

Сверточная нейронная сеть для распознавания образов

Михно Егор Владимирович

студент кафедры интеллектуальных информационных технологий
БрГТУ

mikhnoegor07@gmail.com

Научный руководитель

Головко Владимир Адамович, заведующий кафедрой
интеллектуальных информационных технологий БрГТУ, доктор
технических наук

gva@bstu.by

Введение

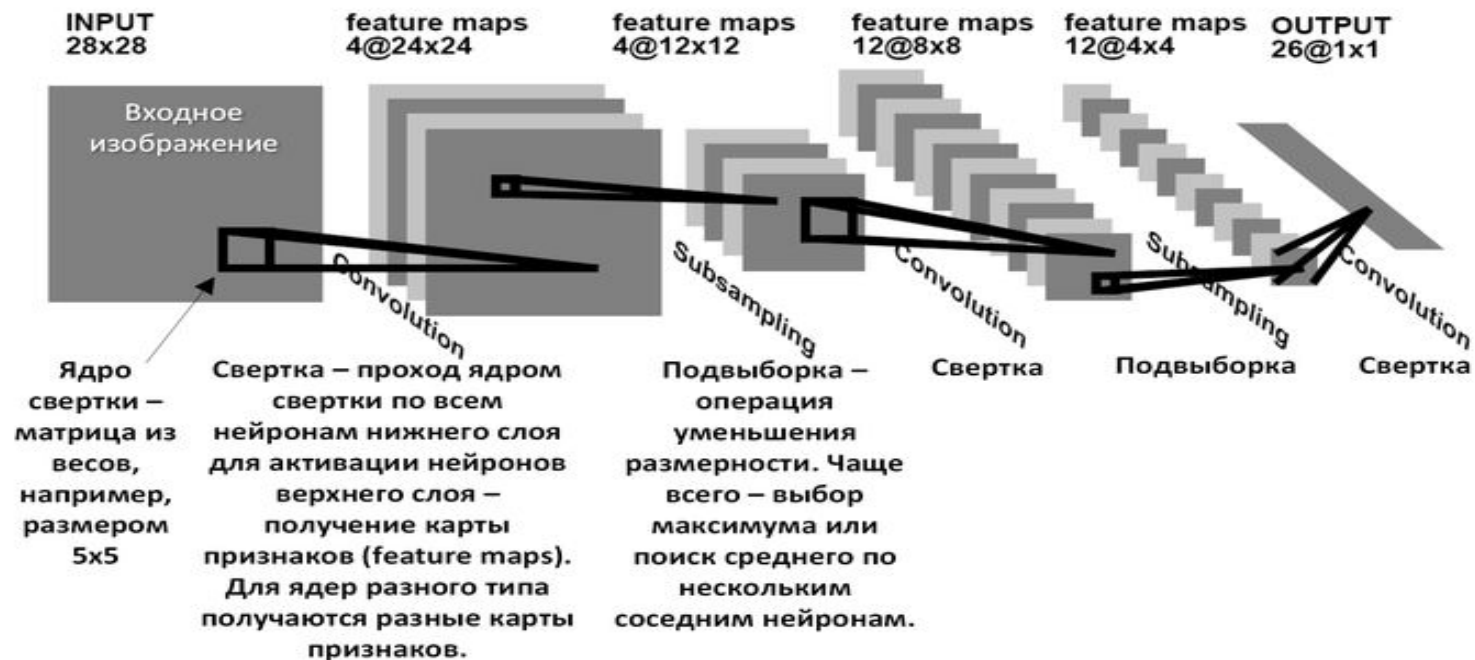
Одной из самых сложных задач в информационных технологиях является задача распознавания объектов на изображениях. Исследования по распознаванию объектов являются одним из приоритетных направлений развития науки и техники.

Практически у всех систем распознавания символов на изображениях точностные характеристики резко снижаются при искажениях входного изображения.

Для решения данной задачи эффективно использовать нейронные сети в связи с тем, что они слабо чувствительны к искажениям входного сигнала и обладают высокой скоростью распознавания.

Сверточная нейронная сеть

Идея классических сверточных нейронных сетей заключается в использовании чередующихся сверточных и субдискретизирующих слоев и многослойного персептрона на выходе.



Особенности

- Локальное восприятие.
- Свертка.
- Субдискретизация.

Локальное восприятие

Подразумевает, что на вход одного нейрона подается не все изображение, а лишь некоторая его область, ограниченная ядром свертки. Каждое ядро свертки формирует собственную карту признаков, делая нейронную сеть многомерной.

Свертка

Операция свертки подразумевает умножение каждого фрагмента изображения поэлементно на ядро свертки, которое выступает в качестве матрицы весовых коэффициентов, и суммирование результата. Полученная в итоге матрица является картой признаков данного изображения.

Данная концепция подразумевает использование небольшого количества весовых коэффициентов для большого количества связей. К примеру, 1 изображению размерности 32×32 , выделенным из него 4-м картам признаков и ядру свертки размерности 5×5 будет соответствовать $4 \times 5 \times 5 = 100$ весовых коэффициентов и 4 пороговых значения, по 1 на каждую карту.

Искусственно введенное ограничение веса положительно влияет на способность сети находить инварианты в изображении и реагировать главным образом на них, не акцентируя внимание на прочие шумы.

Субдискретизация

Субдискретизация уменьшает размерности карт признаков. В качестве операции сжатия используется выбор максимального элемента из ядра обхода или их усреднение.

Данная операция ускоряет дальнейшие вычисления и обеспечивает инвариантность к масштабу.

MNIST

Для сравнения сверточных нейронных сетей будет использоваться MNIST - база данных рукописных цифр, содержащая 60 000 образов для обучения и 10 000 образов для тестирования. Размерность образов 28x28 пикселей.

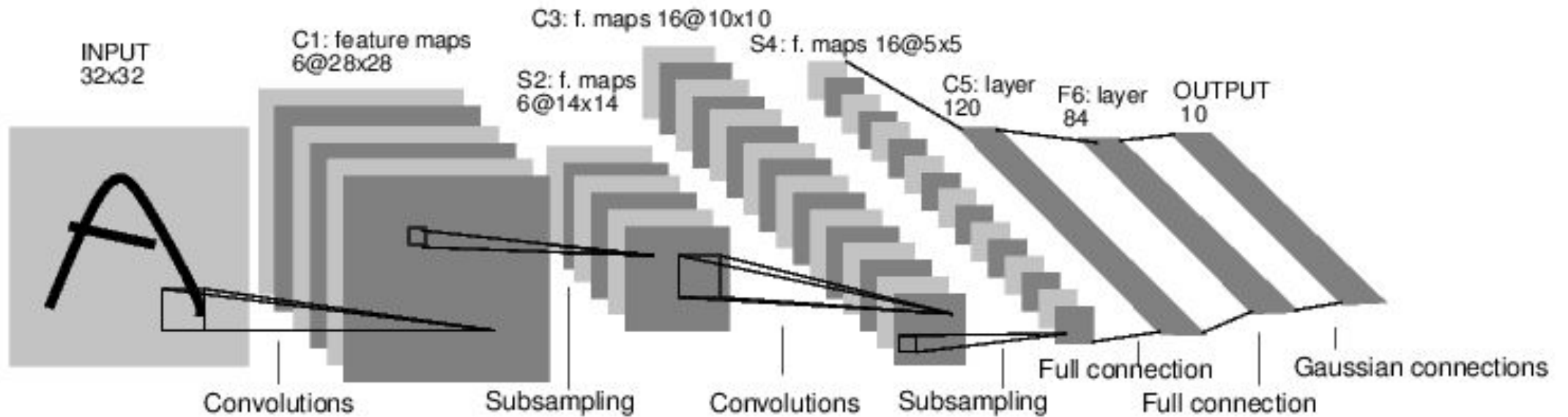
Пример первых 12 цифр из обучающего набора:



Сверточная сеть Яна Лекуна

Сверточная сеть Яна Лекуна для распознавания рукописных цифр из базы данных MNIST представляет собой нейронную сеть из 8-ми слоев (с учетом входного, полносвязного и выходного слоев) использующую чередующиеся сверточные и субдискретизирующие слои и многослойный персептрон на выходе.

Структура сверточной сети Яна Лекуна



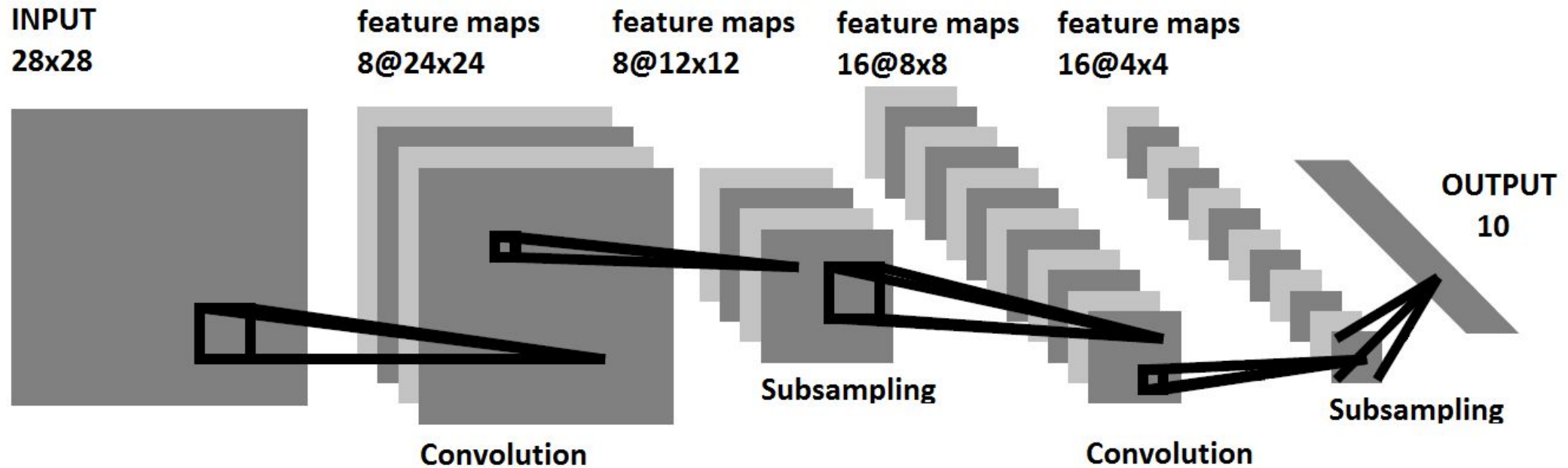
Структура сверточной сети Яна Лекуна

1. Входной слой. Одно изображение размерности 32×32 .
2. Сверточный слой. 6 карт признаков размерности 28×28 (ядро обхода 5×5).
3. Субдискретизирующий слой. 6 карт признаков размерности 14×14 .
4. Сверточный слой. 16 карт признаков размерности 10×10 (ядро обхода 5×5).
5. Субдискретизирующий слой. 16 карт признаков размерности 5×5 .
6. Сверточный слой. 120 карт признаков размерности 1×1 (ядро обхода 6×6).
7. Полносвязный слой. 84 нейрона.
8. Выходной слой. 10 нейронов.

Модифицированная сеть Яна Лекуна

Модифицированная нейронная сеть Я. Лекуна состоит из 6 слоев. Размерность входного изображения 28x28. Убран полносвязный слой. Количество и размерность карт признаков изменены. Функции активации сети сигмоидные. Количество элементов в групповой выборке: 50.

Структура модифицированной сети Яна Лекуна



Структура модифицированной сети Яна Лекуна

1. Входной слой. Одно изображение размерности 28x28.
2. Сверточный слой. 8 карт признаков размерности 24x24 (ядро обхода 5x5).
3. Субдискретизирующий слой. 8 карт признаков размерности 12x12.
4. Сверточный слой. 16 карт признаков размерности 8x8 (ядро обхода 5x5).
5. Субдискретизирующий слой. 16 карт признаков размерности 4x4.
6. Выходной слой (сверточный слой). 10 нейронов (ядро обхода 4x4).

Сравнение сетей

Признаки	Сверточная сеть Я. Лекуна	Модифицированная сеть Я. Лекуна
Число слоев	8	6
Число карт признаков	6 x 16 x 120	8 x 16
Точность тестирования без искажений входных данных	99.05%	99.29%
Точность тестирования с искажениями входных данных	99.2%	-

Как мы видим, модифицированная сеть обладает большей скоростью отклика из-за меньшего числа слоев, занимает меньший объем памяти в связи с меньшим количеством хранимых признаков и показывает более высокую точность распознавания в сравнении с классической (LeNet5) сверточной сетью Я. Лекуна.

Результаты классификации различных сверточных сетей

Классификатор	Предварительная обработка	Количество ошибок при тестировании (%)
Convolution net LeNet-1	Subsampling to 16x16 pixels	1.7
Convolution net LeNet-4	none	1.1
Convolution net LeNet-5, [no distortions]	none	0.95
Convolution net LeNet-5, [distortions]	none	0.8
Convolution net Boosted LeNet-4, [distortions]	none	0.7

Вывод

Модифицированная нейронная сеть Я. Лекуна обладает большей точностью и скоростью работы чем классическая LeNet5, что говорит о целесообразности внесенных изменений.