Добрый день! Меня зовут Григорьев Дмитрий. Как вы помните на прошлом докладе был поднят ряд вопросов, по которым я сделал более детальный рассказ. Напомню об этих: это нулевой паддинг, в чем могут заключаться проблемы при работе с ним; отличие dilated сверток от сверток с шагом (stride) и свертки типа 1 x 1.

**Padding.**

Начнем с более детального представления о паддинге.

Напомню что это: Крайние пиксели никогда не оказываются в центре ядра, потому что тогда ядру не над чем будет скользить за краем. Это совсем не идеальный вариант, так как мы хотим, чтобы размер на выходе равнялся входному.

Padding добавляет к краям поддельные пиксели (обычно нулевого значения, вследствие этого к ним применяется термин “нулевое дополнение” — “zero padding”), чтобы настоящие пиксели при скольжении оказывались в своем центре, а поддельные пиксели попадали в ядро, создавая выходную матрицу того же размера, что и входная. Так поддельные пиксели не вносят свой вклад в результат свертки. Также мы можем использовать зеркальный паддинг - когда дополнительные пиксели добавляются как зеркальные отражения по всем направлениям углового пикселя. Так мы можем акцентировать большее внимание на соседние пиксели, увеличив их вклад в результат свертки.

Паддинг нужен нам для того, чтобы иметь важность анализировать информацию с крайних пикселей. При этом паддинг поможет нам сохранить размерность исходной картинки, что при большом количестве сверточных слоев не даст нам слишком быстро

Следующий момент, о котором я хочу рассказать – это разница между dilated сверток и сверток с шагом. Разница между этими типами сверток неплохо раскрывается на примере действия последовательных сверток. Прежде всего, нужно рассказать о том, что такое рецептивное поле. Рецептивное поле - это область в сверточном слое, которая влияет на каждый элемент следующего слоя.

На картинке мы видим, что последовательно увеличивая параметр dilation, мы можем очень быстро увеличить рецептивное поле, при этом не добавляя каких-то новых параметров и сохраняя разрешение картинки входной области. Подобного эффекта можно достичь и с помощью stride, но сделать это сложнее, мы, во-первых, пропускаем информации за счет такого агрессивного перескакивания, а во-вторых, для сохранения размерности приходится использовать padding. Проще в таком случае использовать dilated convolution.

И наконец последний вопрос – 1 х 1 свертки. Суть их заключается в контроле над числом каналов в результате действия свертки.

Пусть у нас имеется изображение 6 х 6 пикселей с тремя каналами со сверткой размерности 3 х 3 х 3. После применения этой свертки получим результат размерности 4 х 4 х 1 (учитывая padding 1 и отсутствие stride). Видно, что в результате такой свертки число каналов сократилось до единицы.

Теперь возьмем свертку размерности 1 х 1 х 3, число каналов также осталось равным единице, но мы сохранили оставшиеся размерности.

Так, увеличивая число фильтров для свертки 1 х 1 х 3 мы можем контролировать число каналов. Например, для двух фильтров число каналов равно 2, для трех 3 и так далее.

Таким образом, мы используем такие свертки для того, чтобы контролировать число каналов на выходе. Использование сверток 1 х 1 позволяет нам экономить вычислительную мощность.

И так, в своем докладе я постарался более подробно раскрыть ряд вопросов о паддинге, вопрос о различии dilated и strided convolution, а также более подробно рассказать о 1x1 свертках. Спасибо за внимание, постараюсь ответить на ваши новые вопросы =).