

2020 資格スクエア
日本ディープラーニング協会 G 検定対策ビデオ
ディープラーニング 5-09
畳み込みニューラルネットワーク(3)

浅川伸一

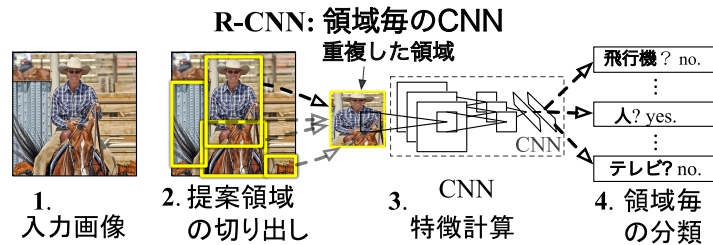
2020

物体位置決め課題 と物体検出課題

入力画像中の

- 「どこ」に 場所位置決め課題 入力画像から対象が写っている矩形領域（四角形）を切り出す バウンディングボックス
- 「何」があるか 物体検出課題 その物体

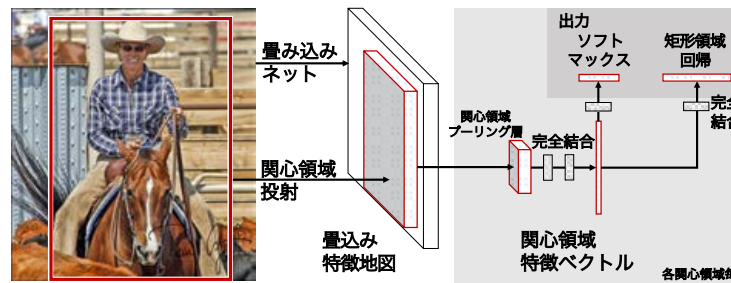
R-CNN



出典: Girshick et al. (2014)

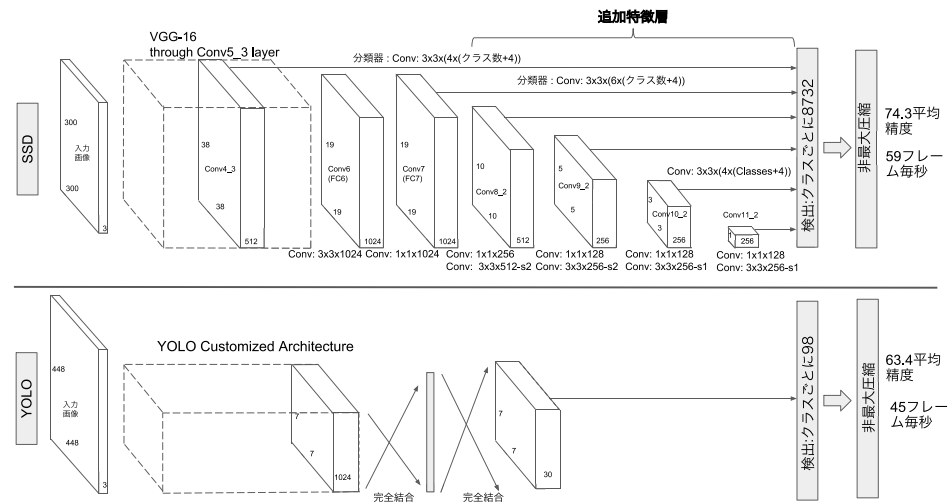
関心領域:ROI (Region of Interest) 切り出しには, CNN ではない従来手法を用いられました。ROI の画像切り出しの後に各領域毎に個別に CNN を呼び出す 2 段階のモデルであるため時間がかかっていた Girshick et al. (2014)

領域の切り出しとは画像上の矩形領域(長方形)の左上の座標と右下の座標で表現することにすれば, 4つの点を予測する 回帰問題 とみなすことができる。矩形領域のことを バウンディングボックス (bounding box) と呼ぶ



出典: Girshick (2015)

YOLO と SSD

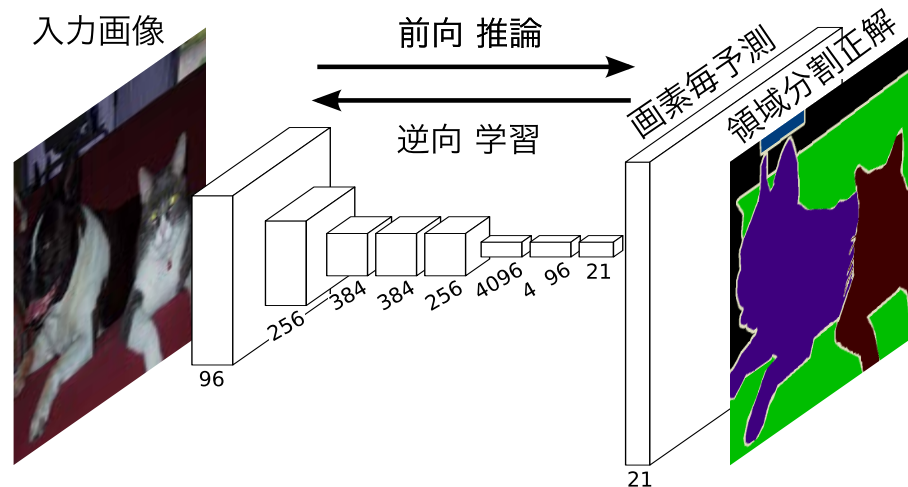


出典: Liu et al. (2016)

領域切り出しと検出を同時に行う手法

完全畳込みネットワーク FCN

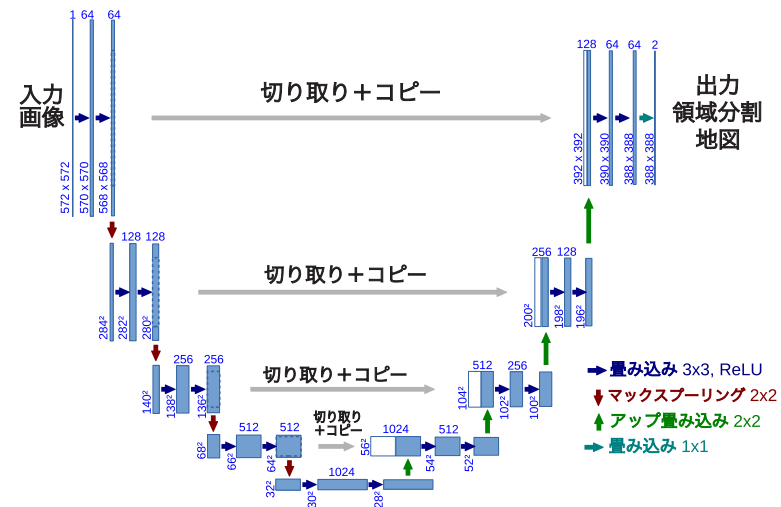
- 通常の CNN では、畳み込み演算、プーリング、を繰り返す。上位層では解像度が低下
- このときプーリング層を介在させないネットワーク 完全畳み込みネットワーク (fully convolutional networks: FCN)



出典: Long, Shelhamer, and Darrell (2015)

U-Net

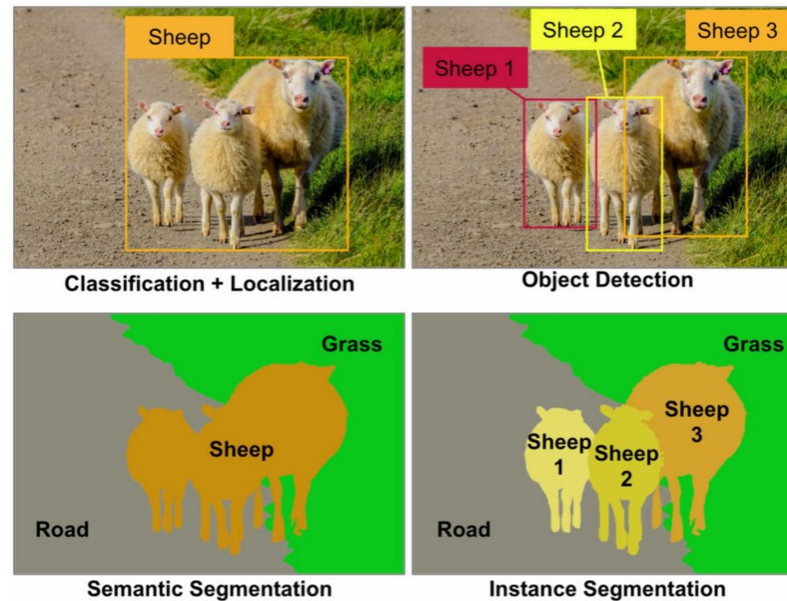
- 領域分割を行う際に、同じ解像度で対応させたスキップ結合を持たる。
- U 字型 Ronneberger, Fischer, and Brox (2015)
- X 線画像 (CT, MRI など) の医療画像診断に用いられている



出典: Ronneberger, Fischer, and Brox (2015)

領域分割課題

- セマンティック
- オブジェクト
- インスタンス
- パノプティック（可算名詞, 不可算名詞両方で表現される資格対象の切り分け）



出典: <https://towardsdatascience.com/detection-and-segmentation-through-convnets-47aa42de27ea>

実習

- 実習 [Open in Colab](#)



左上:元画像, 右上:インスタンスセグメンテーション



左下:パノプティックセグメンテーション, 右下:姿勢キーポイント検出。Wu et al. (2019) を使用

まとめ

- 画像の切り出し: R-CNN, fast-RCNN
- バウンディングボックスの回帰と物体検出とを行う
- 一段階モデル YOLO, SSD と 2 段階モデル
- 完全畳込みネットワークによるセマンティックセグメンテーション（意味的切り分け 20 種類 + 背景）
- U-net
- セマンティックセグメンテーション, インスタンスセグメンテーション, パノプティック（汎視覚的）セグメンテーション

クイズ

領域分割や画像切り分け(セグメンテーション)では、畳み込み演算とプーリングを繰り返すことで上位層において画像の解像度が低下します。解像度の低下は画像切り分けにとって不利になるので、さまざまな工夫がありました。このうち、低次層の解像度を援用することで解像度を保つネットワークで、アルファベット一文字で表されるネットワークは何だったのでしょうか？

クイズの答え

領域分割や画像切り分け(セグメンテーション)では、畳み込み演算とプーリングを繰り返すことで上位層において画像の解像度が低下します。解像度の低下は画像切り分けにとって不利になるので、さまざまな工夫がありました。このうち、低次層の解像度を援用することで解像度を保つネットワークで、アルファベット一文字で表されるネットワークは何だったでしょうか？

U net

文献

Girshick, Ross. 2015. “Fast R-CNN.” *ArXiv:1504.08083*.

Girshick, Ross, Jeff Donahue, Trevor Darrell, and Jitendra Malik. 2014. “Rich Feature Hierarchies for Accurate Object Detection and Semantic Segmentation.” In *Proceedings of Computer Vision and Pattern Recognition Conference (CVPR)*. Columbus, Ohio, USA.

Liu, Wei, Dragomir Anguelov, Dumitru Erhan, Christian Szegedy, Scott Reed, Cheng-Yang Fu, and Alexander C. Berg. 2016. “SSD: Single Shot Multibox Detector.” *arXiv Preprint arXiv:1512.02325*. <https://github.com/weiliu89/caffe/tree/ssd>.

Long, Jonathan, Evan Shelhamer, and Trevor Darrell. 2015. “Fully Convolutional Networks for Semantic Segmentation.” In *Proceedings of Computer Vision and Pattern Recognition Conference (CVPR)*. Boston, MA, USA. <http://arxiv.org/abs/1411.4038>.

Ronneberger, Olaf, Philipp Fischer, and Thomas Brox. 2015. “U-Net: Convolutional Networks for Biomedical Image Segmentation.” *ArXiv Preprint*, arXiv:1505.04597v1 [cs.CV].

Wu, Yuxin, Alexander Kirillov, Francisco Massa, Wan-Yen Lo, and Ross Girshick. 2019. “Detectron2.” <https://github.com/facebookresearch/detectron2>.