

СВЕРТОЧНАЯ НЕЙРОННАЯ СЕТЬ ДЛЯ РАСПОЗНАВАНИЯ ОБРАЗОВ

Рассматривается возможность распознавания рукописных цифр на базе данных MNIST с использованием сверточных нейронных сетей различной структуры

ВВЕДЕНИЕ

Для решения задачи распознавания будет применяться 2 сверточные нейронные сети: классическая Яна Лекуна и модифицированная.

I. СВЕРТОЧНАЯ СЕТЬ ЯНА ЛЕКУНА

Сверточная сеть Яна Лекуна представляет собой 8-ми слойную нейронную сеть с чередующимися сверточными и субдескрипторными слоями и выходным слоем, возвращающим в качестве результата число от 0 до 9. Сверточный слой формирует карты признаков путем последовательного обхода изображения ядром обхода размерности 5x5, являющегося матрицей весовых коэффициентов. Количество матриц весовых коэффициентов на сверточном слое равняется числу изображений или карт признаков поступающих на слой. Субдескрипторный слой сжимает поступающие карты признаков в N раз, усредняя её значения. В качестве функций активации использовались гиперболические тангенсы. Пример структуры данной сети (рис. 1):

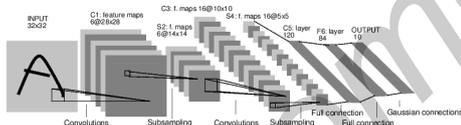


Рис. 1 – Структура сверточной сети Яна Лекуна (LeNet5)

В сети использовался 1 распределяющий (входной) слой, по 2 чередующихся между собой сверточных (ядро обхода 5x5) и субдескрипторных слоев, затем сверточный полносвязный слой (ядро обхода 6x6), полносвязный слой, и выходной. Размерности карт признаков: 1@32x32 - 6@28x28 - 6@14x14 - 16@10x10 - 16@5x5 - 120@1x1 - 86@1x1 - 10@1x1.

Для обучения и тестирования сети использовалась база данных MNIST: 60 000 образов для обучения и 10 000 для тестирования. Итоговая точность распознавания данной сети составляет 99.05% без искажений входных образов и 99.2% с искажениями.

Михно Егор Владимирович, студент кафедры интеллектуальных информационных технологий БрГТУ, mikhnoegor07@gmail.com.

Научный руководитель: Головкин Владимир Адамович, заведующий кафедрой интеллектуальных информационных технологий БрГТУ, доктор технических наук, gva@bstu.by.

II. МОДИФИЦИРОВАННАЯ СЕТЬ ЯНА ЛЕКУНА

В качестве основы используется сверточная нейронная сеть Я. Лекуна, но с некоторыми изменениями. Размерность входного изображения 28x28. Количество карт признаков и их размерность изменены. Функции активации сети сигмоидные. Полносвязные слои опускаются. Скорость обучения сети (коэффициент альфа) изменяется от 0.8 до 0.001 в процессе обучения (чем выше точность, тем ниже альфа). Пример структуры данной сети (рис. 2):

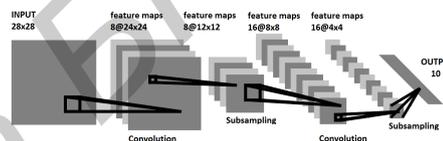


Рис. 2 – Структура модифицированной сверточной сети

В сети использовался 1 распределяющий (входной) слой, по 2 чередующихся между собой сверточных (ядра обхода 5x5) и субдескрипторных слоев и выходной слой (сверточный полносвязный, ядро обхода 4x4). Размерности карт признаков: 1@28x28 - 8@24x24 - 8@12x12 - 16@8x8 - 16@4x4 - 10@1x1.

Итоговая точность распознавания данной сети составляет 99.29% без искажений входных образов.

III. ВЫВОДЫ

Предлагаемая модификация сверточной нейронной сети Яна Лекуна позволяет увеличить точность распознавания образов на 0.1 - 0.24, что позволит с большей вероятностью правильно идентифицировать входные образы.

1. <http://yann.lecun.com/exdb/mnist/> - Ян Лекун, Нейронные сети для распознавания рукописных цифр на базе MNIST