
進捗報告

1 パーツ抽出のためのモデル作成に向けて

ひとまず 1 番最初に思いつきそうなモデルを作成. 図 1 にモデル構造を示す. 今まで使ってきた CAE モデルを 2 つ分岐させるようなモデル構造を定義した. 入力画像としてオリジナル画像を入れ, ターゲットとして目抜き画像と目のみ画像を設定し, それぞれに関して binary cross entropy を最小化する学習を行う. 問題点としては, MaxPooling 層を挟むことによって復元される画像がぼやけてしまうことが挙げられる. そこで MaxPooling 層を除いて学習させたが, 出力の 1 つがオリジナル画像とほぼ同一になる学習をしてしまうことがわかった. 図 2 に入力画像と出力画像を示す.

2 Semantic Segmentation

やりたいことは Semantic Segmentation とほとんど同じ (ピクセルが目に属するものとそれ以外を取ってくればよい). 4 コマ漫画ストーリーデータセットを使えば Semantic Segmentation を行うためのアノテーションとしては十分だと思われる. Semantic Segmentation のためのモデルには Fully Convolutional Networks for Semantic Segmentation [1], SegNet [2] や U-Net [3] があるらしい. 最新の手法は Google の DeepLab-v3 [4].

3 Data Augmentation

現在, 画像オーギュメンテーションの手法については 3 種類の方法を考えている.

1. 回転やスライド等のオーギュメンテーション (keras の ImageDataGenerator を使えばすぐできる)
2. レイヤ抜きによるオーギュメンテーション
3. 上記の両方を用いる

4 来週以降のタスク

1. FCN, SegNet, U-Net の論文を読む
2. 教師ラベル作成のためのジェネレータ定義 (けっこう大変そう)

参考文献

- [1] Jonathan Long, Evan Shelhamer, and Trevor Darrell. Fully convolutional networks for semantic segmentation. *CoRR*, abs/1411.4038, 2014.
- [2] 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.
- [3] Olaf Ronneberger, Philipp Fischer, and Thomas Brox. U-net: Convolutional networks for biomedical image segmentation. *CoRR*, abs/1505.04597, 2015.
- [4] 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.

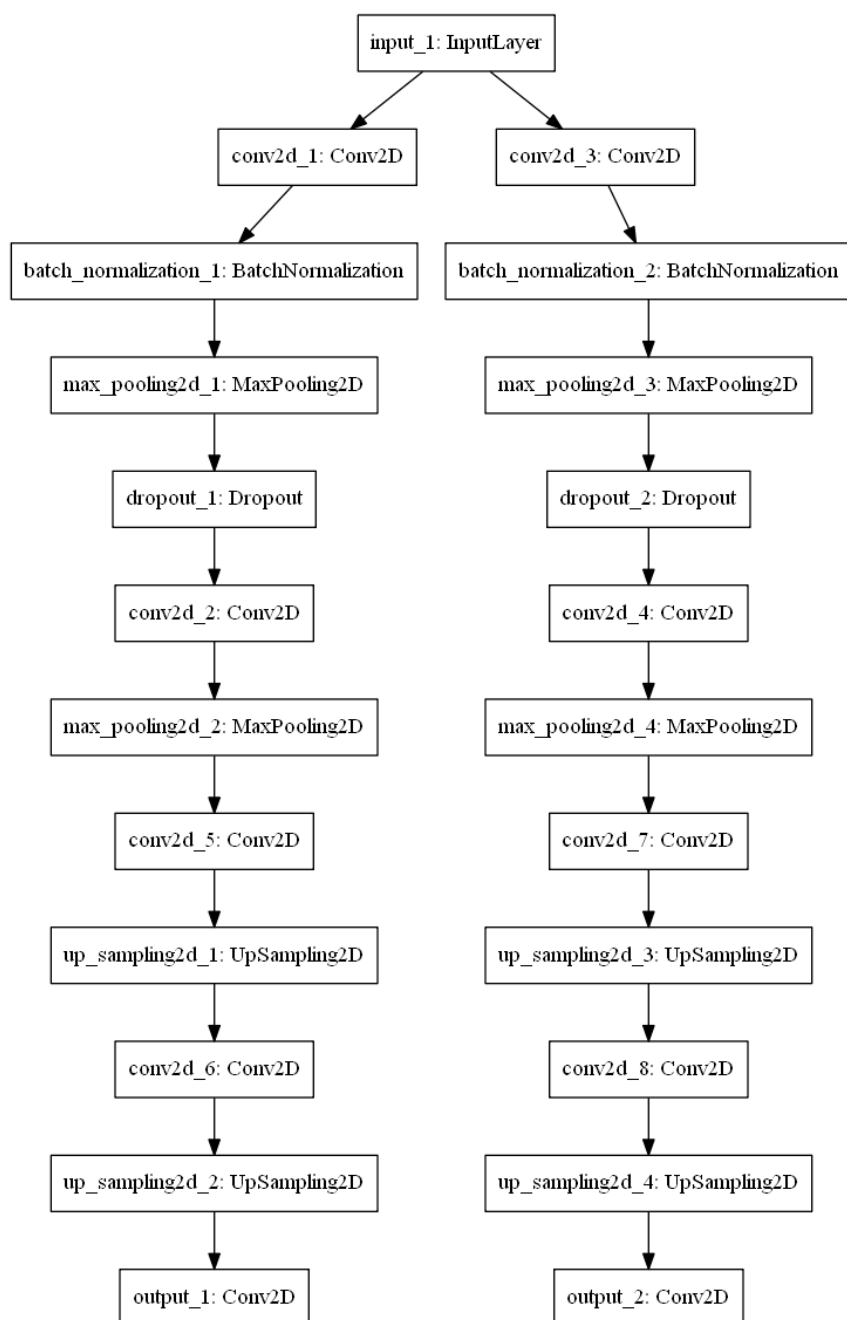


図 1: パーツ抽出のためのモデル定義



(a) オリジナル画像



(b) 出力 1(ターゲット:目のみ画像)



(c) 出力 2(ターゲット:目抜き画像)

図 2: オリジナル画像とその出力