

ПРОЕКТИРОВАНИЕ СВЕРТОЧНЫХ НЕЙРОННЫХ СЕТЕЙ ДЛЯ РАСПОЗНОВАНИЯ РУКОПИСНЫХ СИМВОЛОВ

Попова Е.С.

Научный руководитель Болотова Ю.А.
Томский политехнический университет
esp9@tpu.ru

Введение

На сегодняшний день существует необходимость в распознавании различных образов, а в частности рукописных символов. Рукописные цифры или буквы, очевидны для человека, но для компьютеров идентификация таких символов — очень сложная задача, поэтому распознавание рукописных символов является достаточно актуальной на сегодняшний день.

Для решения этой задачи часто применяются сверточные нейронные сети, так как являются наиболее эффективными в области классификации изображений [1].

Свёрточная нейронная сеть была представлена в 1998 году французским исследователем Яном Лекуном [1], как развитие модели неокогнитрон.

Обоснования к использованию

Свёрточные нейронные сети, в отличие от остальных нейросетевых архитектур, обеспечивают частичную устойчивость к изменениям масштаба, смещениям, поворотам, смене ракурса и прочим искажениям. Они объединяют три архитектурных идеи, для обеспечения инвариантности к изменению масштаба, повороту сдвигу и пространственным искажениям [1]:

- локальные рецепторные поля (обеспечивают локальную двумерную связность нейронов);
- общие синаптические коэффициенты (обеспечивают детектирование некоторых черт в любом месте изображения и уменьшают общее число весовых коэффициентов);
- иерархическая организация с пространственными подвыборками.

Именно поэтому для решения поставленной задачи была выбрана свёрточная нейронная сеть, обучающаяся на основе алгоритма обратного распространения ошибки.

Структура сверточной нейронной сети

Модель свёрточной сети, которая будет рассмотрена в этой статье, состоит из трёх типов слоёв: свёрточные (convolutional) слои, субдискретизирующие (subsampling, подвыборка) слои и слои "обычной" нейронной сети — перцептрона [3].

Принцип функционирования сверточного слоя:

$$y_k^{(i,j)} = b_k + \sum_{s=1}^K \sum_{t=1}^K w_{k,s,t} x^{((i-1)+s,(j+t))}$$

где, $y_k^{(i,j)}$ - нейрон k -ой плоскости свёрточного слоя, b_k - нейронной смещение k -ой плоскости, K -

размер рецептивной области нейрона, $W_{k,s,t}$ - матрица синаптических коэффициентов, X - выходы пейропов предыдущего слоя.

Принцип функционирования подвыборочного слоя:

$$y_k^{(i,j)} = b_k + \frac{1}{4} w_k \sum_{s=1}^K \sum_{t=1}^K x^{((i-1)+s,(j+t))}$$

Затем полученный результат подвыборки, пропускается через активационную функцию:

$$out_k^{(i,j)} = f(y_k^{(i,j)})$$

На рисунке 1 представлена модель свёрточной нейронной сети, где C-Layer – это свёрточный слой, S-Layer – подвыборочный слой.

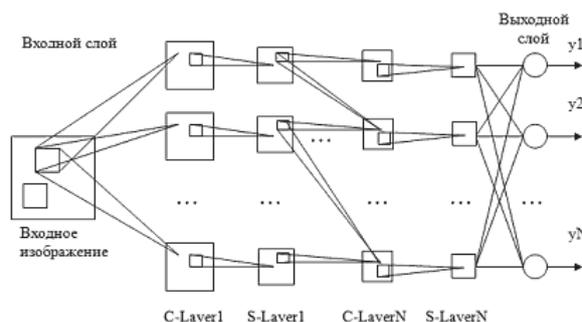


Рис. 3. Модель свёрточной нейронной сети

В данной работе в качестве функции активации использовался гиперболический тангенс, который обладает следующими свойствами:

- Симметричные активационные функции, типа гиперболического тангенса, обеспечивают более быструю сходимость, чем стандартная логистическая функция;
- Функция имеет непрерывную первую производную;
- Функция имеет простую производную, которая может быть вычислена через ее значение, что дает экономию вычислений.

Алгоритм обратного распространения ошибки

Данный алгоритм используется для минимизации отклонения реальных значений выходных сигналов нейронной сети от требуемых. В качестве функции ошибки ИНС рассматривается следующая величина [2]:

$$E(w) = \frac{1}{2} \sum_{i,k} (f_{i,k} - y_{i,k}^{(T)})^2$$

где $f_{i,k}$ – значение выходного сигнала k -го выходного нейрона сети при подаче на её входы i -

го набора обучающих данных, $y_{j,k}^{(T)}$ – требуемое значение выходного сигнала k -го выходного нейрона для i -го набора данных для обучения. Обучение ИНС направлено на минимизацию функции $E(w)$.

Изменение веса связи определяется следующим образом:

$$\Delta w_{ij}^{(q)} = \eta \delta_j x_i$$

где δ_j – значение ошибки j -го нейрона в слое q , x_i – значение i -го входного сигнала для j -го нейрона слоя q . Данная формула применима и для настройки смещений нейронов, только вместо x_i необходимо подставить «1».

Отметим, что значение ошибки нейрона определяется в зависимости от его положения в сети. Для нейронов выходного сло:

$$\delta_j = (f_{i,k}(S))' (f_{i,k} - y_{i,k})$$

где $y_{i,k}$ – требуемое, а $f_{i,k}$ – фактическое значение выходного сигнала k -го нейрона для i -го набора данных из обучающей выборки, $(f_{i,k}(S))'$ – значение производной активационной функции k -го нейрона для i -го набора обучающих данных. Если нейрон принадлежит одному из скрытых слоев, то

$$\delta_i^{(q)} = (f_i^{(q)}(S))' \sum_j w_{ij} \delta_j^{(q+1)}$$

где $\delta_i^{(q)}$ – ошибка i -го нейрона в слое q , $\delta_j^{(q+1)}$ – ошибка j -го нейрона в $(q+1)$ слое, w_{ij} – вес связи, соединяющей эти нейроны, $(f_i^{(q)}(S))'$ – значение производной активационной функции i -го нейрона слоя q .

Анализ полученных результатов

Точность распознавания сети:

$$R = \frac{n}{N} = \frac{4650}{10000} = 0,465,$$

где, R – точность распознавания по всему набору тестовой выборки, n – количество правильно распознанных символов из тестовой выборки, N – количество элементов в тестовые выборки (так как тестирование сети проходила на наборе данных MNIST, $N = 10000$).

Исходя из полученных результатов были сделаны предположения, что для улучшения качества распознавания и ускорения сходимости сети необходимо:

- Усложнить структуру сети, за счет добавления дополнительных сверточных и подвыборочных слоев;
- В качестве функции активации использовать выпрямленную линейную функцию активации (ReLU) : $\sigma(x) = \max(0, x)$;
- Распараллелить процесс обучения сети внутри каждого слоя.

Проектирование

В качестве языка реализации, был выбран язык разработки – C#, относящийся к семейству языков с C-подобным синтаксисом. Данный язык включает

очень многие достоинства языков-предшественников и исключает некоторые модели, зарекомендовавшие себя как проблематичные при разработке программных систем.

Так же основным аспектом при выборе данного языка программирования, была поддержка ООП (объектно-ориентированное программирование).

На рисунке 2 представлена разработанная логическая модель сверточной нейронной сети в виде UML диаграммы классов.

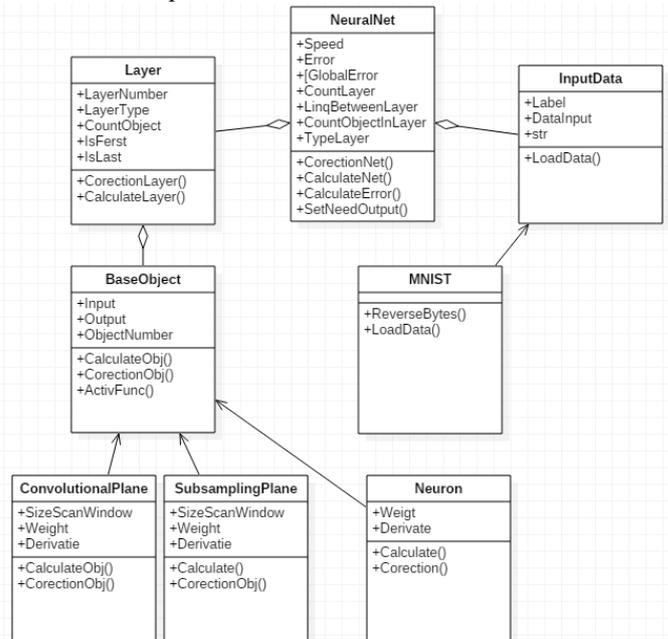


Рисунок 2 – Логическая модель сверточной нейронной сети в виде UML диаграммы классов.

Заключение

В процессе исследования был проведен анализ алгоритмов классификации, выбор наиболее подходящего алгоритма на основе нейронных сетей, исследованы алгоритмы обучения и виды функции активации нейронов, выбраны наиболее подходящие для реализуемой архитектуры нейронной сети.

Список использованных источников

1. Хайкин С. Нейронные сети: полный курс. М.: Вильямс, 2006. - 1104 с.
2. The MNIST database of handwritten digits [Электронный ресурс] / LeCun, Y. – URL: <http://yann.lecun.com/> (дата обращения: 15.10.2016).
3. LeCun, Y. Efficient BackProp in Neural Networks: Tricks of the trade / Y. LeCun, L. Bottou, G. Orr, K. Muller – Springer, 1998.
4. Гонсалес, Р. Цифровая обработка изображений / Р. Гонсалес, Р. Вудс. – М.: Техносфера, 2005. – 1072 с.