

МНОГОМАСШТАБНАЯ СЕГМЕНТАЦИЯ ИЗОБРАЖЕНИЙ В РЕШЕНИИ ЗАДАЧ ОБНАРУЖЕНИЯ РАЗЛИВОВ НЕФТЕПРОДУКТОВ НА ОСНОВЕ СВЕРТОЧНЫХ НЕЙРОННЫХ СЕТЕЙ

DOI: 10.36724/2072-8735-2025-19-6-16-24

Васильева Дина ВладимировнаСанкт-Петербургский Государственный университет
аэрокосмического приборостроения, Санкт-Петербург, Россия,
dolli.dina@mail.ru**Дворников Сергей Викторович**Санкт-Петербургский Государственный университет
аэрокосмического приборостроения, Санкт-Петербург, Россия,
practicdsv@yandex.ru**Дворников Сергей Сергеевич**Санкт-Петербургский Государственный университет
аэрокосмического приборостроения, Санкт-Петербург, Россия,
dvornik92@mail.com**Manuscript received** 30 April 2025;**Accepted** 20 May 2025

Ключевые слова: сверточные нейронные сети,
обнаружение разливов нефтепродуктов,
многомасштабная сегментация изображений,
экологический мониторинг, разлив
нефтепродуктов

Эффективный экологический мониторинг акваторий морей и океанов является сложным из-за огромных размеров контролируемых территорий. Поэтому для обработки получаемого первичного фото и видео материала при решении задач обнаружения разливов нефтепродуктов, как правило, используют методы машинного обучения. Высокая вариативность водной поверхности мирового океана и большой объем каждого кадра изображения затрудняют эффективную работу нейронных сетей. Для этого предлагается использовать многомасштабную сегментацию изображения, в соответствии с алгоритмом К-средних, что позволит получить несколько однородных фрагментов, распознавание которых обеспечит решение задачи обнаружения разливов нефтепродуктов. А для эффективной обработки выделенных объектов распознавания, представленных в виде сегментов, и расширения возможностей фильтров используется пространственная свертка, с последующим комплексным объединением пространственных пирамид. Искомая модель строится на сверточной нейронной сети. Использование представленного решения позволяет снизить вычислительную нагрузку на сверточную нейронную сеть, а также повысить достоверность обнаружения результатов разливов нефтепродуктов при обработке изображений с высокой вариативностью параметров водной поверхности. Представлена сравнительная оценка предложенного подхода по отношению к модели, реализующей технологию бинарной семантической сегментации на основе комплексного пространственного пирамидального объединения, получившей широкую научную апробацию. Представленное решение предлагается реализовать в виде модели, реализуемой сетью CNNs. Предлагаемая модель позволяет снизить вычислительную нагрузку сети примерно в 6 раз по отношению к каноническому подходу, при этом обеспечивая совпадение результатов обнаружения с данными визуального анализа в 67% случаев (эксперимент проводился по 50 снимков водных акваторий, на 30 из которых были запечатлены разливы нефтепродуктов, при этом изображения не содержали сторонних объектов), что почти в 1,5 раза превышает возможности апробированной модели.

Информация об авторах:

Васильева Дина Владимировна, старший преподаватель кафедры радиотехнических систем Института радиотехники и инфокоммуникационных технологий, Санкт-Петербургский государственный университет аэрокосмического приборостроения, Санкт-Петербург, Россия. <http://orcid.org/0009-0009-5343-3356>

Дворников Сергей Викторович, доктор технических наук, профессор кафедры радиотехнических и оптоэлектронных комплексов, Санкт-Петербургский государственный университет аэрокосмического приборостроения; профессор кафедры радиосвязи, Военная академия связи им. С.М. Буденного, Санкт-Петербург, Россия. <http://orcid.org/0000-0002-4889-0001>

Дворников Сергей Сергеевич, кандидат технических наук, доцент кафедры конструирования и технологий электронных и лазерных средств, Санкт-Петербургский государственный университет аэрокосмического приборостроения; научный сотрудник научно-исследовательского центра, Военная академия связи им. Маршала Советского Союза С.М. Буденного, Санкт-Петербург, Россия. <https://orcid.org/0000-0001-7426-6475>

Для цитирования:

Васильева Д.В., Дворников С.В., Дворников С.С. Многомасштабная сегментация изображений в решении задач обнаружения разливов нефтепродуктов на основе сверточных нейронных сетей // Т-Comm: Телекоммуникации и транспорт. 2025. Том 19. №6. С. 16-24.

For citation:

D.V. Vasilieva, S.V. Dvornikov, S.S. Dvornikov, "Multi-scale image segmentation in solving problems of oil spill detection on the basis of convolutional neural networks," T-Comm, 2025, vol. 19, no.6, pp. 16-24. (in Russian)

Введение

Сложность решения задач экологического мониторинга во многом обусловлена громадными размерами контролируемых территорий [1, 2]. В такой ситуации целесообразно использование летно-подъемных платформ (ЛПП) [3,4], для размещения на них аппаратуры видового контроля подстилающей поверхности [5]. При этом получаемый фото и видеоматериал, как правило, в полном объеме посредством телекоммуникационного оборудования транслируется на стационарные пункты сбора информации, что приводит к чрезмерной перегрузке телекоммуникационных каналов.

В такой ситуации рациональным решением видится предварительная селекция получаемых кадров изображений непосредственно на борту ЛПП, что позволит существенно снизить объем транслируемого трафика, представляемого для последующей обработки в интересах обнаружения последствий разлива нефтепродуктов [6].

При этом существенные различия водных поверхностей контролируемых акваторий даже в отсутствии загрязнений, а также структурная вариативность самого разлива, существенно усложняют задачу его обнаружения.

Рассмотренная проблема усугубляется отсутствием в достаточной степени апробированного научно-методический аппарат, применение которого однозначно позволило бы получить искомый результат [8, 9].

Указанные обстоятельства определяют необходимость поиска приемлемого научно-технического подхода, позволяющего успешно разрешить обозначенные проблемы, с учетом последующей автоматизации реализующих его процедур.

В настоящее время на практике выделяют два фундаментальных подхода, обеспечивающих автоматизацию процедур поиска объектов на изображениях [10]:

первый – базируется на результаты статистической обработки параметров изображений;

второй – основан на использовании методов машинного обучения.

И если первый подход достаточно глубоко изучен, благодаря фундаментальной работе К. Фукунаги «Введение в статистическую теорию распознавания образов» [11], то второе направление пока еще динамично развивается.

В рамках настоящей статьи представлена модель работы глубокой сверточной сети, обеспечивающей обнаружение разливов нефтепродуктов по результатам обработки предварительно сегментированных изображений водных акваторий.

Анализ технологий машинного обучения

В соответствии с [12], под технологиями машинного обучения понимают разновидности форм искусственного интеллекта, основанные на применении нейронных сетей.

В свою очередь применение нейронных сетей предполагает разработку и реализацию процедур, направленных на уточнение содержания основных этапов реализации алгоритма ее работы, по результатам текущей обработки выходных данных, получаемых в ходе их функционирования. То есть в процессе функционирования самого алгоритма работы нейронной сети, непосредственно реализуются процедуры по ее «обучению».

Такой подход позволяет базовому алгоритму адаптивно изменять свою работу в ходе обработки текущих данных, в соответствии с особенностями выявленных в них закономерностей. При этом следует понимать, что эффективность работы алгоритма при таком подходе во многом обеспечивается качеством предварительного проведения первичного обучения нейронной сети.

В настоящее время принято выделять следующие методы машинного обучения [12]:

- с учителем;
- без учителя;
- с подкреплением;
- и с частичным привлечением учителя.

Вместе с тем в работе [13] показано, что для обнаружения разливов нефтепродуктов, как правило, используются методы глубокого обучения нейронных сетей, получивших название глубокая нейронная сеть (англ.: *Deep learning*).

А среди класса глубоких нейронных сетей (англ.: *Deep Neural Networks* – DNNs) наилучшая эффективность в вопросах классификации изображений отмечена у следующих типов DNNs:

- многослойные перцептроны (англ.: *Multilayer Perceptron* – MLP);
- сверточные нейронные сети (англ.: *Convolutional Neural Network* – CNNs);
- рекуррентные нейронные сети (англ.: *Recurrent Neural Network* – RNN).

При этом следует понимать, что главное отличие DNNs от обычных нейронных сетей определяется большим количеством задействованных слоев перцептронов, и большим количеством связей на уровне каждого слоя.

Рассмотрим возможности основных типов DNNs по решению задач обнаружения разливов нефтепродуктов.

Сети MLP являются базовыми для DNNs и относятся к классу нейронных сетей с прямой связью. Концепция построения сетей MLP демонстрируется на рисунке 1 (здесь и далее в качестве исходных данных представлены кадры изображения водной поверхности с разливами нефтепродуктов, используемые в эксперименте). (Gaudenz Boesch, «Deep Learning». Deep Neural Network: The 3 Popular Types (MLP, CNN and RNN): <https://viso.ai/deep-learning/deep-neural-network-three-popular-types/> (Свободный доступ на 12.02.2025)).

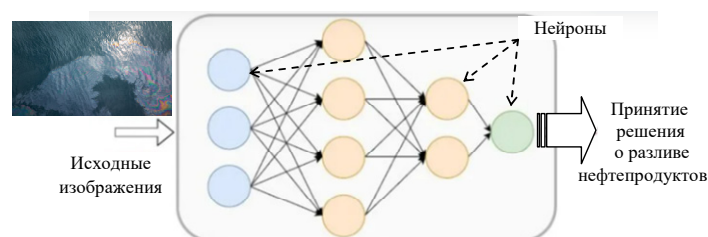


Рис. 1. Концепция построения нейронных сетей типа MLP

Основу построения нейронных сетей типа MLP составляют нейроны, каждый из которых представляет собой математическую модель передаточной функции взвешенного сумматора, которую можно записать в следующем виде:

$$y = f \left[w_0 x_0 + \sum_{i=1}^N w_i x_i \right], \quad (1)$$

где x_i – вектор исходных данных, поступающий на вход нейрона; w_i – значения весовых коэффициентов, определяющих текущий переход вектора данных x_i на последующий слой обработки.

В выражении (1) x_0 и w_0 – вектор данных инициализации нейрона и соответствующий ему весовой коэффициент.

Аргумент передаточной функции $f(*)$ в [14] определен как функция индуцированного локального поля от переменного показателя N :

$$u(N) = w_0 x_0 + \sum_{i=1}^N w_i x_i. \quad (2)$$

При этом область определения передаточной функции $f(*)$ лежит в пределах значений от 0 до 1.

Учитывая, что передаточная функция определяет характер преобразования входного вектора данных в зависимости от значений ее весовых коэффициентов, то выходные значения $f(*)$ будут лежать в интервале $(-1; 1)$.

Инициализация нейрона посредством воздействия начального элемента вектора x_0 на дополнительный вход обеспечивает смещение активационной функции, тем самым формируя его порог чувствительности персептрона.

В MLP сетях нейроны принято разделять на входные, выходные и промежуточные.

И если, как правило, входные и выходные нейроны служат для ввода и вывода данных, то промежуточные нейроны выполняют все расчетные операции.

Характер вычислений определяется типом передаточной функции.

Из всего многообразия используемых передаточных функций наибольшее распространение получила так называемая сигмоида, формируемая на основе логистической функции вида [12]:

$$\sigma(x) = \frac{1}{1 + \exp(-tx)}. \quad (3)$$

Характер изменения функции (3) определяет параметр t , который формирует крутизну спада сигмоиды.

В качестве примера на рисунке 2 показаны изображения сигмоид при различных значениях t , область определения которого лежит в пределах $(0; 1)$.

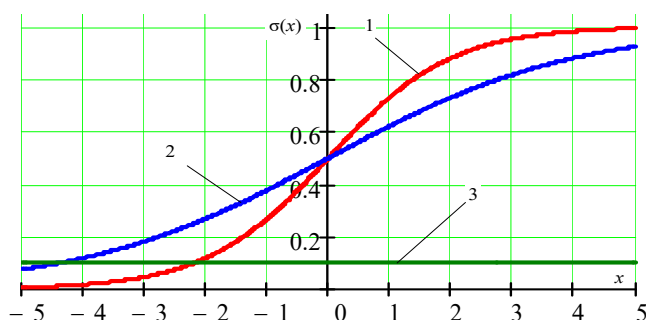


Рис. 2. Изображение сигмоиды при различных значениях параметра t

На рисунке 2 цифрами обозначены значения сигмоиды, соответствующие следующим значениям параметра t : $1 - t = 1$; $2 - t = 0,5$; $3 - t = 0,1$.

Учитывая, что именно параметр t является результатом суммы взвешенных коэффициентов, то, анализируя графики, представленные на рисунке 2, можно утверждать следующее.

При значениях $w_i \rightarrow 0$, обусловленных или малым размером сети, или низкой контрастностью входных данных, передаточная функция искусственного нейрона фактически не вносит изменений после их обработки. Поэтому такая сеть не сможет обеспечить высокую разрешающую способность.

Наибольшая эффективность практических приложений нейронных сетей вида MLP отмечена при обработке структурированных данных, используемых, в частности, в финансовой сфере при оценке кредитных историй, анализе транзакций при предотвращении несанкционированного доступа, а также в вопросах прогнозирования. (Gaudenz Boesch, «Deep Learning». Deep Neural Network: The 3 Popular Types (MLP, CNN and RNN): <https://viso.ai/deep-learning/deep-neural-network-three-popular-types/> (Свободный доступ на 12.02.2025)).

При этом увеличение размеров сети, в свою очередь определяемого размерностью вектора входных данных, ведет к существенному повышению сложности ее структуры, связанной с увеличением взаимосвязей между нейронами. И как результат – к экспоненциальному возрастанию потребного вычислительного ресурса, необходимого для обучения и функционирования таких сетей.

Именно поэтому в настоящее время размеры обрабатываемых изображений посредством нейронных сетей вида MLP существенно ограничены.

Сверточные нейронные сети CNNs успешно применяются в решении задач компьютерного зрения.

Работа таких сетей основана на поиске в потоке данных предварительно заданных признаков, поэтому в отличие от технологии, заложенной в сетях вида MLP, в моделях CNNs реализуется принцип последовательной реализации процедур свертки на каждом уровне для извлечения возможных признаков распознаваемого объекта, см. рис. 3. (Gaudenz Boesch, «Deep Learning». Deep Neural Network: The 3 Popular Types (MLP, CNN and RNN): <https://viso.ai/deep-learning/deep-neural-network-three-popular-types/> (Свободный доступ на 12.02.2025)).

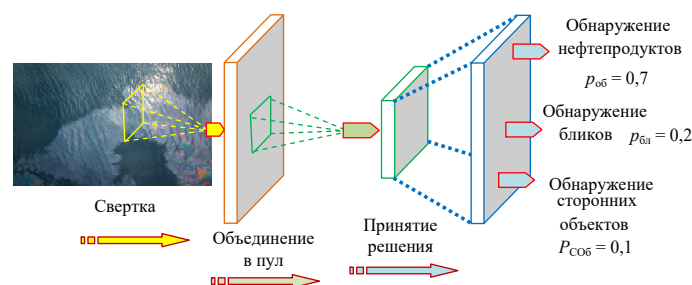


Рис. 3. Особенности построения и функционирования сетей вида CNNs

Так как каждый из уровней обработки в моделях на основе CNNs представляет собой совокупность нелинейных функций предварительно рассчитываемых взвешенных

сумм, то применительно к каждому из последовательно обрабатываемых фрагментов изображения окончательное решение принимается с учетом введенных весовых коэффициентов, а также результатов обработки, полученных на предыдущем уровне.

Рекуррентные нейронные сети RNNs преимущественно ориентированы на решение задач, связанных с обработкой временных рядов данных.

При этом входные данные таких сетей структурируются таким образом, что на каждый последующий слой поступают не только текущие данные, но и данные, полученные на предыдущих этапах обработки. Принцип работы сети RNN, раскрывающий ее особенности, представлен на рисунке 4. (Gaudenz Boesch, «Deep Learning». Deep Neural Network: The 3 Popular Types (MLP, CNN and RNN): <https://viso.ai/deep-learning/deep-neural-network-three-popular-types/> (Свободный доступ на 12.02.2025)).

Следует отметить, что в сетях RNNs каждый последующий уровень перцептронов представляет собой набор нелинейных функций, весовыми коэффициентами которых являются взвешенные суммы входных данных и данных, полученных в предыдущем состоянии сети.

В результате такой архитектуры, каждый базовый элемент сети RNNs состоит из совокупности слоев и ячеек, реализующих последовательную обработку рекуррентных моделей.

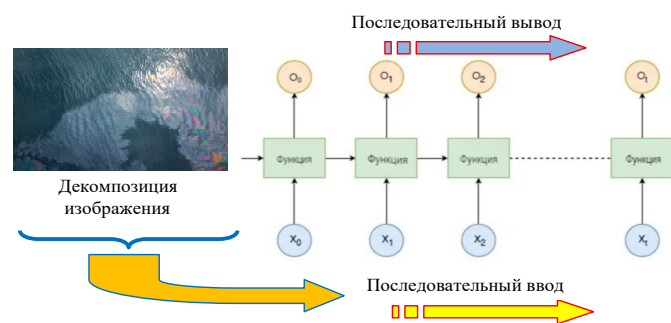


Рис. 4. Принцип обработки данных в моделях на основе сетей типа RNNs

Свое применение сети RNNs находят в задачах распознавания речи, а также синхронного машинного перевода.

Таким образом, проведенный предварительный анализ показал, что среди широко известных моделей глубоких нейронной сети, к решению задач обнаружения разливов нефтепродуктов в большей степени адаптированы сверточные нейронные сети вида CNNs. Поэтому для дальнейшей обработки целесообразно проанализировать известные решения на основе CNNs.

Обоснование выбора модели для решения задач обнаружения разливов нефтепродуктов на основе обучения сетей CNNs

Принципиальное отличие моделей, реализуемых в сетях вида CNNs, от моделей на основе многослойных перцептронов, используемых в сетях MLP, состоит в том, что последние представляют собой полно связную структуру, охватывающую все узлы [14]. Такой подход при работе с изобра-

жениями приводит к размерности, которая вычислительно неосуществима.

Так, для цветного изображения размером 1000×1000 пикселей в канале RGB, реализация модели предполагает вычислением 3 миллионов весовых коэффициентов на каждый полно связный нейрон. Например, в алгоритмах компьютерного зрения CIFAR-10, где изображения имеют размер всего лишь $32 \times 32 \times 3$ (ширина $\times$ высота $\times$ цветовые каналы), подключенный нейрон первого скрытого слоя будет иметь $32 \times 32 \times 3 = 3072$ весовых значения. А для изображения размером 200×200 потребуется использование нейронов с $200 \times 200 \times 3 = 120\,000$ весами.

Кроме того, такая сетевая архитектура построения моделей не учитывает пространственную структуру данных изображения. В результате пиксели, расположенные далеко друг от друга, будут обрабатываться так же, как близко расположенные друг к другу.

В тоже время сверточные нейронные сети представляют многослойные перцептроны. Поэтому модели, построенные на их основе, успешно решают проблемы присущие архитектуре построения сетей MLP за счёт использования сильной пространственно-локальной корреляции, характерной для естественных изображений.

Анализ работ [12-14], посвященных разработке и исследованию возможностей глубоких сверточных нейронных сетей CNNs, позволил обобщить сведения об их эффективности и производительности, которые сведены в табл. 1.

Таблица 1

Параметры эффективности и производительность CNNs

	Тип глубокой нейронной сети				
	AlexNet	VGG-16	GoogLeNet	ResNet	SENet
Количество вычислительных операций,	724×10^6	$15,5 \times 10^9$	$1,43 \times 10^9$	$3,6 \times 10^9$	$5,8 \times 10^9$
Вероятность обнаружения объекта, $p_{об}$	0,84	0,92	0,93	0,96	0,965

Проведенный анализ возможностей глубоких сверточных нейронных сетей CNNs показал, что одной из проблем их эффективного применения является необходимость глубокого обучения сети, которое основано на строгом детерминированном описании признаков распознаваемых объектов. Однако, такой подход требует более 10^9 операций на обработку каждого идентифицируемого изображения. Следовательно, необходима серьезная модификация самого подхода к представлению исходных данных.

В частности, предлагается кадры разливов нефтепродуктов предварительно преобразовывать на основе их семантического сегментирования, с последующим выделением включенных в них объектов так.

То есть вместо распознавания изображения в целом, перейти к распознаванию отдельных выделенных сегментов (меток), аномально выделяющихся среди общего фона подстилающей поверхности.

Оправданность такого подхода обусловлена тем, что разливы нефтепродуктов характеризуются большим разнообразием принимаемых форм, что потребует значительной обучающей базы.

Для эффективного выделения требуемых объектов распознавания на выделенных сегментах и расширения зоны охвата фильтров была использована пространственная свертка, а также комплексное объединение пространственных пирамид (*Atrous Spatial Pyramid Pooling* – ASPP) для использования параллельных фильтров с различной частотой. Затем выделенные сегменты расширялись посредством билинейной интерполяции, для последующего восстановления их первоначального разрешения.

В качестве примера, на рисунке 5 представлена декомпозиция входного изображения на верхнем уровне представлении более высокого уровня глубокой сверточной нейронной сети (Deep Convolutional Neural Network – DCNN) [14].

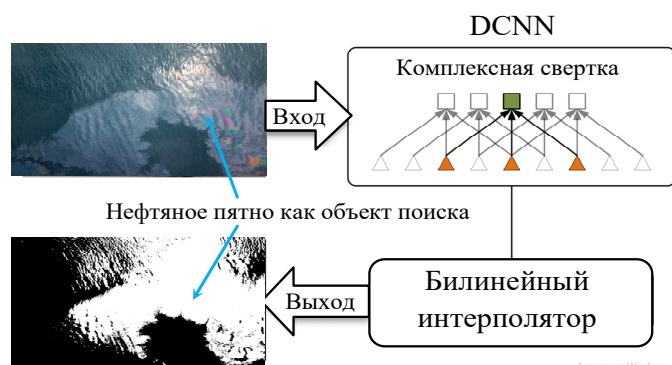


Рис. 5. Принцип функционирования модели «DeepLab» на высшем уровне ее декомпозиции

Для более эффективного извлечения значимых признаков и повышения разрешения фильтров, на верхнем уровне модели использовалась комплексная свертка (показана на рис. 5). Кроме того, в модели реализовано комплексное пространственное пирамидальное объединение (*Atrous Spatial Pyramid Pooling* – ASPP) для использования параллельных фильтров с различными разрешениями. После чего, разрешение кадров восстанавливалось посредством процедур билинейной интерполяции.

В рамках решаемой задачи представленная модель изначально использует технологию DCNN, и при этом базируется на сетевую модель ResNet-101. Выбор сети ResNet-101 обусловлен ее высокой производительностью при реализации семантической сегментации изображений.

В рамках решаемой задачи модель была перестроена для работы с сегментами изображений, что в свою очередь позволило значительно снизить объем вычислительных операций.

Учитывая, что для компенсации снижения размерности матрицы признаков сегмента (разрешения), на конечном этапе использовалась билинейная интерполяция, то общий объем проводимых вычислений увеличивается пропорционально количеству используемых слоев. Поэтому для управления разрешением матриц признаков обрабатываемых объектов применялась комплексная свертка в DCNN.

Несмотря на способность сетей DCNN адекватно представлять масштаб по результатам проведенного обучения при работе с изображениями с различным разрешением, их возможности по масштабированию сегментов могут быть улучшены. Это особенно важно при обнаружении разливов нефтепродуктов, по результатам обработки изображений, полученных со спутников, находящихся на различных орбитах.

Предложения по совершенствованию технологии семантической сегментации изображений в моделях на основе сетей DCNNs при решении задач обнаружении разливов нефтепродуктов

Следует отметить, что рассмотренный подход в [14] подход к семантической сегментации оправдан только при обработке спутниковых снимков лишенных мелких деталей, характеризующихся вариативностью водной поверхности. Поэтому для систем экологического мониторинга, размещаемых на воздушных платформах на высоте до 10 км, бинарная сегментация не обеспечит требуемой контрастности разливов нефтепродуктов, по отношению к окружающей акватории, что приведет к существенному росту ложных срабатываний.

В данном случае перспективным видится применение многомасштабной сегментации, в которой при переходе на каждый новый слой обработки изображения, меняется не только разрешение сегментов, но и их масштаб. Такой подход, во-первых, позволяет многократно сократить объем необходимых вычислений. А во-вторых, повысить структурные различия выделенных сегментов.

Как правило, сегментация изображения позволяет упростить структуру объекта посредством присвоения метки каждому пикселю изображения таким образом, чтобы пиксели с одинаковыми метками обладали определёнными характеристиками, позволяющими объединить их в один сегмент.

Результатом многомасштабной сегментации изображения является набор сегментов, которые в совокупности покрывают всё изображение, или набор контуров, извлечённых из изображения. При этом объединение в сегменты на каждом слое предлагается реализовать по аналогии с канонической сегментацией обычных снимков, при которой объединение происходит по таким параметрам как цвет, интенсивность или текстура. В результате соседние сегменты получают существенные отличия, позволяющие обеспечить не только их локализацию, но и распознавание. Подробно данная технология описана в [15].

В качестве научной основы для реализации многомасштабной сегментации была определена семантическая сегментация, при которой для каждого пикселя был определен соответствующий ему класс. А непосредственное разбиение на сегменты осуществлялось в соответствии с алгоритмом K-средних.

Принцип многомасштабной сегментации изображения 224×224 пиксела демонстрируется на рисунке 6.

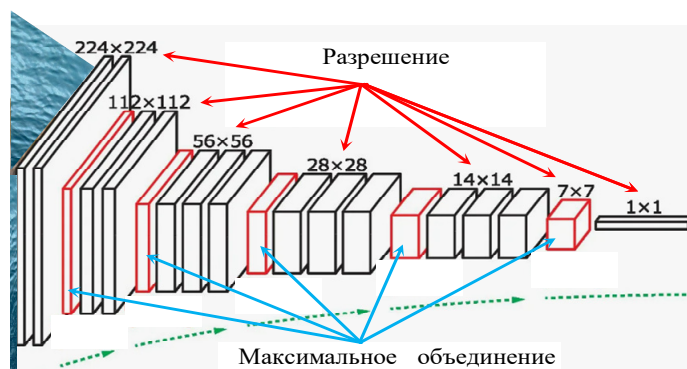


Рис. 6. Принцип многомасштабной сегментации исходного изображения

Для разбиения изображений на сегменты использовался алгоритм максимизации математического ожидания, обеспечивающий итеративную оценку апостериорных вероятностей, характеризующих близость интенсивности цвета объединяемых пикселей.

К сожалению, рекомендованный в [14] сайт, <https://scihub.copernicus.eu/> был закрыт в 2023 г., что не позволило для демонстрации работы разработанного подхода использовать графический материал, представленный в [14].

Ограниченность бинарной сегментации, используемой в [14], не позволяет учесть в полном объеме все аспекты изображений, обусловленные высокой вариативностью водной поверхности без результатов разлива нефтепродуктов. Поэтому ее использование может привести к ложному принятию решения.

Продemonстрируем данное заключение на тестовых изображениях водной поверхности без разливов нефтепродуктов (см. рис. 7).

На рисунке 8 показано бинарная сегментация тестового изображения, и многомасштабная сегментация, в соответствии с представленным подходом.

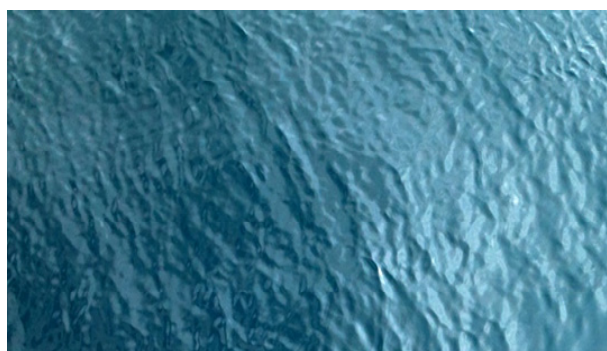


Рис. 7. Тестовое изображение водной поверхности без разливов нефтепродуктов

Если в исследовании [14] в качестве опорных цветов определены черный и красный, то в предложенном подходе в качестве цветовой гаммы были выбраны белый и черный цвет.

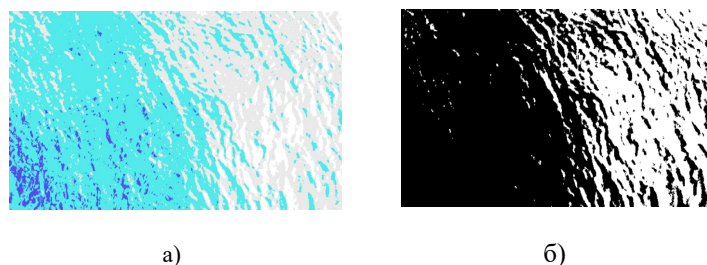


Рис. 8. Тестовое изображение: а) после бинарной сегментации; б) после многомасштабной сегментации

Очевидно, что данные обработки тестового изображения после бинарной сегментации однозначно укажут на наличие разлива нефтепродуктов, что будет являться ложным результатом (поскольку блики водной поверхности были сегментированы как результаты разлива нефтепродуктов).

В тоже время тестовое изображение после многомасштабной сегментации (16 градаций масштаба, которые с учетом низкой динамики изменений $\max(f_i)$ и $\min(f_i)$, были

сведены до четырех) не идентифицируется нейронной сетью как загрязнённая акватория.

Справедливости ради стоит отметить, что размер тестового изображения после бинарной сегментации уменьшился в 15,5 раз по отношению к исходному, в то время как после многомасштабной – всего лишь в шесть раз.

На рисунках 9-10 представлены пример изображения с разливом нефтепродуктов и результаты его бинарной и многомасштабной сегментации (16 градаций).



Рис. 9. Тестовое изображение с разливами нефтепродуктов

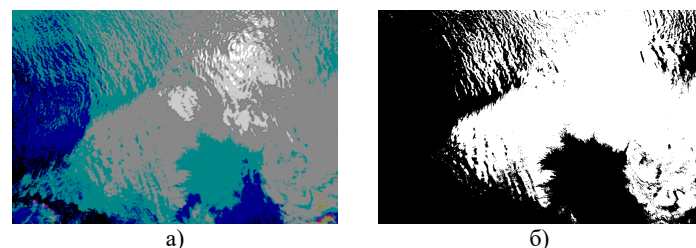


Рис. 10. Тестовое изображение с разливами нефтепродуктов: а) после бинарной сегментации; б) после многомасштабной сегментации

Заметим, что на тестовом рисунке 10а цветовая гамма разливов нефтепродуктов отличается от сегментов изображения (рис. 8а), на котором отсутствуют разливы нефтепродуктов.

Учитывая, что в исследовании [14] обработка велась спутниковых снимков, на которых бликовый эффект был не так ярко выражен, как на тестовых изображениях, полученных с борта беспилотного аппарата, то для сравнительной оценки эффективности разработанного подхода, рассматривался не результат классификации разливов, а точность сегментации изображений.

Результаты сведены в таблице 2.

В качестве исходных данных были отобраны 50 снимков водных акваторий без сторонних объектов, представленных в сети Internet в свободном доступе. На 30 снимках были запечатлены разливы нефтепродуктов. Затем на обрабатываемых изображениях разливы нефтепродуктов предварительно были локализованы в ручном режиме. После чего выделенные контуры сравнивались с результатами сегментации. Положительным результатом считалось не менее чем 80% совпадение контуров.

Таблица 2

Результаты оценки эффективности моделей

2 градации		16 градаций	
Back-ground	Look-alikes	Back-ground	Look-alikes
41%	45%	13%	67%

В таблице 2 введены следующие обозначения:

Back-ground – восприятие фоновой поверхности как объекта распознавания;

Look-alikes – совпадение результатов с данными визуального анализа.

В подтверждение обоснованности малоразмерной (до градации четыре) сегментации, при ее использовании в обработке спутниковых снимков, воспользуемся представленными в [18] дополнительными результатами, определяющими достоверность классификации на уровне 0,89, которые сопоставимы с апробированными ранее данными других исследований.

Таким образом, можно полагать, что полученные результаты, базирующиеся на сетях DCNN, сопоставимы с возможностями апробированных подходов. Но при этом используемая многомасштабная сегментация позволяет почти на порядок снизить вычислительные затраты.

Заключение

Полученные результаты позволяют заключить следующее. Несмотря на высокую эффективность функционирования глубоких нейронных сетей при решении задач распознавания изображений, их применение для решения задач обнаружения разливов нефтепродуктов непосредственно на борту летательных аппаратов связано с определенными трудностями.

Во-первых, обучение сетей CNNs происходит крайне медленно. При этом архитектура сетей громоздка и требует достаточно больших вычислительных ресурсов.

В частности, только исходный код сети VGG16 занимает объем более 533 МБ памяти. И это без баз данных, формируемых в ходе ее обучения, при том, что общее количество реализуемых вычислительных операций превышает 15 млрд. Это делает процесс ее развертывания далеко не тривиальной задачей и предполагает наличие мощных вычислительных серверов.

И хотя модель VGG16 успешно используется для решения широкого класса задач распознавания, ее применение также ограничено. Во многих случаях на практике используют более простые сети, типа SqueezeNet, GoogLeNet и др.

Следует отметить, что сети CNNs, как правило, ориентированы на работу с конкретными изображениями, в то время как последствия разливов нефтепродуктов не имеет четкой детерминированности. При этом даже для объектов с ярко выраженной формализацией, такие сети предполагают наличие оборудования, обладающего значительных вычислительными мощностями.

Вместе с тем многомасштабная сегментация, как результат дальнейшего развития семантической сегментации, адаптированной в [14], открывает широкие возможности по совершенствованию методов экологического мониторинга, связанного с автоматизацией процедур обнаружения результатов разливов нефтепродуктов. В частности, по результа-

там обработки изображений, получаемых с ЛПП, расположенных на высоте до 10 км.

Дальнейшее исследование авторы связывают с созданием автоматических классификаторов обнаружения результатов техногенных катастроф в акватории мирового океана. Для этого планируется использовать подходы, представленные в [15-17]. С исследованием возможностей повышения помехозащищенности линий радиосвязи с борта ЛПП и повышением эффективности процедур принятия решения при обнаружении [18-20], и повышения энергетического потенциала сигнальных конструкций [21-25].

Литература

1. Алексанин А. И., Качур В. А., Орлова Т. Ю., Павлов А. Н., Салюк П. А. Организация исследований по оценке экологического состояния морских акваторий средствами дистанционного зондирования // Современные проблемы дистанционного зондирования Земли из космоса. 2010. Т. 7, № 3. С. 131-138. EDN NCYASH.
2. Сурина Е. Е. Моделирование локальных систем промышленного экологического мониторинга // Образовательные ресурсы и технологии. 2021. № 2(35). С. 65-71. DOI 10.21777/2500-2112-2021-2-65-71. EDN VOUFNL.
3. Иванов М. С., Шушков А. В., Макаренко С. И. Повышение скорости передачи данных в сети воздушной радиосвязи управления летательными аппаратами за счет адаптивного использования энергетического, сигнального и частотного сетевых ресурсов. Часть 1. Модели и методика повышения скорости передачи данных // Системы управления, связи и безопасности. 2023. 1. С. 125-219. DOI 10.24412/2410-9916-2023-1-125-219. EDN WWMDQN.
4. Васильева Д. В., Дворников С. В., Якушенко С. А., Дворников С. С. Автоматизация процедур обнаружения лесных пожаров по результатам обработки видео // Научно-аналитический журнал "Вестник Санкт-Петербургского университета Государственной противопожарной службы МЧС России". 2023. № 4. С. 47-58. DOI 10.61260/2218-130X-2024-2023-4-47-58. EDN TQYCOI.
5. Колосов П. Удаленный мониторинг водных ресурсов с помощью ИИ и БПЛА // Control Engineering Россия. 2022. № 1(97). С. 58-59. EDN WWHHPN.
6. Зиновьев Е. С., Сатишев Р. П., Павленко С. А. Применение ГИС «Оператор» в комплексах с БПЛА для оперативной обработки и анализа геоинформационных данных // Вестник науки. 2023. Т. 1, № 12(69). С. 821-829. EDN FDTNIX.
7. Васильева Д. В., Дворников С. В. Структуризация первичных признаков изображений систем мониторинга окружающей среды // Радиотехника. 2024. Т. 88, № 8. С. 5-17. DOI 10.18127/j00338486-202408-01. EDN CXJBDK.
8. Дворников С. В., Дворников С. С., Васильева Д. В. Автоматизация процедур обнаружения нефтяных разливов на водной поверхности // Информация и космос. 2024. № 1. С. 126-132. EDN GMFJFC.
9. Васильева Д. В., Дворников С. С., Толстуха Ю. Е., Обрезков П. С., Дворников С. В. Формирование векторов признаков для систем видеонаблюдения // Вопросы радиоэлектроники. Серия: Техника телевидения. 2023. № 4. С. 62-68. EDN NZSBEJ.
10. Гибадуллин Р. Ф., Смирнов И. Н., Хевронин Н. В., Никитин А. В., Перухин М. Ю. Разработка аппаратно-программного модуля обнаружения объектов для встраиваемых систем // Вестник Технологического университета. 2018. Т. 21, № 6. С. 118-122.
11. Фукунага К. Введение в статистическую теорию распознавания образов / Пер. с англ. И. Ш. Торговицкого; под ред. А. А. Дорофеева. Москва: Наука, 1979. С. 367.
12. Simonyan K., Zisserman A. Very Deep Convolutional Networks for Large-Scale Image Recognition. arxiv2014, arXiv: 1409.1556.
13. Stockman G., Shapiro L. G. Computer Vision. In Computer Vision, 1st ed.; Prentice Hall PTR: Upper Saddle River, NJ, USA, 2001; pp. 72-89.

14. Orfanidis G., Ioannidis K., Avgerinakis K., Vrochidis S., Kompatsiaris I., A Deep Neural Network for Oil Spill Semantic Segmentation in Sar Images // 2018 25th IEEE International Conference on Image Processing (ICIP), Athens, Greece, 2018, pp. 3773-3777, doi: 10.1109/ICIP.2018.8451113.

15. Nielsen Frank, Richard Nock. On region merging: The statistical soundness of fast sorting, with applications // 2003 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2003. Proceedings. Vol. 2. IEEE. pp. II:19-26. doi:10.1109/CVPR.2003.121144

16. Васильева Д. В., Якушенко С. А., Дворников С. С., Погорелов А. А., Дворников С. В. Обнаружение морских дронов в оптическом диапазоне // Морской вестник. 2023. № 4(88). С. 90-92. EDN KYEBDL.

17. Васильева Д. В., Дворников С. В., Якушенко С. А., Дворников С. С. Автоматизация процедур обнаружения лесных пожаров по результатам обработки видео // Науч.-аналит. журн. Вест. СПбУ ГПС МЧС России. 2023. № 4. С. 47-58.

18. Дворников С. В., Васильева Д. В. Повышение достоверности выявления аномалий на изображениях при формировании их векторов признаков в базисах вейвлетов // Информатика и автоматизация. 2024. Т. 23, № 6. С. 1698-1729. DOI 10.15622/ia.23.6.5. EDN JDDWJS.

19. Бестугин А. Р., Дворников С. В., Дворников С. С., Кирилина И. А. Расчет вероятности ложной тревоги при обнаружении сигнала с несколькими неизвестными параметрами // Датчики и системы. 2023. № 4-2(270). С. 30-33. DOI 10.25728/datsys.2023.4.2.5. EDN ESOQOI.

20. Васильева Д. В. Исследование чувствительности векторов признаков, сформированных на основе кратномасштабных преобразований обрабатываемых изображений // Известия высших учебных заведений России. Радиоэлектроника. 2024. Т. 27. № 3. С. 20-29.

21. Дворников С. В., Якушенко С. А. Управление параметрами устойчивости инфокоммуникационной системы в условиях де-

структивных воздействий // Радиотехника. 2023. Т. 87, № 6. С.23-31. DOI 10.18127/j00338486-202306-03. EDN EXKZOH.

22. Vasilyeva D. V., Dvornikov S. V., Yakushenko S. A., Dvornikov S. S. Automation of detection procedures based on the results of processing images from video surveillance systems. В книге: Математические методы и модели в высокотехнологичном производстве. Сборник тезисов докладов III Международного форума. В 2-х частях. Санкт-Петербург, 2023. С. 157-158.

23. Павлов А. А., Дворников С. С., Кондрашин А. Е., Дворников С. В. Управление параметрами сигналов с амплитудной манипуляцией в радиолиниях морской подвижной службы // Информация и космос. 2023. № 2. С. 12-17. EDN GEDBJE.

24. Dvornikov, S. S., Zheglov K. D., Dvornikov S. V. SSB signals with controlled pilot level // T-Comm: Телекоммуникации и транспорт. 2023. Vol. 17, No. 3. P. 41-47. DOI 10.36724/2072-8735-2023-17-3-41-47. EDN VUZDRH.

25. Афонин И. Е., Макаренко С. И., Петров С. В. Модель оценивания устойчивости системы управления воздушно-космической обороной в конфликте со средствами воздушно-космического нападения // Системы управления, связи и безопасности. 2023. № 3. С. 227-266. DOI 10.24412/2410-9916-2023-3-227-266. EDN CMUQPQ.

26. Чертова О. Г., Чиров Д. С. Построение опорной сети связи на базе малоразмерных беспилотных летательных аппаратов с отсутствием наземной инфраструктуры // Научные технологии в космических исследованиях Земли. 2019. Т. 11. № 3. С. 60-71.

27. Чиров Д. С., Лобов Е. М. Выбор сигнально-кодовой конструкции для командно-телеметрической линии радиосвязи с беспилотными летательными аппаратами средней и большой дальности // T-Comm: Телекоммуникации и транспорт. 2017. Т. 11. № 10. С. 21-28.

MULTISCALE IMAGE SEGMENTATION IN SOLVING PROBLEMS OF OIL SPILL DETECTION BASED ON CONVOLUTIONAL NEURAL NETWORKS

Dina V. Vasilieva, St. Petersburg State University of Aerospace Instrumentation, St. Petersburg, Russia, dolli.dina@mail.ru
Sergey V. Dvornikov, St. Petersburg State University of Aerospace Instrumentation, St. Petersburg, Russia, practicdsv@yandex.ru
Sergey S. Dvornikov, St. Petersburg State University of Aerospace Instrumentation, St. Petersburg, Russia, dvornik92@mail.com

Abstract

Effective ecological monitoring of water areas of seas and oceans is difficult due to the huge size of the monitored areas. Therefore, machine learning methods are usually used to process the received primary photo and video material when solving problems of oil spill detection. High variability of the water surface of the world ocean and the large volume of each image frame complicate the effective work of neural networks. For this purpose it is proposed to use multiscale image segmentation, in accordance with the K-means algorithm, which will allow to obtain several homogeneous fragments, the recognition of which will provide a solution to the problem of oil spill detection. And for effective processing of the selected recognition objects represented as segments and expanding the capabilities of filters, spatial convolution is used, followed by complex integration of spatial pyramids. The desired model is built on a convolutional neural network. The use of the presented solution allows to reduce the computational load on the convolutional neural network, as well as to increase the reliability of detection of oil spill results when processing images with high variability of water surface parameters. A comparative evaluation of the proposed approach in relation to the model realizing binary semantic segmentation technologies on the basis of complex spatial pyramidal association, which has received wide scientific validation, is presented. The presented solution is proposed to be implemented as a model realized by a series of CNNs. The proposed model allows to reduce the computational load of the network approximately 6 times in relation to the canonical approach, while ensuring the coincidence of detection results with the data of visual analysis in 67% of cases (the experiment was conducted on 50 images of water areas, 30 of which depicted spills of oil products, and the images did not contain third-party objects), which is almost 1.5 times higher than the capabilities of the tested model.

Keywords: convolutional neural networks, oil spill detection, multiscale image segmentation, environmental monitoring, oil spill

References

- [1] A.I. Aleksanin, V.A. Kachur, T.Yu. Orlova, A.N. Pavlov, P.A. Saliuk, "Organization of research to assess the ecological state of marine areas by means of remote sensing," *Modern problems of Earth remote sensing from space*. 2010. T. 7, no. 3. Pp. 131-138.
- [2] E.E. Surina, "Modeling of local systems of industrial environmental monitoring," *Educational Resources and Technologies*. 2021. No. 2(35). Pp. 65-71. DOI 10.21777/2500-2112-2021-2-65-71.
- [3] M.S. Ivanov, A.V. Shushkov, S.I. Makarenko, "Increase of data transmission rate in the airborne radio network of aircraft control due to adaptive utilization of energy, signal and frequency network resources. Part 1. Models and methodology of data transmission speed increase," *Control, communication and security systems*. 2023. 1. Pp. 125-219. DOI 10.24412/2410-9916-2023-1-125-219.
- [4] D.V. Vasilieva, S.V. Dvornikov, S.A. Yakushenko, S.S. Dvornikov, "Automation of pro- procedures of forest fire detection based on the results of video processing," *Scientific and analytical journal "Bulletin of St. Petersburg University of the State Fire Fighting Service of the Ministry of Emergency Situations of Russia"*. 2023. No. 4. Pp. 47-58. DOI 10.61260/2218-130X-2024-2023-4-47-58.
- [5] P. Kolosov, "Remote monitoring of water resources using AI and UAVs," *Control Engineering Russia*. 2022. No. 1(97). Pp. 58-59.
- [6] E.S. Zinoviev, R.P. Satishev, S.A. Pavlenko, "Application of GIS "Operator" in com-plexes with UAVs for operational processing and analysis of geoinformation data," *Vestnik nauki*. 2023. Vol. 1, no. 12(69). Pp. 821-829.
- [7] D.V. Vasilieva, S.V. Dvornikov, "Structurization of the primary image features of the environmental monitoring systems," *Radiotekhnika*. 2024. Vol. 88, no. 8. Pp. 5-17. DOI 10.18127/j00338486-202408-01.
- [8] S.V. Dvornikov, S.S. Dvornikov, D.V. Vasilieva, "Automation of the oil spill detection procedures on the water surface," *Information and Space*. 2024. No. 1. Pp. 126-132.
- [9] D.V. Vasilieva, S.S. Dvornikov, Yu.E. Tolstukha, P.S. Obrezkov, S.V. Dvornikov, "Formation of the feature vectors for the video surveillance systems," *Voprosy radioelektroniki. Series: Television Technics*. 2023. No. 4. Pp. 62-68.
- [10] R.F. Gibadullin, I.N. Smirnov, N.V. Khevronin, A.V. Nikitin, M.Y. Perukhin, "Development of hardware-software module of object detection for embedded systems," *Vestnik of Technological University*. 2018. Vol. 21, no. 6. Pp. 118-122.
- [11] K. Fukunaga, "Introduction to the statistical theory of pattern recognition. Moscow: Nauka, 1979. 367 p.
- [12] K. Simonyan, A. Zisserman, "Very Deep Convolutional Networks for Large-Scale Image Recognition," arxiv2014, arXiv: 1409. 1556.
- [13] G. Stockman, L.G. Shapiro, "Computer Vision. In Computer Vision," 1st ed.; Prentice Hall PTR: Upper SaddleRiver, NJ, USA, 2001; pp. 72-89.
- [14] G. Orfanidis, K. Ioannidis, K. Avgerinakis, S. Vrochidis, I. Kompatsiaris, "A Deep Neural Network for Oil Spill Semantic Segmentation in Sar Images," *2018 25th IEEE International Conference on Image Processing (ICIP)*, Athens, Greece, 2018, pp. 3773-3777, doi: 10.1109/ICIP.2018.8451113.
- [15] Nielsen Frank, Richard Nock, "On region merging: The statistical soundness of fast sorting, with applications," *2003 IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 2003. Proceedings. Vol. 2. IEEE. pp. 11:19-26. doi:10.1109/CVPR.2003.121144
- [16] D.V. Vasilyeva, S.A.v Yakushenko, S.S. Dvornikov, A.A. Pogorelov, S.V. Dvornikov, "Detection of marine drones in the optical range," *Marine Bulletin*. 2023. № 4(88). Pp. 90-92.
- [17] D.V. Vasilieva, S.V. Dvornikov, S.A. Yakushenko, S.S. Dvornikov, "Automation of the forest fire detection proce- dures based on the video pro- cessing results," *Nauch-analit. zhurn. vest. SPbU GPS MES of Russia*. 2023. No. 4. Pp. 47-58.
- [18] S.V. Dvornikov, D.V. Vasilieva, "Increase of reliability of anomalies detection on images at formation of their feature vectors in wavelets bases," *Informatics and Automation*. 2024. Vol. 23, no. 6. Pp. 1698-1729. DOI 10.15622/ia.23.6.5.
- [19] A.R. Bestugin, S.V. Dvornikov, S.S. Dvornikov, I.A. Kirshina, "Calculation of the false alarm probability at signal detection with several unknown parameters," *Datchi-ki i sistemy*. 2023. No. 4-2(270). Pp. 30-33. DOI 10.25728/datsys.2023.4.2.5.
- [20] D.V. Vasilyeva, "Investigation of sensitivity of feature vectors formed on the basis of multiple scale transformations of processed images," *Izvestiya vysshee obrazovaniya vysshee obrazovaniya Rossii. Radioelectronics*. 2024. Vol. 27. No. 3. Pp. 20-29.
- [21] S.V. Dvornikov, S.A. Yakushenko, "Control of the infocommunication system stability parameters under the destructive influences," *Radiotekhnika*. 2023. Vol. 87, no. 6. Pp.23-31. DOI 10.18127/j00338486-202306-03.
- [22] D.V. Vasilyeva, S.V. Dvornikov, S.A. Yakushenko, S.S. Dvornikov, "Automation of detection procedures based on the results of processing images from video surveillance systems," In the book: *Mathematical methods and models in high-tech production. Collection of abstracts of reports of the III International Forum*. In 2 parts. St. Petersburg, 2023. Pp. 157-158.
- [23] A.A. Pavlov, S.S. Dvornikov, A.E. Kondrashin, S.V. Dvornikov, "Parameter management of the signals with the amplitude manipulation in the marine mobile service radio-lines," *Information and Space*. 2023. No. 2. Pp. 12-17.
- [24] S.S. Dvornikov, K.D. Zheglov, S.V. Dvornikov, "SSB signals with controlled pilot level," *T-Comm*. 2023. Vol. 17, No. 3. Pp. 41-47. DOI 10.36724/2072-8735-2023-17-3-41-47.
- [25] I.E. Afonin, S.I. Makarenko, S.V. Petrov, "Model of stability estimation of the air-space defense control system in the conflict with the air-space attack means," *Control, communication and security systems*. 2023. No. 3. Pp. 227-266. DOI 10.24412/2410-9916-2023-3-227-266.
- [26] O.G. Chertova, D.S. Chirov, "Construction of a backbone communication network based on small-sized unmanned aerial vehicles with no ground infrastructure," *H&ES Research*. 2019. Vol. 11. No. 3, pp. 60-71.
- [27] D.S. Chirov, E.M. Lobov, "Selection of a signal-code design for a command-telemetry radio communication line with medium- and long-range unmanned aerial vehicles," *T-Comm*. 2017. Vol. 11. No. 10, pp. 21-28.

Information about authors:

Dina V. Vasilieva, Senior Lecturer, Department of Radio Engineering Systems, Institute of Radio Engineering and Info-communication Technologies, St. Petersburg State University of Aerospace Instrumentation, St. Petersburg, Russia. <https://orcid.org/1my-orcid?orcid=0009-0009-5343-3356>

Sergey V. Dvornikov, Doctor of Technical Sciences, Professor of the Department of Radio Engineering and Optoelectronic Complexes, St. Petersburg State University of Aerospace Instrumentation, Professor of the Department of Radio Communications, Budenniy Military Academy of Communications, St. Petersburg, Russia. <https://orcid.org/0000-0002-4889-0001>

Sergey S. Dvornikov, Candidate of Technical Sciences, Associate Professor of the Department of Design and Technology of Electronic and Laser Means, St. Petersburg State University of Aerospace Instrumentation, Scientific Associate of the Research Center, Budenniy Military Academy of Communications, St. Petersburg, Russia. <https://orcid.org/0000-0001-7426-6475>