

МИНИСТЕРСТВО НАУКИ И ВЫСШЕГО ОБРАЗОВАНИЯ РОССИЙСКОЙ ФЕДЕРАЦИИ
ФЕДЕРАЛЬНОЕ ГОСУДАРСТВЕННОЕ АВТОНОМНОЕ
ОБРАЗОВАТЕЛЬНОЕ УЧРЕЖДЕНИЕ ВЫСШЕГО ОБРАЗОВАНИЯ
«САМАРСКИЙ НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ УНИВЕРСИТЕТ
ИМЕНИ АКАДЕМИКА С.П. КОРОЛЕВА»
(САМАРСКИЙ УНИВЕРСИТЕТ)

На правах рукописи

Макаров Андрей Романович

**Метод и нейросетевые алгоритмы повышения точности поэлементной
классификации гиперспектральных изображений**

1.2.1 Искусственный интеллект и машинное обучение

Диссертация на соискание учёной степени
кандидата технических наук

Научный руководитель:

Никоноров Артем Владимирович,
доктор технических наук

Самара – 2026

СОДЕРЖАНИЕ

ВВЕДЕНИЕ.....	7
Глава 1 Формирование и оценка качества гиперспектральных изображений	24
1.1 Формирование гиперспектрального изображения при сканирующей съёмке	24
1.2 Показатели качества гиперспектральных изображений	30
1.2.1 Влияние сдвига на показатели качества ГСИ и точность классификации	31
1.3 Исследование влияния сдвига на показатели качества и точность классификации ГСИ.....	37
1.3.1 Экспериментальные данные	37
1.3.2 Постановка экспериментов	39
1.3.3 Влияние искажений, вносимых в ГСИ	39
1.3.4 Протокол обучения и применения классификаторов.....	42
1.3.5 Влияние линейного сдвига на показатели качества и ОА для ГСИ NTIRE 2022	44
1.3.6 Влияние сдвига на показатели качества и ОА для ГСИ участка реки Чапаевки.....	47
1.3.7 Оценка снижения точности классификации при линейном сдвиге	52
1.4 Оценка зарегистрированных и восстановленных ГСИ по общей точности классификации.....	55
1.5 Методика оценки связи показателей классификации реальных ГСИ и ГСИ печатных сцен	61
1.5.1 Используемые данные в эксперименте и формирование ГСИ печатных сцен.....	63

1.5.2 Отбор пар соответствующих классов по изменениям показателей классификации	67
1.5.3 Связь изменений показателей классификации при добавлении искажений в ГСИ реальных и печатных сцен.....	70
1.6 Формирование калибровочной сцены.....	72
1.7 Выводы к главе 1	76
Глава 2 Выбор параметров формирования и обработки гиперспектральных изображений с учетом точности поэлементной классификации	78
2.1 Настройка спектральной предобработки ГСИ.....	80
2.1.1 Постановка задачи настройки спектральной предобработки	80
2.1.2 Программная поправка уровня чёрного и нормирование по белому эталону	82
2.1.3 Спектральное сглаживание фильтром Савицкого–Голея.....	83
2.1.4 Обучаемый модуль спектральной предобработки	83
2.1.5 Критерии настройки спектральной предобработки	84
2.2 Нейросетевая настройка спектральной предобработки	87
2.2.1 ГСИ калибровочной и тестовой сцен.....	87
2.2.2 Нейросетевые классификаторы и варианты предобработки	90
2.2.3 Протокол обучения классификаторов и настройки предобработки	91
2.2.4 Влияние длины окна фильтра Савицкого–Голея на точность классификации	98
2.2.5 Выбор веса спектральной составляющей комбинированного критерия	101
2.2.6 Результаты поэлементной классификации и оценки спектрального сходства.....	102
2.2.7 Перенос предобработки на SVM, kNN и MLP.....	110

2.2.8 Сравнение с методами DeSpaWN и DeCWN	112
2.2.9 Оценка вычислительных затрат	113
2.2.10 Обсуждение результатов спектральной предобработки.....	114
2.3 Влияние видеокодека и параметров сжатия на точность классификации.....	115
2.3.1 Условия эксперимента и параметры сжатия.....	115
2.3.2 Результаты классификации при различных параметрах сжатия ..	118
2.4 Влияние времени экспозиции и усиления на точность классификации.....	120
2.4.1 Сокращение времени экспозиции с увеличением усиления	123
2.4.2 Изменение усиления при фиксированном времени экспозиции ..	126
2.5 Обсуждение результатов	129
2.6 Выводы к главе 2.....	131
Глава 3. Обучаемая позиционно-зависимая коррекция пространственно- спектральных искажений гиперспектральных изображений	134
3.1 Пространственно-спектральные искажения ГСИ и методы коррекции.....	134
3.1.1 Пространственно-спектральное рассогласование и неоднородность чувствительности	134
3.1.2 Калибровка по объектам с известными спектральными характеристиками.....	139
3.1.3 Условия регистрации и переносимость калибровочных коэффициентов.....	140
3.1.4 Нейросетевые методы коррекции ГСИ.....	141
3.2 Модель и постановки задачи позиционно-зависимой коррекции	142

3.2.1 Модель неоднородности чувствительности и функция коррекции	142
3.2.2 Коррекция по опорным спектрам классов калибровочной сцены	145
3.2.3 Межприборная коррекция по средним спектрам классов	148
3.3 Позиционно-зависимая коррекция ГСИ с обучаемыми калибровочными слоями.....	150
3.3.1 Калибровка по белому объекту и аппроксимация коэффициентов	150
3.3.2 Обучаемые калибровочные слои и их параметризации.....	150
3.3.3 Обучаемая калибровочная матрица	151
3.3.4 Обучаемая полиномиальная параметризация калибровочной матрицы.....	152
3.3.5 Векторная факторизация калибровочной матрицы.....	153
3.3.6 Определение коэффициентов методом гребневой регрессии	154
3.3.7 Фиксация и перенос коэффициентов калибровочного слоя.....	155
3.3.8 Нейросетевая настройка калибровочных слоёв.....	156
3.4 Коррекции позиционной неоднородности ГСИ	156
3.4.1 Калибровочное и тестовое ГСИ	157
3.4.2 Формирование опорных спектров классов	158
3.4.3 Протокол настройки и варианты коррекции.....	158
3.4.4 Оценка позиционной неоднородности сигнала	160
3.4.5 Результаты поэлементной классификации.....	163
3.4.6 Обсуждение результатов коррекции позиционной неоднородности в первой постановке эксперимента	164
3.5 Коррекция ГСИ при межприборном сопоставлении	164

3.5.1 ГСИ разработанного гиперспектрометра и Specim IQ.....	165
3.5.2 Фиксированная спектральная предобработка и расчёт КСЯ	165
3.5.3 Настройка коррекции по калибровочным сценам.....	166
3.5.4 Методика оценки спектрального сходства и точности классификации	167
3.5.5 Результаты сопоставления спектров и поэлементной классификации	170
3.5.6 Обсуждение результатов коррекции при межприборном сопоставлении	173
3.6 Проверка метода коррекции на ГСИ Specim FX	174
3.6.1 ГСИ Specim FX и подготовка выборок.....	174
3.6.2 Настройка калибровочного слоя для Specim FX	175
3.6.3 Методика и результаты классификации тестовых ГСИ	177
3.7 Дополнительная демонстрация нейросетевой калибровки	182
3.8 Выводы к главе 3	185
ЗАКЛЮЧЕНИЕ	188
СПИСОК ИСПОЛЬЗОВАННЫХ ИСТОЧНИКОВ	192
ПРИЛОЖЕНИЕ А	212
ПРИЛОЖЕНИЕ Б.....	213

ВВЕДЕНИЕ

Актуальность темы диссертации

Гиперспектральное изображение (ГСИ) содержит значения зарегистрированного сигнала в узких спектральных поддиапазонах. Каждой пространственной позиции соответствует *спектральный вектор* – последовательность значений по спектральным каналам. Такой вектор далее называется *гиперпикселом*. При *поэлементной классификации* каждому гиперпикселу присваивается метка одного из заданных классов [1].

Гиперспектральная съёмка применяется в дистанционном зондировании Земли [2], сельском хозяйстве [3], экологическом мониторинге [4] и анализе материалов [2]. Регистрация сигнала в узких спектральных поддиапазонах позволяет выявлять различия между объектами, которые могут быть неразличимы на трёхканальных изображениях красного, зелёного и синего диапазонов (RGB) [3].

Для получения ГСИ используется *гиперспектральная аппаратура* (ГСА). В *сканирующей гиперспектральной аппаратуре* (СГСА) пространственные или спектральные отсчёты регистрируются последовательно [4]. В процессе сканирования при помощи СГСА каждый кадр *гиперспектральной видеоинформации* (ГВИ) или отдельные изображения представляют собой *пространственную спектрограмму* (после сборки соответствует строке ГСИ): одна координата соответствует положению вдоль щели, другая – спектральному каналу. При относительном перемещении аппаратуры и сцены последовательность таких кадров образует ГСИ [1].

В работе [1] рассмотрено представление ГСИ в виде матрицы гиперпикселов, каждый из которых является спектральным вектором. В работах [1, 5] используется представление ГСИ в виде тензора третьего порядка с двумя пространственными и одной спектральной осями. В настоящей диссертации ГСИ представляется в виде тензора третьего порядка – трёхмерного массива с двумя пространственными и одной спектральной

осями [6]. Такое представление удобно для описания пространственных спектрограмм и задания калибровочных коэффициентов, зависящих от положения вдоль щели и спектрального канала:

$$S(i, j, k): 1 \leq i \leq I, 1 \leq j \leq J, 1 \leq k \leq K,$$

где $S(i, j, k)$ – значение элемента ГСИ; i – пространственный отсчёт вдоль направления сканирования, соответствующий номеру строки ГСИ (пространственной спектрограмме); j – пространственный отсчёт вдоль изображения щелевой диафрагмы; k – спектральный отсчёт, соответствующий номеру спектрального канала; I, J и K – размеры тензора по этим осям. Спектральный вектор в пространственной позиции (i, j) образован значениями $S(i, j, k)$ по всем каналам k . В зависимости от этапа обработки элементы ГСИ содержат *цифровые отсчёты зарегистрированного сигнала* (DN) или значения, полученные после радиометрического преобразования [7]. Представление ГСИ показано на рисунке Введение.1.

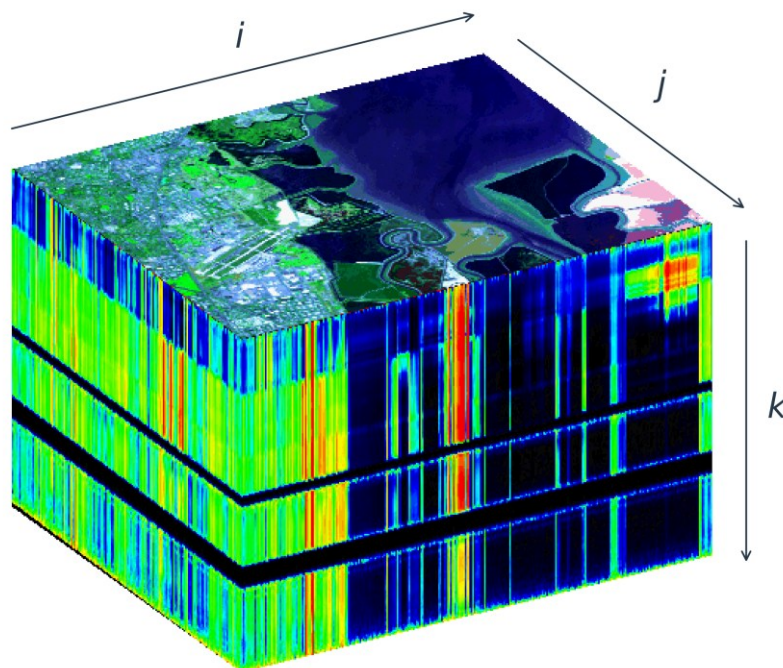


Рисунок Введение.1 – Представление ГСИ с пространственными отсчётами i, j и спектральными отсчётами k

По способу регистрации ГСИ различают *точечное пространственное сканирование* (whiskbroom) [4], *линейное пространственное сканирование*

(pushbroom) [4], *спектральное сканирование* [3] и *полнокадровую съёмку без сканирования* (snapshot) [3] (рисунок Введение.2 [8]). При точечном сканировании последовательно регистрируются спектры отдельных пространственных позиций, при линейном – спектры точек одной строки. При спектральном сканировании последовательно регистрируются двумерные изображения сцены в разных спектральных каналах [3].

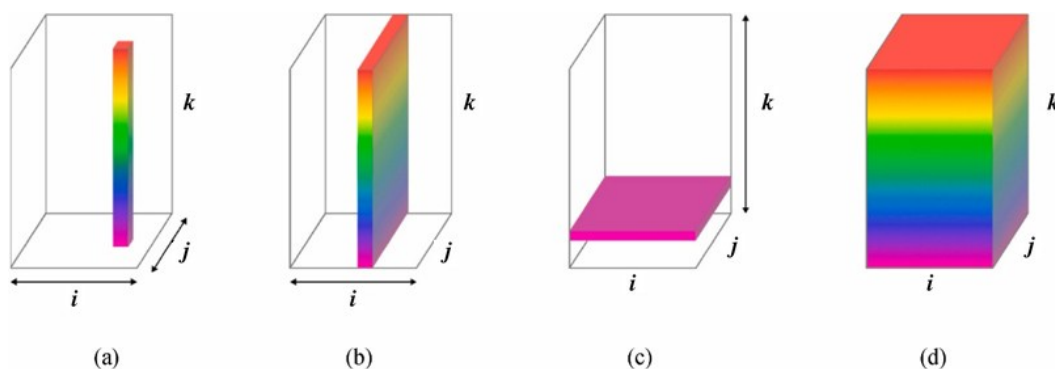


Рисунок Введение.2 – Способы регистрации ГСИ: (а) точечное сканирование; (б) линейное сканирование; (с) спектральное сканирование; (d) полнокадровая съёмка без сканирования

В работе используются ГСИ, полученные щелевыми СГСА на основе схемы Оффнера, разработанными в Самарском университете (далее – *разработанные гиперспектрометры*) [1], и коммерческой СГСА Specim IQ [9] и Specim FX [1]. Экспериментальные исследования также включают открытые наборы ГСИ [10] и ГСИ, реконструированные из RGB-изображений [11].

При линейном сканировании строки ГСИ регистрируются последовательно, поэтому изменение положения и ориентации аппаратуры во время съёмки влияет на их взаимное расположение. Вибрации и колебания ориентации аппаратуры (*джиттер*) могут приводить к геометрическим искажениям и размытию сканирующих изображений [12]. При формировании ГСИ необходимо учитывать смещения и относительные повороты проекций щели на поверхности сцены [13, 14]. Геометрические искажения оптической системы проявляются также в эффектах *smile* и *keystone*: *smile* соответствует изменению центральной длины волны спектрального канала вдоль щели [15], а *keystone* – зависимости положения изображения пространственной точки от

спектрального канала [16]. Неоднородность отклика фотоприёмника и тёмного сигнала вызывает радиометрические искажения, в том числе полосовые артефакты [17, 18].

При оценке качества ГСИ рассматриваются характеристики регистрирующей системы [4] и метрики качества, основанные на сопоставлении с эталоном. К характеристикам системы относятся уровень шума, динамический диапазон, спектральное и пространственное разрешение, точность калибровки и величина межканального рассогласования [19, 20]. Для сопоставления ГСИ с эталоном используются *средний квадрат ошибки* (MSE) [21], *среднеквадратичное отклонение* (RMSE), равное квадратному корню из MSE [21], *пиковое отношение сигнал/шум* (PSNR) [21], *индекс структурного сходства* (SSIM) [22] и *спектральный угол* (SAM) [23].

Классификация характеристик качества ГСИ приведена в ГОСТ Р 59079–2020 [24], а состав показателей качества ГСИ полученных СГСА – в ГОСТ Р 59475–2021 [25]. Для решения конкретной задачи поэлементной классификации эту оценку необходимо дополнять показателями правильности присвоения меток классов.

PSNR, SSIM и SAM вычисляются по значениям элементов сравниваемых ГСИ и не используют метки классов. *Общая точность поэлементной классификации* (Overall Accuracy, OA) определяется как доля гиперпикселей, для которых предсказанная метка совпадает с эталонной на заданной области оценивания [26]. Поэтому связь метрик качества, основанных на сопоставлении с эталоном, с точностью классификации необходимо исследовать для конкретной целевой задачи. Такой подход к оценке качества спектральных изображений рассматривается в работах по прогнозированию результатов классификации [27, 28] и обнаружения объектов [29].

В настоящей работе применение общей точности поэлементной классификации (OA) для оценки и выбора параметров съёмки, предобработки и калибровки по ГСИ печатной калибровочной сцены обозначается как

точность классификации целевой задачи (ТКЦЗ). Такая оценка проводится для заданного состава классов при фиксированном протоколе обучения и проверки классификатора.

Для отдельного класса *полнота* (recall) представляет собой долю правильных решений среди его эталонных гиперпикселей, а *точность* (precision) – среди гиперпикселей, отнесённых к нему классификатором. *F1-мера класса* – гармоническое среднее точности и полноты. *Взвешенная F1-мера* рассчитывается усреднением F1-мер классов с весами, пропорциональными числу их эталонных тестовых гиперпикселей [30].

Время экспозиции [7] и усиление [31] влияют на зарегистрированный сигнал, а спектральная предобработка – на спектры, поступающие на вход классификатора. В работах [32, 33] исследовано влияние спектрального сглаживания и вычисления производных спектров на показатели классификации ГСИ. Эти результаты обосновывают выбор параметров предобработки с учётом целевой задачи. В настоящей работе исследуется совместная настройка спектрального сглаживания и программной поправки уровня чёрного по калибровочной сцене.

Нормирование сигнала объекта по белому эталону после вычитания тёмного сигнала позволяет оценивать *коэффициент спектральной яркости* (КСЯ) относительно эталона [7]. При таком преобразовании необходимо учитывать условия освещения, время экспозиции и характеристики отклика прибора; образец и эталон должны регистрироваться в сопоставимых условиях [7].

Характеристики ГСА и метрики качества, основанные на сопоставлении с эталоном, описывают условия регистрации и различия между ГСИ. Оценка пригодности ГСИ для поэлементной классификации дополнительно требует учёта заданных классов и используемого классификатора [27, 28].

Это определяет необходимость исследования связи показателей качества ГСИ с точностью поэлементной классификации. Результаты такого исследования могут использоваться при сравнении ГСИ и выборе параметров

их получения и обработки. Для повышения точности классификации требуется разработка методов предобработки и коррекции отклика СГСА, настраиваемых с учётом целевой задачи.

Степень разработанности темы

Развитие гиперспектрального дистанционного зондирования связано с методами измерения спектральных характеристик природных объектов и созданием изображающих спектрометров. В работе [34] систематизированы данные о спектральной отражательной способности природных образований. Методы получения, преобразования и анализа многоспектральных и гиперспектральных изображений обобщены в работе [35].

В отечественных исследованиях развиваются компактные щелевые гиперспектрометры на основе схемы Оффнера. В работах [36–38] рассмотрены моделирование такой аппаратуры и формирование изображения в плоскости фотоприёмника. ГСИ, полученные приборами на основе этой схемы, используются в экспериментальной части настоящей работы. Наряду с созданием аппаратуры исследуются методы тематической классификации ГСИ. В работе [39] предложен подход, основанный на оценке близости спектрального вектора к подпространству класса.

В работе [40], выполненной с участием автора диссертации, предложена ансамблевая архитектура на основе пространственно-спектральных свёрточных нейронных сетей для классификации образцов почв по содержанию углерода и кальция.

Спектральная калибровка устанавливает соответствие элементов фотоприёмника длинам волн, а *радиометрическая калибровка* определяет преобразование цифрового сигнала в значения измеряемой величины. В работе [41] эти процедуры рассмотрены для AVIRIS-NG, включая коррекцию дефектных элементов и неоднородности отклика. В работе [42] описано определение характеристик гиперспектральной системы на основе спектрографа.

Для сканирующих щелевых систем разработаны методы коррекции систематических полосовых артефактов. В работе [17] такая коррекция рассмотрена для ГСИ CHRIS/PROBA. В работе [18] исследовано уменьшение радиометрических ошибок, обусловленных различиями характеристик элементов детектора. Эти работы показывают необходимость учитывать неоднородность отклика при преобразовании зарегистрированного сигнала.

В работах [15, 16] предложены методы обнаружения и коррекции спектрального и пространственного рассогласования и исследована предобработка с учётом smile и keystone. В работе [20] рассмотрены лабораторное определение пространственного рассогласования каналов и его влияние на реальные изображения. В работе [43] калибровка и улучшение ГСИ выполняются двумерной деконволюцией с оценкой изменяющейся функции отклика.

Для наземной съёмки используются методы калибровки по опорным объектам. В работе [7] исследована калибровка гиперспектральной системы с помощью доступного белого эталона. В работе [44] рассмотрены определение геометрических и спектральных характеристик пикселей и радиометрическая калибровка, включая нелинейность отклика. Поэтому преобразование сигнала по одному белому уровню следует применять с учётом характеристик конкретного прибора.

В задачах восстановления ГСИ по RGB-изображениям и повышения пространственного разрешения результаты обычно сравниваются с эталонными ГСИ по нескольким показателям. В работе [45] используются PSNR, SSIM и SAM при оценке суперразрешения, а в работе [46] – показатели ошибок восстановления и спектрального соответствия при реконструкции методом многоэтапного спектрального трансформера MST++. Такое сравнение позволяет оценить расхождения значений и спектров, но не заменяет проверку результата классификации.

Для поэлементной классификации ГСИ применяются спектральные и пространственно-спектральные модели. *Одномерная свёрточная нейронная*

сеть 1DCNN обрабатывает спектр отдельного гиперпиксела [47]. Пространственно-спектральные модели используют спектр гиперпиксела и его пространственное окружение для предсказания метки этого гиперпиксела [48, 49]. В главах 2 и 3 используются 1DCNN, *трёхмерная свёрточная сеть* NM3DCNN и *пространственно-спектральный трансформер* SSFTT (Spectral–Spatial Feature Tokenization Transformer). На вход NM3DCNN и SSFTT подаётся *подтензор* – фрагмент тензора ГСИ, содержащий все спектральные каналы классифицируемого гиперпиксела и его пространственного окружения [1]. Таким образом, исследования проводятся с представителями одномерных и трёхмерных свёрточных сетей, а также трансформерных архитектур.

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. Сергеев, В.В. Мясников, 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.

Влияние обработки на точность классификации исследуется отдельно от сопоставления ГСИ с эталоном. В работе [50] рассмотрены методы шумоподавления ГСИ и их применение к последующему анализу. В работе [51] исследовано влияние деградации изображения на классификацию, а в работе [52] – влияние пространственного разрешения при классификации растительности. Эти результаты обосновывают включение точности классификации в экспериментальную оценку предобработки. В качестве базового способа спектрального сглаживания в главе 2 используется *фильтр Савицкого–Голея* [53].

Проводились исследования оценки спектральных изображений с учётом целевой задачи классификации. В работах [27, 28] исследованы показатели спектрального качества применительно к классификации земной поверхности. В работе [29] предложен метод оценки пригодности спектрального

изображения для обнаружения субпиксельных целей. В настоящей диссертации исследуется применение общей точности классификации для сравнительной оценки ГСИ при пространственном рассогласовании и фиксированных классах и классификаторе.

Пространственное рассогласование влияет и на показатели классификации. В работе [54] показано изменение общей точности при попиксельном сравнении с пространственно сдвинутой опорной сеткой. В работах [55, 56] исследовано влияние позиционной точности на оценку классификации по одному пикселю и по блоку пикселей. При пространственном рассогласовании необходимо отдельно исследовать изменение значений показателя классификации и сохранение ранжирования сравниваемых вариантов ГСИ.

Обучаемые нейросетевые методы также применяются на этапе калибровки. В работе [57] предложен нейросетевой метод оценки отражательной способности при пространственно однородном освещении. В опубликованной по теме диссертации работе [1] представлен нейросетевой метод коррекции пространственно-спектральной чувствительности щелевого гиперспектрометра. Метод учитывает зависимость отклика прибора от положения пространственного отсчёта вдоль щели.

Таким образом, актуальность настоящей работы обусловлена необходимостью оценки качества ГСИ с учётом целевой задачи поэлементной классификации и разработки методов предобработки и коррекции, направленных на повышение точности её решения.

Целью работы является повышение точности поэлементной классификации гиперспектральных изображений за счёт выбора параметров их формирования, предобработки и коррекции.

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

1) Разработать методику оценки качества ГСИ, полученных сканирующей ГСА, согласующуюся с точностью поэлементной классификации гиперспектрального изображения.

2) На основе методики оценки качества ГСИ разработать алгоритмы формирования и предобработки ГСИ, в том числе нейросетевые, обеспечивающие повышение точности поэлементной классификации.

3) Разработать алгоритмы коррекции пространственно-спектральных искажений ГСИ, в том числе нейросетевые, ориентированные на повышение точности решения задачи поэлементной классификации.

4) Провести экспериментальную проверку разработанных методов и алгоритмов на гиперспектральных данных, полученных различной гиперспектральной аппаратурой, а также на открытых наборах ГСИ.

Научная новизна работы заключается в следующем:

- 1) Разработана методика формирования калибровочной сцены, позволяющая определять параметры формирования, предобработки и калибровки ГСИ для повышения точности решения задачи поэлементной классификации.
- 2) Разработаны нейросетевые алгоритмы оптимизации параметров формирования и предобработки ГСИ, таких как: уровня черного, компенсации темновых токов, усиления камеры, алгоритмов сглаживания и сжатия по критерию точности классификации целевой задачи (ТКЦЗ).
- 3) Разработаны метод и алгоритмы обучаемой позиционно-зависимой коррекции пространственно-спектральных искажений ГСИ по критерию ТКЦЗ в виде переносимых калибровочных слоёв. Экспериментально подтверждено, что использование калибровочных слоев обеспечивает повышение точности поэлементной классификации по сравнению с предварительной калибровкой по белому эталону.

Практическая значимость

Предложенная методика описывает формирование калибровочной сцены, на основе которой подбираются параметры съёмки, предобработки и калибровки гиперспектрометров по критерию точности классификации

целевой задачи (ТКЦЗ). Для отобранных спектрально близких классов калибровочной сцены и реальных ГСИ установлена согласованность изменений показателей классификации.

Разработан модуль спектральной предобработки, объединяющий спектральную свёртку, вычитание программной поправки уровня чёрного и нормирование по белому эталону. Коэффициенты модуля определяются по калибровочным ГСИ и используются с отдельно обученными классификаторами. Результаты исследования влияния видеокодеков, параметров сжатия, времени экспозиции и усиления на точность классификации позволяют обосновать выбор режимов съёмки, формирования и сохранения ГСИ.

Калибровочные слои позволяют корректировать отклик гиперспектрометра с учётом положения пространственного отсчёта вдоль щели. Дообучение коэффициентов по ошибке классификации позволяет учитывать целевую задачу при настройке коррекции. Применимость метода исследована для разработанного гиперспектрометра в Самарском университете основанного на схеме Оффнера и гиперспектрометра Specim FX.

Результаты работы внедрены в Центр «Интеллектуальная мобильность многофункциональных беспилотных авиационных систем» и использовались в рамках проекта «Разработка фреймворка дистанционного мониторинга на основе интеллектуального анализа мультимодальных данных, полученных на борту беспилотных авиационных систем FW_Monitor»» (Приложение А).

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» (Приложение Б).

Методология и методы исследования. Используются методы математического моделирования формирования ГСИ, корреляционного анализа, обработки изображений и спектральных данных, спектрального сглаживания, гребневой регрессии и обучения нейронных сетей.

Экспериментальная оценка выполняется по метрикам качества, основанным на сопоставлении с эталоном, общей точности поэлементной классификации, полноте и F1-мере класса, а также взвешенной F1-мере.

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) Общая точность поэлементной классификации (ОА) в рассмотренных экспериментах является более устойчивой к малым пространственным рассогласованиям сравниваемых ГСИ, чем метрики сопоставления с эталоном PSNR, SSIM и SAM. Применение ОА в разработанной методике выбора параметров по ГСИ печатной калибровочной сцены обозначается как ТКЦЗ.

2) Разработанный нейросетевой алгоритм спектрального сглаживания и коррекции уровня чёрного в ГСИ позволил повысить точность поэлементной классификации. Для нейросетевых вариантов с наибольшей средней ОА в каждом из пяти рассмотренных сочетаний классификаторов прирост ОА относительно варианта без дополнительной предобработки составил от 2,17 до 6,22 п.п., прирост взвешенной F1-меры – от 2,03 до 7,68 п.п., а прирост PSNR между средними спектральными векторами классов – от

0,268 до 0,695 дБ. По сравнению с предобработкой на основе фильтра Савицкого–Голея соответствующие приросты составили от 1,25 до 5,19 п.п., от 1,44 до 5,53 п.п. и от 0,187 до 0,614 дБ.

3) Разработанные алгоритмы обучаемой позиционно-зависимой коррекции отклика гиперспектрометра в виде переносимых калибровочных слоёв позволили повысить точность поэлементной классификации ГСИ. В первой, внутриприборной постановке на данных разработанного гиперспектрометра для нейросетевых конфигураций с наибольшей средней ОА у каждого классификатора прирост ОА относительно данных без калибровки составил от 0,65 до 8,83 п.п., а прирост взвешенной F1-меры – от 1,15 до 10,43 п.п. По сравнению с калибровкой по белому эталону соответствующие приросты составили от 0,20 до 1,57 п.п. и от 0,51 до 1,56 п.п. Среди результатов двух постановок на данных разработанного гиперспектрометра наибольшая средняя ОА составила 96,78 % для 1DCNN, 98,73 % для NM3DCNN и 98,51 % для SSFTT.

Достоверность полученных результатов подтверждена посредством воспроизводимых вычислительных экспериментов на наборах гиперспектральных изображений, полученных в том числе щелевой ГСА на основе схемы Оффнера, а именно нейросетевой поэлементной классификации. Часть собранных наборов гиперспектральных изображений и разработанных нейросетевых моделей выложена в общий доступ. Описания предложенных методов нейросетевой поэлементной классификации гиперспектральных изображений допускают их воспроизводимость и опубликованы в ведущих научных журналах и трудах конференций по машинному зрению и обработке изображений.

Апробация работы. Результаты работы были представлены на 3 международных конференциях:

1) 35th International Joint Conference on Artificial Intelligence (IJCAI–ECAI 2026), Bremen, Germany, 2026.

2) X International conference on information technology and nanotechnology (ITNT – 2024), Samara, 2024.

3) «19th International Conference on Machine Vision (ICMV 2026), Budapest, Hungary, 2026». Две научные статьи по тематике диссертации приняты в печать 17.08.2026. Дата проведения конференции: с 15 по 18 октября 2026 года.

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 проводились совместно с Н.А. Ивлиевым, В.В. Подлипновым, Д.М. Владимировым и К.Э. Савельевым. Исследованная гиперспектральная аппаратура разработана Н.Л. Казанским, Р.В. Скидановым, Н.А. Ивлиевым и В.В. Подлипновым.

Публикации. Основные результаты по теме диссертации изложены в 5 печатных работах, в том числе в 3 статьях в журналах, индексируемых Scopus и Web of Science, 2 статьях в журналах, рекомендованных ВАК.

Объём и структура работы. Диссертация состоит из введения, трёх глав, заключения. Объём диссертации составляет 212 страниц, включая 70 рисунков и 27 таблиц. Список литературы содержит 144 наименования.

В работе предлагается метод повышения точности поэлементной классификации ГСИ, основанный на выборе параметров их получения, предобработки и коррекции по результатам классификации ГСИ печатной калибровочной сцены.

В первой главе для реализации предложенного метода разрабатывается методика формирования и применения печатной калибровочной сцены, предусматривающая выбор параметров получения и обработки ГСИ по точности поэлементной классификации.

В экспериментах первой главы показано, что при малом пространственном рассогласовании с эталоном показатели PSNR, SSIM и SAM могут указывать на более высокое качество ГСИ с высоким уровнем внесённых искажений по сравнению с ГСИ с средним уровнем искажений. При тех же пространственных смещениях общая точность поэлементной классификации (ОА) для исходного ГСИ остаётся выше, чем для его вариантов с внесёнными искажениями. Поэтому в предлагаемой методике выбор параметров съёмки и обработки ГСИ выполняется по ТКЦЗ, оцениваемой метрикой ОА на калибровочной сцене.

Для обоснования возможности настройки параметров съёмки и предобработки по ГСИ печатной калибровочной сцены необходимо показать, что при изменении уровня шума показатели классификации классов реальных ГСИ и их спектрально близких печатных представлений изменяются согласованно. Для этого сначала проводится вычислительный эксперимент на данных NTIRE с использованием реальных спектров и спектров печатных представлений тех же классов, полученных моделированием печати. Добавление гауссовского шума различных уровней моделирует изменение шумовой составляющей сигнала как одно из возможных последствий варьирования параметров съёмки.

Результаты промежуточного эксперимента показывают, что для значительной части исследованных классов NTIRE показатели классификации реальных спектров и их печатных представлений изменяются согласованно при зашумлении. На основании этого результата предлагается формировать печатную калибровочную сцену из классов, для которых установлено такое соответствие.

Далее предлагается набор классов печатной калибровочной сцены и проводится его экспериментальная проверка. Показано, что при зашумлении показатели классификации выбранных печатных классов изменяются согласованно с показателями спектрально близких классов реальных ГСИ из набора данных NTIRE 2022. При этом одному классу печатной сцены могут

соответствовать несколько реальных классов. Полученные результаты обосновывают применение предложенной сцены для настройки параметров съёмки и обработки ГСИ. На основе предложенной в первой главе методики создана печатная калибровочная сцена.

Во второй главе разрабатываются алгоритмы спектральной предобработки ГСИ, включающей свёртку и поправку уровня чёрного. Для выбора параметров съёмки и формирования ГСИ, а также настройки коэффициентов предобработки используются ГСИ печатной калибровочной сцены, разработанной в первой главе. Рассматриваются аналитическое определение коэффициентов на основе гребневой регрессии и их нейросетевая настройка с учётом задачи поэлементной классификации.

Для нейросетевого обучения исследуются два варианта функции потерь: перекрёстная энтропия и комбинированный критерий, объединяющий перекрёстную энтропию со среднеквадратичной ошибкой между преобразованными и опорными спектрами. В экспериментах сопоставляются результаты аналитической настройки и двух вариантов нейросетевого обучения, а также исследуется влияние параметров съёмки, формирования и предобработки ГСИ на точность поэлементной классификации.

В третьей главе разрабатываются алгоритмы позиционно-зависимой коррекции отклика гиперспектрометра, учитывающие его зависимость от положения вдоль щели и спектрального канала. В рамках предложенного метода для настройки коэффициентов коррекции используются ГСИ печатной калибровочной сцены, сформированной по методике первой главы. Коэффициенты инициализируются методом гребневой регрессии и затем уточняются при нейросетевом дообучении по перекрёстной энтропии или комбинированному критерию. Рассматриваются внутриприборная коррекция пространственной неоднородности отклика и межприборная коррекция различий между откликами гиперспектрометров. В экспериментах исследуется влияние предложенной коррекции на точность поэлементной классификации, а также возможность применения коэффициентов,

настроенных с одним классификатором, совместно с другими классификаторами.

ГЛАВА 1 ФОРМИРОВАНИЕ И ОЦЕНКА КАЧЕСТВА ГИПЕРСПЕКТРАЛЬНЫХ ИЗОБРАЖЕНИЙ

В настоящей главе разрабатывается методика формирования и применения печатной калибровочной сцены для выбора параметров получения и обработки ГСИ по точности поэлементной классификации. Методика является составной частью предложенного в диссертации метода повышения точности классификации.

Сначала рассматривается модель формирования ГСИ при сканирующей съёмке. Для обоснования выбора показателя сравнения экспериментально исследуется влияние пространственного рассогласования на PSNR, SSIM, SAM и общую точность поэлементной классификации (OA). Отдельно проводится сравнение зарегистрированных и восстановленных ГСИ по OA.

Далее обосновывается возможность использования печатной калибровочной сцены для настройки параметров съёмки и обработки ГСИ. Сначала на данных NTIRE исследуется согласованность изменений показателей классификации реальных спектров и их моделируемых печатных представлений при добавлении гауссовского шума. Затем формируется набор классов печатной калибровочной сцены и проверяется согласованность изменений показателей их классификации с показателями спектрально близких классов реальных ГСИ при зашумлении.

Разработанная сцена используется во второй главе для выбора параметров съёмки и формирования ГСИ, а также настройки спектральной предобработки. В третьей главе ГСИ этой сцены используются для настройки позиционно-зависимой коррекции отклика гиперспектрометра.

1.1 Формирование гиперспектрального изображения при сканирующей съёмке

В щелевой СГСА объектив формирует изображение наблюдаемой сцены в плоскости щелевой диафрагмы. Прошедшее через щель излучение

разлагается по длинам волн, после чего оптическая система формирует изображение на матричном фотоприёмнике. В работе [1] зарегистрированный кадр описывается как пространственная спектрограмма: одна координата соответствует положению вдоль изображения щелевой диафрагмы, другая – спектральному каналу. Последовательность пространственных спектрограмм, зарегистрированных при относительном перемещении аппаратуры и сцены, образует гиперспектральное изображение (ГСИ).

Перед формированием цифрового массива данных сигнал проходит оптическое преобразование, интегрирование на фотоприёмнике и аналого-цифровое преобразование [38].

Помимо пространственного рассогласования, в настоящей работе рассматриваются шум и пространственное размытие ГСИ. Для описания шума используется пуассоновско-гауссовская модель [58]: дисперсия пуассоновской составляющей зависит от уровня сигнала, а дисперсия гауссовской составляющей постоянна [59]. Пространственное размытие в оптической системе описывается функцией рассеяния точки (ФРТ), которая может зависеть от длины волны [60-61]. Математические выражения и параметры внесения этих искажений приведены в пункте 1.3.3.

В настоящей диссертации ГСИ представляется в виде тензора третьего порядка – трёхмерного массива с двумя пространственными и одной спектральной осями [1, 6]:

$$S(i, j, k): 1 \leq i \leq I, 1 \leq j \leq J, 1 \leq k \leq K, \quad (1.1)$$

где $S(i, j, k)$ – значение элемента ГСИ; i – пространственный отсчёт вдоль сканирования, соответствующий номеру кадра; j – пространственный отсчёт вдоль изображения щелевой диафрагмы; k – спектральный отсчёт, соответствующий номеру спектрального канала; I , J и K – размеры тензора по соответствующим осям. Под отсчётами понимаются дискретные позиции по соответствующим осям. Зарегистрированный i -й кадр соответствует срезу $S_i(j, k)$. При фиксированном k срез $S_k(i, j)$ задаёт изображение спектрального

канала ГСИ [1], а при фиксированных i и j последовательность $s_{ij}(k) = S(i, j, k)$ образует спектральный вектор пространственного отсчёта (рисунок 1.1).

Сетка длин волн ГСИ [62] представляет собой упорядоченный набор центральных длин волн его спектральных каналов. Для сопоставления спектров, заданных на разных сетках длин волн, их значения интерполируют на общую сетку.

В зависимости от данных и уровня их обработки значение элемента ГСИ может представлять исходное цифровое значение сигнала, спектральную энергетическую яркость или коэффициент отражения [24, 63]. Конкретная величина и способ её получения указываются при описании данных.

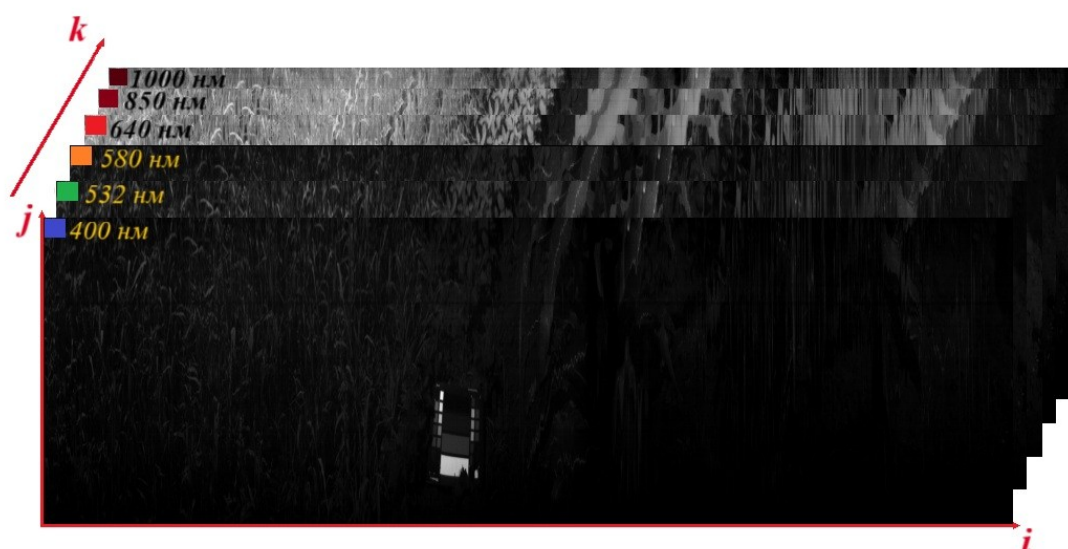


Рисунок 1.1 – Представление ГСИ с пространственными отсчётами i, j и спектральными отсчётами k

Формирование отдельной пространственной спектрограммы $S_i(j, k)$ опишем интегральной моделью из публикации [1]. При регистрации i -го кадра оптическая система формирует в плоскости фотоприёмника распределение освещённости $E(\lambda, x, y)$. Значение элемента спектрограммы определяется интегрированием по светочувствительной площадке элемента фотоприёмника и по диапазону длин волн:

$$S_i(j, k) = \int_{j \cdot b - 0.5\delta}^{j \cdot b + 0.5\delta} \int_{k \cdot b - 0.5\delta}^{k \cdot b + 0.5\delta} \int_{\Delta\lambda} E(\lambda, x, y) G(\lambda) d\lambda dy dx, \quad (1.2)$$

где b – шаг между центрами светочувствительных площадок; δ – размер стороны квадратной светочувствительной площадки; $\Delta\lambda$ – диапазон длин волн, по которому выполняется интегрирование; $E(\lambda, x, y)$ – непрерывное распределение освещённости в плоскости фотоприёмника с учётом спектрального разложения; $G(\lambda)$ – спектральная чувствительность фотоприёмника; λ – длина волны; x и y – физические координаты в плоскости фотоприёмника, оси которых ориентированы соответственно в направлениях возрастания j и k . Координаты x , y и параметры b , δ выражаются в согласованных единицах длины.

При фиксированном i значения элементов для всех j и k образуют одну пространственную спектрограмму. Интеграл (1.2) описывает формирование сигнала на фотоприёмнике. Положение пространственного отсчёта на поверхности сцены зададим координатами центра соответствующего участка поверхности. Для каждого кадра эти участки соответствуют пространственным отсчётам j .

В задачах дистанционного зондирования Земли геометрические искажения сканирующей съёмки описываются с учётом положения строк изображения и ориентации аппаратуры. В работе [13] при коррекции ГСИ положение каждой строки на опорной карте задаётся координатами центра, длиной и углом ориентации. В работе [14] геометрия гиперспектральной сканирующей съёмки уточняется по навигационным данным, совместно полученным кадровым изображениям и цифровой модели рельефа. В работе [64] высокочастотные угловые колебания платформы учитываются в геометрической модели спутниковой панхроматической съёмки для компенсации вызванных ими искажений.

В работе [65] искажения сканирующей съёмки описываются смещениями по двум координатам изображения. Ниже рассматриваются перенос и относительный поворот проекций щелевой диафрагмы в плоскости

сцены. Соответствующие соотношения получены для принятых далее допущений.

Рассмотрим регистрацию двух последовательных кадров i и $i + 1$ щелевой СГСА, установленной на борту беспилотного воздушного судна в составе беспилотной авиационной системы (БАС). Поверхность сцены считаем плоской, высоту съёмки и параметры оптической системы – постоянными. Оптическая ось перпендикулярна поверхности сцены. Допускаются перемещение аппаратуры параллельно этой поверхности и поворот вокруг нормали к ней. Шаг пространственных отсчётов вдоль проекции щели принимается равномерным и одинаковым для обоих кадров. Движением аппаратуры за время экспозиции пренебрегаем.

Шаг пространственной дискретизации на местности (Ground Sampling Distance, GSD) – расстояние между проекциями центров соседних пикселей на поверхность сцены [35]. В работе [14] различаются шаги поперёк и вдоль направления движения, зависящие также от ориентации камеры. Здесь g – постоянный шаг отсчётов вдоль проекции щели, а GSD_y – шаг между строками, принятый при формировании ГСИ. Обе величины выражаются в метрах на пиксель. Постоянство g определяется совокупностью принятых допущений: плоской поверхностью сцены, постоянной высотой съёмки, неизменной оптикой и перпендикулярностью оптической оси поверхности.

В плоскости сцены ось x направим вдоль первой проекции в сторону возрастания номера пространственного отсчёта j , ось y – перпендикулярно ей в плоскости сцены, в сторону увеличения номера строки при сборке ГСИ. В этом геометрическом построении нумерацию j начинаем с нуля. Точки O и A соответствуют отсчёту $j = 0$ первой и второй строк, точки P и B – одному и тому же номеру j в этих строках. Начало координат поместим в O . Поэтому расстояния OP и AB равны jg .

Перенос начала второй строки относительно первой показан стрелкой $O \rightarrow A$ и обозначается на рисунке буквой t . Его составляющие по осям

обозначим $t_{x,i}$ и $t_{y,i}$ (м). Угол θ_i (рад.) отсчитываем от направления первой строки к направлению второй; положительное направление – от x к y .

Из точки B опустим перпендикуляр на прямую, проходящую через A параллельно оси x . Его основание обозначим C . В прямоугольном треугольнике ABC гипотенуза $AB = jg$. Проекции этого отрезка после поворота обозначим $r_{i,j}^x$ и $r_{i,j}^y$: $r_{i,j}^x = jg \cos \theta_i$, $r_{i,j}^y = jg \sin \theta_i$. Для показанного на рисунке положительного угла они равны длинам AC и CB .

Координаты точки B равны $t_{x,i} + r_{i,j}^x$ и $t_{y,i} + r_{i,j}^y$, а координаты точки P – jg и ноль. Чтобы получить смещение между этими точками, вычитаем координаты P из координат B . После перевода смещений в пиксели получаем:

$$d_{i,j}^x = \frac{t_{x,i} + r_{i,j}^x - jg}{g} = \frac{t_{x,i}}{g} + j(\cos \theta_i - 1), \quad (1.3)$$

$$d_{i,j}^y = \frac{t_{y,i} + r_{i,j}^y}{GSD_y} = \frac{t_{y,i}}{GSD_y} + j \frac{g}{GSD_y} \sin \theta_i, \quad (1.4)$$

где $d_{i,j}^x$ – смещение вдоль первой строки, $d_{i,j}^y$ – смещение перпендикулярно ей, выраженные в пикселях. Величины $r_{i,j}^x$ и $r_{i,j}^y$ описывают проекции отрезка AB в метрах. Вклад поворота именно в смещение равен $jg(\cos \theta_i - 1)$ по x и $jg \sin \theta_i$ по y . Вычитание jg в (1.3) учитывает исходное положение отсчёта первой строки.

При нулевых переносе и угле оба смещения равны нулю. При отсутствии поворота и переносе на одну строку без сдвига вдоль неё они равны нулю по x и единице по y . Для оценки отклонения от такого размещения строк из $d_{i,j}^y$ вычитают один пиксель. Неучтённые отклонения вызывают геометрические искажения при формировании ГСИ. Модель рассогласований пространственных спектрограмм при съёмке сканирующим гиперспектрометром представлена на рисунке 1.2.

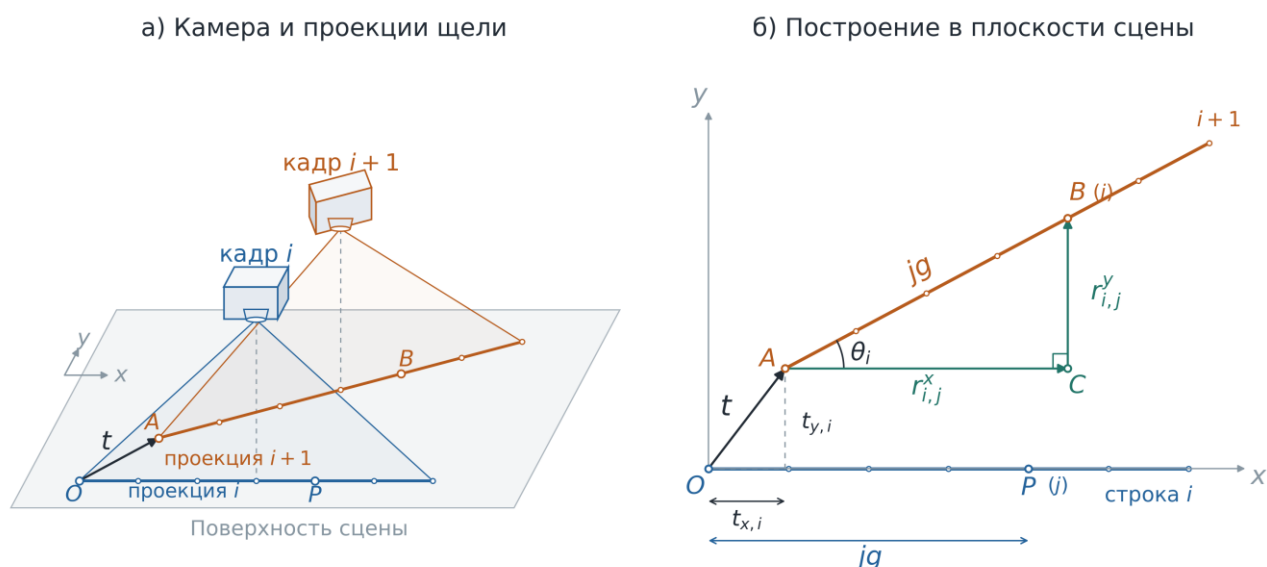


Рисунок 1.2 – Перенос начал строк и разложение отрезка от начала второй строки до отсчёта j на составляющие по осям: а) модель съёмки сцены; б) модель плоскости сцены

1.2 Показатели качества гиперспектральных изображений

Характеристики ГСИ рассматриваются в терминологии стандартов на данные дистанционного зондирования Земли из космоса. ГОСТ Р 59079-2020 содержит определения терминов «пространственное разрешение» и «спектральное разрешение» [24]. Радиометрические показатели приведены в ГОСТ Р 59475-2021 [25], а точность координатной привязки включена в перечень потребительских характеристик ГОСТ Р 59830-2021 [66].

ГОСТ Р 59474-2021 допускает устанавливать критерии оценки пригодности данных и продуктов их обработки для конкретных задач [67]. При оценке ГСИ в настоящей работе также учитываются результаты поэлементной классификации [68].

Для количественной оценки ГСИ далее используются метрики качества, основанные на сопоставлении с эталоном: пиковое отношение сигнала к шуму PSNR, индекс структурного сходства SSIM и спектральный угол SAM. В работе [45] эталонное ГСИ высокого пространственного разрешения сопоставляется с восстановленным: PSNR и SSIM рассчитываются для отдельных спектральных каналов и затем усредняются, а SAM вычисляется между спектрами соответствующих пространственных отсчётов и усредняется

по изображению. В работе [69] эти показатели отнесены к часто используемым объективным показателям оценки результатов сверхразрешения ГСИ. Эти показатели рассчитываются при попарном сравнении значений в одноимённых пространственных позициях эталонного и оцениваемого массивов. Одних значений PSNR, SSIM и SAM недостаточно для оценки пригодности ГСИ к классификации. В работе [70] меньшая ошибка восстановления спектров не всегда соответствовала лучшим результатам классификации по сжатым данным.

В настоящей работе SAM характеризует спектральное сопоставление по углам между спектральными векторами соответствующих пикселей, а SSIM – структурное сходство изображений спектральных каналов. Значения SAM усредняются по области сравнения, значения SSIM – по каналам после усреднения по локальным окнам. Для PSNR вычисляется единый средний квадрат ошибки по всем пространственным позициям и спектральным каналам области сравнения.

Чувствительность показателей к пространственному рассогласованию исследована в ряде работ. В работе [71] приведены примеры уменьшения SSIM при линейном сдвиге изображения. В работе [72] показано, что субпиксельное рассогласование восстановленного и эталонного изображений влияет на значения PSNR и SSIM и может изменить результаты сопоставления методов повышения пространственного разрешения. В работе [73] исследовано пространственное рассогласование между спектральными каналами ГСИ. В приведённых в этой работе экспериментах увеличение межканального смещения сопровождается уменьшением PSNR и SSIM и увеличением SAM.

1.2.1 Влияние сдвига на показатели качества ГСИ и точность классификации

Для исследования чувствительности показателей качества к пространственному рассогласованию рассмотрим постоянный сдвиг ГСИ.

Пусть X – эталонное, а Y – оцениваемое ГСИ, заданные на общей пространственной и спектральной сетке; i и j – пространственные отсчёты, k – номер спектрального канала, K – число каналов. Случай $X = Y$ позволяет оценить влияние сдвига при отсутствии других различий между сравниваемыми ГСИ.

Для исследования частного случая геометрического рассогласования зададим постоянный сдвиг вдоль пространственной оси с отсчётами j : $d_T(i, j) = (d, 0)$. Здесь d – одинаковый для всех пространственных отсчётов и спектральных каналов сдвиг, выраженный в пикселях общей сетки. Рассогласования, обусловленные переменными переносами и поворотами строк, в общем случае зависят от пространственного положения. Постоянный сдвиг позволяет отдельно исследовать чувствительность показателей к пространственному рассогласованию. Схема линейного сдвига ГСИ и маски разметки вдоль оси j представлена на рисунке 1.3.

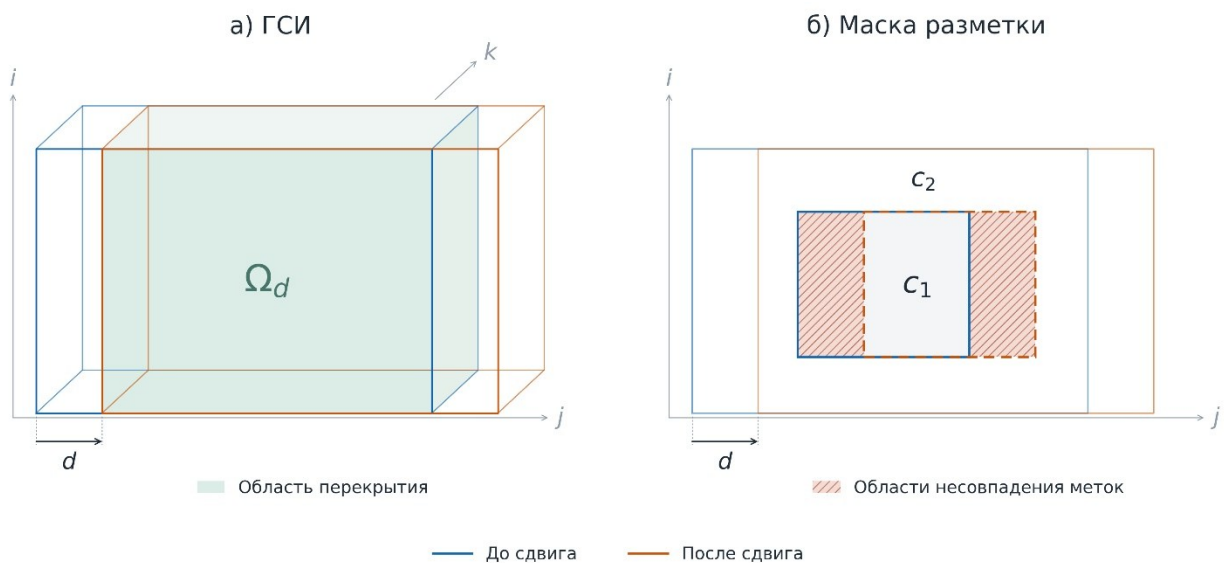


Рисунок 1.3 – Возникновение расхождений при линейном сдвиге на d пикселей вдоль оси j : а) сопоставление значений ГСИ в области Ω_d ; б) несовпадение меток у границ классов c_1 и c_2

Для целого числа пикселей d сдвиг задаётся непосредственно через элементы исходного массива:

$$Y_d(i, j, k) = Y(i, j + d, k), \quad (i, j) \in \Omega_d, \quad (1.5)$$

где $Y_d(i, j, k)$ – значение элемента оцениваемого ГСИ после линейного сдвига; d – целочисленный сдвиг вдоль оси j . При $d > 0$ в (1.5) содержимое массива перемещается в сторону уменьшения j .

Область совместного определения X и Y_d при заданном d обозначим Ω_d . Подстановка (1.5) в определения показателей задаёт зависимости $\text{PSNR}(d)$, $\text{SSIM}(d)$ и $\text{SAM}(d)$. Далее представим их через коэффициент корреляции Пирсона и статистические характеристики сравниваемых значений. Для PSNR они вычисляются по всем элементам ГСИ в области сравнения, для SSIM – в отдельных окнах спектральных каналов, для SAM – по компонентам спектров соответствующих пикселей.

В соответствии с определением MSE в работе [21] средний квадрат ошибки при сдвиге (1.5) имеет вид:

$$\text{MSE}(d) = \frac{1}{K|\Omega_d|} \sum_{(i,j) \in \Omega_d} \sum_{k=1}^K [X(i, j, k) - Y_d(i, j, k)]^2, \quad (1.6)$$

где Ω_d – область сравнения при сдвиге d ; $|\Omega_d|$ – число пространственных позиций в ней; K – число спектральных каналов. Средние, дисперсии и ковариация далее вычисляются по этим же $K|\Omega_d|$ парам значений с делителем $K|\Omega_d|$. Поэтому все статистические характеристики могут зависеть от d , в том числе из-за изменения области сравнения.

В работе [74] полная квадратическая ошибка разделена на квадрат разности средних и ошибку центрированных данных; последняя выражена через стандартные отклонения и корреляцию. Объединяя эти соотношения для X и Y_d , получаем:

$$\text{MSE}(d) = (\mu_X - \mu_{Y_d})^2 + \sigma_X^2 + \sigma_{Y_d}^2 - 2\rho_d \sigma_X \sigma_{Y_d}, \quad (1.7)$$

где μ_X , μ_{Y_d} – средние значения; σ_X , σ_{Y_d} – стандартные отклонения; ρ_d – коэффициент корреляции Пирсона между соответствующими значениями элементов X и Y_d на Ω_d по всем каналам. В обозначениях средних и

стандартных отклонений зависимость от d опущена. Корреляционная запись применима при ненулевых стандартных отклонениях обоих наборов.

Подставляя (1.7) в определение PSNR [21], получаем выражение для расчёта по всему ГСИ:

$$\text{PSNR}(d) = 10\log_{10} \frac{L^2(d)}{(\mu_X - \mu_{Y_d})^2 + \sigma_X^2 + \sigma_{Y_d}^2 - 2\rho_d \sigma_X \sigma_{Y_d}}, \quad (1.8)$$

где $L(d)$ – максимальное значение элемента эталонного ГСИ на Ω_d по всем каналам; остальные обозначения определены в (1.7). Зависимость от d обусловлена изменением сопоставляемых пар значений и, при изменении Ω_d , состава области сравнения.

При дополнительных условиях $\mu_X = \mu_{Y_d}$ и $\sigma_X = \sigma_{Y_d} = \sigma$ выражения (1.7) и (1.8) упрощаются:

$$\text{MSE}(d) = 2\sigma^2(1 - \rho_d), \quad \text{PSNR}(d) = 10\log_{10} \frac{L^2(d)}{2\sigma^2(1 - \rho_d)}, \quad (1.9)$$

где σ – общее стандартное отклонение сравниваемых наборов. При фиксированных σ и L рост корреляции уменьшает MSE и увеличивает PSNR. Равенство и постоянство этих характеристик не следуют из самого линейного сдвига: при выделении области перекрытия состав сравниваемых данных может изменяться.

Для SSIM связь с корреляцией рассматривается в каждом локальном окне q спектрального канала k . В работе [22] структурная компонента выражена через корреляцию нормированных центрированных значений. Подстановка $\sigma_{XY_d} = \rho_{k,q}(d)\sigma_X\sigma_{Y_d}$ в определение SSIM даёт:

$$\text{SSIM}_{k,q}(d) = \frac{(2\mu_X\mu_{Y_d} + C_1)(2\rho_{k,q}(d)\sigma_X\sigma_{Y_d} + C_2)}{(\mu_X^2 + \mu_{Y_d}^2 + C_1)(\sigma_X^2 + \sigma_{Y_d}^2 + C_2)}, \quad (1.10)$$

где $\rho_{k,q}(d)$ – локальный коэффициент корреляции; в пределах (1.10) μ_X , μ_{Y_d} , σ_X , σ_{Y_d} обозначают статистики окна q в канале k , вычисленные с весами локального окна [22]. Корреляция рассчитывается с теми же весами и нормировкой ковариации и дисперсий. Константы C_1 и C_2 определены в [22]. Значение для ГСИ получается усреднением локальных SSIM по окнам и

каналам. При нулевой локальной дисперсии применяется исходная формула SSIM [22].

Для SAM сопоставляются спектры $s = X(i, j, :)$ и $t = Y_d(i, j, :)$. Коэффициент корреляции Пирсона между их компонентами имеет вид [75]:

$$\rho_\lambda(s, t) = \frac{\sum_{k=1}^K (s_k - \bar{s})(t_k - \bar{t})}{\sqrt{\sum_{k=1}^K (s_k - \bar{s})^2} \sqrt{\sum_{k=1}^K (t_k - \bar{t})^2}}, \quad (1.11)$$

где $\rho_\lambda(s, t)$ – корреляция по спектральным компонентам; s_k и t_k – значения элементов спектров в канале k ; $\bar{s}$ и $\bar{t}$ – средние по K каналам. Для дальнейшего преобразования примем $\sigma_s^2 = \frac{1}{K} \sum_{k=1}^K (s_k - \bar{s})^2$ и аналогичное определение σ_t^2 .

В определении SAM используется угол между исходными спектральными векторами [76]. Раскрывая в определении SAM скалярное произведение и квадраты норм через средние, дисперсии и корреляцию (1.11), получаем:

$$\text{SAM}_{ij}(d) = \arccos \frac{\rho_\lambda(s, t) \sigma_s \sigma_t + \bar{s} \bar{t}}{\sqrt{(\sigma_s^2 + \bar{s}^2)(\sigma_t^2 + \bar{t}^2)}}, \quad (1.12)$$

где σ_s и σ_t – стандартные отклонения компонент спектров с делителем K . Формула (1.12) является алгебраическим следствием определений SAM и корреляции Пирсона при ненулевых дисперсиях спектров. Если дисперсия хотя бы одного спектра равна нулю, SAM вычисляется непосредственно по определению угла между спектральными векторами. Средний спектральный угол рассчитывается усреднением по области сравнения.

Соотношение $\theta_c = \arccos \rho_\lambda$ относится к углу между центрированными спектрами [75]. Для исходных спектров формула (1.12) дополнительно учитывает средние и стандартные отклонения. Таким образом, при линейном сдвиге связь показателей с корреляцией задаётся на разных совокупностях значений: по всему ГСИ для PSNR, по локальным окнам каналов для SSIM и по отдельным спектрам для SAM. Формулы (1.7)–(1.12) описывают эту связь для каждой пары X и Y_d .

Полученные выражения описывают изменение PSNR, SSIM и SAM при постоянном линейном сдвиге сравниваемых ГСИ. Далее эта зависимость исследуется экспериментально. Для тех же вариантов ГСИ оценивается общая точность поэлементной классификации при сдвиге массива предсказанных меток относительно эталонной разметки.

Общая точность поэлементной классификации (Overall Accuracy, OA) определяется как доля пикселей, предсказанные метки которых совпадают с эталонными [54]. При линейном сдвиге сравниваемых массивов меток нарушается их пространственное соответствие, что влияет на OA. В работе [55] исследовано влияние позиционной ошибки на тематическую точность данных NLCD 2016 при различных пространственных масштабах и числе классов. Сопоставляемые данные формировались из одного продукта, поэтому изменение точности определялось внесённым рассогласованием. В работе [54] результаты классификации снимков Landsat TM и SPOT HRV сравнивались со своими копиями, сдвинутыми на один пиксель. После сдвига OA составляла соответственно 62,7–77,5 % и 63,2–84,9 %. В работе [77] исследовано влияние смещения независимой эталонной разметки на точность (precision) и полноту (recall) класса «застроено». В работе [78] изучались синтетические бинарные массивы меток при циклическом сдвиге на один пиксель. Оценивание проводилось по F1-мере и нормированному коэффициенту корреляции Мэтьюса; для квадратного объекта установлена зависимость F1-меры от его размера.

Далее исследуется сдвиг на d пикселей вдоль оси j . Значения OA до и после сдвига вычисляются по пикселям в области перекрытия. Изменение показателя определяется разностью $\Delta OA(d) = OA_0(d) - OA(d)$. При отсутствии сдвига исходная разметка совпадает со своей копией и $OA(0) = 1$. Поэтому $\Delta OA(d)$ равно доле пикселей, метки которых перестали совпадать после сдвига. Снижение общей точности в этой постановке определяется

величиной сдвига и пространственной конфигурацией границ классов относительно оси сдвига.

1.3 Исследование влияния сдвига на показатели качества и точность классификации ГСИ

В эксперименте сопоставляется влияние линейного сдвига на показатели PSNR, SSIM и SAM и на общую точность поэлементной классификации (OA). В первом случае сдвигается оцениваемое ГСИ относительно эталонного, во втором случае массив предсказанных меток относительно эталонной разметки. Рассматриваются исходные ГСИ, их зашумлённые варианты и варианты после пространственного размытия и внесения шума. Результаты обоих способов оценки приводятся последовательно для каждой сцены.

1.3.1 Экспериментальные данные

Для экспериментальной оценки использованы ГСИ набора ARAD_1K [10] и ГСИ речных участков [11]. Сопоставление с ГСИ, восстановленными методом MST++ [46], вынесено в подраздел 1.4.

Первую группу составили ГСИ набора ARAD_1K, опубликованного в рамках NTIRE 2022 Spectral Recovery Challenge. Набор содержит 1000 сцен: 900 обучающих, 50 валидационных и 50 тестовых. Для 950 сцен опубликованы гиперспектральные данные, тогда как эталонные ГСИ тестовой части не опубликованы. Каждое ГСИ имеет размер 482×512 пространственных отсчётов и содержит 31 спектральный канал в диапазоне 400–700 нм с шагом 10 нм [10]. Данные доступны через официальный репозиторий ARAD_1K. В эксперименте использовались эталонное ГСИ и сформированные из него зашумлённое ГСИ и ГСИ после пространственного размытия и внесения шума (рисунок 1.4).

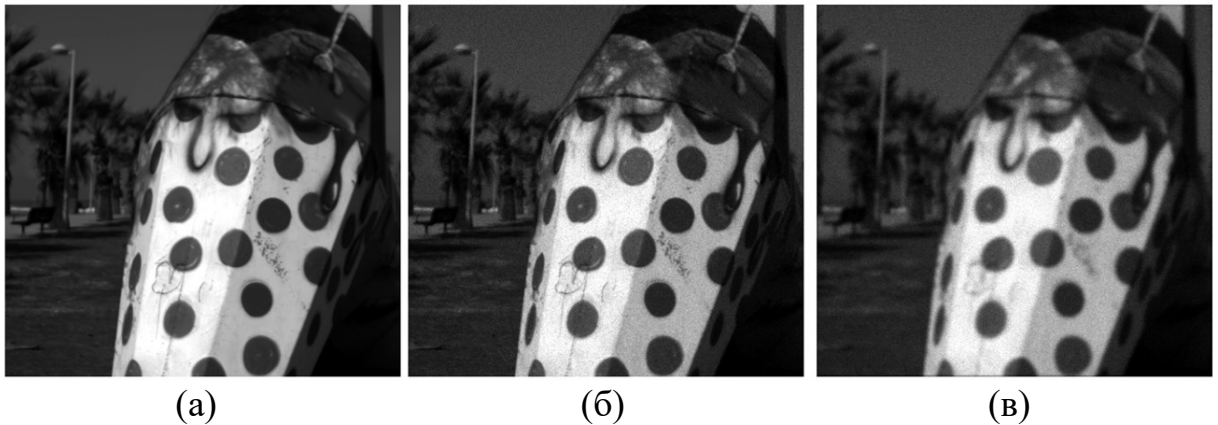


Рисунок 1.4 – Изображение спектрального канала эталонного ГСИ набора ARAD_1K (а), зашумлённого ГСИ (б) и ГСИ после пространственного размытия и внесения шума (в)

Вторую группу составили ГСИ речных участков с различной интенсивностью цветения водорослей [11]. Для их регистрации использовался разработанный в Самарском университете бортовой щелевой гиперспектрометр с оптической схемой Оффнера (далее – бортовой гиперспектрометр). Съёмка с борта беспилотной авиационной системы выполнялась на четырёх участках реки Чапаевки: «Горки», «Причал», «Маяк» и «Малое Томылово»; ГСИ формировались из видеопоследовательности с использованием данных телеметрии [11]. В работе [79] ГСИ указанных участков использованы для классификации интенсивности цветения водорослей методами машинного обучения. При моделировании искажений исходное ГСИ принималось за эталонное. Из него формировались зашумлённое ГСИ и ГСИ после пространственного размытия и внесения шума (рисунок 1.5).

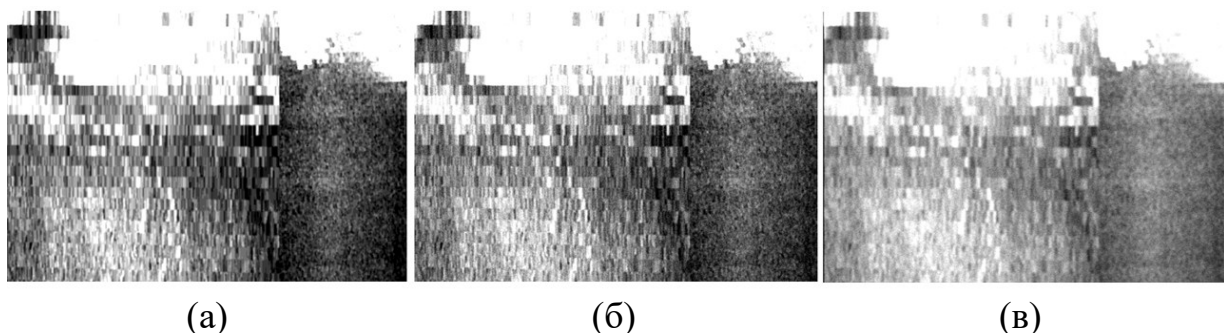


Рисунок 1.5 – Изображение спектрального канала эталонного ГСИ участка реки Чапаевки (а), зашумлённого ГСИ (б) и ГСИ после пространственного размытия и внесения шума (в)

1.3.2 Постановка экспериментов

В эксперименте из эталонного ГСИ формировались два варианта: зашумлённое ГСИ и ГСИ после пространственного размытия и внесения шума. Эта постановка применялась к данным ARAD_1K и ГСИ участков реки Чапаевки. Каждый вариант, включая эталонное ГСИ, сдвигался и сравнивался с несмещённым эталонным ГСИ.

В соответствии с постановкой пункта 1.2.1 выполнялся сдвиг оцениваемого ГСИ вдоль оси j на целое число пикселей d . PSNR, SSIM и SAM вычислялись в области перекрытия эталонного и сдвинутого ГСИ. Для каждого d эта область определялась одинаково для всех вариантов. Соотношения (1.7)–(1.12) выражают показатели через коэффициент корреляции Пирсона, средние и дисперсии сравниваемых данных. Влияние сдвига на ОА оценивалось по предсказанным и эталонным меткам в соответствующей области перекрытия.

По зависимостям $PSNR(d)$, $SSIM(d)$ и $SAM(d)$ и $OA(d)$ оценивалось, как сдвиг влияет на значения показателей и порядок вариантов ГСИ по этим показателям.

1.3.3 Влияние искажений, вносимых в ГСИ

Для неискажённого значения x пуассоновско-гауссовская модель имеет вид [58]:

$$y = \eta_p + \eta_g = \frac{1}{a} \alpha + \beta, \quad \alpha \sim \mathcal{P}(ax), \quad \beta \sim \mathcal{N}(0, b^2), \quad (1.13)$$

где y – наблюдаемое значение; x – неискажённое значение; η_p и η_g – пуассоновская и гауссовская составляющие; a и b – параметры модели; α и β – независимые случайные величины с распределениями $\mathcal{P}(ax)$ и $\mathcal{N}(0, b^2)$ соответственно. Здесь $x \geq 0, a > 0, b \geq 0$. Пуассоновская составляющая имеет математическое ожидание x . Математическое ожидание и дисперсия наблюдаемого значения в модели (1.13) равны [58]:

$$E[y] = x, \quad D[y] = \frac{x}{a} + b^2, \quad (1.14)$$

где $E[y]$ и $D[y]$ – соответственно математическое ожидание и дисперсия наблюдаемого значения y ; остальные обозначения введены в формуле (1.13).

Перед внесением шума значения элементов входного ГСИ нормировались на его максимум. Зашумлённое ГСИ формировалось из эталонного ГСИ S_{GT} , а ГСИ после пространственного размытия и внесения шума – из ГСИ после пространственного размытия S_{PSF} . Максимум определялся отдельно для каждого входного ГСИ. Нормированное значение элемента ГСИ и параметр пуассоновской составляющей определялись выражениями:

$$\tilde{S}_l(i, j, k) = \frac{S_l(i, j, k)}{S_{\max, l}}, \quad S_{\max, l} = \max_{i, j, k} S_l(i, j, k), \quad N_{\text{peak}} = \frac{1}{\sigma_s^2}, \quad (1.15)$$

где $\tilde{S}_l(i, j, k)$ – нормированное значение элемента входного ГСИ; $S_l(i, j, k)$ – исходное значение этого элемента; $S_{\max, l}$ – максимум l -го варианта; N_{peak} – параметр пуассоновской составляющей; σ_s – стандартное отклонение пуассоновской составляющей при единичном нормированном сигнале.

Для каждого элемента ГСИ пуассоновская и гауссовская составляющие задавались независимыми случайными величинами:

$$\alpha_l(i, j, k) \sim \mathcal{P}(N_{\text{peak}} \tilde{S}_l(i, j, k)), \quad \beta_l(i, j, k) \sim \mathcal{N}(0, \sigma_r^2), \quad (1.16)$$

где $\alpha_{l, i, j, k}$ и $\beta_{l, i, j, k}$ – независимые случайные величины, задающие пуассоновскую и гауссовскую составляющие; $P(\cdot)$ и $N(\cdot)$ – соответственно пуассоновское и нормальное распределения; σ_r – стандартное отклонение гауссовской составляющей; остальные обозначения определены формулой (1.15).

При переходе к исходной шкале отрицательные значения заменялись нулями; ограничение значений сверху не применялось:

$$S_{NOISE,l}(i,j,k) = S_{max,l} \max\left(\frac{\alpha_l(i,j,k)}{N_{peak}} + \beta_l(i,j,k), 0\right), \quad (1.17)$$

где $S_{NOISE,l}(i,j,k)$ – значение элемента зашумлённого l -го варианта ГСИ; операция $\max(\cdot, 0)$ исключает отрицательные значения; остальные обозначения введены в формулах (1.15)–(1.16).

В эксперименте для ГСИ ARAD_1K_0930 использованы $\sigma_s = 0,05$ и $\sigma_r = 0,001$, поэтому $N_{peak} = 400$. Для вариантов S_{NOISE} и $S_{PSF+NOISE}$ формировалось по одной реализации шума; усреднение результатов по повторным реализациям не выполнялось. Генератор псевдослучайных чисел инициализировался значением 12345. Сформированные зашумлённые ГСИ сохранялись и далее использовались без повторной генерации шума. Поскольку $S_{max,l}$ определялся отдельно для эталонного ГСИ и ГСИ после пространственного размытия, одинаковые σ_s и σ_r задавали одинаковый относительный, но не обязательно одинаковый абсолютный масштаб шума.

В эксперименте использована изотропная гауссова ФРТ, одинаковая для всех спектральных каналов. Её дискретное ядро имеет вид:

$$h_\sigma(u, v) = \frac{\exp\left[-\frac{u^2+v^2}{2\sigma^2}\right]}{\sum_{p=-r}^r \sum_{q=-r}^r \exp\left[-\frac{p^2+q^2}{2\sigma^2}\right]}, \quad -r \leq u, v \leq r, \quad (1.18)$$

где u и v – дискретные координаты ядра; σ – стандартное отклонение гауссовой ФРТ в пикселях; r – радиус ядра; $h_\sigma(u, v)$ – его нормированное значение. Размер квадратного ядра по каждой оси равен $2r + 1$, а условие нормировки имеет вид:

$$\sum_{u=-r}^r \sum_{v=-r}^r h_\sigma(u, v) = 1, \quad (1.19)$$

где суммирование выполняется по $u, v = -r, \dots, r$; равенство единице обеспечивает сохранение постоянной составляющей при свёртке вне области влияния нулевого дополнения на границах изображения.

Пространственное размытие выполнялось независимо для каждого спектрального канала:

$$S_{\text{PSF}}(i, j, k) = \sum_{u=-r}^r \sum_{v=-r}^r h_{\sigma}(u, v) S_{\text{GT}}(i - u, j - v, k). \quad (1.20)$$

где $S_{\text{PSF}}(i, j, k)$ – значение элемента ГСИ после пространственного размытия; $S_{\text{GT}}(i - u, j - v, k)$ – значение элемента эталонного ГСИ.

В эксперименте использована гауссова ФРТ с параметрами $\sigma = 2$ пикселя и $r = 10$. Размер ядра составлял 21×21 . Значения за пределами изображения принимались равными нулю; после свёртки пространственный размер ГСИ сохранялся. К ГСИ после пространственного размытия добавлялся шум:

$$S_{\text{PSF}+\text{NOISE}} = \text{Noise}(S_{\text{PSF}}), \quad (1.21)$$

где $\text{Noise}(\cdot)$ – оператор внесения шума по формулам (1.15)–(1.17); $S_{\text{PSF}+\text{NOISE}}$ – ГСИ после пространственного размытия и внесения шума.

В данном эксперименте сравнивались эталонное ГСИ S_{GT} , зашумлённое ГСИ S_{NOISE} и ГСИ после пространственного размытия и внесения шума $S_{\text{PSF}+\text{NOISE}}$. ГСИ после пространственного размытия S_{PSF} сохранялось для иллюстрации действия ФРТ; значения PSNR, SSIM и SAM для него не приводились.

1.3.4 Протокол обучения и применения классификаторов

В двух экспериментах, описанных в пунктах 1.3.5 и 1.3.6, использовался единый протокол обучения и применения классификатора. Для поэлементной классификации применялась спектральная одномерная свёрточная нейронная сеть (1D-CNN) [47]. Пакет данных (англ. batch) – группа элементов данных, совместно обрабатываемых нейронной сетью [68]. В настоящих экспериментах сеть содержала три одномерных свёрточных блока с 32, 64 и 96 выходными каналами, пакетной нормализацией (BatchNorm) и функцией активации ReLU, слой глобального усреднения и линейный классификационный слой. На вход сети подавался спектральный вектор отдельного пикселя; спектральные векторы соседних пространственных отсчётов не использовались.

Для каждого из двух наборов классификатор обучался отдельно на размеченной части соответствующего эталонного ГСИ. В обучающую выборку случайным образом включалось 10 % размеченных пикселей каждого класса; в качестве функции потерь использовалась категориальная перекрёстная энтропия. После завершения обучения параметры сети фиксировались. Для каждого набора обученный классификатор без повторного обучения применялся к эталонному и всем оцениваемым вариантам ГСИ: к зашумлённому ГСИ и ГСИ после пространственного размытия и внесения шума.

Результатом применения классификатора являлся двумерный массив предсказанных меток. Общая точность рассчитывалась как доля совпадающих меток в области перекрытия этого массива и эталонной разметки.

Для каждого варианта ГСИ массив предсказанных меток формировался однократно. Затем он сдвигался вдоль оси с отсчётами j относительно неподвижной эталонной разметки согласно постановке пункта 1.2.1. При изменении линейного сдвига повторная классификация не выполнялась. Поэтому изменение ОА характеризовало только влияние пространственного рассогласования предсказанных и эталонных меток.

В дополнительном эксперименте на сцене ARAD_1K_0927 наряду с 1D-CNN использовались модифицированная многомасштабная трёхмерная свёрточная сеть NM3DCNN [48, 68] и пространственно-спектральный трансформер SSFTT [49]. На вход NM3DCNN и SSFTT подавались пространственно-спектральные фрагменты ГСИ; результатом классификации являлась метка центрального пикселя. Общая точность для всех трёх моделей рассчитывалась по массивам предсказанных меток согласно описанной выше процедуре.

1.3.5 Влияние линейного сдвига на показатели качества и ОА для ГСИ NTIRE 2022

Для оценки влияния линейного сдвига на PSNR, SSIM и SAM использовано эталонное ГСИ NTIRE 2022 «ARAD_1K_0930». Рассматривались эталонное ГСИ, зашумлённое ГСИ и ГСИ после пространственного размытия и внесения шума. Каждый вариант сдвигался вдоль оси с отсчётами j на целое число пикселей $d = 0, \dots, 5$ и сопоставлялся с несмещённым эталонным ГСИ. Порядок формирования вариантов ГСИ и правило выделения области перекрытия Ω_d приведены в пунктах 1.3.2–1.3.4. Результаты обучения классификатора для ГСИ ARAD_1K

Для ГСИ «ARAD_1K_0930» классификатор применялся к эталонному, зашумлённому ГСИ и ГСИ после пространственного размытия и внесения шума. Изменение функции потерь и ОА в процессе обучения представлено на рисунке 1.6. Далее для этих же вариантов сопоставляются зависимости PSNR, SSIM, SAM и ОА от величины линейного сдвига.

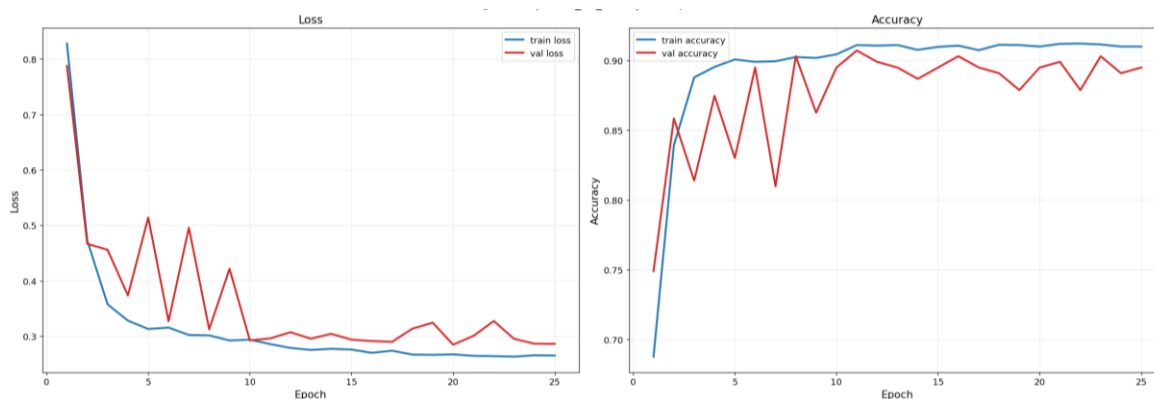


Рисунок 1.6 – Изменение функции потерь и общей точности классификации при обучении 1D-CNN на ГСИ ARAD_1K_0930: слева – функция потерь; справа – общая точность классификации. Синим цветом показана обучающая выборка, красным – валидационная

Зависимости PSNR, SSIM и SAM от величины сдвига представлены на рисунках 1.7–1.9.

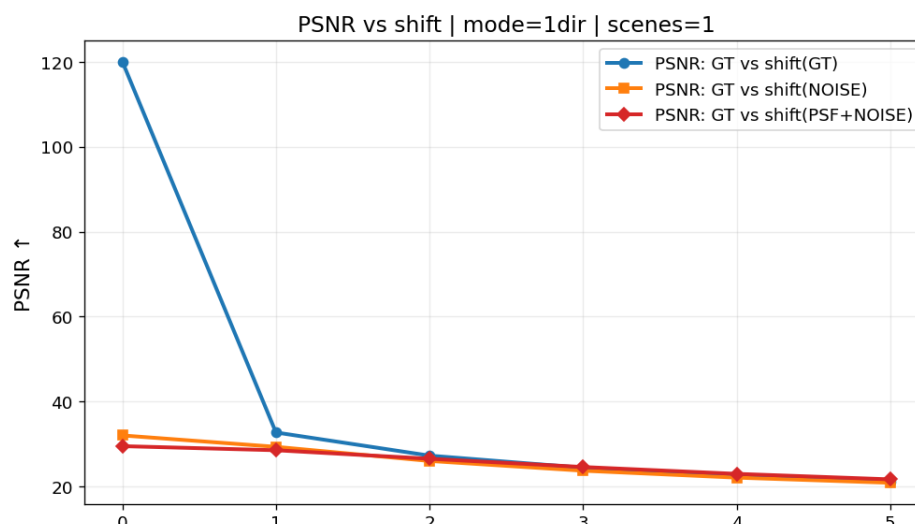


Рисунок 1.7 – Зависимость PSNR от линейного сдвига d вдоль оси s отсчётами j для эталонного ГСИ ARAD_1K_0930, зашумлённого ГСИ и ГСИ после пространственного размытия и внесения шума

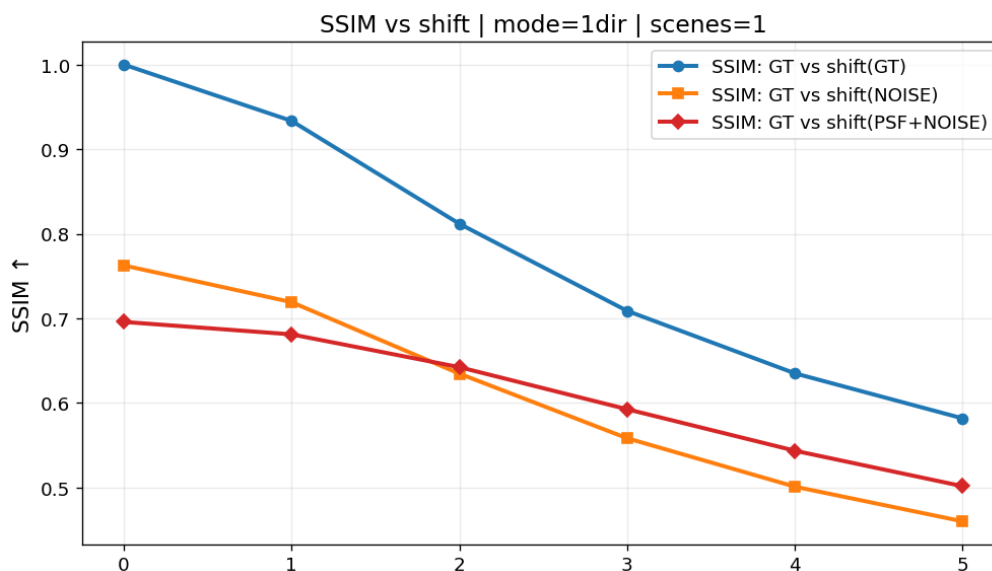


Рисунок 1.8 – Зависимость SSIM от линейного сдвига d вдоль оси s отсчётами j для эталонного ГСИ ARAD_1K_0930, зашумлённого ГСИ и ГСИ после пространственного размытия и внесения шума

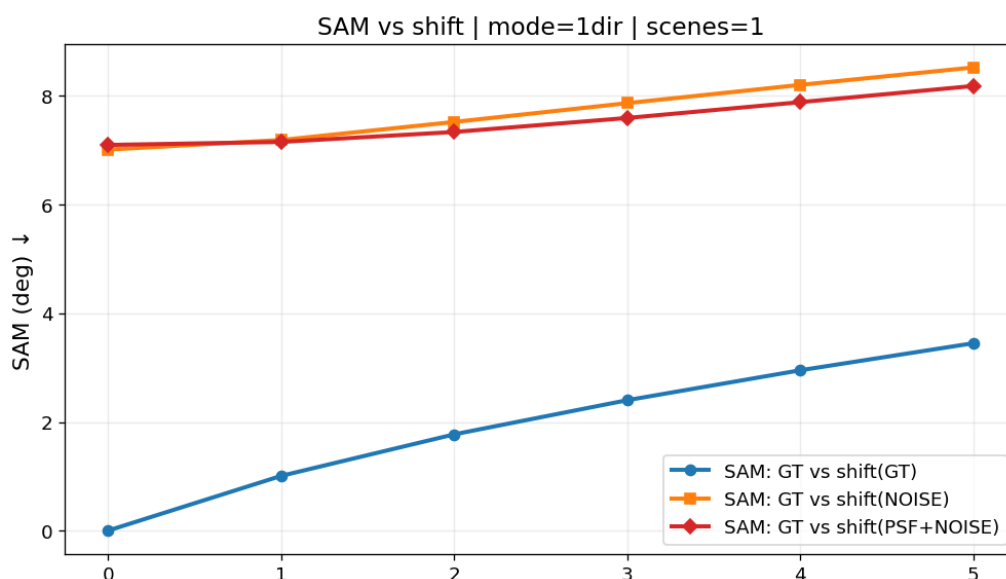


Рисунок 1.9 – Зависимость SAM от линейного сдвига d вдоль оси с отсчётами j для эталонного ГСИ ARAD_1K_0930, зашумлённого ГСИ и ГСИ после пространственного размытия и внесения шума

Зависимости ОА от величины линейного сдвига массивов предсказанных меток представлены на рисунке 1.10.

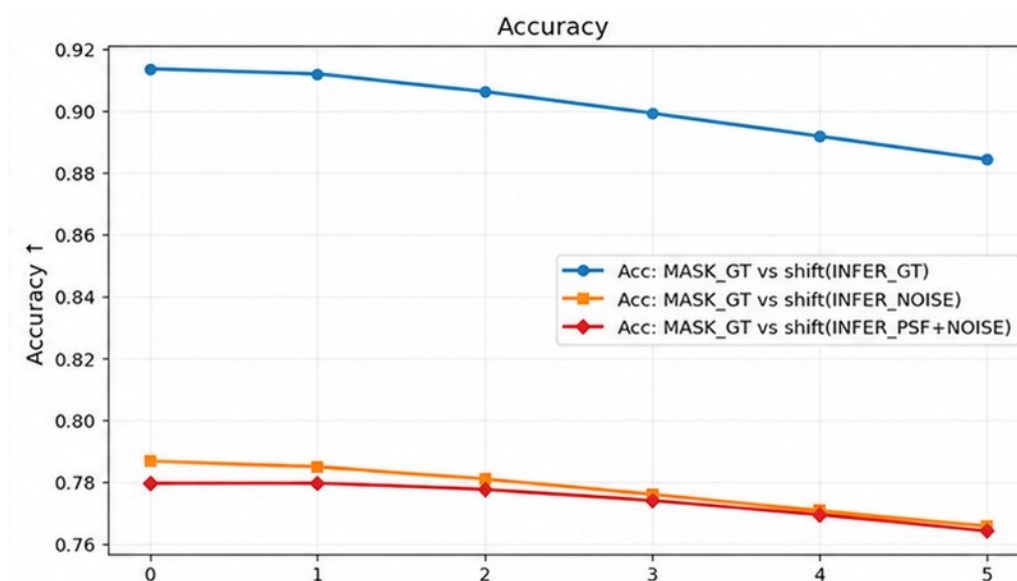


Рисунок 1.10 – Зависимость ОА от линейного сдвига d массива предсказанных меток относительно эталонной разметки для ГСИ ARAD_1K_0930

При увеличении сдвига значения всех трёх показателей изменялись монотонно. Для эталонного ГСИ при увеличении сдвига от $d = 1$ до $d = 5$ значение PSNR уменьшилось с 32,745 до 21,258 дБ, SSIM – с 0,9334 до 0,5818, а SAM увеличился с $1,010^\circ$ до $3,447^\circ$. Для зашумлённого ГСИ в диапазоне $d = 0, \dots, 5$ значение PSNR уменьшилось с 32,050 до 20,907 дБ, SSIM – с 0,7626 до

0,4601, *SAM* увеличился с $7,003^\circ$ до $8,512^\circ$. Для ГСИ после пространственного размытия и внесения шума соответствующие изменения составили 29,530–21,716 дБ, 0,6959–0,5017 и $7,089^\circ$ – $8,172^\circ$.

Начиная с $d = 2$, ГСИ после пространственного размытия и внесения шума имело более высокие значения PSNR и SSIM и меньшие значения *SAM*, чем зашумлённое ГСИ. При $d = 0$ соотношение значений всех трёх показателей было обратным. Следовательно, в данном эксперименте пространственное размытие уменьшает чувствительность показателей к последующему линейному сдвигу и может изменять порядок вариантов ГСИ по PSNR, SSIM и *SAM*.

Полученные результаты подтверждают, что PSNR, SSIM и *SAM* при сопоставлении с эталонным ГСИ зависят не только от внесения шума и пространственного размытия, но и от пространственного рассогласования сравниваемых ГСИ. Поскольку область Ω_d изменяется с d , приведённые зависимости характеризуют совместное влияние линейного сдвига и изменения состава оцениваемых пространственных отсчётов.

При увеличении d общая точность уменьшалась для всех трёх вариантов. Во всём исследованном диапазоне наибольшая ОА соответствовала эталонному ГСИ, промежуточная – зашумлённому ГСИ, наименьшая – ГСИ после пространственного размытия и внесения шума. Следовательно, порядок вариантов по ОА, установленный при отсутствии сдвига, сохранялся при $d = 0, \dots, 5$. В то же время, как показано выше на рисунках 1.7–1.9, PSNR, SSIM и *SAM* при увеличении сдвига изменяли взаимный порядок двух искажённых вариантов.

1.3.6 Влияние сдвига на показатели качества и ОА для ГСИ участка реки Чапаевки

Во втором эксперименте использовалось ГСИ участка реки Чапаевки, зарегистрированное бортовым гиперспектрометром, описанным в пункте 1.3.1. Как и в первом эксперименте, сравнивались эталонное ГСИ,

зашумлённое ГСИ и ГСИ после пространственного размытия и внесения шума. Каждый вариант сдвигался вдоль оси с отсчётами j на целое число пикселей $d = 0, \dots, 5$ и сопоставлялся с несмещённым эталонным ГСИ. Порядок формирования вариантов ГСИ и правило выделения области перекрытия Ω_d приведены в пунктах 1.3.2–1.3.4.

К ГСИ участка реки Чапаевки и его искажённым вариантам применялся тот же протокол классификации. Массивы предсказанных меток сдвигались относительно эталонной разметки на $d = 0, \dots, 5$ пикселей. Результаты обучения представлены на рисунке 1.11, а зависимости ОА от величины линейного сдвига – на рисунке 1.15.

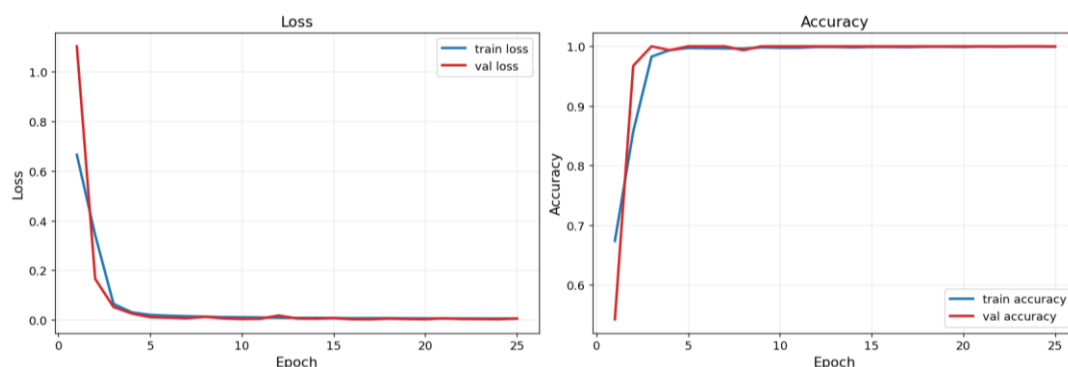


Рисунок 1.11 – Изменение функции потерь и общей точности классификации при обучении 1D-CNN на ГСИ участка реки Чапаевки: слева – функция потерь; справа – общая точность классификации. Синим цветом показана обучающая выборка, красным – валидационная

Зависимости PSNR, SSIM и SAM от величины линейного сдвига представлены на рисунках 1.12–1.14.

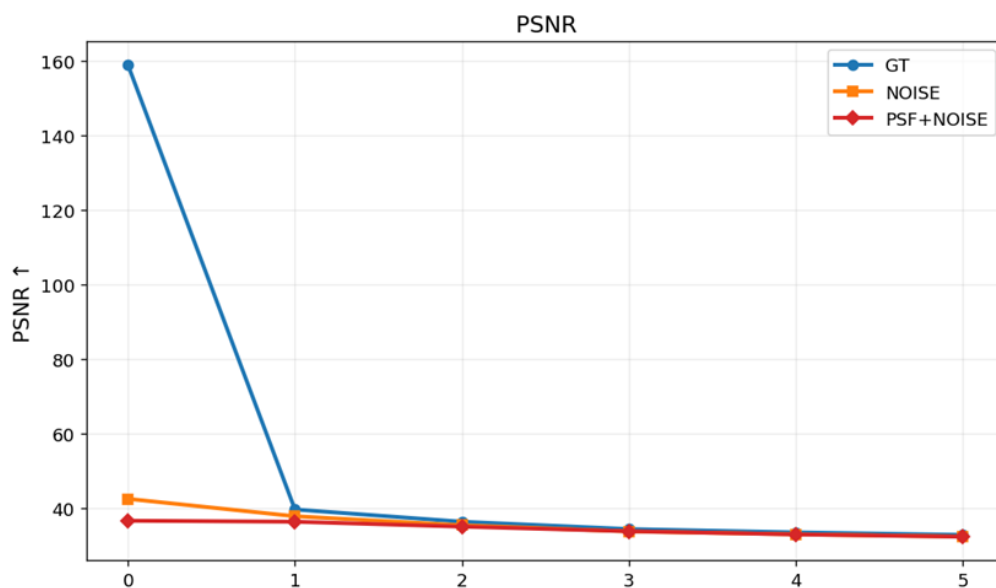


Рисунок 1.12 – Зависимость PSNR от линейного сдвига d для эталонного ГСИ участка реки Чапаевки, зашумлённого ГСИ и ГСИ после пространственного размытия и внесения шума

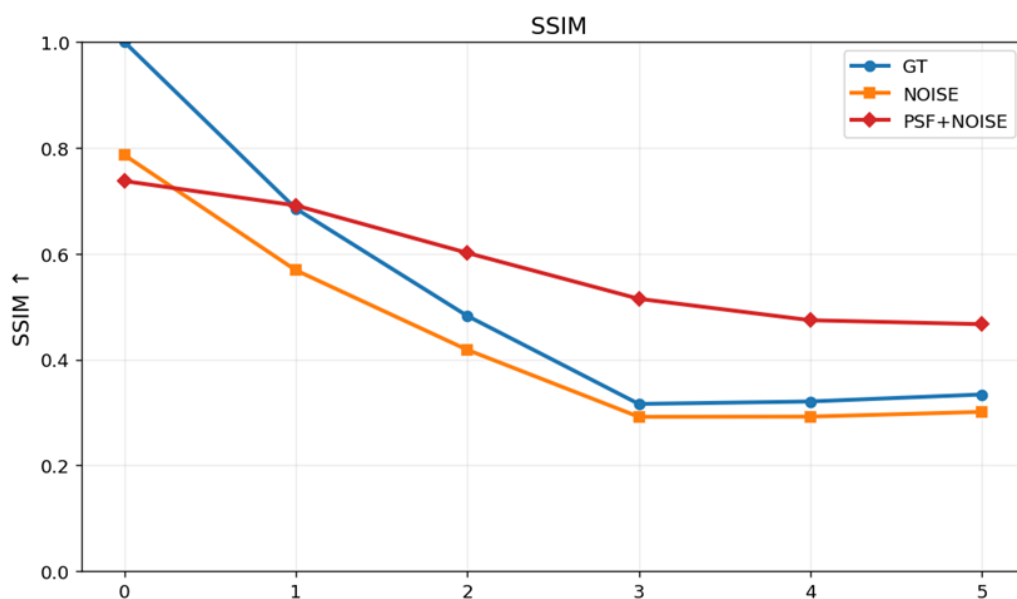


Рисунок 1.13 – Зависимость SSIM от линейного сдвига d для эталонного ГСИ участка реки Чапаевки, зашумлённого ГСИ и ГСИ после пространственного размытия и внесения шума

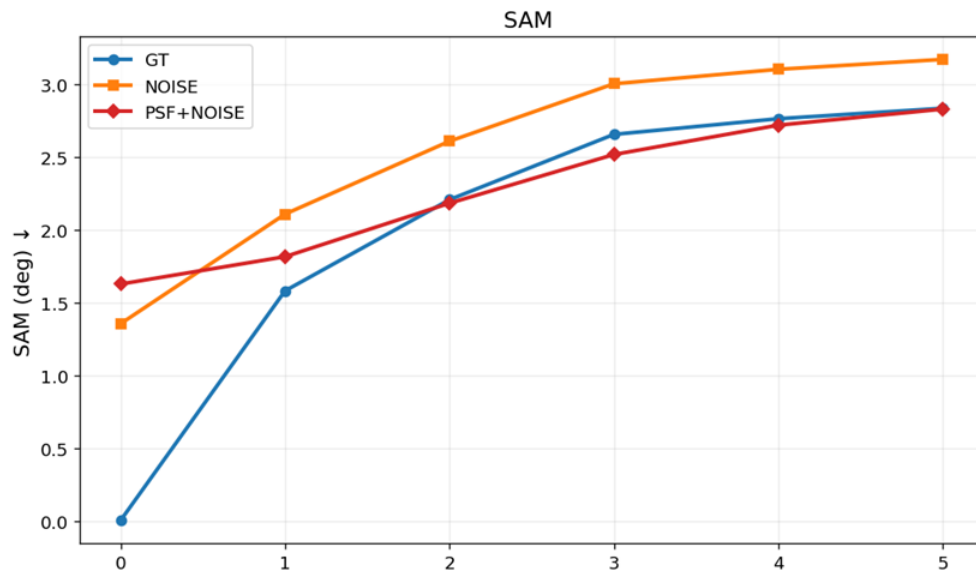


Рисунок 1.14 – Зависимость SAM от линейного сдвига d для эталонного ГСИ участка реки Чапаевки, зашумлённого ГСИ и ГСИ после пространственного размытия и внесения шума

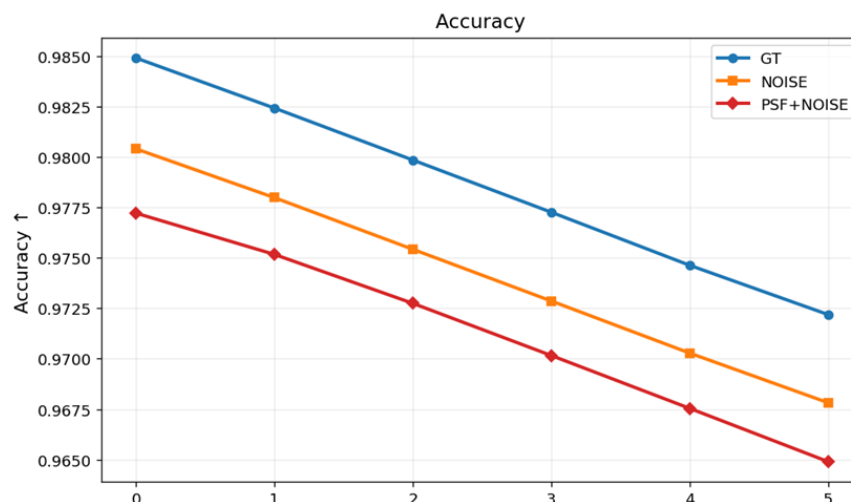


Рисунок 1.15 – Зависимость ОА от линейного сдвига d массива предсказанных меток относительно эталонной разметки для ГСИ участка реки Чапаевки

При увеличении d значение $PSNR$ уменьшалось, а значение SAM возрастало для всех трёх вариантов. Значения $SSIM$ для эталонного и зашумлённого ГСИ резко уменьшались до $d = 3$, после чего изменялись незначительно; для ГСИ после пространственного размытия и внесения шума снижение $SSIM$ было более плавным. Уже при $d = 2$ ГСИ после пространственного размытия и внесения шума имело большее значение $SSIM$, чем эталонное и зашумлённое ГСИ. При $d = 5$ значения $PSNR$ для всех вариантов сблизались, а значения SAM эталонного ГСИ и ГСИ после

пространственного размытия и внесения шума практически совпали. Пространственное рассогласование изменило полученный при $d = 0$ порядок вариантов по PSNR, SSIM и SAM.

Значение ОА монотонно уменьшалось с увеличением линейного сдвига для каждого варианта ГСИ. Зависимости имели близкий характер и не пересекались: наибольшая общая точность соответствовала эталонному ГСИ, промежуточная – зашумлённому ГСИ, наименьшая – ГСИ после пространственного размытия и внесения шума. Таким образом, при $d = 0, \dots, 5$ ОА сохраняла порядок вариантов, установленный при отсутствии сдвига, тогда как для PSNR, SSIM и SAM на рисунках 1.12–1.14 установлено его изменение при пространственном рассогласовании.

Для ГСИ ARAD_1K и участка реки Чапаевки увеличение линейного сдвига приводило к изменению как показателей сравнения с эталоном, так и ОА. При этом взаимный порядок вариантов ГСИ по этим показателям изменялся различным образом.

В обоих экспериментах сдвиг массива предсказанных меток относительно эталонной разметки приводил к уменьшению ОА вследствие нарушения поэлементного соответствия меток. В диапазоне $d = 0, \dots, 5$ порядок вариантов ГСИ по ОА сохранялся, тогда как порядок по PSNR, SSIM и SAM изменялся.

Для проверки сохранения выявленного соотношения точностей при использовании различных классификаторов проведён дополнительный эксперимент на сцене ARAD_1K_0927. Сопоставлялись зашумлённое ГСИ и ГСИ после пространственного размытия и внесения шума. Поэлементная классификация выполнялась с использованием 1D-CNN, NM3DCNN и SSFTT. Для каждой модели массивы предсказанных меток формировались однократно и затем сдвигались относительно неподвижной эталонной разметки без повторной классификации. Показатели PSNR, SSIM и SAM

рассчитывались при сдвиге оцениваемых ГСИ относительно эталонного ГСИ. Результаты представлены на рисунке 1.16.

В диапазоне линейных сдвигов от 0 до 5 пикселей для каждой из трёх моделей ОА зашумлённого ГСИ оставалась не ниже ОА ГСИ после пространственного размытия и внесения шума. При этом пространственное рассогласование изменяло соотношение значений PSNR, SSIM и SAM для двух сравниваемых вариантов ГСИ.

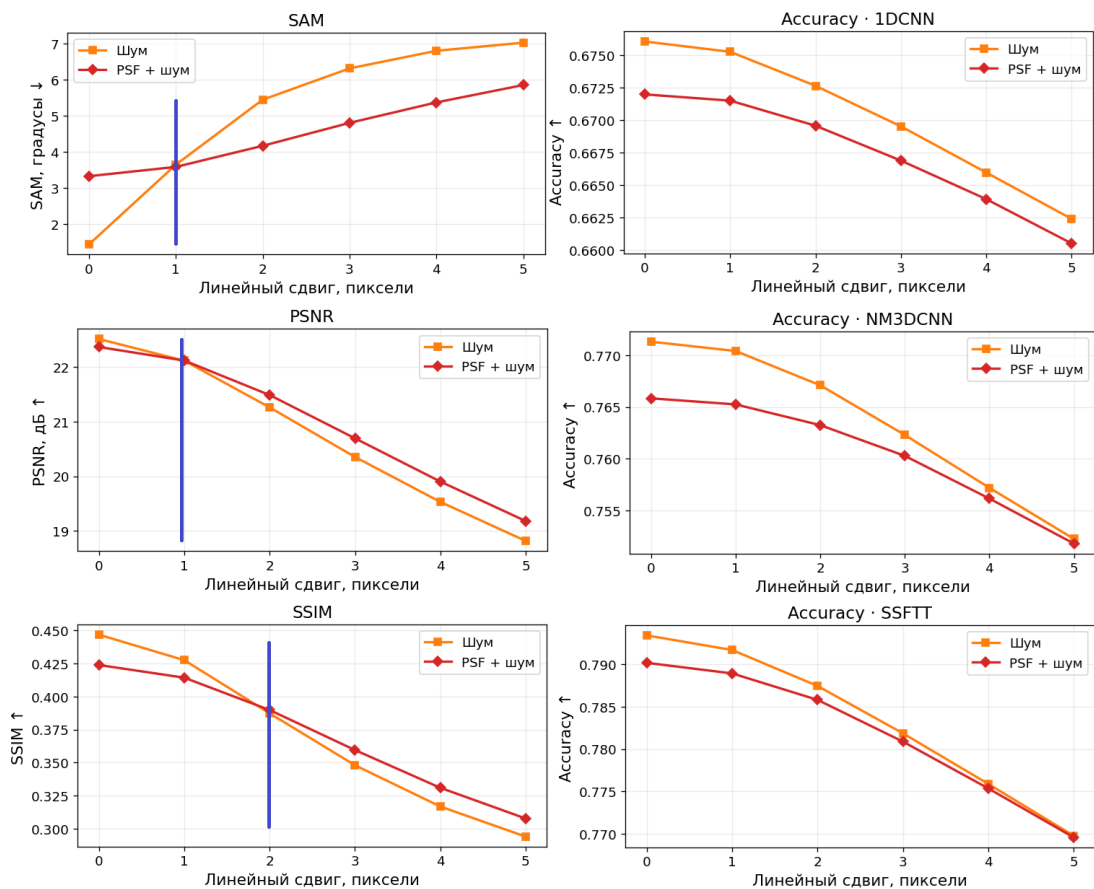


Рисунок 1.16 – Зависимости PSNR, SSIM, SAM и общей точности поэлементной классификации (ОА) от линейного сдвига для двух вариантов ГСИ сцены ARAD_1K_0927: с внесённым шумом и после пространственного размытия и внесения шума. ОА рассчитана с использованием 1D-CNN, NM3DCNN и SSFTT

1.3.7 Оценка снижения точности классификации при линейном сдвиге

Для оценки связи между величиной линейного сдвига и изменением общей точности проведён дополнительный вычислительный эксперимент на 50 сценах набора ARAD_1K [10] (ARAD_1K_0901–ARAD_1K_0950). Для

каждой сцены использовались четыре заранее сформированных массива меток локальных спектральных классов при $C = 2, 3, 4$ и 5 ; всего исследовано 200 массивов меток размером $I \times J = 482 \times 512$ пикселей. Исходные области классов выделялись на RGB-представлении ГСИ с использованием модели Segment Anything [80], после чего их границы уточнялись по спектральному углу [81]. В данном эксперименте исходная разметка локальных спектральных классов использовалась как эталонная и сопоставлялась со своей сдвинутой копией. Тем самым исключалось влияние ошибок классификатора и оценивалось только нарушение пространственного соответствия меток. Пример исходной сцены, изображений отдельных спектральных каналов и соответствующей разметки классов приведён на рисунке 1.17.

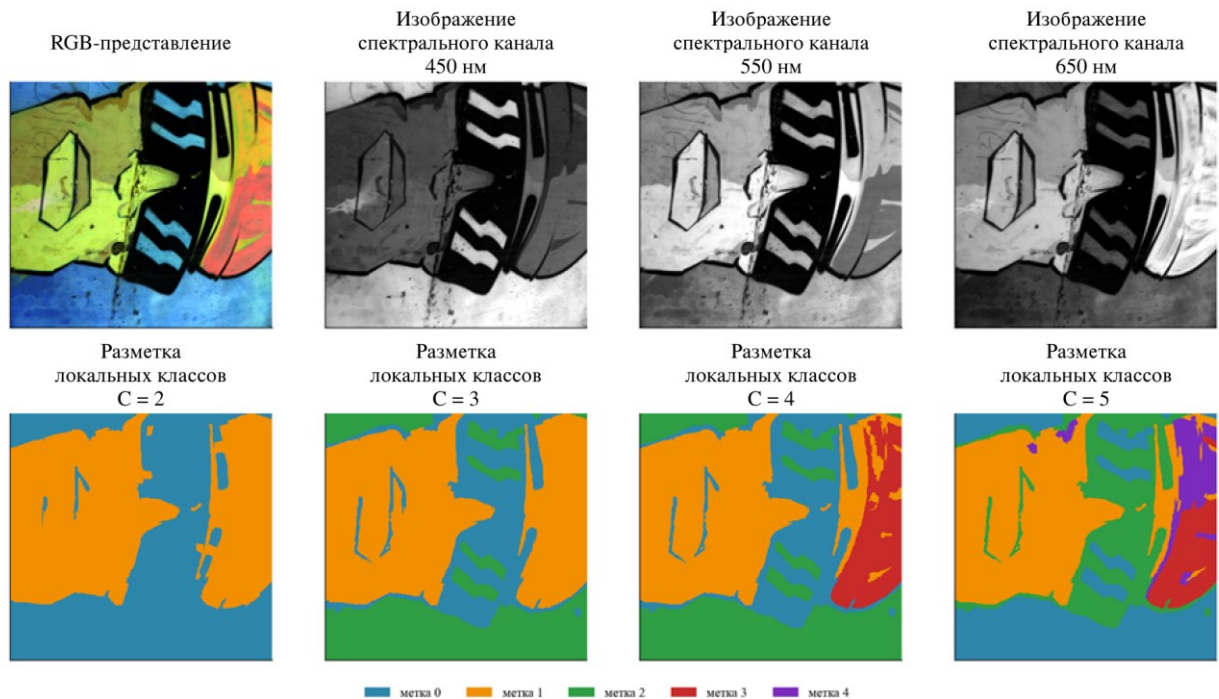


Рисунок 1.17 – Пример исходных данных, использованных при проверке эмпирической аппроксимации снижения общей точности

Сдвиг выполнялся в положительном направлении оси с отсчётами j на $d = 1, \dots, 5$ пикселей. Для каждого d общая точность $OA(d)$ вычислялась поэлементным сопоставлением массивов меток в области перекрытия Ω_d размером $I \times (J - d)$, а её снижение относительно полностью совмещённых

массивов определялось как $\Delta OA(d) = OA(0) - OA(d)$, где $OA(0) = 1$. Для 200 массивов и пяти ненулевых сдвигов получено 1000 значений $\Delta OA(d)$.

Средние экспериментальные значения $\Delta OA(d)$ аппроксимировались зависимостью:

$$\Delta OA_{\text{эмп}}(d) = \frac{2,71 C d}{N}, \quad N = J = 512, \quad (1.22)$$

где C – число локальных спектральных классов, а N – размер массива вдоль направления сдвига; в проведённом эксперименте $N = J = 512$. Значение задаётся в долях единицы; для представления в процентных пунктах правую часть умножают на 100. По первой группе из 35 сцен коэффициент α в зависимости $\Delta OA = \frac{\alpha C d}{N}$ составил 2,71, а 95%-й доверительный интервал – [2,130; 2,939]. На второй группе из 15 сцен средняя абсолютная ошибка для отдельных массивов меток составила 1,89 процентного пункта, после усреднения по каждой паре (C, d) – 0,36 процентного пункта. Аппроксимация воспроизводит усреднённую зависимость (рисунок 1.18), но не предназначена для точного прогнозирования отдельного результата: отклонение зависит от пространственной структуры границ.

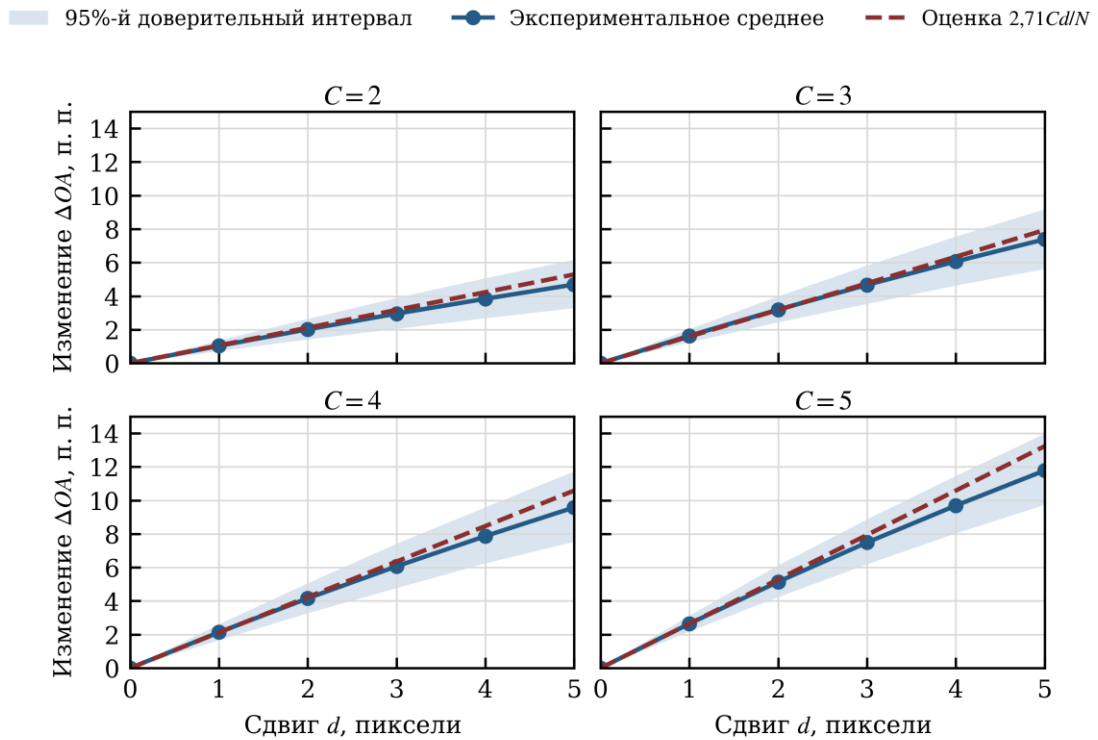


Рисунок 1.18 – Сопоставление экспериментальных средних значений ΔOA с эмпирической аппроксимацией (1.22) при сдвиге массивов меток вдоль оси с отсчётами j ($C = 2, \dots, 5$; $N = J = 512$). Закрашенные области соответствуют 95%-м доверительным интервалам среднего по сценам

1.4 Оценка зарегистрированных и восстановленных ГСИ по общей точности классификации

Цель эксперимента состояла в сравнении ГСИ, полученных различными способами, по результатам поэлементной классификации. Рассматривались три варианта данных: ГСИ, зарегистрированное гиперспектрометром Specim IQ; ГСИ, зарегистрированное разработанным в Самарском университете сканирующим щелевым гиперспектрометром (далее – разработанный гиперспектрометр); ГСИ, восстановленное методом MST++ по RGB-представлению ГСИ Specim IQ.

Первое ГСИ зарегистрировано гиперспектрометром Specim IQ. Данные представлены значениями коэффициента отражения, рассчитанного относительно белого эталона [9], пространственный размер массива составлял 512×512 пикселей, а 204 спектральных канала охватывали диапазон длин волн 397–1003 нм. Среднее спектральное разрешение камеры составляет 7 нм

[82]. Применение Specim IQ для регистрации и классификации спектров растительных объектов рассмотрено в [9], [83].

Второе ГСИ зарегистрировано разработанным гиперспектрометром с оптической схемой Оффнера [84]. Для регистрации сигнала использовалась камера Baumer с матричным сенсором CMV4000. Размер массива составлял $960 \times 2000 \times 252$, спектральный диапазон – 422–978 нм. Нормирование спектрального сигнала выполнялось по результатам регистрации эталонной диффузно отражающей панели Spectralon; после нормирования значения элементов ГСИ представляли коэффициент отражения. На графиках ГСИ разработанного гиперспектрометра обозначено как Baumer.

Третье ГСИ восстановлено методом MST++ по RGB-представлению ГСИ Specim IQ [46]. В работе [85] MST++ используется для восстановления ГСИ по RGB-изображениям при сравнении с методами, сочетающими оптическое кодирование и нейросетевое восстановление. В настоящем эксперименте обучение модели выполнялось на данных набора ARAD_1K, использованного в конкурсе NTIRE 2022 [10]. RGB-представление исходного ГСИ Specim IQ формировалось преобразованием спектров в координаты CIE XYZ [86] и далее в sRGB [87] с использованием стандартного источника D65 [88]. Поскольку спектр освещения при съёмке не измерялся, применение D65 являлось расчётным допущением. Результат MST++ содержал 31 канал диапазона 400–700 нм. Для приближённой оценки коэффициента отражения восстановленный сигнал нормировался на спектр D65 и умножался на единый коэффициент 3,7213, определённый методом наименьших квадратов по 4474 обучающим пикселям исходного ГСИ. Отдельная коррекция спектральных каналов не выполнялась.

Обучение и оценка классификаторов в данном эксперименте проводились в диапазоне длин волн 430–700 нм. Для этого все три ГСИ были приведены к общей сетке длин волн с шагом 10 нм, содержащей 28 каналов. Приведение к общей сетке длин волн запишем в виде:

$$\tilde{S}(i, j, k) = \mathcal{I}_\lambda[S(i, j, :)](\lambda'_k), \quad (1.23)$$

где $\tilde{S}(i, j, k)$ – значение элемента ГСИ на общей спектральной сетке; $S(i, j, :)$ – исходный спектральный вектор в позиции (i, j) , заданный на сетке длин волн λ_ℓ ; λ'_k – длина волны k -го канала общей сетки; $\mathcal{I}_\lambda$ – оператор интерполяции по длине волны. Операция выполняется отдельно для каждого спектра без изменения пространственных координат. Общая сетка выбирается в пределах пересечения спектральных диапазонов сравниваемых ГСИ.

В эталонной разметке выделялись пять классов: живой зелёный лист, зелёная свеча, искусственный зелёный лист из полимерного материала, муляж красного яблока из полимерного материала и красный картон. Для исходного и восстановленного ГСИ использовалась одна эталонная разметка, поскольку MST++ сохранял пространственную сетку исходного ГСИ Specim IQ. Для ГСИ разработанного гиперспектрометра была подготовлена отдельная эталонная разметка тех же классов. RGB-представления данных и разметка приведены на рисунке 1.19.

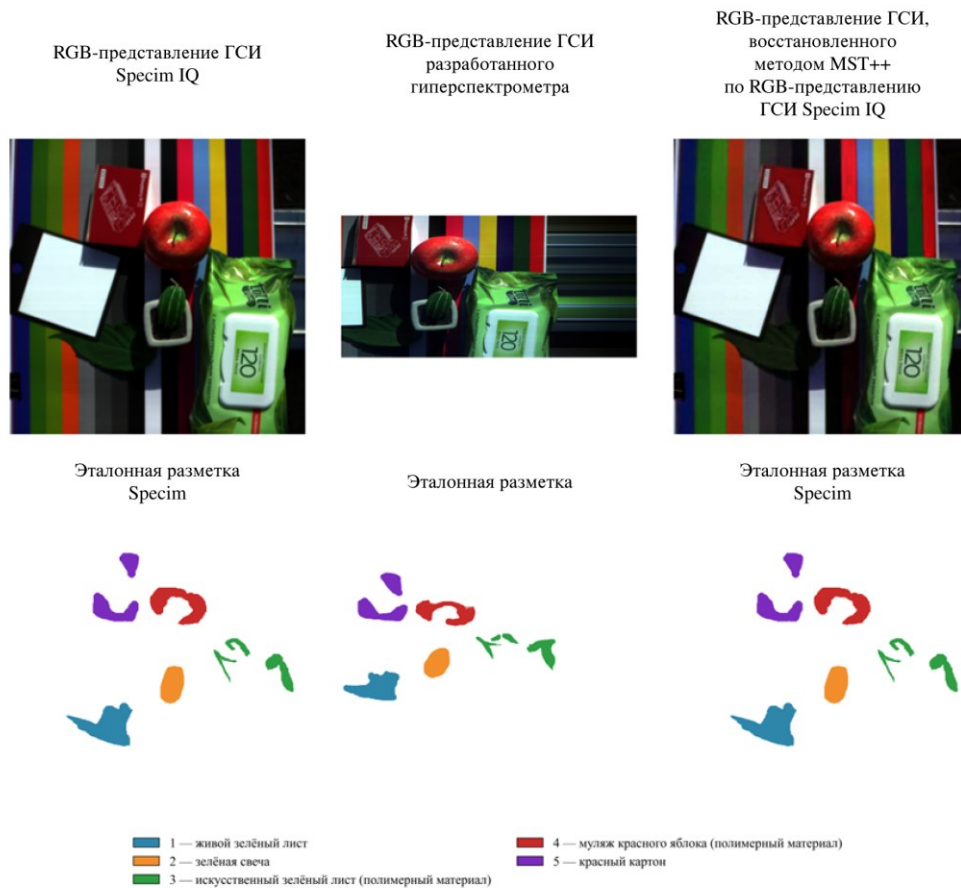


Рисунок 1.19 – RGB-представления и эталонная разметка, использованные при сравнении ГСИ Specim IQ, ГСИ разработанного гиперспектрометра и ГСИ, восстановленного методом MST++

Средние спектры $R(\lambda)$ пяти классов после приведения ГСИ к общей сетке длин волн представлены на рисунке 1.20.

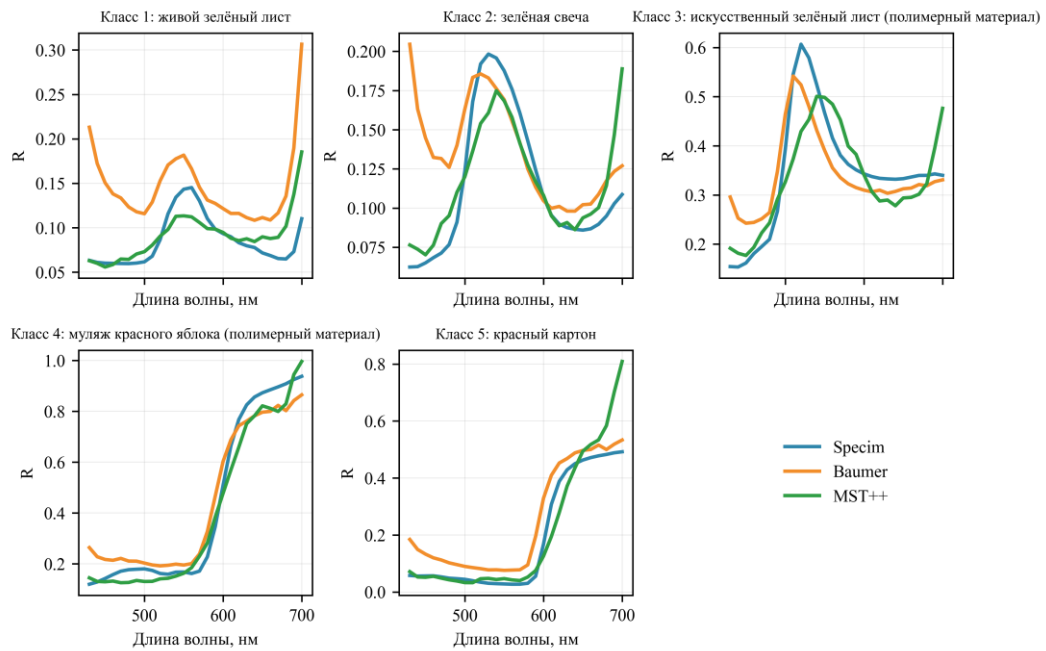


Рисунок 1.20 – Средние спектральные зависимости коэффициента отражения для пяти классов материалов

Размеченные пиксели ГСИ Specim IQ распределялись между обучающей, валидационной и тестовой выборками в соотношении 30 %, 5 % и 65 % соответственно. Для данных разработанного гиперспектрометра, не использовавшихся при обучении классификаторов, валидационная и тестовая выборки составляли приблизительно 10 и 90 % размеченных пикселей. Для восстановленного ГСИ использовались те же пространственные позиции пикселей валидационной и тестовой выборок, что и для исходного ГСИ Specim IQ.

Исследовались одномерная спектральная свёрточная сеть Nu1DCNN [47] и представленные в пункте 1.3.4 модели NM3DCNN и SSFTT. Модели обучались на данных Specim IQ с использованием оптимизатора AdamW [89]. Размер пакета данных составлял 128. Обучение продолжалось не более 100 эпох и прекращалось после 10 последовательных эпох без улучшения валидационного показателя.

Классификаторы, обученные на ГСИ Specim IQ, применялись к ГСИ разработанного гиперспектрометра и восстановленному ГСИ без дообучения. При выборе конфигураций использовались валидационные выборки этих двух

вариантов данных. Изменение функции потерь и общей точности показано на рисунке 1.21.

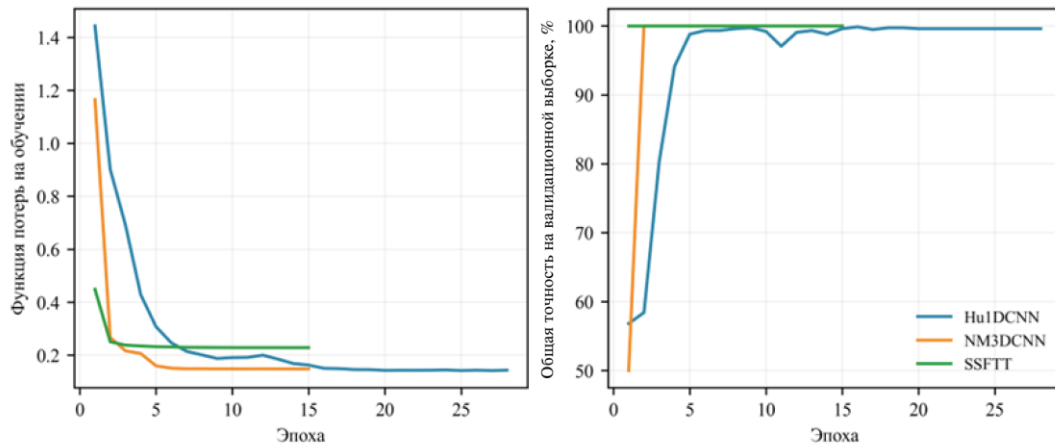


Рисунок 1.21 – Изменение функции потерь и общей точности на валидационной выборке при обучении классификаторов

Общая точность классификации оценивалась на сбалансированных тестовых подвыборках, содержавших по 1388 пикселей каждого класса (по 6940 пикселей для каждого варианта данных). Для Hu1DCNN общая точность составила 94,0 % на исходном ГСИ Specim IQ, 74,5 % на ГСИ разработанного гиперспектрометра и 63,1 % на восстановленном ГСИ. Для NM3DCNN получены значения 95,5, 74,3 и 63,3 %, для SSFTT – 96,4, 81,5 и 69,1 % соответственно. Результаты приведены на рисунке 1.22.

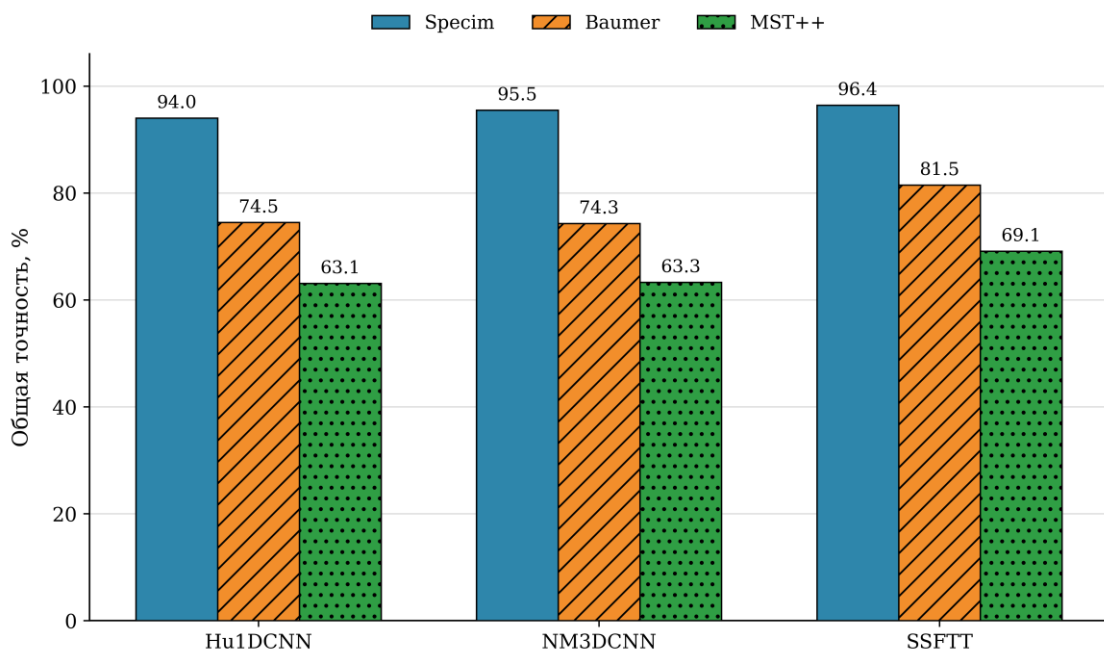


Рисунок 1.22 – Общая точность классификации трёх вариантов ГСИ

Для всех трёх архитектур наибольшая общая точность получена на исходном ГСИ Specim IQ, промежуточная – на данных разработанного гиперспектрометра, наименьшая – на восстановленном ГСИ. Различие между данными разработанного гиперспектрометра и результатом MST++ составило 11,4 процентного пункта для Nu1DCNN, 11,0 процентного пункта для NM3DCNN и 12,4 процентного пункта для SSFTT. Таким образом, при использовании классификаторов, обученных на данных Specim IQ, ГСИ разработанного гиперспектрометра обеспечило более высокую общую точность, чем восстановленное ГСИ.

1.5 Методика оценки связи показателей классификации реальных ГСИ и ГСИ печатных сцен

Эксперимент с линейным сдвигом позволяет проверить, как пространственное рассогласование влияет на сравнительную оценку ГСИ. В рассмотренных условиях ОА сохраняет более высокие значения для ГСИ с меньшими искажениями, тогда как метрики сопоставления с эталоном PSNR, SSIM и SAM могут указывать на более высокое качество сильнее искажённых данных. Полученный результат обосновывает применение ОА для сравнения вариантов получения и обработки ГСИ. Для настройки по печатной калибровочной сцене необходимо дополнительно учитывать различия реальных и печатных спектров. Поэтому далее исследуется согласованность изменений показателей классификации этих двух доменов при зашумлении. Выделяются пары классов со связанными изменениями показателей и формируется библиотека их спектров. На её основе исследуется возможность использовать результаты классификации ГСИ печатных сцен для сравнительной оценки изменений показателей реальных ГСИ.

В данной методике ТКЦЗ используется для выбора параметров по ГСИ печатной калибровочной сцены при заданных классах и фиксированном экспериментальном протоколе. Возможность такого выбора обосновывается в

последующем эксперименте сопоставлением изменений ОА реальных ГСИ и ГСИ печатных сцен.

В работе [27] на реальных ГСИ исследовали влияние отношения сигнал/шум, числа спектральных каналов и пространственного разрешения на поэлементную классификацию земной поверхности методом минимального евклидова расстояния и методом спектрального угла. Коэффициент каппа составлял 0,30–0,99 в зависимости от сцены и классификатора; наибольшее влияние оказало пространственное разрешение. В работе [28] на 27 вариантах зарегистрированных ГСИ сопоставили четыре расчётных показателя качества с результативностью обнаружения целей. Для 67 включённых в анализ сочетаний «изображение–цель» коэффициенты линейной корреляции составили 0,06–0,31.

В работе [29] полезность реальных ГСИ оценивали посредством синтетического внедрения субпиксельных целей и построения ROC-кривых; показано, что она зависит от типа и размера цели, сцены и применяемого детектора.

В работе [90] реальные ГСИ, полученные при бортовой съёмке, и их моделируемые аналоги сравнивали по ковариационной структуре, результатам классификации, разложения спектральных смесей и обнаружения целей. Общая точность классификации составила 95,61 % и 98,18 % соответственно, однако полнота класса искусственных целей – 93,86 % и 55,97 %. В работе [91] классификатор минимального евклидова расстояния, обученный на моделируемых ГСИ, применяли к реальным данным авиационной гиперспектральной съёмки. После преобразования, использующего вручную выбранные соответствия спектров одинаковых материалов, общая точность повысилась с 45 % до 72 %.

Таким образом, в рассмотренных работах поэлементная классификация либо выполнялась только на реальных данных [27], либо использовалась для сопоставления результатов обработки реальных и моделируемых ГСИ [90] и переноса классификатора с моделируемых данных на реальные [91],

остальные исследования были ориентированы на обнаружение целей [28]–[29]. В настоящем исследовании сопоставляются изменения показателей классификации реального ГСИ и ГСИ печатной сцены при добавлении шума.

Для каждого фрагмента реального ГСИ формируется ГСИ печатной сцены. На неискажённых реальном ГСИ и ГСИ печатной сцены отдельно обучаются классификаторы одинаковой архитектуры. Затем к спектрам их тестовых выборок добавляются независимые реализации гауссовского шума одинакового относительного уровня. Оцениваются вызванные шумом изменения полноты и F1-меры классов, а также общей точности и взвешенной F1-меры фрагмента. Правило отбора соответствующих классов настраивается по обучающей и валидационной частям набора фрагментов, фиксируется и проверяется на независимой тестовой части. Проверяется гипотеза о монотонной зависимости между изменениями классификационных показателей реального ГСИ и ГСИ печатной сцены при одинаковом относительном уровне шума.

1.5.1 Используемые данные в эксперименте и формирование ГСИ печатных сцен

В эксперименте использовались 900 опубликованных ГСИ обучающей части набора ARAD_1K, подготовленного в рамках NTIRE 2022 [10]. Каждое изображение имело размер 482×512 пикселей и содержало 31 спектральный канал диапазона 400–700 нм с шагом 10 нм. Из исходных данных сформировано 5382 перекрывающихся фрагмента размером 256×256 пикселей. Исключались фрагменты с недостаточными яркостью, контрастностью, цветовой насыщенностью и спектральной изменчивостью; для уменьшения повторяемости ограничивалось число фрагментов одного исходного ГСИ.

Разметка формировалась автоматически: области выделялись по RGB-представлению с помощью Segment Anything [80], после чего их границы уточнялись по спектральному углу между спектральными векторами пикселей

и средними спектрами классов. Сохранялись фрагменты, содержащие от 3 до 12 классов с достаточным числом пикселей и межклассовой спектральной разделимостью. Полученные области являются локальными спектральными классами внутри фрагмента и не интерпретируются как экспертные тематические классы объектов. После отбора по перечисленным критериям оставлено 399 фрагментов из 305 исходных ГСИ; их распределение приведено в таблице 1.1. Все фрагменты одного исходного ГСИ включались в одну часть набора: обучающую, валидационную или тестовую.

Таблица 1.1 – Состав сформированного набора фрагментов ARAD_1K

Часть набора	Число фрагментов	Число исходных ГСИ
Обучающая	239	183
Валидационная	82	61
Тестовая	78	61
Всего	399	305

Для проверки сопоставимости каждой сцене валидационной части набора назначалась отдельная сцена тестовой части, близкая по среднему спектру, статистикам RGB-представления и процентильным характеристикам уровня сигнала. Средний спектральный угол между подобранными сценами составил $12,11^\circ$, а коэффициент корреляции их средних спектров – 0,766. Эти значения характеризуют близость сцен валидационной и тестовой частей набора. Они не относятся к сопоставлению классификационных показателей реального ГСИ и ГСИ печатной сцены.

Для реального ГСИ и ГСИ печатной сцены использовалась общая эталонная разметка. Классы с одной меткой, занимающие одни и те же пространственные позиции в этих двух массивах, далее называются парой соответствующих классов. Пример реального ГСИ, изображения его спектрального канала, эталонной разметки и ГСИ печатной сцены показан на рисунке 1.23.



Рисунок 1.23 – Фрагмент ГСИ test_frag0052: RGB-представление (а), изображение спектрального канала около 550 нм (б), выделенные локальные классы (в) и RGB-представление соответствующего ГСИ печатной сцены (г)

Формирование ГСИ печатной сцены начиналось с преобразования спектральных векторов 31-канального ГСИ в координаты CIE XYZ [86] и далее в sRGB [87]. Нормированные значения RGB преобразовывались в CMYK с частичной заменой серой составляющей чёрной компонентой [92], коэффициент замены принимался равным 0,85. При моделировании печати CMYK растискивание учитывалось как увеличение эффективной площади покрытия подложки краской. Зависимость эффективной доли площади покрытия $p(u)$ от номинальной доли u задавалась квадратичной функцией:

$$p(u) = u + \gamma u(1 - u), \quad \gamma = 0,56 \quad (1.24)$$

где γ – коэффициент растискивания. При $\gamma = 0,56$ и $u = 0,5$ эффективная доля площади покрытия равна 0,64. По полученным значениям C, M, Y и K с помощью уравнений Демишеля, обобщённых на четыре краски CMYK, [93] вычислялись доли площади w_i шестнадцати первичных сочетаний красок, включая незапечатанную подложку. Коэффициент отражения печатной сцены рассчитывался по спектральной модели Нойгебауэра с поправкой Юла–Нильсена [93-94]:

$$R_{\text{print}}(\lambda) = \left[\sum_{i=1}^{16} w_i R_i(\lambda)^{\frac{1}{n}} \right]^n, \quad n = 7,7 \quad (1.25)$$

где $R_i(\lambda)$ – коэффициент отражения i -го первичного сочетания красок на длине волны λ , w_i – занимаемая им доля площади, n – параметр Юла–Нильсена; в эксперименте $n = 7,7$. Примеры заданных спектров первичных сочетаний и

рассчитанных по средним RGB-значениям спектров классов приведены на рисунке 1.24.

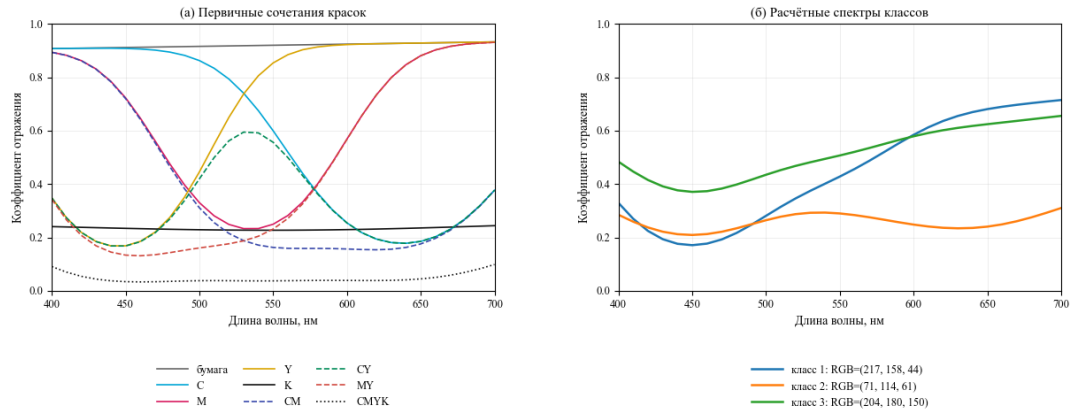


Рисунок 1.24 – Синтетические спектры первичных сочетаний красок (а) и расчётные спектры классов фрагмента test_frag0052 (б)

Для расчёта спектрального сигнала по коэффициенту отражения использовалось табличное относительное спектральное распределение стандартного источника D65 [88]:

$$S_{\text{print}}(\lambda) = \alpha R_{\text{print}}(\lambda) E_{\text{D65}}(\lambda), \quad (1.26)$$

где α – единый масштабный коэффициент. Он определялся только по валидационной части как отношение медиан 99-х перцентилей сигналов реального ГСИ и ГСИ печатной сцены до масштабирования, после чего фиксировался для всех фрагментов; отдельная нормировка каждого фрагмента не выполнялась. Сопоставление средних спектров трёх классов реального ГСИ и ГСИ печатной сцены на рисунке 1.25 показывает, что для одних классов формы спектров близки, а для других заметно различаются. Далее применяется правило отбора пар соответствующих классов по признакам, рассчитанным до внесения шума.

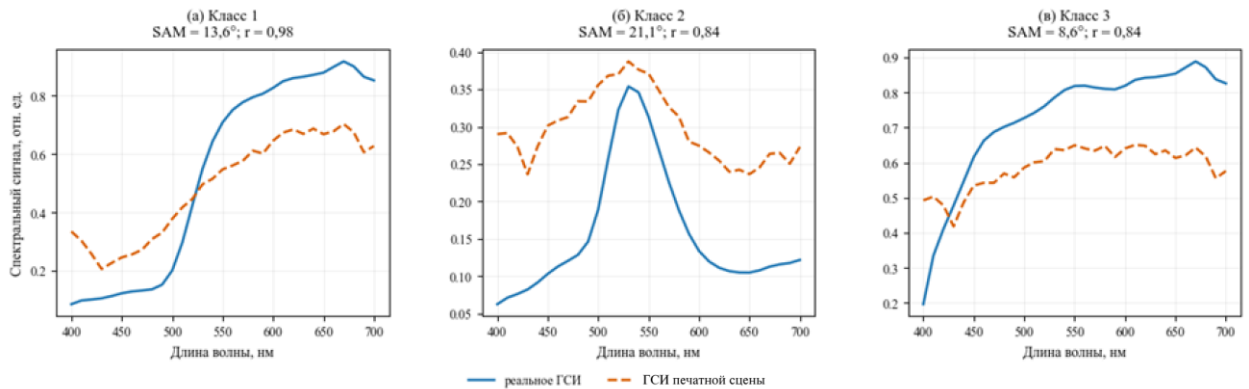


Рисунок 1.25 – Средние спектральные сигналы классов 1–3 фрагмента test_frag0052 для реального ГСИ и ГСИ печатной сцены до внесения шума

Спектры шестнадцати первичных сочетаний задавались вычислительно; ИСС-профиль и измерения конкретного печатного устройства не использовались. ГСИ печатной сцены получено по модели спектрального воспроизведения сцены в системе СМУК. Оно не является результатом съёмки физического отпечатка и не характеризует определённый принтер. В эксперименте проверяется согласованность изменений классификационных показателей реального ГСИ и ГСИ печатной сцены при добавлении шума.

1.5.2 Отбор пар соответствующих классов по изменениям показателей классификации

Для каждого из 399 фрагментов обучались два независимых классификатора одинаковой архитектуры: первый – на неискажённых спектрах реального ГСИ, второй – на спектрах ГСИ печатной сцены. Использовался подход к спектральной классификации на основе одномерной свёрточной сети (1D-CNN) [47]. В данном эксперименте сеть содержала три одномерных свёрточных блока с 32, 64 и 96 выходными каналами, пакетной нормализацией и функцией активации ReLU, слой глобального усреднения и линейный классификационный слой. Внутри каждого фрагмента формировались выборки пикселей для обучения, валидации и тестирования классификаторов. Для каждого класса в обучающую выборку включалось 15 % пикселей, но не менее 20 и не более 300; объёмы валидационной и тестовой выборок ограничивались соответственно 200 и 2500 пикселями.

Обучение выполнялось оптимизатором AdamW [89] при начальном шаге обучения 10^{-3} и коэффициенте уменьшения весов 10^{-4} . Число эпох не превышало 30; досрочная остановка применялась после восьми эпох без повышения макроусреднённой F1-меры – невзвешенного среднего F1-мер по классам [30]. Раздельное обучение двух классификаторов позволяет сопоставить изменения их показателей при добавлении шума. Прямой перенос классификатора с ГСИ печатной сцены на реальное в этом эксперименте не рассматривается.

После фиксации параметров классификаторов к тестовым спектрам добавлялся независимый аддитивный гауссовский шум:

$$S_d^{(\sigma)}(x, \lambda) = S_d(x, \lambda) + \varepsilon(x, \lambda), \quad \varepsilon \sim N\left(0, (\sigma s_f)^2\right) \quad (1.27)$$

где $d \in \{r, p\}$ обозначает реальное ГСИ или ГСИ печатной сцены соответственно, x – пространственная позиция, λ – длина волны, σ – относительный уровень шума. Масштаб s_f определялся по межпроцентильному диапазону значений сигнала фрагмента реального ГСИ f :

$$s_f = P_{99}(S_r, f) - P_1(S_r, f), \quad (1.28)$$

Здесь P_1 и P_{99} – 1-й и 99-й процентиля значений сигнала фрагмента реального ГСИ. Один и тот же масштаб s_f использовался для реального ГСИ и ГСИ печатной сцены, а реализации шума формировались независимо. Параметр σ принимал значения 0,01; 0,03; 0,05; 0,07; 0,10 и 0,12. Влияние шума характеризовалось изменением показателя относительно неискажённых данных:

$$\Delta M_d(\sigma) = M_d(\sigma) - M_d(0), \quad d \in \{r, p\}, \quad (1.29)$$

На уровне классов в качестве M использовались полнота и F1-мера. На уровне фрагмента ГСИ полнота классов усреднялась с весами, пропорциональными числу пикселей этих классов в эталонной разметке тестовой выборки. Полученное значение равно общей точности классификации (ОА) на тестовых пикселях учитываемых классов [30]. F1-мера

усреднялась с теми же весами. Коэффициенты ранговой корреляции Спирмена и Кендалла вычислялись между изменениями показателей реального ГСИ и ГСИ печатной сцены и характеризовали согласованность их ранжирования [95].

Поскольку ГСИ печатной сцены воспроизводит спектры реальной сцены приближённо, для каждой пары соответствующих классов с определялось абсолютное расхождение изменений показателя:

$$e_c(\sigma) = |\Delta M_{r,c}(\sigma) - \Delta M_{p,c}(\sigma)|, \quad (1.30)$$

Правило отбора строилось по признакам, рассчитанным до внесения шума; их группы приведены в таблице 1.2.

Таблица 1.2 – Признаки для отбора пар соответствующих классов реального ГСИ и ГСИ печатной сцены

Группа	Используемые признаки	Назначение
Спектральная близость	SAM, косинус спектрального угла, коэффициент корреляции средних спектров	Сходство формы спектральных зависимостей
Уровень сигнала	Различия средних значений, 99-х перцентилей и межперцентильных диапазонов $P_{99} - P_1$	Сопоставление уровня и межперцентильного диапазона
Объём класса	Число пикселей и доля класса во фрагменте	Проверка достаточности данных для оценки
Показатели до внесения шума	Полнота, F1-мера и их различия для реального ГСИ и ГСИ печатной сцены	Учёт исходных значений полноты и F1-меры

На обучающей части набора фрагментов при $\sigma = 0,12$ по квантильному порогу $e_c(\sigma)$ формировалась метка для обучения правила отбора пар соответствующих классов. Проверялись дерево решений, случайный лес и градиентный бустинг; тип модели, её параметры, квантильный порог и доля сохраняемых пар выбирались по составному критерию, включавшему корреляции Пирсона на обучающей и валидационной частях и штраф за их расхождение. Для полноты применялся второй этап отбора: по валидационной части оценивались устойчивость расхождений между изменениями показателей и направления этих изменений при $\sigma = 0,05; 0,07; 0,10$ и $0,12$.

После выбора параметров правило фиксировалось и без повторной настройки применялось к независимой тестовой части набора фрагментов. Изменения её классификационных показателей не использовались при обучении правила и выборе порога. Отобранные пары соответствующих классов и их средние спектры в реальном ГСИ и ГСИ печатной сцены образовывали библиотеку спектров целевой задачи. При расчёте показателей на уровне фрагмента ГСИ учитывались только классы этой библиотеки.

1.5.3 Связь изменений показателей классификации при добавлении искажений в ГСИ реальных и печатных сцен

Результаты независимой проверки при наибольшем исследованном уровне шума $\sigma = 0,12$ приведены в таблице 1.3. Для каждой пары соответствующих классов сопоставлялись изменения показателя класса реального ГСИ и соответствующего ему класса ГСИ печатной сцены. На уровне фрагментов ГСИ показатели отобранных классов усреднялись с весами, пропорциональными числу пикселей соответствующих классов.

Таблица 1.3 – Ранговая корреляция изменений классификационных показателей реальных ГСИ и ГСИ печатных сцен при уровне шума 12 %

Показатель	Уровень оценки	Число объектов	Спирмен	Кендалл
F1-мера класса	Пары соответствующих классов	100	0,798	0,629
Взвешенная F1-мера	Фрагменты ГСИ	38	0,789	0,619
Полнота класса	Пары соответствующих классов	41	0,924	0,778
ТКЦЗ (ОА)	Фрагменты ГСИ	20	0,889	0,779

Для F1-меры в библиотеку вошли 100 пар соответствующих классов из 38 фрагментов ГСИ. Коэффициент корреляции Спирмена составил 0,798 на уровне пар и 0,789 после взвешенного усреднения по классам каждого фрагмента; соответствующие коэффициенты Кендалла равнялись 0,629 и

0,619. Для полноты отобрана 41 пара соответствующих классов из 20 фрагментов ГСИ. Получены коэффициенты Спирмена 0,924 на уровне пар соответствующих классов и 0,889 для ОА на уровне фрагментов ГСИ; коэффициенты Кендалла составили соответственно 0,778 и 0,779.

Дополнительная проверка устойчивости для F1-меры повысила коэффициент Спирмена на уровне пар соответствующих классов с 0,798 до 0,813, но снизила его после усреднения по классам каждого фрагмента ГСИ с 0,789 до 0,758. Поэтому в итоговом варианте библиотеки для F1-меры сохранено правило, настроенное при уровне шума 12 %. Диаграммы рассеяния изменений показателей приведены на рисунке 1.26.

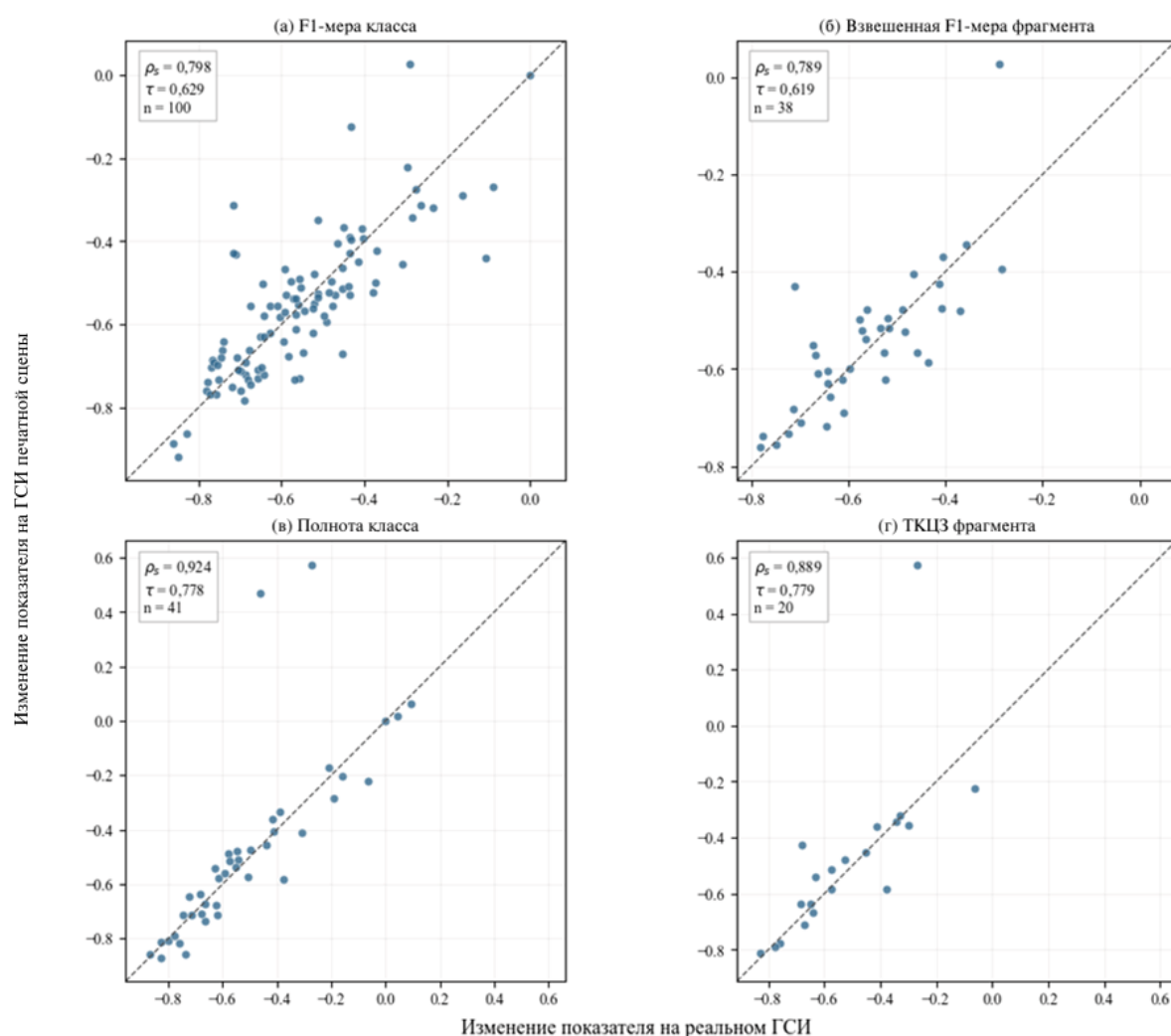


Рисунок 1.26 – Сопоставление изменений классификационных показателей реальных ГСИ и ГСИ печатных сцен при относительном уровне шума 12 %: F1-мера класса (а), взвешенная F1-мера фрагмента ГСИ (б), полнота класса (в) и ОА фрагмента ГСИ (г)

При уровнях шума 10–12 % ранговая корреляция изменений была выше, чем при уровнях 1–7 %. При малом уровне шума абсолютные изменения показателей малы, поэтому их ранжирование может сильнее зависеть от случайных различий в результатах обучения и реализации шума. Полученные коэффициенты относятся к отобранным парам соответствующих классов. Значения корреляции по полной совокупности пар до отбора не представлены. Поэтому результаты таблицы 1.3 характеризуют ранговую связь изменений для отобранных пар, но не позволяют численно оценить прирост корреляции относительно полной совокупности.

Для отобранных пар соответствующих классов выявлена высокая ранговая корреляция изменений классификационных показателей реальных ГСИ и ГСИ печатных сцен. Результат подтверждает согласованность ранжирования по изменениям показателей при добавлении гауссовского шума. Высокая корреляция сама по себе не означает совпадения величин снижения показателей [96]. Область применимости вывода ограничена набором ARAD_1K, диапазоном 400–700 нм, принятой моделью CMYK-печати, аддитивным гауссовским шумом, локальными спектральными классами и зафиксированным протоколом для сети 1D-CNN. Полученная связь свидетельствует о возможности использования ГСИ печатных сцен для предварительной сравнительной оценки изменений общей точности классификации реальных ГСИ для классов сформированной библиотеки в рассмотренных условиях.

1.6 Формирование калибровочной сцены

Установленная согласованность изменений показателей классификации реальных ГСИ и ГСИ печатных сцен позволяет рассмотреть возможность формирования калибровочной сцены для настройки параметров гиперспектральной съёмки и обработки данных. Для этого необходимо подобрать ограниченный набор классов, связанных с классами реальных ГСИ по спектральным характеристикам и изменениям показателей классификации.

Принцип включения в тестовую сцену цветов, характерных для природных объектов, использован при разработке калибровочной шкалы «ColorChecker». В работе [97] предложена шкала, содержащая поля, имитирующие спектральные характеристики кожи человека, листья, голубого неба, а также основные цвета и градации серого. Такой подход служит основанием для подбора состава калибровочной сцены. Связь её классов с классами реальных ГСИ далее проверяется экспериментально.

Вариант калибровочной сцены содержит восемь цветных полос, каждой из которых соответствует отдельный класс. RGB-представление её ГСИ и обозначения классов приведены на рисунке 1.27.

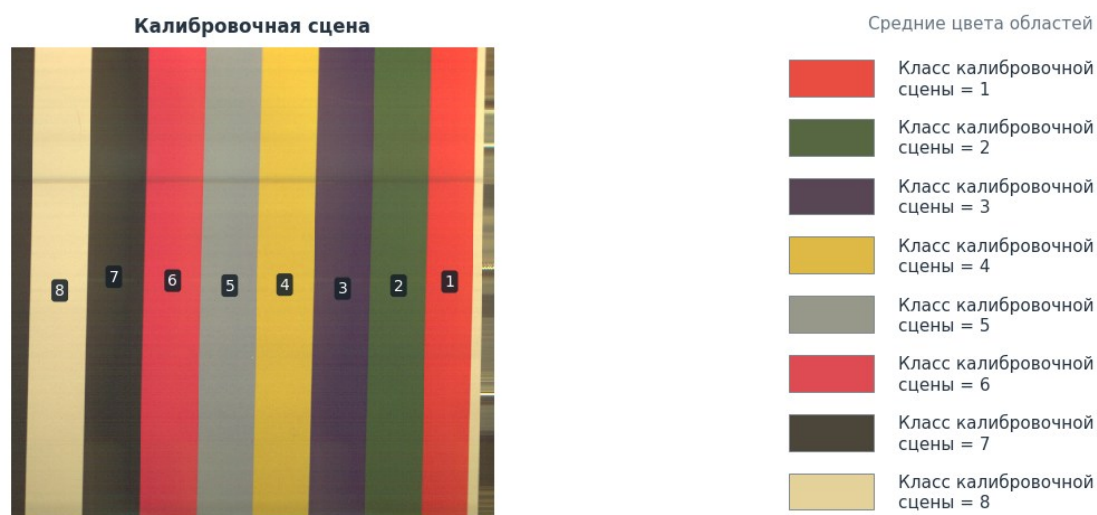


Рисунок 1.27 – RGB-представление ГСИ калибровочной сцены с обозначением классов

Для каждого класса калибровочной сцены из набора NTIRE отбирались спектрально близкие классы. Затем для отобранных пар сопоставлялись изменения показателей классификации при добавлении гауссовского шума. Согласованность оценивалась коэффициентом корреляции Спирмена, рассчитанным по зависимостям показателей классификации от уровня шума. Примеры сопоставления двух пар классов представлены на рисунке 1.28.

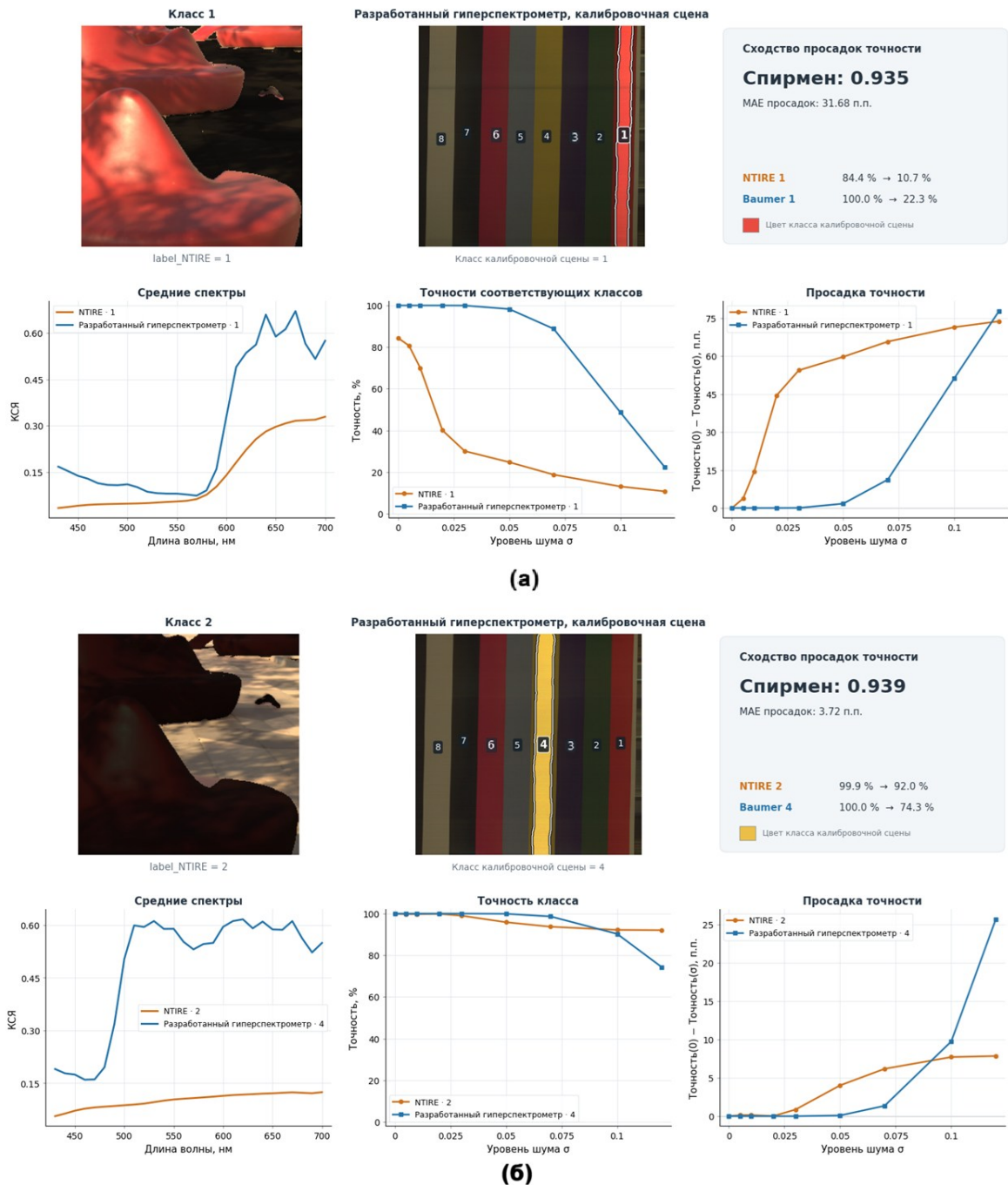


Рисунок 1.28 – Сопоставление изменений показателей классификации двух пар классов калибровочной сцены и реальных ГСИ NTIRE при добавлении гауссовского шума: (а) класс 1 из фрагмента NTIRE 2022, (б) класс 2 из фрагмента NTIRE 2022

На рисунке 1.29 для каждого класса калибровочной сцены приведено число спектрально близких классов NTIRE, для которых коэффициент корреляции Спирмена превышает 0,7. Такое представление позволяет оценить, со сколькими классами реальных ГСИ согласуются изменения показателей классификации каждого класса калибровочной сцены.

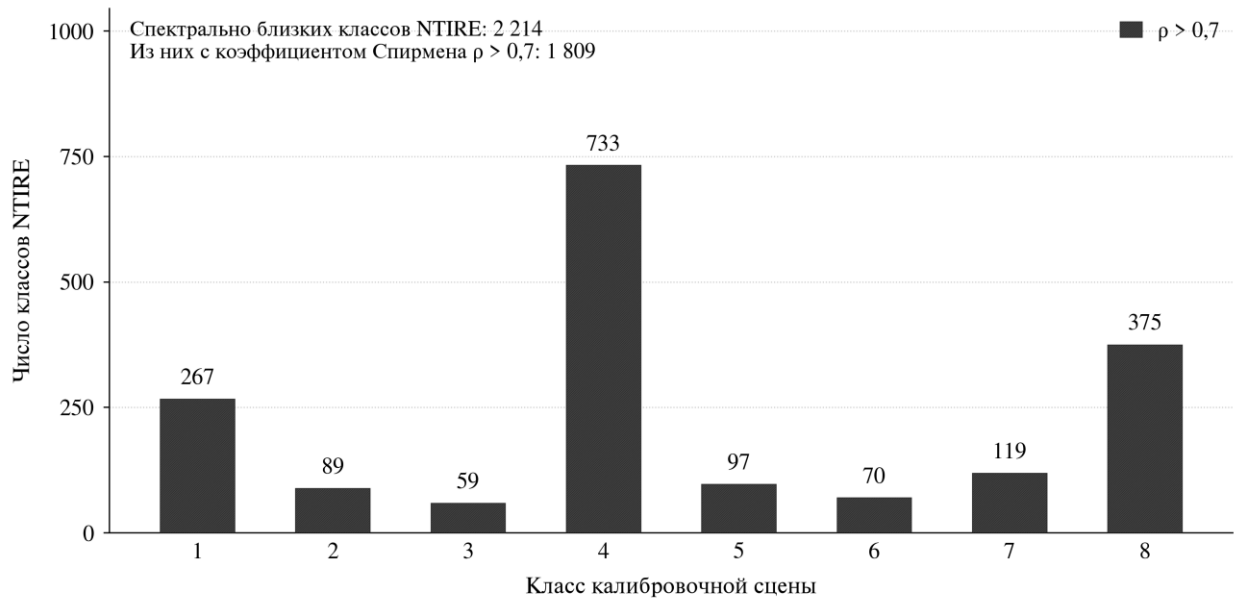


Рисунок 1.29 – Число классов реальных ГСИ NTIRE для каждого класса калибровочной сцены при коэффициенте корреляции Спирмена $\rho > 0,7$

Полученные результаты показывают возможность сопоставления небольшого набора классов калибровочной сцены с множеством классов реальных ГСИ. Это обосновывает использование предложенной сцены в дальнейших экспериментах по настройке параметров съёмки, предобработки и калибровки гиперспектрометров с учётом точности классификации целевой задачи.

Для последующих экспериментов калибровочная сцена дополнена четырьмя градациями серого, вследствие чего число классов увеличено до двенадцати (рисунок 2.3). Такое дополнение позволяет включить в калибровочные данные участки с различными уровнями сигнала. Использование нескольких уровней пропускания нейтральных серых фильтров для определения зависимости цифрового отклика камеры от падающего излучения рассмотрено в работе [98]. В настоящей работе градации серого включены непосредственно в состав сцены, используемой при исследовании методов предобработки и калибровки в главах 2 и 3.

1.7 Выводы к главе 1

1. Получены выражения для пространственных смещений отсчётов при переносе и повороте последовательных проекций щели. Для частного случая линейного сдвига показатели PSNR, SSIM и SAM представлены через средние значения, дисперсии и коэффициенты корреляции сравниваемых данных.

2. В экспериментах на ГСИ сцены ARAD_1K_0930 из набора NTIRE 2022 и ГСИ участка реки Чапаевки при пространственном рассогласовании с эталоном показатели PSNR, SSIM и SAM могли указывать на более высокое качество ГСИ после пространственного размытия и внесения шума по сравнению с зашумлённым ГСИ. При сдвиге предсказанных меток относительно эталонной разметки от 0 до 5 пикселей ОА исходного ГСИ оставалась выше ОА зашумлённого ГСИ, а ОА зашумлённого ГСИ – выше ОА ГСИ после пространственного размытия и внесения шума. В дополнительном эксперименте на сцене ARAD_1K_0927 для 1D-CNN, NM3DCNN и SSFTT в том же диапазоне сдвигов ОА зашумлённого ГСИ также не становилась ниже ОА ГСИ после пространственного размытия и внесения шума.

3. Предложена эмпирическая аппроксимация снижения ОА при сдвиге разметки в зависимости от величины сдвига и числа классов. Для массивов из 15 отложенных сцен ARAD_1K средняя абсолютная ошибка аппроксимации составила 1,89 процентного пункта, а после усреднения по сочетаниям числа классов и величины сдвига – 0,36 процентного пункта.

4. При сопоставлении зарегистрированных и восстановленных ГСИ для всех трёх архитектур классификаторов, обученных на данных Specim IQ, наибольшая ОА получена на ГСИ Specim IQ, промежуточная – на ГСИ разработанного гиперспектрометра, наименьшая – на ГСИ, восстановленном методом MST++ из RGB-представления.

5. Предложена методика сопоставления реальных ГСИ и ГСИ печатных сцен, синтезируемых по RGB-представлениям с использованием модели цветной печати. Для отобранных пар классов установлена согласованность

изменений показателей классификации при добавлении гауссовского шума. При относительном уровне шума 12 % коэффициенты Спирмена составили 0,924 для полноты и 0,798 для F1-меры класса; на уровне фрагментов – 0,889 для ОА и 0,789 для взвешенной F1-меры. Полученные результаты обосновывают возможность предварительной сравнительной оценки изменений точности классификации выбранных классов реальных ГСИ в рассмотренных условиях.

6. Разработана методика формирования и применения печатной калибровочной сцены, предусматривающая отбор классов по спектральной близости к классам реальных ГСИ и согласованности изменений показателей классификации при добавлении гауссовского шума. При сопоставлении с данными NTIRE показано, что одному классу печатной сцены могут соответствовать несколько классов реальных ГСИ с согласованным изменением показателей классификации. На основе методики создана печатная калибровочная сцена для выбора параметров съёмки, предобработки и калибровки гиперспектрометров по точности поэлементной классификации. Разработанная методика применяется при исследовании алгоритмов предобработки и коррекции ГСИ, реализующих предложенный метод повышения точности поэлементной классификации. Во второй главе ГСИ созданной печатной калибровочной сцены используются для выбора параметров съёмки и формирования ГСИ, а также настройки спектральной предобработки. В третьей главе ГСИ этой сцены используются для настройки позиционно-зависимой коррекции отклика гиперспектрометра, учитывающей положение вдоль щели и спектральный канал.

ГЛАВА 2 ВЫБОР ПАРАМЕТРОВ ФОРМИРОВАНИЯ И ОБРАБОТКИ ГИПЕРСПЕКТРАЛЬНЫХ ИЗОБРАЖЕНИЙ С УЧЁТОМ ТОЧНОСТИ ПОЭЛЕМЕНТНОЙ КЛАССИФИКАЦИИ

В первой главе разработана методика формирования и применения печатной калибровочной сцены для выбора параметров получения и обработки ГСИ по точности поэлементной классификации. В настоящей главе эта методика используется при разработке и исследовании алгоритмов спектральной предобработки, реализующих предложенный метод повышения точности классификации. Настройка предобработки выполняется по ГСИ печатной калибровочной сцены, сформированной в подразделе 1.6, а её применение оценивается на отдельной тестовой сцене.

При выборе параметров формирования и обработки ГСИ необходимо учитывать их влияние на точность поэлементной классификации. Время экспозиции и параметр усиления сигнала определяют уровень регистрируемого сигнала и шума; предобработка и сжатие также могут изменять данные, используемые классификатором.

В работе [99] оценено влияние нейросетевого шумоподавления ГСИ на последующую классификацию. Для Indian Pines применение предложенной свёрточной нейронной сети повысило общую точность классификации с 75,9 до 85,4 %. Авторы сопоставили результаты классификации исходного изображения и изображений после применения различных методов шумоподавления.

В работе [70] исследовано влияние сокращения объёма гиперспектральных данных на классификацию. Методы снижения спектральной размерности сопоставлены по ошибке восстановления спектров и точности классификации. Меньшая ошибка восстановления не всегда соответствовала более высокой точности классификации. В работе [100] исследовано сжатие с потерями, допускающее отклонение восстановленных данных от исходных: JPEG для цветных изображений и H.264 для видео.

Влияние параметров сжатия на результаты обработки различалось в зависимости от задачи и использования сжатых данных при обучении нейронной сети.

В работе [101] для исследования шумоподавления при малом времени экспозиции сформирован набор парных ГСИ: опорное изображение регистрировалось с длительной экспозицией, а зашумлённое – с экспозицией, уменьшенной в 50 раз. В работе [102] рассмотрена зависимость шума цифровой камеры от времени экспозиции и усиления. В использованной модели сокращение времени экспозиции с увеличением усиления сохраняет средний цифровой отсчёт при снижении отношения сигнал/шум. Отдельно исследовано изменение усиления при фиксированном времени экспозиции.

Цель настоящей главы – исследовать влияние параметров формирования и обработки ГСИ на точность поэлементной классификации. В разделах 2.1 и 2.2 рассматриваются обучаемый модуль спектральной предобработки, выполняющий спектральное сглаживание и вычитание программной поправки уровня чёрного (см. пункт 2.1.2), и результаты его применения для классификации ГСИ. Исследуются выбор длины окна фильтра Савицкого–Голея и веса спектральной составляющей комбинированного критерия, перенос настроенной предобработки на другие классификаторы и вычислительные затраты её применения. В разделе 2.3 исследуется влияние выбора видеокодека и параметров сжатия ГСИ на точность поэлементной классификации. В разделе 2.4 оценивается точность классификации в двух сериях вычислительных экспериментов: при сокращении времени экспозиции с одновременным увеличением усиления и при изменении усиления в условиях недостаточной экспозиции. Во втором эксперименте время экспозиции оставалось фиксированным. Обсуждение полученных результатов приведено в разделе 2.5.

2.1 Настройка спектральной предобработки ГСИ

Настройка модуля спектральной предобработки заключается в совместном определении коэффициентов спектральной свёртки и программной поправки уровня чёрного. Для этого рассматриваются два способа: нейросетевая настройка с использованием классификатора, предварительно обученного на опорном ГСИ Specim IQ, и аналитическое определение параметров методом гребневой регрессии.

Авторский результат данного раздела состоит в разработке алгоритмов совместной настройки коэффициентов спектральной свёртки и программной поправки уровня чёрного по ГСИ печатной калибровочной сцены, в том числе с учётом классификационного и комбинированного критериев. Для их реализации используются известные математические средства: свёртка, функция softmax, перекрёстная энтропия и среднеквадратичная ошибка [103], фильтр Савицкого–Голея [104, 105] и гребневая регрессия [106, 107]. Ниже эти средства описаны применительно к предлагаемым алгоритмам.

2.1.1 Постановка задачи настройки спектральной предобработки

Используются два ГСИ одной калибровочной сцены, зарегистрированные Specim IQ и разработанным гиперспектрометром. Первое принимается за опорное, второе – за сравниваемое. Цифровые отсчёты сравниваемого ГСИ и спектр белого эталона интерполируются на сетку длин волн Specim IQ. Пространственное совмещение ГСИ не выполняется: средние спектральные векторы соответствующих классов сопоставляются по эталонной разметке, подготовленной отдельно для каждого ГСИ.

Значения коэффициента отражения опорного ГСИ на общей сетке длин волн (см. подраздел 1.4, формулу (1.23)) обозначим:

$$S_{IQ}^*(i, j, k), 1 \leq k \leq K^*, \quad (2.1)$$

где k – номер спектрального канала; K^* – число каналов общей сетки. Сравнимое ГСИ на этой сетке представлено цифровыми отсчётами:

$$DN_B(i, j, k), 1 \leq k \leq K^*. \quad (2.2)$$

Оператор предобработки P_θ с набором настраиваемых параметров θ преобразует сравнимое ГСИ:

$$\hat{R}_B^\theta(i, j, k) = P_\theta(DN_B)(i, j, k), \quad (2.3)$$

где $\hat{R}_B^\theta(i, j, k)$ – значение элемента ГСИ после предобработки. Оператор P_θ включает спектральную свёртку, вычитание программной поправки уровня чёрного, замену отрицательных значений нулём и нормирование по фиксированному спектру белого эталона.

Результат предобработки поступает на вход классификатора F_φ , обученного на опорном ГСИ. При фиксированных параметрах φ предсказанная метка класса определяется выражением:

$$\hat{Y}_B^\theta(i, j) = \arg \max_{1 \leq c \leq C} [F_\varphi(P_\theta(DN_B))(i, j)]_c, \quad (2.4)$$

где $\hat{Y}_B^\theta(i, j)$ – предсказанная метка класса пикселя с пространственными отсчётами (i, j) ; c – номер класса; C – число классов.

При настройке предобработки изменяются только параметры θ . Их значения определяются минимизацией выбранной функции потерь:

$$\theta^* = \arg \min_{\theta} \mathcal{L}(\theta), \quad (2.5)$$

где $L(\theta)$ – функция потерь; θ^* – найденный набор параметров. Критерии настройки предобработки определены в подразделе 2.1.5.

В варианте с фильтром Савицкого–Голея коэффициенты сглаживания фиксированы, а программная поправка уровня чёрного d_0 принимается равной нулю. Схема настройки и применения предобработки представлена на рисунке 2.1.

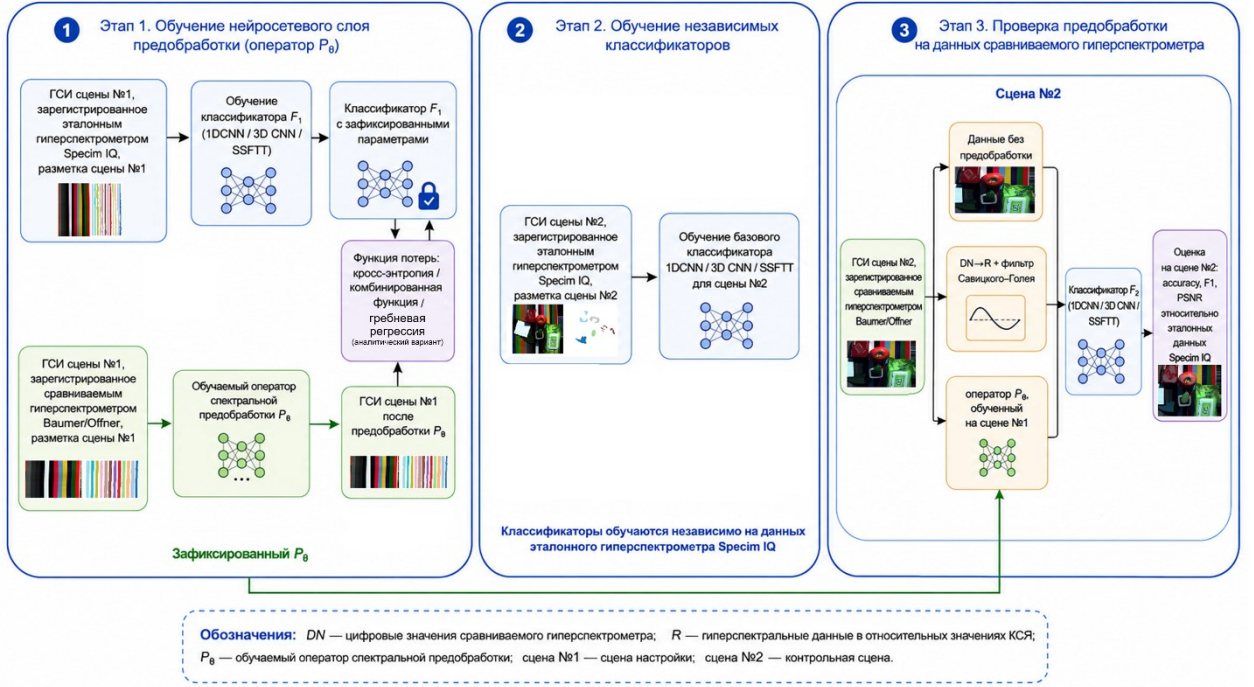


Рисунок 2.1 – Схема настройки и применения спектральной предобработки ГСИ

2.1.2 Программная поправка уровня чёрного и нормирование по белому эталону

В руководстве Baumer параметр уровня чёрного (Black Level) определяет смещение, вносимое в видеосигнал камеры [108]. При регистрации сравниваемого ГСИ аппаратная настройка уровня чёрного фиксирована. В предобработке используется отдельная программная поправка d_0 , вычитаемая из цифровых отсчётов после спектральной свёртки.

Результат спектральной свёртки обозначим:

$$\widetilde{DN}_B(i, j, k) = G(DN_B)(i, j, k), \quad (2.6)$$

где G – оператор свёртки по спектральной оси.

После вычитания программной поправки уровня чёрного и нормирования по белому эталону получаем:

$$R_B(i, j, k; d_0) = \frac{\max(\widetilde{DN}_B(i, j, k) - d_0, 0)}{W_B(k) + \varepsilon}, \quad (2.7)$$

где $W_B(k)$ – значение сигнала белого эталона в k -м канале общей сетки длин волн; ε – малая положительная величина, исключающая деление на нуль.

Спектр белого эталона фиксируется до настройки предобработки. Поправка d_0 одинакова для всех пространственных отсчётов и спектральных каналов.

2.1.3 Спектральное сглаживание фильтром Савицкого–Голея

Фильтр Савицкого–Голея основан на локальной аппроксимации данных полиномом методом наименьших квадратов. Сглаженное значение соответствует значению полинома в центре окна. Это преобразование представляется свёрткой с коэффициентами, определяемыми длиной окна и степенью полинома [104].

Для спектрального вектора при фиксированных пространственных отсчётах (i, j) преобразование имеет вид:

$$\widetilde{DN}_B(i, j, k) = \sum_{t=-r}^r c_t DN_B(i, j, k + t), \quad (2.8)$$

где c_t – коэффициент фильтра для смещения t по спектральной оси; r – число соседних каналов с каждой стороны центрального. Длина окна составляет $2r + 1$ каналов. Формула записана для каналов $r + 1 \leq k \leq K^* - r$, в которых окно целиком находится внутри спектрального диапазона.

Для внутренних каналов сглаживание выполняется независимо для каждого спектрального вектора с одинаковыми фиксированными коэффициентами. Для крайних каналов сглаженные значения вычисляются по полиному, аппроксимирующему значения ближайшего полного окна спектра [105]. Затем применяется преобразование (2.7).

2.1.4 Обучаемый модуль спектральной предобработки

В обучаемом модуле коэффициенты спектральной свёртки β и программная поправка уровня чёрного d_0 определяются по калибровочной сцене. Полный оператор предобработки представим как последовательность преобразований:

$$P_\theta(DN_B) = R_{B,d_0}(G_\beta(DN_B)), \quad (2.9)$$

где G_β – оператор спектральной свёртки с настраиваемыми коэффициентами β ; R_{B,d_0} – преобразование (2.7), включающее вычитание d_0 , замену

оценивается по ОА на размеченных пикселях калибровочной сцены при заданном экспериментальном протоколе.

Пусть Ω – множество размеченных пикселей сравниваемого ГСИ, используемых при настройке предобработки, а $Y_B(i, j)$ – метка класса пикселя $(i, j) \in \Omega$. После предобработки классификатор формирует вектор выходных значений:

$$z^\theta(i, j) = F_\phi(P_\theta(DN_B))(i, j), \quad (2.13)$$

где $z^\theta(i, j)$ – вектор выходных значений классификатора до преобразования в вероятности классов.

Вероятности классов определяются функцией softmax, нормирующей экспоненты выходных значений классификатора [103]:

$$p_c^\theta(i, j) = \frac{\exp(z_c^\theta(i, j))}{\sum_{q=1}^C \exp(z_q^\theta(i, j))}, c = 1, \dots, C, \quad (2.14)$$

где $p_c^\theta(i, j)$ – оценка вероятности принадлежности пикселя (i, j) классу c ; $z_c^\theta(i, j)$ – соответствующее выходное значение классификатора; q – индекс суммирования по классам. Категориальная перекрёстная энтропия имеет вид:

$$\mathcal{L}_{CE}(\theta) = -\frac{1}{|\Omega|} \sum_{(i, j) \in \Omega} \log(p_{Y_B(i, j)}^\theta(i, j)), \quad (2.15)$$

где $|\Omega|$ – число пикселей, участвующих в расчёте; $p_{Y_B(i, j)}^\theta(i, j)$ – вероятность класса, заданного эталонной разметкой.

Спектральное слагаемое вычисляется по средним спектральным векторам соответствующих классов. Для класса c опорного ГСИ:

$$\mu_{IQ, c}(k) = \frac{1}{|\Omega_c^{IQ}|} \sum_{(i, j) \in \Omega_c^{IQ}} S_{IQ}^*(i, j, k), k = 1, \dots, K^*, \quad (2.16)$$

где Ω_c^{IQ} – множество пикселей класса c в опорном ГСИ; $|\Omega_c^{IQ}|$ – число этих пикселей; $\mu_{IQ, c}(k)$ – значение k -го канала среднего спектрального вектора.

Для сравниваемого ГСИ после предобработки:

$$\mu_{B,c}^{\theta}(k) = \frac{1}{|\Omega_{B,c}|} \sum_{(i,j) \in \Omega_{B,c}} \hat{R}_B^{\theta}(i,j,k), k = 1, \dots, K^*, \quad (2.17)$$

где $\Omega_{B,c}$ – множество пикселей класса c в сравниваемом ГСИ; $|\Omega_{B,c}|$ – число этих пикселей; $\mu_{B,c}^{\theta}(k)$ – значение k -го канала среднего спектрального вектора после предобработки. Множества Ω_c^{IQ} и $\Omega_{B,c}$ задаются разметкой соответствующих ГСИ.

Средний квадрат ошибки между средними спектральными векторами классов определяется выражением:

$$\mathcal{L}_{MSE}(\theta) = \frac{1}{|\mathcal{C}|K^*} \sum_{c \in \mathcal{C}} \sum_{k=1}^{K^*} (\mu_{B,c}^{\theta}(k) - \mu_{IQ,c}(k))^2, \quad (2.18)$$

где $\mathcal{C}$ – множество сопоставляемых классов; $|\mathcal{C}|$ – число классов в этом множестве. Критерий $\mathcal{L}_{MSE}(\theta)$ вычисляется непосредственно по преобразованным спектральным векторам.

Комбинированная функция потерь объединяет классификационное и спектральное слагаемые:

$$\mathcal{L}_{CE_MSE}(\theta) = \mathcal{L}_{CE}(\theta) + \lambda \mathcal{L}_{MSE}(\theta), \quad (2.19)$$

где λ – коэффициент, определяющий вклад спектрального слагаемого.

Для аналитического определения коэффициентов спектральной свёртки и программной поправки уровня чёрного используется метод гребневой регрессии. Сумма квадратов ошибок линейной модели дополняется квадратичным штрафом на её коэффициенты. В работе [106] изложен общий принцип добавления регуляризующего слагаемого. В работе [107] гребневая регрессия используется для определения коэффициентов спектрального преобразования при восстановлении спектров.

Преобразование (2.12) без замены отрицательных значений нулём линейно по параметрам β и d_0 . Через $\mu_{B,c}^{DN}(k)$ обозначим среднее значение цифровых отсчётов класса c в k -м канале сравниваемого ГСИ калибровочной сцены, рассчитанное по эталонной разметке. Опорные значения $\mu_{IQ,c}(k)$

определены в (2.16). Для внутренних каналов $r + 1 \leq k \leq K^* - r$ линейная модель среднего спектрального вектора имеет вид:

$$\tilde{\mu}_{B,c}^{\theta}(k) = \frac{\sum_{t=-r}^r \beta_t \mu_{B,c}^{\text{DN}}(k+t) - d_0}{W_B(k) + \varepsilon}. \quad (2.20)$$

Для матричной записи набор параметров (2.10) представим вектором $\theta = (\beta_{-r}, \dots, \beta_r, d_0)^T$. Каждой паре (c, k) соответствует строка матрицы A , содержащая коэффициенты при компонентах θ в (2.20). Элементы вектора y равны соответствующим опорным значениям $\mu_{\text{IQ},c}(k)$. Число сопоставляемых значений составляет $N = |\mathcal{C}|(K^* - 2r)$. Критерий гребневой регрессии запишем как:

$$\theta_R = \arg \min_{\theta} \left[\frac{1}{N} \|A\theta - y\|_2^2 + \eta \|\beta\|_2^2 \right], \quad (2.21)$$

где $\eta > 0$ – параметр регуляризации. В рассматриваемой постановке штраф накладывается на коэффициенты свёртки β , а поправка d_0 определяется без штрафа. Введём диагональную матрицу $D = \text{diag}(1, \dots, 1, 0)$, последний диагональный элемент которой соответствует d_0 . Из условий минимума критерия (2.21) получаем:

$$\theta_R = (A^T A + N\eta D)^{-1} A^T y. \quad (2.22)$$

Полученная запись учитывает исключение d_0 из штрафа; в работе [107] приведено стандартное решение со штрафом на все коэффициенты. Найденные коэффициенты используются в полном преобразовании (2.12). Классификатор в аналитическом определении θ_R не участвует. Критерий (2.21) относится к линейной модели (2.20), тогда как спектральное слагаемое (2.18) рассчитывается после полного преобразования, включающего замену отрицательных значений нулём.

2.2 Нейросетевая настройка спектральной предобработки

2.2.1 ГСИ калибровочной и тестовой сцен

Для калибровочной и тестовой сцен использовались по два ГСИ, зарегистрированные гиперспектрометром Specim IQ и разработанным

гиперспектрометром на основе оптической схемы Оффнера. ГСИ Specim IQ принимались за опорные, а ГСИ разработанного гиперспектрометра – за сравниваемые.

Калибровочная сцена сформирована по методике подраздела 1.6 и использовалась для определения параметров предобработки, тестовая – для оценки переноса предобработки между сценами. RGB-представления ГСИ обеих сцен и соответствующая эталонная разметка приведены на рисунках 2.3–2.6.

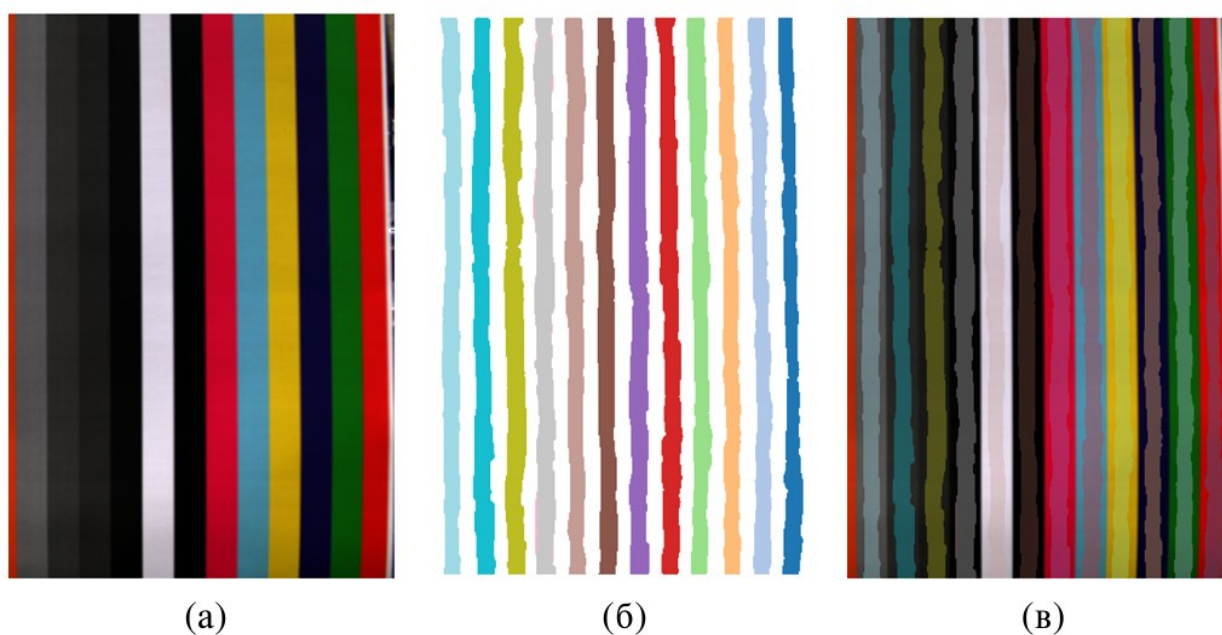


Рисунок 2.3 – Калибровочная сцена по данным Specim IQ: RGB-представление ГСИ (а), эталонная разметка (б), наложение разметки на RGB-представление ГСИ (в)

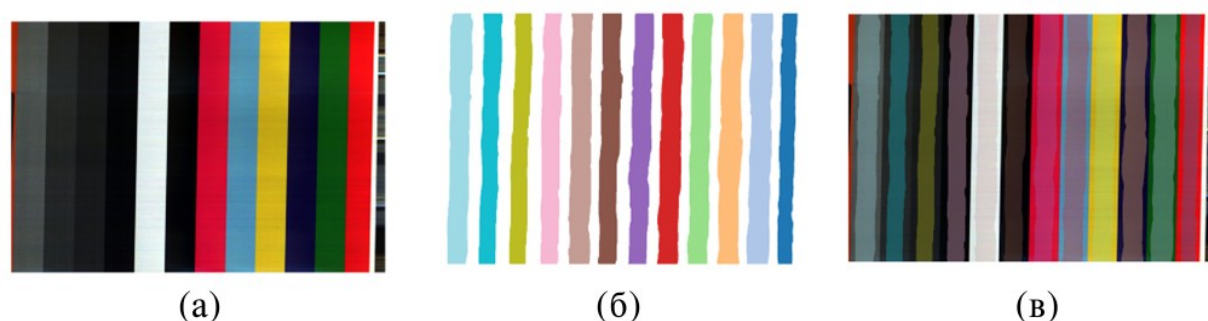


Рисунок 2.4 – Калибровочная сцена по данным разработанного гиперспектрометра: RGB-представление ГСИ (а), эталонная разметка (б), наложение разметки на RGB-представление ГСИ (в)

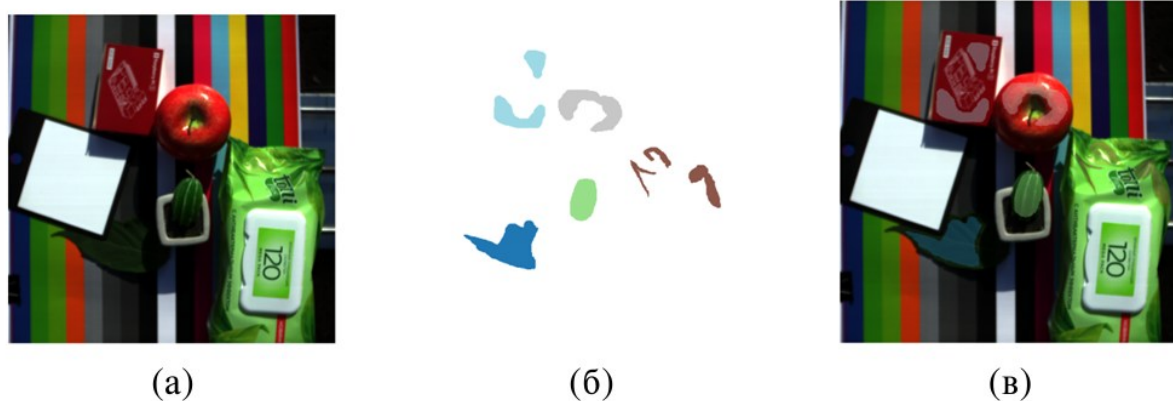


Рисунок 2.5 – Тестовая сцена по данным Specim IQ: RGB-представление ГСИ (а), эталонная разметка (б), наложение разметки на RGB-представление ГСИ (в)

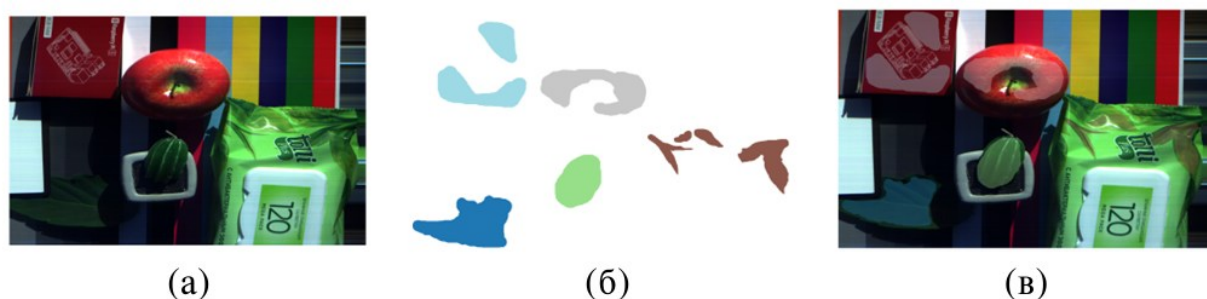


Рисунок 2.6 – Тестовая сцена по данным разработанного гиперспектрометра: RGB-представление ГСИ (а), эталонная разметка (б), наложение разметки на RGB-представление ГСИ (в)

Для каждого ГСИ была подготовлена отдельная эталонная разметка. В калибровочной сцене метки классов 1–12 соответствовали вертикальным полосам цветовой шкалы. Тестовая сцена описана в подразделе 1.4; в ней выделены пять классов:

- 1) живой зелёный лист;
- 2) зелёная свеча;
- 3) искусственный зелёный лист из полимерного материала;
- 4) муляж красного яблока из полимерного материала;
- 5) красный картон.

Пиксели с меткой 0, соответствующей фону, исключались из обучения и расчёта функций потерь. При формировании RGB-представлений ГСИ использовались спектральные каналы с длинами волн 440, 531 и 638 нм для

Specim IQ и 440, 530 и 640 нм для разработанного гиперспектретра, соответствующие синему, зелёному и красному каналам RGB-представления.

Спектральный диапазон обоих ГСИ ограничивался интервалом 450–950 нм. Цифровые отсчёты сравниваемого ГСИ и спектр белого эталона линейно интерполировались на сетку длин волн Specim IQ. Пространственное совмещение ГСИ не проводилось; средние спектральные векторы соответствующих классов рассчитывались по эталонной разметке каждого ГСИ.

2.2.2 Нейросетевые классификаторы и варианты предобработки

Для поэлементной классификации использовались одномерная спектральная свёрточная сеть (Hu1DCNN, далее – 1DCNN), предложенная в работе [47], модифицированная многомасштабная трёхмерная свёрточная сеть NM3DCNN [109] и пространственно-спектральный трансформер SSFTT [49]. Входом 1DCNN служил спектральный вектор отдельного пикселя, а NM3DCNN и SSFTT – локальный пространственно-спектральный фрагмент ГСИ. В SSFTT выделенные свёрточными слоями признаки обрабатываются блоком трансформера.

Схема применения обучаемого модуля спектральной предобработки перед классификатором приведена на рисунке 2.7.

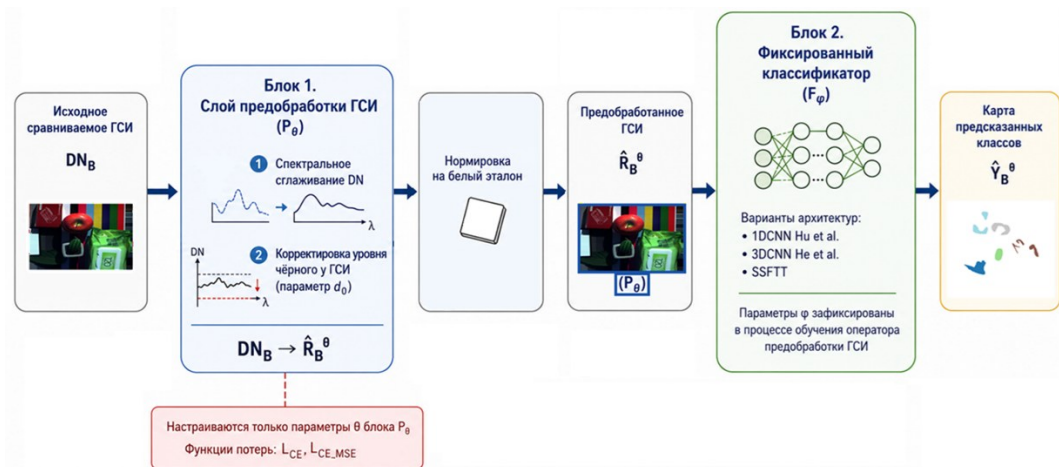


Рисунок 2.7 – Применение обучаемого модуля спектральной предобработки перед классификатором F_φ

Обозначим через F_1 классификатор, используемый при нейросетевой настройке предобработки на калибровочной сцене, а через F_2 – классификатор для оценки её применения на тестовой сцене. Рассмотрены пять сочетаний архитектур F_1 и F_2 , приведённых в таблице 2.1.

Таблица 2.1 – Сочетания архитектур классификаторов F_1 и F_2

Классификатор F_1	Классификатор F_2	Назначение
1DCNN	1DCNN	Настройка и оценка предобработки с 1DCNN
1DCNN	NM3DCNN	Применение предобработки с другой свёрточной архитектурой
1DCNN	SSFTT	Применение предобработки с архитектурой SSFTT
NM3DCNN	NM3DCNN	Настройка и оценка предобработки с NM3DCNN
SSFTT	SSFTT	Настройка и оценка предобработки с SSFTT

В обозначении сочетания первой указывается архитектура F_1 , второй – архитектура F_2 . Наряду с совпадающими архитектурами рассмотрены сочетания 1DCNN–NM3DCNN и 1DCNN–SSFTT, в которых предобработка, настроенная с использованием спектрального классификатора, применялась перед пространственно-спектральным классификатором.

2.2.3 Протокол обучения классификаторов и настройки предобработки

При нейросетевой настройке классификатор F_1 сначала обучался на опорном ГСИ калибровочной сцены. Затем при фиксированных параметрах F_1 по сравниваемому ГСИ этой сцены совместно определялись коэффициенты спектральной свёртки и программная поправка уровня чёрного. Градиент

функции потерь по параметрам предобработки вычислялся через классификатор при фиксированных параметрах слоёв классификатора.

При аналитическом определении параметров использовалась гребневая регрессия по средним спектральным векторам классов калибровочной сцены в соответствии с (2.20)–(2.22). Параметр регуляризации η задавался равным 0,01.

Классификатор F_2 независимо обучался на обучающей выборке опорного ГСИ тестовой сцены без оператора P_θ . Оператор P_θ , настроенный на калибровочной сцене, применялся к сравниваемому ГСИ тестовой сцены перед классификатором F_2 без дополнительной настройки. Сравнимое ГСИ тестовой сцены не использовалось для определения параметров P_θ и F_2 . Для каждого класса опорного ГСИ тестовой сцены доли обучающей, валидационной и тестовой выборок составляли 10, 5 и 85 % соответственно.

Доли выборок калибровочной сцены и параметры спектральных преобразований приведены в таблице 2.2. Для обучения классификаторов и нейросетевой настройки предобработки использовались методы оптимизации Adam [110] и AdamW [89]. Параметры основной серии, зависящие от архитектуры, приведены в таблицах 2.3–2.4. В дополнительных сериях выбора окна фильтра Савицкого–Голея и коэффициента λ использовались отдельные параметры обучения, описанные в подразделах 2.2.4 и 2.2.5 соответственно.

Для варианта без дополнительной предобработки в таблицах 2.6–2.10 сохранены результаты одного запуска основной серии. При нейросетевой настройке по L_{CE} выполнено по десять запусков с использованием базовых значений генератора псевдослучайных чисел (seed) 43, 123 и 2026. Десять запусков выполнялись последовательно в одной вычислительной сессии без перезапуска ядра; указанное число является суммарным для каждого варианта. Результаты фильтра Савицкого–Голея и комбинированного критерия получены в дополнительных сериях, также включавших по десять запусков для каждого варианта. Для этих вариантов приведены средние значения и

выборочные стандартные отклонения (SD). Для гребневой регрессии приведены значения показателей без оценки разброса.

Таблица 2.2 – Общие параметры эксперимента

Название параметра	Значение
Доля обучающей выборки калибровочной сцены для каждого класса	10%
Доля валидационной выборки калибровочной сцены для каждого класса	5%
Доля тестовой выборки калибровочной сцены для каждого класса	85%
Длина окна фильтра Савицкого–Голея, каналов	5
Степень аппроксимирующего полинома	2
Длина ядра обучаемой спектральной свёртки, каналов	7
Начальное значение d_0 при нейросетевой настройке	0

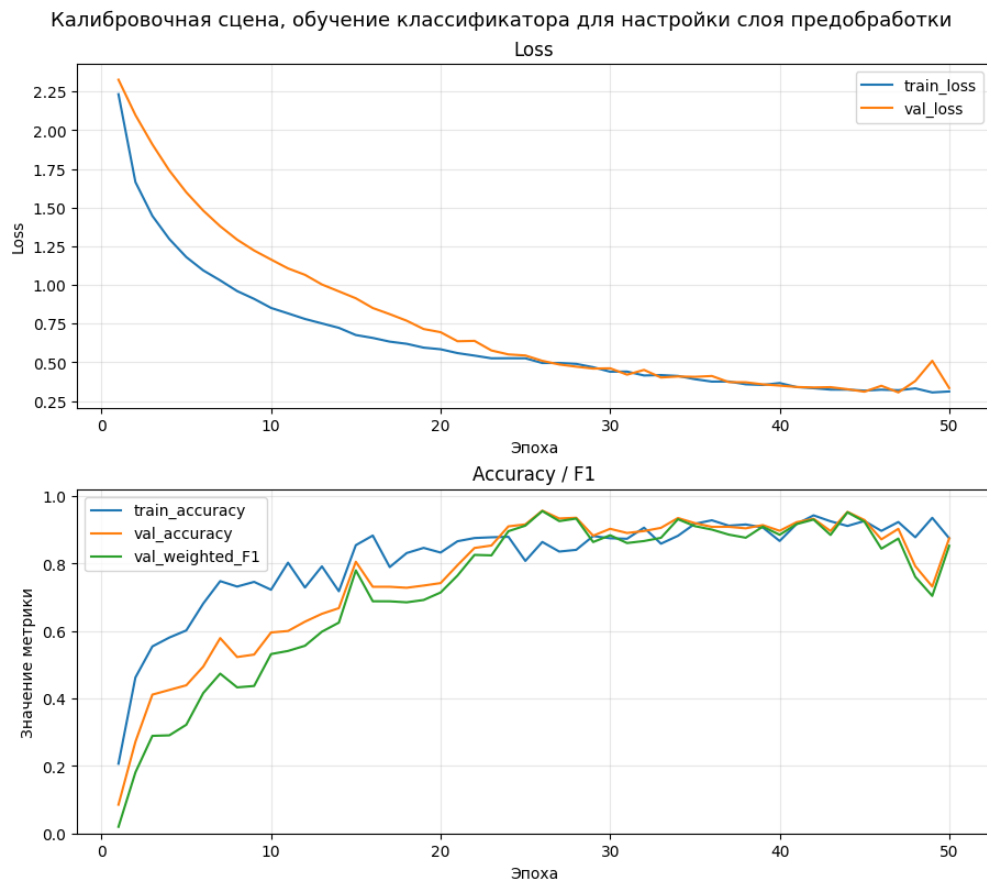
Таблица 2.3 – Параметры нейросетевой настройки предобработки

Классификатор	Размер входного фрагмента, пикселей	Число эпох	Размер пакета данных	Оптимизатор	Шаг обучения
$P_\theta + 1\text{DCNN}$	1	50	128	Adam	10^{-2}
$P_\theta + \text{NM3DCNN}$	5x5	35	256	AdamW	$5 \cdot 10^{-3}$
$P_\theta + \text{SSFTT}$	7x7	35	256	AdamW	$5 \cdot 10^{-4}$

Таблица 2.4 – Параметры обучения классификаторов

Классификатор	Размер входного фрагмента, пикселей	Число эпох	Размер пакета данных	Оптимизатор	Шаг обучения
1DCNN	1	50	128	Adam	10^{-2}
NM3DCNN	5x5	35	256	AdamW	$5 \cdot 10^{-3}$
SSFTT	7x7	35	256	AdamW	$5 \cdot 10^{-4}$

Процесс обучения классификаторов контролировался по значениям функции потерь, общей точности классификации и взвешенной F_1 -меры на обучающей и валидационной выборках. Для сочетания 1DCNN–NM3DCNN приведены графики обучения F_1 на опорном ГСИ калибровочной сцены и F_2 на опорном ГСИ тестовой сцены (рисунки 2.8–2.9).

Рисунок 2.8 – Обучение классификатора F_1 (пример 1DCNN) на опорном ГСИ калибровочной сцены Sresim IQ

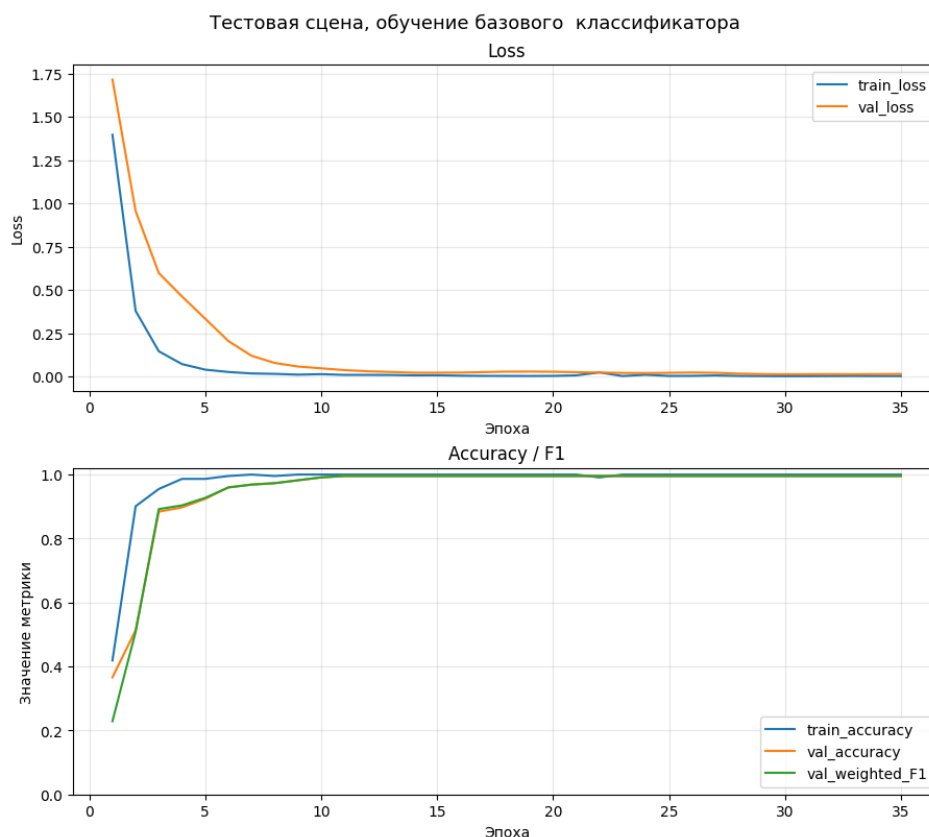


Рисунок 2.9 – Обучение классификатора F_2 (пример NM3DCNN) на опорном ГСИ тестовой сцены Sresim IQ

На рисунках 2.10–2.11 приведены примеры зависимостей функции потерь, общей точности поэлементной классификации, взвешенной F1-меры и программной поправки уровня чёрного от номера эпохи при настройке P_θ с классификатором 1DCNN по критериям L_{CE} и L_{CE_MSE} . Пример для комбинированной функции потерь относится к исходной основной серии при $\lambda = 2,5$. В итоговых таблицах используются значения λ , выбранные в дополнительном исследовании подраздела 2.2.5.

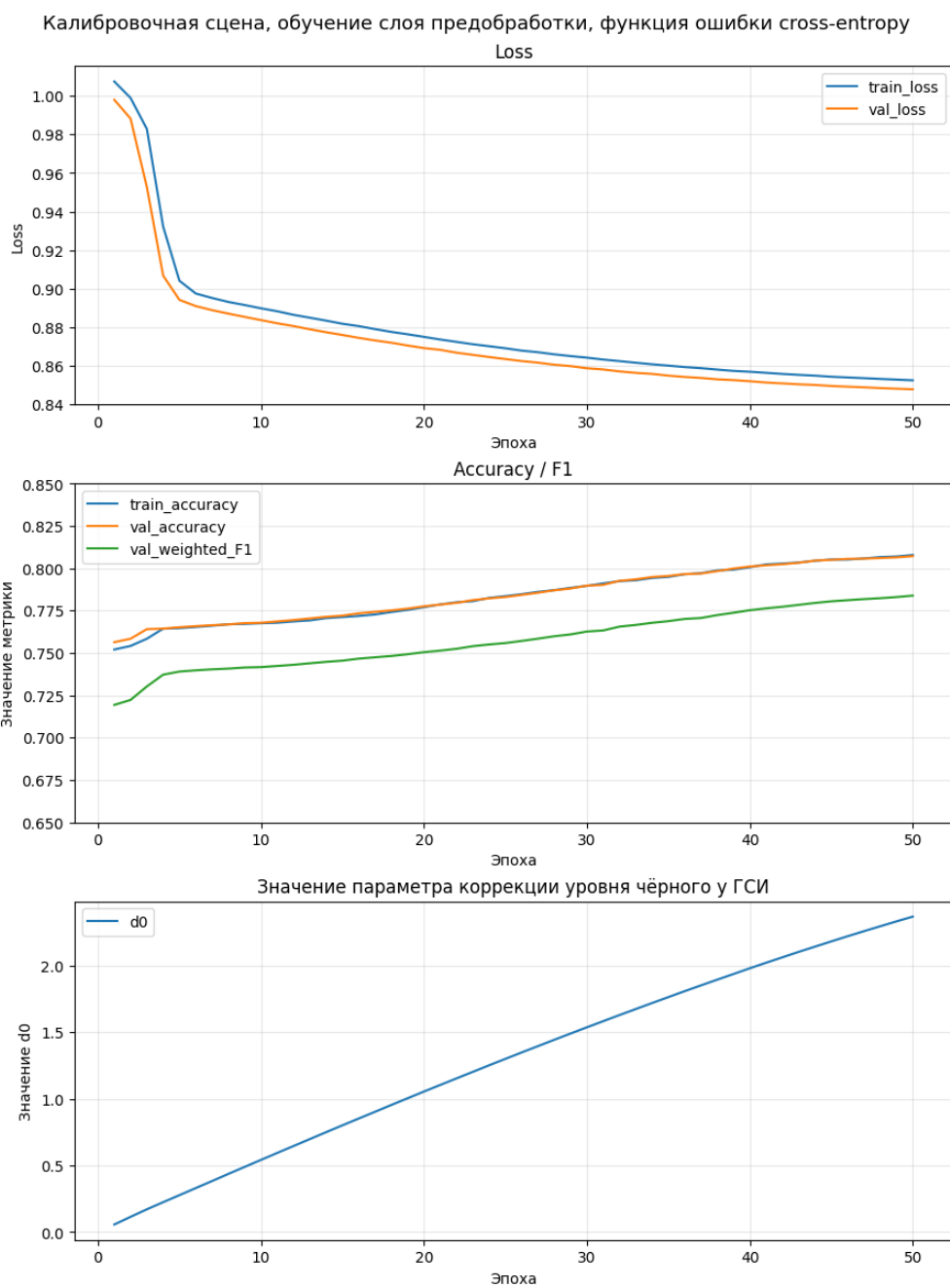


Рисунок 2.10 – Настройка оператора P_θ на калибровочной сцене при фиксированном классификаторе F_1 (пример 1DCNN) по категориальной перекрёстной энтропии $L_{\text{CE}}(\theta)$ (2.15)

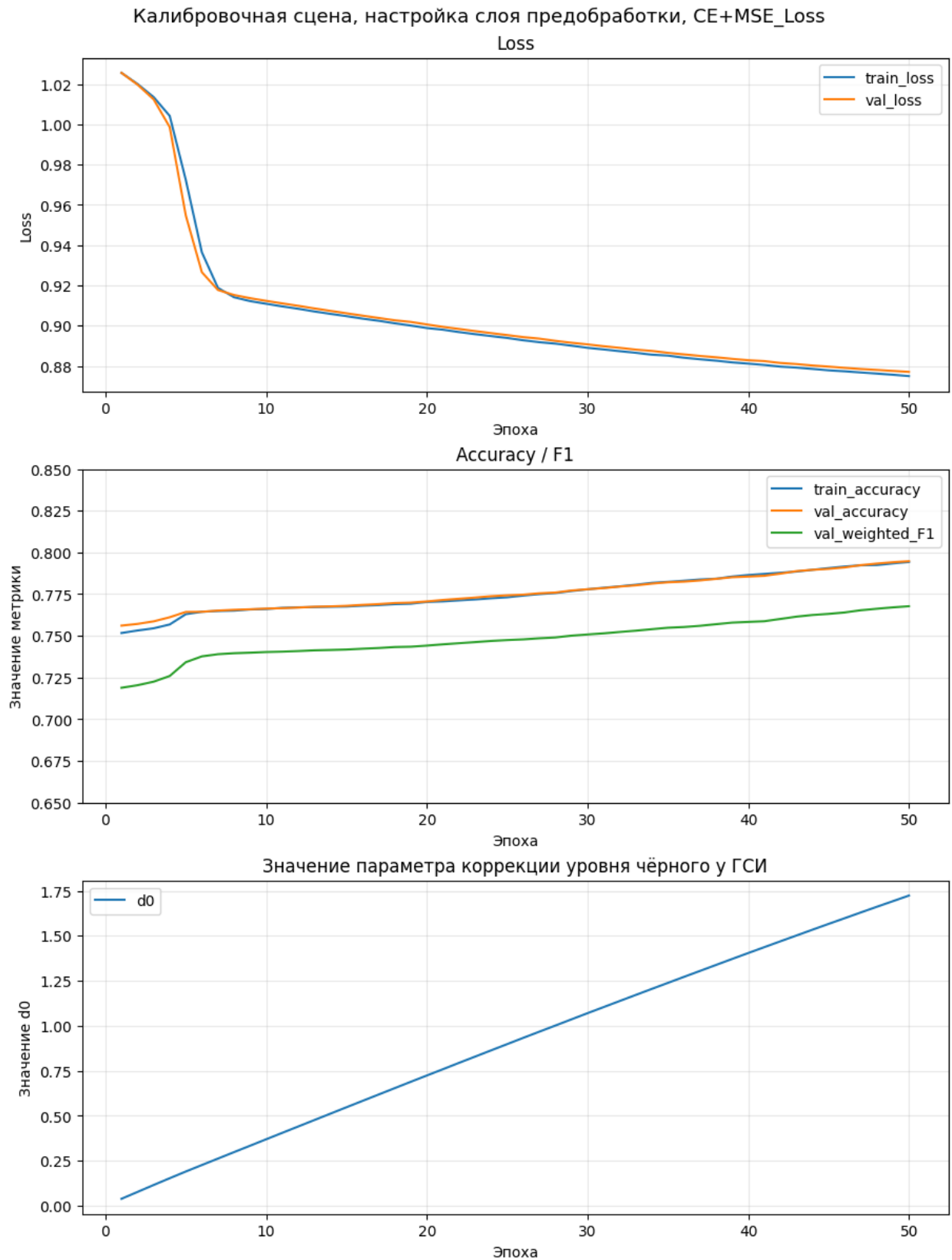


Рисунок 2.11 – Настройка оператора P_θ на калибровочной сцене при фиксированном классификаторе F_1 (пример 1DCNN) по комбинированной функции потерь $L_{CE_MSE}(\theta)$ (2.19)

2.2.4 Влияние длины окна фильтра Савицкого–Голея на точность классификации

Для выбора длины окна проведена дополнительная серия экспериментов с классификаторами 1DCNN, NM3DCNN и SSFTT. Проверены окна 3, 5, 7, 9 и 11 спектральных каналов при степени полинома 2 и вариант без фильтра. Фильтр применялся однократно вдоль спектральной оси после интерполяции цифровых отсчётов сравниваемого ГСИ на общую сетку длин волн и до нормирования по фиксированному спектру белого эталона. Обучающая выборка опорного ГСИ Specim IQ фильтром не обрабатывалась.

Для каждой архитектуры выполнено по десять запусков обучения на опорном ГСИ Specim IQ с использованием базовых значений генератора 43, 123 и 2026 – всего тридцать запусков обучения. В этой дополнительной серии 1DCNN и NM3DCNN обучались в течение 25 эпох, SSFTT – 35 эпох. Размеры пространственных фрагментов составляли соответственно 1×1 , 5×5 и 7×7 , размеры пакетов данных – 256, 32 и 32. Для 1DCNN использовался Adam с шагом обучения 0,001; для NM3DCNN и SSFTT – AdamW с шагом обучения 0,0005 и коэффициентом регуляризации весов 0,0001. Эпоха выбиралась по валидационной выборке опорного ГСИ Specim IQ. Затем каждый классификатор с фиксированными весами оценивал все окна и вариант без фильтра на одних и тех же пикселях сравниваемого ГСИ. Валидационная и тестовая выборки не содержали общих пикселей, но относились к одной сцене; пространственные окрестности соседних пикселей могли перекрываться.

Общее для трёх архитектур окно выбиралось по средней валидационной ОА на ГСИ печатной калибровочной сцены с одинаковым весом каждой архитектуры и каждого запуска. Наибольшая средняя валидационная ОА для 1DCNN, NM3DCNN и SSFTT получена при окне 5: 0,8901; 0,9154 и 0,9495 соответственно. Поэтому это окно принято для итогового сопоставления вариантов предобработки. Тестовая ОА при окне 5 составила соответственно

$0,8777 \pm 0,0127$; $0,9116 \pm 0,0102$ и $0,9497 \pm 0,0086$, где указаны среднее значение и выборочное стандартное отклонение (SD) по десяти запускам.

Средние парные приросты тестовой ОА относительно варианта без фильтра в этой дополнительной серии составили +0,24; +0,27 и +0,96 процентного пункта для 1DCNN, NM3DCNN и SSFTT соответственно. Они вычислены по одним и тем же запускам и не относятся к однократным базовым строкам таблиц 2.6–2.10. Окно 3 при степени полинома 2 практически воспроизводит вариант без фильтра. Увеличение окна свыше 5 не улучшило среднюю валидационную ОА ни у одной архитектуры.

Результаты представлены на рисунках 2.12–2.14. Точки показывают средние значения показателей. Сплошные отрезки при окне 5 на панелях (b) и (c) соответствуют среднему $\pm$ SD по десяти запускам. Пунктирные отрезки имеют условную ширину, заданную по масштабу SD для окна 5; они не являются оценками разброса соответствующих вариантов или доверительными интервалами. На панели (d) приведены средние парные приросты без интервалов.

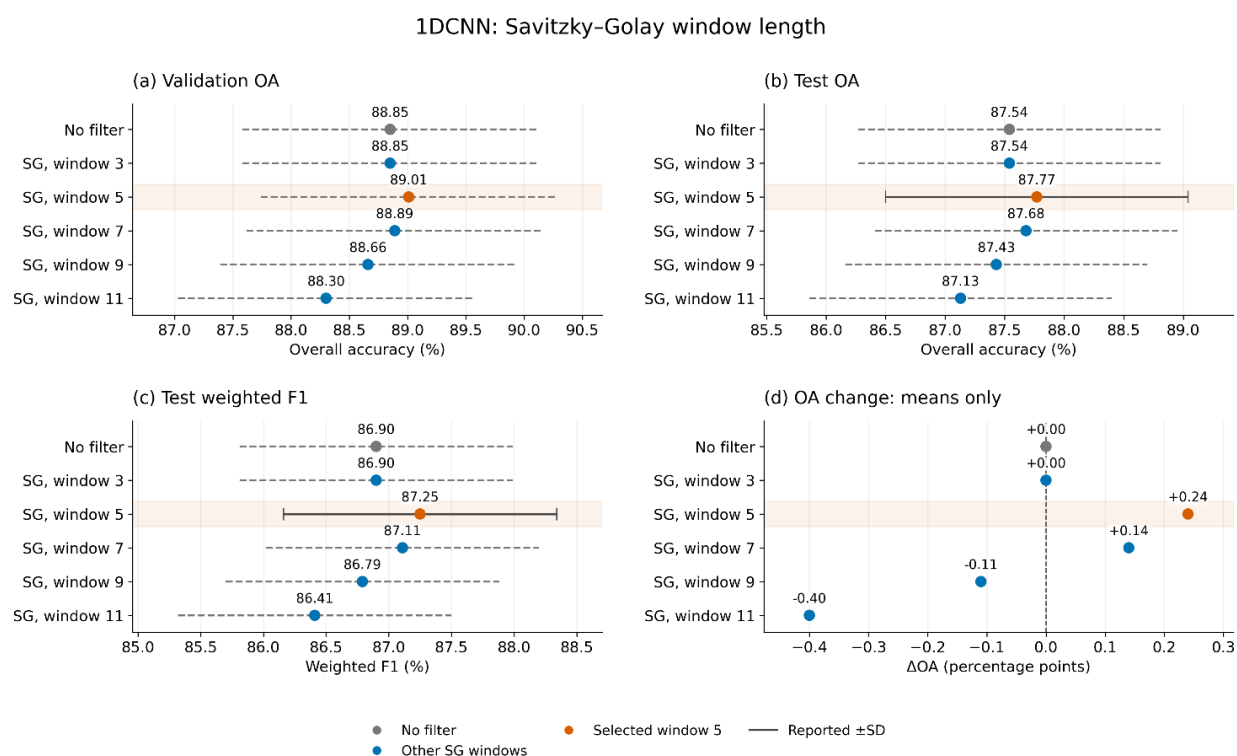


Рисунок 2.12 – Влияние длины окна фильтра Савицкого–Голя на результаты 1DCNN (тестовая сцена): а – валидационная ОА; б – тестовая ОА; в – взвешенная F1-мера; д – парный прирост тестовой ОА. Сплошные отрезки при окне 5 – $\pm$ SD; пунктирные – условные иллюстративные полосы

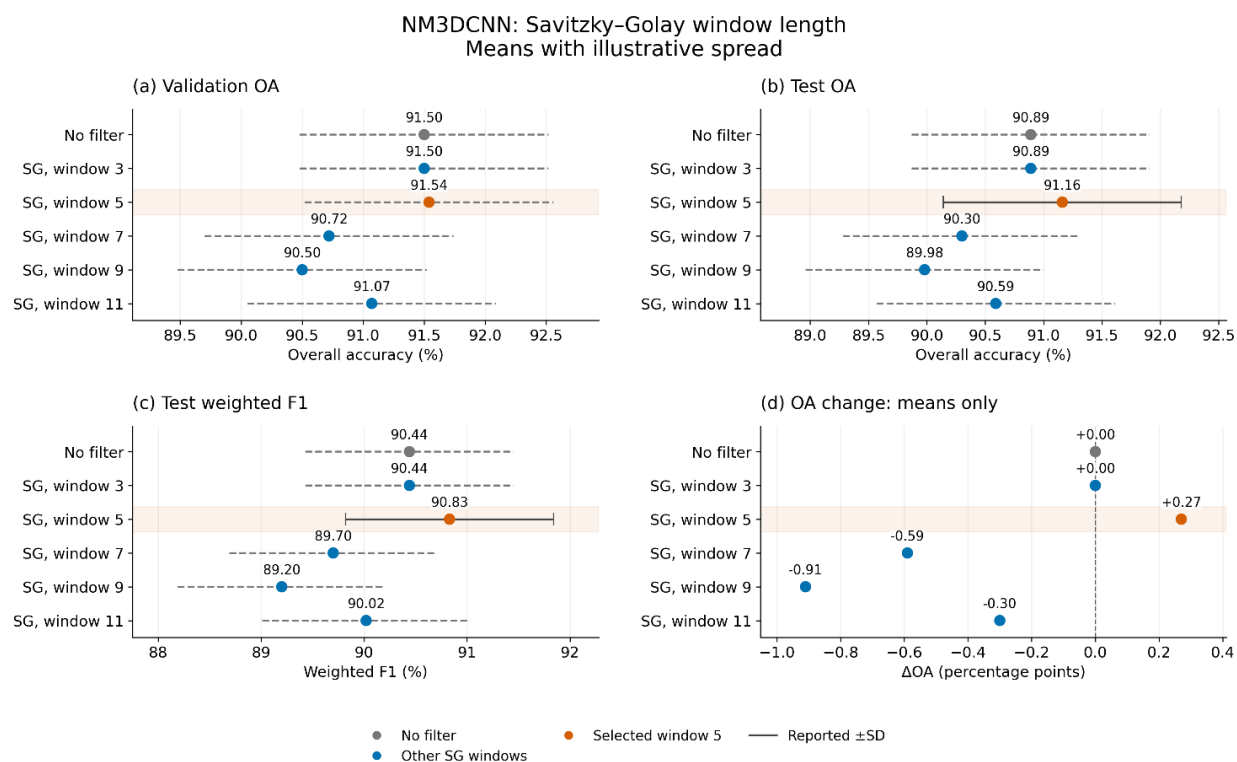


Рисунок 2.13 – Влияние длины окна фильтра Савицкого–Голя на результаты NM3DCNN (тестовая сцена): а – валидационная ОА; б – тестовая ОА; в – взвешенная F1-мера; д – парный прирост тестовой ОА. Сплошные отрезки при окне 5 – $\pm$ SD; пунктирные – условные иллюстративные полосы

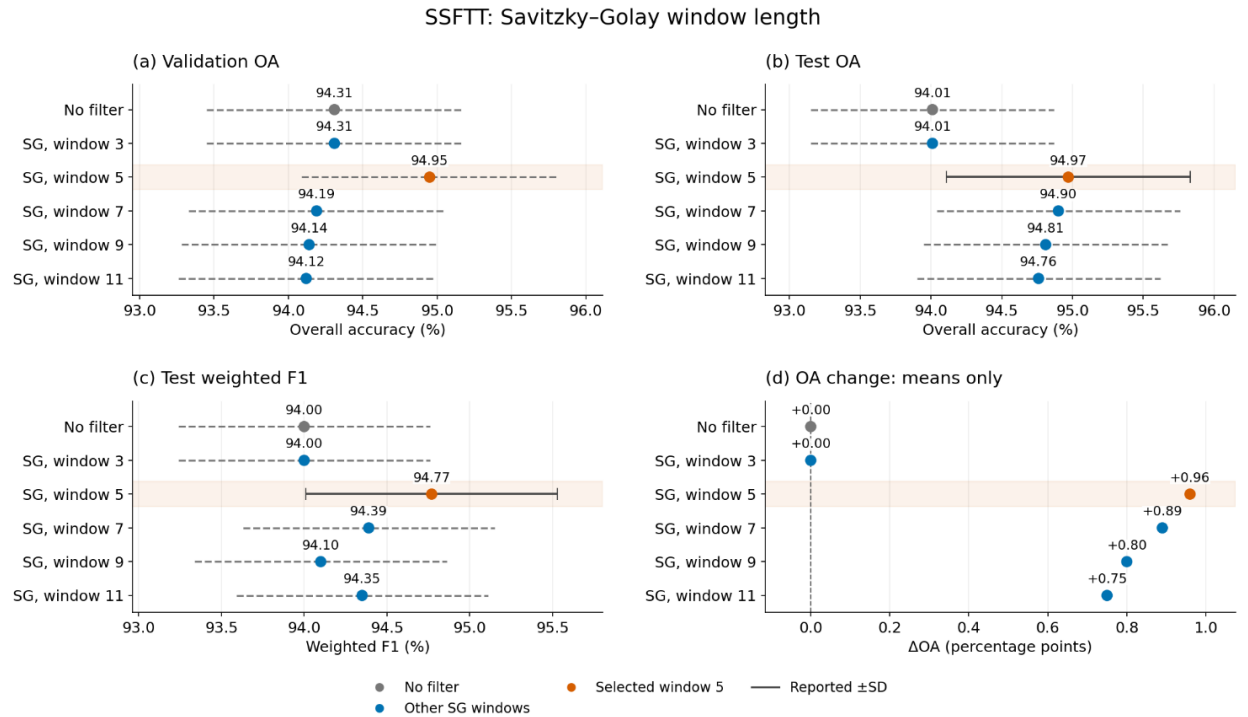


Рисунок 2.14 – Влияние длины окна фильтра Савицкого–Голея на результаты SSFTT (тестовая сцена): а – валидационная ОА; б – тестовая ОА; с – взвешенная F1-мера; д – парный прирост тестовой ОА. Сплошные отрезки при окне 5 – $\pm$ SD; пунктирные – условные иллюстративные полосы

2.2.5 Выбор веса спектральной составляющей комбинированного критерия

В дополнительной серии исследовано влияние коэффициента λ комбинированной функции потерь (2.19). Проверены значения 0,01; 0,1; 1; 2,5 и 5. Для каждого значения и каждой архитектуры классификатора настройки P_0 выполнено десять запусков с использованием базовых значений генератора 43, 123 и 2026; всего выполнено 150 запусков настройки. Внутри каждой архитектуры начальные параметры P_0 , классификатор и разбиение данных при одинаковом начальном значении генератора сохранялись для всех λ . Параметры классификатора оставались фиксированными. Настройка P_0 ограничивалась 20 эпохами; использовался Adam с шагом обучения 0,002 для 1DCNN и 0,0002 для NM3DCNN и SSFTT.

Коэффициент выбирался отдельно для архитектуры классификатора F_1 по максимуму средней ОА на валидационной части калибровочной сцены по десяти запускам. При равенстве ОА выбиралось меньшее λ . Значение для

1DCNN применялось во всех трёх сочетаниях оценки: 1DCNN–1DCNN, 1DCNN–NM3DCNN и 1DCNN–SSFTT. Для NM3DCNN–NM3DCNN и SSFTT–SSFTT использовались значения, выбранные для соответствующих архитектур F_1 . Тестовая ОА в критерий выбора не входила. Выбранные значения приведены в таблице 2.5.

Таблица 2.5 – Значения коэффициента λ в итоговых экспериментах

Классификатор при настройке	Выбранное λ
1DCNN	0,1
NM3DCNN	0,01
SSFTT	1

Переход от общего коэффициента к выбору по архитектуре был уточнён после просмотра промежуточных тестовых результатов, поэтому данный анализ имеет исследовательский характер. Параметры обучения дополнительной и основной серий различаются. Их сопоставление носит описательный характер и не позволяет выделить влияние только функции потерь.

2.2.6 Результаты поэлементной классификации и оценки спектрального сходства

Перенос предобработки после настройки по калибровочной сцене оценивался на независимой тестовой сцене по общей точности поэлементной классификации (ОА) и взвешенной F_1 -мере. Для всех вариантов использовались одни и те же тестовые пиксели сравниваемого ГСИ; пиксели фона с меткой 0 исключались. Спектральное сопоставление с опорным ГСИ оценивалось по PSNR средних спектральных векторов соответствующих классов, рассчитанных по эталонной разметке каждого ГСИ. В таблицах приведено среднее арифметическое PSNR по классам.

В варианте без дополнительной предобработки цифровые отсчёты нормировались по спектру белого эталона в соответствии с (2.7). В варианте с

фильтром Савицкого–Голея перед нормированием выполнялось спектральное сглаживание. В обоих базовых вариантах принималось $d_0 = 0$.

В таблицах 2.6–2.10 сопоставлены результаты без дополнительной предобработки, с фильтром Савицкого–Голея и с оператором P_θ , параметры которого определялись гребневой регрессией либо нейросетевой настройкой по L_{CE} и L_{CE_MSE} . Для нейросетевых вариантов и фильтра запись «среднее $\pm$ SD» обозначает среднее значение и выборочное стандартное отклонение по десяти запускам. Для варианта без дополнительной предобработки приведён результат одного запуска основной серии, для гребневой регрессии – значения без оценки разброса. Классификатор F_1 участвовал только в нейросетевой настройке; F_2 использовался для оценки всех вариантов.

Строки фильтра Савицкого–Голея с окном 5 получены в дополнительной серии подраздела 2.2.4, а строки комбинированного критерия – в серии подраздела 2.2.5. Параметры обучения этих серий отличаются от основной серии, поэтому сравнение носит описательный характер. При одинаковом классификаторе F_2 результаты фильтра повторяются независимо от F_1 ; PSNR фильтра во всех пяти таблицах равен 23,490 дБ. Полужирным выделены наибольшие значения каждого показателя.

Таблица 2.6 – Результаты на тестовой сцене для сочетания 1DCNN–1DCNN; $\lambda = 0,1$ для комбинированного критерия

Вариант предобработки	ОА	Взвешенная F1-мера	PSNR, дБ
Без дополнительной предобработки	0,8754	0,8690	23,409
Фильтр Савицкого–Голея (окно 5)	0,8777 $\pm$ 0,0127	0,8725 $\pm$ 0,0109	23,490
P_θ, L_{CE}	0,9155 $\pm$ 0,0110	0,9120 $\pm$ 0,0142	24,037 $\pm$ 0,027
Гребневая регрессия	0,9104	0,9059	24,016
P_θ, L_{CE_MSE}	0,9130 $\pm$ 0,0085	0,9074 $\pm$ 0,0075	24,104 $\pm$ 0,009

Таблица 2.7 – Результаты на тестовой сцене для сочетания 1DCNN–NM3DCNN; $\lambda = 0,1$ для комбинированного критерия

Вариант предобработки	ОА	Взвешенная F1-мера	PSNR, дБ
Без дополнительной предобработки	0,9013	0,8868	23,409
Фильтр Савицкого–Голея (окно 5)	$0,9116 \pm 0,0102$	$0,9083 \pm 0,0101$	23,490
P_{θ}, L_{CE}	$0,9489 \pm 0,0057$	$0,9482 \pm 0,0040$	$24,037 \pm 0,027$
Гребневая регрессия	0,9495	0,9425	24,016
P_{θ}, L_{CE_MSE}	$0,9635 \pm 0,0073$	$0,9636 \pm 0,0055$	$24,104 \pm 0,009$

Таблица 2.8 – Результаты на тестовой сцене для сочетания 1DCNN–SSFTT; $\lambda = 0,1$ для комбинированного критерия

Вариант предобработки	ОА	Взвешенная F1-мера	PSNR, дБ
Без дополнительной предобработки	0,9405	0,9418	23,409
Фильтр Савицкого–Голея (окно 5)	$0,9497 \pm 0,0086$	$0,9477 \pm 0,0076$	23,490
P_{θ}, L_{CE}	$0,9619 \pm 0,0120$	$0,9610 \pm 0,0139$	$24,037 \pm 0,027$
Гребневая регрессия	0,9509	0,9501	24,016
P_{θ}, L_{CE_MSE}	$0,9737 \pm 0,0096$	$0,9736 \pm 0,0063$	$24,104 \pm 0,009$

Таблица 2.9 – Результаты на тестовой сцене для сочетания NM3DCNN–NM3DCNN; $\lambda = 0,01$ для комбинированного критерия

Вариант предобработки	ОА	Взвешенная F1-мера	PSNR, дБ
Без дополнительной предобработки	0,9013	0,8868	23,409
Фильтр Савицкого–Голея (окно 5)	$0,9116 \pm 0,0102$	$0,9083 \pm 0,0101$	23,490
P_{θ}, L_{CE}	$0,9469 \pm 0,0027$	$0,9458 \pm 0,0031$	$24,066 \pm 0,034$
Гребневая регрессия	0,9495	0,9425	24,016
P_{θ}, L_{CE_MSE}	$0,9450 \pm 0,0078$	$0,9452 \pm 0,0027$	$23,395 \pm 0,011$

Таблица 2.10 – Результаты на тестовой сцене для сочетания SSFTT–SSFTT; $\lambda = 1$ для комбинированного критерия

Вариант предобработки	ОА	Взвешенная F1-мера	PSNR, дБ
Без дополнительной предобработки	0,9405	0,9418	23,409
Фильтр Савицкого–Голея (окно 5)	$0,9497 \pm 0,0086$	$0,9477 \pm 0,0076$	23,490
$P_{\theta, LCE}$	$0,8466 \pm 0,0167$	$0,8570 \pm 0,0441$	$21,543 \pm 0,021$
Гребневая регрессия	0,9509	0,9501	24,016
P_{θ, LCE_MSE}	$0,9622 \pm 0,0121$	$0,9621 \pm 0,0243$	$23,677 \pm 0,035$

На рисунках 2.15–2.19 приведены примеры средних спектральных векторов пяти классов тестовой сцены для сочетания 1DCNN–NM3DCNN; на рисунке 2.20 – соответствующие результаты поэлементной классификации. В этих примерах использованы фильтр Савицкого–Голея с окном длиной 5 каналов и комбинированная функция потерь при $\lambda = 0,1$. Иллюстрации характеризуют форму спектров и пространственное распределение предсказанных классов в отдельных запусках. Численные результаты сопоставления вариантов предобработки приведены в таблицах 2.6–2.10.

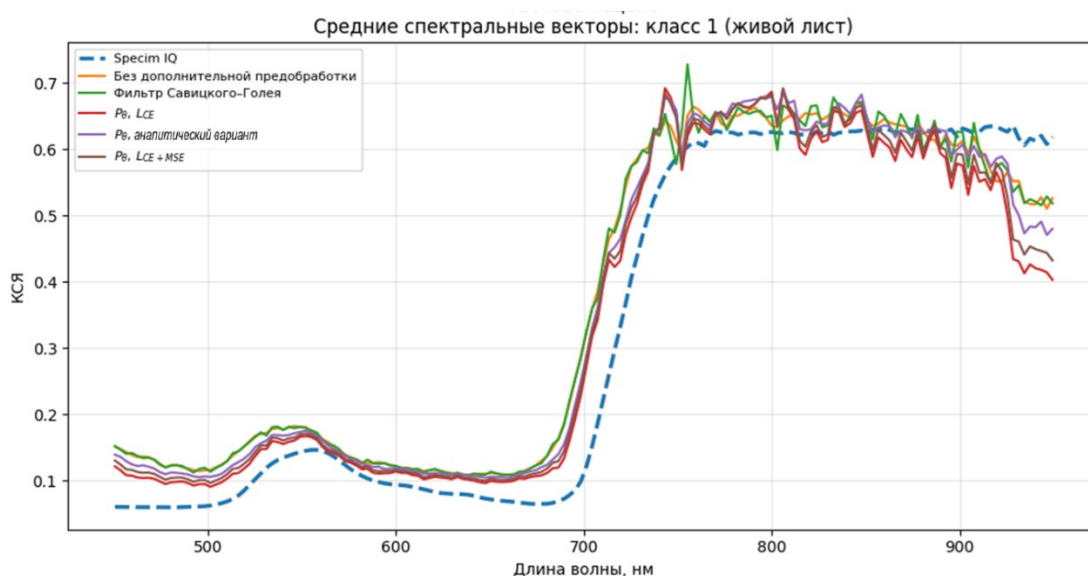


Рисунок 2.15 – Средние спектральные векторы класса 1 тестовой сцены для опорного ГСИ Specim IQ и сравниваемого ГСИ при различных вариантах предобработки (пример отдельного запуска: окно 5, $\lambda = 0,1$)

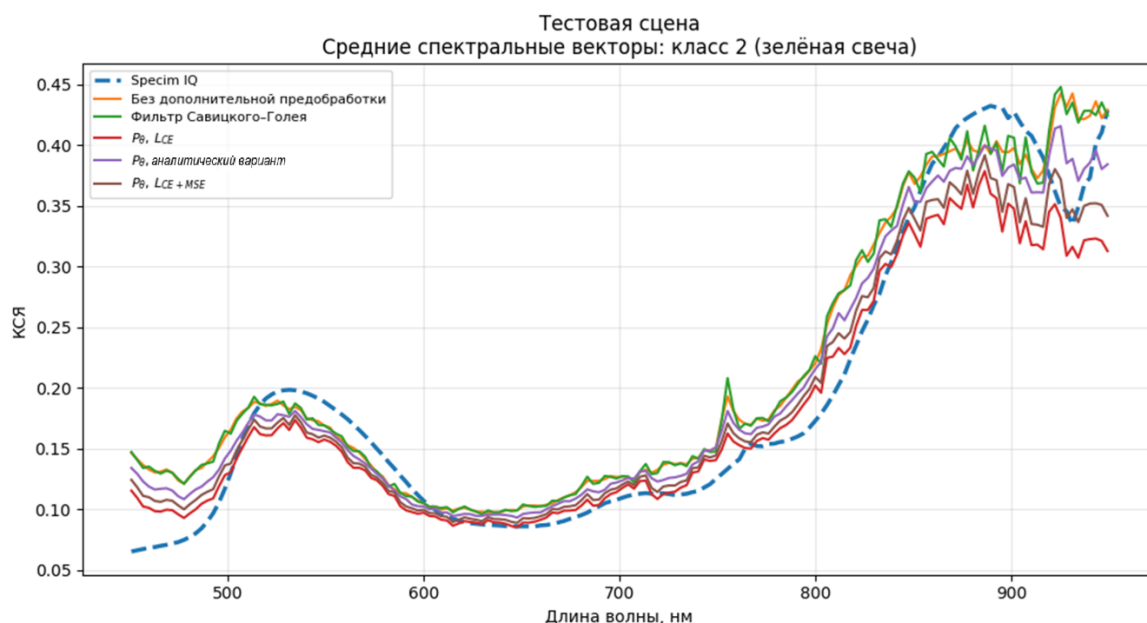


Рисунок 2.16 – Средние спектральные векторы класса 2 тестовой сцены для опорного ГСИ Specim IQ и сравниваемого ГСИ при различных вариантах предобработки (пример отдельного запуска: окно 5, $\lambda = 0,1$)

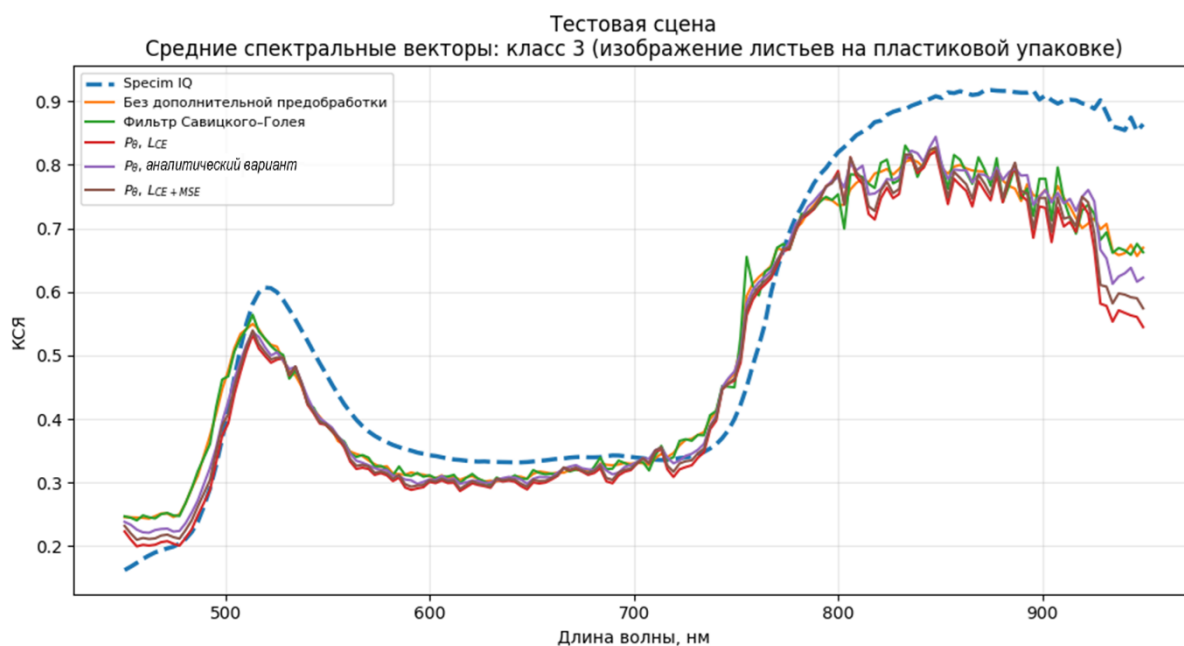


Рисунок 2.17 – Средние спектральные векторы класса 3 тестовой сцены для опорного ГСИ Specim IQ и сравниваемого ГСИ при различных вариантах предобработки (пример отдельного запуска: окно 5, $\lambda = 0,1$)



Рисунок 2.18 – Средние спектральные векторы класса 4 тестовой сцены для опорного ГСИ Specim IQ и сравниваемого ГСИ при различных вариантах предобработки (пример отдельного запуска: окно 5, $\lambda = 0,1$)

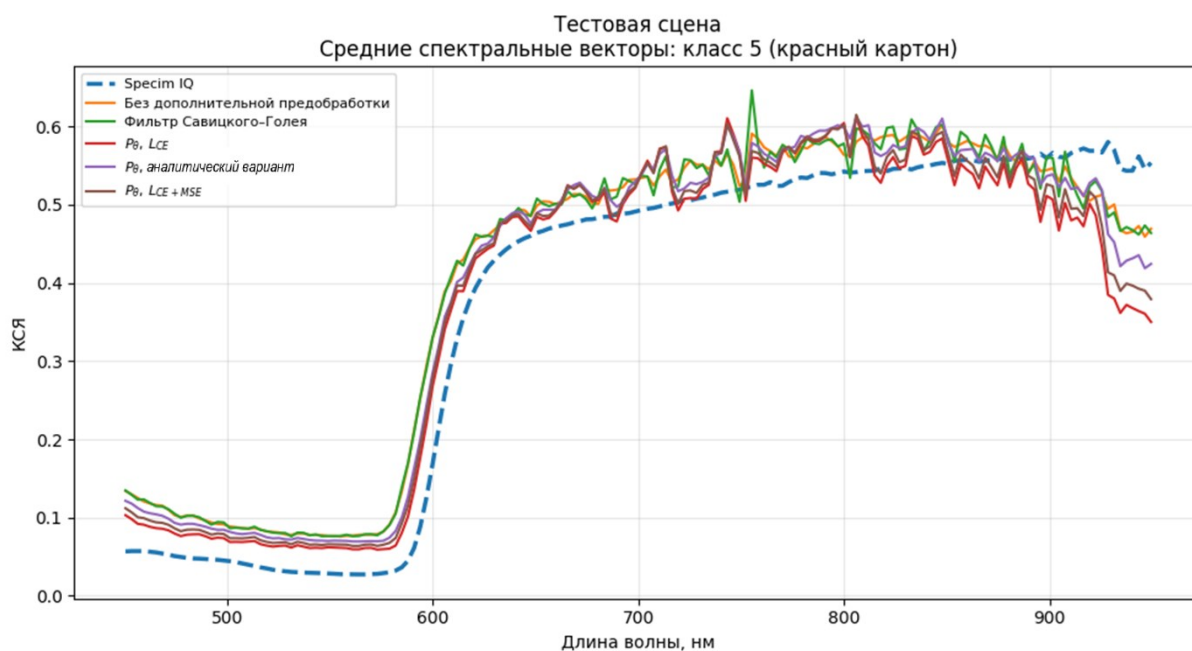


Рисунок 2.19 – Средние спектральные векторы класса 5 тестовой сцены для опорного ГСИ Specim IQ и сравниваемого ГСИ при различных вариантах предобработки (пример отдельного запуска: окно 5, $\lambda = 0,1$)

Тестовая сцена. Пример результатов поэлементной классификации

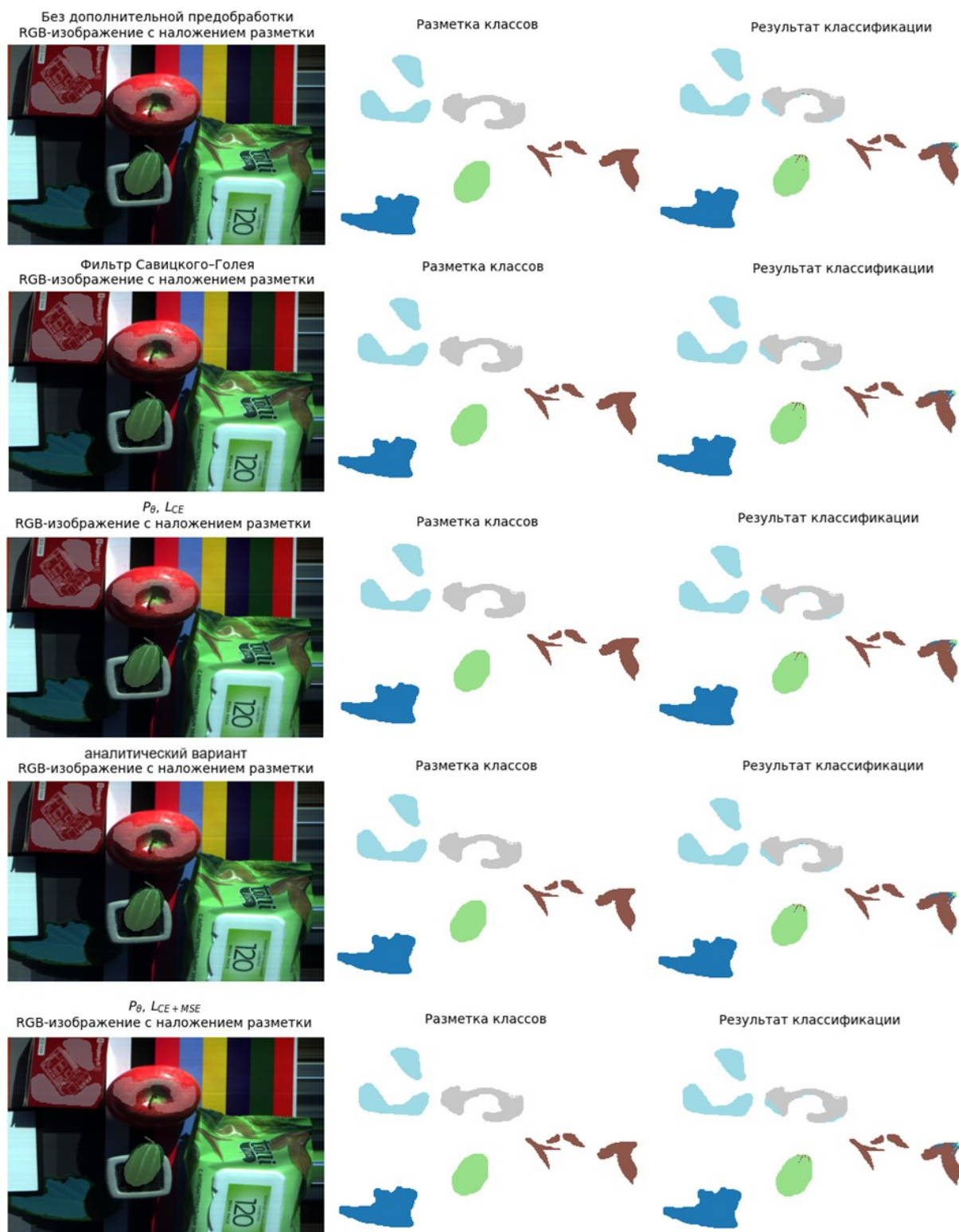


Рисунок 2.20 – Результаты поэлементной классификации тестовой сцены (пример отдельного запуска: окно 5, $\lambda = 0,1$)

В каждом из пяти сочетаний классификаторов хотя бы один вариант обучаемой предобработки имел средние ОА и взвешенную F1-меру выше

приведённых значений без дополнительной предобработки и после фильтра Савицкого–Голея с окном 5. Для сочетания 1DCNN–1DCNN наибольшие классификационные показатели получены при L_{CE} : 0,9155 и 0,9120, что выше базовых значений на 4,01 и 4,30 процентного пункта соответственно. Для сочетаний 1DCNN–NM3DCNN и 1DCNN–SSFTT наибольшие показатели относятся к комбинированному критерию при $\lambda = 0,1$; приросты ОА составили 6,22 и 3,32 процентного пункта, а взвешенной F1-меры – 7,68 и 3,18 процентного пункта соответственно (таблицы 2.6–2.8).

Для сочетания NM3DCNN–NM3DCNN наибольшее значение ОА в таблице 2.9 приведено для гребневой регрессии: 0,9495, что на 4,82 процентного пункта выше базового значения. Наибольшие средние значения взвешенной F1-меры и PSNR получены при L_{CE} : 0,9458 и 24,066 дБ; превышения над базовыми значениями составляют 5,90 процентного пункта и 0,657 дБ соответственно. Для сочетания SSFTT–SSFTT вариант L_{CE} имел среднюю ОА 0,8466, что ниже базового результата на 9,39 процентного пункта. В дополнительной серии комбинированный критерий при $\lambda = 1$ дал средние значения ОА 0,9622 и взвешенной F1-меры 0,9621, то есть превышения на 2,17 и 2,03 процентного пункта соответственно (таблица 2.10).

Среди результатов таблиц 2.6–2.10 наибольшие средние значения ОА и взвешенной F1-меры составили 0,9737 и 0,9736 для сочетания 1DCNN–SSFTT с комбинированным критерием при $\lambda = 0,1$. По сравнению с фильтром Савицкого–Голея с окном 5 эти значения выше на 2,40 и 2,59 процентного пункта; средний PSNR выше на 0,614 дБ. Наибольшее среднее значение PSNR среди всех пяти таблиц, равное 24,104 дБ, получено для комбинированного критерия при $\lambda = 0,1$ в сочетаниях 1DCNN–1DCNN, 1DCNN–NM3DCNN и 1DCNN–SSFTT (таблицы 2.6–2.8). Таким образом, для сочетания 1DCNN–SSFTT наибольшие средние значения всех трёх показателей получены в одном варианте предобработки.

2.2.7 Перенос предобработки на SVM, kNN и MLP

В дополнительном эксперименте оценён перенос настроенного оператора предобработки P_θ на метод опорных векторов (SVM), метод k ближайших соседей (kNN) и многослойный перцептрон (MLP). В качестве классификатора F_1 при настройке P_θ использовались 1DCNN, NM3DCNN и SSFTT. Настройка выполнялась по комбинированной функции потерь при выбранных значениях λ : 0,1 для 1DCNN, 0,01 для NM3DCNN и 1 для SSFTT. Параметры F_1 оставались фиксированными; обновлялись только коэффициенты спектральной свёртки и программная поправка уровня чёрного.

После настройки оператор P_θ применялся к тестовой сцене без дополнительной оптимизации. Модели SVM, kNN и MLP выступали в роли классификатора F_2 : каждая модель однократно обучалась на опорном ГСИ Specim IQ тестовой сцены и затем применялась с фиксированными параметрами ко всем вариантам предобработки, не участвуя в её настройке.

Результаты переноса приведены в таблицах 2.11–2.13. Для каждого обучаемого варианта выполнено десять запусков настройки предобработки; это число указано в столбце «Количество запусков». Для обучаемой предобработки представлены средние значения ОА и взвешенной F1-меры $\pm$ выборочное стандартное отклонение (SD). Для варианта без дополнительной предобработки приведены результаты однократно обученного классификатора без оценки разброса; настройка предобработки в этом варианте не выполнялась.

Таблица 2.11 – Результаты переноса предобработки на SVM при настройке по комбинированной функции потерь

Вариант предобработки	λ	Количество запусков	ОА	Взвешенная F1-мера
Без дополнительной предобработки	—	—	0,7728	0,7553
P_θ , настройка с 1DCNN	0,1	10	0,8913 $\pm$ 0,0336	0,8891 $\pm$ 0,0255

Вариант предобработки	λ	Количество запусков	ОА	Взвешенная F1-мера
P_0 , настройка с NM3DCNN	0,01	10	$0,8085 \pm 0,0806$	$0,7913 \pm 0,0895$
P_0 , настройка с SSFTT	1	10	$0,8181 \pm 0,1232$	$0,7956 \pm 0,1426$

Таблица 2.12 – Результаты переноса предобработки на kNN при настройке по комбинированной функции потерь

Вариант предобработки	λ	Количество запусков	ОА	Взвешенная F1-мера
Без дополнительной предобработки	—	—	0,8993	0,8992
P_0 , настройка с 1DCNN	0,1	10	0,9065 $\pm$ 0,0019	$0,8984 \pm 0,0019$
P_0 , настройка с NM3DCNN	0,01	10	$0,8932 \pm 0,0157$	$0,8932 \pm 0,0156$
P_0 , настройка с SSFTT	1	10	$0,8692 \pm 0,0351$	$0,8685 \pm 0,0362$

Таблица 2.13 – Результаты переноса предобработки на MLP при настройке по комбинированной функции потерь

Вариант предобработки	λ	Количество запусков	ОА	Взвешенная F1-мера
Без дополнительной предобработки	—	—	0,6971	0,6528
P_0 , настройка с 1DCNN	0,1	10	0,7093 $\pm$ 0,0067	$0,6520 \pm 0,0091$
P_0 , настройка с NM3DCNN	0,01	10	$0,6867 \pm 0,0126$	$0,6375 \pm 0,0197$
P_0 , настройка с SSFTT	1	10	$0,6987 \pm 0,0307$	0,6548 $\pm$ 0,0438

Для всех трёх классификаторов F_2 наибольшая средняя ОА получена при настройке P_0 с 1DCNN: 0,8913 для SVM, 0,9065 для kNN и 0,7093 для MLP. Приросты относительно варианта без дополнительной предобработки составили 11,85, 0,72 и 1,22 процентного пункта соответственно (таблица 2.14). Для SVM средняя взвешенная F1-мера повысилась с 0,7553 до 0,8891, то есть на 13,38 процентного пункта. Для kNN и MLP этот показатель несколько снизился: с 0,8992 до 0,8984 и с 0,6528 до 0,6520 соответственно, то есть на 0,08 процентного пункта в обоих случаях. Таким образом, повышение ОА не во всех случаях сопровождается повышением взвешенной F1-меры.

Таблица 2.14 – Прирост ОА при переносе предобработки, настроенной с 1DCNN

Модель	$\Delta\text{ОА}$, п.п.	Стандартное отклонение, п.п.
SVM	+11,850	3,360
kNN	+0,720	0,190
MLP	+1,220	0,670

В таблице 2.14 приросты средней ОА рассчитаны относительно фиксированных табличных базовых значений. Приведённые стандартные отклонения совпадают со стандартными отклонениями обучаемых вариантов, выраженными в процентных пунктах, и не учитывают возможную вариабельность базового классификатора.

2.2.8 Сравнение с методами DeSpaWN и DeCWN

В работе [111] предложены модели DeSpaWN и DeCWN, использующие обучаемое вейвлетное преобразование и пороговое подавление коэффициентов одномерных сигналов. В DeSpaWN фильтр определяется отдельно для каждого уровня разложения, а в DeCWN один фильтр используется на всех уровнях. В настоящей работе модели адаптированы к спектральной предобработке ГСИ на основе открытой программной реализации [112].

Модули настраивались по ГСИ печатной калибровочной сцены и применялись к отдельной тестовой сцене по протоколу подраздела 2.2.3 с использованием SSFTT. Параметры классификаторов оставались фиксированными; обучалась только предобработка. Рассматривались перекрёстная энтропия (CE) и комбинированный критерий (CE совместно с MSE). На тестовой сцене предобработка повторно не настраивалась.

Для каждого сочетания модели и критерия выполнено десять запусков настройки с использованием базовых значений seed 43, 123 и 2026; число запусков является суммарным для каждого варианта. Результаты в таблице 2.15 представляются как среднее $\pm$ выборочное стандартное отклонение (SD).

Взвешенная F1-мера, как и ОА. Число обучаемых параметров спектрального модуля в рассмотренных конфигурациях составило 52 для DeSpaWN и 20 для DeCWN.

Таблица 2.15 – Результаты применения DeSpaWN и DeCWN с классификатором SSFTT, настройка при использовании 1DCNN

Модель	Функция ошибки	ОА	Взвешенная F1-мера	PSNR, дБ	Число параметров
DeSpaWN	CE	$0,9643 \pm 0,0023$	$0,9642 \pm 0,0046$	$23,887 \pm 0,015$	52
	CE + MSE	$0,9642 \pm 0,0057$	$0,9642 \pm 0,0033$	$24,004 \pm 0,008$	
DeCWN	CE	$0,9592 \pm 0,0054$	$0,9589 \pm 0,0042$	$23,890 \pm 0,021$	20
	CE + MSE	$0,9621 \pm 0,0073$	$0,9614 \pm 0,0061$	$24,056 \pm 0,011$	

2.2.9 Оценка вычислительных затрат

Дополнительно оценены число параметров нейросетевых классификаторов и время их прямого прохода (таблица 2.16). Для входа из 168 спектральных каналов рассмотрены конфигурации с 13 и 6 выходами, включая фоновый выход. Пространственный размер входа составлял 1×1 для 1DCNN, 5×5 для NM3DCNN и 7×7 для SSFTT.

Измерения выполнены на NVIDIA GeForce RTX 3050 Laptop GPU с 4 ГБ видеопамяти в PyTorch 2.4.1, CUDA 12.1, с точностью FP32. Время прямого прохода измерялось для конфигураций с шестью выходами при размере пакета данных 32, на случайно сформированных входах и моделях с начальной инициализацией весов. После десяти прогревочных вызовов выполнялись 15 серий по десять вызовов с синхронизацией CUDA на границах каждой серии. В таблице приведены среднее время одного вызова и выборочное стандартное отклонение средних времён серий для входа, размещённого на GPU, и для режима CPU $\rightarrow$ GPU $\rightarrow$ CPU.

Таблица 2.16 – Число параметров и время прямого прохода классификаторов

Модель	Параметров, 13 выходов	Параметров, 6 выходов	Вход на GPU, мс	CPU → GPU → CPU, мс
1DCNN	48 709	48 254	$0,452 \pm 0,111$	$0,578 \pm 0,151$
NM3DCNN	57 885	35 030	$1,365 \pm 0,172$	$1,839 \pm 0,244$
SSFTT	788 861	788 406	$2,321 \pm 0,337$	$2,802 \pm 0,497$

Для предобработки с ядром спектральной свёртки длиной семь каналов время применения к пакету данных из 4096 спектров составило $0,534 \pm 0,027$ мс при размещении входа на GPU и $1,757 \pm 0,112$ мс с передачей входа и результата между оперативной и видеопамятью. Классификатор в этот замер предобработки не включался.

Приведённые измерения характеризуют вычислительные операции в указанных конфигурациях. В измерения не включены чтение ГСИ, построение пространственных фрагментов и обучение. Сопоставимое время применения базовых вариантов и полное время экспериментальной серии в этом тесте не измерялись, поэтому результаты не устанавливают выигрыш полного вычислительного процесса относительно базовых вариантов.

2.2.10 Обсуждение результатов спектральной предобработки

В проведённых экспериментах предложенная спектральная предобработка позволила повысить ОА и PSNR по сравнению с вариантом без дополнительной предобработки. Результаты зависели от критерия настройки и используемых классификаторов. При настройке предобработки с использованием 1DCNN и последующей классификации ГСИ тестовой сцены трансформерным классификатором SSFTT получена более высокая средняя ОА, чем при применении предобработки на основе моделей DeSpaWN и DeCWN.

Предобработка, настроенная с использованием 1DCNN, позволила повысить ОА при последующем применении как с нейросетевыми классификаторами, так и с SVM, kNN и MLP. Таким образом, показана

возможность применения настроенных коэффициентов предобработки с классификаторами, отличными от использованного при настройке.

2.3 Влияние видеокодека и параметров сжатия на точность классификации

2.3.1 Условия эксперимента и параметры сжатия

Влияние сжатия на спектральные характеристики изображений и точность классификации рассматривалось в ряде исследований. В работе [113] исследовано сжатие данных сканирующего гиперспектрометра видеокодеками, включая H.264 и H.265. Оценивались изменение положения, интенсивности и ширины спектральных линий, возможность выполнения спектральной калибровки, PSNR и коэффициент корреляции.

В работе [114] использована оценка сжатия по результатам последующей классификации. Внутрикадровое кодирование HEVC и JPEG 2000 сопоставлено на многоспектральных изображениях Sentinel-2 при классификации сельскохозяйственных культур. Для каждого варианта сжатия классификатор случайного леса обучался и проверялся на соответствующих восстановленных данных. Были определены диапазоны коэффициента сжатия, в которых сокращение объёма данных сопровождалось небольшим изменением общей точности классификации.

В настоящем эксперименте оценивается влияние видеокодека и параметров сжатия на точность поэлементной классификации ГСИ. Классификаторы обучаются на ГСИ без видеосжатия и затем применяются к восстановленным ГСИ при фиксированных параметрах. Сопоставляются два способа преобразования цифровых отсчётов в 8-битное представление. Для обоснования выбора параметров по ГСИ печатной калибровочной сцены сопоставляются изменения общей точности поэлементной классификации (ОА) на этой сцене и отдельной тестовой сцене при одинаковом варьировании параметров сжатия и способа преобразования цифровых отсчётов.

Использовались ГСИ печатной калибровочной сцены и отдельной тестовой сцены, полученные разработанным гиперспектрометром. Тестовая сцена описана в подразделе 1.4. Её эталонная разметка и состав пяти классов приведены в пункте 2.2.1. Использовались все 252 спектральных канала. ГСИ тестовой сцены представлялось последовательностью из 2000 пространственных спектрограмм: ширина кадра составляла 960 отсчётов вдоль изображения щелевой диафрагмы, высота – 252 спектральных отсчёта. Порядок кадров соответствовал порядку сканирования.

Перед сжатием цифровые отсчёты нормировались на максимальное значение во всём ГСИ. При линейном преобразовании нормированные значения умножались на 255 и округлялись до ближайшего целого. Во втором варианте перед масштабированием и округлением применялось степенное преобразование с показателем $\frac{1}{\gamma}$, где $\gamma = 2,2$. В спецификации PNG такое преобразование нормированных значений перед квантованием используется для гамма-кодирования (gamma encoding) [115]. После декодирования значения делились на 255; для варианта с гамма-кодированием дополнительно выполнялось обратное степенное преобразование с показателем γ . Базовым вариантом служило ГСИ после преобразования в 8-битное представление и обратно без видеосжатия.

Сопоставлялись Motion-JPEG (MJPEG), H.264/AVC и H.265/HEVC. В RFC 2435 термином Motion-JPEG обозначается последовательность независимо закодированных JPEG-изображений [116]. Стандарты H.264/AVC и H.265/HEVC предусматривают предсказание блоков по уже восстановленным данным текущего или других кадров, а также преобразование и квантование разности между исходными и предсказанными значениями [117], [118].

Кодирование выполнялось средствами FFmpeg. Для H.264 и H.265 использовались реализации libx264 и libx265 в режиме CRF (Constant Rate Factor) с заданным уровнем качества [119]. Для обеих реализаций был выбран

набор настроек `preset = medium`. Для MJPEG изменялся параметр качества `q:v` [120]. Заданные значения параметров приведены в таблице 2.17.

Таблица 2.17 – Видеокодеки и исследуемые значения параметров

Видеокодек	Реализация FFmpeg	Параметр	Заданные значения
H.264/AVC	libx264	CRF	18, 26, 34, 42, 48
H.265/HEVC	libx265	CRF	18, 26, 34, 42, 48
MJPEG	mjpeg	q:v	2, 8, 16, 32, 45

Входные пространственные спектрограммы передавались в FFmpeg в виде одноканальных 8-битных изображений в формате «gray8» [121]. Для H.264 и H.265 использовалось представление «yuv444p», а для MJPEG – «yuvj444p». В обоих случаях изображение содержит яркостную компоненту Y и две цветоразностные компоненты Cb и Cr. В представлении YCbCr компоненты Cb и Cr определяются через разности между синей и яркостной, красной и яркостной компонентами соответственно [122].

Обозначение 4:4:4 указывает на одинаковое разрешение трёх компонент: каждому отсчёту Y соответствуют по одному отсчёту Cb и Cr, «p» обозначает раздельное хранение компонент в отдельных плоскостях. В рассматриваемых форматах каждая компонента представлена 8-битными значениями [121]. Спектральные каналы ГСИ при этом располагаются по вертикальной оси кадра; компоненты Y, Cb и Cr относятся к формату видеокодирования.

Для H.264 и H.265 параметр `color_range = rc` задавал полный диапазон яркостных значений от 0 до 255. В формате «yuvj444p» полный диапазон JPEG предусмотрен самим форматом [121]. Обозначения `rc`, `jrcg` и `full` в FFmpeg соответствуют полному диапазону [119]. Видеопоследовательность записывалась с частотой 30 кадров/с. После декодирования средствами OpenCV использовалась первая компонента BGR-кадра – синяя компонента B; порядок B, G, R принят в OpenCV для представления цветных изображений

[122]. Из полученных двумерных кадров восстанавливался исходный порядок осей ГСИ.

Для классификации использовались 1DCNN, NM3DCNN и SSFTT, описанные в пункте 2.2.2. В каждой сцене размеченные пиксели каждого класса случайно разделялись на обучающую, валидационную и тестовую выборки в долях 10, 5 и 85 % соответственно. Разбиение сохранялось для всех вариантов преобразования и сжатия. Пиксели фона с меткой 0 исключались из обучения и расчёта классификационных показателей.

Для каждой сцены при линейном преобразовании и гамма-кодировании каждый классификатор независимо обучался на соответствующем базовом ГСИ без видеосжатия. Обучение выполнялось в течение 25 эпох методом Adam [110]. Для 1DCNN размер пакета данных составлял 512, шаг обучения – 10^{-2} ; для NM3DCNN и SSFTT – 256 и 10^{-3} соответственно. Пространственно-спектральные классификаторы использовали входные фрагменты размером 5×5 пикселей. Параметры классификатора выбирались по наибольшей взвешенной F1-мере на валидационной выборке.

Классификатор применялся к базовому ГСИ и к ГСИ, восстановленным после сжатия каждым видеокодеком при всех заданных значениях параметров. В пределах каждой сцены оценка выполнялась на одной и той же тестовой выборке по общей точности классификации ОА и взвешенной F1-мере, определённым в пунктах 1.2.1 и 1.5.2 соответственно. Представленные значения ОА являются средними по трём запускам. Показатели разброса не приводятся.

2.3.2 Результаты классификации при различных параметрах сжатия

На рисунке **Ошибка! Источник ссылки не найден.** представлены зависимости ОА от параметров сжатия для ГСИ печатной калибровочной сцены (а) и тестовой сцены из пяти классов (б). Показаны сочетания 1DCNN с H.264, NM3DCNN с H.265 и SSFTT с MJPEG. Для каждого сочетания сопоставлены линейное преобразование и гамма-кодирование цифровых

отсчётов. Обозначение «с гамма-коррекцией» на графиках соответствует описанному в пункте 2.3.1 варианту с гамма-кодированием и обратным преобразованием после декодирования.

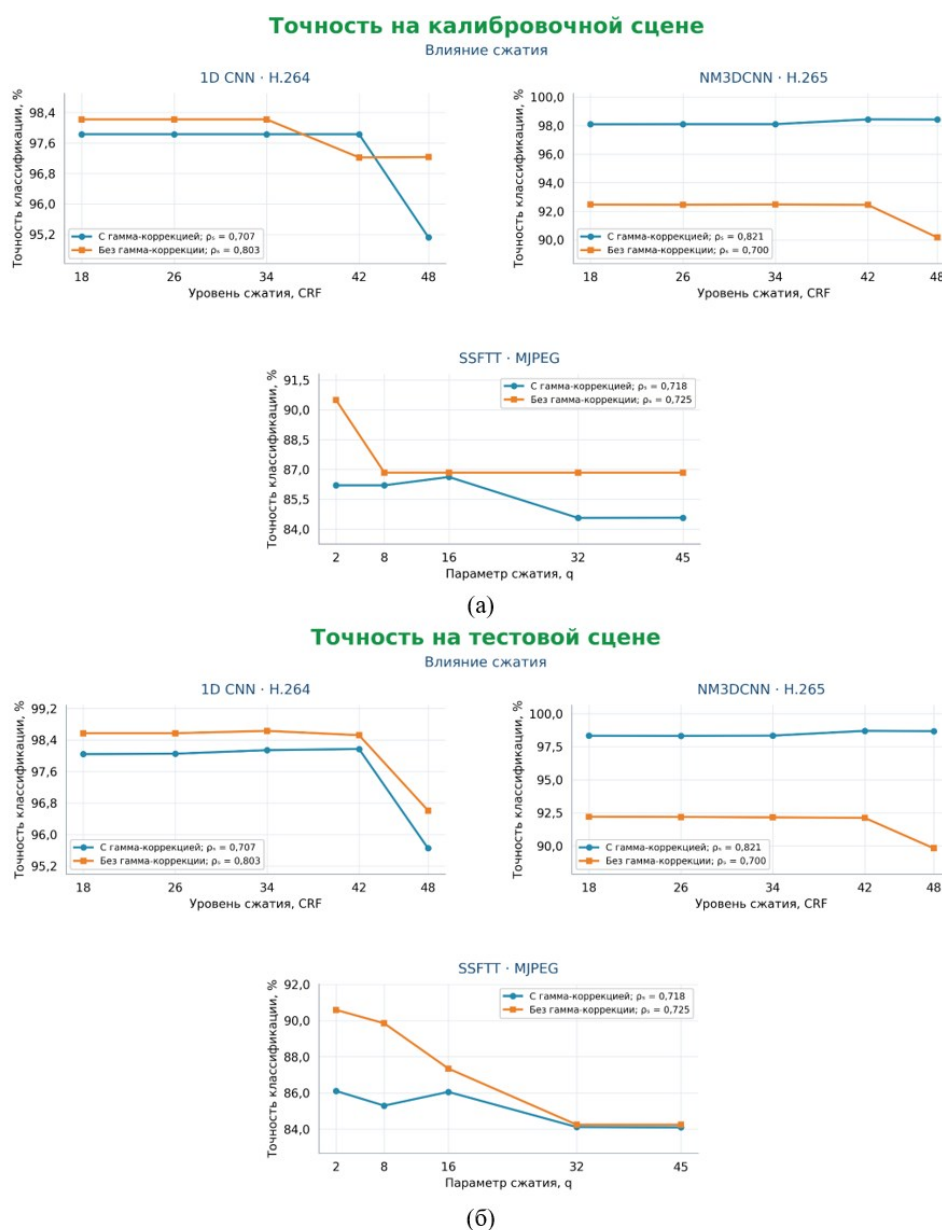


Рисунок 2.21 – Зависимости ОА от параметров сжатия для ГСИ печатной калибровочной (а) и тестовой (б) сцен; средние значения по трём запускам

Для сочетания 1DCNN с H.264 при гамма-кодировании ОА на обеих сценах изменяется слабо при увеличении CRF от 18 до 42 и снижается при CRF = 48. При линейном преобразовании зависимости различаются: на калибровочной сцене снижение наблюдается уже при CRF = 42, после чего точность практически сохраняется; на тестовой сцене наиболее выраженное снижение происходит при CRF = 48. Для сочетания NM3DCNN с H.265 при

гамма-кодировании ОА на обеих сценах слабо возрастает при переходе к CRF = 42 и далее практически не изменяется. При линейном преобразовании на обеих сценах наблюдается снижение ОА при CRF = 48.

Для сочетания SSFTT с MJPEG при линейном преобразовании ОА на калибровочной сцене снижается преимущественно при переходе от $q = 2$ к $q = 8$, а затем практически сохраняется. На тестовой сцене снижение продолжается до предпоследнего исследованного значения q . При гамма-кодировании зависимости на обеих сценах немонотонны; при переходе к двум последним значениям q точность становится ниже, чем при $q = 16$. На каждой сцене значения ОА в двух последних режимах практически совпадают. Таким образом, на графиках прослеживается сходство общих тенденций, однако положение отдельных участков снижения и их выраженность различаются.

Сопоставление представленных зависимостей позволяет выделить общие особенности изменения ОА на калибровочной и тестовой сценах при варьировании параметров сжатия. Возможность выбора параметров по ГСИ печатной калибровочной сцены оценивается по воспроизведению этих особенностей на отдельной тестовой сцене; совпадение абсолютных значений ОА при этом не предполагается.

2.4 Влияние времени экспозиции и усиления на точность классификации

В работе [102] исследовано совместное влияние времени экспозиции и усиления на отношение сигнал/шум: рассмотрены сокращение времени экспозиции с увеличением усиления при сохранении среднего цифрового отсчёта и изменение усиления при постоянном времени экспозиции. В работе [123] сопоставлены увеличение усиления, увеличение времени экспозиции и их совместное изменение при видеосъёмке в условиях недостаточной освещённости. Результаты оценивались после нейросетевого восстановления видео с учётом шума и размытия движения.

В настоящей главе рассматриваются две серии вычислительных экспериментов: сокращение времени экспозиции с увеличением усиления и изменение усиления при фиксированном времени экспозиции. Исходными данными служили ГСИ печатной калибровочной сцены и отдельной тестовой сцены из пяти классов. Тестовая сцена описана в подразделе 1.4; состав классов и эталонная разметка приведены в пункте 2.2.1. Для обеих сцен использовались ГСИ разработанного гиперспектрометра с 252 спектральными каналами после компенсации тёмного сигнала и пространственной коррекции.

Для обеих сцен моделировались одинаковые изменения времени экспозиции и усиления и оценивалась соответствующая точность поэлементной классификации. Сопоставление зависимостей ОА проводилось для проверки того, воспроизводится ли на тестовой сцене характер изменения точности, наблюдаемый на печатной калибровочной сцене. Такое сопоставление позволяет оценить возможность выбора параметров съёмки по ГСИ печатной калибровочной сцены.

Для моделирования использовалась пуассоновско-гауссовская модель шума. В работе [59] пуассоновская составляющая описывает статистику регистрации фотонов, а гауссовская – независимые от сигнала шумовые составляющие. Также различаются шум до усиления и шум после усиления.

Через x_0 обозначим исходный цифровой отсчёт, полученный при времени экспозиции $t_0 = 100$ мс и принятом относительном коэффициенте усиления $G_0 = 1$. Для каждого отсчёта ожидаемое число фотоэлектронов при времени экспозиции t рассчитывалось как:

$$\mu_e(t) = q_e x_0 \frac{t}{t_0}, \quad (2.23)$$

где q_e – модельный коэффициент пересчёта цифровых отсчётов в число электронов. Значение сигнала до замены отрицательных отсчётов нулём определялось выражением:

$$\tilde{x}_{t,G} = \frac{G}{G_0 q_e} (N_t + n_{\text{pre}}) + n_{\text{post}}, \quad (2.24)$$

где N_t – случайная величина с распределением Пуассона и математическим ожиданием $\mu_e(t)$; n_{pre} и n_{post} – независимые гауссовские шумовые составляющие до и после усиления с нулевыми средними и стандартными отклонениями σ_{pre} и σ_{post} соответственно; G – относительный линейный коэффициент усиления.

В моделировании принимались $q_e = 16$ электронов на единицу цифрового отсчёта, $\sigma_{\text{pre}} = 3$ электрона и $\sigma_{\text{post}} = 0,5$ единицы цифрового отсчёта. Тёмный ток принимался равным нулю. Отрицательные значения сигнала заменялись нулём. Квантование и ограничение верхним пределом АЦП не применялись. Параметры шума задавались одинаковыми для всех спектральных каналов и значений усиления. Исходное ГСИ использовалось как оценка среднего сигнала для формирования шумовых реализаций.

В каждой из двух серий для каждой сцены и каждой пары времени экспозиции и усиления классификаторы 1DCNN, NM3DCNN и SSFTT обучались независимо с новой инициализацией параметров. Это относилось и к общему для двух серий режиму на печатной калибровочной сцене с экспозицией 0,357 мс и усилением 28. В каждой сцене размеченные пиксели каждого класса случайно разделялись на обучающую, валидационную и тестовую выборки в долях 10, 5 и 85 % соответственно, как в эксперименте со сжатием (пункт 2.3.1). Разбиение сохранялось для всех режимов моделирования. Параметры классификатора выбирались по наибольшей взвешенной F1-мере на валидационной выборке. На тестовой выборке рассчитывались ОА и взвешенная F1-мера. Далее приведены зависимости ОА для печатной калибровочной и тестовой сцен. Значения ОА являются средними по трём запускам; показатели разброса не приводятся. Примеры изображений спектрального канала и средних спектральных векторов относятся к тестовой сцене из пяти классов.

2.4.1 Сокращение времени экспозиции с увеличением усиления

В первой постановке экспериментов для обеих сцен время экспозиции последовательно уменьшалось с 10 до 0,357 мс, а коэффициент усиления увеличивался от 1 до 28. Значение времени экспозиции 0,357 мс для усиления 28 приведено с округлением. Пары параметров выбирались из условия $\frac{t}{t_0} \frac{G}{G_0} =$

1. До замены отрицательных значений нулём это обеспечивает сохранение математического ожидания цифрового отсчёта в модели (2.24). Исследованные пары приведены в таблице 2.18.

Таблица 2.18 – Время экспозиции и коэффициент усиления в первой постановке

Время экспозиции, мс	Относительный коэффициент усиления
10	1
5	2
2,5	4
1,25	8
0,625	16
0,357	28

После моделирования цифровые отсчёты нормировались по спектру белого эталона Spectralon, приведённому к той же паре времени экспозиции и усиления. Обозначив исходный спектр эталона через $W_0(k)$, его значения для выбранного режима запишем как $W_{t,G}(k) = \frac{tG}{t_0G_0} \max\{W_0(k), 2\}$. В данном эксперименте шум белого эталона дополнительно не моделировался. Значение элемента ГСИ после нормирования по белому эталону рассчитывалось по формуле:

$$R_{t,G}(i, j, k) = \frac{\max\{\tilde{x}_{t,G}(i, j, k), 0\}}{W_{t,G}(k) + \varepsilon}, \quad (2.25)$$

где $\varepsilon = 10^{-8}$ – положительная постоянная в знаменателе. Перед подачей в классификатор значения элементов ГСИ нормировались и приводились к диапазону $[0, 1]$.

Для каждой сцены при каждой паре параметров формировалась одна шумовая реализация ГСИ, по которой обучались и проверялись три классификатора. Обучение выполнялось в течение 15 эпох методом Adam; размеры пакетов данных, шаги обучения и размеры входных фрагментов соответствовали указанным в пункте 2.3.1.

На рисунке 2.22 для тестовой сцены сопоставлены изображения одного спектрального канала и средние спектральные векторы двух классов. При сокращении времени экспозиции с увеличением усиления возрастает неоднородность изображения спектрального канала.

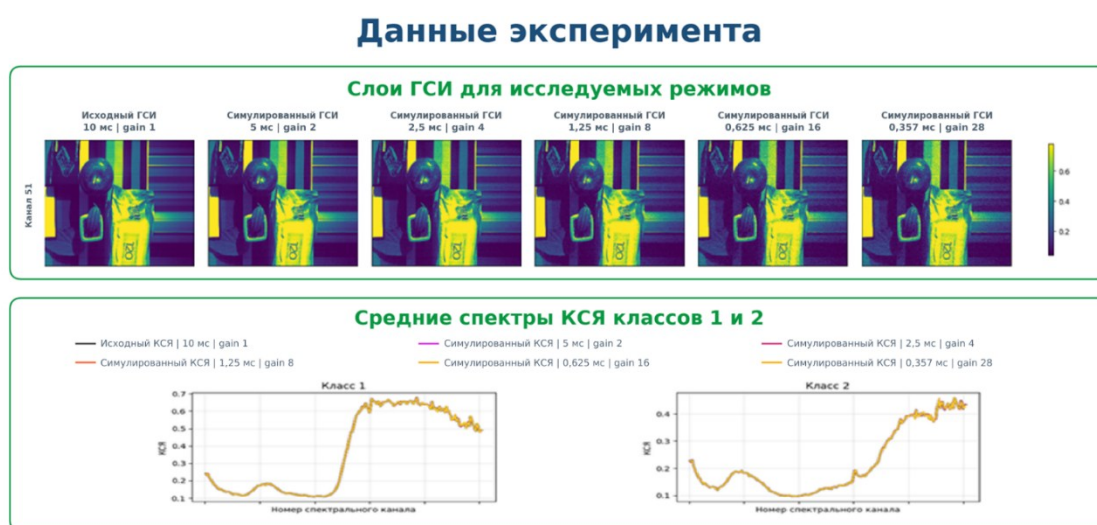
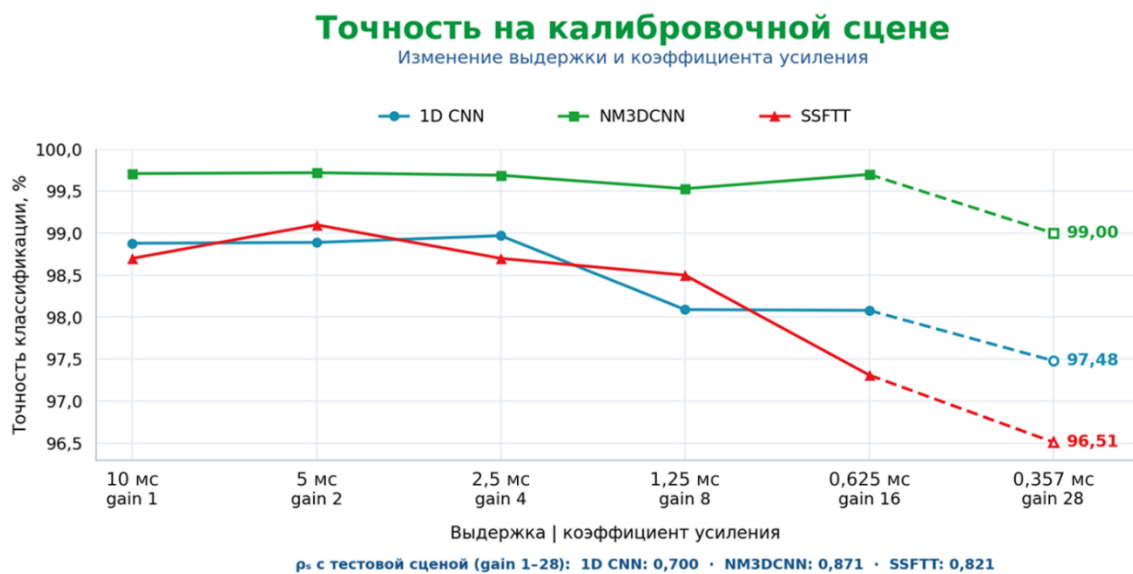


Рисунок 2.22 – Изображения спектрального канала и средние спектральные векторы при сокращении времени экспозиции с увеличением усиления

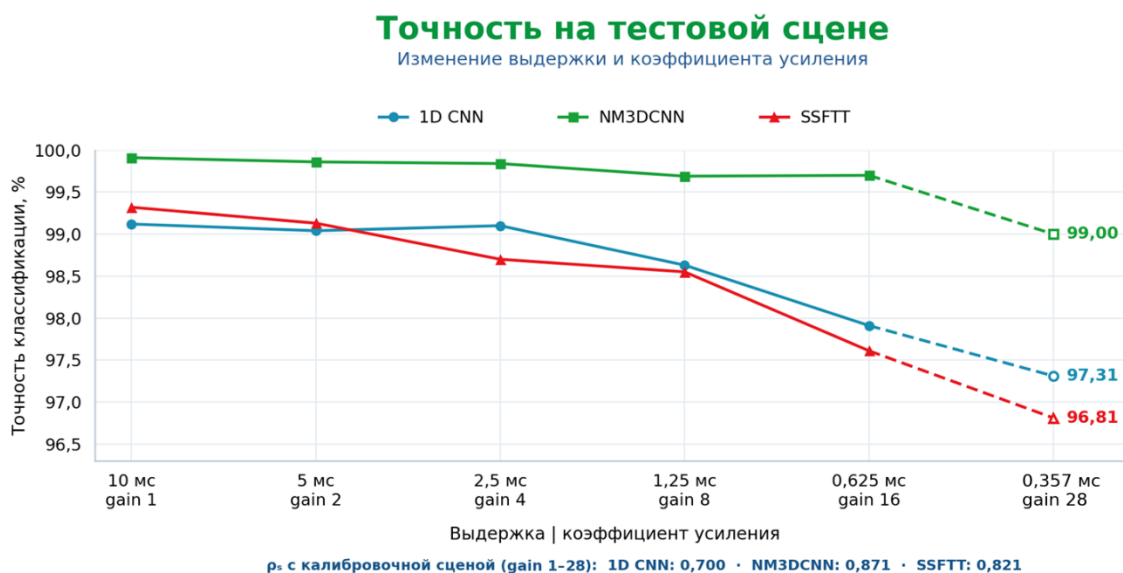
Зависимости ОА от времени экспозиции и усиления представлены на рисунке 2.23 для печатной калибровочной сцены (а) и тестовой сцены (б). На обеих сценах наиболее выраженное снижение точности при переходе к коротким экспозициям наблюдается для 1DCNN и SSFTT. При этом отдельные участки зависимостей различаются: для SSFTT на тестовой сцене

ОА последовательно снижается, тогда как на калибровочной сцене сначала возрастает при переходе к экспозиции 5000 мс. Для 1DCNN на калибровочной сцене снижение происходит преимущественно при переходе от 2500 к 1250 мс; на тестовой сцене оно продолжается и при дальнейшем сокращении экспозиции. ОА классификатора NM3DCNN изменяется слабее на обеих сценах.

При экспозиции 0,625 мс и усилении 16 на графиках для калибровочной сцены указаны значения ОА 98,08; 99,70 и 97,31 % для 1DCNN, NM3DCNN и SSFTT соответственно. Для тестовой сцены соответствующие значения составляют 97,91; 99,70 и 97,61 %. Представленные зависимости имеют сходную общую тенденцию при различиях между отдельными режимами.



(a)



(б)

Рисунок 2.23 – Зависимости ОА от времени экспозиции и усиления для ГСИ печатной калибровочной (а) и тестовой (б) сцен при усилении 1–28; средние значения по трём запускам

При экспозиции 0,357 мс и усилении 28 средняя ОА на печатной калибровочной сцене составила 97,48; 99,00 и 96,51 % для 1DCNN, NM3DCNN и SSFTT соответственно, а на тестовой сцене – 97,31; 99,00 и 96,81 %. По сравнению с режимом 0,625 мс и усилением 16 ОА снизилась на 0,60; 0,70 и 0,80 процентного пункта для 1DCNN, NM3DCNN и SSFTT соответственно на каждой из двух сцен.

2.4.2 Изменение усиления при фиксированном времени экспозиции

Во втором эксперименте время экспозиции составляло 0,357 мс, а относительный коэффициент усиления принимал значения 1, 2, 4, 8, 16 и 28. По модели (2.23), (2.24) при неизменной экспозиции ожидаемое число фотоэлектронов остаётся постоянным, а средний цифровой отсчёт до замены отрицательных значений нулём возрастает пропорционально усилению.

Нормирование по белому эталону и подготовка данных к классификации выполнялись по схеме пункта 2.4.1.

На рисунке 2.24 средние спектральные векторы двух классов представлены в общей фиксированной шкале. С увеличением усиления их амплитуда возрастала при сохранении основных особенностей формы.

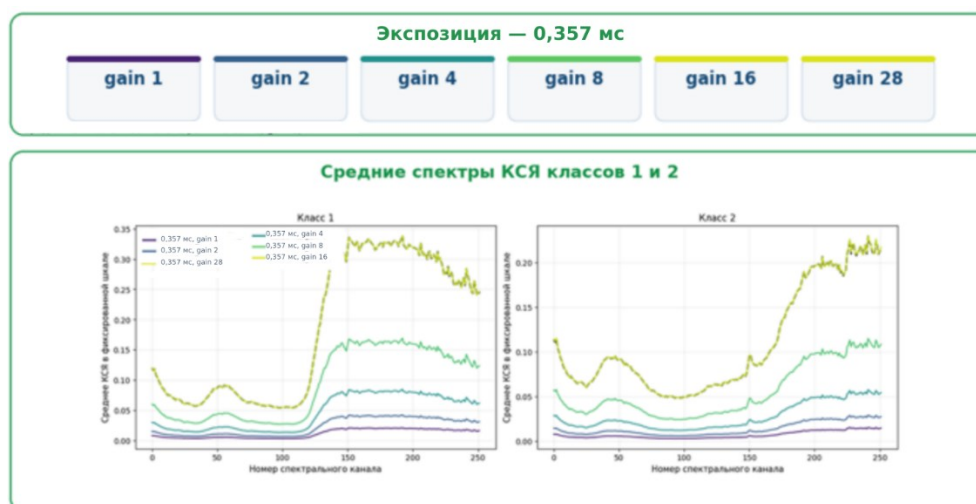
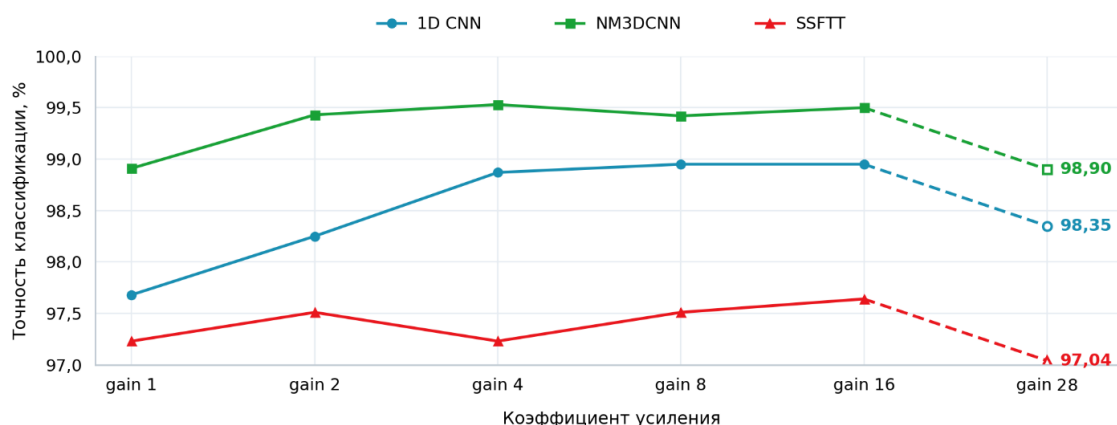


Рисунок 2.24 – Средние спектральные векторы при времени экспозиции 0,357 мс и коэффициентах усиления 1, 2, 4, 8, 16 и 28

На рисунке 2.25 приведены зависимости ОА от коэффициента усиления при фиксированной экспозиции 0,357 мс для печатной калибровочной сцены (а) и тестовой сцены (б). Для NM3DCNN основной прирост наблюдается при увеличении усиления с 1 до 2, после чего до усиления 16 точность изменяется слабо. Для 1DCNN точность возрастает до усиления 8; при дальнейшем увеличении усиления до 16 на калибровочной сцене она практически сохраняется, а на тестовой снижается. Для SSFTT на тестовой сцене при увеличении усиления от 1 до 16 наблюдается последовательное повышение ОА, тогда как на калибровочной сцене при усилении 4 имеется локальное снижение. При переходе от усиления 16 к усилению 28 ОА всех трёх классификаторов снижается на обеих сценах.

Точность на калибровочной сцене

Влияние усиления при недостаточной экспозиции

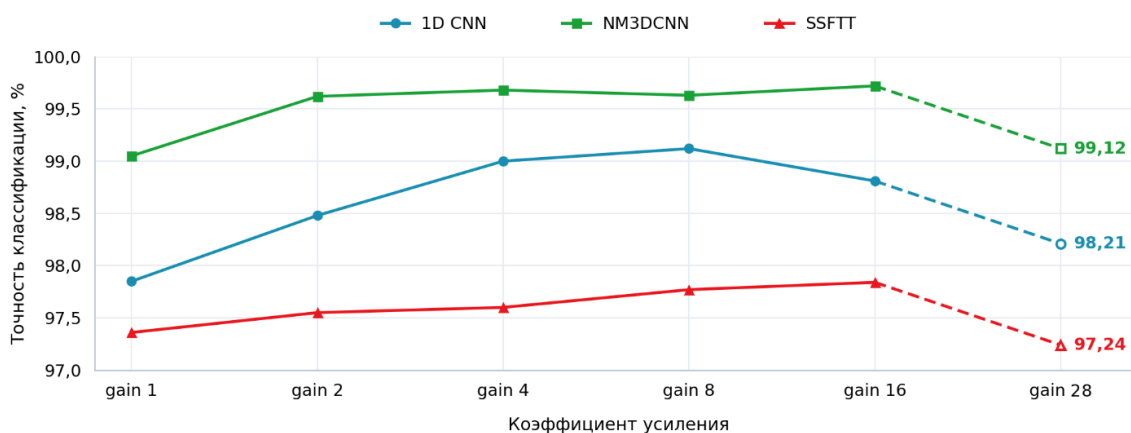


ρ с тестовой сценой (gain 1-28): 1D CNN: 0,821 · NM3DCNN: 0,830 · SSFTT: 0,738

(a)

Точность на тестовой сцене

Влияние усиления при недостаточной экспозиции



ρ с калибровочной сценой (gain 1-28): 1D CNN: 0,821 · NM3DCNN: 0,830 · SSFTT: 0,738

(б)

Рисунок 2.25 – Зависимости ОА от усиления при экспозиции 0,357 мс для ГСИ печатной калибровочной (а) и тестовой (б) сцен при коэффициентах усиления 1, 2, 4, 8, 16 и 28; средние значения по трём запускам

При усилении 16 на графиках для калибровочной сцены указаны значения ОА 98,95; 99,50 и 97,64 % для 1DCNN, NM3DCNN и SSFTT соответственно. Для тестовой сцены соответствующие значения составляют 98,81; 99,72 и 97,84 %.

При усилении 28 средняя ОА на калибровочной сцене составила 98,35; 98,90 и 97,04 % для 1DCNN, NM3DCNN и SSFTT соответственно, а на тестовой сцене – 98,21; 99,12 и 97,24 %. Для каждого классификатора на

каждой сцене снижение относительно усиления 16 составило 0,60 процентного пункта. В рассмотренных условиях усиление 28 не рекомендуется для поэлементной классификации. Выбор усиления целесообразно ограничить диапазоном от 1 до 16, определяя конкретное значение по точности классификации; при этом для 1DCNN наибольшая ОА получена при усилении 8.

В обеих сериях графики отражают сходство общих тенденций изменения ОА на печатной калибровочной и тестовой сценах. Отдельные режимы могут сопровождаться различающимися локальными изменениями точности. При оценке возможности выбора параметров съёмки по печатной калибровочной сцене учитывается характер всей исследованной зависимости ОА от параметров.

2.5 Обсуждение результатов

Эксперимент со сжатием данных видеокодеками характеризует устойчивость предварительно обученного классификатора к изменениям входных данных. Представленные зависимости ОА на печатной калибровочной и тестовой сценах имеют сходные общие тенденции, хотя выраженность снижения точности и отдельные участки зависимостей различаются. В экспериментах с временем экспозиции и усилением классификаторы обучались независимо для каждой сцены и каждого режима. Сокращение экспозиции с компенсирующим увеличением усиления сопровождалось общим снижением ОА для 1DCNN и SSFTT, тогда как точность NM3DCNN изменялась слабее. При фиксированной экспозиции 0,357 мс переход от единичного усиления к усилению 16 сопровождался повышением ОА всех трёх классификаторов на обеих сценах, при этом зависимости не во всех случаях были монотонными. Дальнейшее увеличение усиления до 28 снизило ОА каждого классификатора на 0,60 процентного пункта на обеих сценах. В серии с сокращением экспозиции переход от режима 0,625 мс и усиления 16 к режиму 0,357 мс и усиления 28 также

сопровождался снижением ОА на калибровочной сцене: на 0,60; 0,70 и 0,80 процентного пункта для 1DCNN, NM3DCNN и SSFTT соответственно. Для рассмотренных условий значение 16 принято как «верхняя граница рекомендуемого диапазона усиления». Конкретное значение внутри этого диапазона выбирается с учётом классификатора и времени экспозиции по результатам классификации ГСИ печатной калибровочной сцены. Сопоставление характера изменения ОА на двух сценах позволяет оценивать возможность выбора параметров сжатия и съёмки по ГСИ печатной калибровочной сцены. Результаты представлены средними значениями ОА по трём запускам.

2.6 Выводы к главе 2

1. Разработаны алгоритмы настройки спектральной предобработки ГСИ, совместно определяющие коэффициенты спектральной свёртки и программную поправку уровня чёрного аналитически либо с учётом задачи поэлементной классификации. В дополнительных исследованиях по валидационной ОА выбраны длина окна фильтра Савицкого–Голея (5 каналов) и значения λ комбинированной функции потерь: 0,1 при настройке с 1DCNN, 0,01 с NM3DCNN и 1 с SSFTT. Сопоставление результатов основной и дополнительных серий носит описательный характер, поскольку параметры обучения этих серий различались.

2. Среди рассмотренных вариантов спектральной предобработки наибольшие средние ОА и взвешенная F1-мера получены при настройке с 1DCNN по комбинированному критерию при $\lambda = 0,1$ и оценке с SSFTT: 97,37 % и 97,36 % соответственно. Приросты относительно варианта без дополнительной предобработки составили 3,32 и 3,18 процентного пункта. PSNR между средними спектральными векторами классов увеличился на 0,695 дБ до 24,104 дБ. При указанном сочетании классификаторов и комбинированном критерии настройки получена более высокая средняя ОА, чем при применении нейросетевых аналогов на основе моделей DeSpaWN и DeCWN в рассмотренных соответствующих конфигурациях.

3. При аналитическом определении параметров гребневой регрессией точность и взвешенная F1-мера превышали показатели варианта без дополнительной предобработки во всех пяти рассмотренных сочетаниях классификаторов. PSNR составил 24,016 дБ, что на 0,607 дБ выше значения без дополнительной предобработки. Для сочетания SSFTT–SSFTT гребневой регрессии соответствовало более высокое значение PSNR, чем комбинированному критерию при $\lambda = 1$ (24,016 и 23,677 дБ), при меньшей точности (0,9509 и 0,9622). Следовательно, наибольшему PSNR не всегда соответствовала наибольшая точность.

4. При применении настроенной предобработки совместно с SVM, kNN и MLP наибольшая средняя ОА для каждой модели получена после настройки с 1DCNN по комбинированному критерию при $\lambda = 0,1$. Приросты ОА относительно варианта без дополнительной предобработки составили 11,85; 0,72 и 1,22 процентного пункта соответственно. Взвешенная F1-мера повысилась на 13,38 процентного пункта для SVM и снизилась на 0,08 процентного пункта для kNN и MLP. Таким образом, показана возможность применения коэффициентов предобработки с классификаторами, отличными от использованного при настройке, с повышением ОА.

5. Оценены число параметров классификаторов и время их прямого прохода в заданных вычислительных конфигурациях. Для модуля предобработки с ядром спектральной свёртки длиной 7 каналов время применения к пакету данных из 4096 спектров составило $0,534 \pm 0,027$ мс при размещении входа на GPU и $1,757 \pm 0,112$ мс с передачей входа и результата между CPU и GPU. Классификатор в измерение времени предобработки не включался. Приведённые измерения не включают обучение, чтение ГСИ и построение пространственных фрагментов.

6. При изменении параметров сжатия на печатной калибровочной и тестовой сценах наблюдались сходные общие тенденции изменения точности. В большинстве рассмотренных сочетаний усиление сжатия сопровождалось снижением точности; для NM3DCNN с H.265 при гамма-кодировании наблюдалось небольшое повышение ОА. Выраженность и положение отдельных участков снижения точности на двух сценах различались.

7. В вычислительном эксперименте сокращение времени экспозиции с 10 до 0,357 мс при увеличении усиления с 1 до 28 сопровождалось снижением ОА на печатной калибровочной и тестовой сценах. На тестовой сцене снижение составило 1,81; 0,90 и 2,51 процентного пункта для 1DCNN, NM3DCNN и SSFTT соответственно. Сохранение среднего цифрового отсчёта при сокращении экспозиции не обеспечило сохранения точности классификации.

8. При фиксированном времени экспозиции 0,357 мс увеличение усиления с 1 до 16 повысило ОА на тестовой сцене на 0,96; 0,67 и 0,48 процентного пункта для 1DCNN, NM3DCNN и SSFTT соответственно. При дальнейшем увеличении усиления до 28 ОА всех трёх классификаторов на каждой из двух сцен снизилась на 0,60 процентного пункта относительно усиления 16. Наибольшая ОА для 1DCNN получена при усилении 8. В рассмотренных условиях выбор усиления целесообразно ограничить диапазоном 1–16, определяя конкретное значение по результатам классификации ГСИ печатной калибровочной сцены.

9. Методика формирования и применения печатной калибровочной сцены, разработанная в первой главе, использована для настройки спектральной предобработки ГСИ и исследования возможности выбора параметров съёмки и сжатия по точности классификации. Протокол настройки предобработки по калибровочной сцене и оценки её применения на отдельной тестовой сцене служит основой для экспериментального исследования обучаемой позиционно-зависимой коррекции отклика гиперспектрометра в третьей главе.

ГЛАВА 3. ОБУЧАЕМАЯ ПОЗИЦИОННО-ЗАВИСИМАЯ КОРРЕКЦИЯ ПРОСТРАНСТВЕННО-СПЕКТРАЛЬНЫХ ИСКАЖЕНИЙ ГИПЕРСПЕКТРАЛЬНЫХ ИЗОБРАЖЕНИЙ

В настоящей главе представлен метод обучаемой позиционно-зависимой коррекции ГСИ с использованием калибровочных слоёв. Разработанная в первой главе методика формирования и применения печатной калибровочной сцены используется для настройки коэффициентов коррекции с учётом точности поэлементной классификации. Корректирующие коэффициенты зависят от положения вдоль изображения щелевой диафрагмы и спектрального канала и определяются по спектральным и классификационным критериям. Коррекция выполняется умножением цифровых отсчётов на соответствующие калибровочные коэффициенты. Рассматриваются две постановки: коррекция относительно опорных спектров классов калибровочной сцены разработанного гиперспектрометра и коррекция по средним спектрам соответствующих классов Specim IQ при межприборном сопоставлении. В экспериментах влияние коррекции на точность поэлементной классификации оценивается на независимых тестовых сценах. Дополнительно первая постановка проверяется на ГСИ Specim FX.

3.1 Пространственно-спектральные искажения ГСИ и методы коррекции

3.1.1 Пространственно-спектральное рассогласование и неоднородность чувствительности

Спектральное рассогласование – изменение центральной длины волны канала вдоль координаты x изображения щели. Относительно опорной координаты x_{ref} смещение равно $d_{\text{sm},\lambda}(x) = \lambda(x) - \lambda(x_{\text{ref}})$ [16] (рисунок 3.1).

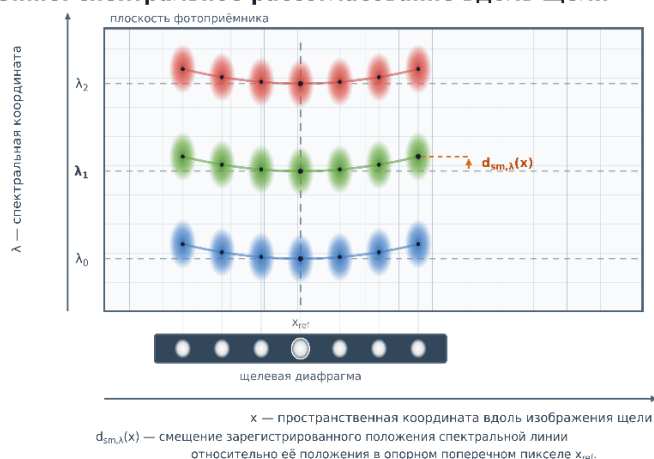
Smile: спектральное рассогласование вдоль щели

Рисунок 3.1 – Схема изменения центральной длины волны канала вдоль щели

Для гиперспектрометра Hyperion на спутнике EO-1 спектральное рассогласование в 41-м канале определяли по полосе поглощения кислорода при 762,6 нм [16]. Для лабораторных камер использовали спектральные особенности тестовых изображений [124] или положения эмиссионных линий [125], во втором случае учитывали также наклон линий из-за несоосности щели [125].

Межканальное пространственное рассогласование — смещение изображения участка сцены вдоль щели при изменении длины волны. Для координаты центра ФРТ $x_c(\lambda)$ и опорной длины волны λ_0 оно равно $d_{key}(\lambda) = x_c(\lambda) - x_c(\lambda_0)$ [16] (рисунок 3.2).

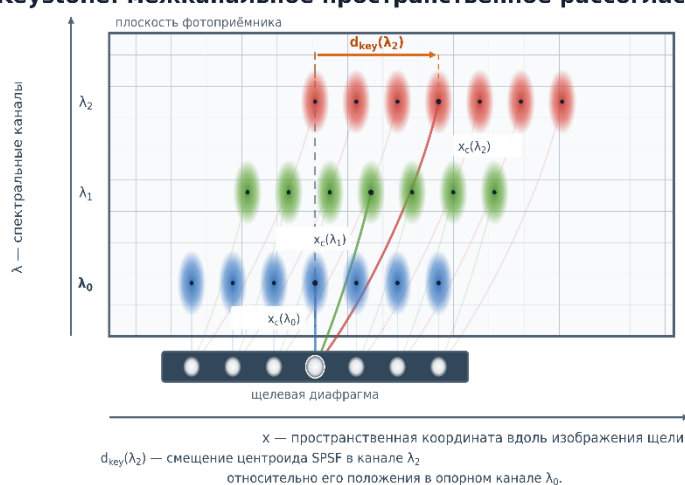
Keystone: межканальное пространственное рассогласование

Рисунок 3.2 – Схема межканального пространственного смещения

Для спутникового гиперспектрометра Hyperion смещения определяли относительно канала 31 (660,9 нм) [16], для лабораторных камер – по изображениям чёрных полос [124]. В вычислительном эксперименте кубическая сплайн-интерполяция дала лучший результат коррекции, чем линейная [126].

Пространственное рассогласование определяется не только смещением центра ФРТ, но и изменением её ширины и формы между спектральными каналами [20]. В работе [19] для совместной характеристики этих составляющих исследовано изображение полихроматического точечного источника, сформированное на фотоприёмной матрице щелевого гиперспектрометра. В работе [1] для разработанного гиперспектрометра по пространственной спектрограмме точечного источника установлены изменение формы ФРТ и отклонение расположения её центров от прямой при разных положениях источника вдоль щели.

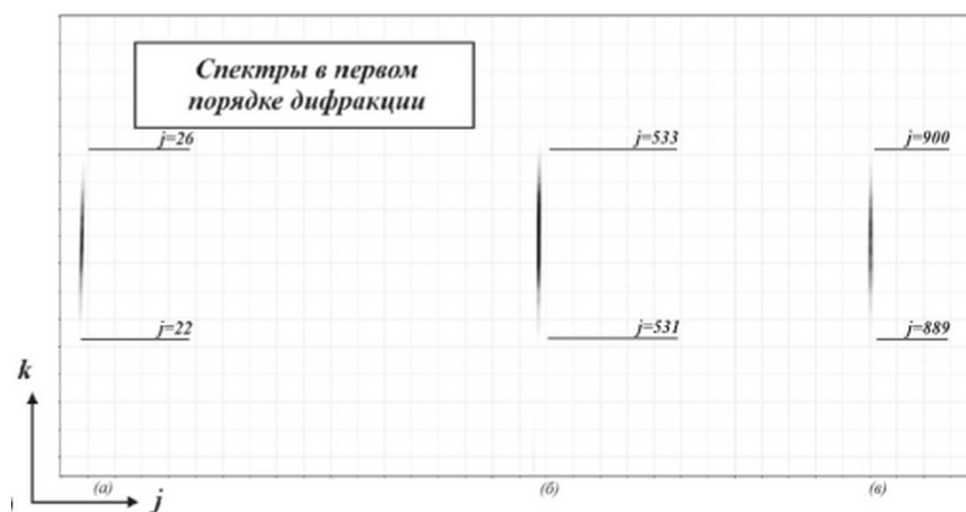


Рисунок 3.3 – Изменение формы ФРТ и отклонение расположения её центров от прямой при различных положениях точечного источника вдоль щели [1]

На рисунке 3.3 показано, что при изменении спектрального канала k изменяются координата j центра и форма пространственной ФРТ [1, 20]. Такое пространственное рассогласование следует отличать от изменения центральной длины волны канала в зависимости от координаты j вдоль щели [16].

Неоднородность чувствительности исследована по спектрам равномерно освещённого белого тестового объекта и монохроматического источника с длиной волны 640 нм. Положения локальных экстремумов спектров белого тестового объекта, усреднённых по 50 строчным срезам (рисунок 3.4), зависели от координаты j , тогда как положение максимума монохроматического источника от неё не зависело [1].

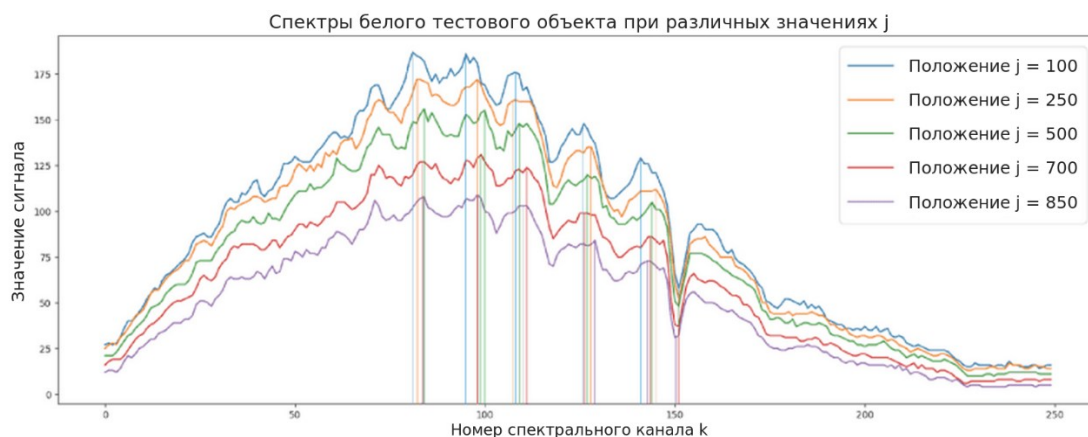


Рисунок 3.4 – Спектры белого тестового объекта при различных значениях координаты j с отмеченными локальными экстремумами; результат усреднения 50 строчных срезов (по [1])

После нормировки спектров в диапазон от 0 до 1 различия их формы и положений локальных экстремумов сохранялись: для каналов 60–145 смещение составляло 10–12 нм, а в области линии Фраунгофера А с длиной волны 760 нм – не более 2,5 нм.

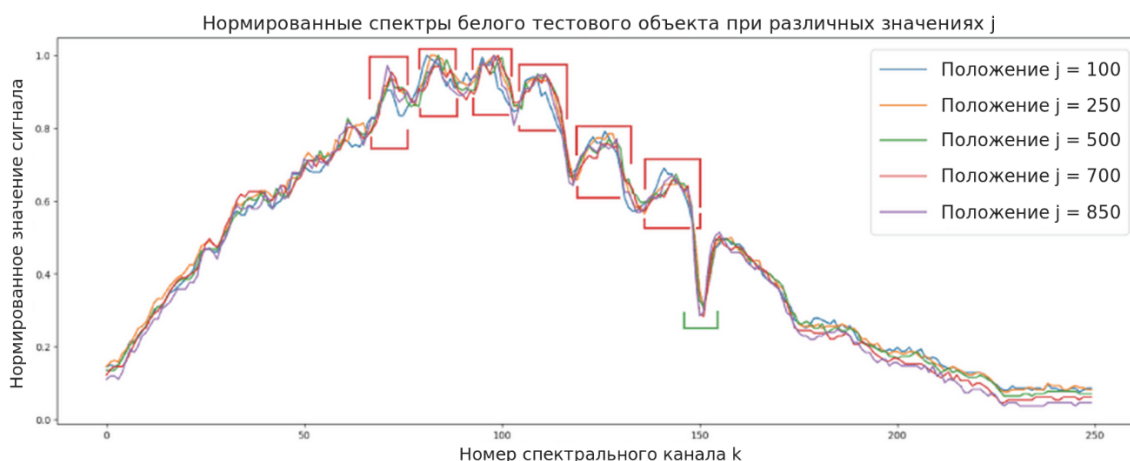


Рисунок 3.5 – Нормированные в диапазон от 0 до 1 спектры белого тестового объекта при различных значениях координаты j : красным выделены области значительного смещения локальных экстремумов, зелёным – область без смещений [1]

В модели [127] свет каждого элемента сцены равномерно заполняет соответствующий элемент щели. Пространственные смещения и размытие учитываются коэффициентами смешения q_{mn} :

$$E_m^R = \sum_{n=1}^N q_{mn} E_n, \quad m = 1, \dots, M, \quad (3.1)$$

где E_m^R – энергия, зарегистрированная m -м пикселем; q_{mn} – доля энергии n -го элемента сцены, регистрируемая этим пикселем; E_n – энергия, поступающая от данного элемента сцены; $n = 1, \dots, N$, $m = 1, \dots, M$; N и M – числа элементов сцены и пикселей приёмника. Верхний индекс R обозначает зарегистрированную энергию. Систему решают отдельно для каждого канала.

В работе [42] геометрическое отображение объединено с позиционно-зависимым размытием:

$$I(u, w) = \iint_{\partial D(u, w)} [T(g(\theta, \lambda)) * \text{SRF}(v, \omega) * \text{SiRF}(v, \omega)](v, \omega) dv d\omega + \varepsilon(u, w), \quad (3.2)$$

где $I(u, w)$ – измеренное значение; $\partial D(u, w)$ – область соответствующего пикселя; $g(\theta, \lambda)$ – распределение сигнала объекта; T – отображение в координаты v, ω плоскости изображения; $\text{SRF}(v, \omega)$ и $\text{SiRF}(v, \omega)$ – спектральная и пространственная функции отклика; $*$ – свёртка; $\varepsilon(u, w)$ – ошибка модели.

Спектральная и пространственная функции отклика считаются постоянными лишь локально [42], их оценивают по калибровочным изображениям и используют при устранении размытия [43].

Рассмотренные модели описывают смещения спектральной и пространственной шкал, пространственное смешение и изменение функций отклика.

3.1.2 Калибровка по объектам с известными спектральными характеристиками

Для авиационного гиперспектрометра AVIRIS-NG радиометрическое преобразование имеет вид [41]:

$$L_{\text{FPA}}(\lambda, x) = [D(\lambda, x) - \alpha(\lambda, x)]\beta(\lambda, x), \quad (3.3)$$

где $L_{\text{FPA}}(\lambda, x)$ – энергетическая яркость; $D(\lambda, x)$ – цифровой отсчёт; $\alpha(\lambda, x)$ – аддитивная поправка; $\beta(\lambda, x)$ – коэффициент радиометрического преобразования. Здесь $\alpha(\lambda, x)$ учитывает тёмный сигнал и сдвиг нулевого уровня, а $\beta(\lambda, x)$ – спектральную чувствительность и её позиционно-зависимую составляющую $\beta_{\text{flat}}(\lambda, x)$ [41].

По изображению белого эталона учитывают совместное влияние освещения и чувствительности прибора, а по тёмному кадру – аддитивный сигнал. Нормировка имеет вид [128]:

$$R_{\text{rel}}(j, k) = \frac{S(j, k) - D(j, k)}{W(j, k) - D(j, k)}, \quad (3.4)$$

где $R_{\text{rel}}(j, k)$ – сигнал объекта, нормированный по белому эталону после вычитания тёмного сигнала; j – номер пространственного пикселя вдоль щели; k – номер спектрального канала; $S(j, k)$, $W(j, k)$ и $D(j, k)$ – отсчёты исследуемого объекта, белого эталонного объекта и тёмного кадра соответственно. При одинаковых условиях освещения и регистрации объекта и эталона и известном паспортном коэффициенте отражения белого эталона $\rho_W(k)$ оценка спектрального коэффициента отражения объекта получается умножением правой части (3.4) на $\rho_W(k)$; такая запись приведена в [129] и в [130].

В лабораторных камерах Specim FX10 и FX17 изображение белого эталона и тёмный кадр усредняли по выборке до 200 кадров при экспозиции объекта [128]. В 14 системах межлабораторного исследования эталон заполнял

поле зрения [131], а в промышленной установке использовалась белая полоса по ширине конвейера [132].

Цветовые шкалы используют для оценки спектральной и колориметрической точности по десяти полям [133] и остаточной паразитной засветки по серым полям [134], а объект с тремя полями основных цветов и измеренными спектрами – для калибровки спектральной шкалы [129].

Для выравнивания чувствительности вдоль щели использован равномерно освещённый белый объект, заполняющий поле зрения; коэффициент отражения определён по светло-серому полю цветовой шкалы [130].

3.1.3 Условия регистрации и переносимость калибровочных коэффициентов

Калибровочные коэффициенты зависят от режима регистрации. Для щелевого спектрометра показана зависимость отклика от экспозиции и диафрагмы, а при температуре выше 20 °C – зависимость тёмного сигнала от экспозиции и температуры [135].

Для гиперспектрометра CHRIS на спутнике PROBA по 387 изображениям при 1–9 °C установлена температурная зависимость профиля изображения щели [17]. Для камеры Specim FX17e отмечена зависимость тёмного сигнала от экспозиции; объект, эталон и тёмный кадр следует регистрировать при одной экспозиции [7].

Нормирование по белому эталону не устраняет различия освещения объекта и эталона при съёмке растений [136] и коры деревьев [137], а также остаточную паразитную засветку камеры Specim IQ [134].

Линейную нормировку (3.4) следует применять после проверки отклика: у камеры Specim FX17e обнаружены отклонения вблизи насыщения [7], у камеры HySpex VNIR-1600 нелинейность правой половины приёмника превышала 15 % [44]. Для гиперспектрометра спутника EnMAP

предусмотрена калибровка примерно по 10–100 уровням входного излучения и дополнительная поправка при остаточной нелинейности более 5 % [138].

В работе [18] для AISA DUAL и Hyperion поправки, найденные по разным сценам, были коррелированы, однако рекомендовано выполнять радиометрическое масштабирование каждой сцены, не предполагая неизменность коррекции. Перенос калибровочных коэффициентов требует проверки соответствия режима регистрации, температуры, освещения и линейности отклика условиям их определения.

Переносимость фиксированной радиометрической калибровки непосредственно проверена для щелевого гиперспектрометра AisaEAGLE-II. Параметры радиометрической калибровки Vic2, определённые по наземным опорным объектам с измеренными характеристиками отражения на полётной линии Msh2, без повторной оценки применяли к линии Msh1 и к изображению Por4, зарегистрированному в другой местности через неделю [139]. Для Msh1 точность соответствовала калибровке Vic1, определённой по этой же линии, а для Por4 применение Vic2 несколько повысило точность относительно лабораторной калибровки [139]. Атмосферные параметры при этом определяли отдельно для каждого изображения; следовательно, проверялся перенос параметров радиометрической калибровки датчика, но не всей процедуры коррекции [139].

3.1.4 Нейросетевые методы коррекции ГСИ

В работе [57] автоматическая компенсация пространственно однородного освещения представлена как оценивание отражательной способности по исходному ГСИ. Эталон для Specim IQ формировали последовательной съёмкой сцены и белой панели с вычитанием темнового сигнала. Для оценивания использована нейросетевая модель на основе трансформера, учитывающая зависимости между спектральными каналами и общие характеристики освещения сцены; критерий обучения – L1-ошибка.

В работе [140] предложено прогнозировать по некалиброванному ГСИ пространственно-разрешённое изображение белого эталона и выполнять поэлементное деление входного сигнала на прогноз. Трёхмерный свёрточный автоэнкодер обучали по среднему квадрату ошибки восстановления белого эталона. При независимой проверке по ColorChecker среднее косинусное сходство с эталонными спектрами составило 0,9905; для физического белого эталона – 0,9945.

В рассмотренных нейросетевых методах критерием обучения служит ошибка воспроизведения опорной радиометрической величины: коэффициента отражения в методе, предложенном в работе [57], либо пространственно-разрешённого изображения белого эталона для трёхмерной свёрточной сети [140].

В настоящей главе предлагается метод позиционно-зависимой коррекции ГСИ с использованием обучаемых калибровочных слоёв. Коэффициенты коррекции зависят от положения вдоль щелевой диафрагмы и спектрального канала. Рассматриваются две постановки калибровки: относительно опорных спектров классов калибровочной сцены разработанного гиперспектрометра и относительно средних спектров соответствующих классов Specim IQ. В подразделе 3.3 описаны аналитическое определение коэффициентов методом гребневой регрессии и нейросетевая настройка калибровочных слоёв. Полученные коэффициенты фиксируются и применяются к независимым сценам для оценки влияния коррекции на точность поэлементной классификации ГСИ.

3.2 Модель и постановки задачи позиционно-зависимой коррекции

3.2.1 Модель неоднородности чувствительности и функция коррекции

В обозначениях подраздела 1.1 ГСИ представляется тензором третьего порядка $S(i, j, k)$, где i – пространственный отсчёт вдоль сканирования, соответствующий номеру кадра, j – индекс пространственного отсчёта вдоль изображения щелевой диафрагмы, а k – номер спектрального канала.

Спектральный вектор, соответствующий одному пространственному положению ГСИ, называется гиперпикселом [1]. При формировании i -го строчного среза ГСИ (пространственной спектрограммы) в плоскости фотоприёмной матрицы образуется непрерывное распределение освещённости $E(\lambda, x, y)$. Без учёта неоднородности чувствительности гиперспектрометра отсчёт определяется выражением, приведённым в подразделе 1.1 (формула (1.2)):

$$S_i(j, k) = \int_{jb-0,5\delta}^{jb+0,5\delta} \int_{kb-0,5\delta}^{kb+0,5\delta} \int_{\Delta\lambda} E(\lambda, x, y) G(\lambda) d\lambda dy dx, \quad (3.5)$$

где $S_i(j, k)$ – отсчёт в отсутствие неоднородности чувствительности, полученный при формировании i -го строчного среза ГСИ элементом фотоприёмника с индексами j, k ; $E(\lambda, x, y)$ – распределение освещённости в плоскости фотоприёмника; x и y – координаты, направленные вдоль отсчётов j и k соответственно; $\Delta\lambda$ – регистрируемый диапазон длин волн; δ – размер стороны светочувствительной площадки; b – шаг между центрами площадок; $G(\lambda)$ – спектральная чувствительность фотоприёмника.

Неоднородность чувствительности гиперспектрометра задаётся функцией $H(x, y)$, представленной произведением двух составляющих [1]:

$$H(x, y) = V(x)T(x, y), \quad (3.6)$$

где $V(x)$ – составляющая, описывающая изменение пропускной способности вдоль щелевой диафрагмы; $T(x, y)$ – составляющая, зависящая от координат вдоль щелевой диафрагмы и направления дисперсии.

С учётом функции $H(x, y)$ зарегистрированный сигнал имеет вид [1]:

$$S_i^1(j, k) = \int_{jb-0,5\delta}^{jb+0,5\delta} \int_{kb-0,5\delta}^{kb+0,5\delta} \int_{\Delta\lambda} E(\lambda, x, y) G(\lambda) V(x) T(x, y) d\lambda dy dx, \quad (3.7)$$

где $S_i^1(j, k)$ – значение зарегистрированного строчного среза с учётом неоднородности чувствительности прибора; остальные обозначения соответствуют выражениям (3.5) и (3.6).

При допущении о локальном постоянстве функций $V(x)$ и $T(x, y)$ в пределах светочувствительной площадки их влияние представляется

дискретными множителями $v(j)$ и $t(j, k)$. При этом интегральная модель (3.7) принимает вид [1]:

$$S_i^1(j, k) = v(j)t(j, k)S_i(j, k), \quad (3.8)$$

где $S_i(j, k)$ – значение строчного среза ГСИ без рассматриваемой неоднородности чувствительности; $S_i^1(j, k)$ – соответствующее значение зарегистрированного строчного среза с учётом этой неоднородности; $v(j)$ и $t(j, k)$ – локальные дискретные значения составляющих $V(x)$ и $T(x, y)$ соответственно. В дальнейшем функцию $t(j, k)$ будем записывать в виде $t_k(j)$, рассматривая k как номер спектрального канала. Множитель $v(j)$ описывает изменение пропускной способности вдоль щелевой диафрагмы, обусловленное изменением её ширины; множитель $t_k(j)$ характеризует искажение интенсивности сигнала, предположительно связанное с изменением формы ФРТ, виньетированием, спектральным пропусканием оптической системы и спектральной чувствительностью фотоприёмника [1].

Позиционно-зависимая коррекция зарегистрированных спектральных интенсивностей выполняется поэлементным умножением среза спектрального канала на функцию коррекции [1]:

$$\hat{S}_k(i, j) = S_k(i, j) \odot l_k(j), \quad (3.9)$$

где $S_k(i, j)$ и $\hat{S}_k(i, j)$ – значения исходного и скорректированного срезов k -го спектрального канала соответственно; $l_k(j)$ – корректирующий коэффициент, то есть значение функции коррекции для спектрального канала k и положения j вдоль изображения щелевой диафрагмы; $\odot$ – поэлементное умножение.

Поскольку функция $l_k(j)$ не зависит от индекса строки i , для фиксированных j, k один и тот же коэффициент применяется ко всем строкам ГСИ [1].

Функция коррекции представлена произведением двух составляющих [1]:

$$l_k(j) = l_k^{(t)}(j) \odot l_k^{(v)}(j), \quad (3.10)$$

где $l_k^{(t)}(j)$ предназначена для компенсации изменения интенсивности, зависящего от координат j и k и представленного составляющей $t_k(j)$ модели неоднородности чувствительности; $l_k^{(v)}(j)$ предназначена для компенсации составляющей $v(j)$, характеризующей изменение пропускной способности вдоль щелевой диафрагмы [1].

Выбор опорных спектров и способ определения коэффициентов $l_k(j)$ зависят от постановки эксперимента подбора коэффициентов калибровки. В пункте 3.2.2 рассматривается коррекция относительно опорных спектров классов калибровочной сцены, а в пункте 3.2.3 – относительно данных другого гиперспектрометра.

3.2.2 Коррекция по опорным спектрам классов калибровочной сцены

В первой постановке функция коррекции определяется по размеченному ГСИ калибровочной сцены разработанного гиперспектрометра. Для каждого класса формируется опорный спектр: выбирается наиболее яркий из средних спектров этого класса, рассчитанных для разных положений вдоль изображения щелевой диафрагмы. Опорные спектры вычисляются по

исходным цифровым отсчётам калибровочного ГСИ и фиксируются до настройки коррекции.

При фиксированных классе c и положении j среднее значение в каждом спектральном канале определяется усреднением по соответствующим пикселям вдоль направления сканирования:

$$\bar{S}_{c,j,k} = \frac{1}{|\Omega_{c,j}|} \sum_{i \in \Omega_{c,j}} S(i,j,k), \quad |\Omega_{c,j}| > 0, \quad (3.11)$$

где $\bar{S}_{c,j,k}$ – значение среднего спектра класса c в положении j и канале k ; $\Omega_{c,j} = \{i \mid y(i,j) = c\}$ – множество индексов строк, для которых пиксель в положении j имеет метку класса c ; $y(i,j)$ – метка класса пикселя; $|\Omega_{c,j}|$ – число таких пикселей.

Опорное положение выбирается отдельно для каждого класса по максимуму суммы значений его среднего спектра по всем каналам:

$$j_{\text{ref},c} = \arg \max_{\substack{1 \leq j \leq J \\ |\Omega_{c,j}| > 0}} \sum_{k=1}^K \bar{S}_{c,j,k}, \quad (3.12)$$

где $j_{\text{ref},c}$ – опорное положение для класса c ; J – число положений вдоль изображения щелевой диафрагмы; K – число спектральных каналов.

Значения выбранного среднего спектра образуют опорный спектр соответствующего класса:

$$S_{c,k}^{\text{ref}} = \bar{S}_{c,j_{\text{ref}},c,k}, \quad (3.13)$$

где $S_{c,k}^{\text{ref}}$ – значение опорного спектра класса c в канале k , общее для всех положений вдоль щели. Подтензор ГСИ представляет собой часть исходного тензора, соответствующую пространственной области заданного размера и содержащую все спектральные каналы [1]. При нейросетевой настройке по спектральному критерию опорный спектр выбирается по классу центрального гиперпиксела входного подтензора и повторяется во всех пространственных позициях опорного подтензора. Скорректированный входной подтензор сопоставляется с этим опорным подтензором. Одна функция коррекции $l_k(j)$ определяется совместно по всем классам и после настройки применяется без их меток. Этот способ задания спектральной цели используется при настройке по комбинированной функции потерь (далее – COMBO); обозначение CE соответствует настройке по перекрёстной энтропии. Критерии и способы определения коэффициентов рассматриваются в подразделе 3.3.

Базовым вариантом служит коррекция по белому объекту. В нём усредняются срезы, содержащие белый объект; из полученных спектров выбирается наиболее яркий по сумме значений каналов. Коэффициенты задаются отношением значений этого опорного спектра к значениям спектра белого объекта в каждом положении вдоль щели. Такая схема соответствует аналитической коррекции по белому объекту, описанной в статье [1].

3.2.3 Межприборная коррекция по средним спектрам классов

Во второй постановке средний опорный спектр каждого класса формируется по данным другого гиперспектрометра. Для определения параметров функции коррекции используются калибровочные сцены разработанного гиперспектрометра и Specim IQ, содержащие один и тот же набор классов $\mathcal{C}$. Соответствие между сценами устанавливается по меткам классов; попиксельное пространственное совмещение двух ГСИ не выполняется. Каждому классу c соответствует средний опорный спектр, не зависящий от положения вдоль изображения щелевой диафрагмы.

Последовательность формирования сопоставляемых спектров показана на рисунке 3.6.



Рисунок 3.6 – Схема межприборного согласования среднего спектра класса c для положений j вдоль изображения щелевой диафрагмы

Спектры Specim IQ предварительно приводятся к сетке длин волн разработанного гиперспектрометра линейной интерполяцией как в (1.23). Пусть Ω_c^{IQ} – множество пикселей класса c в калибровочном ГСИ Specim IQ. Средний опорный спектр класса определяется выражением:

$$\bar{R}_{c,k}^{\text{IQ}} = \frac{1}{|\Omega_c^{\text{IQ}}|} \sum_{(u,v) \in \Omega_c^{\text{IQ}}} R_{u,v,k}^{\text{IQ}}, \quad (3.14)$$

где $R_{u,v,k}^{\text{IQ}}$ – значение коэффициента спектральной яркости в пространственных координатах u, v и спектральном канале k после приведения к сетке длин волн разработанного гиперспектрометра; Ω_c^{IQ} – множество пикселей класса c ; $|\Omega|$ – число элементов соответствующего множества.

Во второй постановке функция коррекции $l_k(j)$, введённая в выражении (3.9), записывается в параметризованном виде $G_\theta(j, k) = l_k(j)$, где θ – совокупность настраиваемых параметров. Предварительно скорректированные по тёмному сигналу цифровые отсчёты умножаются на $G_\theta(j, k)$. Затем к каждому спектру применяется общий оператор спектральной предобработки P_φ ; его состав и предварительная настройка описаны в пункте 3.5.2. После этого сигнал нормируется по зарегистрированному спектру белого эталонного объекта Spectralon. Полученный коэффициент спектральной яркости (КСЯ) обозначим через $\hat{R}_{i,j,k}(\theta)$. Для каждой пары «положение вдоль щели – класс» средний скорректированный спектр определяется выражением:

$$\bar{R}_{c,j,k}(\theta) = \frac{1}{|\Omega_{c,j}|} \sum_{i \in \Omega_{c,j}} \hat{R}_{i,j,k}(\theta), \quad (3.15)$$

где $\Omega_{c,j} = \{i \mid y(i, j) = c\}$ – множество индексов строк i ; $\bar{R}_{c,j,k}(\theta)$ – значение среднего скорректированного спектра этого класса в k -м спектральном канале и в положении j вдоль изображения щелевой диафрагмы. При нейросетевой настройке по спектральному критерию параметры θ определяются из условия приближения среднего скорректированного спектра каждой представленной группы (j, c) к опорному спектру $\bar{R}_{c,k}^{\text{IQ}}$ соответствующего класса; отклонение при этом вычисляется для средних спектров групп, а не для отдельных пикселей. При настройке по классификационному критерию функция коррекции $G_\theta(j, k)$ включается в модель классификации. Используемые критерии и параметризации рассматриваются в подразделе 3.3.

3.3 Позиционно-зависимая коррекция ГСИ с обучаемыми калибровочными слоями

3.3.1 Калибровка по белому объекту и аппроксимация коэффициентов

Совокупность коэффициентов $l_k(j)$, используемых в выражении (3.9) для всех положений вдоль изображения щелевой диафрагмы и спектральных каналов, далее представим матрицей калибровочных коэффициентов $L \in R^{J \times K}$, $L_{j,k} = l_k(j)$. При коррекции по белому объекту элементы калибровочной матрицы вычисляются как отношения значений опорного спектра белого объекта к его зарегистрированным значениям в соответствующем положении и не являются параметрами нейросетевой модели [1].

Для формирования малопараметрического представления матрицы L зависящая от положения вдоль щели и спектрального канала составляющая $l_k^{(t)}(j)$, введённая в выражении (3.10), аппроксимируется для каждого канала квадратичным полиномом. Его параметры определяются по критерию суммы квадратов отклонений [1]:

$$\sum_{j=1}^J [l_k(j) - l_k^{(t)}(j)]^2 \rightarrow \min, \quad (3.16)$$

$$l_k^{(t)}(j) = a_{k,2}j^2 + a_{k,1}j + a_{k,0}, \quad (3.17)$$

где $l_k(j)$ – исходные отсчёты корректирующей функции; $l_k^{(t)}(j)$ – их квадратичная аппроксимация; $a_{k,2}$, $a_{k,1}$, $a_{k,0}$ – коэффициенты полинома для k -го спектрального канала; J – число отсчётов вдоль изображения щелевой диафрагмы.

3.3.2 Обучаемые калибровочные слои и их параметризации

В отличие от матрицы L , элементы которой предварительно рассчитываются по калибровочному сигналу, в нейросетевой реализации корректирующее преобразование задаётся обучаемыми параметрами, определяемыми по выбранной функции потерь. Параметризации калибровочных слоёв описаны в пунктах 3.3.3–3.3.5. Порядок настройки слоя

и классификатора описан в пункте 3.3.8. Для обучаемой матрицы параметров далее используется обозначение M [1].

Схема включения калибровочного слоя в модели классификации 1DCNN, NM3DCNN и SSFTT показана на рисунке 3.7.

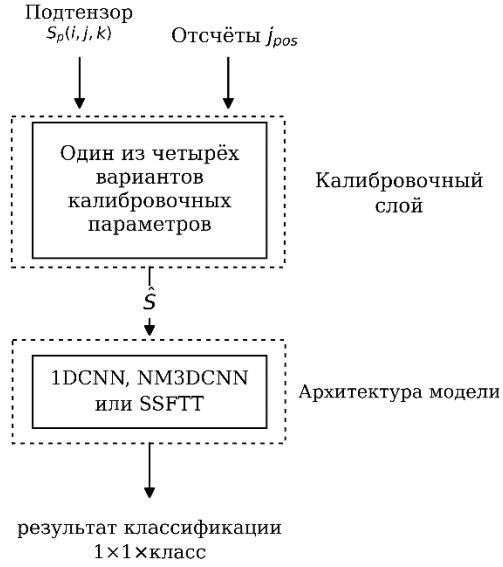


Рисунок 3.7 – Схема включения калибровочного слоя в модель классификации ГСИ

3.3.3 Обучаемая калибровочная матрица

Для компенсации неоднородности чувствительности, представленной множителями $t_k(j)v(j)$ в (3.8), используются обучаемые калибровочные слои. Поэлементное изменение интенсивностей не выполняет геометрического преобразования координат отсчётов. Матрицу корректирующих коэффициентов представим обучаемой матрицей $M \in R^{J \times K}$ (3.18), где каждому сочетанию координат (j, k) в ГСИ сопоставляется индивидуальный корректирующий коэффициент [1]:

$$S_m(i, j, k) = S(i, j, k) \odot M(j, k), \quad (3.18)$$

где $M(j, k)$ – коэффициент коррекции, соответствующий спектральному каналу k и положению вдоль щелевой диафрагмы с индексом j [1].

На этапе подготовки данных для обучения нейросети исходное ГСИ разбивается на набор фрагментов, каждый из которых интерпретируется как отдельный подтензор фиксированного размера [1]:

$$S_p(i', j', k) \in R^{H_p \times W_p \times K}, \quad (3.19)$$

где H_p и W_p – пространственные размеры подтензора S_p вдоль направлений с локальными индексами i' и j' соответственно; K – число спектральных каналов, сохраняемых при выделении фрагмента [1].

Калибровочные слои используют глобальные координаты вдоль изображения щелевой диафрагмы. Для подтензора S_p вводится вектор $j_{\text{pos}} \in \{1, \dots, J\}^{W_p}$, содержащий координаты его столбцов в исходном ГСИ. Компонента $j_{\text{pos}}(j')$ соответствует локальному столбцу j' . Для каждого значения подтензора коррекция записывается следующим образом [1]:

$$S_m(i', j', k) = S_p(i', j', k) \cdot M(j_{\text{pos}}(j'), k), \quad (3.20)$$

3.3.4 Обучаемая полиномиальная параметризация калибровочной матрицы

В отличие от аппроксимации предварительно рассчитанных коэффициентов по (3.16) и (3.17), в данном слое коэффициенты полинома являются обучаемыми параметрами. Слой параметризует матрицу корректирующих коэффициентов (3.18). Для каждого спектрального канала $k \in \{1, \dots, K\}$ вводится набор параметров $a_k, b_k, c_k \in R$, задающих квадратичную зависимость корректирующего коэффициента от нормированной координаты вдоль щели [1]:

$$y_k(j) = a_k \xi(j)^2 + b_k \xi(j) + c_k, \quad (3.21)$$

где $y_k(j)$ – корректирующий коэффициент для канала k и положения j ; $\xi(j) = \frac{j-1}{J-1}$ – нормированная координата при $J > 1$ и $j \in \{1, \dots, J\}$; a_k, b_k, c_k – обучаемые параметры полинома. Нормированная координата изменяется в диапазоне от 0 до 1 [1].

Для локального столбца j' подтензора коэффициент полиномиального слоя вычисляется по глобальной координате $j_{pos}(j')$. Коррекция значения спектрального канала имеет вид [1]:

$$\hat{S}_{p,k}(i', j') = S_{p,k}(i', j') \cdot y_k(j_{pos}(j')). \quad (3.22)$$

3.3.5 Векторная факторизация калибровочной матрицы

Для уменьшения числа параметров калибровочной матрицы предлагается использовать векторную факторизацию, где комбинированный коэффициент коррекции для каждой пары (j, k) определяется как [1]:

$$M(j, k) = \alpha(j) \cdot \beta(k), \quad (3.23)$$

где $\alpha(j)$ и $\beta(k)$ – компоненты обучаемых векторов $\alpha \in R^J$ и $\beta \in R^K$ соответственно.

Векторная факторизация уменьшает число обучаемых параметров с JK до $J + K$, ограничивая матрицу коррекции произведением пространственного и спектрального множителей. Для отдельного значения подтензора получаем [1]:

$$\hat{S}_v(i', j', k) = S_p(i', j', k) \cdot \alpha(j_{pos}(j')) \cdot \beta(k), \quad (3.24)$$

Далее используются следующие обозначения конфигураций калибровочных слоёв: «m» – обучаемая калибровочная матрица; «v» – векторная факторизация калибровочной матрицы; «pm» – полиномиальный слой и матрица параметров; «pv» – полиномиальный слой и два вектора параметров [1].

Схема обучаемой нейросетевой калибровки ГСИ представлена на рисунке 3.8.

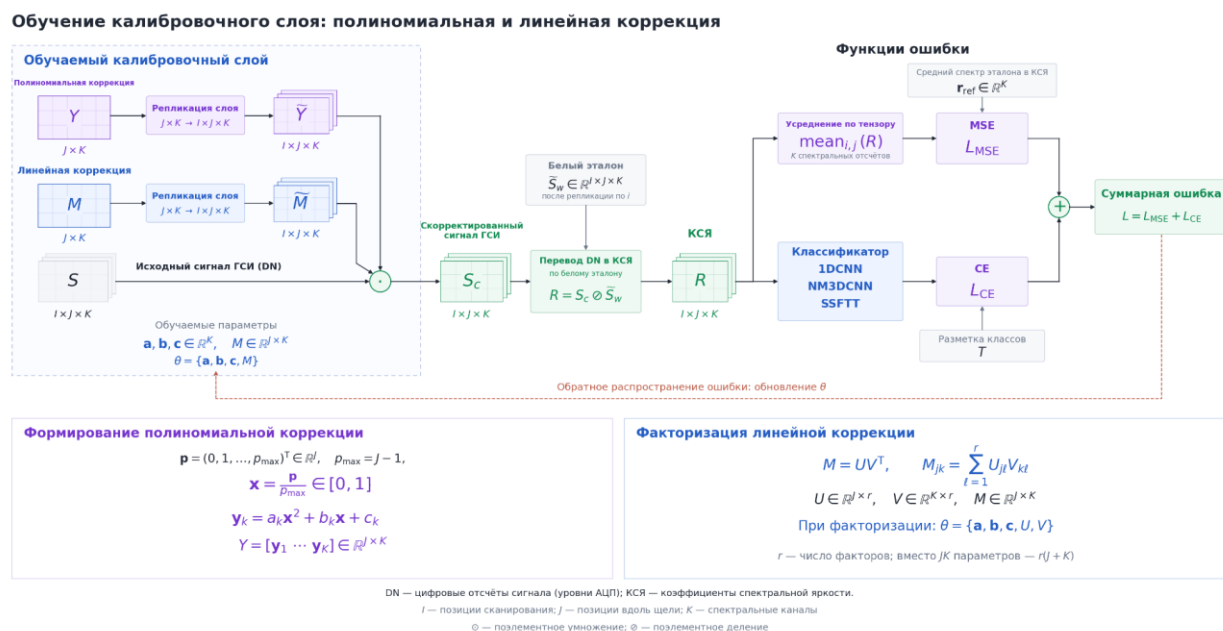


Рисунок 3.8 – Схема обучаемой нейросетевой калибровки ГСИ

3.3.6 Определение коэффициентов методом гребневой регрессии

Метод гребневой регрессии [141], [107], [142] рассмотрен в пункте 2.1.5, формулы (2.21), (2.22). В настоящей работе гребневая регрессия используется для независимого определения мультипликативного коэффициента $l_{j,k}^R$ для каждой пары индексов j, k . Обозначим через $x_{n,j,k}$ исходное значение n -го калибровочного наблюдения, а через $y_{n,k}$ — соответствующее опорное значение. Множество сопоставляемых наблюдений для положения j обозначим I_j . Поскольку единичный множитель не изменяет сигнал, регуляризатор ограничивает отклонение коэффициента от единицы. Коэффициент определяется из условия:

$$l_{j,k}^R = \arg \min_l [\sum_{n \in I_j} (lx_{n,j,k} - y_{n,k})^2 + \eta_k(l-1)^2], \quad (3.25)$$

где n – номер сопоставляемого наблюдения; $\eta_k > 0$ – параметр регуляризации для соответствующего канала; он может быть одинаковым для всех каналов. В отличие от стандартного штрафа на величину коэффициента, слагаемое $\eta_k(l-1)^2$ соответствует используемому в данной работе приоритету

единичного корректирующего множителя. Критерий (3.25) разделяется по индексам j и k , поэтому для каждого элемента матрицы получается аналитическое решение:

$$l_{j,k}^R = \frac{\eta_k + \sum_{n \in I_j} x_{n,j,k} y_{n,k}}{\eta_k + \sum_{n \in I_j} x_{n,j,k}^2}. \quad (3.26)$$

Выражения (3.25), (3.26) задают общий аналитический критерий двух постановок, но исходные наблюдения в них формируются по-разному. В первой постановке используются значения средних спектров классов калибровочной сцены при фиксированном j для положений, где класс присутствует; опорными служат значения $S_{c,k}^{\text{ref}}$ из (3.13). Во второй постановке каждое наблюдение соответствует отдельному пикселю обучающей части калибровочной сцены. Его исходные цифровые отсчёты, компенсированные по тёмному сигналу, переводятся в КСЯ нормировкой по Spectralon при единичной коррекции и без оператора P_φ . Опорой служит средний спектр соответствующего класса Specim IQ (3.14). Таким образом, гребневая регрессия сопоставляет величины в одинаковом представлении и не минимизирует ошибку после полной последовательности преобразований, описанной в пункте 3.2.3. После аналитического расчёта применяются предусмотренные протоколом ограничения диапазона коэффициентов; выражение (3.26) задаёт решение до такого ограничения. Полученная матрица L^R используется в корректирующем преобразовании (3.9).

3.3.7 Фиксация и перенос коэффициентов калибровочного слоя

После аналитического расчёта или нейросетевой настройки коэффициенты коррекции фиксируются и применяются к независимым сценам.

В первой постановке используются данные калибровочной сцены разработанного гиперспектрометра. При гребневой регрессии и вычислении спектрального слагаемого функции потерь опорой служат спектры классов

(3.11)–(3.13); в варианте СЕ используются метки классов. Протокол эксперимента приведён в подразделе 3.4.

Во второй постановке параметры определяются по калибровочным сценам разработанного гиперспектрометра и Specim IQ. Для классификации независимой сцены разработанного гиперспектрометра используется классификатор, отдельно обученный на соответствующей сцене Specim IQ. Протокол эксперимента приведён в подразделе 3.5.

Дополнительно исследуется применение одного и того же калибровочного преобразования с классификаторами различных архитектур.

3.3.8 Нейросетевая настройка калибровочных слоёв

Настройка калибровочных слоёв по ТКЦЗ выполняется на калибровочной сцене по критериям СЕ и COMBO, рассмотренным в пункте 2.1.5 (формулы (2.15), (2.18), (2.19)). В первой постановке оба варианта предусматривают совместное обучение калибровочного слоя и классификатора. В COMBO к перекрёстной энтропии добавляется средний квадрат отклонения скорректированного подтензора от опорного, сформированного по спектру класса (пункт 3.2.2).

Во второй постановке обучается только калибровочный слой. Параметры спектральной предобработки и классификатора, предварительно обученного на опорных данных Specim IQ, фиксированы. В варианте СЕ штрафные слагаемые ограничивают изменения коэффициентов и увеличение расхождения спектров относительно начального состояния. В COMBO на протяжении всего обучения учитывается взвешенное расхождение между средним скорректированным спектром каждого класса в положении j и соответствующим опорным средним спектром Specim IQ (3.14), (3.15).

Протоколы экспериментов приведены в подразделах 3.4 и 3.5.

3.4 Коррекции позиционной неоднородности ГСИ

В первой постановке проверяется коррекция позиционной неоднородности сигнала относительно опорных спектров классов,

полученных разработанным гиперспектрометром. Эксперимент включает оценку профилей сигнала на калибровочной сцене и проверку классификации независимой тестовой сцены после применения зафиксированной коррекции. Эти оценки характеризуют соответственно изменение неоднородности вдоль щели и пригодность скорректированных данных для распознавания классов.

3.4.1 Калибровочное и тестовое ГСИ

Для настройки преобразования используется ГСИ печатной калибровочной сцены, сформированной по методике главы 1 (подраздел 1.6), а для проверки его переноса – ГСИ отдельной тестовой сцены. Оба ГСИ зарегистрированы разработанным гиперспектрометром. Каждое ГСИ содержит 960 положений вдоль изображения щелевой диафрагмы, 2000 отсчётов по направлению сканирования и 252 спектральных канала в диапазоне 422–978 нм. Калибровочная сцена включает 12 размеченных классов поверхностей, тестовая – пять. Пиксели фона исключаются из обучающих и тестовых выборок.

Параметры позиционной коррекции определяются только по калибровочной сцене. После настройки они фиксируются и применяются ко всем положениям сканирования тестового ГСИ. Метки классов тестовой сцены используются для вычисления показателей классификации; при применении самого преобразования они не требуются. Состав классов при настройке слоя и при итоговой классификации различается, поэтому эти задачи рассматриваются отдельно.

В тракте итоговой классификации используются общий фиксированный оператор спектральной предобработки P_ϕ (сглаживание, подавление выбросов и компенсация смещения) и перевод цифровых отсчётов к коэффициенту спектральной яркости. Параметры оператора предварительно определяются по калибровочным данным разработанного гиперспектрометра и Specim IQ. Спектры Specim IQ интерполируются на сетку длин волн разработанного гиперспектрометра (пункт 3.5.1). Этот общий этап применяется ко всем

сравниваемым вариантам, включая вариант без позиционной коррекции. Опорные спектры для настройки позиционной коррекции формируются по данным разработанного гиперспектрометра согласно пункту 3.2.2. Изображения калибровочной и тестовой сцен и их эталонная разметка приведены в пункте 2.2.1 (рисунки 2.4 и 2.6).

3.4.2 Формирование опорных спектров классов

Для каждого класса калибровочной сцены сначала вычисляются средние спектры при фиксированных положениях вдоль щели по выражению (3.11). Затем по выражениям (3.12), (3.13) выбирается средний спектр с наибольшей суммой значений по всем спектральным каналам. Такой выбор выполняется независимо для каждого класса, поэтому его опорное положение может отличаться от опорных положений других классов. Полученные спектры фиксируются до настройки преобразования.

При гребневой регрессии средние спектры классов в разных положениях сопоставляются с соответствующими опорными спектрами. При комбинированном дообучении спектральная составляющая функции потерь задаётся относительно опорного подтензора: спектр класса центрального гиперпиксела повторяется во всех пространственных позициях этого подтензора. Все классы участвуют в определении одной общей функции коррекции $l_k(j)$. Коррекция по белому объекту составляет самостоятельный базовый вариант и определяется по узкой полосе белой поверхности согласно пункту 3.2.2.

3.4.3 Протокол настройки и варианты коррекции

Сравниваются пять вариантов: без позиционной коррекции, коррекция по белому объекту, гребневая регрессия, SE и COMBO. Коэффициенты гребневой регрессии, рассчитанные по (3.25), (3.26), служат начальными значениями для независимого дообучения по SE и COMBO.

При нейросетевом дообучении в первой постановке совместно изменяются параметры калибровочного слоя и классификатора. В варианте SE

используется перекрёстная энтропия, в COMBO к ней добавляется средний квадрат отклонения скорректированного подтензора от опорного. Регуляризация ограничивает отклонения коэффициентов от единицы и их резкие изменения.

Классификация исходных и скорректированных ГСИ выполняется сетями 1DCNN, NM3DCNN и SSFTT. На вход 1DCNN подаётся спектр отдельного гиперпиксела [47]. Входом NM3DCNN и SSFTT служит подтензор ГСИ, содержащий классифицируемый гиперпиксел и его пространственное окружение [1].

Настройка слоёв с SSFTT и NM3DCNN выполнялась в течение 20 эпох при размере пакета данных 256. Для SSFTT использовались подтензоры размером $5 \times 5 \times 252$. В варианте CE со слоем «*pt*» формировалось по 3500 примеров каждого класса, в COMBO со слоями «*t*» и «*v*» – по 2500. В этих запусках COMBO вес спектрального слагаемого составлял $5 \cdot 10^{-8}$, вес перекрёстной энтропии – 1. Для NM3DCNN скорость обучения составляла 0,001. Обозначения слоёв приведены в пункте 3.3.5.

После настройки коэффициенты коррекции фиксируются и применяются к независимому тестовому ГСИ совместно с общим спектральным оператором. Итоговые классификаторы обучаются отдельно, без встроенного калибровочного слоя. Для обучения использовались 5 % размеченных гиперпикселов каждого класса соответствующей сцены Specim IQ – 743 центра. Проверка проводилась на 181 284 размеченных гиперпикселах тестовой сцены разработанного гиперспектрометра. Внутри одного запуска каждый итоговый классификатор применялся ко всем вариантам обработки с неизменными параметрами. Начальное значение генератора псевдослучайных чисел задавалось равным «43» (*seed* = 43). Для каждого варианта эксперимента выполнено по три запуска.

3.4.4 Оценка позиционной неоднородности сигнала

Позиционная неоднородность оценивается по калибровочной сцене в цифровых отсчётах до применения общего спектрального оператора. Для каждого варианта обработки a средние профили $\bar{S}_{c,j,k}^{(a)}$ вычисляются по выражению (3.11). В множество J_c включаются только положения вдоль щели, в которых присутствуют пиксели класса c ; отсутствующие положения при вычислении метрик не заполняются.

Для профильной оценки в каждом обработанном ГСИ заново выбирается наиболее яркий отдельный гиперпиксел каждого класса по сумме значений всех каналов:

$$(i_{c,a}^*, j_{c,a}^*) = \arg \max_{(i,j): y_{i,j}=c} \sum_{k=1}^K S_{i,j,k}^{(a)}, \quad (3.27)$$

где $i_{c,a}^*$ и $j_{c,a}^*$ – координаты выбранного гиперпикселя; a – вариант обработки; $y_{i,j}$ – метка класса. Значение его спектра в канале k обозначим $b_{c,k}^{(a)} = S_{i_{c,a}^*, j_{c,a}^*, k}^{(a)}$. Такая опора отличается от зафиксированного до обучения среднего спектра $S_{c,k}^{\text{ref}}$ в (3.13): при вычислении профильной метрики используется отдельный гиперпиксел, выбранный уже после соответствующей обработки.

Среднеквадратичное отклонение профиля от выбранного уровня определяется выражением:

$$\text{RMSE}_{c,k}^{(a)} = \sqrt{\frac{1}{|J_c|} \sum_{j \in J_c} \left(\bar{S}_{c,j,k}^{(a)} - b_{c,k}^{(a)} \right)^2}, \quad (3.28)$$

где $|J_c|$ – число учитываемых положений вдоль щели. Метрика выражается в цифровых отсчётах и вычисляется при наличии не менее двух таких положений. Для сопоставления классов с разной яркостью используется относительное значение в процентах:

$$r_{c,k}^{(a)} = \frac{100 \text{RMSE}_{c,k}^{(a)}}{|b_{c,k}^{(a)}| + \varepsilon}, \quad (3.29)$$

где $\varepsilon = 10^{-8}$ – добавка к знаменателю в единицах цифрового отсчёта, предотвращающая деление на нуль. Нормировка выполняется по модулю

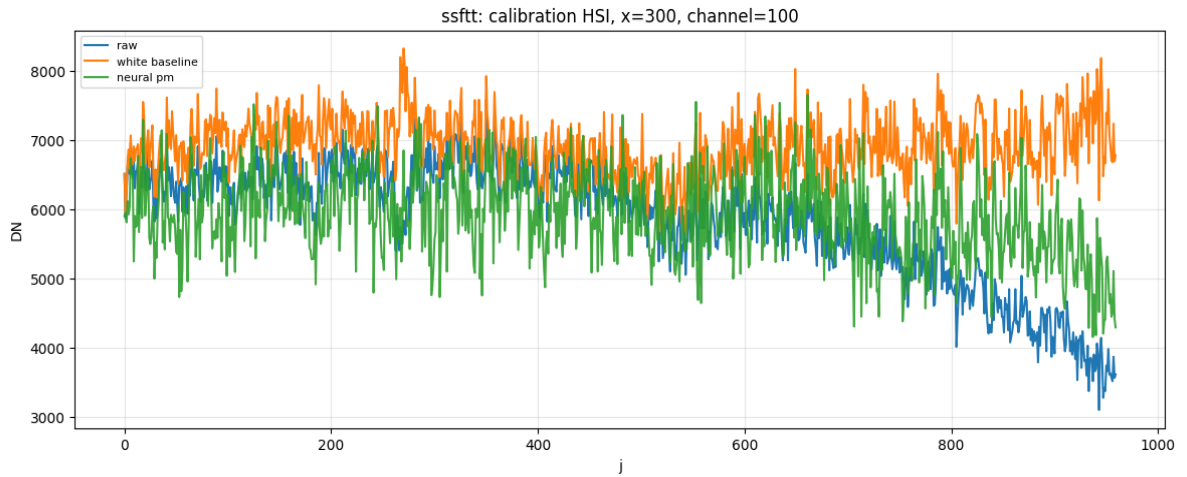
значения индивидуального опорного спектра в рассматриваемом канале. Итоговый показатель равен арифметическому среднему по выбранным классам и каналам:

$$\bar{r}^{(a)} = \frac{1}{|C||\mathcal{K}_0|} \sum_{c \in C} \sum_{k \in \mathcal{K}_0} r_{c,k}^{(a)}, \quad (3.30)$$

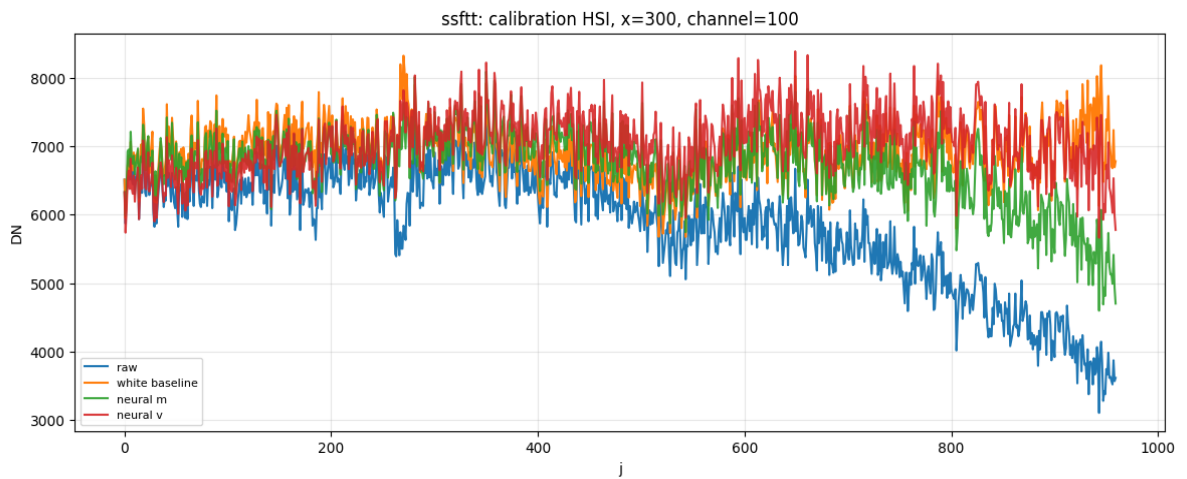
где C – множество 12 классов калибровочной сцены; $\mathcal{K}_0 = \{31, 101, 171\}$ – три выбранных канала при нумерации с единицы. Таким образом, усредняются 36 относительных ошибок с одинаковым весом каждой пары «класс – канал». Средняя абсолютная RMSE вычисляется таким же усреднением значений (3.28).

На калибровочной сцене средняя относительная RMSE составила 24,63 % без позиционной коррекции и 15,72 % после коррекции по белому объекту. Для слоёв, настроенных с использованием классификатора SSFTT, получены 18,79 % в варианте CE со слоем pm и 17,20 % в варианте COMBO со слоем m . Все три варианта снизили показатель относительно исходных данных; среди перечисленных конфигураций наименьшее значение дала коррекция по белому объекту.

По выражению (3.28) оценивается отклонение усреднённого по классу сигнала от опорного уровня на калибровочной сцене. Опорный уровень определяется отдельно для каждого варианта коррекции. На рисунке 3.9 приведены срезы сигнала вдоль щели при фиксированных положении сканирования i и спектральном канале k .



(a) CE: слой pm



(б) COMBO: слои m и v

Рисунок 3.9 – Пример цифровых отсчётов вдоль щелевой диафрагмы на калибровочной сцене при различных вариантах позиционно-зависимой коррекции. Настройка калибровочных слоёв с использованием SSFTT: а – по критерию CE, слой pm; б – по критерию COMBO, слои m и v

3.4.5 Результаты поэлементной классификации

На независимой тестовой сцене рассчитываются ОА (пункт 1.3.6) и взвешенная F1-мера. Последняя определяется усреднением F1-мер классов с весами, равными их долям в тестовой выборке.

В таблице 3.1 для каждого классификатора приведены средние значения ОА и взвешенной F1-меры по трём запускам. Для гребневой регрессии, СЕ и COMBO отдельно выбраны конфигурации с наибольшей средней ОА.

Таблица 3.1 – Результаты классификации в первой постановке (средние значения по трём запускам)

Вариант	1DCNN		NM3DCNN		SSFTT	
	ОА, %	Взвеш. F1-мера, %	ОА, %	Взвеш. F1-мера, %	ОА, %	Взвеш. F1-мера, %
Без калибровки	78,38	76,55	96,85	96,83	97,86	97,36
По белому эталону	86,66	86,39	97,16	97,17	98,31	98,00
Гребневая регрессия	79,68	79,05	96,88	96,88	98,33	98,32
«СЕ»	86,83	86,61	98,73	98,73	98,36	98,36
«COMBO»	87,21	86,98	97,89	97,88	98,51	98,51

Таблица 3.2 – Конфигурации нейросетевых вариантов с наибольшей средней ОА в первой постановке

Итоговый классификатор	Критерий	Классификатор при настройке слоя	Калибровочный слой
1DCNN	CE	NM3DCNN	Матрица с полиномом (pm)
1DCNN	COMBO	SSFTT	Матрица с полиномом (pm)
NM3DCNN	CE	SSFTT	Матрица с полиномом (pm)
NM3DCNN	COMBO	SSFTT	Матрица (m)
SSFTT	CE	SSFTT	Матрица с полиномом (pm)
SSFTT	COMBO	NM3DCNN	Матрица с полиномом (pm)

В таблице 3.2 приведены параметризации слоёв и классификаторы, использованные при их настройке. Классификатор при настройке слоя может

отличаться от итогового: например, лучший результат NM3DCNN в варианте CE получен со слоем *rm*, настроенным с использованием SSFTT.

3.4.6 Обсуждение результатов коррекции позиционной неоднородности в первой постановке эксперимента

Среди проверенных конфигураций наибольшая ОА на независимой тестовой сцене получена при нейросетевой настройке: для 1DCNN – 87,21 % по критерию COMBO, для NM3DCNN – 98,73 % по критерию CE, для SSFTT – 98,51 % по критерию COMBO. Эти значения превышают соответствующие результаты гребневой регрессии.

При этом на калибровочной сцене коррекция по белому объекту обеспечила меньшую среднюю относительную RMSE (3.30), чем слои «*rm*» и «*m*», обученные совместно с SSFTT по критериям CE и COMBO соответственно.

Результаты подтверждают возможность переноса калибровочного преобразования на независимую сцену разработанного гиперспектрометра и его применения с отдельно обученными классификаторами. Среди проверенных конфигураций наибольшие значения ОА и взвешенной F1-меры получены при нейросетевой настройке.

3.5 Коррекция ГСИ при межприборном сопоставлении

Во второй постановке исследуется коррекция спектрального сигнала разработанного гиперспектрометра по опорным данным Specim IQ. Опорными данными при настройке служат средние спектры классов калибровочной сцены Specim IQ, определённые в пункте 3.2.3. После фиксации коэффициентов сопоставляются средние спектры соответствующих классов независимых тестовых сцен и оценивается точность классификации тестового ГСИ разработанного гиперспектрометра моделями, отдельно обученными по соответствующей сцене Specim IQ.

3.5.1 ГСИ разработанного гиперспектрометра и Specim IQ

Во второй постановке печатная калибровочная сцена, разработанная в первой главе, используется для настройки межприборной коррекции по данным разработанного гиперспектрометра и Specim IQ. Калибровочные и независимые тестовые ГСИ описаны в пункте 2.2.1 (рисунки 2.3–2.6). Спектры Specim IQ линейно интерполируются с исходной сетки 397–1003 нм из 204 каналов на сетку разработанного гиперспектрометра 422–978 нм из 252 каналов.

Калибровочные сцены содержат 12 соответствующих классов, тестовые – пять. После интерполяции опорный спектр каждого класса рассчитывается по всем его размеченным гиперпикселям калибровочного ГСИ Specim IQ согласно (3.14). Сопоставление выполняется по меткам классов без попиксельного совмещения ГСИ.

На калибровочной сцене разработанного гиперспектрометра формируются выборки для обучения, выбора модели и итоговой проверки. Для каждого класса в каждом положении j выделяются соответственно 12, 4 и 4 центра подтензоров; объёмы выборок составляют 138 240, 46 080 и 46 080 центров. Выборки формируются в разных полосах по направлению сканирования. Промежутки между полосами исключают перекрытие подтензоров из разных выборок.

3.5.2 Фиксированная спектральная предобработка и расчёт КСЯ

После компенсации тёмного сигнала цифровые отсчёты умножаются на калибровочные коэффициенты, затем выполняются спектральная предобработка и нормирование по спектру белого эталона Spectralon для получения КСЯ (пункт 3.2.3).

Спектральная предобработка включает сглаживание с окном 11 каналов и подавление выбросов с окном 5 каналов. Её параметры предварительно настраиваются по обучающим калибровочным данным при фиксированной

матрице коэффициентов, полученной методом гребневой регрессии (подраздел 2.1).

В варианте без калибровки выполняется только нормирование по Spectralon. При коррекции по белому объекту, гребневой регрессии, SE и COMBO применяется один и тот же фиксированный спектральный оператор. Сравнение с исходным вариантом поэтому характеризует совокупное влияние калибровки и спектральной предобработки.

3.5.3 Настройка коррекции по калибровочным сценам

Коэффициенты гребневой регрессии рассчитываются по обучающей части калибровочной сцены для каждого положения j и канала k согласно (3.25), (3.26). Используются КСЯ отдельных гиперпикселов до спектральной предобработки и опорные средние спектры соответствующих классов Specim IQ. Для всех каналов параметр регуляризации $\eta_k = 0,01$. Полученные коэффициенты ограничиваются диапазоном 0,50–2,60. При оценке этого варианта к скорректированным данным также применяется спектральный оператор P_φ .

При коррекции по белому объекту спектры в каждом положении j усредняются по семи отсчётам вдоль направления сканирования. Опорным выбирается спектр с наибольшим средним значением по каналам. Коэффициенты определяются поканальным отношением опорного спектра к спектру в каждом положении и ограничиваются тем же диапазоном.

Регрессионное решение используется для начальной настройки конфигураций «m», «v», «pm» и «pv» (пункт 3.3.5). Начальные параметры каждой конфигурации одинаковы для SE и COMBO. Коэффициенты коррекции зависят от положения j и канала k и являются общими для всех классов.

Калибровочный слой настраивается с использованием 1DCNN и NM3DCNN, предварительно обученных на калибровочной сцене Specim IQ. Параметры классификатора и оператора P_φ фиксированы. В варианте SE

перекрёстная энтропия дополняется регуляризацией коэффициентов и ограничением увеличения расхождения спектров. В COMBO на протяжении всего обучения учитывается средний квадрат расхождения средних скорректированных спектров каждого класса в каждом положении j с опорными спектрами (3.14), (3.15). Спектральное слагаемое нормируется на начальное значение; его вес составляет 0,35 для 1DCNN и 0,50 для NM3DCNN.

Для оптимизации используется AdamW: до 40 эпох с 1DCNN и до 30 эпох с NM3DCNN. Пакеты данных формируются из групп по 12 спектров одного класса в одном положении j . В каждом запуске выбираются параметры слоя с наибольшей ТКЦЗ (ОА) на выборке калибровочной сцены для отбора модели при соблюдении спектральных ограничений. При равенстве ОА предпочтение отдаётся большей взвешенной F1-мере, при равенстве обоих показателей – более раннему шагу обучения. Полученные коэффициенты фиксируются и применяются к независимой тестовой сцене разработанного гиперспектрометра. Начальное значение генератора псевдослучайных чисел задавалось равным 43 ($\text{seed} = 43$). Для каждого варианта эксперимента выполнено по три запуска.

3.5.4 Методика оценки спектрального сходства и точности классификации

После фиксации преобразований для итоговой оценки отдельно обучаются классификаторы 1DCNN, NM3DCNN и SSFTT по независимой сцене Specim IQ с пятью классами. Каждый из них используется без изменения при оценке всех вариантов обработки тестового ГСИ разработанного гиперспектрометра. Классификаторы, использованные для настройки слоя по 12 классам калибровочной сцены, в итоговую проверку не переносятся. SSFTT в этой постановке используется как итоговый классификатор, не участвовавший в настройке слоёв.

Для обучения и выбора состояния каждого итогового классификатора выделено по 5 % размеченных гиперпикселей каждого класса Specim IQ – по

747 центров в каждой части. Пространственная поддержка обучающих и проверочных подтензоров не пересекается. Для 1DCNN используется спектр отдельного гиперпиксела, для NM3DCNN и SSFTT – подтензоры с пространственными размерами 7×7 и 5×5 соответственно и всеми 252 каналами. Параметры моделей выбираются по ОА и взвешенной F1-мере на их проверочной части Specim IQ, после чего фиксируются.

Проверка на разработанном гиперспектрометре проводится по одному набору из 181 284 допустимых тестовых центров для всех конфигураций и всех трёх классификаторов. При его формировании центральный гиперпиксел должен принадлежать рабочему классу, а окно максимального используемого размера не должно содержать другой рабочий класс; пиксели фона в окружении допускаются. Коэффициенты переносятся без изменения и без интерполяции, поскольку размеры по щели и спектральной оси у калибровочного и тестового ГСИ совпадают. Внутри каждого подтензора коэффициент применяется по фактическому положению j соответствующего гиперпиксела.

Для классификации вычисляются ОА и взвешенная F1-мера, в которой F1-меры отдельных классов усредняются с весами, пропорциональными их численности в тестовом наборе. Спектральная оценка использует истинные метки и средние опорные спектры, заново рассчитанные по всем размеченным пикселям классов независимой сцены Specim IQ. Эти тестовые опоры служат только для оценки уже зафиксированных преобразований.

Для сопоставления спектров используются RMSE, PSNR и спектральный угол SAM, применяемые также при исследовании калибровки ГСИ [57]. Для этих показателей используются разные уровни усреднения. Все средние по (3.14), (3.15) при сопоставлении спектров рассчитываются по независимым тестовым сценам; опорный спектр Specim IQ остаётся одинаковым для всех вариантов обработки.

Для вычисления PSNR обозначим через $\bar{R}_{c,k}^\theta$ средний спектр класса c по всем допустимым тестовым центрам разработанного гиперспектрометра после соответствующей обработки. При его расчёте каждый гиперпиксел имеет одинаковый вес. Средний квадрат ошибки этого спектра определяется выражением:

$$e_c = \frac{1}{K} \sum_{k=1}^K \left(\bar{R}_{c,k}^\theta - \bar{R}_{c,k}^{\text{IQ}} \right)^2, \quad (3.31)$$

где $\bar{R}_{c,k}^{\text{IQ}}$ – опорный средний спектр того же класса независимой сцены Spresim IQ; $K = 252$ – число спектральных каналов. PSNR рассчитывается для каждого класса, после чего значения в децибелах усредняются с равными весами:

$$\text{PSNR} = \frac{1}{|\mathcal{C}_t|} \sum_{c \in \mathcal{C}_t} 10 \log_{10} \left(\frac{\Delta R^2}{\max(e_c, \varepsilon)} \right), \quad (3.32)$$

где $\mathcal{C}_t$ – множество пяти тестовых классов; $\Delta R = 1$ – фиксированный диапазон для расчёта PSNR; $\varepsilon = 10^{-8}$ – нижняя граница знаменателя. Этот диапазон не определяется по наблюдаемым минимуму и максимуму и не означает обрезания значений КСЯ до интервала от нуля до единицы.

Для сопоставления спектров с учётом положения вдоль щели используются средние спектры групп (j, c) , определённые в (3.15). Пусть $\mathcal{Q}_t$ – множество непустых групп в тестовой выборке; в данном эксперименте оно содержит 1140 элементов. Групповой RMSE вычисляется как корень из среднего квадрата расхождения по всем таким группам и каналам:

$$\text{RMSE} = \sqrt{\frac{1}{K|\mathcal{Q}_t|} \sum_{(j,c) \in \mathcal{Q}_t} \sum_{k=1}^K \left(\bar{R}_{c,j,k}^\theta - \bar{R}_{c,k}^{\text{IQ}} \right)^2}, \quad (3.33)$$

где $\bar{R}_{c,j,k}^\theta$ – среднее значение КСЯ в канале k для группы (j, c) . Группы имеют одинаковый вес независимо от числа входящих в них гиперпикселов. Корень извлекается после общего усреднения квадратов ошибок; среднее отдельных групповых RMSE здесь не используется.

Обозначим векторы из всех K значений соответствующих средних спектров через $\bar{\mathbf{R}}_{c,j}^{\theta}$ и $\bar{\mathbf{R}}_c^{\text{IQ}}$. Средний спектральный угол по группам, выраженный в градусах, определяется формулой:

$$\text{SAM} = \frac{180}{\pi|Q_t|} \sum_{(j,c) \in Q_t} \arccos \left(\text{clip}_{[-1,1]} \left(\frac{\langle \bar{\mathbf{R}}_{c,j}^{\theta}, \bar{\mathbf{R}}_c^{\text{IQ}} \rangle}{\max(\|\bar{\mathbf{R}}_{c,j}^{\theta}\|, \|\bar{\mathbf{R}}_c^{\text{IQ}}\|, \varepsilon)} \right) \right), \quad (3.34)$$

где $\langle \cdot, \cdot \rangle$ – скалярное произведение; $\|\cdot\|$ – евклидова норма; $\text{clip}_{[-1,1]}$ – ограничение аргумента интервалом от -1 до 1 . В (3.34) используется та же численная защита $\varepsilon = 10^{-8}$. Показатели RMSE и SAM рассчитываются по группам, а PSNR – по общим средним спектрам классов; поэтому значение PSNR в таблице 3.5 нельзя получить непосредственно из соседнего табличного RMSE.

3.5.5 Результаты сопоставления спектров и поэлементной классификации

В таблице 3.3 приведены средние значения ОА и взвешенной F1-меры по трём запускам для пяти вариантов обработки. Для СЕ и COMBO у каждого итогового классификатора выбраны конфигурации с наибольшей средней ОА; средние значения F1-меры относятся к тем же конфигурациям. Конфигурации отобраны после тестовой оценки, а состояния отдельных слоёв зафиксированы ранее по калибровочным данным.

Таблица 3.3 – Результаты классификации во второй постановке (средние значения по трём запускам)

Вариант	1DCNN		NM3DCNN		SSFTT	
	ОА, %	Взвеш. F1-мера, %	ОА, %	Взвеш. F1-мера, %	ОА, %	Взвеш. F1-мера, %
Без калибровки	78,70	75,35	93,62	93,54	80,48	80,98
По белому эталону	75,05	67,73	83,99	84,65	86,86	86,61
Гребневая регрессия	78,29	74,22	94,55	94,60	92,32	92,11
«СЕ»	96,78	96,77	94,70	94,68	87,49	87,01
«COMBO»	91,72	91,68	95,86	95,85	91,11	90,79

В таблице 3.4 приведены параметризации выбранных слоёв для вариантов CE и COMBO и классификаторы, использованные при их настройке. Итоговые классификаторы применяются к данным после фиксации калибровочного преобразования.

Таблица 3.4 – Конфигурации нейросетевых вариантов с наибольшей средней ОА во второй постановке

Итоговый классификатор	Критерий	Классификатор при настройке слоя	Калибровочный слой
1DCNN	CE	1DCNN	Матрица с полиномом (pm)
1DCNN	COMBO	1DCNN	Векторная (v)
NM3DCNN	CE	NM3DCNN	Векторная с полиномом (pv)
NM3DCNN	COMBO	1DCNN	Матрица с полиномом (pm)
SSFTT	CE	1DCNN	Матрица (m)
SSFTT	COMBO	1DCNN	Векторная с полиномом (pv)

В таблице 3.5 приведены средние по трём запускам значения спектральных показателей базовых вариантов и шести конфигураций CE и

COMBO из таблицы 3.4. Эти показатели характеризуют конкретное преобразование и не зависят от архитектуры итогового классификатора.

Таблица 3.5 – Сопоставление спектров для выбранных вариантов второй постановки (средние значения по трём запускам)

Классификатор при настройке слоя	Слой	Критерий	PSNR, дБ	RMSE по группам	SAM, ° по группам
Без калибровки			19,54	0,1497	9,52
По белому эталону			22,19	0,1169	9,13
Гребневая регрессия			24,22	0,1067	6,55
1DCNN	pm	CE	21,19	0,1381	6,67
1DCNN	v	COMBO	23,99	0,1130	6,28
NM3DCNN	pv	CE	22,14	0,1272	6,99
1DCNN	pm	COMBO	24,02	0,1125	6,04
1DCNN	m	CE	21,99	0,1286	6,66
1DCNN	pv	COMBO	24,09	0,1127	6,23

Группа – пара «положение вдоль щели – класс». PSNR усредняется по средним спектрам классов.

Пример сопоставления средних спектров классов тестовой сцены с опорными спектрами Specim IQ приведён на рисунке 3.10.

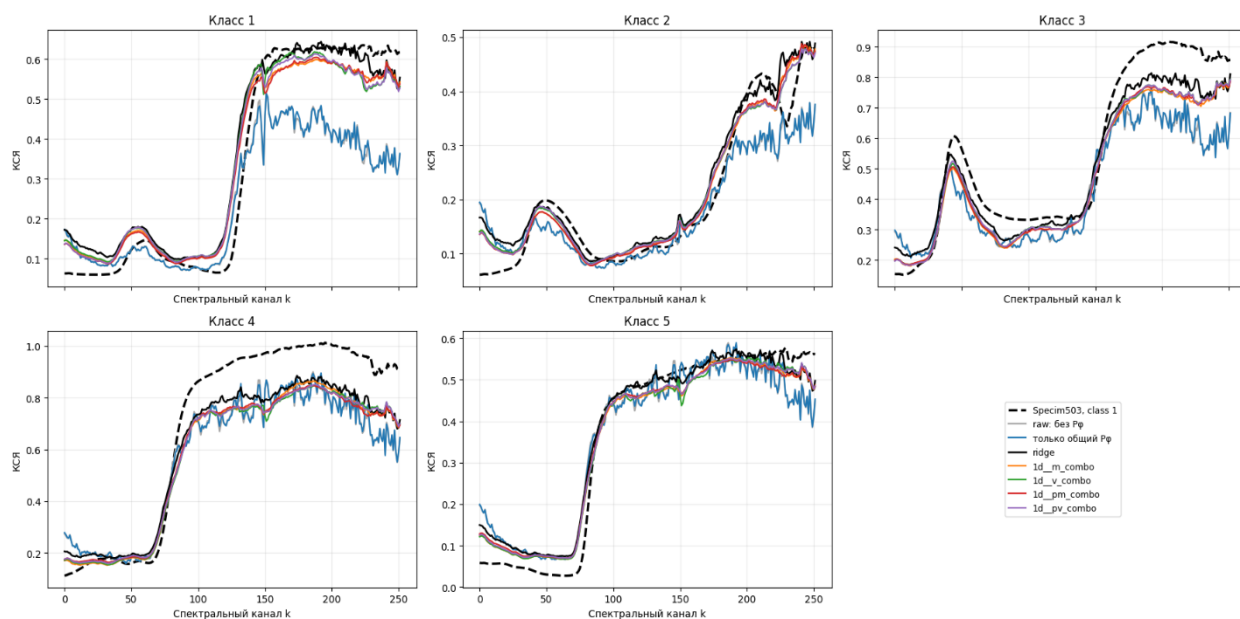


Рисунок 3.10 – Сопоставление средних спектров классов тестовой сцены при различных вариантах позиционно-зависимой коррекции. Нейросетевая настройка калибровочных слоёв с использованием 1DCNN по критерию COMBO

3.5.6 Обсуждение результатов коррекции при межприборном сопоставлении

Во второй постановке среди проверенных конфигураций наибольшая ОА для 1DCNN составила 96,78 % в варианте СЕ, для NM3DCNN – 95,86 % в варианте COMBO. В обоих вариантах слои «pm» настраивались с использованием 1DCNN и применялись с итоговыми классификаторами без изменения коэффициентов. Для SSFTT наибольшая ОА получена при гребневой регрессии – 92,32 %; Среди обучаемых конфигураций лучшие результаты СЕ и COMBO составили 87,49 и 91,11 % соответственно.

Среди вариантов таблицы 3.5 гребневая регрессия обеспечила максимальный PSNR (24,22 дБ) и минимальное значение RMSE (0,1067). У варианта СЕ с наибольшей ОА для 1DCNN эти показатели составили 21,19 дБ и 0,1381 соответственно. Минимальный средний спектральный угол SAM – $6,04^\circ$ – получен для слоя «pm», настроенного с использованием 1DCNN по критерию COMBO.

Сопоставление с данными Specim IQ выполняется по средним спектрам классов, рассчитанным для сцены в целом и отдельных положений j . Прямое

сравнение этих значений RMSE с результатами пункта 3.4.4 некорректно: в первой постановке использованы цифровые отсчёты, другая опора и другой порядок усреднения.

3.6 Проверка метода коррекции на ГСИ Specim FX

Дополнительный эксперимент воспроизводит первую постановку на данных гиперспектральной камеры Specim FX. Калибровочный слой настраивается по ТКЦЗ с использованием калибровочной сцены того же типа, что и в первой постановке, после чего его коэффициенты фиксируются для обработки двух других сцен. Настройка и применение коррекции выполняются в пространстве масштабированных цифровых отсчётов (DN), без перевода в КСЯ.

В камерах Specim FX предусмотрена встроенная калибровка «Automatic Image Enhancement» (AIE) [143], компенсирующая спектральное и межканальное пространственное рассогласование. Проявления пространственно-спектральных искажений рассмотрены в пункте 3.1.1 и проиллюстрированы на рисунках 3.3–3.5. В настоящем эксперименте используются ГСИ с включённой AIE; обучаемая коррекция применяется дополнительно к ней.

3.6.1 ГСИ Specim FX и подготовка выборок

Для настройки использовалась калибровочная сцена с 11 размеченными классами поверхностей. Тестовая сцена №1 и тестовая сцена №2 содержали соответственно два и семь классов. Гиперпикселы с нулевой меткой фона не использовались как центры обучающих примеров и не учитывались при расчёте классификационных метрик.

ГСИ содержали 224 спектральных канала в диапазоне 900–1700 нм [144]. Число строк по направлению сканирования составляло 618 или 619. Перед обучением цифровые отсчёты каждого ГСИ делились на единый для всего изображения масштабный коэффициент. Он определялся как 99,5-й процентиль случайной выборки отсчётов из размеченных областей.

Входными примерами служили подтензоры ГСИ, выделенные вокруг размеченных гиперпикселей и содержащие все спектральные каналы. Метка примера определялась классом центрального гиперпикселя. Для выбора коэффициентов калибровочного слоя использовались положения столбцов подтензора вдоль изображения щелевой диафрагмы [1].

3.6.2 Настройка калибровочного слоя для Specim FX

Опорные спектры формировались по масштабированным цифровым отсчётам калибровочной сцены согласно пункту 3.2.2. Для каждого класса рассчитывались средние спектры отдельно для каждого положения вдоль щели по (3.11). Опорным выбирался наиболее яркий из них – с максимальной суммой значений по спектральным каналам согласно (3.12):

$$S_{c,k}^{\text{ref}} = \bar{S}_{c,j_{\text{ref},c},k}, \quad (3.35)$$

где $S_{c,k}^{\text{ref}}$ – значение опорного спектра класса c в канале k ; $j_{\text{ref},c}$ – опорное положение вдоль щели для данного класса. Опорные спектры фиксировались до обучения и использовались для всех положений вдоль щели.

Для формирования обучающей выборки отдельно отбирались по 3000 случайных центральных гиперпикселей каждого из 11 классов. Всего использовалось 33 000 примеров, сформированных согласно пункту 3.6.1.

Исследовались конфигурации «m», «v», «rm» и «rv», описанные в пунктах 3.3.3–3.3.5 [1]. Коэффициенты коррекции задавались выражениями (3.18), (3.21) и (3.23). Начальные коэффициенты определялись методом гребневой регрессии по (3.25), (3.26): средние спектры классов в каждом положении вдоль щели сопоставлялись с опорными спектрами (3.35). Полученное решение использовалось для инициализации калибровочных слоёв перед нейросетевым дообучением.

Нейросетевое дообучение выполнялось по схеме первой постановки (пункт 3.3.8). Далее описан вариант с комбинированной функцией потерь COMBO, при котором параметры калибровочного слоя и SSFTT

оптимизировались совместно. Обозначим функцию коррекции через $G_\theta(j, k) = l_k(j)$, где θ – обучаемые параметры слоя. Критерий имел вид:

$$\mathcal{L} = 10\mathcal{L}_{\text{MSE}} + 0,05\mathcal{L}_{\text{CE}} + 5 \cdot 10^{-5}\mathcal{R}_{\text{id}} + 5 \cdot 10^{-4}\mathcal{R}_{\text{sm}}, \quad (3.36)$$

где $\mathcal{L}_{\text{MSE}}$ – средний квадрат отклонения скорректированного подтензора от опорного, усреднённый по пространственным позициям, спектральным каналам и примерам пакета данных; $\mathcal{L}_{\text{CE}}$ – перекрёстная энтропия для меток центральных гиперпикселей, усреднённая по примерам того же пакета. Опорный подтензор формировался повторением спектра класса центрального гиперпикселя во всех пространственных позициях (пункт 3.2.2). На вход SSFTT подавался скорректированный подтензор целиком.

Регуляризующие слагаемые рассчитывались по всей матрице корректирующих коэффициентов:

$$\begin{aligned} \mathcal{R}_{\text{id}} &= \frac{1}{JK} \sum_{j=1}^J \sum_{k=1}^K (G_\theta(j, k) - 1)^2 \\ \mathcal{R}_{\text{sm}} &= \frac{1}{(J-1)K} \sum_{j=1}^{J-1} \sum_{k=1}^K (G_\theta(j+1, k) - G_\theta(j, k))^2, \end{aligned} \quad (3.37)$$

Первое слагаемое ограничивает отклонение коэффициентов от единицы, второе – различия между коэффициентами в соседних положениях вдоль изображения целевой диафрагмы.

Нейросетевое дообучение выполнялось методом AdamW [89] в течение 20 эпох с размером пакета данных 128. Скорость обучения параметров калибровочного слоя составляла $8 \cdot 10^{-4}$, параметров SSFTT – 10^{-4} . Параметр уменьшения весов в оптимизаторе равнялся нулю.

После каждого шага оптимизации элементы матричных и векторных множителей, а также свободные члены полиномов ограничивались интервалом $[0,35; 5]$. Для остальных коэффициентов полиномов использовался интервал $[-4; 4]$.

Для обработки тестовых сцен использовалась матрица корректирующих коэффициентов размером 590×224 , полученная после последней, 20-й эпохи. Её элементы ограничивались интервалом $[0,35; 5]$, после чего матрица фиксировалась. Начальное значение генератора псевдослучайных чисел

задавалось равным 43 ($\text{seed} = 43$). Для каждого варианта эксперимента выполнено по три запуска.

3.6.3 Методика и результаты классификации тестовых ГСИ

На тестовой сцене №1 и тестовой сцене №2 сравнивались ГСИ Specim FX с включённой AIE и те же ГСИ после дополнительной обучаемой коррекции в конфигурациях m , v , pm и pv . Исходным служило ГСИ после масштабирования согласно пункту 3.6.1. Коэффициенты всех вариантов коррекции определялись по калибровочной сцене и применялись к двум тестовым сценам без перенастройки. Каждый отсчёт умножался на коэффициент, соответствующий его положению j вдоль изображения щелевой диафрагмы и спектральному каналу k ; для всех строк сканирования использовалась одна и та же матрица коэффициентов.

Для каждого сочетания тестовой сцены, варианта обработки и архитектуры заново создавался и обучался итоговый классификатор 1DCNN, NM3DCNN или SSFTT. Его параметры определялись по выделенной обучающей части соответствующей тестовой сцены. Калибровочное преобразование к этому моменту уже было выполнено; его коэффициенты при обучении итогового классификатора не изменялись. Таким образом, независимость тестовых сцен относится к настройке калибровочного слоя, а итоговые классификаторы обучаются на выделенных частях каждой новой сцены.

Обучающие и тестовые центры выбирались случайно отдельно в каждом ненулевом классе маски. В обучение включалось около 10 % размеченных гиперпикселей класса с округлением их числа до целого, остальные использовались для оценки. Дополнительное ограничение численности этих выборок не задавалось. Разбиение формировалось при начальном значении генератора случайных чисел 42, поэтому внутри одной сцены наборы обучающих и тестовых центров совпадали для всех вариантов обработки и

архитектур. Отдельная проверочная выборка для выбора состояния классификатора не выделялась.

Для 1DCNN использовались спектральные векторы отдельных гиперпикселей, для NM3DCNN и SSFTT - пространственно-спектральные фрагменты ГСИ. Обучающие и тестовые центры не совпадали, однако пространственные фрагменты могли перекрываться.

Итоговые классификаторы обучались по перекрёстной энтропии методом AdamW [89] в течение 20 эпох с размером пакета данных 256 и параметром уменьшения весов 10^{-4} . Скорость обучения NM3DCNN и SSFTT составляла $5 \cdot 10^{-4}$. Для 1DCNN проверялись два значения: 10^{-3} и 10^{-1} . Оценивалось состояние после последней эпохи.

Качество классификации оценивалось по ОА и взвешенной F1-мере на тестовых центрах. Обучающие центры и фон при расчёте метрик не учитывались.

В таблице 3.6 представлены средние значения ОА и взвешенной F1-меры по трём запускам на данных с калибровкой гиперспектрометра Specim FX AIE без нейросетевой коррекции и с нейросетевой коррекцией, настроенной по COMBO (пункт 3.6.2). Для каждого классификатора и тестовой сцены в каждом варианте выбрана конфигурация с наибольшей средней ОА; при равенстве средней ОА учитывалась средняя взвешенная F1-мера. Для обучаемой коррекции отбор проводился среди конфигураций «m», «v», «rm» и «rv». Средние значения F1-меры относятся к тем же конфигурациям.

Таблица 3.6 – Сравнение вариантов 1) «AIE» и 2) «AIE с обучаемой коррекцией» по точности классификации тестовых сцен Specim FX (средние значения по трём запускам)

Тестовая сцена	Классификатор	AIE		AIE с обучаемой коррекцией		
		ОА, %	Взвеш. F1-мера, %	Слой	ОА, %	Взвеш. F1-мера, %
Тестовая сцена №1	1DCNN	89,20	89,34	pv	87,65	87,58
	NM3DCNN	99,06	99,06	v	99,87	99,87
	SSFTT	96,72	96,75	m	97,93	97,94
Тестовая сцена №2	1DCNN	64,09	58,12	pv	70,99	66,74
	NM3DCNN	96,67	96,67	pm	97,58	97,59
	SSFTT	94,25	94,07	m	95,86	95,83

На обеих тестовых сценах наибольшая ОА среди представленных результатов получена для NM3DCNN (таблица 3.6).

Для 1DCNN на обеих сценах среди обучаемых конфигураций лучшие результаты получены со слоем «pv», для SSFTT – со слоем «m». Для NM3DCNN на тестовой сцене №1 лучшим оказался слой «v», на тестовой сцене №2 – слой «pm».

На рисунках 3.11 и 3.12 показаны спектральные срезы тестовых сцен.

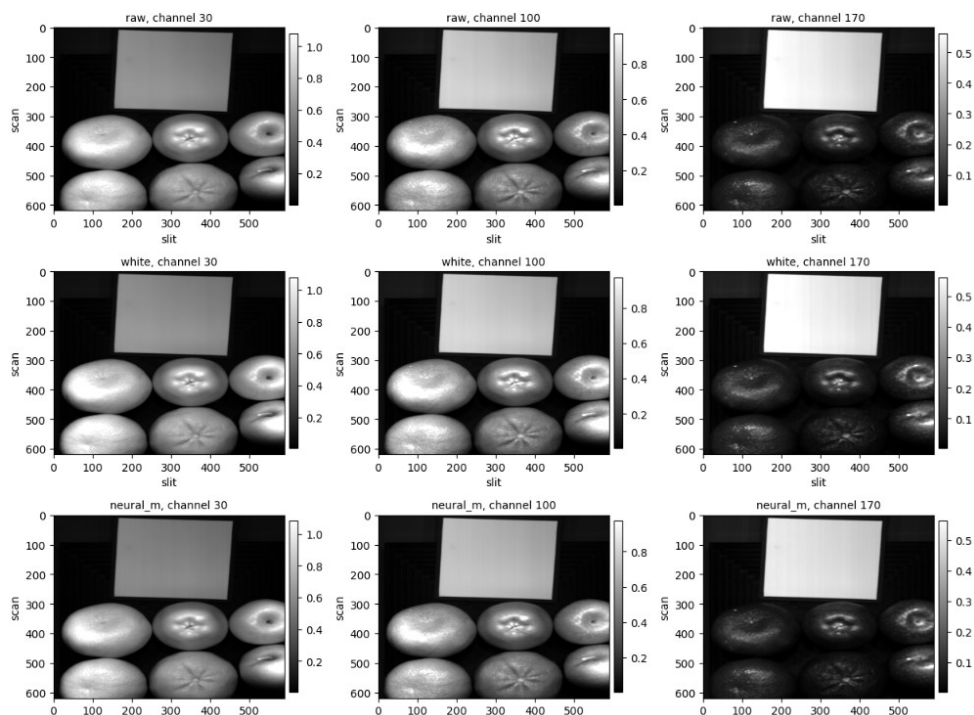


Рисунок 3.11 – Спектральные срезы ГСИ Specim FX для тестовой сцены № 1.
Слева направо: спектральные каналы 30, 100 и 170 (нумерация с нуля)

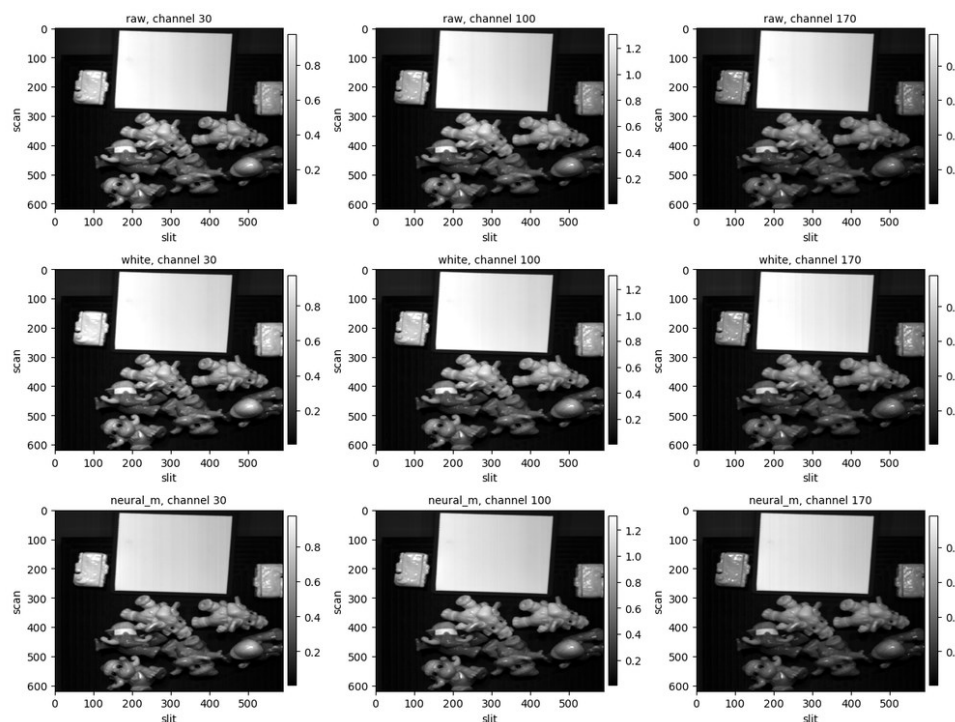


Рисунок 3.12 – Спектральные срезы ГСИ Specim FX для тестовой сцены № 2.
Слева направо: спектральные каналы 30, 100 и 170 (нумерация с нуля)

Применение обучаемой коррекции повысило ОА и взвешенную F1-меру в пяти из шести рассмотренных сочетаний тестовой сцены и классификатора (таблица 3.6). Наибольший прирост получен для 1DCNN на тестовой сцене №2: 6,90 и 8,62 процентного пункта соответственно. Однако для той же архитектуры на тестовой сцене №1 ОА снизилась с 89,20 до 87,65 %, а взвешенная F1-мера – с 89,34 до 87,58 %. Таким образом, влияние коррекции на показатели классификации зависит от сцены и архитектуры классификатора; повышение точности наблюдалось не во всех рассмотренных случаях.

В обеих постановках параметры позиционно-зависимой коррекции определялись по ГСИ печатной калибровочной сцены и после фиксации применялись к другим сценам без повторной настройки (разделы 3.4 и 3.5). В первой постановке опорные спектры формировались по данным разработанного гиперспектрометра, во второй – по средним спектрам соответствующих классов Specim IQ.

В обеих постановках результаты позиционно-зависимой коррекции оценивались с использованием классификаторов 1DCNN, NM3DCNN и SSFTT. Для 1DCNN наибольшая ОА среди рассмотренных вариантов получена во второй постановке и составила 96,78 %, тогда как в первой постановке максимальное значение составило 87,21 %. Для NM3DCNN и SSFTT, напротив, наибольшие значения ОА получены в первой постановке: 98,73 и 98,51 % соответственно; во второй постановке они составили 95,86 и 92,32 %. Таким образом, соотношение результатов двух постановок различается в зависимости от архитектуры классификатора.

В межприборном эксперименте гребневая регрессия обеспечила наибольший PSNR и наименьший RMSE, а наибольшая ОА для 1DCNN и NM3DCNN получена при нейросетевой настройке.

Дополнительная проверка на Specim FX воспроизводила первую постановку на другой камере: коэффициенты определялись по собственной калибровочной сцене и применялись к двум тестовым сценам. Конфигурации

с наибольшей средней ОА зависели от сцены и архитектуры классификатора. В итоговых таблицах приведены средние значения показателей по трём запускам для выбранных конфигураций.

3.7 Дополнительная демонстрация нейросетевой калибровки

Для наглядной демонстрации действия калибровочных слоёв рассматриваются ранее исследованные калибровочная и тестовая сцены, а также дополнительная сцена с сельскохозяйственными культурами. Сопоставляются спектры в выбранных точках и изображения одного спектрального канала до и после коррекции. На дополнительной сцене также рассматривается изменение значений элементов ГСИ вдоль щели.

На рисунке 3.13 приведены спектры одного класса калибровочной сцены в трёх положениях вдоль щели. После коррекции различия КСЯ между спектрами в выбранных точках уменьшаются, а также уменьшаются спектральные искажения.

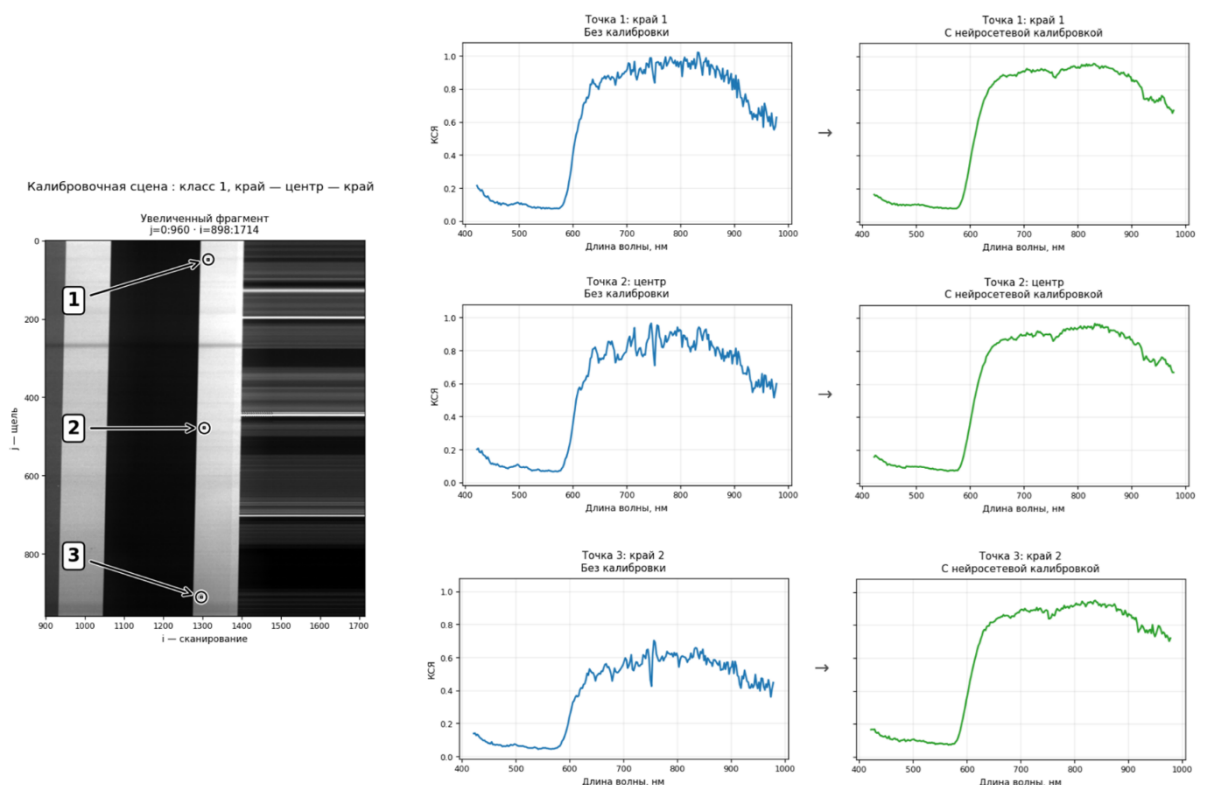


Рисунок 3.13 – Спектры одного класса калибровочной сцены в различных положениях вдоль щели до и после коррекции

На рисунке 3.14 показан пример изменения спектра в отмеченной точке тестовой сцены. После обработки локальные колебания спектра становятся менее выраженными.

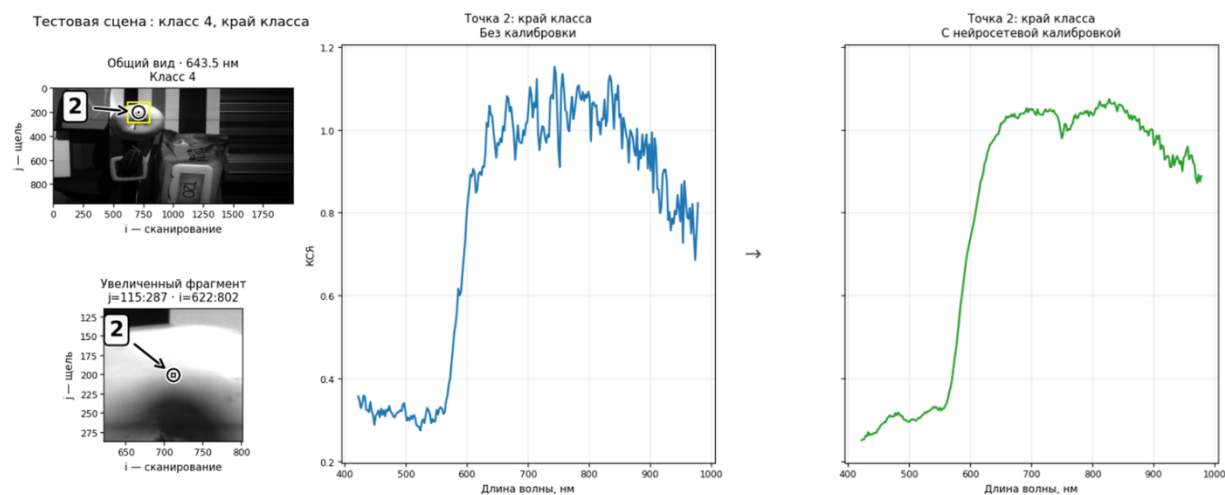


Рисунок 3.14 – Спектр в выбранной точке тестовой сцены до и после коррекции

На дополнительной сцене с сельскохозяйственными культурами сопоставляются изображения спектрального канала 100 в цифровых отсчётах (DN) до и после коррекции (рисунок 3.15).

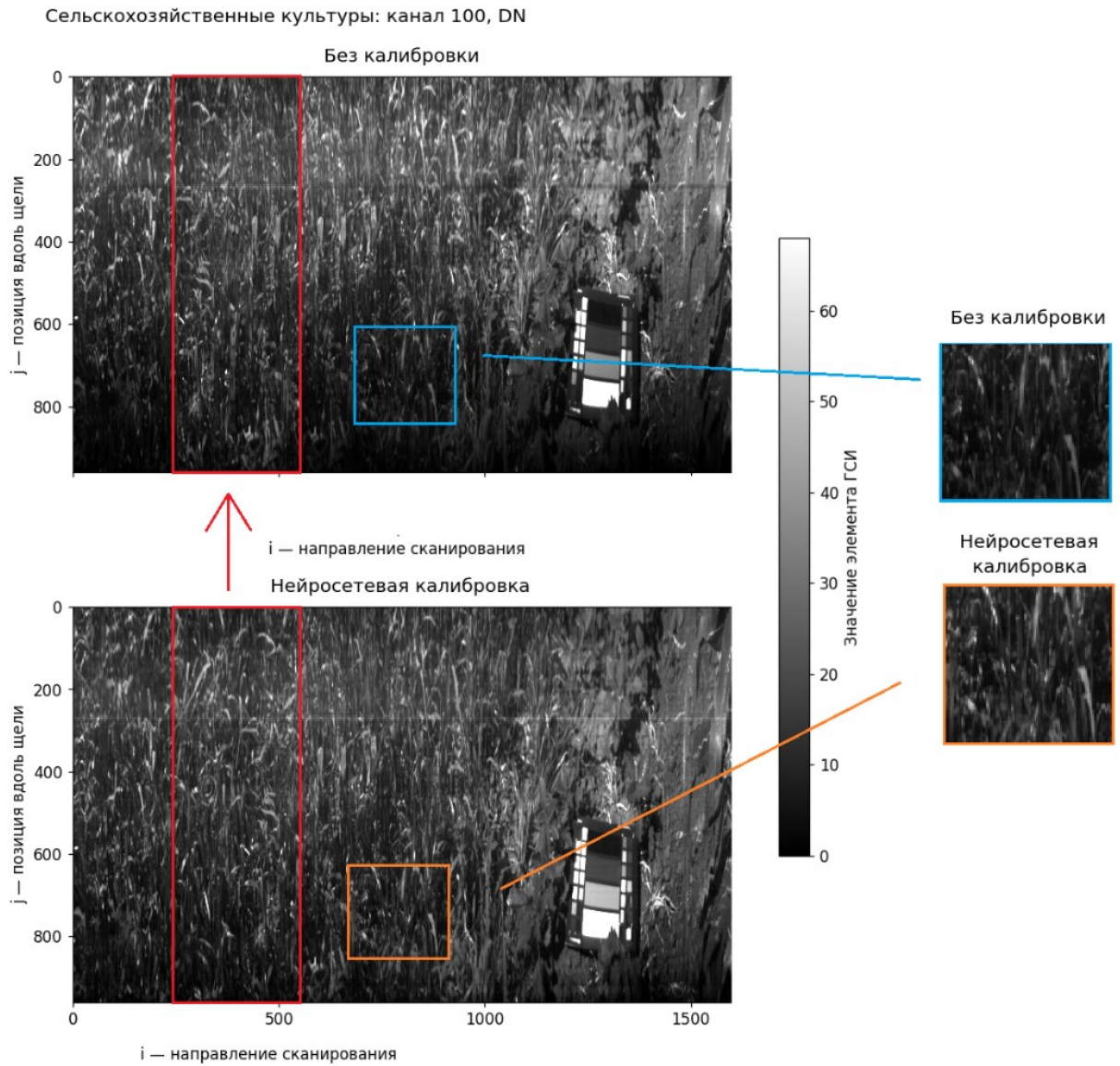


Рисунок 3.15 – Изображение спектрального канала 100 сцены с сельскохозяйственными культурами до и после нейросетевой калибровки.

В верхнее изображение для наглядного сравнения помещён фрагмент скорректированных данных, выделенный красной рамкой. Справа представлены увеличенные фрагменты до и после коррекции, иллюстрирующие компенсацию виньетирования. На нижнем фрагменте после коррекции ослаблено затемнение, наблюдаемое на исходном верхнем фрагменте вблизи края изображения щелевой диафрагмы. Для фиксированного положения по оси сканирования $i = 1000$ приведён пространственный профиль значений элементов ГСИ вдоль щели в том же канале (рисунок 3.16). Линейный тренд после коррекции становится близким к горизонтальному.



Рисунок 3.16 – Пространственный профиль значений элементов ГСИ вдоль щели при $i = 1000$ в спектральном канале 100 до и после коррекции.

Пунктиром показаны линейные тренды

Приведённые примеры дополняют результаты экспериментов, показывая действие калибровочного слоя на отдельные спектры и пространственное распределение значений элементов ГСИ.

3.8 Выводы к главе 3

1. Разработан метод позиционно-зависимой коррекции ГСИ с обучаемыми калибровочными слоями. Рассмотрены матричная и векторная параметризации и их сочетания с полиномиальной составляющей. Коэффициенты зависят от положения j вдоль изображения щелевой

диафрагмы и спектрального канала k . Для их определения использованы гребневая регрессия и нейросетевая настройка по ТКЦЗ на калибровочной сцене с использованием критериев СЕ и COMBO.

2. Рассмотрены две постановки калибровки: по наиболее ярким спектрам классов разработанного гиперспектрометра и по средним спектрам классов Specim IQ, приведённым к сетке длин волн разработанного гиперспектрометра. Во второй постановке сопоставление выполняется по средним спектрам классов опорного и сравниваемого ГСИ.

3. В первой постановке на тестовой сцене наибольшая ОА среди проверенных конфигураций составила 87,21 % для 1DCNN при настройке по COMBO, 98,73 % для NM3DCNN при настройке по СЕ и 98,51 % для SSFTT при настройке по COMBO (таблицы 3.1 и 3.2).

4. Во второй постановке наибольшая ОА для 1DCNN составила 96,78 % со слоем « μ », настроенным по СЕ, для NM3DCNN – 95,86 % со слоем « μ », настроенным по COMBO. Для SSFTT наибольшая ОА получена при гребневой регрессии – 92,32 % (таблицы 3.3 и 3.4).

5. В межприборном эксперименте гребневая регрессия обеспечила наибольший PSNR по средним спектрам классов – 24,22 дБ – и наименьший RMSE по группам «положение вдоль щели – класс» – 0,1067. Минимальный средний SAM по группам – 6,04° – получен для слоя « μ », настроенного с использованием 1DCNN по COMBO (таблица 3.5).

6. Коэффициенты калибровочных слоёв, полученные при настройке по ТКЦЗ на калибровочной сцене, снятой на гиперспектрометр Specim FX, применены к двум другим сценам того же прибора. Обучаемая коррекция повысила ОА и взвешенную F1-меру в пяти из шести рассмотренных сочетаний тестовой сцены и классификатора. Наибольшая ОА достигнута с NM3DCNN: 99,87 % на тестовой сцене №1 и 97,58 % на тестовой сцене №2 (таблица 3.6).

7. Показана возможность применения обучаемого калибровочного слоя с отдельно обученным классификатором другой архитектуры (таблицы 3.2 и 3.4).

8. Методика формирования и применения печатной калибровочной сцены, разработанная в первой главе, использована в обеих постановках позиционно-зависимой коррекции ГСИ. Коэффициенты настроены по калибровочной сцене, а результат их применения оценён по точности поэлементной классификации независимых тестовых сцен.

ЗАКЛЮЧЕНИЕ

В диссертационной работе разработан метод повышения точности поэлементной классификации ГСИ, основанный на выборе параметров их формирования, предобработки и коррекции по результатам классификации ГСИ печатной калибровочной сцены. Для реализации метода разработаны методика формирования и применения этой сцены, алгоритмы спектральной предобработки и обучаемой позиционно-зависимой коррекции отклика гиперспектрометра. Проведена экспериментальная проверка предложенных решений на данных разработанного сканирующего гиперспектрометра основанного на схеме Оффнера, гиперспектрометров Specim IQ и Specim FX, а также открытых наборов ГСИ. Основные результаты работы заключаются в следующем.

1. Получены выражения для пространственных смещений отсчётов при переносе и повороте последовательных проекций щелевой диафрагмы. Для линейного сдвига показатели PSNR, SSIM и SAM представлены через статистические характеристики сравниваемых данных. В экспериментах на ГСИ набора NTIRE и участков речных массивов порядок сравниваемых вариантов ГСИ по точности поэлементной классификации сохранялся при сдвиге предсказанных меток относительно эталонной разметки на 0–5 пикселей, тогда как порядок этих вариантов по PSNR, SSIM и SAM при пространственном рассогласовании изменялся, при разной степени искажений ГСИ.

2. Разработана методика формирования и применения печатной калибровочной сцены с учётом спектральной близости её классов к классам реальных ГСИ и согласованности изменений показателей их классификации. При сопоставлении отобранных пар классов реальных ГСИ и ГСИ печатных сцен при относительном уровне гауссовского шума 12 % коэффициенты корреляции Спирмена составили 0,924 для полноты и 0,798 для F1-меры класса. Выполнено сопоставление классов сформированной калибровочной

сцены с классами набора NTIRE. Калибровочная сцена использована для настройки спектральной предобработки и калибровочных слоёв по точности классификации целевой задачи – ТКЦЗ.

3. Разработаны алгоритмы настройки спектральной предобработки, объединяющие определение коэффициентов спектральной свёртки и программной поправки уровня чёрного. Рассмотрены аналитическое определение параметров методом гребневой регрессии и нейросетевая настройка по ТКЦЗ с использованием критериев «CE» и «COMBO». Для нейросетевых вариантов с наибольшей средней ОА в каждом из пяти рассмотренных сочетаний классификаторов прирост ОА на независимой тестовой сцене составил 2,17–6,22 процентного пункта относительно варианта без дополнительной предобработки. Наибольшая средняя ОА составила 97,37 % при настройке с 1DCNN по комбинированному критерию при $\lambda = 0,1$ и оценке с SSFTT. Сопоставление результатов основной и дополнительных серий носит описательный характер, поскольку параметры обучения этих серий различались. При указанном сочетании классификаторов и комбинированном критерии настройки получена более высокая средняя ОА, чем при применении предобработки на основе моделей DeSpaWN и DeCWN в рассмотренных конфигурациях. Показана возможность применения предобработки, настроенной с 1DCNN, совместно с SVM, kNN и MLP с повышением ОА.

4. Исследовано влияние видеокодеков, параметров сжатия, времени экспозиции и усиления на точность поэлементной классификации. В вычислительном эксперименте сокращение экспозиции с 10 до 0,357 мс при увеличении усиления с 1 до 28 сопровождалось снижением ОА на тестовой сцене на 1,81; 0,90 и 2,51 процентного пункта для 1DCNN, NM3DCNN и SSFTT соответственно. Сохранение среднего цифрового отсчёта при сокращении экспозиции не обеспечило сохранения точности классификации. При фиксированной экспозиции 0,357 мс увеличение усиления с 16 до 28 снизило ОА каждого классификатора на 0,60 процентного пункта на печатной

калибровочной и тестовой сценах. В рассмотренных условиях выбор усиления целесообразно ограничить диапазоном 1–16, определяя конкретное значение по результатам классификации ГСИ печатной калибровочной сцены. При сжатии видеокодеками H.264, H.265 и MJPEG выявлена зависимость точности от степени сжатия, видеокодека и способа преобразования цифровых отсчётов. На калибровочной и тестовой сценах наблюдались сходные общие тенденции изменения ОА при различиях в выраженности и положении отдельных участков снижения точности.

5. Разработаны алгоритмы позиционно-зависимой коррекции ГСИ с обучаемыми калибровочными слоями, коэффициенты которых зависят от положения отсчёта вдоль щели и спектрального канала. Рассмотрены внутриприборная и межприборная постановки, аналитическая настройка методом гребневой регрессии и нейросетевая настройка по ТКЦЗ. Во внутриприборном эксперименте для лучших исследованных конфигураций прирост ОА относительно данных без позиционной коррекции составил 0,65–8,83 процентного пункта; наибольшая ОА достигла 98,73 % при классификации сетью NM3DCNN. В межприборном эксперименте наибольшая ОА составила 96,78 % при использовании 1DCNN.

6. Показана возможность применения коэффициентов коррекции, определённых по калибровочной сцене, к независимым сценам того же прибора, а также использования фиксированного калибровочного слоя с классификатором другой архитектуры. При проверке первой постановки на данных Specim FX настройка по ТКЦЗ обеспечила повышение ОА и взвешенной F1-меры в пяти из шести сочетаний тестовой сцены и классификатора. Наибольшая ОА получена с NM3DCNN: 99,87 % на тестовой сцене №1 и 97,58 % на тестовой сцене №2.

ОПРЕДЕЛЕНИЯ, ОБОЗНАЧЕНИЯ И СОКРАЩЕНИЯ

ГСА – гиперспектральная аппаратура

СГСА – сканирующая гиперспектральная аппаратура

ГСИ – гиперспектральное изображение

БАС – беспилотная авиационная система

КСЯ – коэффициент спектральной яркости

ТКЦЗ – точность классификации целевой задачи

ФРТ – функция рассеяния точки.

DN – цифровой отсчёт зарегистрированного сигнала

ОА – общая точность поэлементной классификации

MSE – средний квадрат ошибки

RMSE – среднеквадратичное отклонение

PSNR – пиковое отношение сигнал/шум

SSIM – индекс структурного сходства

SAM – спектральный угол

CE – категориальная перекрёстная энтропия

COMBO – комбинированная функция потерь, включающая CE и MSE

m – обучаемая калибровочная матрица

v – векторная факторизация калибровочной матрицы

pm – полиномиальный слой и матрица параметров

pv – полиномиальный слой и два вектора параметров

СПИСОК ИСПОЛЬЗОВАННЫХ ИСТОЧНИКОВ

1. Макаров А. Р., Никоноров А. В. Нейросетевой метод коррекции пространственно-спектральной чувствительности целевого гиперспектрометра // Компьютерная оптика. – 2025. – Т. 49, № 6. – С. 947–960. – DOI: 10.18287/COJ1812.
2. Bioucas-Dias J. M., Plaza A., Camps-Valls G., Scheunders P., Nasrabadi N., Chanussot J. Hyperspectral Remote Sensing Data Analysis and Future Challenges // IEEE Geoscience and Remote Sensing Magazine. – 2013. – Vol. 1, No. 2. – P. 6–36. – DOI: 10.1109/MGRS.2013.2244672.
3. Bhargava A., Sachdeva A., Sharma K., Alsharif M. H., Uthansakul P., Uthansakul M. Hyperspectral imaging and its applications: A review // Heliyon. – 2024. – Vol. 10, No. 12. – Art. e33208. – DOI: 10.1016/j.heliyon.2024.e33208.
4. Stuart M. B., McGonigle A. J. S., Willmott J. R. Hyperspectral imaging in environmental monitoring: A review of recent developments and technological advances in compact field deployable systems // Sensors. – 2019. – Vol. 19, No. 14. – Art. 3071. – DOI: 10.3390/s19143071.
5. Ren Y., Liao L., Maybank S. J., Zhang Y., Liu X. Hyperspectral Image Spectral-Spatial Feature Extraction via Tensor Principal Component Analysis. – 2024. – arXiv: 2412.06075.
6. Kolda T. G., Bader B. W. Tensor Decompositions and Applications // SIAM Review. – 2009. – Vol. 51, No. 3. – P. 455–500. – DOI: 10.1137/07070111X.
7. Shaikh M. S., Jaferzadeh K., Thörnberg B., Casselgren J. Calibration of a Hyper-Spectral Imaging System Using a Low-Cost Reference // Sensors. – 2021. – Vol. 21, No. 11. – Art. 3738. – DOI: 10.3390/s21113738.
8. Valme D. et al. From ADAS to Material-Informed Inspection: Review of Hyperspectral Imaging Applications on Mobile Ground Robots // Sensors. – 2025. – Vol. 25, No. 8. – Art. 2346. – DOI: 10.3390/s25082346.
9. Behmann J., Acebron K., Emin D., Bennertz S. et al. Specim IQ: Evaluation of a New, Miniaturized Handheld Hyperspectral Camera and Its Application for

Plant Phenotyping and Disease Detection // *Sensors*. – 2018. – Vol. 18, No. 2. – Art. 441. – DOI: 10.3390/s18020441.

10. Arad B., Timofte R., Yahel R., Morag N., Bernat A. et al. NTIRE 2022 Spectral Recovery Challenge and Data Set // *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops*. – 2022. – P. 863–881. – DOI: 10.1109/CVPRW56347.2022.00102.

11. Novikov I. A., Makarov A. R., Pirogov A. V., Podlipnov V. V., Nikonov A. V. et al. Analysis of Hyperspectral Images of River Waters // *Optical Memory and Neural Networks*. – 2024. – Vol. 33, Suppl. 2. – P. S386–S397. – DOI: 10.3103/S1060992X24700668.

12. Chen Z., Zhang Z., Li H., Li M., Chen Y., Li Q., Feng H., Xu Z., Chen S. Deep Linear Array Pushbroom Image Restoration: A Degradation Pipeline and Jitter-Aware Restoration Network. – 2024. – arXiv: 2401.08171. – DOI: 10.48550/arXiv.2401.08171.

13. Strakhov P., Badasen E., Shurygin B., Kondranin T. A New Control Points Based Geometric Correction Algorithm for Airborne Push Broom Scanner Images without On-board Data // *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences*. – 2016. – Vol. III-1. – P. 99–104. – DOI: 10.5194/isprsannals-III-1-99-2016.

14. Barbieux K. Pushbroom Hyperspectral Data Orientation by Combining Feature-Based and Area-Based Co-Registration Techniques // *Remote Sensing*. – 2018. – Vol. 10, No. 4. – Art. 645. – DOI: 10.3390/rs10040645.

15. Yokoya N., Miyamura N., Iwasaki A. Preprocessing of Hyperspectral Imagery with Consideration of Smile and Keystone Properties // *Proceedings of SPIE: Multispectral, Hyperspectral, and Ultraspectral Remote Sensing Technology, Techniques, and Applications III*. – 2010. – Vol. 7857. – Art. 78570B. – DOI: 10.1117/12.870437.

16. Yokoya N., Miyamura N., Iwasaki A. Detection and Correction of Spectral and Spatial Misregistrations for Hyperspectral Data Using Phase

Correlation Method // *Applied Optics*. – 2010. – Vol. 49, No. 24. – P. 4568–4575. – DOI: 10.1364/AO.49.004568.

17. Gómez-Chova L., Alonso L., Guanter L., Camps-Valls G., Calpe J., Moreno J. Correction of systematic spatial noise in push-broom hyperspectral sensors: Application to CHRIS/PROBA images // *Applied Optics*. – 2008. – Vol. 47, No. 28. – P. F46–F60. – DOI: 10.1364/AO.47.000F46.

18. Rogaß C., Spengler D., Bochow M., Segl K., Lausch A., Doktor D., Roessner S., Behling R., Wetzel H.-U., Kaufmann H. Reduction of Radiometric Miscalibration–Applications to Pushbroom Sensors // *Sensors*. – 2011. – Vol. 11, No. 6. – P. 6370–6395. – DOI: 10.3390/s110606370.

19. Høyе G., Løke T., Fridman A. Method for Quantifying Image Quality in Push-Broom Hyperspectral Cameras // *Optical Engineering*. – 2015. – Vol. 54, No. 5. – Art. 053102. – DOI: 10.1117/1.OE.54.5.053102.

20. Høyе G., Fridman A. Spatial Misregistration in Hyperspectral Cameras: Lab Characterization and Impact on Data Quality in Real-World Images // *Optical Engineering*. – 2020. – Vol. 59, No. 8. – Art. 084103. – DOI: 10.1117/1.OE.59.8.084103.

21. Wang Z., Bovik A. C. Mean Squared Error: Love It or Leave It? A New Look at Signal Fidelity Measures // *IEEE Signal Processing Magazine*. – 2009. – Vol. 26, No. 1. – P. 98–117. – DOI: 10.1109/MSP.2008.930649.

22. Wang Z., Bovik A. C., Sheikh H. R., Simoncelli E. P. Image Quality Assessment: From Error Visibility to Structural Similarity // *IEEE Transactions on Image Processing*. – 2004. – Vol. 13, No. 4. – P. 600–612. – DOI: 10.1109/TIP.2003.819861.

23. Carvalho O. A. de, Jr., Meneses P. R. Spectral Correlation Mapper (SCM): An Improvement on the Spectral Angle Mapper (SAM) // *AVIRIS Proceedings 2000 (JPL Publication 00-18)*. – 2000. – 9 p.

24. ГОСТ Р 59079-2020. Данные дистанционного зондирования Земли из космоса. Типы данных дистанционного зондирования Земли из космоса. – Москва : Стандартинформ, 2020. – URL:

<https://protect.gost.ru/gost/details/7e77e812-19be-4fe1-bdfc-9746ba8af759> (дата обращения: 19.08.2026).

25. ГОСТ Р 59475-2021. Данные дистанционного зондирования Земли из космоса. Качество данных дистанционного зондирования Земли из космоса. Перечень показателей качества данных дистанционного зондирования Земли из космоса, получаемых с космических аппаратов оптико-электронного наблюдения в видимом и ближнем инфракрасном диапазоне. – Москва: Стандартинформ, 2021. – URL: <https://protect.gost.ru/gost/details/fe8ae1af-22d4-4b6a-9542-dec3923cc905> (дата обращения: 19.08.2026).

26. Zhang W., Hu R., Wang J., Zhang L., Zhu C. Spectral Prototype Attention Domain Adaptation for Hyperspectral Image Classification // Remote Sensing. – 2025. – Vol. 17, No. 23. – Art. 3901. – DOI: 10.3390/rs17233901.

27. Kerekes J. P., Hsu S. M. Spectral Quality Metrics for Terrain Classification // Proceedings of SPIE: Imaging Spectrometry X. – 2004. – Vol. 5546. – P. 382–389. – DOI: 10.1117/12.560025.

28. Kerekes J. P., Cisz A. P., Simmons R. E. A Comparative Evaluation of Spectral Quality Metrics for Hyperspectral Imagery // Proceedings of SPIE: Algorithms and Technologies for Multispectral, Hyperspectral, and Ultraspectral Imagery XI. – 2005. – Vol. 5806. – P. 469–480. – DOI: 10.1117/12.605916.

29. Stefanou M. S., Kerekes J. P. A Method for Assessing Spectral Image Utility // IEEE Transactions on Geoscience and Remote Sensing. – 2009. – Vol. 47, No. 6. – P. 1698–1706. – DOI: 10.1109/TGRS.2008.2006364.

30. scikit-learn. Показатели классификации: recall_score и f1_score : официальная документация. – URL: https://scikit-learn.org/stable/modules/generated/sklearn.metrics.recall_score.html; https://scikit-learn.org/stable/modules/generated/sklearn.metrics.f1_score.html (дата обращения: 10.09.2026).

31. Swaminathan V., Thomasson J. A., Hardin R. G., Rajan N. Selection of appropriate multispectral camera exposure settings and radiometric calibration

methods for applications in phenotyping and precision agriculture. – 2024. – arXiv: 2402.18553.

32. Loggenberg K., Strever A., Greyling B., Poona N. Modelling Water Stress in a Shiraz Vineyard Using Hyperspectral Imaging and Machine Learning // Remote Sensing. – 2018. – Vol. 10, No. 2. – Art. 202. – DOI: 10.3390/rs10020202.

33. Gutiérrez S., Fernández-Novales J., Diago M. P., Tardaguila J. On-The-Go Hyperspectral Imaging Under Field Conditions and Machine Learning for the Classification of Grapevine Varieties // Frontiers in Plant Science. – 2018. – Vol. 9. – Art. 1102. – DOI: 10.3389/fpls.2018.01102.

34. Кринов Е. Л. Спектральная отражательная способность природных образований / под ред. К. В. Чибисова. – Москва ; Ленинград : Издательство Академии наук СССР, 1947. – 270, [2] с.

35. Шовенгердт Р. А. Дистанционное зондирование. Модели и методы обработки изображений. – Москва : Техносфера, 2010. – 560 с.

36. Расторгуев А. А., Харитонов С. И., Казанский Н. Л. Моделирование распределения освещённости в плоскости регистратора космического гиперспектрометра, основанного на схеме Оффнера // Компьютерная оптика. – 2017. – Т. 41, № 3. – С. 399–405. – DOI: 10.18287/2412-6179-2017-41-3-399-405.

37. Казанский Н. Л., Харитонов С. И., Досколович Л. Л., Павельев А. В. Моделирование работы космического гиперспектрометра, основанного на схеме Оффнера // Компьютерная оптика. – 2015. – Т. 39, № 1. – С. 70–76. – DOI: 10.18287/0134-2452-2015-39-1-70-76.

38. Расторгуев А. А., Харитонов С. И., Казанский Н. Л. Моделирование формирования изображения космическим гиперспектрометром по схеме Оффнера // Компьютерная оптика. – 2020. – Т. 44, № 1. – С. 12–21. – DOI: 10.18287/2412-6179-CO-644.

39. Фурсов В. А., Бибииков С. А., Байда О. А. Тематическая классификация гиперспектральных изображений по показателю

сопряжённости // Компьютерная оптика. – 2014. – Т. 38, № 1. – С. 154–158. – DOI: 10.18287/0134-2452-2014-38-1-154-158.

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

41. Chapman J. W., Thompson D. R., Helmlinger M. C., Bue B. D., Green R. O., Eastwood M. L., Geier S., Olson-Duvall W., Lundeen S. R. Spectral and Radiometric Calibration of the Next Generation Airborne Visible Infrared Spectrometer (AVIRIS-NG) // Remote Sensing. – 2019. – Vol. 11, No. 18. – Art. 2129. – DOI: 10.3390/rs11182129.

42. Kosec M., Bürmen M., Tomažević D., Pernuš F., Likar B. Characterization of a spectrograph-based hyperspectral imaging system // Optics Express. – 2013. – Vol. 21, No. 10. – P. 12085–12099. – DOI: 10.1364/OE.21.012085.

43. Jemec J., Pernuš F., Likar B., Bürmen M. Push-broom hyperspectral image calibration and enhancement by 2D deconvolution with a variant response function estimate // Optics Express. – 2014. – Vol. 22, No. 22. – P. 27655–27668. – DOI: 10.1364/OE.22.027655.

44. Baumgartner A. Traceable Imaging Spectrometer Calibration and Transformation of Geometric and Spectral Pixel Properties : dissertation. – Osnabrück : Universität Osnabrück, 2022. – 154 p. – DOI: 10.48693/38.

45. He Z., Liu L. Hyperspectral Image Super-Resolution Inspired by Deep Laplacian Pyramid Network // Remote Sensing. – 2018. – Vol. 10, No. 12. – Art. 1939. – DOI: 10.3390/rs10121939.

46. Cai Y., Lin J., Lin Z., Wang H., Zhang Y., Pfister H., Timofte R., Van Gool L. MST++: Multi-Stage Spectral-Wise Transformer for Efficient Spectral Reconstruction // Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops. – 2022. – P. 745–755. – DOI: 10.1109/CVPRW56347.2022.00090.
47. Hu W., Huang Y., Wei L., Zhang F., Li H. Deep Convolutional Neural Networks for Hyperspectral Image Classification // Journal of Sensors. – 2015. – Vol. 2015. – Art. 258619. – DOI: 10.1155/2015/258619.
48. He M., Li B., Chen H. Multi-Scale 3D Deep Convolutional Neural Network for Hyperspectral Image Classification // 2017 IEEE International Conference on Image Processing. – 2017. – P. 3904–3908. – DOI: 10.1109/ICIP.2017.8297014.
49. Sun L., Zhao G., Zheng Y., Wu Z. Spectral–Spatial Feature Tokenization Transformer for Hyperspectral Image Classification // IEEE Transactions on Geoscience and Remote Sensing. – 2022. – Vol. 60. – Art. 5522214. – DOI: 10.1109/TGRS.2022.3144158.
50. Rasti B., Scheunders P., Ghamisi P., Licciardi G., Chanussot J. Noise Reduction in Hyperspectral Imagery: Overview and Application // Remote Sensing. – 2018. – Vol. 10, No. 3. – Art. 482. – DOI: 10.3390/rs10030482.
51. Li C., Li Z., Liu X., Li S. The Influence of Image Degradation on Hyperspectral Image Classification // Remote Sensing. – 2022. – Vol. 14, No. 20. – Art. 5199. – DOI: 10.3390/rs14205199.
52. Miao L., Yu T., Gu X., Sun Z., Yang J., Zhang Z., Mi X., Cao W., Li J. The Impact of Spatial Resolution on the Classification of Vegetation Types in Highly Fragmented Planting Areas Based on Unmanned Aerial Vehicle Hyperspectral Images // Remote Sensing. – 2020. – Vol. 12, No. 1. – Art. 146. – DOI: 10.3390/rs12010146.
53. Savitzky A., Golay M. J. E. Smoothing and Differentiation of Data by Simplified Least Squares Procedures // Analytical Chemistry. – 1964. – Vol. 36, No. 8. – P. 1627–1639. – DOI: 10.1021/ac60214a047.

54. Verbyla D. L., Hammond T. O. Conservative Bias in Classification Accuracy Assessment Due to Pixel-by-Pixel Comparison of Classified Images with Reference Grids // *International Journal of Remote Sensing*. – 1995. – Vol. 16, No. 3. – P. 581–587. – DOI: 10.1080/01431169508954424.
55. Gu J., Congalton R. G. Analysis of the Impact of Positional Accuracy When Using a Single Pixel for Thematic Accuracy Assessment // *Remote Sensing*. – 2020. – Vol. 12, No. 24. – Art. 4093. – DOI: 10.3390/rs12244093.
56. Gu J., Congalton R. G. Analysis of the Impact of Positional Accuracy When Using a Block of Pixels for Thematic Accuracy Assessment // *Geographies*. – 2021. – Vol. 1, No. 2. – P. 143–165. – DOI: 10.3390/geographies1020009.
57. Du Z., You S., Cheng C., Wei S. Automatic Spectral Calibration of Hyperspectral Images: Method, Dataset and Benchmark // *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. – 2025. – P. 28081–28090. – DOI: 10.1109/CVPR52734.2025.02615.
58. Bähler N., El Helou M., Objois É., Okumuş K., Süsstrunk S. PoGaIN: Poisson-Gaussian Image Noise Modeling From Paired Samples // *IEEE Signal Processing Letters*. – 2022. – Vol. 29. – P. 2602–2606. – DOI: 10.1109/LSP.2022.3227522.
59. Foi A., Trimeche M., Katkovnik V., Egiazarian K. Practical Poissonian-Gaussian Noise Modeling and Fitting for Single-Image Raw-Data // *IEEE Transactions on Image Processing*. – 2008. – Vol. 17, No. 10. – P. 1737–1754. – DOI: 10.1109/TIP.2008.2001399.
60. Fu Q., Souza M., Choi E., Shin S., Baek S.-H., Heidrich W. Limitations of Data-Driven Spectral Reconstruction: An Optics-Aware Analysis // *IEEE Transactions on Computational Imaging*. – 2025. – Vol. 11. – P. 1505–1520. – DOI: 10.1109/TCI.2025.3622928.
61. Jeon D. S., Baek S.-H., Yi S., Fu Q., Dun X., Heidrich W., Kim M. H. Compact Snapshot Hyperspectral Imaging with Diffracted Rotation // *ACM Transactions on Graphics*. – 2019. – Vol. 38, No. 4. – Art. 117. – DOI: 10.1145/3306346.3322946.

62. Ruddick K. G., Brando V. E., Corizzi A. et al. WATERHYPERNET: a prototype network of automated in situ measurements of hyperspectral water reflectance for satellite validation and water quality monitoring // *Frontiers in Remote Sensing*. – 2024. – Vol. 5. – Art. 1347520. – DOI: 10.3389/frsen.2024.1347520.

63. NASA Jet Propulsion Laboratory. AVIRIS – Airborne Visible / Infrared Imaging Spectrometer. Data Processing. – URL: https://aviris.jpl.nasa.gov/aviris/data_facility.html (дата обращения: 10.09.2026).

64. Wang M., Fan C., Pan J., Jin S., Chang X. Image Jitter Detection and Compensation Using a High-Frequency Angular Displacement Method for Yaogan-26 Remote Sensing Satellite // *ISPRS Journal of Photogrammetry and Remote Sensing*. – 2017. – Vol. 130. – P. 32–43. – DOI: 10.1016/j.isprsjprs.2017.05.004.

65. Wang Z., Zhang Z., Dong L., Xu G. Jitter Detection and Image Restoration Based on Generative Adversarial Networks in Satellite Images // *Sensors*. – 2021. – Vol. 21, No. 14. – Art. 4693. – DOI: 10.3390/s21144693.

66. ГОСТ Р 59830-2021. Дистанционное зондирование Земли из космоса. Данные дистанционного зондирования Земли из космоса. Перечень потребительских характеристик данных, получаемых с космических аппаратов оптико-электронного наблюдения в видимом и инфракрасном диапазоне. – Москва : ФГБУ «РСТ», 2022. – 12 с. – URL: https://meganorm.ru/mega_doc/norm/gost-r_gosudarstvennyj-standart/9/gost_r_59830-2021_natsionalnyy_standart_rossiyskoy.html (дата обращения: 19.08.2026).

67. ГОСТ Р 59474-2021. Данные дистанционного зондирования Земли из космоса. Качество данных дистанционного зондирования Земли из космоса. Оценка качества данных дистанционного зондирования Земли из космоса и продуктов их обработки. Общие положения. – Москва : Стандартинформ, 2021. – 16 с. – URL: <https://protect.gost.ru/gost/details/04d5b548-24cd-45b3-b4e3-327a971bea94> (дата обращения: 19.08.2026).

68. Фирсов Н. А. Метод и алгоритмы нейросетевой классификации гиперспектральных изображений : дис. ... канд. техн. наук : 1.2.1. – Самара, 2025. – 144 с. – URL: https://ssau.ru/storage/pages/6809/file_68da34e6b29b09.75837130.pdf.
69. Chen C., Wang Y., Zhang N., Zhang Y., Zhao Z. A Review of Hyperspectral Image Super-Resolution Based on Deep Learning // Remote Sensing. – 2023. – Vol. 15, No. 11. – Art. 2853. – DOI: 10.3390/rs15112853.
70. Mantripragada K., Dao P. D., He Y., Qureshi F. Z. The Effects of Spectral Dimensionality Reduction on Hyperspectral Pixel Classification: A Case Study // PLOS ONE. – 2022. – Vol. 17, No. 7. – Art. e0269174. – DOI: 10.1371/journal.pone.0269174.
71. Wang Z., Simoncelli E. P. Translation Insensitive Image Similarity in Complex Wavelet Domain // IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP). – 2005. – P. II/573–II/576. – DOI: 10.1109/ICASSP.2005.1415469.
72. Kim W.-H., Lee J.-S. Framework for Fair Objective Performance Evaluation of Single-Image Super-Resolution Algorithms // Electronics Letters. – 2015. – Vol. 51, No. 1. – P. 42–44. – DOI: 10.1049/el.2014.2784.
73. Zhang X., Cheng X., Xue T., Wang Y. Linear Spatial Misregistration Detection and Correction Based on Spectral Unmixing for FAHI Hyperspectral Imagery // Sensors. – 2022. – Vol. 22, No. 24. – Art. 9932. – DOI: 10.3390/s22249932.
74. Taylor K. E. Summarizing Multiple Aspects of Model Performance in a Single Diagram // Journal of Geophysical Research: Atmospheres. – 2001. – Vol. 106, No. D7. – P. 7183–7192. – DOI: 10.1029/2000JD900719.
75. Batyrshin I., Kreinovich V. One More Geometric Interpretation of Pearson's Correlation // Thailand Statistician. – 2015. – Vol. 13, No. 1. – P. 125–126.
76. Kruse F. A., Lefkoff A. B., Boardman J. W., Heidebrecht K. B., Shapiro A. T., Barloon P. J., Goetz A. F. H. The Spectral Image Processing System (SIPS) –

Interactive Visualization and Analysis of Imaging Spectrometer Data // Remote Sensing of Environment. – 1993. – Vol. 44, No. 2–3. – P. 145–163. – DOI: 10.1016/0034-4257(93)90013-N.

77. Uhl J. H., Leyk S. Spatially Explicit Accuracy Assessment of Deep Learning-Based, Fine-Resolution Built-Up Land Data in the United States // International Journal of Applied Earth Observation and Geoinformation. – 2023. – Vol. 123. – Art. 103469. – DOI: 10.1016/j.jag.2023.103469.

78. Rossi M. W. Evaluating the Accuracy of Binary Classifiers for Geomorphic Applications // Earth Surface Dynamics. – 2024. – Vol. 12. – P. 765–782. – DOI: 10.5194/esurf-12-765-2024.

79. Новиков И. А., Макаров А. Р., Подлипнов В. В., Платонов В. И., Рыськова Д. Д., Калашникова О. В., Хабибуллин Р. М., Скиданов Р. В., Илларионова С. В., Выборнова Ю. В., Никоноров А. В., Шадрин Д. Г., Подладчикова Т. В. Классификация интенсивности цветения водорослей на основе методов машинного обучения и гиперспектральных данных БПЛА // Компьютерная оптика. – 2025. – Т. 49, № 6. – С. 972–985. – DOI: 10.18287/2412-6179-CO-1539.

80. Kirillov A., Mintun E., Ravi N., Mao H., Rolland C., Gustafson L. et al. Segment Anything // Proceedings of the IEEE/CVF International Conference on Computer Vision. – 2023. – P. 4015–4026. – DOI: 10.1109/ICCV51070.2023.00371.

81. Yuhas R. H., Goetz A. F. H., Boardman J. W. Discrimination among Semi-Arid Landscape Endmembers Using the Spectral Angle Mapper (SAM) Algorithm // Summaries of the Third Annual JPL Airborne Geoscience Workshop. Vol. 1: AVIRIS Workshop. JPL Publication 92-14. – 1992. – P. 147–149.

82. Specim, Spectral Imaging Ltd. Specim IQ: Hyperspectral Goes Mobile. Data Sheet. – URL: <https://www.specim.com/wp-content/uploads/2024/11/Specim-IQ-Datasheet-Long-08.pdf> (дата обращения: 23.08.2026).

83. Fu J., Liu J., Zhao R., Chen Z., Qiao Y., Li D. Maize Disease Detection Based on Spectral Recovery from RGB Images // *Frontiers in Plant Science*. – 2022. – Vol. 13. – Art. 1056842. – DOI: 10.3389/fpls.2022.1056842.

84. Фирсов Н. А., Подлипов В. В., Ивлиев Н. А., Николаев П. П., Машков С. В., Ишкин П. А., Скиданов Р. В., Никоноров А. В. Нейросетевая классификация гиперспектральных изображений растительности с формированием обучающей выборки на основе адаптивного вегетационного индекса // *Компьютерная оптика*. – 2021. – Т. 45, № 6. – С. 887–896. – DOI: 10.18287/2412-6179-CO-1038.

85. Pronin A., Vladimirov D., Korepanov A., Muzyka A., Makarov A., Podtikhova S., Ershov E., Rastorguev A., Skidanov R., Nikonorov A. Diffract and Conquer: Hyperspectral Imaging from Any RGB Camera via Optical Encoding and Learning // *Proceedings of the Thirty-Fifth International Joint Conference on Artificial Intelligence (IJCAI-26)*. – 2026. – P. 1523–1531. – DOI: 10.24963/ijcai.2026/170.

86. CIE. Colour-matching Functions of CIE 1931 Standard Colorimetric Observer [Data set]. – Vienna : International Commission on Illumination, 2019. – DOI: 10.25039/CIE.DS.xvudnb9b.

87. International Color Consortium. How to Interpret the sRGB Color Space (Specified in IEC 61966-2-1) for ICC Profiles. – 2015. – URL: <https://registry.color.org/rgb-registry/files/sRGB.pdf> (дата обращения: 08.09.2026).

88. CIE. CIE standard illuminant D65. – Vienna : International Commission on Illumination, 2019. – DOI: 10.25039/CIE.DS.hjfm59. – URL: <https://cie.co.at/datatable/cie-standard-illuminant-d65> (дата обращения: 08.09.2026).

89. Loshchilov I., Hutter F. Decoupled Weight Decay Regularization // *International Conference on Learning Representations*. – 2019. – arXiv: 1711.05101.

90. Cui Z., Kerekes J. P., DeAngelis C., Brown S. D., Nance C. E. A Comparison of Real and Simulated Airborne Hyperspectral Imagery // 2014 IEEE Western New York Image and Signal Processing Workshop. – 2014. – P. 19–22. – DOI: 10.1109/WNYIPW.2014.6999478.

91. Bue B. D., Merényi E., Csathó B. An Evaluation of Class Knowledge Transfer from Synthetic to Real Hyperspectral Imagery // 2011 3rd Workshop on Hyperspectral Image and Signal Processing: Evolution in Remote Sensing. – 2011. – P. 1–4. – DOI: 10.1109/WHISPERS.2011.6080872.

92. Bandyopadhyay S., Mandal S. Effect of Gray Component Replacement on Color Reproduction // Proceedings of IS&T's PICS 2000. – 2000. – P. 188–191. – URL: <https://www.imaging.org/common/uploaded%20files/pdfs/Papers/2000/PICS-0-81/1613.pdf>.

93. Hersch R. D., Crété F. Improving the Yule–Nielsen Modified Neugebauer Model by Dot Surface Coverages Depending on the Ink Superposition Conditions // Proceedings of SPIE: Color Imaging X: Processing, Hardcopy, and Applications. – 2005. – Vol. 5667. – P. 434–445. – DOI: 10.1117/12.588907.

94. Wyble D. R., Berns R. S. A Critical Review of Spectral Models Applied to Binary Color Printing // Color Research & Application. – 2000. – Vol. 25, No. 1. – P. 4–19. – DOI: 10.1002/(SICI)1520-6378(200002)25:1<4::AID-COL3>3.0.CO;2-X.

95. SciPy Developers. `scipy.stats.spearmanr`; `scipy.stats.kendalltau` // Документация SciPy. – URL: <https://docs.scipy.org/doc/scipy/reference/generated/scipy.stats.kendalltau.html> (дата обращения: 10.09.2026).

96. Bland J. M., Altman D. G. Statistical Methods for Assessing Agreement between Two Methods of Clinical Measurement // The Lancet. – 1986. – Vol. 327, No. 8476. – P. 307–310. – DOI: 10.1016/S0140-6736(86)90837-8.

97. McCamy C. S., Marcus H., Davidson J. G. A Color-Rendition Chart // Journal of Applied Photographic Engineering. – 1976. – Vol. 2, No. 3. – P. 95–99.

98. Rychtáriková R., Souček P., Štys D. Colorimetric Calibration of a Digital Camera. – 2017. – 14 p. – arXiv: 1708.04685. – URL: <https://arxiv.org/abs/1708.04685>.
99. Zhang Q., Yuan Q., Li J., Liu X., Shen H., Zhang L. Hybrid Noise Removal in Hyperspectral Imagery With a Spatial–Spectral Gradient Network // IEEE Transactions on Geoscience and Remote Sensing. – 2019. – Vol. 57, No. 10. – P. 7317–7329. – DOI: 10.1109/TGRS.2019.2912909.
100. Poyser M., Atapour-Abarghouei A., Breckon T. P. On the Impact of Lossy Image and Video Compression on the Performance of Deep Convolutional Neural Network Architectures // 2020 25th International Conference on Pattern Recognition (ICPR). – 2021. – P. 2830–2837. – DOI: 10.1109/ICPR48806.2021.9412455.
101. Zhang T., Fu Y., Li C. Hyperspectral Image Denoising With Realistic Data // Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV). – 2021. – P. 2248–2257.
102. Hasinoff S. W., Durand F., Freeman W. T. Noise-Optimal Capture for High Dynamic Range Photography // 2010 IEEE Computer Society Conference on Computer Vision and Pattern Recognition. – 2010. – P. 553–560. – DOI: 10.1109/CVPR.2010.5540167.
103. Goodfellow I., Bengio Y., Courville A. Deep Learning. – Cambridge, MA : MIT Press, 2016. – 800 p. – URL: <https://www.deeplearningbook.org/>.
104. Schmid M., Rath D., Diebold U. Why and How Savitzky–Golay Filters Should Be Replaced // ACS Measurement Science Au. – 2022. – Vol. 2, No. 2. – P. 185–196. – DOI: 10.1021/acsmeasuresciau.1c00054.
105. SciPy Developers. `scipy.signal.savgol_filter` // Документация SciPy. – URL: https://docs.scipy.org/doc/scipy/reference/generated/scipy.signal.savgol_filter.html (дата обращения: 12.09.2026).

106. Тихонов А. Н. О решении некорректно поставленных задач и методе регуляризации // Доклады АН СССР. – 1963. – Т. 151, № 3. – С. 501–504. – URL: <https://www.mathnet.ru/rus/dan28329>.
107. Lin Y.-T., Finlayson G. D. On the Optimization of Regression-Based Spectral Reconstruction // Sensors. – 2021. – Vol. 21, No. 16. – Art. 5586. – DOI: 10.3390/s21165586.
108. Baumer. Operating Manual: VCXG.2 / VCXG.2.XC / VCXG.2.I / VCXU.2. – Version 1.7. – 2024. – URL: https://media.baumer.com/mamfile_2084982.pdf (дата обращения: 28.05.2026).
109. Lobanov V., Firsov N., Myasnikov E., Khabibullin R., Nikonorov A. HyperKAN: Kolmogorov-Arnold Networks make Hyperspectral Image Classifiers Smarter. – 2024. – arXiv: 2407.05278. – URL: <https://arxiv.org/abs/2407.05278>.
110. Kingma D. P., Ba J. Adam: A Method for Stochastic Optimization // International Conference on Learning Representations. – 2015. – arXiv: 1412.6980.
111. Michau G., Frusque G., Fink O. Fully learnable deep wavelet transform for unsupervised monitoring of high-frequency time series // Proceedings of the National Academy of Sciences. – 2022. – Vol. 119, No. 8. – Art. e2106598119. – DOI: 10.1073/pnas.2106598119.
112. Michau G. DeSpaWN: программная реализация обучаемого вейвлетного преобразования. – GitHub. – URL: <https://github.com/MichauGabriel/DeSpaWN> (дата обращения: 10.09.2026).
113. Oudijk A. E., Sigernes F., Mulders H. C. J., Bakken S., Johansen T. A. Quality Assessments of Standard Video Compression Techniques Applied to Hyperspectral Data Cubes // 2021 11th Workshop on Hyperspectral Imaging and Signal Processing: Evolution in Remote Sensing (WHISPERS). – 2021. – P. 1–5. – DOI: 10.1109/WHISPERS52202.2021.9483989.
114. Radosavljević M., Brkljač B., Lugonja P., Crnojević V., Trpovski Ž., Xiong Z., Vukobratović D. Lossy Compression of Multispectral Satellite Images with Application to Crop Thematic Mapping: A HEVC Comparative Study //

Remote Sensing. – 2020. – Vol. 12, No. 10. – Art. 1590. – DOI: 10.3390/rs12101590.

115. W3C. Portable Network Graphics (PNG) Specification (Third Edition). – W3C Recommendation, 24 June 2025. – Разд. 12.1: Encoder gamma handling. – URL: <https://www.w3.org/TR/2025/REC-png-3-20250624/> (дата обращения: 08.09.2026).

116. Berc L., Fenner W., Frederick R., McCanne S., Stewart P. RTP Payload Format for JPEG-compressed Video. – RFC 2435. – IETF, October 1998. – DOI: 10.17487/RFC2435. – URL: <https://www.rfc-editor.org/rfc/rfc2435.html> (дата обращения: 08.09.2026).

117. ITU-T. Recommendation H.264 (11/2007): Advanced video coding for generic audiovisual services. – Geneva : ITU, 2007. – URL: https://www.itu.int/rec/dologin_pub.asp?id=T-REC-H.264-200711-S!!PDF-E&lang=e&type=items (дата обращения: 08.09.2026).

118. ITU-T. Recommendation H.265 (04/2013): High efficiency video coding. – Geneva : ITU, 2013. – URL: https://www.itu.int/rec/dologin_pub.asp?id=T-REC-H.265-201304-S!!PDF-E&lang=e&type=items (дата обращения: 08.09.2026).

119. FFmpeg Developers. FFmpeg Codecs Documentation: libx264, libx264rgb, libx265. – URL: <https://ffmpeg.org/ffmpeg-codecs.html> (дата обращения: 11.09.2026).

120. FFmpeg Developers. ffmpeg Documentation: Main options. – URL: <https://ffmpeg.org/ffmpeg.html> (дата обращения: 11.09.2026).

121. FFmpeg Developers. FFmpeg 7.1 API Documentation: libavutil/pixfmt.h; libavutil/pixdesc.c. – URL: https://ffmpeg.org/doxygen/7.1/pixdesc_8c_source.html (дата обращения: 11.09.2026).

122. OpenCV Developers. OpenCV 4.13.0 Documentation: Color conversions. – URL:

https://docs.opencv.org/4.13.0/de/d25/imgproc_color_conversions.html (дата обращения: 11.09.2026).

123. Jiang H., Zhong Z., Zheng Y. Gain-first or Exposure-first: Benchmark for Better Low-light Video Photography and Enhancement // Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR) Workshops. – 2024. – P. 1345–1356. – URL: https://openaccess.thecvf.com/content/CVPR2024W/PBDL/html/Jiang_Gain-first_or_Exposure-first_Benchmark_for_Better_Low-light_Video_Photography_and_CVPRW_2024_paper.html.

124. Bakker W. H., van der Werff H. M. A., van der Meer F. D. Determining smile and keystone of hyperspectral lab cameras // 2019 10th Workshop on Hyperspectral Imaging and Signal Processing: Evolution in Remote Sensing (WHISPERS). – 2019. – P. 1–5. – DOI: 10.1109/WHISPERS.2019.8921045.

125. Riihiahio K. A., Eskelinen M. A., Pölönen I. A Do-It-Yourself Hyperspectral Imager Brought to Practice with Open-Source Python // Sensors. – 2021. – Vol. 21, No. 4. – Art. 1072. – DOI: 10.3390/s21041072.

126. Fridman A., Høye G., Løke T. Resampling in hyperspectral cameras as an alternative to correcting keystone in hardware, with focus on benefits for optical design and data quality // Optical Engineering. – 2014. – Vol. 53, No. 5. – Art. 053107. – DOI: 10.1117/1.OE.53.5.053107.

127. Høye G., Fridman A. Method for calibrating the image from a mixel camera based solely on the acquired hyperspectral data // EARSeL eProceedings. – 2013. – Vol. 12, No. 2. – P. 174–181. – DOI: 10.12760/01-2013-2-09.

128. Morales A., Horstrand P., Guerra R., Leon R., Ortega S., Díaz M., Melián J. M., López S., López J. F., Callico G. M., Martel E., Sarmiento R. Laboratory Hyperspectral Image Acquisition System Setup and Validation // Sensors. – 2022. – Vol. 22, No. 6. – Art. 2159. – DOI: 10.3390/s22062159.

129. Davies M., Stuart M. B., Hobbs M. J., McGonigle A. J. S., Willmott J. R. Image Correction and In Situ Spectral Calibration for Low-Cost, Smartphone

Hyperspectral Imaging // Remote Sensing. – 2022. – Vol. 14, No. 5. – Art. 1152. – DOI: 10.3390/rs14051152.

130. Makarov A., Mirpulatov I., Firsov N., Lobanov V., Illarionova S., Podlipnov V., Vybornova Y., Shadrin D., Rastorguev A., Skidanov R., Burnaev E., Nikonorov A. Deep Spectral-Spatial Transformer for Robust Hyperspectral Image Segmentation in Varying Field Conditions // IEEE Access. – 2025. – Vol. 13. – P. 97454–97471. – DOI: 10.1109/ACCESS.2025.3575699.

131. Pillay R., Picollo M., Hardeberg J. Y., George S. Evaluation of the Data Quality from a Round-Robin Test of Hyperspectral Imaging Systems // Sensors. – 2020. – Vol. 20, No. 14. – Art. 3812. – DOI: 10.3390/s20143812.

132. Henriksen M. L., Pedersen W. N., Klarskov P., Hinge M. One Step Calibration of Industrial Hyperspectral Cameras // Chemometrics and Intelligent Laboratory Systems. – 2022. – Vol. 227. – Art. 104609. – DOI: 10.1016/j.chemolab.2022.104609.

133. Vitorino T., Casini A., Cucci C., Gebejesje A., Hiltunen J., Hauta-Kasari M., Picollo M., Stefani L. Accuracy in Colour Reproduction: Using a ColorChecker Chart to Assess the Usefulness and Comparability of Data Acquired with Two Hyper-Spectral Systems // Computational Color Imaging. – Cham : Springer, 2015. – (Lecture Notes in Computer Science; Vol. 9016). – P. 225–235. – DOI: 10.1007/978-3-319-15979-9_21.

134. Signoroni A., Conte M., Plutino A., Rizzi A. Spatial–Spectral Evidence of Glare Influence on Hyperspectral Acquisitions // Sensors. – 2020. – Vol. 20, No. 16. – Art. 4374. – DOI: 10.3390/s20164374.

135. Huang C., Zhang L., Fang J., Tong Q. A Radiometric Calibration Model for the Field Imaging Spectrometer System // IEEE Transactions on Geoscience and Remote Sensing. – 2013. – Vol. 51, No. 4. – P. 2465–2475. – DOI: 10.1109/TGRS.2012.2211026.

136. Mertens S., Verbraeken L., Sprenger H., Demuynck K., Maleux K., Cannoot B., De Block J., Maere S., Nelissen H., Bonaventure G., Crafts-Brandner S. J., Vogel J. T., Bruce W., Inzé D., Wuyts N. Proximal Hyperspectral Imaging

Detects Diurnal and Drought-Induced Changes in Maize Physiology // *Frontiers in Plant Science*. – 2021. – Vol. 12. – Art. 640914. – DOI: 10.3389/fpls.2021.640914.

137. Juola J., Hovi A., Rautiainen M. Practical recommendations and limitations for pushbroom hyperspectral imaging of tree stems // *Remote Sensing of Environment*. – 2023. – Vol. 298. – Art. 113837. – DOI: 10.1016/j.rse.2023.113837.

138. Krawczyk H., Gerasch B., Walzel T., Storch T., Müller R., Sang B., Chlebek C. EnMAP Radiometric Inflight Calibration // *Proceedings of IEEE International Geoscience and Remote Sensing Symposium (IGARSS) 2014*. – 2014. – P. 1–5. – URL: https://elib.dlr.de/90656/1/IGARSS_2014_Krawczyk_Harald_EnMAP-v1.pdf.

139. Markelin L., Honkavaara E., Takala T., Pellikka P. Calibration and validation of hyperspectral imagery using a permanent test field // *2013 5th Workshop on Hyperspectral Image and Signal Processing: Evolution in Remote Sensing (WHISPERS)*. – 2013. – P. 1–4. – DOI: 10.1109/WHISPERS.2013.8080708.

140. Baumann A. et al. Neural Illumination Calibration for Surgical Workflow-Optimized Spectral Imaging // *International Journal of Computer Assisted Radiology and Surgery*. – 2026. – Vol. 21, No. 4. – P. 665–675. – DOI: 10.1007/s11548-025-03525-8.

141. Darrodi M. M., Finlayson G. D., Goodman T., Mackiewicz M. Reference Data Set for Camera Spectral Sensitivity Estimation // *Journal of the Optical Society of America A*. – 2015. – Vol. 32, No. 3. – P. 381–391. – DOI: 10.1364/JOSAA.32.000381.

142. Boucher T., Dyar M. D., Mahadevan S. Proximal Methods for Calibration Transfer // *Journal of Chemometrics*. – 2017. – Vol. 31, No. 4. – Art. e2877. – DOI: 10.1002/cem.2877.

143. Specim, Spectral Imaging Oy Ltd. Specim FX17: User Manual. – Version 2.2. – 2020. – 50 p. – URL: <https://ftp.stemmer-imaging.com/webdavs/docmanager/152613-Specim-FX17-Reference-Manual.pdf> (дата обращения: 08.09.2026).

144. Specim, Spectral Imaging Ltd. Specim FX17. – URL: <https://www.specim.com/products/specim-fx17/> (дата обращения: 08.09.2026).

ПРИЛОЖЕНИЕ А

АКТ О ВНЕДРЕНИИ

результатов исследования по теме «Метод и нейросетевые алгоритмы повышения точности поэлементной классификации гиперспектральных изображений»

Настоящий Акт свидетельствует, что результаты исследований А.Р. Макарова по теме «Метод и нейросетевые алгоритмы повышения точности поэлементной классификации гиперспектральных изображений» внедрены в Центр «Интеллектуальная мобильность многофункциональных беспилотных авиационных систем» Самарского университета им. Королева и используются в работах по направлению «Разработка фреймворка дистанционного мониторинга на основе интеллектуального анализа мультимодальных данных, полученных на борту беспилотных авиационных систем FW_Monitor».

Разработки А.Р. Макарова по выбору параметров съёмки и калибровке сканирующих гиперспектрометров, а также нейросетевой спектральной предобработке и позиционно-зависимой коррекции данных используются в задачах поэлементной классификации гиперспектральных изображений. Указанные разработки позволяют настраивать параметры съёмки и обработки данных с учётом точности классификации, уменьшать влияние шумов и неоднородности отклика регистрирующего устройства и повышать точность поэлементной классификации.

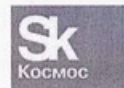
Данные разработки применяются при сборе, подготовке и анализе гиперспектральных данных в рамках проекта «Разработка фреймворка дистанционного мониторинга на основе интеллектуального анализа мультимодальных данных, полученных на борту беспилотных авиационных систем FW_Monitor».

Заместитель директора
Центра «Интеллектуальная
мобильность многофункциональных
беспилотных авиационных систем»
к.т.н.



Ю.Д. Выборнова

ПРИЛОЖЕНИЕ Б

**ООО ЛОКУС**

443051, Самарская область, г. Самара, Кировский район,
ул. Республиканская, дом 65, кв. 38
ИНН 7731315489 КПП 631201001 ОГРН 1167746437990
ОКПО 2337363 ОКВЭД 72.19.2, 72.19.3
Тел./факс: +7 (846) 2702385
artniko@gmail.com

АКТ о внедрении

**результатов исследования по теме «Метод и нейросетевые алгоритмы повышения
точности поэлементной классификации гиперспектральных изображений»**

Настоящий Акт свидетельствует, что результаты исследований А.Р. Макарова по теме «Метод и нейросетевые алгоритмы повышения точности поэлементной классификации гиперспектральных изображений» внедрены в ООО «Локус» и используются в рамках проекта «Разработка открытой платформы сбора и анализа гиперспектральных изображений для решения прикладных задач цифровой экономики».

Разработки А.Р. Макарова по выбору параметров съёмки и калибровке сканирующих гиперспектрометров, а также нейросетевой спектральной предобработке и позиционно-зависимой коррекции данных используются в задачах поэлементной классификации гиперспектральных изображений. Указанные разработки позволяют настраивать параметры съёмки и обработки данных с учётом точности классификации, уменьшать влияние шумов и неоднородности отклика регистрирующего устройства и повышать точность поэлементной классификации.

Данные разработки применяются при сборе, подготовке и анализе гиперспектральных данных в рамках проекта «Разработка открытой платформы сбора и анализа гиперспектральных изображений для решения прикладных задач цифровой экономики».

Директор по науке
ООО «Локус», д.ф.-м.н.



Р.В. Скиданов