Graspability を導入した DCNN による物体把持位置検出

EP13004 荒木諒介　　指導教授：藤吉弘亘

1. はじめに

産業用ロボットや生活支援ロボットにおいて工業部品や日用品などの物体を正確に把持する (掴む) ことは様々な作業の基本となるため，重要なタスクとして挙げられる．このタスクは，ロボットに搭載されているカメラで撮影した画像から把持可能な領域を検出する．本研究では，Deep Convolutional Neural Network(DCNN) を利用して，ロボットハンドの開き幅や回転を考慮した物体把持位置を検出する．また，把持可能性を同時に出力することで，高精度な物体把持位置検出を実現する．

2. Deep Convolutional Neural Network

DCNN は，重みフィルタの畳み込みとプーリングを複数回繰り返して得られた特徴マップを全結合層へ入力して，回帰や識別問題を解く機械学習法である．

畳み込み層では，入力された画像または特徴マップに $n \times n$ の重みフィルタを畳み込み，得られた値 v を活性化関数 $f(v)$ へ入力する．活性化関数の出力値を用いて特徴マップを作成する．提案手法では，活性化関数に Rectified Linear Unit (ReLU) を用いる．

プーリング層では，一定間隔で局所領域を設定し，局所領域内の最大値でサンプリングするマックスプーリングを行う．

3. 提案手法

提案手法で用いる DCNN の構成の全体構成を図 1 に示す．提案手法では，把持位置だけでなく把持可能性を合わせて出力する．以下に出力内容の詳細を示す．

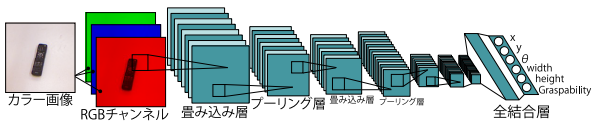

図 1：提案手法の DCNN の構成

3.1. 把持位置の 5 次元表現

本研究では，Lenz ら [1] が用いた Five-Dimensional Representation (5 次元表現) と同様の情報を DCNN で出力する．この出力情報は，物体把持位置の x および y 座標，把持矩形の回転角 (θ)，把持矩形の高さ (height)，把持矩形の幅 (width) の 5 つの要素から成り，ロボットハンドの開き幅と回転を表現することができる．

3.2. Graspability の導入

本研究では，把持位置の検出と同時に，その把持可能性 (Graspability) を出力する．Graspability は物体の把持が可能であるかを表す指標で，0 から 1 の値となる．「物体が存在しない」または「物体が存在するが，把持できない把持位置」の場合は 0,「物体であり，把持が可能である把持位置」の場合は 1 を出力するように学習する．

学習時の教師信号には，図 2 のように物体領域と背景領域の二値画像に対してガウシアンフィルタを畳み込んで作成した Graspability マップを用いる．物体の中心に近い把持位置の場合に 1.0 となり，中心から遠ざかるほど 0.0 に近い値となる．ただし，把持できない把持位置は物体の中心に近い場合でも 0.0 とする．

図 2：物体画像から Graspability マップを作成する手順

3.3. 把持位置の検出

把持位置を検出する際，画像中をラスタスキャンして検出ウィンドウを逐次 DCNN に入力する．検出ウィンドウ 1 つにつき 1 つの把持位置が得られるが，これらの把持位置のうち最も Graspability の高い把持位置を最終的な把持位置として出力する．

4. 評価実験

Graspability を導入した DCNN による物体把持位置検出の有効性を示すために，評価実験を行う．

4.1. 実験概要

把持位置検出の成功判定には，正解把持矩形と検出結果の矩形の重なり率を用いる．正解把持矩形と検出結果の矩形の重なった面積が，両矩形の合計面積の 30%以上であれば検出成功とする．

学習および評価には Cornell Grasping Dataset を使用する．データセットには 280 種類のアイテムの RGB 画像が 885 枚含まれている．また，各 RGB 画像に対して，距離点群データ，Positive 教師信号，Negative 教師信号が用意されている．

4.2. 実験結果

提案手法と従来手法の 2 段階 DNN[1]，Graspabiliy を導入しない場合の DCNN の検出結果を表 1 に示す．提案手法は従来手法と比べ検出率を 22.76 %向上させることができた．また，Graspability を導入しない DCNN と比べ，Graspability を導入することで検出率を 35.26 %向上させることができた．ラスタスキャンによる検出ウィンドウ数が 240 個のとき，検出時間は 1 枚あたり約 30 秒である．また，GPU による高速演算を用いることでリアルタイムに検出することができる．検出結果の一例を図 3 に示す．

表 1：検出率の比較

検出手法	従来手法	DCNN	提案手法
検出率 [%]	54.02	41.52	76.78

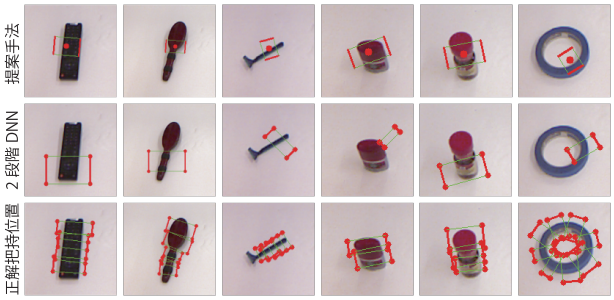

図 3：検出結果

5. おわりに

本研究では，Graspability を導入した DCNN による高精度な物体把持位置の検出法を提案した．DCNN を利用して把持位置の候補を検出するだけではなく，Graspability を導入することで未知物体の適切な把持位置検出を実現した．今後は，3 次元ベースでの把持位置検出，特定物体認識と把持位置の同時検出に取り組んでいく．

参考文献

[1] I. Lenz, *et al.*, “Deep Learning for Detecting Robotic Grasps”, IJRR, vol.34, no.4-5, pp.705-724, 2015.

[2] Y. Domae, *et al.*, “Fast Graspability Evaluation on Single Depth Maps for Bin Picking with General Grippers”, ICRA, pp.1997-2004, 2014.