

ОБЕСПЕЧЕНИЕ УСТОЙЧИВОСТИ СВЕРТОЧНОЙ НЕЙРОННОЙ СЕТИ ПРИ АНАЛИЗЕ РУКОПИСНЫХ ТЕКСТОВ

Белорусский государственный университет информатики и радиоэлектроники
г. Минск, Республика Беларусь

Попитич А. Я.

Серебряная Л. В. – к.т. техн. наук, доцент

Для решения задач распознавания, связанных с использованием нейронных сетей, наиболее актуальным и важным свойством является их обобщающая способность и инвариантность параметров сети к различным отклонениям. В данном докладе описан метод улучшения качества распознавания при использовании сверточной нейронной сети.

В ходе решения задачи распознавания рукописных цифр с использованием сверточной нейронной сети были предприняты попытки модификации классической архитектуры для повышения устойчивости к масштабу распознаваемых рукописных цифр.

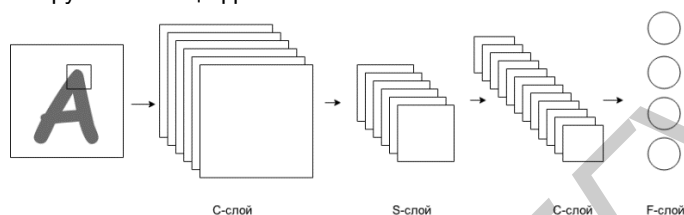


Рис. 1 – Архитектура сверточной нейронной сети

Если рассмотреть классическую архитектуру сети данного вида (рисунок 1), то можно отметить следующие особенности: сеть состоит из чередующихся сверточных и субдискретизирующих слоев. Сверточный слой выделяет инварианты на изображении путем применения операции свертки части изображения с ядром фиксированного размера (в результате получается карта признаков). Субдискретизирующий слой выполняет функцию уменьшения пространства признаков.

После детального анализа, был сделан вывод о том, что применение свертки с ядром фиксированного размера приводит к тому, что успешность распознавания будет зависеть от масштаба анализируемого объекта и размера свертки. На рисунке 2 приведены иллюстрации, показывающие вышеописанную проблему.

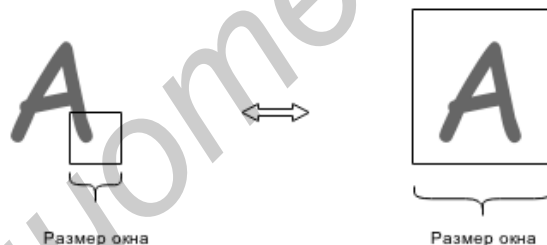


Рис. 2 – Иллюстрация проблемы выбора размера окна для операции свертки

Для придания сети инвариантности к масштабу было решено использовать ядра разных размеров. В результате чего будут получены карты признаков одинакового размера, но полученные при помощи ядер разного размера, что позволит находить больше признаков уже на первом слое нейронной сети.

Однако, как выяснилось, использование ядер разного размера привело к появлению перекрывающихся связей для соседних нейронов одной карты признаков, что в свою очередь приводит к увеличению количества связей-дубликатов для каждого нейрона. Данный недостаток, как показала практика, отражается в небольшом падении быстродействия.

Можно заметить, что при увеличении размера окна увеличивается и количество входов для нейрона следующего уровня. Соответственно, чем больше входных связей у нейрона, тем меньший вклад в формирование результата вносит каждая из них. Это предположение и было использовано для уменьшения связей нейрона путем удаления некоторого числа связей от общего количества.

Таким образом, были внесены изменения в архитектуру сверточной сети, путем применения ядер разных размеров и уменьшения количества связей-дубликатов для каждого нейрона. Данные изменения позволили обеспечить устойчивость алгоритма обучения и распознавания к масштабу, сохраняя при этом практически прежний уровень быстродействия.

Список использованных источников:

1. Y. Le Cun, Handwritten digit recognition with a back-propagation network / Y. LeCun, L. Bottou, Y. Bengio, and P. Haffner // AT & T Bell Laboratories, Homdel, - 1990 – 9 p.
2. Y. Taigman, DeepFace: Closing the Gap to Human-Level Performance in Face Verification / Y. Taigman, M. Yang, Marc'Aurelio Ranzato, L. Wolf // Facebook AI Research, Menlo Park, USA, 2014 – 8 p.