ПРИМЕНЕНИЕ НЕЙРОСЕТЕЙ

В РАСПОЗНАВАНИИ ИЗОБРАЖЕНИЙ

Аннотация: в статье рассматривается применение нейронных сетей в задачах распознавания

изображений.

Ключевые слова: нейросети, нейронные сети, распознавание образов.

Среди методов распознавания, можно выделить широкий класс методов, методы,

заслуживающие отдельного рассмотрения нейросетевые методы. В их основе лежат

нейронные сети вычислительные модели, принцип функционирования которых сходен

с сетями биологических нейронов головного мозга. Благодаря заимствованию

принципов организации у биологических структур мозга, нейросети демонстрируют

многие их свойства, такие как обучение на основе предыдущего опыта, извлечение

существенных свойств из поступающей информации, обобщение имеющихся

прецедентов на новые случаи.

Возможности, предоставляемые нейронными сетями, были использованы для

решения задач распознавания и классификации образов во множестве исследований и

прикладных разработок. В данной главе рассматриваются некоторые основные

разновидности нейронных сетей, их возможности и способы обучения.

«Обычные» нейросети

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

53

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

Международный научный журнал «ВЕСТНИК НАУКИ» № 7 (16) Т.3…..……………………….. ИЮЛЬ 2019 г.

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

Рисунок 1 Нейросеть

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. Будем

называть такую сеть сокращенно ПНС. Сперва необходимо решить, как подавать

данные на вход. Самое простое и почти безальтернативное решение для ПНС — это

выразить двумерную матрицу изображения в виде одномерного вектора. Дальше

происходит выбор архитектуры. До сих пор не существует методов, позволяющих

однозначно определить структуру и состав нейросети исходя из описания задачи. Для

трудно формализуемых задач вряд ли когда-либо такой метод будет создан. Кроме

того, существует множество различных методик редукции сети (например OBD), а

также разные эвристики и эмпирические правила. Одно из таких правил гласит, что

количество нейронов в скрытом слое должно быть хотя бы на порядок больше

количества входов.

Но вот если задуматься, на ум приходит что, когда мы преобразуем изображение в

линейную цепочку байт, мы что-то безвозвратно теряем. Причем с каждым слоем эта

потеря только усугубляется. Так и есть — мы теряем топологию изображения, т.е.

взаимосвязь между отдельными его частями. Кроме того, задача распознавания

подразумевает умение нейросети быть устойчивой к небольшим сдвигам, поворотам и

изменению масштаба изображения, т.е. она должна извлекать из данных некие

инварианты, не зависящие от почерка того или иного человека. Так какой же должна

быть нейросеть, чтобы быть не очень вычислительно сложной и, в тоже время, более

инвариантной к различным искажениям изображений?

Сверточные нейронные сети

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

54

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

Международный научный журнал «ВЕСТНИК НАУКИ» № 7 (16) Т.3…..……………………….. ИЮЛЬ 2019 г.

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

Решение этой проблемы было найдено американским ученым французского

происхождения Яном ЛеКуном, вдохновленным работами нобелевских лауреатов в

области медицины Torsten Nils Wiesel и David H. Hubel. Эти ученые исследовали

зрительную кору головного мозга кошки и обнаружили, что существуют так

называемые простые клетки, которые особо сильно реагируют на прямые линии под

разными углами и сложные клетки, которые реагирую на движение линий в одном

направлении. Ян ЛеКун предложил использовать так называемые сверточные

нейронные сети. Идея сверточных нейронных сетей заключается в чередовании

сверточных слоев (C-layers), субдискретизирующих слоев (S-layers) и наличии

полносвязных (F-layers) слоев на выходе.

Рисунок 2 Архитектура

Такая архитектура заключает в себе 3 основных парадигмы:

1. Локальное восприятие.

2. Разделяемые веса.

3. 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 32х32 пикселя, то каждый из нейронов следующего слоя

примет на вход только небольшой участок этого изображения размером, к примеру,

5х5, 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

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

55

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

Международный научный журнал «ВЕСТНИК НАУКИ» № 7 (16) Т.3…..……………………….. ИЮЛЬ 2019 г.

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

применен ко всему изображению. Такие наборы часто называют ядрами. Нетрудно

посчитать, что даже для 10 ядер размером 5х5 для входного изображения размерами

32х32 количество связей окажется равным примерно 256000 (сравниваем с 10 млн.), а

количество настраиваемых параметров всего 250. А как же это скажется на качестве

распознавания? В лучшую сторону. Дело в том, что такое искусственно введенное

ограничение на веса улучшает обобщающие свойства сети, что в итоге позитивно

сказывается на способности сети находить инварианты в изображении и реагировать

главным образом на них, не обращая внимания на прочий шум. Можно посмотреть на

этот подход немного, с другой стороны. Те, кто занимался классикой распознавания

изображений и знает, как это работает на практике (например, в военной технике)

знают, что большинство таких систем строятся на основе двумерных фильтров.

Фильтр представляет собой матрицу коэффициентов, обычно заданную вручную. Эта

матрица применяется к изображению с помощью математической операции,

называемой сверткой. Суть этой операции в том, что каждый фрагмент изображения

умножается на матрицу (ядро) свертки поэлементно и результат суммируется и

записывается в аналогичную позицию выходного изображения. Основное свойство

таких фильтров заключается в том, что значение их выхода тем больше, чем больше

фрагмент изображения похож на сам фильтр. Таким образом, изображение, свернутое

с неким ядром, даст нам другое изображение, каждый пиксел которого будет означать

степень похожести фрагмента изображения на фильтр. Иными словами, это будет

карта признаков.

Каждый фрагмент изображения поэлементно умножается на небольшую матрицу

весов (ядро), результат суммируется. Эта сумма является пикселом выходного

изображения, которое называется картой признаков. Здесь я опустил тот факт, что

взвешенная сумма входов еще пропускается через функцию активации (как в любой

другой нейросети). На самом деле это может происходить и в S-слое, принципиальной

разницы нет. Следует сказать, что в идеале не разные фрагменты проходят

последовательно через ядро, а параллельно все изображение проходит через

идентичные ядра. Кроме того, количество ядер (наборов весов) определяется

разработчиком и зависит от того какое количество признаков необходимо выделить.

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

56

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

Международный научный журнал «ВЕСТНИК НАУКИ» № 7 (16) Т.3…..……………………….. ИЮЛЬ 2019 г.

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

Еще одна особенность сверточного слоя в том, что он немного уменьшает

изображение за счет краевых эффектов. Суть субдискретизации и S-слоев заключается

в уменьшении пространственной размерности изображения. Т.е. входное изображение

грубо (усреднением) уменьшается в заданное количество раз. Чаще всего в 2 раза, хотя

может быть и не равномерное изменение, например, 2 по вертикали и 3 по

горизонтали. Субдискретизация нужна для обеспечения инвариантности к масштабу.

Чередование слоев позволяет составлять карты признаков из карт признаков, что

на практике означает способность распознавания сложных иерархий признаков.

Обычно после прохождения нескольких слоев карта признаков вырождается в вектор

или даже скаляр, но таких карт признаков становится сотни. В таком виде они подаются

на один-два слоя полносвязной сети. Выходной слой такой сети может иметь различные

функции активации. В простейшем случае это может быть тангенциальная функция,

также успешно используются радиальные базисные функции.