

УДК 004.93'1/.8

## ПРИМЕНЕНИЕ НЕЙРОННОЙ СЕТИ ДЛЯ СОЗДАНИЯ СИСТЕМЫ РАСПОЗНАВАНИЯ ИЗОБРАЖЕНИЙ

<sup>1</sup>Маркова С.В., <sup>2,3</sup>Жигалов К.Ю.

<sup>1</sup>Московский университет МВД им. В.Я. Кикотя, Москва;

<sup>2</sup>Институт проблем управления Российской академии наук, Москва;

<sup>3</sup>НОУ ВО Московский технологический институт, Москва, e-mail: kshakalov@mail.ru

Нейронная сеть представляет собой математическую модель и ее воплощение в виде программной либо программно-аппаратной реализации, которая выполнена на основе моделирования деятельности биологических нейронных сетей, которыми являются сети нейронов в биологическом организме. Научный интерес к данной структуре возник потому, что исследование ее модели позволяет получать информацию о некоторой системе. То есть подобная модель может иметь практическую реализацию в ряде отраслей современной науки и техники. В статье рассматриваются вопросы, затрагивающие тему применения нейронных сетей для построения систем идентификации изображений, имеющих широкое применение в системах обеспечения безопасности. Подробно исследованы вопросы, связанные с темой алгоритма распознавания изображения и его применения. Кратко приведена информация о методике обучения нейронных сетей.

**Ключевые слова:** нейронные сети, обучение с помощью нейронных сетей, распознавание изображений, парадигма локального восприятия, системы безопасности

## NEURAL NETWORK APPLICATION FOR PURPOSES OF RECOGNITION OF IMAGES

<sup>1</sup>Markova S.V., <sup>2,3</sup>Zhigalov K.Yu.

<sup>1</sup>Moscow University of the Ministry of the Interior of Russia, Moscow;

<sup>2</sup>V.A. Trapeznikov Institute of Control Sciences of Russian Academy of Sciences, Moscow;

<sup>3</sup>Moscow Technological Institute, Moscow, e-mail: kshakalov@mail.ru

A neural network is a mathematical model and its embodiment in the form of software or software and hardware implementation, which is based on modeling the activities of biological neural networks, which are networks of neurons in a biological organism. The scientific interest in this structure has arisen from the fact that the study of the model allows to obtain information about a system. That is, such a model can have practical implementation in several branches of modern science and technology. In the article the questions touching on a theme of application of neural networks for construction of systems of identification of the images, having wide application in systems of maintenance of safety are considered. The questions connected with the theme of the algorithm of image recognition and its application are studied in detail. Brief information is given on the method of training neural networks.

**Keywords:** neural networks, training by using of neural networks, image understanding, a paradigm of local perception, a security arrangement

На сегодняшний день технологический и научно-исследовательский прогресс охватывает все новые горизонты, стремительно прогрессируя. Один из них – это моделирование окружающего естественного мира при помощи математических алгоритмов. В данном аспекте существуют тривиальные, например моделирование колебаний моря, и крайне сложные, нетривиальные, многокомпонентные задачи, например моделирование функционирования человеческого мозга. В процессе исследования данного вопроса было выделено отдельное понятие – нейронная сеть. Нейронная сеть представляет собой математическую модель и ее воплощение в виде программной либо программно-аппаратной реализации, которая выполнена на основе моделирования деятельности биологических нейронных сетей, которыми являются сети нейронов в биологическом организме. Научный интерес к данной структуре возник потому,

что исследование ее модели позволяет получать информацию о некоторой системе. То есть подобная модель может иметь практическую реализацию в ряде отраслей современной науки и техники.

### Краткая история развития нейронных сетей

Стоит отметить, что изначально понятие «нейронная сеть» берет свое начало в работе американских математиков, нейролингвистов и нейропсихологов У. Маккалока и У. Питтса (1943 г.), где авторы впервые упоминают о ней, дают ей определение и производят первую попытку построения модели нейронной сети [1]. Уже в 1949 г. Д. Хебб предлагает первый алгоритм обучения. Далее был ряд исследований в области нейронного обучения, и первые рабочие прототипы появились примерно в 1990–1991 гг. прошлого столетия. Тем не менее вычислительных мощностей оборудования

того времени не хватало для достаточно быстрой работы нейронных сетей. К 2010 году мощности GPU видеокарт сильно увеличились и появилось понятие программирования непосредственно на видеокартах, что существенным образом (в 3–4 раза) увеличило производительность компьютеров. В 2012 г. нейросети впервые победили на чемпионате ImageNet, чем и ознаменовалось их дальнейшее бурное развитие [2, 3] и появление термина Deep Learning.

В современном мире нейронные сети имеют колоссальный охват, ученые считают исследования, проводимые в области изучения поведенческих особенностей и состояний нейронных сетей, крайне перспективными. Перечень областей, в которых нейронным сетям нашлось применение, огромен. Это и распознавание и классификация образов, и прогнозирование, и решение аппроксимационных задач, и некоторые аспекты сжатия данных, анализа данных и, конечно, применение в системах безопасности различного характера.

Исследование нейронных сетей сегодня активно происходит в научных сообществах разных стран. При подобном рассмотрении она представлена в качестве частного случая ряда методов распознавания образов, дискриминантного анализа, а также методов кластеризации.

Следует также отметить, что в течение последнего года на стартапы в области систем распознавания изображений было выделено финансирование более чем за предыдущие 5 лет [4], что говорит о достаточно большой востребованности данного типа разработок на конечном рынке.

### Применение нейронных сетей для распознавания изображений

Рассмотрим стандартные задачи, решаемые нейронными сетями в приложении к изображениям [4]:

- идентификация объектов;
- распознавание частей объектов (например, лиц, рук, ног и т.д.);
- семантическое определение границ объектов (позволяет оставлять только границы объектов на картинке);
- семантическая сегментация (позволяет разделять изображение на различные отдельные объекты);
- выделение нормалей к поверхности (позволяет преобразовывать двумерные картинки в трехмерные изображения);
- выделение объектов внимания (позволяет определять то, на что обратил бы внимание человек на данном изображении).

Стоит отметить, что задача распознавания изображений имеет яркий характер,

решение данной задачи является сложным и неординарным процессом. При выполнении распознавания в качестве объекта может быть человеческое лицо, рукописная цифра, а также множество других объектов, которые характеризуются рядом уникальных признаков, что существенно осложняет процесс идентификации.

В настоящем исследовании будет рассмотрен алгоритм создания и обучения распознаванию рукописных символов нейронной сети. Изображение будет считано одним из входов нейронной сети, а один из выходов будет задействован для вывода результата.

На данном этапе необходимо кратко остановиться на классификации нейронных сетей. На сегодняшний день основных видов три [5]:

- сверточные нейронные сети (CNN);
- рекуррентные сети (deep learning);
- обучение с подкреплением.

Один из наиболее частых примеров построения нейронной сети – это классическая топология нейронной сети. Такая нейронная сеть может быть представлена в качестве полносвязного графа, характерной ее чертой является прямое распространение информации и обратное распространение сигнализации об ошибке. Данная технология не обладает рекурсивными свойствами. Иллюстративно нейронную сеть с классической топологией можно изобразить на рис. 1.

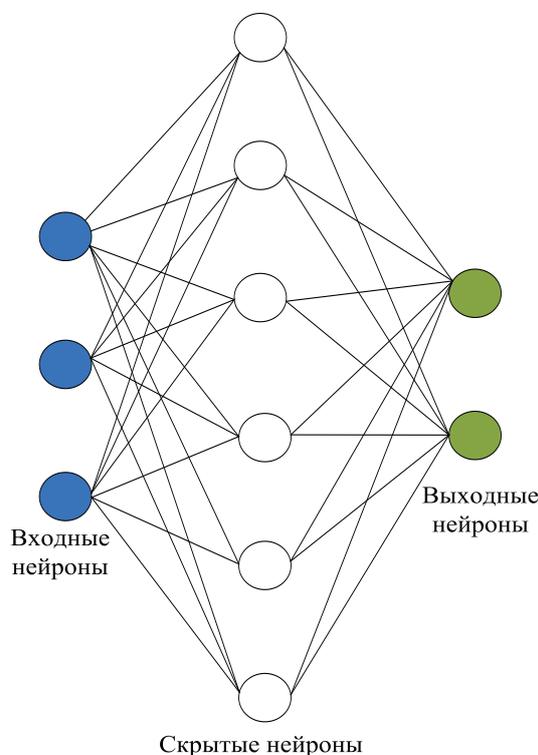


Рис. 1. Нейронная сеть с простейшей топологией

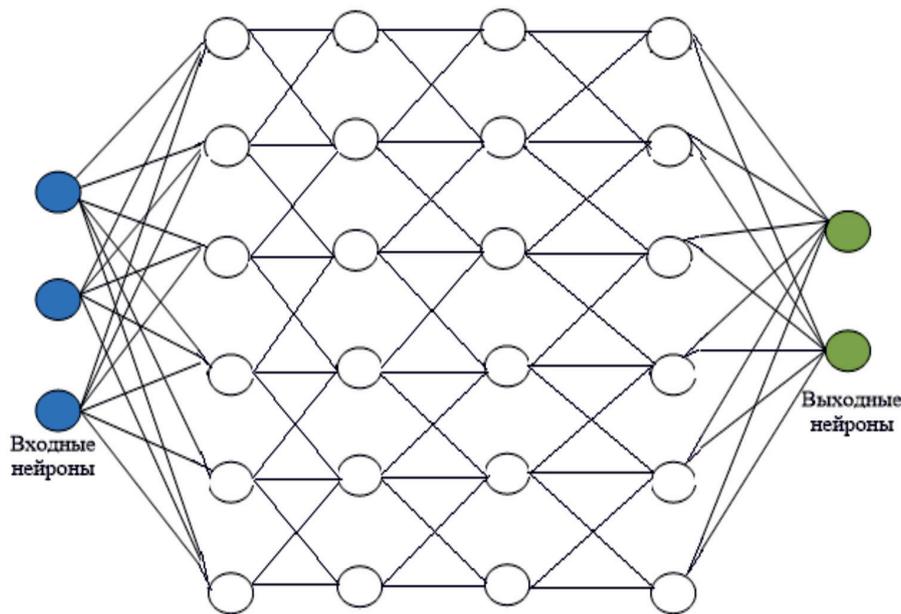


Рис. 2. Нейронная сеть с 4-мя слоями скрытых нейронов

Один из явно существенных минусов данной топологии сети – избыточность. За счет избыточности при подаче данных в виде, например, двумерной матрицы на вход можно получить одномерный вектор. Так, для изображения рукописной латинской буквы, описанной при помощи матрицы размером  $34 \times 34$ , потребуется 1156 входов. Это говорит о том, что вычислительные мощности, затрачиваемые на реализацию программно-аппаратного решения данного алгоритма окажутся слишком большими.

Проблема была решена американским ученым Яном Ле Куном, который провел анализ работ лауреатов Нобелевской премии в области медицины Т. Wtesel и D. Hubel [1, 5, 6]. В рамках исследования, проведенного ими, в качестве объекта исследования выступила зрительная кора головного мозга кошки. Анализ результатов показал, что в коре присутствует ряд простых клеток, а также ряд сложных клеток. Простые клетки реагировали на полученное от зрительных рецепторов изображение прямых линий, а сложные – на поступательное движение в одном направлении. В результате был выработан принцип построения нейронных сетей, называемый сверточным. Идея данного принципа заключалась в том, что для реализации функционирования нейронной сети используется чередование сверточных слоев, которые принято обозначать *C – Layers*, субдискретизирующих слоев *S – Layers* и полносвязных слоев *F – Layers* на выходе из нейронной сети.

В основе построения сети подобного рода лежат три парадигмы – это парадигма локального восприятия, парадигма разделяемых весов и парадигма субдискретизации.

Суть парадигмы локального восприятия заключается в том, что на каждый входной нейрон подается не вся матрица изображения, а ее часть. Остальные части подаются на другие входные нейроны. В данном случае можно наблюдать механизм распараллеливания, при помощи подобного метода можно сохранять топологию изображения от слоя к слою, многомерно обрабатывая его, то есть в процессе обработки может использоваться некоторое множество нейронных сетей.

Парадигма разделяемых весов говорит о том, что для множества связей может быть использован небольшой набор весов. Данные наборы имеют также название «ядра». Для конечного результата обработки изображений можно сказать, что разделяемые веса положительно влияют на свойства нейронной сети, при исследовании поведения которой повышается способность нахождения инвариантов в изображениях и фильтровать шумовые компоненты, не производя их обработку.

Исходя из вышеизложенного можно сделать вывод о том, что при применении процедуры свертывания изображения на базе ядра явится выходное изображение, элементы которого будут главной характеристикой степени соответствия фильтру, то есть произойдет генерация карты признаков. Данный алгоритм приведен на рис. 3.

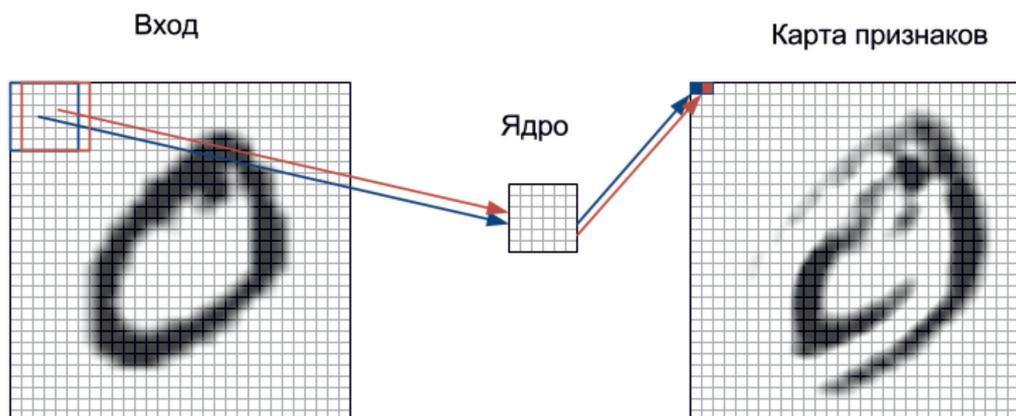


Рис. 3. Алгоритм генерации карты признаков

Парадигма субдискретизации состоит в том, что происходит уменьшение входного изображения за счет уменьшения пространственной размерности его математического эквивалента –  $n$ -мерной матрицы. Необходимость субдискретизации выражается в инвариантности к масштабу исходного изображения. При применении методики чередования слоев появляется возможность генерации новых карт признаков из уже имеющихся, то есть практическая реализация данного метода заключается в том, что будет приобретена способность вырождения многомерной матрицы в векторную, а затем и вовсе в скалярную величину.

#### Реализация обучения нейронной сети

Существующие сети делятся на 3 класса архитектур с точки зрения обучения:

- обучение с учителем (перцептон);
- обучение без учителя (сети адаптивного резонанса);
- смешанное обучение (сети радиально-базисных функций).

Один из наиболее важных критериев оценки работы нейронной сети в случае распознавания изображения – это качество распознавания изображений. Стоит отметить, что для количественной оценки качества распознавания изображения при помощи функционирования нейронной сети чаще всего применяется алгоритм среднеквадратической ошибки [6]:

$$E^p = \frac{1}{2} (D^p - O(I^p, W))^2 \quad (1)$$

В данной зависимости  $E^p$  –  $p$ -ая ошибка распознавания для пары нейронов,

$D^p$  – ожидаемый выходной результат нейронной сети (обычно сеть должна стремиться к распознаванию на 100%, но

этого на практике пока не происходит), а конструкция  $O(I^p, W)$  – квадрат выхода сети, который зависит от  $p$ -ого входа и набора коэффициентов веса  $W$ . В данную конструкцию входят и ядра свертки, и весовые коэффициенты всех слоев. Просчет ошибки заключается в вычислении среднего арифметического значения для всех пар нейронов.

В результате анализа была выведена закономерность, что номинальное значение веса, когда значение ошибки будет минимальным, можно рассчитать исходя из зависимости (2):

$$W_{\min} = W_c - \left( \frac{d^2 E(W_c)}{dW^2} \right)^{-1} \frac{dE(W_c)}{dW} \quad (2)$$

Из данной зависимости можно сказать, что задача вычисления оптимального веса есть арифметическая разность производной функции первого порядка ошибки по весу, разделенная на производную функции ошибки второго порядка [5].

Приведенные зависимости дают возможность тривиального вычисления ошибки, которая находится в выходном слое. Вычисление ошибки в скрытых слоях нейронов можно реализовать при помощи метода обратного распространения ошибки. Основная идея метода заключается в распространении информации, в виде сигнализации об ошибке, от выходных нейронов к входным, то есть по направлению, обратному по отношению к распространению сигналов по нейронной сети [6].

Стоит также отметить, что обучение сети производится на специально подготовленных базах изображений, классифицированных на большое количество классов, и занимает достаточно большое время.

На сегодняшний день крупнейшей базой является ImageNet ([www.image\\_net.org](http://www.image_net.org)) [6]. Она имеет бесплатный доступ для академических учреждений.

### Заключение

В результате вышеизложенного стоит отметить, что нейронные сети и алгоритмы, реализованные на принципе их функционирования, могут находить свое применение в системах распознавания дактилоскопической карты для органов внутренних дел. Часто именно программная компонента программно-аппаратного комплекса, направленного на распознавание такого уникального сложного изображения, как рисунок, являющийся идентификационными данными, решает возложенные на нее задачи не в полной мере. Программа, реализованная на базе алгоритмов, в основу которых входит нейронная сеть, будет значительно эффективнее.

Подводя итог можно резюмировать следующее:

- нейронные сети могут находить применение, как в вопросе распознавания изображений, так и текстов;
- данная теория дает возможность говорить о создании нового перспективного класса моделей, а именно моделей на основе интеллектуального моделирования;
- нейронные сети способны к обучению, что говорит о возможности оптимизации процесса из функционирования. Такая возможность является крайне важной опцией для практической реализации алгоритма;
- оценка алгоритма распознавания образов при помощи исследования нейронной сети может иметь количественное значение, соответственно, имеются механизмы подстройки параметров до необходимого значения за счет вычисления нужных весовых коэффициентов.

На сегодняшний день, дальнейшее исследование нейронных сетей представляется перспективной областью исследования, которая будет успешно применяться в еще большем количестве отраслей науки и техники, а также деятельности человека. Основной упор в развитии современных систем распознавания сейчас сдвигается в область семантической сегментации 3D изображений в геодезии, медицине, прото-

типировании и прочих сферах деятельности человека – это достаточно сложные алгоритмы и это связано:

- с отсутствием достаточного количества баз данных эталонных изображений;
- отсутствие достаточного количества свободных экспертов для первоначального обучения системы;
- изображения хранятся не в пикселях, что требует дополнительных ресурсов как от компьютера, так и от разработчиков.

Следует также отметить, что на сегодняшний день существует большое количество стандартных архитектур построения нейронных сетей, что существенно облегчает задачу построения нейронной сети с нуля и сводит ее к подбору подходящей конкретной задаче структуре сети.

В настоящее время на рынке существует достаточно большое количество инновационных компаний, занимающихся распознаванием изображений с использованием нейросетевых технологий обучения системы. Доподлинно известно, что они достигли точности распознавания изображений в районе 95 % при использовании базы данных в 10 000 изображений. Тем не менее все достижения относятся к статичным изображениям, с видеорядом в настоящий момент все гораздо сложнее.

### Список литературы

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). – 2000. – 100 p.
2. Жигалов К.Ю. Методика фотореалистичной векторизации данных лазерной локации с целью дальнейшего использования в ГИС // *Известия высших учебных заведений. Геодезия и аэрофотосъемка*. – 2007. – № 6. – С. 285–287.
3. Ranzato Marc'Aurelio, Christopher Poultney, Sumit Chopra and Yann LeCun: Efficient Learning of Sparse Representations with an Energy-Based Model, in J. Platt et al. (Eds), *Advances in Neural Information Processing Systems* (NIPS 2006). – 2010. – 400 p.
4. Жигалов К.Ю. Подготовка техники к использованию в системах автоматизированного управления строительства автодорог // *Естественные и технические науки*. – М., 2014. – № 1 (69). – С. 285–287.
5. Y. LeCun and Y. Bengio: Convolutional Networks for Images, Speech, and Time-Series, in Arbib, M. A. (Eds) // *The Handbook of Brain Theory and Neural Networks*. – 2005. – 150 p.
6. Y. LeCun, L. Bottou, G. Orr and K. Muller: Efficient BackProp, in Orr, G. and Muller K. (Eds) // *Neural Networks: Tricks of the trade*. – 2008. – 200 p.