

НЕЙРОННАЯ СЕТЬ ДЛЯ ИДЕНТИФИКАЦИИ ОБЪЕКТОВ НА ИЗОБРАЖЕНИЯХ

Проводится обзор и сравнение нейронных сетей для распознавания объектов на изображениях.

ВВЕДЕНИЕ

Основные задачи, которые ставятся перед нейронными сетями, относятся к задачам распознавания образов. Они заключаются в том, чтобы классифицировать входной образ, то есть отнести его к какому-либо известному сети классу. Изначально сети даются эталонные образы – такие образы, принадлежность которых к определенному классу известна. Затем на вход сети подается некоторый неизвестный образ, и сеть пытается по определенному алгоритму соотнести его с каким-либо эталонным образом. Можно сказать, что нейронные сети проводят кластеризацию образов.

I. «ОБЫЧНЫЕ» НЕЙРОННЫЕ СЕТИ

Под «обычными» нейронными сетями понимают полносвязные нейронные сети прямого распространения с обратным распространением ошибки. Как следует из названия в такой сети каждый нейрон связан с каждым, сигнал идет только в направлении от входного слоя к выходному, нет никаких рекурсий.

Для подачи изображения на вход нейронной сети его необходимо выразить в виде одномерного вектора. В таком случае должно быть $M \times N$ входных нейронов, где M и N – ширина и высота изображения в пикселях. Количество входных нейронов будет достаточно большим, поэтому сеть будет очень сложной. К тому же при переводе изображения в последовательность байт теряется его топология, т.е. взаимосвязь между отдельными его частями. Кроме того, задача распознавания подразумевает умение нейронной сети быть устойчивой к небольшим сдвигам, поворотам и изменению масштаба изображения, т.е. она должна извлекать из данных некие инварианты. Решение этой проблемы было найдено американским ученым французского происхождения Яном ЛеКуном, который предложил использовать так называемые свёрточные нейронные сети.

II. СВЁРТОЧНЫЕ НЕЙРОННЫЕ СЕТИ

Идея свёрточных нейронных сетей заключается в чередовании свёрточных слоев (S-

layers), субдискретизирующих слоев (S-layers) и наличия полносвязных (F-layers) слоев на выходе. Такая архитектура включает в себе 3 основных парадигмы:

- локальное восприятие;
- разделяемые веса;
- субдискретизация.

Локальное восприятие подразумевает, что на вход одного нейрона подается не все изображение (или выходы предыдущего слоя), а лишь некоторая его область. Такой подход позволил сохранять топологию изображения от слоя к слою.

Концепция *разделяемых весов* предполагает, что для большого количества связей используется очень небольшой набор весов. Т.е. если у нас имеется на входе изображение размерами 32×32 пикселя, то каждый из нейронов следующего слоя примет на вход только небольшой участок этого изображения размером, к примеру, 5×5 , причем каждый из фрагментов будет обработан одним и тем же набором. Важно понимать, что самих наборов весов может быть много, но каждый из них будет применен ко всему изображению. Такие наборы часто называют ядрами (kernels).

Суть *субдискретизации* и S-слоев заключается в уменьшении пространственной размерности изображения. Т.е. входное изображение грубо (усреднением) уменьшается в заданное количество раз. Субдискретизация нужна для обеспечения инвариантности к масштабу.

III. ВЫВОДЫ

Свёрточная нейронная сеть является одним из лучших решений для распознавания и классификации изображений. Её использование позволяет не только уменьшить сложность системы, но и улучшить её обобщающие свойства.

1. Yann LeCun, J. S. Denker, S. Solla, R. E. Howard and L. D. Jackel: Optimal Brain Damage, in Touretzky, David (Eds), Advances in Neural Information Processing Systems 2 (NIPS*89), Morgan Kaufman, Denver, CO, 1990

Ткач Кирилл Юрьевич, магистрант кафедры информационных технологий автоматизированных систем БГУИР, kiryltkach@hotmail.com.

Научный руководитель: Навроцкий Анатолий Александрович, заведующий кафедрой информационных технологий автоматизированных систем БГУИР, кандидат физико-математических наук, доцент, navrotsky@bsuir.by.