

CT 画像と深層学習を用いた骨格標本上の形態学的変異の可視化と発見

森田 堯 (大阪大学)

概要

骨格標本は形態学・進化生物学・考古学における重要なデータであり、性・生物種・生活環境等に起因する骨格上の変異を発見することは大きな学術的意義を持つ。本研究の究極目的は、骨格標本を分析するニューラルネットワークの「着目点」を可視化し、これまで人間の研究者が見落としていた未知の形態学的変異を発見することである。その実現のための基盤技術として、2022年度は骨格標本の効果的観測視点の学習手法の開発を実施した。微分可能レンダリング技術を活用し、標本分類に最適な仮想カメラの位置と向き(焦点)の機械学習を試行した。形態学における教科書の問題であるニホンザルの下顎の雌雄分類課題で学習を実行したところ、164検体という限られた訓練データから、各カメラで73.43-99.94%の分類正答率を達成し、さらに形態学的に妥当な観測視点を自動抽出した。

1 共同研究に関する情報

1.1 共同研究を実施した拠点名

- 大阪大学 サイバーメディアセンター

1.2 課題分野

- データ科学・データ利活用課題分野

1.3 共同研究分野 (HPCI 資源利用課題のみ)

- 超大規模データ処理系応用分野

1.4 参加研究者の役割分担

- 森田堯 (大阪大学産業科学研究所・代表者)
役割: 研究総括。機械学習技術の開発。
- 西村剛 (京都大学・副代表者)
役割: 機械学習結果の形態学的考察。

2 研究の目的と意義

骨格標本は形態学・進化生物学・考古学における重要なデータであり、性・生物種・生活環境等に起因する骨格上の変異を発見することは大きな学術的意義を持つ。既存の骨格標本分析手法では、専門家が骨格標本上で同定した特徴的領域やランドマーク点に基づいて変異を検出してきた(図1左)。しかし、人間の目視に依存した分析は常に客観性・再現性の担保の問題が付きまとい、新たな形態学的発見をもたらす独創的な視点との両立は容易ではない。

本研究の究極目的は骨格標本を分析するニューラルネットワークの「着目点」を可視化し、これまで人間の研究者が見落としていた未知の形態学的変異を発見することである(図1右)。専門家が持つ知見(変異を含む箇所的位置情報

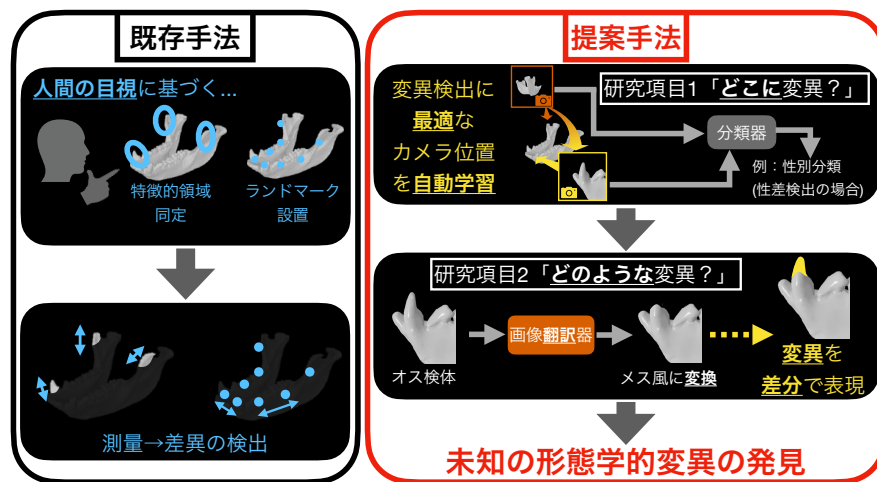


図 1: 既存の分析手法 (左) と提案手法 (右)。

等) をニューラルネットワークの訓練に用いず、ニューラルネットワーク自身が独自の根拠に基づいて標本を分析するように設計する。究極目的の達成に向け、最新の画像処理技術を応用した骨格標本分析用ニューラルネットワークを構築し、既に形態学的変異に関する知見が蓄積されているデータを用いて当該手法の性能評価を行うことが本研究における至近目的である。

2022 年度は、2 つの研究項目を設定した (図 1 右上)。第 1 に、骨格標本の効果的観測視点の学習手法の開発を主要目的に設定した。研究代表者らの先行研究では事前設定した複数視点を用いたモデルを採用したが、この手法では不要な視点の画像処理にも計算資源 (GPU メモリ) を割くため探求可能なモデルの選択肢が狭まる。この問題を解決すべく、本研究では微分可能レンダリング技術 (Kato et al., 2018, *CVPR* 等) を応用した視点最適化学習を試みた。

さらに、画像翻訳技術を用いた骨格標本分析手法の探求を第 2 の目的として設定した (図 1 右中)。第 1 の目的である観測視点の最適化学習では、骨格標本上の「どこに」形態学的変異が存在するのかを人間が理解可能な形で可視化することが可能となるが、それが「どのような」形態

学的変異なのかについては可視化できない。後者の問いに答えるためには、形態学的変異をシミュレートし、人工的に骨格標本を生成することが有効である。このシミュレーション・標本生成を、実データからの画像翻訳によって実現する (Zhu et al., 2017, *ICCV*)。変換前後の標本を比較することで、形態学的変異の質的可視化が可能となる。

3 当拠点公募型研究として実施した意義

本研究は形態学的新発見をもたらす機械学習技術を開発する学際的研究である。また、研究体制も、機械学習を専門とする代表・森田と形態学の専門家である西村が協力する、学際性の高いものである。したがって、本研究は、情報科学分野と応用分野の研究者による学際共同研究の推進を目指す学際大規模情報基盤共同利用・共同研究拠点の公募課題趣旨に合致している。

また、本研究では 3D データを用いた画像処理の機械学習を行うため、大規模な GPU メモリを必要とする。今年度使用した大阪大学の SQUID システムでは、40GB の GPU を 8 枚搭載した計算ノードが提供されており、本研究実

施に最適であった。

4 前年度までに得られた研究成果の概要

本研究に先立ち、研究代表者らは畳み込みニューラルネットワーク (CNN, 具体的には 18 層 ResNet; He et al., 2016, *CVPR*) を用いたマカク属の下顎の骨格標本をデータとして用いた性別分類 (図 2)、及び分類時の CNN の着目領域の可視化を行った (Morita et al., 2022, *AJBA*)。当該手法は未知のデータに対して 90% 以上の分類精度を達成し、また可視化された着目領域は専門家のそれと同様に犬歯や筋突起・下顎枝部分に分布していた (図 3)。以上のことから機械学習を用いた形態解析が一定の実用性を秘めていることが分かった。一方で、CNN の着目領域可視化手法は解像度を著しく落とした形でしか得られず、将来的に細部の分析が必要となった際に適用できないことが予見される。(例えば、標準的な解像度である 224×224 ピクセルを入力した場合、可視化される着目領域は 7×7 ピクセルの出力をアップサンプルして得ることになる。) また離散的分類に特化した手法であるため、年齢や各種生理学的指標等の連続量の変化が骨格標本上でどのように反映されるかを可視化するには新たな手法の開発が必要となる。本研究が探求する新手法は、分析解像度・連続量分析双方の問題を解決する可能性を秘めている。

2021 年度の萌芽課題では、本研究項目 2 にあたる敵対的骨格標本形質変換を、複数視点及びランダム視点を用いた形で取り組んだ。当初取り組んだ複数視点学習では申請時に見積もったよりも GPU メモリの消費量が大きく、大阪大学 SQUID システムを用いてもメモリ不足となった。この結果を受けて単一のランダム視点を用いた手法に切り替えたところ、今度は骨格

標本上の特徴領域が移らない視点が多く選択されてしまい、形質変換がなされない恒等写像に陥ってしまった (変換なしでも分類器を騙すことができる状態になってしまった)。これらの失敗を受け、敵対的形質変換以前に観測視点の学習技術開発 (本研究項目 2) が急務であると判断し、2022 年度主要課題に設定した。

5 今年度の研究成果の詳細

骨格標本上の「どこに」形態学的変異が存在するのかの自動同定 (研究項目 1) について、微分可能レンダリング技術 (Kato et al., 2018, *CVPR* 等) を応用し、骨格標本上の変異検出に最適な仮想カメラの位置・焦点の自動学習技術を開発した。3 次元空間上設置された 3 台の仮想カメラの位置と向き (焦点) の座標をそれぞれ学習可能パラメータとして設定し、このパラメータを基に骨格標本のレンダリング画像を得た。このレンダリング画像を入力として、CNN にニホンザルの下顎の雌雄分類課題を解かせ、分類学習の一環として、カメラの位置と向きを最適化した。

本研究のデータには、京都大学 Digital Morphology Museum で公開されているニホンザルの下顎の骨格標本の Computed Tomography (CT) を使用した。CT データをポリゴンメッシュデータに変換し、このメッシュデータに対して微分可能レンダリング技術による視点学習を行った。使用可能な標本が全 178 検体 (内訓練用データ 164 検体) と、深層学習用データとしては極めて少ないため、過学習防止のためのデータ拡張が必要であった。研究代表者らの先行研究 (Morita et al., 2022, *AJBA*) では、仮想カメラの位置に ± 5 度程度の水平摂動を与えることで過学習を防いだが、本研究はカメラの位置の学習を目的とするため、当該手法は利用できない。そのため、標本の表面色や背景色 (グレースケールでの白さ)、標本への光の当て方

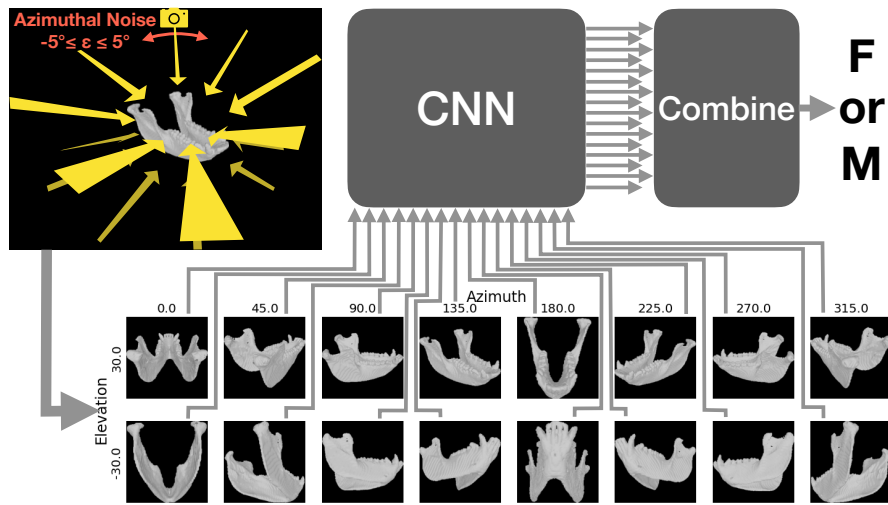


図 2: 先行研究で用いた多視点モデル (Morita et al., 2022, *AJBA*)。

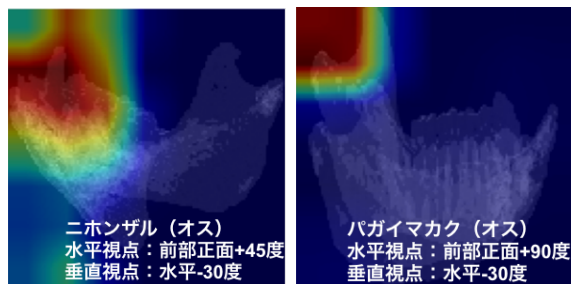


図 3: 図 2 のモデルによる雌雄分類時の着目領域の可視化 (Morita et al., 2022, *AJBA*)。

等、レンダリング時のパラメータをランダムに設定するデータ拡張を試行した。また、カメラ位置と焦点を結ぶ直線に対する直行面での回転は、理論的には分類精度に影響しないはずであるため、この回転をランダムに行うデータ拡張も実施した。

骨格標本上には複数箇所の形態学的変異が存在する可能性があるため、本研究では 3 台の仮想カメラの同時学習を試みた。しかし、3 台のカメラに自由な学習を許すと、着目領域が重複する可能性があるため、他のカメラの方が近い距離にある標本部分をマスクする実装を行った。また、CNN はレンダリング画像の角の情報も捉えることができるため、単に分類精度の向上

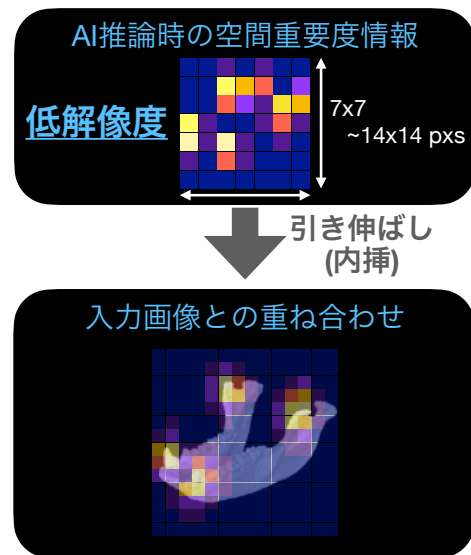


図 4: Class Activation Map (CAM; Zhou et al., 2016, *CVPR* 他) を用いる CNN 着目領域可視化手法の概要。圧縮された空間特徴量を内挿して元の画像サイズに引き伸ばすため、可視化結果はぼやける。

を学習した際に、特徴的領域が中心に来るとは限らない。そのため、着目領域可視化で用いられる Class Activation Map(CAM; Zhou et al., 2016, *CVPR*) を算出し、中央部分が周辺よりも大きくなるよう制御する損失関数を加えた (CNN には影響せず、カメラのみに影響するよ

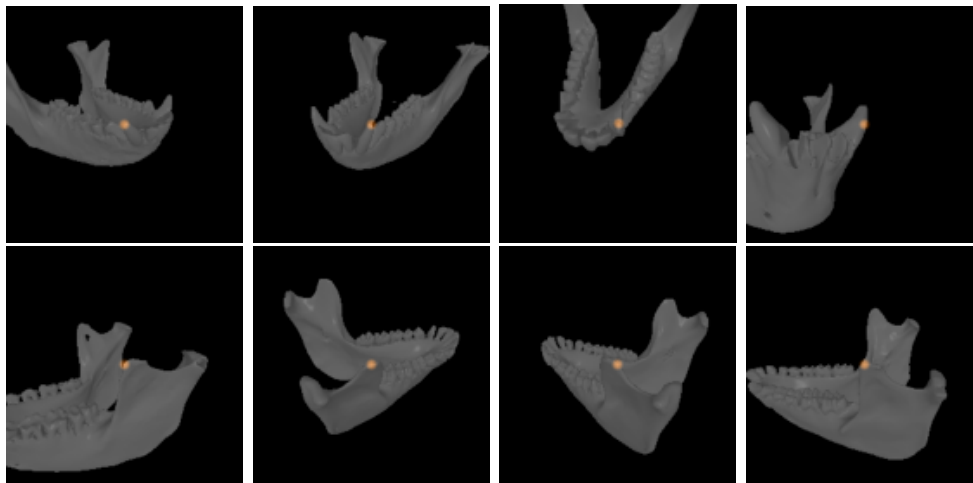


図 5: 微分可能レンダリングによる仮想カメラの位置・向きの学習結果。橙色で示した点はカメラの焦点 (= 画像中心点)。上段: 犬歯への着目結果の例。下段: 筋突起への着目結果の例。

うに設計)。

機械学習プログラムの実装には、PyTorch (深層学習基盤; Paszke et al., 2017, *NIPS*)、及び PyTorch3D (微分可能レンダリング; Ravi et al., 2020, *arXiv*; Jognson et al., 2020, *SIG-GRAPH*) を採用した。

全 178 検体中 14 検体をテストデータに、残り 164 検体を訓練データに使用する cross validation を 10 回実施した結果、各カメラあたり 73.43–99.94% の正答率を達成した。さらに、カメラの位置・向きの学習結果は、犬歯や筋突起、下顎枝といった、形態学的に妥当な結果が多く見られた (図 5)。このことから、微分可能レンダリングによる骨格標本の観測視点学習は、形態学特徴領域の同定に有用性を持つことが示唆された。

一方で、3 台のカメラの相補学習のために導入した標本マスクングの結果、テスト時にマスクを外した際の分類精度が約 45–75% 程度まで著しく低下する結果を得た。この結果を受け、マスク部分を完全に不可視とせず、ガウシアンぼかしで曖昧にする手法を現在試行中である。ぼかし具合をランダムに変更することで、完全な

検体を観測した場合に近い画像の学習が可能となり、テスト時の頑健性向上が期待される。

6 今年度の進捗状況と今後の展望

今年度の進捗状況

研究項目 1 として設定した骨格標本上の「どこに」形態学的変異が存在するのかの自動同定については、技術開発が概ね完了し、先述の非近傍標本面のぼかし処理学習を試行しつつ、論文発表に向けて考察をまとめている状況である。進捗状況は、90% である。

一方で、研究項目 2 として設定した「どのような」変異が存在するのかの可視化については、2022 年度後期 (11 月以降) の実施を計画していたが、電気代高騰の影響で後期の利用可能計算資源が実質 1/3 になった影響で、当該研究項目の実施に至る前に配分された計算資源を使い切ってしまった。そのため、研究項目 2 はコード開発のみの実施となり、進捗状況は 10% である。

今後の展望

2022 年度実施に至らなかった研究項目 2 (「どのような」変異が存在するのかの可視化) について、2023 年度学際大規模情報基盤共同利用・共

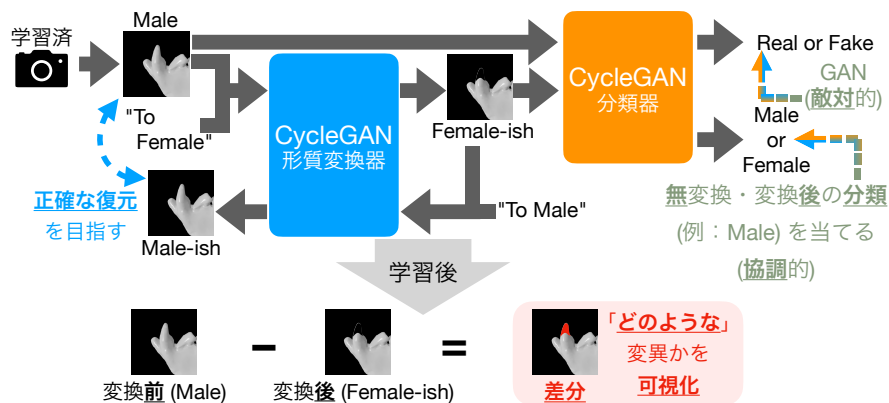


図 6: CycleGAN による骨格標本分析。

同研究拠点公募研究の継続課題として採択され、画像翻訳技術を応用した骨格標本分析技術の開発を進めている。

具体的な画像翻訳技術には、CycleGAN を採用する方針で実装を進めている (Zhu et al., 2017, ICCV)。当該手法は、入力となる骨格標本の形質を変換するモジュール (図内青色) と、変換前後の標本の形質及び変換の有無を推定するモジュール (図内橙色) からなる ANN を訓練する。形質変換モジュールは入力標本の分析対象形質を任意のカテゴリに変換する (図ではオス → メス)。このとき、(1) 変換後の標本画像は無変換画像と区別不可能となるほど現実的なものとなり、且つ (2) 個体差等の分析対象形質に無関係な要素は保持していることが望ましい。(1) の学習目標達成には敵対的学習 (GAN) の枠組みを用いる。(2) の学習目標については、変換後の画像を再変換し (図ではメス → オス)、元の入力の復元を目指すことで実装する。

7 研究業績一覧 (発表予定も含む)

学術論文 (査読あり)

該当なし。

国際会議プロシーディングス (査読あり)

- T. Morita, R. O. Tachibana, K. Okanoya, H. Koda. “On the relation between context dependency and vocabulary in human language and birdsong”, Proceedings of the Joint Conference on Language Evolution 2022, pp. 517–524, 5 September, 2022.

国際会議発表 (査読なし)

該当なし。

国内会議発表 (査読なし)

- “CT 画像と深層学習を用いた骨格標本上の形態学的変異の可視化と発見”, 第 14 回 JHPCN シンポジウム, 東京, 7 July, 2022.
- “Towards biological discoveries through the eyes of AI”, 室蘭工業大学異分野研究セミナー, 室蘭, 25 May, 2022.

公開したライブラリ等

該当なし。

その他 (特許, プレス発表, 著書等)

- “CT 画像と深層学習を用いた骨格標本上の形態学的変異の可視化と発見”, サイバーメディア HPC ジャーナル, 12, pp. 33–36, September, 2022.

- “動物音声認識のための教師なし機械学習”,
生産と技術, 74 (3), pp. 45-47, June, 2022.