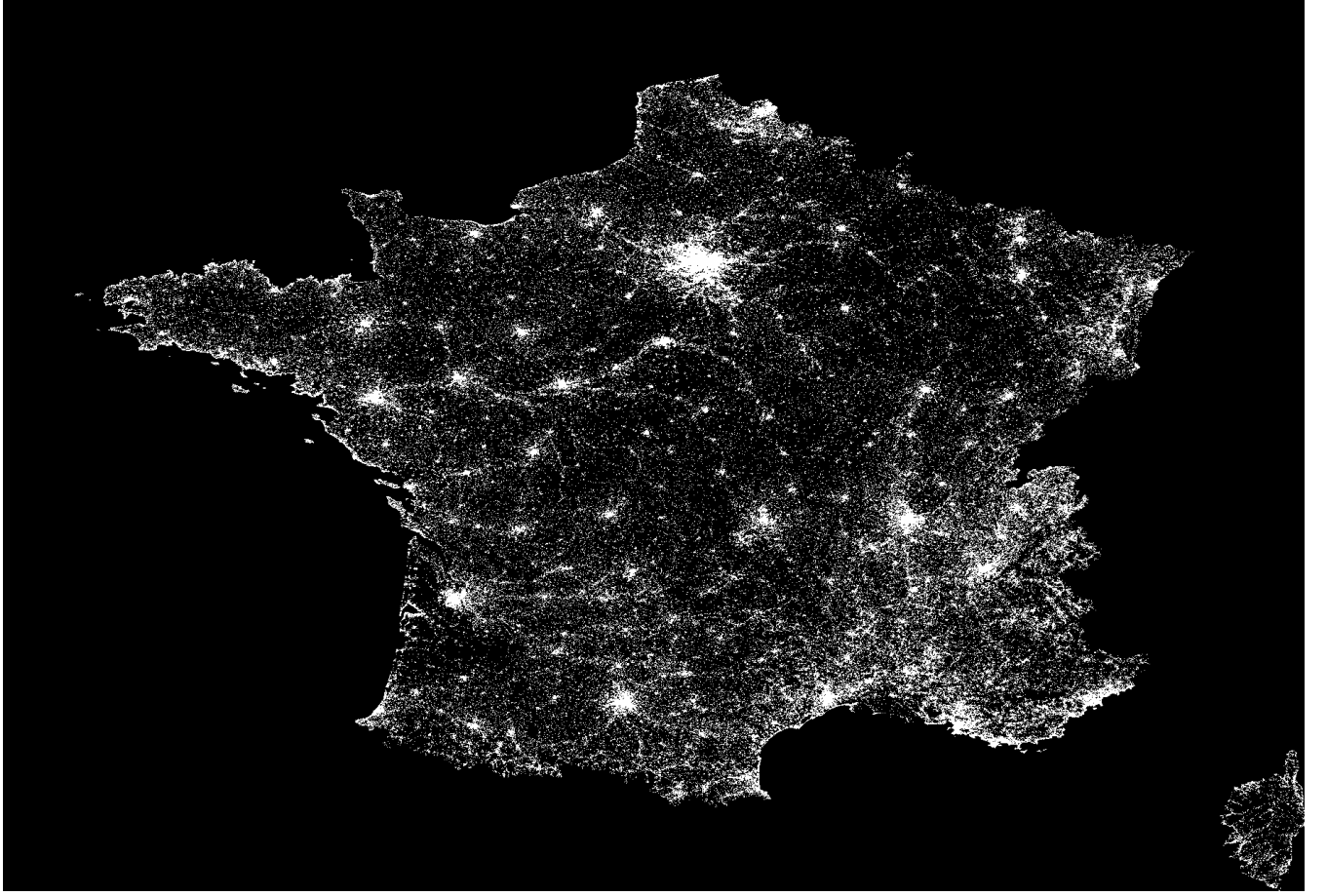


図 4 ラベルマップの例

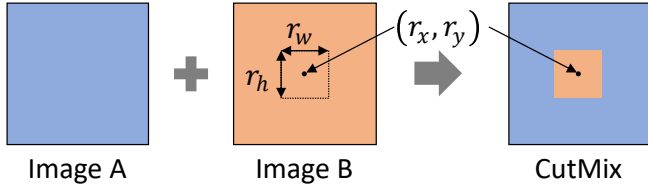


図 5 CutMix

0 を取った場合、 $r_w/2$  の範囲が画像外にはみ出すため、生成される訓練画像  $\tilde{x}$  において元となる訓練画像の割合がそれぞれ  $\lambda, (1 - \lambda)$  から離れてしまう．このようにして生成された訓練画像を使用すると、訓練画像とラベルの対応に齟齬が生じ、ノイズとしてネットワークへの悪影響が生じる可能性が存在する．

そこで、これらのパラメータ  $r_x, r_y$  が  $r_w, r_h$  を考慮するように変更を加える．変更は式 4 に従う．

$$\begin{aligned} r'_x &\sim \text{unif}\left(\frac{r_w}{2}, W - \frac{r_w}{2}\right) \\ r'_y &\sim \text{unif}\left(\frac{r_h}{2}, H - \frac{r_h}{2}\right) \end{aligned} \quad (4)$$

これにより、合成の際に切り取る領域が画像外にはみ出すことを防ぎ、元の CutMix と比較して精度の向上が期待される．

### 3.4 CrossEntropyLoss の拡張

先に述べたようにノイズにより学習が阻害される可能性が存在する．また、本研究で使用するデータセットはマルチラベル

として学習させることが好ましいと考えられるが、タスクそのものはマルチクラス分類問題として用意されているため、ネットワークの出力の総和は SoftMax 関数等により 1 とすることが好ましい．そのため、通常の Cross Entropy Loss により学習した場合、予測が特定のクラスに偏り、精度の低下を招くと考えられる．まず、今回使用する Cross Entropy Loss を式 5 に定義する．

$$\begin{aligned} \text{CrossEntropyLoss}(\mathbf{x}, \text{class}) &= -\log \left( \frac{\exp(\mathbf{x}_{\text{class}})}{\sum_j \exp(\mathbf{x}_j)} \right) \\ &= -\mathbf{x}_{\text{class}} + \log \left( \sum_j \exp(\mathbf{x}_j) \right) \end{aligned} \quad (5)$$

ここで、 $\mathbf{x}$  は各要素がそれぞれのクラスに対応するベクトルであり、 $\text{class}$  は正解クラスを表す．この Cross Entropy Loss では、唯一の正解クラスの出力が Loss の計算に大きく影響していることが伺える．

しかし、ある環境において複数の生物が生息している場合が存在することから、本研究では予測上位のクラス全ての出力が同様に重要であると考えられる．そこで、予測上位  $k$  クラスの内に一つでも正解クラスが含まれていれば、正解クラス以外を含め、上位  $k$  個の出力を全て正解であると仮定する．以下に

Cross Entropy Loss を拡張した Cross Entropy Loss' の定義を式 6 に示す．

$$\begin{aligned} & \text{CrossEntropyLoss}'(\mathbf{x}, \text{class}, k) \\ &= -\log \left( \frac{\sum_{j \in \{j | \text{rank}_j \leq k\}} \exp(\mathbf{x}_j)}{\sum_j \exp(\mathbf{x}_j)} \right) \\ &= -\log \left( \sum_{j \in \{j | \text{rank}_j \leq k\}} \exp(\mathbf{x}_j) \right) + \log \left( \sum_j \exp(\mathbf{x}_j) \right) \end{aligned} \quad (6)$$

ここで， $k$  は上位  $k$  個の出力を使用して，正解クラスと仮定するための定数である．また， $\text{rank}_{\text{class}}$  はクラス  $\text{class}$  の予測順位を表し， $\{j | \text{rank}_j \leq k\}$  は予測順位が  $k$  位以上であるクラスの集合を表す．これにより，上位  $k$  個のクラスの予測値を加算し 1 クラスにまとめる．そして，これら Cross Entropy Loss と Cross Entropy Loss' を用いた TopK Cross Entropy Loss を提案し，その式を式 7 に定義する．

$$\begin{aligned} & \text{TopKCrossEntropyLoss}(\mathbf{x}, \text{class}, k) \\ &= \begin{cases} \text{CrossEntropyLoss}'(\mathbf{x}, \text{class}, k) & (\text{rank}_{\text{class}} \leq k) \\ \text{CrossEntropyLoss}(\mathbf{x}, \text{class}) & (\text{otherwise}) \end{cases} \end{aligned} \quad (7)$$

また，TopK Cross Entropy Loss を通常の Cross Entropy Loss と式 8 により合成した Composed Loss を提案する．

$$\begin{aligned} & \text{ComposedLoss}'(\mathbf{x}, \text{class}, k, \mu) \\ &= \mu \text{CrossEntropyLoss}(\mathbf{x}, \text{class}) \\ &+ (1 - \mu) \text{TopKCrossEntropyLoss}(\mathbf{x}, \text{class}, k) \end{aligned} \quad (8)$$

ここで， $\mu$  は 0 から 1 までの実数値を取る Loss の係数であり，0.5 で平均をとる．これにより，通常の Cross Entropy Loss と TopK Cross Entropy Loss' の特徴を兼ね備えた Loss として用いる事ができる．

## 4 評価実験

各手法の精度を評価するために，評価実験を行った．以下に，使用したデータセットと実験の条件，結果について述べる．提案手法として，3 節で述べた手法を使用する．

### 4.1 データセット

GeoLifeCLEF2019 [2] のデータセットを使用する．本データセットは，地理情報として標本数が 13,263,287 件，クラス数は 34,720 種類の生物（動植物）で構成される．環境情報としては，フランス領土の環境情報 33 種類を含み，その内訳を表 2 に示す．また，環境情報の一部を視覚化したものを例として図 6 に示す．図は 33 種類の環境情報をそれぞれ視覚化したものであり，縦軸と横軸がそれぞれ緯度と経度（北緯）を表し，各図の右側のカラーマップは小領域内での数値を正規化したも

表 2 環境情報の詳細

環境情報	チャンネル数
年間平均気温（月平均）	1
最高気温と最低気温の差	1
等温線	1
温度季節性	1
最も暖かい月の最高気温	1
最も寒い月の最低気温	1
温度年間範囲	1
最も雨の多い四半期の平均気温	1
最も乾燥した四半期の平均気温	1
最も暖かい四半期の平均気温	1
最も寒い四半期の平均気温	1
年間降水量	1
最も雨が多い月の降水量	1
最も乾燥している時期の降水量	1
降水量の季節性	1
最も雨の多い四半期の降水量	1
最も乾燥した四半期の降水量	1
最も暖かい四半期の降水量	1
最も寒い四半期の降水量	1
潜在的蒸発蒸散	1
標高	1
表土利用可能な水容量	1
表土の基本飽和度	1
表土陽イオン交換容量	1
土壌分類クラス	1
グライ層までの深さ	1
不透水層までの深さ	1
土壌侵食性クラス	1
表土の有機炭素含有量	1
表土の充填密度	1
支配的な土性クラス	45
淡水まで 50 メートル未満の場所	1
地上占領	1
合計	77

表 3 使用データセット

データセット	標本数
訓練データ	224,756
テストデータ	56,189
合計	280,945

のである．

本研究では，これらのうちコンペティションで使用された 844 クラスをサブセットとして用いる．使用するサブセットとして，地理情報として標本数 280,945 件を訓練データとテストデータにそれぞれ 224,756 件と 56,189 件で分割した．その内訳を表 3 に示す．

### 4.2 評価方法

主な評価指標として，Top1，Top30 及び Mean Reciprocal Rank (MRR) を用いる．いずれの指標も 0 から 1 の範囲を取り，数値が大きいほど精度が高い．

Top1 は各予測のうち，上位 1 位が正解クラスであれば正解



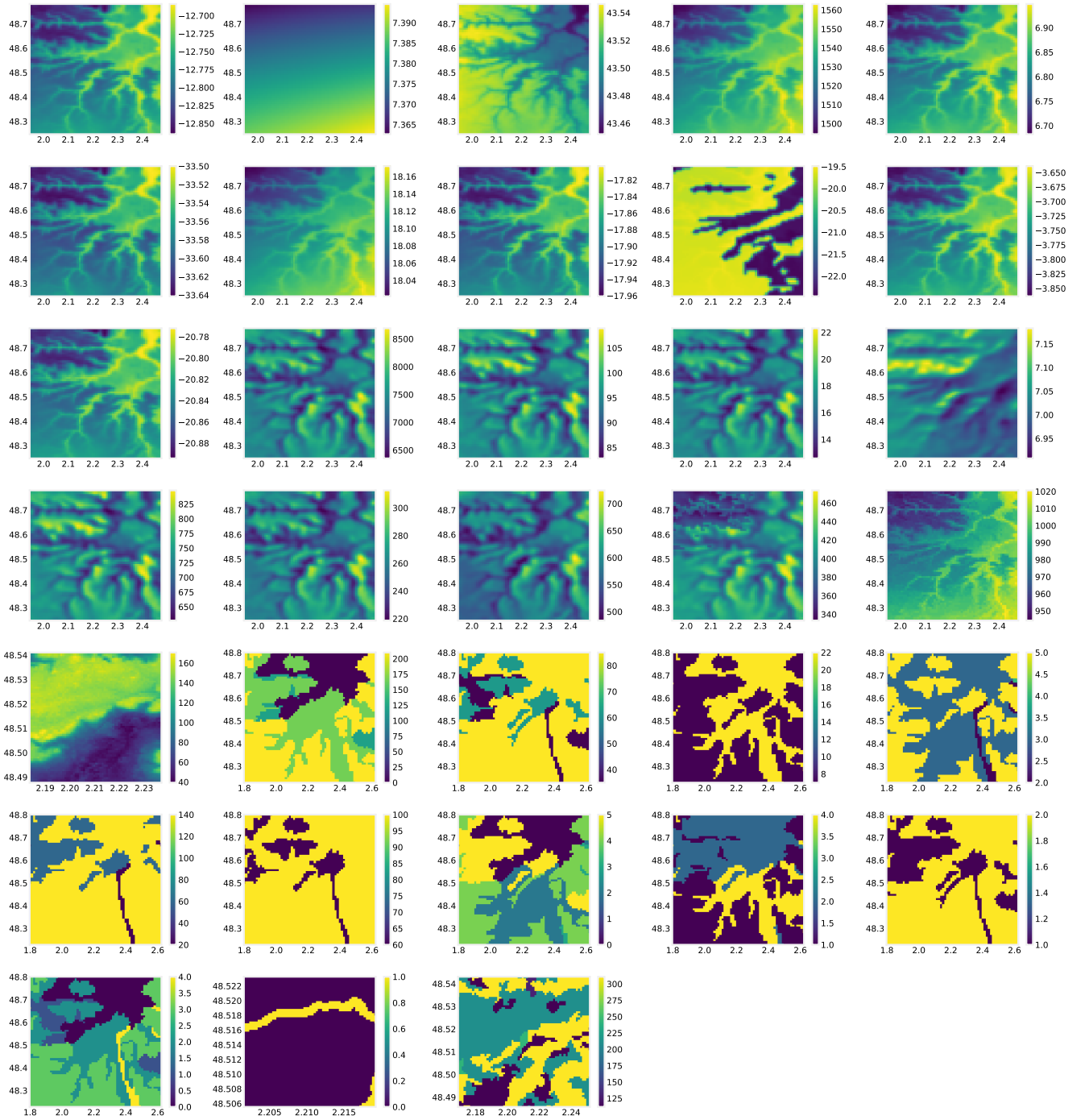


図 6 環境情報の例

として扱う。

Top30 は各予測のうち、上位 30 位までに正解クラスが含まれている予測を正解として扱う。そのため、予想 1 位でなくとも予測上位に正解が含まれていれば精度が高くなる。

MRR は式 9 により求められ、各予測のうち正解クラスの順位によって値が変化し、正解クラスが上位に存在するほど高精度となる。

$$\text{MRR} = \frac{1}{|Q|} \sum_{i=1}^{|Q|} \frac{1}{\text{rank}_i} \quad (9)$$

ここで、 $\text{rank}_{class}$  はクラス  $class$  の予測順位を表す。

#### 4.3 実験条件

実際にはベースラインに加え、各提案手法を適用したモデルを用いる。ベースラインには ResNet18 を用いた図 2 に示すモデルを用いる。

提案手法として、ベースラインにラベルマップを適用したものを提案手法 1 とする。次に、CutMix を拡張したものを提案手法 2 とする。また、Cross Entropy Loss を拡張したものを使用したものを提案手法 3 とし、さらに通常の Cross Entropy Loss を加えたものを提案手法 3' とする。最後に、すべての手法を組み合わせたものを提案手法 4 とする。

最適化には Adam を使用し、学習率は 0.001 とした。学習はバッチサイズ 256 で 100 エポック行った。

4.4 結果

これまでに述べた条件で実験を行った。各手法による実験の結果を表 4 に示す。

表 4 各提案手法で実施した実験結果

Method	Top1	Top30	MRR
ベースライン	0.1496	0.5157	0.2236
提案 1 (ラベルマップ)	<b>0.2144</b>	0.5372	<b>0.2882</b>
提案 2 (CutMix の拡張)	0.1769	0.5332	0.2578
提案 3 (Loss の拡張)	0.0607	<b>0.5643</b>	0.1462
提案 3' (Loss の拡張)	0.2086	0.5432	0.2846
提案 4 (1 + 2 + 3')	0.2051	0.5420	0.2815

結果より、提案手法 3 を除いた提案手法はベースラインの精度を上回ることを確認した。特に、提案手法 1 は、全手法の中で最も高い精度を記録した。

5 考察

表 4 から、ベースラインと提案手法の精度を比較して考察を行う。

提案手法 1 では、ラベルマップを作成したことによりベースラインの精度を上回ることを確認した。このことから、擬似的なマルチラベルの作成という戦略が本タスクに有効であると考えられる。

提案手法 2 では、CutMix を拡張することにより精度の向上が確認できた。特に、学習時の Loss の増減が安定することを確認したため、初期パラメータの影響を受け難くなったと考えられる。しかしながら、学習の助けにはなるものの、直接的に精度の向上につながらない点には再検討の余地がある。

提案手法 3 では、Top30 のみがベースラインと比べて大きく上回った。これは、ベースラインと比較して提案手法 3 では正解クラスが予測の上位に来るように学習できたためと思われる。ただし、その他の指標がベースラインよりも大きく下回っているため、Top30 による評価のみを目的とする場合でしか機能しない。

一方、提案手法 3' では Top30 が提案手法 3 から僅かに低下したが、すべての指標がベースラインを大きく上回った。これは、提案した Loss が他の指標を犠牲に Top30 のみを改善したのに対し、その弱点を Cross Entropy Loss と組み合わせることで補うことができたためであると考えられる。特に提案手法 3 では学習が安定せず、序盤の学習に失敗した場合その影響が長く残り続けることを確認した。対して、Cross Entropy Loss は学習が安定するものの、ノイズの影響を受ける可能性が存在するため、これらを組み合わせることにより互いの弱点を補い合うことができたと考えられる。

最後に、提案手法 4 では全ての手法を組み合わせた手法である。全ての指標がベースラインの精度を上回ったが、それぞれの手法を単体で適用した場合と比較して精度が想定通り向上

しなかった。これは、各提案手法がデータ拡張に類する処理を行っているため、拡張されたデータが実データから乖離したためではないかと考えられる。特に、提案手法 2 と提案手法 3、3' は学習後半に有効な手法であるため、これらを同時に学習前半から適用した影響が現れた可能性が存在する。改善案としては、学習前半と学習後半で提案手法の組み合わせを変化させることなどが考えられる。

6 おわりに

本研究では、特定の場所で最も観察される可能性が高い種の予測を行い、提案手法としてラベルマップや CutMix の拡張、Loss の拡張を提案した。各提案手法に対して実験を行った結果、提案手法の精度がベースラインを上回ることが確認できた。

問題点としては、提案手法を組み合わせた際に、想定通りの効果が得られなかったことが挙げられる。一方で、ベースラインの精度を上回ったことは確認できたため、学習時のエポックを基準に適用する提案手法をそれぞれ変化させるなどの方法により改善する可能性が存在する。また、予備実験として入力画像を縦横方向にシャッフルして学習を行った所、大幅な精度の低下が見られたが、一部の予測は正確に行っていることを確認した。そのため、画像畳み込み処理が十分に空間方向の情報を活用できていない可能性が存在し、ネットワーク構造に改善の余地があると考えられる。

今後の課題としては、生物種推定に適したネットワーク構造の検討が挙げられる。そして、既知の生物の分布モデルや種の分類体系を用いた特徴の導入が有効であると考えられるが、それには専門的な知識も必要となるため専門家の協力などが求められる。また、本研究において予測に使用する環境情報の素性は 33 種類と多く、データセットの作成に必要なコストを考慮するとより少ない素性を用いた予測方法の検討が必要であると考えられる。そして、素性には一部類似する素性が存在するため、使用する素性の組み合わせにより予測が改善される可能性が存在する。最後に、他のデータセットを用いての手法の有効性のさらなる検証も必要であると考えられる。

謝辞

本研究の一部は、科研費基盤 (B) (課題番号 17H01746) の支援を受けて遂行した。

文献

[1] Camilo Mora, Derek P. Tittensor, Sina Adl, Alastair G. B. Simpson, and Boris Worm. How many species are there on earth and in the ocean? *PLOS Biology*, 9(8):1–8, 08 2011.

[2] Christophe Botella, Maximilien Servajean, Pierre Bonnet, and Alexis Joly. Overview of geolifeclef 2019: Plant species prediction using environment and animal occurrences. 2019.

[3] F John Odling-Smee, Kevin N Laland, and Marcus W Feldman. Niche construction. *The American Naturalist*, 147(4):641–648, 1996.

[4] Arnaud Pocheville. The ecological niche: history and recent controversies. In *Handbook of evolutionary thinking in the*

*sciences*, pages 547–586. Springer, 2015.

- [5] Christophe Botella, Pierre Bonnet, François Munoz, Pascal Monestiez, and Alexis Joly. Overview of geolifeclef 2018: location-based species recommendation. 2018.
- [6] Benjamin Deneu, Maximilien Servajean, Christophe Botella, and Alexis Joly. Location-based species recommendation using co-occurrences and environment-geolifeclef 2018 challenge. 2018.
- [7] Leo Breiman. Random forests. *Machine learning*, 45(1):5–32, 2001.
- [8] Stefan Taubert, Max Mauermann, Stefan Kahl, Danny Kowanko, and Maximilian Eibl. Species prediction based on environmental variables using machine learning techniques. In *CLEF (Working Notes)*, 2018.
- [9] Tianqi Chen and Carlos Guestrin. Xgboost: A scalable tree boosting system. In *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, KDD '16, page 785794, New York, NY, USA, 2016. Association for Computing Machinery.
- [10] Jerome H. Friedman. Greedy function approximation: A gradient boosting machine. *Ann. Statist.*, 29(5):1189–1232, 10 2001.
- [11] Chandrabose Aravindan and SM Jaisakthi. Convolutional long short-term memory neural networks for hierarchical species prediction. 2018.
- [12] F.A. Gers. Learning to forget: continual prediction with lstm. *IET Conference Proceedings*, pages 850–855(5), January 1999.
- [13] Saining Xie, Ross Girshick, Piotr Dollár, Zhuowen Tu, and Kaiming He. Aggregated residual transformations for deep neural networks. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 1492–1500, 2017.
- [14] Mathilde Negri, Maximilien Servajean, Benjamin Deneu, and Alexis Joly. Location-based plant species prediction using a cnn model trained on several kingdoms - best method of geolifeclef 2019 challenge. In *CLEF*, 2019.
- [15] Christian Szegedy, Vincent Vanhoucke, Sergey Ioffe, Jonathon Shlens, and Zbigniew Wojna. Rethinking the inception architecture for computer vision. *CoRR*, abs/1512.00567, 2015.
- [16] Sara Si-Moussi, Mickael Hedde, and Wilfried Thuiller. Plant recommendation using environment and biotic associations. In *CLEF*, 2019.
- [17] Nanda H Krishna, Praveen Kumar, Ram Kaushik, Mirualini Palaniappan, Chandrabose Aravindan, and S.M.Jai Sakthi. Species recommendation using machine learning - geolifeclef 2019. 07 2019.
- [18] Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. Deep residual learning for image recognition. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 770–778, 2016.
- [19] Geoffrey E. Hinton, Nitish Srivastava, Alex Krizhevsky, Ilya Sutskever, and Ruslan Salakhutdinov. Improving neural networks by preventing co-adaptation of feature detectors. *ArXiv*, abs/1207.0580, 2012.
- [20] Yoshua Bengio, Patrice Y. Simard, and Paolo Frasconi. Learning long-term dependencies with gradient descent is difficult. *IEEE transactions on neural networks*, 5 2:157–66, 1994.
- [21] Ross B. Girshick, Jeff Donahue, Trevor Darrell, and Jitendra Malik. Rich feature hierarchies for accurate object detection and semantic segmentation. *2014 IEEE Conference on Computer Vision and Pattern Recognition*, pages 580–587, 2013.
- [22] Karen Simonyan and Andrew Zisserman. Very deep convolutional networks for large-scale image recognition. *arXiv 1409.1556*, 09 2014.
- [23] Siang Thye Hang and Masaki Aono. Residual network with delayed max pooling for very large scale plant identification. In *CLEF (Working Notes)*, 2017.
- [24] Sangdoo Yun, Dongyoon Han, Seong Joon Oh, Sanghyuk Chun, Junsuk Choe, and Youngjoon Yoo. Cutmix: Regularization strategy to train strong classifiers with localizable features. *CoRR*, abs/1905.04899, 2019.
- [25] Hongyi Zhang, Moustapha Cissé, Yann Dauphin, and David Lopez-Paz. mixup: Beyond empirical risk minimization. *ArXiv*, abs/1710.09412, 2017.
- [26] Terrance Devries and Graham W. Taylor. Improved regularization of convolutional neural networks with cutout. *ArXiv*, abs/1708.04552, 2017.