

深層学習に基づく漫画における 小さいオブジェクトの検出

仲 偉博[†] 古田 諒佑^{††} 谷口 行信[†] 日並 遼太^{†††} 石渡 祥之佑^{†††}
[†] 東京理科大学 ^{††} 東京大学 ^{†††} Mantra 株式会社

1. はじめに

近年、デジタル化されたコミックが広く普及しているのに伴い、漫画に対する画像解析技術が重要になっている。例えば、漫画内のテキストやキャラクターなどのオブジェクトを自動で検出できれば、自動翻訳等に应用可能である。しかし、漫画には小さく書かれたテキストなどの検出の難しい小さいオブジェクトが多数存在している。本研究では、漫画内の小さなオブジェクトの検出精度を向上させることを目指す。

2. 小さいオブジェクト

実験用のデータベースは Manga109 を用いる。Manga109 についているアノテーションは図 1 に示すように、face, text, body, frame 四つの種類がある。MSCOCO データセットに倣い、オブジェクトのうち 1 ページの面積 (ピクセル数) の 1% 以下の大きさのものを小さいオブジェクトと定義する。表 1 に示すように、face と text は小さいオブジェクトの占める割合が大きい。

3. 従来手法

深層学習に基づいて漫画のオブジェクト検出に取り組んだ研究例としては、SSD[1]と呼ばれるモデルを基にした SSD-fork[2]という手法がある。漫画ではクラス間でオブジェクトのオーバーラップが多いため、アンカーボックスをクラスごとに用意することで対処している。しかし[2]では小さいオブジェクトの扱いは考慮されていない。

4. 小さなオブジェクト検出

小さなオブジェクトの検出精度向上のために以下の 3 つの方法を取り入れた。

①Oversampling[3]: 学習中、小さいオブジェクトを含む画像をオーバーサンプリングすることで、学習サンプルを増加させる。

②ネットワークモデルの調整: セマンティック情報をより考慮するため、Deconvolution 層を含む DSSD[4]モデルを導入する。VGG をより深い層を持つ ResNet-101 に置き換える。

③ネットワーク入力サイズの増加: SSD の入力サイズを 300 から 512 に増やす。DSSD の入力サイズを 321 から 513 に増やす。

5. 実験

Manga109 でアノテーションが付いている画像 10,130 枚を用いた。各冊で 6:4 の割合で学習、テストに分けた。SSD の実装と同じ設定で学習した。オーバーサンプリングで、トレーニング画像を 3 倍に増加させた。

表 2 に示すように DSSD, Oversampling の導入により small オブジェクト、特に face と text の精度が向上したことがわかる。また、入力サイズが大きくなると、精度が上がるのが検証された。

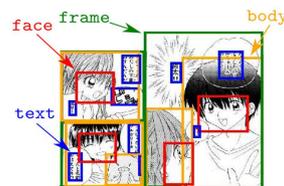


図 1 アノテーションの一例¹

表 1 小さいオブジェクトの割合

	small
frame	0.58%
body	37.57%
face	85.82%
text	90.07%

表 2 検出精度の比較

Method	mAP	AP				
		frame	text	face	body	small
SSD300 (without Oversampling)	75.5	89.1	78.5	61.2	73.1	27.0%
SSD300 (with Oversampling)	79.1	91.5	80.2	69.6	75.2	31.5%
SSD512 (with Oversampling)	79.7	93.1	80.7	71.0	74.0	32.1%
DSSD321 (without Oversampling)	79.5	90.3	81.3	72.0	74.2	32.7%
DSSD321 (with Oversampling)	82.0	93.5	82.2	74.5	77.8	36.2%
DSSD513 (with Oversampling)	81.9	92.4	82.5	74.7	75.1	37.4%

6. まとめ

漫画の face や text の部分は小さなオブジェクトの割合が大きく、それらの検出精度を向上させることができる。学習データとモデルの改善により、検出精度が向上した。

7. 今後の課題

データ増量以外のデータ処理方法と学習方法を改善する予定である。検出ネットワークをさらに改良する予定である。

参考文献

- [1] Liu, Wei et al. (2016). SSD: Single Shot MultiBox Detector. In ECCV.
- [2] Ogawa, Toru et al. (2018). Object Detection for Comics using Manga109 Annotations. ArXiv, abs/1803.08670.
- [3] Kisantant, Mate et al. (2019). Augmentation for small object detection. ArXiv, abs/1902.07296.
- [4] Fu, Cheng-Yang et al. (2017). DSSD : Deconvolutional Single Shot Detector. ArXiv, abs/1701.06659.
- [5] Aizawa, Kiyoharu et al. (2020). Building a Manga Dataset "Manga109" with Annotations for Multimedia Applications.
- [6] Matsui, Yusuke et al. (2017). Sketch-based Manga Retrieval using Manga109 Dataset.

¹ "BakuretsuKungFuGirl" © Ueda Miki