

# R-CNN を用いたマンガキャラクター検出に関する一検討

## A Study on Character Detection for Comic Images with R-CNN

柳澤 秀彰 渡辺 裕  
Hideaki Yanagisawa Hiroshi Watanabe

早稲田大学大学院 基幹理工学研究科 情報通信専攻  
Graduate School of Fundamental Science and Engineering,  
Waseda University

**Abstract:** 近年の電子コミックの普及に伴い、マンガコンテンツのメタデータを利用して作品の検索や要約の作成を行なうといった新たなサービスが提案されている。しかし、現状ではマンガ画像より手動でメタデータ抽出を行なう必要があるため、効率化のためにメタデータを自動抽出する技術が必要である。本研究では、マンガ画像からの登場キャラクターを検出について検討を行なう。過去の研究において、HOG 特徴量や Deformable Part Model (DPM)などの一般物体検出に用いられる手法がマンガ画像に対しても有効であることが示されている。一方、現在の一般物体検出の分野において、畳み込みニューラルネットワーク(CNN)を用いた手法が既存手法を上回る検出精度を示していることが示されている。本稿では、CNN で画像特徴量学習し、物体検出を行なう手法である Regions with Convolutional Neural Network Features (R-CNN)のマンガ画像への適用について検討を行なった。その結果 R-CNN は DPM と同様に、高い精度でキャラクターの検出が可能であることが分かった。

## 1 はじめに

近年の電子書籍の普及に伴い、文字情報やオブジェクトといったマンガ内のコンテンツをコード化することによって、電子コミックに新たなサービスを提供するといった提案がなされている。しかし現状の電子コミックコンテンツの多くは単に紙媒体のマンガをスキャンして電子化したものであるため、コード化には人手でメタデータを付与する必要があるため、コスト面が問題となっている。よって電子コミックサービス実用化のために、マンガコンテンツよりメタデータを自動抽出する技術が必要となっている。

本研究では、メタデータ抽出技術の中でマンガ画像内のキャラクター位置検出について検討する。本稿では、畳み込みニューラルネットワークで学習された画像特徴量を用いて物体検出を行なう手法である Regions with Convolutional Neural Network Features (R-CNN)のマンガ画像への適用について検討した。

## 2 従来研究

マンガ画像よりメタデータを抽出する技術は、現在コマの識別や、キャラクター・フキダシの抽出について研究が行なわれている。

キャラクターの抽出において、マンガキャラクターは主に2値の線画によって表現されるため、一般物体と比較すると、認識に使用できる特徴が少ないことや、場面ごとの見た目の変化が大きいといった問題がある。過去研究では、HOG 特徴量と対象物体の各パーツの位置関係を利用した物体検出手法である Deformable Part Model (DPM)がマンガキャラクター検出に有効であることが示されている[1][2]。

一方、近年の一般物体検出分野では、ニューラルネットワークを用いた手法が高い性能を示しているが、この手法がマンガ画像に対しても有効であるかとい

った検討はまだ行なわれていない。

## 3 R-CNN

R-CNN は Girshick らによって提案された一般物体検出のアルゴリズムであり、従来の手法と比べ、物体検出の精度と速度が向上していることが示されている[3]。R-CNN による物体検出の流れは以下のようになっている。

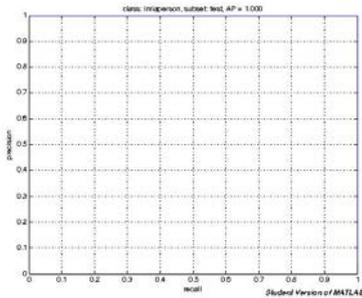
1. Selective Search によって画像内より似たような領域をセグメンテーション化し、物体の候補領域を抽出する。
2. 抽出された候補領域をそれぞれ規定の大きさにリサイズし、畳み込みニューラルネットワークに入力する。
3. 畳み込みニューラルネットワークの出力を特徴量として SVM で分類する。
4. 矩形の座標を回帰して、候補領域のズレを補正する。

R-CNN は一般物体のデータセットである PASCAL VOC 2007 および PASCAL VOC 2010 における実験において、DPM 等の従来手法を上回る精度を記録しており、D-CNN で抽出した特徴量が HOG 特徴量より高い記述能力を持つことが示唆されている。

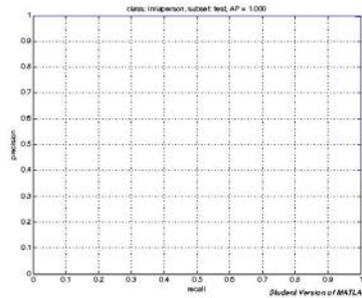
## 4 実験

R-CNN と既存手法である DPM について、マンガキャラクターに対する検出率の比較を行なった。マンガ作品ではキャラクターごとの特徴変化が大きいため、学習およびテスト画像に含まれるキャラクターの種類によって検出率が影響される。本実験では、3種類のキャラクターA, B, C それぞれについて R-CNN と DPM の顔検出器を作成し、各キャラクターに対する検出率を比較した。

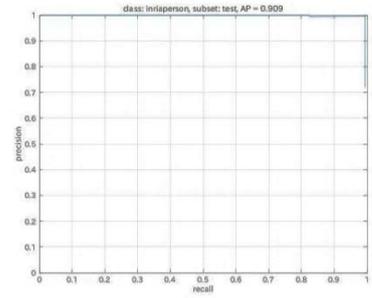
### 4.1 実験条件



キャラクターA

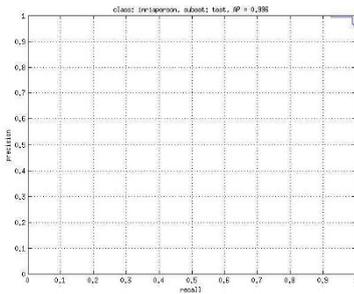


キャラクターB

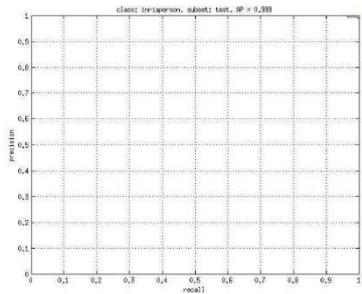


キャラクターC

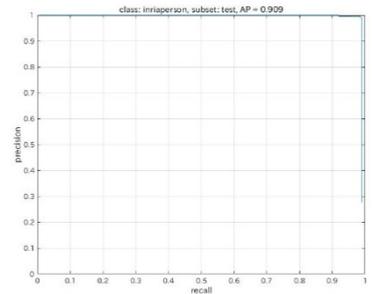
図 1: R-CNN による顔検出結果



キャラクターA



キャラクターB



キャラクターC

図 2: DPM による顔検出結果

表 1: R-CNN と DPM の比較(AP)

	A	B	C
R-CNN	1.000	1.000	0.909
DPM	0.996	0.999	0.909

R-CNN と DPM のプログラムには, Fast R-CNN [4] および voc-release 5 [5]をそれぞれ使用した. ここで Fast R-CNN は R-CNN を改良し検出精度の向上や計算量の削減などを行なった手法である.

検出器の学習では, キャラクターの顔領域を指定したマンガ画像を正例, 顔以外の領域を切り出した画像を負例として, 各キャラクターの顔検出器を作成した. 学習に使用する画像の枚数はそれぞれ正例 200 枚, 負例 1000 枚とした.

また, 顔領域検出のテストに使用する画像は正例 200 枚, 負例 500 枚であり, 全て学習に使用したものと異なる画像とした.

## 4.2 実験結果

キャラクターA, B, C について, R-CNN による顔検出結果を図 1 に, DPM による顔検出結果を図 2 に示す. 図 1, 2 について, 縦軸は適合率, 横軸は再現率を表している. また, R-CNN と DPM による平均適合率(AP)の比較を表 1 に示す.

実験結果より, R-CNN と DPM とともに 90%以上の精度でキャラクターの検出が可能であることが示された. また, キャラクターA, B については R-CNN の検出率がわずかに DPM を上回った. このことからニューラルネットを用いた物体検出手法がマンガ画像に対しても有効であることが分かった.

## 5 まとめ

本稿では, R-CNN をマンガ画像に適用した際のキャラクター検出精度について検討を行なった. DPM との比較実験から R-CNN がマンガ画像に対しても有効であることを確認した. 今後は両手法のより正確な検出性能を検討するために, 学習枚数を変化させた場合における性能の比較を行いたい.

## 参考文献

- [1] 石井大祐, 渡辺裕: マンガからの自動キャラクター検出に関する検討, 情報処理学会研究報告, Vol.2012-AVM-76, No.1, pp.1-5 (2012).
- [2] H. Yanagisawa, D. Ishii, H. Watanabe: “Face detection for comic images with deformable part model,” IEVC2014, 4A-1, (2014).
- [3] R. Girshick, J. Donahue, T. Darrell, J. Malik: Rich feature hierarchies for accurate object detection and semantic segmentation in IEEE conference on Computer Vision and Pattern Recognition, (2014).
- [4] R. Girshick: Fast R-CNN, arXiv preprint arXiv:1504.08083, (2015).
- [5] R. B. Girshick, P. F. Felzenszalb, D. McAllester: Discriminatively Trained Deformable Part Models, Release 5, <http://people.cs.uchicago.edu/~rbg/latent-release5/>,

早稲田大学大学院 基幹理工学研究科 情報通信  
専攻 渡辺研究室  
〒169-0072 東京都新宿区大久保 3-14-9 早大シル  
マンホール 401  
Phone: 03-5286-2509  
E-mail: bule-cosmo@ruri.waseda.jp