**Основные сведения о сети VGG16**

VGG16 — модель сверточной нейронной сети, предложенная K. Simonyan и A. Zisserman из Оксфордского университета в статье “Very Deep Convolutional Networks for Large-Scale Image Recognition”. Она получила свою известность благодаря выдающимся результатам, которые она показала в соревновании ILSVRC 2014-го года.

VGG16 является улучшенной версией AlexNet, в которой заменены большие фильтры (размера 11 и 5 в первом и втором сверточном слое, соответственно) на несколько фильтров размера 3х3, следующих один за другим.

Сеть VGG16 обучалась на протяжении нескольких недель при использовании видеокарт NVIDIA TITAN BLACK.

**Архитектура**

На вход первого слоя подаются RGB изображения размера 224х224. Далее изображения проходят через стек сверточных слоев, в которых используются фильтры с очень маленьким окном размера 3х3 (который является наименьшим размером для получения представления о том,где находится право/лево, верх/низ, центр). Свертка происходит с шагом в 1 пиксель - без пробелов. Пространственное дополнение (padding) входа сверточного слоя выбирается таким образом, чтобы пространственное разрешение сохранялось после свертки, то есть дополнение равно 1 для 3х3 сверточных слоев. Пространственный пулинг осуществляется при помощи пяти max-pooling слоев, которые следуют за одним из сверточных слоев (не все сверточные слои имеют последующие max-pooling). Операция max-pooling выполняется на окне размера 2х2 пикселей с шагом 2.

После стека сверточных слоев идут три полносвязных слоя: первые два имеют по 4096 каналов, третий — 1000 каналов (число может варьироваться в зависимости от количества классов, которое необходимо распознать, один нейрон соответствует одному классу). Последним идет soft-max слой. Конфигурация полносвязных слоев одна и та же во всех нейросетях. Все скрытые слои снабжены функцией активации ReLU.

**Недостатки**

К сожалению, сеть VGG имеет два серьезных недостатка:

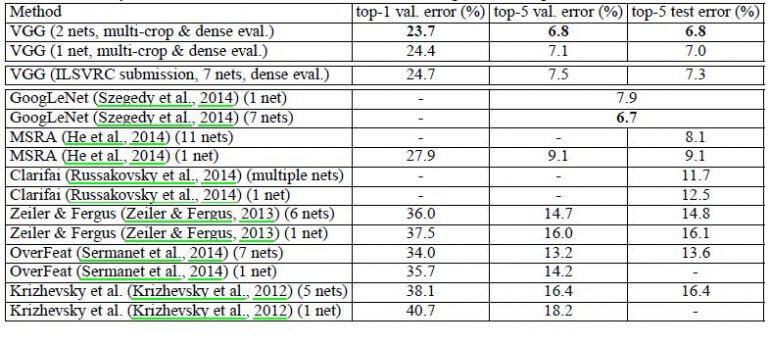
* Очень медленная скорость обучения.
* Сама архитектура сети весит слишком много (появляются проблемы с диском и пропускной способностью)

Из-за глубины и количества полносвязных узлов, VGG16 весит более 533 МБ. Это делает процесс развертывания VGG утомительной задачей.

Несмотря на недостатки, данная архитектура является отличным строительным блоком для обучения, так как её легко реализовать.

**Результаты (ILSVRC 2014)**

VGG16 существенно превосходит в производительности прошлые поколения моделей в соревнованиях ILSVRC-2012 and ILSVRC-2013. Достигнутый VGG16 результат сопоставим с победителем соревнования по классификации (GoogLeNet с ошибкой 6.7%) в 2014 году и значительно опережает результат Clarifai победителя ILSVRC-2013, который показал ошибку 11.2% с внешними тренировочными данными и 11.7% без них. Что касается одной сети, архитектура VGG16 достигает наилучшего результата (7.0% ошибки на тесте), опережаю одну сеть GoogLeNet на 0.9%:

****

Было показано, что глубина представления положительно влияет на точность классификации.