

マンガキャラクターのマルチビュー顔検出に関する検討

A Study of Multi-view Face Detection for Characters in Comic Images

柳澤秀彰
Hideaki Yanagisawa

渡辺 裕
Hiroschi Watanabe

早稲田大学大学院 基幹理工学研究科
Graduate School of Fundamental Science and Engineering, WASEDA University

1. まえがき

電子コミックは電子書籍市場の売上の 80%以上を占める重要なコンテンツである。マンガ画像からメタデータを抽出し、タグ付けすることによって、高度なマンガ検索システムや自動要約作成などの様々な機能を付加できる。

マンガキャラクターの正面顔の検出について DPM や R-CNN の有効性が示されている[1]。しかし、横顔を含んだ顔検出に関する研究は行なわれていない。本稿では、DPM および R-CNN を用いたマンガキャラクターのマルチビュー顔検出について検討を行なった。

2. 検出手法

(1) Deformable Part Model (DPM) [2] [3]

DPM は物体の姿勢変化に頑強な検出手法であり、対象を複数のパーツから構成されるツリー構造として捉え、物体の全体および各パーツの形状とパーツ位置のずれに対するコストから評価する。

(2) Regions with CNN feature (R-CNN) [4]

R-CNN はニューラルネットを用いた物体検出手法である。まず、Selective Search によって入力画像から物体の候補領域を抽出し、リサイズした候補領域を畳み込みニューラルネットワーク (CNN) に入力する。次に、CNN の出力を SVM で分類し、候補領域の判定を行なう。

マルチビュー顔検出において、Farfadi らは Selective Search の代わりにスライディングウィンドウを使用することで検出率が向上すると報告している[5]。

3. 実験

DPM と R-CNN について、それぞれキャラクターの正面および横顔を学習した検出器の比較を行なった。本研究では、マンガ画像からバウンディングボックスで顔領域を指定したものを正例、顔以外の領域を切り抜いたものを負例とした。学習セットは正例を正面顔 400 枚・横顔 200 枚、負例を 1000 枚とした。テストセットは正例を正面顔 400 枚・横顔 200 枚、負例を 2000 枚とした。

DPM について、正面および横顔の左右方向に対応する 4 種類の検出器を作成し、パートフィルタ枚数を 4 枚、Non-maximal suppression (NMS) を 0.1 と設定した。また、R-CNN では正面と横顔を分類せずに学習を行ない、候補領域の抽出に Selective Search を用いた検出器とスライディングウィンドウを用いた検出器の 2 種類を作成した。NMS の値は DPM と同様に 0.1 に設定した。

実験結果を図 1 に示す。Selective Search を用いた R-CNN の検出率は、DPM やスライディングウィンドウを用いた

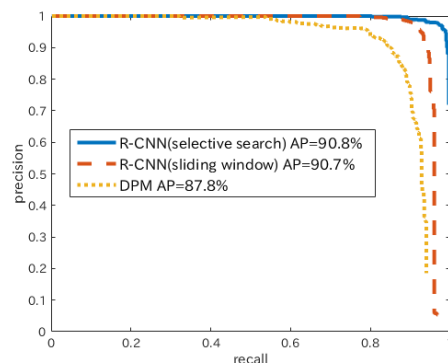


図 1: 顔検出結果の比較

手法より高くなった。この理由として、マンガ画像は線画より構成されるので、Selective Search によるセグメンテーションで候補領域を正確に抽出できるためと考えられる。

4. まとめ

本稿では DPM および R-CNN を用いたマンガキャラクターのマルチビュー顔検出について検討を行なった。その結果、Selective Search を利用した R-CNN が最もマンガ画像に適していることが分かった。

謝辞 本研究は JSPS 科研費 25330137 の助成を受けたものである。

参考文献

- [1] 柳澤秀彰, 渡辺裕: “R-CNN を用いたマンガキャラクター検出に関する一検討”, 映像メディア処理シンポジウム, Nov. 2015.
- [2] P. Felzenszalb, D. McAllester, D. Ramanan, “A Discriminatively Trained, Multiscale, Deformable Part Model,” IEEE Conference on Computer Vision and Pattern Recognition, 2008.
- [3] P. F. Felzenszalb, R. B. Girshick, D. McAllester, and D. Ramanan, “Object detection with discriminatively trained part-based models”, IEEE Trans. Pattern Anal. Mach. Intell., vol.32, no.9, pp.1627-45, Sep. 2010.
- [4] R. Girshick, J. Donahue, T. Darrell, J. Malik: “Rich feature hierarches for accurate object detection and semantic segmentation,” in IEEE Conference on Computer Vision and Pattern Recognition, 2014.
- [5] S.S. Farfadi, M. Saberian, L. Li: “Multi-view Face Detection Using Deep Convolutional Neural Networks”, in International Conference on Multimedia Retrieval, Apr. 2015.