

セマンティックセグメンテーションによる 不動産物件外観画像からの建物抽出：建物構造・築年代推定への応用

小川芳樹*・沖拓弥**・陳聖隆***・関本義秀*

Building Extraction from Images of Real Estate by Semantic Segmentation: Application to Built-age and Structure Estimation

Yoshiki Ogawa*, Takuya Oki**, Shenglong Chen***, Yoshihide Sekimoto*

Abstract: The built-age and structure of buildings are essential for the accurate estimation of earthquake damage. On the other hand, recent years have seen the accumulation of images of building by real estate companies, and deep learning have proposed a method of estimating classification with high accuracy using images. Although such a deep learning model for classifying specific classes of objects is capable of learning and classifying images as they are, it is expected to improve the classification accuracy by extracting only the target objects. However, the characteristics of estimation of accuracy of built-age and structure in the previous studies are not well understood by using building extraction. In this study, we applied semantic segmentation to our proposed building-age and structure estimation model and tried to develop a deep learning model for building-age and structure estimation by extracting buildings from images. In addition, the accuracy of semantic segmentation for built-age and structure class was verified and discussed.

Keywords: 不動産データセット (real estate dataset), 深層学習 (deep learning), 建物抽出 (building detection), 築年代 (built-age), 構造 (structure)

1. はじめに

建物構造・築年代の違いは、地震時の建物被害の程度に大きく影響する。そのため、建物構造・築年代は地震の被害想定・建物倒壊危険度を算出する際にも考慮される。したがって、正確な建物構造・築年代に関するデータ整備は、学術的観点からだけでなく、災害軽減の観点からも重要な課題である。

近年、ニューラルネットワークに基づく機械学習手法である深層学習が注目を集めており、さまざまな画像認識タスクに適用されている。筆者らは、深層学習を不動産外観画像に適用して建物構造・築年代の推定モデルを構築してきた(小川ほか 2020)。この研究では、建物を大まかに捉えた画像に深層学習を使用している。しかしながら、画像を建物クラスと建物以外の背景クラスに判別し、建物クラスのピクセルのみを用いて建物構造・築年代を推定して精度検証した研究事例は見られない。

以上の背景から本研究では、まず不動産データセ

ットに含まれる建物外観画像データを使用して、深層学習アプローチの1つであるセマンティックセグメンテーションにより、画像内の建物クラスと背景クラスを判別し建物を抽出する。その上で、建物外観画像ごとの築年代・構造の推定を深層畳み込みニューラルネットワーク(以下、CNN)を用いて試みる。また、建物構造・築年代の推定においては、建物外観画像から以下のクラス区分を推定する。建物構造に関しては、木造・S造・RC造の計3クラス分類と、築年代に関しては1962年以前・1963～1971年・1972～1980年・1981～1989年・1990～2001年・2002～2011年・2012年以降の計7クラス分類を推定するモデルを構築する。最後に建物抽出の有無による建物構造・築年代モデルの精度をそれぞれ比較・検証することで、セマンティックセグメンテーションの有用性を検討する。

2. データ

本研究で用いる不動産データは株式会社 LIFULL

* 正会員 東京大学生産技術研究所 (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)

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

表 1 建物画像データの構造・築年代別の枚数

	クラス	枚数
築年代	-1962 年	16,542
	1963-1971 年	95,329
	1972-1980 年	385,628
	1981-1989 年	1,457,054
	1990-2001 年	2,975,915
	2002-2011 年	2,951,056
構造	2012 年-	1,434,706
	RC 造	4,114,702
	S 造	2,943,367
	木造	2,258,224



図 1. 本研究で用いる建物外観画像の例

(2015 年) の日本全国を網羅しているデータを用いる。このデータには、不動産物件ごとに画像データと建物構造・築年代に関する正解ラベルがアノテーションされている。収録されているクラス別の画像枚数は、表 1 の通りである。

データセットに含まれる多くの外観画像は、図 1 に示すように建物の全体像を写している。一方、外観以外の画像や、建物の一部を拡大した画像、あるいは新築物件の完成イメージ図なども混在している。また、複数の建物が写っており、画像中のどれが対

象建物であるかが判断しにくいものや、植栽や自動車などによって大部分が覆われている画像も一部に見られる。こうした画像が学習データセットに含まれると、適切な学習が行われずに推定精度を低下させる要因となるため、本研究では、建物以外が大部分を覆っている画像を取り除く。またクラス間の画像枚数の違いによるバイアスを取り除くため、最小枚数のクラスに合わせてアンダーサンプリングする。その結果得られた 17,415 枚の建物外観画像を用いて実験を行う。

3. 手法

3.1. セマンティックセグメンテーションによる建物抽出

植栽や自動車などの建物以外の要素が予測精度に与える影響を軽減するために、セマンティックセグメンテーションを用いて画像内の建物クラスと建物以外のクラスの 2 つに判別し、建物を抽出する。セマンティックセグメンテーションは、画像の中の同じオブジェクトクラスに属する部分をピクセル単位でクラスタリングすることが可能である。本研究では、ADE20K データセットで事前に学習したセグメンテーションモデル (PSPNet) を適用する (Zhao et al. 2016)。PSPNet のアーキテクチャを図 2 に示す。この PSPNet を用いることで、最終的なピクセル単位のクラス予測値を得ることができる。

3.2. CNN

建物外観画像を用いた建物構造・築年代モデルには CNN を用いる (Krizhevsky et al. 2012)。CNN は畳み込み層とプーリング層と呼ばれる 2 種類の層を交

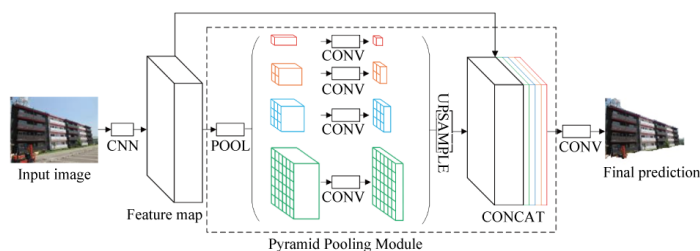


図 2. PSPNet のアーキテクチャ

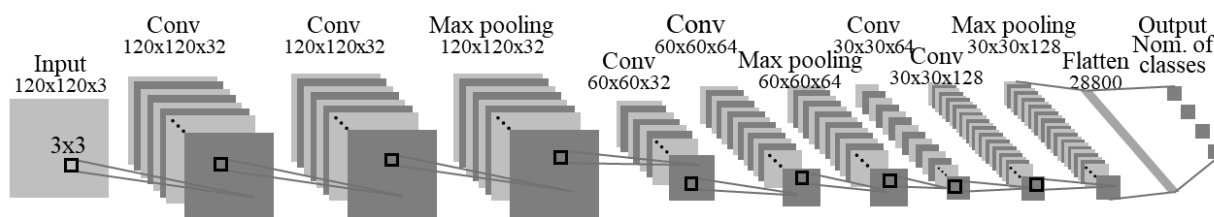


図 3. 本研究で用いる CNN のアーキテクチャ

互に積み重ねた構造の多層ニューラルネットワークである（図3）．CNNはこのような構造をもつことにより，得られる特徴に不変性があることで知られている．また，CNNの最後の出力にソフトマックス関数を用いることにより，入力画像がそれぞれのクラス O_i である確率 $P(O_i)$ を求めることができる．クラスの判別は，この確率に基づいて行われる．

4. 結果

4.1. セマンティックセグメンテーション

建物外観画像にセマンティックセグメンテーションを適用し，建物クラス以外の背景クラスにマスキングした画像の例を図4に示す．また，正しく建物クラスを抽出できていることを確かめるために，建物構造別に目視による検証を行った結果を表2に示す．いずれの構造においても建物クラスを95%以上の高精度で抽出できている，PSPNetによるセマンティックセグメンテーションが有効であるといえる．その一方で，一部の画像では建物を背景クラスに過小抽出した例や，電柱を建物クラスに過大抽出した画像が見られた．これらは，対象の建物外観画像が低いことに起因している可能性が高く，解像度を上げてクラス境界をより正確に判別できれば，改善できる可能性がある．

4.2. 構造・築年代の精度検証

ここでは，まず，建物構造・築年代モデルにおける正解クラスに対する確率が高く，正しく判別できた建物外観画像を紹介する．図5は，誤差が最小であるエポック時の学習モデルを用いて各構造と築年代を正しく判別できた画像と各クラスの確率を示している．

続いて，構造モデルの結果を見ると，その精度は建物抽出した場合に総合精度81.5%（3クラス），抽出しない場合に85.1%（3クラス）の精度で良好に判別できた（表3）．木造と非木造の2クラスに集計した場合には，建物抽出した場合に総合精度71.1%，抽出しない場合に85.1%であり，建物抽出をしない方が良好な結果が得られた．建物抽出した場合は，しない場合と比較して，S造においてF値が極端に低い結果となった．これは，建物抽出することで建

物周辺領域の情報が欠落すること起因する可能性があるが，今後の精査が必要である．また，建物抽出した場合において80.0%以上の確率で正しく判別出来た画像は，構造の場合に全体の51.5%あり，建物外観の特徴を的確に捉え，高い判別確率をもって推定していることが確認された．

次に築年代モデルの場合，総合精度は建物抽出をした場合に木造:33.4%，S造:34.5%，RC造:25.9%，抽出しない場合に木造:33.8%，S造:31.7%，RC造:30.4%（いずれも7クラス）であり，建物抽出の有無



図4. 建物クラスの抽出結果の例

表2 建物抽出精度

	成功画像	過大抽出画像	過小抽出画像	精度 (%)
木造	5,761	146	7	97.4
S造	5,822	82	11	98.4
RC造	5,832	78	5	99.0
合計	17,415	306	23	98.1

木造: 96.6% S造: 3.2% RC造: 0.2%
木造: 2.0% S造: 93.2% RC造: 48.3%
木造: 4.2% S造: 10.2% RC造: 85.6%



木造		S造		RC造			
-1962	55.1%	-1962	9.9%	-1962	1.1%	-1962	7.0%
1963-1971	36.7%	1963-1971	81.6%	1963-1971	9.1%	1963-1971	17.5%
1972-1980	3.1%	1972-1980	5.9%	1972-1980	84.6%	1972-1980	15.6%
1981-1989	0.9%	1981-1989	0.9%	1981-1989	4.1%	1981-1989	54.1%
1990-2001	0.9%	1990-2001	1.7%	1990-2001	1.0%	1990-2001	3.1%
2002-2011	0.8%	2002-2011	0.0%	2002-2011	0.1%	2002-2011	0.8%
2012-	2.4%	2012-	0.0%	2012-	0.0%	2012-	1.9%



-1962



1963-1971



1972-1980



1981-1989

-1962	0.4%	-1962	0.0%	-1962	0.1%
1963-1971	3.3%	1963-1971	0.0%	1963-1971	0.8%
1972-1980	8.1%	1972-1980	0.0%	1972-1980	0.9%
1981-1989	1.2%	1981-1989	0.1%	1981-1989	3.1%
1990-2001	85.9%	1990-2001	4.2%	1990-2001	1.8%
2002-2011	1.0%	2002-2011	194.2%	2002-2011	4.7%
2012-	0.0%	2012-	0.0%	2012-	88.6%



1990-2001



2002-2011



2012-

図5. 正しく判定できた各クラスの建物画像と各クラス確率の例（上段：建物構造，下段：築年代）

による精度の違いは小さかった（表 4）。また、耐震構造が大きく変わる 1981 年前後の 2 クラスに集計した場合の精度は、建物抽出した場合に総合精度 61.5-71.5%，抽出しない場合に 64.4-73.5%であり、どちらも精度は良好で差異のない結果であった。以上のように、今回の実験では建物抽出による効果は限定的で一部クラスの精度は高くなった一方で、精度が下がったクラスも見られる結果となった。

4.3. 今後の展望

画像に複数の建物が写っている画像については、さらに詳細に建物判別をする必要があり、インスタンスセグメンテーションの適用が課題である。また、建物構造・築年代の推定精度を高めるためには、今回の実験で用いた画像よりも高精度な画像を多数用いて学習を行うことで、建物表面の状態なども学習できる可能性がある。画像から建物構造・築年代を精度良く推定できれば、建物を包括的に調査できる車両カメラ画像に適用して、網羅的に建物構造・築年代をマッピングするのに有効と考えられる。

謝辞

本研究を行うにあたり、JSPS 科研費 JP19K0494 および JP20K15001 の助成を受けた。また、国立情報学研究所の IDR データセット提供サービスにより株式会社 LIFULL から提供を受けた「LIFULL HOME'S データ

セット」を利用した。関係各位に謝意を表する。

参考文献

- 小川芳樹・沖拓弥・関本義秀・柴崎亮介（2020）不動産に関する地理空間情報と建物画像を用いた建物の構造・築年代推定手法の検討，人工知能学会全国大会論文集 第 34 回全国大会。
- A. Krizhevsky, I. Sutskever, and G.E. Hinton (2012) ImageNet Classification with Deep Convolutional Neural Networks, in Advances in neural information processing systems, 1097–1105.
- H. Zhao, J. Shi, X. Qi, X. Wang, and J. Jia (2016) Pyramid scene parsing network, Proceedings of the IEEE conference on computer vision.

表 3 建物抽出の有無による構造モデルの精度（%）

	建物抽出あり			建物抽出無し		
	適合率	再現率	F 値	適合率	再現率	F 値
木造	71.4	90.4	79.8	77.1	79.8	78.5
S 造	66.4	17.2	27.3	62.3	56.9	59.5
RC 造	72.5	91.1	80.7	79.5	83.2	81.3
総合精度 (3 クラス)	71.5			73.6		
木造	71.4	90.4	79.8	77.1	79.8	78.5
非木造	92.1	75.6	83.0	89.5	87.8	89.6
総合精度 (2 クラス)	81.5			85.1		

表 4 建物抽出の有無による築年代モデルの構造別の精度（%）

	木造						S 造						RC 造					
	建物抽出あり			建物抽出無し			建物抽出あり			建物抽出無し			建物抽出あり			建物抽出無し		
	適合率	再現率	F 値	適合率	再現率	F 値	適合率	再現率	F 値	適合率	再現率	F 値	適合率	再現率	F 値	適合率	再現率	F 値
-1962	38.1	8.2	13.6	31.3	58.6	40.8	28.6	25.0	26.7	27.8	49.1	35.5	36.4	5.8	10.0	30.9	71.7	43.2
1963-1971	24.7	60.6	35.1	12.0	1.9	3.2	33.1	63.5	43.5	24.3	32.0	27.6	21.9	75.2	33.9	25.0	3.7	6.5
1972-1980	21.3	34.9	26.4	75.0	1.8	3.6	26.7	24.2	25.4	27.3	16.8	20.8	14.3	1.0	1.9	22.7	48.8	31.0
1981-1989	31.7	33.7	32.7	36.7	13.2	19.4	37.0	17.9	24.1	26.7	43.3	33.1	0.0	0.0	0.0	21.1	2.5	4.5
1990-2001	35.0	17.4	23.2	25.3	52.9	34.3	25.0	4.4	7.5	30.7	12.8	18.0	29.1	38.2	33.1	34.4	6.7	11.3
2002-2011	47.5	42.1	44.6	37.0	64.9	47.1	62.5	9.8	16.9	44.7	26.9	33.6	30.2	19.3	23.5	36.8	24.1	29.2
2012-	84.0	36.2	50.6	68.5	40.9	51.3	40.8	75.3	53.0	55.0	45.8	50.0	34.1	37.3	35.6	39.9	52.8	45.5
総合精度 (7 クラス)	33.4			33.8			34.5			31.7			25.9			30.4		
-1980	59.6	86.4	70.5	77.4	54.0	63.6	64.0	82.1	71.9	62.4	75.7	75.7	52.2	73.3	61.0	55.6	86.7	67.8
1981-	87.7	62.2	72.8	71.9	88.2	79.2	78.8	59.1	67.6	78.0	65.4	65.4	74.1	53.2	62.0	82.5	47.6	60.3
総合精度 (2 クラス)	71.1			73.5			69.9			69.8			61.5			64.4		