**Тема 2.5 Остаточные связи**

**План**

1. Идея остаточных связей
2. Архитектуры ResNet
3. Проблема излишне глубоких нейронных сетей

**Идея остаточных связей**

В 2015 году был предложен принцип идентичных связей для обучения нейронных сетей увеличенной глубины. По существу, идея таких связей представляет собой попытку создания некоторой отрицательной обратной связи или дифференцирующего звена в блоке сверточной сети. Идентичные связи должны быть параллельны основному слою нейронной сети. Слой (или блок слоев) нейронной сети с идентичными связми принято называть skip-connection, identity-connection, residual-connection или residual-layer, остаточный слой, тождественный слой.

**Архитектуры ResNet**

Благодаря использованию остаточной связи удалось увеличить глубину нейронных сетей до 152 слоев. Архитектуры сверточных нейронных сетей с остаточными связями, принято называть ResNet-18, ResNet-34, ResNet-50, ResNet-101, ResNet-152, где цифра обозначает число слоев. В соревнованиях ILSVRC 2015 модель ResNet-152 достигла погрешности 4,6 % для одной сети и 3,6 % для ансамбля сетей. При этом значения ошибки ниже, чем для эксперта (у человека ошибка для тех же условий порядка 5 %).

В основе работы остаточных связей лежат следующие идеи. Если во время тренировки нейронной сети в основном слое или блоке слоев возникает вымывание или взрыв градиента, то остаточная связь действует как регуляризация, компенсируя проблему. Компенсация происходит за счет суммирования информации (входных данных) перед блоком и результата работы самого блока. В случае, если блок не нужен нейронной сети (например, слишком глубокий слой), то по задумке авторов сеть должна обучиться так, чтобы влияние остаточной связи преобладало в блоке (говоря на языке электронной схемотехнике – шунтировало блок). Таким образом остаточная связь является приемом регуляризации обучения нейронных сетей. Благодаря остаточным связям стало возможным увеличивать глубину нейронных сетей практически до бесконечности.

**Проблема излишне глубоких нейронных сетей**

Наиболее глубокие нейронные сети с остаточными связями имеют более 1000 слоев. Однако, на практике оказалось, что при использовании остаточных связей слои выше некоторого порога почти полностью или полностью становятся идентичными. То есть слишком большое рецептивное поле тоже может оказаться не нужным сети для достижения поставленных целей.

**Резюме**

Идея остаточных связей стала революционной для всего глубокого обучения нейронных сетей. Это, пожалуй, наиболее популярный прием регуляризации. Архитектура ResNet остается наиболее популярной в свертоных нейронных сетях компьютерного зрения по настоящее время. Также эксперименты с архитектурой ResNet положили конец тенденции увеличения числа слоев нейронной сети. Все дальнейшие попытки оптимизации сверточных слоев сводились к попыткам повышения их эффективности при сравнительно небольшом общем числе слоев. Мы рассмотрим эти попытки в нескольких следующих лекциях.