

3D データを学習データとする深層学習を用いた分類器の現実空間への適用 - 人間には判別困難な差異を持つ物体の個体識別 -

Application of Classifier with Deep Learning Which Uses 3D Data as Learning Data to Real Space - Individual identification of objects which have indistinguishable different parts for people -

○越智 広樹^{*1}, 山田 悟史^{*2}
Hiroki OCHI^{*1} and Satoshi YAMADA^{*2}

*1 立命館大学大学院 理工学研究科

Graduate Student, Graduate School of Science and Engineering, Ritsumeikan University

*2 立命館大学理工学部 准教授 博士 (工学)

Associate professor, Department of Science and Engineering, Ritsumeikan University, Eng.Dr.

キーワード：深層学習；画像分類；3D モデル；現実空間

Keywords: Deep learning; image classification; 3D model; real space.

1. はじめに

近年 AI 分野では、ロボットを仮想環境上で学習させることで運用までに要していた時間を大幅に改善する事例が見られる。建築分野においても BIM 等の 3D モデルによるデータベース環境があり、同種の活用手法を実現する下地があると考えられる。特に画像分類においては、3D モデルによって学習データの収集が容易になる点や、設備機器など外見上は同一の物体を管理する上で個体の識別が必要になる点から相性がよいと考える。

そこで本研究は、将来的に BIM モデル等の仮想環境を学習データとして活用することを見据え、3D モデルから作成した学習データを用いて現実空間の物体の分類可能性を検証する。また、その違いとして人間には識別が困難な微細な差異を与えた場合の深層学習の適用可能性も合わせて検証する。

2. 研究概要

2-1. 研究の流れ

本研究では、まず、深層学習を用いて微細な差異を持つ複数の 3D モデルを撮影した画像群を分類する分類器を作成（以降、学習フェーズと表記）する。なお、学習は CNN（畳み込みニューラルネットワーク）の VGG16 を用いてファインチューニングで行う。

次に、作成した分類器に 3D モデルと同一の形態の現実空間の立体を撮影した画像群を与えて正しく分類できるか確認（以降、確認フェーズと表記）する。これは各画像が本来属するクラスに対して分類器によって予想されたクラスの度数をクロス集計表の形で算出する。この双方のフェーズにおいて CAM を用いて分類の根拠を可視化し、模様が AI の判断に寄与しているか否か確認する。

こうして 3D モデルを元に作成した AI の現実空間への適用可能性と微細な差異の分類可能性を検証する。

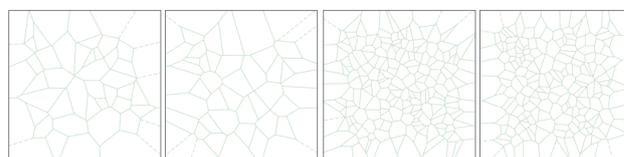


Fig.1 Voronoi diagram of 50pt (left side) and 200pt (right side)

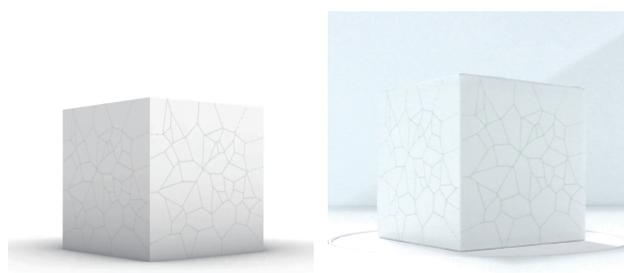


Fig2. "3D Cube" (left) and "R CUbe" (right)

2-2. 分類対象

本研究では微細な差異としてボロノイ図を分類対象の模様を用いる。その例を図 1 に示す。これはボロノイ図が母点のランダム性と分割の規則性から形成される図であり、人間から見て違いを認識できる一方で、その識別は困難だと考えたためである。なお、母点は 50 個の場合と 200 個の場合の 2 パターン（以降、50pt と 200pt と表記）があり、それぞれ 4 種類の模様を用意して各パターン 4 クラスの分類を実施する。この画像を 3D と現実上の双方で分類対象である 10 cm の立方体（以降、3D 立方体と R 立方体と表記）の全面に貼り付けたものを図 2 に示す。

2-3. データセットの作成

まず、学習フェーズのデータセットとして図 2 に示した 3D 立方体を水平方向に回転させて撮影する。その動画から静止画切り出しで各 360 枚（訓練用:288 枚、検証用:72 枚）の画像を作成する。また、確認フェーズのデータセットとして、図 2 に示した R 立方体を 3D 空間を模した撮影

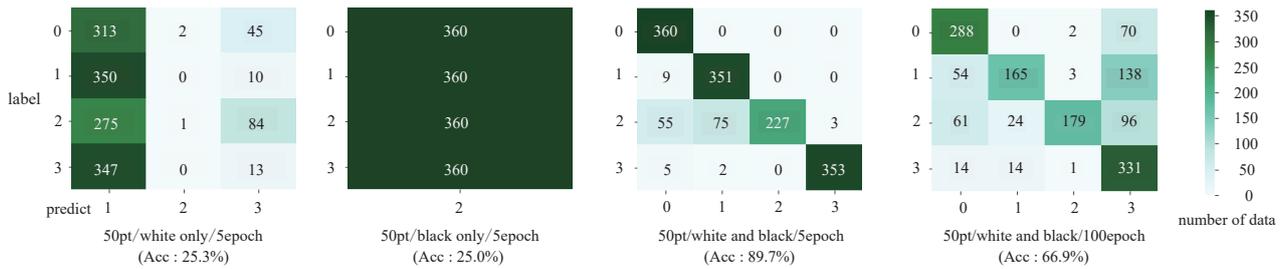


Fig3. Cross-tabulation table in a phase of verifying

ブースで回転台上で回転させて撮影する。同様に、静止画切り出しで各クラス 360 枚の画像を作成する。

3. 現状の結果と考察

研究の初期段階では図 2 のように 3D 空間の背景が白色の場合のみでデータセットを作成し、epoch 数を 100 回で学習をさせた。この際、学習フェーズの分類精度は 100% に達したが、確認フェーズの分類精度はおよそ 25% と学習をせずランダムに分類した場合と変わらない低い精度であった。この後、R 立方体の画像群を学習に用いた分類器に手ブレのある動画から出力した R 立方体の画像群を与えたところ、学習フェーズでは分類精度が 100% に達したが、確認フェーズでは分類精度が 80% 近くまで低下していた。これらのことから、学習が繰り返される中で模様ではなく背景で分類するように分類器が特化した結果、背景がフレーム単位で一致しなければ分類が不可能なほど大幅に汎化性能が低下したと考えた。

上記の問題を解決する手段として、3D 空間の背景色を黒に変更した場合の画像群を既存の学習データに加え、背景色が白と黒で混在した状態で学習を実施した。加えて、学習を epoch 数を 5 回と浅い段階で止めることで汎化性能を維持することを試みた。また、比較として背景色が黒単体で epoch 数 5 回の場合の学習と、背景色が混在で epoch 数 100 回の場合の学習も実施した。

その結果を図 3 から図 5 に示す。これらから本研究では背景色が白と黒で混在していると確認フェーズの精度が向上し、双方の背景色の場合で 3D 立方体や R 立方体の模様への着目度合も向上することがわかる。また、epoch が進むと汎化性能が低下していることも読み取れる。

4. まとめと今後の展望

本研究は、3D モデルでの学習を現実空間での分類においても適用可能であることに加え、対象のテクスチャにおける微細な差異によって、従来のコード類と異なり特定の識別子に依存しない対象の分類可能性を提示した

これは、差が微細であるため表出しても建築物の美観を損なわない手法になり得ると考えられる。今後は更に他の背景色を混在させるほか、背景の多色化など様々な条件での検証を通して、より現実空間で高い精度と汎化性能をもつ分類器の作成を試みる。

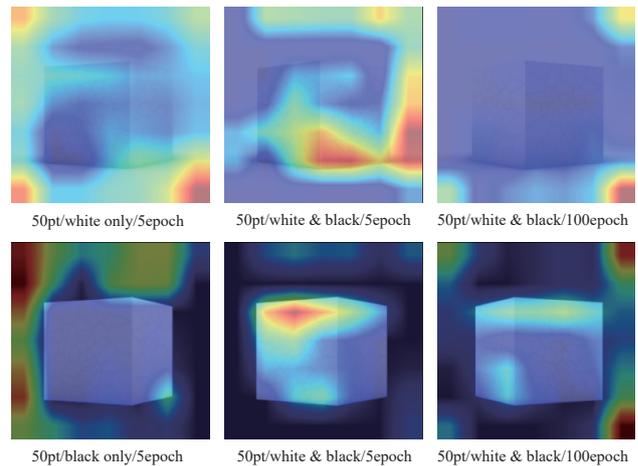


Fig4. CAM image in a phase of studying

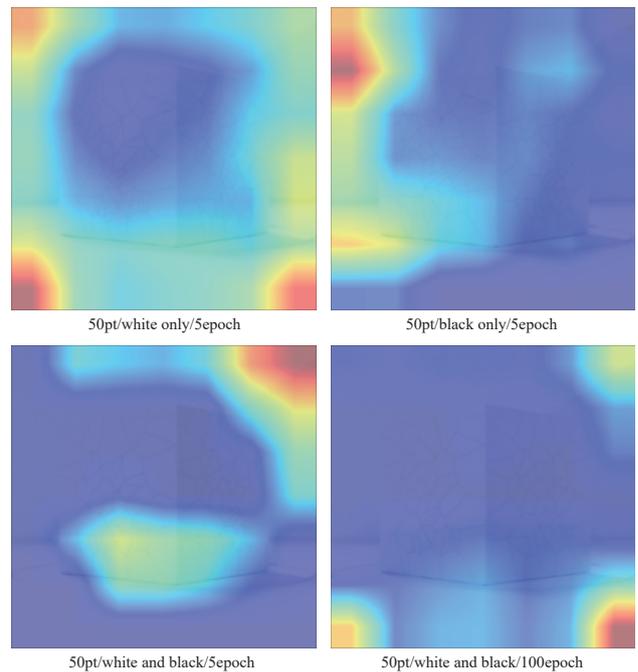


Fig5. CAM image in a phase of verifying

[参考文献]

- 産経新聞, 歩行するロボットの進化を、仮想環境でのシミュレーションが加速する [https://www.sankei.com/article/20211031-KUZDPOYRGVPKLBBDBGHE3NO4PQ2/], 2021
- Karen Simonyan, Andrew Zisserman, Very Deep Convolutional Networks for Large-Scale Image Recognition. [https://arxiv.org/pdf/1409.1556.pdf], International Conference on Learning Representations 2015, 2015