

Д. С. Казанцев^{1,2}, В. Г. Эдвабник¹, М. В. Орлова¹*

Сверточная нейронная сеть как метод принятия решения об обнаружении

¹ Научно-исследовательский институт электронных приборов, г. Новосибирск, Российская Федерация

² Новосибирский государственный технический университет, г. Новосибирск, Российская Федерация

* e-mail: denis-kazanchev@yandex.ru

Аннотация. В данной работе предлагается метод принятия решения об обнаружении, также имеющий потенциал к классификации обнаруживаемого объекта для локационных систем. В статье дано краткое описание сверточной нейронной сети и ее применения в настоящий момент, а также дана характеристика возможности ее применения для решения задач локации, и выделены проблемы, возникающие при обработке отраженного сигнала. В работе обсуждается метод зондирования пачкой импульсов. Основная проблема, которую призван решить предлагаемый метод – это распознавание полезного сигнала на фоне помех, а также различение помехового сигнала и сигнала, отраженного от объекта. Также данная статья является теоретической, описывающей метод и проблематику, а последующая работа предполагает практическую реализацию, основанную на описанной системе, обученной на синтетических данных.

Ключевые слова: сверточная нейронная сеть, обнаружение, распознавание

D. S. Kazantsev^{1,2}, V. G. Edvabnik¹, M. V. Orlova¹*

The Convolutional Neural Network as a Detection Decision-making Technique

¹ The Electronic Device Research and Development Center, Novosibirsk, Russian Federation

² Novosibirsk State Technical University, Novosibirsk, Russian Federation

* e-mail: denis-kazanchev@yandex.ru

Abstract. This paper proposes a detection decision-making technique that also has the potential to classify the detected object for radar systems. The paper briefly describes the convolutional neural network and its current applications, as well as characterize the feasibility of its application to location tasks and highlight the problems encountered in processing the echo signal. The paper discusses the pulse packet zoning method. The main problem that the proposed method is designed to solve is the recognition of a useful signal on the background of interference, as well as distinguishing between the interfering signal and the signal reflected from the target. Also, this paper is a theoretical paper describing the method and problematics, and the subsequent work involves a practical implementation based on the described system trained on synthetic data.

Keywords: convolution neural networks, detection, recognition

Введение

В современных условиях воздействия разного рода электромагнитных излучений проблема выделения полезного сигнала набирает все большую актуальность, в связи с этим необходимо рассматривать новые подходы к решению данной задачи. Большими перспективами среди методов идентификации сигналов обладают нейронные сети [1]. В настоящее время нейронные сети хорошо проявляют себя в решении задач распознавания и идентификации.

Нейронная сеть – это структура, имитирующая часть нейронной системы биологических организмов [2].

Главными преимуществами нейронных сетей являются:

1) многослойная высоко-параллельная логическая структура – способность работать с большим числом независимых элементов [3];

2) обучаемость – способность обучаться, настраивать свободные параметры нейронной сети, симулируя рабочую среду [2].

Используемая в данной работе архитектура сверточных нейронных сетей была впервые предложена Яном Лекуном в 1998 году в статье «Gradient-Based Learning Applied to Document Recognition» [4], но первые результаты качественного применения данных методов глубокого обучения были продемонстрированы Дэном Киресаном на ICDAR 2011 и IJCNN 2011 в соревнованиях по распознаванию китайских иероглифов и немецких дорожных знаков [5], соответственно, а в 2012 году, применяя сверточные нейронные сети, группа Хинтона выиграла крупномасштабное состязание по распознаванию изображений из набора ImageNet, показав точность 84,6 %, в дальнейшем эта характеристика продолжала расти.

Основная часть

На вход приемного устройства поступает случайный процесс в виде суперпозиции:

$$y(t) = A \cdot s(t) + n(t), \quad (1)$$

где $s(t)$ – полезный сигнал, отраженный от объекта; $n(t)$ – помеховый сигнал, представленный случайным процессом.

Коэффициент A принимает значения 1 или 0, в зависимости от состава поступившего сигнала.

Используем зондирующий сигнал в виде пачки импульсов:

$$x(t) = \sum_{k=1}^M X_k(t - (k-1) \cdot T) \cdot \exp(i \cdot 2 \cdot \pi \cdot f_0 \cdot (t - (k-1) \cdot T)), \quad (2)$$

где $X_k(t - (k-1) \cdot T) = A_k(t - (k-1) \cdot T) \cdot \exp(i \cdot \varphi(t - (k-1) \cdot T))$ – комплексная огибающая; k – номер импульса в пачке; $A_k(t - (k-1) \cdot T)$ – закон амплитудной

модуляции; T – период повторения импульсов; M – число импульсов в пачке; f_0 – несущая частота; $\varphi(t - (k - 1) \cdot T)$ – закон фазовой модуляции.

Тогда полезный сигнал, отраженный от объекта, будет иметь вид:

$$s(t) = \sum_{k=1}^M S_k(t - (k - 1) \cdot T - \tau_k) \cdot \exp(i \cdot 2 \cdot \pi \cdot (f_0 + f_D) \cdot (t - (k - 1) \cdot T - \tau_k)), \quad (3)$$

где

$$S_k(t - (k - 1) \cdot T - \tau_k) =$$

$$= Z_k(t - (k - 1) \cdot T - \tau_k) A_k(t - (k - 1) \cdot T - \tau_k) \cdot \exp(i \cdot \varphi'_k(t - (k - 1) \cdot T - \tau_k))$$

– комплексная огибающая отраженного сигнала; $Z_k(t - (k - 1) \cdot T - \tau_k)$ – закон пространственной модуляции; f_D – частота Доплера; τ_k – время запаздывания k -го импульса [6]:

$$\tau_k = \frac{2 \cdot R}{c} - \frac{2 \cdot U_r \cdot R}{c \cdot (c + U_r)} + (k - 1) \cdot T \cdot \left(1 - \frac{2 \cdot U_r}{c + U_r} \right), \quad (4)$$

где R – расстояние от излучателя до отражающего объекта, измеряемое по фронту первого импульса; c – скорость света; U_r – радиальная скорость движущегося объекта, отразившего сигнал.

Тогда период повторения импульсов в пачке отраженного от объекта сигнала:

$$T_{pr} = T - \frac{2 \cdot U_r \cdot T}{c + U_r}. \quad (5)$$

Сдвиг фазы каждого следующего импульса внутри пачки отраженного от подвижного объекта сигнала:

$$\Delta\varphi = 2 \cdot \pi \cdot (f_0 + f_D) \cdot \frac{2 \cdot U_r \cdot T}{c + U_r}. \quad (6)$$

Для принятия решения об обнаружении в данной работе предлагается применять сверточную нейронную сеть.

Сверточная нейронная сеть (CNN) – это многослойный персептрон [2], обученный для распознавания двумерных поверхностей инвариантно к деформациям.

Одним из важнейших свойств структуры CNN является ее блочный тип, т.е. разработчик имеет возможность изменять структуру в соответствии с выполняе-

мыми задачами, а также корректировать ее, исходя из полученных результатов. Далее в работе будет дано описание каждого из основных блоков.

Еще одним достоинством данного вида сетей является его слабая восприимчивость к различным искажениям обрабатываемого сигнала, т.е. система выделяет признаки, по которым в дальнейшем сможет распознать сигнал при его флуктуациях. Можно выделить три основных идеи устройства сверточных нейронных сетей [4]:

- локальность выделяемых признаков [1];
- общие обучаемые весовые коэффициенты для всех элементов обрабатываемого сигнала (возможность распознавать признаки по всему обрабатываемому сигналу и, в сравнении с другими видами нейронных сетей, небольшое число обучаемых элементов) [1];
- иерархичность структуры.

Первые параметры, которые необходимо задать при обучении нейронной сети, это объем выборки – конечный набор прецедентов, выбранных некоторым способом из множества всех возможных прецедентов. Выборка делится на следующие части [7]:

1) обучающая выборка, представляющая часть, применяемую для настройки (оптимизации параметров) нейронной сети; при завышенном объеме сеть будет работать некорректно, пытаясь искать шаблоны, а не признаки, при неполноте данных сети не будет достаточно выделенных признаков;

2) тестовая выборка – часть, применяемая для оценки качества обученной модели; для наиболее качественной оценки эти сигналы не должны совпадать с сигналами из обучающей выборки;

3) валидационная выборка – выборка сигналов для сравнения различных моделей, построенных по обучающей выборке.

Также необходимо установить размер сигнала, поступающего на вход, размер входного слоя $M \times N$ влияет на затрачиваемые вычислительные мощности, и слишком большой объем может повлечь за собой долгий отклик сети. В нашем случае это матрица размером $M \times N$, где M – это число импульсов в пачке, а также количество интервалов, на которые разбивается сигнал для дискретизации. N – это количество дискретных значений на интервале.

$$N = \frac{T \cdot M}{2 \cdot \max(f_0 + f_D; f_0 - f_D)}. \quad (7)$$

Главный критерий дискретизации – соответствие теореме Котельникова [8]. Также, при выборе интервалов дискретизации необходимо учитывать сдвиг периода повторения импульсов (5).

Однако, перед поступлением на входной слой элементы $a_{m,n}$ матрицы $A_{M \times N}$ входного сигнала необходимо нормализовать по интервалу $[0;1]$, это упростит вычислительный процесс [5].

Функционал входного слоя заключается в принятии на обработку сигнала и подаче его на основное тело сети.

Основной функционал возложен на сверточные слои, представляемые картами признаков и соответствующими им фильтрами (ядрами). Фильтр – это матрица, состоящая из настраиваемых элементов – весов.

Изначально, весам присваиваются случайные значения из узкого диапазона, однако, они являются обучаемыми параметрами, и в ходе обучения сети принимают новые значения, в соответствии с оптимальными параметрами функции потерь. Обучение сети следует понимать, как процесс минимизации функции ошибки в результате настройки весов. Функция ошибки – это разность между реакцией с выхода системы и значением, установленным в качестве порога для принятия решения. Одним из первых методов обучения и основным при решении прикладных задач является алгоритм обратного распространения ошибки.

На рис. 1 продемонстрирован принцип выделения локальных признаков, фильтр «скользит» по матрице, поступившей с предыдущего слоя, с заданным шагом, вычисляя скалярное произведение, которое устанавливается в качестве соответствующего элемента карты признаков. Размер карт признаков определяется размером обрабатываемого массива, размером фильтра и шагом свертки.

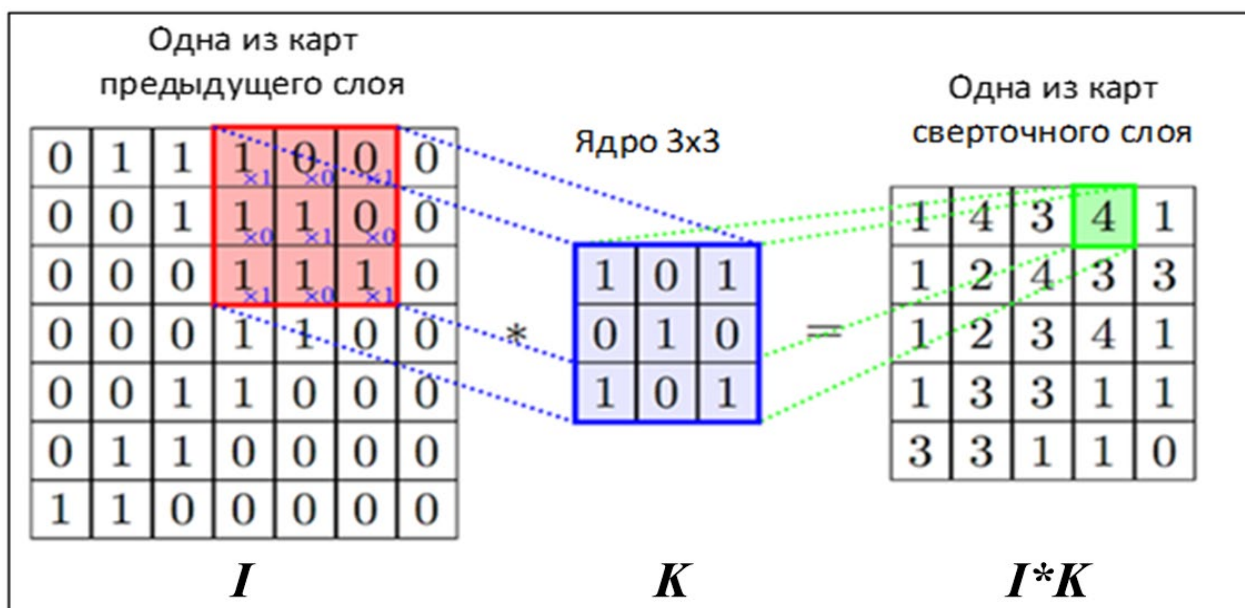


Рис. 1. Пример выполнения операции свертки

Таким образом, основными параметрами сверточного слоя являются: количество карт признаков, влияющее на качество распознавания сигнала и затрачиваемые вычислительные ресурсы; размер ядра фильтра, влияет на качество выделяемых признаков; шаг свертки, влияет на качество выделяемых признаков, также, меняя шаг свертки, можно влиять на размерность карт признаков; пара-

метр дополнения (padding), применяется для сохранения исходной размерности путем дополнения обрабатываемой матрицы контуром из нулей.

Схожую со сверточными слоями функцию имеют субдискретизирующие (подвыборочные) слои. Главная задача таких слоев – выделение наиболее значимых признаков в результате уменьшения карт, полученных с предыдущего слоя. Фильтрация менее значимых признаков позволяет сети не переобучаться, что сокращает вероятность работы сети по шаблонам вместо признаков.

Окно субдискретизирующего слоя скользит по матрице, полученной с предыдущего слоя, и в рамках этого окна вычисляется значение соответствующего элемента выходного сигнала в соответствии с выбранной функцией. Наиболее эффективной функцией является max-pooling – выбор максимального значения.

Основными параметрами подвыборочных слоев являются правило формирования значений элементов карт признаков, влияющее на их качество, а также размерность скользящего окна, определяющая размерность карт признаков.

Важным типом слоев, который часто не выделяют как отдельный, а считают частью других, является активационный слой. Слои активации применяются для введения нелинейности, что приводит к расширению пространства гипотез, невозможному из-за его ограниченности при аффинных преобразованиях входных данных [5].

Получая на вход матрицу, после слоя активации мы имеем матрицу того же размера, каждый элемент которой является значением функции активации в качестве аргумента, получившей соответствующий элемент обрабатываемой матрицы. Главным параметром данного слоя является функция активации, выбираемая, исходя из поставленных задач. Существует большое множество готовых функций, а также имеется возможность применять собственные функции. Однако, закона оптимального выбора функции активации нет, и необходимо выбирать, исходя из результатов обучения сетей.

После выделения карт признаков следует полносвязный слой, каждый нейрон этого слоя связан с каждым элементом карты признаков, поступившей на него. Количество нейронов является главным параметром этого слоя.

Таким образом, выходной нейрон полносвязного слоя равен:

$$\lambda_{n,i} = W_{n,i}^{l \times y} \otimes H_i^{l \times y} + w_0, \quad (9)$$

где $\lambda_{n,i}$ – значение n -го нейрона, i -ой карты признаков; $W_{n,i}^{l \times y}$ – весовая матрица для n -го нейрона и i -го входного тензора, размером $l \times y$; $H_i^{l \times y}$ – матрица i -й карты признаков, размером $l \times y$, из слоя, расположенного перед полносвязным, w_0 – добавочный весовой коэффициент.

Важно сказать, что после полносвязного слоя обязательно должен следовать активационный.

Завершает нейронную сеть выходной слой, он работает так же, как и полносвязный, однако количество нейронов на нем равно количеству выделяемых классов.

Заключение

В данной работе был предложен метод принятия решения об обнаружении, основанный на сверточной нейронной сети, представлена математическая модель предлагаемого к применению сигнала, описана структура сверточной нейронной сети, а также дано краткое описание ее элементов.

По итогам можно выделить следующие основные задачи, которые необходимо решить при создании нейронных систем, применяемых в локациях. Определение выборки. Большое значение имеет как объем, так и разнообразие выборки, а также качество сигналов и полнота отображения вероятных условий работы системы. Сеть выделяет абстрактные признаки, свойственные отраженному сигналу в разных условиях. Далее, необходимо установить степень влияния вида сигнала, а также сравнить результаты при подаче на вход сети сигнала с предобработкой и без нее. Требуется рассмотреть случаи с различными параметрами входного слоя, что означает изменение количества импульсов в пачке и увеличение частоты дискретизации. Основываясь на описании структурных блоков, изменяя их количество, порядок сверточных и слоев подвыборки, а также их параметры, необходимо найти оптимальную структуру сети. Еще одна важная задача, возникающая перед разработчиком – это контроль за вычислительными нагрузками и учет их допустимости при получаемой точности распознавания сигнала.

Также, стоит отметить потенциал не только бинарной классификации, но и мультизадачной. Это возможно при достаточном объеме данных, применяемых при генерации выборки, для каждого класса. Главное отличие – количество нейронов на выходном слое, соответствующее количеству классов.

Дальнейшая работа предполагает генерирование обучающей выборки, состоящей из синтетических сигналов, основанных на математической модели (2), и обучение на ее основе сверточной нейронной сети при различных структуре и параметрах. Также необходимо рассмотреть применение сверточных нейронных сетей для иных систем обнаружения и возможность модернизации сети путем дополнения ее другими сетями, например, рекуррентной.

БИБЛИОГРАФИЧЕСКИЙ СПИСОК

1. Татузов А. Л., Нейронные сети в задачах радиолокации. Кн. 28. – М.: Радиотехника, 2009. – 432 с: ил. (Научная серия «Нейрокомпьютеры и их применение»). ISBN 978-5-88070-244-2.
2. Хайкин, С. Нейронные сети: полный курс / Саймон Хайкин; [пер. с англ. и ред. Н.Н. Кассуль]. – Изд. 2-е, испр. – Москва [и др.]: Вильямс, 2017. – 1103 с.: ил.
3. Большая российская энциклопедия : [в 35 т.] / гл. ред. Ю. С. Осипов. – М. : Большая российская энциклопедия, 2004–2017.
4. LeCun, Y., Bottou, L., Bengio, Y., & Haffner, P. (1998). Gradient-based learning applied to document recognition. Proc. IEEE, 86, 2278-2324.
5. Шолле Франсуа, Глубокое обучение на Python. – СПб.: Питер, 2018. – 400 с.: ил.

6. Чепкасов А.В., Сухоплюев А.В. Радиолокационный прием пачки импульсов большой длительности при обнаружении движущихся целей. Вестник Концерна ВКО «Алмаз – Антей». 2019;(1):30-34. DOI 10.38013/2542-0542-2019-1-30-34.

7. Гудфеллоу Я., Бенджио И., Курвилль А. Глубокое обучение / пер. с англ. А. А. Слинкина. – 2-е изд., испр. – М.: ДМК Пресс, 2018. – 652 с.: цв. ил.

8. Оппенгейм А., Шафер Р. Цифровая обработка сигналов / пер. с англ. С. А. Кулешова под ред. А. Б. Сергиенко. 2-е изд., испр. М. : Техносфера, 2007. 856 с.

© Д. С. Казанцев, В. Г. Эдвабник, М. В. Орлова, 2024