

Машинное обучение (Machine Learning)

Сегментация изображений

Уткин Л.В.

Санкт-Петербургский политехнический университет Петра Великого



Содержание

- 1 Patch-by-patch scanning
- 2 Полносверточная сеть (full convolutional neural networks - FCNN)
- 3 Deconvnet
- 4 SegNet

Patch-by-patch scanning

Основная идея

- Метка класса определяется для каждого пиксела: задача рассматривается как задача “попиксельной” классификации
- Окно (patch) с центром в каждом пикселе подается на вход сверточной сети для генерации его класса
- Точность сегментации повышается с увеличением размера окна, так как оно охватывает больше контекстной информации
- Главный недостаток - большое перекрытие соседних окон и большой объем избыточных вычислений.

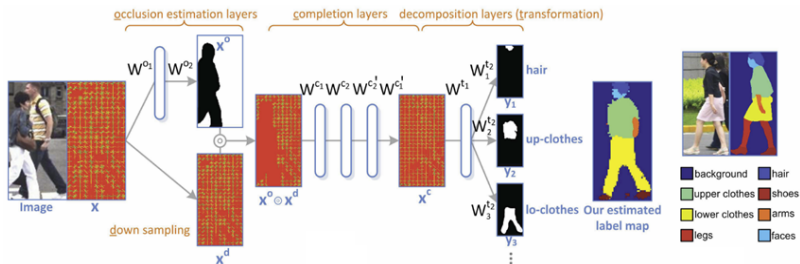
Полносвязанные слои (fully connected layers)

- Без **patch-by-patch scanning**
- Вся картинка - входное данные для реализации прямого и обратного распространения с многоуровневой полносвязанной сетью
- Каждый нейрон связан с каждым нейроном на предыдущем слое.

Полносвязанные слои (fully connected layers)

- Каждый выходной нейрон (прогнозирование класса) зависит только от подмножества входных пикселей.
- В CNN сеть “решает”, является ли изображение элементов класса 1 в зависимости от первых нескольких патчей, затем класса 2 на основе нескольких следующих патчей и т.д.
- Нельзя классифицировать, например, кота, используя только первые патчи и игнорируя остальные.
- Однако, если ввести полносвязанный слой, то сеть становится способной смешивать, так как каждый нейрон связан с каждым нейроном на следующем слое, т.е. есть связь между каждым пикселем и каждым выходным классом так, что решение основано на всем изображении.

Пример полносвязанных слоев



Сеть сегментирует картинку на фон и различные части тела
 (X. Wang. Deep Learning in Object Recognition, Detection, and Segmentation,
 Foundations and Trends in Signal Processing, 8(4), 2014, pp. 217–382)

Полносверточная сеть

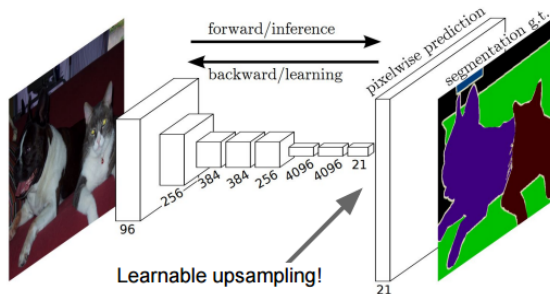
Полносверточная сеть

Полносверточная сеть была предложена для устранения patch-by-patch scanning и для повышения эффективности FCNN заменяет полносвязанные слои в сверточной сети на 1×1 сверточные ядра.

FCNN в качестве входа использует всю картинку и получает сегментационную карту на выходе одним проходом прямого распространения

FCNN для сегментации

- FCNN обычная CNN, но последний полносвязанный слой заменяется другим сверточным слоем
- J. Long, E. Shelhamer, T. Darrell. Fully Convolutional Networks for Semantic Segmentation. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 3431-3440, 2015.



FCNN

- Каждый слой данных - трехмерный массив размера $h \times w \times d$, где h и w - пространственные координаты, d - цвета
- Первый слой - изображение с пикселями размера $h \times w$, и d цветами
- Положение на слоях более высокого уровня соответствуют положениям на изображении, с которыми они связаны путями, называемыми *рецептивными полями*

FCNN

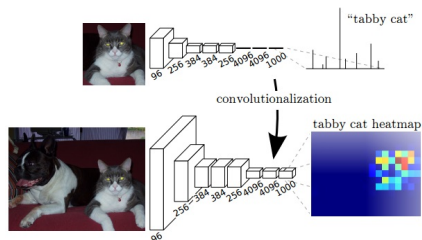
Как конвертировать классификационную сеть в FCNN, которая получает на выходе “грубое” представление?

Для попиксельной классификации необходимо соединить эти грубые выходы обратно с пикселями.

Последний сверточный слой выбирается достаточно большим.

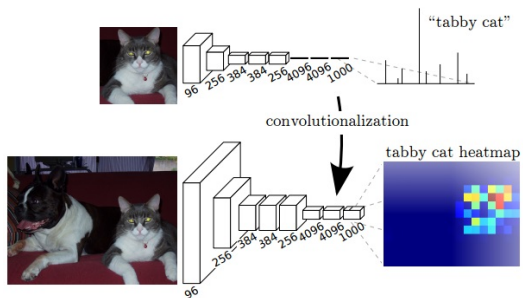
CNN

- Для обычной CNN, которая классифицирует имеются несколько полносвязанных уровней на конце сети
- Выход полносвязанных уровней обрабатывается soft-max функции так, что несколько вероятностей генерируются
- Входное изображение классифицируется в класс, который имеет наибольшую вероятность

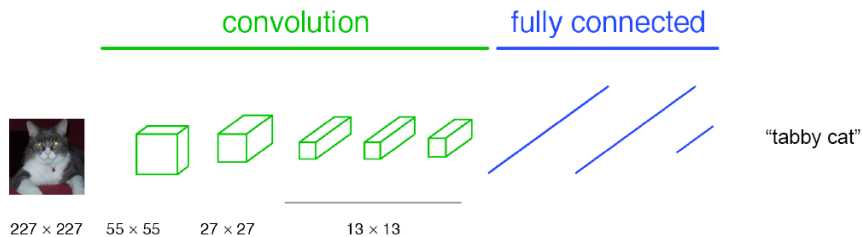


FCNN

- Для FCNN полносвязанные слои заменяются сверточными слоями



Классификационная сеть



Переход к полносверточной

convolution



227×227



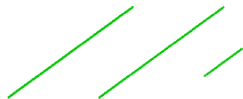
55×55



27×27



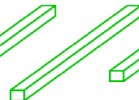
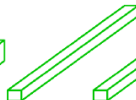
13×13



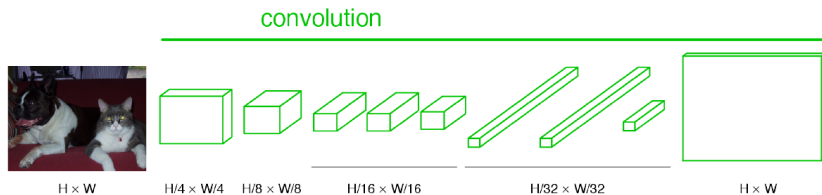
1×1

Или

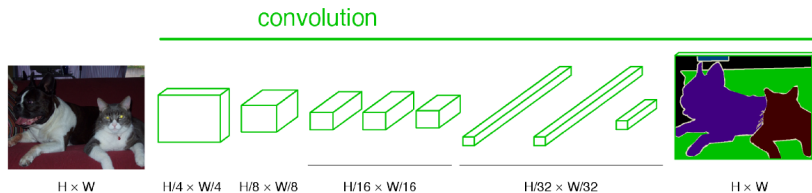
convolution

 $H \times W$  $H/4 \times W/4$  $H/8 \times W/8$  $H/16 \times W/16$  $H/32 \times W/32$

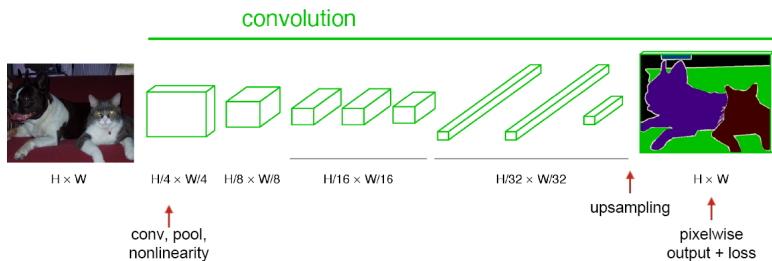
Upsampling



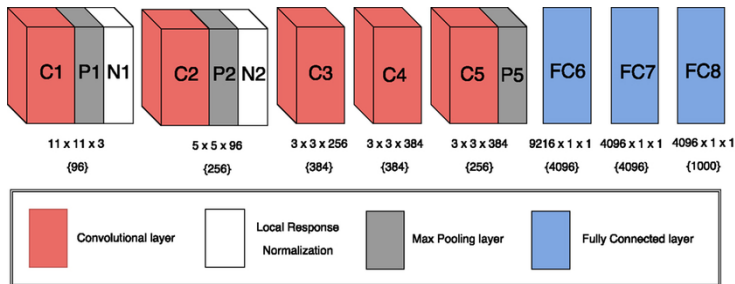
Пиксел-to-пиксел сеть



Пиксел-to-пиксел сеть



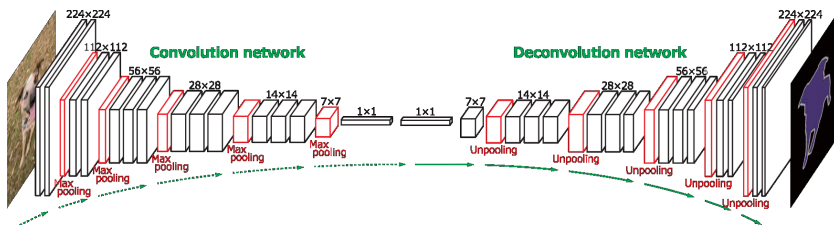
Переход от стандартной к полносверточной FCNN (Alexnet to a FCNN)



Deconvnet

Deconvolution network (Deconvnet)

H. Noh S. Hong B. Han, Learning Deconvolution Network for Semantic Segmentation



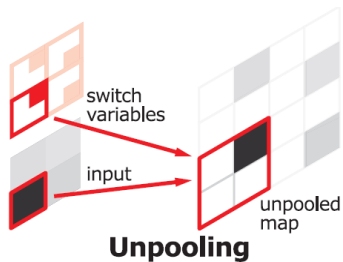
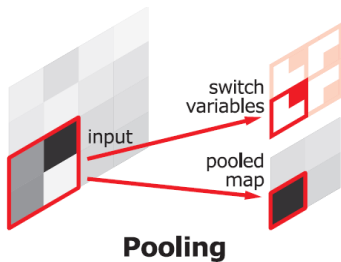
Deconvolution network

Реализует две основные процедуры

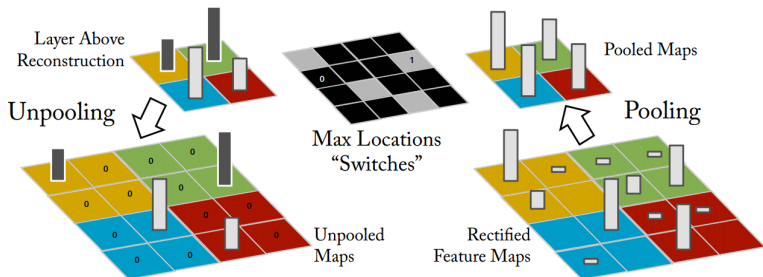
- 1 Unpooling
- 2 Deconvolution

Unpooling (upsampling)

Обратная операция pooling



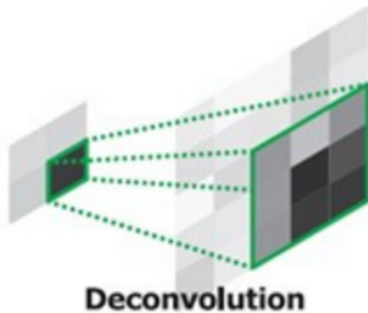
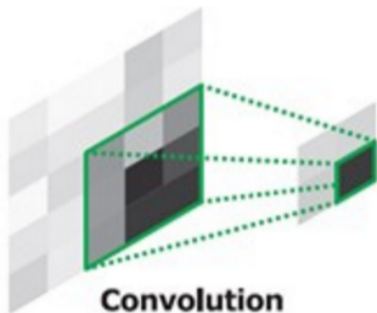
Unpooling (upsampling)



Deconvolution

- Deconvolution слой layers уплотняет разреженный слой, полученный в unpooling, с использованием операций свертки
- В отличие от сверточных слоев, которые соединяют множественные входные значения активации нейронов внутри окна фильтра в одно значение, обратная операция свртки ассоциирует одиночное входное значение активации с множественным выходом

Deconvolution (схема)



Deconvolution (операция)

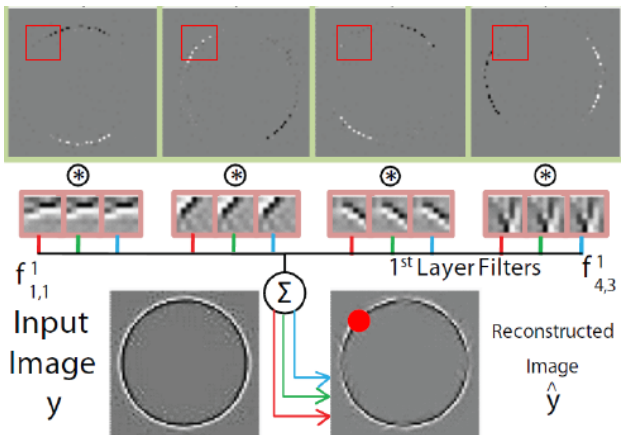
Задача оптимизации

$$C = \frac{\lambda}{2} \|Fz - y\|_2^2 + |z|_1$$

Известно: y - вход, F - веса фильтра

Найти: z - карта признаков

Deconvolution (пример)



Deconvolution (более сложный пример)

2nd layer pooled maps

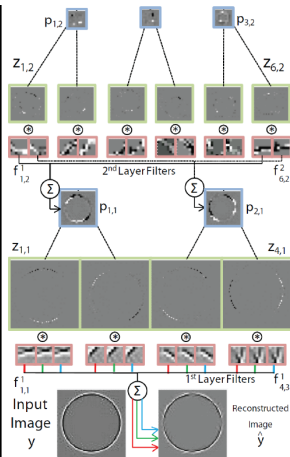
2nd layer feature maps

2nd layer filters

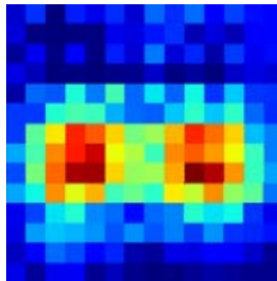
1st layer pooled maps

1st layer feature maps

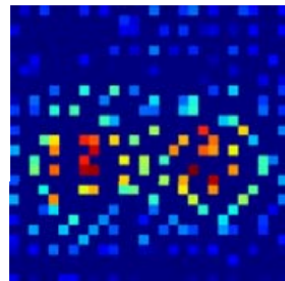
1st layer filters



Deconvolution and unpooling



Deconv: 14x14

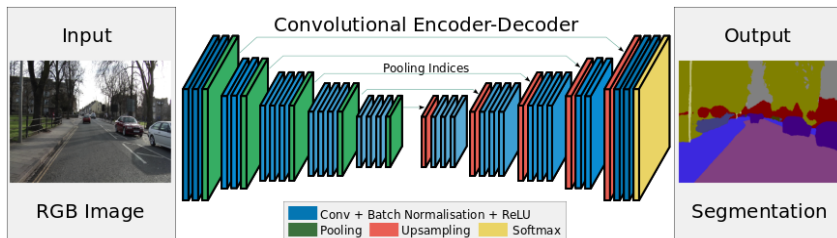


Unpool: 28x28

SegNet

SegNet

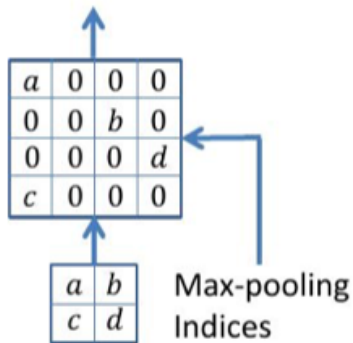
V. Badrinarayanan, A. Kendall and R. Cipolla "SegNet: A Deep Convolutional Encoder-Decoder Architecture for Image Segmentation." arXiv preprint arXiv:1511.00561, 2015



Особенность SegNet

- Сеть состоит из :
 - последовательности слоев нелинейного преобразования (кодер) и
 - соответствующего множества декодеров, за которыми следует “попиксельный” классификатор
- **Ключевой элемент SegNet** - декодер использует индексы, полученные на этапах max-pooling в кодере, для реализации upsampling
- **Индексы** - местоположения признаков с максимальными значениями в каждом окне pooling (запоминаются для каждой карты признаков в кодере)
- Это позволяет избежать обучения процедуры upsampling.
- Сеть в целом обучается, используя градиентный спуск

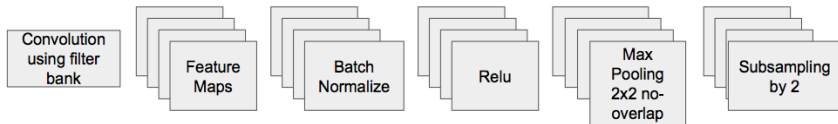
Индексы max-pooling в SegNet



SegNet

Кодеры SegNet

13 сверточных слоев:



Декодеры SegNet

- Для каждого из 13 кодеров имеется соответствующий декодер, который выполняет upsampling, используя индексы max-pooling
- Разреженные карты признаков более высокого разрешения образуются
- Последний декодер соединен с softmax-классификатором, который классифицирует каждый пиксел

Отличие SegNet от FCNN

- SegNet использует индексы max-pooling для upsampling (без обучения) карт признаков
- FCNN выполняет upsampling, обучая для “обратной” свертки входные карты признаков, и добавляет карту признаков кодера для реализации выхода декодера

Полезные ссылки (описание инструментов сегментации)

- https://leonardoaraujasantos.gitbooks.io/artificial-intelligence/content/image_segmentation.html
- <https://wiki.tum.de/display/Ifdv/Image+Semantic+Segmentation>

Полезные ссылки (программное обеспечение)

- [http://warmspringwinds.github.io/tensorflow/tf-slim/2017/01/23/fully-convolutional-networks-\(fcns\)-for-image-segmentation/](http://warmspringwinds.github.io/tensorflow/tf-slim/2017/01/23/fully-convolutional-networks-(fcns)-for-image-segmentation/)
Fully Convolutional Networks (FCNs) for Image Segmentation trained on PASCAL VOC
- <http://warmspringwinds.github.io/tensorflow/tf-slim/2016/12/18/image-segmentation-with-tensorflow-using-cnns-and-conditional-random-fields/>
Image Segmentation with Tensorflow using CNNs and Conditional Random Fields
- <http://warmspringwinds.github.io/tensorflow/tf-slim/2016/10/30/image-classification-and-segmentation-using-tensorflow-and-tf-slim/>
Image Classification and Segmentation with Tensorflow and TF-Slim

Полезные ссылки (программное обеспечение)

- <https://github.com/MarvinTeichmann/tensorflow-fcn>
An Implementation of Fully Convolutional Networks in Tensorflow
- <https://github.com/alexgkendall/caffe-segnet>
Implementation of SegNet: A Deep Convolutional Encoder-Decoder
Architecture for Semantic Pixel-Wise Labelling
- <http://mi.eng.cam.ac.uk/projects/segnet/>
SegNet

Вопросы

?