

深層特徴量を用いた類似間取り図検索

Similar Floor Plan Retrieval using Deep Convolutional Neural Networks

高田 祐樹 井上 直人 山崎 俊彦 相澤 清晴
Yuki Takada Naoto Inoue Toshihiko Yamasaki Kiyoharu Aizawa

東京大学
The University of Tokyo

我々は、間取りをクエリとした新しい不動産物件検索の枠組みを提案している。これまでのシステム中では、間取りをグラフとして表現し、グラフ類似度を用いて検索を行っていたが、間取り画像からグラフ構造を抽出する部分はクラウドワーカーによる人手で行っていたため、手間やコストが問題になっていた。本論文では間取りタイプの認識と特定の機能を持った部屋の有無の認識という2つのタスクを同時学習することで検索に適した深層特徴量を抽出することに成功した。22140枚の間取り図を用いた実験では、precision@5で15.7%の精度を実現した。

1. はじめに

近年、ReTech (Real Estate Tech) と呼ばれる不動産分野にITを導入する動きが盛んである。例えばフランス・カンヌで毎年3月に開催されている世界最大の不動産見本市 MIPIM では2年連続で ReTech がテーマとして取り上げられた [1]。情報化の例として、ユーザーが自分の好みの見つけるための検索システムなどがあげられるが、不動産においては現状駅からの距離や部屋の広さなど基本的なメタデータを用いた検索しかできないのが現状である。そのためメタデータに依存しない別の方法での検索が求められている。

そこで本研究では、類似している間取り図を画像のみから検索することを目指す。従来の検索システムはメタデータを用いていることは先に述べた。メタデータ以外を用いる検索として大原ら [4] は、クエリとして間取り図をグラフ構造に落とし込んだものを提案した。しかし、この手法は例えば、1件あたり20円と安価な値段でラベル付が可能であることを示したものの、物件は入れ替わりが激しく、新規物件が大量に発生すること、オンタイムにクラウドワーカーにラベル付してもらう仕組みづくりが困難であることなどの問題があった。

本研究の提案手法では、間取り図のタイプや、各部屋タイプのあるなしなどのメタデータがついた間取り図のデータセットを用いて間取り図を深層学習を用いて学習する。学習したモデルを用い深層特徴量を用いることで、メタデータに頼らない検索が可能になる。

2. 関連研究

不動産物件の間取り図に関する研究は既に存在する。花里ら [2] は、間取り図486枚を元にその間取り図のグラフ構造を作成し、廊下を中心としてLDや私室など全ての部屋がそこに繋がっている「個室群化傾向」が進んでいると指摘している。Liu [3] らは国立情報学研究所で公開されている株式会社ネクストが提供した HOMES データセットを用いて間取り図とその物件内部の写真とのペアリングを Siamese Network を用いて行い、人間より高い精度で推定に成功している。

Table 1: 間取りのタイプとその物件数

間取りの種類	物件数	間取りの種類	物件数
2LDK	5194	1LDK	758
2DK	5083	2SLDK	507
3LDK	2985	1DK	388
1K	2419	4LDK	214
2K	2065	3K	140
3DK	1112	3SLDK	121
ワンルーム	974	others	180

3. 提案手法

本研究の目標は、間取り図画像を入力として、入力画像と構造的に類似した画像を提示することである。本研究では間取り図のメタデータを用いて深層学習を行い、そのモデルを使い深層特徴量を取り出すことで間取り図の特徴ベクトルを作成した。その特徴ベクトルを用いて最近傍法を用いて構造的に類似した画像の提示を行った。本研究では、大原ら [4] が作成した間取り図のデータセットを用いた。今回使用した間取り図は SUMMO のサイトからスクレイピングした千代田区・文京区・世田谷区・府中市の物件と HOMES データセットの世田谷区の物件である。このデータセットには、間取り図画像とそのメタデータ、グラフ構造が存在する。データセットから、グラフ構造がない間取り図、住所・家賃・床面積が同じ重複していると考えられる間取りを取り除くと22140件となった。このデータセットを間取り図の種類ごとに分類すると以下の Table 1 の14種類になる。但し100件に満たないタイプは others にまとめた。また、グラフ構造には Table 2 に示す24種類のノードが存在する。本研究では、そのうち太字・下線にした14種類を用いた。

本手法では次のネットワーク Fig. 1 を用いて学習を行った。データセット内の画像22140枚のうち12140枚を用いてネットワークの training を行った。このネットワークは Imagenet で学習済みの VGG-16 の fc7 層の後ろに、間取りタイプを学習するための14次元の全結合層と各部屋の有無を学習するための28次元の全結合層を繋げたものである。このネットワークを用いて間取りタイプ、部屋の有無の2種類の学習を同時

連絡先: 東京大学大学院情報理工学系研究科 〒113-8656
東京都文京区本郷 7-3-1 TEL: 03-5841-6761 mail:
takada@hal.t.u-tokyo.ac.jp

Table 2: 間取りグラフにあるノードの種類

ノード名	説明	ノード名	説明	ノード名	説明
Loft	ロフト	BR	寝室	DR	衣装室
WR	洋室	UB	ユニットバス	L	リビング
Bal	バルコニー	Ba	浴室	D	ダイニング
UPDN	階段	WC	トイレ	K	キッチン
JR	和室	Hall	廊下	DK	ダイニングキッチン
WIC	ウォークインクローゼット	PR	化粧室	LD	リビングダイニング
Ver	ベランダ	CL	クローゼット	LDK	リビングダイニングキッチン
R	部屋	E	玄関	Other	その他

Table 3: 混合行列

		判定														
		2LDK	2DK	3LDK	1K	2K	3DK	1R	1LDK	2SLDK	1DK	4LDK	3K	3SLDK	others	accuracy
正解	3LDK	435	32	16	1	4	1	0	6	1	0	0	0	0	1	0.88
	2DK	23	379	3	2	13	3	0	5	0	3	0	0	0	1	0.88
	3LDK	15	4	271	0	0	6	0	1	5	0	0	0	3	0	0.85
	1K	2	0	0	192	5	0	16	5	0	0	0	0	0	0	0.87
	2K	2	28	0	6	126	0	2	3	0	1	0	0	0	0	0.75
	3DK	5	7	8	1	1	82	0	0	1	0	0	0	0	0	0.78
	1R	0	0	0	24	1	0	42	1	0	0	0	0	0	0	0.62
	1LDK	11	4	0	13	2	0	1	37	0	3	0	0	0	0	0.52
	2SLDK	3	0	5	0	0	0	0	0	29	0	0	0	0	0	0.78
	1DK	2	1	0	11	4	0	1	7	0	11	0	0	0	0	0.30
	4LDK	0	0	3	0	0	0	0	0	0	0	9	0	0	3	0.60
	3K	0	3	0	0	2	2	0	0	0	0	0	5	0	1	0.38
	3SLDK	1	0	8	0	0	0	0	0	1	0	0	0	5	0	0.33
	others	1	3	3	0	1	0	0	2	0	0	1	0	0	3	0.21

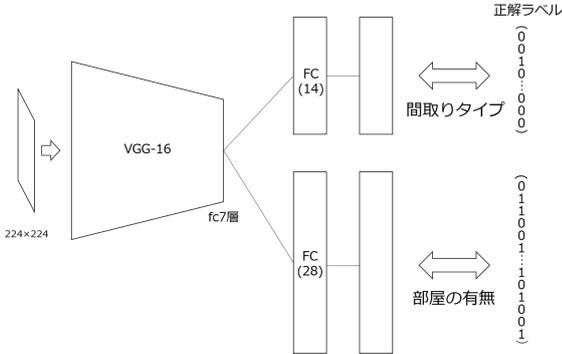


Fig. 1: 使用したネットワーク

4. 実験

本実験は以下の条件で行った。まず、テスト用の間取り画像 10000 枚についてデータの前処理を行った。本研究で使用したデータセットには、住所や家賃などのメタデータが異なっているにも関わらず、間取り図が同じものが存在している。そこで本実験では、間取り図画像が完璧に同じものを除外した。そうした結果テスト画像は 7960 枚となった。これをクエリ画像 100 枚、検索対象 7860 枚に分割した。クエリとした画像に対し、先に述べた手法で検索対象の画像の中から類似画像を取得した。本実験の評価手法には precision@5 を用いた。これは検索結果上位 5 件のうち正解データが含まれる割合を示す。本実験では構造類似度 p 以上の間取り画像を検索出来た場合正解と定義し、 $p = 0.5, 0.6, 0.7, 0.8, 0.9$ でそれぞれ評価を行った。次の 6 つの手法の比較を行った。

- ランダム:ランダムに間取り画像を取得した。
- メタデータ:間取りのタイプ情報を用いて間取りのタイプが同じものからランダムに間取り画像を取得した。
- 学習済みモデル:Imagenet で学習済みの VGG-16 を用いた。
- 間取りのタイプ:間取りタイプのみを学習したネットワークを用いた。
- 部屋の有無:部屋の有無のみを学習したネットワークを用いた。
- 提案手法:間取りタイプ、部屋の有無を同時に学習したネットワークを用いた。

また、Fig. 1 のネットワークの一部を用い、間取りタイプのみの学習、各部屋の有無のみの学習を行いそれぞれ学習の精度を計算した。この実験は training に使った画像は 20140 枚、test に使った画像は 2000 枚で行った。

間取りタイプ識別の結果は Table 3 のようになった。これは test に使った間取り図画像 2000 枚に対し、正解のラベルと学習したモデルが出力したラベルの分布を表にしたものである。accuracy は、そのタイプの間取りをどれくらい正確に正しいタイプと認識できたかを示す。この表から以下のことが分かる。まず出現回数が多いクラスの方が、正答率が高い傾向にあることが分かる。画像数が少ないため出現回数が多いクラスでは十分に学習できていない可能性や、出現回数が多いクラスを重要視して学習が行われている可能性が考えられる。4LDK が画像数が少ないクラスの中では比較的 accuracy が高いのは 4LDK だけ部屋の数が 5 つと、他の間取りタイプとの差が大きいためと考えられる。一方で間違えやすい 2 タイプの間取りを見てみると、2DK と 2K、1K とワンルームなどがあげられ、これは間取りタイプの差分が他の間取りタイプ間の差と比べて小さいためと考えられる。

部屋タイプごとの各部屋の有無の認識の精度をまとめたものが以下の Table 4 である。例えば”FP”の行は実際に部屋は存在しないが、存在すると間違えてしまった例を示している。accuracy は正しく判定出来た割合、 f 値は以下の式 (2)(3) で

に行った。同時に学習を行うことで、間取りの特徴をよりとらえた深層特徴量を得ることができる。

類似画像の取得は以下のように行った。Fig. 1 で学習したネットワークに、間取り画像を入力して得られる fc7 層の値を深層特徴量として用いて 4096 次元の特徴ベクトルを作成した。クエリとした間取り画像の特徴ベクトルと近い特徴ベクトルを最近傍法を用いて探索し類似画像の取得を行った。ベクトル同士の距離計算にはユークリッド距離を用いた。得られた類似間取り図に対し、各間取り図画像と対応するグラフ構造同士の類似度を計算し、それを間取り画像間の構造類似度とした。グラフ構造間の類似度の計算は、大原ら [4] の手法を参考にした。まず、2 つのグラフの最大共通部分グラフ (MCS) を計算する。得られる MCS の 1 例を Fig. 2 に示す。得られた最大共通部分グラフを用いて 2 グラフ間の類似度を以下の式 (1) で計算した。 $|q|, |g|$ はそのグラフのエッジの数とノードの数の和を示す。 $MCS(q, g)$ は、2 グラフの MCS を表す。2 つのグラフが完全に一致している時 1、全く共通部分がない時 0 とする。

$$sim(q, g) = \frac{|MCS(q, g)|}{|q| + |g| - |MCS(q, g)|} \quad (1)$$

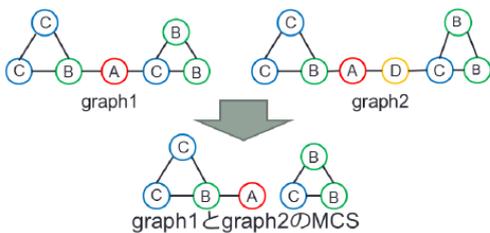


Fig. 2: 最大共通部分グラフの例 [4]

Table 4: 部屋タイプごとの正答率

部屋の種類	TP	FP	TN	FN	accuracy	f 値
ロフト	20	4	1964	12	0.99	0.71
洋室	1538	56	328	78	0.93	0.96
バルコニー	1500	116	264	120	0.88	0.93
階段	173	16	1798	13	0.99	0.92
和室	716	29	1217	38	0.97	0.96
ウォークインクローゼット	101	77	1764	58	0.93	0.6
ベランダ	131	67	1763	39	0.95	0.71
Room	89	49	1793	69	0.94	0.6
寝室	178	30	1754	38	0.97	0.84
ユニットバス	142	65	1561	232	0.85	0.49
浴室	1409	122	259	210	0.83	0.89
トイレ	1768	123	84	25	0.93	0.96
廊下	1056	131	697	116	0.88	0.9
洗面室	1194	79	527	200	0.86	0.9

Table 5: precision@5 の比較

p	ランダム	メタデータ	学習済み	間取りタイプ	部屋の有無	提案手法
0.5	0.084	0.240	0.294	0.400	0.470	0.494
0.6	0.038	0.140	0.224	0.294	0.368	0.396
0.7	0.022	0.065	0.151	0.202	0.259	0.275
0.8	0.013	0.032	0.112	0.151	0.174	0.200
0.9	0.008	0.019	0.107	0.112	0.120	0.157

表せる 2 クラス分類の評価に用いられる値である。

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}} \quad \text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}} \quad (2)$$

$$f \text{ 値} = \frac{2\text{Recall} \cdot \text{Precision}}{\text{Recall} + \text{Precision}} \quad (3)$$

accuracy は全体として高い値が出ているものの、ユニットバス、ウォークインクローゼット、Room の 3 つの f 値が低くなっている。この原因について考察する。ユニットバスは浴室との分類、ウォークインクローゼットは通常のクローゼットとの分類が上手く行かなかったためと考えられる。また、Room は他の部屋のタイプに比べて情報量が少ないためと考えられる。

検索結果を Table 5 に示す。これを見ると提案手法が全て勝っている。このことから提案手法が有効であるといえる。また、間取りタイプのみ学習と部屋の有無のみ学習では部屋の有無のみ学習の方が精度が高い。これから間取り図の構造当てには、部屋の有無の方が重要であると考えられる。また、最も提案手法の精度向上率が大きいのは $p = 0.9$ の時である。これには以下の理由で考えられる。部屋の有無を学習するネットワークでは、部屋同士の繋がりがや間取りの構造は全く考慮できていないが、グラフ類似度の定義から考えると、部屋ノードを当てるだけでもある程度構造類似度が高い間取り画像を検索することができる。つまり、 p が小さくなると間取りタイプを考慮しなくても正解の画像を持ってこれる可能性が上がる。間取りのタイプも一緒に学習することで構造類似度が高い物件の検索が可能になるが、学習する項目が増えた分正しく学習できる確率は部屋の有無のみを学習したネットワークより下がるので、相対的に提案手法が弱くなってしまったと考えられる。

次に個別の検索結果を見てみる。構造類似度が 0.8 以上の間取り画像を類似間取りの検索に成功したとみなし、赤く囲んでいる。Fig. 3a は提案手法が上手く行った例である。部屋の有無で学習したものでは、クエリ画像に和室、洋室、バルコニーがあることが学習できているが 2DK であることは学習できていないことがわかる。一方提案手法では上手く 2DK であることも学習できていることが分かる。Fig. 3b も提案手法が上手く行った例である。部屋の有無のみ学習だと、クエリ画像に洋室が 2 個あることや、間取り画像が 3LDK であること

まで学習できていないことが分かる。一方提案手法では、間取り画像が 3LDK であることまで学習できていることが分かる。Fig. 4 は提案手法が失敗した例である。提案手法ではクエリ画像が 2LDK であることの学習に失敗し 2DK の画像を提示してしまっている。これは間取りタイプで見た時 2LDK と 2DK は類似しているため間違いやすいといえる。LDK のノードは間取り図をグラフ構造に落とし込んだ時、中央部分に来るのでそのノードを間違えた結果、構造類似度としては低くなってしまった。以上のことから、間取りタイプ、部屋の有無の両方を学習することで初めて構造的に類似した間取り画像を検索できると考えられる。

5. まとめ

本研究ではまず、間取り図画像に対して間取りタイプを推定する学習、各部屋タイプの有無を推定する学習の 2通りの学習を行いその精度の評価を行った。そして、その 2種類の学習を同時に行い学習したモデルから深層特徴量を得た。その深層特徴量を用いて類似した間取り図画像の検索を行う手法を提案し、メタデータのみを用いた手法など 5 つの手法と比較した。それにより以下の成果が得られた。

- 間取り図の分類において、間取りのタイプの学習や、部屋の有無の学習が有用であることを示した。
- 間取りのタイプの学習や、部屋の有無の学習を同時に学習する提案手法が最も優秀であることを示した。precision@5 で 15.7 % の精度で構造類似度 0.9 以上の間取り画像を検索することができた。

次に本研究の課題を述べる。まず同じ間取り図でも表記が違う問題がある。本研究では、同じ間取り図を取り除くときに、画像が完全に同じもののみしか取り除いていない。同じ画像を提示するのはユーザーにとって好ましくないと考えられるので、検索結果の要約を今後の課題としたい。また、本研究の学習には間取り図画像をグラフ構造に落とし込んだもののうち、ノードのみを学習に用いておりエッジの部分も学習していない。部屋同士の繋がりの部分は学習できていないことが分かる。エッジも考慮した学習の方法として Siamese Network を用いる方法が考えられる。

謝辞

本研究の一部は科学研究費助成事業 (26700008)、不動産流通経営協会研究助成の支援を受けて行われた。

参考文献

- [1] <https://www.keieiken.co.jp/monthly/2016/0405/index.html> (2017/02/06 アクセス)。
- [2] 花里俊廣, 平野雄介, 佐々木誠. 首都圏で供給される民間分譲マンション 100m² 超住戸の隣接グラフによる分析. 日本建築学会計画系論文集, No. 591, pp. 9–16, 2005.
- [3] Chen Liu, Jiajun Wu, Pushmeet Kohli, and Yasutaka Furukawa. Deep multi-modal image correspondence learning. *arXiv preprint arXiv:1612.01225*, 2016.
- [4] 大原康平, 俊彦, 相澤清晴. 間取りや広さをクエリとする直感的な不動産検索システム. 情報処理学会第 78 回全国大会, 2016.



(a)



(b)

Fig. 3: 成功例



Fig. 4: 失敗例