

1. Что такое пространственное объединение в пирамиду в СНС?

Один из возможных методов решения проблемы объекта в разных масштабах: использование пространственного объединения в пирамиду, входящий сигнал преобразуется фильтрами с разными степенями расширения и тем самым разными полями зрения, таким образом, усваивая изображения различного масштаба. Обычно используются четыре параллельных расширенных с.с. с различными степенями расширения, затем все выходы объединяются в один. Но с увеличением степени расширения, количество валидных весов фильтра уменьшается, т.е. фильтр может превратиться в 1×1 , для решения этой проблемы выделяются дополнительные признаки на уровне изображения. Сначала применяется глобальный усредненный пулинг к последней карте признаков, затем 1×1 свертка с 256 фильтрами и билинейная интерполяция до требуемого пространственного разрешения.

2. Как обычная свертка раскладывается на поканальную свертку?

Обычная свертка раскладывается на поканальную свертку в 2 этапа:

1. Входящий тензор сворачивается ядром $1 \times 1 \times n_c$, где n_c – количество каналов входящего тензора.
2. В получившемся результате сворачивается каждый канал по отдельности сверткой $3 \times 3 \times 1$.

При поканальной раздельной свертки имеется в виду, что сначала делается свертка 3×3 по каждому каналу, а затем 1×1 .

3. Приведите ситуацию, когда в сети необходимо использовать несколько выходных слоев.

Когда требуются предсказания двух или более числовых значений, например, прогнозирование координат x и y . Или мы хотим предсказать возраст, пол, расу человека на изображении. Или у нас есть 3 класса с 3 фигурами разных цветов. Каждый выходной нейрон отделяет "свой класс" от

остальных двух. Например, первый нейрон распознаёт объекты, помеченные синими кружочками, второй нейрон распознаёт красные крестики, а третий - зелёные квадратики.

4. В отчете сказано, что увеличение ядра на первом слое не позволяет сети выделить нужные признаки. На основании чего был сделан данный вывод?

Потому что из-за большого ядра свёртки идёт слишком сильное обобщение пикселей и выходной шейп размывается. Чем больше фильтр, тем больше пикселей вокруг оно задействует и тем сильнее размытие.

5. Распишите, что значат параметры в данной строке "`conv_2 = Convolution2D(32, 3, 3, border_mode='same', activation='relu')(conv_1)`"

- 32, 3, 3 – input_shape (кортеж целых чисел, не включая ось выборки), input_shape = (n_images , x_shape, y_shape).
- border_mode='same' – возвращает то же измерение, что и входное изображение.
- activation='relu' – название функции активации, в данном случае 'relu'.

6. Для чего нужен слой Flatten в Вашей модели?

Мы создаем модель классификации. Обработанные данные должны быть в виде одномерного линейного вектора. Прямоугольные или кубические формы не могут быть прямыми входами. Слой Flatten преобразовывает данные в одномерный массив для ввода их в следующий слой. Мы сглаживаем вывод сверточных слоев, чтобы создать один длинный вектор объектов. И это связано с окончательной классификационной моделью. Другими словами, мы помещаем все данные пикселей в одну строку и устанавливаем связи с последним слоем.