

進捗報告

1 今週やったこと

先週に引き続き Semantic Segmentation 関連の論文を読んでまとめた。はやく実装のほうに移りたい...

2 SegNet [1]

FCN で最後の全結合層を Convolution 層にしていたところを省略して省メモリおよび高速化を目指したモデル。結局最後のほうの特徴マップの情報って粗すぎて Semantic segmentation であんまり有効に使われてないよねというモチベーション。プーリング層によって情報が欠落する問題には、Encoder のプーリングでどの index のものをとってきたかを記憶しておき、Decoder でその情報を復元することによって対策している。前処理としては局所コントラスト正規化 (LCN) を行っている。Encoder と Decoder のペアを 1 つずつ Tuning する Layer-Wise Pretraining の手法を用いている。使用しているデータセットは CamVid [2, 3], KITTI [4], NYU [5]。

3 U-Net [6]

医療用のセグメンテーションのために生み出されたモデル。FCN を拡張して作成しており、Encoder と Decoder のモデル構造が U の形をしているためこの名前になっている。U-Net においても skip 構造を採用しており、U-Net ではチャンネルに対する連結をしている。細胞をアノテーションしたデータセット [7] を用いている。クラスとしては細胞と背景の 2 クラスであり、指標として IoU やハミング距離を用いている。データオーギュメンテーションはシフトや回転の他に弾性変形を用いている。

4 来週以降のタスク

引き続き論文を読み進める。あと DeepLab くらいは読みたい。ある程度読めたら github にいろいろなフレームワークでの実装が紹介されているのでひとまず動かせるかどうか試してみる。

参考文献

- [1] Vijay Badrinarayanan, Ankur Handa, and Roberto Cipolla. Segnet: A deep convolutional encoder-decoder architecture for robust semantic pixel-wise labelling. *arXiv preprint arXiv:1505.07293*, 2015.
- [2] Gabriel J. Brostow, Jamie Shotton, Julien Fauqueur, and Roberto Cipolla. Segmentation and recognition using structure from motion point clouds. In *ECCV (1)*, pages 44–57, 2008.
- [3] Gabriel J. Brostow, Julien Fauqueur, and Roberto Cipolla. Semantic object classes in video: A high-definition ground truth database. *Pattern Recognition Letters*, xx(x):xx–xx, 2008.
- [4] Andreas Geiger, Philip Lenz, Christoph Stiller, and Raquel Urtasun. Vision meets robotics: The kitti dataset. *International Journal of Robotics Research (IJRR)*, 2013.
- [5] Nathan Silberman and Rob Fergus. Indoor scene segmentation using a structured light sensor. *2011 IEEE International Conference on Computer Vision Workshops (ICCV Workshops)*, pages 601–608, 2011.
- [6] Olaf Ronneberger, Philipp Fischer, and Thomas Brox. U-net: Convolutional networks for biomedical image segmentation. *CoRR*, abs/1505.04597, 2015.
- [7] Martin Ma?ka, Vladim?r Ulman, David Svoboda, Pavel Matula, Petr Matula, Cristina Ederra, Ainhua Urbiola, Tom?s Espa?a, Subramanian Venkatesan, Deepak M.W. Balak, Pavel Karas, Tereza Bolckov?, Mark?ta ?treitov?, Craig Carthel, Stefano Coraluppi, Nathalie Harder, Karl Rohr, Klas E. G. Magnusson, Joakim Jald?n, Helen M. Blau, Oleh Dzyubachyk, Pavel K???ek, Guy M. Hagen, David Pastor-Escuredo, Daniel Jimenez-Carretero, Maria J. Ledesma-Carbayo, Arrate Mu?oz-Barrutia, Erik Meijering, Michal

Kozubek, and Carlos Ortiz-de Solorzano. A benchmark for comparison of cell tracking algorithms. *Bioinformatics*, 30(11):1609–1617, 02 2014.

- [8] Liang-Chieh Chen, Yukun Zhu, George Papandreou, Florian Schroff, and Hartwig Adam. Encoder-decoder with atrous separable convolution for semantic image segmentation. *CoRR*, abs/1802.02611, 2018.