

СВЕРТОЧНАЯ НЕЙРОННАЯ СЕТЬ ДЛЯ КЛАССИФИКАЦИИ БЕЗ ИСПОЛЬЗОВАНИЯ ДОПОЛНИТЕЛЬНЫХ БИБЛИОТЕК.

В.С. Адаричев, Ю.А. Иванова
Томский политехнический университет
E-mail: shoutsoul97@mail.ru, jbolotova@tpu.ru

Аннотация

В работе представляется сверточная нейронная сеть для классификации цифр на изображениях реализованная без использования дополнительных библиотек машинного обучения. Для оценки качества сегментации применяются различные метрики, такие как: ошибка по модулю и процентная точность. Проведено сравнение различных топологий нейронной сети, на выборке данных, находящейся в библиотеке машинного обучения «Scikit-learn» [1]. По полученным результатам тестирования сделан вывод о наилучшей из рассматриваемых топологий нейронной сети.

Введение

Сверточная нейронная сеть (convolutional neural network, CNN) - специальная архитектура искусственных нейронных сетей, предложенная Яном Лекуном в 1988 году и нацеленная на эффективное распознавание образов, входит в состав технологий глубокого обучения. Использует некоторые особенности зрительной коры, в которой были открыты так называемые простые клетки, реагирующие на прямые линии под разными углами, и сложные клетки, реакция которых связана с активацией определённого набора простых клеток. Таким образом, идея сверточных нейронных сетей заключается в чередовании слоев свертки и подвыборки. [2]

Название архитектура сети получила из-за наличия операции свёртки, суть которой в том, что каждый фрагмент изображения умножается на матрицу (ядро) свёртки поэлементно, а результат суммируется и записывается в аналогичную позицию выходного изображения [2,3].

В данной работе исследуется качество классификации цифр на изображениях сверточной нейронной сетью с различной топологией.

Программа реализована на языке Python, с использованием Google Colaboratory - сервиса, который обеспечивает бесплатный доступ к вычислительным ресурсам.

Описание выборки данных

В данной работе используется выборка данных, находящаяся в библиотеке машинного обучения «Scikit-learn» [1]. Выборка содержит 1797 изображений размером 8x8 пикселей. Выборка была разделена на 2 части. Первая часть содержит 1437 изображений, что составляет 80% от всей выборки, вторая содержит 360 изображений, что составляет 20% от всей выборки. Первая часть

используется для обучения нейронной сети, вторая для тестирования.

На данных изображениях необходимо классифицировать цифры от 0 до 9.

Пример входных изображений из выборки можно увидеть на рисунке ниже (рис 1).



Рис. 1. Примеры изображений из выборки.

Размерность входных изображений может быть различной, и от размера изображения будет зависеть только скорость вычислений и время, требуемое для обучения сети.

Топологии исследуемых нейронных сетей

Каждая из топологий нейронной сети имела логистическую функцию активации (сигмоида) на выходном слое:

$$f(x) = \frac{1}{1 + e^{-x}},$$

и линейный выпрямитель с «утечкой» (Leaky ReLU) на всех остальных слоях:

$$f(x) = \begin{cases} 0.01x & x < 0 \\ x & x \geq 0. \end{cases}$$

В работе рассматривается 3 различных топологии.

«Нейронная сеть 1» содержала в себе 1 сверточный слой с 15 картами признаков, и на выходе полносвязный слой с 10 нейронами, представляющими собой 10 цифр.

«Нейронная сеть 2» содержала в себе 2 сверточных слоя с 25 и 15 картами признаков, и на выходе полносвязный слой с 10 нейронами, представляющими собой 10 цифр.

«Нейронная сеть 3» содержала в себе 2 сверточных слоя с 25 и 15 картами признаков, полносвязный слой с 25 нейронами и на выходе полносвязный слой с 10 нейронами, представляющими собой 10 цифр.

Результаты тестирования

Было проведено обучение [4] и тестирование нейронных сетей. Каждая нейронная сеть обучалась на 100 эпохах, с ядром свертки размером 3x3, и скоростью обучения «0.01» на каждом слое.

Качество классификации изображений сравнивается по следующим метрикам качества: ошибка по модулю (module) и процентная точность (percent). Где module – это средняя ошибка

сегментации пикселя по модулю, а percent – это количество верно классифицированных изображений (в процентах).

Результаты тестирования «Нейронная сеть 1» представлены в таблице ниже (таблица 1):

Таблица 1. Результаты тестирования «НС 1»

Выборка	module	percent
Обучающая	0.024	99.582%
Тестовая	0.064	96.944%

Результаты тестирования «Нейронная сеть 2» представлены в таблице ниже (таблица 2):

Таблица 2. Результаты тестирования «НС 2»

Выборка	module	percent
Обучающая	0.013	99.652%
Тестовая	0.057	97.778%

Результаты тестирования «Нейронная сеть 3» представлены в таблице ниже (таблица 3):

Таблица 3. Результаты тестирования «НС 3»

Выборка	module	percent
Обучающая	0.014	99.443%
Тестовая	0.064	96.667%

В таблице приведено сравнение времени обучения с использованием динамической компиляции из библиотеки Numba [5], и обычной компиляции.

Таблица 4. Время обучения нейронной сети в секундах.

	Обычная компиляция	Динамическая компиляция
НС1	1537	114
НС2	2911	260
НС3	3182	271

Сравнение полученных результатов

По результатам, полученным на основе 10 запусков взят средний результат каждого варианта топологии нейронной сети.

По полученным результатам можно сделать вывод, что наилучшую точность показала «Нейронная сеть 2» со следующей топологией: сверточный слой с 25 нейронами, слой свертки с 15 нейронами, и на выходе полносвязный слой с 10 нейронами.

Динамическая компиляция ускоряет обучение в среднем в 12 раз.

Заключение

Разработанная программа позволяет обучать нейронную сеть, которая позволяет классифицировать изображения. При корректном

подборе гиперпараметров и обучающей выборки нейронная сеть отлично справляется с данной задачей.

Минусом данного подхода является, высокие затраты вычислительных мощностей при использовании большого количества эпох из-за трудоемкой операции свертки, что в какой-то мере было решено использованием динамической компиляции.

С увеличением количества слоев увеличивается необходимое количество эпох для обучения нейронной сети.

Исследование выполнено при финансовой поддержке РФФИ в рамках научного проекта № 18-08-00977 А и в рамках Программы повышения конкурентоспособности ТПУ.

Список использованных источников:

1. База данных цифр для классификации. [Электронный ресурс] / Библиотека машинного обучения Scikit-learn — URL: https://scikit-learn.org/stable/modules/generated/sklearn.datasets.load_digits.html#sklearn.datasets.load_digits (дата обращения 10.12.2019)
2. Что такое сверточная нейронная сеть [Электронный ресурс] / IT форум Хабр – URL: <https://habr.com/ru/post/309508> (дата обращения 20.11.2019)
3. Сверточная нейронная сеть, часть 1: структура, топология, функции активации и обучающее множество [Электронный ресурс] / IT форум Хабр – URL: <https://habr.com/ru/post/348000> (дата обращения 21.11.2019)
4. Сверточная нейронная сеть, часть 2: обучение алгоритмом обратного распространения ошибки. [Электронный ресурс] / IT форум Хабр – URL: <https://habr.com/ru/post/348028> (дата обращения 26.11.2019)
5. A JIT Compiler for Numerical Functions in Python. [Электронный ресурс] – URL: <https://github.com/numba/numba> (дата обращения 20.12.2019)