

# Deep Learning を用いた心理評価に基づくデザイン生成 AI

## Design Generation AI based on Psychological Evaluation using Deep Learning

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

Tatsuya NAKANISHI<sup>\*1</sup>, Kotaro ONO<sup>\*2</sup>, Syota NAKAMURA<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: DeepLearning; 感性; 町並み; GAN; pix2pix

Keywords: DeepLearning; Sensibility; City landscape; GAN; pix2pix.

### 1. はじめに

近年, 人工知能 (AI) が社会の中で急速に関心を集めるようになった. それに伴い Deep Learning という用語が人工知能や情報技術, 中には音楽やイラストといった娯楽分野にまで広く普及するようになった. Deep Learning とは, 人工知能を構成する要素の一つであり, 画像の分類を例にあげれば, 人の手による特徴量の設定を必要としない, 特徴量の自動学習を行う機械学習ネットワークのことを指す. Deep Learning を用いた人工知能は, 近年様々な分野において活かされており, 特に「画像認識」「音声認識」「自然言語処理」の分野において著しい精度と成果を挙げている. 建築の分野においても Deep Learning の応用は進んでおり, 山田と大野<sup>1)</sup>による既往研究では, 都市の町並み画像における感性・印象評価の推定を行う AI の検証が進められており, 精度の確立と証明にも既に成功している.

また, この Deep Learning は画像の分野において, 「画像認識」だけでなく「画像生成」の面でも著しい成果を挙げている. この画像生成で成果を挙げている人工知能は, GAN<sup>注1)2)</sup> (敵対性ネットワーク) と呼ばれる Deep Learning のネットワークの一種が主であり, 生成器と呼ばれる学習画像に近い偽物の画像を生成するネットワークと, 分類器と呼ばれる生成器によって作られた画像とそうでない画像を分類するネットワークの組み合わせによって構成される. GAN において, この生成器と分類器は互いの精度が向上し合うように学習するため, 生成器によって生成される学習画像群を真似た画像は, 最終的に分類器の判定を騙すほどの精度を持った偽物の画像になる. この GAN による画像の生成は, 山田と大野による既往研究<sup>3)4)</sup> や山田と池之上による既往研究<sup>5)</sup> においても

建築デザインへの応用が検証されており, Deep Learning を用いた GAN の, 建築や都市のデザインソースへの適用が今後も期待される. しかし, GAN は既往の研究で著しい結果を出してはいるものの, 未だに未解明の部分も多く, 発展途上であると言える. また, GAN の建築を含めた他分野への適応は, 未だに大きな余地を残しており, 今後も大いにその探求が期待される. GAN による建築や都市の町並み画像の生成において, 人の感性を反映させることが可能であることを提示できれば, 今後の建築の設計・デザイン分野における Deep Learning の発展に繋がっていく. 上記を踏まえ本研究では GAN, 特に pix2pix<sup>6)</sup> を用いた人の感性による都市への添景配置を学習した AI による画像の生成とその検証を行なう.

### 2. 研究概要

#### 2.1. 研究の流れ

今回は pix2pix という GAN の一種を用いて添景の生成を行い, その生成の仕方に我々人間の感性が影響を与えることができるかどうか検証を行う. pix2pix とは先述のように GAN を利用した画像生成ネットワークの1つで, それぞれ別の属性を持つ2つの画像群の関係性を学習し, 片方の属性を持つ画像をもう片方の属性の画像のように変換することができるネットワークのことである. pix2pix のネットワーク図は図1に示す. 本研究では pix2pix を用いて, 生成する画像に人間の感性を学習させることが可能か検証を行う.

まず今回の試行では, pix2pix を用いて, 添景 (人と樹木) の無い都市の画像 (3DCG を用いた架空のもの) の画像から, 添景の描写された都市画像の生成を行う. この時 pix2pix には, 添景の無い都市の画像 (x) と添景のあ

る都市の画像 (y) の関係性を学習させ、「x から y」といった流れで画像を変換できるように学習させる。この時 y の画像において、y1 (乱数によってランダムに添景を配置した都市の画像) と y2 (人間の感性によって好ましいと思われる場所に添景を配置した都市の画像) の二つを用意する。そして pix2pix によって  $x \rightarrow y1$  と  $x \rightarrow y2$  の二通りを別々に学習させる。学習が完了した後この二つの AI に、学習に用いなかった添景の無い都市の画像から、それぞれ乱数と感性による添景配置のなされているであろう都市の画像を生成し、生成されたこの二種の画像を被験者 (y2 の作成において添景の感性配置を行った者) に評価してもらう。また今回の研究は個人のデザインセンスの学習に焦点に当てているため、被験者の数は一人とした (建築学科の4回生)。この全体の流れを図2に示す。

## 2.2. 研究対象・学習データセット・学習モデル

研究対象となる都市は図3に示される架空の都市を扱う。図3内の右画面端に続く道の先には駅があると仮定して、被験者には添景の配置を行なってもらう。この都市モデルに配置する添景は、人間 (学習の都合上、街路に置く場合のみ白の添景とした) と樹木の二種 (図5) であり、それぞれ別に乱数配置と感性配置を行い、学習・検証を行う。人の添景の配置量は道の広さとのバランスを考慮して、街路上に 0.02 人/m<sup>2</sup> の密度で配置するものとした。公園内は街路の歩行者密度とのバランスを考慮して 0.01 人/m<sup>2</sup> とし、図4のように配置量と配置範囲を指定した。もう一つの試行で行う樹木の配置は、歩行者の通行を妨げない適量として街路と公園、両者とも配置密度を 0.02 本/m<sup>2</sup> とし、図4のように配置量と配置範囲を指定した。人間と樹木の同時配置を行う試行においても両者の密度は上記と同じものとした。都市の構成物の中には、添景を感性配置する際に影響を与える要因になるように、店 (カフェやコンビニ等) と公園用の噴水、そして歩道に飛び出し部を作成した。AI が学習時にこれらのオブジェクトを認識しやすいように店は赤色 (R255, G0, B0 値)、噴水は青色 (R0, G0, B255 値) の 3D モデルとした。

データセットは先述の都市を図4の経路で歩くウォークスルーのカメラから作成し、街路のみを歩くパターン A と公園のみを歩くパターン B の二種を作成した。今回はこれらのパターンを別々に学習し結果を検証していく。ウォークスルーのカメラは学習データセットに多様性を持たせるため、首を振りながら歩くものとした。また y の学習画像内において添景が描写されていない画像はデータセットから除くものとした。学習に用いたデータセットの例は図7に示す。カメラの設定については図6に示す。学習に用いる訓練用画像は添景の無い都市の画像 (x) と添景のある都市の画像 (y) でそれぞれ 5000 枚ずつ、テスト用画像は 200 枚ずつとして学習させた。

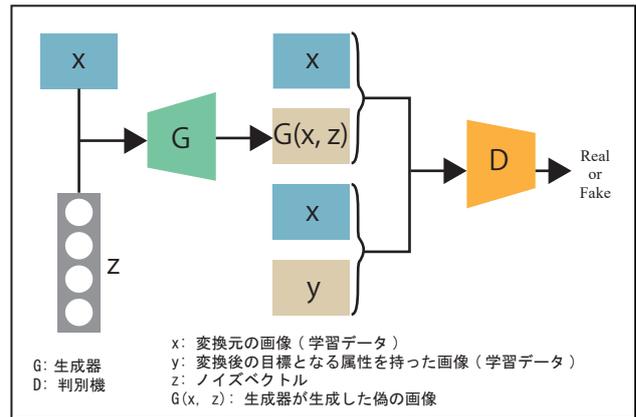


Fig.1 pix2pix network

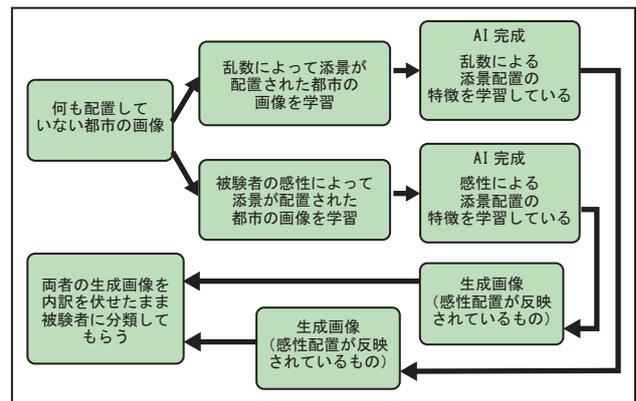


Fig.2 Research flow

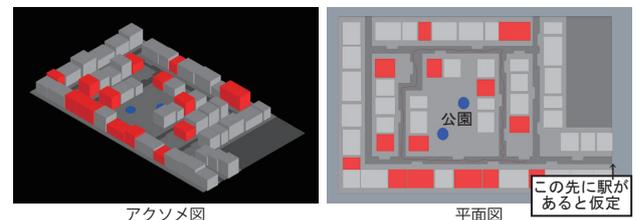


Fig.3 Image of city used for learning

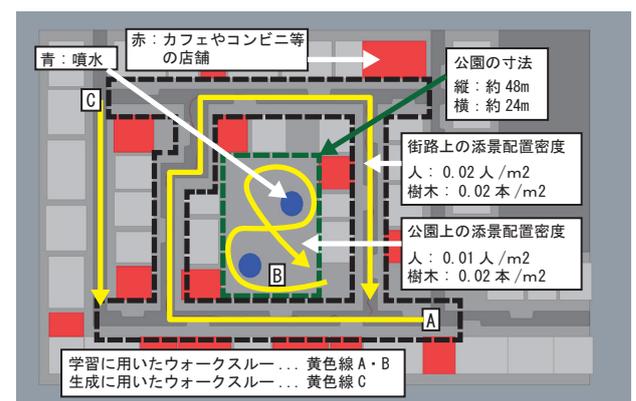


Fig.4 Arrangement amount and range of the scenery

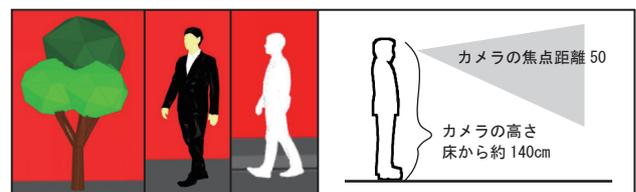


Fig.5 Scenery

Fig.6 Camera settings

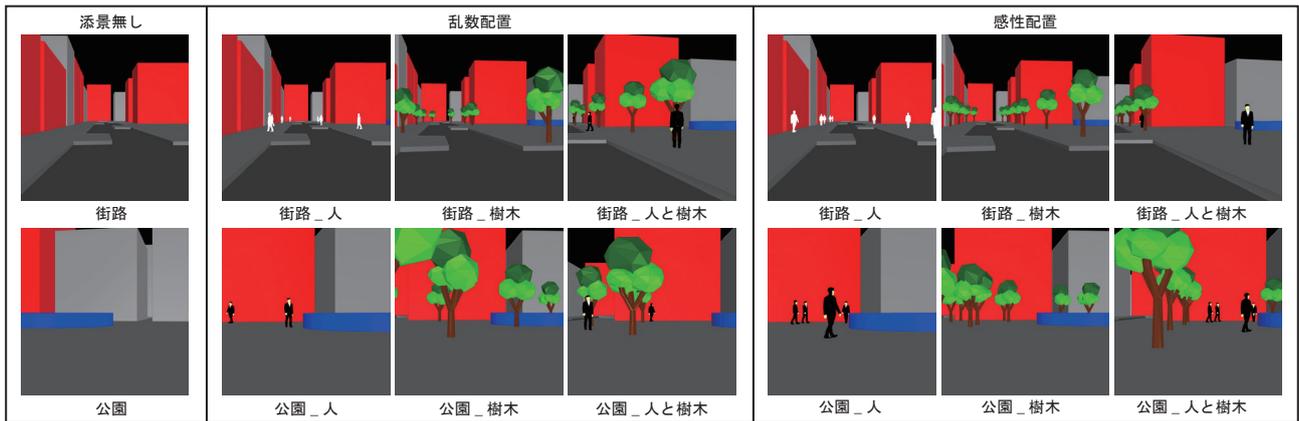


Fig.7 Dataset example

### 3. 町並み画像の学習とその生成結果

#### 3.1. 都市の町並み画像の印象評価

学習後の各 AI による生成の結果は図 8 のようになった。この画像の生成に用いた町並みの画像は図 4 の C (街路) の経路のウォークスルーと, A (街路) B (公園) のウォークスルーを逆方向に歩いて見える町並み, そして学習画像を左右に反転させた画像を用いており, 学習に用いた画像と違いながらもある程度の類似性を持った町並みの画像となっている。生成画像の精度をある程度確保するために, 今回はこのような仕様にした。

人を町並みに配置した場合の学習結果は図 8-A と図 8-B のようになっており, 図 8-A が乱数配置学習 AI による生成の結果。図 8-B が被験者による感性配置を学習した AI の生成結果である。公園にて黒のスーツを着た人の添景を学習し生成した結果, 生成がほとんど上手くいかなかったため, 街路の生成では添景を白の人の添景に差し替えて, 学習し生成を行なった。その結果生成は成功し, A と B とともに街路の上に白い人の添景が生成されていた。街路において A と B の両者を比較すると, 同じ景色でも添景の生成位置に違いがあることが分かる。

人間ではなく樹木を町並みに配置した場合の学習結果は図 8-C と図 8-D のようになっており, 図 8-C が乱数配置 AI による生成の結果。図 8-D が被験者による感性配置を学習した AI の生成結果である。結果, 街路のパターンにおいては, 乱数配置 AI によって生成された画像 (C) と感性配置 AI によって生成された画像 (D) の両者において, 樹木と認識できるものが明確に生成されており成功という結果に終わった。今回は公園のパターンでも生成が成功しており, 街路と公園共に, 乱数配置 AI と感性配置 AI による生成画像を比較した時, その配置の仕方にある程度の差を見受けることができ, 噴水周りの樹木にも配置位置にある程度の差を確認することができた。樹木のモデルは人のモデルよりも大きく色彩が豊かであったことが公園での生成の成功に繋がった一因と思われる。

人間と樹木の両方を配置した場合の結果は図 8-E と図 8-F となっており, 図 8-E が乱数配置による生成の結果,

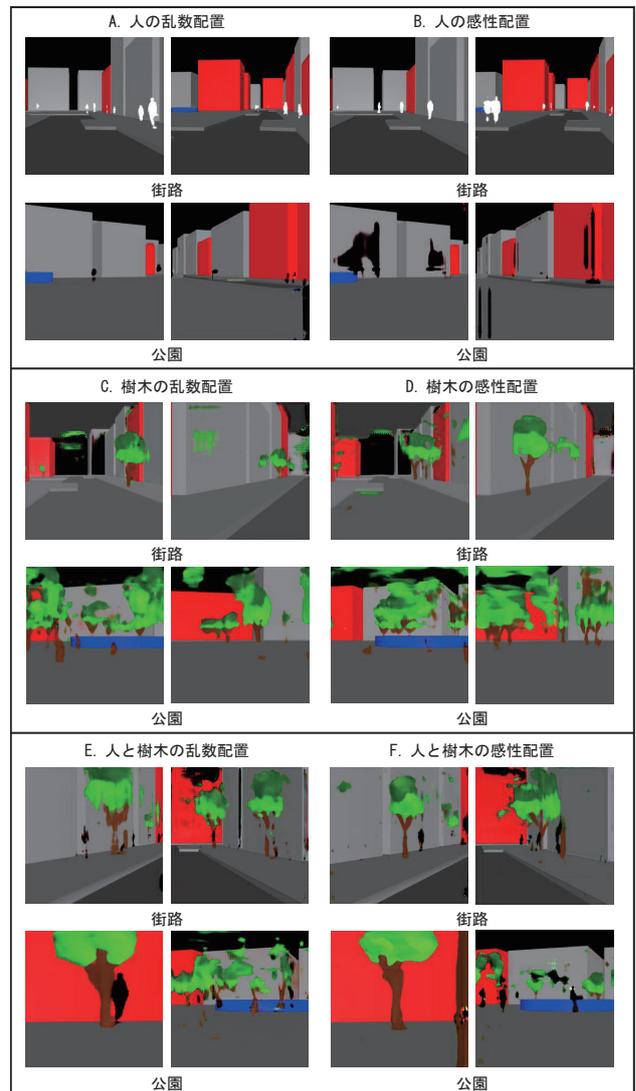


Fig.8 Generation result

図 8-F が被験者による感性配置の生成結果である。この試行では黒のスーツを着た人間の添景を用いている。結果街路と公園共に生成が成功していた。黒のスーツを着た人のみを公園に配置した A, B の場合では生成が失敗していたのに対し, 人と樹木の同時配置を行なった今回の試行では人の生成がある程度成功していた。樹木を足し

たことで二者の画像間の違いが大きくなったことが、学習と生成の成功の要因になったと思われる。

### 3.2. 生成の失敗例

上記では主に生成の成功例を挙げたが、成功パターンの中には成功と呼べない生成画像が幾つか存在していた。図9がその例である。添景が何も生成されていない画像と、小さいノイズしか生成されていない画像が見受けられた。しかし前者においては「この位置に添景を配置するべきではない」とAIが判断して何も生成が行われなかった可能性がある。小さなノイズの生成も「添景の成り損ない」であると捉えれば、添景の配置位置をある程度読み取ることが可能である。これらの失敗例は、生成に用いた元画像が極端に壁に近い視点の場合に多く見受けられた。原因は学習に用いた訓練画像の中に、上記のような視点で添景が配置されている画像が少なかったためであると考察する。

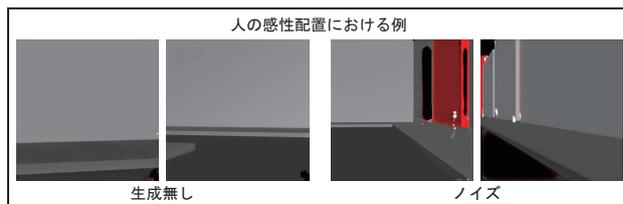


Fig.9 Example of generation failure

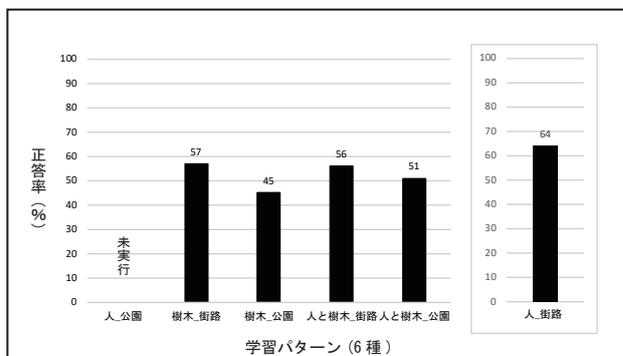


Fig.10 Results of sensitivity evaluation by subjects

## 4. 生成された町並み画像の検証

### 4.1. 被験者による感性評価の条件

先述の試行で生成された都市の町並み画像の感性評価を被験者に行なってもらった。被験者には生成を行なったそれぞれのパターンにつき、乱数配置AIの生成画像は100枚、感性配置の生成画像も100枚。計200枚の生成画像を被験者に分類してもらった。この時見せる生成画像は全て、添景の生成に成功したものを用いる。分類後は200枚の画像が100枚と100枚に等しく分かれるように分類してもらった。またこの時、街路の人添景配置における感性評価では、同じ位置の視点から見た乱数配置生成画像と感性配置生成画像同士の二者択一で分類できるように行なった。それに対し街路の人以外のパターンでは二者択一ではなく無作為にまとめられた200枚の中から自分の感性が反映されていると思うものを選ぶものとした。

### 4.2. 被験者による感性評価の結果

実際に生成画像の生成元を被験者に分類してもらった結果、その正答率は図10のようになった。なお公園における人の生成パターンの感性評価は、生成が失敗に終わってしまったため行っていない。感性評価の結果、無作為にまとめた200枚から感性反映画像を分類した全てのパターンにおいては正答率が50%辺りで収まる結果となった。しかし、二者択一で感性評価を行った街路\_人のパターンにおいては正答率が64%となり、被験者の感性が僅かに生成画像へ反映されていることがわかった。

前者の感性評価において正答率が50%に落ち着いてしまった考えられる原因の1つとして、今回の生成画像の精度が挙げられる。人と樹木の添景の生成は共にある程度の精度を得ていたが(人\_公園のパターンを除く)、一部生成画像では木の幹や人の体が途切れてしまい、添景の配置位置が認識し辛くなっていた。より精度の高い感

性評価を得るには、生成画像の精度の向上が望まれる。

## 5. 終わりに

今回の感性評価では生成画像に被験者の感性が反映されているか明確には分からなかった。しかしpix2pixによる添景の生成には概ね成功と言える結果で終わることとなった。今後は感性評価の精度の向上に努めたい。

### [謝辞]

本研究に用いたデータセットは立命館大学理工学部4回生の安達直登の協力で作成されました。また株式会社梓設計AXチームにもご助言頂きました。この場を借りて深く御礼申し上げます。

### [注釈]

注1)Generative Adversarial Networksの略。

### [参考文献]

- 1)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-pp1331, 2019. 5, 日本建築学会
- 2)Ian J Goodfellow, Jean Pouget-Abadie, MehdiMirza, Bing Xu, David Warge-Farley, Sherjil Ozair, Aaron Couville, Yoshua Bengio:Generative AdversarialNets, <http://papers.nips.cc/paper/5423-generative-adversarial-nets>, 2014, 6
- 3)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
- 4)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, 日本建築学会・情報システム技術委員会第42回情報, システム利用, 技術シンポジウム, 2019
- 5)池之上慎吾, 大野耕太郎, 山田悟史:ボリュームモデルに任意のデザインソースのデザイン性を反映させるファサードデザインAI, 本建築学会・情報システム技術委員会第42回情報, システム利用, 技術シンポジウム, 2019
- 6)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