(cc) BY

Надёжкин Александр Игоревич
магистрант
ФГАОУ ВО «Южно-Уральский государственный университет» (НИУ)
г. Челябинск, Челябинская область
Мокеев Владимир Викторович
д-р техн. наук, старший научный сотрудник
Высшая школа экономики и управления
ФГАОУ ВО «Южно-Уральский государственный университет» (НИУ)
г. Челябинск, Челябинская область

НЕЙРОННЫЕ СЕТИ КАК ПРИОРИТЕТНОЕ НАПРАВЛЕНИЕ В РЕШЕНИИ ЗАДАЧ РАСПОЗНАВАНИЯ ОБРАЗОВ НА ИЗОБРАЖЕНИИ

Аннотация: *в данной статье проанализировано применение нейронной сети для распознавания образов на изображении. Рассмотрены основные особенности применения нейронных сетей для распознавания образов. Отмечены недостатки существующих сетей. Представлена структура свёрточной нейронной сети как наиболее устойчивой к искажениям изображений сети.*

Ключевые слова: *обработка изображений, распознавание образов, свёрточная нейронная сеть.*

Введение

В настоящее время нейронные сети успешно применяются в распознавании объектов на изображении. Они показывают гораздо лучший результат по сравнению с классическими методами распознавания. При обнаружении образа и его распознавании достаточно часто возникают трудности, связанные с большим разнообразием искажений, таких как различные сдвиги, повороты, освещенность и т. д. Преимущество нейронных сетей для распознавания образов в том, что с этими проблемами они справляются хорошо и в том, что они обучаемы для вы-

Content is licensed under the Creative Commons Attribution 4.0 license (CC-BY 4.0)

деления ключевых характеристик образов из обучающих выборок, что тем самым увеличивает точность распознавания по сравнению с остальными методами.

Существует около десятка разновидностей нейронных сетей, используемых в задачах распознавания и идентификации изображений. Наиболее известная нейронная сеть – это персептрон. Она позволяет классифицировать поданное на вход изображение в соответствии с предварительной настройкой и обучением сети. Классический персептрон имеет малое количество слоев и при расширении количества слоев обучение этой сети становится очень трудной задачей.

Многослойная сеть MADALINE обладает недостатком быстрого роста размера при попытке её обучения к большому числу инвариантных преобразований. Она состоит из множества пластин, состоящих из набора нейронов и классификаторов, которые могут быть инвариантны к сдвигу, повороту, изменению масштаба исходного изображения.

Неокогнитрон является многослойной сетью и состоит из двух типов слоёв: сложный (C-слой) и простой (S-слой). Простые клетки реагируют на входной вектор и имеют общие веса. Сложные клетки сжимают информацию, полученную от простых клеток. Преимущество такой сети в инвариантности к масштабированию и сдвигам. Недостатками является слишком большое количество настраиваемых параметров и очень медленная работа при последовательной реализации.

Как видно, главными недостатками существующих сетей являются большое количество настраиваемых параметров обучения сети и высокий рост размеров сети при увеличении слоев. Поэтому, в задачах распознавания образов на изображении применяется так называемая свёрточная нейронная сеть, которая лишена данных недостатков и является удачной серединой между биологически правдоподобными сетями и многослойным персептроном [1].

Описание архитектуры свёрточной нейронной сети

Свёрточная нейронная сеть относится к многослойным сетям и состоит из двух частей: выделитель признаков и классификатор (рисунок 1). Выделение

Содержимое доступно по лицензии Creative Commons Attribution 4.0 license (CC-BY 4.0)

признаков осуществляется частью сети, которая состоит из чередования слоёв свертки (С- слои) и слоёв усреднения (S- слои). Пластины с нейронами в каждом слое воспринимают информацию из предыдущего слоя посредством своего рецептивного поля. Рецептивное поле представляет собой квадратное окно с нейронами на предыдущем слое, которое связано с помощью настраиваемых параметров с нейроном в текущем слое. Классификатором обычно бывает двухслойный персептрон.

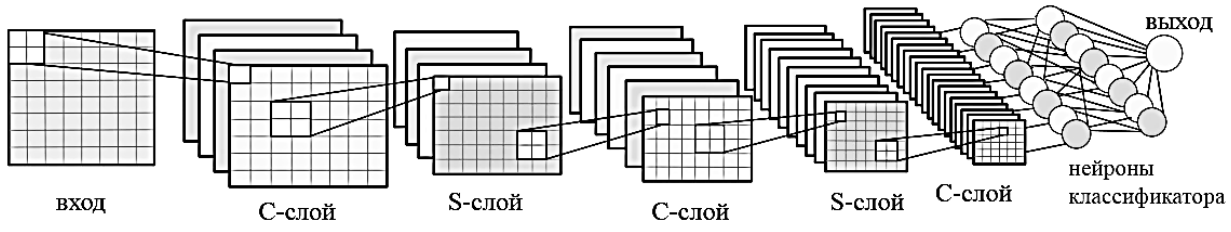

Рис. 1. Архитектура свёрточной нейронной сети

Особенность С-слоя заключается в том, что для всех нейронов используются одни и те же настраиваемые параметры (веса). Благодаря этому, такие веса, которые называются связанными, позволяют реагировать на некоторый признак в любом месте карты. Каждая карта С-слоя может быть связана с несколькими картами предыдущего слоя, что позволяет реагировать на информацию из разных модальностей. У рецептивных полей имеется шаг пересечения, чем меньше этот шаг, тем больше рецептивных полей будет наложено на карту и тем больше нейронов будет в следующем слое. Таким образом, выделение признаков происходит за счет настраиваемых общих весов и размера рецептивного поля.

Реакция нейрона определяется функцией активации. Функция активации берется одинаковой для всех элементов сети и представляет собой функцию, непрерывно дифференцируемую в каждой точке и симметричную относительно оси абсцисс (в основном, гиперболический тангенс). На вход этой функции подается свёртка входных данных и настраиваемых весов нейрона p. Наилучшей функцией активации, показывающей наименьшую ошибку обобщения

Content is licensed under the Creative Commons Attribution 4.0 license (CC-BY 4.0)

и высокую скорость сходимости ошибки обучения, является гиперболический тангенс вида [3]:

$$f(p) = 1.7159 * \tanh(\frac{2}{3} * p) \quad (1)$$

Основной задачей S-слоя является облегчение выделения общих признаков за счет уменьшения масштаба отфильтрованного отображения, получившегося на карте предыдущего C-слоя. Рецептивные поля нейронов S-слоя не пересекаются, и каждая карта S-слоя связана только с одной картой предыдущего C-слоя. Настраиваемые веса также одинаковые для всех нейронов.

В конечном итоге, карты слоев вырождаются в единичные нейроны, при этом потеря точности отображений компенсируется ростом количества карт и связанностью между ними. Единичные нейроны считаются входом для классификатора. Классификатор состоит из слоя обычных нейронов, где каждый нейрон связан со всеми нейронами входа. Количество нейронов в классификаторе соответствует количеству классов, к которым нужно отнести входное изображение. Параметры для каждого нейрона в классификаторе уникальные.

Обучение свёрточной нейронной сети происходит методом обратного распространения ошибки. Этот метод подразумевает корректировку весов за счет полученного вектора ошибки – разностью между полученными и известными данными. То есть, если нейронная сеть неправильно классифицирует поданное на вход изображение, синаптические веса обновляются в соответствии с ошибкой, и при следующем проходе того же изображения, сеть верно классифицирует объект.

Заключение

В последнее время наука значительно продвинулась в изучении искусственных нейронных сетей. Свёрточные нейронные сети показывают лучшие результаты, по сравнению с классическими статистическими методами (в среднем на 10%–15% [2]). Использование свёрточных нейронных сетей позволяет распозна-

Содержимое доступно по лицензии Creative Commons Attribution 4.0 license (CC-BY 4.0)

вать образы на изображениях с различными искажениями, сдвигами, изменениями масштаба, а обучение таких сетей отличается простотой, благодаря связанности весов.

Список литературы

1. Face Recognition: A Convolutional Neural-Network Approach / S. Lawrence, C. Lee Giles, Ah Chung Tsoi, A.D. Back // IEEE transactions on neural networks. – Vol. 8. – 1997. – №1.

2. Назаров Л.Е. Нейросетевые алгоритмы обнаружения, классификации и распознавания объектов на изображениях [Текст] / Л.Е. Назаров, Н.С. Томашевич, А.Н. Балухто // Нейрокомпьютеры в прикладных задачах обработки изображений. – Кн. 25. – 2007. – С. 25–54.

3. Немков Р.М. Экспериментальное исследование и анализ влияния базовых параметров сверточных нейронных сетей на качество их обучения / Р.М. Немков, О.С. Мезенцева // Вестник Северо-Кавказского федерального университета. – Ставрополь, 2013. – №3 (36). – С. 21–26.

Content is licensed under the Creative Commons Attribution 4.0 license (CC-BY 4.0)