

ОБЗОР ПРИМЕНЕНИЯ НЕЙРОННЫХ СЕТЕЙ

Свито И.Л., Свито А.И.

Белорусский государственный университет
информатики и радиоэлектроники

Abstract

This article presents a review of neural network usage in various areas of human activities. The article describes various areas of neural network applications and achievements. A particular focus is placed on the usage of these technologies in the field of the object segmentation image, semantic text analysis, etc.

Key words: *Нейронные сети, рекуррентные нейронные сети (Recurrent neural network (RNN). Перцептрон, сегментация объектов на изображениях; Семантический анализ текста; распознавание речи.*

Введение

Классификация нейронных сетей. Нейронные сети – это своеобразная имитация некоторых видов мозговой деятельности. С их помощью можно решать так называемые «нечеткие задачи» – распознавания образов, речи, рукописного текста, выявления закономерностей, классификации, прогнозирования и т.п. [1-6]. Можно провести следующую классификацию нейронных сетей:

Нейронные сети классифицируются по следующим видам обучения:

1. нейронные сети проходящие обучение с учителем;
2. нейронные сети проходящие обучение без учителя.

Нейронные сети так же делятся на следующие группы с учетом настройки весовых коэффициентов. С **фиксированными связями**, веса которых выбираются заранее исходя из задачи и с **динамическими связями**, которые перестраивают свои веса в процессе обучения.

По типу входных данные сети так же делятся на сети с **аналоговыми входными данными**, которые представлены в виде действительных чисел и **двоичные**, информация которых представляется в виде нулей и единиц. На данный момент чаще всего используются следующие модели нейронной сети:



Рисунок 1. Классификация нейронных сетей

1. **Сети прямого распространения** – все связи этой сети имеют строгое направление от входных нейронов к их выходам. Среди таких сетей хочется отметить: **простейший персептрон**, автором которого является **Фрэнк Розенблатт** и **многослойный персептрон**.

2. **Рекуррентные нейронные сети** – данные с выходных нейронов или из скрытого слоя передаются частично обратно на входные нейроны.

3. **Радиально-базисные функции** – это нейронная сеть, в основе которой является наличие скрытого слоя из радиальных элементов и выходного слоя из линейных элементов. Такие сети довольно компактны и обучаются достаточно быстро. **Радиально-базисная сеть** пользуется следующими уникальными свойствами: один скрытый слой, нейроны только скрытого слоя имеют нелинейную функцию активации и синаптические веса скрытого и входного слоев являются единицей.

4. **Самоорганизующиеся карты** или **Сети Кохонена** – этот класс сетей обычно обучается без помощи учителя и часто применяется в задачах связанных с распознаванием изображений. Такие сети способны определять новые элементы во входных данных. **Сеть Кохонена** имеет всего два слоя: выходной и входной, составленный из радиальных элементов.

Что такое рекуррентные нейронные сети? Рекуррентные нейронные сети (Recurrent neural network; RNN) (рис. 2) – это сети, где связи меж-

ду элементами образуют направленную последовательность. Благодаря этому появляется возможность обрабатывать серии событий во времени или последовательные пространственные цепочки. В отличие от многослойных перцептронов, рекуррентные сети могут использовать свою внутреннюю память для обработки последовательностей произвольной длины. Поэтому сети RNN применимы в таких задачах, как [распознавание рукописного текста](#) или [распознавание речи](#). В последнее время наибольшее распространение получили [сеть с долговременной и кратковременной памятью](#) (LSTM). Вероятно, первой RNN была сеть Хопфилда (впервые упомянута в 1974 году), которая реализовывала на практике ячейку ассоциативной памяти. От современных RNN она отличается тем, что работает с последовательностями фиксированного размера.

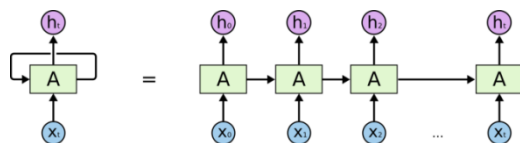


Рисунок 2. Рекуррентные сети

Как работает *нейронная сеть*? Простые нейронные сети получили большую популярность в последние десятилетия. Их математическая модель успешно строит аналогии между работой нейрона и перцептрона (см. рис. 3) – компьютерной интерпретации нейрона и позволяет создавать сложные системы по распознаванию сигналов, звука, речи и изображений. Это позволило использовать их в системах распознавания почерка при анализе чеков в банковских и платёжных системах и даже в качестве распознавателя рукописных адресов.

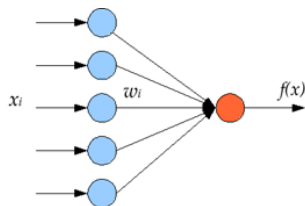


Рисунок 3. Структура перцептрона в нейронной сети

Несмотря на схожесть в работе с человеческим мозгом, нейронные сети в их общем виде по-прежнему не способны решать некоторые специфические задачи анализа и распознавания, в частности, задачи нахождения темы и тональности текста.

Как видно, перцептрон принимает несколько входных сигналов (битов информации, пикселей), свой вес и при помощи функции активации преобразует все входные данные в один ответ, который подаётся на выход либо передаётся другим перцептронам. Данная модель отлично работает, когда нам необходимо оценить какую-либо вероятность (например, принадлежность рукописного символа определённому алфавиту или вероятность того, что заёмщик вернёт выданный ему кредит), однако не эффективна, например, при семантическом анализе текста. Связано это с тем, что данная модель нейронной сети не обладает памятью.

Применение современных нейронных сетей для сегментации объектов на изображениях. Технологии компьютерного зрения применяются для распознавания лиц, пешеходов, объектов, для медицинского анализа, навигации автономных автомобилей и в других сферах. В задаче распознавания образов особого успеха достигли сверточные нейронные сети (Convolutional Neural Networks; CNN). С появлением больших баз изображений стало возможным обучать нейронные сети с большим числом скрытых слоёв.

Классических задач для CNN не так много. Их условно можно разделить на 3 уровня: от самых низкоуровневых задач до самых высокоуровневых задач:

- **Определение границ** – это самая низкоуровневая задача, для которой уже классически применяются сверточные нейронные сети.
- **Определение вектора к нормали** позволяет нам реконструировать трёхмерное изображение из двухмерного.
- **Saliency, определение объектов внимания** – это то, на что обратил бы внимание человек при рассмотрении этой картинки.
- **Семантическая сегментация** позволяет разделить объекты на классы по их структуре, ничего не зная об этих объектах, то есть еще до их распознавания.
- **Семантическое выделение границ** – это выделение границ, разбитых на классы.
- **Выделение частей тела человека.**

И самая высокоуровневая задача – **распознавание самих объектов**.

Можно привести примеры нескольких нейронных сетей для задачи классификации изображений, таких как AlexNet, VGG16, ResNet, CycleGAN, UNet, Mask RCNN.

AlexNet – сверточная нейронная сеть, которая оказала большое влияние на развитие машинного обучения, в особенности – на алгоритмы компьютерного зрения. Сеть с большим отрывом выиграла конкурс по распознаванию изображений ImageNet ILSVRC 2012. AlexNet содержит восемь слоев с весовыми коэффициентами. Первые пять из них сверточные, а остальные три – полносвязные. Результаты показывают, что большая, глубокая сверточная [нейронная сеть](#) способна достигать рекордных результатов на очень сложных датасетах, используя только [обучение с учителем](#). AlexNet была первой реализацией сверточных нейронных сетей и открыла новую эру исследований.

VGG16 – модель сверточной нейронной сети, предложенная К. Simonyan и А. Zisserman из Оксфордского университета в статье “Very Deep Convolutional Networks for LargeScale Image Recognition”. Модель достигает точности 92,7% при тестировании на ImageNet в задаче распознавания объектов на изображении. Этот датасет состоит из более чем 14 миллионов изображений. VGG16 – одна из самых знаменитых моделей, отправленных на соревнование ILSVRC2014. Она является улучшенной версией AlexNet, в которой заменены большие фильтры (размера 11 и 5 в первом и втором сверточном слое, соответственно) на несколько фильтров размера 3x3, следующих один за другим. К сожалению, сеть VGG имеет два серьезных недостатка: 1) очень медленная скорость обучения; 2) сама архитектура сети весит слишком много (появляются проблемы с диском и пропускной способностью). Несмотря на недостатки, данная сеть является отличным строительным блоком для обучения, так как её легко реализовать.

ResNet – это сокращенное название для Residual Network (дословно – «остаточная сеть»), но что такое residual learning («остаточное обучение»)? Глубокие сверточные нейронные сети превзошли человеческий уровень классификации изображений в 2015 году. Когда более глубокая сеть начинает сворачиваться, возникает проблема: с увеличением глубины сети точность сначала увеличивается, а затем быстро ухудшается. Снижение точности обучения показывает, что не все сети легко оптимизировать. Чтобы преодолеть эту проблему, Microsoft ввела глубокую «остаточную» структуру обучения. Модель ResNet имеет меньше фильтров и сложность меньше, чем сети VGG. Сеть ResNet сходится быстрее, чем ее простой аналог, при этом более глубокие ResNet дости-

гают лучших результатов обучения по сравнению с неглубокой сетью.

CycleGAN – это тип генеративно состязательной сети, используемой для переноса стиля изображения. CycleGAN можно обучить конвертировать изображения из одного домена (например, Fortnite) в другой, например, PUBG.

UNet считается одной из стандартных архитектур CNN для задач сегментации изображений, когда нужно не только определить класс изображения целиком, но и сегментировать его области по классу, т. е. создать маску. Архитектура сети представляет собой последовательность слоёв, которые сначала уменьшают пространственное разрешение картинки, а потом увеличивают его, предварительно объединив с данными картинки и пропустив через другие слои свёртки. Таким образом, сеть выполняет роль своеобразного фильтра. Архитектура свёрточной сети UNet может использоваться не только для сегментации, но и для детектирования объектов на изображениях.

Mask RCNN – сеть с такой архитектурой позволяет выделять на фотографиях контуры (маски) экземпляров разных объектов, даже если таких экземпляров несколько, они имеют различный размер и частично перекрываются. Сеть так же способна к распознаванию поз людей на изображении. Mask RCNN развивает архитектуру Faster RCNN путём добавления ещё одной ветки, которая предсказывает положение маски, покрывающей найденный объект, и, таким образом решает уже задачу instance segmentation.

Таким образом, сегментация объектов одна из перспективных сфер для развития нейронных сетей. Данную технологию можно использовать на дорогах, в аэропортах, на улицах и т. д.

Применение RNN при семантическом анализе текста и распознавании речи. На примере распознавания темы текста это можно выразить в следующем: читая параграф, каждое новое прочитанное слово мы воспринимаем исходя из контекста всех ранее прочитанных слов. Для того, чтобы успешно выделять основную тему текста или определять его интонацию, нейронной сети необходимо определённым образом запоминать результаты анализа предыдущих слов и предложений, чтобы иметь контекст, на который можно опираться. Таким образом, появились модификации нейронных сетей под общим названием рекуррентные нейронные сети (RNN – Recurrent Neural Networks). В частности для анализа текста используется разно-видность рекуррентной нейронной сети – сеть долгой краткосрочной памяти (LSTM – Long Short Term Memory). LSTM блоки могут обладать большим числом вентиляей, однако общая структура перцептрона в общих чертах остаётся такой же.

Помимо обычного линейного входа и выхода рекуррентные сети также обладают возможностью «забывать» ненужную информацию, что улучшает результаты анализа.

Каждый перцептрон в данной сети обладает выходом на самого себя, что при развёртывании создаёт цепь нейронов, запоминающих результаты анализа предыдущих слов. Формально, для LSTM, это даёт возможность нейронной сети обратиться к себе самой и получить информацию о смысловом и тональном содержании предшествующего текста в анализируемом корпусе.

LSTM архитектура нейронных сетей позволяет запоминать очень большие контексты информации и делать выводы, основываясь на словах, употреблённых в самом начале корпуса. Именно их эффективность и близость к алгоритмам распознавания текста человеческим мозгом сделала их основным средством семантического анализа и анализа тональности текстов. Сейчас RNN сети активно используются для классификации корпуса текстов по темам. Также сети способны определить эмоциональный тон текста, его лексическую наполненность или даже количество полезной информации в тексте, обучаясь как с учителем, так и без.

Применение современной нейронной сети для машинного перевода. Машинный перевод похож на языковое моделирование поскольку вектор входных параметров представляет собой последовательность слов на исходном языке (например, на польском). Мы хотим получить последовательность слов на целевом языке (например, на русском). Ключевое различие заключается в том, что мы получим эту последовательность только после того, как увидим все входные параметры, поскольку первое слово переводимого предложения может потребовать информации всей последовательности вводимых слов.

Нестандартные применения нейронных сетей. Ученые Стэнфорда недавно придумали очень необычное применение нейронной сети CNN для предсказания бедности. В Африке уровень бедности зашкаливает за все мыслимые и немыслимые пределы. Нет возможности собирать социальные демографические данные, а с 2005 года вообще нет никаких данных о том, что там происходит. Учёные собирали дневные и ночные карты со спутников и скармливали их нейронной сети в течение некоторого времени. Нейронная сеть была настроена на ImageNet, то есть первые слои фильтров были настроены так, чтобы она умела распознавать уже какие-то совсем простые вещи, например, крыши домов, для поиска поселения на дневных картах. Затем дневные карты были сопоставлены с картами ночной освещённости того же

участка поверхности для того, чтобы сказать, насколько есть деньги у населения, чтобы хотя бы освещать свои дома в течение ночного времени. Нейронная сеть составила достаточно точный прогноз, даже с небольшим сдвигом с 2005 года.

Применение нейронной сети для семантической сегментации 3D изображений в медицине. Это область, с которой очень сложно работать и она активно развивается. Для этого есть несколько причин:

- Очень мало баз данных. Даже если есть картинка поврежденного мозга, нужно взять медика и заставить его вручную размещать все многослойные изображения, что очень долго и крайне неэффективно.
- Нужна очень высокая точность и надежность системы. Если мы не распознали опухоль, то это уже не очень хорошо.
- Изображения в трехмерных элементах – вокселях, не в пикселях, что доставляет дополнительные сложности разработчикам систем.

Здесь нейронная сеть может применяться для определения повреждений после удара, для поиска опухоли в мозгу, в кардиологии для определения того, как работает сердце.

Заключение. Существует еще множество областей, где можно применять нейронные сети. Появляются все новые стартапы в ближайшее время. Для многих задач, таких как распознавание образов, пока не создано доминирующих подходов. Между тем нейронные сети уверенно продолжают проникать в нашу жизнь. Множество надежд в отношении нейронных сетей сегодня связывают именно с аппаратными реализациями. На их разработку тратится значительное время, за которое программные реализации на самых последних компьютерах оказываются лишь на порядок менее производительными, что делает использование нейронных процессоров нерентабельным. Но все это только вопрос времени – нейронным сетям предстоит пройти тот же путь, по которому еще совсем недавно развивались компьютеры, увеличивая свои возможности и производительность, захватывая новые сферы применения по мере возникновения новых задач и развития технической основы для их разработки.

Литература

1. Graves, A. and Schmidhuber, J. (2005). Framewise phoneme classification with bidirectional LSTM and other neural network architectures. *Neural Networks*, 18(56):602–610.
2. Graves, A. and Schmidhuber, J. (2009). Offline handwriting recognition with multidimensional recurrent neural networks. In Koller, D., Schuurmans, D., Bengio, Y., and Bottou, L., editors, *Advances in Neural Information Processing Systems 21*, pages 545–552.
3. Hochreiter, S. and Schmidhuber, J. (1997). Long shortterm memory. *Neural Computation*, 9(8):1735–1780.
4. Jarrett, K., Kavukcuoglu, K., Ranzato, M., and LeCun, Y. (2009). What is the best multistage architecture for object recognition? In *Computer Vision, 2009 IEEE 12th International Conference on*, pages 2146–2153. IEEE.
5. Kearns, M. and Vazirani, U. (1994). *An introduction to computational learning theory*. MIT Press.