Обнаружение малых объектов на радиолокационных изображениях с использованием графово-свёрточных нейросетевых реализаций квадродеревьев

А. М. Достовалова, А. К. Горшенин

Федеральный исследовательский центр «Информатика и управление» РАН
Москва, 119333, Россия
E-mails: adostovalova@frccsc.ru, agorshenin@frccsc.ru

Предложена архитектура сбалансированного нейросетевого квадродерева для решения задач сегментации небольших объектов на радиолокационных изображениях в условиях недостаточного количества обучающих данных. Эта архитектура является ансамблевой и состоит из предварительно обученного свёрточного кодировщика, формирующего дополнительные признаки пикселей изображения, а также графово-свёрточной нейронной сети с улучшенной обработкой пространственных взаимосвязей в данных за счёт интегрированной в неё вероятностной модели квадродерева. Графовая часть также содержит специальный блок обрезки ветвей для выделения сходства между пикселями в разном пространственном разрешении. Кроме того, для повышения точности разделения несбалансированных классов предусмотрен блок обработки признаков масштабированных изображений и в процессе обучения использована пользовательская функция потерь специального вида. Эта архитектура со свёрточным кодировщиком U-Net была применена для сегментации нескольких радиолокационных изображений (источники — Sentinel-1 и HRSID (*англ.* High Resolution Synthetic Aperture Radar Images Dataset)). Разработанная архитектура продемонстрировала более высокое качество сегментации небольших объектов в задаче как многоклассовой, так и двухклассовой сегментации (т. е. выделения объекта на фоне) в сравнении с базовой реализацией нейросетевого квадродерева, а также относительно сети U-Net. Так, прирост значений метрики F_1 для классов малых объектов в сравнении с результатами U-Net достигает 3,59 % в задаче двухклассовой сегментации и 47,42 % в случае многоклассовой. Разработанные подходы имеют потенциал для применения в задаче обнаружения малоразмерных объектов с использованием суперпиксельных представлений на снимках высокого разрешения.

Ключевые слова: нейросетевое квадродерево, вероятностно-информированные нейронные сети, графово-свёрточные сети, сегментация радиолокационных снимков, малые объекты

Одобрена к печати: 07.08.2025
DOI: 10.21046/2070-7401-2025-22-5-51-62

Введение

Радиолокационные (РЛ) изображения часто используются в задачах мониторинга поверхности Земли. Нейронные сети (НС) стали популярным инструментом для выделения объектов на снимках, например, в контексте сегментации. Пиксели разделяются на две или более групп, образуя набор классов K. Нейронная сеть распознаёт объект, определяя для каждого пикселя изображения его наиболее вероятный класс $x \in K$:

$$x_s = \arg\max_{x \in K} \mathbb{P}\left(s \middle| x\right).$$

Основная сложность НС-обработки РЛ-снимков заключается в ограниченности обучающих данных, возникающая из-за сильной вариативности снимков (по разрешению, географическому региону, объектам съёмки и т. д.) и трудоёмкости их разметки, требующей специальных экспертных навыков. Эти факторы ограничивают объём и качество размеченных обучающих наборов: в практических задачах он может состоять всего лишь из нескольких снимков. Небольшие объёмы обучающих датасетов усложняют получение результатов высокого качества с помощью НС, особенно в случае сильного дисбаланса классов, характерного для задачи поиска малых объектов. Для повышения эффективности НС-обработки в таком

случае необходим источник дополнительной информации об объектах. Подход физически-информированного машинного обучения (Karniadakis et al., 2021) предлагает использовать для этих целей физические модели данных, которые разнообразными способами могут быть интегрированы в архитектуру нейронной сети. Для учёта влияния различных случайных факторов перспективным представляется задействование вероятностных моделей, см., например, работы (Gorshenin, Kuzmin, 2022; Gorshenin, Vilyaev, 2022). В частности, значительные успехи вероятностно-информированного подхода продемонстрированы в сфере современных телекоммуникационных сетей (Gorshenin et al., 2024; Kochetkova et al., 2023), обработки разнообразных изображений (Dostovalova, Gorshenin, 2024, 2025), в том числе радиолокационных, и геофизических временных рядов (Gorshenin, Vilyaev, 2024).

Случайное поле Маркова в форме квадродерева — вероятностная модель, описывающая связи между пикселями в различных пространственных разрешениях (Dostovalova, 2023; Pastorino et al., 2021, 2022). В статье (Достовалова, 2024) была предложена нейросетевая реализация, позволяющая учесть базовые принципы, содержащиеся в математической модели, в рамках архитектурного блока графовой свёртки. Она была названа нейросетевым квадродеревом (НКД). В настоящей статье развиваются ранее предложенные подходы для реализации эффективной сегментации малоразмерных объектов в условиях существенного дисбаланса разделяемых классов.

Основной вклад статьи состоит в следующем:

– развиты принципы вероятностно-информированной сегментации небольших объектов на РЛ-изображениях с использованием новой архитектуры, названной сбалансированным нейросетевым квадродеревом (СНКД);

– сбалансированное нейросетевое квадродерево применено для обнаружения малых объектов в условиях несбалансированных классов в задачах двух- и многоклассовой сегментации на одинаковых наборах данных. При этом для двухклассового случая, по сути соответствующего задачи выделения некоторого объекта на фоне, дисбаланс классов выражен особенно сильно, поэтому дополнительно использована функция потерь специального вида;

– сетью СНКД продемонстрировано повышение точности выделения малых объектов в задачах много- и двухклассовой сегментации: в первом случае прирост значений метрики F_1 достигал 47,42 % (в среднем 13,73 %), а во втором — 3,59 %.

Оставшаяся часть статьи организована следующим образом. В разд. «Современные подходы использования графовых сетей для сегментации изображений» представлен обзор известных подходов к использованию квадродерева в НС-архитектурах. Раздел «Архитектура сбалансированного нейросетевого квадродерева» содержит описание разработанной в настоящей статье новой эффективной архитектуры. В разд. «Тестирование на радиолокационных изображениях» представлены результаты экспериментов с использованием снимков из открытых наборов данных Sentinel-1 (Potin et al., 2012) и HRSID (*англ.* High Resolution Synthetic Aperture Radar Images Dataset) (Wei et al., 2020). В заключительном разделе кратко обсуждаются полученные результаты и возможные направления будущих исследований.

Современные подходы использования графовых сетей для сегментации изображений

Квадродерево — это графовая структура для описания пространственных данных, каждая вершина в которой имеет ровно четырёх потомков, исключая самый нижний уровень. Она может быть представлена пирамидой изображений разного разрешения $S_0 \ldots S_n$, где S_0 — верхний слой квадродерева, а S_n — нижний, но относящихся к одной обрабатываемой области поверхности Земли (*рис. 1*, см. с. 53). Каждая вершина $s \in S_l$ соответствует определённому пикселю в пирамиде с единственным родительским узлом s^- в предыдущем слое и четырьмя дочерними узлами s^+ (Pastorino et al., 2021). В статье (Pastorino et al., 2022) случайное поле со структурой квадродерева использовалось для улучшения сегментации изображений НС. В предлагаемом подходе сначала изображение сегментируется НС, затем из этих результатов

формируются вероятности классов для изображений низкого разрешения, наконец, полученные вероятности пересчитываются с учётом переходов между узлами квадродерева. Недостаток этого подхода заключается в том, что вероятности перехода определяются эмпирически, что снижает способность выделять детали, например очертания малых объектов (Dostovalova, 2023).

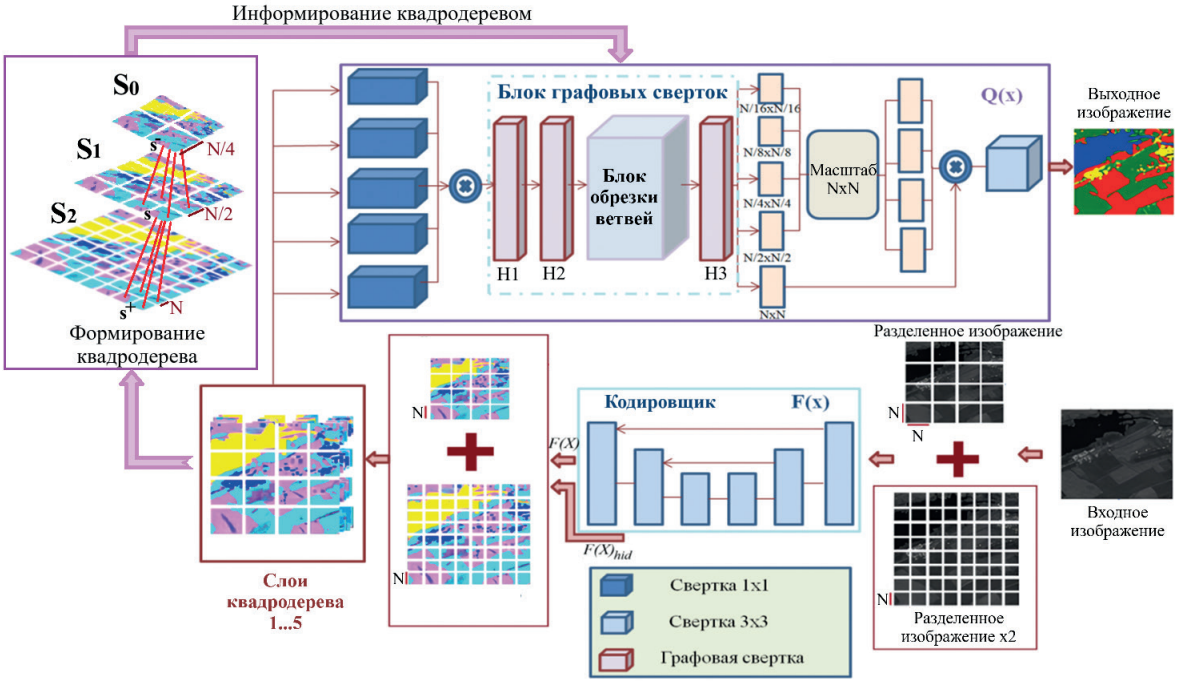


Рис. 1. Архитектура СНКД для работы с несбалансированными данными

Нейросетевая реализация квадродерева выступает естественным расширением существующего вероятностного алгоритма. Более того, существует несколько НС-архитектур, использующих элементы структуры квадродерева, например, для оптимизации свёртки (Jayaraman et al., 2018), токенизации изображений (Jewsbury et al., 2021; Ronen et al., 2023) или построения многомасштабного внимания (Ke et al., 2022; Tang et al., 2022). Однако перечисленные НС не применяют графовые представления (Chitta et al., 2020) данных для учёта пространственных связей между пикселями. Графовая свёртка — это способ обработки графовых структур с помощью НС. Этот подход был успешно применён для сегментации спутниковых изображений (Ma et al., 2019). Однако существующие методы обычно не рассматривают графы, работающие с изображениями в нескольких разрешениях. Квадродерево предоставляет структуру взаимосвязей между пикселями в разных масштабах. Данное свойство существенным образом используется для повышения эффективности соответствующих НС-решений в этой статье.

Архитектура сбалансированного нейросетевого квадродерева

Пусть $\mathbb{X}$ — малый набор изображений (в рамках данной статьи его размер считается равным порядка 1000 элементов), $F(\cdot)$ — свёрточная НС, которая формирует для изображения $X \in \mathbb{X}$ размером $N \times N$ признаки $F(X)$ и $X^+ = F(X)_{hid}$, взятые из скрытых слоёв $F(\cdot)$ с разным пространственным разрешением. Задача заключается в повышении точности сегментации небольших объектов (целей), полученной с помощью $F(X)$. При этом, естественно, необходимо избежать переобучения.

Для её решения предлагается использовать ансамбль $F(X)$ и графовую НС $Q(X)$, принимающую на вход X и X^+, которая моделирует пространственные связи между пикселями обраба-

тываемого изображения с помощью модели поля Маркова в форме квадродерева. Архитектура $Q(\cdot)$ должна улучшить сегментацию классов малых объектов в метрике $L(\cdot)$: $L\left(Q\left(F(X^+),F(X)\right)_{small},Y_{small}\right)<L\left(F(X)_{small},Y_{small}\right)$, где индекс *small* обозначает класс малого объекта, а Y — реальные метки классов пикселей изображения X.

Архитектура СНКД (см. *рис. 1*), рассматриваемая в этой статье, обобщает ранее предложенную авторами схему НКД (Достовалова, 2024). Базовая архитектура НКД представляет собой ансамбль из предобученного кодировщика $F(\cdot)$ (для него используется хорошо зарекомендовавшая себя в разнообразных задачах сеть U-Net (Ronneberger et al., 2015)) и графовой сети $Q(\cdot)$, состоящей из блока графовых свёрток, информированного системой связей между элементами квадродерева, и свёрточных слоёв, выполняющих сжатие данных. Графовый блок состоит из трёх слоёв графовой свёртки, аналогично исходной вероятностной модели квадродерева (Pastorino et al., 2021). Первые два слоя имеют одинаковую выходную размерность $H_1 = H_2$, а третий слой обладает размером $H_3 = KM$, где $M \in N$. Графовый блок также включает блок обрезки ветвей, усиливающий сходство между пикселями в соседних слоях. Новые значения признаков S_l^* для слоя S_l вычисляются по формуле:

$$S_l^* = 0{,}5\left(S_l + S_{l-1}\left(1 - \frac{\mathrm{Re}\,LU\left(soft\max\left(\left(S_l - Upsample(S_{l-1})\right)^2\right) - q\right)}{1-q}\right)\right),$$

где $q \in [0{,}3;1{,}0]$ — порог значимости коэффициента подобия. Выходные значения графового блока разбиваются на подвекторы, каждый из которых соответствует одному из слоёв квадродерева. Затем подвекторы объединяются и обрабатываются свёрточным классификатором для получения выходных вероятностей классов пикселей изображения.

Модификации, введённые в архитектуре СНКД, ориентированы на повышение точности сегментации малых объектов в условиях выраженного дисбаланса разделяемых классов. А именно, в графовый блок $Q(\cdot)$ была добавлена обработка свёрточными слоями признаков масштабированных изображений, полученных из исходного снимка с помощью линейной интерполяции. Изображение размером $2N \times 2N$ делится на четыре части, обрабатывается сетью $F(\cdot)$, после чего из него формируется изображение размером $N \times N$. Для этого изображение разбивается на участки размером 2×2, из которых формируются дополнительные четыре канала признаков, объединяющиеся со входными данными сети $Q(\cdot)$. Кроме того, когда дисбаланс классов выражен особенно сильно (ранее упоминавшийся двухклассовый случай с выделением объекта на фоне), для обучения $Q(\cdot)$ используется взвешенная функция потерь на основе кросс-энтропии (Özdemir, Sönmez, 2020) с настраиваемым весом классов $\mathbf{w}_{y_n}$:

$$L(\mathbf{x},\mathbf{y}) = \frac{1}{n}\sum_{i=1}^{n} -\mathbf{w}_{y_n} \log\left(\frac{\exp\left(\mathbf{x}_{n,y_n}\right)}{\sum_{j=1}^{K}\exp(\mathbf{x}_{n,j})}\right),$$

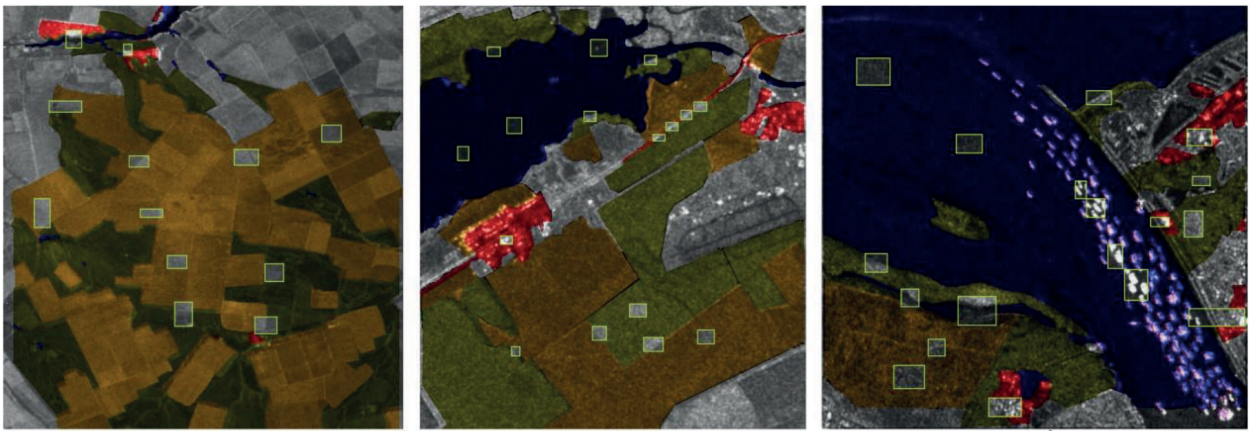
где n — это количество элементов в векторе предсказаний $\mathbf{x} \in R^{n \times K}$ для n обучающих примеров, а $\mathbf{y} \in N^n$ — вектор реальных меток классов для n обучающих примеров. Веса классов вводятся для предотвращения переобучения сети.

Тестирование на радиолокационных изображениях

Используемые данные и метрики

Для тестирования решения задач двух- и многоклассовой сегментации с помощью СНКД использованы три РЛ-изображения. Два из них получены с помощью радиолокатора спутника Sentinel-1 с разрешением 20 м, а третье взято из набора данных HRSID с разрешением

3 м. Каждое из них рассматривается как самостоятельный набор данных. Размер тестовых изображений составляет от 800×800 до 3000×3000 пикселей. Для обучения и тестирования сетей проводилась ручная разметка изображений (см. пример на *рис. 2*). Выделенные прямоугольные области с жёлтыми контурами на размеченных участках рассматривались как обучающий набор сети, а остальные размеченные участки использовались для тестирования. К малым объектам на трёх изображениях относились классы поселений (отмечены красным цветом), кораблей (фиолетовые области), а также водная поверхность (синяя заливка на *рис. 2а*). Тестовый набор имеет больший размер по сравнению с обучающим: пиксели выделенных жёлтыми прямоугольниками на цветном фоне областей составляют около 7 % от всех размеченных пикселей изображения. Тестирование на более обширном и разнообразном наборе позволяет лучше оценить обобщающие способности сети. Обучающие области — малые: их ширина и высота составляют около 50–100 пикселей для сохранения разнообразия поверхностей в обучающем наборе.



а *б* *в*

Рис. 2. Тестовые изображения Sentinel-1 (*а*, *б*) и HRSID (*в*) с масками классов, содержащие пять типов выделенных поверхностей: вода (синий цвет), обрабатываемые поля (оранжевый), необрабатываемые поля (оливковый), поселения (красный), корабли (фиолетовый)

Каждое изображение сегментировалось по участкам (патчам) заданного размера 32×32 пикселя, что превышает ранее апробированное, например, в работе (Достовалова, 2024) он составляет всего 16×16 пикселей. Данная модификация потребовалась для усиления эффекта классового дисбаланса в обучающем наборе. Тренировочный датасет после аугментации (повороты на произвольный угол и сдвиг) достигал размера от 800 до 6000 элементов. Он использовался для предобучения длительностью 100 эпох кодировщика U-Net, формирующего признаки исходных и масштабированных изображений.

Конфигурация сети СНКД определялась при переборе следующих гиперпараметров: размера входных данных графового блока $C \in \{100;32\}$, размерности $H_1 \in \{11;32\}$ и $M \in [1;10]$. Оптимальные значения определялись с помощью метода поиска по сетке (Bergstra, Bengio, 2012).

Результаты СНКД и НКД также сравнивались с сопоставимыми графово-свёрточными архитектурами без информирования. В таких архитектурах в качестве графа изображения используется обыкновенная двумерная решётка, которая не обрабатывает связи между элементами в разных масштабах. Сети НКД и СНКД без информирования (НКД-БИ и СНКД-БИ) повторяют архитектуры СНКД и НКД, представленные на *рис. 1*, но с тем отличием, что блоки, реализующие обработку изображения в масштабах меньших исходного, из этих сетей исключены.

Точность Prec (*англ.* Precision), полнота Rec (*англ.* Recall) и F_1-мера использовались для оценки точности сегментации каждого класса объектов (Powers, 2011):

$$\mathrm{Prec} = \frac{\mathrm{TP}}{\mathrm{TP}+\mathrm{FP}}, \quad \mathrm{Rec} = \frac{\mathrm{TP}}{\mathrm{TP}+\mathrm{FN}}, \quad F_1 = \frac{2\mathrm{Rec}\cdot\mathrm{Prec}}{\mathrm{Rec}+\mathrm{Prec}},$$

где TP (*англ.* True Positive) — число пикселей, верно классифицированных как выбранный класс; FP (*англ.* False Positive) — неверно классифицированные как выбранный класс пиксели; FN (*англ.* False Negative) — число пикселей выбранного класса, классифицированных неверно.

Результаты многоклассовой сегментации

Метрики точности для классов малых объектов, полученные в ходе многоклассовой сегментации изображений на *рис. 2* с помощью сетей U-Net; НКД-БИ; СНКД-БИ; НКД и СНКД приведены в *табл. 1.* Лучшие достигнутые значения выделены полужирным шрифтом.

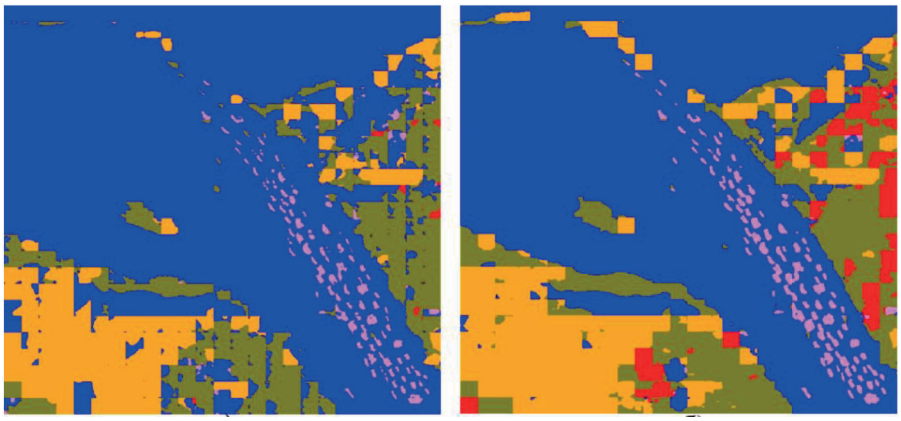
Таблица 1. Оценки точности сегментации (в %) классов целей для всех рассматриваемых изображений в задаче многоклассовой сегментации

НС	Поселения *рис. 2а*			Поселения *рис. 2б*			Вода *рис. 2а*			Поселения *рис. 2в*			Корабли *рис. 2в*		
	Rec	Prec	F_1	Rec	Prec	F_1	Rec	Prec	F_1	Rec	Prec	F_1	Rec	Prec	F_1
U-Net	63,11	84,68	72,33	**90,55**	80,73	85,36	54,70	**94,49**	69,29	1,48	**69,55**	2,91	78,33	**87,28**	82,57
НКД-БИ	**89,06**	9,06	16,44	89,98	82,98	86,34	76,88	12,41	21,37	42,56	29,00	34,49	93,39	69,4	79,64
НКД	88,04	10,63	18,97	90,06	83,25	86,52	70,0	15,40	25,25	**49,99**	35,34	41,41	**94,52**	71,05	81,12
СНКД-БИ	71,71	**93,49**	81,16	88,55	83,39	85,89	**81,55**	47,52	60,05	17,58	53,11	26,42	0,848	81,10	82,90
СНКД	76,90	89,98	**82,93**	88,40	**85,75**	**87,05**	72,87	82,29	**77,29**	43,14	60,40	**50,33**	89,64	81,79	**85,53**

Наилучшие точности для всех трёх изображений были получены при следующих значениях гиперпараметров: $C = 32$, $H_1 = 16$ и $M = 8$. Сбалансированное нейросетевое квадродерево продемонстрировало наилучшие результаты сегментации для всех трёх изображений по метрике F_1 как для классов крупных объектов, так и для малых. В первом случае прирост составляет от 0,41 до 9,12 % (в среднем 2,45 %), а во втором — от 1,69 до 47,42 % (в среднем 15,66 %). Информирование квадродеревом повышает точность сегментации малых объектов: прирост значений F_1 для СНКД и НКД относительно соответствующих им сетей без информирования достигает 23,91 %. Сбалансированное НКД и НКД во всех случаях превосходят по точности свои неинформированные аналоги. Для проверки статистической значимости различий результатов F_1 для разных архитектур использован критерий Фридмана (Friedman, 1940) при заданном уровне значимости 0,05. Полученное P-значение (0,005) свидетельствует об отклонении нулевой гипотезы в пользу альтернативы, что говорит о статистической значимости различий между полученными значениями.

Пример визуализации решения задачи многоклассовой сегментации приведён на *рис. 3* (см. с. 57). Сопоставление значений F_1 для СНКД и НКД с базовыми величинами, полученными классической сетью U-Net при обработке изображений с *рис. 2*, приведено на *рис. 4* (см. с. 57). Размеры столбцов диаграммы соответствуют значениям прироста (для классического НКД в ряде случаев получено ухудшение результатов) метрики, усреднённым по классам малых и крупных объектов.

Увеличение размера патча, по которому проводилась обработка изображения, на одних и тех же данных приводит к уменьшению размера тренировочного набора и, соответственно, усилению несбалансированности классов и изменчивости тренировочных примеров. Поэтому НКД, ранее продемонстрировавшее лучшие результаты (Достовалова, 2024), на более несбалансированном наборе уступает в двух из трёх случаев сети U-Net. Однако предложенная в данной статье архитектура СНКД для каждого изображения продемонстрировала хотя бы небольшой прирост для обоих классов.



а *б*

Рис. 3. Результаты сегментации изображения HRSID (*рис. 2в*) с использованием архитектуры: *а* — U-Net; *б* — СНКД

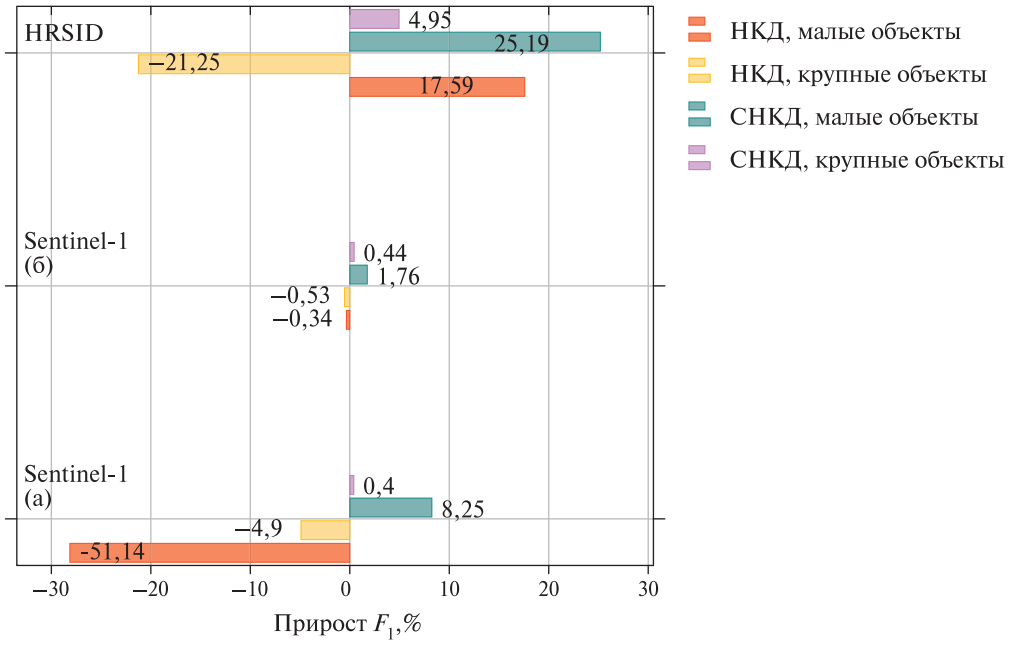
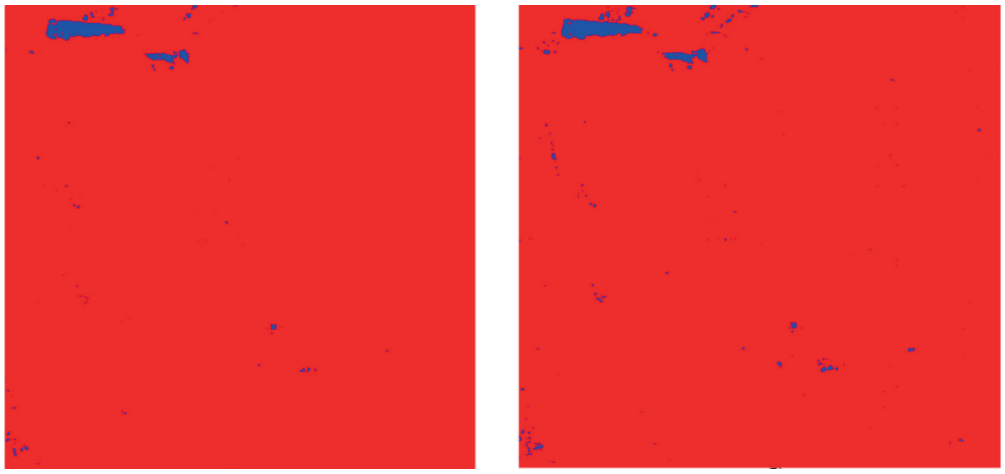


Рис. 4. Среднее изменение метрики F_1 для малых и крупных объектов при обработке СНКД и НКД в сравнении с результатами U-Net

Результаты двухклассовой сегментации

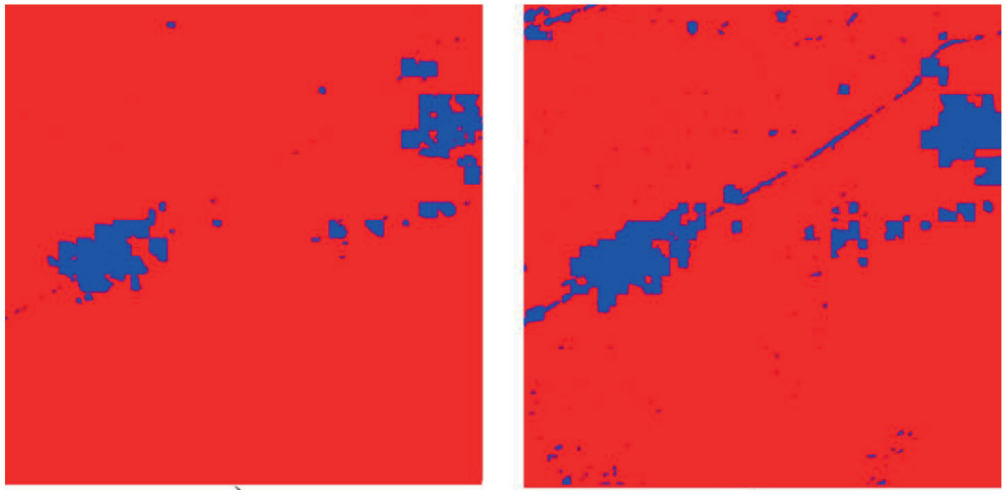
В задаче двухклассовой сегментации, соответствующей задаче выделения объекта на фоне, в качестве классов целей рассматривались поселения и дороги (см. *рис. 2а*, *б*) и корабли на *рис. 2в*. Остальные поверхности были отнесены к классу фона. Сильная несбалансированность разделяемых классов приводит к необходимости использования для обучения взвешенной кросс-энтропии. Вес класса целей был равен 0,8, а фонового — 0,2 при обучении U-Net. При обучении СНКД для *рис. 2б* вес класса цели составлял 0,2 (фона — 0,8), а для *рис. 2а*, *в* уже 0,01 (вес класса фона — 0,99). Вес класса целей при обучении графовой сети выбирался меньшим, чем вес класса фона, чтобы избежать переобучения при использовании признаков увеличенных изображений.

Значения метрик точности, полученные для классов целей U-Net, НКД-БИ, СНКД-БИ, НКД и СНКД, приведены в *табл. 2.* Результаты сегментации показаны на *рис. 5–7* (см. с. 58). Пиксели класса целей отмечены синим цветом.



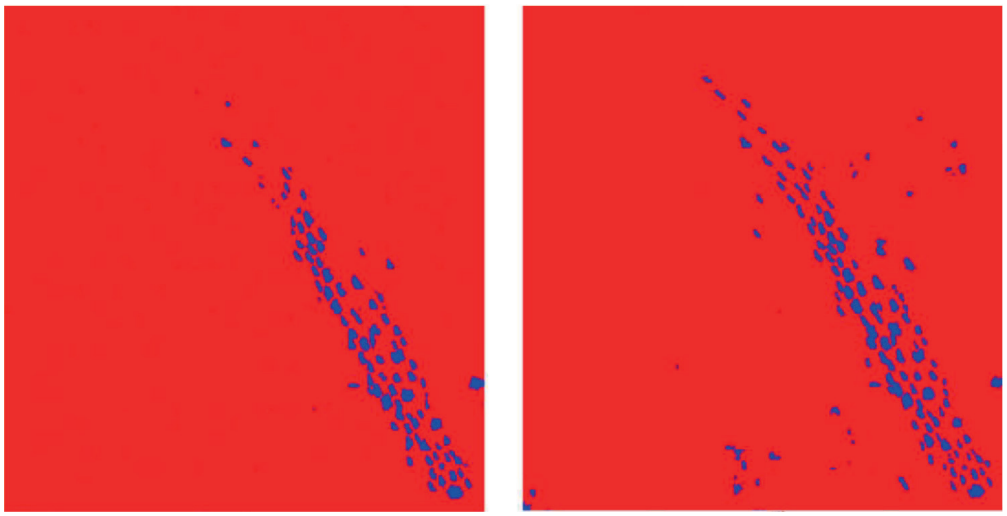
а *б*

Рис. 5. Результаты сегментации изображения Sentinel-1 (*рис. 2а*): *а* — U-Net; *б* — СНКД



а *б*

Рис. 6. Результаты сегментации изображения Sentinel-1 (*рис. 2б*): *а* — U-Net; *б* — СНКД



а *б*

Рис. 7. Результаты сегментации изображения HRSID (*рис. 2в*): *а* — U-Net; *б* — СНКД

Сбалансированное нейросетевое квадродерево демонстрирует наилучшие результаты сегментации для всех трёх изображений. Для *рис. 2б* и *в* максимальная точность была получена при следующих значениях гиперпараметров: $C = 24$, $H_1 = 11$ и $M = 1$. Для *рис. 2а* $C = H_1 = 32$ и $M = 8$. Сбалансированное НКД превосходит U-Net и классическое НКД, а также архитектуры без информирования в задаче двухклассовой сегментации: прирост значений F_1 составляет от 1,61 до 3,59 %.

Таблица 2. Оценки точности сегментации (в %) классов целей для всех рассматриваемых изображений в задаче двухклассовой сегментации

НС	Поселения *рис. 2а*			Поселения *рис. 2б*			Корабли *рис. 2в*		
	Rec	Prec	F_1	Rec	Prec	F_1	Rec	Prec	F_1
U-Net	65,32	**92,67**	76,62	70,60	84,69	76,99	64,70	**84,16**	73,15
НКД-БИ	70,96	44,19	54,46	**88,51**	19,62	32,13	**98,56**	14,07	24,62
НКД	66,09	86,20	74,81	83,41	68,01	74,92	88,42	48,85	62,93
СНКД-БИ	68,13	87,15	76,47	55,11	**90,99**	68,65	70,83	77,01	73,79
СНКД	**72,88**	84,43	**78,23**	86,36	75,54	**80,58**	83,34	70,33	**76,28**

Нейросетевое квадродерево и СНКД во всех случаях превосходят по значениям метрики F_1 аналоги без информирования: прирост значений достигает 42,79 %. Согласно критерию Фридмана при заданном уровне значимости 0,05 различия значений F_1 для пяти рассмотренных архитектур и в этой ситуации являются статистически значимыми (P-значение равно 0,00069). Сбалансированное НКД снова демонстрирует лучшую стабильность при обработке малых объектов по сравнению с НКД, результаты которого характеризуются высоким приростом (до 23,72 %) метрики Rec при значимом уменьшении значений Prec.

Заключение

В статье представлена графово-свёрточная архитектура СНКД для сегментации малоразмерных объектов в условиях выраженного дисбаланса разделяемых классов. Блок графовых свёрток в СНКД информирован структурой квадродерева, которая обеспечивает НС дополнительной информацией о пространственных и иерархических связях между пикселями в разных пространственных разрешениях, необходимой в случае работы с малыми наборами данных.

Сбалансированное нейросетевое квадродерево продемонстрировало успехи при сегментации небольших объектов на различных по характеристикам РЛ-изображениях. Оно детектирует небольшие объекты эффективнее свёрточной сети и базового НКД в условиях дисбаланса классов в задаче как много-, так и двухклассовой сегментации (выделение цели на фоне). В первом случае прирост значений F_1 для малых объектов составляет от 1,76 до 47,42 % (в среднем 13,73 %), а во втором — до 3,59 %.

В качестве направления дальнейших исследований можно выделить применение архитектуры СНКД и её модификаций в задаче обнаружения малоразмерных объектов с использованием суперпиксельных (Sasmal, Dhal, 2023) представлений обрабатываемых изображений. Реализация такого подхода позволит снизить сложность операции графовой свёртки и применять метод для обработки снимков высокого разрешения.

Работа выполнена при поддержке проекта № 075-15-2024-544 Министерства науки и высшего образования Российской Федерации. Для обучения нейронных сетей использована инфраструктура Центра коллективного пользования «Высокопроизводительные вычисления и большие данные» (ЦКП «Информатика») ФИЦ ИУ РАН (Москва).

Литература

1. *Достовалова А. М.* Нейросетевое квадродерево и его применение для сегментирования спутниковых изображений // Информатика и её применения. 2024. Т. 18. № 3. С. 77–85. DOI: 10.14357/19922264240410.
2. *Bergstra J.*, *Bengio Y.* Random search for hyper-parameter optimization // J. Machine Learning Research. 2012. V. 13. Iss. 10. P. 281–305.
3. *Chitta K.*, *Álvarez J. M.*, *Hebert M.* Quadtree generating networks: Efficient hierarchical scene parsing with sparse convolutions // Proc. 2020 IEEE Winter Conf. Applications of Computer Vision (WACV). Snowmass, CO, USA, 2020. P. 2009–2018. DOI: 10.1109/WACV45572.2020.9093449.
4. *Dostovalova A. M.* Using a model of a spatial–hierarchical quadtree with truncated branches to improve the accuracy of image classification // Izvestiya, Atmospheric and Oceanic Physics. 2023. V. 59. Iss. 12. P. 1255–1262. DOI: 10.1134/S0001433823120071.
5. *Dostovalova A. M.*, *Gorshenin A. K.* Neural network image classifiers informed by factor analyzers // Doklady Mathematics. 2024. V. 110. P. S35–S41. DOI: 10.1134/S106456242460204X.
6. *Dostovalova A.*, *Gorshenin A.* Small sample learning based on probability-informed neural networks for SAR image segmentation // Neural Computing and Applications. 2025. V. 35. P. 8285–8308. DOI: 10.1007/s00521-025-10997-x.
7. *Friedman M.* A comparison of alternative tests of significance for the problem of m rankings // The Annals of Mathematical Statistics. 1940. V. 11. Iss. 1. P. 86–92. DOI: 10.1214/aoms/1177731944.
8. *Gorshenin A. K.*, *Kuzmin V. Yu.* Statistical feature construction for forecasting accuracy increase and its applications in neural network based analysis // Mathematics. 2022. V. 10. Article 589. DOI: 10.3390/math10040589.
9. *Gorshenin A. K.*, *Vilyaev A. L.* Finite normal mixture models for the ensemble learning of recurrent neural networks with applications to currency pairs // Pattern Recognition and Image Analysis. 2022. Iss. 32. P. 780–792. DOI: 10.1134/S1054661822040058.
10. *Gorshenin A. K.*, *Vilyaev A. L.* Machine learning models informed by connected mixture components for short- and medium-term time series forecasting // AI. 2024. V. 5. Iss. 4. P. 1955–1976. DOI: 10.3390/ai5040097.
11. *Gorshenin A.*, *Kozlovskaya A.*, *Gorbunov S.*, *Kochetkova I.* Mobile network traffic analysis based on probability-informed machine learning approach // Computer Networks. 2024. V. 247. Article 110433. DOI: 10.1016/j.comnet.2024.110433.
12. *Jayaraman P. K.*, *Mei J.*, *Cai J.*, *Zheng J.* Quadtree convolutional neural networks // ECCV 2018. Lecture Notes in Computer Science. 2018. V. 11210. P. 554–569. DOI: 10.1007/978-3-030-01231-1_34.
13. *Jewsbury R.*, *Bhalerao A.*, *Rajpoot N. M.* A quadtree image representation for computational pathology // Proc. 2021 IEEE/CVF Intern. Conf. Computer Vision Workshops (ICCVW). Montreal, BC, Canada, 2021. P. 648–656. DOI: 10.1109/ICCVW54120.2021.00078.
14. *Karniadakis G. E.*, *Kevrekidis I. G.*, *Lu L. et al.* Physics-informed machine learning // Nature Reviews Physics. 2021. V. 3. P. 422–440. DOI: 10.1038/s42254-021-00314-5.
15. *Ke L.*, *Danelljan M.*, *Li X. et al.* Mask Transfiner for high-quality instance segmentation // Proc. 2022 IEEE/CVF Conf. Computer Vision and Pattern Recognition (CVPR). New Orleans, LA, USA, 2022. P. 4402–4411. DOI: 10.1109/CVPR52688.2022.00437.
16. *Kochetkova I.*, *Kushchazli A.*, *Burtseva S.*, *Gorshenin A.* Short-term mobile network traffic forecasting using seasonal ARIMA and Holt-Winters models // Future Internet. 2023. V. 15. Iss. 9. Article 290. DOI: 10.3390/fi15090290.
17. *Ma F.*, *Gao F.*, *Sun J. et al.* Attention Graph Convolution Network for image segmentation in big SAR imagery data // Remote Sensing. 2019. V. 11. Iss. 21. Article 2586. DOI: 10.3390/rs11212586.
18. *Özdemir Ö.*, *Sönmez E. B.* Weighted Cross-Entropy for unbalanced data with application on COVID X-ray images // 2020 Innovations in Intelligent Systems and Applications Conf. (ASYU). Istanbul, Turkey, 2020. P. 1–6. DOI: 10.1109/ASYU50717.2020.9259848.
19. *Pastorino M.*, *Montaldo A.*, *Fronda L. et al.* Multisensor and multiresolution remote sensing image classification through a causal hierarchical Markov framework and decision tree ensembles // Remote Sensing. 2021. V. 13. Iss. 5. Article 849. DOI: 10.3390/rs13050849.
20. *Pastorino M.*, *Moser G.*, *Serpico S. B.*, *Zerubia J.* Semantic segmentation of remote-sensing images through fully convolutional neural networks and hierarchical probabilistic graphical models // IEEE Trans. Geoscience and Remote Sensing. 2022. V. 60. P. 1–16. DOI: 10.1109/TGRS.2022.3141996.
21. *Potin P.*, *Bargellini P.*, *Laur H. et al.* Sentinel-1 mission operations concept // Proc. IEEE Intern. Geoscience and Remote Sensing Symp. Munich, Germany, 2012. P. 1745–1749.
22. *Powers D. M. W.* Evaluation: from precision, recall and F-measure to ROC, informedness, markedness and correlation // Intern. J. Machine Learning Technology. 2011. V. 2. Iss. 1. P. 37–63.
23. *Ronen T.*, *Levy O.*, *Golbert A.* Vision transformers with mixed-resolution tokenization // Proc. 2023 IEEE/CVF Conf. Computer Vision and Pattern Recognition (CVPRW). Vancouver, BC, Canada, 2023. P. 4613–4622. DOI: 10.1109/CVPRW59228.2023.00486.

24. *Ronneberger O.*, *Fischer P.*, *Brox T.* U-Net: Convolutional networks for biomedical image segmentation // MICCAI 2015. Lecture Notes in Computer Science. 2015. V. 9351. P. 234–241. DOI: 10.1007/978-3-319-24574-4_28.
25. *Sasmal B.*, *Dhal K. G.* A survey on the utilization of Superpixel image for clustering based image segmentation // Multimedia Tools and Applications. 2023. V. 82. P. 35493–35555. DOI: 10.1007/s11042-023-14861-9.
26. *Tang S.*, *Zhang J.*, *Zhu S.*, *Tan P.* Quadtree attention for vision transformers // https://arxiv.org/. arXiv:2201.02767. 2022. 16 p. DOI: 10.48550/arXiv.2201.02767.
27. *Wei Sh.*, *Zeng X.*, *Que Q. et al.* HRSID: A high-resolution SAR images dataset for ship detection and instance segmentation // IEEE Access. 2020. V. 8. P. 120234–120254. DOI: 10.1109/ACCESS.2020.3005861.

Segmentation of small objects in radar images using a graph-convolutional neural network implementation of quadtrees

A. M. Dostovalova, A. K. Gorshenin

Federal Research Center "Computer Science and Control" RAS, *Moscow 119333*, *Russia*
E-mails: adostovalova@frccsc.ru, *agorshenin@frccsc.ru*

The paper proposes an architecture for a balanced neural network quadtree to solve the problem of segmenting small objects in radar images with insufficient training data. This architecture includes a pretrained convolutional encoder that generates additional features for each pixel in the image, as well as a graph-convolutional neural network that enhances the processing of spatial relationships in the data through an integrated probabilistic quadtree model. The graph part contains a branch-pruning block that detects similarities between pixels across different spatial resolutions. Additionally, there is a feature-processing unit designed for scaled images, and a custom loss function to improve the accuracy of separating unbalanced classes. The U-Net architecture was used for segmenting radar images from Sentinel-1 and High Resolution Synthetic Aperture Radar Images Dataset (HRSID). The balanced neural network quadtree demonstrated improved segmentation quality for smaller objects in both multi-class and binary segmentation (i.e., object detection against the background) compared to a basic neural network quadtree and vanilla U-Net. As a result, F_1 scores for small object classes increased by 3.59% for binary segmentation and by 47.42% for multi-class segmentation. The developed approaches have potential to be applied to the problem of detection of small-size objects using superpixels on high-resolution images.

Keywords: neural quadtrees, probability-informed neural networks, graph convolution networks, SAR image segmentation, small objects

Accepted: 07.08.2025
DOI: 10.21046/2070-7401-2025-22-5-51-62

References

1. Dostovalova A. M., Neural quadtree and its applications for SAR imagery segmentation, *Informatika i ee primeneniya*, 2024, V. 18, Iss. 3, P. 77–85 (in Russian), DOI: 10.14357/19922264240410.
2. Bergstra J., Bengio Y., Random search for hyper-parameter optimization, *J. Machine Learning Research*, 2012, V. 13, Iss. 10, pp. 281–305.
3. Chitta K., Álvarez J. M., Hebert M., Quadtree generating networks: Efficient hierarchical scene parsing with sparse convolutions, *Proc. 2020 IEEE Winter Conf. Applications of Computer Vision* (*WACV*), Snowmass, CO, USA, 2020, pp. 2009–2018, DOI: 10.1109/WACV45572.2020.9093449.
4. Dostovalova A. M., Using a model of a spatial–hierarchical quadtree with truncated branches to improve the accuracy of image classification, *Izvestiya*, *Atmospheric and Oceanic Physics*, 2023, V. 59, Iss. 12, pp. 1255–1262, DOI: 10.1134/S0001433823120071.
5. Dostovalova A. M., Gorshenin A. K., Neural network image classifiers informed by factor analyzers, *Doklady Mathematics*, 2024, V. 110, pp. S35–S41, DOI: 10.1134/S106456242460204X.

6. Dostovalova A., Gorshenin A., Small sample learning based on probability-informed neural networks for SAR image segmentation, *Neural Computing and Applications*, 2025, V. 35, pp. 8285–8308, DOI: 10.1007/s00521-025-10997-x.
7. Friedman M., A comparison of alternative tests of significance for the problem of m rankings, *The Annals of Mathematical Statistics*, 1940, V. 11, Iss. 1, pp. 86–92, DOI: 10.1214/aoms/1177731944.
8. Gorshenin A. K., Kuzmin V. Yu., Statistical feature construction for forecasting accuracy increase and its applications in neural network based analysis, *Mathematics*, 2022, V. 10, Article 589, DOI: 10.3390/math10040589.
9. Gorshenin A. K., Vilyaev A. L., Finite normal mixture models for the ensemble learning of recurrent neural networks with applications to currency pairs, *Pattern Recognition and Image Analysis*, 2022, Iss. 32, pp. 780–792, DOI: 10.1134/S1054661822040058.
10. Gorshenin A. K., Vilyaev A. L., Machine learning models informed by connected mixture components for short- and medium-term time series forecasting, *AI*, 2024, V. 5, Iss. 4, pp. 1955–1976, DOI: 10.3390/ai5040097.
11. Gorshenin A., Kozlovskaya A., Gorbunov S., Kochetkova I., Mobile network traffic analysis based on probability-informed machine learning approach, *Computer Networks*, 2024, V. 247, Article 110433, DOI: 10.1016/j.comnet.2024.110433.
12. Jayaraman P. K., Mei J., Cai J., Zheng J., Quadtree convolutional neural networks, *ECCV 2018. Lecture Notes in Computer Science*, 2018, V. 11210, pp. 554–569, DOI: 10.1007/978-3-030-01231-1_34.
13. Jewsbury R., Bhalerao A., Rajpoot N. M., A quadtree image representation for computational pathology, *Proc. 2021 IEEE/CVF Intern. Conf. on Computer Vision Workshops* (*ICCVW*), Montreal, BC, Canada, 2021, pp. 648–656, DOI: 10.1109/ICCVW54120.2021.00078.
14. Karniadakis G. E., Kevrekidis I. G., Lu L. et al., Physics-informed machine learning, *Nature Reviews Physics*, 2021, V. 3, pp. 422–440, DOI: 10.1038/s42254-021-00314-5.
15. Ke L., Danelljan M., Li X. et al., Mask Transfiner for high-quality instance segmentation, *Proc. 2022 IEEE/CVF Conf. Computer Vision and Pattern Recognition* (*CVPR*), New Orleans, LA, USA, 2022, pp. 4402–4411, DOI: 10.1109/CVPR52688.2022.00437.
16. Kochetkova I., Kushchazli A., Burtseva S., Gorshenin A., Short-term mobile network traffic forecasting using seasonal ARIMA and Holt-Winters models, *Future Internet*, 2023, V. 15, Iss. 9, Article 290, DOI: 10.3390/fi15090290.
17. Ma F., Gao F., Sun J. et al., Attention Graph Convolution Network for image segmentation in big SAR imagery data, *Remote Sensing*, 2019, V. 11, Iss. 21, Article 2586, DOI: 10.3390/rs11212586.
18. Özdemir Ö., Sönmez E. B., Weighted Cross-Entropy for unbalanced data with application on COVID X-ray images, *2020 Innovations in Intelligent Systems and Applications Conference* (*ASYU*), Istanbul, Turkey, 2020, pp. 1–6, DOI: 10.1109/ASYU50717.2020.9259848.
19. Pastorino M., Montaldo A., Fronda L. et al., Multisensor and multiresolution remote sensing image classification through a causal hierarchical Markov framework and decision tree ensembles, *Remote Sensing*, 2021, V. 13, Iss. 5, Article 849, DOI: 10.3390/rs13050849.
20. Pastorino M., Moser G., Serpico S. B., Zerubia J., Semantic segmentation of remote-sensing images through fully convolutional neural networks and hierarchical probabilistic graphical models, *IEEE Trans. Geoscience and Remote Sensing*, 2022, V. 60, pp. 1–16, DOI: 10.1109/TGRS.2022.3141996.
21. Potin P., Bargellini P., Laur H. et al., Sentinel-1 mission operations concept, *Proc. IEEE Intern. Geoscience and Remote Sensing Symp.*, Munich, Germany, 2012, pp. 1745–1749.
22. Powers D. M. W., Evaluation: from precision, recall and F-measure to ROC, informedness, markedness and correlation, *Intern. J. Machine Learning Technology*, 2011, V. 2, Iss. 1, pp. 37–63.
23. Ronen T., Levy O., Golbert A., Vision transformers with mixed-resolution tokenization, *Proc. 2023 IEEE/CVF Conf. Computer Vision and Pattern Recognition* (*CVPRW*), Vancouver, BC, Canada, 2023, pp. 4613–4622, DOI: 10.1109/CVPRW59228.2023.00486.
24. Ronneberger O., Fischer P., Brox T., U-Net: Convolutional networks for biomedical image segmentation, *MICCAI 2015. Lecture Notes in Computer Science*, 2015, V. 9351, pp. 234–241, DOI: 10.1007/978-3-319-24574-4_28.
25. Sasmal B., Dhal K. G., A survey on the utilization of Superpixel image for clustering based image segmentation, *Multimedia Tools and Applications*, 2023, V. 82, pp. 35493–35555, DOI: 10.1007/s11042-023-14861-9.
26. Tang S., Zhang J., Zhu S., Tan P., Quadtree attention for vision transformers, *https://arxiv.org/*, arXiv:2201.02767, 2022, 16 p., DOI: 10.48550/arXiv.2201.02767.
27. Wei Sh., Zeng X., Que Q. et al., HRSID: A high-resolution SAR images dataset for ship detection and instance segmentation, *IEEE Access*, 2020, V. 8, pp. 120234–120254, DOI: 10.1109/ACCESS.2020.3005861.