

# 衛星画像と所得に関する統計情報を用いた建物単位の所得水準の推定

山野寺瞭太\*・秋山祐樹\*\*・宮崎浩之\*\*\*

## Estimation of Income Levels per Building Unit Using Satellite Image and Income Statistics

Ryota YAMANOTERA, Yuki AKIYAMA, Hiroyuki MIYAZAKI

Many cities in developing countries are currently facing the rapid population growth. However, the lack of comprehensive statistics on current population distribution makes it difficult to devise effective urban planning. In recent years, efforts have been made to understand population distribution by estimating buildings where people reside using satellite images. However, since the number of household members is strongly correlated with household income, there is a demand for methods to estimate the income level of each building. While there are existing studies that attempt to estimate household income using interview survey results as training data, these efforts have not been validated using reliable statistics. Therefore, this study proposes and verifies a method to estimate the income level of individual buildings using satellite images, focusing on Japan where income-related statistics are well-maintenance.

**Keywords:** 衛星画像 (satellite image), 深層学習 (deep learning), 所得水準 (income level), 建物 (building), 統計 (statistics)

### 1. はじめに

開発途上国の多くの都市では、近年急激な人口増加という課題に直面している。そのような中で、都市の持続的な開発・発展を続けていくためには、都市の実態を定量的かつ正確に把握する必要がある。そして、それを可能にする情報の一つとして人口統計が挙げられる。しかし、開発途上国の人口統計は地理的にカバーする範囲が限られていたり、更新の頻度が低い、あるいは不定期であったりする場合もある。そのため、現状の人口分布を正確に把握することが困難である。

そこで、これらの国では高い解像度で人口分布を把握する技術が求められている。この問題を解決するために、建物データを用いて建物単位の人口を推定する手法が挙げられる。例えば、Akiyama et al. (2019) は、詳細な建物データと人口統計が整備された先進国で開発した、建物面積と階数に基づく人口推定モデルを途上国に外挿する手法を提案している。

しかしこの手法は、対象地域の建物データが十分に整備されていることが前提となる。そのため、建物データが十分に整備されていない都市や地域には、この手法を適用することは困難である。また、衛星画像から人口が分布する建物を推定することで、人口分布を把握する手法の研究が進められている。Kaiser et al. (2017) は、深層学習を用いて衛星画像から建物の分布を把握する手法を開発しており、同手法により人がどこに集中して住んでいるのかを大まかに把握することができるようになった。

一方、世帯人数は建物の大きさや分布だけでなく、世帯所得とも大きな関係があることが知られているため、建物ごとの所得水準を推定する手法も求められている。しかし、開発途上国には正確な所得水準に関する統計が存在しない場合が多い。そのため、まずは建物単位の所得水準を把握する手法を開発する必要がある。建物単位の所得水準の把握の取り組みとしては奥田ほか (2019) による、現地での聞き

\* 学生会員 東京都市大学総合理工学研究科建築・都市専攻 (Tokyo City University)

〒158-0087 東京都世田谷区玉堤1丁目28-1 E-mail: [g2381644@tcu.ac.jp](mailto:g2381644@tcu.ac.jp)

\*\* 正会員 東京都市大学建築都市デザイン学部都市工学科 (Tokyo City University)

〒158-0087 東京都世田谷区玉堤1丁目28-1 E-mail: [akiyamay@tcu.ac.jp](mailto:akiyamay@tcu.ac.jp)

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

〒277-8568 千葉県柏市柏の葉5-1-5 E-mail: [heromiya@csis.u-tokyo.ac.jp](mailto:heromiya@csis.u-tokyo.ac.jp)

1 取り調査から推定した建物単位の所得水準の情報を  
2 用いて教師データを作成し、衛星画像と深層学習を  
3 用いて建物単位の所得レベル推定がある。しかし、  
4 聞き取り調査はごく限られた地域を対象としており、  
5 また所得に関する精緻なデータを用いた検証は行わ  
6 れていない。

7

### 8 1.1. 本研究の目的

9 以上の課題を解決するためには、所得に関する大  
10 規模かつ信頼性の高い統計が利用可能な地域におい  
11 て、建物単位の所得水準を推定するモデルを構築し、  
12 その信頼性を検証する必要がある。そこで本研究で  
13 は、所得に関する信頼性の高い統計が全国規模で整  
14 備されている日本を対象に、衛星画像を用いて建物  
15 単位の所得水準を推定する手法の提案とその検証を  
16 行うことを目的とする。

17

## 18 2. 研究手法

### 19 2.1. 概要

20 まず、住宅土地統計調査と国勢調査を用いて、個々  
21 の小地域における世帯の平均所得データ（以下「小  
22 地域世帯所得データ」）を作成した。次に、ゼンリン  
23 の住宅地図から得られた建物ポリゴンデータと小地  
24 域世帯所得データを空間結合することで、個々の建  
25 物にその建物が立地する小地域の平均世帯所得を属  
26 性として付与した。さらに、家計調査を参考に所得  
27 を3つのランクに分類し、所得水準ラベルとすること  
28 とで、教師データのラベルとして使用した。また、  
29 それらの建物が中心に来るように衛星画像を切り出  
30 し、作成した所得水準ラベルと対となるようにする  
31 ことで、教師データを作成した。続いて、VGG-16と  
32 呼ばれる学習済み深層学習モデルと教師データを用  
33 いて、所得水準レベルの分類モデルを構築した。こ  
34 の時の学習方法としては、転移学習を用いた。最後  
35 に、構築したモデルと学習に使用していないデータ  
36 を用いて精度の検証を行った。

37

### 38 2.2. 対象地域

39 本研究の対象地域は東京都新宿区（以下「新宿区」）  
40 とした。理由は、様々な高さの建物が高密度に分布

41 しているという地理的条件が、開発途上国の都市部  
42 に見られる条件に類似しているためである。また、  
43 日本では信頼性の高い所得に関する統計が利用可能  
44 であることも、新宿区を対象地域とした理由である。  
45 対象地域全域の衛星画像を図-1に示す。

46

## 47 3. 使用したデータ

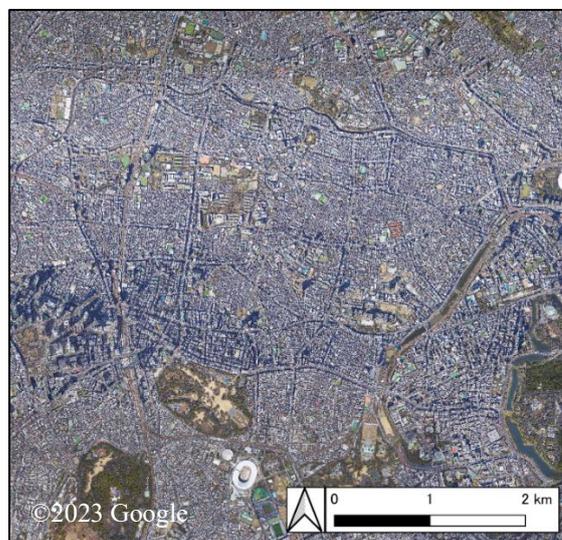
### 48 3.1. 小地域世帯所得データ

49 小地域世帯所得データは、個々の建物に経済レベ  
50 ルの属性を付与するために使用した。国勢調査(2015  
51 年)と住宅・土地統計調査(2018年)を用いて作成  
52 した。手法の詳細は山中ほか(2021)を参照された  
53 い。なお、山中ほか(2021)では地域メッシュ単位  
54 の世帯所得を推定しているが、本研究では山中ほか  
55 (2021)の手法を小地域に適用することで、小地域  
56 世帯所得データを作成した。図-2に新宿区における  
57 小地域世帯所得データを示す。

58

### 59 3.2. 建物ポリゴンデータ

60 株式会社ゼンリンの住宅地図(Zmap TOWN II:  
61 2020年)に収録された建物レイヤが持つ建物の形状  
62 を表現するポリゴンデータを建物ポリゴンデータと  
63 して使用した。なお、建物ポリゴンデータは建物用  
64 途に関する属性も含まれている。そこで、本研究で  
65 は居住用の建物を対象とするため、表-1に示す用途  
66 の建物のみを使用した。



67

68

図-1 東京都新宿区の様子

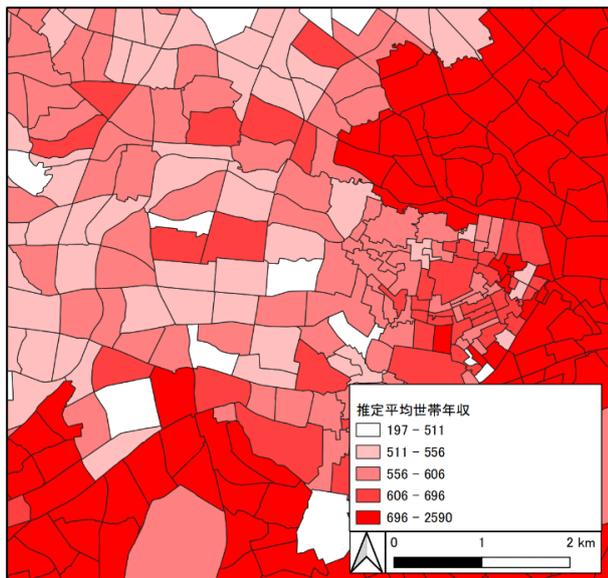


図-2 東京都新宿区の小地域世帯所得データ

表-1 建物ポリゴンデータの用途別件数

属性種別 コード	建物用途	建物数
1363	ビル・アパート等	37,424
1364	戸建て住宅	34,567
1365	事業用建物	14,283



図-3 本研究で使用した衛星画像の例

表-2 所得水準レベルの内訳と該当する建物数

所得水準 レベル	世帯所得帯 (万円)	建物数
0	350 未満	37,424
1	350 以上 750 未満	34,567
2	750 以上	14,283

### 3.3. 衛星画像

GoogleMap の背景画像を使用した。大きさは 256px で、解像度としては 1px あたり 0.3m のものを使用している。図-3 に本研究で使用した衛星画像の例を示す。

## 4. 建物単位の所得水準レベルの分類

### 4.1. 教師データの作成

まず、深層学習モデルを構築するための教師データを作成した。小地域世帯所得データから得られる世帯所得を 0~2 の所得水準レベルに分類した。本研究ではレベル 0 を 350 万円未満、レベル 1 を 350 万円以上 750 万円未満、レベル 2 を 750 万円以上とした。各レベルに分類された建物数を表-2 に示す。この分類は家計調査の年間収入と世帯人員の関係を参考に設定しており、居住人数に 1 人以上の変化が出る平均収入を閾値として採用した。

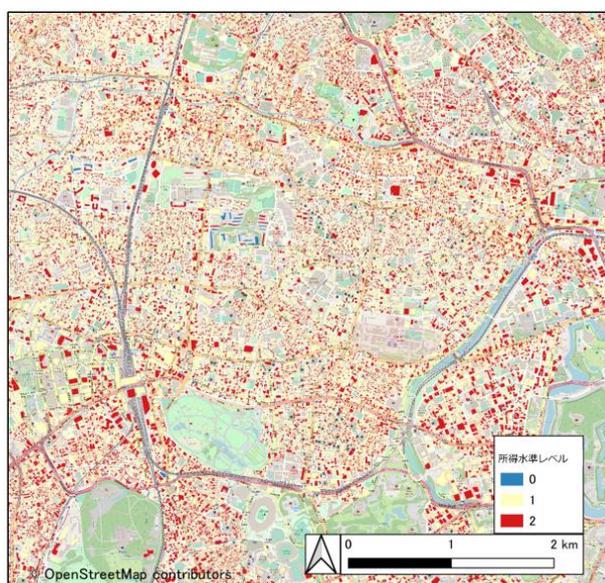
そして、小地域世帯所得データを建物ポリゴンデータに空間結合することで、個々の建物に世帯の平均所得と、所得水準レベルを付与した。なお、複数の小地域に跨って立地する建物の場合、交差している小地域の面積を重みにして平均世帯年収を按分した。以上の処理により、まず正解ラベルを作成した。

次に、正解ラベルと対になる画像データを作成した。各建物の中心座標を 256px 四方の正方形で囲う領域で衛星画像を切り出し、建物と対になるようにすることで、モデル構築用の教師データを作成した。以上の処理により、合計 86,274 件の建物に関する教師データを作成することができた。図-4 に教師データとなる建物ごとの所得水準レベルを示す。

### 4.2. 深層学習モデルの構築

次に、先程作成した学習データを用いて深層学習モデルの構築を行った。本研究では VGG-16 と呼ば

1 れる学習済み深層学習モデルを用いて、所得水準レ  
 2 ベルの分類モデルを実装した。VGG-16 (Simonyan &  
 3 Zisserman, 2014) は物体検出と分類のアルゴリズム  
 4 であり、1,000 種類の画像を 92.7%の精度で分類する  
 5 ことができ、特に画像分類の分野において非常に優  
 6 れたパフォーマンスを発揮している。また、アーキ  
 7 テクチャも理解しやすい構造になっているため、モ  
 8 デルの調整がしやすく転移学習にも頻繁に用いられ  
 9 る。そこで、本研究では VGG-16 を転移学習という  
 10 手法を用いて構築した。モデル構築には、教師デー  
 11 タの 8 割となる 69,019 枚の画像を使用し、残りの 2  
 12 割となる 17,255 枚を精度検証に使用した。以上の方  
 13 法で所得水準レベルを分類するモデルを構築した。  
 14 なお、学習回数は 20 回とした。



16 図-4 教師データの所得水準レベル

### 18 4.3. 分類結果の出力

19 以上の手法で作成した所得水準レベルの分類モデ  
 20 ルを検証用データに適用することで、建物ごとの所  
 21 得水準レベルを出力した。図-5 に正解データとモデ  
 22 ルによる予測結果、また誤分類された建物の分布を  
 23 可視化した結果を示す。

## 25 5. 精度の検証

### 26 5.1. 検証方法

27 本研究では、それぞれの所得水準レベルに対して、  
 28 5.1.1~5.1.4 の4つの指標を用いて精度の検証を行っ  
 29 た。このように多くの指標を用いることで、クラス  
 30 ごとの予測品質を詳細に理解することができ、どの  
 31 所得水準の地域において分類精度に課題があるのか  
 32 を明確にできる。また、学習傾向に偏りが無いかも  
 33 確認することが可能になる。

34 さらに、モデルが学習データに偏っていないかど  
 35 うかを確認するために、損失曲線と正解率曲線を学  
 36 習データと検証用データそれぞれを用いて作成した。

#### 38 5.1.1 正解率 (Accuracy)

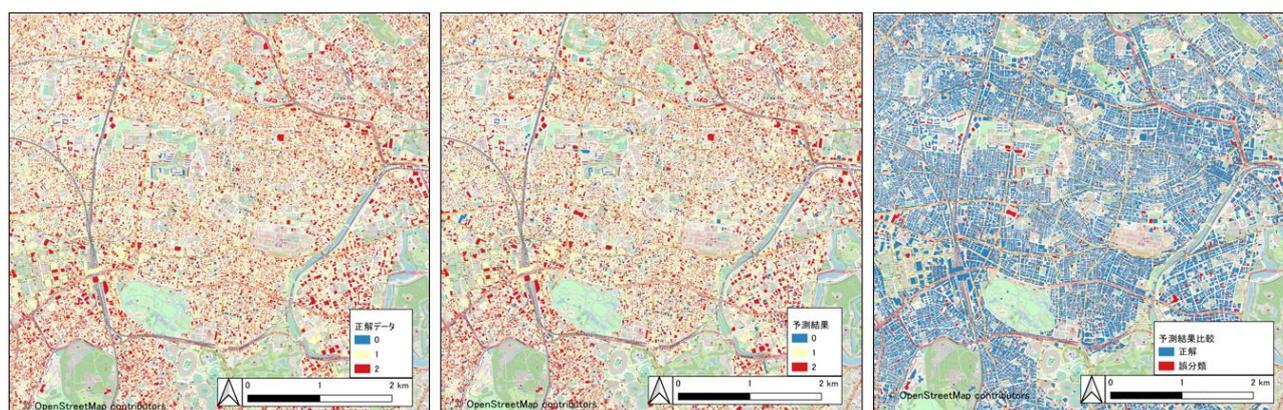
39 すべての予測に対する正しい予測の割合。

$$40 \text{ Accuracy} = \frac{TP + TN}{TP + TN + FP + FN}$$

#### 42 5.1.2 適合率 (Precision)

43 正と予測した中で実際に正だった割合。

$$44 \text{ Precision} = \frac{TP}{TP + FP}$$



17 図-5 正解データ (左)、予測結果 (中央) および誤分類された建物の分布 (右)

## 5.1.3 再現率 (Recall)

実際に正の中で正と予測できた割合.

$$\text{Recall} = \frac{TP}{TP + FN}$$

## 5.1.4 F 値 (F-measure)

適合率と再現率の調和平均.

$$\text{F-measure} = \frac{2 \times \text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}}$$

## 5.2. 検証結果

各所得水準レベルの各指標における結果を表-3に、分類結果の混同行列と、モデル学習時の損失曲線および正解率曲線を図-6に示す.

正解率曲線から、学習データと検証用データの両方で90%以上の高い正解率で分類できていることが分かる. また、学習データと検証用データの正解率や損失の変化の傾向も類似していることから、汎化性能を保ちながら学習できていることがわかる. 各指標の結果を確認すると、所得水準レベル1と2ではすべての指標で90%以上と非常に高い水準で分類ができていた. 一方、所得水準レベル0ではF値が

74%と、他の所得水準レベルと比べて精度が低いことが確認された. この結果から、中所得や高所得の地域では正確な分類が可能だが、低所得の地域に対しての精度に課題があることが分かった.

## 6. 考察

## 6.1. 画像の目視確認による考察

5章で得られた精度によると、衛星画像から経済レベルを分類することは、かなり高い精度で行うことが可能であるということが確認できた. ただし、低所得帯での分類精度に課題があることも確認することができた.

そこでまず、全体としてどのような地域で誤分類が多かったのかを、衛星画像を目視で確認することでその傾向を探った. その結果、複数の機能を有する複合建物のような形状が複雑な建物や、高所得な地域で建物同士の隣棟間隔にゆとりがある地域において誤分類が多くなることが確認された. 図-7に複雑な形状の建物の例を、図-8に所得が高く隣棟間隔も大きい地域の例を示す. これらの地域で精度が下がった原因は、これらの条件に合致するサンプル数が少なかったため、モデルが正しく判断しきれな

表-3 各所得水準レベルにおける各指標の結果

所得水準 レベル	正解率 (Accuracy) (%)	適合率 (Precision) (%)	再現率 (Recall) (%)	F 値 F-measure (%)
0	98.39	92.00	62.16	74.19
1	98.39	99.09	98.48	98.79
2	98.39	97.13	98.45	97.78

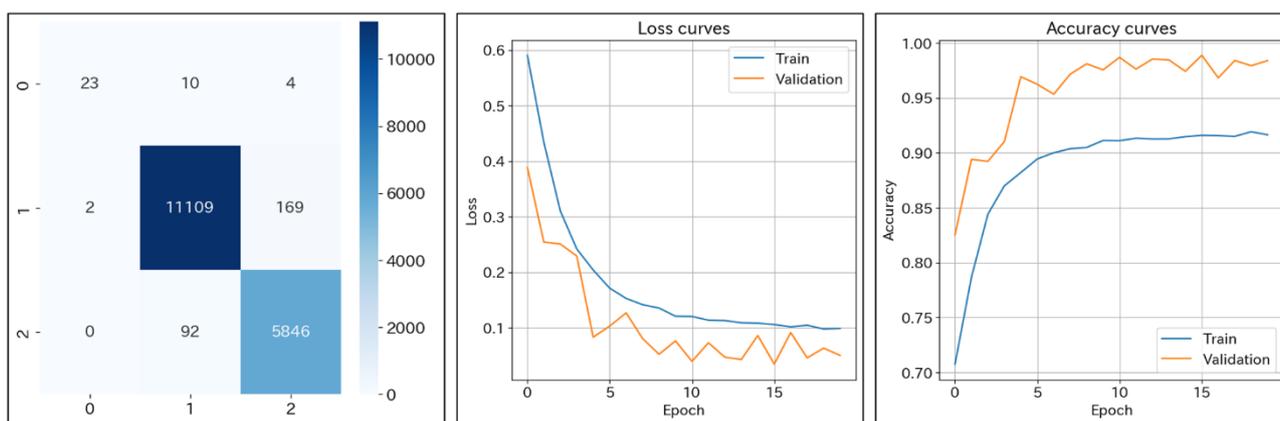


図-6 検証用データにおける混同行列 (左) と損失曲線 (中央) および正解率曲線 (右)

1 かったことが原因であると考えられる。これは低所  
2 得帯における分類精度が低いことにも共通する原因  
3 であると考えられる。本研究の対象地域である新宿  
4 区は平均世帯所得が350万円以上の世帯が95%以上  
5 とデータ自体に偏りが見られた。すなわち、本研  
6 究の分類精度はデータの件数にある程度依存する結果  
7 になっていることが分かる。

8  
9 **6.2. Grad-CAMを用いた特徴量の可視化による考察**  
10 次に、本研究で開発したモデルが建物のどのよう  
11 な特徴を用いて所得水準レベルを判断しているのか  
12 を探るために、Grad-CAM (Selvaraju. et al.) と呼ば  
13 れる手法を用いて学習の可視化を行った。Grad-  
14 CAMとは、畳み込みニューラルネットワーク(CNN)  
15 ベースのモデルがクラスを決定する際の視覚的説明  
16 を生成するための技術である。最終畳み込み層に流  
17 れる任意のターゲット概念の勾配を使用して、概念  
18 を予測するために画像内で重要な領域を強調表示す  
19 るローカリゼーションマップを生成し、判断に使用  
20 している領域とその重要度を可視化することができる。  
21 この技術を用いて、本研究で作成したモデルが  
22 何を重要視してクラス分類を行っているのかを確認  
23 した。

24 図-9にGrad-CAMを用いて分類を可視化した結果  
25 を示す。これらの結果によると、建物自体を根拠に

26 判断しているわけではなく、建物同士の隙間を判断  
27 材料にして分類を行っていることが分かった。実際  
28 に高所得と呼ばれる地域は日本では建物同士が密集  
29 している傾向が強いため、その特徴を十分に捉えら  
30 れていると考えられる。また、大規模な工場などの  
31 面積が非常に大きく形状も複雑な建物の場合、道路  
32 と建物の境界を適切に判断することができず、誤分  
33 類を引き起こしている可能性があることも確認でき  
34 た。

### 35 6.3. 精度改善に向けた取り組みと今後の展開

36 まず、低所得帯や複雑な建物を含んだ地域での分  
37 類精度の改善は、対象地域を拡大することで解決で  
38 きるものと考えられる。新宿区は日本の中でも所得  
39 水準がかなり高く、また様々な高さや形状の建築物  
40 が高密度に集積する、日本国内でも特異な地域であ  
41 ると言える。そこで、今後は対象地域を東京都全域、  
42 さらには関東全域と拡大していくことで、モデルの  
43 分類精度の向上や、都市としての成長の段階を図る  
44 ことができるモデルになっていくものと考えられる。

45 また、面積の大きく複雑な形状の建物に対しては、  
46 一定以上の面積を閾値にして一旦省いた状態でモデ  
47 ルを構築する方法も考えられる。実際に複合施設や  
48 大規模な事業所に人が定住している可能性は小さい  
49 ため、それらを省いてもモデルの推定精度への影響  
50



図-7 複雑な形状を有する建物の例

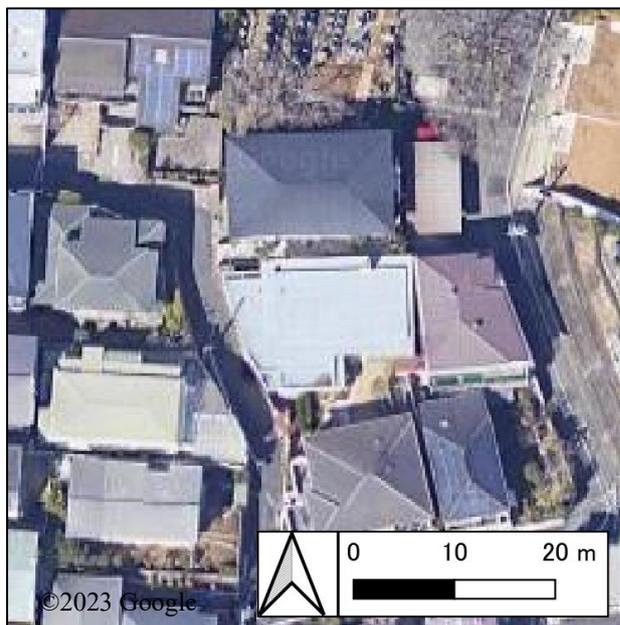


図-8 建物同士の隣棟間距離が大きい地域の例

1 は小さいものと考えられる。  
 2 また、今後の展開としては以上の改善点のほかに、  
 3 海外の開発途上国に展開した際にどのような影響が  
 4 あるのかを確認してみる必要がある。新宿区では高  
 5 所得な地域になるほど建物が密集する傾向が見られ  
 6 たが、海外で実装した際にも正しく分類ができるの  
 7 か検証する必要がある。

8 また、本研究全体の最終目標である衛星画像から  
 9 人口分布を推定する技術の実現に向けて、衛星画像  
 10 からの建物の自動抽出や、建物用途推定にも取り組  
 11 んでおり、これらの技術と組み合わせた際に、全体  
 12 としてどれほどの精度を発揮することができるのか  
 13 も、今度の取り組みに向けて検証していきたい。

## 14 7. まとめ

15 本研究では、所得に関する統計が整備された日本  
 16 を対象に、衛星画像を用いて建物単位の所得水準を  
 17 推定する手法の提案とその検証を行った。その結果、  
 18 高い精度で建物単位の所得水準の分類を行うことが  
 19 できることが確認された。一方、データの件数が少  
 20

21 なかった低所得帯での分類精度や、複合建物などの  
 22 件数が全体を通して少ない特徴の建物に対する精度  
 23 に課題が残った。今後はこれらの特徴を有する建物  
 24 に対しても精度が発揮できるよう教師データの充実  
 25 とモデルの改良を図りつつ、国内の他の地域や海外  
 26 などへの対象地域の拡大や、建物単位の人口推定に  
 27 向けた取り組みを進めていく。

## 28 謝辞

29 本研究は JSPS 科研費 (JP20H01483) の助成を受  
 30 けて実施しています。また、本研究は、東京大学 CSIS  
 31 共同研究 (No. 1170) による成果の一部です (利用デ  
 32 ータ: Zmap TOWN II (2020 年度 Shape 版) 東京都  
 33 データセット (株式会社ゼンリン提供)。ここに記し  
 34 て謝意を表します。

## 35 参考文献

36 奥田康平・川崎昭如・濱口竜平 (2019) 衛星画像と  
 37 世帯調査データを用いた建物ごとの収入レベルの  
 38 推定, 「GIS—理論と応用」, Vol. 27, No.2, pp.9-18.  
 39  
 40

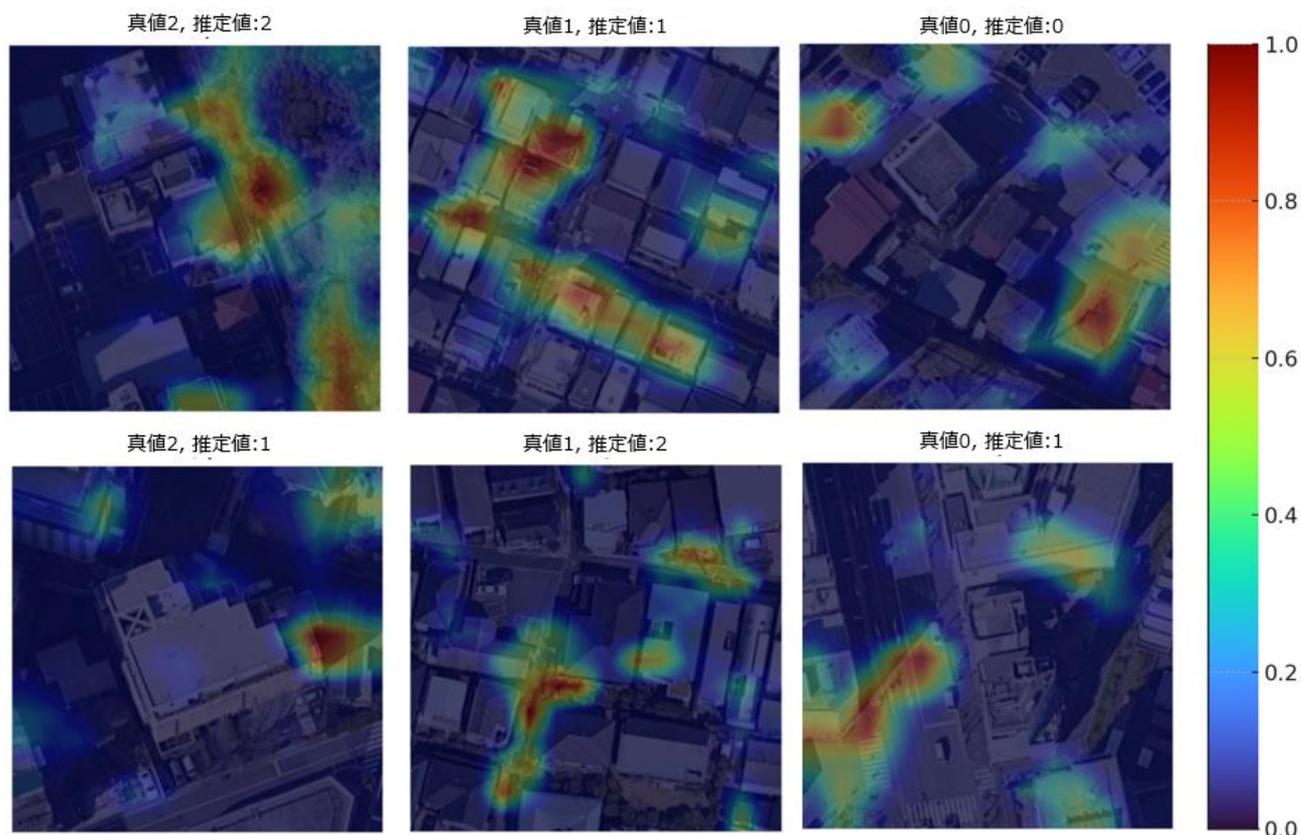


図-10 Grad-CAM を用いた分類時に使用している領域とその重要度を可視化した結果

## 【審査用原稿用紙】

8/10

- 1 Akiyama, Y., Miyazaki, H. and Sirikanjanaanan, S.,  
2 (2019) Development of Micro Population Data for Each  
3 Building: Case Study in Tokyo and Bangkok, 2019  
4 *First International Conference on Smart Technology  
5 and Urban Development (STUD)*, pp.1-6.
- 6 Kaiser, P., Wegner, J., Lucchi, A., Jaggi, M., Hofmann, T.  
7 and Schindler, K. (2017) Learning Aerial Image  
8 Segmentation from Online Maps, *IEEE Transactions  
9 on Geoscience and Remote Sensing*, Vol.55, No.11, 201.
- 10 Simonyan, K. and Zisserman, A. (2014) Very Deep  
11 Convolutional Networks for Large-Scale Image  
12 Recognition, *arXiv preprint arXiv:1409.1556*.
- 13 Selvaraju, R. R., Cogswell, M., Das, A., Vedantam, R.,  
14 Parikh, D., & Batra, D. (2017), Grad-CAM: Visual  
15 Explanations from Deep Networks via Gradient-based  
16 Localization, *Proceedings of the IEEE international  
17 conference on computer vision*, pp. 618-626.