

## Раздел 4 АВТОМАТИЗАЦИЯ, АНАЛИЗ И ОБРАБОТКА ИНФОРМАЦИИ, УПРАВЛЕНИЕ ТЕХНОЛОГИЧЕСКИМИ ПРОЦЕССАМИ В СОЦИАЛЬНЫХ И ЭКОНОМИЧЕСКИХ СИСТЕМАХ

УДК 303.732.4

DOI: 10.34046/aumsuomt94/29

### НЕЙРОСЕТЕВЫЕ СИСТЕМЫ АВТОМАТИЧЕСКОГО РАСПОЗНАВАНИЯ МОРСКИХ ОБЪЕКТОВ

*Н.А. Полковникова, кандидат технических наук*

В статье рассмотрены технологии компьютерного зрения на основе глубоких свёрточных нейронных сетей. Применение нейронных сетей особенно эффективно для решения трудно формализуемых задач. Разработана архитектура свёрточной нейронной сети применительно к задаче распознавания и классификации морских объектов на изображениях. В ходе исследования выполнен ретроспективный анализ технологий компьютерного зрения и выявлен ряд проблем, связанных с применением нейронных сетей: «исчезающий» градиент, переобучение и вычислительная сложность. При разработке архитектуры нейросети предложено использовать функцию активации RELU, обучение некоторых случайно выбранных нейронов и нормализацию с целью упрощения архитектуры нейросети. Сравнение используемых в нейросети функций активации ReLU, LeakyReLU, Exponential ReLU и SOFTMAX выполнено в среде Matlab R2019b. На основе свёрточной нейронной сети разработана программа на языке программирования Visual C# в среде MS Visual Studio 2019 для распознавания морских объектов. Программа предназначена для автоматизированной идентификации морских объектов, производит детектирование (нахождение объектов на изображении) и распознавание объектов с высокой вероятностью обнаружения.

**Ключевые слова:** глубокое машинное обучение, свёрточные нейронные сети, распознавание образов, обработка изображений, машинное зрение, системы автоматического слежения за объектом, функция активации.

The article considers computer vision technologies based on deep convolutional neural networks. Application of neural networks is particularly effective for solving difficult formalized problems. As a result convolutional neural network architecture to the problem of recognition and classification of marine objects on images is implemented. In the research process a retrospective analysis of computer vision technologies was performed and a number of problems associated with the use of neural networks were identified: vanishing gradient, overfitting and computational complexity. To solve these problems in neural network architecture development, it was proposed to use RELU activation function, training some randomly selected neurons and normalization for simplification of neural network architecture. Comparison of ReLU, LeakyReLU, Exponential ReLU, and SOFTMAX activation functions used in the neural network implemented in Matlab R2019b. The computer program based on convolutional neural network for marine objects recognition implemented in Visual C# programming language in MS Visual Studio 2019 integrated development environment. The program is designed for automated identification of marine objects, produces detection (i.e., presence of objects on image), and objects recognition with high probability of detection.

**Key words:** deep machine learning, convolutional neural networks, pattern recognition, image processing, computer vision, object tracking systems, activation function.

За последние годы компьютерное зрение стало ключевой технологией искусственного интеллекта во многих областях и разрабатывается для:

- беспилотных автомобилей (Tesla, Volvo, Audi, BMW, Mercedes-Benz, Google, Yandex, Alphabet Waymo, Baidu Apollo, Uber);
- безэкипажных морских судов (Kongsberg, Rolls-Royce, Hyundai, Transas);
- робототехники (Honda, Google Boston Dynamics, NASA, Sony, Aldebaran Robotics, Hanson Robotics и др.);

– сервисов для программных приложений (Amazon Rekognition, Google Vision AI, Facebook DeepFace, IBM Watson);

- интеллектуальных диалоговых систем (Apple Siri, Amazon Alexa, Google Assistant, Яндекс Алиса, Samsung Bixby, Microsoft Cortana, Mail Group Маруся);
- распознавания и синтеза речи (Google Voice Search, Google WaveNet);
- машинного перевода (Google Neural Machine Translation);

- семантического анализа текста и изображений (IBM Project Debater, StackGAN);
- классификации и распознавания изображений (Google SketchRNN);
- повышения качества фотографий (Google GAN);
- розничной торговли, электронной коммерции и автоматизированных складов (Alibaba Group, Amazon, Jingdong Mall, Walmart, Coca-Cola, Kimberly-Clark, McDonald's);
- поддержки принятия инвестиционных решений (Alibaba KunPeng Parameter Server);
- медицины (Tencent Miying, Infervision);
- автоматического слежения за объектами, ситуационного анализа (распознавания действий и намерений);
- технологий «Интернета вещей» (Samsung Smart Home, Xiaomi Mi Smart Home, Disney MagicBand);

Технологии компьютерного зрения применяются также для автоматизированных торговых точек, контроля качества и технологических процессов, технологий дополненной реальности, систем безопасности и видеонаблюдения [1-5]. В

2018 году в США открылись первые автоматизированные супермаркеты Amazon Go без касс и продавцов. Покупатели заходят с помощью смартфона, система контроля распознаёт покупки и добавляет их в виртуальную корзину, а при выходе из магазина со счёта клиента автоматически списывается стоимость всех выбранных товаров. Компьютерное зрение позволяет ответить на следующие вопросы: на каком расстоянии находится объект, движется ли объект, сколько человек присутствуют на изображении и присутствует ли на нём конкретный человек. Для решения задач компьютерного зрения применяются сверточные искусственные нейронные сети – это исключительно мощный метод имитации процессов и явлений, позволяющий обрабатывать чрезвычайно сложные зависимости [6-9].

**Ретроспективный анализ развития нейронных сетей.** Первая модель искусственной нейронной сети разработана в 1943 году американским нейрофизиологом Уорреном МакКаллоком и нейролингвистом и математиком Уолтером Питтсом. На рис. 1 представлена структурная схема модели искусственного нейрона.

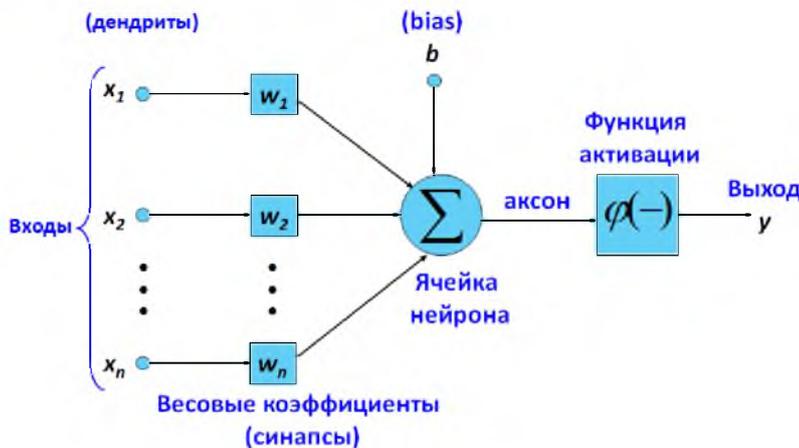


Рисунок 1 – Структурная схема модели искусственного нейрона

Текущее состояние нейрона определяется, как взвешенная сумма его входов:

$$s = \sum_{i=1}^n x_i \cdot w_i \quad (1)$$

где  $x = (x_1, x_2, \dots, x_n)$  – множество входных сигналов, поступающих на искусственный нейрон;  $w = (w_1, w_2, \dots, w_n)$  – множество весовых коэффициентов (каждый вес соответствует «силе» одной биологической синаптической связи). Выход нейрона есть функция его состояния  $y = f(s)$ . Таким образом, искусственный нейрон обладает группой синапсов – однонаправленных входных связей, соединенных с выходами других нейронов, а также имеет аксон – выходную связь данного нейрона, с которой сигнал поступает на синапсы следующих нейронов.

В 1957 году американский учёный Фрэнк Розенблатт (Frank Rosenblatt) разработал перцептрон – модель восприятия информации мозгом на основе обучающейся нейронной сети. На вход перцептрона подаются весовые коэффициенты  $w$  и исходные данные  $x$ , а их произведение корректируется смещением  $b$ :

$$f(x) = \begin{cases} 1 & \text{если } \sum_{i=1}^n w_i \cdot x_i + b > 0, \\ 0 & \text{если } \sum_{i=1}^n w_i \cdot x_i + b \leq 0 \end{cases} \quad (2)$$

В выражении (2) выходное значение перцептрона равно 1, если взвешенная сумма  $\sum_{i=1}^n w_i \cdot x_i + b$  больше нуля, а в остальных случаях значение перцептрона равно нулю.

В 1958 году Ф. Розенблатт создал первый в мире нейрокомпьютер «Марк-1», способный распознавать с помощью камер некоторые буквы, написанные на карточках. В 1986 году был разработан алгоритм обратного распространения ошибки, позволивший обучать многослойные нейронные сети прямого распространения (многослойные персептроны).

Однако несовершенство компьютеров и отсутствие на тот момент необходимой вычислительной мощности приостановило развитие технологий машинного обучения. Фактически всю вторую половину XX века машинное обучение оставалось исключительно математической дисциплиной, которой занимались учёные в академической среде. Стремительный переход к практике случился лишь в начале XXI века. В период с 1998 по 2010 гг развитию нейронных сетей препятствовали технические ограничения: высокая стоимость вычислений, памяти, цифровой фотографии и громоздкость фото- и видеокамер.

С 2010 года рост вычислительных возможностей, снижение стоимости аппаратных ресурсов и качество камер мобильных телефонов привели к накоплению цифровых данных для обучения нейросетей. С 2015 года получили широкое распространение графические процессоры, которые ускоряют и удешевляют параллельные вычисления. Все эти причины привели к разработке множества новых алгоритмов глубокого машинного обучения [10].

**Анализ проблем обучения нейронных сетей.** Развитие алгоритмов и технологий глубоких нейронных сетей открыло ряд проблем: «исчезающий» градиент, переобучение и вычислительная сложность.

Проблема «исчезающего» градиента («vanishing gradient problem») – заключается в быстрой потере информации с течением времени. В алгоритме обратного распространения ошибка обучения распространяется в направлении от результирующего слоя по сети и тем самым происходит настройка весовых коэффициентов. Если ошибка обучения не достигает всех скрытых слоёв нейросети, обучение не может быть выполнено успешно. «Исчезающий» градиент приводит к трудности обучения и настройке параметров первых слоёв нейросети. Проблема усугубляется по мере увеличения общего количества слоёв.

Проблема «исчезающего» градиента для глубоких нейронных сетей может быть решена применением функции активации ReLU (Rectified

Linear Unit или Rectifier – «выпрямитель»). Функция ReLU принимает максимальное значение между нулём и заданным входным значением:

$$f(x) = \max(0, x) = \begin{cases} x, & \text{при } x \geq 0 \\ 0, & \text{при } x < 0 \end{cases} \quad (3)$$

Функция возвращает значение  $x$ , если  $x$  положительно, и  $0$  – в противном случае. Диапазон вывода функции от  $0$  до  $+\infty$ . Производная функции ReLU:

$$f'(x) = \begin{cases} 1, & x \geq 0 \\ 0, & x < 0 \end{cases} \quad (4)$$

Функция активации («activation function») – это способ нормализации входных данных, является одним из важнейших аспектов нейронных сетей и привносит в нейросети нелинейность, делая их универсальными аппроксиматорами функций. Если обозначить функцию активации нейрона как  $f$ , то выходной сигнал  $Y$  зависит от вектора входных значений  $[x_1, x_2, \dots, x_n]$  и вектора весовых коэффициентов  $[w_1, w_2, \dots, w_n]$  следующим образом:

$$Y = f(x_1 \cdot w_1 + x_2 \cdot w_2 + \dots + x_n \cdot w_n) \quad (5)$$

Функция ReLU вычисляется без применения сложных математических преобразований, что позволяет значительно ускорить обучение. Градиент (наклон) функции ReLU остаётся постоянным и позволяет значительно ускорить обучение (по сравнению с другими функциями). Функция ReLU и её производная вычисляется без применения сложных математических преобразований, менее требовательна к вычислительным ресурсам, чем гиперболический тангенс (tanh) или сигмоида (sigmoid). Функции активации tanh и sigmoid «утяжеляют» работу нейросети, поскольку выполняют преобразования при любом значении  $x$ , тогда как ReLU – только при  $x \geq 0$ . Поэтому ReLU позволяет произвести процесс обучения быстрее. Выбор функции активации во многом зависит от типа решаемой задачи. Однако ReLU хорошо работает для широкого круга задач и используется для скрытых слоёв глубоких нейронных сетей.

Недостаток функции ReLU заключается в том, что когда на вход нейрона поступают отрицательные числа ( $x < 0$ ), на выходе нейрона получается число  $0$  вплоть до конца процесса обучения. Это явление называется «проблемой угасающей ReLU».

Для решения «проблемы угасающей ReLU» введена функция Leaky ReLU, которая для отрицательных значений аргумента возвращает аргумент умноженный на малый положительный коэффициент  $a$  (например,  $a = 0,01$ ):

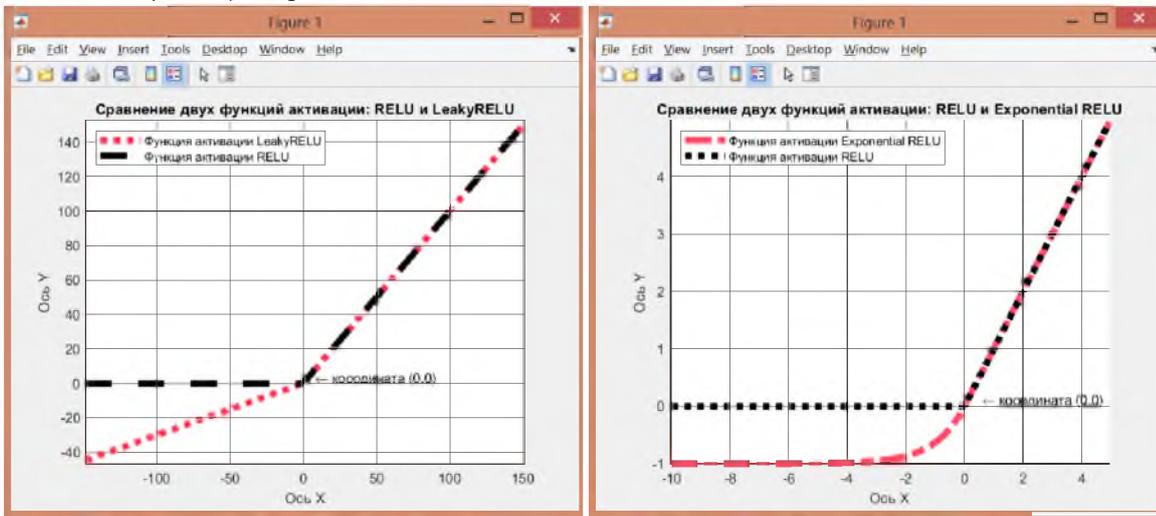
$$f(x) = \begin{cases} x, & \text{при } x \geq 0 \\ a \cdot x, & \text{при } x < 0 \end{cases} \quad (6)$$

Для улучшения функций ReLU и Leaky ReLU введена экспоненциальная ReLU (Exponential ReLU). Экспоненциальная ReLU определяется как:

$$f(x) = \begin{cases} x, & \text{при } x \geq 0 \\ \alpha(e^x - 1), & \text{при } x < 0 \end{cases} \quad (7)$$

где коэффициент  $\alpha \geq 0$  – гиперпараметр.

На рис. 2 представлено графическое сравнение функций активации ReLU, Leaky ReLU и Exponential ReLU в среде Matlab R2019b.



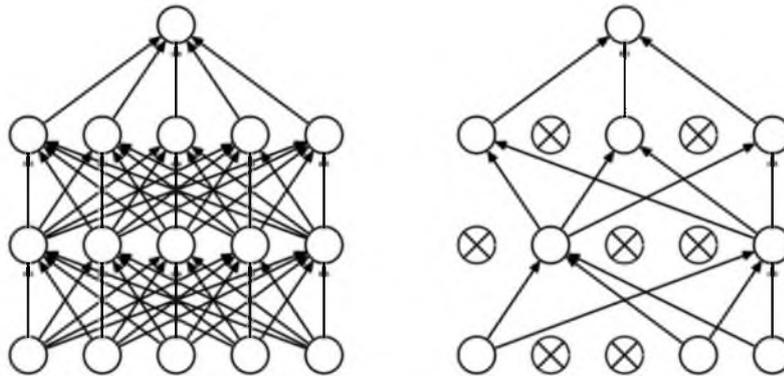
а) функции ReLU и LeakyReLU

б) функции ReLU и Exponential ReLU

Рисунок 2 – Сравнение графиков функций активации

Проблема переобучения («overfitting») – излишне точное соответствие нейронной сети конкретному набору обучающих примеров, при котором сеть теряет способность к обобщению.

Решение проблемы переобучения связано с применением метода «dropout» – обучение некоторых случайно выбранных нейронов (обычно применяя стохастический градиентный спуск), а не сразу всей нейросети (рис. 3).



а) стандартная нейросеть

б) нейросеть после «dropout»

Рисунок 3 – Результат применения метода «dropout»

На каждой итерации выбирается случайным образом часть нейронов для обучения, а остальные нейроны устанавливаются в значение ноль (отключаются). Аналогичным образом на второй итерации обучение производится с другими нейронами. Эксперименты показывают, что отключение 50% нейронов на скрытых слоях и отключение 25% нейронов на результирующем слое дают при обучении приемлемые результаты.

Для решения проблемы вычислительной сложности («computational complexity») используется нормализация («regularization») – упрощение

архитектуры нейросети насколько это позволяет поставленная задача. Количество нейронных связей увеличивается в геометрической зависимости от количества скрытых слоёв нейросети. Чем больше слоёв в нейросети, тем больше необходимо данных для обучения. Большие объёмы данных для обучения в свою очередь увеличивают время обучения нейросети, например, некоторые задачи требуют непрерывного обучения нейросети в течение нескольких дней.

**Свёрточные нейронные сети.** Свёрточные нейронные сети (Convolutional neural

networks) – класс глубоких искусственных нейронных сетей прямого распространения, позволяющий эффективно обрабатывать изображения. Свёрточные нейронные сети построены по аналогии со зрительной корой головного мозга. На зрительной коре имеются небольшие участки клеток нейронов, которые связаны с определенными местами зрительного поля. За это открытие учёные Хьюбел Д.Х. (D.H. Hubel) и Визель Т.Н. (T.N. Wiesel) удостоились Нобелевской премии по медицине 1981 года. В 1962 году они провели эксперимент, в котором показали, что отдельные клетки нейронов откликались исключительно при наблюдении границ конкретной ориентации. В частности, отдельные нейроны срабатывали, когда подаваемые изображения имели вертикальные линии, а другие нейроны срабатывали при наличии горизонтальных или диагональных линий на изображении. Хьюбел Д.Х. и Визель Т.Н. определили, что все эти нейроны основывают человеческое визуальное восприятие.

В 1988 году французский учёный Ян Лекун (Yann LeCun), используя особенности данного открытия – идею того, что определённые нейроны имеют конкретную задачу, разработал первую архитектуру свёрточной нейронной сети. Название «свёрточная» происходит из-за наличия операции свёртки, суть которой в том, что каждый фрагмент изображения умножается на матрицу (ядро) свёртки поэлементно, а результат суммируется и записывается в аналогичную позицию выходного изображения. В архитектуру свёрточной нейросети заложены априорные знания из предметной области компьютерного зрения: пиксель изображения сильнее связан с соседним пикселем (локальная корреляция) и объект на изображении может встретиться в любой части изображения.

В конце 1990-х гг в компании AT&T Bell Laboratories исследовательская группа во главе с

Яном Лекуном разработала свёрточную нейросеть для распознавания банковских чеков. С помощью данной технологии компания NEC производила распознавания более 10% банковских чеков в США. В настоящее время существует множество разновидностей данной нейросети: LeNet-1, LeNet-4 and Boosted LeNet-4 и LeNet-5. Архитектура LeNet5 стала фундаментальной для глубокого обучения [11].

Углублённой и расширенной версией LeNet является архитектура свёрточной нейросети AlexNet, разработанная в 2012 году канадскими учёными А. Крижевским (Alex Krizhevsky) и И. Сатскевером (Ilya Sutskever). В AlexNet результаты вычислений LeNet масштабированы в гораздо более крупную нейросеть, которая способна изучать намного более сложные объекты и их иерархии. Нейросеть AlexNet оказала большое влияние на развитие машинного обучения и алгоритмы компьютерного зрения. В 2012 году нейросеть AlexNet с большим отрывом выиграла конкурс «ImageNet Large-Scale Visual Recognition Competition» по классификации объектов на изображениях с количеством ошибок 15,3% (против 26,2% у второго места). Нейросеть AlexNet содержит 11 обучаемых слоев (5 свёрточных, 3 слоя субдескрипции и 3 полносвязных), 650 000 нейронов и 60 миллионов параметров [12]. На результирующем слое данные пропускаются через функцию SOFTMAX, которая формирует распределение 1000 меток классов. Функция SOFTMAX является экспоненциальной и представляется в следующем виде:

$$f(x_i) = \frac{e^{x_i}}{\sum_{l=0}^n e^{x_l}}, \text{ где } i = 0, 1, 2, \dots, n. \quad (8)$$

На рис. 4 представлен график функции активации SOFTMAX в среде Matlab R2019b.

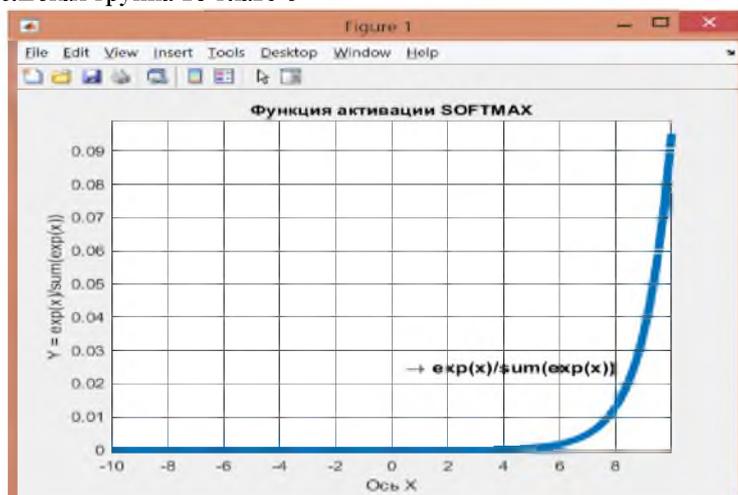


Рисунок 4 – График функции активации SOFTMAX

Функция SOFTMAX применяется в машинном обучении для задач классификации, когда количество возможных классов больше двух. С помощью данной функции формируется вероятностное распределение по всем классам.

Обработка данных в свёрточных нейронных сетях имеет следующие особенности: 1) применение локальных рецепторных полей – на вход одного нейрона подаётся не целое изображение (выходы предыдущего слоя), а только его часть, что позволяет сохранять топологию изображения от слоя к слою; 2) использование разделяемых весовых коэффициентов – ограниченный набор весовых коэффициентов для большого количества нейронных связей; 3) применение пространственной подвыборки – уменьшение размерности, пространственная размерность изображения постепенно уменьшается за счёт слоёв субдискретизации (subsampling).

Свёрточные слои имеют ряд преимуществ перед полносвязными нейросетями:

- для вычисления результата свёрточного слоя требуется меньше объёмы вычислительных ресурсов, чем для полносвязного слоя с такими же размерами входного и выходного массивов;

- свёрточный слой имеет меньше настраиваемых параметров, что уменьшает риск переобучения;

- использование свёрточных слоёв позволяет учесть специфику обрабатываемых данных. Свёрточные слои на первых уровнях детектируют простые объекты (прямолинейные границы, перепады цвета) в разных участках изображения, а на более глубоких уровнях происходит выделение более сложных объектов.

**Разработка архитектуры нейронной сети.** Главный принцип работы скрытых слоёв свёрточной нейросети заключается в разделении признаков: от простых (базовых, локальных) на первых слоях до более сложных на дальних слоях. Простыми признаками являются контуры и углы изображения, а сложные признаки строятся на сочетании простых признаков. Важная особенность глубокого обучения заключается в том, что скрытые слои с признаками не создаются вручную, а извлекаются из обучающих данных постепенно с помощью алгоритма обучения. Увеличение количества слоёв в нейросети позволяет распознавать на изображениях большее количество объектов и элементов, что повышает точность сети. Поскольку изображения в обучающей выборке достаточно сложные и содержат несколько объектов, то для распознавания требуются десятки и

даже сотни скрытых слоёв, поэтому такое машинное обучение называется «глубоким». Однако увеличение количества слоёв в нейросети имеет ряд технических ограничений: слишком большое количество параметров в нейросети может существенно затруднить процесс обучения. На результирующем слое происходит классификация и исходное изображение получает метку выходного класса с определённым коэффициентом уверенности. Этот процесс можно рассматривать как иерархическое обучение, поскольку каждый слой в сети использует выходные данные предыдущих уровней в качестве «строительных блоков» для построения всё более сложных концепций на более высоких уровнях [13-15].

Исследования показали, что применение нейронных сетей особенно эффективно для решения трудно формализуемых задач. Например, во время рейса судоводителю приходится решать задачу идентификации целей для последующего выбора манёвра, который позволит максимально безопасно разойтись с опознаваемым объектом. Поэтому актуальной является разработка системы автоматизированной идентификации морских объектов на основе свёрточной нейронной сети. На рис. 5 представлена разработанная архитектура свёрточной нейронной сети для распознавания пяти типов судов: балкер (bulk cargo carrier), сухогруз (general cargo ship), контейнеровоз (container ship), рыболовная лодка (fishing boat) и пассажирское судно (passenger ship).

На вход нейросети подаётся изображение размера  $N \cdot M$  пикселей. Соответственно количество нейронов на входном слое нейросети для распознавания изображения размера  $400 \cdot 300$  пикселей составляет:

$$InputNeurons = N \cdot M = 400 \cdot 300 = 120000 \quad (9)$$

При цветном 8-битном изображении (RGB) размером  $400 \cdot 300$  пикселей количество параметров одного нейрона составляет:

$$N \cdot M \cdot 3 + bias = 400 \cdot 300 \cdot 3 + 1 = 360001 \quad (10)$$

Если всего 120000 нейронов, тогда общее количество параметров, которое обрабатывает вся нейросеть составляет:

$$InputNeurons \cdot (N \cdot M \cdot 3 + bias) = 120000 \cdot 360001 = 43200120000 \quad (11)$$

На результирующем слое (рис. 5) количество нейронов соответствует типам судов, которые нейросеть должна идентифицировать ( $Y_1, Y_2, Y_3, Y_4, Y_5$ ). Для обучения нейросети выбран алгоритм обратного распространения (back-propagation algorithm), ошибка обучения распространяется по сети (в направлении от результирующего слоя) и тем самым происходит настройка весовых коэффициентов  $w_{N \cdot M}$ . Для входов  $x =$

$(x_1, x_2, \dots, x_{N \cdot M})$  и связанных с ними весовыми коэффициентами  $w = (w_1, w_2, \dots, w_{N \cdot M})$  при  $w_i \geq 0$  функция суммирования определяется следующим образом:

$$u = f(x, w) = \sum_{i=1}^{N \cdot M} (w_i \cdot x_i) \quad (12)$$

При этом функция активации  $\alpha = s(u)$  используется для вычисления уровня активности нейрона. Выходные значения нейрона находятся в диапазоне  $[0 \dots 1]$ . Целью обучения нейронной

сети является нахождение весовых коэффициентов  $w_i$ , где  $i = 1, 2, \dots, N \cdot M$ , которые минимизируют ошибку обучения или целевую функцию нейросети:

$$\varepsilon = \frac{1}{2} \sum_i (\text{output}_i - g(w_i, x_i))^2, \quad (13)$$

где  $\text{output}_i$  – выходное значение нейросети,  $w_i$  – весовой коэффициент,  $g(w_i, x_i)$  – функция обучения.

Обучение нейросети прекращается, когда ошибка больше не уменьшается.

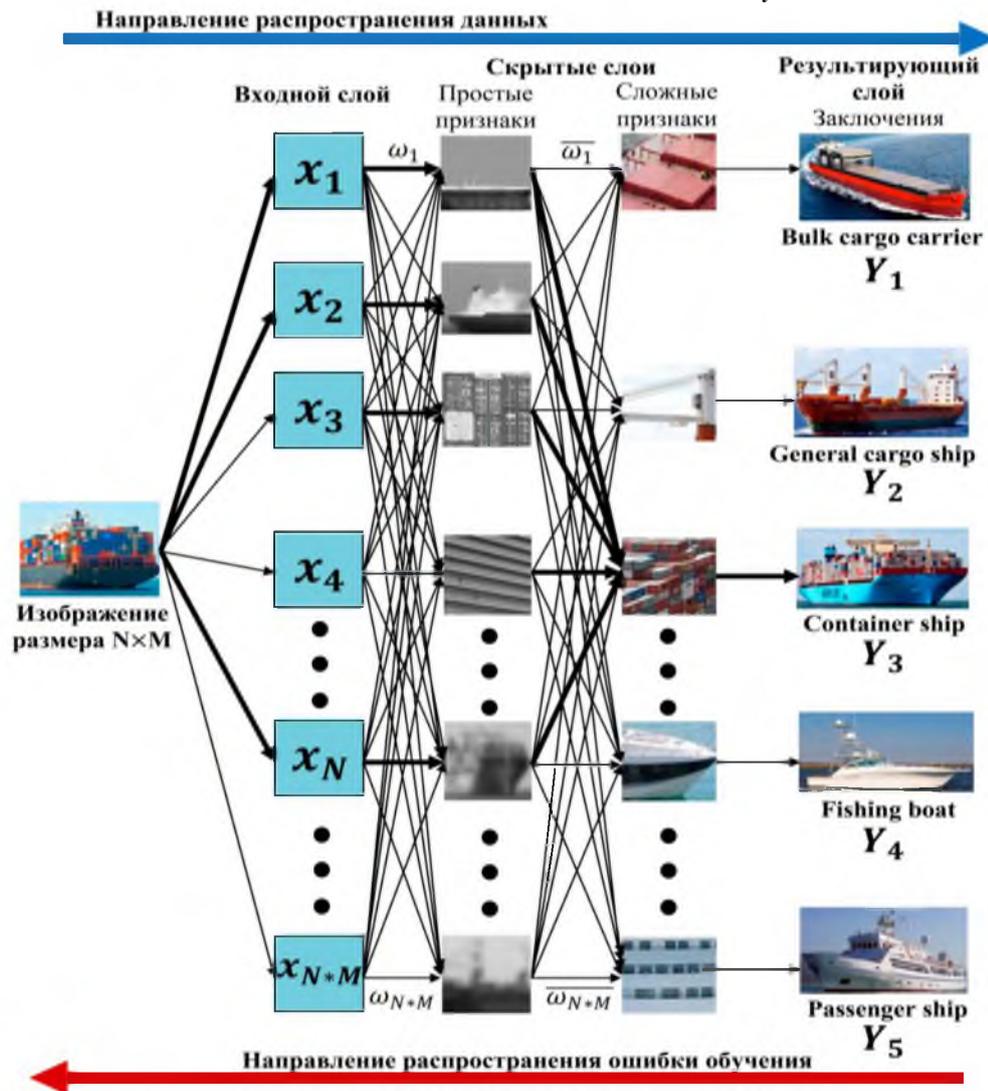


Рисунок 5 – Принцип распознавания пяти типов судов свёрточной нейросетью

**Цветовое пространство RGB.** Предположим, что  $0 \leq R, G, B \leq G_{max}$ . Рассмотрим многоканальное изображение со значениями пикселей  $u = (R, G, B)$ . Число комбинаций цветов составляет  $n^k$ , где  $n$  – количество состояний,  $k$  – количество цветов. Поскольку количество состояний равно 256 (включая 0), количество цветов равно 3, поэтому  $256^3 = 16\,777\,216$  различных цветов, например:  $u = (255, 0, 0)$  – красный,  $u = (255, 255, 0)$  – жёлтый и т. д.

Распознавание объектов выполняется

независимо от масштаба. При этом факт наличия признака важнее места его точного расположения на изображении. Свёрточная нейронная сеть состоит из последовательно расположенных слоёв свёртки и субдискретизации. Это позволяет формировать карты признаков на текущем слое на основании обработки карт признаков предыдущего слоя.

Этапы работы свёрточной нейронной сети:

1. Для каждого поступающего изображения из обучающей выборки окно фиксированного

размера перемещается по всей площади этого изображения, извлекая локальные образцы. Перемещение происходит дискретными шагами.

2. Самоорганизующаяся карта обучается на векторах, полученных на предыдущем этапе. Таким образом, она квантует  $n$ -мерный вектор в одно из топологических значений.

3. Сканирующее окно перемещается по всем изображениям обучающей и тестирующей выборки.

4. Свёрточная нейросеть обучается на новой полученной обучающей выборке.

**Алгоритмы распознавания объектов на изображениях.** В настоящее время существует множество алгоритмов для распознавания объектов на изображениях: SPP-net (Spatial Pyramid Pooling), R-CNN (Region-based Convolutional Neural Networks), Fast R-CNN, Faster R-CNN, Mask R-CNN, SSD (Single Shot MultiBox Detector), R-FCN (Region-based Fully Convolutional Network), YOLO и др. Также используются разновидности алгоритма YOLO: YOLO v2, YOLO9000, YOLO v3.

Алгоритм YOLO (сокращение от «You Only Look Once») – один из наиболее быстрых алгоритмов распознавания объектов на изображениях разработан американскими учёными Д. Редмоном (Joseph Redmon) и А. Фаради (Ali Farhadi) в 2015 году [16]. Данный алгоритм позволяет не только распознавать множество объектов на изображениях, но и в видеопотоке в режиме реального времени со скоростью распознавания 45 кадров в

секунду. Принцип распознавания объектов с помощью алгоритма YOLO:

1) исходное изображение разделяется сеткой ячеек размером  $S \times S$  (каждая ячейка «ответственна» за объект, который в неё попадает);

2) каждая ячейка имеет  $m$  ограничивающих рамок (bounding boxes);

3) для каждой ограничивающей рамки нейросеть выводит класс вероятности и значения смещений;

4) ограничивающие рамки, имеющие вероятность класса выше порогового значения, выбираются и используются для определения местоположения объекта на изображении.

В алгоритме YOLO одна свёрточная нейросеть идентифицирует ограничивающие блоки и вероятности класса для этих блоков. Таким образом, в отличие от других алгоритмов YOLO анализирует изображение целиком без разбиения на области для локализации объекта. Работа алгоритма YOLO начинается с предварительно обученной свёрточной нейросети с целью извлечения характеристик для карты признаков (feature map). Свёрточная нейронная сеть представляет собой чередование свёрточных слоев (convolution layers) и субдискретизирующих слоев (subsampling layers) при наличии полносвязных слоев (fully-connected layer) на выходе. Все три вида слоёв могут чередоваться в произвольном порядке (рис. 6).

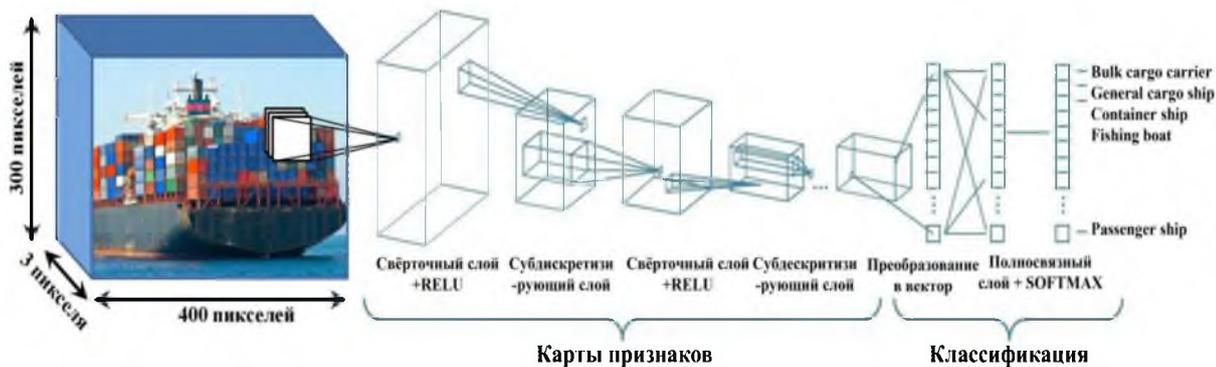


Рисунок 6 – Архитектура свёрточной нейронной сети для распознавания пяти типов судов

В качестве первого слоя нейросети всегда выступает свёрточный слой. Название «свёрточный» происходит из-за сходства операции свёртки в функциональном анализе, который часто применяется в обработке сигналов и теории вероятности. Математическая операция свёртки (convolution operation) выполняется для двух функций  $f$  и  $g$  записывается как  $(f \cdot g)$ . Результатом выполнения операции свёртки для области определения  $n$ -мерного тензора  $t$  определяется как:

$$(f \cdot g) \cdot (t) = \int_{-\infty}^{\infty} f(\tau)g(t - \tau)d\tau = \int_{-\infty}^{\infty} f(t - \tau)g(\tau)dt \tag{14}$$

Рассмотрим дискретную версию выполнения операции свёртки:

$$(f \cdot g) \cdot (i) = \sum_{k=-\infty}^{\infty} f(k)g(i - k) = \sum_{k=-\infty}^{\infty} f(i - k)g(k) \tag{15}$$

Часто операция свертки рассматривается как использование функции  $g$  для преобразования функции  $f$ . В этом контексте функция  $g$  иногда называется ядром (англ. «kernel»). Когда ядро конечно (ограничено), то  $g(k)$  определяется только для  $k = 0, 1, \dots, m - 1$  и выражение упрощается до следующего вида:

$$(f \cdot g)(i) = \sum_{k=0}^{m-1} f(i - k)g(k) \quad (16)$$

Вводимое изображение представляет матрицу размера, например,  $400 \cdot 300 \cdot 3$  с пиксельными значениями. Фильтр представляет собой матрицу (называемую также матрицей весов или матрицей параметров) размером, например,  $5 \cdot 5 \cdot 3$ . Фильтр производит операцию свертки, т.е. умножает элементы фильтра на значения пикселей исходного изображения (поэлементное умножение). После умножения все числа суммируются, и в итоге получается одно число (для фильтра  $5 \cdot 5 \cdot 3$  всего 75 поэлементных умножений). Фильтр движется по всей области вводного изображения, а области, на которые взаимодействует фильтр, называются рецептивным полем (полем восприятия). В итоге после прохода фильтра по всей области (движение с шагом один) получается новая матрица. Особенностью свёрточного слоя является сравнительно небольшое количество параметров, устанавливаемое при обучении. Если при исходном изображении размером  $400 \cdot 300$  пикселей по трём RGB каналам (120000 входных нейронов), свёрточный слой использует фильтры размером  $5 \cdot 5$  пикселя с выходом на 5 каналов, тогда в процессе обучения определяется только 25 весовых коэффициентов фильтра. По

всем сочетаниям каналов ( $5 \cdot 5 \cdot 3 \cdot 5 = 375$ ) данный слой требует нахождения только 375 параметров, что существенно меньше количества искомым параметров ( $400 \cdot 300 \cdot 3 \cdot 5 = 1800000$ ) полносвязной нейронной сети.

В архитектуре свёрточной нейронной сети применяются слои субдескрипции (pooling layer), чередующиеся со свёрточными слоями. Основная задача слоёв субдескрипции состоит в последовательном уменьшении пространственных габаритов (разрешения) изображения. Это необходимо для сокращения количества входных параметров для следующего слоя, а также для снижения вычислительных операций и контроля обучаемости нейросети. Последним слоем свёрточной нейронной сети является полносвязная нейронная сеть. Входными данными для полносвязной нейронной сети являются предыдущие слои. В итоге свёрточная нейросеть на каждом слое преобразования трансформирует исходное изображение. Преобразование начинается с первоначальных значений исходного изображения и заканчивается определением класса изображения (типа судна): bulk cargo carrier, general cargo ship, container ship, fishing boat, passenger ship.

**Обучение нейронной сети.** Существуют несколько парадигм обучения нейронных сетей: обучение с учителем, обучение с подкреплением, обучение без учителя, обучение с частичным привлечением учителя и предобучение без учителя. На рис. 7 представлены три модели обучения, когда: 1) данные для обучения размечены, 2) большая часть данных не размечена, 3) все данные для обучения не размечены.



Рисунок 7 – Парадигмы обучения нейронных сетей

Обучение с учителем (supervised learning) – обучение с обратной связью, при котором исходные данные для обучения размечены и известны правильные ответы. Весовые коэффициенты нейросети настраиваются так, чтобы полученные

ответы были как можно ближе к известным значениям. Ошибка обучения нейросети должна быть минимальной.

Обучение с подкреплением (reinforcement learning) – является частным случаем обучения с

учителем, но учителем является среда или её модель. Откликом среды на принятые решения являются сигналы подкрепления. Обучение с подкреплением считается одним из основных путей к созданию «сильного» искусственного интеллекта. Однако, недостаток заключается в том, что современные алгоритмы обучения с подкреплением требуют для решения задач примерно столько же времени, как и обычный случайный поиск [17]. Обучение без учителя или самообучение (unsupervised learning) – данные для обучения не размечены и не известны правильные ответы. Обучение с частичным привлечением учителя или полуавтоматическое обучение (semi-supervised learning) – является смешанной формой обучения, разновидность обучения с учителем, нечто среднее между обучением с учителем и без учителя. Данный тип обучения использует большое количество неразмеченных данных в сочетании с небольшим количеством размеченных данных, что позволяет значительно улучшить точность

обучения. Предобучение без учителя (unsupervised pretraining) – позволяет получить модель, которую для получения лучших решений необходимо «дообучить» на размеченных данных [18].

**Источники данных для обучения нейросети.** Для эффективного обучения нейросети необходимо иметь большое количество изображений. В настоящее время существует множество открытых баз с размеченными изображениями (датасетами), например, ImageNet, PASCAL VOC, COCO (Common Objects in Context), CIFAR-10, CIFAR-100, MASATI, Caltech-256 и др. База изображений ImageNet содержит набор из 15 миллионов размеченных изображений с высоким разрешением, разделённых на 22 000 категорий [19-21]. На рис. 8 представлены примеры изображений базы ImageNet из категории «container ship» (содержит 1613 изображений контейнерных судов).



Рисунок 8 – Примеры изображений из базы ImageNet

База изображений MASATI (MAritime SATellite Imagery dataset) содержит цветные изображения морских аэрофотоснимков и используется для оценки методов обнаружения судов (рис. 9).

Каждое изображение может содержать одну или несколько морских объектов в разных

погодных условиях и условиях освещения. База MASATI состоит из 7389 спутниковых изображений (на июнь 2019 года), помеченных в соответствии со следующими семью классами: суша, побережье, море, судно, несколько морских объектов, берег-судно и детализация судов.



Рисунок 9 – Примеры изображений из базы MASATI

**Реализация и результаты работы свёрточной нейронной сети.** В результате исследования разработана архитектура свёрточной нейронной сети применительно к задаче распознавания и классификации морских объектов на изображениях. На основе свёрточной нейронной сети разработана компьютерная

программа для распознавания морских объектов. На рис. 10 представлен результат разработанной программы по распознаванию морских объектов у памятника «Малая Земля» в г. Новороссийске. Программа разработана в среде MS Visual Studio 2019 на языке программирования Visual C#

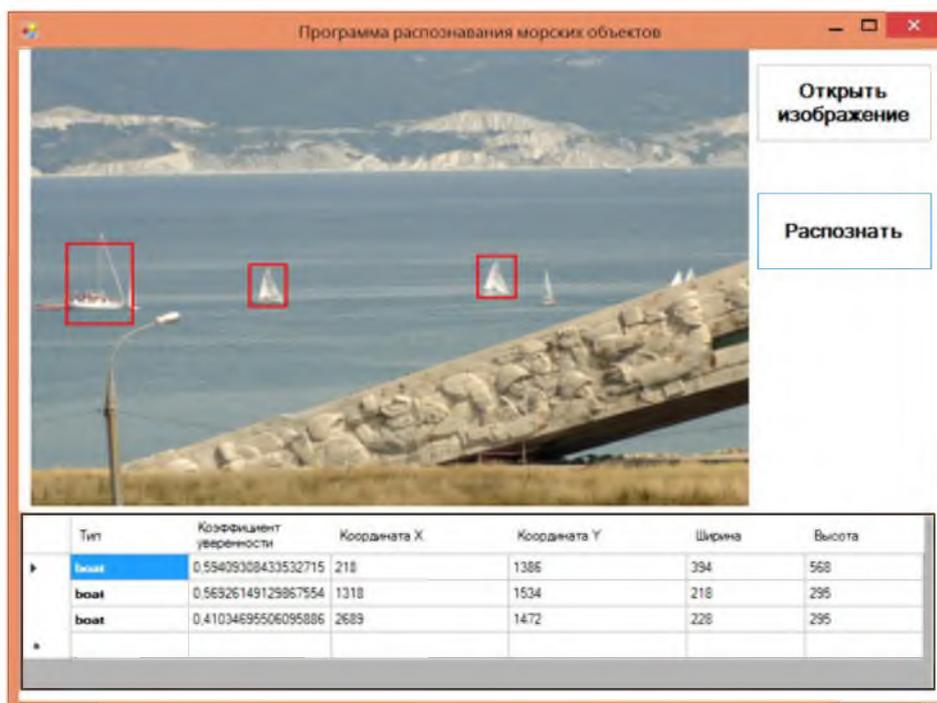


Рисунок 10 – Результат распознавания морских объектов

Программа позволяет загрузить изображение, а кнопка «Распознать» запускает процесс распознавания с помощью обученной нейронной сети. Обнаруженные морские объекты на изображении выделяются в прямоугольник, а в таблице выводятся типы распознанных объектов, коэффициенты уверенности и координаты объектов на изображении. Таким образом, программа производит детектирование (нахождение экземпляров объектов на изображении) и распознавание объектов (определение наличия объектов и их расположение на изображении). К достоинствам программы можно отнести низкую ресурсоёмкость применяемых алгоритмов, а также высокую вероятность обнаружения объектов при низкой вероятности ложных срабатываний.

**Заключение.** В ходе исследования выполнен ретроспективный анализ технологий компьютерного зрения на основе глубоких свёрточных нейронных сетей. Развитие алгоритмов и технологий обучения глубоких нейронных сетей открыло ряд проблем: «исчезающий» градиент, переобучение и вычислительная сложность. Применение нейронных сетей особенно эффективно для решения трудно формализуемых задач. Поэтому

разработана архитектура свёрточной нейронной сети применительно к задаче распознавания и классификации морских объектов на изображениях. Сравнение используемых в нейросети функций активации ReLU, LeakyReLU, Exponential ReLU и SOFTMAX выполнено в среде Matlab R2019b. На основе свёрточной нейронной сети разработана программа на языке программирования Visual C# в среде MS Visual Studio 2019. Программа предназначена для автоматизированной идентификации морских объектов, производит детектирование (т.е. нахождение объектов на изображении) и распознавание объектов с высокой вероятностью обнаружения.

#### Литература

1. Астреин В.В., Кондрачев С.И., Хекерт Е.В. Алгоритм самоорганизации групп судов для предупреждения столкновений // Эксплуатация морского транспорта. Гос. морской университет им. адмирала Ф.Ф. Ушакова, Новороссийск. – 2016.– №2(79). – С. 45-50.
2. Полковникова Н.А., Полковников А.К. Система поддержки принятия решений для выбора режима и прогнозирования отказов главного судового двигателя // Эксплуатация морского

- транспорта. Гос. морской университет им. адмирала Ф.Ф. Ушакова, Новороссийск. – 2019. – №3(92). – С. 170-180.
3. Пятакович В.А., Пятакович Н.В., Пашкеев С.В. Применение нейросетевых технологий искусственного интеллекта в структуре системы мониторинга и контроля морской среды средствами военно-морского флота // Стратегическая стабильность. – 2018. – № 3 (84). – С. 53-62.
  4. Marr B., Ward M. Artificial intelligence in practice. – Wiley, 2019. – 340 p.
  5. Клетте Р. Компьютерное зрение. Теория и алгоритмы / пер. с англ. А.А. Слинкин. – М.: ДМК Пресс, 2019. – 506 с.
  6. Krohn J., Grant B., Aglae B. Deep learning illustrated: a visual interactive guide to artificial intelligence. – Addison-Wesley Professional, 2019. – 416 p.
  7. Rebal G., Ravi A., Sanjay C. An introduction to machine learning. – Springer, 2019. – 263 p.
  8. Harley A.W. An interactive node-link visualization of convolutional neural networks. *International Symposium on Visual Computing, Springer*, 2015. – 867-877 pp.
  9. Wei Qi Yan. Introduction to intelligent surveillance. – Springer, 3<sup>rd</sup> edition, 2019. – 222 p.
  10. LeCun Y. Deep Learning Hardware: Past, Present, and Future. *IEEE International Solid-State Circuits Conference-(ISSCC)*, 2019. – 12-19 pp.
  11. LeCun Y., Bengio Y., Hinton G. Deep learning. *Nature*, vol. 521 (7553), 2015. – 436-444 pp.
  12. Krizhevsky A., Sutskever I., Hinton G.E. Imagenet classification with deep convolutional neural networks. *Advances in neural information processing systems*, 2012. – 1097-1105 pp.
  13. Shanmugamani R. Deep Learning for Computer Vision: Expert techniques to train advanced neural networks using Tensorflow and Keras. – Packt Publishing, 2018. – 312 p.
  14. Thakkar M. Beginning machine learning in iOS: CoreML Framework. – Apress, 2019. – 157 p.
  15. Leskovec J., Rajaraman A., Ullman J.D. Mining of Massive Datasets. – Cambridge University Press, 3<sup>rd</sup> edition, 2019. – 583 p.
  16. Redmon J., Farhadi A. YOLO9000: Better, Faster, Stronger. *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2017. – 7263-7271 pp.
  17. Deep Reinforcement Learning Doesn't Work Yet <https://www.alexirpan.com/2018/02/14/rl-hard.html> (Дата обращения 16.12.2019)
  18. Николенко С., Кадурич А., Архангельская Е. Глубокое обучение. – СПб.: Питер, 2018. – 480 с.
  19. Kwanghyun K., Sungjun H., Baehoon C., Euntai K. Probabilistic Ship Detection and Classification Using Deep Learning. *Applied Sciences*, vol. 8, no. 936, 2018. – 1-17 pp.
  20. Zhenfeng Shao, Wenjing Wu, Zhongyuan Wang, Wan Du, Chengyuan Li. SeaShips: a large-scale precisely annotated dataset for ship detection. *IEEE transactions on multimedia*, vol. 20, no. 10, 2018. – 2593-2604 pp.
  21. Recent advances on memetic algorithms and its applications in image processing. Editors: Hemanth D.J., Kumar B.V., Karpagam Manavalan G.R. – Springer, 2020. – 199 p.

## REFERENCES

1. Astrein V.V., Kondrat'ev S.I., Khekert E.V. Algoritm samoorganizatsii grupp sudov dlya preduprezhdeniya stolknoveniy // Ekspluatatsiya morskogo transporta. Gos. morskoy universitet im. admirala F.F. Ushakova, Novorossiysk. – 2016, №2. – S. 45-50.
2. Polkovnikova N.A., Polkovnikov A.K. Sistema podderzhki prinyatiya resheniy dlya vybora rezhima i prognozirovaniya otkazov glavnogo sudovogo dvigatelya // Ekspluatatsiya morskogo transporta. Gos. morskoy universitet im. admirala F.F. Ushakova, Novorossiysk. – 2019, №3(92). – S. 170-180.
3. Pyatakovich V.A., Pyatakovich N.V., Pashkeev S.V. Primenenie neyrosetevykh tekhnologiy iskusstvennogo intellekta v strukture sistemy monitoringa i kontrolya morskoy sredy sredstvami voenno-morskogo flota // Strategicheskaya stabil'nost', 2018, № 3 (84). – S. 53-62.
4. Marr B., Ward M. Artificial intelligence in practice. – Wiley, 2019. – 340 p.
5. Klette R. Komp'yuternoe zrenie. Teoriya i algoritmy / per. s angl. A.A. Slinkin. – M.: DMK Press, 2019. – 506 s.
6. Krohn J., Grant B., Aglae B. Deep learning illustrated: a visual interactive guide to artificial intelligence. – Addison-Wesley Professional, 2019. – 416 p.
7. Rebal G., Ravi A., Sanjay C. An introduction to machine learning. – Springer, 2019. – 263 p.
8. Harley A.W. An interactive node-link visualization of convolutional neural networks. *International Symposium on Visual Computing, Springer*, 2015. – 867-877 pp.
9. Wei Qi Yan. Introduction to intelligent surveillance. – Springer, 3<sup>rd</sup> edition, 2019. – 222 p.
10. LeCun Y. Deep Learning Hardware: Past, Present, and Future. *IEEE International Solid-State Circuits Conference-(ISSCC)*, 2019. – 12-19 pp.
11. LeCun Y., Bengio Y., Hinton G. Deep learning. *Nature*, vol. 521 (7553), 2015. – 436-444 pp.
12. Krizhevsky A., Sutskever I., Hinton G.E. Imagenet classification with deep convolutional neural networks. *Advances in neural information processing systems*, 2012. – 1097-1105 pp.
13. Shanmugamani R. Deep Learning for Computer Vision: Expert techniques to train advanced neural networks using Tensorflow and Keras. – Packt Publishing, 2018. – 312 p.
14. Thakkar M. Beginning machine learning in iOS: CoreML Framework. – Apress, 2019. – 157 p.

15. Leskovec J., Rajaraman A., Ullman J.D. Mining of Massive Datasets. – Cambridge University Press, 3rd edition, 2019. – 583 p.
16. Redmon J., Farhadi A. YOLO9000: Better, Faster, Stronger. *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2017. – 7263-7271 pp.
17. Deep Reinforcement Learning Doesn't Work Yet <https://www.alexirpan.com/2018/02/14/trl-hard.html> (Data obrashcheniya 16.12.2019)
18. Nikolenko S., Kadurin A., Arkhangel'skaya E. Glubokoe obuchenie. – SPb.: Piter, 2018. – 480 s.
19. Kwanghyun K., Sungjun H., Baehoon C., Euntai K. Probabilistic Ship Detection and Classification Using Deep Learning. *Applied Sciences*, vol. 8, no. 936, 2018. – 1-17 pp.
20. Zhenfeng Shao, Wenjing Wu, Zhongyuan Wang, Wan Du, Chengyuan Li. SeaShips: a large-scale precisely annotated dataset for ship detection. *IEEE transactions on multimedia*, vol. 20, no. 10, 2018. – 2593-2604 pp.
21. Recent advances on memetic algorithms and its applications in image processing. Editors: Hemanth D.J., Kumar B.V., Karpagam Manavalan G.R. – Springer, 2020. – 199 p.

УДК 621.3

DOI: 10.34046/aumsuomt94/30

## ВЕРОЯТНОСТНАЯ МОДЕЛЬ ОЦЕНКИ ЭКСПЛУАТАЦИОННОЙ НАДЁЖНОСТИ БЕРЕГОВОЙ СТАНЦИИ АИС КАК ПОДСИСТЕМЫ СУДС В УСЛОВИЯХ ВНЕДРЕНИЯ КОНЦЕПЦИИ E-NAVIGATION

*Н.В. Старжинская, кандидат технических наук, доцент  
А.И. Чернова, кандидат технических наук, доцент*

Современные информационные системы связи являются составной частью береговой инфраструктуры, обеспечивающей безопасность мореплавания и являющейся неотъемлемым компонентом внедряемой концепции e-Navigation, которую можно считать перспективным путем развития информационных технологий на морском транспорте. В связи с этим обеспечение надёжности берегового радиооборудования связи, входящего в состав таких систем как СУДС, является одной из важнейших эксплуатационных задач. В статье рассмотрено оборудование береговой станции АИС, являющейся компонентом СУДС, проведён анализ особенностей его эксплуатации и функционирования. Предложена математическая модель, позволяющая провести количественную оценку влияния основных параметров, характеризующих процесс функционирования и условия применения оборудования базовой станции АИС на его готовность.

**Ключевые слова:** надёжность, математическое моделирование, береговое радиооборудование, АИС, e-Навигация, готовность к применению, динамика функционирования, марковский процесс.

Modern information communication systems are an integral part of the coastal infrastructure that ensures the safety of navigation and is an integral component of the implemented e-Navigation concept, which can be considered a promising way to develop information technologies in Maritime transport. In this regard, ensuring the reliability of coastal radio communication equipment, which is part of such systems as VTS, is one of the most important operational tasks. The article considers the equipment of the coastal AIS station, which is a component of the VTS, and analyzes the features of its operation and functioning. A mathematical model is proposed that allows us to conduct a quantitative assessment of the influence of the main parameters that characterize the process of functioning and conditions for using the AIS base station equipment on its readiness.

**Key words:** reliability, mathematical modeling, coastal radio equipment, AIS, e-Navigation, readiness for use, dynamics of functioning, Markov process.

Развитие современных информационных технологий и систем связи охватило практически все сферы человеческой деятельности. Они стали неотъемлемой частью современного мира, существенно изменив роль человека-оператора в сложных технических системах. В том числе такие системы были внедрены и на морском транспорте для выполнения функций предупреждения и предотвращения аварийных ситуаций на море, что особенно актуально в условиях стремительного роста интенсивности судоходства, увеличения тоннажа морских судов и грузооборота морских портов. К таким системам можно отнести:

Глобальную морскую систему связи при бедствии и для обеспечения безопасности (GMDSS – Global Maritime Distress Safety System), автоматические идентификационные системы (AIS – Automatic Identification System), Системы управления движением судов (СУДС) и др. В совокупности эти системы являются составной частью береговой инфраструктуры, обеспечивающей безопасность мореплавания и являющейся неотъемлемым компонентом внедряемой концепции e-Navigation, которую можно считать перспективным путем развития информационных технологий на морском транспорте [1].