

САНКТ-ПЕТЕРБУРГСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ

КУРСОВАЯ РАБОТА НА ТЕМУ
“СВЕРТОЧНЫЕ НЕЙРОННЫЕ СЕТИ И ИХ ИСПОЛЬЗОВАНИЕ В АНАЛИЗЕ И РАСПОЗНАВАНИИ
ИЗОБРАЖЕНИЙ”

Студент 231 группы:
Кононов Николай Алексеевич

Научный руководитель:
к.ф.-м.н. Ананьевский М.С.

27 декабря 2019 г.

1 Введение

Нейронные сети(neural networks) – одно из направлений в области искусственного интеллекта, построенное на основе математической модели биологических нейронных сетей – сетей нервных клеток живого организма. Последовательность нейронов, соединенная **синапсами**.

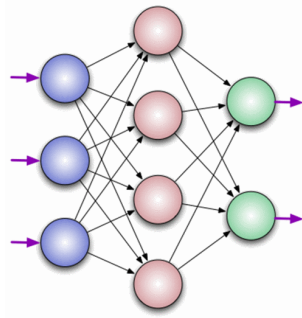


Рис. 1: Пример простой нейронной сети: синим цветом обозначен входной слой(input layer), красным – скрытый слой(hidden layer), зеленым – выходной слой(output layer)

Компьютерное зрение(computer vision) – одна из областей машинного обучения, занимающаяся созданием технологий, которые могут производить обнаружение, отслеживание и классификацию различных объектов на изображениях и видео.

Задачи компьютерного зрения можно условно поделить на 4 типа:

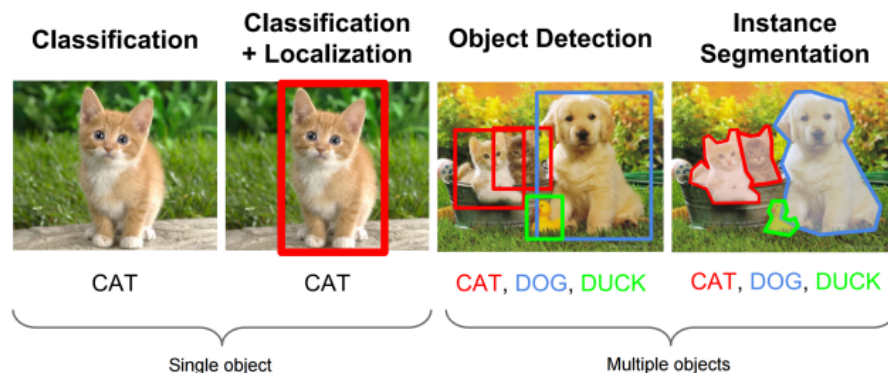


Рис. 2: Пример задач компьютерного зрения(computer vision).

Целью классификации(classification) является распознавание каждого объекта на изображении, целью локализации – выделение этого объекта. Задача по обнаружению объектов(object detection) подразумевает нахождение заданных объектов на фотографии и их выделение.

Сверточные нейронные сети(convolutional neural networks, CNN) – специальная архитектура нейронных сетей, нацеленная на эффективное распознавание и обработку изображений. Однонаправленная и многослойная сеть.

До того, как технологии нейронных сетей были применены к компьютерному зрению, задача о распознавании объектов казалась невероятно трудной. Технологии сверточных нейронных сетей(далее СНС) активно применяются во многих задачах обработки изображений. Например, в системах распознавания лиц, распознавании эмоций и т.д. В следующей части будет написана упрощенная модель СНС на примере распознавания лиц(face detection).

2 Работа СНС.

Идея СНС заключается в следующем: мы разбиваем изображение на пиксели, либо некоторые области(патчи) по $n \times n$ пикселей значение n зависит от входных изображений и от удобства для создателей системы. Далее сигналы этих входных слоёв передаются от слоя к слою с помощью синапсов. Каждый из слоев имеет свои собственные коэффициенты. Так мы двигаемся по слоям, пока мы не получим сообщение о том, что система распознала какой-то объект.

2.1 Принципы сверточных нейронных сетей

- **Локальное восприятие:** если в полносвязной сети каждый нейрон подключается к каждому пикселю входного изображения, то в сверточных сетях нейрон получает на вход только ограниченное количество пикселей. Обычно это участки изображения размером 3×3 , 5×5 , 7×7 пикселей.

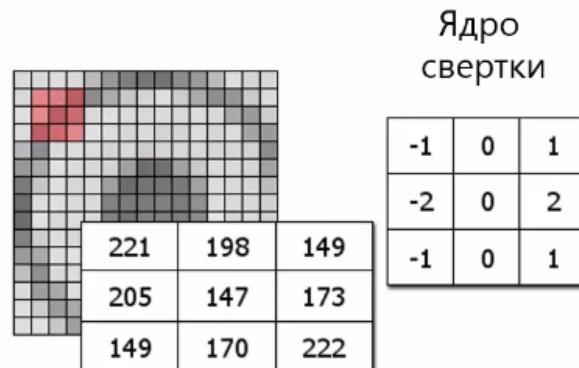
Следующий нейрон получает участок, который может пересекаться с участком соседнего нейрона. Операция, которая выполняется нейроном называется сверткой. Идея свертки заключается в следующем: мы рассматриваем некоторую область изображения(на Рис.3 эта область имеет размер 3×3) и по ядру свертки(некоторая матрица такого же размера, которая вычисляется в процессе обучения) вычисляем новое значение центрального пикселя, с учетом его соседних пикселей. Заметим, что размер блока должен быть нечетным, потому что ядро свертки как бы "накладывается" своим центром на наш пиксель, а окружающие пиксели также покрываются ядром. Если применить свертку к каждому пикселю, то получится некий эффект, зависящий от ядра свертки.

Свертка широко используется для обработки изображений независимо от нейронных сетей, с помощью нее можно накладывать различные эффекты на изображения, такие как размытие, выделение границ и повышение четкости. Приведем эти ядра свертки в качестве примера:

$$\begin{pmatrix} 1/9 & 1/9 & 1/9 \\ 1/9 & 1/9 & 1/9 \\ 1/9 & 1/9 & 1/9 \end{pmatrix} - \text{размытие}, \begin{pmatrix} 0 & -1 & 0 \\ -1 & 4 & -1 \\ 0 & -1 & 4 \end{pmatrix} - \text{выделение границ}, \begin{pmatrix} 0 & -1 & 0 \\ -1 & 5 & -1 \\ 0 & -1 & 4 \end{pmatrix} - \text{повышение четкости}$$

- **Разделяемые веса:** все нейроны используют одно и то же ядро свертки для обработки разных регионов изображения.
- **Уменьшение размерности:** важно уметь распознавать объекты вне зависимости от того, какую часть изображения они занимают. На практике чаще всего важнее факт наличие какого-то объекта или признака, а не его местоположение на изображении.

Для уменьшения размерности в СНС используются **слои подвыборки(sub-sampling, подвыборочные слои, пулингвые слои)**: один нейрон слоя подвыборки подключен к нескольким нейронам предыдущего слоя(обычно к фрагменту 2×2). Значения нейрона следующего слоя определяется значением нейронов предыдущего слоя.



2.2 Архитектура построения СНС

СНС состоят из чередующихся слоев свертки и подвыборки(convolution and sub-sampling layers), соединенных между собой.

Разберемся как работают сверточные нейронные сети на примере сети для распознавания рукописных цифр LeNet-5. Подробнее об этой сети можно почитать в статье: *"Back-Propagation Applied to Handwritten Zip Code Recognition / Y. LeCun, B.Boser, J.S. Denker et al. 1989"*

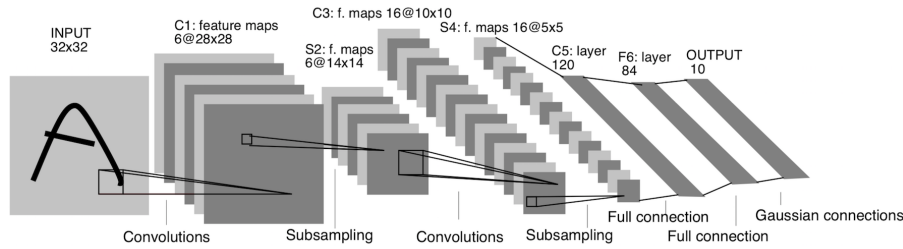


Рис. 3: Сеть LeNet-5 для распознавания рукописных цифр, предложенная Яном Лекунем

LeNet-5 состоит из пяти слоев и работает с изображениями размера 32×32 пикселя. Первый слой - слой свертки. Используется 6 разных ядер свертки размера 5×5 для поиска разных признаков изображений.

Определение 1. Карта признаков (feature map) – набор слоев нейронов, каждый из которых использует разное ядро свертки для поиска различных признаков

Следующий слой – слой подвыборки. Использует усреднение с окном размера 2×2 . Затем повторяется еще один слой свертки, который использует 16 ядер свертки. И пятый слой - снова слой подвыборки с усреднением окна 2×2 . Для того, чтобы решить задачу классификации на основе этих признаков используется еще 2 слоя. Выходной слой содержит 10 нейронов, каждый из которых говорит о наличии одной из цифр.

2.3 Сверточный слой

Как и говорилось выше, сверточный слой нейронной сети представляет из себя применение операции свертки к выходам предыдущего слоя. Слой свертки включает в себя ядро свертки, весовые коэффициенты которого неизвестны и устанавливаются в процессе обучения.

В одном сверточном слое может быть несколько сверток. В этом случае для каждой свертки на выходе получится свое изображение. Например, если вход имел размерность $w \times h$, а в слое было n сверток с ядром размерность $i \times j$, то выход будет иметь размерность $n \times (w - i + 1) \times (h - j + 1)$

Применение операции свертки уменьшает изображение. Также пиксели, которые находятся на границе изображения вносят вклад в меньшее количество клеток, чем другие. Поэтому в сверточных слоях используется дополнение изображения (padding). Выходные данные предыдущего слоя дополняются пикселями так, чтобы после свертки сохранился размер изображения. Такие свертки называют одинаковыми (same convolution), а свертки без дополнения изображения называются правильными (valid convolution). Среди способов, которыми можно заполнить новые пиксели, можно выделить следующие:

- *zero shift*: 00|ABC|00
- *border extension*: AA|ABC|CC
- *mirror shift*: BA|ABC|CB
- *cyclic shift*: BC|ABC|AB

Обычно свертка обходит все пиксели подряд, но иногда еще одним параметром сверточного слоя является *stride* (шаг).

2.4 Пулинговый слой

Этот слой снижает размерность изображения. Исходное изображение делится на блоки размером $w \times h$ и для каждого блока вычисляется функция – обычно это функция максимума (*max pooling*) или (взвешенного) среднего (*weighted*) (*average pooling*). У слоя нет обучаемых параметров. **Цели пулингового слоя:**

- уменьшение размерности изображения, чтобы следующие свертки работали над большей частью исходного изображения
- ускорение вычислений

3 Используемая литература

- Salman Khan, Hossein Rahmani, Syeq Afaq Ali Shah, Mohammed Bennamoun "A Guide to Convolutional Neural Networks for Computer Vision"