



# 深層学習による物体検出のパラメータから 算出した類似度を用いた博物館における展示手法

An exhibition method using similarity calculated  
with parameters from object detection using deep learning for museums

小嶺愛紀菜<sup>1)</sup>, 赤嶺有平<sup>2)</sup>, 根路銘もえ子<sup>3)</sup>

Akina KOMINE, Yuhei AKAMINE, and Moeko NEROME

- 1) 琉球大学 理工学研究科 (〒 903-0213 沖縄県中頭郡西原町字千原 1 番地, kggakmu@acs.u-ryukyuu.ac.jp)  
2) 琉球大学 工学部 (〒 903-0213 沖縄県中頭郡西原町字千原 1 番地, kggakmu@acs.u-ryukyuu.ac.jp)  
3) 沖縄国際大学 経済学部 (〒 901-2701 沖縄県宜野湾市宜野湾二丁目 6 番 1 号, genchr@okiu.ac.jp)

概要: 筆者らは拡張現実技術や深層学習を用いた展示手法によって来館者の理解をより深めることや興味関心の向上を目的としたデジタルガイダンスシステムを開発している。博物館には多数の展示物があり、来館者にとってそれら展示物同士の形状や模様等の細かな類似点、共通点を一度の見学で把握することは難しい。そこで深層学習モデルを活用し来館者が見ている展示物と類似した展示物を提示することで、展示物同士の関連性を見だしやすい手法を提案する。

キーワード: デジタルガイダンスシステム, 深層学習, 博物館

## 1. はじめに

近年、博物館等の展示施設においてモバイルデバイスを用いたデジタルガイダンスシステムが注目されている。筆者らは博物館におけるデジタルガイダンスシステムの開発を行っており、拡張現実技術や深層学習を用いた展示手法によって来館者の理解をより深めることや興味関心の向上を目的としている。博物館には多数の展示物があり、用途は同じでも描かれている模様や大きさ、形状が若干異なるものなど様々な展示物が存在する。しかし、来館者にとってそれら展示物同士の形状や模様などの細かな類似点、共通点を一度の見学で把握することは難しい。そこで深層学習モデルを活用して来館者が現在見ている展示物と類似した展示物を提示し、展示物同士の関連性を見だしやすい展示手法を提案する。このように関連性のある展示物をデジタルガイダンスシステムにより関連付けて表示し、来館者の展示物に対する理解がより深まるアプリケーションを開発する。

## 2. 提案手法

まず、展示物一つ一つを様々な方向から撮影した画像を用意する。そして、撮影した写真の背景をマスク処理によって削除する。背景を削除することによって、深層学習モデルが展示物そのものに着目するようにしている。画像ごとのマスクは、展示物と背景の 2 クラスを分類するセマンティックセグメンテーションモデルを学習させ、そのモデルを用いて自動生成する。背景が削除された画像に対して、それらを学習データとして分類またはクラスタリングを行う。そ

して、それぞれ画像ごとにモデルの中間層または最終層の出力を取得し、特徴ベクトルとする。次に、特徴ベクトル同士の距離を算出して距離が近いものから順に類似した展示物として提示する。

## 3. 実験方法

### 3.1 写真の準備

はじめに、展示物一つ一つを様々な方向から撮影した画像を用意する。今回の実験では、16 種類の展示物を撮影した。撮影した展示物の一部を図 1 に示す。

### 3.2 セマンティックセグメンテーションによる背景削除

続いて、セマンティックセグメンテーションを行い、撮影した画像の背景を削除する。セマンティックセグメンテーションは、展示物と背景の 2 クラス分類を行う。今回の実験では、学習データ 60 枚とテストデータ 15 枚を用意し、アノテーションを行った。アーキテクチャは FPN[1]、エン



図 1: 展示物

コーダーは resnext50[2] を用いて学習を行った。結果として、テストデータに対する IoU score(0.5) は 0.9402 となった。テストデータによる分類結果の一部を図 2 に示す。

次に、セマンティックセグメンテーションによって取得したマスクを用いて、元の画像から背景を削除する。背景削除後の画像を図 3 に示す。学習データに含まれている展示物が他の展示物画像の背景に映り込んでいると上手くマスクが生成されず、背景を削除しても残ってしまうことが多い。しかし、学習データに含まれている展示物が背景に映り込んでいない場合は、ほとんど画像で背景の削除が行えている。

### 3.3 分類またはクラスタリング

背景を削除した画像を学習データとして、分類またはクラスタリングを行う。今回の実験では、分類を行った場合とクラスタリングを行った場合の両方の実験を行った。

分類モデルには、resnet18[4] を用いた。学習時には、回転・シフト・拡大縮小・明るさ・コントラスト・色相などの Augmentation を行った。

クラスタリングでは、IIC(Invariant Information Clustering)[3] を用いて実験を行った。

### 3.4 類似度の算出

それぞれの画像をモデルに入力しモデルの中間層の出力を取得する。今回の実験では、分類モデルに関してはモデルの出力に近い全結合層から、クラスタリングでは最終層から特徴ベクトルを取得した。そして、取得した出力を特

徴ベクトルとして、ユークリッド距離やマンハッタン距離またはコサイン類似度を算出する。

## 4. 実験結果

分類モデルを用いた場合とクラスタリングを用いた場合の両方の実験結果を示す。

### 4.1 分類モデル

はじめに、分類モデルから得られた特徴ベクトルでの結果を示す。分類モデルの後ろから二番目の全結合層の出力を取得し、ユークリッド距離を算出した。そして、距離が近い順に類似した展示物として列挙している。その結果が図 4 である。図 4 の一番左に表示されているものが入力画像で、その右に類似していると思われる展示物が表示されている。1 番に表示されているものは最も特徴ベクトルの距離に近い画像である。

### 4.2 クラスタリング

クラスタリングの結果を図 5 に示す。散布図に割り当てられている番号に対応する展示物は図 6 の通りである。

## 5. 考察

本稿では、筆者らが似ていると感じた展示物群を「グループ」としている。分類モデルを用いた結果である図 4 を見ると、同じグループに属するものが表示されやすくなっている。図 4 に載っている同じグループに属する展示物には、三種類の模様が異なる酒器や二種類の木地などがある。しかし、これらが入力された場合に同じグループのものが上位に挙がるとは限らず、人間が見て全く異なる展示物が上位に表示される場合も多くあった。また、この方法では入力画像の撮影した位置や角度の違いによって、類似展示物の結果が大きく変わってしまい、全体的にかなり不安定な結果だった。

図 5 を見るとクラスタリング結果では展示物 1, 2, 7, 8 と展示物 5, 6 が近い位置にあることが分かる。また皿の展示物である 4, 3 も近くに位置する。展示物 13 はほぼ独立している結果となった。展示物 1, 7, 8 と展示物 5, 6 は模様や材質は異なるが同じグループに属するものであるため、同一グループに属するサンプルの特徴が比較的近くに集まっている。

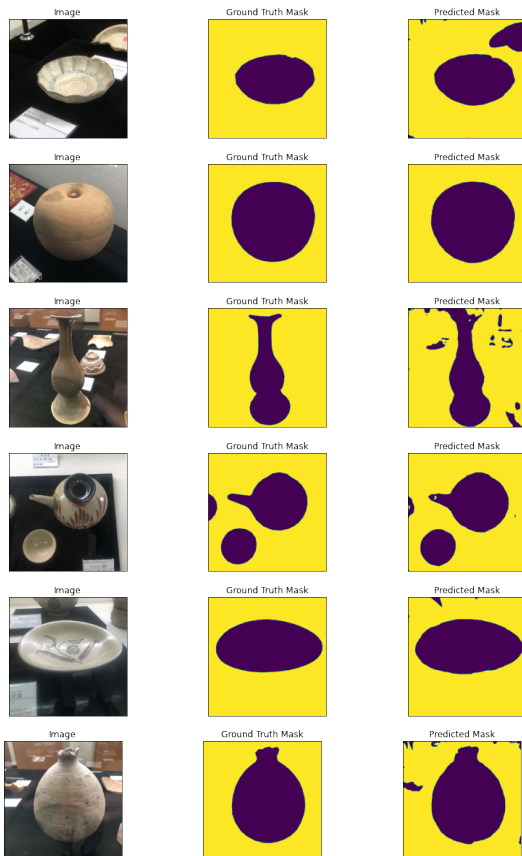


図 2: セマンティックセグメンテーションのテスト結果

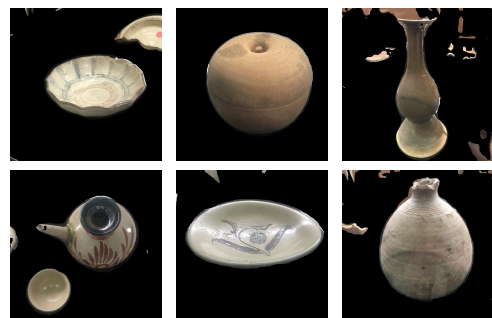


図 3: 背景削除後の画像

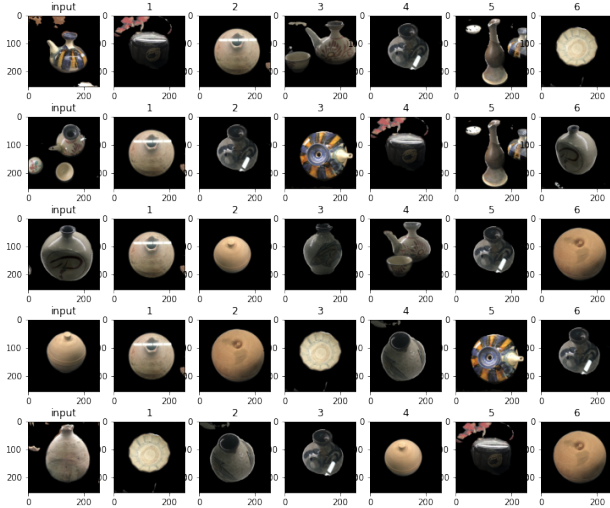


図 4: 分類モデルの中間層出力による類似した展示物の提示

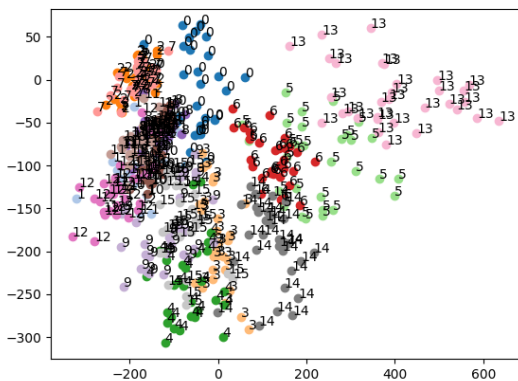


図 5: IIC の中間層出力による類似した展示物の提示

## 6. まとめと今後の課題

本研究では、分類モデルやクラスタリングなど深層学習を用いて類似した展示物を提示するための実験を行った。結果として、同じグループに属するものは特徴ベクトルが近い位置にあるということが分かった。しかし、人間から見て全く似ていない展示物も特徴ベクトルが近い位置に存在する 경우가多々ある。そのような場合、深層学習モデルがどの部分を見ているのかを可視化し、より重要な特徴を捉えられるようにすることが今後の課題である。

謝辞 本研究は JSPS 科研費 JP19K01142 の助成によるものである。

## 参考文献

- [1] Tsung-Yi Lin and Piotr Dollr and Ross Girshick and Kaiming He and Bharath Hariharan and Serge Be-



図 6: 展示物のリスト

longie: Feature Pyramid Networks for Object Detection, arXiv, 2017.

- [2] Saining Xie and Ross Girshick and Piotr Dollr and Zhuowen Tu and Kaiming He: Aggregated Residual Transformations for Deep Neural Networks, arXiv, 2017.
- [3] Xu Ji and Joo F. Henriques and Andrea Vedaldi: Invariant Information Clustering for Unsupervised Image Classification and Segmentation, arXiv, 2019.
- [4] Kaiming He and Xiangyu Zhang and Shaoqing Ren and Jian Sun: Deep Residual Learning for Image Recognition, arXiv, 2015.