

# 街路全方位画像ビッグデータを用いた 街路景観の主観的印象評価モデルの構築

小川芳樹\*・沖拓弥\*\*・趙琛渤\*\*\*・関本義秀\*・清水千弘\*\*\*\*

## Development of a Model for Evaluating Subjective Impressions of Streetscapes Using Omnidirectional Street Image Big Data

Yoshiki OGAWA\*, Takuya OKI\*\*, Chenbo ZHAO\*\*\*, Yoshihide SEKIMOTO\*, Chihiro SHIMIZU\*\*\*\*

**Abstract:** The development of a model to evaluate urban streetscapes based on people's subjective impressions are significant for quantitatively understanding people's preferences for streetscapes and for urban planning that people prefer more. In this work, we proposed a deep learning-based multi-label classification model which can estimate each subjective impression from omnidirectional images using the results of a Web questionnaire survey of 22 subjective impressions of omnidirectional images. The proposed model achieved high accuracy for all subjective impressions. Furthermore, applying the developed model to Setagaya-ku, Tokyo to map the distribution of the wide-scale people's subjective impressions at a high-spatial level.

**Keywords:** 街路景観 (street scape) , 深層学習 (deep learning) , マルチラベル分類 (multi-label classification)

### 1. はじめに

都市の街路景観は、住民の行動、犯罪や健康に因果関係があると考えられてきた (Keizer et al. 2008). 例えば、ポイ捨てが更なるポイ捨てを呼ぶ割れ窓理論はその1つである (Wilson and Kelling 1982). これらは街路景観が人の主観的印象評価に影響を及ぼしてきたことを意味する (Goodchild 2011). 街路景観の主観的印象を評価することができれば、都市景観の連続的な変化と都市機能の影響を明らかにすることが可能になる. 加えて、主観的印象評価は土地利用評価 (Levering et al. 2021) や住宅価格 (Xu et al. 2022) の分析などの多様な分野に役立っており、その有用性は高い.

これまでの街路景観に関する主観的印象評価に関する研究は、アンケートや現地調査などのデータ収集方法により実施されてきたが、費用と時間がかかるため広域展開が難しかった. もし同じ指標で様々

な都市の街路景観の主観的印象評価が可能になれば、都市の再開発や都市の変化をモニタリング・評価することができる. そのため、人々がどのように景観を視覚として捉えて、そしてどのようにその情報を知覚として評価しているのかを大規模な都市地域で詳細に把握することは、都市デザインの設計において非常に重要である.

以上の背景から本研究では、全方位街路画像を用いて、道路沿いの撮影地点ごとの街路景観に対する人間の 22 の主観的印象評価を高精度に推定可能なモデルを構築する. さらに構築したモデルを東京都世田谷区の全域に適用して、主観的印象評価の地理的傾向を広域で評価可能なことを示す. さらに関連分析により 22 の主観の関係を明らかにする. 本研究により、街路景観画像を入力するだけで 22 もの主観的印象評価を推定することが可能になり、人間の主観的印象評価を考慮した効率的な都市デザインに寄与するものと考えられる.

\* 正会員 東京大学空間情報科学研究センター (The University of Tokyo)

〒153-8505 東京都目黒区駒場 4-6-1 Tel : 03-5452-6406

\*\* 正会員 東京工業大学環境・社会理工学院 (Tokyo Institute of Technology)

\*\*\* 学生会員 東京大学生産技術研究所 (The University of Tokyo)

\*\*\*\*非会員 一橋大学ソーシャル・データサイエンス教育研究推進センター (Hitotsubashi University)

## 2. 既往研究

本研究が目的とする人間の主観的印象評価は、街路景観に対する個々の人間の視覚からの知覚による主観的反応を定量化するものである。しかし、一般にこのような主観的印象評価は、データ収集コスト・手法の課題から定量化することが困難であるといった欠点がある。この問題に対して Dubey et al. (2016) は、2 枚の画像を比較する形式でアンケート調査を行い、画像ペアの比較を学習可能な深層学習モデルを構築し、一般化を図った。具体的には、110,988 枚の街路景観画像を用いて、6 つの主観（「安全な (Safe)」、「活気のある (Lively)」、「退屈な (Boring)」、「裕福な (Wealthy)」、「気が滅入る (Depressing)」、「美しい (Beautiful)」) に沿った 1,170,000 枚のペア比較を含むクラウドソースデータセットを構築した。そして、定量化タスクを分類タスクに変換するために、二つの画像の似ている、似ていないという指標となる距離を学習するシヤム CNN(Convolutional neural network)を利用したニューラルネットワークを構築した。このシヤム CNN の出力はソフトマックス関数値であり、それは2つの入力画像間の各主観の反応比較である。すなわち、画像 1 が画像 2 よりも「安全な」などに見えるかどうかの相対値を算出することを可能にした。6 つの主観しか対象としなかった問題に対して、沖・木澤 (2022) は、Google street view を用いて、Dubey et al. (2016)によって提案された手法により、22 の主観的印象評価項目に拡張することに成功した。しかし、いずれも既往研究は画像の主観的印象評価スコアを算出する際に物体検出の 1000 クラスを特徴量として用いてモデル化をしているが、一般的にその前のニューロン層の特徴量を用いる方が精度は高いと報告している。また、既往研究で提案されたモデルでは、1 つ 1 つの主観的印象評価項目ごとに個別に学習を行っているために、項目数分のモデルを構築する必要があるため、学習コストが非常に高く、学習モデルを適用する際のコストも大きいために、実用的でない。さらに、個別に学習することで項目間の関係を無視したモデルとなっている。本研究は、これらの一連の問題を

解決し、すべての項目を同時に入力して項目間の関係を考慮して学習することで、より精度の高いモデルの開発を目指す。本研究で提案しているモデルは、マルチラベル分類モデルとして、すべての項目を同時に学習することで1つのモデルからすべての主観的印象評価を出力することが可能になり、より実用的なモデルとして構築できたものと考えている。

## 3. データ

### 3.1 街路全方位画像

本研究で用いる緯度経度付きの街路全方位画像は、株式会社ゼンリン (2013 年) の東京都世田谷区の 610,017 枚を用いる。また、東京都世田谷区では、都市部の景観に限られるため、高知県高知市の山岳や田園風景を含む 50 枚をアンケート調査データ用に追加した。画像は、車両の屋根上に搭載された全方位 (360 度) カメラが 7.5m 間隔で街路を走行しながら撮影したものである。このデータには、画像ごとに GPS により測位された緯度・経度、車両方向、撮影日時などのテキスト情報が付属している。オリジナル画像は、パノラマ化された縦 2,700 pixel、横 5,400pixel の jpeg tar フォーマットである (表 1)。また画像の下部約 35%もしくは 45%の領域は、撮影車

表 1 対象とする街路全方位画像データの仕様

項目	仕様
フォーマット	パノラマ化した Jpeg tar フォーマット
サイズ	横 5,400 × 縦 2,700 画素
撮影高	200cm
撮影ピッチ	250cm
位置情報	緯度、経度、方位
対象地域	東京都世田谷区

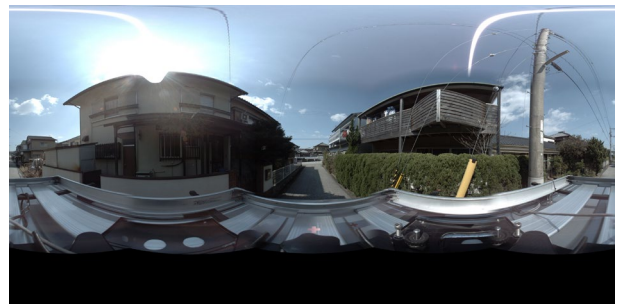


図 1 本研究で用いる街路全方位画像の例

両の屋根部分が映り込んでいる。街路全方位画像の例を図1に示す。

### 3.2 主観的印象評価付き街路全方位画像データ

街路景観に関する主観的印象評価データを集めるために、Web アンケート調査を実施した (図2)。調査方法は、沖・木澤 (2022) の手法を参考にした。被験者には、ランダムに並べて表示された2枚の街路画像ペアを比較しながら、画面中央にランダムに提示される22問の評価項目に対する印象を、5段階で評価するというプロセスを繰り返してもらった。なおアンケート調査で用いる画像は、街路全方位画像の中央部分を切り出し、レンズの歪みを補正した。本研究で対象とする評価項目は、既往研究を網羅的にレビューし、評価項目対象として取り扱われることの多い22項目を抽出した (図2)。被験者の利用する端末は、スマートフォンまたはタブレット、画面は縦向きに限定した。アンケート調査に用いた画像枚数は1,000枚であり、多様な土地利用の画像を対象とするために、国土数値情報の用途地域データを用いて、世田谷区に存在する全10種類の用途地域の画像枚数が均等になるように画像を抽出した (図3)。画像の組み合わせ数は10,000ペア、対象となる属性区分は20代以下男女、30代男女、40代男女、50代男女、60代以上男女の合計10セグメントであり、1画像ペアあたりのチェック人数は40人 (各セグメントにつき4人) になる。以上により主観データセットには、 $10,000 \times 22 \times 40 = 8,800,000$  サンプルの回答総数が含まれる。

## 4. 手法

### 4.1 街路景観の主観的印象評価モデル

街路景観を撮影した街路全方位画像を用いて人間の主観的印象評価をモデル化するには、ペア画像をモデルが学習して、各街路画像の主観的印象評価スコアを22の主観印象評価項目 (図2) に沿って予測し、スコアの差異を算出する必要がある。そのため本研究では、人々の主観的印象評価の予測問題を

をマルチラベル分類タスクとして定式化する。つまり、ペア画像ごとに相対的な22の主観的印象評価スコアを1つのモデルから予測する必要がある。アンケート調査の10,000ペアの画像は、22の主観的

- 22の主観的印象項目
- 景観に関する質問
  - (1) 開放的な
  - (2) 親しみのある
  - (3) 活気のある
  - (4) 快適な
  - (5) 緑が豊かな
  - (6) 落ち着いた
  - (7) 明るい
  - (8) 昔ながらの
  - (9) 安全な
  - (10) すっきりした
  - (11) 生活感のある
  - (12) 居心地の良い
  - (13) 清潔な
  - (14) 美しい
  - (15) 裕福な
  - (16) 退屈な
  - (17) 気が滅入る
  - 魅力に関する質問
  - (18) 好きな
  - (19) 面白い
  - (20) 住みたい
  - (21) 通りたい
  - (22) 魅力的な



図2 本研究でのWebアンケート調査画面と対象とする主観的印象評価の調査項目



図3 アンケート調査に用いた街路全方位画像の分布。なお地図に表示されていないが、都市部だけでなく田舎の風景を考慮するために高知県の画像を用いた。

印象評価項目のすべてをカバーしているため、主観的印象評価は、1つのモデルのソフトマックス関数の出力値からマルチラベル分類に変換することが可能である。各主観的印象評価は、「美しい」1つを取っても、「とても美しい」や「やや美しい」に代表されるように大きさがあるため、バイナリ分類モデルよりもソフトマックス関数を用いて、主観的印象スコアの大小を確率的に示す方が実用的であると考えられる。

マルチラベル分類モデルでは、22の主観的印象評価を一度に学習し、22の出力値が得られる。本研究で用いるマルチラベル分類モデルは、シャムCNNであり、このタイプのアーキテクチャは画像の変化や比較を対象とする問題タスクで使用されている (Daudt et al. 2018; Dong and Shen 2018)。そのため、2つのペア画像間の主観的印象評価を比較するために使用するのに適している。本研究で構築した人間の主観的印象評価モデルのアーキテクチャを図4に示す。

まず、ImageNet データセットで事前学習済みモデルを用いて、全レイヤーの重みを固定した Freezing バックボーン (Brock et al. 2017) を利用した上で2枚の画像を別々に入力する。このバックボーンに画像を入力することで1024次元の2つのベクトル (Feature vector  $D=1024$ ) を出力することができる。次に、この2つのベクトルを重み共有層 (Weights shared) に入力することで、主観的印象評価スコアを意味する2つの22次元ベクトル

(subjective score  $D=22$ ) を学習する。最後に、画像2の22次元ベクトルを画像1のベクトルから差分を計算し、最終的な22次元の主観的印象評価結果を得ることができる。すなわち、各画像ペアにおい

て、画像1が画像2よりも「開放的」、「魅力的な」などの22の主観的印象評価スコアを得ることができる。なお損失値は、22次元ベクトルの結果と22の主観的印象評価アンケート結果との間のバイナリクロスエントロピー (BCE) として計算される。またバックボーンは、CNNから代表的な最先端バックボーンである ConvNeXt (Liu et al. 2022) の基本バージョン (ConvNeXt-B) を選択した。

学習が収束した後、このシャムCNNの2つの枝は重みを共有しているため、いずれかの枝を主観的印象評価スコアの推定モデルとすることができる。

## 4. 2 景観要素と主観的印象評価の関係分析

### 4. 2. 1 画素単位の景観要素の推定

街路全方位画像の景観要素を画素単位で推定するためにセマンティックセグメンテーションを行う。本研究では、アノテーションが不要かつ、セマンティックセグメンテーションのオープンソースのデータセット (CityScapes データセット (Cordts et al. 2016)) で事前学習済みモデルの SegFormer (Xie et al. 2021) を用いて、セマンティックセグメンテーションを実装した。SegFormer のアーキテクチャを図5に示す。Vision transformers (Dosovitskiy et al. 2020) に属する SegFormer の変換器は、画像の自己相関の学習をブロックする自己注意メカニズムにより、エンコーダ部の性能を強化し、SegFormer は2021年に Cityscapes データセットで 83.48% mIoU という最高精度を達成した。また、CityScapes データセットは、道路、歩道、建物、壁、塀、ポール、信号、交通標識、植生、地形、空、人物、ライダー、車、トラック、バス、列車、バイク、自転車と

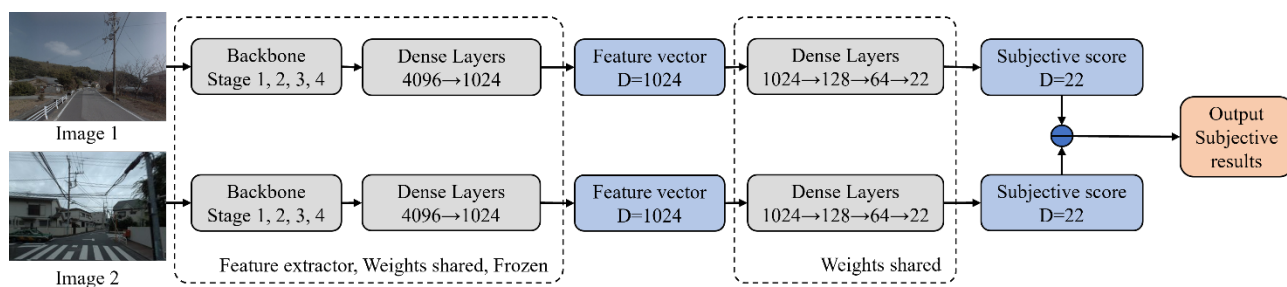


図4 主観的印象評価モデルのアーキテクチャ

いう 19 の画素単位のセマンティクスを含む大規模データセットである。この 19 の画素カテゴリは、都市の街路景観のほとんどをカバーしており、景観と人間の知覚スコアの定量的関係を分析する研究に十分な情報を得ることができる。

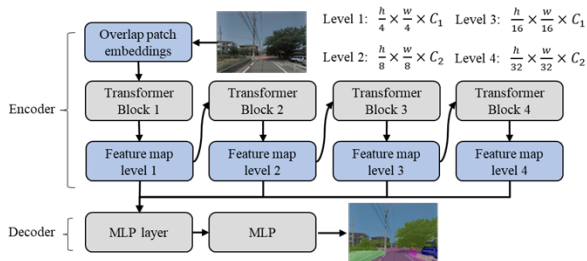


図 5 SegFormer のアーキテクチャ

#### 4. 2. 2 ℓ1 正則化スパースモデリングに基づく主観的印象評価の関係分析

ここでは、画像ごとの主観的印象評価項目のソフトマックス関数の出力値（主観的印象評価スコア）を目的変数とし、また 19 種類の画素単位の景観要素を説明変数として重回帰分析を行う。重回帰分析は、通常の線形重回帰モデルに ℓ1 正則化のペナルティ項を追加し、負の対数尤度を最小化するアプローチを用いる (Hastie et al. (2015))。ℓ1 正則化の重要な特性は、スパース解が得られることである。スパース解は、重回帰係数ベクトルの非ゼロ成分が比較的少ない解のことであり、この性質を利用してコヒーレント成分を同定し、通常の線形重回帰分析では考慮出来なかった多重共線性を考慮することができる。スパース性は、回帰係数の 0 値成分の割合として定義され、ペナルティが大きくなるにつれて増加する。この手法により、主観的印象評価項目の潜在的に寄与する景観要素を抽出し、その影響を評価することが可能になる。

### 5. 実験と結果

#### 5. 1 主観的印象評価推定と精度

この実験では、主観データセットの 10,000 ペアを 60%、20%、20%の割合で訓練用、検証用、テスト用のデータに分割し実験を行った。なお、計算

は、データ活用社会創成プラットフォーム mdx

(Suzumura et al. 2022) 上で行った。結果を表 2 に示す。

表 2 の Conv NeXt-B, 22 に示すように、22 の主観的印象評価推定において正解率 0.7958~0.9112 の高い精度を達成した。主観的印象評価項目により精度のばらつきが見られ、0.900 以上の高精度で推定できた項目としては、「開放的な」、「活気のある」、「緑が豊かな」、「明るい」、「昔ながらの」、「すっきりした」、「生活感のある」、「清潔な」であった。一方、精度が 0.800 未満の項目は、「面白い」だけであった。このように主観的印象項目によって精度にばらつきがある理由としては、人間がこれらの主観をどのように視覚として理解し、知覚に変換しているかといった測定上の揺らぎと、そもそも人々の嗜好の異質性が生まれやすい対象と同質的な嗜好が表明されやすい対象があると考えられる。特に本研究におけるアンケート調査の被験者は、年齢性別を幅広く等量に選んでいるため、被験者のこれまでの経験や生まれた土地の文化等の属性が知覚に大きく影響している可能性がある。例えば、「開放的な」という主観的印象評価項目は被験者間で一致する傾向があるが、「面白い」とは一致しない傾向がある。これは、人間が年齢とともに景観に対する嗜好が変化することと同様の傾向であり、ばらつきが発生する 1 つの要因であると考えられる (Kienast et al. 2012)。

また、既往研究のモデルとの精度比較も行った (表 2)。Dubey et al. (2016) と沖・木澤 (2022) のモデル (VGG16) は、主観的印象評価項目ごとにモデルを構築する必要があり、また評価項目が 6 項目であったため 22 項目に拡張し、項目ごとの精度を算出した。さらに彼らのモデルを本研究と同じバックボーンの ConvNeXt に変更したモデルを構築し比較することで、個別分類モデルと本提案手法であるマルチラベル分類モデルの比較結果を示すことにした。その結果、本提案のマルチラベル分類モデルは、すべての主観的印象評価項目に対して最も精度が高い結果が得られた。この結果は、個別の主観的印象評価項目にフィッティングさせることなく、22

表 2 提案モデル（マルチラベル分類モデル（ConvNeXt-B,22））と  
 既往研究モデル（個別項目ごとの分類モデル（VGG16, ConvNeXt-B））の精度

モデル	開放的な	親しみの ある	活気のある	快適な	緑が豊かな	落ち着いた	明るい	昔ながらの	安全な	すっきり した	生活感の ある
VGG-16, 個別	0.8763	0.8373	0.8798	0.8455	0.8803	0.8518	0.8623	0.8633	0.8448	0.8675	0.8685
ConvNeXt-B, 個別	0.9060	0.8618	0.8957	0.8652	0.9000	0.8815	0.8892	0.8873	0.8615	0.8947	0.8940
ConvNeXt-B, マルチ	<b>0.9112</b>	<b>0.8712</b>	<b>0.9020</b>	<b>0.8842</b>	<b>0.9062</b>	<b>0.8995</b>	<b>0.9082</b>	<b>0.9057</b>	<b>0.8630</b>	<b>0.9137</b>	<b>0.9038</b>

モデル	居心地の 良い	清潔な	美しい	裕福な	退屈な	気が滅入る	好きな	面白い	住みたい	通りたい	魅力的な
VGG-16, 個別	0.8373	0.8603	0.8563	0.8398	0.8105	0.8313	0.8230	0.7590	0.8385	0.8240	0.8233
ConvNeXt-B, 個別	0.8585	0.8848	0.8925	0.8608	0.8522	0.8648	0.8587	0.7947	0.8635	0.8665	0.8595
ConvNeXt-B, マルチ	<b>0.8783</b>	<b>0.9000</b>	<b>0.8928</b>	<b>0.8640</b>	<b>0.8540</b>	<b>0.8730</b>	<b>0.8692</b>	<b>0.7958</b>	<b>0.8725</b>	<b>0.8765</b>	<b>0.8660</b>

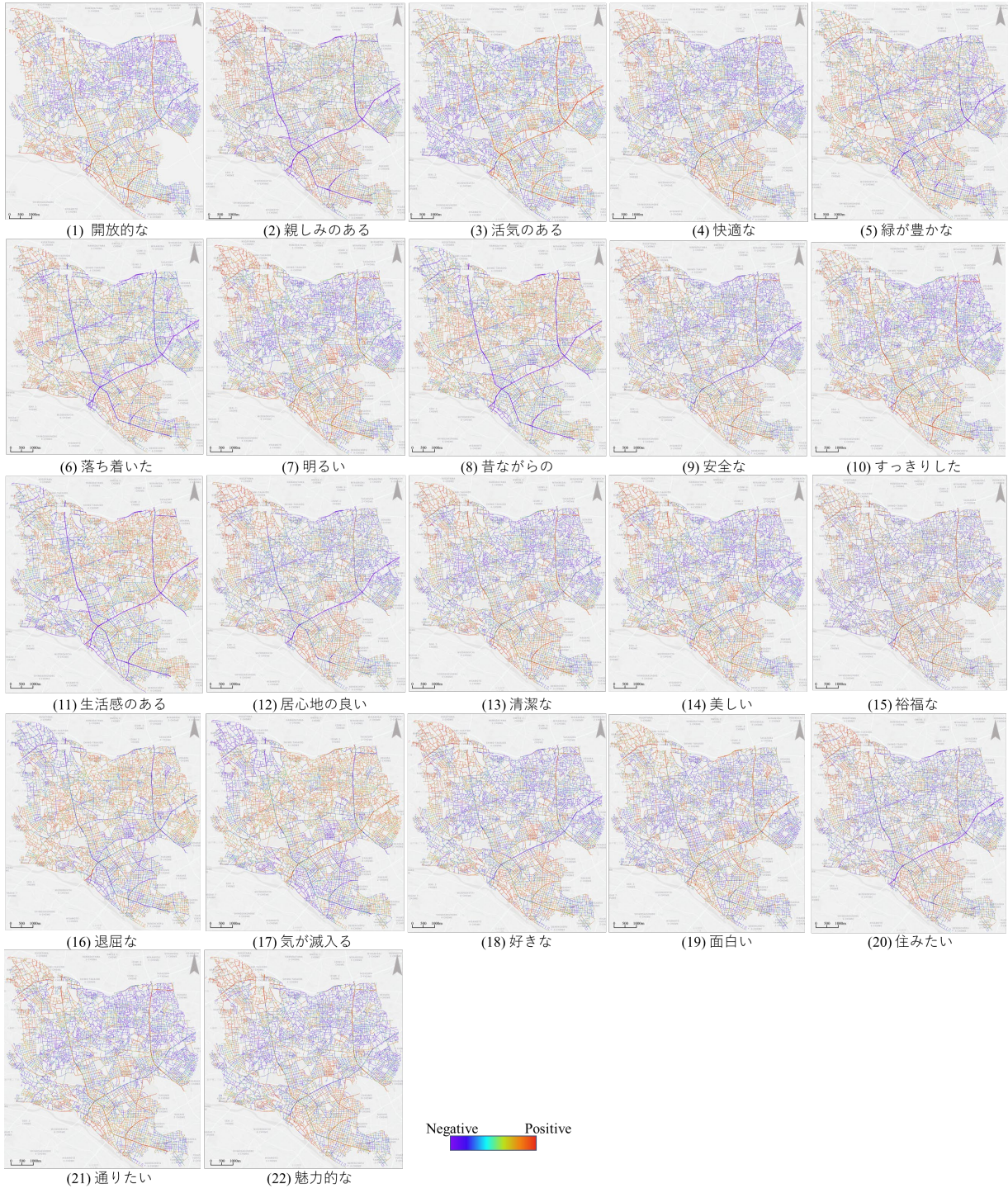


図 6 東京都世田谷区における 22 の主観的印象評価推定結果の可視化

の全ての主観的印象評価項目の相互作用を学習し、低計算コストで高精度に推定できる強力なモデルであることを示唆している。単一のモデルが個別モデルより高精度な理由としては、22の主観的印象評価項目全てを同時に入力して学習することで、22項目間の関係性を保持したまま個別項目より多くの情報を学習できたことが挙げられる。

## 5. 2 東京都世田谷区における主観的印象評価マップ

構築した学習済みの主観的印象評価モデルを用いて、東京都世田谷区全域の街路全方位画像に適用して画像主観的印象評価の分布を推定した。推定結果は連続値のスコアであり、スコアが大きい程、対象の主観的印象評価項目に対してポジティブであることを示している。本研究では、道路を7.5m地点毎に撮影された街路全方位画像、約61万枚を用いて、空間解像度の高い予測結果を獲得する。図6は、世田谷区における22の主観的印象評価項目のマップを示している。

この図より、大きく2つのエリアに分かれていることが見てとれる。1つ目のエリアは、地図の北東部（下北沢駅、明大前、三軒茶屋駅周辺）で、渋谷駅からのアクセスが良く、商店が多く建ち並び、道路と建物が密で若者に人気なエリアである。2つ目のエリアは、地図の北西部（千歳烏山駅、仙川駅）と南部のエリア（二子玉川駅周辺）で、駅周辺が比較的新しく落ち着きがあり、比較的面積が大きな住宅が立ち並ぶ住宅街エリアである。

賑やかなエリアが住宅街エリアよりもスコアが高かった主観的印象評価項目は、「昔ながらの」、「生活感がある」、「気が滅入る」であった。一方で、住宅街エリアが北東部エリアよりもスコアが高かった主観的印象評価項目は、「開放的な」、「快適な」、「緑が豊かな」、「落ち着いた」、「明るい」、「安全な」、「すっきりした」、「居心地の良い」、「清潔な」、「美しい」、「裕福な」、「好きな」、「住みたい」、「通りたい」、「魅力的な」などのポジティブな主観的印象評価項目であった。これは北西部と南部

エリアの住宅街が北東部エリアよりも道路幅員も大きく、比較的新しい建物が立ち並んでいることが要因と考えられる。2つのエリアで差が見られなかった主観的印象評価項目としては、「親しみがある」、「活気のある」であった。「親しみがある」は、郊外の落ち着いたエリアだけでなく下北沢駅周辺のような賑やかなエリアでのスコアが高く、幹線道路沿いにおいて低い傾向があった。また「活気のある」は、各駅周辺の密集した道路を含む商業地域と幹線道路沿いにおいて高い傾向があった。この結果は、Jacobs (1992)や Fan et al. (2008)の中国の北京・上海での結果と一致しており、活気がある街作りをする際に街区を小さく設計する計画手法とも一致する。

## 5. 3 主観的印象評価項目間の相関関係とクラスタリング結果

5.2節の結果より主観的印象評価項目間を可視化することで似た地理的傾向があることが分かった。そこで、22項目の主観的印象評価に対して相関分析を行った。図7は、主観的印象評価項目間のデータをプロットした散布図とその相関係数を示している。地理的傾向と同様に、「開放的な」に対して、「明るい」、「安全」、「すっきりとした」、「清潔」、「美しい」、「好きな」、「通りたい」は、相関係数が大きく(0.8以上)、「裕福な」、「退屈な」、「面白い」は相関係数が小さい(0.3以下)結果を示した。また、「昔ながら」、「生活感のある」、「面白い」などのいくつかの項目は、他のいずれの項目とも相関が弱く、独立していることがわかる。この結果は、沖・木澤 (2022)とも類似した結果である。

次に、主観的印象評価項目間の関係を視覚的に理解するために主成分分析を適用して22次元の相関係数行列を用いて2次元に削減し、第一主成分(PC1)と第二主成分(PC2)の各主観的印象評価項目の主成分得点を可視化した。第一主成分の寄与率は約84.1%であり第二主成分の寄与率は約9%であった。主観的印象評価構造が類似した項目は、近くにプロットされるため、「魅力的な」、「住みたい」、「居心地の良い」などのポジティブな主観的印象

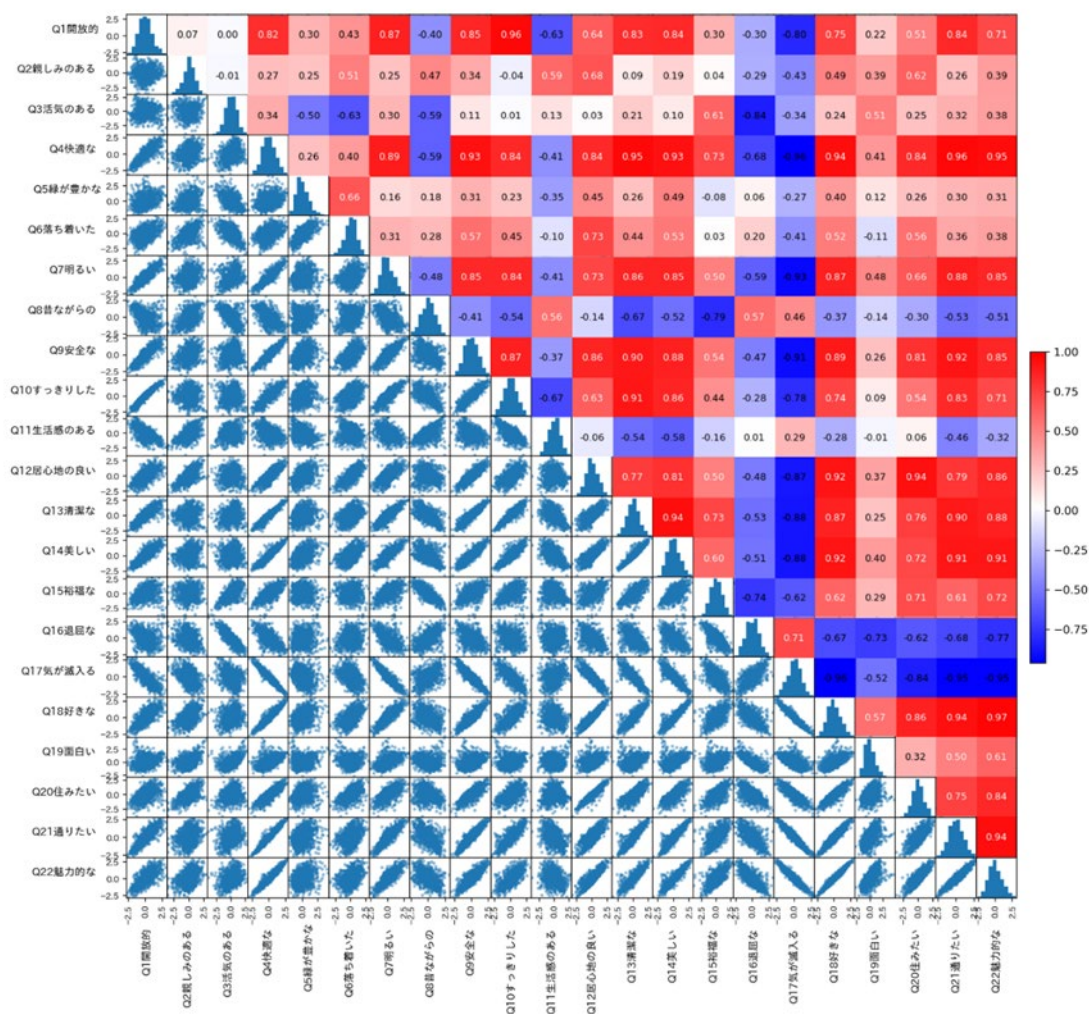


図7 街路景観の主観的印象評価項目間の散布図と相関係数

象評価項目と「気が滅入る」、「退屈な」、「昔ながらの」、「生活感のある」などのネガティブな主観的印象評価項目は、それぞれ近い位置にまとまっており、対立関係にある指標であることが見てとれる。そこで、主観的印象評価項目を数的にクラスタリングすることで類似構造の主観的印象評価項目をグループ分けした(図8)。その内訳をみると、Class0は最も大きなクラスターを形成しており、「好きな」を中心に、景観に対してポジティブな主観的印象評価項目が含まれている。そこで、主観的印象評価項目を数的にクラスタリングすることで類似構造の主観的印象評価項目をグループ分けした(図8)。その内訳をみると、Class0は最も大きなクラスターを形成しており、「好きな」を中心に、景観に対してポジティブな主観的印象評価項目が含まれている。Class1

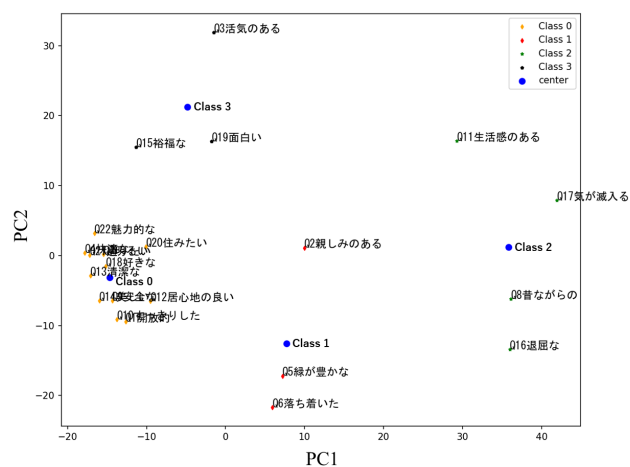


図8 主成分分析とクラスタリングの結果



は、「緑が豊かな」を中心に、景観に対して癒しの主観的印象評価項目が含まれている。Class 2 は、「昔ながらの」を中心に、景観に対して「退屈な」などのネガティブな主観的印象評価項目も含まれている。Class 3 は、「面白い」を中心に、景観に対してユニークな主観的印象評価項目が含まれている。

#### 5. 4 主観的印象評価項目と景観要素の関係分析

図9には、クラスタリングのClass 0-3ごとの主観的印象評価と景観要素との間の $\ell_1$ 正則化スパースモデリングによる重回帰分析結果を示す。ここでは、19の景観要素の中で、主観的印象評価に正または負に寄与した景観要素であるスパース解を示す。Classによって影響の受ける景観的要素は、様々であることが分かる。また $\ell_1$ 正則化に基づく重回帰分析によって得られたスパース解は、Class 0, 2, 3において7種類の景観要素が得られ、Class 1において5種類の景観要素であった。Class 0のポジティブな主観的印象評価項目に対しては、道路(Road)、歩道(Sidewalk)、空(Sky)、地面(Terrain)、緑(Vegetation)が正に寄与し、建物(Building)と車(Car)は負に寄与している。Class 1の癒しの主観的印象評価項目は、Class 0と似た傾向があるが、道路の寄与率が大きく異なり、その分、緑が大きく正に寄与しており直感に一致する結果である。Class 2のネガティブな主観的印象評価項目は、Class 0, 1と対照的な結果になっている。またClass 3の「面白い」などのユニークな主観的印象評価項目は、道路、車、建物、歩道といった人工的な景観要素が正に寄与しており、空が負に寄与する結果となった。全体として、空と緑の自然的要素が主観的印象評価項目に大きく影響していると言える。例えば、緑視率は快適性の増加に大きく影響していることと一致する。また天空率は、商店街のような建物が込み入ったところでは低いため、「活気のある」に対しては、負に大きく影響していることが考えられる。一方で、空の面積比率によってポジティブな主観的印象評価に相関するだけでなく、Class 2のようなネガティブな主観的印象評価にも相関する場合があることが判明した。これらの

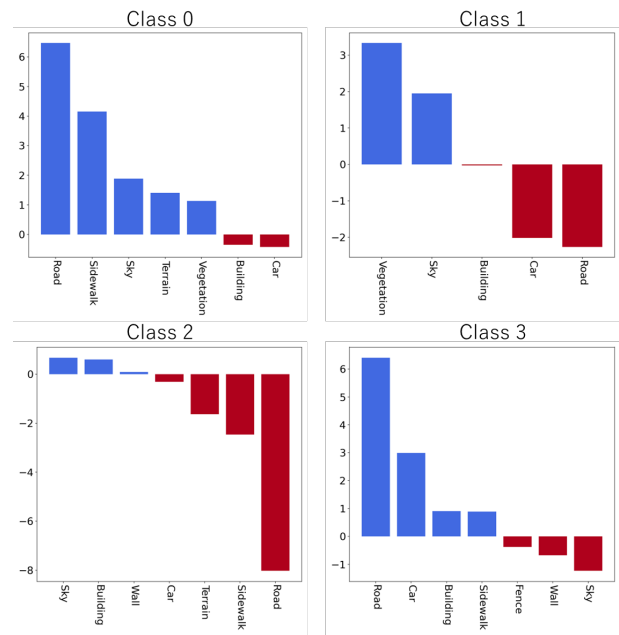


図9 主観的印象評価のClassと景観的要素の関係分析結果

結果は、都市をデザインする際に、どのような主観的印象評価項目を目的にするかを明らかにして、緑と空の面積比率を調整することが有効であることを示唆している。

## 6. まとめ

本研究では、街路景観に対する主観的印象評価を行うために、街路全方位画像用いて22項目の主観的印象評価を推定するモデルを提案した。その結果、提案モデルは一つのマルチラベル分類モデルであるがすべての項目に対して、従来の個別分類モデルと比べて良好な精度で主観的印象評価を推定できることを確認できた。今後は、年齢・性別属性を考慮したモデルを構築すると共に、外国人の被験者に対してもアンケート調査を実施し、文化や国籍の違いによって主観的印象評価が変化するか等を検証する予定である。

## 謝辞

本研究を行うにあたり、JSPS20K15001、JSPS22K04490 科研費の助成を受けた。株式会社ゼンリンには本研究で用いた街路全方位画像データを提

供して頂いた。関係各位に謝意を表する。

## 参考文献

- 沖拓弥・木澤佐椰茄 (2022) 画像ビッグデータと大規模被験者アンケートに基づく住宅地における街路印象評価推定モデル. 日本建築学会計画系論文集, 87 (800), 2102-2113.
- Brock, A., Lim, T., Ritchie, J. M., and Weston, N. (2017) Freezeout: Accelerate training by progressively freezing layers, arXiv preprint arXiv:1706.04983.
- Cordts, M., Omran, M., Ramos, S., Rehfeld, T., Enzweiler, M., Benenson, R., and Schiele, B. (2016) The cityscapes dataset for semantic urban scene understanding, In Proceedings of the IEEE conference on computer vision and pattern recognition, 3213-3223.
- Daudt, R. C., Le Saux, B., & Boulch, A. (2018) Fully convolutional siamese networks for change detection, In 2018 25th IEEE International Conference on Image Processing (ICIP), 4063-4067.
- Dong, X., and Shen, J. (2018) Triplet loss in siamese network for object tracking, In Proceedings of the European conference on computer vision (ECCV), 459-474.
- Dosovitskiy, A., Beyer, L., Kolesnikov, A., Weissenborn, D., Zhai, X., Unterthiner, T., and Houshy, N. (2020) An image is worth 16x16 words: Transformers for image recognition at scale, arXiv preprint arXiv:2010.11929.
- Dubey, A., Naik, N., Parikh, D., Raskar, R., and Hidalgo, C. A. (2016) Deep learning the city: Quantifying urban perception at a global scale, In European conference on computer vision, 196-212, Springer.
- Fan, Z., Zhou, B., Liu, L., Liu, Y., Fung, H.H., Lin, H., Ratti, C., (2018) Measuring human perceptions of a large-scale urban region using machine learning, *Landscape and Urban Planning*, 180, 148-160.
- Goodchild, M. F. (2011) Formalizing place in geographic information systems, In Communities, neighborhoods, and health, 21-33, Springer.
- Hastie, T., Tibshirani, R., and Wainwright, M. (2015) Statistical Learning With Sparsity: The Lasso and Generalizations, Chapman & Hall/CRC.
- Jacobs, J. (1992) The death and life of great American cities. 1961. New York: Vintage, 321-325.
- Keizer, K., Lindenberg, S., and Steg, L. (2008) The spreading of disorder, *Science*, 322(5908), 1681-1685.
- Kienast, F., Degenhardt, B., Weilenmann, B., Wäger, Y., and Buchecker, M. (2012) GIS-assisted mapping of landscape suitability for nearby recreation, *Landscape and Urban Planning*, 105(4), 385-399.
- Levering, A., Marcos, D., and Tuia, D. (2021) On the relation between landscape beauty and land cover: A case study in the UK at Sentinel-2 resolution with interpretable AI, *ISPRS Journal of Photogrammetry and Remote Sensing*, 177, 194-203.
- Liu, Z., Mao, H., Wu, C. Y., Feichtenhofer, C., Darrell, T., and Xie, S. (2022) A ConvNet for the 2020s, arXiv preprint arXiv:2201.03545.
- Suzumura, T., Sugiki, A., Takizawa, H., Imakura, A., Nakamura, H., Taura, K., et al. (2022) mdx: A cloud platform for supporting data science and cross-disciplinary research collaborations, The 8th IEEE International Conference on Cloud and Big Data Computing.
- Wilson, J. Q., and Kelling, G. L. (1982) Broken windows, *Atlantic monthly*, 249(3), 29-38.
- Xie, E., Wang, W., Yu, Z., Anandkumar, A., Alvarez, J. M., and Luo, P. (2021) SegFormer: Simple and efficient design for semantic segmentation with transformers, *Advances in Neural Information Processing Systems*, 34, 12077-12090.
- Xu, X., Qiu, W., Li, W., Liu, X., Zhang, Z., Li, X., and Luo, D. (2022) Associations between Street-View Perceptions and Housing Prices: Subjective vs. Objective Measures Using Computer Vision and Machine Learning Techniques, *Remote Sensing*, 14(4), 891.