

ドローン搭載カメラ画像を用いた物体認識

安岡 里都¹, 藺頭 元春², 飯山 将晃^{3*}

¹ 京都府立洛北高等学校, ² 京都大学大学院情報学研究科, ³ 京都大学学術情報メディアセンター

要旨

本研究では、室内環境を対象として環境内に存在する物体を自律移動型のロボット(ドローン)を用いて認識するシステムの構築を行う。画像認識などの分野で活用が盛んに研究されている深層学習を用い、ドローンから撮影した画像に何が写っているのかを判別すること、すなわち物体認識を行うことに対して、その精度を高めることを目的とした。研究手法としては、まずドローンを室内で飛行させ、ドローン前方に付けられているカメラで室内の様子を撮影した。その後、撮影した画像を用いて室内に何があるのか物体認識を行った。物体認識にはVGG16を用いPython製の深層学習ライブラリであるKerasにより実装した。画像内に複数の物体が存在する場合もしくは何も存在しない場合などに対して、尤度を活用して安定して認識できるようにした。

重要語句: 物体認識, 深層学習, 畳み込みニューラルネットワーク, ドローン, 尤度

1. 序論

画像処理技術の発展に伴い、民生用カメラやスマートフォンを用いた様々なアプリケーションが開発・利用されている。例えば、監視カメラにおける人物検出や自動車の自動運転、医療診断などが挙げられる。物体認識、すなわち画像に映っている物体が何かを識別する処理は、人間における視覚での認識処理に対応し、これを計算機上で実現することにより、様々な応用が可能となる。

本研究では、特に室内環境での物体認識を目的として、自律移動型のロボットに搭載されたカメラを利用した物体認識技術を提案し、その有効性を評価する。人間にとっては、目に入った物体が何の物体かを瞬時に認識することは難しくない。しかしながら、そのような人の優れた視覚機能を計算機で実現することは容易ではない。逆にいえばその技術が確立されれば、ロボットも人間のように自ら状況を把握し、適切な判断・行動が可能になるはずである。これを室内環境下で実現できれば、家庭内での生活支援ロボットなどへの応用が期待できる。

自律移動型のロボットとして、本研究ではドローンを

用いる。ドローンは特に災害時などに活用することができ、ドローンから撮影した画像を用いて、災害の状況を上空から広範囲にわたって把握したり、人が立ち入ることができない場所の状況を調べたりすることができる。また室内環境下においても、床上の障害物などの影響を受けずに観測が可能であるという利点がある。

画像からの物体認識についてはVGG⁽¹⁾やAlexnet⁽²⁾など畳み込みニューラルネットを利用した手法が提案されている。しかしながら、ドローンで撮影された画像に対して物体認識を行う場合、撮影された画像内に認識対象となり得るような物体が写っていない場合が多く存在し、そのような画像に対して物体認識を行うことにより、誤認識、特に本来そこに存在しない物体が存在すると誤判定されてしまう場合が多く生じる。これに対処するため、我々は物体認識の尤度を利用することで、このような誤認識の問題を克服する。

2. 提案手法

2.1. ドローンによる室内環境の撮影

ドローンとは人が搭乗しない無人飛行機の総称であり、遠隔操縦や自動操縦ができる。本研究に用いたドローンは図1のように4つのプロペラがついた小型のもので、プログラムによって飛行を制御できる。地面に着地した状態で起動させると、プロペラが回転して浮上する。また4方向に飛行でき、進路を変えるために回転も可能である。さらに本体の前頭部に小型のカメラがついており、動画を撮影できる。本研究では、あらかじめ作られたプログラムに従い、決められた動作を行って飛行させた。



図1. ドローン

* 内容に関する連絡先: iiyama@mm.media.kyoto-u.ac.jp

2.2. 畳み込みニューラルネットを用いた物体認識

本研究の物体認識では、Python製の深層学習ライブラリであるKerasを用いた。深層学習とは近年研究が進み画像認識分野へ応用されている技術である。深層学習の主要な技術は畳み込みニューラルネットワークである。畳み込みニューラルネットワークとは人間の脳内の神経構造を元にしたニューラルネットワークを多層化したものであり、畳み込み層とプーリング層から成る。本研究で用いたVGG16は2014年ILSVR (ImageNet Large Scale Visual Recognition) で提案されたもので⁽¹⁾、畳み込み層13層とフル結合3層の合計16層からなる。また、出力層は1,000ユニット、1,000クラスを分類するニューラルネットワークである。物体認識に用いる入力画像には認識対象が1つだけ含まれているという仮定がある。

撮影後、画像を保存して物体認識を行う。撮影した動画から1枚ずつ画像を取り出し、物体認識を行う入力画像とする。物体認識の結果、どんな物体がどれくらいその物体らしいのか、すなわち、画像中に存在すると思われる物体名と物体らしさを表す尤度が得られる。

2.3. 物体認識尤度を用いた誤判定への対処

第2.2節の処理により物体認識を行う際、物体が存在しないような場合にでも何かしらの認識結果が得られてしまう。そのため、物体が存在しないことを判定する処理が必要となる。第2.2節の処理では、画像中に物体が存在する場合はその物体の尤度が他の候補と比較して高くなる。また画像中に物体が存在しない場合はどの候補の物体の尤度も低くなる。ここで、1,000個の物体に対する尤度の総和が1という制約があるため、それぞれの尤度は物体が画像中に存在する場合に比べて比較的小さい値になる。すなわち、物体が存在しない場合、1,000個の物体に対する尤度は、どの値も小さくすると期待できる。そこで、閾値 t を定めて、どの認識結果も t 以下であるような画像に対しては、物体が存在しないと判定する。

3. 実験と考察

ドローンは室内で飛行させる。このドローンはPythonのプログラムで制御されており、地上から約1mの高さを飛行する。まずドローンを地上に置いた状態で起動し、真上に上昇させた後、ドローン前方についているカメラによる室内の撮影を開始する。ドローンは3秒間空中で静止（ホバリング）し、その後少しずつ回転する。そして静止と回転を繰り返しながら室内180度を撮影する。ここで画像のブレを少なくするため、撮影はホバリング中のみ行い、回転中は撮影を停止する。

撮影した画像に対する物体認識手法としてVGG16を用い、PythonのKerasライブラリを使って実装した。認識対象候補としてILSVRCの1,000種類の物体を選んだ。

物体認識の結果として撮影画像に存在すると思われる



図2. 冷蔵庫と認識されたシーン



図3. 何もないと認識されたシーン

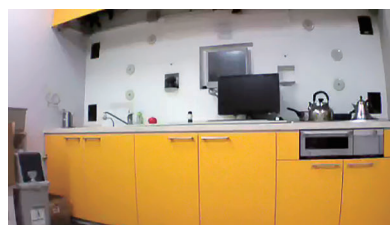


図4. 誤認識の例。冷蔵庫と誤認識

表1. 図2の尤度上位3位の物体と尤度

	物体名	尤度
1位	冷蔵庫	0.78038335
2位	電子レンジ	0.1134325
3位	食器洗い機	0.041876428

物体名と物体らしさを数値的に表した尤度を得られる。本研究では得られた結果から、物体認識の確率が高い上位3位の物体の名前と尤度を用いて考察した。図2, 3, 4は実験に用いた画像の例である。図2の画像を用いた物体認識の結果は表1に示す通り、尤度が1位だった物体は冷蔵庫で、尤度の値は0.78と高くなった。すなわち図2においては画像中に存在する冷蔵庫が正しく認識されている。また図3を用いた物体認識の結果、閾値0.1を超える尤度をもつ物体はなく、画像中には何もないと判定された。すなわち、画像中に特に認識対象となるような物体が存在しない場合、提案手法により正確に物体が存在しないと判定することができた。しかしながら、図4の場合には、誤認識が生じることがあった。この時の尤度の高い上位3つの認識候補を表2に示す。1位の候補（冷蔵庫）の尤度が0.14と閾値以上となっているが、2位（引き戸）や3位（薬箱）の候補についても尤度がそれぞれ0.13, 0.12と、閾値に近い値となっている。このような場合、1位の尤度を用いるだけでなく、1位と2位、1位と3位の尤度の差が十分に大きいかどうかを利用して誤認識を回避することなどが可能ではないかと考えられる。

表2. 図4の尤度上位3位の物体名と尤度

	物体名	尤度
1位	冷蔵庫	0.13631952
2位	引き戸	0.12500748
3位	薬箱	0.1241414

実際に、表1の結果において、1位の冷蔵庫の尤度と2位の電子レンジの尤度との比が0.145であるのに対し、表2の結果においては0.917であり、尤度の比に対して閾値を設定することにより物体の有無を判定できるのではないかと考えられる。

4. 結論

本研究では、ドローンから撮影した画像を用い、尤度を活用することで誤認識に対処できる物体認識のシステムを提案した。実験結果より、画像中に存在する物体を正しく認識できることが確認できた。今後の課題としては、1位と2位、また1位と3位の尤度の差を利用した誤認識の改善や、複数の物体が同一画像中に存在する場合への対処などが挙げられる。

謝辞

本研究はJSTグローバルサイエンスキャンパスの支援のもとおこなった。ELCAS事務局の方々には様々な方面で支えていただいた。貴重な機会を私に与えてくださった多くの方々にこの場を借りてお礼申し上げる。

参考文献

1. Karen Simonyan, Andrew Zisserman. “Very Deep Convolutional Networks for Large-Scale Image Recognition”. arXiv: 1409.1556. (2015).
2. Alex Krizhevsky, Ilya Sutskever, Geoffrey E. Hinton. “ImageNet classification with deep convolutional neural networks”. Advances in Neural Information Processing Systems 25(NIPS 2014).

An Object Recognition with Photographs That Was Taken By a Drone’ Camera

RITO YASUOKA¹, MOTOHARU SONOGASHIRA² & MASAOKI IYAMA^{3*}

¹Kyoto Prefectural Rakuhoku high school, ²Graduate School of informatics, Kyoto University, ³Academic Center for Computing and Media Studies, Kyoto University

Abstract

In this paper, we report on a system that recognizes objects indoors with an autonomous mobile robot drone. A contribution of our research is that we improve the accuracy of object recognition in such scenes by employing a deep learning technique that has been studied in the field of image recognition. Our system takes images using a flying drone, and recognizes an object in the images. We developed our system using Python with a Keras framework. In order to recognize a situation with no objects in a scene, we use the likelihood of the object recognition.

Key words: Objects recognition, Deep learning, Convolutional neural network, Drone, Particular degree

*Corresponding Researcher: iiyama@mm.media.kyoto-u.ac.jp