

# Обзор Решений Нейронных Сетей Применяемых для Распознавания Образов

Александр Олейник  
кафедра Программной Инженерии  
Харьковский национальный университет  
радиоэлектроники  
Харьков, Украина  
oleksandr.oliinyk1@nure.ua

Екатерина Зыбина  
кафедра Программной Инженерии  
Харьковский национальный университет  
радиоэлектроники  
Харьков, Украина  
kateryna.zybina@nure.ua

Елена Олейник  
кафедра Программной Инженерии  
Харьковский национальный университет  
радиоэлектроники  
Харьков, Украина  
olena.oliinik@nure.ua

## Review of Neural Network Solutions Used for Pattern Recognition

Oleksandr Oliinyk  
Department of System Engineering  
Kharkiv National University of Radio Electronics  
Kharkiv, Ukraine  
Oleksandr.oliinyk1@nure.ua

Ekaterina Zybina  
Department of Software Engineering  
Kharkiv National University of Radio Electronics  
Kharkov, Ukraine  
kateryna.zybina@nure.ua

Elena Oleinik  
Department of Software Engineering  
Kharkiv National University of Radio Electronics  
Kharkov, Ukraine  
olena.oliinik@nure.ua

**Анотация**—В данном тексте рассмотрены основные идеи легшие в основу применения нейронных сетей для распознавания образов и ряд примеров их использования для решения задач в этой области.

**Abstract**—This text discusses the main ideas that formed the basis for the use of neural networks for pattern recognition and a number of examples of their use to solve problems in this area.

**Ключовые слова**—нейронные сети, распознавание образов, сверточные нейронные сети

**Keywords**—neural networks, pattern recognition, convolutional neural networks

### I. ВВЕДЕНИЕ

Искусственные нейронные сети (ИНС) – это математическая модель функционирования традиционных для живых организмов нейросетей, которые представляют собой сети нервных клеток. Как и в биологическом аналоге, в искусственных сетях основным элементом выступают нейроны, соединенные между собой и образующие слои, число которых может быть разным в зависимости от сложности нейросети и ее назначения (решаемых задач) [1].

Нейросети используются для распознавания и классификации многих объектов. Например, их используют для распознавания рукописного текста,



Інформаційні системи та технології ICT-2020  
Секція 4.  
Розпізнавання образів, цифрова обробка зображень і сигналів.

выделения важного на изображении, определения номеров машин, водители которых нарушают правила дорожного движения, распознавания лиц людей в видеопотоке, опознания самих людей по изображению, подсчета количества предметов определенного типа на изображении и многого другого. Как вы можете заметить, нейросети довольно многогранны и позволяют сделать многое, что не под силу обычным алгоритмам. Давайте рассмотрим их более пристально.

## II. ОБЩЕЕ ОПИСАНИЕ

Нейросеть – это математическая модель в виде программного и аппаратного воплощения, строящаяся на принципах функционирования биологических нейросетей [1].

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 [2]. Уже в 1949 г. Д. Хемб предлагает первый алгоритм обучения. Далее был ряд исследований в области нейронного обучения, и первые рабочие прототипы появились 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.

Существует два вида нейронных сетей: «обычные» и сверточные. Рассмотрим немного подробнее.

«Обычные» или «классические» нейросети [4] – это полносвязные нейронные сети (ПНС) прямого распространения с обратным распространением ошибки (см. Рис. 1).

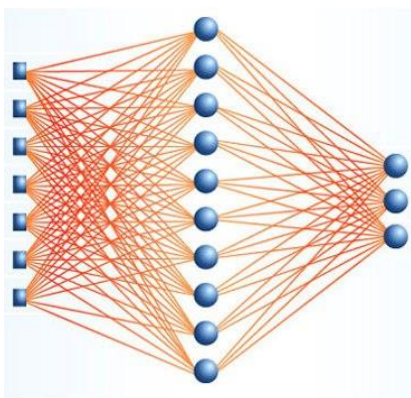


Рис. 1. Примерная схема ПНС [4]

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.

Сначала 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. Кроме того существует множество различных методик редукции сети (например OBD [3]), 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. Если принять во внимание что само по себе преобразование из изображения в индикатор класса довольно сложное и существенно нелинейное, одним слоем тут не обойтись. Задача распознавания подразумевает умение нейросети быть устойчивой к небольшим сдвигам, поворотам и изменению масштаба изображения.

Решение этой проблемы было найдено американским ученым французского происхождения Яном ЛеКуном, который предложил использовать так называемые сверточные нейронные сети [5].

Идея сверточных нейронных сетей заключается в чередовании сверточных слоев (C-layers), субдискретизирующих слоев (S-layers) и наличии полносвязных (F-layers) слоев на выходе (см. Рис. 2).

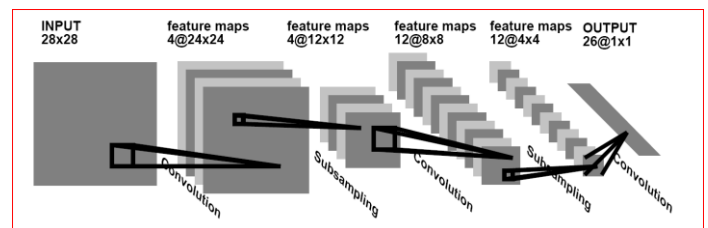


Рис. 2. Примерная схема сверточной нейронной сети [4]

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.

Концепция разделяемых весов предполагает, что для большого количества связей используется очень небольшой набор весов. Важно понимать, что самих наборов весов может быть много, но каждый из них будет применен ко всему изображению. Такие наборы часто называют ядрами.

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.



Процесс распространения сигнала в S-слое изображен на рисунке 3.

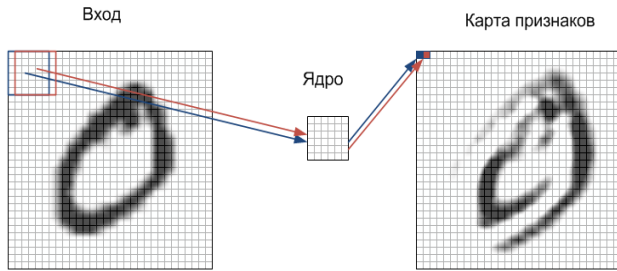


Рис. 3. Процесс распространения сигнала в S-слое [4]

Каждый фрагмент изображения поэлементно умножается на небольшую матрицу весов (ядро), результат суммируется. Эта сумма является пикселем выходного изображения, которое называется картой признаков. 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. 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 [4].

### III. ПРИМЕРЫ ИСПОЛЬЗОВАНИЯ

Применение подвижных технических средств позволяет добиться автоматизации и качества их решения. Примером выступают задачи, связанные с быстрым осмотром или обходом территории без присутствия человека, проникновением на недоступные или угрожающие жизни людей территории. Обход территории сопровождается составлением ее плана, нанесением на план точек интереса, распознавание найденных объектов, Двигаясь по известному маршруту помещения или

местности, техническое средство выделяет и распознает встречающиеся на пути объекты, тем самым создавая план помещения с нанесенными на него знаками, в качестве которых могут выступать особые объекты на местности, условные обозначения, нанесенные на стены, пол или потолок. Особые знаки могут контролировать взлет и посадку БПЛА.

Например, в статье [6] распознавание объектов основано на контурном представлении изображений и измерении сходства этих контуров. Изначально входное изображение преобразуется в контурное бинарное изображение. Процесс сопоставления и вычисления меры сходства контурного эталона с текущим, преобразованным в дистантное изображение объекта, сводится к процедуре вычисления локальной суммы пикселей дистантного изображения, “накрытых” контурными пикселями эталонного изображения.

В работе [7] предложен и описан метод позиционирования БПЛА без использования систем спутниковой навигации, с применением данных, полученных с бортового фото-видео регистратора, и заранее загруженных данных о местности. В статье [8] представлен быстрый алгоритм нахождения маркеров на изображениях. Данный алгоритм полезен при нахождении маркеров, которые имеют четкие границы с фоном, и устойчив к неравномерному освещению и деформации маркеров.

В работе [9] описывается система, которая является ассистентом для водителя. Система позволяет выделять разметку и границы дороги, тем самым корректируя движение автомобиля. Определение линий разметки не чувствительно к освещению и масштабу. В работе [10] представлен алгоритм автоматического поиска и распознавания особых объектов на фото и видеокдрах, получаемых с борта БПЛА и ПТС, совершающих обход помещения. Он настолько хорошо описан, что хотелось бы его процитировать: «Применительно к фото и видеокдрам (далее изображениям) используется подход на основе сканирующего окна: окном поиска сканируется изображение, а затем применяется классификатор к каждому положению. Далее СНС выбирает положения с наиболее значимыми признаками (с наименьшей ошибкой классификации). Алгоритм сканирования окна с признаками выглядит так: (1) дано исследуемое изображение, выбрано окно сканирования размером  $40 \times 40$  пикселей; (2) далее окно сканирования начинает последовательно двигаться по изображению с шагом в одну ячейку окна (размер самого окна есть размер ячейки); (3) сканирование производится последовательно для различных масштабов; (4) масштабируется не само изображение, а сканирующее окно (изменяется размер ячейки: 70, 100, 130, 170, 210, 250, 300 пикселей); (5) все найденные фрагменты формируют выборку для обучения СНС, которая «выносит вердикт», к какому классу относится объект во фрагменте».



#### IV. ВЫВОДЫ

В данной работе рассмотрены идеи, легшие в основу нейронных сетей, применения нейронных сетей и сверточных нейронных сетей для распознавания образов. А так же рассмотрены примеры решения задач распознавания образов с использованием нейронных сетей.

#### ЛИТЕРАТУРА REFERENCES

- [1] Нейронные сети: распознавание образов и изображений с помощью ИИ // Центр2М URL: <https://center2m.ru/ai-recognition> (дата обращения: 10/21).
- [2] Маркова С.В., Жигалов К.Ю. Применение нейронной сети для создания системы распознавания изображений // Фундаментальные исследования. – 2017. – № 8-1. – С. 60-64; URL: <http://www.fundamental-research.ru/ru/article/view?id=41621> (дата обращения: 02.11.2020).
- [3] 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.
- [4] Применение нейросетей в распознавании изображений // Хабр URL: <https://habr.com/ru/post/74326/> (дата обращения: 10/21).
- [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, MIT Press, 1995.
- [6] Павлова В.А., Крюков С.Н., Каркаева Р.К., Созинова М.В. Автоматическое компьютерное распознавание наземных и морских объектов. Известия ЮФУ. Технические науки, 2010, №3, с. 73–77.
- [7] Амелин К.С. «Метод ориентирования сверхлегкого БПЛА при редком обновлении данных о его местоположении», Стохастическая оптимизация в информатике, 10:2 (2010), с. 3–14.
- [8] M. Hirzer. Marker Detection for Augmented Reality Applications, Technical Report ICG-TR-08/05, Inst. for Computer Graphics and Vision, Graz University of Technology, Graz, Austria, 2008, 27 p., URL: [http://studierstube.icg.tugraz.at/thesis/marker\\_detection.pdf](http://studierstube.icg.tugraz.at/thesis/marker_detection.pdf).
- [9] N. Arshad, K.-S. Moon, S.-S. Park, J.-N. Kim. “Lane Detection with Moving Vehicles Using Color Information”, Proceedings of the World Congress on Engineering and Computer Science. V. 1, WCECS 2011 (San Francisco, USA, October 19–21, 2011), 2011, pp. 499–502, URL: [http://www.iaeng.org/publication/WCECS2011/WCECS2011\\_pp499-502.pdf](http://www.iaeng.org/publication/WCECS2011/WCECS2011_pp499-502.pdf)
- [10] Кирюшина А.Е. Выделение и классификация знаков пожарной безопасности с использованием нейронных сетей. Программные Системы: Теория и приложения №4(31), 2016, с. 317–329.

