

На правах рукописи



**Достовалова Анастасия Михайловна**

**Вероятностно-информированные нейросетевые модели  
анализа изображений при ограниченных обучающих  
данных**

Специальность 1.2.1 —  
«Искусственный интеллект и машинное обучение»

**Автореферат**  
диссертации на соискание учёной степени  
кандидата физико-математических наук

Москва — 2026

Работа выполнена в отделе 64 Федерального исследовательского центра «Информатика и управление» Российской академии наук (ФИЦ ИУ РАН).

Научный руководитель: доктор физико-математических наук, доцент, главный научный сотрудник ФИЦ ИУ РАН  
**Горшенин Андрей Константинович**

Официальные оппоненты: **Визильтер Юрий Валентинович**, доктор физико-математических наук, профессор РАН, Государственный научно-исследовательский институт авиационных систем, директор по направлению – руководитель научного комплекса «Искусственный интеллект и техническое зрение»

**Мясников Владислав Валерьевич**, доктор физико-математических наук, профессор, Самарский национальный исследовательский университет имени академика С.П. Королева, профессор кафедры геоинформатики и информационной безопасности

Ведущая организация: Федеральное государственное бюджетное учреждение науки **Институт системного программирования им. В.П. Иванникова Российской академии наук**

Защита состоится 25 июня 2026 г. в «        » часов на заседании диссертационного совета 24.1.224.03 при Федеральном государственном учреждении «Федеральный исследовательский центр «Информатика и управление» Российской академии наук» по адресу: 119333, Москва, Вавилова, д.42.

С диссертацией можно ознакомиться в библиотеке ФИЦ ИУ РАН и на официальном сайте [https://www.frccsc.ru/diss-council/00207305/diss/list/dostovalova\\_am](https://www.frccsc.ru/diss-council/00207305/diss/list/dostovalova_am).

Отзывы на автореферат в двух экземплярах, заверенные печатью учреждения, просьба направлять по адресу: 119333, Москва, Вавилова, д.42, ученому секретарю диссертационного совета 24.1.224.03.

Автореферат разослан \_\_\_\_\_ 2026 года.

Телефон для справок: +7 (499) 135-62-60.

Ученый секретарь  
диссертационного совета  
24.1.224.03,  
кандидат технических наук



Рейер Иван Александрович

## Общая характеристика работы

**Актуальность темы.** Извлечение новых знаний из массивов накопленных данных в таких областях, как молекулярная биология, инженерия, медицина и обработка данных космических систем и беспилотных летательных аппаратов (БПЛА), является одной из важнейших современных исследовательских задач. Часто анализ данных подразумевает распознавание образов или сигналов, включающее в себя математическое моделирование их характерных закономерностей, например, на основе мозаичных представлений<sup>1</sup> однородных участков цифровых изображений, развиваемых в работах Ю.В. Визильтера, или описание специфичной изменчивости с помощью аппарата теории вероятностей и математической статистики. Фундаментальные основы таких подходов были заложены в классических трудах К. Пирсона<sup>2</sup>, К. Фукунаги<sup>3</sup>, академиков А.Н. Колмогорова<sup>4</sup>, А.А. Харкевича<sup>5</sup>, Ю.И. Журавлева<sup>6</sup> и К.В. Рудакова<sup>7</sup>. Ряд из этих основ также стал фундаментом алгоритмов нейронных сетей.

Искусственные нейронные сети (перцептрон) впервые были представлены еще в 1960 годах. Но только в последние десятилетия развитие средств вычислительной техники, алгоритмов оптимизации весов, создание более сложных архитектур, таких как сверточные и рекуррентные сети в работах нобелевского лауреата Дж. Хинтона<sup>8</sup>, Я. Лекуна<sup>9</sup>, Д. Бенджио<sup>10</sup> и др., позволило эффективно анализировать с их помощью объекты окружающего мира. В условиях, требующих высокой эффективности,

---

<sup>1</sup> *Визильтер Ю., Желтов С., Бусурин В.* Современный морфологический анализ и его применение в авиационных системах технического зрения. Москва : Московский авиационный институт (национальный исследовательский университет), 2020. С. 160.

<sup>2</sup> *Pearson K.* The Grammar of Science. Adam, Charles Black, 1900. С. 548.

<sup>3</sup> *Fukunaga K.* Introduction to Statistical Pattern Recognition. New York : Academic Press, 1972. С. 369.

<sup>4</sup> *Колмогоров А. Н.* Теория вероятностей и математическая статистика. Москва : Наука, 1986. С. 534.

<sup>5</sup> *Харкевич А. А.* Теория информации и ее приложения: Сборник переводов иностранных статей по теории информации и ее приложениям к связи. Москва : Гос. изд. физико-математической литературы, 1959. С. 329.

<sup>6</sup> *Журавлев Ю. И., Гуревич И. Б.* Распознавание образов и распознавание изображений // Распознавание, классификация, прогноз. Математические методы и их применение. М. : Наука, 1989. С. 5—72.

<sup>7</sup> *Рудаков К.* Об алгебраической теории универсальных и локальных ограничений для задач классификации // Распознавание, классификация, прогноз. М.: Наука. 1989. С. 176—201.

<sup>8</sup> *Rumelhart D. E., Hinton G. E., Williams R. J.* Learning representations by back-propagating errors // Nature. 1986. Т. 323, № 6088. С. 533.

<sup>9</sup> *Lecun Y.* [и др.]. Gradient-based learning applied to document recognition // Proceedings of the IEEE. 1998. Т. 86, № 11. С. 2278—2324.

<sup>10</sup> *Bengio Y.* [и др.]. A neural probabilistic language model // Journal of Machine Learning Research. 2003. Т. 3. С. 1137—1155.

адаптивности и точности анализа, алгоритмы искусственного интеллекта (ИИ) оказываются незаменимы.

Сверточные нейронные сети (НС) являются традиционными архитектурами, применяемыми для обработки изображений. С момента появления они претерпели множество модификаций, которые позволили повысить точность распознавания сложных сцен и увеличить масштаб обработки изображения – с  $32 \times 32$  до  $2048 \times 2048$  пикселей. Их основным недостатком является ограниченная способность понимания глобальных взаимосвязей между элементами изображения, что сильно усложняет анализ снимков высокого разрешения. Альтернативой сверточным сетям стали трансформерные архитектуры. В сети Vision Transformer изображение впервые было представлено как набор токенов, связи между которыми обрабатывались с использованием механизма самовнимания (англ. self-attention) для выявления наиболее важных связей между элементами векторов данных.

Трансформеры успешно адаптированы для решения обработки специализированных типов изображений, например, медицинских<sup>11</sup>. Однако их существенными недостатками являются вычислительная сложность и потребность в большом количестве обучающих данных. Множество модификаций трансформеров были направлены на повышение производительности сети и интерпретируемости результата. Однако в прикладных и научных задачах, например, в сфере обработки спутниковых изображений, сложные модифицированные сверточные или трансформерные архитектуры нередко демонстрируют более слабые результаты, чем при анализе обычных данных. Причиной этого являются характерные особенности предметной области, такие как сложность проведения наблюдений и экспериментов или уникальность исследуемых явлений, из-за которых доступные обучающие наборы оказываются ограниченными по некоторым характеристикам, а именно являются несбалансированными, малыми по числу элементов или сильно неоднородными.

Несбалансированность обучающего набора означает, что элементов некоторой категории объектов в нем существенно меньше, чем остальных. Примером такого типа данных являются изображения, содержащие малоразмерные<sup>12</sup> или фоновые объекты – на панорамах улиц или аэрокосмических (спутниковых и БПЛА) снимках поверхности Земли. При этом несбалансированный набор может содержать довольно большое количество наблюдений. Малым же набор считается, если содержит всего

---

<sup>11</sup> *Gliner V., Makarov V., Avetisyan A. I.* [и др.] Using domain adaptation for classification of healthy and disease conditions from mobile-captured images of standard 12-lead electrocardiograms // *Scientific Reports*. 2023. Т. 13, № 1. С. 14023.

<sup>12</sup> *Sang S.* [и др.]. Small-Object Sensitive Segmentation Using Across Feature Map Attention // *IEEE Transactions on Pattern Analysis and Machine Intelligence*. 2023. Т. 45, № 5. С. 6289–6306.

порядка нескольких сотен или тысяч наблюдений<sup>13</sup> (в литературе обычно границы определены нечетко и изменяются от 100 до 2-3 тысяч элементов). Наконец, под неоднородными или изменчивыми в рамках диссертации понимаются наборы с выраженной внутриклассовой дисперсией. Неоднородность может быть обусловлена особенностями источников данных или оборудования, как, например, географическая и суточная изменчивость при аэрокосмической съемке. Нередко неоднородность является следствием ограниченности и несбалансированности обучающего набора.

Общей чертой неоднородных, несбалансированных и малых по числу элементов датасетов является отсутствие достаточного количества информативных признаков для построения правильных закономерностей. Строящееся на таком наборе решающее правило не воспроизводит вариативность реальных данных, что, аналогично случаю малых выборок в статистике, повышает подверженность правила к искажениям и переобучению. В этом случае универсальная стратегия добавления данных из открытых датасетов является неэффективной. Объединение данных, различных по свойствам, зачастую приводит к усилению внутриклассовой дисперсии, особенно если исходный датасет ограничен.

Проблема построения решающих правил в условиях ограниченности обучающих признаков характерна для научных и технических задач, в которых требуется разработка высокоточных, надежных, доверенных<sup>14</sup> методов ИИ. Высокая потребность в автоматизации обработки ограниченных специфических датасетов привела к созданию множества методов, альтернативных традиционному расширению набора открытыми данными. Наиболее популярными являются методы семплирования, аугментации, целенаправленно увеличивающие количество обучающих примеров малочисленных классов, и даже генеративные подходы. Развита альтернативные подходы, в том числе различные техники регуляризации. Перспективными являются архитектурные модификации – специальные модули внимания или графово-сверточные решения. Наконец, применяются методы переноса обучения, настраиваемые на больших и разнообразных открытых датасетах.

В научных и технических задачах в качестве отдельного направления выделяют подход информирования, предлагающий использовать математические модели данных при обучении НС для учета особенностей и ограничений предметной области без привлечения новых наблюдений. Впервые

---

<sup>13</sup> *Safonova A.* [и др.]. Ten deep learning techniques to address small data problems with remote sensing // International Journal of Applied Earth Observation and Geoinformation. 2023. Т. 125. С. 103569.

<sup>14</sup> *Турдаков Д., Аветисян А.* [и др.] Доверенный искусственный интеллект: вызовы и перспективные решения // Доклады Российской Академии Наук. Математика, информатика, процессы управления. 2022. Т. 508, № 1. С. 13–18.

идея привлечения<sup>15,16</sup> математических моделей для повышения эффективности методов ИИ возникла в 90-х годах XX века: НС использовались для решения обыкновенных дифференциальных уравнений и уравнений в частных производных. Термин «физически-информированные нейронные сети» ввел Дж. Карниадакис в 2017-2019 годах, формализовав подход и сделав его более удобным для приложений<sup>17</sup>. При информировании математические модели или выполняют роль источника дополнительных признаков, или вводят ограничения на область поиска решения, справедливые для предметной области, которые в условиях ограниченных наборов не могут быть выведены напрямую из данных. По способу реализации можно выделить три основных способа информирования НС (рисунок 1). При информировании на уровне признаков вектор входных данных расширяется его характеристиками, полученными из математической модели. При информировании на уровне функции потерь к стандартным обучающим функционалам добавляются регуляризирующие слагаемые, гарантирующие выполнение свойств модели, например, учет граничных условий. Наконец, при информировании на уровне архитектуры в сеть добавляются специальные блоки, в которых этапы обработки данных математической моделью повторяются с помощью нейросетевых слоев.

Подход информирования продемонстрировал высокую эффективность в задачах математической физики, требующих решения систем дифференциальных уравнений в частных производных<sup>18</sup> – Бюргерса, Курамото-Сивашинского, Навье-Стокса и др.. Успехи продемонстрированы в области гидродинамики, а также для моделирования климата, тепловых и сверхзвуковых потоков, колебательных движений, оценке надежности конструкций и медицине.

В задачах обработки изображений применение подхода информирование сопряжено со сложностями, из-за отсутствия для них в общем случае универсальных адекватных физических моделей. Известные применения подхода, как правило, реализованы на уровне признаков: используются спектральные представления снимков, полученные преобразованием Фурье, условия фотометрической согласованности и геометрические характеристики объектов на изображениях, такие как их расположение или азимутальный угол. В качестве дополнительных данных, как показывают

---

<sup>15</sup> *Dissanayake M. W. M. G., Phan-Thien N.* Neural network-based approximations for solving partial differential equations // *Communications in Numerical Methods in Engineering*. 1994. Т. 10, № 3. С. 195–201.

<sup>16</sup> *Lagaris I., Likas A., Fotiadis D.* Artificial Neural Networks for Solving Ordinary and Partial Differential Equations // *IEEE Transactions on Neural Networks*. 1998. Окт. С. 987–1000.

<sup>17</sup> *Karniadakis G.* [и др.]. Physics-Informed Machine Learning // *Nature Reviews Physics*. 2021. Т. 3. С. 422–440.

<sup>18</sup> *Raissi M., Perdikaris P., Karniadakis G.* Physics-Informed Neural Networks: A Deep Learning Framework for Solving Forward and Inverse Problems Involving Nonlinear Partial Differential Equations // *Journal of Computational Physics*. 2019. Т. 378. С. 686–707.

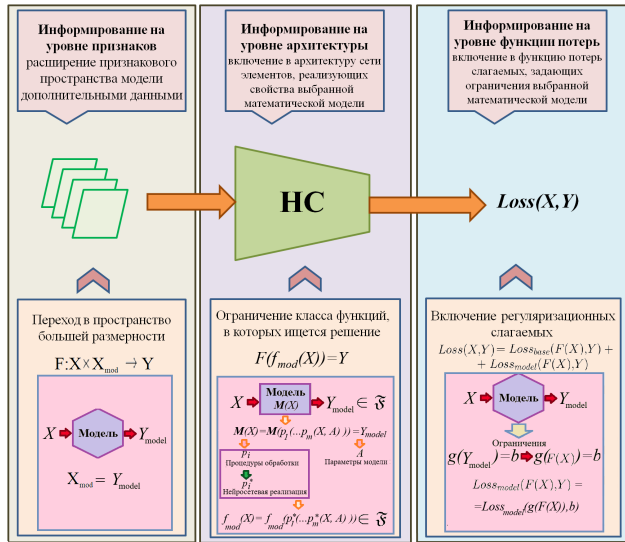


Рисунок 1 — Способы реализации информирования нейронных сетей

работы Мясникова В.В., могут выступать также данные другой модальности, например, радиолокационные снимки для обработки оптических изображений<sup>19</sup>, а также результаты специализированной предобработки признаков разными математическими моделями<sup>20</sup>.

Для изображений как многомерных сигналов зачастую справедливыми оказываются вероятностно-статистические представления, например, описывающие поведение шумовой составляющей сигнала. Информирование сетей вероятностными моделями продемонстрировало заметные успехи в задачах обработки последовательностей и временных рядов, например при оценке неопределенности предсказаний<sup>21</sup>, надежности инженерных систем и функций риска. Вероятностные модели также могут описывать пространство скрытых признаков или процесс их слияния.

<sup>19</sup> Konovalov V., Myasnikov V., Sergeev V. Reconstruction of Optical Remote Sensing Images Using Synthetic Aperture Radar Data and Diffusion Neural Networks // Optoelectronics, Instrumentation and Data Processing. 2025. Т. 61, № 4. С. 449–457.

<sup>20</sup> Kozlova Y., Myasnikov V. Head model reconstruction and animation method using color image with depth information // Computer Optics. 2024. Т. 48, № 1. С. 118–122.

<sup>21</sup> Tyralis H., Papacharalampous G. A review of predictive uncertainty estimation with machine learning // Artificial Intelligence Review. 2024. Т. 57, № 4. С. 94.

Как демонстрируют работы А.К. Горшенина<sup>22,23,24</sup>, моделирование структурных особенностей стохастических процессов с помощью модели конечной смеси вероятностных распределений (формирования на ее основе компонент связности временного ряда) и использование ее характеристик (значений математического ожидания, дисперсии и др.) в качестве дополнительных входных признаков существенно повышает точность предсказаний нейронных сетей, в том числе для анализа малых наборов – данных о состоянии турбулентной плазмы и информационных потоков. Встраивание элементов вероятностной модели в архитектуру нейронной сети также демонстрирует эффективность в задачах предсказания трафика и геофизических временных рядов.

Основным преимуществом подхода информирования является его универсальность и применимость для сильно ограниченных и специфичных данных. Например, методы переноса обучения часто сталкиваются с трудностями из-за различия предметных областей базового и целевого наборов, кроме того, часто они требуют использования дополнительных метаданных, которые не всегда существуют или доступны. Дополнение набора синтетическими данными решает задачу перераспределения вероятностей признаков, выделенных нейронной сетью, и сталкивается с проблемой воспроизведения вариативности реальных данных<sup>25</sup>. По этой причине задача развития методов информирования нейронных сетей для обработки ограниченных наборов изображений является весьма актуальной.

**Целью** данного научного исследования является развитие теоретически обоснованных вероятностно-информированных нейросетевых моделей для решения задач обработки неоднородных, несбалансированных и малых наборов изображений.

Для достижения поставленной цели необходимо было решить следующие **задачи**:

1. Развить теоретические основы, определяющие выбор вероятностной модели и способа информирования в задачах обработки ограниченных наборов изображений.
2. Выполнить аналитическое исследование разработанных моделей и информированных архитектурных блоков нейронных сетей в условиях ограниченных данных.

---

<sup>22</sup> Gorshenin A. [и др.]. Mobile network traffic analysis based on probability-informed machine learning approach // Computer Networks. 2024. Т. 247. С. 110433.

<sup>23</sup> Gorshenin A., Kuzmin V. Statistical Feature Construction for Forecasting Accuracy Increase and Its Applications in Neural Network Based Analysis // Mathematics. 2022. Т. 10. С. 589.

<sup>24</sup> Gorshenin A., Vilyaev A. Machine Learning Models Informed by Connected Mixture Components for Short- and Medium-Term Time Series Forecasting // AI. 2024. Т. 5, № 4. С. 1955–1976.

<sup>25</sup> Zhou Y. [и др.]. SAR Target Classification with Limited Data via Data Driven Active Learning // IGARSS 2020–2020 IEEE International Geoscience and Remote Sensing Symposium. 2020. С. 2475–2478.

3. Разработать методы информирования вероятностными моделями различных нейросетевых архитектур.
4. Апробировать разработанные подходы для решения задач классификации и сегментации несбалансированных, малых и неоднородных наборов изображений.

**Методология и методы исследования.** В диссертации предложены оригинальные подходы и процедуры, в том числе:

- метод информирования нейронных сетей с помощью дополнительных входных признаков, построенных с помощью смешанных вероятностных моделей для повышения точности сегментации неоднородных наборов изображений;
- метод информирования нейронных сетей композицией вероятностных моделей конечной смеси распределений и случайного поля Маркова для повышения точности сегментации малых и неоднородных наборов изображений;
- метод архитектурного информирования моделью факторного анализатора построения блока слияния признаков для повышения точности классификации ограниченных по числу элементов наборов изображений;
- метод архитектурного информирования моделью случайного Марковского поля для сегментации сильно несбалансированных наборов изображений;
- метод повышения точности выделения на изображениях мало-размерных объектов с помощью информированных графовых ансамблей .

Применяются и такие классические методы исследования, как аналитический аппарат теории вероятностей и математической статистики для смешанных распределений, методы параметрического и непараметрического статистического оценивания, проверка статистических гипотез, методы линейной алгебры, в том числе для оценки сложности предлагаемых вычислительных процедур, алгоритмы машинного обучения и нейронные сети.

Для создания комплекса программных решений, предназначенных для анализа данных, использованы языки программирования C++ и Python, а также современные высокопроизводительные вычислительные ресурсы ЦКП «Информатика» ФИЦ ИУ РАН.

**Научная новизна:** В диссертации впервые выполнена адаптация подходов вероятностного информирования нейронных сетей для решения задач обработки изображений, включающая:

1. теоретическое доказательство свойств вероятностных моделей изображений, включающих случайное поле Маркова в виде квадродерева и факторный анализатор с импульсно-аддитивным шумом, обосновывающих выбор способа информирования

- и возможность повышения с их помощью точности обработки ограниченных наборов данных;
2. теоретическое доказательство свойств информированных архитектурных блоков, включающих оценку вычислительной сложности, повышение точности обработки малых наборов и ускорение обучения сети;
  3. теоретически обоснованный метод архитектурного информирования моделью случайного Марковского поля графовых нейросетей для повышения точности обработки несбалансированных наборов изображений;
  4. теоретически обоснованный метод архитектурного информирования моделью факторного анализатора сверточных нейросетей для повышения точности классификации малых наборов изображений;
  5. теоретически обоснованный метод информирования композицией вероятностных моделей смеси и Марковского случайного поля для учета признаков разной природы для обработки сильно неоднородных наборов данных.

#### **Основные положения, выносимые на защиту:**

1. Метод архитектурного информирования моделью Марковского случайного поля нейронных сетей с доказательством теорем о более быстром убывании функции потерь.
2. Метод архитектурного информирования моделью факторного анализатора с импульсно-аддитивным шумом в блоке слияния признаков с доказательством теорем о его аналитических свойствах и оценке вычислительной сложности.
3. Метод комбинированного информирования на уровне признаков и архитектуры сети композицией моделей конечной смеси вероятностных распределений и случайного поля Маркова с доказательством теоремы о повышении точности обработки неоднородных наборов данных.
4. Аналитические свойства модели Марковского случайного поля в виде квадродрева, в том числе теорема о связи с графовыми нейронными сетями.

Работа соответствует следующим пунктам паспорта специальности 1.2.1: 15, 5, 4 и 17.

**Практическая значимость** Результаты диссертации являются одновременно фундаментальными и прикладными. Эффективность разработанных методов анализа данных и вычислительных процедур определяется полученными в диссертации математическими результатами, включающими обоснование выбора способа информирования сети, оценки вычислительной сложности и скорости обучения вероятностно-информированных нейросетевых блоков. Эти результаты подтверждаются

и путем всестороннего тестирования созданных методов на реальных ограниченных наборах изображений, что продемонстрировано в диссертации на примерах анализа аэрокосмических снимков земной поверхности (полученных как с помощью спутников, так и БПЛА) и открытых наборов обычных изображений.

**Достоверность** полученных результатов обеспечивается аналитическими доказательствами свойств предлагаемых методов и обоснованием их эффективности для работы с ограниченными наборами данных, всесторонним эмпирическим тестированием предложенных методов на разнообразных по свойствам открытых наборах данных, а также соответствием с результатами, полученными другими авторами.

**Апробация работы.** Основные результаты работы докладывались на следующих научных мероприятиях:

- Международная конференция по искусственному интеллекту AI Journey: 2024 г. [3], 2025 г. [5];
- Международная конференция «Интеллектуальные системы» (INTELS'24): 2024 г. [9];
- Конференция молодых учёных «Фундаментальные и прикладные космические исследования»: 2025 г.;
- Международная научная конференция студентов, аспирантов и молодых учёных «Ломоносов»: 2024 г. [12], 2025 г. [13];
- Научная конференция «Ломоносовские чтения»: 2025 г. [10];
- Научная конференция «Тихоновские чтения»: 2024 г. [11];
- Научный семинар факультета компьютерных наук НИУ ВШЭ под руководством профессора А.А. Наумова: 2026 г.;
- Научный семинар кафедры математической статистики ВМК МГУ «Интеллектуальные методы вычислительной статистики»: 2023-2025 гг.
- Научный семинар кафедры математической статистики ВМК МГУ «Теория риска и смежные вопросы»: 2026 г.;
- Научные семинары отделений 1, 5 и 6 ФИЦ ИУ РАН: 2025-2026 гг.

Основные результаты диссертации получены в процессе выполнения работ по проекту №075-15-2024-544 «Математические модели и численные методы как основа для разработки робототехнических комплексов, новых материалов и интеллектуальных технологий конструирования» Министерства науки и высшего образования Российской Федерации.

**Публикации.** Основные результаты по теме диссертации изложены в 13 печатных изданиях, 9 из которых – статьи в рецензируемых изданиях из перечня ВАК и учитываемых наравне с ними международных баз Web of Science и Scopus, 4 – в тезисах докладов.

**Личный вклад.** Основные результаты диссертации получены автором самостоятельно под научным руководством д.ф.-м.н. А.К. Горшенина. В работах [3; 8; 11; 13] А.М. Достоваловой изучены существующие решения

и разработаны новые подходы вероятностного информирования для повышения точности обработки малых наборов данных. В работах [2; 4; 10; 12] А.М. Достоваловой разработаны подходы вероятностного информирования для решения задач обработки неоднородных и несбалансированных наборов изображений, а также проведен всесторонний анализ полученных результатов. В работах [1–7; 9] А.М. Достоваловой доказан ряд теорем, обосновывающих эффективность предлагаемых способов информирования сетей в конкретных задачах, а также развиты и исследованы математические модели, методы и вычислительные алгоритмы анализа неоднородных и несбалансированных данных с реализацией в виде программных решений и их приложениями к обработке аэрокосмических изображений Земной поверхности.

## Содержание работы

Диссертация состоит из трех глав, каждая из которых посвящена решению задачи обработки одного из типов ограниченных наборов изображений – малого, неоднородного или несбалансированного датасета. Кратко остановимся на содержании работы. Во **введении** обосновывается актуальность исследований, проводимых в рамках данной диссертационной работы, приводится обзор научной литературы по изучаемой проблеме, формулируется цель, ставятся задачи работы, излагается научная новизна и практическая значимость представляемой работы.

**Первая глава** посвящена информированной нейросетевой модели слияния глобальных признаков изображений для повышения точности классификации снимков в условиях малых наборов. В разделе 1.1 содержится постановка задачи. В разделе 1.2 представлена новая вероятностная модель факторного анализатора с импульсным и аддитивным шумами для моделирования процесса слияния признаков. Предполагается, что импульсная помеха полностью искажает полезную информацию в сигнале и все искаженные элементы могут быть приравнены к постоянному значению  $y_0$ :

$$Z = \eta \cdot (A \cdot \Theta + \xi_0) + y_0 \cdot (I - \eta), \quad (1)$$

где  $Z \in \mathbb{R}^{1 \times Kz}$  – наблюдаемое значение,  $\eta$  – матрица, на диагоналях которой расположены независимые одинаково распределенные биномиальные случайные величины с вероятностью успеха  $d_{prob}$ , описывающие импульсную помеху,  $I$  – единичная матрица,  $A$  – матрица нагрузок,  $\Theta \in \mathbb{R}^{M_0 \times K_\theta}$  – матрица факторов модели и  $\xi_0$  – стандартная нормально распределенная случайная величина, описывающая аддитивную помеху.

Теоремы 1–3 обосновывают выбор модели (1) для информирования: такие НС могут быть способны с более высокой точностью восстанавливать распределение реальных данных даже в случаях ограниченных наборов.

**Теорема 1.** Система случайных величин, составляющих факторную модель  $Z = \eta \cdot (A \cdot \Theta + \xi_0) + y_0 \cdot (I - \eta)$  с импульсным и аддитивным шумами, идентифицируема.

**Теорема 2.** Пусть факторы (строки матрицы  $\Theta$ )  $\theta_i$ ,  $i = 1 \dots M_0$ , центрированы,  $\xi_0 \sim N(0, \Sigma_{\xi_0})$ ,  $y_0 = 0$ , а вероятность  $d_{\text{гроб}}$  известна. Тогда достаточное условие однозначной определенности матриц  $A, \Sigma_{\xi_0}$  состоит в том, что размерности  $Z$  и  $\Theta$  связаны соотношением  $K_Z > 2 \cdot K_\theta + 1$ .

**Теорема 3.** Пусть  $\mathbb{E}(\Theta) = 0$ ,  $\eta$ , случайные величины  $\Theta$ ,  $\xi_0$  независимы, а значения  $y_0$  и  $d_{\text{гроб}}$  известны. Тогда:

- если  $\xi_0 \sim N(y_0, \Sigma_{\xi_0})$ ,  $\Sigma_{\xi_0} = \sigma^2 \cdot I$ , оценки метода наименьших квадратов (МНК-оценки) матрицы  $A$  являются несмещенными и состоятельными;
- если  $\xi_0 \sim N(0, \Sigma_{\xi_0})$ ,  $\Sigma_{\xi_0} = \sigma^2 \cdot I$ , то МНК-оценка  $A$  является ее оценкой максимального правдоподобия (ОМП-оценкой);
- если  $\xi_0 \sim N(0, \Sigma_{\xi_0})$ ,  $\Sigma_{\xi_0} = \sigma^2 \cdot I$ , и оценки параметров  $A, \Sigma_{\xi_0}$  получены при минимизации кросс-энтропии, то эти оценки являются несмещенными и состоятельными.

Теорема 1 утверждает, что параметры факторного анализатора с аддитивным и импульсным шумами могут быть определены однозначно. В теореме 2 установлены ограничения на количество факторов модели, следующие из теоремы Андерсона-Рубина<sup>26</sup>. Теорема 3 устанавливает, что оценки параметров модели (1), полученные в ходе минимизации кросс-энтропии, являются несмещенными и состоятельными, что означает отсутствие подверженности систематическим искажениям.

В разделе 1.3 представлена информированная факторами в блоке слияния признаков нейросеть Factor Fusion Neural Network (FtFNN). Поскольку модель (1) использовалась для объединения глобальных признаков, информирование было реализовано на уровне архитектуры сети. FtFNN состоит из кодировщика и классификатора  $G(\cdot)$ , реализующего слияние глобальных признаков снимка в разных пространственных разрешениях, которые рассматриваются в качестве факторов модели. Обработка представлений снимка в  $G(\cdot)$  разделена на три этапа, повторяющих этапы обработки факторным анализатором: отображение в пространство факторов, учет шумовой составляющей и умножение на матрицу нагрузок.

Конкретный вид архитектуры FtFNN определяется гиперпараметрами модели. Доказано (теорема 4), что при определенных значениях параметров классификатор FtFNN оказывается вычислительно проще аналогов, построенных в виде композиции полносвязных и сверточных слоев

<sup>26</sup>Anderson T. W., Rubin H. Statistical inference in factor analysis // In Proceedings of the Third Berkeley Symposium on Mathematical Statistics and Probability. Т. 15. 1956. С. 111–150.

(величины  $T_{base}^{(1)}$  и  $T_{base}^{(2)}$  означают вычислительную сложность полносвязного и сверточного классификаторов).

**Теорема 4.** Пусть число факторов в скрытом пространстве  $M \in \mathbb{N}$ , параметры  $hid = K$ ,  $M \cdot scale \cdot K \in \mathbb{N}$  и веса факторов  $a_i$  являются фиксированными параметрами, а для уменьшения размерности выходного слоя кодировщика  $N_{M_0} \times N_{M_0} \times H_{M_0}$  ( $Pool_{sub} = Pool - N_{M_0}^2 \cdot H_{M_0}$ , где  $Pool$  – размерность вектора после пулинга) используется пулинг в среднем. Пусть также  $H = \sum_{i=1}^3 H_i$ ,  $F_{scale}^{(2)}$  – размерность вектора признаков после обработки его сверточным слоем. Тогда число операций, выполняемых в классификаторе FtFNN, меньше, чем в базовом классификаторе на основе полносвязного слоя (формула (1.13)) если:

1.  $M < \frac{H_{M_0} - \frac{Pool_{sub}}{K}}{(H - (1+K \cdot scale)) \cdot scale}$ , когда первый этап реализован полносвязным слоем, а третий – блоками (I) или (II), реализующими умножение на матрицу нагрузок, из архитектуры FtFNN;
2.  $M < \frac{\sqrt{(H+1)^2 + 4 \cdot (H_{M_0} - \frac{Pool_{sub}}{K}) \cdot K} - (H+1)}{2 \cdot K \cdot scale}$ , когда первый этап реализован полносвязным слоем, а третий – блоками (III) или (IV), реализующими умножение на матрицу нагрузок, из архитектуры FtFNN;
3.  $scale < \frac{H_{M_0} \cdot K - Pool_{sub} - (\frac{9 \cdot (H) \cdot (F_{scale}^{(2)} - 2)^2}{(F_{scale}^{(2)})^2} + (F_{scale}^{(2)})^2)}{((F_{scale}^{(2)})^2 + K) \cdot K}$  и  $K > 9 \cdot \frac{(F_{scale}^{(2)} - 2)^2}{(F_{scale}^{(2)})^2}$ , когда первый этап реализован сверточным слоем, а третий – блоками (I) или (II);
4.  $scale < \frac{H_{M_0} \cdot K - Pool_{sub} - (\frac{9 \cdot H \cdot (F_{scale}^{(2)} - 2)^2}{(F_{scale}^{(2)})^2} + (F_{scale}^{(2)})^2)}{(M \cdot (F_{scale}^{(2)})^2 + K) \cdot K}$  и  $K > 9 \cdot \frac{(F_{scale}^{(2)} - 2)^2}{(F_{scale}^{(2)})^2}$ , когда первый этап реализован сверточным слоем, а третий – блоками (III) или (IV).

Также число операций, выполняемых в классификаторе FtFNN, меньше, чем в базовом классификаторе на основе сверточных слоев (формула (1.14)) если:

1.  $M < \frac{T_{base}^{(2)} - Pool_{sub} - scale \cdot K^2}{(H+1+K \cdot scale) \cdot K \cdot scale}$ , когда первый этап реализован полносвязным слоем, а третий – блоками (I) или (II);
2.  $M < \frac{\sqrt{(H+1)^2 + 4(T_{base}^{(2)} - Pool_{sub} - scale \cdot K^2)} - (H+1)}{2 \cdot K \cdot scale}$ , когда первый этап реализован полносвязным слоем, а третий – блоками (III) или (IV);
3.  $scale < \frac{T_{base}^{(2)} - Pool_{sub} - (\frac{9 \cdot H \cdot (F_{scale}^{(2)} - 2)^2}{(F_{scale}^{(2)})^2} + (F_{scale}^{(2)})^2)}{((F_{scale}^{(2)})^2 + K) \cdot K}$ , когда первый этап реализован сверточным слоем, а третий – блоками (I) или (II);

$$4. \text{ scale} < \frac{T_{base}^{(2)} - Pool_{sub} - \left( \frac{9 \cdot H \cdot (F_{scale}^{(2)} - 2)^2}{(F_{scale}^{(2)})^2} + (F_{scale}^{(2)})^2 \right)}{(M \cdot (F_{scale}^{(2)})^2 + K) \cdot K}, \text{ когда первый этап ре-}$$

ализован сверточным слоем, а третий – блоками (III) или (IV).

В разделе 1.4 представлены результаты обработки FtFNN тринадцати малых датасетов. FtFNN сравнивалась с семью сверточными классификаторами, такими как EfficientNet B0, Xception и др. Предобученные и неподобученные кодировщики этих архитектур использовались в качестве вариантов реализации кодировщика FtFNN. Результаты сравнивались по максимальным и средним значениям. Для гарантии общности получаемых результатов была проведена перекрестная проверка.

FtFNN превосходит по точности все рассмотренные базовые сети. Прирост максимального значения метрики Top-1 Accuracy достигает 16.9%, а прирост ее средних значений (см. рис. 2) – 19.8%. Максимальные приросты классификационных метрик Top-3 и Top-5 Accuracy составляют 10.23% и 5.67% соответственно. Средний прирост Top-1 Accuracy, полученный при перекрестной проверке, достигает 14.4% (см. приросты для разных размеров обучающего набора на рис. 3). Показано, что согласно статистическому критерию Фридмана разница в значениях Top-1 Accuracy является статистически значимой при уровне значимости 0.01. Также установлено, что прирост точности классификации тем выше, чем меньше элементов в обучающем наборе и чем более данные неоднородны.

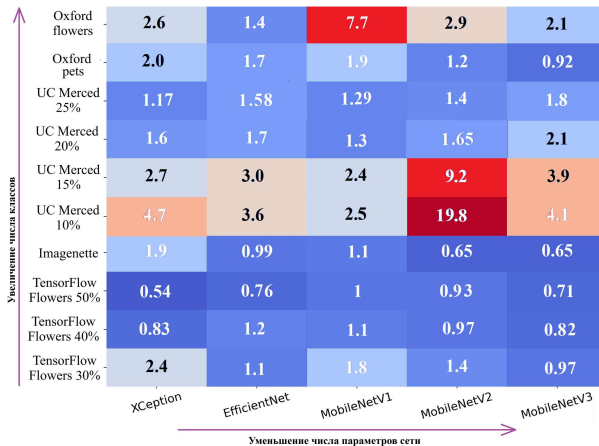


Рисунок 2 — Приросты средних значений Top-1 Accuracy, полученные FtFNN при использовании предобученных кодировщиков

Моделирование шума в FtFNN повышает точность классификации в сравнении с базовыми архитектурами: в 45 из 72 тестов при моделировании

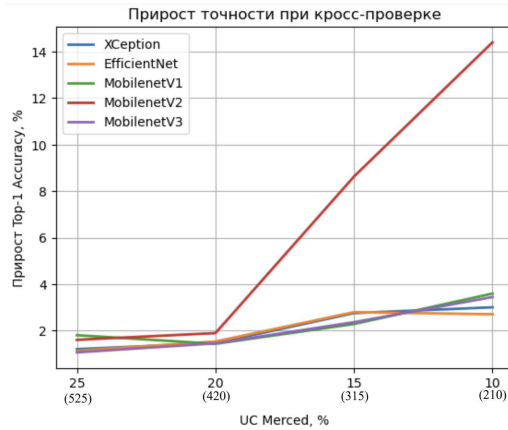


Рисунок 3 — Приросты средних значений Top-1 Accuracy, полученные FtFNN при перекрестной проверке на наборах UCM

аддитивного шума, а в 65 из 72 – импульсного. Меньшая вычислительная сложность классификатора FtFNN продемонстрирована экспериментально (см. теорему 4) – уменьшение числа выполняемых операций достигает 1.3 MFLOPS. Количество параметров сети увеличивается по сравнению с базовым классификатором только в 10 из 72 тестов, в остальных случаях снижение числа параметров достигает 496 тысяч.

**Вторая глава** посвящена нейросетевой модели сегментации неоднородных наборов изображений, информированной композицией вероятностных моделей. Поскольку рассматривалась задача сегментации, были выбраны модели, описывающие локальные характеристики снимка: конечная смесь нормальных законов и случайное поле Маркова в форме квадродерева:  $K$ -компонентная смесь ( $K$  – число выделяемых классов) моделирует яркости отдельных пикселей снимка, тогда как поле Маркова – пространственные взаимосвязи между ними. В разделе 2.1 представлена постановка задачи. В модели смеси для каждого пикселя яркостью  $X_i$  вычисляется вектор  $p^*(X_i) = \left( p_1 \varphi\left(\frac{X_i - a_1}{\sigma_1}\right), \dots, p_K \varphi\left(\frac{X_i - a_K}{\sigma_K}\right) \right)$  вероятностей соответствия каждой из  $K$  компонент смеси нормальных распределений с плотностью  $\varphi(\cdot)$  ( $a_j$  и  $\sigma_j$  – параметры сдвига и масштаба компонент смеси).

В разделе 2.2 доказывается теорема 5, демонстрирующая, что информирование смесью следует реализовать на уровне входных признаков, поскольку такой подход позволяет уменьшить ошибку восстановления целевой функции по сравнению с сетью без информирования, если целевой набор данных не однороден.

**Теорема 5.** Пусть  $f(x)$  – полностью связная однослойная нейронная сеть, использующая для восстановления дважды дифференцируемой функции

$F(x)$ , а  $h$  – среднее расстояние между нормализованными элементами обучающего набора. Тогда обогащение входных данных с помощью конечной смеси нормальных распределений, а именно представление их в виде вектора  $p^*(x)$ , уменьшает  $\mathbb{E}(\int (F(x) - f(p^*(x)))^2 dx) \leq M \cdot h^4$  – погрешность восстановления  $F(x)$  ( $M$  – константа), которая в случае наборов неоднородных данных, соответствующем  $h \geq 1$ , имеет меньший порядок в сравнении с ошибкой неинформированной полностью связной сети.

В разделе 2.3 представлено описание случайного поля Маркова в форме квадродерева, моделирующего пространственные взаимосвязи между пикселями. Квадродерево можно представить в виде пирамиды изображений разного пространственного разрешения  $S_0 \dots S_{h-1}$ , где  $S_0$  – нижний слой дерева. Каждый узел  $s \in S_l$  в пирамиде изображений и связан с одним родительским узлом  $s^-$  в предыдущем слое и с четырьмя дочерними узлами  $s^+$  в следующем слое. Также для каждого узла дополнительно определяется множество  $\{s^* \in S_l, s \in S_l, s \preceq s^*\}$  из ближайших соседних узлов в слое  $S_l$ . Таким образом, квадродерево представляет собой структуру, одновременно описывающую пространственные и иерархические взаимосвязи между элементами изображения.

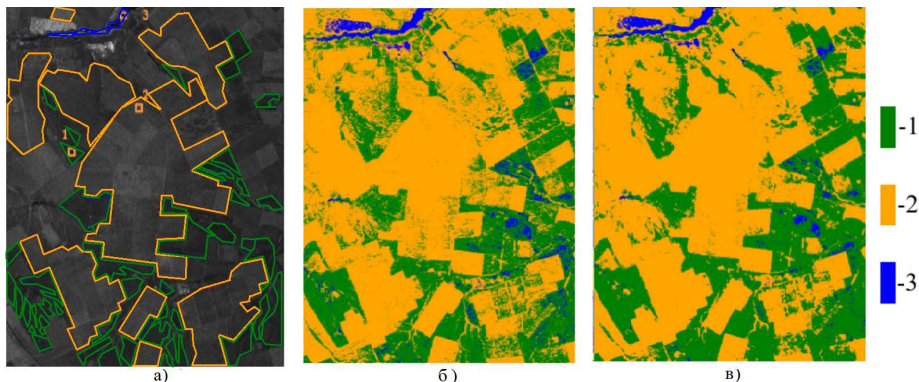


Рисунок 4 — Исходное изображение Sentinel-1 (а) и его сегментация с помощью неинформированной сети (б) и архитектурой PrINN (в). Цвет пикселей соответствует номеру класса: зеленый (1), оранжевый (2) и синий (3)

В разделе 2.3 доказываются свойства марковской модели квадродерева. Согласно теореме 6, обработка изображений с использованием квадродерева может быть интерпретирована как обработка графово-сверточной нейронной сетью. Теорема 6, а также доказанная теорема 7 об эргодичности, обосновывают выбор архитектурного типа информирования.

**Теорема 6.** *Вычисление вероятностей классов вершин квадродерева для элементов изображения  $X$  в байесовском алгоритме обхода вершин квадродерева эквивалентно применению к ним предобученной графово-сверточной нейронной сети с фиксированными весами.*

**Теорема 7.** *Поле Маркова, построенное по неориентированному графу пространственно-иерархического квадродерева  $Q_G$ , обладает свойством эргодичности.*

В разделе 2.4 представлено описание вероятностно-информированной нейросетевой архитектуры Probability Informed Neural Network (PrINN). Сеть состоит из трех блоков: блока моделирования входных изображений моделью смеси, блока сегментации базовой сверточной или трансформерной сетью, и блока пост-обработки с помощью квадродерева. В разделе 2.5 представлены результаты тестирования PrINN на семи датасетах, построенных на основе изображений, полученных радиолокаторами Sentinel-1, ESAR, Capella, и снимка из набора HRSID. Обучающие наборы сильно неоднородны, поскольку из-за зашумленности и особенностей рельефа данные характеризуются малыми межклассовыми различиями. Пример сегментации изображения Sentinel-1 приведен на рисунке 4.

PrINN существенно превосходит по точности все рассмотренные базовые сверточные и трансформерные сети без информирования: прирост точности достигает 20.31% по метрике Ассигасу и 19.24% – по  $F_1$ -мере. Согласно результатам, полученным в разделе 2.5, информирование композицией двух вероятностных моделей значительно повышает точность обработки неоднородных наборов в сравнении с информированием каждой моделью по отдельности (см. табл. 1). Так для сетей, информированных только конечной смесью (столбец «FNM + vNN»), прирост среднего значения Ассигасу относительно результатов сетей без информирования достигает 8.02%, а  $F_1$ -меры – до 18.14%. Для сетей, информированных только квадродеревом (столбец «vNN + QTree») прирост Ассигасу достигает 5.4%, а  $F_1$ -меры – до 5.89%.

**Третья глава** посвящена разработке информированной НС модели для повышения точности обработки пространственных связей между внутренними признаками снимка в нейросетевом сегментаторе в условиях сильно несбалансированных наборов. В качестве примера такой задачи рассмотрена сегментация снимков, включающих разномасштабные и мало-размерные объекты. В разделе 3.1 представлена постановка задачи.

Эргодичность поля Маркова в виде квадродерева, доказанная в теореме 7, позволяет переносить закономерности, выделенные для крупных объектов в слоях квадродерева  $S_1, \dots, S_{h-1}$  на малые объекты в слоях более высокого разрешения и наоборот. При этом, согласно теореме 6, доказанной в главе 2, информирование реализовано на уровне архитектуры сети с помощью обучаемых графово-сверточных слоев.

Таблица 1 — Приросты метрик точности сегментации (диапазоны изменения и средние/медианные значения) для PrINN, FNM + vNN и vNN+QTree

Снимок	PrINN		FNM + vNN		vNN + QTree	
	$F_1$	Accuracy	$F_1$	Acc	$F_1$	Acc
Sentinel-1 (1)	<b>19.24/8.55</b> <b>(0.26-60.65)</b>	<b>3.82/3.93</b> <b>(0.16-4.2)</b>	18.14/7.79 (0.5-56.59)	2.88/2.94 (1.54-4.15)	1.93/1.35 (0.19-6.01)	2.52/2.5 (1.08-4.01)
Sentinel-1 (2)	<b>6.27/7.29</b> <b>(0.54-16.26)</b>	<b>7.2/8.0</b> <b>(2.9-10.7)</b>	5.60/6.45 (0.0-15.25)	5.56/6.9 (0.5-9.30)	3.25/2.76 (0.6-6.68)	3.3/3.0 (1.6-5.4)
Sentinel-1 (3)	<b>13.00/9.24</b> <b>(1.34-47.3)</b>	<b>4.7/4.93</b> <b>(2.5-7.56)</b>	13.78/6.97 (0.97-47.03)	2.3/2.28 (1.91-2.75)	4.94/6.17 (1.78-10.06)	2.88/3.39 (1.4-3.65)
ESAR (1)	<b>6.75/7.23</b> <b>(3.05-10.51)</b>	<b>6.27/6.25</b> <b>(5.0-7.6)</b>	2.41/1.92 (0.46-7.68)	2.03/2.0 (0.7-3.4)	5.89/6.07 (1.73-10.79)	5.4/5.8 (3.6-7.2)
ESAR (2)	<b>12.2/9.6</b> <b>(0.45-45.6)</b>	<b>8.8/11.04</b> <b>(6.5-11.05)</b>	9.73/4.04 (0.0-34.00)	4.48/7.9 (1.07-8.01)	2.38/2.32 (0.73-4.41)	2.07/1.89 (1.53-2.53)
Capella	<b>12.19/9.12</b> <b>(4.39-28.8)</b>	<b>8.9/9.91</b> <b>(2.1-11.6)</b>	10.11/9.11 (0.28-22.90)	7.05/7.99 (7.9-11.1)	4.48/4.92 (0.20-8.66)	4.7/5.36 (0.94-7.81)
HRSID	<b>12.74/13.4</b> <b>(1.8-35.04)</b>	<b>10.48/7.14</b> <b>(3.99-20.3)</b>	9.58/10.06 (0.37-30.46)	8.02/4.52 (0.7-10.58)	5.22/6.71 (0.51-16.25)	5.22/4.26 (0.87-8.52)

В разделе 3.2 представлены свойства нейросетевой модели, информированной квадродеревом. Доказана теорема 8, согласно которой информированный квадродеревом графовый блок может обучаться быстрее, чем архитектура, выполняющая графовую свертку по двумерной решетке, традиционно применяющейся для моделирования изображений.

**Теорема 8.** Пусть  $G(\mathbf{x}) = A \cdot \mathbf{x} \cdot B$ , — однослойная линейная графовая сеть с матрицей смежности  $A \in R^{\frac{H_X \cdot W_X}{M^2} \times \frac{H_X \cdot W_X}{M^2}}$ , соответствующей двумерной пиксельной решетке, и матрицей линейного преобразования  $B$ . При этом  $G_Q(\cdot)$  — линейная графовая сеть, реализующая свертку по графу квадродерева с матрицей смежности  $A_Q$ . Тогда:

$$\Delta_t (L_t(G(\mathbf{x}^{(sp)}), Y)) < \Delta_t (L_t(G_Q(\mathbf{x}_Q^{(sp)}), Y)), \quad (2)$$

где  $L_t(\cdot, y)$  — произвольная дифференцируемая функция потерь,  $Y$  — истинные метки классов, а  $t$  — номер эпохи обучения,  $\Delta_t$  — изменение величины за одну эпоху обучения (разностная производная),  $H_X, W_X$  — высота и ширина обрабатываемого изображения, а  $\mathbf{x}^{(sp)}$  — вектор пикселей изображения, разбитый на подобласти (суперпиксели) размера  $M \times M$ .

Также в разделе 3.2 был создан новый двухветочный графовый блок для обработки снимков высокого разрешения. Одна его ветвь обрабатывает глобальные признаки входного снимка с помощью квадродерева, а

вторая – локальные взаимосвязи внутри подобласти (суперпикселя), моделируемые двумерной решеткой. Аналитически доказано (теоремы 9 и 10), что двухветочный графовый блок обучается быстрее, чем сопоставимые линейные графовые и сверточные сети, в том числе с вниманием.

**Теорема 9.** Пусть заданы двухветочная сеть  $G_Q^*(\cdot)$  с линейной обработкой глобальных признаков, сверточная  $F(\mathbf{x}) = X * w$  и линейная графовая  $G(\mathbf{x}) = A \cdot \mathbf{x}^{(sp)} \cdot B_G$ , причем  $G_Q^*(\cdot)$  и  $G(\cdot)$  обрабатывают изображение по суперпикселям размера  $M$ . Пусть  $L_t(\cdot, \mathbf{y})$  – дифференцируемая функция потерь. Тогда справедливы следующие неравенства:

$$\Delta_t (L_t(G(\mathbf{x}), \mathbf{y})) < \Delta_t (L_t(G_Q^*(\mathbf{x}), \mathbf{y})), \quad \Delta_t (L(F(\mathbf{x}), \mathbf{y})) < \Delta_t (L(G_Q^*(\mathbf{x}), \mathbf{y})),$$

где  $\mathbf{y}$  – вектор истинных значений, а  $t$  задает шаг обучения сети, то есть  $G_Q^*(\cdot)$  обучается быстрее, чем сопоставимые по размеру графовые и сверточные решения.

**Теорема 10.** Пусть заданы нейронные сети  $G_Q^*(\cdot)$ , глобальные признаки которой  $\mathbf{x}_{gRes}$  в первой ветви вычисляются с использованием механизма самовнимания, и графовая сеть с вниманием  $G_{Att}(\cdot)$ , причем обе сети обрабатывают изображение по суперпикселям размера  $M$ . Пусть  $L_t(\cdot, \mathbf{y})$  – дифференцируемая функция потерь. Тогда

$$\Delta_t (L_t(G_{Att}(\mathbf{x}), \mathbf{y})) < \Delta_t (L_t(G_Q^*(\mathbf{x}), \mathbf{y})),$$

то есть  $G_Q^*(\cdot)$  обучается быстрее, чем  $G_{Att}(\cdot)$ .

В разделе 3.3 представлена новая, построенная на основе теорем 9 и 10, информированная квадродеревом графовая архитектура с вниманием Fused-Nested Quadtree-informed Graph Self-Attention Network (FN-QiGSAN) с двухветочным объединенным графовым блоком. FN-QiGSAN также включает сверточный или трансформерный кодировщик, блок сжатия входных признаков кодировщика и блок объединения результатов двух ветвей в выходное изображение.

В разделе 3.4 представлены результаты тестирования FN-QiGSAN в задачах двухклассовой и многоклассовой сегментации четырех наборов аэрокосмических снимков высокого разрешения – HRSID, SSDD, UDD и UAVid. В качестве кодировщика рассматривались различные сверточные (DeepLabV3, ENet, PSPNet и др.) и трансформерные (SegFormer, LWGANet и др.) архитектуры. Для подтверждения общности полученных результатов выполнялась кросс-валидация.

Результаты многоклассовой сегментации показали, что FN-QiGSAN повышает точность обработки изображений в сравнении со всеми базовыми сегментаторами – пример представлен на рисунке 5. Средний прирост значений  $F_1$ -меры для крупных объектов относительно трансформеров

составляет 6.58%, а относительно сверточных сетей – 14.67%. Для малых объектов (автомобили) средний прирост  $F_1$ -меры составляет 9.36% относительно трансформеров и 11.89% – относительно сверточных сетей. FN-QiGSAN демонстрирует существенное улучшение точности обработки изображений в сравнении с State-of-the-Art (SOTA) моделью LWGANet – для отдельных классов крупных объектов прирост точности распознавания достигает 15.11%. FN-QiGSAN также демонстрирует более высокую точность сегментации в сравнении с альтернативными реализациями графового ансамбля, прирост точности у которых в среднем составляет всего 6.5% относительно базовых сегментаторов.

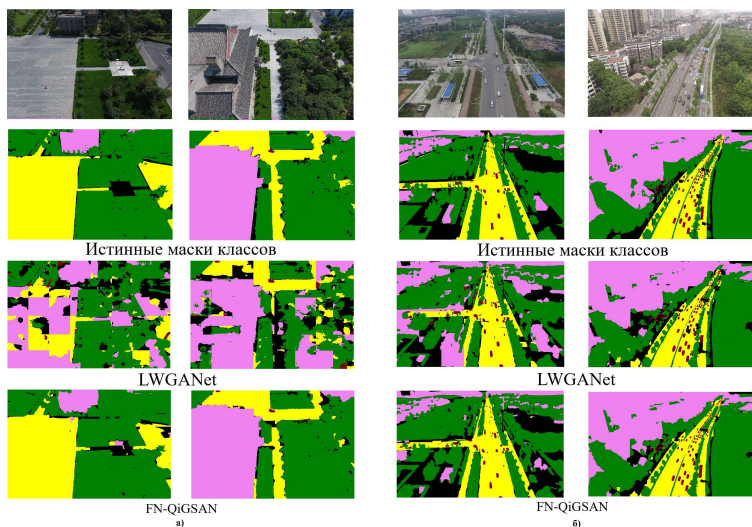


Рисунок 5 — Пример обработки наборов UDD (а) и UAVid (б). В столбцах: исходные изображения, маски классов (здания сиреневого цвета, дороги – желтого, растительность – зеленого, автомобили – бордового, и шумы – черные), результаты сегментации с помощью LWGANet и FN-QiGSAN на основе LWGANet

В задаче двухклассовой сегментации FN-QiGSAN также демонстрирует превосходящую точность выделения малоразмерных объектов в сравнении со всеми рассмотренными базовыми архитектурами. Прирост точности сегментации кораблей по  $F_1$ -мере достигает 66.24% в сравнении с трансформерными сетями, и 62.05% – в сравнении со сверточными. Прирост точности сегментации автомобилей на БПЛА изображениях достигает 32.81% по  $F_1$ -мере в сравнении с трансформерами, и 23.76% – в сравнении со сверточными архитектурами.

Лучшие по точности конфигурации FN-QiGSAN демонстрирует результаты превосходящей или сопоставимой с трансформерами (SegFormer

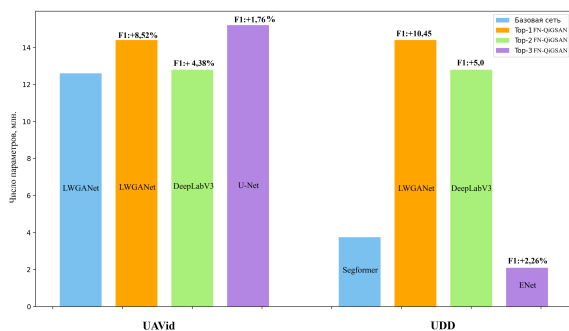


Рисунок 6 — Лучшие по метрике  $F_1$  конфигурации базовых сегментаторов (названия внутри столбцов) и три лучшие по точности конфигурации FN-QiGSAN для наборов UAVid и UDD (средний прирост метрики  $F_1$  над базовой сетью указаны вверху столбца). Названия сетей и кодировщиков в FN-QiGSAN указаны внутри столбцов

и LWGANet) точности при использовании простых сверточных кодировщиков, например U-Net, ENet или DeeplabV3, а уменьшение числа параметров сети достигает 1.78 раз при многоклассовой сегментации (см. рис. 6), и до 13.4 раз – при двухклассовой.

**Благодарности.** Автор выражает искреннюю признательность своему научному руководителю доктору физико-математических наук, доценту Андрею Константиновичу Горшенину за полезные обсуждения, ценные рекомендации и плодотворные совместные исследования.

## Публикации автора по теме диссертации

1. *Dostovalova A.* Using a Model of a Spatial-Hierarchical Quadtree with Truncated Branches to Improve the Accuracy of Image Classification // *Izvestiya, Atmospheric and Oceanic Physics*. — 2023. — Т. 59, № 10. — С. 1255–1262. — (**Q3 Scopus, Q4 WoS**).
2. *Достовалова А.* Нейросетевое квадродерево и его применение для сегментирования спутниковых изображений // *Информатика и ее применения*. — 2024. — Т. 18, № 4. — С. 77–85. — (**Q3 Scopus, K1 BAK**).
3. *Dostovalova A., Gorshenin A.* Neural Network Image Classifiers Informed by Factor Analyzers // *Doklady Mathematics*. — 2024. — Т. 110, Suppl.1. — S35–S41. — (**Q2 Scopus, Q3 WoS**).
4. *Dostovalova A., Gorshenin A.* Small sample learning based on probability-informed neural networks for SAR image segmentation // *Neural Computing and Applications*. — 2025. — Т. 37. — С. 8285–8308. — (**Q1 Scopus**).

5. *Gorshenin A., Dostovalova A.* MMRFiGN: An Ensemble Graph Segmentation Model for Imbalanced High-Resolution Images Informed by Multicomponent Markov Random Fields // *Doklady Mathematics*. — 2025. — Т. 112. — С. 308–318. — (**Q2 Scopus, Q3 WoS**).
6. *Gorshenin A., Dostovalova A.* QiGSAN: A Novel Probability-Informed Approach for Small Object Segmentation in the Case of Limited Image Datasets // *Big Data and Cognitive Computing*. — 2025. — Т. 9, № 9. — Art. No. 239. — (**Q1 Scopus, Q1 WoS**).
7. *Достовалова А., Горшенин А.* Обнаружение малых объектов на радиолокационных изображениях с использованием графово-свёрточных нейросетевых реализаций квадродеревьев // *Современные проблемы дистанционного зондирования Земли из космоса*. — 2025. — Т. 22, № 5. — С. 51–62. — (**Q3 Scopus**).
8. *Достовалова А.* [и др.]. Сравнительный анализ модификаций нейросетевых архитектур U-Net в задаче сегментации медицинских изображений // *Digital diagnostics*. — 2024. — Т. 5, № 4. — С. 833–853. — (**Q4 Scopus, K3 ВАК**).
9. *Dostovalova A., Gorshenin A.* QuadTree-Based Graph Convolutional Networks for Small Object Segmentation // *Communications in Computer and Information Science*. — 2026. — Т. 2604. — С. 31–44. — (**Q4 Scopus**).
10. *Горшенин А., Достовалова А.* О композиции графово-сверточных нейронных сетей и квадродеревьев в задаче сегментации кораблей на радиолокационных изображениях // *Ломоносовские чтения*. — 2025. — С. 93–94. — (**РИНЦ**).
11. *Горшенин А., Достовалова А.* Нейросетевые аналоги случайных полей маркова в задачах сегментации объектов на спутниковых снимках // *Тихоновские чтения*. — 2024. — С. 114. — (**РИНЦ**).
12. *Достовалова А.* Совместное применение нейронных сетей и вероятностных моделей для сегментирования радиолокационных изображений // *Материалы Международного молодежного научного форума "ЛОМОНОСОВ-2024"*. Секция "Вычислительная математика и кибернетика". — 2024. — ISBN: 978-5-901640-42-5. (**РИНЦ**).
13. *Достовалова А.* О применении глубоких гауссовских смешанных моделей в задачах классификации и регрессии временных рядов и табличных данных // *Материалы Международного молодежного научного форума "ЛОМОНОСОВ-2025"*. Секция "Вычислительная математика и кибернетика". — 2025. — ISBN: 978-5-317-07418-0. (**РИНЦ**).

*Достовалова Анастасия Михайловна*

Вероятностно-информированные нейросетевые модели анализа изображений  
при ограниченных обучающих данных

Автореф. дис. на соискание ученой степени канд. физ.-мат. наук