

## 不動産間取り図画像の特徴量による物件類似性の導出

佐嘉田, 悠樹  
九州大学大学院システム情報科学府

伊東, 栄典  
九州大学情報基盤研究開発センター

<https://hdl.handle.net/2324/2544136>

---

出版情報 : SIG-KBS. B5 (2), pp.19-24, 2019-11-10. 人工知能学会

バージョン :

権利関係 : Notice for the use of this material. The copyright of this material is retained by the Japanese Society for Artificial Intelligence (JSAI). This material is published on this web site with the agreement of the authors and the JSAI. Please be complied with Copyright Law of Japan if any users wish to reproduce, make derivative work, distribute or make available to the public any part or whole thereof. / All Rights Reserved, (c) The Japanese Society for Artificial Intelligence.





# 不動産間取り図画像の特徴量による物件類似性の導出

## Derivation of Room Similarity from Floor Plan Image Features

佐嘉田悠樹<sup>1\*</sup> 伊東栄典<sup>2</sup>  
Yuki Sakata<sup>1</sup> Eisuke Ito<sup>2</sup>

<sup>1</sup> 九州大学 システム情報科学府

<sup>1</sup> Graduate School of ISEE, Kyushu University

<sup>2</sup> 九州大学 情報基盤研究開発センター

<sup>2</sup> Research Institute for Information Technology, Kyushu University

**Abstract:** On a real estate portal site in Japan, users can search for text information but not image information. This is because textual data entered by property management company is in an available format for searching, but image data is not. The floor plans which each property management company made in their respective formats are not compatible unlike the character provided in each fixed entry field. Therefore, the user needs to infer what kind of structure from the input character strings (number of rooms, facilities, area, etc.). In this study, we consider a system to search for similar floor plan images by giving a photograph of a construction toy structure as a search query in order to search for floor plan images. Deep CNN model (VGG16) is fine-tuned separately for the floor plan image and the query image. Subsequently, each feature vector is extracted from each learning layer of VGG16. Siamese Network is constructed to calculate similarity between the floor plan image and the query image by CNN learning the match / mismatch between the floor plan image and the query image. This paper gives an overview of the experimental procedures and results performed. This experiment deals with preprocessing of floor plan images and learning of floor plan images. In another experiment, the character information was extracted from the floor plan images.

## 1 はじめに

近年、経済界での FinTech に代表される IT 導入が、各業界へ波及している。不動産業界でも、テクノロジー活用「不動産テック (Real Estate Tech: ReTech)」が注目されている [1]。とりわけ日本の不動産業界は、取引の不透明性や複雑な制度によりデジタル化が妨げられていたものの、法改正や不動産ベンチャーの参入増大など徐々に緩和の動きがある。不動産テックの普及によって、情報量が増大し、取引が活性化し、ビッグデータ活用が可能となり、さらに潜在的な需要を明らかにする可能性がある。

現在、日本の不動産情報サイトは各不動産管理会社がそれぞれ入力した文字情報について、利用者が検索することができるものの、画像情報を検索に用いることができない。各入力欄が用意されている文字列情報と異なり、特に間取り図は各不動産管理会社が各々の

形式で作成しているため、互換性がない。そのため、利用者は入力された文字列情報 (部屋数・設備・面積) などから検索して類推することが強いられている。

本稿では、間取り図画像を対象に検索可能にする手法について取り扱う。本研究では検索クエリとして利用者が間取りを想定して組み立てたブロック模型の写真を与えて、類似した間取り図画像を表示できるシステムを考える。Deep CNN モデルの VGG16 を用いて、間取り図画像とクエリ画像についてそれぞれ転移学習させる。続いて、それぞれの VGG16 から特徴量を抽出して、CNN でクエリ画像と間取り図画像の一致・不一致を学習させる。

2 章では、間取り図の画像を対象に、多様な形式を同一の形式で扱えるように機械的処理を行う手法についての研究を 3 編紹介する。3 章では、本研究の目的と手法について取り上げる。最後に学習モデルを評価する。4 章では、実際に行った実験手順と、結果の概要について取り上げる。今回の実験では間取り図画像の前処理と間取り図画像の学習について取り上げる。また、間取り図画像の処理によって文字情報の取得する。

\*連絡先：九州大学システム情報科学府  
〒 819-0395 福岡県福岡市西区元岡 744  
E-mail: y.sakata.222@s.kyushu-u.ac.jp

取得した結果についても取り上げる。5章では、今後の研究の方針について示す。

## 2 関連研究

本研究の提案手法に関連する先行研究を述べる。まず、間取り図から情報の抽出を行い、部屋の自動検出および部屋の自動ラベリングを行う手法について提案している文献を紹介する。間取り図に対して3段階の処理を行い、部屋へラベリングを行っている。次に、間取り図と室内の写真の自動対応付けについて研究している文献を紹介する。間取り図と写真との間の類似性を求め、間取り図に対応する写真のマッチングや、間取り図の中で対応する領域の可視化を行っている。最後に間取り図同士の構造的類似度を求める手法について提案している文献を紹介する。間取り図とメタデータから間取りや部屋の有無などについて深層学習を行い、入力に類似する間取り図を提示させている。

### 2.1 間取り図画像からの構造抽出・ラベル抽出

Ahmedら[2]は、間取り図から部屋を検知して自動でラベリングする手法について研究している。白黒2色の間取り画像に対して、情報的セグメンテーション (Information Segmentation), 構造的解析 (Structural Analysis), 意味論的分析 (Semantic Analysis) の3段階の処理を行い、部屋へラベリングを行う手法を提案している。

情報的セグメンテーションでは画像から文字列領域の抽出、直線の抽出、線の太さによる分類など画像的な処理を行っている。構造的解析では線の情報に基づいた、壁の抽出、部屋の領域の決定など図形的な処理を行っている。意味論的分析では、文字列の抽出、扉や窓のシンボル抽出、部屋のラベリングなど意味に基づいた処理を行っている。

$2,479 \times 3,508$ , 80枚の間取り図画像を用いて評価実験をしたとき、部屋の認識精度は意味論的分析を用いていない手法[3]を上回る結果を得られている。このため、文字列の抽出などの意味論的分析は認識精度の向上に効果がある。

### 2.2 間取り図画像と部屋の写真の対応付け

Chenら[4]は、部屋の内装の写真と間取り図画像という異なる種類の画像を対応手法を実験している。間取り図と写真との間の類似性を求め、間取り図に対

応する写真のマッチングや、間取り図の中で対応する領域の可視化を行っている。

写真と間取り図の対応をさせるクロスモーダル画像マッチング問題を解決するために以下のような手順で行う。まず、ニューラルネットワーク (VGG16) を用いて各写真、間取り図の画像について学習させる。次に、VGG16のベクトルの連結をして完全連結回帰ネットワークからシヤムネットワーク (Siamese Network) を構築して類似度を学習させる。これを部屋の種類を区別せず学習させた場合、部屋の種類 (台所, 浴場, 居間) ごとに写真を学習させた場合、部屋の写真のセット (台所, 浴場, 居間の各ベクトルの連結) で学習させた場合について、間取り図のマッチングでの類似度を評価している。類似度を用いて逆に間取り図に一致する内装写真の分類をさせる実験も行っている。

結果としては、いずれの実験においても高い精度を記録し、いずれの場合でも人間の手作業 (Amazon Mechanical Turk) によるものより高い結果を出している。このため、形式が全く異なる種類の画像であってもシヤムネットワークで深層学習することで、人間より高い精度での対応付けができる。

### 2.3 間取り図画像と文字列情報を用いた類似間取り図画像の抽出

Takadaら[5]は、従来の間取り図画像の教師あり学習でのセマンティックセグメンテーションの手法[6]を簡易化するために、不動産メタデータ (文字列情報) と間取り図画像を用いてディープニューラルネットワーク学習を行い、間取り図画像の間取りタイプ自動分類・部屋の有無の自動判定を行う手法を提案している。さらにこの2つを組み合わせると類似する間取り図の提案する手法を研究している。

間取りのグラフ構造類似度  $p$  以上の画像を正解として、評価手法には precision@5 (検索結果上位5件のうち正解データが含まれる割合) を用いて実験している。20140枚の画像を用いて学習を行い、2000枚の画像に対してテストを行っている。出現数の多い間取りタイプの正答率が高く、間取りタイプの差が少ないものは低くなっている。部屋の有無の自動判定では情報量が少ない部屋の精度が低くなっている。間取り図の提案では間取りタイプのみ、部屋の有無のみの場合よりも、組み合わせた提案手法のほうが高い精度を出している。このため、間取り図タイプ、部屋の有無など複数の種類のラベリングを用いた学習が効果がある。

### 3 研究の目的

本研究では間取り図画像を検索するために、利用者の曖昧なイメージから間取り図を検索する手法を提案する。今回、利用者はイメージを表現するためにブロック模型を作成した後、上部から撮影し、写真をクエリとして入力する(図1)。クエリをもとに間取り図の検索を行い、利用者の意図した部屋の間取り図を出力することができるシステムを考える。検索システムを実現するための手順について以下で説明する。(図2)

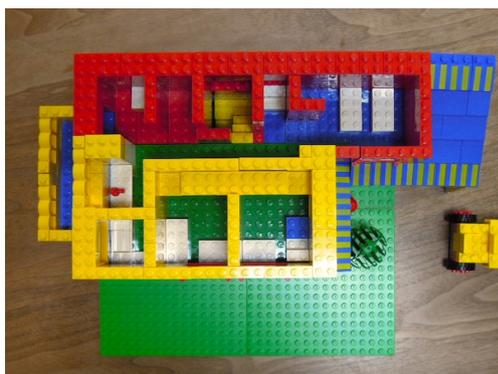


図1: ブロック模型の写真例

#### 3.1 前処理

まず、不動産情報サイトから各物件について間取り図と文字情報を取得する。文字情報は間取り(1DK, 2LDKなど)、面積、設備などが含まれる。

次に、間取り図の前処理を行い、間取り図を同一の学習モデルで扱いやすい形式に整形する。間取り図画像に含まれる、ファイル圧縮の過程で混入したノイズ、不動産会社によって異なるフォント、色、部屋の構造などの特徴を考慮した処理を行う。この処理は、色の置換、文字抽出、支配色ポスターライゼーション(Dominant Color Posterization)、平滑化、正規化、ZCA Whitening、ぼかしなどを組み合わせる。続いて、手作業で各間取り図に対応したブロック模型を作成し撮影して訓練用データを用意する。クエリの写真についても光源の明るさやノイズなどを考慮した前処理を行い、学習しやすい形式に整形する。

#### 3.2 画像学習

次に、間取り図に Python ライブラリ Keras の Deep-CNN 学習済みモデル VGG16 を用いて、不動産情報サイトから取得した間取り図に対応した間取り・部屋の有

無などを学習ラベルとして転移学習を行う。同様に写真クエリについても同様の学習ラベルを用いて VGG16 で別途転移学習を行う。間取り図および写真クエリを学習した各 VGG の学習層のベクトルを抽出して、間取り図の特徴量および写真クエリの特徴量とする。これによって画像情報をベクトルとして扱えるようにする。

#### 3.3 特徴量学習

その次に、間取り図と写真クエリの各特徴量を用いて CNN を用いて学習を行い、Siamese Network[7] を作成する。間取り図と写真クエリの一致(1) / 不一致(0) を学習ラベルとして学習させる。これによって間取り図と写真クエリについての類似度を計算できるようになる。(図3)

3.1~3.3 で述べた手順で検索システムを実現することができる。例えば、このシステムに対して利用者が新たなブロック模型写真を与えたとき、既存の間取り図との類似度を計算することで、最も類似した間取り図を検索できる。

#### 3.4 検索システムの評価

最後に、訓練用データ以外の検証用の写真クエリの特徴量を学習モデルに与えて、対応する間取り図の類似度が高くなるかを検証する。

## 4 実験

### 4.1 実験データ

実験用のデータとして間取り図を用意する必要がある。今回、実験データとして SUUMO<sup>1</sup> と LIFULL HOME'S<sup>2</sup> の Web ページ、および LIFULL HOME'S のデータセット [8] が候補として存在した。表1に3つの仕様について示す。3つのデータの比較の結果、間取り図の画像サイズが最も大きい SUUMO の Web ページを用いて実験を行うことに決定した。

物件情報検索サイト SUUMO をスクレイピングして福岡市西区の物件から間取り図画像 1702 件および付属する文字情報を取得し、以下実験データとして用いる。

<sup>1</sup><https://suumo.jp/>

<sup>2</sup><https://www.homes.co.jp/>

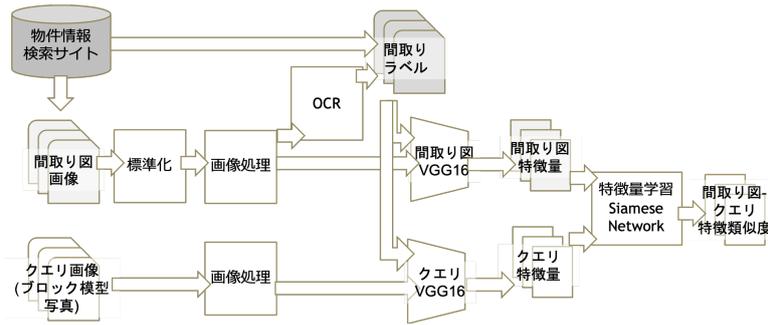


図 2: 実験の流れ

	SUUMO Web ページ	LIFULL HOME ' S Web ページ	LIFULL HOME ' S データセット
物件数	650 万	770 万	530 万
間取り図画像サイズ	800x800 以内	568x426 以内	120x120 以内
画像ファイル取得	スクレイピング	スクレイピング	NII 提供
メタデータ取得	スクレイピング	スクレイピング	NII 提供
時期	最新	最新	2015 年 9 月時点

表 1: 実験データ

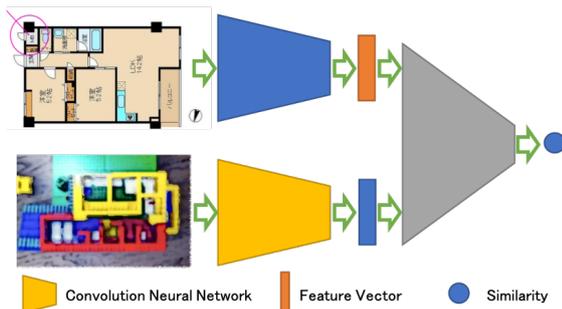


図 3: Siamese Network

## 4.2 実験 1 : 間取り図画像の処理と学習

実験データの間取り図画像に色の置き換えの処理 (Col Ex) を適用して整形を行った。比較対象として元のデータのまま実験を行うものも用意した。

その後、間取り図画像を間取りに応じて 1706 枚 22 クラス、部屋数に応じて 972 枚 4 クラスに分類して実験を行う。訓練用データ (Training) と検証用データ (Validation) に分割した。学習データ・バリデーションデータの分割比 (Val : Trn) の比較として、0.2(学習データ 80 % : バリデーションデータ 20 %), 0.5(学習データ 50 % : バリデーションデータ 50 %) の 2 通りを行った。

バッチ学習時の画像拡張処理の比較 (以下の 2 種類) として、A (変形): Shear(引き伸ばし) / Zoom(拡大・縮小), B (反転+変色): Flip(反転) / Channel Shift(色温度変化) / Standard Normalization(正規化), の 2 通

りを行った。

Imagenet を用いたもの、学習モデルの重みを用いなかったものの比較を行って転移学習をさせた。

学習済みモデル VGG16 を転移学習して学習を行い、結果を Loss と Accuracy 評価した。

## 4.3 結果 1

結果は表 2 の通りとなった。VGG16 の Weight に Imagenet を用いた場合は None より Acc が高いが過学習も起きやすい。画像拡張に A より B を用いた方が Val acc が高いという結果になった。画像の前処理として Col Ex を用いたものは 200epoch では Val loss が高い結果になった。このため Val loss を下げるように Epoch 数を増やして実験する必要がある。

## 4.4 実験 2 : 間取り図画像から文字列の抽出

不動産情報サイトに掲載されている文字情報だけでなく、間取り図画像内に含まれる文字を抽出して利用することを考える。ある間取り図画像に対して Python3 の OpenCV ライブラリを用いて手作業によって色の置き換えを行う。文字色以外の色を RGB(254,254,254) に置き換えを行って文字色のみを表示色にする。その後、TesseractOCR[9] を用いて文字情報と座標情報を抽出した。

画像処理	画像拡張	Weights	Ctgr	Val:Trn	Loss	Acc	Val loss	Val acc
None	A	Imagenet	22	0.2	0.4477	0.8628	1.2566	0.6167
None	B	Imagenet	22	0.5	0.2713	0.918	0.7931	0.7746
Col Ex	B	None	22	0.2	2.2588	0.1966	2.2362	0.1833
None	A	None	4	0.2	0.8997	0.6927	0.8939	0.6987
None	B	None	4	0.2	0.9124	0.6906	0.8501	0.716
Col Ex	B	None	4	0.2	1.1707	0.4234	1.1725	0.4321

表 2: 実験 1:VGG16 による間取り図の学習結果

## 4.5 結果 2

図 4 が実験前の間取り図画像, 図 5 が文字色抽出した間取り図画像である. 比較するためにそれぞれについて Tesseract OCR を適用したところ, 加工前では座標情報は取れていても文字列情報が取得できていないが, 加工後は文字列情報と座標情報がある程度取得できているという結果が得られた.



図 4: 未加工の画像



図 5: 加工後の画像

加工前	加工後
((66, 140), (169, 186))	
((33, 204), (657, 207))	
((37, 204), (657, 207))	
((98, 217), (657, 219))	
((98, 229), (953, 231))	
((98, 242), (956, 243))	
((98, 254), (954, 255))	((70, 123), (147, 124))
((302, 254), (954, 255))	Entrance ((53, 153), (167, 175))
((37, 264), (954, 266))	Toile ((173, 405), (246, 428))
((314, 266), (954, 268))	Living ((412, 245), (503, 276))
((169, 278), (954, 280))	9 ((446, 292), (467, 318))
((169, 291), (958, 292))	Dining,Kitchen ((343, 338), (569, 370))
((169, 303), (956, 304))	(9.0) ((422, 382), (490, 414))
((320, 314), (956, 316))	CL ((733, 170), (768, 190))
((319, 327), (954, 329))	Roomn ((769, 314), (864, 340))
((270, 376), (954, 378))	(6.(.)) ((781, 360), (852, 393))
((270, 388), (954, 390))	
((270, 401), (956, 402))	
((270, 425), (926, 427))	
((270, 437), (921, 439))	
((270, 450), (899, 451))	

表 3: 実験 2 : 文字列抽出結果

について述べた.

SUUMO のデータを用いた前処理と画像学習の予備実験を行った.

今後の流れとして, 間取り図画像の適当な処理を行う, VGG16 で精度の高くなるまで学習を行う. その後, 各間取り図に対応したクエリ用のブロック模型を作成する. ブロック模型の撮影をしてクエリ用の写真にして, 適当な処理を行い, 同様に VGG16 で学習させる. その後, 各学習モデルから特徴量を抽出して CNN で学習させる. 最後に学習済み CNN を用いてクエリ写真での汎化性能を検証する.

## 参考文献

- [1] Y. Kiyota, T. Yamasaki, H. Suwa, and C. Shimizu, "Real estate and AI," *Artificial Intelligence*, vol. 32, no. 4, pp. 529–535, 7 2017.
- [2] S. Ahmed, M. Liwicki, M. Weber, and A. Dengel, "Automatic room detection and room labeling from architectural floor plans," *2012 10th IAPR International Workshop on Document Analysis Systems, Gold Cost, QLD*, pp. 339–343, 2012.

## 5 おわりに

ICT 及び ReTech の進展に伴い, 文字や数値以外の情報を不動産サービスに用いることが検討されている. 本研究では不動産物件の間取り図に着目し, ブロック模型で作る間取り図写真に類似した物件を抽出することの実現を目指す.

本稿では関連する先行研究を調査した結果を述べた. また我々が提案する類似物件検索システムの実現手法

- [3] S. Macé, H. Locteau, E. Valveny, and S. Tabbone, “A system to detect rooms in architectural floor plan images,” 06 2010, pp. 167–174.
- [4] C. Liu, J. Wu, P. Kohli, and Y. Furukawa, “Deep multi-modal image correspondence learning,” 2016.
- [5] Y. Takada, N. Inoue, T. Yamasaki, and K. Aizawa, “Similar floor plan retrieval featuring multi-task learning of layout type classification and room presence prediction,” *2018 IEEE International Conference on Consumer Electronics (ICCE), Las Vegas, NV*, pp. 1–6, 2018.
- [6] 大原康平, 山崎俊彦, and 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, 5Y-08*, 2016.
- [7] G. Koch, R. Zemel, and R. Salakhutdinov, “Siamese neural networks for one-shot image recognition,” *ICML Deep Learning Workshop. Vol. 2*, 2015.
- [8] 情報学研究データリポジトリ lifull home’s データセット. [Online]. Available: <https://www.nii.ac.jp/dsc/idr/lifull/>
- [9] Tesseract. [Online]. Available: <https://github.com/tesseract-ocr/tesseract>