

# 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（中部大学）

## 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。一方で、画像分類器に基づく既存の着目点可視化手法は分析解像度が低く、微細な変異の検出に不向きであるという問題を抱えている。この問題を解決すべく、本研究は骨格標本の形態カテゴリを変換する機械学習技術の開発に取り組んだ。当初、画像変換技術の応用を行う計画であったが、学術的意義から3次元データを直接編集する技術開発へと方針転換し、最終的に骨格標本表面の点群データを分析対象とした。特に形態カテゴリ変換の基盤となる事前学習を考案・実装したが、その精度には改善の必要があり、今後の研究プロジェクトへの課題を残した。

## 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

### 研究計画全体の目的

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。

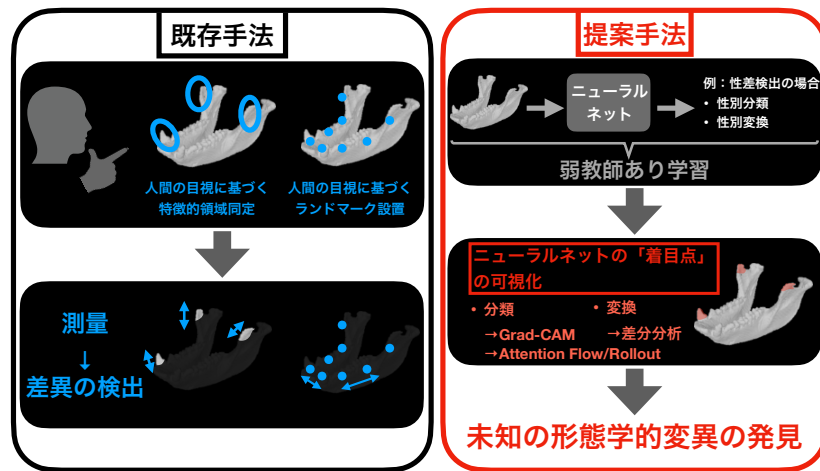


図1 既存の分析手法 (左) と提案手法 (右)。

### 今年度の目的

2023年度は、画像変換技術の活用による、高解像度な形態学的変異の検出技術開発を計画した。一般的な画像処理の推論根拠可視化手法は、圧縮された空間重要度情報を元の画像サイズに引き伸ばす低解像度な手法であり、骨格標本上の詳細な変異を同定できない (例: Grad-CAM; Selvaraju et al., 2017, *ICCV*; 図2)。そこで、本研究では、骨格標本の自然な形態学的変異を画像変換技術によってシミュレートし、変換前後の差分による詳細な変異の可視化を目指した。

### 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。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。これは、分野横断・融合的学際研究を推進する JHPCN 公募型共同研究の趣旨に合致する。

また、本研究では3Dデータを扱う機械学習を行うため、大規模なGPUメモリを必要とす

る。JHPCN 公募型共同研究で提供される大阪大学の SQUID システムは、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

前年度は、微分可能レンダリング技術 (Kato et al., 2018, *CVPR* 等) を活用し、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 (図1右上)。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。

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

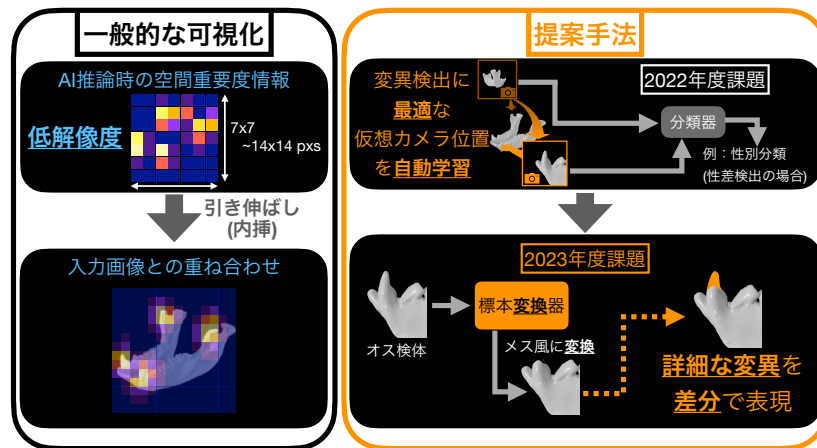


図2 一般的な推論根拠可視化手法(左)と本研究の開発手法(右)。

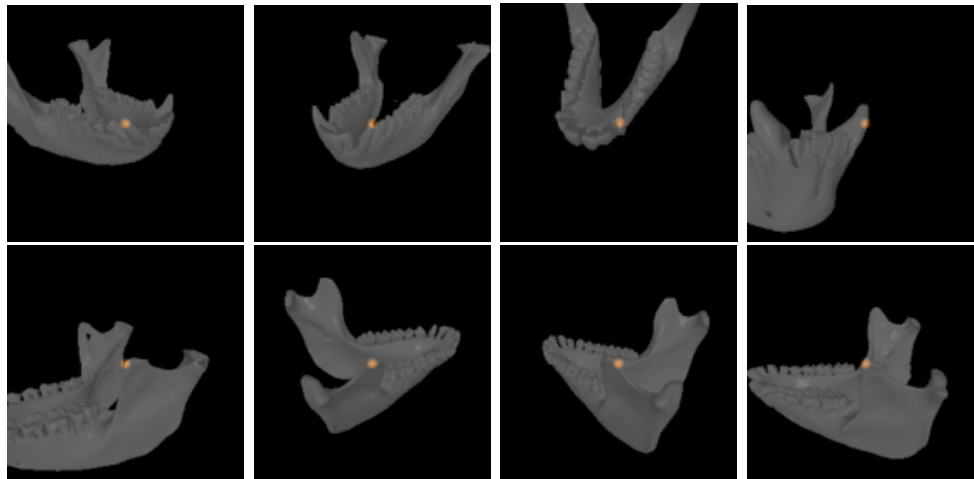


図3 微分可能レンダリングによる仮想カメラの位置・向き学習結果。橙色で示した点はカメラの焦点(=画像中心点)。上段：犬歯への着目結果の例。下段：筋突起への着目結果の例。

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 $\pm 5$ 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・向き学習の結果は、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 3)。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キングの結果、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。

## 5 今年度の研究成果の詳細

研究提案時、本研究では、昨年度開発技術から得られる 3D ポリゴン標本データの 2D レンダリング画像に対して画像変換を行う計画であった。学習を 2 段階に分割する計画は、主に機械学習訓練時の GPU メモリ負荷の軽減が目的であったが、基盤となる画像変換技術の提案論文 (Zhu et al., 2017, *ICCV*) がバッチサイ

ズを 1 に設定しているため、3D 標本を直接編集する End-to-End な変換は、計算資源の観点において不可能ではない。本研究に関する昨年度の評価でも 3D データへの対応の重要性が指摘されていたため、研究計画を見直し、3D データの変換技術の開発へと方針を転換した。

上半期 (中間報告時点まで) は、ポリゴンの各頂点に対する変形差分を計算するニューラルネットワークを探索した。提案モデルを白紙状態から訓練することは困難であるため、標本変形の基礎となるような変形による事前学習を試みた。具体的には、標本の一部にせん断写像等の基礎的な変形を行い、変形に用いた情報 (せん断行列と適用範囲; ファインチューニング時には形態学的変異カテゴリの埋込に置換予定だった) から変形後の標本を回帰推測するモデルを検証した。

しかしながら、ポリゴン (頂点) 変形モデルは、事前学習時点での回帰精度が悪かったため、下半期 (中間報告以降) では変換対象をポリゴンデータから標本表面の点群へと変更した。深層ボリュームレンダリング技術 (Jun & Nichol, 2023, *arXiv* 他) を参考に、点群データの形状を標本表面への「投射写像」として表現するモデルを設計した。入力データ (点群) を「標本表面から外れた点を標本表面上に投射する写像」という動的な形で表現し、形態カテゴリの変形操作を組み込むことで、実標本表面上の点群を変形する写像として扱う狙いであった。また、「写像」という形状表現を採用することで、大規模な点群データをコンパクトに表現・変形することが可能になるという利点もある。

実装では、まず事前学習として、3 次元ランダム点群を標本表面上に投射する多層パーセプトロン (MLP) の生成を訓練した。モデルは、Attention 機構を用いて実標本上の点群を読み

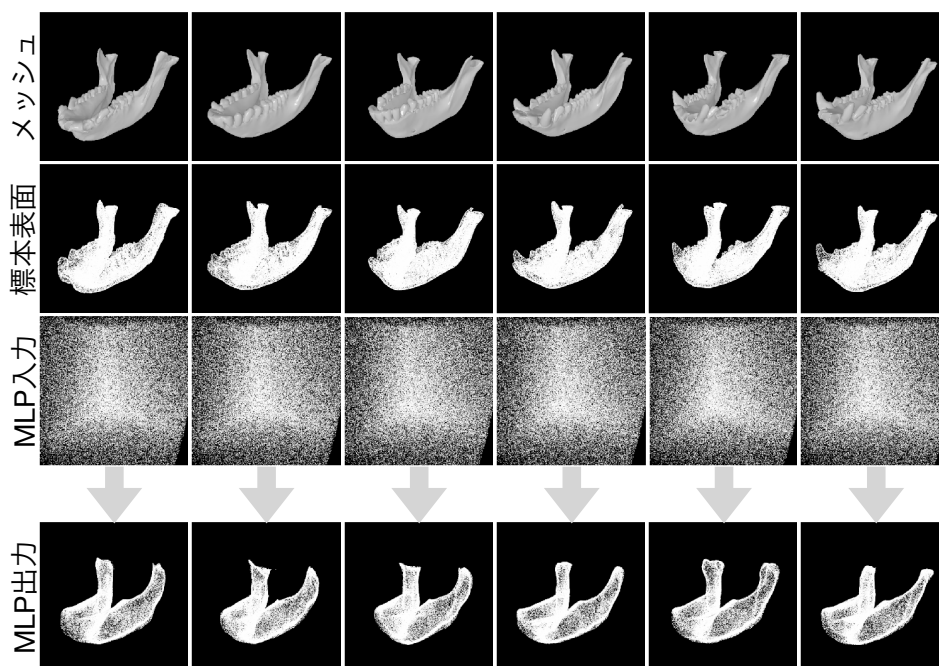


図4 投射事前学習の結果。1 段目：3D メッシュデータのレンダリング画像。2 段目：標本表面からサンプルした点群 (10 万点)。3 段目：MLP への入力。立法空間からサンプルした点群 (10 万点)。4 段目：MLP 出力。

取り、各 MLP 層の重みを出力する。そして、標本を含む立法空間内のランダム座標を MLP に入力し、当該座標から標本上への投射点を MLP の出力目標として回帰学習を行った。また、この回帰学習と同時に、点群データの形態カテゴリ分類 (本開発では雌雄分類) を訓練し、後の形態変換用目的関数の事前学習とした。

事前学習の結果、ランダム座標を標本の大きな概形へと移す写像が得られた (図 43 段目から 4 段目)。しかしながら、歯や顎関節等の末端部分詳細な形状は表現できておらず、最終的な目標である形態変換の事前学習としては精度に課題が残った。

一方で、同時に事前学習した標本の雌雄分類精度は高く、学習時未使用の標本の表面上 16,000 点のデータから 92.85% (13/14 検体) の精度での弁別を達成した。したがって、変異評価に必要な情報は点群データに十分含有されて

いることが確認された。

## 6 今年度の進捗状況と今後の展望

### 6.1 今年度の進捗状況

モデル設計に関する研究計画を申請時から変更したが、「形態学的変異をシミュレートする」という最終目的に対する達成率は 40% と評価する。標本形状の情報を投射写像という動的表現でエンコードする方針には手応えを感じているものの、事前学習時点での投射精度が不十分であるため、ファインチューニングへの移行前に事前学習の改善が必要である。

中間報告以降の開発が遅延したのは、2023 年末から 2024 年始めにかけて研究代表者が異動し、十分な研究時間を確保できなかったことが大きな要因である。加えて、年度末に近づくにつれてスーパーコンピュータシステムの混雑が激化し、ジョブ実行まで 1 週間以上の待機が

求められることもあったため、十分な探求を行うことができなかった。

## 6.2 今後の展望

今後は 2023 年度の方針を継続し、点群変形モデルの開発に取り組む。まずは事前学習における投射精度向上を目標とし、特に MLP 入力のサンプリング方法を改善する。現在、MLP への入力立法空間からの一様サンプル (図 43 段目) と、標本表面からの一様サンプル (図 42 段目; この場合の投射は恒等写像) を採用しているが、この学習では標本の概形しか表現できておらず、歯や顎関節といった末端部分の情報が欠損している。これらの微細情報のエンコードを促進すべく、立法空間をボクセルに区切り、投射精度の低いボクセルからの重点的なサンプリングを行う等の工夫によって、改善に挑む。

## 7 研究業績一覧 (発表予定も含む)

### 国内会議発表 (査読なし)

- 森田堯. ‘動物音声研究のための機械学習’, 第 57 回関西合同音声ゼミ, 2023 年 7 月 8 日, 大阪府吹田市.
- 森田堯. ‘動物行動分析のための教師なし機械学習’, 東京工業大学-基礎生物学研究所-生理学研究所-中部大学合同マッチングワークショップ「生命と情報の新たなる融和: 超階層生物学と AI・数理」, 2024 年 2 月 21 日, 愛知県岡崎市.