

機械学習によるゴリラの個体識別の評価

2020SC056 中辻友孝

指導教員：河野浩之

1 個体識別の重要性

2023年現在哺乳類は約27%が絶滅危惧種に指定されていて、国際自然保護連合(IUCN)では近い将来約49%が絶滅危惧種になる[1]と予想されている。ゴリラは哺乳類の中でも個体数が極端に少なく、ゴリラの保護団体である特定非営利活動法人ゴリラのはなうた[2]によると、哺乳類の中で一番絶滅してしまう可能性が高い。ゴリラの種類は主に4種類に分けられていて、2020年時点で一番少ない種類だと約300頭、一番多い種類でも約35000頭なので少ないことが分かる。ゴリラの個体数の減少の原因としては、密猟やゴリラ自身のストレスが考えられる。ゴリラの保護、観察をするためには一頭一頭の体調や、食べ物を食べる事が出来ているのかを知る必要がある。

動物の個体を管理する方法として色のついたタグを腕につける方法や、マイクロチップ等があるがその方法だと動物にストレスを与えてしまう。動物にストレスを与えることなく保護、観察できる方法として遠隔カメラやカメラトラップがある。この方法だとカメラを設置するだけでデータを得ることが出来るので、データを集めるのに労力はかからない。しかしこの方法だと画像の枚数が多くなってしまふので、人の手で個体の識別を行うには膨大な時間と手間がかかってしまう。日本の動物園では個体数が少ないので個体識別は必要ないが、国立公園や野生環境では個体数が多いので個体識別が必要になってくる。国立公園や野生環境では木などの障害物で顔が隠れてしまっていることが多いので、より汎用的なモデルが要求される。

2 個体識別に関する先行研究

顔の画像を用いて個体識別を行った先行研究を表1に示す。Andrewら[3]は京都大学によって設立された長期

表1 個体識別に関する先行研究

研究者	手法
Andrewら[3]	CNNを用いてモデルを作成
Ottoら[4]	YOLOv3を用いてモデルを作成

チンパンジーフィールドサイトの西アフリカ、ギニア南東部のポソウ森林で収集された新生児から57歳のチンパンジー23頭の個体識別を行った。フィールドサイトの遠隔カメラから収集した動画データをフレームごとに切り取り、KLT法を用いて切り取った画像のグループ化を行いラベルを付けた画像を使用し、CNNを用いてモデルを作成した。全ての画像を用いると精度が92%、正面の画像のみだと95%と高い精度を得ることが出来た。

Ottoら[4]は動物園の生後6ヵ月から37歳のゴリラ7

頭の個体識別を行った。動物園の遠隔カメラからフレームごとに切り取り顔の境界ボックスで注釈が付けられた画像を使用し、YOLOv3を用いて個体識別のモデルを作成した。モデルを訓練する際には事前学習済みのオープンソースを利用した。精度は全体で92.1%と高い精度を得ることが出来た。Ayana以外の個体は91.3% 98.5%の精度で認識できていたが、Ayanaは74.9%の精度でしか判別が出来ていなかった。Ayanaのデータセットの画像はライトの逆光や、遊具のタイヤなどで顔が隠れてしまっている画像が多かったことが原因と先行研究[4]では考えられていた。

3 個体識別手法の提案

3.1 CNNでの個体識別のモデルの構築

Ottoらの研究[4]ではYOLOv3でモデルを作成していたが、本研究で構築するモデルは入力層、3つの畳み込み層、3つのプーリング層、平坦化層、全結合層、出力層の全10層のCNNのモデルを使用する。畳み込み層では活性化関数のReLUを用いて畳み込みを行い、最終的な出力層ではsoftmax関数を使い7つのクラスに関する確率分布を出力する。本研究ではOttoらの研究[4]の精度を上回ることと、より汎用的なモデルを作成することを目標にする。

3.2 各層の説明

入力層では全ての画像のピクセルを 150×150 に揃えるという作業を行う。畳み込み層の次元数は32, 64, 128, となっていて特徴量を抽出する際は 3×3 のカーネルを使用し特徴量を抽出する。畳み込み層の次元数を増やすことにより、より高度な特徴を抽出する。畳み込み層で特徴量を抽出した後は次のプーリング層でダウンサンプリングを行い、サイズを 2×2 の半分のサイズに圧縮する。ダウンサンプリングを行うことにより、局所的な特徴量を保持したまま、計算量を減らすことが出来る。この作業を3回繰り返す。畳み込み層、プーリング層で特徴量の抽出を行った後は、全結合層に入力するために平坦化層で3Dデータを1次元のベクトルに変換する作業を行う。全結合層では前の層のすべてのニューロン(ユニット数は512)と接続され、最後の出力層では7つのクラスに関する正解率(accuracy)を出力する。

4 本研究での課題

本研究ではOttoらの研究[4]と同じ画像のデータセットを使用して、CNNを用いて個体識別のモデルを作成する。Ottoらの研究[4]でYOLOv3を使用した際全体で90%の精度を出力するのに学習回数が40000回以上、個

体名 Ayana の精度は 74.9 % と他の個体に比べて 20 % 近く低い精度になっていた。本研究では YOLOv3 よりも学習回数が少なく、全体的に高い精度で汎用性が高いモデルを作成することを目標にする。

5 CNN による個体識別の評価

本研究では、Otto らの研究 [4] と同様に BristolGorillas2020 で提供されている、動物園のゴリラ 7 頭の画像を用いる。トレーニングセットは 4303 枚、テストセットは 1105 枚で実験を行う。データセットには破損している画像も含まれていたため Otto らの研究 [4] より画像の枚数は 30 枚ほど少なくなっている。Training accuracy (Training acc) は訓練時にどのくらいロスなく学習できているのかを表す値であり、Validation accuracy (Validation acc) はテストセットで最終的にどのくらいの精度で判別できているのかを表す。Validation acc とともに Training acc を出力することにより、モデルがどの程度学習出来ているのかを知ることが出来るので、学習がうまく行くことが出来ているか確認しながら実験を進めていく。実験結果のグラフは図 1、値は表 3 に示す。

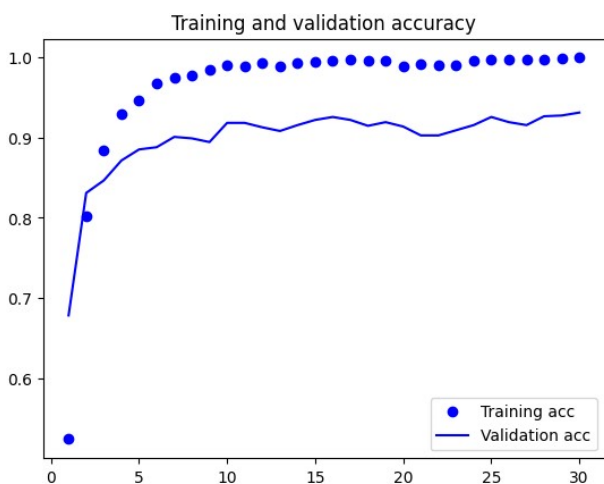


図 1 本研究の実験結果のグラフ (横軸は epoch 数を表す)

個体別の精度では Otto らの先行研究 [4] で一番精度が高かった個体の Jock と、一番精度が低かった個体の Ayana のみ表 3 に示した。

表 2 本研究の実験結果

	Otto らの研究 [4][%]	本研究 [%]
Training acc	97.3	99.2
Validation acc	92.1	94.3
Jock	98.5	98.7
Ayana	74.9	89.2

6 考察

全体の精度の平均や、個体別の精度でも Otto らの研究 [4] を上回る結果が得られた。Validation accuracy や Jock

の精度はあまり変わらないが、Ayana の精度は Otto らの研究 [4] を大きく上回る結果が得られた。本研究の CNN では大きく上回る結果を得ることが出来たので、Otto らの研究 [4] よりも汎用的なモデルを作成することが出来たといえる。学習回数でも Otto らの研究 [4] では最大で 45000 回学習を行っていて、精度が 90 % を超え始めたのが学習回数 40000 回からだったが、本研究では学習回数が 10 回を超えたあたりで精度が 90 % を超え始めたので、処理速度も CNN の方がいいといえる。

7 まとめ

本研究で使用した CNN のモデルは YOLOv3 の精度よりも全体で 2.5 % 上回る結果を得ることが出来た。個体別の精度でも他の個体の精度を下げることなく、Ayana の精度を YOLOv3 よりも 15.7 % 上回る結果を得ることが出来たので、CNN のモデルの方が汎用性が高いといえる。処理速度の面でも CNN が YOLOv3 を上回っている。顔が隠れてしまっている画像が多かった個体の Ayana も高い精度を得ることが出来たので、ある程度顔が隠れていても個体の識別が可能なが分かった。本研究の課題としていた YOLOv3 よりも必要な学習回数が少なく、汎用性が高いモデルを作成することができた。

8 参考文献

- [1] IUCN 絶滅危惧種レッドリスト, “The iucn red list of threatened species,” 2022/02, <https://www.iucnredlist.org/ja>, (2023-8-20).
- [2] 特定非営利活動法人ゴリラのはなうた, “Humming of Gorillas,” 2021, <https://www.hummingofgorillas.org/lp/>, (2023-10-25).
- [3] Daniel Schofield, Andrew Zisserman, Misato Hayashi, Tetsuro Matsuzawa, Dora Biro, Susana Carvalho: “Chimpanzee face recognition from videos in the wild using deep learning,” 科学の進歩, vol.5, no.9, pp.1-9, Sep 2019.
- [4] Otto Brookes, Tilo Burghardt: “A Dataset and Application for Facial Recognition of Individual Gorillas in Zoo Environments,” pp.1-4, May 2021.
- [5] Alexander Loos, Marin Pfister, Laura Aporius: *IDENTIFICATION OF GREAT APES USING FACE RECOGNITION* In 19th European Signal Processing Conference (EUSIPCO), pp.1-5, Aug 2011
- [6] 岡 幸秀, 岡留 剛「少数画像からの生物個体識別手法に関する研究 —ゴリラの個体識別を題材として—」2020 年度 情報処理学会関西支部 支部大会 講演論文集, vol.2020, pp.1-4, Sep 2020.