

3種の深層学習を用いた建築デザインの参照関係の可視化と近似性の解釈 Visualization of Reference Relationships and Interpretation of Approximation of Architectural Design Using Three Types of Deep Learning

○越智 広樹^{*1}, ○谷川 奈央^{*1}, 山田 悟史^{*2}
Hiroki OCHI^{*1}, Nao TANIKAWA^{*1}, Satoshi YAMADA^{*2}

^{*1} 立命館大学 理工学部 建築都市デザイン学科

Undergraduate, Department. of Architecture and Urban Design, Ritsumeikan University.

^{*2} 立命館大学 理工学部 建築都市デザイン学科 任期制講師・博士(工学)

Lect, Dept. of Architecture and Urban Design, Ritsumeikan University. Dr.Eng

キーワード：深層学習；建築デザイン；画像分類；データビジュアライゼーション

Keywords: deep learning; architectural design; image classification; data visualization.

1. はじめに

データビジュアライゼーションはグラフなどの情報の視覚的な表現のことである。また、天気予報も位置情報を用いたビジュアライゼーションの一種である。このようにデータビジュアライゼーションは日常生活において多くの場面で活用されている。その理由として、情報が持っている特徴を瞬時に多くの人に理解可能な視覚的情報として伝達可能な点が挙げられる。近年では機械学習などを利用した高度なデータビジュアライゼーションも盛んに行われている。これにより、ただの大量のデータを価値ある知見に昇華することができる。そのため、従来とは異なる特徴や視点を見出せる可能性もある。

本研究においても数値データを視覚化することで傾向や規則性を見出すことを目的としている。今回は対象の建築家の作品の参照関係の可視化(ビジュアライゼーション)を試みる。

本研究における利点が2つある。1つ目は既存の建築作品の評価に'参照された'という新たな視点を与えることができるという点である。これは深層学習により新たな参照関係を見出せる可能性があるためである。2つ目は各企業の建築デザインデータベースを形成することが可能になるという点である。企業のデータ類はクラウド化され企業内に共有されることが一般的になった。しかし、建築デザイン分野においては人に直接、または膨大な資料の中から確認をするという状況が続いている。そこで、各企業で建築デザインのアーカイブを作成し、分類して管理することができれば手早くデザイン引用をすることができると考えた。これにより日々の業務の効率化に加え、企業の実績を簡単に整理、継承することができると考えられる。

建築デザインでは用途別に分類・系統化されていることが多い。しかし、今回は深層学習によりデザイン類似別に分類することを目指す。その出力結果を可視化する

ことで価値ある知見へと昇華させることを目的とする。以下に具体的な手法を示す。

2. 研究概要

本研究では、建築家間の類似度判定の指標として深層学習で画像分類を行った際の出力に着目する。具体的には、対象建築家の作品の特徴を学習させた画像分類AIを作成し、その入力画像に対する出力である正解クラス以外のクラス(以下、他クラスと表記)への推測値を利用する。この値は入力それぞれのクラスに属し得る度合いを示しており、そのうち他クラスへの値は、入力が部分的に他クラスの特徴に近いことを示している。その値を集約することでクラス間の特徴量の近似性を表せると考え、データビジュアライズを活用して表現し分析を行った。

2.1. データセット

データセットの元となる画像はスクレイピングによってYahoo!検索(画像)[<https://search.yahoo.co.jp/image>]の検索結果に表示されるサムネイルから収集した。その後、AIによる各建築家の作品の特徴抽出が正しく行われるように、独自で設定した条件に従い画像の選定を行った(画像選定条件は図1内に記載)。この際、建築物の角度や種類は統制していないが、これは建築家の特徴を包括的に抽出することを目的としている。また、調査の対象には世界的に著名とされる3名の建築家に加え、ここ10年でプリツカー賞を受賞した10の建築家・建築家集団¹⁾を選出し、合計13のクラスを定義した。上記の手法で集めた画像群から各クラス無作為に250枚を選出し、8:2の比率で学習用と検証用に分割した。ただし、本研究は今回収集したデータセット内でのクラス間の近似性の抽出を目的としているため、汎用性の調査に用いるテストデータは不要と判断した。その後、後述の学習モデルの入力サイズ(224×224)になるように、画像の形状の比率を保つ

1	Alejandro Aravena	2016	2	Lacaton & Vassal	2021
3	Balkrishna Doshi	2018	4	Shigeru Ban	2014
5	Frank Lloyd Wright		6	Frei Otto	2015
7	Grafton Architects	2020	8	Arata Isozaki	2019
9	Toyo Ito	2013	10	Le Corbusier	
11	Mies van der Rohe		12	RCR Arquitectes	2017
13	Wang Shu	2018			

- ・カラー画像
- ・文字を含んでいない
- ・画像内の対象建築物の全体に占める割合が目測でおよそ 50% 以下
- ・写り込んでいる人物の全体に占める割合が目測でおよそ 5% 以下
- ・過度にデフォルメされていない模型や CG 画像

条件をもとに収集した画像の総枚数 :5617 枚
total number of images collected under conditions

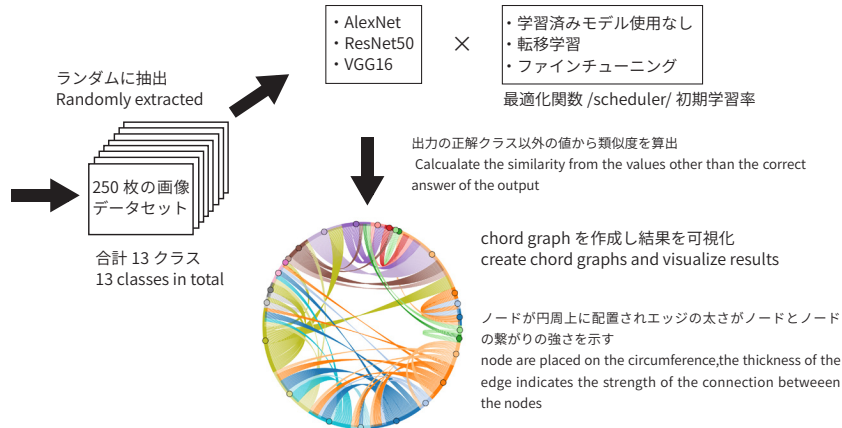


Figure 1 Overall flow of this research

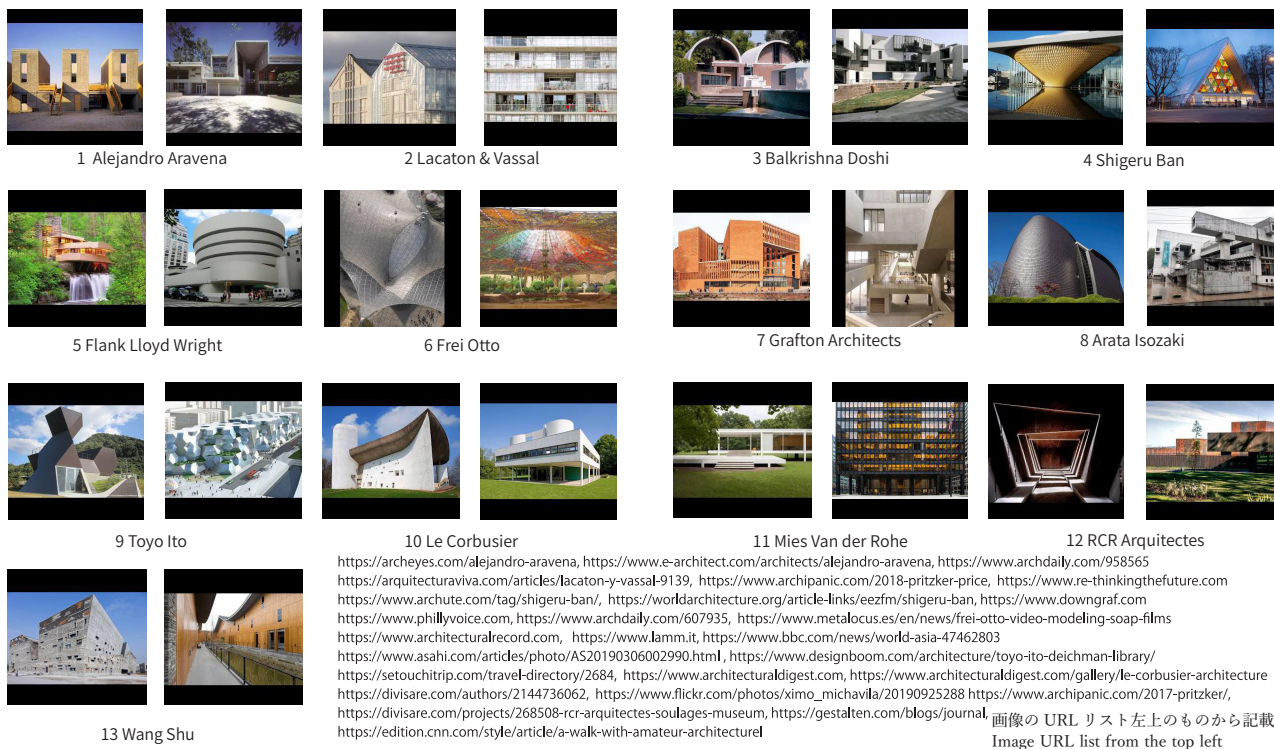


Figure 2 Image example of each class

たままを値 0 でパディングした。図 2 に編集後の画像を一部示す。また、学習前に Data Augmentation として学習用画像には左右反転処理を施した。

2.2. 学習モデル

学習には画像認識モデルの AlexNet²⁾, ResNet50³⁾, VGG16⁴⁾ を利用した。これら 3 種のモデルは全ての学習モデルが同じ結果を示すかどうかは不明であり、アンサンブル学習の観点から採用した。これらを複数の条件下で学習させた結果を比較し、それぞれ最も高い精度を示した条件における出力を後述のビジュアライゼーションに利用した。この際、分類時に対象が建築となっているかは CAM⁵⁾ を利用して確認をした。

2.3 学習条件

条件の設定は検証の過程で 2 度行った。1 度目は 2.2 の

3 種のモデルに対し、それぞれ、学習済みモデル不使用、転移学習⁶⁾、ファインチューニング⁷⁾ の 3 種類の学習方法を適用した計 9 種類の学習条件を設定した。この時 epoch 数はいずれも 100 とし、学習率などのハイパーパラメータや最適化関数、scheduler を変更させ様々な学習条件を設定した。条件の一覧は表 1 に示す。このうち各モデルで最も高い精度を示した条件の学習モデルを建築家間の近似性の抽出に利用する。

2.4 推論値の処理と解釈

2.3 で学習させたモデルを利用して入力画像に対する推論値を取得する。推論値は入力画像が各クラスに属し得る度合いを示す値である。そのため正解クラスの値を取り除いてから同一クラスごとに合算してから出力することで、入力クラスの各クラスに対する特徴量の近似性を

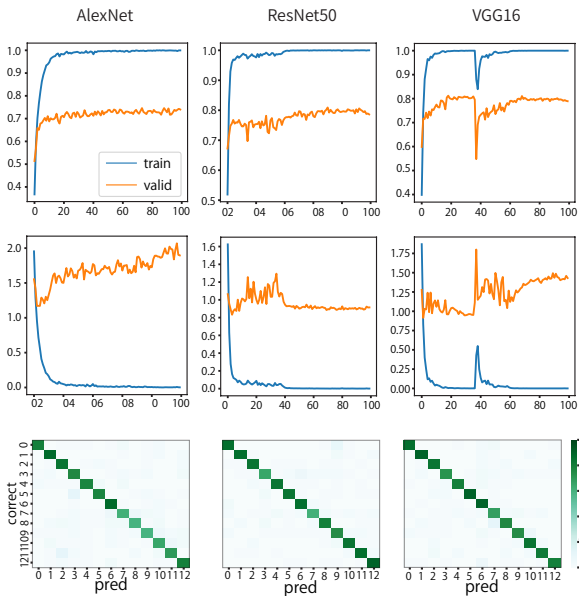


Figure 3 Learning transition and Summary table of the best pattern in each learning model

Table 1 Learning conditions and their achievements

学習タイプ	最適化関数	scheduler	初期学習率	epoch	正答率	損失	
AlexNet	学習済みモデル使用なし	SGD ^{注1)}	StepLR ^{注3)}	0.001	500	0.5030	1.9114
				0.0001		0.1461	2.5254
				0.001		0.7385	1.2430
				0.001		0.7415	1.1324
				0.0001		0.7153	1.1407
	転移学習	SGD ^{注1)}	StepLR ^{注3)}	0.00001	200	0.6184	1.2783
				0.001		0.7461	1.1704
				0.0001		0.7153	1.1814
				0.001		0.0769	2.5649
				0.0001		0.7307	1.0967
ResNet50	学習済みモデル使用なし	SGD	StepLR	0.001	100	0.1646	0.2446
				0.0001		2.4874	2.271
				0.001		0.6169	1.2982
				0.001		0.7784	0.7905
				0.0001		0.7476	0.9028
	転移学習	SGD	StepLR	0.00001	100	0.3938	2.2919
				0.001		0.7876	0.7789
				0.001		0.6215	1.7552
				0.001		0.8092	0.8324
				0.0001		0.5707	1.7278
VGG16	学習済みモデル使用なし	SGD	StepLR	0.0001	300	0.5707	1.6260
				0.001		0.5569	1.8695
				0.001		0.5538	1.6363
				0.001		0.6369	1.2251
				0.001		0.7923	0.8785
	転移学習	SGD	StepLR	0.001	150	0.7476	0.8750
				0.001		0.4246	2.1649
				0.0001		0.7892	0.8486
				0.001		0.5138	2.6077
				0.001		0.8076	0.8121
ファインチューニング	AdamW	ReduceLR	ReduceLR(patience=5) ^{注5)}	0.0001	0.8030	0.9219	
					ReduceLR(mode=max) ^{注6)}	0.8015	0.8628
					ExponentialLR ^{注7)}	0.7984	0.9286
					CosineAnnealingWarmRestarts ^{注8)}	0.8123	0.9136
					0.8030	0.9219	

示す値と見なした。この際、後述の chord graph への変換に適した形になるように出力に softmax 関数を適用して 0 から 1 の範囲に値が収まるように調整した。

2.5 ビジュアライゼーション

ネットワークグラフ "chord graph"(コードグラフ)を作成する。コードグラフは各ノードの繋がりを表現するグラフである。コードグラフ作成のために python の holoviews ライブラリー⁸⁾を使用する。今回は得点のエッジを重みにもつグラフを作成する。つまり、各クラス間の得点はコードグラフ上において各ノードに接続するエッジの太さに反映される。エッジが太くなるほど近似性が高いことを示している。

3 結果と考察

3.1 結果

表 1 に今回実施した学習条件とその成果一覧をまとめた。図 3 に各学習モデルにおいて最高精度が出力された際の学習推移、画像の分類集計結果をまとめた。AlexNet ではファインチューニング、最適化関数 SGD, scheduler:ReduceLR, 初期学習率 0.001 のパターンにおいて 0.7461 ,ResNet50 ではファインチューニング、最適化関数 AdamW,scheduler:ReduceLR, 初期学習率 0.0001 のパターンにおいて 0.8092 ,VGG16 ではファインチューニング、最適化関数 AdamW,scheduler:CosineAnnealingWarm Restarts, 初期学習率 0.0001 において 0.8123 と 7 割から 8 割ほどの精度を獲得することができた。その結果におけるコードグラフが図 4 である。

3.2 考察

図 4 で示した 3 種のコードグラフと、コードグラフ作成

に使用した 3 種のモデルからの出力データを元にアンサンブル的に考察を行う。この際、予想クラスに対する正解クラスごとの出力を予想スコアとし、その分布とスコアの合計値の分布を考察に利用する。

スコアの合計に関しては、まず Ludwig Mies van der Rohe (以下 Mies と表記) は各モデルにおいても各ノードから入るエッジが太く、特に AlexNet で顕著であった。Le Corbusier (以下 Corbusier と表記) は ResNet 時にエッジが太く、それ以外のモデルにおいても一定の太さを保持している。このようにモデル間で差はあるが、各クラスとも共通する特徴量を有することを示している。この結果は Mies と Corbusier が同年代に活動した近代建築を代表する建築家であり、多くの建築家に影響を与えたとする言説と一致する。次に、上記の 2 名と同様に近代建築の巨匠とされる Frank Lloyd Wright (以下 Wright と表記) についてはいずれのモデルにおいても目立ったエッジが確認できず、今回の対象の中では独特な特徴量を持つことを示している。これは Wright が有機的建築を設計の主軸にしていたためと考えられる。他、特質できる点として Grafton Architects は全てのグラフにおいてもエッジが太く、他クラスと共通する特徴量を持っているという発見があった。

個々のクラス内では、スコアの分布から 3 種のモデルで一貫する特徴量の近似性は表 2 のように示された。この中でも特に「Arata Isozaki - Toyo Ito」, 「Corbusier - Mies」, 「Shigeru Ban - Toyo Ito」, 「RCR Arquitectes - Mies」, 「Corbusier - Shigeru Ban」は互いにスコアが高く、特に近い特徴量を有することが示唆される。特にスコア合計時に着目した三大巨匠に関しては Corbusier と Mies の特徴量が近く、Wright は上記 2 名とは特徴量的には

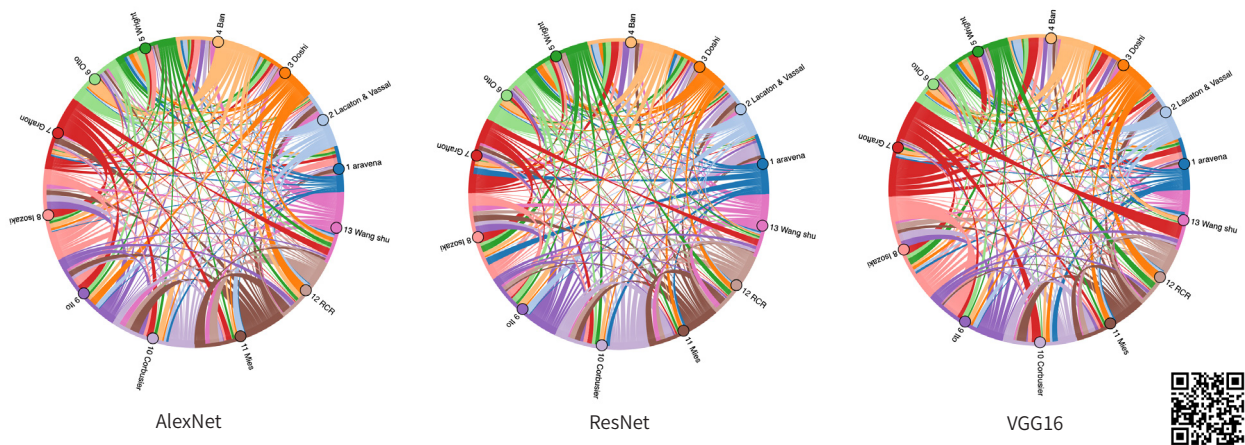


Figure 4 chord graph generated from each learning model

Table 2 correct answer class with the highest predicted score for each class

予想クラス	正解クラス
Alejandro Aravena	Le Corbusier
Lacaton & Vassal	Shigeru Ban, Toyo Ito
Balkrishna Soshi	RCR Arquitectes, Frank Lloyd Wright
SHigeru Ban	Toyo Ito, Le Corbusier
Frank Lloyd Wright	Arata Isozaki, RCR Arquitectes
Frei Otto	Toyo Ito
Grafton Architects	Arata Isozaki, Wang Shu, Le Corbusier
Arata Isozaki	Toyo Ito, Shigeru Ban
Toyo Ito	Arata Isozaki, Shigeru Ban
Le Corbusier	Mies van der Rohe, Shigeru Ban
Mies van der Rohe	Le Corbusier, RCR Architectes, Lacaton & Vassal
RCR Arquitectes	RCR Arquitectes, Mies van der Rohe
Wang Shu	Arata Isozaki

離れていることが示されている。また、既知の関係性を有する事例として Corbusier の元で働いていた Doshi が挙げられるが、スコアの分布上では形態としての特徴量の近似性は低い結果となった。

4 おわりに

本研究では、画像分類モデルにおける出力に着目し、chord graph によってビジュアライズすることで、複数の建築家間における特徴量の近似性を分析した。結果としては、一般的に知られている建築家間関係性に沿った特徴量の近似性を発見したほか、従来では見出されていない建築家間に共通する特徴量が存在することが示唆された。一方で今回はモデルの過学習を抑えることができず、無統制の画像の分類という観点においては分類精度に課題が残る結果となった。加えて、交差検証などの実施も課題としてあげられる。今後の展望としては、学習精度を向上に加え、他モデルの使用や分類器における出力のメカニズムを踏まえた分析などを通して、結果の妥当性を向上させることを考えている。

なお、今回作成したコードグラフは以下のリンクおよび、図 4 内の QR コードより確認が可能である。

[<http://satoshi-bon.jp/2021/10/01/approximation/>]

[注釈]

- 注 1) 確率的勾配降下法に基づく最適化関数
- 注 2) Adam の重みに関するものを書き換えたもの
- 注 3) 指定された epoch ごとに学習率を任意の割合で減衰させる
- 注 4) 学習の中で最小の損失が任意の回数 (デフォルトは 10) 以上更新されなかった場合、学習率を減衰させる
- 注 5) 注 4) のにおける回数を 5 に変更したもの
- 注 6) 注 4) において損失の代わりに正答率を基準にしたもの
- 注 7) (任意の値)^{***epoch} で学習率を更新するもの
- 注 8) 半周期分の余弦波を一定の倍率で波長を引き延ばし、連続させるかたちで学習率を推移させるもの

[参考文献]

- 1) The Pritzker Architecture Prize Laureates, The Pritzker Architecture Prize, The Hyatt Foundation, 2021.9.27, <https://www.pritzkerprize.com/laureates>
- 2) Alex Krizhevsky, Ilya Sutskever, Geoffrey E. Hinton: ImageNet Classification with Deep Convolutional Neural Networks, NIPS'12 Proceedings of the 25th International Conference on Neural Information Processing System - Volume 1, pp.1097-1105, 2012.12
- 3) Keiming He, Xiangyu Zhang, Shaoqing Ren, Jian Sun: Deep Residual Learning for Image Recognition, IEEE Conference on computer vision and pattern recognition, pp. 770-778, 2016
- 4) Karen Simonyan, Andrew Zisserman: Very deep convolutional networks for large-scale image recognition, arXiv 1409.1556, 2015
- 5) Ramprasaath R. Selvaraju, Michael Cogswell, Abhishek Das, Ramakrishna Vedantam, Devi Parikh, Dhruv Batra: Grad-CAM: Visual Explanations from Deep Networks via Gradient-based Localization, arXiv:1610.02391v4, 2019
- 6) Chuanqi Tan, Fuchun Sun, Tao Kong, Wenchang Zhang, Chao Yang, Chunfang Liu: A survey on Deep Transfer Learning, arXiv:1808.01974, 2018
- 7) Zhizhong Li, Derek Hoiem: Learning without Forgetting, arXiv:1606.09282, 2017
- 8) Chord, Holoviews, Anaconda Inc, 2021.9.29, <https://holoviews.org/reference/elements/bokeh/Chord.html>
- 9) 大野耕太郎, 山田悟史, 宗本晋作, 細波哲也: Deep Learning を用いた街並み画像に対する平均訪問推定意欲推定 AI の作成と検証, 日本建築学会大会学術講演梗概集, pp.119-pp.120, 2020, 9
- 10) Sasank Chilamkurthy: transfer learning for computer vision tutorial, PyTorch, 2021.9.30, https://pytorch.org/tutorials/beginner/transfer_learning_tutorial.html