

На правах рукописи

Фирсов Никита Александрович

**МЕТОД И АЛГОРИТМЫ НЕЙРОСЕТЕВОЙ КЛАССИФИКАЦИИ  
ГИПЕРСПЕКТРАЛЬНЫХ ИЗОБРАЖЕНИЙ**

1.2.1. Искусственный интеллект и машинное обучение

**АВТОРЕФЕРАТ**  
диссертации на соискание учёной степени  
кандидата технических наук

Самара – 2025

Работа выполнена на кафедре киберфотоники федерального государственного автономного образовательного учреждения высшего образования «Самарский национальный исследовательский университет имени академика С.П. Королева».

**Научный руководитель:**

**Никоноров Артем Владимирович**, доктор технических наук, доцент, федерального государственного автономного образовательного учреждения высшего образования «Самарский национальный исследовательский университет имени академика С.П. Королева», профессор кафедры киберфотоники.

**Официальные оппоненты:**

**Обухова Наталья Александровна**, доктор технических наук, доцент, федеральное государственное автономное образовательное учреждение высшего образования «Санкт-Петербургский государственный электротехнический университет «ЛЭТИ» им. В.И. Ульянова (Ленина)», декан факультета радиотехники и телекоммуникаций;

**Юдин Дмитрий Александрович**, кандидат технических наук, федеральное государственное автономное образовательное учреждение высшего образования «Московский физико-технический институт (национальный исследовательский университет)», заведующий лабораторией интеллектуального транспорта.

**Ведущая организация:**

Федеральное государственное учреждение «Федеральный исследовательский центр «Информатика и управление» Российской академии наук», г. Москва.

Защита состоится «19» декабря 2025 г. в 10 часов 00 минут на заседании диссертационного совета 24.2.379.08, созданного на базе федерального государственного автономного образовательного учреждения высшего образования «Самарский национальный исследовательский университет имени академика С.П. Королева», по адресу: 443086, Россия, г. Самара, ул. Московское шоссе, д. 34.

С диссертацией можно ознакомиться в библиотеке и на сайте федерального государственного автономного образовательного учреждения высшего образования «Самарский национальный исследовательский университет имени академика С.П. Королева»: [https://ssau.ru/resources/dis\\_protection/firsov](https://ssau.ru/resources/dis_protection/firsov).

Автореферат разослан «\_\_\_\_» октября 2025 г.

Ученый секретарь диссертационного совета,  
доктор физико-математических наук, доцент

А.В. Дорошин

## Общая характеристика работы

**Актуальность.** Гиперспектральное изображение (ГСИ) с математической точки зрения представляет собой матрицу пикселей, где каждый пиксель является вектором, элементы которого определяют мощность (интенсивность) зарегистрированного излучения для конкретной длины волны (Гашников и др., 2015):

$$\bar{x}(n_1, n_2) \equiv \begin{pmatrix} x_0(n_1, n_2) \\ x_1(n_1, n_2) \\ \dots \\ x_{n-1}(n_1, n_2) \end{pmatrix}, n_1 = \overline{0, N_1 - 1}, n_2 = \overline{0, N_2 - 1}. \quad (1)$$

Для цифровых изображений она указывается в условных единицах. Учитывая наличие в формальном представлении ГСИ трёх независимых координат, где  $n_1$ ,  $n_2$  отвечают за пространственную составляющую ГСИ, а условный индекс  $n$  характеризует число спектральных каналов, ГСИ часто интерпретируют как трёхмерный массив данных и называют «гиперкубом» (Гашников и др., 2015).

Каждый материал, имеющий уникальную спектральную сигнатуру, может быть однозначно идентифицирован на ГСИ (Lu and Fei, 2014; Кринов, 1947; Klein et al., 2008). Это позволяет выполнять такие аналитические задачи, как семантическая сегментация (поэлементная классификация), кластеризация, отслеживание изменений во временных рядах и т. д. (Гашников и др., 2015; Lu and Fei, 2014; Klein et al., 2008; Барталев и др., 2023), с более высоким качеством по сравнению с обычными цветными или мультиспектральными изображениями (Барталев и др., 2023).

В 2015 году научный коллектив под руководством Н.Л. Казанского представил прототип компактной щелевой гиперспектральной аппаратуры (ГСА) на основе схемы Оффнера для размещения на борту космического аппарата (КА) (Kazanskiy et al., 2015). Она характеризовалась высоким пространственным и спектральным разрешением, оставаясь в то же время более компактной, чем аналогичные ГСА. Данная ГСА применялась не только как бортовая ГСА КА, но и в качестве наземной штативной аппаратуры в различных прикладных задачах (Podlipnov et al., 2015; Hamza et al., 2023).

В рамках настоящей работы использовались открытые наборы ГСИ, а также ГСИ высокого пространственного разрешения, полученные изображающими щелевыми ГСА, основанными на схеме Оффнера (Kazanskiy et al., 2020; Ivliev et al., 2023). Актуальность развития методов и алгоритмов анализа ГСИ, в том числе поэлементной классификации, обусловлена развитием малых космических аппаратов, оборудованных компактными ГСА (Kazanskiy et al., 2020; Ivliev et al., 2023), а также беспилотных авиационных систем (БАС), применяемых в широком спектре задач (Zhang et al., 2025, Fumin et al., 2021), и распространением изображающих спектрометров в бытовых приборах, таких как мобильные телефоны (Tominaga et al., 2021).

### Степень разработанности темы

Анализ данных ГСА включает в себя широкий спектр задач: снижение размерности ГСИ (Yu, 2005), поэлементной классификации (Zhang et al., 2025), кластеризации (Ahmed et al., 2017), оценки временных рядов (Ju et al., 2022), индексного анализа (Шовенгердт, 2010; Ahmed et al., 2017) и т.д.

В рамках настоящей работы рассматривается задача поэлементной классификации ГСИ, где каждому фрагменту изображения присваивается определенная метка класса на основе его отдельных спектральных, пространственных или пространственно-спектральных характеристик. Такого рода задачи широко востребованы в дистанционном зондировании, особенно для сельскохозяйственных задач (Wang et al., 2021), мониторинга окружающей среды (Sun et al., 2021), картографии (Liu et al., 2020) (например, определение границ водохранилищ (Liu et al., 2020) и ледников (Sood et al., 2022)) и т. д.

Поскольку ГСИ содержат как пространственную, так и подробную спектральную информацию о захваченной сцене, анализ, а особенно поэлементная классификация таких изображений, требует специальных инструментов (Nalera, 2021). Для решения задач анализа ГСИ используются классические методы, к примеру метод спектрального угла (Spectral Angle Mapper – SAM) (Chakravarty et al., 2021), спектральные индексы (NDVI, NDWI и т. д.) (Шовенгердт, 2010; Ju et al., 2022), методы машинного обучения общего назначения, такие как метод опорных векторов (Support Vector Machine – SVM) (Гашников и др., 2015; Wang et al., 2020), деревья решений (Xu et al., 2021), случайный лес (Wang et al., 2022) и т. д. Однако такие методы не учитывают пространственную информацию, содержащуюся в ГСИ, и, более того, не могут извлекать более глубокие зависимости в спектральных данных. Для более эффективного извлечения глубоких признаков из спектральных данных использовались глубокие архитектуры, основанные на свёрточных нейронных сетях (Huang et al., 2022) и сетях-трансформерах (Hong et al., 2021).

Большой вклад в решение задач поэлементной классификации ГСИ внесли такие исследователи, как В.А. Сойфер, Е.А. Лупян, Р.А. Шовенгердт, И.В. Оселедец, Е.В. Бурнаев, Н.А. Обухова, Н.И. Глумов, А.В. Кузнецов, В.В. Мясников, В.В. Сергеев, Е.В. Мясников, В.А. Фурсов, А.В. Куприянов, Н.Л. Казанский, Р.В. Скиданов, М. Хе, В. Хейдрих, Л. Сун, Ю. Чен, Ж. Лин, Г. Кэмпс-Вальс, Ф. Мелгани, Ч.-И. Чанг, Д.К. Ду, У. Стилл.

Для сканирующих ГСА на основе схемы Оффнера (Kazanskiy et al., 2020; Ivliev et al., 2023, Kazanskiy et al., 2015), описанных выше, до текущего времени решались задачи, связанные с определением спектральных индексов, в частности индекса влажности (Blank et al., 2021), задачи, связанные с калибровкой самой ГСА (Подлипов и др., 2017) и повышением качества регистрируемого ГСИ (Nikonov et al., 2018). Задачи поэлементной классификации ГСИ для данного типа ГСА не были исследованы на текущий момент. Разработке методов, алгоритмов и инструментария поэлементной классификации с учетом пространственной составляющей ГСИ, необходимых для дальнейшего прикладного использования указанного типа ГСА, и посвящена настоящая работа. Также в работе решается проблема нормализации вариаций освещения, которая, как показано в ряде работ (Mishra et al., 2020; Mishra et al., 2019; Asaari et al., 2018), является ключевой для прикладного использования ГСА при съемке вне помещения.

Для обучения классификаторов ГСИ необходима разметка (аннотирование) обучающей выборки, заключающаяся в соотношении фрагментов ГСИ со значением целевой метки класса. Проведение такой разметки вручную оператором требует существенных затрат времени, особенно в случае высокого пространственного разрешения (Puustinen et al., 2023; Yao et al., 2016). При ручной разметке также возникают затруднения визуальной идентификации принадлежности пикселя вследствие визуальных шумов. Для сканирующих гиперспектрометров могут

наблюдаться неоднородные пространственные искажения, вызванные изменением положения снимаемого объекта во время съемки, неоднородностями освещенности сцены, а также особенностями регистрирующей аппаратуры (Gaidel et al., 2023; Schlapfer et al., 2007; Wang et al., 2024). Такие искажения значительно усложняют разметку данных и увеличивают трудозатраты на ее выполнение (Jia et al., 2021).

Сложность получения и разметки данных ГСИ высокого разрешения приводит к недостатку открытых наборов данных. Такая нехватка качественных данных для обучения ограничивает практическое применение глубоких нейронных сетей и приводит к ухудшению качества классификации традиционными методами, требующими больших объемов данных для обучения (Luca et al., 2022). Таким образом, для задачи анализа ГСИ актуален поиск новых методов поэлементной классификации ГСИ, которые могут преодолеть указанные проблемы (Shenming et al., 2022), а также разработка алгоритмов повышения эффективности разметки гиперспектральных наборов данных (Paoletti et al., 2019).

Таким образом, тематика настоящей работы является актуальной и посвящена решению важных проблем в сфере поэлементной классификации и анализа ГСИ.

**Целью** является разработка и исследование нейросетевых методов и алгоритмов поэлементной классификации гиперспектральных изображений высокого разрешения, в том числе полученных ГСА, основанными на схеме Оффнера.

В соответствии с поставленной целью сформулированы следующие **основные задачи**:

- 1) Разработать алгоритм поэлементной классификации ГСИ, учитывающий пространственно-спектральные признаки в гиперспектральных изображениях высокого разрешения, полученные в условиях вариаций освещения сканирующими гиперспектрометрами на основе схемы Оффнера.
- 2) Разработать метод повышения качества поэлементной классификации гиперспектральных изображений классических свёрточных сетей и сетей-трансформеров.
- 3) Разработать алгоритм подготовки наборов данных для обучения нейросетевых классификаторов гиперспектральных изображений высокого разрешения, полученных в том числе сканирующими гиперспектрометрами на основе схемы Оффнера.

**Научная новизна** работы состоит в следующих впервые полученных результатах:

- 1) Разработана пространственно-спектральная нейросеть-трансформер DSST с блоками параллельных трёхмерных свёрток с пакетной нормализацией и основанный на ней алгоритм поэлементной классификации гиперспектральных изображений, в том числе полученных в различных условиях освещения сканирующим гиперспектрометром на основе схемы Оффнера.
- 2) Впервые предложен метод повышения точности классификации, основанный на модификации существующих глубоких полносвязных, свёрточных и трансформерных архитектур на основе сетей Колмогорова–Арнольда с PReLU активацией в блоке сквозной связи.
- 3) Разработан адаптивный вегетационный индекс ANDVI и основанный на нем алгоритм подготовки обучающих наборов данных по гиперспектральным

изображениям растительности полученным сканирующим гиперспектрометром на основе схемы Оффнера.

**Практическая значимость** работы состоит в повышении качества анализа ГСИ, получаемых как с наземных ГСА, так и с установленных на беспилотные летательные аппараты или космические аппараты, а также возможности подготовки наборов данных, содержащих ГСИ, полученные сканирующими ГСА, основанными на схеме Оффнера.

Работы по теме диссертации выполнялись в рамках: гранта Фонда Содействия Инновациям КодИТ-231563 и гранта Аналитического Центра при правительстве Российской Федерации 000000D730324P540002 № 70-2023-001317.

Результаты работы по теме диссертации запатентованы как «Программный модуль для анализа гиперспектральных изображений на базе свёрточной нейронной сети» (свидетельство №2023686406).

Результаты работы внедрены в ООО «Локус» и используются в рамках проекта «Разработка открытой платформы сбора и анализа гиперспектральных изображений для решения прикладных задач цифровой экономики».

Результаты работы внедрены в учебный процесс кафедры киберфотоники и включены в курсы «Нейронные сети», «Нейронные сети и глубокое обучение», «Технологии искусственного интеллекта».

**Методология и методы исследования.** В диссертации используются: математические методы обработки изображений, методы обучения искусственных нейронных сетей. В ходе выполнения диссертационной работы разработаны алгоритмы решения поставленных задач, а также программные средства, реализующие данные алгоритмы. Обучение нейронных сетей проводилось с использованием открытой программной библиотеки машинного обучения PyTorch.

**Соответствие паспорту специальности.** Диссертация соответствует паспорту научной специальности 1.2.1 Искусственный интеллект и машинное обучение и охватывает следующие области исследования, входящие в данную специальность:

4. Разработка методов, алгоритмов и создание систем искусственного интеллекта и машинного обучения для обработки и анализа текстов на естественном языке, для изображений, речи, биомедицины и других специальных видов данных.

13. Методы и средства формирования массивов данных и прецедентов, включая «большие данные», необходимых для решения задач искусственного интеллекта и машинного обучения. Проблемно-ориентированные коллекции данных для важных прикладных областей.

17. Исследования в области многослойных алгоритмических конструкций, в том числе – многослойных нейросетей.

**Основные положения, выносимые на защиту:**

- 1) Разработанный алгоритм поэлементной классификации ГСИ на основе пространственно-спектральной свёрточной сети NM3DCNN, компенсирующей вариации освещения за счёт применения пакетной нормализации, обеспечил повышение точности классификации ГСИ образцов почвы, полученных гиперспектрометром на основе схемы Оффнера, по содержанию минеральных веществ до 3 %, достигнув лучшей точности 95,8 %, а для ГСИ растительности – до 0,9 %, достигнув лучшей среди рассмотренных свёрточных сетей до точности 94,1%.
- 2) Разработанный алгоритм поэлементной классификации ГСИ на основе

пространственно-спектральной сети-трансформера DSST, обеспечил лучшую точность классификации ГСИ образцов почвы, полученных гиперспектрометром на основе схемы Оффнера, по содержанию минеральных веществ в 98,5 %, а для ГСИ растительности в двух экспериментах – до 94,6 %.

3) Разработанный метод повышения качества классификации ГСИ основанный на модификации существующих глубоких полносвязных, свёрточных и трансформерных архитектур на основе сетей Колмогорова-Арнольда позволил повысить качество классификации различных архитектур нейронных сетей от 0,31 до 5,05 % без изменения их структуры и достичь точности классификации на открытых наборах данных в 99,28%, а на ГСИ, полученных сканирующими гиперспектрометрами на основе схемы Оффнера, до 94,8%.

4) Разработанный алгоритм автоматизированной разметки на основе адаптивного вегетационного индекса ANDVI позволил сократить затрачиваемое время на подготовку наборов данных по ГСИ растительности, полученные сканирующими гиперспектрометрами на основе схемы Оффнера, в 16,65 раза, сохраняя точность в 98,45% относительно полностью ручной разметки.

**Достоверность полученных результатов** подтверждена посредством воспроизводимых вычислительных экспериментов на наборах ГСИ, полученных в том числе щелевым гиперспектрометром на основе схемы Оффнера, а именно нейросетевой классификации. Часть собранных наборов ГСИ и разработанных нейросетевых моделей выложена в общий доступ. Описания предложенного метода и алгоритмов нейросетевой классификации ГСИ допускают их воспроизводимость и опубликованы в ведущих научных журналах и трудах конференций по машинному зрению, обработке изображений и искусственному интеллекту.

**Апробация работы.** Результаты работы были представлены на 3 международных конференциях: VIII International conference on information technology and nanotechnology (ITNT – 2022), Samara, 2022; IX International conference on information technology and nanotechnology (ITNT – 2023), Samara, 2023; 3) IOP Conference Series: Earth and Environmental Science, 2023.

**Авторский вклад** в работу является определяющим. Все изложенные в работе и выносимые на защиту результаты получены автором лично. Постановка задач и формализация результатов проводились совместно с научным руководителем. В диссертации использованы результаты совместных исследований: формирование и подготовка наборов ГСИ выполнена А.Р. Макаровым совместно с Н.А. Ивлевым, В.В. Подлипновым; гиперспектральная аппаратура, исследованная в работе, разработана Н.Л. Казанским, Р.В. Скидановым, Н.А. Ивлевым, В.В. Подлипновым.

**Публикации.** По теме диссертации опубликовано 8 статей, 6 из которых опубликованы в изданиях, индексируемых международными базами данных Scopus и WoS (две работы в изданиях Q1); 2 – в изданиях, рекомендованных ВАК России; получено одно свидетельство о регистрации программ для ЭВМ.

**Объем и структура работы.** Диссертация состоит из введения, 4 глав, заключения и 3 приложений. Объём диссертации составляет 144 страницы, включая 89 рисунков и 37 таблиц. Список литературы содержит 141 наименование.

## Основное содержание работы

**Во введении** описано состояние рассматриваемой проблемы и обоснована актуальность темы диссертации. Сформулированы цели и задачи исследования, новизна полученных результатов и защищаемые положения.

**В первой главе** описываются ГСИ, приведены модель наблюдения ГСИ, методы формирования ГСИ, полученных сканирующей ГСА, и обзор существующих методов анализа, в том числе разметка ГСИ, описана задача поэлементной классификации ГСИ и введены критерии оценки качества классификации ГСИ.

**В разделе 1.1** приводится описание модели наблюдения ГСИ и способы их получения. Описан процесс формирования ГСИ, полученного щелевым (сканирующим) гиперспектрометром. Модель наблюдения ГСИ (1) можно представить следующим образом:

$$x_i(n_1, n_2) = \int_{\lambda_i}^{\lambda_i + \Delta\lambda} I(\lambda) R(n_1, n_2, \lambda) T(\lambda) d\lambda, \quad (2)$$

где  $\lambda_i$  – длина волны  $i$ -го регистрирующего канала;  $\Delta\lambda_i$  – ширина  $i$ -го регистрирующего канала ГСА;  $I(\lambda)$  – интенсивность падающего излучения;  $T(\lambda)$  – чувствительность датчика регистрирующей камеры ГСА (спектральный отклик сенсора и пропускающая способность оптической системы);  $R(n_1, n_2, \lambda)$  – значение отражательной способности объекта, зарегистрированной ГСА.

**В разделе 1.2** приводится описание процесса разметки (аннотирования) ГСИ для задачи классификации. Рассматриваются ручные и автоматизированные методы разметки. При решении задач поэлементной классификации ГСИ, снятых с высоким пространственным разрешением (в частности, сканирующими ГСА на основе схемы Оффнера), процедура разметки существенно усложняется, так как недостаточно выполнить разметку (аннотирование) отдельных гиперпикселей изображения, требуется фактически выполнить семантическую сегментацию ГСИ.

**В разделе 1.3** приводится описание анализа ГСИ, в частности постановка задачи поэлементной классификации. Рассмотрены методы классификации ГСИ и метрики оценки качества классификации. Задача поэлементной классификации ГСИ представляет собой соотнесение целевой метки из конечного набора классов с определенным гиперпикселем на ГСИ.

Алгоритмы глубокого обучения продемонстрировали высокую эффективность в этой задаче, поскольку такие подходы работают с высокоуровневыми признаками, что дает высокую устойчивость к нелинейностям в гиперспектральных данных (Prasad et al., 2020).

**Во второй главе** представлен алгоритм поэлементной классификации ГСИ, полученных ГСА на основе схемы Оффнера, использующий пространственно-спектральную нейросеть-трансформер. Данная архитектура нейросети учитывает пространственно-спектральные признаки за счёт применения глубокого экстрактора признаков, обладающего блоками параллельных 3D-сверток с батч-нормализацией, совместно с энкодером трансформера.

**В разделе 2.1** приводится описание существующих нейросетевых алгоритмов поэлементной классификации ГСИ. Представлены описания одномерных, двумерных и трехмерных свёрточных архитектур, а также сетей-трансформеров. Показано, что применение пространственно-спектральных сетей и

сетей-трансформеров позволило достичь высоких показателей качества классификации.

В разделе 2.2 приводится описание совместного применения пространственно-спектрального экстрактора признаков и трансформера для задачи поэлементной классификации ГСИ на примере архитектуры SSFTT (Sun et al., 2022). Алгоритм поэлементной классификации ГСИ, основанный на данной сети, показывал высокие значения метрик классификации на открытых наборах данных, достигая значения в 99 % точности в оригинальной работе.

В разделе 2.3 приводится описание развития идеи из раздела 2.2 для задачи поэлементной классификации ГСИ, полученных ГСА на основе схемы Оффнера. Сначала проводится выбор наиболее эффективного пространственно-спектрального свёрточного классификатора среди 3DCNN Luo (Luo et al., 2018), 3DCNN Li (Li et al., 2017) и 3DCNN He (He et al., 2017). Сравнение производится на ГСИ образцов почв с различным минеральным составом (рисунок 1а), полученном ГСА на основе схемы Оффнера. Разметка данного ГСИ представлена на рисунках 1б и 1в. Для сравнения были использованы также одномерные алгоритмы на основе SVM, KNN и одномерной СНС 1DCNN (Hu et al., 2015) для демонстрации важности учёта пространственных признаков.

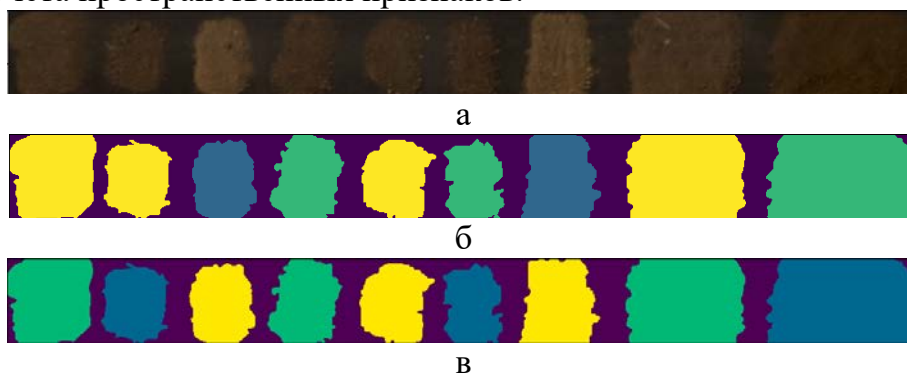


Рисунок 1 – Цветосинтезированное изображение из ГСИ образцов почв (а); Разметка ГСИ по содержанию углерода (б) и кальция (в). Желтым цветом выделено высокое содержание, зелёным – среднее, синим – низкое.

Численные результаты предсказания классификаторов представлены в таблице 1. Из них видно, что наилучшие значения точности классификации рассмотренного ГСИ достигаются 3DCNN He – 92,8 % для углерода и 91,5 % – для кальция.

Таблица 1 – Средние значения точности классификации ГСИ образцов почв

| Элементы | SVM   | KNN  | 1DCNN,<br>Hu et al. | 3DCNN,<br>Lou et al. | 3DCNN,<br>Li et al. | 3DCNN,<br>He et al. |
|----------|-------|------|---------------------|----------------------|---------------------|---------------------|
| Углерод  | 29,32 | 84,1 | 84,9                | 87,2                 | 81,6                | <b>92,8</b>         |
| Кальций  | 36,9  | 78,9 | 76,9                | 90,5                 | 86,7                | <b>91,5</b>         |

Представлено подробное описание архитектуры 3DCNN He, особенностью которой является наличие блоков параллельных свёрток с различным размером ядра. Предложена модификация данной архитектуры, называемая далее NM3DCNN, заключающаяся в добавлении слоев пакетной нормализации, позволяющих учитывать вариации освещенности на ГСИ, полученных ГСА на основе схемы Оффнера. Данная модификация показывает более высокие значения метрики качества классификации, чем исходная архитектура (таблица 2), увеличив значения точности на 2–3 %.

Представлено описание предложенной архитектуры DSST, основанной на

применении глубокого экстрактора признаков на основе экстрактора признаков NM3DCNN совместно с блоком трансформера (рисунок 2). В качестве сравнения представлены результаты классификации ГСИ почв архитектурами NM3DCNN, SSFTT и DSST, где DSST показывает лучшие результаты поэлементной классификации ГСИ высокого пространственного разрешения, полученного сканирующей ГСА на основе схемы Оффнера (таблица 2).

Таблица 2 – Средние значения точности классификации ГСИ образцов почв

| Элементы | 3DCNN | NM3DCNN | SSFTT | DSST        |
|----------|-------|---------|-------|-------------|
| Углерод  | 92,8  | 95,8    | 97,5  | <b>97,8</b> |
| Кальций  | 91,5  | 93,5    | 96,7  | <b>98,5</b> |

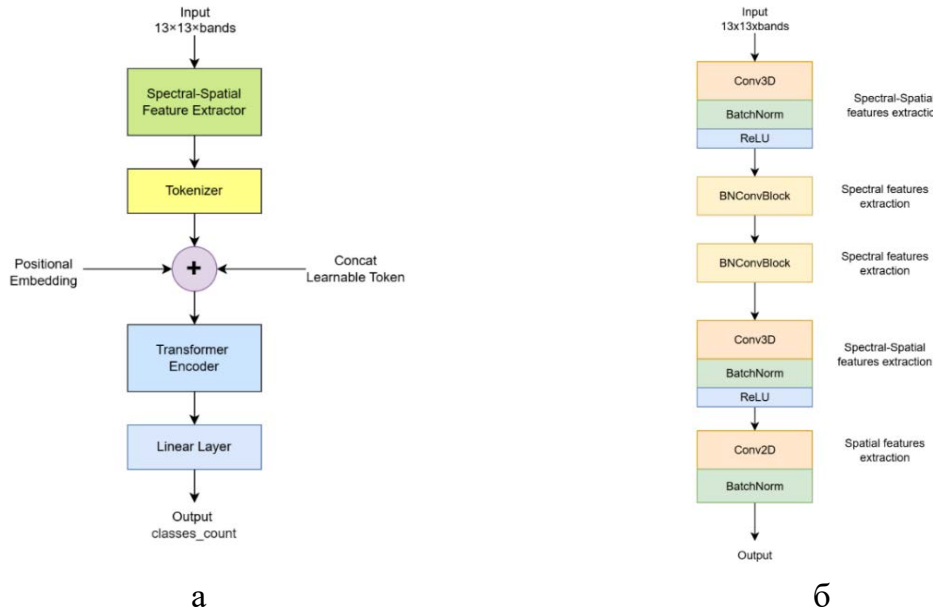


Рисунок 2 – Архитектура DSST (а); Архитектура экстрактора признаков на основе NM3DCNN (б)

Эксперименты подтвердили, что применение сетей-трансформеров совместно с глубокими пространственно-спектральными экстракторами признаков для задачи поэлементной классификации позволяет достигнуть больших значений точности, чем при использовании свёрточных сетей. Лучшие результаты показал предложенный в настоящей работе алгоритм на основе сети-трансформера DSST, достигнув показателей точности 97,8 % для углерода и 98,5 % для кальция.

**В третьей главе** представлен метод повышения качества классификации ГСИ на основе сетей Колмогорова–Арнольда (KAN – Kolmogorov-Arnold Network). Представлены модификации различных нейросетевых архитектур на основе KAN-аналогов полносвязных слоёв, слоёв свёртки, и впервые представленного в данной работе KAN-аналога механизма внимания и блока трансформера.

**В разделе 3.1** представлено описание сетей Колмогорова–Арнольда, их особенности и связь с исходной теоремой Колмогорова–Арнольда:

$$f(x) = f(x_1, \dots, x_n) = \sum_{q=1}^{2n+1} \Phi_q \left( \sum_{p=1}^n \phi_{q,p}(x_p) \right). \quad (3)$$

В исходной работе (Liu et al., 2024) предложено представить (3) в виде суперпозиции одномерных B-сплайнов.

Слой KAN размерностью  $n_{in}$  и  $n_{out}$  можно определить как матрицу одномерных функций

$$\Phi = \{\phi_{q,p}\}, p = 1, 2, \dots, n_{in}, q = 1, 2, \dots, n_{out}, \quad (4)$$

где функции  $\phi_{q,p}$  имеют обучаемые параметры.

**В разделе 3.2** описаны KAN-аналоги для полносвязных и свёрточных слоёв, а также предложенный в настоящей работе блок трансформера с KAN модификацией.

Аналог полносвязного слоя может быть представлен как

$$x_h^l = \sum_{k=1}^{H_{l-1}} \phi_{kh}^l(x_k^{l-1}), \quad (5)$$

где каждая функция активации  $\phi(u)$  представляет собой комбинацию базисной функции  $b(u)$ , являющейся аналогом сквозной связи в residual-блоке, и линейной комбинации В-сплайнов с соответствующими весами  $w_b$  и  $w_s$ :

$$\phi(u) = w_b b(u) + w_s \sum_{m=0}^{G+k-1} c_m B_m(u). \quad (6)$$

Основное отличие свёрток KAN от свёрток, используемых в CNN, заключается в ядре, где каждый элемент,  $\phi$ , является обучаемой нелинейной функцией, которая использует В-сплайны. KAN-аналоги для слоёв свёртки можно представить следующим образом:

$$1D: (K^{KAN} * T)_i = \sum_{l=-A}^A \phi_l(T_{i+l}),$$

$$2D: (K^{KAN} * T)_{i,j} = \sum_{m=-B}^B \sum_{l=-A}^A \phi_{l,m}(T_{i+l,j+m}), \quad (7)$$

$$3D: (K^{KAN} * T)_{i,j,k} = \sum_{n=-C}^C \sum_{m=-B}^B \sum_{l=-A}^A \phi_{l,m,n}(T_{i+l,j+m,k+n}), \quad (8)$$

где  $K^{KAN}$  — ядро KAN соответствующей размерности.

В данной работе предлагается заменить матрицы проекций в механизме внимания  $W_Q$ ,  $W_K$  и  $W_V$  на KAN слои (рисунок 3):

$$Q_{KAN} = \Phi_Q(X), K_{KAN} = \Phi_K(X), V_{KAN} = \Phi_V(X). \quad (9)$$

Модифицированный механизм внимания принимает вид

$$ka\_attention(Q, K, V) = \text{Softmax}\left(\frac{\Phi_Q(X)\Phi_K(X)^T}{\sqrt{d_k}}\right)\Phi_V(X) \quad (10)$$

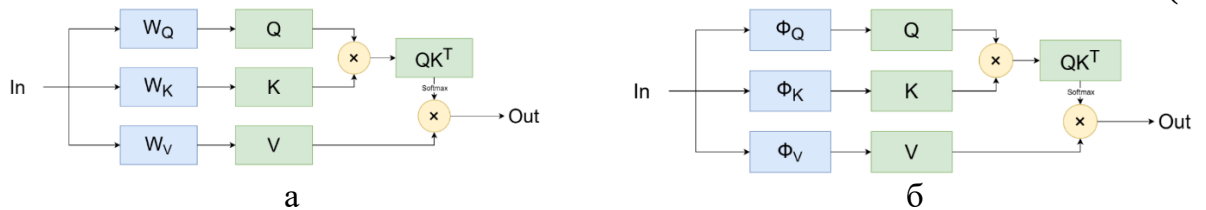


Рисунок 3 – Механизм внимания (а) и предлагаемая KAN-модификация (б)

**В разделе 3.3** представлено описание вычислительных экспериментов, в рамках которых проводились сравнения модифицированных архитектур на основе KAN с исходными. Набор данных состоял из 7 ГСИ, представленных в открытом доступе, – Pavia University, Pavia Center, Salinas, Indian Pines, Houston 13, Houston 18, KSC. Такой выбор ГСИ обусловлен их вариативностью по количеству каналов, диапазонам длин волн и разнообразию классов.

Были исследованы различные вариации MLP и KAN-сетей, в которых последние показывали уверенный прирост точности, хотя наблюдался заметный рост числа параметров сети. Замена традиционных слоев аналогами KAN оказалась результативной для свёрточных сетей и сетей-трансформеров – во всех случаях наблюдалось улучшение качества классификации, для СНС среднее улучшение составило до 5,05. Хотя сети KAN на основе трансформеров SSFTT и DSST показали более скромное среднее увеличение менее 1 %, они, тем не менее, достигли наивысшего качества классификации (таблица 3).

Таблица 3 – Точность классификации (ОА) для различных моделей нейронных сетей и их KAN-аналогов, %

| Показатели               | 1DCNN |              | 2DCNN |              | 3DCNN |              | 3DCNN |              | NM3DCNN |              | SSFTT |              | DSST  |              |
|--------------------------|-------|--------------|-------|--------------|-------|--------------|-------|--------------|---------|--------------|-------|--------------|-------|--------------|
|                          | -     | KAN          | -     | KAN          | -     | KAN          | -     | KAN          | -       | KAN          | -     | KAN          | -     | KAN          |
| Точность                 | 91,42 | <b>94,33</b> | 92,61 | <b>95,05</b> | 92,50 | <b>95,50</b> | 92,02 | <b>97,08</b> | 94,63   | <b>97,20</b> | 98,37 | <b>99,20</b> | 98,97 | <b>99,28</b> |
| Средний прирост точности | 2,92  |              | 2,44  |              | 3,00  |              | 5,05  |              | 2,57    |              | 0,83  |              | 0,31  |              |

В четвертой главе рассматривается прикладное применение разработанных алгоритма и метода для задачи классификации типов растительности на ГСИ, полученных сканирующей ГСА на основе схемы Оффнера. Также представлен алгоритм подготовки наборов данных, основанный на адаптивном вегетационном индексе ANDVI, позволяющий ускорить процесс разметки данных. Проведены вычислительные эксперименты и сравнение с аналогичными нейросетевыми классификаторами.

В разделе 4.1 представлено описание набора данных, содержащего 3 ГСИ нескольких видов растений, полученных ГСА на основе схемы Оффнера (рисунок 4).



Рисунок 4 – Цветосинтезированные изображения из ГСИ, содержащих различные растительные культуры: ГСИ 1 (а); ГСИ 2 (б); ГСИ 3 (в)

В разделе 4.2 представлено описание ручной и автоматизированной разметки описанных ранее ГСИ растительности. Получена маска разметки для ГСИ 1 при помощи ручной разметки, установлено, что такая процедура занимает около 2,5 часа реального времени.

При наличии на ГСИ пикселей, не относящихся к растениям (такowymi могут быть почва или иные объекты, попавшие на изображение), первым шагом, позволяющим ускорить процесс разметки, может быть бинарная сегментация растений от иных объектов. Для этого могут применяться спектральные индексы, как часть автоматизированной разметки, позволяющей существенно сократить время разметки ГСИ. В настоящей работе предложен алгоритм на основе адаптивного вегетационного индекса, в котором подбираются границы диапазонов красного и ближнего инфракрасного спектров. Такой индекс в точке ГСИ с координатами  $(i,j)$  можно представить как

$$ANDVI(HSI^{(i,j)}) = \frac{ANIR^{i,j} - ARED^{i,j}}{ANIR^{i,j} + ARED^{i,j}}, \quad (12)$$

где:

$$ANIR^{i,j} = \frac{1}{N_{NIR}} \sum_{k=NIR_L}^{NIR_R} C_k^{i,j}, \quad ARED^{i,j} = \frac{1}{N_{RED}} \sum_{k=RED_L}^{RED_R} C_k^{i,j}, \quad (13)$$

где  $C_k^{i,j}$  – значение пикселя в  $k$ -м канале;  $NIR_L$  и  $NIR_R$  – левый и правый индексы границ ближней инфракрасной области;  $RED_L$  и  $RED_R$  – левый и правый индексы границ красной области спектра,  $N_{NIR}$  и  $N_{RED}$  количество каналов в заданных диапазонах.

Идея адаптивного индекса заключается в подборе границ диапазонов так, чтобы максимизировать разницу между гиперпикселями двух классов «иное» и «растительность». Как критерий качества для адаптивного подбора границ спектральных диапазонов используется значение t-статистики, рассчитываемой между средними значениями яркости каждого канала гиперпикселей классов «иное» и «растительность». Для каждого диапазона  $[RED_L, RED_R]$  и  $[NIR_L, NIR_R]$  рассчитывается индексная маска, по которой формируются два набора гиперпикселей:

$$\begin{cases} X_S = \{HSI^{(i,j)}\}_{ANDVI(HSI^{(i,j)}) < tr}; \\ X_V = \{HSI^{(i,j)}\}_{ANDVI(HSI^{(i,j)}) \geq tr}, \end{cases} \quad (14)$$

где  $X_S$  и  $X_V$  – множества гиперпикселей классов «иное» и «растительность» соответственно,  $tr$  – порог бинарной обработки. По полученным наборам производится вычисление t-статистики и соответствующее ему значение p-value для каждого канала:

$$t_i = \frac{\overline{X}_{S_i} - \overline{X}_{V_i}}{\sqrt{\frac{\sigma_{S_i}^2}{n_S} + \frac{\sigma_{V_i}^2}{n_V}}}, \quad (15)$$

где  $\overline{X}_{S_i}$  и  $\overline{X}_{V_i}$  – средние значения гиперпикселей представленных классов в  $i$ -м канале;  $\sigma_{S_i}^2$  и  $\sigma_{V_i}^2$  – оценки дисперсии классов в  $i$ -м канале;  $n_S$  и  $n_V$  – объем выборок. Полученные для  $N$  каналов значения t-статистики усреднялись. Значение границ диапазонов выбиралось исходя из минимальной величины уровня значимости нулевой гипотезы (p-value).

Численное сравнение алгоритмов представлено в таблице 4. Из полученных данных видно, что адаптивный вегетационный индекс показывает наилучшие результаты бинарной сегментации классов «растительность» и «иное» в 98,45 % точности, хотя затрачивает больше всех времени на этап сегментации. Большая разница во времени работы на этапе автоматизированной разметки компенсируется минимальными правками на этапе уточнения, и наоборот. Таким образом, предложенный алгоритм подготовки наборов данных позволяет проводить точную (до 98,45 % точности) и быструю (ускорение относительно обычной ручной разметки в 16,65 раза – 9 минут против 2,5 часа) разметку.

Таблица 4 – Сравнение различных алгоритмов разметки ГСИ растений

| Показатели                                       | NDVI<br>(670, 800) | NDVI<br>(597-750, 750-980) | Kmeans<br>(2 кластера) | ANDVI         |
|--------------------------------------------------|--------------------|----------------------------|------------------------|---------------|
| Accuracy                                         | 0,8317             | 0,7629                     | 0,5287                 | <b>0,9845</b> |
| F1                                               | 0,8742             | 0,7476                     | 0,7527                 | <b>0,9866</b> |
| Время работы, с                                  | ~1                 | ~1                         | ~50                    | ~240          |
| Требуемое время ручного<br>уточнения разметки, с | ~1200              | ~2100                      | -                      | ~300          |

| Показатели                                          | NDVI<br>(670, 800) | NDVI<br>(597-750, 750-980) | Kmeans<br>(2 кластера) | ANDVI |
|-----------------------------------------------------|--------------------|----------------------------|------------------------|-------|
| Суммарно затраченное время, с                       | ~1200              | ~2100                      | -                      | ~540  |
| Ускорение относительно<br>полностью ручной разметки | 7,5                | 4,28                       | -                      | 16,65 |

Далее полученная бинарная маска может быть скорректирована для каждого отдельного класса растений (рисунок 5) путём зануления лишних пикселей, назначением ненулевым пикселям целевой метки класса и последующего их объединения между собой (рисунок 6).

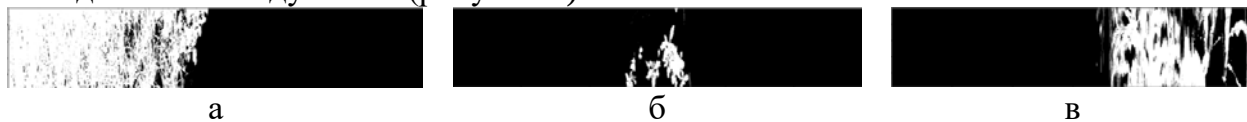


Рисунок 5 – Скорректированные маски разметки для каждого из конкретных классов на ГСИ 1: овёс (а), сорняк (б), кукуруза (в)

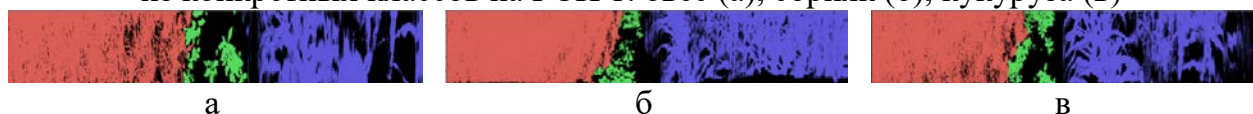


Рисунок 6 – Результирующие маски разметки ГСИ 1 (а); ГСИ 2 (б); ГСИ 3 (в), где синий цвет – овёс, зелёный – сорное растение, красный – кукуруза

**В разделе 4.3** представлено экспериментальное исследование точности классификации с учетом вариаций освещения. Для обучения классификаторов использовались фрагменты всех трёх ГСИ, описанных выше, тестирование точности проводилось на фрагментах, не попавших в обучающий набор. Такой подход к формированию обучающего набора позволяет учесть вариативность освещенности сцены как вдоль оси съемки, так и вдоль щели ГСА.

Для сравнения качества классификации были использованы пространственно-спектральные сверточные сети (3DCNN Li, 3DCNN Hamida, 3DCNN He, NM3DCNN) и сети-трансформеры (SSFTT и DSST). Наилучшие показатели качества классификации показала сеть-трансформер DSST в 94,7 %. Ее модификация при помощи замены блока трансформера на KAN-трансформер позволила увеличить значение точности классификации до 94,9 % (таблица 5). Таблица 5 – Среднее значение метрики точности (Ассигасу, %) классификации типов растительности с учётом вариаций освещения

| 3DCNN     |           |               |             | Трансформеры |      |             |
|-----------|-----------|---------------|-------------|--------------|------|-------------|
| Li et al. | He et al. | Hamida et al. | NM3DCNN     | SSFTT        | DSST | DSST KAN    |
| 91,2      | 93,7      | 92,6          | <b>94,1</b> | 94,4         | 94,7 | <b>94,9</b> |

**В разделе 4.4** представлено экспериментальное исследование точности классификации без учета вариаций освещения, в рамках которого классификатор был обучен на ГСИ с одним освещением, а предсказание выполнялось на другом, так, чтобы классификатор не мог выучить особенности освещения тестового ГСИ. Эксперимент проводился как без предобработки ГСИ, так и с приведением значений к коэффициентам спектральной яркости (КСЯ). Для сравнения качества классификации использовались классические методы машинного обучения CatBoost и RandomForest (RF), свёрточные сети CLOLN, 1DCNN, NM3DCNN и сети-трансформеры GSCViT, SSFTT, DSST. Наилучшие показатели качества классификации показала сеть-трансформер DSST в 88,3 % и 89,8 % без калибровки и с приведением к КСЯ. Ее модификация при помощи замены блока трансформера на KAN-трансформер позволила увеличить значение точности классификации до 88,4 % и 90,0 % (таблица 6).

Таблица 6 – Среднее значение метрики точности (Ассигасу, %) классификации типов растительности без учёта вариаций освещения

| CatBoost       | RF   | CLOLN | GSCViT | 1DCNN | NM3DCNN | SSFTT | DSST | DSST KAN    |
|----------------|------|-------|--------|-------|---------|-------|------|-------------|
| Без калибровки |      |       |        |       |         |       |      |             |
| 77,3           | 74,9 | 82,4  | 82,7   | 85,5  | 86,4    | 84,3  | 88,3 | <b>88,4</b> |
| КСЯ            |      |       |        |       |         |       |      |             |
| 77,3           | 75,0 | 86,9  | 84,1   | 87,1  | 87,2    | 85,4  | 89,8 | <b>90,0</b> |

### Заключение

В данном исследовании разработан алгоритм подготовки наборов данных для обучения нейросетевых классификаторов на основе адаптивного вегетационного индекса, метод повышения качества классификации ГСИ на основе сетей Колмогорова–Арнольда и алгоритм поэлементной классификации ГСИ на основе пространственно-спектральной сети-трансформера для классификации ГСИ, полученных ГСА на основе схемы Оффнера.

**Основные результаты работы** заключаются в следующем.

1. Разработанный алгоритм поэлементной классификации ГСИ на основе пространственно-спектральной свёрточной сети NM3DCNN, учитывающей вариации освещения за счёт применения пакетной нормализации, для классификации ГСИ, полученных ГСА на основе схемы Оффнера, позволил добиться повышения качества классификации на ГСИ образцов почвы по содержанию минеральных веществ на 2–3 %, достигнув лучшей точности 95,8 % для углерода и 93,5 % для кальция, а для ГСИ растительности – на 0,1–0,9 %, достигнув лучшей точности в двух экспериментах в 94,1 % и 87,2 % среди рассмотренных свёрточных сетей.

2. Разработанный алгоритм поэлементной классификации ГСИ на основе пространственно-спектральной сети-трансформера для классификации ГСИ, полученных ГСА на основе схемы Оффнера, позволил добиться лучшей точности классификации ГСИ образцов почвы по содержанию минеральных веществ в 97,8 % для углерода и 98,5 % для кальция, а для ГСИ растительности в двух экспериментах – в 94,6 % и 89,05 %

3. Разработанный метод повышения качества классификации ГСИ на основе сетей Колмогорова–Арнольда позволил повысить качество классификации различных архитектур нейронных сетей от 0,31 до 5,05 % без изменения их структуры и достичь точности классификации на открытых наборах данных в 99,28 %, а на ГСИ, полученных сканирующими ГСА на основе схемы Оффнера, до 94,8 и 89,2 %.

4. Разработанный алгоритм автоматизированной разметки на основе адаптивного вегетационного индекса позволил сократить затрачиваемое время на подготовку наборов данных в 16,65 раза с точность в 98,45 %.

5. Совокупное применение предложенных в настоящей работе разработок позволило решить задачу поэлементной классификации типов растительности и достичь показателей качества классификации в двух экспериментах до 94,8 % и 89,2 % соответственно.

### Основные публикации автора по теме диссертации

В изданиях, входящих в международные базы цитирования

**Web of Science и Scopus:**

1. **Firsov, N.** HyperKAN: Kolmogorov–Arnold Networks Make Hyperspectral Image Classifiers Smarter / **N. Firsov, N. E. Myasnikov, V. Lobanov, R. Khabibullin, N. Kazanskiy, S. Khonina, M.A. Butt, A. Nikonov** // Sensors. – 2024. – Vol. 24, no. 7683.

2. Makarov, A. Deep Spectral-Spatial Transformer for Robust Hyperspectral Image Segmentation in Varying Field Conditions / A. Makarov, I. Mirpulatov, **N. Firsov**, V. Lobanov, S. Illarionova, A. Nikonorov, V. Podlipnov, Y. Vybornova, D. Shadrin, A. Rastorguev, R. Skidanov, E. Burnaev // IEEE Access. – 2025. – Vol. 13.

3. Podlipnov V.V., **Firsov N.A.**, Ivliev N.A. etc. Spectral-spatial neural network classification of hyperspectral vegetation images // IOP Conference Series: Earth and Environmental Science. – 2023. – Vol. 1138. Issue 1. № 1.

4. Ryskova, D. Neural network classification of soils with different carbon and calcium content based on hyperspectral data / D. Ryskova, A. Nikonorov, A. Muzyka, A. Makarov, A. Pirogov, N. Ivliev, V. Podlipnov, **N. Firsov**, R. Skidanov // 2023 IX International Conference on Information Technology and Nanotechnology (ITNT), Samara, Russian Federation. – 2023. – P. 1-4.

5. Pirogov, A.V. Hyperspectral images neural network analysis of unstained micropreparations / A. Pirogov, A. Nikonorov, A. Muzyka, A. Makarov, D. Ryskova, N. Ivliev, V. Podlipnov, **N. Firsov**, P. Boriskin // 2023 IX International Conference on Information Technology and Nanotechnology (ITNT), Samara, – 2023. – P. 1-5.

6. Makarov, A.R. et al. Neural network classification of coffee varieties on hyperspectral images / A.R. Makarov, V.V. Podlipnov, N.A. Ivliev, A.V. Nikonorov, D.I. Ulyanov, **N.A. Firsov** // IEEE 2022 VIII International Conference on Information Technology and Nanotechnology (ITNT). – 2022. – P. 1-3.

#### **В изданиях, рекомендуемых ВАК России:**

7. **Фирсов, Н.А.** Ансамбли спектрально-пространственных сверточных нейросетевых моделей для задачи классификации типов почв на гиперспектральных изображениях / **Н.А. Фирсов**, В.В. Подлипнов, Н.А. Ивлиев, Д.Д. Рыськова, А.В. Пирогов, А.А. Музыка, А.Р. Макаров, В.Е. Лобанов, В.И. Платонов, А.Н. Бабичев, В.А. Монастырский, В.И. Ольгаренко, П.П. Николаев, Р.В. Скиданов, А.В. Никоноров, Н.Л. Казанский, В.А. Сойфер // Компьютерная оптика. – 2023. – Т. 47, № 5. – С. 795-805. – DOI: 10.18287/2412-6179-СО-1260.

8. **Фирсов, Н.А.** Нейросетевая классификация гиперспектральных изображений растительности с формированием обучающей выборки на основе адаптивного вегетационного индекса / **Н.А. Фирсов**, В.В. Подлипнов, Н.А. Ивлиев, П.П. Николаев, С.В. Машков, П.А. Ишкин, Р.В. Скиданов, А.В. Никоноров // Компьютерная оптика. – 2021. – Т. 45, № 6. – С. 887-896.

#### **Свидетельство о регистрации программы для ЭВМ**

9. Свидетельство о государственной регистрации программы для ЭВМ №2023686406. Программный модуль для анализа гиперспектральных изображений на базе сверточной нейронной сети / **Н.А. Фирсов**, А.Р. Макаров, В.Е. Лобанов, Л.А. Таскина, Р.М. Хабибуллин, А.В. Никоноров. – Заявка № 2023684183. Дата поступления 14 ноября 2023 г. Зарегистрировано в Реестре программ для ЭВМ 06 декабря 2023 г.