

# 深層学習による街路景観の広域分析技術の開発

## Development of a Method for Wide-area Analysis of Streetscapes Using Deep Learning

佐野 祐士 Yuji Sano\*1 安藤 邦明 Kuniaki Ando\*2

### 梗概

まち規模の広域を対象とする街路景観分析が可能なニューラルネットワークを開発した。景観印象評価用の教師データ構築のため、約1,000枚の景観画像データベースの印象アノテーションを、約100名の建築設計者・都市計画プランナーらにより行った。渋谷・久屋大通公園を対象とした街路景観分析により、エリアごと・時刻ごとの景観印象の変化とその妥当性を確認した。

キーワード：街路景観，印象評価，深層学習，まちづくり

### Summary

We have developed a neural network capable of analyzing streetscapes over a wide area on a town scale. In order to construct teacher data for landscape impression evaluation, a database of approximately 1,000 landscape images was annotated by 100 architectural designers and urban planning planners. The change of landscape impression by area and by time and its validity were confirmed by analyzing the streetscapes of Shibuya and Hisaya-odori Park.

Keywords: streetscape, impression evaluation, deep learning, urban planning

## 1 はじめに

昨今、人口減少や少子高齢化に起因する地域の活力低下が懸念される中、都市の魅力向上とにぎわい創出が、多くの都市に共通して求められている。人々が訪れたいくなるウォークブルで魅力的なまちなか空間を作るためには、歩行空間の魅力向上に資する効果的な評価手法と設計技術が必要である。

国土交通省の推進する「居心地が良く歩きたくなる」まちなかづくりの施策では、グランドレベルデザインに必要な「空間デザイン」の項目として、景観に関するガイドライン<sup>1)</sup>を設定しており、景観に基づくまちづくりの重要性は高い。そこで、景観について求められる評価手法・設計技術の仕様を整理した結果、重要と考えられる下記の2点をターゲットとし、Fig. 1に示す街路景観の評価ツールを考案した。

- ①計画者が実務で使いやすい評価手法・設計技術であること
- ②計画者が印象評価とそれに紐づく要素・特徴の相関性を分析可能なこと

本来、主観評価である印象の定量的な評価基準作成は困難であるが、近年発達してきた深層学習技術により実現が容易となった。これまでに街路名や訪問意欲を対象としたニューラルネットワークは既に検討されているが<sup>2)</sup>、どのような景観構成要素が印象に寄与しているかは明らかでない。本ツールを用いて、街路景観画像から印象評価の土台と考えられる要素・特徴を画像から抽出し、人間の印象判断を学習したAIにより印象を評価したのち、抽出された要素・特徴を印象評価結果と突合させることで、設計がどのように与える印象に寄与するかを検討できる。また、単一設建物だけでなく広域に適用することで、まちの魅力や課題から開発計画を洗練させることが可能である。

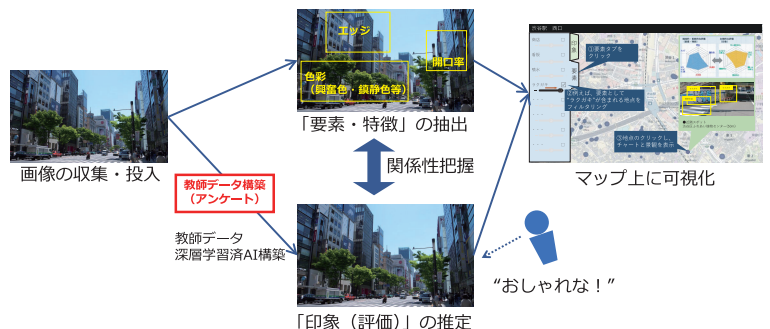


Fig. 1 考案した街路景観評価ツールの概要  
Overview of the proposed streetscape analysis tool

\*1 技術研究所 研究主任 博士(人間情報学) Senior Researcher, Research & Development Institute, Dr. Hum. Info.

\*2 技術研究所 主任研究員 Chief Researcher, Research & Development Institute

本ツールの開発に先立ち、事前に計画担当者に機能ニーズをヒアリングしたところ、景観に基づくまちづくりを実現するためには、来街者の属性ごとに印象評価が可能で、景観以外のまちの魅力を評価するデータと突き合わせられることの2点が強いニーズとして確認された。本ツールの開発からニーズ対応までの研究全体の構想をFig. 2のように3 STEPに分割し、まずSTEP 1の「深層学習による街路画像を用いた景観分析ツールの開発」における開発概要と印象評価用教師データの作成、そして街路景観分析AI学習と可視化を実施した。

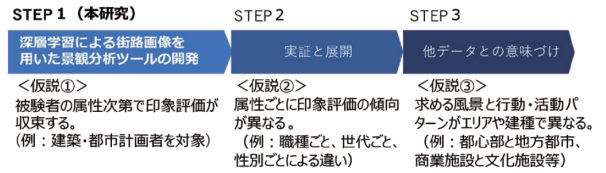


Fig. 2 研究全体の構想  
Overall research concept

## 2 開発概要

### 2.1 街路景観評価ツールの概要と景観画像収集エリアの選定

本ツールでは、撮影した景観画像、または設計者が作成したパースを入力することで、「要素・特徴」の抽出と「印象」の評価を実施する。位置情報が含まれる景観画像は、マップ上に可視化する。

本研究では、渋谷区の渋谷駅周辺部、竹下通り、表参道、奥渋谷エリアを対象とした。渋谷区は住宅街と繁華街がコンパクトにまとまっており、マップ上に可視化した際に効果的な比較が可能だと考えている。

景観画像の評価項目をTable 1に示す。評価項目は、キャプション評価法<sup>3)</sup>に倣い設定した大項目と、当社の建築設計者・都市計画プランナーへのヒアリングをもとに設定した中・小項目からなる。既往研究<sup>4)</sup>では「地域らしさ」を評価項目としていたことから、印象項目に今回対象とした渋谷エリアに対応した「渋谷らしい」を追加した。

### 2.2 画像収集方法

Fig. 3に画像撮影装置を示す。撮影範囲は当社都市設計プランナーが重要視する渋谷駅周辺のストリートとし、台車型の撮影装置で360°画像を手動で撮影した。撮影期間の2021/7/21~9/13のうち、午前・午後ともに晴天日となった4日間と、曇天の1日間の計5日間の日中で撮影を行った。撮影日は7/21 (最高気温33.6℃)、8/3 (同32.9℃)、8/4 (同34.5℃)、8/5 (同34.7℃) が晴天、9/13 (同30.7℃) が曇天であった。撮影地点数は合計で約11,000地点となった。撮影された360°画像は既往研究<sup>5)</sup>を参考に水平方向90°、鉛直方向74°でトリミングし、前方方向のものを利用した。

## 3 印象評価用教師データの構築

ニューラルネットワークの学習のためには、手本となる教師データが必要である。印象評価用教師データ構築のため、既往研究<sup>5)</sup>を参考に、アンケート調査による画像の印象ラベリング（以下、アノテーション）を実施した。Fig. 4のようなアンケート回答システムを構築し、約100名の建築設計者・都市計画プランナーを被験者として、約4,000枚の景観画像の印象を評価させた。ランダムに表示される2枚の景観画像を見比べて印象尺度が高いものを選択させるタスクを、1つの印象評価項目ごとに100回、8項目で計800回回答させ、合計約80,000回の回答を得た。

データセット内の画像がどの程度アノテーションされているか確認するため、アノテーションされた回数をヒス



Fig. 3 街路景観撮影装置  
Street Scene Imaging Equipment



Fig. 4 印象評価アノテーションを行うアンケート回答システム  
Questionnaire response system for impression evaluation annotation



トグラムとした。例として，“おしゃれな・センスのある”のアノテーション時のヒストグラムをFig. 5に示す。

今回のデータセットにおいて、すべての画像の組み合わせにアノテーションするには表示枚数を4,000以上、総計800万以上のアノテーションが必要であるが、アノテーション結果に基づいた画像序列が特徴量化可能な明確な傾向のあるものであれば、十分にアノテーションされていると判断できる。Fig. 6に示した“おしゃれな・センスのある”と選択された上位画像を示す。樹木の木漏れ日と建築のファサードが調和していたり、適度な人数の歩行者が含まれていたりすることなどから選択されたと思われる、アノテーションした設計者・計画者からも納得感のある結果であるとフィードバックされた。これより、教師データは十分にアノテーションされ、印象評価に深層学習で特徴量化が可能な明確な傾向があると判断した。

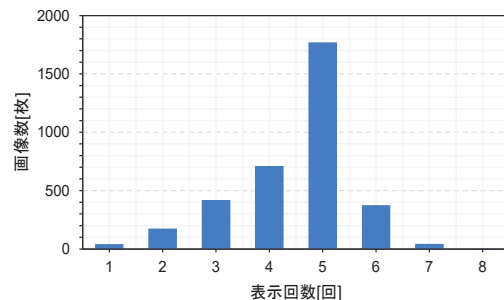


Fig. 5 アノテーションにおける画像の表示回数分布  
Distribution of the number of times an image appears in the annotation



Fig. 6 印象アノテーションされた画像例  
Example of impression-annotated image

Table 1 景観画像の分析・評価項目  
List of landscape image analysis and evaluation items

大項目：＜要素＞		大項目：＜特徴＞	
中項目	小項目	中項目	小項目
人	人	人	アクティビティ構成 (歩く、立つ、座る)
建築	窓	建築	窓開口比率
	外壁		建物凹凸
	路面店舗 (飲食・物販)		エッジ
	テント・オーニング		文字 (数・言語)
	ショップボード		色多様性
	店頭陳列物		材質構成
	看板 (袖・壁面)	インフラ	破損度合
	室外機		ラクガキ (面積比率等)
物産	自然・地形	緑視率・天空視率	
自販機			
インフラ	車道 (舗装)・歩道	大項目：＜印象＞	
	電柱・電線	個別評価	
	路駐・放置自転車	中項目	小項目
	街灯	個性	おしゃれな・センスのある
	ごみ置き場・ごみ	日常性	落ち着きがある
	ガードレール・ボラード	整然性	ごちゃごちゃしている
	ストリートファニチャー (テーブルイス)	親近性	親しみやすい
		植物	総合評価
自然・地形	空	中項目	小項目
	水面	開放性	賑わいのある
		快活性	不安になる
		その他	居心地が良い
			渋谷らしい

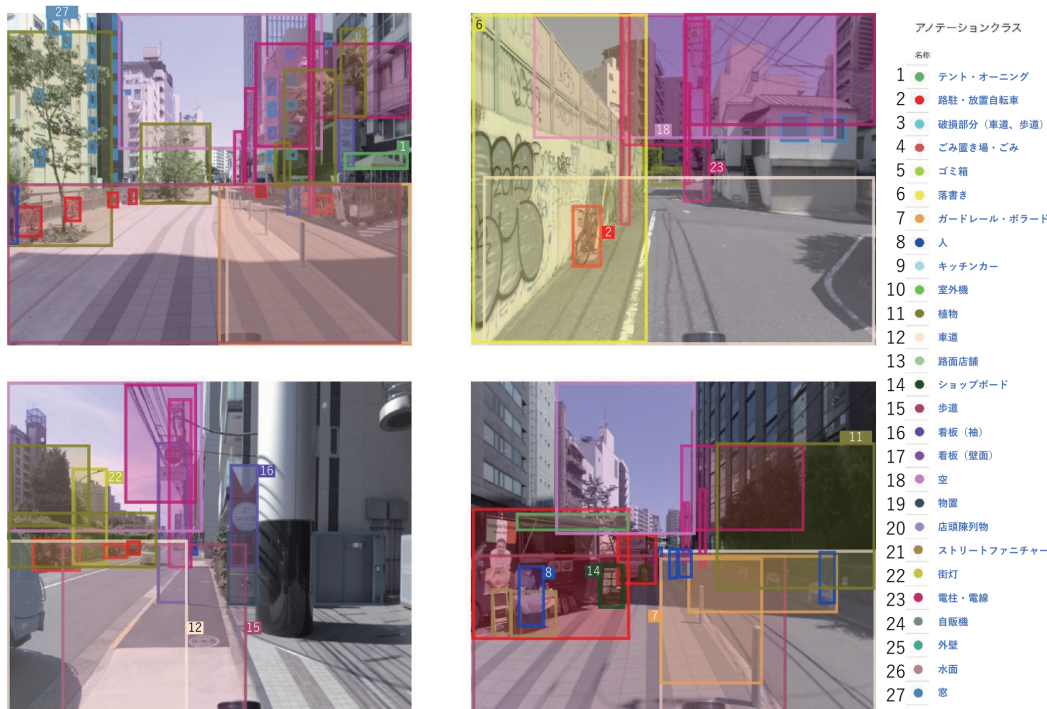


Fig. 7 風景画像に対する要素アノテーションの例  
Example of element annotation for landscape images

## 4 街路景観分析AIの学習

Table 1に示す出力とする街路景観分析AIの学習について述べる。要素項目については、Table 1に示した要素項目を2.2節で収集した画像約11,000枚に対してFig. 7のようにアノテーションし、事前学習済みのEfficientNet<sup>6)</sup>に追加で学習させた。教師データ数が100に満たない項目は既存の公開画像データセットで補完した。

特徴項目のうち色多様度は、解像度16の色相ヒストグラムを、シン普森の多様性指数<sup>7)</sup>を用いて表した。窓開口比率は外壁検出矩形面積に対する窓検出矩形面積の比とした。緑視率と天空率、建物凹凸は、画像内の検出物を塗り分けるネットワークDeepLab v3+<sup>8)</sup>を用いて塗り分け、それぞれ植物と空の面積率、建物領域の輪郭の標準偏差とした。エッジはソーベルフィルタ<sup>9)</sup>で抽出し、抽出後画像のエントロピーを算出した。

印象項目は、画像を比較し評価可能なモデルに3章でのアノテーションのように各印象項目で比較した際の基準画像に対する勝率が出力となるよう学習させた。基準画像は、教師データ中からランダム抽出した500枚の平均画像 (Fig. 8) を用いた。



Fig. 8 街路画像を500枚重ね合わせて平均化した基準画像  
Reference image averaged from 500 superimposed street images

## 5 街路景観分析AIの出力分析

Table 2 要素の抽出精度 (再現率)  
Element extraction accuracy (reproducibility)

要素	再現率	要素	再現率	要素	再現率	要素	再現率
キッチンカー	1.00	歩道	0.96	路面店	0.84	電柱電線	0.58
ゴミ箱	1.00	バイク	0.95	落書き	0.82	袖看板	0.52
自転車	1.00	オーニング	0.90	自動車	0.80	窓	0.51
自販機	1.00	空	0.90	店頭陳列物	0.77	人間	0.45
ストリートファニチャー	1.00	壁面	0.89	ショップボード	0.72	ゴミ	0.00
破損 (路面)	1.00	植物	0.88	ガードレール	0.60	水面	0.00
室外機	0.97	街灯	0.87	壁面看板	0.60	物置	0.00
車道	0.96	ゴミ置き場	0.84				

### 5.1 要素の抽出

要素の物体検出精度はAverage Precision (AP)などで検証するのが一般的であるが、設計・計画においては検出位置より個数が重要であるため、抽出された各要素の個数が教師データのアノテーション個数と一致した画像の割合を示す再現率で検証した。

Table 2に示す各要素項目の再現率は、バイクやオーニングなどは再現率が高い一方で、ゴミや物置、水面などは再現率が0となっている。これは、収集した画像に上記の要素が数枚~数十枚と極端に少なく、他のデータセットから補完し100枚以上を学習したものの、景観画像内に含まれる要素としての教師データが不足したことが主な原因と推測される。ゴミについては定義のあいまいさによるアノテーション不足も影響があると考えられる。不足要素を含む景観画像をさらにアノテーションするほか、水面などの面的要素についてはセマンティックセグメンテーションを用いることで改善が期待される。

Fig. 9に要素の抽出結果を示す。画像内の枠に記載された数値は、物体認識の信頼度 (AIの自信度合)を示している。要素抽出の傾向として、抽出結果に一定の妥当性はあるものの、画像内の中央から左上領域に偏って物体認識がなされていることがわかる。このため、画像全体を満遍なく抽出できるようモデルを改善する必要がある。

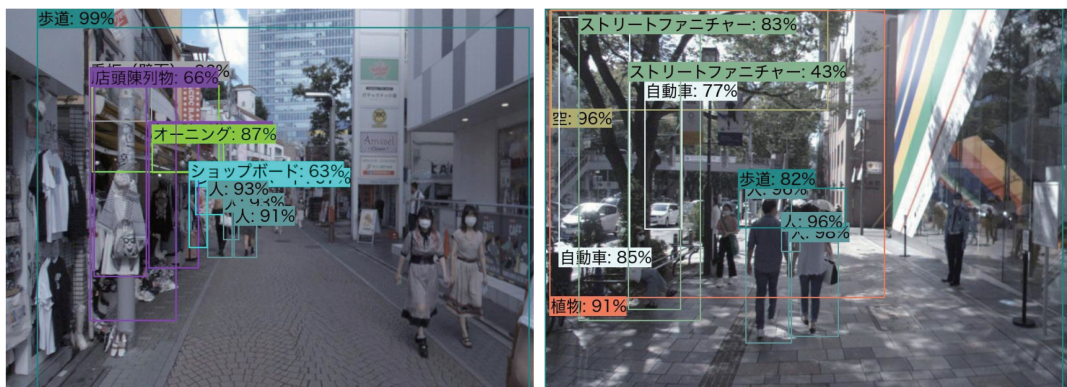


Fig. 9 要素抽出AIによる街路景観要素の検出例  
Example of detection of streetscape elements by element extraction AI



### 5.2 印象評価の結果

4章で学習したAIの出力分析を行うため、渋谷区のストリートを対象に撮影した約11,000枚の街路景観画像を入力した。Fig. 10にAIに入力した全画像の勝敗表から順位表を作成した際の印象評価の最上位・最下位の例を示す。本AIは印象値を直接数値化する回帰モデルではないため、Fig. 8の基準画像に対する勝率を印象値と定義することで疑似的に回帰モデルのような結果を得ることができる。

一方で、基準画像が真に標準的な景観画像であるかどうかは未知であるため、分類モデルとしてのふるまいと疑似的な回帰モデルとしてのふるまいを、入力画像群の全組み合わせにおいて一致するか確認することで、基準画像としての妥当性を検証した。教師データと推論結果の一致率を比較正答率と定義し、各印象項目で評価した (Table 3)。概ね70%の正答率となったが、印象評価がほぼ同値の画像についても比較正答率を検証しているため、印象評価に十分な差がある画像についてはより高い比較正答率であると考えられる。

### 5.3 印象評価と要素・特徴との関係

2節の出力のうち、主な25項目の相関係数をFig. 11に示す。Fig. 11の結果から、印象項目に相関 ( $|R| \geq 0.5$ ) がみられる要素・特徴項目は、人、植物、緑視率、エッジ (エントロピー) であった。人については、賑わいや渋谷らしさに正の、不安になる度合いに負の相関が出ており、人流の多さが与える印象と対応している。

植物の要素はおしゃれ度と親しみやすさに相関がみられた。緑視率は植物の検出面積に影響されるため、同様におしゃれ度と親しみやすさに相関がみられたが、エッジ (エントロピー) にも同様の相関に加え、落ち着き度にも相関がみられた。整備された植栽がポジティブな印象を与えるとともに、木漏れ日がエッジ成分を増加させていたと推測される。また、印象項目内でも相関がみられ、不安と渋谷らしさ・賑わい、おしゃれな・センスのあると親しみやすさに特に相関があらわれた。

## 6 街路景観分析マップによる可視化

「計画者が実務で使いやすい評価手法・設計技術」とするためには、街およびストリートに分布する景観の特徴を捉えることが重要であると考え、5章で得られた結果をWebアプリケーションで可視化した。可視化した街路景観分析マップをFig. 12に示す。本アプリでは左側の評価項目を選択してフィルタリングができ、右側のウィンドウではレーダーチャートにより地点間比較が可能である。

今回対象としたストリートの中で一際特徴的であったのは表参道で、Fig. 10に見られるように、多



Fig. 10 AIによる印象評価の最上位・最下位の例  
 Examples of top and bottom impression ratings by AI

Table 3 印象評価の比較正答率  
 Comparative Correctness Rate for Impression Rating

印象項目	比較正答率
おしゃれな・センスのある	67%
落ち着きがある	66%
ごちゃごちゃしている	67%
なじみやすい	65%
賑わいのある	74%
不安になる	72%
居心地がいい	64%
渋谷らしい	65%

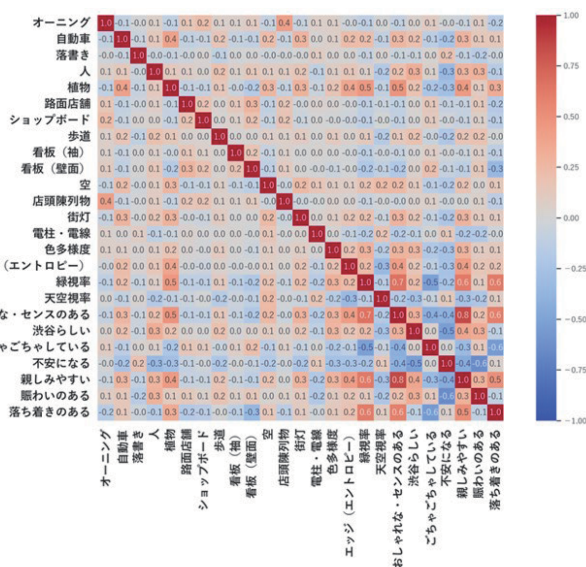


Fig. 11 要素・特徴・印象項目間の相関係数グリッド  
 Correlation coefficient grid among elements, characteristics, and impressions



くの印象項目で最上位または最下位を占めていた。表参道の印象と相関のある要素・特徴項目を考えると、例えばエッジと“おしゃれな・センスのある”に相関が見られることがFig. 11からわかる。エッジの特徴量をマップで可視化したところ、確かに表参道では高い値を示していた。

選択地点の詳細

評価項目

選択地点+プロット  
(評価値で色付け)

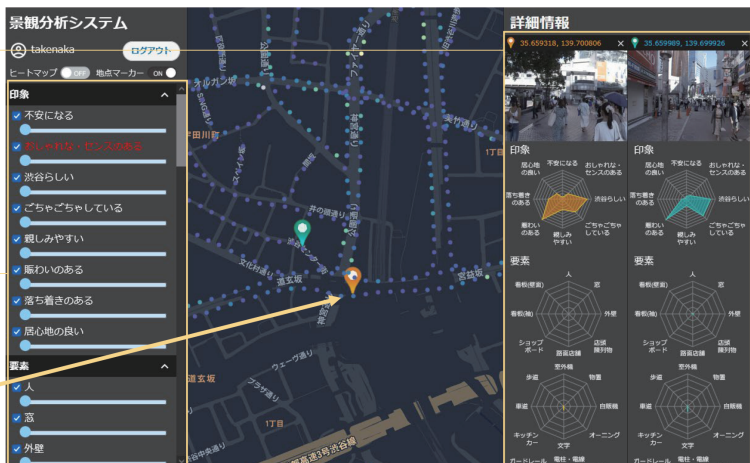


Fig. 12 景観分析結果の可視化 Web アプリケーション  
Web application for visualization of landscape analysis results

表参道の写真は木漏れ日を多く含んでいたため、エッジ特徴量を表すエントロピー増大に影響したと考えられる。

このように計画者が本ツールを用いることで、街やストリートの景観に対する印象を捉えながら、相関する要素や特徴について分析が可能である。

## 7 街路景観分析AIの性能評価のための実地検証

開発した街路景観分析AIの性能評価のため、開発事業の従前従後比較の用途を想定し、2022年10月にイベントが開催されていた名古屋市内の広場にて、イベント開催期間の印象の変化を評価した。本イベントの景観印象を測定・評価するにあたって、定点撮影と移動撮影の2手法で撮影を行った。

定点撮影では、Brinno社製カメラ TLC200を用いて、Fig. 13に示す3ヶ所（A、B、C地点）で10分/枚のインターバル撮影を行った。A地点はイベント空間の出入口の一つ、B地点は音楽ライブなどメインイベントが行われるスペース、C地点は飲食可能な滞留スペースである。撮影期間のうち平日、休日から2日ずつを選出し、日の出から消灯までに撮影された画像を対象として、撮影地点の用途（移動往来、イベント、滞留）によって景観印象の時間変化を捉えられているか評価分析を行った。

移動撮影では、Insta360社製のInsta360 ONE RSを用いてイベント空間とその周辺を、イベント期間中の土曜日の10時・13時・16時の3回、360°動画モードで撮影した（Fig. 13）。撮影された動画の正面方向を水平角80°、垂直角60°でトリミングし、分析対象とした。

本研究では、分析手法①：朝（6～11時）・昼（11～17時30分）・夜（17時30分～22時）の画像群をAIに読み込ませ、広場内の時間帯ごとを分析（定点撮影の画像）、分析手法②：画像群をAIに読み込ませ、経路による変動を分析（移動撮影の画像）、の2つの手法をもとに、“おしゃれな・センスのある”、“落ち着いたのがある”、“ごちゃごちゃしている”、“親しみやすい”、“賑わいのある”、“不安になる”、“居心地の良い”の計7項目で評価を行った。

定点撮影で撮影された画像群の印象評価値を、撮影された画像を母集団として偏差値を算出した。イベントにおいて重視されていた“賑わいのある”についてFig. 14に示す。

A地点とB地点において休日昼の時間帯（赤枠）に、偏差値51～60に偏りが見られ、休日昼は平日の同時時間帯と比較して高い評価値と



Fig. 13 景観評価の実地検証のための定点・移動撮影  
Fixed-point and mobile photography for verification of landscape assessment

なった。また、B地点は照明のあるエリアを映しており、夕方のイベント参加者が集まったことで、A地点よりも“賑わいのある”評価値が高くなったと考えられる。

Fig. 14では、C地点で偏差値に偏りが生じているが、これはFig. 13で示したような仮設物の多い公園がAIの学習データから外れていたため、AIが景観変化を評価できなかったと考えられる。

また、Fig. 15に示す平日と休日の印象評価の比較結果から、什器や出店などの要素が、他地点よりも多かったB地点において“ごちゃごちゃしている”という印象が一貫して高い評価値になっている。

分析手法②の結果についても、分析手法①と同様に偏差値を算出した。全3回の撮影のうち、特に広場周辺にて利用者が多かった回である、13時から撮影した画像の“賑わいのある”に関して、イベント空間とその周辺のエリアで評価の高かった画像をFig. 16に、全体を通してのシーケンス評価をFig. 17に示す。Fig. 16 (a) はイベントの中心となった地点で出店の商品が並び参加者が多く写り込んでいる画像、Fig. 16 (b) はベンチやパラソルが置かれた商業施設の前の画像であり、どちらも人が多く“賑わいのある”という主観に一致した画像であるといえる。Fig. 17からは“賑わいのある”の評価は下側でバラツキの大きい分布であったため、60以上の偏差値の変動を見ることはできなかったが、随所で建物の裏や閑散とした広場などで低い偏差値となった。

Fig. 15, 16に示したように平日と休日、さらに時間帯ごとに評価を行い、Fig. 17に示したような連続した景観画像を評価する中で、印象の変化を捉え比較することができた。連続した場所の撮影を行うことで、Fig. 18のように特定の場所における印象の特徴を表現することも可能になった。一方、Fig. 14ではC地点にてAIによる評価結果が不自然に偏っている箇所が見受けられた。

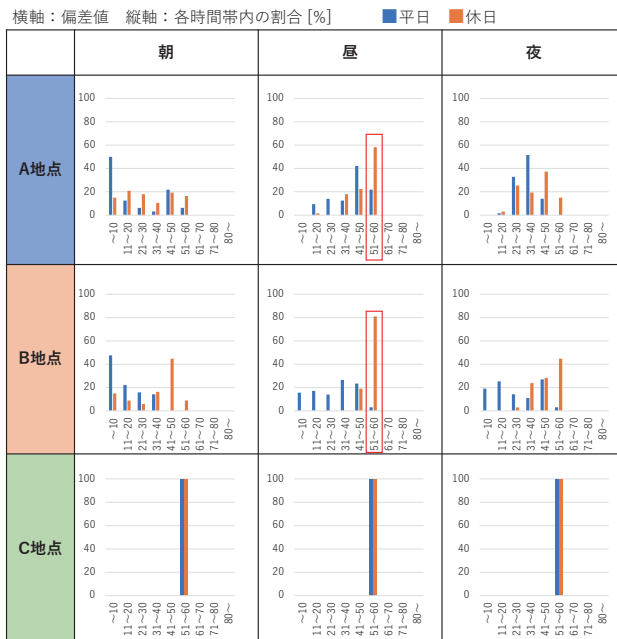


Fig. 14 イベント空間の“賑わいのある”印象についての印象分析結果  
Results of impression analysis of the “liveliness” of the event space

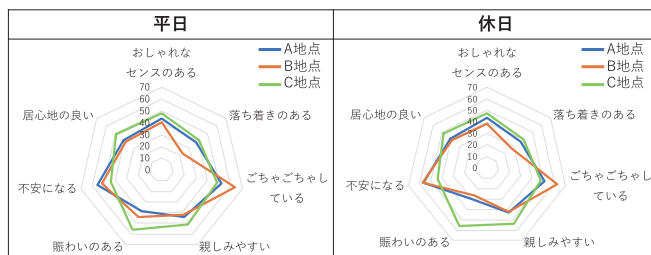


Fig. 15 イベント空間の定点撮影の印象評価結果  
Evaluation results of fixed-point photography of event space

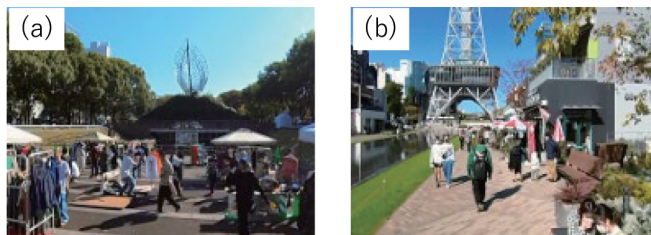


Fig. 16 “賑わいのある”と高く評価された景観画像  
Landscape images highly rated for “liveliness”

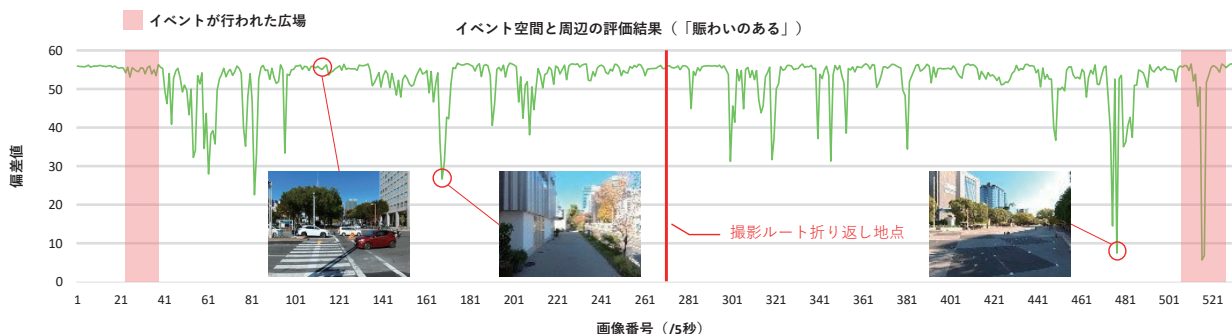


Fig. 17 “賑わいのある”項目についてのイベント空間と周辺の評価結果  
Results of the evaluation of the event space and surrounding area for “liveliness”

これらを踏まえて、現状の限られた学習データに、多様な構図の景観画像を盛り込むことで、AIによる評価の偏りを改善していくことができれば、開発事業の従前従後比較といった用途においても景観印象を定量的に評価することが可能になると考えられる。

## 8 まとめ

本研究では、歩行空間の魅力向上に資する景観の効果的な評価手法として、「①計画者が実務で使いやすい評価手法・設計技術であること」「②印象評価とそれに紐づく要素・特徴の相関性を分析可能なこと」をターゲットに、

街路景観画像から要素・特徴を抽出し、画像に対する印象を推論する評価ツールを提案した。深層学習による景観分析を実現すべく、約100名の建築設計者・都市計画プランナーによる景観印象評価を学習させ、要素・特徴・印象の合計44項目の出力が得られるAIを構築した。その結果、対象エリアにおける景観印象8項目に相関(|R| ≥ 0.5)がみられる要素・特徴項目は、人、植物、緑視率、エッジであることが明らかとなり、②を実現した。また、得られた結果を街路景観分析マップで可視化し、ストリートの特徴を捉えられていることを確認し、①を実現した。

開発した街路景観分析AIの性能評価のため、開発事業の従前従後比較の用途を想定し、実証試験として名古屋市内光塩で開催されていたイベントの開催期間中の印象の変化を評価した。その結果、時間帯による印象の特徴や連続した景観画像によるシーケンシャルな印象の変化を捉えることができた。一方で、教師データ不足と思われる印象評価の失敗もみられ、仮設物や公園を含む多様な景観を追加で学習することで改善する可能性がある。

本ツールの開発により、計画者が街の景観の特徴を広域に捉えることができようになり、街になじみ、より街の魅力を引き立てる街路景観を計画可能となる。また、計画者の考える街路景観に対する印象に対して、本ツールによって必要な要素や特徴項目を検討可能になると期待される。

## 参考文献

- 1) 国土交通省：「居心地が良く歩きたくなる」まちなか創出に向けた道路空間利活用に関するガイドライン，2022年4月
- 2) 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，Vol.84，No.759，pp.1323-1331，2019年5月
- 3) 小島隆矢，古賀誉章，宗方淳，平手小太郎：多変量解析を用いたキャプション評価法データの分析：都市景観の認知と評価の構造に関する研究 その2，日本建築学会計画系論文集，2002年12月
- 4) 百里美和：街路景観における印象評価指標の体系化－夜間街路景観からの考察，東京大学大学院新領域創成科学研究科修士論文，2006年3月
- 5) P. Salesses, K. Schechtner, and C. A. Hidalgo. “The collaborative image of the city: mapping the inequality of urban perception,” *PloS one*, Vol.8, No.7, p.e68400, 2013.
- 6) M. Tan, and Q. V. Le. “EfficientNet: Rethinking model scaling for convolutional neural networks.” International conference on machine learning. PMLR, 2019.
- 7) E. H. Simpson. “Measurement of diversity.” *Nature*, No.163, Vol.4148, p.688, 1949.
- 8) R. Azad, et al. “Attention deeplabv3+: Multi-level context attention mechanism for skin lesion segmentation.” European conference on computer vision. Springer, Cham, 2020.
- 9) N. Kanopoulos, N. Vasanthavada, and R. L. Baker. “Design of an image edge detection filter using the Sobel operator.” *IEEE Journal of solid-state circuits*, No.23, Vol.2, pp.358-367, 1988.

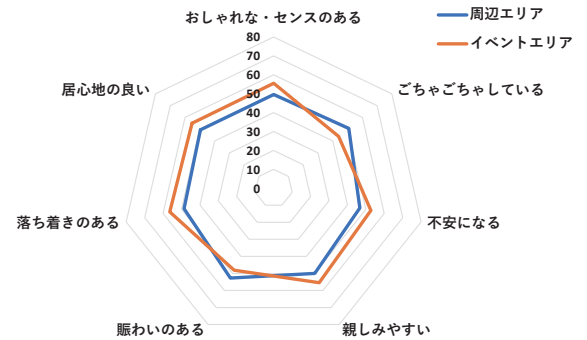


Fig. 18 イベント空間と周辺エリアの評価結果  
 Results of evaluation of event space and surrounding area