

Bottleneck 特徴量を用いた不動産画像の分類

Classification of Real Estate Images using Bottleneck Features

豊原 優^{*1} 門 洋一^{*2} 山崎俊彦^{*3} 藤森 進^{*1} 太原育夫^{*1}
 Suguru Toyohara Youiti Kado Toshihiko Yamasaki Susumu Fujimori Ikuo Tahara

^{*1} 東京理科大学 ^{*2} アットホーム株式会社 ^{*3} 東京大学

Tokyo University of Science At Home Co., Ltd. The University of Tokyo

Image processing enhanced by deep learning is now being actively applied to real-world issues. The real estate industry, which owns lots of property information, also expects to develop new services using deep learning. In this paper, we present the classification methods for property images using bottleneck features generated by convolution networks (CNN), and experimentally show that we can achieve 88% accuracy in Top-1.

1. はじめに

ILSVRC2012において、Convolutional Neural Network(以下、CNN)が優勝して以来、CNNは世界中の研究者たちに大いに注目され、それに伴い画像解析技術はさらに飛躍的に進歩を遂げている。さらに、この技術をビジネスに活かす為の研究開発も盛んに行われ、自動車業界や医療分野にまで広がっている。

一方で不動産業界においては、未だアナログな部分が多く残っているが、Real-Estate-Techと呼ばれる不動産と情報テクノロジーを掛け合わせて新しい価値を生み出す動きが始まっている。具体的には、機械学習理論を用いた価格推定や、新たな推薦システムの仕組みなどが新たに検討されており、これからも大いに注目される分野と言える。

本論文では、CNNを用いた画像分類技術を不動産画像に対し適用した際の分類手法を提案し、不動産会社の負担を減らす様々な業務利用可能な学習データ収集を可能にするシステムを、実運用までを視野に入れ検討する。

2. 物件画像の分類

2.1 CNNを用いた特徴抽出

CNNを用いた深層特徴量の一つに、Bottleneck特徴量と呼ばれる特徴量がある。これはCNNにおける畳み込み最終層の出力する特徴量のことである。ILSVRCでの学習データであるImageNet [Russakovsky 2014]ベースに学習を行ったCNNでは、この特徴量が画像の特徴を非常によく表していることが分かっている為、多くのCNN研究においてImageNetベースの重みを利用している。

当論文ではGoogLeNet (Inception-v3) [Szegedy 2015]におけるImageNet学習済み重みを用いてBottleneck特徴量を計算している。これに多層ペーセptronを用いて解析することで画像の種別判定を行うこととする。

2.2 問題設定

画像のタグ付けを自動化するシステムを考える上で重要なのが、Top1-Errorでの評価指標で考えなければならないという点

連絡先: 豊原 優、東京理科大学工学研究科経営工学専攻、
〒125-8585 東京都葛飾区新宿6-3-1 太原研究室、
E-mail:4417619@ed.tus.ac.jp

にある。不動産業者がシステムを利用した際に、画像登録時にシステムが自動でタグ付けをすることを想定しており、確信度が最も高いクラスを自動で登録する必要がある。また、ランニングコストの観点から、複数の畳み込みモデルによる複数のBottleneck特徴量生成を用いたアンサンブル学習は用いることができない。

分類するカテゴリは、外観・玄関・エントランス・室内・リビング・お風呂など、システム上付けることのできる全20クラスとする。その内訳をTable1に示す。

2.3 提案手法

実運用をする際には、精度とコストのバランスが重要であり、少ない計算量でいかに精度を上げるかが課題となる。本論文で提案する手法は、共通のBottleneck特徴量を用いて、複数の全結合層に解析させることで精度を向上させるものである。提案する学習モデルは、1つの共通の畳み込み層を用いて、データセットの組み合わせを変え学習した各全結合層のモデルと、畳み込み層を除いた各全結合層のみでアンサンブル化したモデルである。

それぞれの全結合層では、データセットの組み合わせを変えることで精度がどのように変わるのがかを観察し、最終的にアンサンブル化した場合とで、どのように性能が変わるのがかを調べた。

さらに、実験条件にて述べるD5に関しては、間違いやすいカテゴリに対して寝室や外階段といった、さらに粒度の細かいカテゴリを用意することで、どのように精度が向上するか、アンサンブルにおいて多様性を発揮するかを調べた。また、本論文のアンサンブル方式は、各全結合層の出力における最大スコアのカテゴリに対して投票を行う方式である。

Table 1: 20 カテゴリの詳細とその出現頻度

カテゴリ	出現比率	カテゴリ	出現比率
外観	12.2%	洗面所	4.8%
バルコニー	3.5%	間取り図	10.5%
お風呂	6.5%	リビング	5.5%
共有部	2.8%	地図	0.1%
エントランス	2.0%	駐車場	1.3%
玄関	5.4%	収納	5.6%
設備	6.4%	周辺環境	5.0%
庭	0.2%	トイレ	5.5%
室内	12.7%	地形図	0.6%
キッチン	8.0%	眺望	1.6%

