

Deep Learning を用いた AI の歴史都市分野への適用可能性

Applicability to the Historical City Sector with AI Using Deep Learning

大野耕太郎¹・山田悟史²

Kotaro Ono and Satoshi Yamada

¹ 立命館大学 理工学研究科建築都市デザインコース (〒525-8577 滋賀県草津市野路東 1-1-1)
Grad. Student, Graduate School of Eng., Architecture and Urban Design Course, Ritsumeikan Univ

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) (〒525-8577 滋賀県草津市野路東 1-1-1)
Lecturer, Dept. of Architecture and Urban Design, Ritsumeikan Univ., Dr. Eng.

At present, artificial intelligence (AI) based on Deep Learning is attracting attention in various fields, and world-renowned research institutes and the world's leading information companies are focusing on R & D from a large source. Attempts to implement society have also begun, and although there are aspects of excessive expectations, the social interest is extremely high not only for researchers, but also for the possibility of changing our lives and work areas. Taking one of the advantages of Deep Learning as an example of image classification, there is a point that the characteristic amount itself can be learned so as to approach a previously prepared classification without a person specifying the characteristic amount representing the image. The results that exceeded the existing results were reported and produced a big development. However, it can be said that the application to the historical city area is still underdeveloped. So, in this research, we try the application possibility to the historical city field of AI technology using such Deep Learning.

Keywords: Deep Learning, AI, CNN, CAM, GAN

1. はじめに

現在、様々な分野で Deep Learning を基盤とする人工知能 (AI) 技術が注目を集めており、それに付随して大学や研究機関における AI 技術の開発が進んでいる。こうした技術の一部はすでに社会にも浸透しており過度な期待もあるのは否定できないが、人間の生活領域に変化を及ぼす可能性に対しての社会的な関心度は高い。契機の一つには、世界的な統計競技サイト「ILSVRC (ImageNet large scale visual recognition challenge)」での成果がある。画像をクラス分類する競技において、Deep Learning を用いた AI がそれまでの精度を大きく上回ったため Deep Learning に対する注目を集めた契機となった。Deep Learning の利点として、画像分類を例として挙げると、学習させる画像が持つ特徴量を人が指定せずに、予め用意した画像のクラス分類に近づくように特徴量自体を学習させられる教師なし学習がある。このような特徴量自体を学習する特徴を活かし、対象の持つ特徴量と及び結果との因果関係の解明が複雑であった「画像認識」「音声認識」「自然言語処理」などの分野において、既存の結果を上回る成果が報告され大きな展開を生み出した。しかしながら Deep Learning を基盤とする AI の歴史都市分野への適用はまだ未発達といえる。そこで、本研究ではこうした Deep Learning を用いた AI 技術の歴史都市分野への適用可能性を試行する。

Deep Learning がその特性を発揮しやすい機能は対象と「分類・認識・推定・予測」の因果関係が曖昧で複雑な事象の学習である。こうした学習にはニューラルネットワークと呼ばれる人間の脳内のネットワークを模した数理的モデルが用いられる。その中でも特に主流なものとして、畳み込みニューラルネットワーク (CNN) がある。本研究でも CNN を用いて、画像などに対する分類・推定を行う AI の歴史都市分野への適用可能性の検討を行う。具体的には、街並み画像から街路名を推定する AI、訪問意欲の有無とその度合を

推定する AI の開発を行った。

また、AI が画像を分類・推定する際に注目する領域を可視化する手法も提案されている。具体的な技術には Class Activation Mapping (CAM) がある。例えば、瀧澤ら¹⁾は VR 上に作られた街並みに対する印象評価を行う AI を作成し、CAM を用いた AI 評価の考察を行っている。そこでこうした CAM の技術を用いて街並み画像から訪問意欲を推定する際に、AI が街並み画像のどこに注目しているのかの考察を行う。

「分類・認識・推定・予測」が着目される一方で、「生成」といった創造的な分野においても AI の適用が進められている。使われている技術には、「敵対性生成ネットワーク (GAN)²⁾」がある。GAN はコンテンツ生成型の Deep Learning であり、生成器 (generator) と判別器 (discriminator) という 2 つのニューラルネットワークで構成されている。近年、この GAN の仕組みを応用した様々な手法が提案されており、2 つのペアの画像から画像間の関係を学習することで、1 枚の画像からその関係を考慮した補間をしてペアの画像を生成する pix2pix³⁾ や、簡単な 3D モデルから徐々に段階を踏んでゆき、最終的に本物に近い 3D データを生成する 3D GAN⁴⁾ などが挙げられる。そこで本研究でも GAN による街並みの画像生成を行い、歴史・都市のデザイン分野に対する適用可能性を試行する。

2. 街並み画像の分類・訪問意欲の推定 AI

(1) 概要

人間の視覚的な空間の特徴量の認識と、それに対する感性・印象の因果関係という複雑な関係は、いまだに未解明な部分が多い。そのため、入力と出力の構造的な把握に基づくデザインに対する感性・印象評価の推定が困難な側面は残っている。このような課題を解決するための手段として、Deep Learning を用いて成果を挙げている分野を俯瞰すると、前掲の分野を例に、対象の特徴量選択と結果の因果関係の構造的な把握が複雑であるという似たような性質を持つ分野がある。そのため、建築・都市のデザインにかかわる感性・印象に対しても、Deep Learning を用いた AI 技術の適用可能性があり得ると考えられる。上述をふまえ、本研究では萌芽性の検討対象を、「街並み画像」から「街路名と訪問意欲」を「推定する AI」の作成、とする。

研究対象と街並み画像は大都市や観光情報から選定した表 1 に示す 21 街路から作成した。画像は街並みの画像であるため、画像が単一の建築物・地面・空に占められることのないように配慮しながら Google Earth のストリートビューから作成した。枚数は既往事例と予備実験を参考に 1 街路 100 枚ずつの 2100 枚とした。図 1 に 21 都市中 4 都市の対象街路と画像例を示す。なお印刷の都合でグレースケール画像を掲載しているが実際には RGB カラー画像である。学習モデルには小さいサイズのフィルタによる畳み込みを繰り返すことで精度を向上させる Network In Network 型の「VGG」を参考に学習を実行した。

(2) 街路名の分類

印象評価の推定に先立ち、建築・都市に関する画像への適用可能性の基礎分析として、他分野で成果を挙

表 1 対象とした街路一覧

Class	Country	City	Name
0	America	New York	Wall Street
1	Australia	Melbourne	Hosier Ln
2	Canada	St John's	Cabot Street
3	Czech	Prague	Mostecka
4	England	London	Oxford Street
5	Hong Kong	Mong Kok	Shandong Street
6	Italy	Firenze	Via dei Calzaiuoli
7	Japan	Kyoto	Sanneizaka Street
8	Japan	Osaka	Dotonbori Street
9	Japan	Tokyo	Takeshita Street
10	Mexico	Guanajuato	Pocitos
11	Peru	Cusco	Questa de Santa Anna
12	Portugal	Porto	R.de Sao Joao
13	Russia	Moscow	Tuverskaya Street
14	Scotland	Edinburgh	Cockburn Street
15	South Africa	Cape Town	Dorp Street
16	Korea	Seoul	Bukchon Hanok Village
17	Spain	Barcelona	Plaça de Catalunya - Gran Teatre del Liceu
18	Taiwan	Chiufen	Chiufen area
19	Thailand	Bangkok	phraeng Nara Rd
20	Arab	Dubai	Al Worood 1 Street

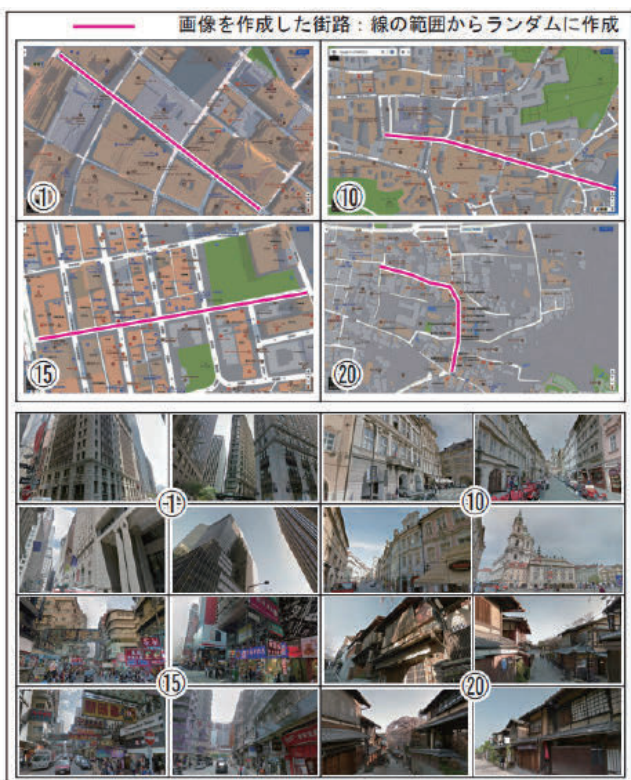


図 1 対象街路とストリートビューの例

表 2 街路名の AI 分類結果

[y_] is the correct class, [y'_] is the Deduction class

	y_0	y_1	y_2	y_3	y_4	y_5	y_6	y_7	y_8	y_9	y_10	y_11	y_12	y_13	y_14	y_15	y_16	y_17	y_18	y_19	y_20	Recall(再現率)
y_0	57	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	0	0	0	0.9661
y_1	1	52	0	0	0	0	0	0	0	3	1	0	2	0	0	0	0	1	1	0	0	0.8524
y_2	0	0	51	0	0	0	0	0	0	0	0	1	2	0	0	1	0	0	0	0	0	0.9272
y_3	4	0	0	53	6	0	4	0	0	0	1	3	1	2	0	0	0	0	0	0	0	0.7162
y_4	0	0	0	4	39	0	1	0	0	0	0	0	2	2	1	0	0	0	0	0	0	0.7959
y_5	0	0	0	1	0	51	0	0	1	2	1	0	1	0	0	0	0	0	1	1	0	0.8644
y_6	4	0	0	6	0	0	48	0	0	0	1	0	5	0	2	0	0	4	0	0	0	0.6857
y_7	0	0	0	0	0	0	0	55	0	0	0	0	0	0	0	0	0	3	1	0	0	0.9322
y_8	1	0	1	5	0	2	0	0	47	6	1	0	0	0	0	0	0	0	0	0	0	0.746
y_9	0	0	0	1	0	2	0	0	0	55	1	0	1	0	0	0	0	0	1	0	0	0.9016
y_10	2	0	1	0	0	0	0	0	0	0	51	4	0	0	0	3	0	0	0	0	0	0.836
y_11	0	0	0	1	0	0	0	0	0	0	0	60	2	0	0	0	0	0	0	0	0	0.9523
y_12	1	0	0	2	0	1	3	0	0	0	0	0	53	0	0	0	0	3	0	0	0	0.8412
y_13	1	0	0	0	0	0	1	0	0	0	0	0	0	44	2	0	0	1	0	0	0	0.8979
y_14	0	0	0	0	0	2	3	0	0	0	0	0	1	1	52	0	0	2	0	0	0	0.8524
y_15	0	0	3	0	0	0	0	0	0	0	1	3	0	0	0	52	0	0	0	0	0	0.8813
y_16	0	0	0	0	0	0	0	1	0	0	0	1	0	2	0	0	55	0	0	0	0	0.9322
y_17	5	0	0	1	1	0	1	0	0	0	0	0	3	3	0	0	0	46	0	0	0	0.7666
y_18	0	2	2	0	0	3	0	0	1	0	0	0	1	0	0	0	2	2	42	0	0	0.7636
y_19	0	0	1	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	59	0	0.9672
y_20	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	59	1
Precision(予測適合率)	0.75	0.9629	0.8644	0.7162	0.8478	0.8225	0.7868	0.9821	0.9591	0.8333	0.8793	0.8333	0.7162	0.8148	0.9122	0.9285	0.9649	0.7187	0.913	0.9833	1	
F-Measures	0.8444	0.9042	0.8946	0.7162	0.821	0.8429	0.7327	0.9564	0.8392	0.8661	0.8571	0.8888	0.7736	0.8543	0.8812	0.9042	0.9482	0.7418	0.8316	0.9751	1	
Accuracy(正確性)	0.8579		Avg.Precision		0.8661		Avg.Recall		0.8609		Avg.F-Measures		0.8607		K-Factor(Non-Weight)		0.8579		P-Value		2.2e-16	

げている物体認識が本分野でも可能であるかを検証した。具体的には、街並み画像を対象とした 21 街路のクラス分類についての適用可能性を検証した。学習データセットは左右反転により水増し処理を行った 4200 枚の 3 チャンネル・200 ピクセル× 338 ピクセルの街並み画像である。最適モデルを用いて推定を行った具体的な結果を表 2 に示す。各クラスの Recall と Precision の値を見ると約 70% から約 90% となっておりクラス毎の精度に約 20% の違いがみられる。全体の精度を見ると、Recall の平均値、Precision の平均値、Accuracy の平均値は全て約 86% である。各指標の値の散らばりが小さいことから作成した推定 AI の精度は約 86% であると考えられる。人でも学習を行えば同程度の精度で街並み推定は可能である。特に専門家であれば同程度以上の推定も可能であると考えられる。しかし約 30 分で学習し、人に依存せずに推定 AI を誰でも再現・利用できるという観点をふまえると、比較的高い精度と考えられる。つまり街並み画像の街路名推定という物体認識に対しても推定 AI は明示した程度には適用可能であると考えられる。そこで次に訪問意欲の推定 AI を作成する。

(3) 訪問意欲の有無の推定

次に印象評価の推定への適用可能性について検証を行う。画像は前節に使用したものと同様であるが、本節のクラスは「訪問したい」「訪問したいとは思わない」という印象評価の 2 クラスである。教師となる分類クラスには 1 名の被験者（建築系学科所属の 4 回生）による街路を対象とした被験者実験の結果を用いた。被験者実験では画像を一覧した後に直感でクラスに分類して貰った。通常このような被験者実験には複数人の代表値を用いることが多いが、個人（例えば建築家）のデザイン感に基づく感性・印象評価を学習・推定できるかという観点から被験者は 1 人とした。

表 3 訪問意欲の AI 推定結果

[y_] is the correct class, [y'_] is the Deduction class

	y'=0 (Want)	y'=1 (Don't want)	Recall(再現率)
y=0 (Want)	328	9	0.9732
y=1 (Don't want)	25	268	0.9146
Precision(予測適合率)	0.9291	0.9675	
F-Measures	0.9506	0.9403	

Accuracy (正確性)	0.9460	Effect Size	0.0539
Avg.Precision	0.9483	K-factor	0.8920
Avg.Recall	0.9439	P-Value	2.2E-16
Avg.F-Measures	0.9454		

Plot diagram of correct value(Human) and deduction value(Model)

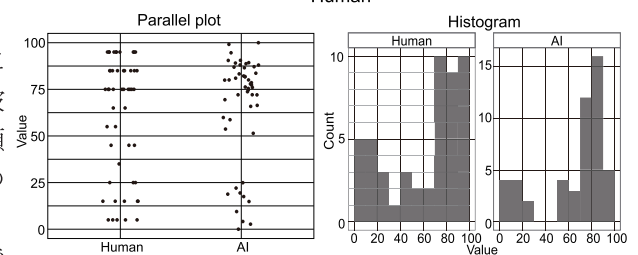
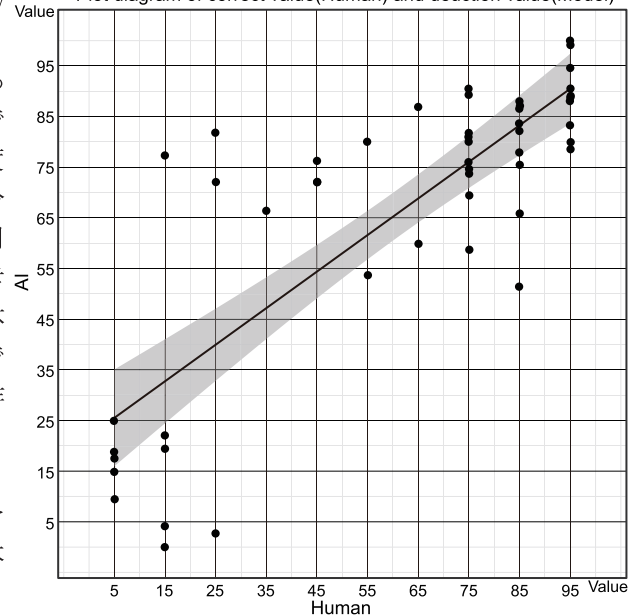


図 2 訪問意欲の人と AI の比較

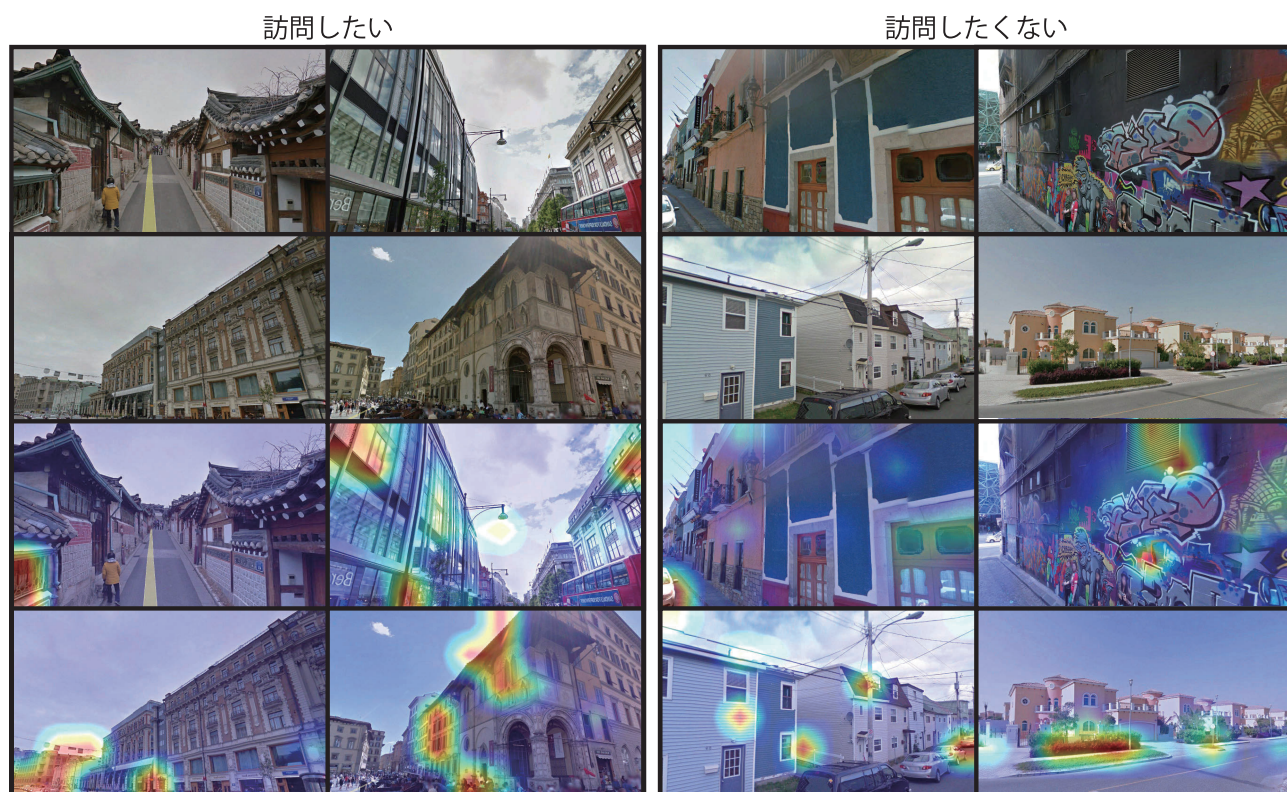


図3 訪問意欲の人とAIの比較

学習精度が最も高かったモデルを用いて推定を行った結果を表3に示す。実際の訪問意欲の有無の画像の割合は「有」が56.66%であり、推定した割合56.03%との差は小さい（比率の差の効果量（対応あり）で表現すれば約0.05）。画像ごとの対応を含めて確認するために、RecallとPrecisionを算出すると表のような値になった。全体の精度として、判定が一致した割合であるAccuracyをみると約94%（630枚中の596枚が一致）であり、Recall・Precisionの平均値も約94%となった。

以上より、前章と同じくより高精度な推定AI作成の余地はあり得るが、本研究は、被験者の訪問意欲の有無と高い一致度をもつ訪問意欲の有無を推定するAIを作成することができたと言える。

(4) 訪問意欲の度合いの推定

本節では前節にて作成した推定AIを援用することで、訪問意欲の度合いの推定が可能かを検証する。前章ではモデルの誤差に二乗誤差である「Squared Error」を採用した。これを用いると、訪問意欲の有無のクラス分類の推定に際して、「訪問したい」を0、「訪問したいとは思わない」を1とする0から1までの連続変数が得られる。本章ではこれを訪問意欲の度合いとして解釈可能であるかを検証した。検証方法は推定値と新たな被験者実験の比較である。被験者実験では、前章と同じ被験者に「訪問したい」度合いを10段階で評価して貰った。画像枚数は、2100枚に10段階の度合いを回答することは被験者の負担が大きいと考え、ランダムに各都市から抽出した50枚とした。本章はこの回答値（Humanと表記）と推定値（AIと表記）の比較である。値の並行ジッタープロット、ヒストグラム、プロット図を図2に示す。ヒストグラムを見ると、値の分布の範囲はおおよそ近似していることが読み取れる。サンプル毎のHumanとAIの値の対応を示すプロット図からは、全体的に正の相関関係であるが、Humanの回答値が低いサンプルをAIが実際より高く推定していることが読み取れる。この度合いの高低の傾向の推定の精度について相関係数を算出した。Spearmanの順位相関係数は0.7564となった。心理学分野の研究においては0.5以上をLargeと言及する。これを引用すれば、相関関係を高く推定できていると言及することを肯定できる。以上より、前章と同じく改善の余地はあり、記した実験条件下ではあるが、被験者の訪問意欲の度合いと差が小さく、高い相関関係を有する度合いの推定が可能なAIを作成することができたと言える。

3. AI 注視領域の可視化

前章では街並み画像を分類・訪問意欲を推定するAIの作成に成功した。しかしながらAIが画像を分類・判定する際に画像のどの部分を見ているかは人の目には見ることができない。特に近年ではAIのモデル構



図4 街並みの学習画像と生成画像の一例

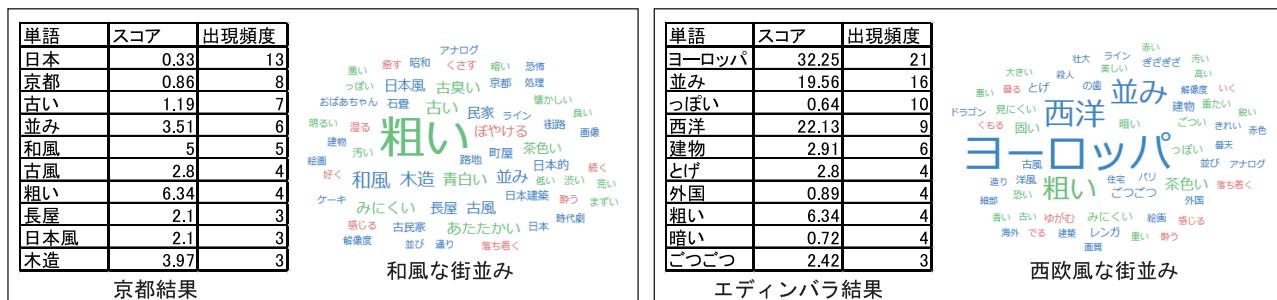


図5 ワードクラウド結果

造が複雑化し、人の手によりこうした判断を行うのが困難となる。そこでこのようにAIが持つ「ブラックボックス」要素を可視化する技術にCAM(Class Activation Mapping)がある。CAMはCNNベースの分類問題をAIが解いたときに、どの領域を注目しているかを可視化することができる手法である。本研究ではGrad-CAM⁵⁾と呼ばれる手法を用いる。Grad-CAMはニューラルネットワークの最終層の出力を、対象クラスのみ1、その他を0とし、その信号を誤差逆伝播により対象の特徴マップにおける勾配を計算することで各チャンネルの重み付けを行いAIが注視している領域ほど赤くヒートマップ化する技術である。全章で人間の訪問意欲の度合いを高い精度で推定するAIの作成に成功した。そこで本章ではAIが街並みの訪問意欲を注目して推定を行っているかをGrad-CAMを用いて可視化を行なった。なおモデル構造は全章(4)と同じである。

図3に人間が訪問したい、したくないと分類した画像とそれに対するCAMでヒートマップ化された画像を載せる。これを見ると、AIが建物の壁の模様や窓の形などに注目して訪問意欲の推定を行っていることがわかる。こうした結果を踏まえて、人間の視覚的な空間の特徴量の認識とAIの画像に対する特徴量の認識の差などについても今後考察を行う予定である。

4. 特定の街並みを生成するAI

建築行為の中でデザインという段階は、極めて重要であると同時に、伝統的な手法を用いて人に依存する労力が非常に多く費やされる側面を持つ。そのため、この段階を発展させることの意義は大きいと言える。現在、このように重要ではあるが、人特有の行為であり機械による代替や支援が難しいと考えられてきた多くの分野でDeep Learningを基盤とする人工知能(AI)の研究が進んでいる。デザインのような創造性に関するAI技術の適用分野には、音楽・絵画などがある。使われている技術の例には、「敵対性生成ネットワーク(GAN)」がある。そこで本章でも学習データが持つデザイン要素を閲覧者が感じるような画像を生成するAIを作成することができるかを検証する。具体的な対象は著名な街並みである。なお、任意のデザイン要素を持ったベクトル生成の可否は被験者実験により確認する。今回用いた学習モデルはDCGAN⁶⁾とよばれる手法である。モデルの構築には深層学習のフレームワークであるchainerを使用した。今回のモデルでは100次元のランダムノイズから最終的に128×128ピクセルの画像を生成している。モデルの概念図を図

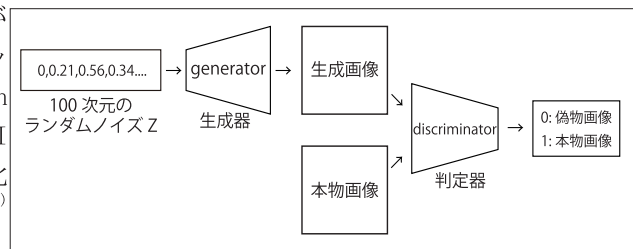


図6 画像生成AIの概念図

6に示す。本章の目的は、街並みらしい画像が生成できるかに加え、AI 作成者の意図を反映した街並みを生成できるかを検証することである。学習データとして選定した街並みは、和風な街並みとして「京都「産寧坂」」、西洋的な街並みとして「スコットランド「エディンバラ」」を選定した。画像はGoogleのストリートビューからスクリーンショットを撮りそれぞれ200枚用意した。ただし学習データは反転による水増しを行った400枚である。図4に学習データの画像の一例を示す。

生成した画像に対する印象調査は、建築系学科に所属する大学生32名を対象に行った。被験者実験では、GANで生成した画像であることを告げずに画像を被験者に提示し、画像が何を表現していると思うか、またそれらの画像に対してどのようなイメージを抱くかを自由記述方式で回答して貰った。提示した画像は予め選定したそれぞれ4枚である。GANを用いて生成した画像の一部を図4に示す。画像を見ると京都の街並み生成画像では軒や瓦屋根などの特徴をAIが学習し、描画に反映させているように読み取れる。またエディンバラの街並み生成画像では、茶色のレンガ壁や屋根の形などの描写ができてるように読み取れる。

アンケートで得られた記述の抜粋と、アンケート結果をテキストマイニングの手法の1つであるワードクラウドを用いて分析した結果を図5である。なお今回のアンケートではユーザーローカル テキストマイニングツール⁷⁾による分析を行った。この図を見ると、京都の画像群では、「和風」や「古風」といった言葉が、エディンバラの画像群では「ヨーロッパ」や「西洋」といった言葉においてスコアが高い。このことからAI作成者の意図・実際の街並みに近似な印象を被験者に抱かせる画像を生成できていると考えられる。しかしながら「粗い」などの言葉のスコアも高い。このことからGANで生成した画像には解像度やリアリティなどの面において課題があると考えられる。

5. まとめ

Deep Learningを用いたAI技術の歴史都市分野への適用可能性について検証と考察を行った。第2章では、街並み画像からクラス分類と訪問意欲の推定を行うAIを作成し、その精度について検証を行った。第3章ではGrad-CAMと呼ばれる手法を用いてAIが街並みの訪問意欲を推定する際に画像のどこに注視しているかを考察した。第4章ではGANを用いて分類や推定のみならずデザインといった分野に対するAIの適用を街並み画像を対象に生成・被験者実験を行った。今後は街並みだけでなく、歴史的建造物や防災分野に対してもAI技術の発展的な段階の検証を行う予定である。

参考文献

- 1) Atsushi Takizawa and Airi Furuta:3D Spatial Analysis Method with First-Person Viewpoint by Deep Convolutional Neural Network with Omnidirectional RGB and Depth Images, eCAADe 2017, Sapienza University of Rome, Rome, Italy, pp.693-702, 2017.10
- 2) Ian J Goodfellow, Jean Pouget-Abadie, Mehdi Mirza, Bing Xu, David Warde-Farley, Sherjil Ozair, Aaron Couville, Yoshua Bengio: "Generative Adversarial Nets" <http://papers.nips.cc/paper/5423-generative-adversarial-nets>
- 3) <https://phillipi.github.io/pix2pix/>
- 4) <http://3dgan.csail.mit.edu/>
- 5) Ramprasaath R. Selvaraju, Michael Cogswell, Abhishek Das, Ramakrishna Vedantam, Devi Parikh, Dhruv Batra, Virginia Tech, Georgia Institute of Technology: Grad-CAM: Visual Explanations from Deep Networks via Gradient-based Localization, 2017
- 6) Alec Radford, Luke Metz, Soumith Chintala: Unsupervised Representation Learning with Deep Convolutional Generative Adversarial Networks, <https://arxiv.org/abs/1511.06434>
- 7) <http://textmining.userlocal.jp/>
- 8) 山田悟史, 大野耕太郎: Deep Learningを用いた印象評価推定AIの作成と検証, 日本建築学会計画系論文集, 第84巻, 第759号, 掲載決定, 2019, 5
- 9) 大野耕太郎, 山田悟史: Deep Learningを用いた画像生成AIの建築都市デザイン分野への適用可能性, 第40回情報・システム・利用・技術シンポジウム論文集, pp.246-249, 2018, 12
- 10) 坂本俊之: Chainerで作るコンテンツ自動生成AIプログラミング入門, 2017.2, C&R研究所