Конспект лекции

А.Н. Шевляков

Введение в нейронные сети

Распознавание изображений, свёрточные нейронные сети и фильтры. Задача сегментации изображений. Сеть U-net

# Картинки как объектдля нейронных сетей

Как подать картинку на вход нейросети? Надо её представить в виде массива (матрицы) чисел!

Сначала рассмотрим **монохромные картинки.** В них каждый пиксел содержит оттенки серого. Логичнее каждый пиксел представить числом от 0 (белый) до 1 (черный). После этого мы картинку превращаем в массив чисел и подаём в качестве значений входного слоя НС.

**Цветную картинку** тоже легко превратить в массив чисел. Для этого нужно вспомнить, что такое формат RGB. Фактически каждый пиксел — это **тройка** целых чисел из интервала [0,255]. А сама картинка — это набор из трёх числовых таблиц (матриц). Все эти таблицы можно «вытянуть» в линию, получится один большой вектор. Его и подаём на вход в нейросеть. На самом деле **стандартная архитектура полносвязной нейросети** неудобна для распознавания картинок!

Многие связи между нейронами последней НС **избыточны**.

Это связано с тем, что практических задачах нейрону не нужно знать информацию обо **всей картинке**.

Что делать? Достаточно картинку разделить на области и каждому нейрону дать небольшую область для изучения.

# Свёрточная архитектура

По какому принципу работают клетки в зрительном центре биологического мозга? Ни один из нейронов нижних отделов зрительного центра не имеет информации обо всём изображении.

**Идея свёрточной архитектуры** идёт из анатомии.Каждый слой нейросети будет не одномерным (в виде столбика), а двумерным (в виде таблицы), ибо сетчатка и слои (!) зрительного центра двумерные.

Все слои свёрточной нейросети (СНС) представляются в виде матриц. Каждый нейрон последующего слоя соединён только с нейронами определённой области предыдущего слоя. Размер области определяется при создании СНС. Нейроны второго слоя могут иметь пересекающиеся области с предыдущего слоя.

Связи в нейрон следующего слоя ведут лишь из выделенной области предыдущего слоя.

Нейроны в свёрточном слое — это как слепые мудрецы, ощупывающие слона!

Проблема слепых мудрецов: нет «мудреца 2го уровня», который бы смог обработать информацию от каждого мудреца. А в нейросетях мы можем добавить ещё один слой, анализирующий информацию с предыдущего слоя

# Матричные фильтры

Матричные фильтры применялись и до эпохи НС в задачах обработки изображений:

* нахождение границ предметов на изображении;
* размытие изображений
* увеличение контрастности, и т.д.

**Как работает фильтр:**

1. **Фильтр** — это матрица с числами (это **маска фильтра**).
2. **Фильтр прикладывается** к каждой области изображения (изображение тоже представляется в виде матрицы чисел) и результат **операции свёртки** записывается в итоговое изображение. И операция свёртки полностью аналогична преобразованию данных в НС!!!

Какой размер после свёртки? Свёртка уменьшает размер изображения. Например, если было изображение 6х6 и маска фильтра 3х3, то результат будет размера 4х4!

А в общем виде сможете подсчитать?У вас есть изображение nxn, фильтр mxm. Тогда ответ будет размера **(n-m+1)x(n-m+1)**.

# Параметры свёртки

Фильтр прикладывается к областям изображения. Этот процесс **можно подкрутить** с помощью параметров, отвечающих за:

* **величину шага** (как сильно смещается фильтр по сравнению с предыдущей итерацией). По умолчанию шаг равен 1 (как по вертикали, так и по горизонтали);
* **добавление пустой рамки вокруг изображения** (это нужно для того, чтобы фильтр приложился к крайним точкам изображения).

**Величину шага задают два параметра: stride\_x, stride\_y** — насколько пикселей сдвигается по горизонтали и вертикали фильтр после очередной итерации.

По умолчанию stride\_x=stride\_y=1, то есть фильтр в ряду изображения каждый раз сдвигается на 1 пиксел, а при окончании ряда фильтр смещается на 1 пиксел вниз.