

# Deep Learning を用いた都市空間における人の行動の予測

## —居住人口及び滞留人口を対象に—

### Prediction of the action using the deep learning

#### -targetting residential population and congestion population-

○中村翔太<sup>\*1</sup>, 大野耕太郎<sup>\*2</sup>, 中西達也<sup>\*1</sup>, 山田悟史<sup>\*3</sup>

Shota NAKAMURA<sup>\*1</sup>, Kotaro ONO<sup>\*2</sup>, Tatsuya NAKANISHI<sup>\*1</sup>, Satoshi YAMADA<sup>\*3</sup>

\*1 立命館大学 理工学部 建築都市デザイン学科

Undergraduate, Dept. of Architecture and Urban Design, Ritsumeikan Univ.

\*2 立命館大学 理工学研究科 環境都市専攻 博士前期課程

Graduate, Dept. of Architecture and Urban Design, Ritsumeikan Univ.

\*3 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）

Lecturer, Dept. of Architecture and Urban Design, Ritsumeikan Univ., Dr.Eng.

キーワード：人工知能，深層学習，都市空間，人口分布，画像生成

**Keywords:** Artificial intelligence, Deep Learning, Urban space, Population distribution, Image Generation.

## 1. はじめに

近年、人工知能（Artificial intelligence、以降 AI と呼称）への関心・期待は様々な分野で高まっている。広く知られている自動運転に限らず、建築・都市分野においても AI を取り入れる動きがみられる。例にはロボット施工や AI とビッグデータの都市への応用<sup>①</sup>などがある。

上記に限らず多くの可能性があるなか、本研究は Deep Learning を基盤とする AI（以降、本研究の AI は Deep Learning の学習済みモデルを指す）を用いた都市における人々の行動予測を試行した。具体的には居住人口及び滞留人口の予測である。都市における人の行動分析はこれまでにマルチエージェントシステム<sup>注1)</sup>や GIS を用いたネットワーク解析<sup>注2)</sup>により広く行われてきた。それら手法では、研究者が行動との因果モデルやシミュレーションアルゴリズムを仮説的に用意する必要があった。一定の成果が報告されている一方で人が因果関係を仮説的に用意することに精度の限界がある可能性がある。そこで本研究では研究者が用意する現象の特徴量や因果仮説への依存度が低い Deep Learning を基盤とする AI の画像生成技術を利用した新たな予測手法を提案する。具体的には、地図と人口分布の関係性を学習した AI による未学習地域の人口分布の予測である。

## 2. 研究手法

### 2.1 対象地域

対象地域は都市部・住宅地・山間部があり多様な人口分布をもつ大阪府を選定した。学習画像には大阪市を中心とし北西は高槻市、南は堺市までの地域を使用した。

### 2.2 深層学習モデルとデータセット

本研究はコンテンツ生成型 AI である「敵対的生成ネッ

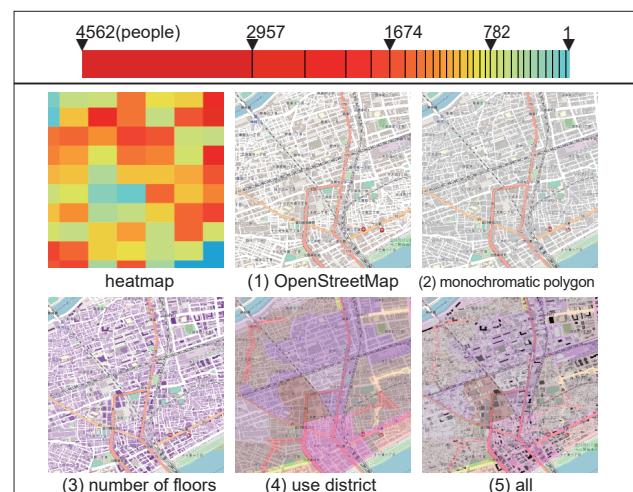


Fig.1 training images

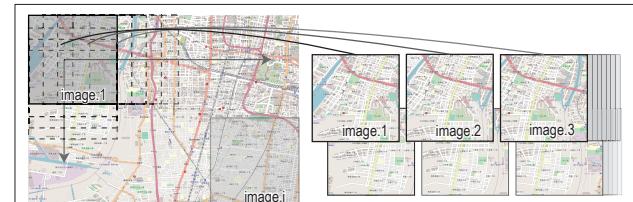


Fig.2 way of making training images

トワーク (Generative Adversarial Network、以降は GAN と呼称)」<sup>注3)</sup> という深層学習モデルを採用した。GAN は 2014 年にイアン・グッドフェロー氏らが「Generative Adversarial Nets」<sup>②)</sup> という論文で発表したニューラルネットワーク構造である。詳細は後述するが本研究はこの GAN を応用した技術の一種である pix2pix を用いた。

学習画像は地図画像と人口分布より作成したヒートマップ画像を用いた。これらの画像は ArcGISpro を使用し作成した。ヒートマップ画像は国勢調査の在住人口データ<sup>注4)</sup>を 250m メッシュごとに色分けしたものを用いた。等級は自然分類を用い、色は図 1 上部のように人口に応

じて指定した。地図画像はAIが学習する際、地図上に画像情報が多い方が予測精度が向上すると仮定し、以下の5種類を作成した。(1) オープンストリートマップ<sup>注5)</sup>、(2) (1)にZmapTOWN II<sup>注6)</sup>より取得した単色の建物ポリゴンを配置した地図画像、(3)(2)の建物ポリゴンを階数ごとに色分けした地図画像、(4)(2)に用途地域区分データ<sup>注7)</sup>を不透明度30%で重ね合わせた地図画像、(5)(3)に用途地域区分データを同様に重ね合わせた地図画像。図1に学習画像の例を示す。これらを縮尺1/15000で地域ごとにに出力した画像を画像左上2km<sup>2</sup>の範囲をトリミングし、縦横200mずつスライドさせトリミングを繰り返すことで学習画像を作成した。図2に学習画像の作成方法の図解を示す。この方法で作成した地図画像とヒートマップ画像各2520枚(1)-(5)の5種類のデータセット(計12600枚)をそれぞれAIに学習させた。

### 2.3 検証

検証には未学習の大阪市西成区～阿倍野区周辺の画像から学習画像と同様に作成した各90枚のデータセットを用いた。予測精度は実際の検証地域のヒートマップ画像とAIが生成したヒートマップ画像の画素ごとのRGB値の相関係数をペア毎に求めその平均を指標とした。

### 3. pix2pixの概要

pix2pixは2016年にフィリップ・イソラ(Phillip Isola)氏らが「Image-to-Image Translation with Conditional Adversarial Networks」<sup>3)</sup>にて発表したGANを利用した画像生成アルゴリズムである。pix2pixは教師画像Aと教師画像Bのペアの画像から、AからBへの変換方法を学習し、未学習の教師画像Aを入力することで学習した変換方法に応じて予測される変換画像Bを生成する技術である。識別器は教師画像Aと生成器が生成した変換画像Bのペアと教師画像Aと教師画像Bのペアを正しく判別するように学習する。生成器は判別結果とともに変換方法を更新し識別器を誤認させるよう学習する。図3にpix2pixの基本構造を示す。本研究において画像Aが地図画像、画像Bがヒートマップ画像にあたる。

本研究ではpix2pixHD<sup>4)</sup>型のネットワークを用いる。pix2pixは通常、生成器の学習にU-Net<sup>5)</sup>型のネットワークを用いるが、今回はResidual Network<sup>6)</sup>(Res-Net)を用いる。また識別器の学習にはPatchGAN<sup>3)</sup>を用いる。図4にネットワーク図<sup>7)</sup>を示す。また使用した学習フレームワークはPyTorchである。

## 4. 生成結果

### 4.1 相関係数

図5は検証地域のヒートマップ画像とAIが生成したヒートマップ画像の相関係数を示した散布図である。相

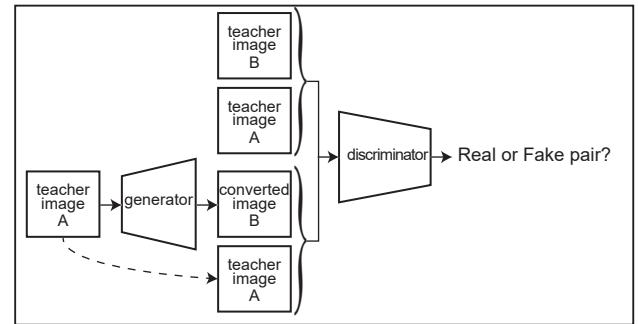


Fig.3 structure of pix2pix

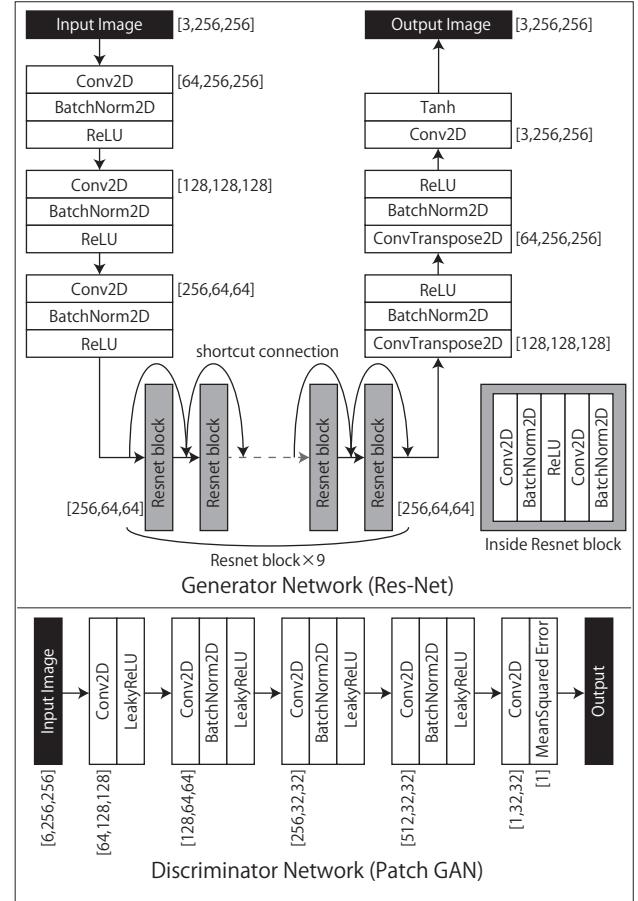


Fig.4. network of pix2pix

関係係数の平均も併記している。いずれも相関係数の平均は0.6以上であり中程度の類似性を持った予測に成功したといえる。また(3)階数ごとに色分けした建物ポリゴンを配置した地図画像を学習させたAIの生成結果は0.8138と最も高く、分散も小さいこと、高い精度で予測に成功したといえる。しかし、(4)用途地域区分を重ね合わせた地図画像を学習させたAIの生成結果は0.6864と(1)オープンストリートマップを用いた場合より精度が下がり、用途地域区分については仮説と異なる結果となつた。(5)建物階数・用途地域区分の両方を地図に表したAIの生成結果は0.7819である。(5)が(3)より数値が低い理由としては用途地域区分の情報が追加されていることが考えられ、在住人口の予測において用途地域区分の情報はあまり有効ではないことが確認できた。

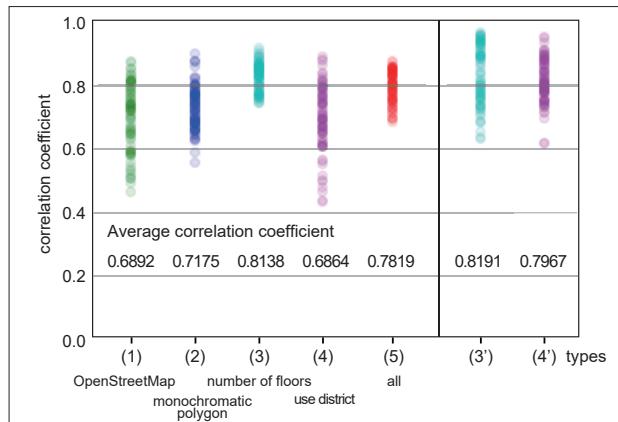


Fig.5 scatter diagram of correlation coefficient

#### 4.2 生成画像の比較・考察

図6に各データセットごとの生成結果を示す。入力画像はデータセットごとに作成した地図画像、正解画像は実際の人口分布のヒートマップ画像、出力画像はAIが生成したヒートマップ画像である。すべての出力画像に共通して川や公園といった場所が反映されている。しかし、人口密集地である黄色～赤色の部分においては各データセットごとに出力画像と正解画像との類似性は異なる。

図7は最も精度の良かった(3)の入力画像と入力画像に出力画像を重ね合わせた画像である。他のデータセットと同様に川や公園は少人口地域として識別されている。また建築面積が小さい高層の建築物が密集している地域が赤色に着色され、建築面積が大きく低層の建築物や建築物が疎らな地域は少人口地域として識別されていることが見て取れる。地図内の建築物の用途について確認すると、この地域において建築面積が小さい高層建築物の多くはマンションであり、建築面積が大きい低層建築物は学校や工場、倉庫などであった。この結果から地図に建物階数情報を付与することでAIが建築物の用途を予測し人口分布に反映させていることがわかる。

#### 5. 滞留人口の予測

人の流れデータ<sup>8)</sup>を用いて滞留人口の予測を試行した。これは、居住より一時的な人の行動を予測が可能な検討である。人口分布には平日正午のデータを用いた。これにより正解画像は4章とは異なる範囲の人口が高くなつた。居住地ではなく人々が活動する場所、駅や商業施設、オフィスビルなどの人口が多くなつたためである。

##### 5.1 学習条件

人の流れデータから平日正午の滞留人口を取り出し、500m メッシュごとの人口で色分けをしたヒートマップ画像を学習に用いる。人の流れデータは点群であるため、メッシュサイズを小さくすると1メッシュあたりに存在する点が少なくなり不正確なメッシュが作成される恐れがあるため今回は500m メッシュを用いた。また地図画像

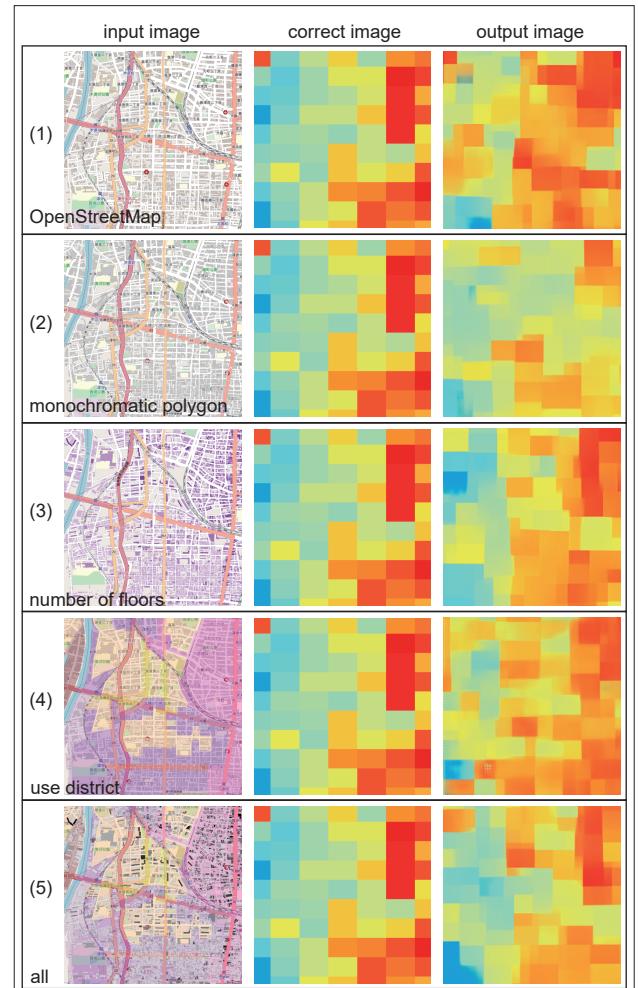


Fig.6 generated images

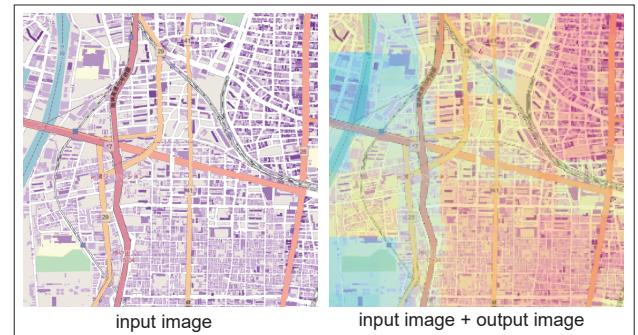


Fig.7 generated images of type(3)

には(3), (4)を用い、作成したデータセットを(3'), (4')とする。その他の条件は2章と同様である。

#### 5.2 生成結果

図8に生成結果の一部を示す。(3'), (4')の生成結果の上段は図6と同じ地域の生成結果(相関係数は約0.8)、下段は相関係数がどちらも約0.6と中程度となった生成結果である。平均相関係数は(3')が0.8191、(4')が0.7967であり、どちらも非常に高い割合で正解画像と出力画像が類似しているといえる。図5に相関係数の散布図を示す。(4')は(4)と比較すると平均相関係数が約0.1高く、用途地域区分情報は滞留人口の予測においては有効である。

ることが確認できた。また居住人口の場合とは異なり地図上で色付けされた幹線道路沿いに商業ビルなどの建築物が並んでいる地域が人口密集地域と識別され、低層の住宅地域は少人口地域と識別されているのが見て取れる。この結果から滞留人口を用いることでAIが都市における人々の時間ごとの活動地域を予測することに成功したといえる。

## 6. おわりに

以上のようにAIの画像生成技術を用いて大阪府の居住人口と滞留人口の高い精度で予測することに成功した。

画像を生成するGANを用いて人の現象を画像として予測する試みは先行研究が少ない。そのため精度は未知であった。そのような現況のなか、本研究の成果は、人口に限らず画像表現可能な現象の予測にAIが適用可能であるという新たな萌芽性を示唆している。今後はこの萌芽性について建築物という異なるスケールにおいても人々の行動が予測かを試行する。

このような予測は人間が計画した地図・図面に対しても適用できる可能性がある。このような予測を計画者にフィードバックすることは実務においても有用であると考えられる。本研究の延長には、人間では困難な評価をフィードバックするAIと人間の共創が期待される。

### [謝辞]

本研究で用いたデータの一部は東大CSIS共同研究No.859として借用しました。また株式会社 梓設計 AX teamから助言を賜りました。記して深謝を申し上げます。

### [注釈]

- 1)エージェントと呼ばれる理論的モデルを複数設定し、人工社会を構成しそれぞれのエージェントの相互作用をシミュレーションするシステム。
- 2)さまざまな対象を点と線からなるネットワークとして表現し、その構造的な特徴を探る研究方法。
- 3)画像生成を例にGANの概要を示す。GANは生成器と識別器の2つのニューラルネットワークで構成されている。生成器は多次元のランダムベクトルのノイズから偽作画像を生成する。識別器は生成された偽作画像が本物であるか判別する。そして生成器は識別器に偽作画像を元画像と誤認させることを目的として学習し、逆に識別器は元画像と偽作画像を正しく判別することを目的として学習する。このように対立した2つのニューラルネットワークが競い合うように学習することで精度の高い画像が生成される構造となっている。この学習の構造から敵対的生成ネットワークと呼称されている。
- 4)政府統計の総合窓口e-Stat(<https://www.e-stat.go.jp/>)より国勢調査の統計データを取得した。
- 5)誰もが自由に利用でき、なおかつ編集機能のある世界地図を作る共同作業プロジェクトにより作成された地図。
- 6)ゼンリンが提供する詳細なデータを書き込んだ住宅地図をデータ化し、コンピュータによって活用できる住宅地図データベース。東京大学空間情報科学研究センターより借用。
- 7)国土交通省のウェブサイトより取得。[\(https://nlftp.mlit.go.jp/ksj/gml/datalist/KsjTmplt-A29.html\)](https://nlftp.mlit.go.jp/ksj/gml/datalist/KsjTmplt-A29.html)
- 8)人々の流动について、何らかのデータソースをもとに粒度を揃えた時空間位置でデータ化したもの。具体的には「2000年京阪神都市圏 人の流れデータセット」を用いた。

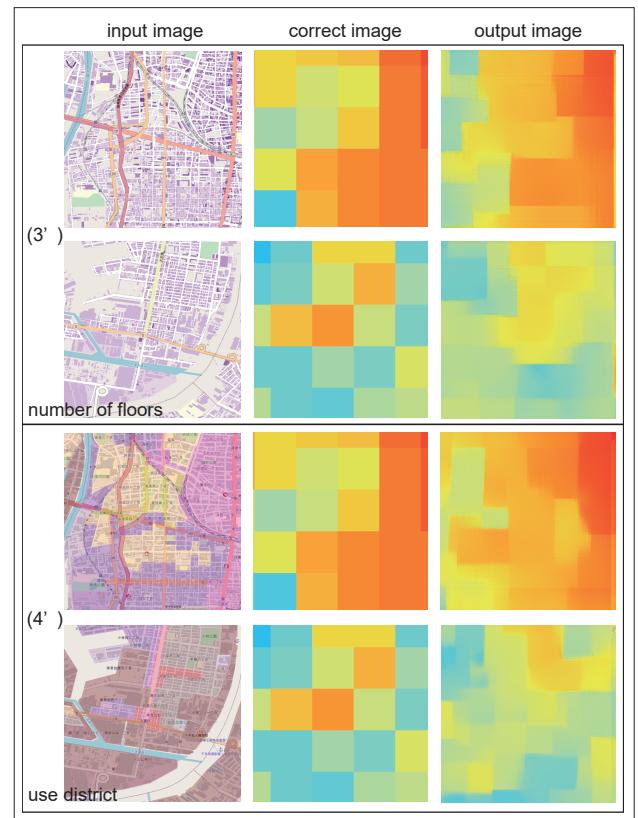


Fig.8 generated images of type(3'),(4')

### [参考文献]

- 1)吉村有司, 建築討論「『ビッグデータは都市理論を変えるか?』をめぐって」, <https://is.gd/lHe4V4>, 最終閲覧日: 2020年9月29日
- 2)Ian J Goodfellow, Jean Pouget-Abadie, Mehdi Mirza, Bing Xu, David Warde-Farley, Sherjil Ozair, Aaron Courville, Yoshua Bengio: Generative Adversarial Nets, <http://papers.nips.cc/paper/5423-generative-adversarial-nets>, 2014, 6
- 3)Phillip Isola, Jun-Yan Zhu, Tinghui Zhou, Alexei A. Efros: Image-to-Image Translation with Conditional Adversarial Networks, <https://arxiv.org/abs/1611.07004>, 2016, 11
- 4)Ting-Chun Wang Ming-Yu Liu Jun-Yan Zhu Andrew Tao Jan Kautz Bryan Catanzaro: High-Resolution Image Synthesis and Semantic Manipulation with Conditional GANs, <https://arxiv.org/pdf/1711.11585.pdf>, 2017, 11
- 5)Olaf Ronneberger, Philipp Fischer, Thomas Brox: U-Net Convolutional Networks for Biomedical Image Segmentation, <https://arxiv.org/abs/1505.04597>, 2015, 5
- 6)Kaiming He, Xiangyu Zhang, Shaoqing Ren, Jian Sun: Deep Residual Learning for Image Recognition, <https://arxiv.org/abs/1512.03385>, 2015, 12
- 7)大野耕太郎, 池之上慎吾, 山田悟史: 目的に応じたジェネレーティブデザインAIの体系化, 第41回情報・システム利用・技術シンポジウム論文集, 2019
- 8)山田悟史, 大野耕太郎: Deep Learningを用いたデザインAIの作成と検証-街並みと建築物外観の画像生成を対象に-, 日本建築学会 計画系論文集, 第85巻, 第770号, pp.987-995, 2020.4, 日本建築学会
- 9)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-街並み画像の街路名と訪問意欲を対象に-, 日本建築学会計画系論文集, 第84巻, 第759号, pp.1323-1331, 2019.5, 日本建築学会