

須恵器3Dデータの深層学習クラスター解析による型式・年代分類基準の検証

専門分野

人文情報学

キーワード

考古資料 須恵器3Dデータ 深層学習
疑似ラベル自己学習型クラスター解析

研究目的・概要

クラスター解析はデータサイエンス手法の一つで、これまでヒトがさまざまな情報から行ってきた主観的な要素が入る分類ではなく、純粋に数値処理だけで未知のクラスター（似たもの同士のグループのこと）を検知する技術である。本研究では考古資料（発掘された遺物）である出土土器の3Dデータを使い、クラスタリングによる未知のグループの疑似ラベルをループでモデル（処理プログラム）の途中に戻し、教師あり畳み込み学習（数理解析手法のひとつ）とクラスタリングを繰り返すことにより、専門家がこれまで分類したことのない基準でクラスターを創出できる独創的な解析モデルによる「ヒトの判断基準の妥当性の検証」を目的とする。

東京国立博物館所蔵6世紀の完形・略完形（欠けた部分が皆無か、ほぼない完全に近い状態のもの）須恵器坏（すえきつき・朝鮮渡来のロク口技術で成形、高温で焼かれた蓋付きの容器）、蓋49点、身59点の計108点を1mmのスキャン精度の光学スキャナーで計測し解像度 128^3 の3D-RGBA（RGBは色情報、Aは透明度）ボクセルデータ（2Dであるピクセルに奥行き情報を加えたもの）に変換し（図1、ボクセル単位は約1mmの立方体）、須恵器資料ごとに平面上に須恵器を置いた場合の垂直方向を軸として 30° ずつ回転し、データ拡張（水増し）を行った。

色情報RGBを含まない形状だけの3D-Aデータ解析によると、まずデータ拡張によるクラス（グループ）の離散が発生していないことを主成分分析で確認した。すなわちデータ拡張により作成された同じクラスのデータは全て同一クラスターに含まれることの確認である。さらに正規化（データを0~1の数値に置き換えること）されたデータをエルボー法により最適なクラスター数は6である結果を得た。

須恵器編年（年代ごとの形状変化）を大系化した中村浩氏の型式分類とクラスター解析により導出された6クラスターの混同行列（セルの数値が大きいほど適合度が高い）を図2に示した。図2の横方向は中村浩氏による型式分類、縦方向はクラスター群を表す。クラスターID3には専門家分類型式II-4、特にII-5が多く集中し、同様にクラスターID5には専門家分類型式II-1とII-2が集中する傾向が明らかとなった。一方その他のクラスターでは専門家分類型式が離散し、クラスターと専門家分類の大きな乖離がみられた。

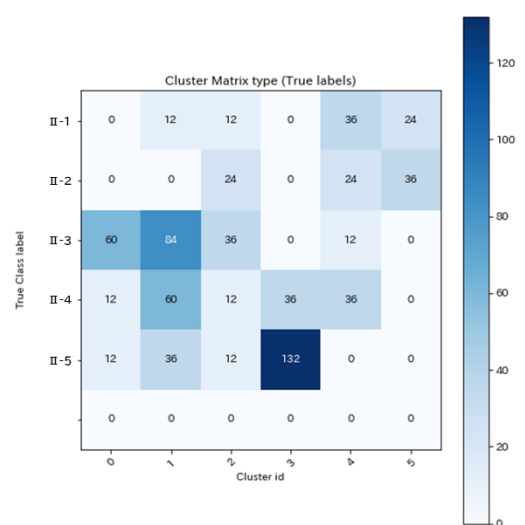
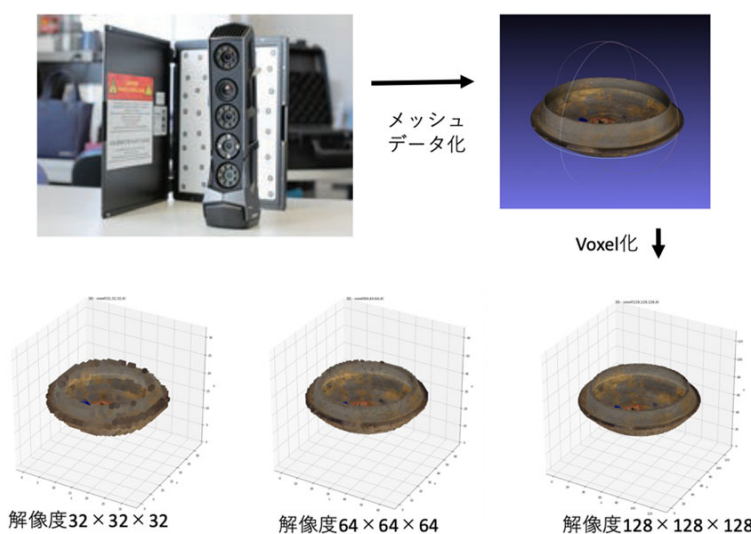


図1 光学スキャナーによる須恵器容器部分の3Dデータ作成およびボクセル化

図2 専門家分類型式とクラスターの混同行列



経営情報学部 経営学科

藤田 晴啓 教授

担当科目：AI・データサイエンス入門、機械学習

HP

https://www.nuis.ac.jp/teacher_fuzita/

Researchmap

https://researchmap.jp/haru_fujita