

АВТОНОМНАЯ НЕКОММЕРЧЕСКАЯ ОБРАЗОВАТЕЛЬНАЯ
ОРГАНИЗАЦИЯ ВЫСШЕГО ОБРАЗОВАНИЯ «СКОЛКОВСКИЙ
ИНСТИТУТ НАУКИ И ТЕХНОЛОГИЙ»

На правах рукописи

Пронина Валерия Сергеевна

**ВОССТАНОВЛЕНИЕ ИЗОБРАЖЕНИЙ С ПОМОЩЬЮ
ОБУЧАЕМЫХ
ОПТИМИЗАЦИОННО-НЕЙРОСЕТЕВЫХ
АЛГОРИТМОВ**

Специальность 1.2.2 —

«Математическое моделирование, численные методы и комплексы программ»

Диссертация на соискание учёной степени
кандидата технических наук

Научный руководитель:
PhD

Дылов Дмитрий Владимирович

Москва — 2022

Оглавление

	Стр.
Введение	6
Глава 1. Деконволюция зашумлённых изображений в микроскопии	14
1.1 Формирование изображений в микроскопии	14
1.2 Общий подход к восстановлению изображений в микроскопии . . .	16
1.3 Восстановление изображений, искажённых пуассоновским шумом	18
1.3.1 Вычисление обучаемых регуляризационных фильтров . . .	20
1.3.2 Вычисление регуляризационных фильтров с помощью нейронной сети	22
1.3.3 Аппроксимация регуляризационного члена с помощью нейронной сети	23
1.3.4 Подготовка набора изображений, искажённых размытием и пуассоновским шумом	24
1.3.5 Результаты восстановления изображений, искажённых размытием и пуассоновским шумом, с помощью алгоритма Ричардсона-Люси с обучаемыми параметрами регуляризации	26
1.4 Обучение моделей для восстановления изображений	29
1.5 Восстановление изображений, искажённых гауссовским шумом . .	29
1.5.1 Вычисление обучаемых регуляризационных фильтров . . .	29
1.5.2 Вычисление регуляризационных фильтров с помощью нейронной сети	31
1.5.3 Вычисление регуляризационных фильтров для каждого пикселя с помощью нейронной сети	32
1.5.4 Аппроксимация регуляризационного члена с помощью нейронной сети	33
1.5.5 Подготовка набора изображений, искажённых размытием и гауссовским шумом	36

1.5.6	Результаты восстановления изображений, искажённых размытием и гауссовским шумом, с помощью фильтра Винера с обучаемыми параметрами регуляризации	37
1.6	Восстановление изображений, искажённых пуассоновским шумом, с помощью фильтра Винера с обучаемыми параметрами регуляризации	41
1.6.1	Подготовка набора изображений, искажённых пуассоновским шумом	42
1.6.2	Результаты восстановления изображений, искажённых размытием и пуассоновским шумом, с помощью фильтра Винера с обучаемыми параметрами регуляризации	44
1.7	Заключение главы	44
Глава 2. Вычисление функции рассеяния точки микроскопа с помощью методов глубокого обучения		47
2.1	Функция рассеяния точки и размытие в микроскопии	47
2.2	Оценка функции рассеяния точки с использованием псевдо-эталонных изображений	49
2.2.1	Применение генеративно-состязательных сетей для создания псевдо-эталонных изображений	50
2.2.2	Оценка функции рассеяния точки	51
2.3	Вычисление функции рассеяния точки и искомого изображения с помощью комбинации оптимизационного алгоритма и тренируемого агента	54
2.3.1	Математическая модель функции рассеяния точки	54
2.3.2	Итеративная схема вычисления функции рассеяния точки и изображения	56
2.3.3	Алгоритмы обучения с подкреплением	58
2.3.4	Применение тренируемого агента для поиска параметров математической модели функции рассеяния точки	60
2.4	Заключение главы	70
Глава 3. Восстановление гиперспектральных изображений, полученных с помощью однопиксельной визуализации		72
3.1	Однопиксельная визуализация	72

3.2	Формирование однопиксельных измерений	75
3.3	Гиперспектральные изображения	76
3.4	Восстановление изображений из однопиксельных измерений . . .	77
3.4.1	Обобщение задачи восстановления для случая гиперспектральных изображений	81
3.4.2	Решение с помощью двумерной свёрточной нейронной сети	81
3.4.3	Решение с помощью трёхмерной свёрточной нейронной сети	82
3.4.4	Вычисление математического ожидания и матриц ковариаций	83
3.5	Архитектуры свёрточных нейронных сетей	85
3.6	Моделирование тренировочных гиперспектральных данных из RGB-изображений	87
3.7	Обучение моделей для восстановления изображений	88
3.8	Экспериментально полученные гиперспектральные однопиксельные измерения	90
3.9	Результаты восстановления гиперспектральных изображений из однопиксельных измерений	92
3.9.1	Результаты для смоделированных однопиксельных измерений	92
3.9.2	Результаты для экспериментальных однопиксельных измерений	93
3.10	Заключение главы	100

Глава 4. Разделение материалов в изображениях спектральной компьютерной томографии с помощью методов глубокого обучения	102
4.1 Спектральная компьютерная томография	102
4.2 Формирование изображений в спектральной компьютерной томографии	105
4.2.1 Прямая модель формирования измерений в спектральной компьютерной томографии	105
4.2.2 Виртуальная модель формирования измерений спектральной компьютерной томографии	106
4.3 Восстановление томографических изображений для отдельных материалов	108

4.3.1	Разделение материалов в области проекций	108
4.3.2	Разделение материалов в области изображений	110
4.4	Моделирование данных спектральной компьютерной томографии	111
4.5	Результаты восстановления томографических изображений для отдельных материалов	113
4.5.1	Метрики оценки результатов разделения материалов . . .	113
4.5.2	Результаты восстановления изображений из небольшого набора данных	113
4.5.3	Результаты восстановления изображений из большого набора данных	116
4.6	Заключение главы	120
Заключение		121
Список сокращений и условных обозначений		123
Список литературы		130
Список рисунков		148
Список таблиц		152
Приложение А. Вывод выражений для алгоритма обратного распространения ошибки		154
A.0.1	Производная по α	155
A.0.2	Производная по \mathbf{g}_d	156
Приложение Б. Пример данных для обучения ESRGAN		159
Приложение В. Дополнительные результаты для фильтра Винера с обучаемыми параметрами регуляризации		160
Приложение Г. Дополнительные результаты реконструкции гиперспектральных изображений из однопиксельных измерений		163

Введение

Восстановление изображений – проблема, играющая ключевую роль во многих приложениях [1; 2]. Для качественного анализа изображений и минимизации возможной ошибки, например, при постановке медицинского диагноза, необходимо, чтобы изображения как можно более точно соответствовали искомому объекту. В настоящее время благодаря развитию технологий машинного обучения и в частности глубокого обучения, решения многих проблем восстановления изображений были получены с использованием свёрточных нейронных сетей. Такие решения представляют собой результат обучения параметров нейросетевых моделей, непосредственно переводящих имеющееся искажённое изображение в ожидаемый результат реконструкции. При наличии модели, обученной на доступных тренировочных данных, использование методов глубокого обучения позволяет получать быстрый результат реконструкции. Тем не менее, основным недостатком такого подхода является отсутствие у модели информации о процессе формирования и искажения результирующего изображения [3]. По этой причине для решения различных задач восстановления изображений модель необходимо переобучить на соответствующих заданию данных.

С другой стороны, классические методы математического моделирования позволяют решать обратные задачи восстановления изображений, сформулированные исходя из физического процесса формирования наблюдений [1; 4]. Для решения задачи необходимо уточнить вид прямой модели формирования измерения. Благодаря этому методы, основанные на таких моделях, могут быть применены к различным видам искажённых данных. Тем не менее классические методы математического моделирования требуют включения дополнительной априорной информации об ожидаемом решении, а также зачастую являются времязатратными [3; 5].

Из этого следует, что актуальным является создание метода, объединяющего в себе классическое решение обратной задачи и тренируемых моделей. Первое призвано учитывать особенности физического процесса формирования изображений, в то время как вторые позволяют получить оценку регуляризационных параметров непосредственно из имеющихся данных.

В различных областях науки и техники визуализация представлена разнообразными приложениями, среди них, например, микроскопия, компьютерная томография, однопиксельная визуализация, цифровая фотография. Для этих приложений процесс получения изображения с использованием визуализирующей системы может быть описан с помощью общей математической формулировки

$$\mathbf{y} = \mathcal{H}(\mathbf{x}), \quad (1)$$

в которой \mathbf{x} соответствует искомому, или истинному, объекту, \mathbf{y} обозначает зарегистрированное измерение или наблюдение, а \mathcal{H} – оператор, осуществляющий отображение из одного векторного пространства в другое [1–3]. В настоящей работе рассмотрены задачи, в которых оператор \mathcal{H} осуществляет линейное преобразование. В различных задачах искомое изображение обычно представлено в виде вектора $\mathbf{x} \in \mathbb{R}^N$. Полученное измерение может представлять собой двух- или трёхмерное изображение такой же размерности, как и искомое изображение \mathbf{x} , и также быть представлено в виде вектора $\mathbf{y} \in \mathbb{R}^N$ – например, в задачах микроскопии. С другой стороны, в задачах однопиксельной визуализации измерение может быть представлено вектором иной размерности – $\mathbf{y} \in \mathbb{R}^P$. В то же время линейный оператор \mathcal{H} обычно представлен матрицей $\mathbf{H} \in \mathbb{R}^{P \times N}$, размер которой так же зависит от приложения. Форма матрицы определяется устройством, осуществляющим формирование изображения, в частности физическим процессом формирования изображения. Например, в случае микроскопии \mathbf{H} соответствует матрице оператора свёртки, представленного функцией рассеяния точки микроскопа (point spread function, PSF) [6; 7]; в случае однопиксельной визуализации \mathbf{H} – матрица бинарных паттернов или матрица Адамара [8]; в случае магнитно-резонансной томографии \mathbf{H} соответствует совокупности преобразования Фурье с последующей субдискретизацией [9].

Процесс формирования изображения обычно называется *прямым процессом*, в то время как задача восстановления искомого изображения является *обратной задачей*. Решение обратной задачи может состоять в оптимизации некоторого критерия, например, в поиске такой оценки $\hat{\mathbf{x}}$, при которой $\mathbf{H}\hat{\mathbf{x}}$ аппроксимирует \mathbf{y} по методу наименьших квадратов [2; 10]. Другими словами, решение должно минимизировать ошибку вида

$$\hat{\mathbf{x}} = \arg \min_{\mathbf{x}} \|\mathbf{H}\mathbf{x} - \mathbf{y}\|_2^2. \quad (2)$$

Обратные задачи восстановления изображений чаще всего являются *некорректно поставленными*, поскольку как минимум одно из свойств корректно поставленной задачи по Адамару [11] нарушено:

1. Решение существует;
2. Решение уникально;
3. Решение устойчиво – решение непрерывно зависит от входных данных, при незначительном изменении входных данных решение также меняется незначительно.

Например, решение, полученное в соответствии с Формулой 2, может не существовать, а в случае, если оно существует, оно неустойчиво к малым изменениям в измерении \mathbf{y} . В работе [12] был предложен метод регуляризации Тихонова, согласно которому задача состоит в поиске решения, минимизирующего функционал следующего вида:

$$\hat{\mathbf{x}} = \arg \min_{\mathbf{x}} \|\mathbf{H}\mathbf{x} - \mathbf{y}\|_2^2 + \gamma\Omega[\mathbf{x}]. \quad (3)$$

Здесь второй член является стабилизирующим функционалом, и часто используется функционал вида $\Omega[\mathbf{x}] = \|\mathbf{x}\|_2^2$. Предложенный метод регуляризации широко используется в различных задачах восстановления изображений [13–16].

С другой стороны, с развитием технологий глубокого обучения решения многих проблем восстановления изображений были получены с использованием свёрточных нейронных сетей. Большое количество алгоритмов, основанных на нейронных сетях, было предложено для решения задач обесшумливания изображений [17; 18], преобразования RGB массива в цветное изображение [19; 20], увеличения разрешения изображений [21–23]. Методы глубокого обучения также нашли применение в задачах устранения размытия изображений или деконволюции. Например, в работе [24] была предложена CNN, осуществляющая восстановление естественных изображений. Авторами работы [25] было предложено применение нейронной сети для решения задачи слепой деконволюции, то есть восстановления изображения без информации о функции размытия. Шаг в сторону объединения традиционных оптимизационных алгоритмов восстановления изображений с методами глубокого обучения был сделан в работе [26] с использованием итеративного алгоритма деконволюции, основанного на быст-

ром преобразовании Фурье (БПФ, fast Fourier transform, FFT). Отличительной особенностью этого алгоритма является то, что регуляризационные фильтры являются обучаемыми параметрами. Аналогичный подход был рассмотрен в работах [27; 28]. В работе [29] авторы предложили получение с помощью нейронной сети обесшумленных вертикальных и горизонтальных градиентов изображения, интегрированных затем в алгоритм деконволюции изображения. В работе [3] было предложено обесшумливание изображения с помощью нейронной сети и затем включение результата в решение обратной задачи восстановления изображения. Наконец, в работе [30] авторами была использована выучиваемая с помощью нейронной сети априорная информация об изображении, описанная сглаженным гауссовой функцией распределением вероятностей для естественных изображений.

В то время как нейросетевые модели, используя доступные тренировочные данные для реконструкции, зачастую не учитывают особенности физического процесса формирования изображения, классические методы, выведенные на основе Формулы 1, располагают приблизительной априорной информацией об ожидаемом решении. В связи с этим настоящая диссертационная работа сфокусирована на преодолении ограничений двух подходов. **Целью** диссертационной работы является разработка методов и алгоритмов математического моделирования для решения обратных задач восстановления изображений. Разработанные схемы должны объединять в себе достоинства классических алгоритмов решения обратных задач и новых алгоритмов глубокого обучения. Также необходимо создание комплекса программ, реализующего разработанные алгоритмы.

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

1. Разработка серии алгоритмов восстановления изображений, полученных с помощью микроскопа с известной функцией рассеяния точки. Разработанные алгоритмы должны включать регуляризацию решения, аппроксимированную с помощью тренируемых моделей.
2. Обучение набора нейронных сетей, осуществляющих решение задачи п.1, и публикация их в открытом доступе.
3. Разработка алгоритмов восстановления изображений, полученных с помощью микроскопа, при отсутствии информации о функции размытия

точки. Вычисление функции размытия точки должно быть осуществлено с помощью современных алгоритмов глубокого обучения.

4. Разработка алгоритма восстановления многоканальных (гиперспектральных) изображений, полученных методом однопиксельной визуализации. Регуляризационный член в решении должен быть аппроксимирован с помощью нейронной сети.
5. Исследование методов разделения сигнала, полученного с помощью спектральной компьютерной томографии, на отдельные для каждого составляющего материала изображения. Разработка решения задачи декомпозиции материалов с помощью нейронных сетей.

Основные положения, выносимые на защиту:

1. Разработан набор из последовательно усложнённых схем аппроксимации регуляризационного члена для решения задачи восстановления изображений, полученных с помощью микроскопа.
2. Показано, что объединение оптимизационного алгоритма восстановления изображений и нейросетевого подхода к аппроксимации регуляризационного члена позволяет получать более высокие метрики реконструкции, чем при использовании указанных подходов по отдельности.
3. Обучены и опубликованы оригинальные нейросетевые модели для восстановления изображений, искажённых гауссовским и пуассоновским шумами.
4. Разработан алгоритм, реализующий итеративную схему совместного поиска искомого изображения, полученного с помощью микроскопа, и функции рассеяния точки микроскопа. Вычисление параметров функции рассеяния точки осуществлено с помощью тренируемых агентов.
5. Разработан алгоритм восстановления гиперспектральных изображений, полученных методом однопиксельной визуализации, в котором нейронная сеть применена для регуляризации решения в пространственной и спектральной областях.
6. Показано, что регуляризация решения в пространственной и спектральной областях в задаче восстановления гиперспектральных изображений приводит к лучшим результатам реконструкции благодаря межканальной корреляции.

7. Исследованы методы разделения сигналов, полученных с помощью спектральной компьютерной томографии. Показано применение нейронной сети для разделения сигналов как в области проекций, так и в области изображений.

Научная новизна заключается в следующем:

1. Разработан новый метод аппроксимации регуляризационного члена в решении задачи восстановления изображений, полученных с помощью микроскопа, с использованием тренируемых моделей. В том числе разработан новый алгоритм аппроксимации индивидуального для каждого пикселя регуляризационного члена.
2. Предложены и опубликованы оригинальные нейросетевые модели для восстановления изображений на основе комбинации классических оптимизационных схем (фильтр Винера, алгоритм Ричардсона-Люси) и тренируемых моделей аппроксимации регуляризационного члена.
3. Разработан новый метод восстановления многоканальных изображений, полученных с помощью однопиксельной камеры, с использованием композиции нейронных сетей. В основе метода лежит комбинация классического решения обратной задачи и тренируемой модели регуляризации решения в пространственной и спектральной областях.
4. Разработан новый метод восстановления функции рассеяния точки микроскопа путём поиска оптимальных параметров математической модели функции с помощью алгоритма обучения с подкреплением.

Практическая ценность результатов диссертации заключается в том, что в работе рассмотрены задачи восстановления изображений, актуальные для различных областей науки и техники. В том числе полученные автором результаты имеют научно-практическую ценность для задач анализа медико-биологических изображений. Некоторые предложенные в работе методы и модели опубликованы в открытом доступе.

Обоснованность и достоверность результатов и выводов. Выводы диссертации обоснованы аналитическими расчётами и подтверждены данными воспроизводимых экспериментов на основе разработанных алгоритмов. Теоретическую и методологическую основу проведённых исследований составили труды отечественных и зарубежных авторов в области восстановления изображений. Для анализа полученных результатов применены показатели ка-

чества, традиционно используемые для оценки решения задач восстановления изображений. Результаты, полученные в процессе подготовки диссертации, были представлены на международных научных конференциях и опубликованы в рецензируемых научных изданиях.

Основные результаты работы докладывались на следующих научных конференциях и семинарах:

- Applied Inverse Problems Conference, доклад “Nonlinear material decomposition in spectral CT using deep learning” JFPJ Abascal, 8-12 июля 2019.
- 16th European Conference on Computer Vision (ECCV), постер “Microscopy Image Restoration with Deep Wiener-Kolmogorov Filters”, Пронина Валерия, 23-28 августа 2020.
- 2020 IEEE 17th International Symposium on Biomedical Imaging (ISBI), доклад “Material decomposition in spectral CT using deep learning”, JFPJ Abascal, 3-7 апреля 2020.
- Научный семинар в Институте Вычислительной Математики им. Г.И. Марчука РАН, доклад “Восстановление изображений с помощью обучаемых алгоритмов реконструкции”, Пронина Валерия, 15 февраля 2022.
- Научный семинар в Институте Автоматики и Электростроения Сибирского отделения РАН, доклад “Восстановление изображений с помощью обучаемых алгоритмов реконструкции”, Пронина Валерия, 4 апреля 2022.

Содержание диссертации соответствует паспорту специальности 1.2.2 – Математическое моделирование, численные методы и комплексы программ, в частности, пунктам: п.1 – Разработка новых математических методов моделирования объектов и явлений; п.3 – Разработка, обоснование и тестирование эффективных вычислительных методов с применением современных компьютерных технологий; п.4 – Реализация эффективных численных методов и алгоритмов в виде комплексов проблемно-ориентированных программ для проведения вычислительного эксперимента.

Публикации. Основные результаты по теме диссертации изложены в 4 печатных изданиях, индексируемых Web of Science и Scopus.

Личный вклад. соискателя в работах с соавторами заключается в следующем: В работе [31] автором были разработаны методы обучения регуляри-

зационных фильтров, их аппроксимации с помощью нейронной сети, а также метод моделирования регуляризационного члена целиком с помощью нейронной сети и алгоритм получения решения. Подготовка данных была осуществлена автором. Экспериментальные результаты были получены автором самостоятельно. В работе [32] автором был разработан метод, обобщающий решение задачи восстановления изображений, полученных с помощью однопиксельной камеры, для случая многоканальных изображений. Также автором был разработан алгоритм генерирования многоканальных изображений из RGB-изображений, подготовлен тренировочный набор данных, обучен набор нейронных сетей и получены результаты экспериментов. В работах [33; 34] автором был разработан алгоритм разделения изображений спектральной КТ на изображения составных материалов с использованием нейронной сети в области проекций. Автор участвовала в сборе и подготовке тренировочных данных, а также в проведении первой серии экспериментов.

Глава 1. Деконволюция зашумлённых изображений в микроскопии

1.1 Формирование изображений в микроскопии

Микроскопия является методом визуализации, имеющим большое значение в биологических и медицинских исследованиях и позволяющим анализировать как органические, так и неорганические вещества и фундаментальные биологические явления в клеточном масштабе. В процессе формирования наблюдения искомое изображение объекта подвергается двум основным искажениям: размытию, вызванному наличием предела разрешающей способности микроскопа, и шуму, вносимому несовершенством системы формирования изображения [35–37]. Оба этих процесса влияют на качество полученного изображения и, как следствие, на восприятие его деталей и достоверность последующего анализа.

С математической точки зрения процесс формирования изображения в микроскопии следует общей формулировке, представленной во введении, и может быть описан моделью

$$\mathbf{y} = \mathbf{H}\mathbf{x}, \quad (1.1)$$

в которой $\mathbf{x} \in \mathbb{R}^N$ – искомое изображение объекта, $\mathbf{H} \in \mathbb{R}^{N \times N}$ – матрица оператора свёртки – функции рассеяния точки микроскопа (point spread function, PSF). Полученное изображение $\mathbf{y} \in \mathbb{R}^N$, так же, как и искомое изображение \mathbf{x} , представляет собой двумерную матрицу, но для удобства осуществления математических операций их принято представлять в виде векторов размерности N .

Популярным инструментом визуализации клеток и тканей является флуоресцентная микроскопия. С помощью этого метода возможно исследование как фиксированных, так и живых биологических образцов. Принцип действия флуоресцентной микроскопии основан на явлении флуоресценции – излучении поглощённого света электронно-возбуждённой частицей при переходе в основное состояние [38]. По причине несовершенства оптической системы в целом и датчика регистрации изображения в частности получаемое изображение подвергается зашумлению [37]. Флуктуации в электрической сети вносят аддитивный шум, который может быть описан с помощью распределения Гаусса [37; 39].

Шум, полученный вследствие процесса испускания фотонов, может быть описан с помощью распределения Пуассона [37]. Чем ниже уровень освещённости и, следовательно, меньше количество фотонов, попадающих на образец, тем сильнее пуассоновский шум [13]. В настоящей работе рассмотрены пуассоновская и гауссовская статистики шума, искажающего полученное изображение.

В то время как зачастую в задачах деконволюции PSF визуализирующей системы также является неизвестной функцией, которую необходимо найти для последующего восстановления искомого изображения, в задачах микроскопии точное измерение PSF является более простой задачей, поскольку её возможно аппроксимировать с помощью наночастиц [40]. По этой причине в данном разделе работы проблема вычисления PSF не рассмотрена и задача восстановления изображений решена исходя из предположения о наличии известной функции размытия точки.

Свёртка как матрично-векторное произведение. В описании математической модели, выраженной Формулой 1.1, указано, что изображение объекта подвергается свёртке с функцией рассеяния точки. Функция рассеяния точки обычно представлена вектором $\mathbf{h} \in \mathbb{R}^{K^2}$, и процедура свёртки обозначена как $\mathbf{h} * \mathbf{x}$. В Формуле 1.1 свёртка выражена через матрично-векторное произведение матрицы \mathbf{H} и вектора \mathbf{x} . Для преобразования вектора \mathbf{h} в матрицу \mathbf{H} удобнее выразить вектор в двумерном виде, то есть $h_{k_x, k_y}, 1 \leq k_x \leq K, 1 \leq k_y \leq K$. Используя обычное для обработки изображений предположение, что функция рассеяния точки \mathbf{h} периодична, матрица, соответствующая ей, имеет вид блочной матрицы [6]

$$\mathbf{H} = \begin{bmatrix} \mathbf{H}_0 & \mathbf{H}_{N-1} & \mathbf{H}_{N-2} & \cdots & \mathbf{H}_1 \\ \mathbf{H}_1 & \mathbf{H}_0 & \mathbf{H}_{N-1} & \cdots & \mathbf{H}_2 \\ \mathbf{H}_2 & \mathbf{H}_1 & \mathbf{H}_0 & \cdots & \mathbf{H}_3 \\ \vdots & & & \ddots & \vdots \\ \mathbf{H}_{N-1} & \mathbf{H}_{N-2} & \cdots & \mathbf{H}_1 & \mathbf{H}_0. \end{bmatrix} \quad (1.2)$$

Здесь каждый блок \mathbf{H}_p также является матрицей,

$$\mathbf{H}_p = \begin{bmatrix} h_{0,p} & h_{N-1,p} & h_{N-2,p} & \cdots & h_{1,p} \\ h_{1,p} & h_{0,p} & h_{N-1,p} & \cdots & h_{2,p} \\ h_{2,p} & h_{1,p} & h_{0,p} & \cdots & h_{3,p} \\ \vdots & & & \ddots & \vdots \\ h_{N-1,p} & h_{N-2,p} & \cdots & h_{1,p} & h_{0,p} \end{bmatrix} \quad (1.3)$$

Первый столбец матрицы \mathbf{H}_p соответствует p -му столбцу двумерного дополненного нулями представления вектора \mathbf{h}_e , определённого как

$$h_{e k_x, k_y} = \begin{cases} h_{k_x, k_y} & 1 \leq k_x \leq K, 1 \leq k_y \leq K \\ 0 & K + 1 \leq k_x \leq \sqrt{N}, K + 1 \leq k_y \leq \sqrt{N} \end{cases} \quad (1.4)$$

Матрица \mathbf{H} является циркулянтной матрицей, то есть частным случаем матрицы Тёплица [41], в которой каждая строка является сдвинутой на один элемент вправо версией предыдущей строки. Представление свёртки как матрично-векторного произведения в силу его удобства чаще всего используется при записи математических формулировок. Здесь и далее, если не указано иное, для удобства принято обозначение “ \mathbf{H} – функция рассеяния точки”.

1.2 Общий подход к восстановлению изображений в микроскопии

Вследствие воздействия функции размытия точки визуализирующей системы и наличия шума, результирующие изображения, полученные с помощью микроскопа, получаются размытыми и зашумлёнными, что влияет на качество их последующего анализа. Задача восстановления таких изображений может состоять в поиске такой оценки $\hat{\mathbf{x}}$, при которой $\mathbf{H}\hat{\mathbf{x}}$ аппроксимирует \mathbf{y} по методу наименьших квадратов [6; 10]. Решение в таком случае должно минимизировать ошибку вида

$$\hat{\mathbf{x}} = \arg \min_{\mathbf{x}} \|\mathbf{H}\mathbf{x} - \mathbf{y}\|_2^2. \quad (1.5)$$

При обращении производной выпуклой функции в некоторой точке в нуль, функция в этой точке достигает своего минимума [42]. Из этого следует, что

производная функции 1.5 и, соответственно, решение имеют вид

$$\begin{aligned}\mathbf{H}^\top (\mathbf{H}\mathbf{x} - \mathbf{y}) &= 0 \\ \hat{\mathbf{x}} &= (\mathbf{H}^\top \mathbf{H})^{-1} \mathbf{H}^\top \mathbf{y}.\end{aligned}\tag{1.6}$$

Если матрица \mathbf{H} является квадратной и невырожденной, то решение в Формуле 1.6 может быть выражено как $\hat{\mathbf{x}} = \mathbf{H}^{-1}\mathbf{y}$ [6]. Решение также может быть получено с помощью преобразования Фурье [12]

$$\mathcal{F}(\hat{\mathbf{x}}) = \frac{\mathcal{F}(\mathbf{y})}{\mathcal{F}(\mathbf{H})}.\tag{1.7}$$

Если в системе присутствует шум, искажающий изображение, решение 1.7 будет иметь вид

$$\mathcal{F}(\hat{\mathbf{x}}) = \frac{\mathcal{F}(\mathbf{y})}{\mathcal{F}(\mathbf{H})} + \frac{\mathcal{F}(\mathbf{n})}{\mathcal{F}(\mathbf{H})} = \mathcal{F}(\mathbf{x}) + \frac{\mathcal{F}(\mathbf{n})}{\mathcal{F}(\mathbf{H})}.\tag{1.8}$$

В работе [12] показано, что решение уравнения 1.8 с применением обратного преобразования Фурье может не существовать, а в случае, если оно существует, оно является неустойчивым к малым отклонениям \mathbf{n} . Задача восстановления изображения в таком случае не удовлетворяет критериям корректно поставленной задачи по Адамару [11]. Для решения некорректно поставленных обратных задач авторами [12] был предложен метод регуляризации Тихонова, заключающийся в минимизации функционала вида

$$L(\mathbf{x}) = \frac{1}{2} \|\mathbf{y} - \mathbf{H}\mathbf{x}\|_2^2 + \gamma \Omega[\mathbf{x}].\tag{1.9}$$

В Формуле 1.9 $\Omega[\mathbf{x}]$ является стабилизирующим решение функционалом, $\Omega[\mathbf{x}] = \|\mathbf{x}\|_2^2$, а коэффициент γ – регуляризационным коэффициентом [12; 43; 44]. Второй член в Формуле 1.9 является регуляризационным членом.

С развитием методов глубокого обучения многие обратные задачи восстановления изображений были решены с применением свёрточных нейронных сетей [17–25]. Свёрточные нейронные сети в частности позволяют преобразовывать искажённое изображение в исправленную версию путём сравнения полученного результата с известным эталонным изображением. Ряд работ применяет методы глубокого обучения для регуляризации в широком диапазоне задач восстановления изображений [3; 20; 29; 45–48]. Такая парадигма регуля-

ризации, в свою очередь, позволила качественно превзойти методы, осуществляющие непосредственно отображение искажённого изображения в восстановленное; однако ни один из этих методов не был протестирован на проблемах устранения размытия в микроскопии.

В данной главе описан метод восстановления изображений, полученных с помощью микроскопа и искажённых размытием и шумом. Согласно этому методу, априорная, или регуляризирующая, информация об ожидаемом решении получена непосредственно из доступных тренировочных данных контролируемым образом с включением стратегий глубокого обучения. Предложен набор техник аппроксимации регуляризационного члена или его частей с помощью тренируемых моделей, что в сочетании с классической оптимизацией позволяет находить компромисс между вычислительной эффективностью и точностью реконструкции изображений.

1.3 Восстановление изображений, искажённых пуассоновским шумом

Распространённым методом решения задачи восстановления изображения в микроскопии является алгоритм Ричардсона-Люси [13; 49; 50]. Этот метод основан на максимизации функции правдоподобия и сформирован из предположения, что изображение искажено размытием и пуассоновским шумом [13]. Плотность вероятности наблюдения y_n в n -м пикселе при регистрации изображения объекта \mathbf{x} с помощью оптической системы с функцией рассеяния точки в матричной форме \mathbf{H} выражена как

$$p(y_n | \mathbf{x}) = \frac{(\mathbf{H}\mathbf{x})_n^{y_n} e^{-(\mathbf{H}\mathbf{x})_n}}{y_n!}. \quad (1.10)$$

Таким образом, плотность вероятности наблюдения \mathbf{y} при визуализации объекта \mathbf{x} выражена как

$$p(\mathbf{y} | \mathbf{x}) = \prod_{n=1}^N \left[e^{y_n \ln((\mathbf{H}\mathbf{x})_n)} \times e^{-(\mathbf{H}\mathbf{x})_n} \times \frac{1}{y_n!} \right]. \quad (1.11)$$

Используя свойство логарифмической функции монотонно возрастать на всей области определения, гарантирующее, что логарифм функции достигнет максимума в той же точке, что и сама функция, а также пользуясь тем, что логарифм произведения равен сумме логарифмов, максимум функции можно выразить как $\max(\ln(p(\mathbf{y}|\mathbf{x})))$ или $\min(-\ln(p(\mathbf{y}|\mathbf{x})))$:

$$\begin{aligned}\ln p(\mathbf{y}|\mathbf{x}) &= \sum_{n=1}^N y_n \ln((\mathbf{H}\mathbf{x})_n) - \sum_{n=1}^N (\mathbf{H}\mathbf{x})_n - \sum_{n=1}^N \ln(y_n!) \\ &= (\ln(\mathbf{H}\mathbf{x}))^\top \mathbf{y} - (\mathbf{H}\mathbf{x})^\top \mathbf{1} - \sum_{n=1}^N \ln(y_n!).\end{aligned}\tag{1.12}$$

Так как $\sum_{n=1}^N \ln(y_n!)$ является постоянной величиной относительно \mathbf{x} , в дальнейших вычислениях ею можно пренебречь. Таким образом, целевая функция, которую необходимо минимизировать, имеет вид

$$L(\mathbf{x}) = -(\ln(\mathbf{H}\mathbf{x}))^\top \mathbf{y} + (\mathbf{H}\mathbf{x})^\top \mathbf{1}.\tag{1.13}$$

Минимизация Формулы 1.13 приводит к выражению вида

$$\mathbf{H}^\top \mathbf{1} - \mathbf{H}^\top \text{diag} \frac{\mathbf{1}}{\mathbf{H}\mathbf{x}} \mathbf{y} = 0.\tag{1.14}$$

Здесь символ $\text{diag}(\cdot)$ обозначает матрицу, на главной диагонали которой находятся элементы $\frac{\mathbf{1}}{\mathbf{H}\mathbf{x}}$, $\mathbf{1}$ – вектор такого же размера, как и $\mathbf{H}\mathbf{x}$, состоящий из единиц. Процедура деления осуществлена поэлементно. Далее необходимо учесть два важных условия. Первое заключается в том, что суммарное значение всех пикселей PSF должно быть равно единице, из чего следует, что $\mathbf{H}^\top \mathbf{1} = \mathbf{1}$ [13]. Второе условие состоит в том, что при достижении оптимального значения $\hat{\mathbf{x}}$ дальнейшие вычисления не будут изменять это значение, то есть на i -й итерации алгоритма, $1 \leq i \leq I$, $x_{n,i+1}/x_{n,i} = 1$. Всё это приводит к формулировке итеративной схемы Ричардсона-Люси в виде

$$\hat{\mathbf{x}}_{i+1} = \mathbf{H}^\top \text{diag} \frac{\mathbf{1}}{\mathbf{H}\mathbf{x}_i} \mathbf{y} \odot \mathbf{x}_i.\tag{1.15}$$

Оператор \odot обозначает поэлементное перемножение векторов.

1.3.1 Вычисление обучаемых регуляризационных фильтров

Рассматривая задачу восстановления изображения, искажённого размытием и пуассоновским шумом, с точки зрения теоремы Байеса, согласно которой $p(\mathbf{x} | \mathbf{y}) = \frac{p(\mathbf{y}|\mathbf{x})p(\mathbf{x})}{p(\mathbf{y})}$, можно сформулировать выражение для оценки апостериорного максимума как

$$\hat{\mathbf{x}} = \arg \max_{\mathbf{x}} [p(\mathbf{y} | \mathbf{x}) p(\mathbf{x})], \quad (1.16)$$

что эквивалентно минимизации отрицательного логарифма:

$$\hat{\mathbf{x}} = \arg \min_{\mathbf{x}} [-\ln p(\mathbf{y} | \mathbf{x}) - \ln p(\mathbf{x})]. \quad (1.17)$$

В Формуле 1.16 первый множитель в правой части представляет собой плотность вероятности, описывающую полученное при визуализации объекта \mathbf{x} изображение \mathbf{y} , а второй множитель – плотность вероятности, описывающую априорную информацию об объекте или искомом изображении \mathbf{x} . При использовании $p(\mathbf{x}) = e^{-\gamma \sum_{d=1}^D \|\mathbf{G}_d \mathbf{x}\|_2^2}$ [13] целевая функция в Формуле 1.13 преобразуется в

$$\begin{aligned} L(\mathbf{x}) &= -(\ln \mathbf{H}\mathbf{x})^\top \mathbf{y} + (\mathbf{H}\mathbf{x})^\top \mathbf{1} + e^\alpha r(\mathbf{x}) = \\ &= -(\ln \mathbf{H}\mathbf{x})^\top \mathbf{y} + (\mathbf{H}\mathbf{x})^\top \mathbf{1} + \gamma \sum_{d=1}^D \|\mathbf{G}_d \mathbf{x}\|_2^2. \end{aligned} \quad (1.18)$$

Здесь последний член выполняет роль регуляризационного или стабилизирующего члена по методу Тихонова, $r(\mathbf{x}) = \sum_{d=1}^D \|\mathbf{G}_d \mathbf{x}\|_2^2$. Решение для Формулы 1.18 имеет вид

$$\hat{\mathbf{x}}_{i+1} = \left[\mathbf{H}^\top \operatorname{diag} \frac{\mathbf{1}}{\mathbf{H}\hat{\mathbf{x}}_i} \mathbf{y} \right] \operatorname{diag} \frac{\mathbf{1}}{\mathbf{1} + e^\alpha \sum_{d=1}^D \mathbf{G}_d^\top \mathbf{G}_d \hat{\mathbf{x}}_i} \odot \hat{\mathbf{x}}_i. \quad (1.19)$$

В Формуле 1.19 матрицы $\mathbf{G}_d \in \mathbb{R}^{N \times N}$ могут быть выражены с помощью свёрточных фильтров $\mathbf{g}_d \in \mathbb{R}^{L^2}$. Процедура деления осуществляется поэлементно. Отметим, что операции свёртки $\mathbf{g}_d * \mathbf{x}$ и $\mathbf{h} * \mathbf{x}$ здесь и далее записаны с помощью матрично-векторного произведения.

На первом этапе работы для реализации решения согласно Формуле 1.19 применён набор из D регуляризационных фильтров \mathbf{g}_d размерностью L^2 или $L \times L$ в двумерной формулировке векторов. Для реализации предлагаемой схемы, в которой априорная информация об ожидаемом решении получена из имеющихся искажённых данных, фильтры \mathbf{g}_d заданы как обучаемые параметры. Это означает, что их значения обновляются в процессе тренировки алгоритма. Инициализация фильтров выполнена с помощью дискретного косинусного преобразования. Этот метод инициализации широко применяем при обработке изображений, поскольку с его помощью возможно извлекать полезную априорную информацию об изображении [51]. В этом и последующих экспериментах принято $\gamma = e^\alpha$, и α также является обучаемым параметром. Вычисление регуляризационного коэффициента и фильтров происходит с помощью набора пар искажённых размытием и шумом и эталонных изображений путём сопоставления эталонного изображения с восстановленным и вычисления разницы между двумя изображениями. Задача тренируемого алгоритма состоит в минимизации величины этой разницы, что достигается путём поиска оптимальных значений регуляризационных фильтров и коэффициента с помощью алгоритма градиентного спуска [52]. Таким образом, процесс вычисления значений регуляризационных фильтров и коэффициента выглядит следующим образом: D регуляризационных фильтров размера $L \times L$ инициализированы с помощью дискретного косинусного преобразования, а коэффициент α равен нулю; искажённое изображение \mathbf{y} поступает на вход итеративного алгоритма Ричардсона-Люси 1.19; после I итераций алгоритма получена первичная оценка решения $\hat{\mathbf{x}}$; с помощью целевой функции или так называемой функции потерь производится сравнение значений $\hat{\mathbf{x}}$ с эталонным изображением \mathbf{x} . Для того, чтобы на следующем этапе получить меньшее значение разницы между восстановленным и эталонным изображениями, производится вычисление градиента функции потерь и необходимых для её минимизации градиентов обучаемых параметров, в данном случае регуляризационных фильтров и коэффициента α . Затем значения фильтров и коэффициента α с помощью алгоритма обратного распространения ошибки обновляются на вычисленную величину градиента. С помощью обновлённых значений параметров происходит вычисление новой оценки решения $\hat{\mathbf{x}}$. Этот цикл повторяется до тех пор, пока значение функции потерь или значения обучаемых параметров не перестанут значительно изменяться. Описанная

схема вычисления значений обучаемых параметров алгоритма является общей для всех описанных в Разделе 1.3 алгоритмов с некоторыми изменениями, которые далее указаны явно. Все варианты алгоритма Ричардсона-Люси содержат $I = 3$ итерации. Здесь и далее, если не указано иное, $D = 8$ и $L = 3$.

Описанный метод, реализующий итеративный алгоритм Ричардсона-Люси для восстановления изображений, искажённых пуассоновским шумом, и использующий регуляризацию с применением обучаемых фильтров, назван TRL-K (**T**rainable **R**ichardson-**L**ucy with learnable **K**ernels), и его схема представлена на рисунке 1.1а.

1.3.2 Вычисление регуляризационных фильтров с помощью нейронной сети

Основным недостатком предыдущей формулировки решения с помощью алгоритма Ричардсона-Люси является то, что регуляризационные фильтры являются одинаковыми для всех изображений. Поэтому при обработке не учитываются различия в искомым изображениях. В связи с этим предложено расширить формулировку схемы Ричардсона-Люси с использованием нейронной сети, предсказывающей регуляризационные фильтры для каждого отдельного изображения. В предложенном методе решение также выражено согласно Формуле 1.19, но в отличие от предыдущего метода TRL-K, регуляризационные фильтры \mathbf{g}_d сформированы с помощью нейронной сети, предсказывающей фильтры (Kernel Prediction Network, KPN [18]). Предложенный метод представлен на рисунке 1.1б и обозначен TRL-KPN (**T**rainable **R**ichardson **L**ucy with **K**ernel **P**rediction Network).

Для Kernel Prediction Network выбрана компактная архитектура UNet [53], содержащая около 470,000 обучаемых параметров (весов сети), на вход которой поступает набор (batch) из искажённых одноканальных изображений. На выходе сети формируется матрица такого же пространственного размера, но содержащая $L^2 D$ каналов или карт признаков. Для этого на выходе последнего свёрточного слоя производится выборка максимального значения для каждого канала, в результате чего формируется вектор размером $L^2 D \times 1$. Далее полу-

ченный вектор подвергается преобразованию в D регуляризационных фильтров размером $L \times L$. Таким образом, для каждого отдельного изображения KPN предсказывает D регуляризационных фильтров \mathbf{g}_d размером $L \times L$.

Архитектура сети UNet. Архитектура описываемой в данной главе 2D UNet состоит из сужающегося пути и расширяющегося пути. Начиная с первого свёрточного слоя с 12 каналами или картами признаков с каждой процедурой подвыборки с определением максимального значения, это число увеличивается в 2 раза, достигая значения 96 на последнем свёрточном слое сужающегося пути. Размер карт признаков, соответствующих входному изображению на первом слое сети, уменьшается. На расширяющемся пути количество каналов соответственно уменьшается. Этапы расширяющегося пути 2D UNet состоят из билинейной повышающей подвыборки карт признаков, объединения с картами признаков из сужающегося пути и двух свёрточных слоев. Каждый свёрточный слой, кроме последнего, использует свёрточные фильтры размером 3×3 , за которыми следует этап нормализации, которой может быть различным в зависимости от алгоритма, и слой ReLU. Последний слой, производящий выходные данные сети, состоит из одного свёрточного слоя с размером фильтра 1×1 . В описываемом в данном разделе алгоритме в архитектуре UNet применён метод instance normalization, или нормализация значений каждого экземпляра в наборе.

Процесс вычисления решения $\hat{\mathbf{x}}$ соответствует процессу, описанному в Разделе 1.3.1, за исключением того факта, что в процессе минимизации разности между восстановленным изображением $\hat{\mathbf{x}}$ и эталонным изображением \mathbf{x} происходит обновление значений не самих регуляризационных фильтров \mathbf{g}_d , а весов сети KPN, предсказывающей эти фильтры.

1.3.3 Аппроксимация регуляризационного члена с помощью нейронной сети

Наконец, предложен метод, идея которого состоит в аппроксимации нейронной сетью всего регуляризационного члена $r(\mathbf{x})$. В таком случае целевая функция также соответствует Формуле 1.18. Но при такой формулировке невоз-

можно сделать предположение о форме регуляризационного члена. Минимизация Формулы 1.18 приводит к выражению вида

$$\mathbf{H}^\top \mathbf{1} - \mathbf{H}^\top \operatorname{diag} \frac{\mathbf{1}}{\mathbf{H}\mathbf{x}} \mathbf{y} + e^\alpha \nabla r(\mathbf{x}) = 0. \quad (1.20)$$

Для получения решения использован упомянутый в Разделе 1.2 подход к формированию итеративной схемы Ричардсона-Люси. Ключевым отличием от предыдущих алгоритмов TRL-K и TRL-KPN является то, что вместо D регуляризационных фильтров размера $L \times L$ в формулировке решения фигурирует градиент регуляризационного члена $\nabla r(\mathbf{x}) = \mathcal{G}(\mathbf{x})$. Таким образом, решение имеет вид

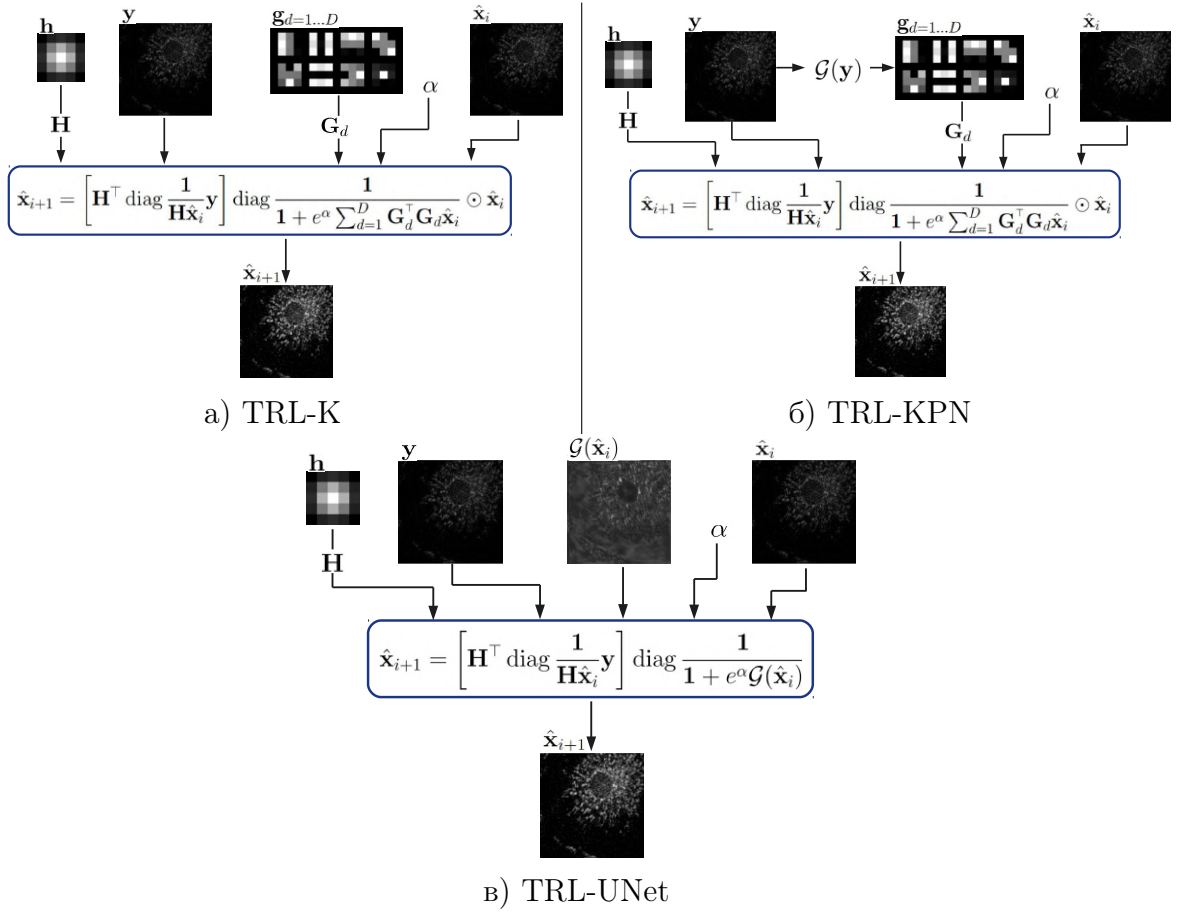
$$\hat{\mathbf{x}}_{i+1} = \left[\mathbf{H}^\top \operatorname{diag} \frac{\mathbf{1}}{\mathbf{H}\hat{\mathbf{x}}_i} \mathbf{y} \right] \operatorname{diag} \frac{\mathbf{1}}{\mathbf{1} + e^\alpha \mathcal{G}(\hat{\mathbf{x}}_i)}. \quad (1.21)$$

Как и в предыдущей схеме TRL-KPN, для аппроксимации регуляризационного члена, а именно его градиента, использована сеть UNet, но в данном алгоритме последний свёрточный слой сети не подвергается дополнительным модификациям – на вход сети поступает полученная на каждой итерации алгоритма Ричардсона-Люси оценка восстановленного изображения $\hat{\mathbf{x}}_i$, а на выходе формируется градиент регуляризационного члена $r(\hat{\mathbf{x}}_i)$. Этот градиент имеет такие же размеры и количество каналов, что и входное изображение. Также после каждого свёрточного слоя нейронной сети применена нормализация значений каждого экземпляра в наборе.

Описанный метод, реализующий итеративный алгоритм Ричардсона-Люси для восстановления изображений, искажённых пуассоновским шумом, и использующий аппроксимацию регуляризационного члена с помощью нейронной сети, назван TRL-UNet (**T**rainable **R**ichardson-**L**ucy with **U**Net), и его схема представлена на Рисунке 1.1.

1.3.4 Подготовка набора изображений, искажённых размытием и пуассоновским шумом

На основе существующих наборов реальных снимков флуоресцентной микроскопии, использованных в работах [54; 55], создан датасет из пар эталонных



На Рисунке 1.1а представлен алгоритм Ричардсона-Люси с обучаемыми регуляризационными фильтрами, на Рисунке 1.1б изображён алгоритм Ричардсона-Люси с регуляризационными фильтрами, вычисленными для каждого изображения с помощью нейросети. На Рисунке 1.1в отображена схема, в которой градиент регуляризационного члена для каждого изображения вычислен с помощью нейронной сети.

Рисунок 1.1 — Схемы трёх алгоритмов на основе схемы Ричардсона-Люси, реализующих аппроксимацию регуляризационного члена тренируемыми моделями.

и искажённых размытием и пуассоновским шумом изображений. Эталонные изображения из обоих датасетов поделены на изображения размером 256×256 . Таким образом создано 1405 эталонных (ground truth) изображений, разделённых далее на 975 тренировочных, 200 валидационных и 230 тестовых.

Для того, чтобы получить искажённые размытием и шумом изображения, с помощью плагинов Diffraction PSF 3D [56] и PSF Generator [57], реализованных в программе Fiji [58], создано 35 двумерных функций размытия точки размерами 5×5 . Из полученных PSF 25 использованы для создания тренировочных

изображений, 5 для создания данных для валидации и оставшиеся 5 для генерации изображений для тестирования. Все PSF нормированы таким образом, что сумма всех элементов равна единице.

Для генерирования изображений с различными уровнями пуассоновского шума использован распространённый подход [59; 60], согласно которому эталонные изображения нормализованы таким образом, чтобы максимальное значение в изображении соответствовало одному из значений в диапазоне (5, 25, 100, 255). Интенсивность пуассоновского шума определяется как $\text{SNR} = \sqrt{y_n}$, где y_n – значение изображения в пикселе n , а SNR – отношение сигнала к шуму, signal-to-noise ratio. Таким образом, с увеличением интенсивности изображения уменьшается уровень пуассоновского шума и наоборот. При использовании описанного подхода к нормализации возможно получение изображений с широким диапазоном уровней пуассоновских шумов.

Во время обучения модели эталонное изображение размером 256×256 подвергается нормализации к случайно выбранному из упомянутого выше диапазона значению. Затем изображение подвергается свёртке со случайно выбранной PSF из 25 тренировочных, после чего из размытого изображения генерируется зашумлённое. Тестовый набор создан из 230 зарезервированных для этих целей изображений, при этом для каждого изображения сгенерировано 4 искажённых – по одному для каждого уровня пуассоновского шума, в результате чего получается 920 размытых и зашумлённых изображений.

1.3.5 Результаты восстановления изображений, искажённых размытием и пуассоновским шумом, с помощью алгоритма Ричардсона-Люси с обучаемыми параметрами регуляризации

Для оценки результатов, полученных с помощью предложенных схем реконструкции, и сравнения их с классической схемой, изображения, искажённые четырьмя уровнями пуассоновского шума, также восстановлены с помощью алгоритма Ричардсона-Люси, описанного в работе [13]. В этом алгоритме априорная информация об искомом изображении описана с помощью функционала, заданного как ℓ_1 -норма градиента изображения [61]. Этот алгоритм назван авто-

рами схемой Ричардсона-Люси с полной вариацией (Richardson-Lucy with total variation) и далее обозначен как RL+TV. В качестве алгоритма, выполняющего реконструкцию изображения исключительно с помощью нейронной сети, выбрана сеть UNet. Архитектура этой сети совпадает с архитектурой, использованной в алгоритме TRL-UNet, за исключением того, что применена нормализация по всему тренировочному набору (batch normalization) вместо нормализации значений каждого экземпляра в наборе (instance normalization).

Для оценки результатов восстановления изображений с помощью всех представленных в настоящей главе методов использованы стандартные для задач обработки изображений метрики, а именно отношение пикового сигнала к шуму (peak signal-to-noise ratio, PSNR) и индекс структурного сходства (structural similarity index measure, SSIM).

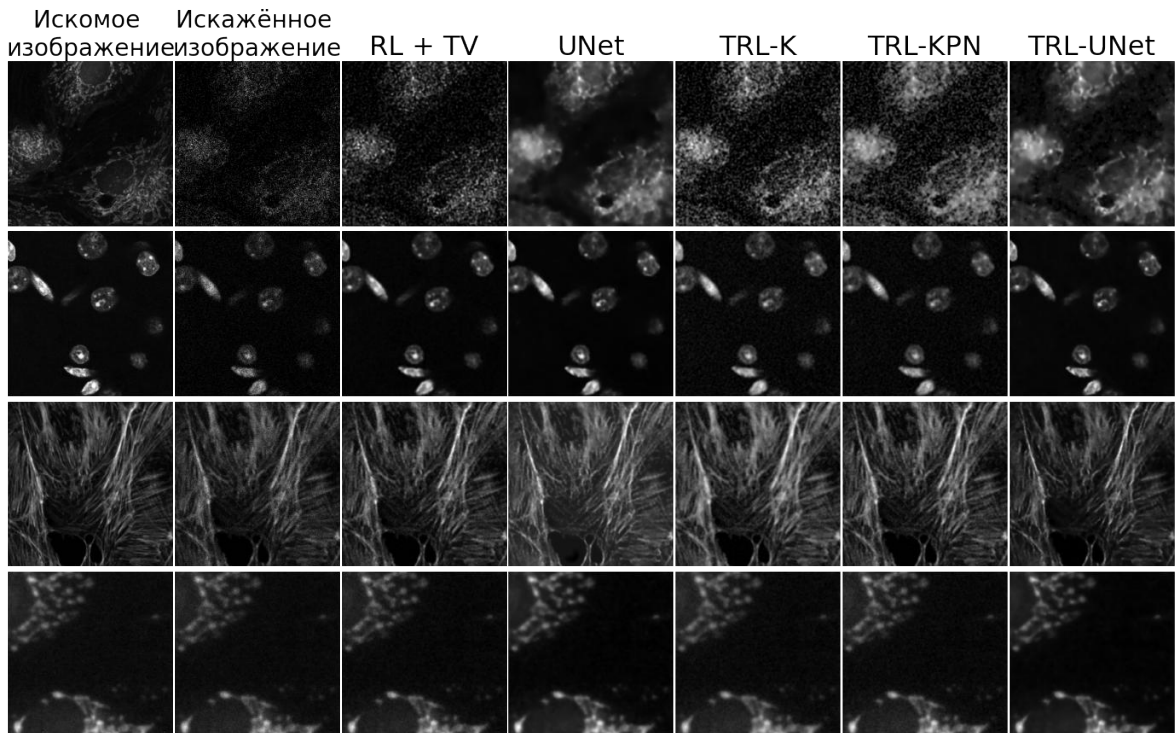
На Рисунке 1.2 представлены результаты восстановления изображений, искажённых размытием и пуассоновским шумом. Из представленных результатов видно, что UNet и TRL-UNet достаточно точно восстанавливают детали изображений даже в случае сильного пуассоновского шума, в отличие от классической схемы Ричардсона-Люси и схем с обучаемыми регуляризационными фильтрами. В случае более слабого шума разница между двумя алгоритмами на базе UNet практически незаметна с визуальной точки зрения. Алгоритм TRL-K в целом хуже справляется с устранением размытия в изображениях. Результаты, представленные в Таблице 1, показывают, что схема, объединяющая алгоритм Ричардсона-Люси и аппроксимацию градиента регуляризационного члена с помощью нейронной сети, превосходит в качестве восстановления изображений все остальные методы, за исключением нейронной сети UNet на самом сильном уровне шума, соответствующем пиковому значению 5. Схемы, реализующие вычисление и предсказание регуляризационных фильтров, показывают более низкие значения метрик, при этом TRL-K показывает на 2 дБ лучшие результаты, чем TRL-KPN, для самого сильного уровня шума.

Разработанная схема, основанная на итеративном алгоритме Ричардсона-Люси с тренируемыми параметрами регуляризации, показала удовлетворительные результаты восстановления изображений, искажённых размытием и слабым пуассоновским шумом, соответствующим пиковым значениям эталонного изображения в диапазоне (25, 100, 255). Тем не менее, в задаче восстановления изображения, искажённого сильным пуассоновским шумом, соответствующим

пиковому значению эталонного изображения, равному 5, предложенные схемы показали худшие по сравнению с методом глубокого обучения – end-to-end нейронной сетью UNet – результаты.

Таблица 1 — Сравнительная таблица метрик качества восстановления изображений, искажённых размытием и пуассоновским шумом четырёх уровней интенсивности. Наилучший результат выделен жирным шрифтом.

	Максимальное значение пикселя в изображении							
	5		25		100		255	
	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM
Искажённое изображение	16.67	.1653	23.25	.3970	28.77	.6574	32.15	.7952
RL+TV	22.09	.3686	28.73	.6701	33.90	.8574	36.65	.9166
UNet	30.90	.8370	33.91	.8895	36.20	.9205	37.35	.9339
TRL	25.13	.4639	30.37	.7383	33.61	.8730	34.92	.9121
TRL-KPN	22.96	.5568	29.36	.7588	34.39	.8759	36.62	.9190
TRL-UNet	30.18	.7425	34.24	.8826	36.81	.9217	38.21	.9365



Сверху вниз: изображения нормализованы к максимальным значениям, равным 5, 25, 100 и 255.

Рисунок 1.2 — Результаты восстановления размытых изображений, искажённых пуассоновским шумом различной интенсивности.

1.4 Обучение моделей для восстановления изображений

Обучение моделей, входящих в состав разработанных алгоритмов, происходит путём восстановления изображения из искажённого и последующего сравнения реконструированного изображения с эталонным и минимизации ℓ_1 функции потерь между изображениями и градиентами изображений

$$L = \|\hat{\mathbf{x}} - \mathbf{x}\|_1 + \|\nabla\hat{\mathbf{x}} - \nabla\mathbf{x}\|_1. \quad (1.22)$$

В работе [62] показано, что выбор ℓ_1 функции потерь приводит к более чёткому восстановлению деталей на изображениях. Все модели оптимизированы с помощью алгоритма Adam [63] с использованием скорости обучения, равной 10^{-3} . Для тренировки всех методов использованы тренировочные наборы, состоящие из 25 изображений, и обучение моделей проводилось на протяжении 300 эпох. Метод WF-KPN-SA, описанный далее, обучен с использованием тренировочного набора, состоящего из трёх изображений.

1.5 Восстановление изображений, искажённых гауссовским шумом

1.5.1 Вычисление обучаемых регуляризационных фильтров

Рассматривая задачу восстановления изображения, искажённого размытием и гауссовским шумом, с точки зрения теоремы Байеса, по аналогии с алгоритмом, описанным в Разделе 1.3.1, можно сформулировать выражение для оценки апостериорного максимума с помощью минимизации отрицательного логарифма:

$$\hat{\mathbf{x}} = \arg \min_{\mathbf{x}} [-\ln p(\mathbf{y} | \mathbf{x}) - \ln p(\mathbf{x})]. \quad (1.23)$$

При наличии аддитивного гауссовского шума плотность вероятности, описывающая полученное при визуализации объекта \mathbf{x} изображение \mathbf{y} , может быть

описана функцией Гаусса, то есть

$$p(\mathbf{y} | \mathbf{x}) = \prod_{n=1}^N \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{((\mathbf{H}\mathbf{x})_n - y_n)^2}{2\sigma^2}}. \quad (1.24)$$

Учитывая Формулу 1.24, а также применяя аналогичную описанной в Разделе 1.3.1 формулировку априорного члена, – а именно $p(\mathbf{x}) = e^{-\gamma \sum_{d=1}^D \|\mathbf{G}_d \mathbf{x}\|_2^2}$ [13], – решение для Формулы 1.23 можно сформулировать, как

$$\hat{\mathbf{x}} = (\mathbf{H}^\top \mathbf{H} + \gamma \sum_{d=1}^D \mathbf{G}_d^\top \mathbf{G}_d)^{-1} \mathbf{H}^\top \mathbf{y}. \quad (1.25)$$

Процедура свёртки здесь также выражена через матрично-векторное произведение. Формула 1.25 соответствует фильтрации Винера-Колмогорова [64] в области изображений. В этой формуле \mathbf{H}^\top и \mathbf{G}_d^\top являются транспонированными матрицами \mathbf{H} и \mathbf{G}_d соответственно. Получение транспонированной матрицы может потребовать большого количества вычислений в случае, если исходная матрица имеет достаточно большие размеры. Тем не менее, преимуществом фильтра Винера является возможность его формулировки с помощью быстрого преобразования Фурье, что делает этот фильтр эффективным методом и позволяет восстановить искомое изображение с низкой вычислительной сложностью. Матрицы \mathbf{H} и \mathbf{G}_d принято рассматривать как циркулянтные матрицы вещественных чисел (см. Раздел 1.1), которые могут быть диагонализированы в области Фурье как

$$\mathbf{H} = \mathbf{F}^H \mathbf{D}_H \mathbf{F}, \mathbf{D}_H = \text{diag}(\mathbf{F} \mathbf{S}_H \mathbf{P}_H \mathbf{h}); \quad \mathbf{G}_d = \mathbf{F}^H \mathbf{D}_{G_d} \mathbf{F}, \mathbf{D}_{G_d} = \text{diag}(\mathbf{F} \mathbf{S}_{G_d} \mathbf{P}_{G_d} \mathbf{g}_d). \quad (1.26)$$

В Формуле 1.26 $\mathbf{F} \in \mathbb{C}^{N \times N}$ является матрицей, осуществляющей дискретное преобразование Фурье [65], $\mathbf{F}^H \in \mathbb{C}^{N \times N}$ – соответствующая эрмитово-сопряжённая матрица, $\mathbf{h} \in \mathbb{R}^{K^2}$ – вектор функции размытия точки, $\mathbf{g}_d \in \mathbb{R}^{L^2}$ – вектор линейного фильтра регуляризации, $\mathbf{S}_H, \mathbf{S}_{G_d} \in \mathbb{R}^{N \times N}$ – соответствующие матрицы, осуществляющие линейную операцию сдвига векторов \mathbf{h} и \mathbf{g}_d , $\mathbf{P}_H \in \mathbb{R}^{N \times K^2}$, $\mathbf{P}_{G_d} \in \mathbb{R}^{N \times L^2}$ – соответствующие матрицы, осуществляющие линейную операцию заполнения нулями векторов \mathbf{h} и \mathbf{g}_d , $\mathbf{D}_H, \mathbf{D}_{G_d} \in \mathbb{C}^{N \times N}$ – соответствующие диагональные матрицы, получающиеся в результате операций заполнения нулями и сдвига, произведённых над векторами \mathbf{h} и \mathbf{g}_d . Отме-

тим, что операция, осуществляемая матрицами \mathbf{FSP} , преобразует PSF в OTF – optical transfer function, оптическую передаточную функцию системы. Для всех экспериментов значения регуляризационного коэффициента γ также приняты равным e^α для обеспечения положительности коэффициента. С учётом всего вышеизложенного, решение Формулы 1.25 имеет вид

$$\hat{\mathbf{x}} = \mathbf{F}^H \left(\frac{\mathbf{D}_{\mathbf{H}}^* \mathbf{F} \mathbf{y}}{|\mathbf{D}_{\mathbf{H}}|^2 + e^\alpha \sum_{d=1}^D |\mathbf{D}_{\mathbf{G}d}|^2} \right), \quad (1.27)$$

где $\mathbf{D}_{\mathbf{H}}^*$ – комплексно-сопряжённая к $\mathbf{D}_{\mathbf{H}}$ матрица, а процедура деления осуществляется поэлементно. Схема предложенного метода, обозначенного здесь и далее WF-K (**W**iener **F**ilter with **K**ernels), представлена на Рисунке 1.3а. Для осуществления решения согласно Формуле 1.27 применён набор из D обучаемых фильтров \mathbf{g}_d размерностью $L \times L$, инициализированных с помощью дискретного косинусного преобразования. Аналогично методам, основанным на алгоритме Ричардсона-Люси, коэффициент α здесь и далее является обучаемым параметром.

В то время как функция потерь, минимизируемая в процессе тренировки моделей, вещественнозначная, решение, полученное в соответствии с формулой 1.27, включает в себя комплекснозначные величины. В связи с этим невозможно применение готовых алгоритмов вычисления обратного распространения ошибки, реализованных в существующих библиотеках алгоритмов глубокого обучения. По этой причине вычисление градиентов функции потерь относительно обучаемых параметров было проведено вручную. Результаты вычисления представлены в Приложении А.

1.5.2 Вычисление регуляризационных фильтров с помощью нейронной сети

Как и в методе TRL-K, основным недостатком предыдущей формулировки фильтра Винера является то, что регуляризационные фильтры являются одинаковыми для всех изображений. Поэтому предложена формулировка фильтра Винера с использованием нейронной сети, предсказывающей регуляризацион-

ные фильтры для каждого отдельного изображения. В предложенном методе решение также сформулировано согласно Формуле 1.27, но в отличие от алгоритма WF-K, регуляризационные фильтры \mathbf{g}_d сформированы на выходе нейронной сети Kernel Prediction Network (KPN). Предложенный метод представлен на Рисунке 1.3б и обозначен WF-KPN (**W**iener **F**ilter with **K**ernel **P**rediction **N**etwork).

По аналогии с методом TRL-KPN, описанным в Разделе 1.3.2, для Kernel Prediction Network выбрана компактная архитектура UNet, содержащая около 470,000 обучаемых параметров, или весов сети, на вход которой поступает набор из искажённых изображений, а на выходе формируется матрица такого же размера, но содержащая L^2D каналов. Далее полученный вектор подвергается преобразованию в D регуляризационных фильтров размером $L \times L$. В данном алгоритме в сети KPN использована нормализация значений каждого экземпляра в наборе. Таким образом, для каждого отдельного изображения KPN предсказывает D регуляризационных фильтров размера $L \times L$, которые затем, согласно Формуле 1.26, формируют диагональные матрицы $\mathbf{D}_{\mathbf{G}_d}$, входящие в состав решения 1.27.

Изображения регуляризационных фильтров, полученных с помощью алгоритмов WF-K и WF-KPN представлены в Приложении В.

1.5.3 Вычисление регуляризационных фильтров для каждого пикселя с помощью нейронной сети

В то время как предсказание регуляризационных фильтров для каждого изображения обеспечивает модели WF-KPN способность адаптироваться к различной контекстуальной информации, на следующем этапе работы исследована возможность адаптации не только к разным изображениям, но и к различным текстурам в одном изображении. С этой целью разработан алгоритм, объединяющий фильтр Винера с регуляризационными фильтрами, предсказанными для каждого пикселя каждого изображения. Эти фильтры также далее называются пространственно адаптируемыми. Предложенный метод представлен на Рисунке 1.3в и обозначен WF-KPN-SA (**W**iener **F**ilter with **K**ernel **P**rediction **N**etwork

Spatially Adaptive). Для реализации этого алгоритма UNet, описанная в Разделе 1.3.2, модифицирована следующим образом: на выходе UNet формируется изображение такого же пространственного размера, как и входное искажённое изображение, но содержащее L^2 каналов. Эмпирическим путём было обнаружено, что нормализация данных в слоях нейросети приводит к худшим результатам для WF-KPN-SA, поэтому во всех экспериментах с этим методом нормализация отключена. Затем полученное на выходе сети изображение подвергается преобразованию в отдельные для каждого пикселя входного изображения \mathbf{y} регуляризационные фильтры размера $L \times L$. После этого фильтры преобразованы в матрицу \mathbf{G} , не являющуюся циркулярной в отличие от \mathbf{G}_d в Формуле 1.25. По этой причине выражение фильтра Винера с помощью БПФ в данном случае невозможно. Таким образом, решение задачи восстановления изображения имеет вид

$$\hat{\mathbf{x}} = (\mathbf{H}^\top \mathbf{H} + e^\alpha \mathbf{G}^\top \mathbf{G})^{-1} \mathbf{H}^\top \mathbf{y}. \quad (1.28)$$

Процедура вычисления оценки изображения $\hat{\mathbf{x}}$ производится с использованием метода сопряжённых градиентов [66] – итеративного метода для оценки локального экстремума целевой функции с использованием информации о её градиенте.

1.5.4 Аппроксимация регуляризационного члена с помощью нейронной сети

На последнем этапе данной работы по аналогии с методом TRL-UNet предложено использование нейронной сети для аппроксимации всего регуляризационного члена, а не отдельных регуляризационных фильтров. Применяя такую формулировку задачи, невозможно сделать предположение о форме регуляризационного члена $r(\mathbf{x})$, поэтому вывод решения в аналитической форме аналогично Формуле 1.25 невозможен. Вместо этого предложено воспользоваться альтернативным способом поиска оптимального решения – с помощью алгоритма градиентного спуска [52]. Идея состоит в том, чтобы найти точку, в которой целевая функция

$$L(\mathbf{x}) = \frac{1}{2} \|\mathbf{y} - \mathbf{H}\mathbf{x}\|_2^2 + e^\alpha r(\mathbf{x}) \quad (1.29)$$

достигает своего минимума, двигаясь в направлении обратном наибольшему изменению целевой функции. Иными словами, производится последовательный поиск точки, в которой значение разности между наблюдением и полученным решением достигает наименьшего значения. Этот процесс выражен формулой

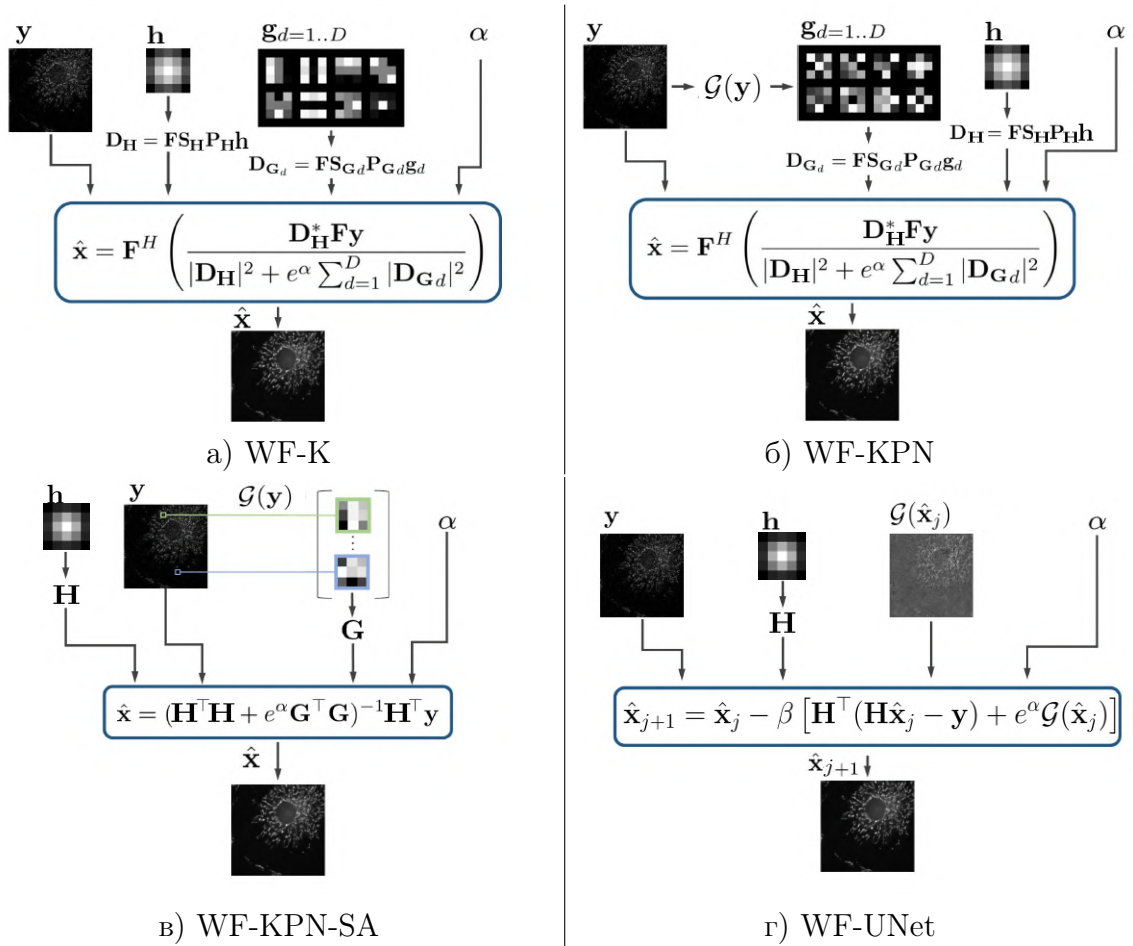
$$\hat{\mathbf{x}}_{j+1} = \hat{\mathbf{x}}_j - \beta \nabla L(\mathbf{x}). \quad (1.30)$$

Здесь $\hat{\mathbf{x}}_{j+1}$ представляет собой оценку решения Формулы 1.29, обновление величины которой происходит на каждой итерации j алгоритма градиентного спуска, а коэффициент β является шагом градиентного спуска и определяет скорость алгоритма. Подставляя в Формулу 1.30 целевую функцию $L(\mathbf{x})$, определённую в соответствии с Формулой 1.29, получаем решение в виде

$$\hat{\mathbf{x}}_{j+1} = \hat{\mathbf{x}}_j - \beta [\mathbf{H}^\top (\mathbf{H}\hat{\mathbf{x}}_j - \mathbf{y}) + e^\alpha \mathcal{G}(\hat{\mathbf{x}}_j)]. \quad (1.31)$$

Здесь нейронная сеть \mathcal{G} используется для аппроксимации градиента регуляризационного члена $\nabla r(\hat{\mathbf{x}}_j)$. С этой целью также, как и в других алгоритмах, использована UNet, описанная в Разделе 1.3.2. В отличие от алгоритмов WF-KPN и WF-KPN-SA, в данном алгоритме последний свёрточный слой сети не подвергается дополнительным модификациям – на входе сеть получает искажённое изображение \mathbf{y} и формирует на выходе градиент регуляризатора $r(\mathbf{x})$, имеющий такие же пространственные размеры и количество каналов, как и входное изображение. Важно уточнить, что на каждой итерации алгоритма градиентного спуска нейронная сеть получает на входе решение $\hat{\mathbf{x}}_j$, полученное на предыдущей итерации алгоритма. После каждого свёрточного слоя нейронной сети применена нормализация значений каждого экземпляра в наборе. Количество итераций J для алгоритма градиентного спуска выбрано равным 10. Коэффициент β , определяющий шаг в алгоритме градиентного спуска, является обучаемым параметром.

Хотя описанный метод не является фильтром Винера-Колмогорова ввиду отсутствия решения в аналитической форме, он разделяет общую для всех описанных методов идею восстановления изображений с использованием поиска оптимального решения целевой функции $L(\mathbf{x})$, описанной Формулой 1.29, и обучаемых параметров регуляризации. Поэтому было принято решение использовать общий подход в наименовании всех разработанных алгоритмов для вос-



На Рисунке 1.3а представлен фильтр Винера с обучаемыми

регуляризационными фильтрами, на Рисунке 1.3б изображён фильтр Винера с регуляризационными фильтрами, вычисленными с помощью нейронной сети

для каждого изображения. На Рисунке 1.3в представлен фильтр Винера с регуляризационными фильтрами, вычисленными с помощью нейронной сети

для каждого пикселя. На Рисунке 1.3г отображена схема, в которой градиент регуляризационного члена вычислен с помощью нейронной сети.

Рисунок 1.3 — Схемы четырёх алгоритмов на основе фильтра Винера, реализующих аппроксимацию регуляризационного члена тренируемыми моделями.

становления изображений, искажённых гауссовским шумом, и назвать последний метод WF-UNet (**W**iener **F**ilter with **U**Net). Схема алгоритма представлена на Рисунке 1.3г.

1.5.5 Подготовка набора изображений, искажённых размытием и гауссовским шумом

Для обучения моделей для восстановления изображений, искажённых размытием и гауссовским шумом, аналогичным описанному в Разделе 1.3.4 образом создан набор тренировочных данных. В отличие от случая пуассоновского шума, для симуляции случая широкопольной микроскопии с помощью плагинов Diffraction PSF 3D и PSF Generator, реализованных в программе Fiji, создано 35 различных двумерных функции размытия точки размерами 7×7 , 9×9 , 11×11 и 13×13 . Из полученных PSF 25 использованы для создания тренировочных изображений, 5 для создания данных для валидации и оставшиеся 5 для генерации изображений для тестирования. Примеры тестовых PSF и их свёрток с изображениями представлены на Рисунке 1.5.

Синтез размытых и искажённых гауссовским шумом изображений происходит следующим образом: сначала эталонные изображения размером 256×256 подвергаются свёртке с PSF, а затем к ним добавляется гауссовский шум. Во время тренировки каждое эталонное изображение из 975 подвергается свёртке со случайно выбранной из 25 зарезервированных для этой цели PSF, затем к результату добавлен гауссовский шум со случайно выбранным из диапазона (0.001, 0.005, 0.01, 0.05, 0.1) стандартным отклонением.

Для создания тестового набора данных каждое из 230 зарезервированных изображений изображение подвергнуто размытию с применением фиксированной PSF из отложенного для этих целей набора, а затем из каждого полученного размытого изображения получено 5 зашумлённых изображений путём добавления шума с каждым стандартным отклонением из диапазона (0.001, 0.005, 0.01, 0.05, 0.1). Таким образом получено 1150 размытых и зашумлённых изображений, объединённых в 5 наборов в зависимости от использованного стандартного отклонения гауссовского шума. Благодаря этому мы имеем возможность оценить разработанные алгоритмы на широком диапазоне шумов.

1.5.6 Результаты восстановления изображений, искажённых размытием и гауссовским шумом, с помощью фильтра Винера с обучаемыми параметрами регуляризации

Для объективной оценки полученных результатов и производительности предложенных алгоритмов разработанные методы сравнены с тремя передовыми (state-of-the-art, SOTA) алгоритмами деконволюции изображений, искажённых гауссовским шумом, – IRCNN [3], FDN [47] и DMSP [30]. Выбранные методы также реализуют схемы, объединяющие классическое решение задачи поиска оптимального решения с методами глубокого обучения для уточнения априорной информации об ожидаемом решении. Именно по этой причине было решено использовать эти алгоритмы для сравнения. Авторы последнего алгоритма предусмотрели два варианта его использования: в первом варианте модель способна самостоятельно адаптироваться к любому уровню шума (noise adaptive, NA); во втором же модели необходима априорная информация об уровне шума (non-blind, NB). Кроме этого для сравнения использован фильтр Винера, в котором регуляризационный член задан с использованием двумерного дискретного фильтра Лапласа (WF-T, **W**iener **F**ilter with **T**ikhonov regularization). Также сравнение проведено с непосредственно моделью UNet, которая входила в состав большинства разработанных алгоритмов. Применение такого базового метода позволяет сделать заключение о том, насколько оправдано объединение классических схем оптимизации с методами глубокого обучения и насколько оно способствует получению лучших результатов восстановления изображений по сравнению с результатами, полученными исключительно с помощью обученной сети, отображающей искажённое изображение в “чистое”.

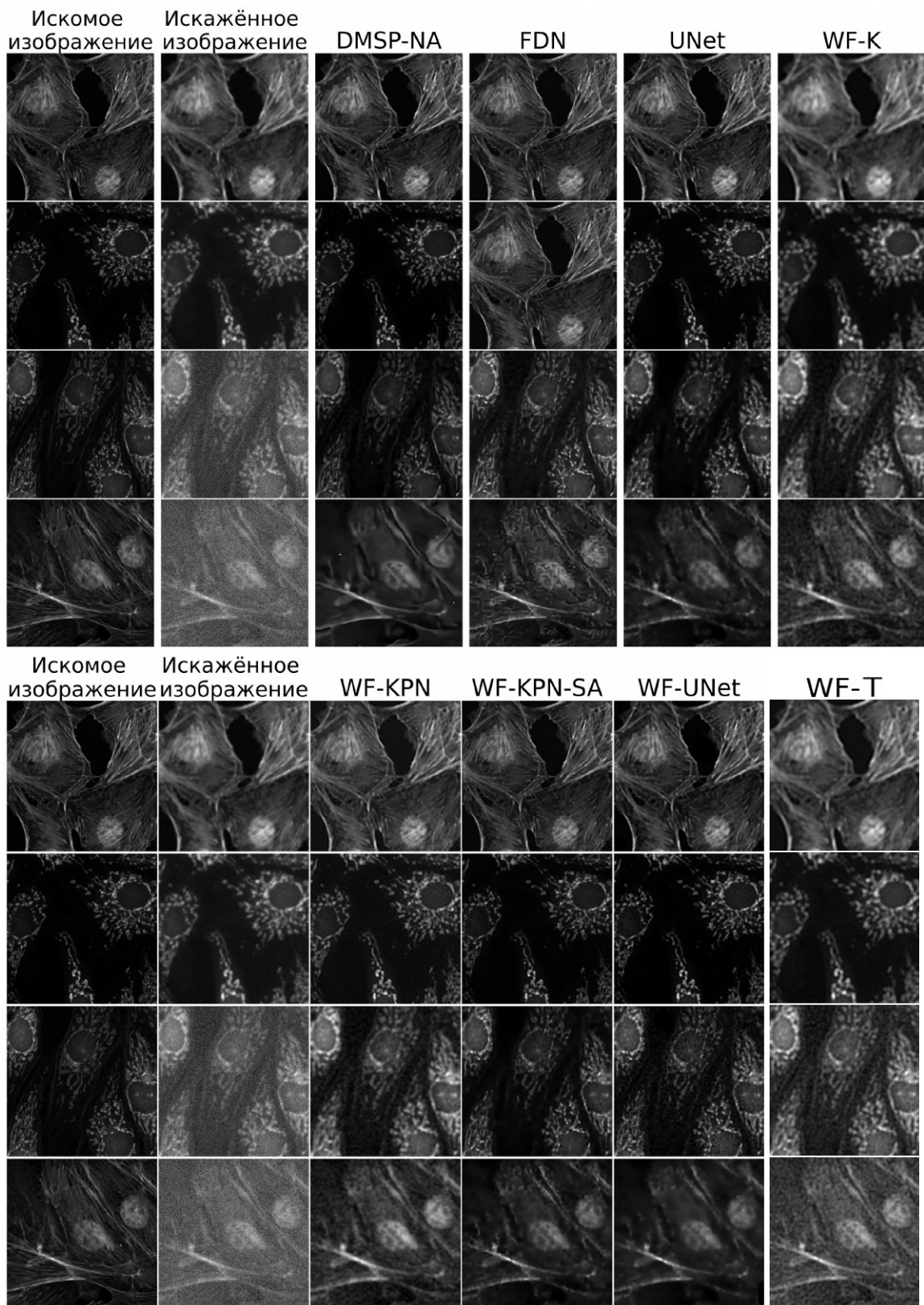
Для оценки всех алгоритмов использован тестовый датасет, описанный в Разделе 1.5.5. Результаты, представленные в Таблице 2, показывают, что предложенная схема пространственно адаптируемых регуляризационных фильтров WF-KPN-SA превосходит остальные разработанные, а также все SOTA методы, за исключением результатов, полученных для самых низких уровней гауссовского шума. Преимущество относительно алгоритмов FDN, IRCNN и DMSP различается от 0.08 дБ для среднего уровня шума до 0.4 дБ для высокого уровня шума, что соответствует $\sigma = 0.01$ и $\sigma = 0.1$ соответственно. Только в случае

самого низкого уровня шума, то есть для $\sigma = 0.001$, SOTA показывают более высокие значения метрик оценки реконструкции изображений, чем WF-KPN-SA. Важно отметить, что так же, как и DMSP-NA, разработанные алгоритмы не получают дополнительной информации об уровне шума в искажённом изображении, то есть являются “слепыми” относительно уровня шума. В свою очередь алгоритмы IRCNN, FDN и DMSP-NB используют информацию о стандартном отклонении сигнала, полученную с помощью функции, оценивающей это значение и реализованной в библиотеке `scikit-image` [67].

Из результатов, представленных в Таблице 2, видно, что фильтр Винера с обучаемыми регуляризационными фильтрами WF-K позволяет получить удовлетворительные результаты восстановления изображений, соперничающие по качеству с другими алгоритмами реконструкции, основанными на глубоком обучении. Кроме того, благодаря реализации в области Фурье и вычислению только регуляризационных фильтров и коэффициента, этот метод имеет высокую скорость выполнения (Таблица 10 сравнения времени, затраченного на выполнение алгоритма, представлена в Приложении В). Из результатов, представленных на Рисунке 1.4, видно, что WF-K и WF-KPN теряют некоторые детали при восстановлении изображений, в то время как WF-KPN-SA и WF-UNet позволяют достигать высокую точность в реконструкции деталей.

Дополнительные результаты для предложенных алгоритмов восстановления изображений, искажённых размытием и гауссовским шумом, представлены в Таблице 8 Приложения В.

Исследование влияния изменения числа и размера регуляризационных фильтров. Предложенные схемы WF-K и WF-KPN осуществляют решение согласно Формуле 1.27, и в обеих схемах используется D регуляризационных фильтров размера $L \times L$. В данном разделе представлены результаты исследования влияния количества и размера регуляризационных фильтров на восстановление изображений. Для обеих схем применено два варианта: в первом варианте $D = 8$ и $L = 3$, во втором – $D = 24$ и $L = 5$. Сравнение проведено для случая восстановления изображений, искажённых размытием и гауссовским шумом, из набора данных, описанного в Разделе 1.3.4. Результаты, представленные в Таблице 3, показывают, что в случае алгоритма WF-K увеличение числа и размера регуляризационных фильтров приводит к увеличению метрики PSNR на 0.15 дБ в случае слабого шума. В то же время, в случае сильного зашум-



Сверху вниз на обеих фигурах: изображения искажены гауссовским шумом со стандартным отклонением 0.005, 0.01, 0.05, и 0.1.

Рисунок 1.4 — Результаты восстановления размытых изображений, искажённых гауссовским шумом различной интенсивности.

Таблица 2 — Сравнительная таблица метрик качества восстановления изображений, искажённых размытием и гауссовским шумом пяти уровней интенсивности. Наилучший результат выделен жирным шрифтом.

	Значение стандартного отклонения									
	0.001		0.005		0.01		0.05		0.1	
	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM
Input	36.23	.8955	35.37	.8791	33.93	.8339	26.03	.3858	21.14	.1718
IRCNN	33.33	.8604	36.88	.8972	36.80	.9013	32.44	.7932	28.95	.6835
FDN	40.31	.9424	38.6	.9239	37.33	.9086	33.50	.8406	30.49	.7550
DMSP-NA	40.44	.9402	39.16	.9290	37.73	.9123	34.31	.8589	32.19	.8312
DMSP-NB	40.27	.9411	37.23	.9027	36.59	.8930	34.29	.8582	32.20	.8313
UNet	37.73	.9177	37.27	.9136	36.64	.9064	34.04	.8556	32.40	.8187
WF-T	36.72	.8964	36.62	.8948	36.33	.8902	32.89	.8029	29.45	.6646
WF-K	35.66	.8849	35.61	.8834	35.45	.8787	32.74	.7950	29.27	.6835
WF-KPN	38.72	.9253	37.98	.9176	36.80	.9028	32.33	.8022	29.20	.7259
WF-KPN-SA	39.86	.9390	38.76	.9275	37.81	.9157	34.58	.8688	32.60	.8363
WF-UNet	38.08	.9102	37.55	.9053	36.77	.8966	33.89	.8442	32.01	.8096

ления изображения, увеличение размера и количества фильтров приводит к незначительному понижению метрики на 0.06 дБ. В случае схемы WF-KPN, использование 24 фильтров размера 5×5 позволяет улучшить качество восстановленных изображений на почти 0.4 дБ в случае сильного шума ($\sigma = 0.05$) и на 1.23 дБ в случае слабого шума. Таким образом, в целом увеличение числа и размера регуляризационных фильтров закономерно улучшает качество восстановленных изображений, но увеличивает время работы алгоритма вследствие большего количества вычисляемых параметров.

Таблица 3 — Сравнительная таблица метрик качества восстановления изображений, искажённых размытием и гауссовским шумом пяти уровней интенсивности для различных размеров и количества регуляризационных фильтров.

		Значение стандартного отклонения									
		0.001		0.005		0.01		0.05		0.1	
		PSNR	SSIM	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM
Искажённое изображение		36.23	.8955	35.37	.8791	33.93	.8339	26.03	.3858	21.14	.1718
$D = 8$	WF-K	35.66	.8849	35.61	.8834	35.45	.8787	32.74	.7950	29.27	.6835
$L = 3$	WF-KPN	38.72	.9253	37.98	.9176	36.80	.9028	32.33	.8022	29.20	.7259
$D = 24$	WF-K	35.81	.8861	35.75	.8846	35.58	.8798	32.75	.7945	29.21	.6807
$L = 5$	WF-KPN	39.95	.9368	38.41	.9218	37.20	.9057	32.69	.8095	29.20	.7386

1.6 Восстановление изображений, искажённых пуассоновским шумом, с помощью фильтра Винера с обучаемыми параметрами регуляризации

Разработанные на первом этапе работы алгоритмы Ричардсона-Люси для изображений, искажённых пуассоновским шумом, показали неудовлетворительные результаты в случае сильного зашумления, в то время как алгоритмы фильтра Винера с обучаемыми параметрами показали превосходные результаты в задаче восстановления изображений, искажённых даже сильным гауссовским шумом. Одним из распространённых методов восстановления сигнала, искажённого пуассоновским шумом, является преобразование его в сигнал, подчиняющийся нормальному распределению, с помощью стабилизирующего дисперсию сигнала преобразования [68]. Такой подход позволяет использовать хорошо изученные методы восстановления сигналов, подчиняющихся гауссовскому распределению [39; 54; 68; 69], например, основанные на фильтре Винера алгоритмы. В настоящей работе предложено применение такого подхода и восстановление изображений, искажённых пуассоновским шумом, с помощью разработанных модификаций фильтра Винера.

В данной работе этот подход реализован с помощью широко известного преобразования, стабилизирующего дисперсию, – преобразования Анскомба [70], применённого к искажённым изображениям

$$y \rightarrow 2\sqrt{y + \frac{2}{3}}. \quad (1.32)$$

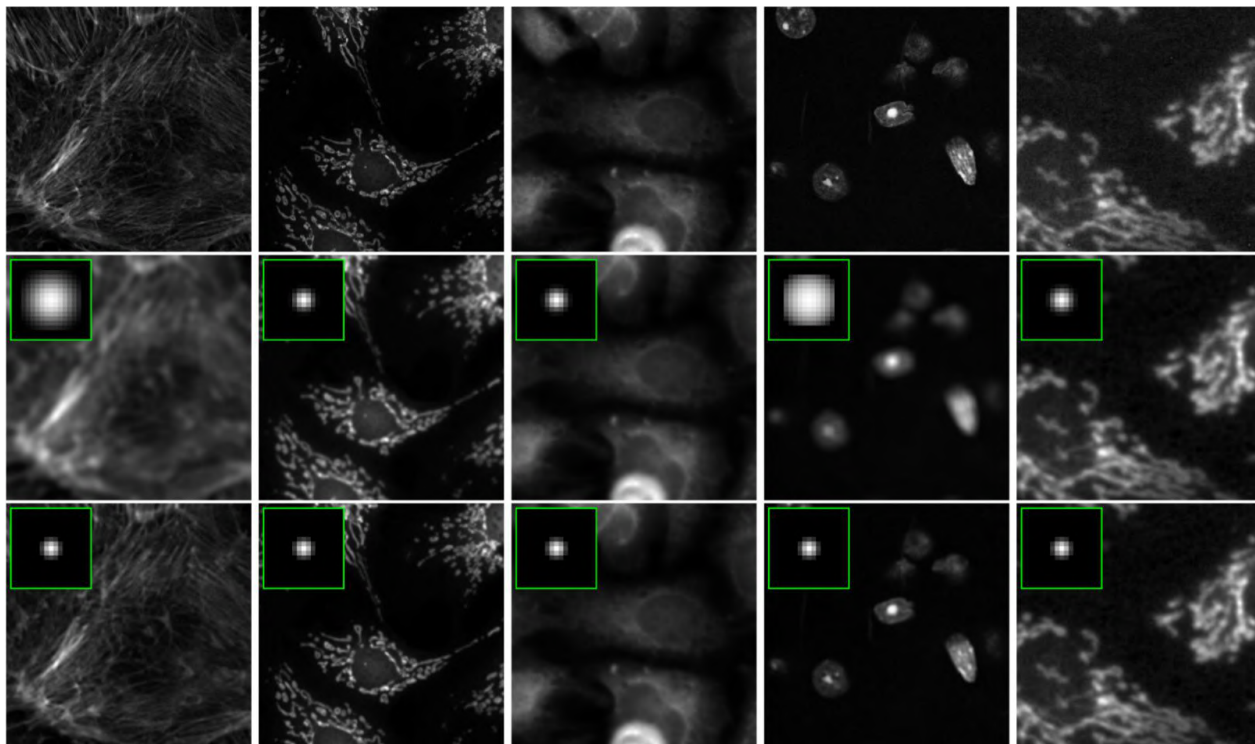
В результате преобразования Анскомба шум, присутствующий в искажённом изображении, может быть описан как независимый от сигнала процесс с дисперсией, равной 1. Таким образом, для восстановления полученного изображения уместно применение фильтра Винера. После того, как изображение $\hat{\mathbf{x}}$ восстановлено, к нему необходимо применить обратное преобразование Анскомба, чтобы вернуть данные в область исходного распределения. Использование непосредственно обратного преобразования обычно приводит к смещению результата, поэтому в настоящей работе применено несмещённое обратное преобразование [71]

$$\hat{\mathbf{x}} \rightarrow \left(\frac{\hat{\mathbf{x}}}{2}\right)^2 - \frac{1}{8} + \frac{1}{4}\sqrt{\frac{3}{2}}\hat{\mathbf{x}}^{-1} - \frac{11}{8}\hat{\mathbf{x}}^{-2} + \frac{5}{8}\sqrt{\frac{3}{2}}\hat{\mathbf{x}}^{-3}. \quad (1.33)$$

1.6.1 Подготовка набора изображений, искажённых пуассоновским шумом

Для оценки эффективности алгоритмов для решения задачи в случае пуассоновского шума, все модели обучены с применением преобразования Анскомба к входным данным \mathbf{y} согласно Формуле 1.32 и обратного преобразования полученных изображений $\hat{\mathbf{x}}$ в соответствии с Формулой 1.33. В этом случае использован набор изображений, описанный ранее в Разделе 1.3.4. С целью исследования способности метода восстанавливать изображения, искажённые более широким диапазоном пуассоновского шума, чем представлено в Разделе 1.3.4, эталонные изображения нормализованы таким образом, чтобы максимальное значение изображения соответствовало одному из значений в диапазоне (1, 2, 5, 10, 25, 50). Для симуляции размытия изображений создано 35 двумерных PSF размером 5×5 , 25 из которых использованы для тренировочных данных,

5 для валидационных и 5 для тестовых. Примеры тестовых PSF и их свёрток с изображениями представлены на Рисунке 1.5.



Верхняя строка – примеры эталонных изображений. Средняя строка – 5 PSF, использованных для симуляции размытия и искажения гауссовским шумом, и результирующие искажённые изображения. Нижняя строка – 5 PSF, использованных для симуляции размытия и искажения пуассоновским шумом, и результирующие искажённые изображения.

Рисунок 1.5 — Примеры функций размытия точки и результатов её свёртки с изображениями.

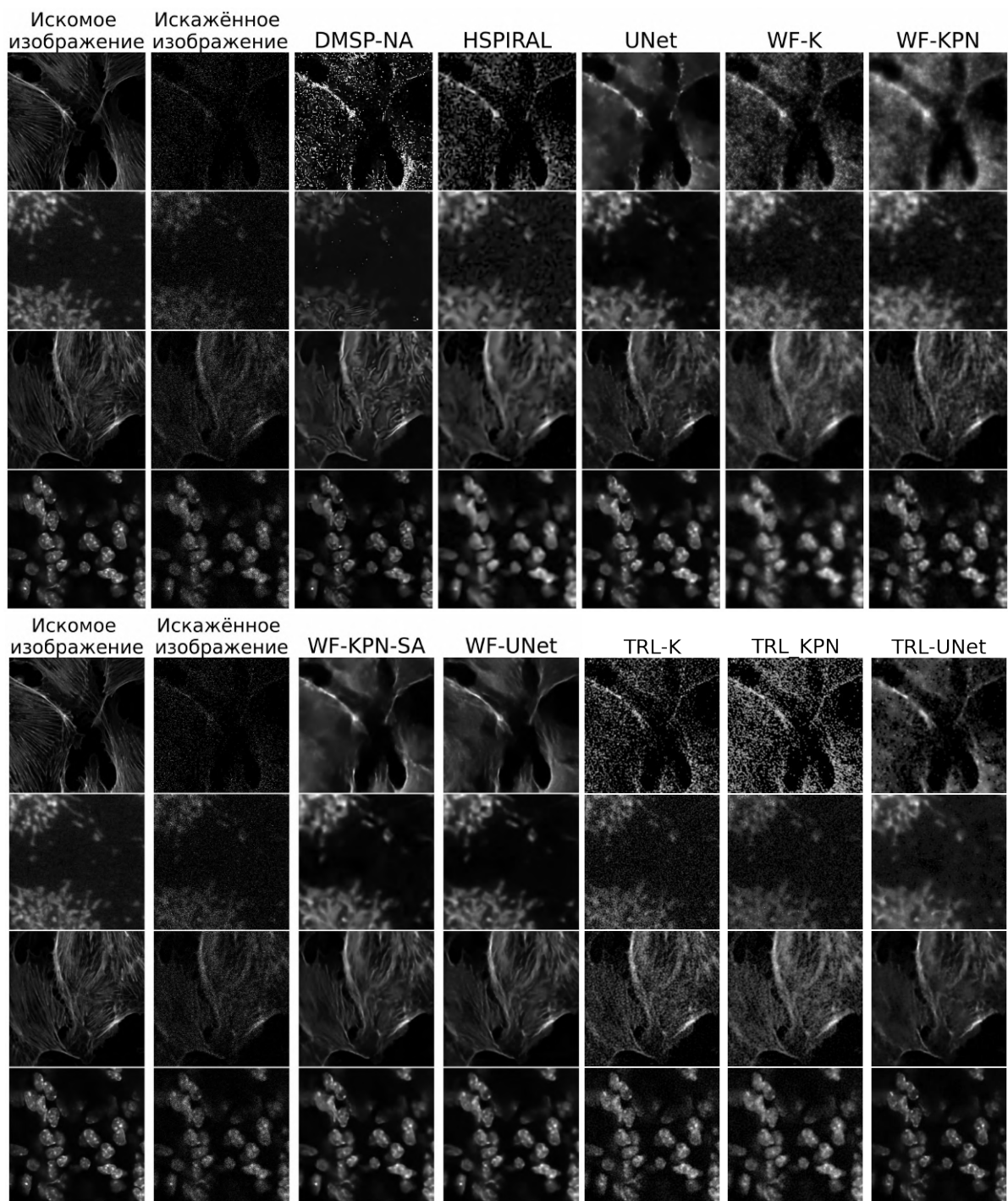
Во время обучения модели эталонное изображение подвергается нормализации к случайно выбранному из упомянутого выше диапазона значению. Затем изображение подвергается свёртке со случайно выбранной PSF из 25 тренировочных, после чего из размытого изображения генерируется зашумлённое. Тестовый набор создан из 230 зарезервированных для этих целей изображений, при этом для каждого изображения сгенерировано 6 изображений – по одному для каждого уровня пуассоновского шума.

1.6.2 Результаты восстановления изображений, искажённых размытием и пуассоновским шумом, с помощью фильтра Винера с обучаемыми параметрами регуляризации

Для объективной оценки разработанные алгоритмы сравнены с SOTA методами, разработанными для деконволюции изображений, искажённых пуассоновским шумом: GILAM [59], HSPiRAL [60] и PURE-LET [72]. Кроме того, в представленные ранее SOTA методы для восстановления изображений, искажённых гауссовским шумом, также внедрено преобразование Анскомба, и с их помощью также получены реконструированные изображения. Из результатов, представленных в Таблице 4 и на Рисунке 1.6 видно, что использование WF-UNet приводит к получению наиболее высоких метрик реконструкции изображений и наиболее высокой точности восстановления деталей по сравнению с остальными методами. Алгоритм WF-KPN-SA показывает близкие к WF-UNet результаты, лишь незначительно уступая последнему. В целом WF-UNet и WF-KPN-SA показывают преимущество перед SOTA методами в 2.6-5.8 дБ на высоких уровнях шума и 1.2-7 дБ на низких уровнях шума.

1.7 Заключение главы

В настоящей главе исследована задача восстановления изображений, полученных с помощью микроскопа и искажённых размытием и пуассоновским и гауссовским шумами. Для решения задачи восстановления таких изображений предложен набор алгоритмов деконволюции, объединяющих классические методы оптимизации, основанные на статистике наблюдений, и современные методы глубокого обучения для аппроксимации регуляризационных параметров. Предложена серия методов, реализованных исходя из пуассоновской статистики наблюдаемых изображений, – алгоритмы Ричардсона-Люси с обучаемыми параметрами регуляризации. Несмотря на удовлетворительные результаты восстановления изображений, полученные с помощью схем, основанных на алгоритме Ричардсона-Люси, они уступают результатам, полученным с помощью нейронной сети UNet. Далее предложен набор методов, реализующих фильтрацию Винера-Колмогорова, основанную на гауссовской статистике наблюдаемых



Сверху вниз на обеих фигурах: изображения нормализованы к максимальным значениям, равным 1, 5, 10 и 25.

Рисунок 1.6 — Результаты восстановления размытых изображений, искажённых пуассоновским шумом различной интенсивности.

данных. Этот набор также использует обучаемые параметры регуляризации. Предложены новые схемы аппроксимации регуляризационных параметров с помощью тренируемых моделей, в том числе предсказание индивидуальных регуляризационных фильтров с помощью нейронной сети.

Таблица 4 — Сравнительная таблица метрик качества восстановления изображений, искажённых размытием и пуассоновским шумом шести уровней интенсивности. Наилучший результат выделен жирным шрифтом.

	Максимальное значение пикселя в изображении											
	1		2		5		10		25		50	
	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM
Искажённое изображение	10.93	.0981	13.22	.1105	16.67	.1639	19.45	.2425	23.21	.3930	26.00	.5247
GILAM	24.57	.4977	25.66	.6350	26.64	.5451	28.24	.6304	29.92	.6975	30.97	.7463
HSPIRAL	23.18	.4101	26.54	.5809	30.18	.7576	31.94	.8229	33.39	.8578	34.38	.8729
PURE-LET	26.18	.7318	26.44	.7494	26.77	.7612	27.67	.7799	28.77	.0.8065	28.77	.8114
FDN	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	31.16	.8176
DMSP-NA	16.07	0.4243	18.86	.5648	23.80	.7126	27.76	.7880	31.74	.8490	33.65	.8770
DMSP-NB	10.86	.2079	12.97	.2761	20.49	.5691	26.12	.7356	30.25	.8136	33.01	.8633
IRCNN	6.81	.0902	11.67	.1809	17.37	.3695	22.14	.5323	29.46	.8038	33.10	.8660
UNet	28.50	.7916	29.79	.8143	31.47	.8428	32.72	.8633	34.20	.8871	35.17	.9008
WF-K	25.81	.5821	27.85	.6947	29.94	.7888	31.13	.8288	32.06	.8557	32.43	.8654
WF-KPN	27.09	.7503	28.19	.7811	29.89	.8177	31.34	.8460	33.22	.8770	34.50	.8945
WF-KPN-SA	28.80	.7949	30.12	.8185	31.72	.8468	32.92	.8672	34.48	.8909	35.60	.9056
WF-UNet	29.04	.8005	30.27	.8220	31.83	.8492	33.06	.8694	34.59	.8928	35.67	.9069
TRL	18.90	.2138	21.14	.2867	24.27	.4284	26.60	.5548	29.28	.7116	30.90	.8028
TRL-KPN	18.22	.2931	20.85	.3935	23.63	.5401	25.99	.6374	28.94	.7557	30.50	.8258
TRL-UNet	25.09	.4616	27.93	.6232	30.76	.7741	32.44	.8387	34.24	.8837	35.37	.9021

Серия проведённых экспериментов показала, что разработанные методы, в частности основанные на фильтрации Винера-Колмогорова, обеспечивают превосходное качество восстановления изображений по сравнению с методами, основанными как исключительно на классических оптимизационных схемах, так и только на методах глубокого обучения. Кроме того, совместно с преобразованием, стабилизирующем дисперсию, фильтр Винера с обучаемыми регуляризационными параметрами показал превосходные результаты восстановления изображений, искажённых пуассоновским шумом. Разработанные алгоритмы превзошли в результатах восстановления другие современные методы, объединяющие глубокое обучение для регуляризации и классические подходы к решению обратных задач.

Глава 2. Вычисление функции рассеяния точки микроскопа с помощью методов глубокого обучения

2.1 Функция рассеяния точки и размытие в микроскопии

Целью исследований, проводимых с помощью оптического микроскопа, является получение изображения, наиболее полно и точно отражающего структуры микроскопического объекта. Способность микроскопа визуализировать мелкие структуры с большой чёткостью определяется его разрешающей способностью, которая соответствует минимальному расстоянию между двумя близко расположенными объектами, на котором возможно различить эти объекты при регистрации изображения. Разрешающая способность оптической системы ограничена явлениями абберации и дифракции света [7]. Абберации представляют собой последствия отклонения света от заданного направления вследствие особенностей геометрии оптической системы. Дифракция света определяется его волновой природой. При прохождении света через линзу оптической системы вследствие интерференции света на плоскости формирования изображения создаётся дифракционная картина [7; 73]. *Функция рассеяния точки* (функция размытия точки, общепринятое обозначение PSF от англ. point spread function) является фундаментальным понятием в оптической микроскопии, описывающим распределение освещённости от координат в плоскости изображения как картину дифракции света, излучаемого точечным источником [73–75].

Обычно реальная форма PSF в задачах восстановления изображений неизвестна. Восстановить её можно опираясь на физические предположения о поведении света в оптической системе и используя известные математические модели, такие как модель Борна и Вольфа [7], модель Ричардса и Вольфа [76], модель Гибсона и Ланни [77]. Другим широко используемым способом восстановления формы PSF является визуализация флуоресцентных наночастиц, имитирующих точечный объект [40]. Однако из-за наличия шума, вносимого оптической системой в изображение, визуализация наночастиц требует большого количества измерений с последующим усреднением полученных значений. Кроме того, невозможно полностью соблюсти условия, при которых были по-

лучены искомые изображения образцов, что может привести к искажению результатов восстановления изображений с использованием полученной функции рассеяния точки. Вдобавок полученная PSF имеет низкое отношение полезного сигнала к шуму [78]. Также для воспроизведения формы PSF были предложены алгоритмы, использующие полученные с её помощью искажённые изображения. Например, в работе [75] PSF определена комбинацией четырёх базисных функций, выраженных с помощью интегральной формулы Кирхгофа, и получена непосредственно из зарегистрированных изображений. Аналогичный подход использован в программе EpiDEMIC [79]. В этом случае использованы другие базисные функции, а именно, полиномы Цернике. Оценка формы функции размытия точки с использованием граничной информации в искажённом изображении была предложена в работе [80].

Методы глубокого обучения также нашли применение в задачах восстановления функции рассеяния точки. Например, в работе [25] свёрточная нейронная сеть была использована для выделения карт признаков, используемых затем для вычисления PSF методом обратной фильтрации в Фурье-области. Авторами [81] была предложена модификация модели CycleGAN [82], в которой слой, осуществляющий моделирование PSF, был реализован с помощью дополнительной генеративной сети. Тем не менее, одной из проблем, ограничивающих применение нейронных сетей для оценки PSF, является отсутствие наборов тренировочных данных, содержащих достаточное количество эталонных изображений функции размытия точки [83].

В настоящей работе рассмотрено два подхода к решению задачи восстановления функции рассеяния точки микроскопа для дальнейшей реконструкции с её помощью изображения. Идея обоих подходов заключается в решении задачи деконволюции относительно неизвестной функции рассеяния точки. В первом случае вычисление PSF осуществлено с помощью так называемого *неконтролируемого фильтра Винера* – метода, основанного на оценке апостериорного максимума и предложенного в работе [84]. Для реализации деконволюции использованы искажённые изображения, а также псевдо-эталонные изображения, сгенерированные с помощью методов глубокого обучения. Согласно второму подходу восстановление функции рассеяния точки является одним из этапов итеративного алгоритма слепой деконволюции – метода восстановления искомого изображения при отсутствии информации о PSF, – основанного на схеме

Ричардсона-Люси [49; 50]. Для оценки функции рассеяния точки использован алгоритм обучения с подкреплением. Оба подхода ориентированы на обход проблемы недостаточного количества тренировочных данных для непосредственного применения алгоритмов глубокого обучения к восстановлению PSF. С другой стороны, предлагаемые методы основаны на классических схемах поиска оптимального решения некорректно поставленной обратной задачи восстановления изображений. Это позволяет им учитывать особенности физических процессов формирования изображения.

2.2 Оценка функции рассеяния точки с использованием псевдо-эталонных изображений

Напомним, что с математической точки зрения процесс формирования изображения при визуализации с помощью микроскопа выражен как

$$\mathbf{y} = \mathbf{h} * \mathbf{x} = \mathbf{H}\mathbf{x}, \quad (2.1)$$

где $\mathbf{x} \in \mathbb{R}^N$ – искомое изображение, $\mathbf{y} \in \mathbb{R}^N$ – полученное в процессе визуализации изображение, а $\mathbf{h} \in \mathbb{R}^{K^2}$ и $\mathbf{H} \in \mathbb{R}^{N \times N}$ – функция рассеяния точки и соответствующая ей матрица, процесс формирования которой описан в Разделе 1.1 Главы 1. Исходя из Формулы 2.1, при наличии известного искажённого изображения \mathbf{y} и соответствующего ему изображения \mathbf{x} , задачу деконволюции можно решить относительно функции рассеяния точки \mathbf{h} . В большинстве задач искомое изображение неизвестно, что приводит к формированию проблемы *слепой деконволюции*. В работе [83] было предложено использование методов глубокого обучения для генерирования приближённой оценки искомого изображения $\tilde{\mathbf{x}}$ из имеющихся искажённых изображений \mathbf{y} .

2.2.1 Применение генеративно-сопоставительных сетей для создания псевдо-эталонных изображений

Помимо искажений, вызванных предельной разрешающей способностью микроскопа, существует ещё один фактор, ограничивающий возможность оптической системы формировать неразмытое изображение. Этот фактор связан с дискретной структурой пикселей матрицы приборов с зарядовой связью, используемых в качестве детекторов. Для визуализации мелких деталей на изображении требуется большее количество пикселей, что приводит к увеличению размера изображения при сохранении требований к его реалистичности и детализации [85; 86]. С недавних пор технологии глубокого обучения успешно справляются с задачей генерирования изображений в высоком разрешении из изображений в низком разрешении с помощью свёрточных нейронных сетей. В частности, такая архитектура, как свёрточные генеративно-сопоставительные сети (generative adversarial network, GAN), была создана в том числе для генерирования изображений [87]. Архитектура GAN состоит из двух нейронных сетей – генератора и дискриминатора, – тренируемых “сопоставляться” друг с другом. Задачей генератора является создание образцов (изображений) из некоторого, чаще всего нормального, распределения, а задача дискриминатора состоит в том, чтобы отличить истинные изображения от сгенерированных. Первые результаты использования GAN в задаче улучшения разрешения изображения были представлены в работе [22], сеть была названа *super-resolution GAN*, SRGAN. Позже в работе [23] была предложена усиленная версия SRGAN – ESRGAN (enhanced).

Для того, чтобы реализовать решение задачи поиска оптимальных значений функции рассеяния точки, необходимо располагать информацией об искомом изображении \mathbf{x} или хотя бы о его приближенной оценке $\tilde{\mathbf{x}}$. Пользуясь способностью генеративно-сопоставительных сетей, в частности используемой в данной работе модели ESRGAN, генерировать изображения в высоком разрешении из изображений в низком разрешении, можно получить такую приближённую оценку искомого изображения из имеющегося размытого. Таким образом, при наличии псевдо-эталонных изображений $\tilde{\mathbf{x}}$ и размытых изображений \mathbf{y} можно найти $\hat{\mathbf{h}}$. Использование ESRGAN для генерирования псевдо-эталонных изображений

жений рассмотрено в работе [83]. Оригинальная версия ESRGAN претренирована на естественных изображениях, поэтому автором было принято решение обучить модель на данных микроскопии. Для тренировки ESRGAN были использованы необработанные трёхмерные изображения мозга взрослых мышей, полученные с помощью светлопольного микроскопа. Датасет состоит из 572 трёхмерных 16-битных чёрно-белых изображений высокой чёткости размером 7680×9600 . Тренировочный набор данных был создан следующим образом: двумерные изображения высокой чёткости различного размера были скадрированы из срезов трёхмерного изображения, а затем преобразованы к размеру 128×128 – эти изображения являются эталонными для тренировки ESRGAN. Искажённые изображения, то есть изображения в низком разрешении, были получены путём преобразования скадрированных изображений к размеру 32×32 [83]. Для повышения эффективности метода был применён подход интерполяции весов, то есть веса оригинальной и обученной моделей ESRGAN были усреднены, и полученная модель была обозначена как “интерполированный ESRGAN” [83]. Пример изображений из тренировочного набора данных представлен на Рисунке Б.1 в Приложении Б.

2.2.2 Оценка функции рассеяния точки

В настоящей работе для решения задачи деконволюции и вычисления неизвестной функции рассеяния точки предложено применение алгоритма *неконтролируемого фильтра Винера (unsupervised Wiener filter)*. Данный алгоритм основан на теореме Байеса и описан в работе [84]. Особенность этого алгоритма заключается в том, что с его помощью производится оценка параметров априорных распределений, входящих в состав апостериорного распределения, описывающего решение. Это позволяет учитывать неизвестную априорную информацию об ожидаемом решении. Алгоритм реализован в библиотеке `scikit-image` языка Python [67].

Итак, процесс восстановления функции рассеяния точки состоит из следующих этапов: имеющееся искажённое изображение \mathbf{u} с помощью претренированной модели интерполированного ESRGAN преобразовано в улучшенное псевдо-

эталонное изображение $\tilde{\mathbf{x}}$. Одновременно с этим изображение \mathbf{y} подвергается бикубической интерполяции для увеличения его размера до соответствующего изображению, полученному с помощью ESRGAN. Затем с помощью искажённого изображения \mathbf{y} и псевдо-эталонного изображения $\tilde{\mathbf{x}}$ и неконтролируемого фильтра Винера выполняется деконволюция функции размытия точки $\hat{\mathbf{h}}$. Полученная PSF подвергается процедуре нормализации для того, чтобы сумма всех элементов $\hat{\mathbf{h}}$ не превышала 1. На Рисунке 2.1 представлена схема восстановления трёхмерной функции рассеяния точки с использованием трёхмерных микроскопических изображений. Процесс осуществлён послойно, то есть деконволюция производится для каждого среза трёхмерной PSF. После деконволюции нормализованные двумерные срезы объединены в трёхмерный стек, который также подвергается финальной нормализации.

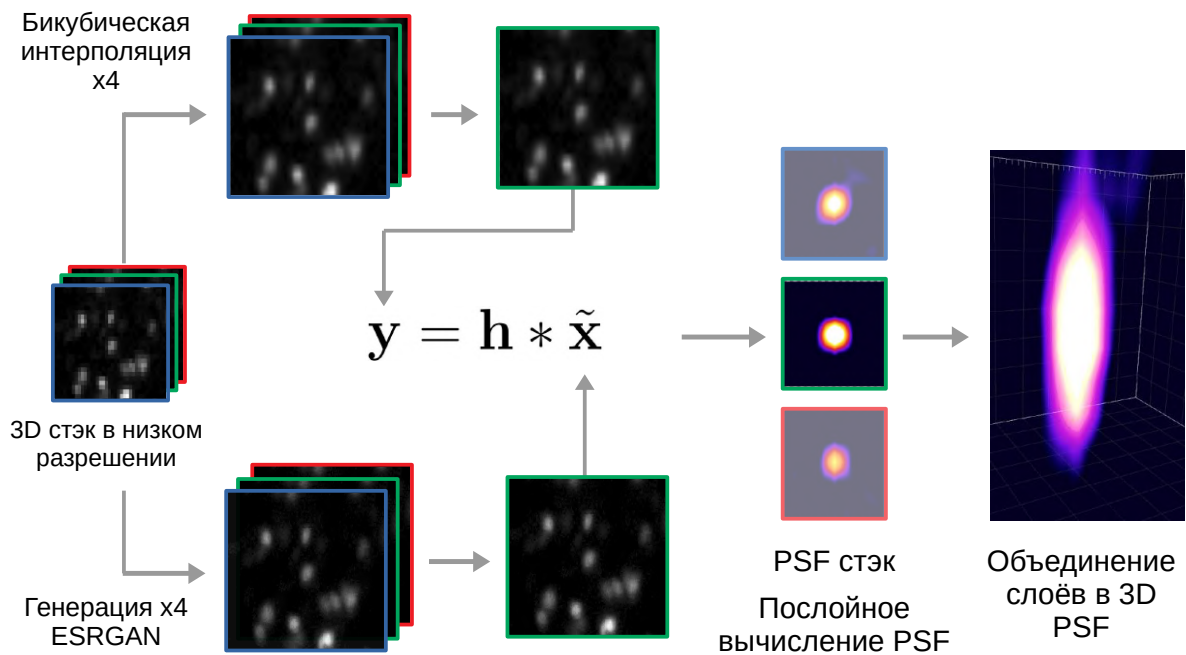


Рисунок 2.1 — Схема восстановления функции рассеяния точки с помощью искажённых и псевдо-эталонных изображений.

Для валидации предложенного метода сгенерировано трёхмерное изображение, содержащее расположенные в нём случайным образом точечные элементы. Это изображение подвергается свёртке со сгенерированной трёхмерной искомой функцией рассеяния точки, а также созданию из него псевдо-эталонного изображения. На Рисунке 2.2 представлены результаты послойного восстановления искомой функции рассеяния точки с помощью искажённых и псевдо-эталонных срезов изображения. Помимо этого для решения задачи слепой деконво-

люции предложенный метод восстановления PSF применён к двумерным изображениям из датасета, описанного в Разделе 1.3.4 Главы 1. Для создания искажённого изображения искомое изображение подвергается свёртке с искомой же функцией размытия точки. Затем искажённое изображение подвергается увеличению, а также с помощью ESRGAN из него генерируется псевдо-эталонное изображение. С помощью этих двух изображений алгоритмом неконтролируемого фильтра Винера производится оценка искомой PSF. Затем искомое изображение восстановлено с использованием вычисленной PSF. Для реконструкции изображения выбран фильтр Винера с обучаемыми регуляризационными фильтрами для каждого пикселя изображения – WF-KPN-SA, описанный в Разделе 1.5.3 Главы 1. Из результатов, представленных на Рисунке 2.3, видно, что реконструкция искажённых изображений с помощью вычисленной PSF показывает лучший результат как визуально, так и с точки зрения метрики отношения пикового сигнала к шуму (PSNR). Эта метрика оценки качества изображения также вычислена и для псевдо-эталонного изображения, предсказанного с помощью ESRGAN. Для этого искомое изображение подвергнуто увеличению с помощью бикубической интерполяции. Полученные результаты показывают, что предложенный метод вычисления неизвестной PSF применим к задачам слепой деконволюции и позволяет улучшить искажённые вследствие размытия изображения.

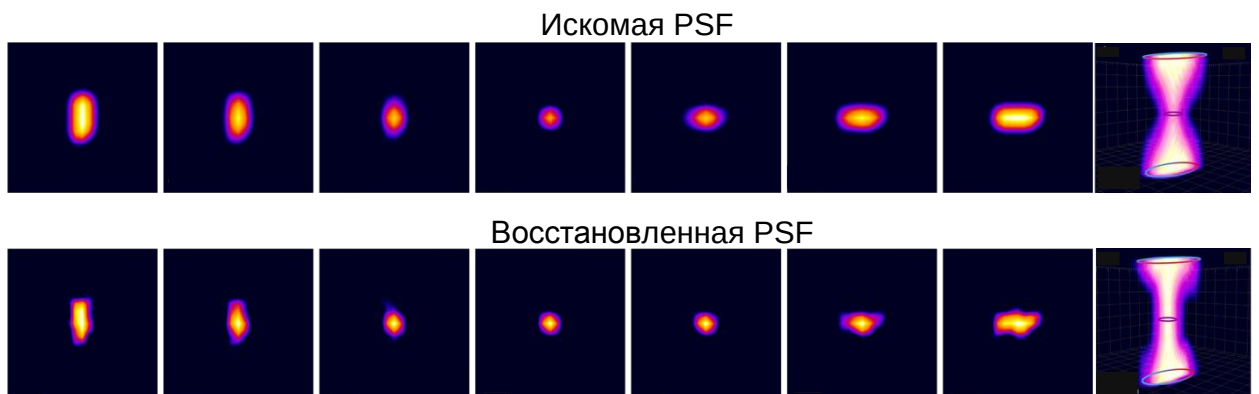
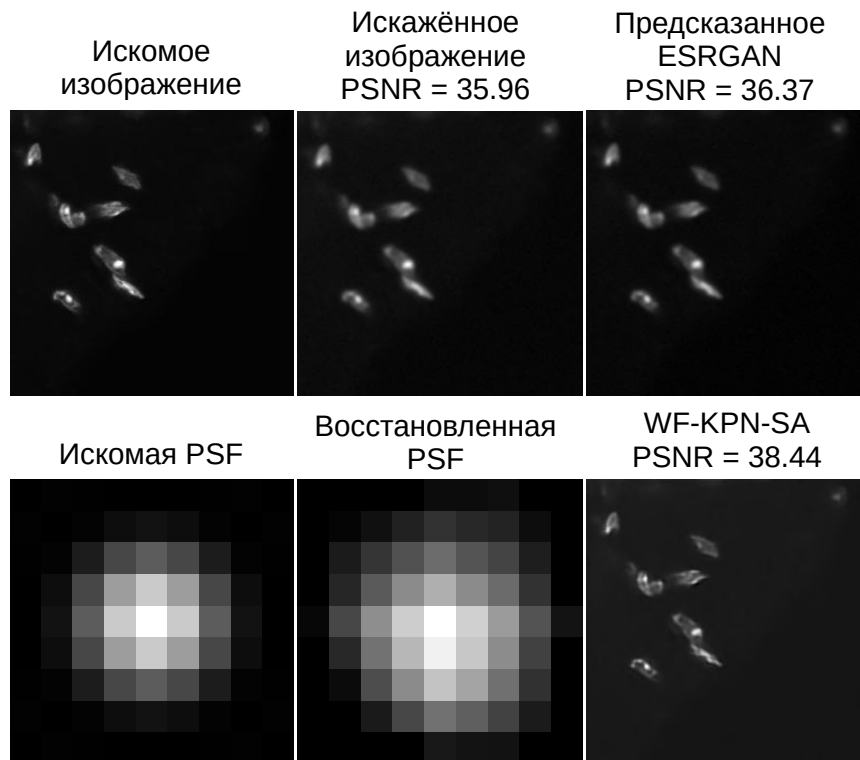


Рисунок 2.2 — Послойное сравнение искомой трёхмерной функции рассеяния точки и её восстановленной с помощью предложенного подхода версией.



Изображение восстановлено с помощью алгоритма WF-KPN-SA(Раздел 1.5.3), PSF восстановлена с помощью неконтролируемого фильтра Винера.
Рисунок 2.3 — Сравнение искомого изображения и функции рассеяния точки с их восстановленными версиями.

2.3 Вычисление функции рассеяния точки и искомого изображения с помощью комбинации оптимизационного алгоритма и тренируемого агента

2.3.1 Математическая модель функции рассеяния точки

Как было указано в Разделе 2.1, математическое моделирование является одним из инструментов для формулировки функции рассеяния точки, характеризующей оптическую систему. При прохождении света через линзу оптической системы, чаще всего имеющую круглую форму, световые волны интерферируют друг с другом, создавая на плоскости изображения картину из концентрических колец, которая носит название узора Эйри [35; 88]. Формула узора Эйри имеет

вид

$$h(r) = \left| C \int_0^1 J_0(k_0 r \text{NA} \rho) d\rho \right|^2. \quad (2.2)$$

Здесь $h(r)$ – функция рассеяния точки в зависимости от радиуса в плоскости изображения r , $k_0 = 2\pi/\lambda n_i$ соответствует волновому числу для длины волны λ , C – максимальная интенсивность в центральной части узора, ρ – нормализованный радиус объектива оптической системы, или выходного отверстия, NA – числовая апертура объектива, J_0 – функция Бесселя первого рода нулевого порядка. Центральное пятно в этом узоре называется диском Эйри.

Более сложные математические модели, такие как модель Борна и Вольфа, учитывают доступную информацию об оптических свойствах объекта и среды, а также об условиях эксперимента, позволяющую им моделировать распространение света не только на плоскости, но и в пространстве [7; 35]. Модель Борна и Вольфа имеет вид

$$h(r, \delta z) = \left| C \int_0^1 J_0(k_0 r \text{NA} \rho) \exp\left[\frac{-ik_0 \delta z \text{NA}^2 \rho^2}{2n_i}\right] \rho d\rho \right|^2. \quad (2.3)$$

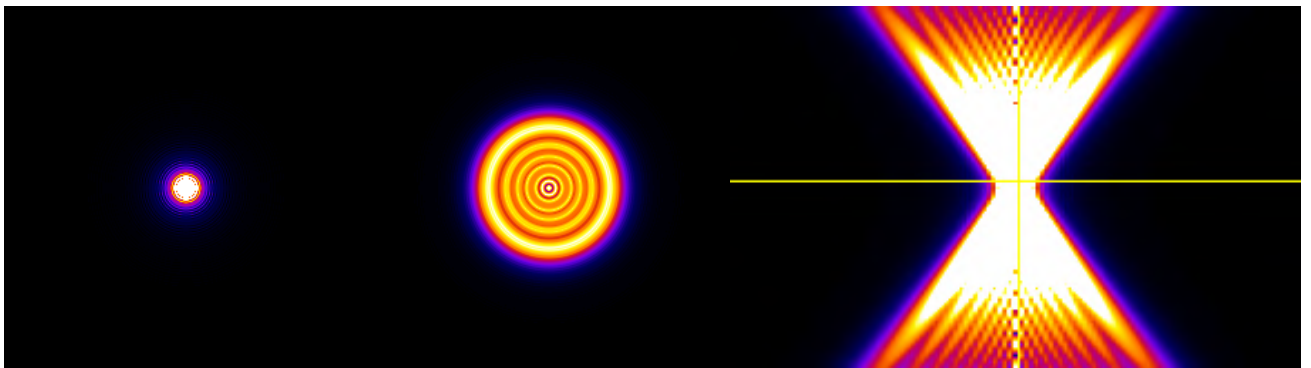
В Формуле 2.3 $\delta z = z_d - z_f$ характеризует степень отклонения плоскости изображения от фокальной плоскости, z_f является фокусным расстоянием, z_d – расстояние от объектива оптической системы до плоскости изображения, $k_0 = 2\pi/\lambda n_i$ соответствует волновому числу для длины волны λ , n_i – иммерсионный показатель преломления, C – максимальная интенсивность в центральной части, r – радиус в плоскости изображения, ρ – нормализованный радиус объектива оптической системы, NA – числовая апертура объектива, J_0 – функция Бесселя первого рода нулевого порядка, $i = \sqrt{-1}$ [35]. Стоит отметить, что при $\delta z = 0$, то есть при нахождении плоскости изображения в фокусе объектива, модель Борна и Вольфа сходится к узору Эйри. Пример PSF, построенной в соответствии с моделью Борна и Вольфа, представлен на Рисунке 2.4.

2.3.2 Итеративная схема вычисления функции рассеяния точки и изображения

В данном разделе настоящей главы удобнее воспользоваться векторной формой записи свёртки функции рассеяния точки и искомого изображения, $\mathbf{h} * \mathbf{x}$. Поэтому здесь и далее в настоящей главе используется обозначение “ \mathbf{h} – функция рассеяния точки”.

Методы решения задачи слепой деконволюции, осуществляющие поиск функции размытия точки и искомого изображения, были ранее исследованы, например, в работе [89]. Авторами была предложена итеративная схема, на каждом шаге которой сначала с помощью метода обратной фильтрации осуществляется оценка функции размытия точки, а затем с её использованием оценка искомого изображения. Позже авторами работы [90] был предложен аналогичный подход, согласно которому вместо обратной фильтрации применён фильтр Винера. Наконец, авторами работ [91] и [92] было предложено итеративное решение задачи слепой деконволюции с использованием схемы Ричардсона-Люси.

Как было описано в Разделе 1.2 Главы 1, алгоритм Ричардсона-Люси сформулирован из предположения, что наблюдаемое изображение описано с помощью распределения Пуассона, и вероятность получения изображения \mathbf{y} при



а) Центральный поперечный срез PSF б) Поперечный срез PSF в) Центральный продольный срез PSF

Рисунок 2.4 — Пример трёхмерной функции рассеяния точки, полученной с помощью модели Борна и Вольфа. Для наглядности визуализации концентрических колец максимальное значение функции ограничено.

визуализации объекта \mathbf{x} определена как

$$p(\mathbf{y} | \mathbf{x}) = \prod_{n=1}^N \left[e^{y_n \ln((\mathbf{h} * \mathbf{x})_n)} \times e^{-(\mathbf{h} * \mathbf{x})_n} \times \frac{1}{y_n!} \right]. \quad (2.4)$$

Таким образом, оценка искомого изображения $\hat{\mathbf{x}}_{i+1}$, полученная на i -й итерации алгоритма Ричардсона-Люси ($1 \leq i \leq I$), выражена как

$$\hat{\mathbf{x}}_{i+1} = \mathbf{h} *^\top \frac{\mathbf{y}}{\mathbf{h} * \mathbf{x}_i} \odot \mathbf{x}_i. \quad (2.5)$$

В Формуле 2.5 символ $*^\top$ обозначает так называемую транспонированную свёртку, символ \odot обозначает поэлементное перемножение и процедура деления также осуществляется поэлементно. Транспонированная свёртка получается путём свёртки с пространственно отражённой версией вектора \mathbf{h} . Рассматривая двумерное представление вектора \mathbf{h} из Раздела 1.1 Главы 1, его отражённую версию можно выразить, как $h_{k_x, k_y}^{\text{rev}} = h_{k_y, k_x}$. Транспонированная свёртка также может быть представлена с помощью матрично-векторного произведения $\mathbf{H}^\top \mathbf{x}$. Слепая деконволюция на основе алгоритма Ричардсона-Люси также является итеративной и предполагает наличие двух шагов на каждой своей итерации – на первом шаге задача деконволюции решена для функции рассеяния точки, а на втором шаге – для изображения. Таким образом, на t -й итерации слепого алгоритма ($1 \leq t \leq T$) известна оценка искомого изображения, полученная на $(t - 1)$ -й итерации за I шагов схемы Ричардсона-Люси 2.5, то есть существует $\hat{\mathbf{x}}_I^{t-1}$. Тогда функция рассеяния точки на t -й итерации также может быть найдена с помощью I итераций алгоритма Ричардсона-Люси как

$$\hat{\mathbf{h}}_{i+1}^t = \mathbf{x}_I^{t-1} *^\top \frac{\mathbf{y}}{\mathbf{h}_i^t * \mathbf{x}_I^{t-1}} \odot \mathbf{h}_i^t, \quad (2.6)$$

и оценка искомого изображения на итерации t в свою очередь может быть получена с использованием вычисленной PSF. Общая итеративная схема алгоритма слепой деконволюции имеет вид

$$\hat{\mathbf{h}}_{i+1}^t = \mathbf{x}_I^{t-1} *^\top \frac{\mathbf{y}}{\mathbf{h}_i^t * \mathbf{x}_I^{t-1}} \odot \mathbf{h}_i^t, \quad (2.7a)$$

$$\hat{\mathbf{x}}_{i+1}^t = \mathbf{h}_I^t *^\top \frac{\mathbf{y}}{\mathbf{h}_I^t * \mathbf{x}_i^t} \odot \mathbf{x}_i^t. \quad (2.7b)$$

Итеративный процесс слепой деконволюции продолжается до тех пор, пока не будет выполнен некоторый выбранный критерий останова.

2.3.3 Алгоритмы обучения с подкреплением

Обучение с подкреплением (reinforcement learning, RL) – это относительно новая технология машинного обучения, в которой так называемый агент обучается принимать последовательные решения и взаимодействовать со средой, наблюдая её состояния s_t , таким образом, чтобы максимизировать суммарную награду, полученную в результате выполнения того или иного действия a_t на t -м шаге эпизода. В то время как ранние работы в области RL были сосредоточены на решении задач роботизированного контроля [93], проблемы восстановления изображений также были рассмотрены специалистами в этой области. В работе [94] предложена рекуррентная нейронная сеть, обученная с использованием методов обучения с подкреплением и применяемая для классификации изображений. Затем эта модель была улучшена в работе [95]. Для коррекции видеозаписей, в частности удаления некорректных кадров и распознавания лиц, авторами работы [96] также было предложено использование метода обучения с подкреплением. Задача повышения чёткости изображений была рассмотрена авторами работ [97] и [98]. Набор инструментов для улучшения изображения и алгоритм их выбора с помощью обучения с подкреплением были предложены в работе [99]. В работе [100] был предложен RL-метод, использующий карты действий и наград, соответствующих каждому пикселю изображения, для устранения шума в изображении. Позже эта идея была применена авторами [101] для задачи устранения размытия. Наконец, авторы работы [102] сформулировали проблему шумоподавления на изображении как частично наблюдаемый марковский процесс принятия решений, где действия агента определены как совокупность операций, применённых к каждому пикселю.

В настоящей работе применены два распространённых алгоритма обучения с подкреплением. Первый носит название *глубокого Q-обучения (deep Q-learning, DQN)*, где Q является так называемой функцией полезности, определяющей, насколько выигрышным в долгосрочной перспективе будет соверше-

ние действия a_t при нахождении агента в состоянии s_t . Классическое Q-обучение определяет функцию полезности следующим образом [103]:

$$Q^{\text{new}}(s_t, a_t) \leftarrow Q(s_t, a_t) + \beta \cdot \left(r_t + \nu \max_a Q(s_{t+1}, a) - Q(s_t, a) \right), \quad (2.8)$$

где β – скорость обучения алгоритма, а ν – коэффициент, гарантирующий, что награде, полученной на ранних этапах, отдано предпочтение. В ходе алгоритма производится заполнение Q-таблицы, то есть для каждого сочетания состояний s_t и действий a_t вычисляется своя функция полезности – ожидаемая максимальная награда, полученная агентом за все дальнейшие шаги, если в данном состоянии s_t будет выполнено конкретное действие a_t . Изначально значения в таблице инициализированы случайным образом, но в ходе выполнения алгоритма они обновляются согласно Формуле 2.8. В дальнейшем агент при нахождении в состоянии s_t может как совершить случайно выбранное действие, так и обратиться к таблице с целью выбора действия a_t , максимизирующего значение Q. Такая дилемма называется компромиссом между исследованием неизученных областей и применением имеющихся знаний или *exploration-exploitation trade-off*. Во время тренировки агента приоритет отдаётся процессу исследования, в то время как во время тестирования агент стремится применить полученные знания на практике [103; 104].

Главным отличием глубокого Q-обучения от обычного Q-обучения является то, что функция полезности для всех комбинаций действий и состояний аппроксимирована свёрточной нейронной сетью. Такой подход особенно полезен в случае высокой размерности Q-таблицы. В алгоритме DQN используются две нейронные сети, одна из которых предсказывает целевые значения Q-функции, а вторая – фактически применяемые. Алгоритм обучается посредством “воспроизведения опыта” (*experience replay*), то есть для обучения используется случайная выборка предыдущих действий вместо самого последнего действия – считается, что таким образом данные в текущем наборе на входе алгоритма будут менее коррелированы [104].

Другой алгоритм обучения с подкреплением, использованный в настоящей работе, носит название *proximal policy optimization, PPO* – проксимальная оптимизация политики [105]. Основное отличие этого метода от DQN состоит в том, что, в отличие от последнего, целью здесь является поиск оптимальной политики, то есть определённого правила или стратегии, согласно которой

агент будет совершать те или иные действия, приводящие к максимизации награды в долгосрочной перспективе. В общем случае методы, ориентированные на политику, и PPO в частности, производят поиск в пространстве политик, а не ищут значение функции полезности. Применяя этот термин к алгоритму DQN, можно сказать, что оптимальная стратегия состоит в том, чтобы сообщить текущее состояние s_t функции Q и выбрать действие a_t , гарантирующее максимальное значение Q . Таким образом, алгоритм PPO производит обновление параметров θ функции π_θ , определяющей политику, согласно которой агент принимает решения о действиях. В частности для обновления параметров используется алгоритм стохастического градиентного подъёма и, согласно одному из вариантов, описанных в работе авторов метода, целевая функция имеет вид

$$L^{CLIP}(\theta) = \hat{\mathbb{E}}_t \left[\min \left(\frac{\pi_\theta(a_t|s_t)}{\pi_{\theta_{\text{old}}}(a_t|s_t)} \hat{A}_t, \text{clip}(\pi_{\theta_{\text{old}}}(a_t|s_t), 1 - \varepsilon, 1 + \varepsilon) \hat{A}_t \right) \right]. \quad (2.9)$$

Здесь \hat{A}_t – функция преимущества, которая оценивает, насколько выгодным оказалось выполнение действия a_t в состоянии s_t по сравнению с другими действиями, а $\varepsilon = 0.2$ – ограничивающий степень отличия параметров старой политики от новой фактор. В целом исследование в алгоритме PPO осуществляется путём выбора действий в соответствии с последней версией стохастической политики. В ходе обучения стратегия обычно становится всё менее случайной.

2.3.4 Применение тренируемого агента для поиска параметров математической модели функции рассеяния точки

В настоящей работе предложено применение метода обучения с подкреплением для поиска оптимальных параметров модели, аппроксимирующей двумерную функцию рассеяния точки оптической системы, с помощью которой осуществляется восстановление искомого изображения. Для реализации предложенной схемы использован алгоритм слепой деконволюции, описанный в Разделе 2.3.2, на первом шаге каждой итерации которого происходит оценка PSF, а на втором – оценка изображения. В предлагаемом методе функция рассеяния точки аппроксимирована известной моделью, параметры которой обновляются

на каждом шаге итеративной схемы в соответствии с найденными с помощью тренируемого агента значениями. В общем виде схема объединения метода обучения с подкреплением и итеративной схемы слепой деконволюции, основанной на алгоритме Ричардсона-Люси, представлена на Рисунке 2.5.

Аппроксимация функции рассеяния точки 2D функцией Гаусса.

На первом этапе работы применена одна из самых распространённых математических моделей, используемых для аппроксимации функции рассеяния точки в микроскопии, – функция Гаусса. Главный параметр, определяющий ширину пятна размытия, – стандартное отклонение σ . Поиск значения этого параметра осуществляется с помощью тренируемого агента. Выражая вектор функции рассеяния точки $\mathbf{h} \in \mathbb{R}^{K^2}$ в двумерном виде h_{k_x, k_y} , $1 \leq k_x \leq K$, $1 \leq k_y \leq K$, и записывая его через формулу двумерной функции Гаусса, двухэтапная оптимизационная схема Ричардсона-Люси из выражения 2.7 трансформируется в схему вида

$$\hat{\mathbf{h}}^t = h_{k_x, k_y}, \quad 1 \leq k_x \leq K, \quad 1 \leq k_y \leq K \quad (2.10a)$$

$$h_{k_x, k_y}(\sigma_t) = \frac{1}{\sigma_t \sqrt{2\pi}} e^{-\frac{((k_x - k_{x_0})^2 + (k_y - k_{y_0})^2)}{2\sigma_t^2}} \quad (2.10b)$$

$$\hat{\mathbf{x}}_{i+1}^t = \mathbf{h}^t * \top \frac{\mathbf{y}}{\mathbf{h}^t * \mathbf{x}_i^t} \odot \mathbf{x}_i^t. \quad (2.10c)$$

Здесь k_{x_0} и k_{y_0} – координаты центра PSF. В Формуле 2.10b для удобства записи двумерной функции Гаусса использована двумерная форма представления вектора \mathbf{h} . В Формуле 2.10c использовано прежнее обозначение $\mathbf{h} \in \mathbb{R}^{K^2}$. Для

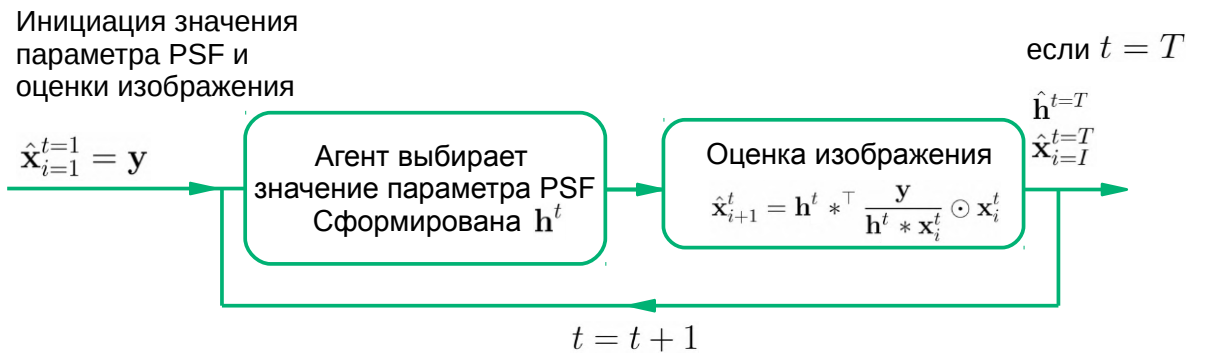


Рисунок 2.5 — Итеративная схема алгоритма слепой деконволюции, объединяющего метод обучения с подкреплением для поиска параметров PSF и алгоритм Ричардсона-Люси для оценки искомого изображения.

тренировки и оценки работы алгоритма использовано одно чёрно-белое изображение размером 256×256 из набора эталонных изображений, описанного в Разделе 1.3.4. Для проверки способности тренируемого агента находить оптимальное значение параметра σ задан диапазон значений поиска $[3.4, \dots, 6.0]$ с шагом 0.1. Процесс обучения агента осуществлён следующим образом. В начале каждого эпизода искомое изображение подвергается размытию гауссовой PSF с σ_{true} , случайным образом выбранной из указанного диапазона. Стартовое значение $\sigma_{t=1}$, с которого агент начинает свой поиск, выбрано из того же диапазона случайным образом. На каждом шаге t эпизода агент, тренируемый в данном случае с помощью алгоритма PPO, согласно выбранной политике совершает действие a_t и выбирает значение σ_t из диапазона $[3.4, \dots, 6.0]$. Действие a_t может быть определено значением 0 или +1, где 0 соответствует выбору предыдущего значения из диапазона, а +1 – следующего. Далее с помощью PSF, сгенерированной с выбранной σ_t , за $I = 3$ итерации алгоритма Ричардсона-Люси производится оценка искомого изображения $\hat{\mathbf{x}}_I^t$. В конце каждого шага t восстановленное изображение с помощью метрики отношения пикового сигнала к шуму (PSNR) сравнивается с искомым изображением. Значение указанной метрики служит наградой агенту как

$$r_t = (\text{PSNR}_t - \text{PSNR}_{t-1}) \cdot \psi. \quad (2.11)$$

Здесь ψ – коэффициент масштабирования награды. Эмпирическим путём было определено, что оптимальное значение $\psi = 10$. Таким образом, чем ближе восстановленное на t шаге эпизода изображение к искомому, чем на шаге $t - 1$, тем корректнее агент выбрал значение параметра PSF σ_t и тем выше награду он получает. Кроме награды агенту необходимо наблюдать текущее состояние среды, чтобы соотнести полученную награду с результатами его действий. На данном этапе состояние среды, передаваемое агенту, представляет собой восстановленное изображение $s_t = \{\hat{\mathbf{x}}_I^t\}$. Выполнение эпизода заканчивается при достижении количества шагов $t = 40$. Затем значения всех параметров снова выбираются случайным образом, генерируется новое искажённое изображение \mathbf{u} , и начинается новый эпизод обучения агента. Кроме того, на агента накладываются штрафы за некоторые нежелательные действия. В частности, в случае попытки выхода из заданного диапазона, агент получает отрицательную награду $r_t = -5$. Результаты восстановления изображения и функции размытия

точки в ходе выполнения алгоритма представлены на Рисунке 2.6. Из представленных результатов видно, что предложенная схема позволяет алгоритму сходиться к истинному значению σ_{true} .

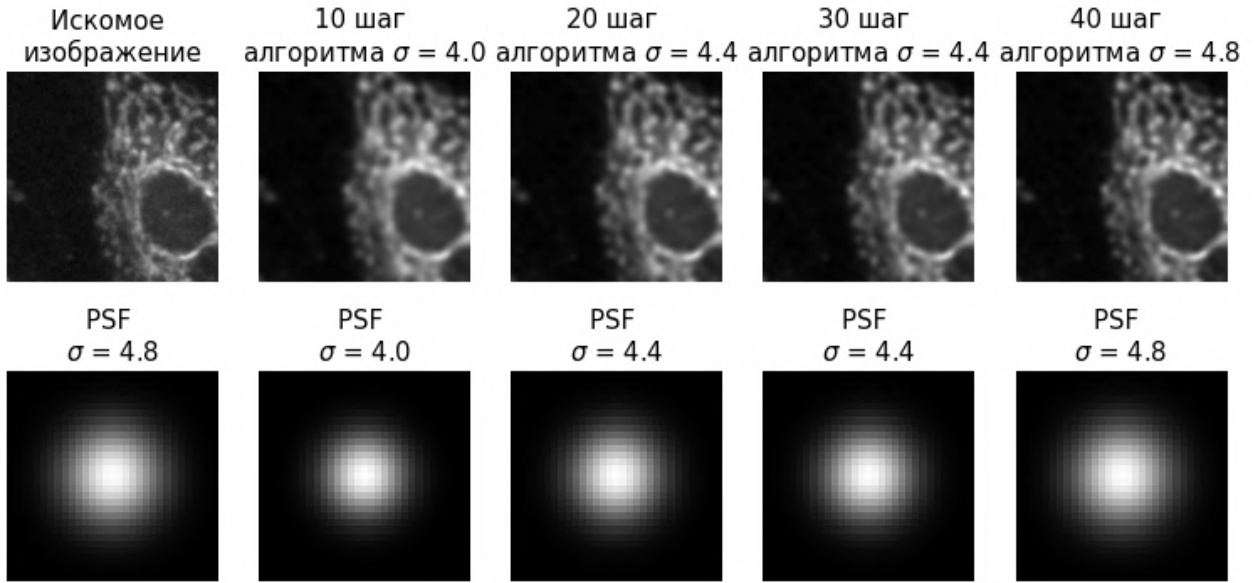
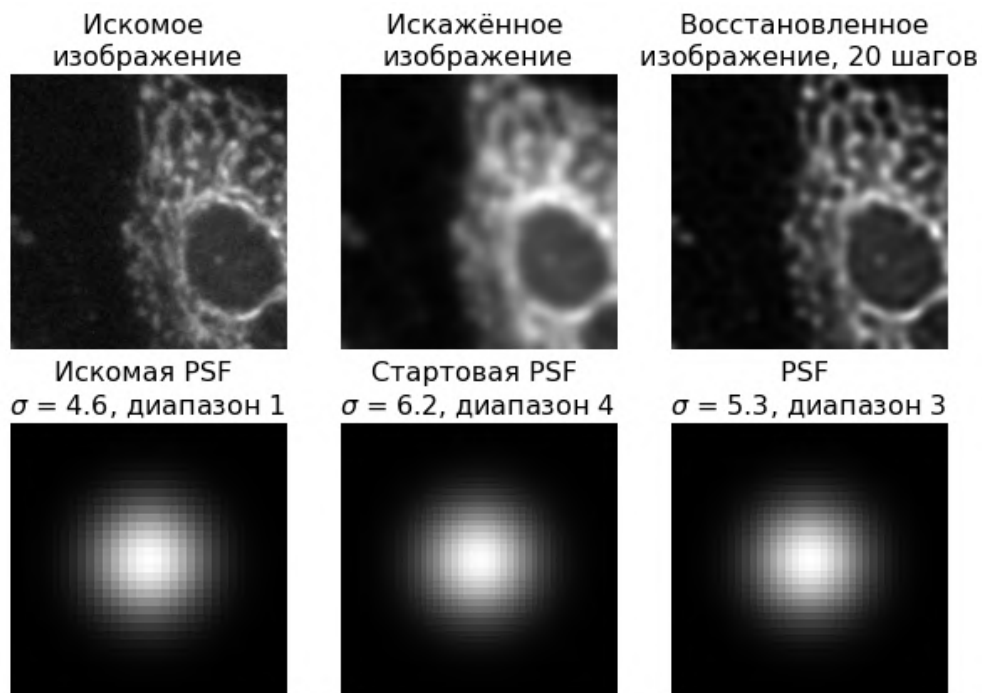


Рисунок 2.6 — Изменение восстановленной PSF (нижняя строка) и соответствующего ей изображения (верхняя строка) в зависимости от шага итеративного алгоритма слепой деконволюции. Одиночный агент осуществляет поиск параметра σ .

Следующим этапом алгоритма было применение иерархической схемы обучения с подкреплением, в которой глобальная проблема разбита на подзадачи, каждая из которых может быть выполнена отдельным агентом. В данном случае диапазон поиска значения $\sigma = [3.5, \dots, 7.0]$ с шагом 0.1 разбит на мини-диапазоны по 6 значений в каждом. Так называемый старший агент с помощью действий $a_t = \{0, +1\}$ осуществляет выбор мини-диапазона, внутри которого младший агент также с помощью действий $a_t = \{0, +1\}$ производит выбор значения σ . Оба агента обучены с помощью алгоритма PPO. Общее количество шагов в эпизоде, равное максимальному количеству итераций алгоритма слепой деконволюции, выбрано равным $t = 150$. В начале эпизода старший агент выбирает мини-диапазон, внутри которого на протяжении $t = 5$ шагов младший агент осуществляет поиск значения параметра σ . По истечении 5 шагов младшего агента старший агент выбирает новый диапазон. Эпизод считается законченным либо при достижении $t = 150$, либо в случае выполнения условия $\text{PSNR}_t - \text{PSNR}_1 > 0.1 \cdot \text{PSNR}_1$, где PSNR_1 является метрикой близости размытого изображения \mathbf{y} к искомому изображению \mathbf{x} . В ходе обучения используется

та же система штрафов, наград и состояний среды для младшего агента, как и в случае обучения одиночного агента. Старшему же агенту в качестве состояния среды и награды сообщаются значения, полученные младшим агентом на последнем – пятом – шаге внутри выбранного мини-диапазона. По результатам восстановления изображения, представленным на Рисунке 2.7, видно, что предложенная иерархическая схема работает хуже, чем схема поиска значения σ , основанная на одиночном агенте. В ходе эксперимента было выявлено, что, несмотря на то, что младший агент стремится выбрать ближайшее к σ_{true} значение параметра, старший агент ошибается в выборе мини-диапазона.



Слева направо: искомое изображение и PSF, стартовые и восстановленные с помощью итеративного алгоритма изображение и PSF. На изображениях PSF указаны номера диапазонов, выбранные старшим агентом, и значения σ , выбранные младшим агентом.

Рисунок 2.7 — Результаты восстановления искомого изображения и PSF с помощью алгоритма слепой деконволюции и иерархического алгоритма обучения с подкреплением для поиска параметра σ .

Аппроксимация функции рассеяния точки моделью Борна и Вольфа. В то время как двумерная функция Гаусса во многих случаях является удовлетворительной аппроксимацией функции размытия точки в микроскопии, она не учитывает концентрические кольца узора Эйри. В случае если регистрируемый объект расположен не в фокальной плоскости, искажения, вноси-

мые в результирующее изображение оптической системой, могут быть описаны трёхмерной функцией рассеяния точки, а именно её срезами вдоль оптической оси. Математическая модель Борна и Вольфа описывает трёхмерную функцию рассеяния точки.

Модель Борна и Вольфа состоит из большого количества параметров, осуществить последовательный поиск которых может быть сложной с вычислительной точки зрения задачей. Параметрами, непосредственно влияющими на форму PSF, являются:

- λ - длина волны излучаемого света;
- r - радиус в плоскости изображения, то есть размер PSF;
- NA - числовая апертура объектива;
- δz - параметр, определяющий положение двумерного среза трёхмерной PSF относительно фокальной плоскости.

Напомним, что эти параметры представлены в модели Борна и Вольфа как

$$h(r; \delta z) = \left| C \int_0^1 J_0 \left(\frac{2\pi}{\lambda} r \text{NA} \rho \right) \exp \left[\frac{-i \frac{2\pi}{\lambda} \delta z \text{NA}^2 \rho^2}{2n_i} \right] \rho d\rho \right|^2. \quad (2.12)$$

Существует ещё два параметра, не фигурирующих непосредственным образом в Формуле 2.12, но тем не менее влияющих на размеры и форму функции размытия точки – этими параметрами являются поперечная и осевая разрешающая способность микроскопа, здесь и далее обозначаемые Δ_x и Δ_z соответственно. Эти параметры определяют частоту выборки измерений в поперечном и осевом направлениях и связаны с радиусом в плоскости изображения r и параметром δz соответственно [106]. Однако в соответствии с частотой дискретизации Найквиста оптимальные значения этих параметров могут быть определены как (для случая широкопольной микроскопии) [107–110]

$$\Delta_x = \frac{\lambda}{4n_i \sin(\alpha)} \quad (2.13a)$$

$$\Delta_z = \frac{\lambda}{2n_i(1 - \cos(\alpha))}. \quad (2.13b)$$

В выражениях 2.13 $\alpha = \arcsin(\text{NA} / n_i)$ [107–110]. Таким образом, параметр r зависит от числовой апертуры NA и поперечных размеров PSF k_x, k_y , которые в данной работе выбираются вручную, $r = \sqrt{(k_x - k_{x_0})^2 + (k_y - k_{y_0})^2} \cdot \Delta_x, k_{x_0}$

и k_{y_0} – координаты центра PSF. Параметр δz также зависит от NA и номера среза PSF d , $\delta z = \Delta_z \cdot d$. Отметим, что $d = 0$ соответствует центральному срезу функции рассеяния точки. Иммерсионный показатель преломления обычно принимается равным 1.515 для наиболее часто используемого для иммерсии в оптической микроскопии вещества – масла [111]. Наконец, экспериментальным образом было выявлено, что при использовании выражений 2.12 и 2.13 для формирования функции рассеяния точки, длина волны излучаемого света λ не оказывает влияния на форму PSF. Таким образом, для поиска были выбраны два параметра – числовая апертура NA и номер среза PSF d . Используя аналогичный Формуле 2.10 подход и выражая радиус $r = \sqrt{(k_x - k_{x_0})^2 + (k_y - k_{y_0})^2} \cdot \Delta_x$, выражения для алгоритма слепой деконволюции можно записать как

$$\hat{\mathbf{h}}^t = h_{k_x, k_y}, \quad 1 \leq k_x \leq K, \quad 1 \leq k_y \leq K \quad (2.14a)$$

$$h_{k_x, k_y}(d_t, \text{NA}_t) = \left| \int_0^1 J_0(k_0 r \text{NA}_t \rho) \exp \left[\frac{-ik_0 \Delta_z d_t \text{NA}_t^2 \rho^2}{2n_i} \right] \rho d\rho \right|^2 \quad (2.14b)$$

$$\hat{\mathbf{x}}_{i+1}^t = \mathbf{h}^t *^\top \frac{\mathbf{y}}{\mathbf{h}^t * \mathbf{x}_i^t} \odot \mathbf{x}_i^t. \quad (2.14c)$$

Для восстановления формы PSF с помощью модели Борна и Вольфа использована иерархическая схема обучения с подкреплением. Старший агент осуществляет поиск параметра σ гауссовой функции и передаёт информацию о значении параметра младшему агенту, а тот в свою очередь уточняет форму PSF с помощью значения числовой апертуры NA.

Эксперименты с использованием итеративного алгоритма слепой деконволюции с аппроксимацией PSF функцией Борна и Вольфа реализованы в нескольких вариантах. В первом варианте агенты обучены восстанавливать различные PSF для реконструкции одного и того же изображения. В среде заданы диапазоны поиска значений NA, d и σ . Экспериментальным путём было выявлено, что параметры σ и d связаны приблизительно как $d = \sigma/2$. Старший агент выполняет поиск параметра σ двумерной гауссовой функции и через зависимость этих двух параметров осуществляет оценку параметра d . Затем младший агент уточняет форму функции размытия точки путём поиска значения параметра NA. В качестве состояний агентам передаётся вектор $s_t = \{\hat{\mathbf{x}}^t, \hat{\mathbf{h}}^t, \text{NA}_t, d_t\}$. Действия, совершаемые агентами, заданы дискретными значениями $a_t = \{-1, 0, +1\}$, соответствующими выбору предыдущего по вели-

чине значения из диапазона, пропуску действия, то есть выбору того же значения, что и на предыдущем шаге, и выбору следующего по величине значения. Награда, которую получает младший агент за совершаемое действие, определена в соответствии с Формулой 2.11. За выход из заданных диапазонов оба агента получают отрицательную награду, равную $r_t = -5$. Максимальное количество итераций в эпизоде ограничено числом 150, эпизод заканчивается либо при достижении заданного количества итераций T , либо при выполнении условия $\text{PSNR}_t - \text{PSNR}_1 > 0.1 \cdot \text{PSNR}_1$, где PSNR_1 является метрикой близости размытого изображения \mathbf{y} к искомому изображению \mathbf{x} . Для реализации описанного эксперимента выбрано два алгоритма обучения с подкреплением – PPO и DQN, а также метод слепой деконволюции, основанный на схеме Ричардсона-Люси и выраженный в соответствии с Формулой 2.7. Для последнего метода количество итераций алгоритма Ричардсона-Люси выбрано равным 3, а число итераций алгоритма слепой деконволюции – 150. Сохранены изображения с наибольшим значением метрики PSNR. Результаты восстановления одного изображения и различных PSF с помощью иерархического метода обучения с подкреплением, реализованного с помощью алгоритмов PPO и глубокого Q-обучения, и метода слепой деконволюции представлены на Рисунке 2.8. Из результатов видно, что, будучи применённым к тому же изображению, на котором алгоритм иерархического обучения с подкреплением был натренирован, он позволяет достаточно точно восстанавливать форму искомой функции размытия точки и, следовательно, искомое изображение. При применении же алгоритма к новому для агентов изображению видно, что они не способны восстановить истинную форму PSF и изображение. В случае применения метода слепой деконволюции восстановленное изображение имеет более низкие значения PSNR по сравнению с результатами реконструкции с использованием обучения с подкреплением. В то же время восстановленная функция рассеяния точки не соответствует форме искомой PSF.

Следующим этапом является обучение агентов на тренировочном наборе данных, состоящем из 64 изображений из датасета, описанного в Разделе 1.3.4 Главы 1, с последующей оценкой алгоритма на отложенных тестовых данных. Для реализации этого подхода в начале каждого эпизода искажению подвергается случайно выбранное изображение из 64 тренировочных. Результаты восстановления различных тестовых изображений и различных PSF с помощью

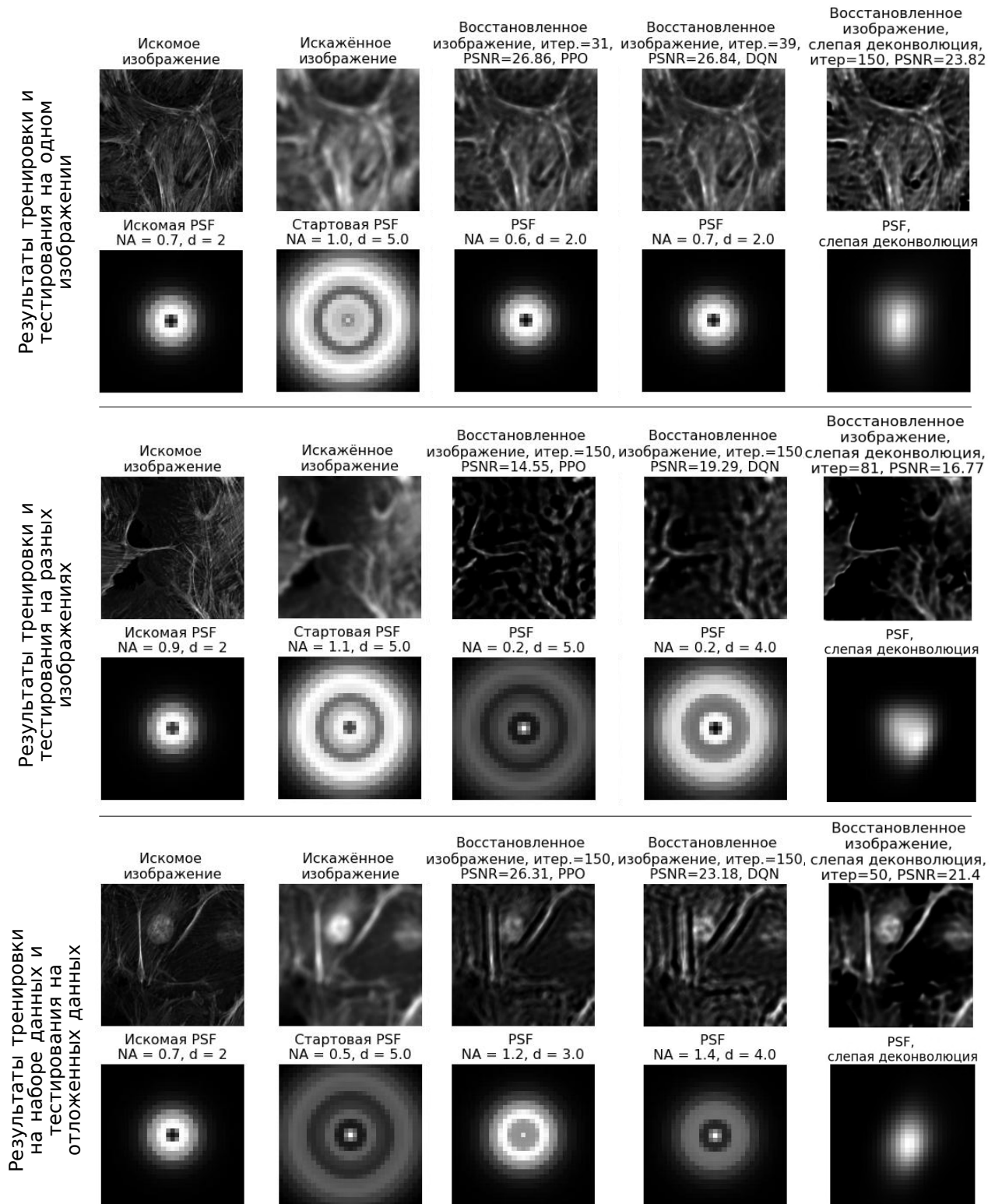


Рисунок 2.8 — Результаты восстановления искомого изображения и PSF с помощью алгоритма слепой деконволюции и иерархического алгоритма обучения с подкреплением для поиска параметров d и NA. Вектор состояний

$$s_t = \{\hat{\mathbf{x}}^t, \hat{\mathbf{h}}^t, \text{NA}_t, d_t\}.$$

иерархического метода обучения с подкреплением, реализованного с помощью алгоритмов PPO и глубокого Q-обучения, и с помощью метода слепой деконволюции представлены на Рисунке 2.8. Как видно из результатов, агенты плохо адаптируются к различным изображениям, передаваемым им в составе вектора состояний s_t . Это выражается в том, что даже при достижении максимального количества шагов в эпизоде не удаётся восстановить форму PSF, а вместе с ней и изображение. Для преодоления этой проблемы следующим шагом предложено использование другого вектора состояний, передаваемого агентам. В этом случае набор действий, наград и штрафов остаётся прежним, а вектор состояний представлен $s_t = \{[a_{t-p+1}, \dots, a_t], [r_{t-p+1}, \dots, r_t]\}$, то есть каждый агент получает вектор из p последних действий и соответствующих им наград, полученных агентом. Таким образом агент получает информацию не о самих изображениях, которые могут сильно отличаться, но о результате “принятых” им решений. На двух верхних строках Рисунка 2.9 представлены результаты восстановления изображения и функции размытия точки иерархическими агентами, обученными на одном изображении с помощью алгоритмов PPO и DQN, а также методом слепой деконволюции (Формула 2.7). На двух нижних строках Рисунка 2.9 отображены результаты восстановления изображения и функции размытия точки иерархическими агентами, обученными на тренировочном наборе изображений, и методом слепой деконволюции, применёнными к отложенным тестовым изображениям.

Из результатов, представленных на Рисунке 2.9, следует, что агенты способны адаптироваться к задаче восстановления PSF для различных изображений в случае, если информация о состоянии среды не зависит от изображений. В случае применения метода слепой деконволюции восстановленное изображение имеет более низкие значения PSNR по сравнению с результатами реконструкции с использованием обучения с подкреплением. В то же время восстановленная функция рассеяния точки не соответствует форме искомой PSF. Интересно, что агенты, обученные с помощью метода глубокого Q-обучения, восстанавливают искомую функцию рассеяния точки за меньшее количество шагов t , чем обученные с помощью PPO. Сформулированный таким образом алгоритм слепой деконволюции способен генерализоваться на различные изображения.

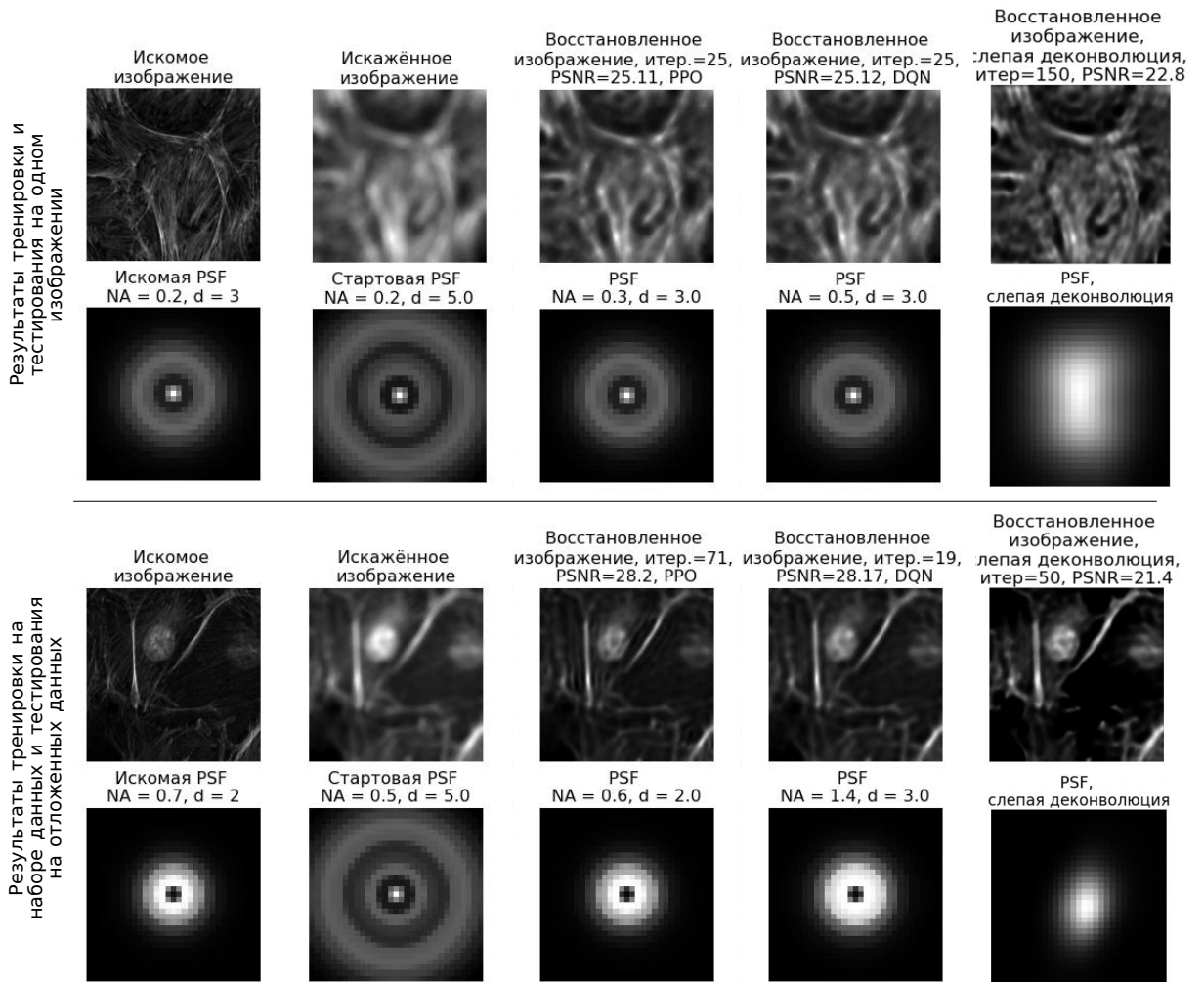


Рисунок 2.9 — Результаты восстановления искомого изображения и PSF с помощью алгоритма слепой деконволюции и иерархического алгоритма обучения с подкреплением для поиска параметров d и NA. Вектор состояний

$$s_t = \{[a_{t-p+1}, \dots, a_t], [r_{t-p+1}, \dots, r_t]\}.$$

2.4 Заключение главы

В настоящей работе рассмотрено два подхода к решению задачи восстановления функции рассеяния точки микроскопа. Предложено две схемы для деконволюции функции рассеяния точки, объединяющие классический аппарат математического моделирования для оценки неизвестной функции и современные методы глубокого обучения для оптимизации процесса поиска решения путём использования доступных тренировочных данных.

Первый метод предлагает решение задачи поиска оптимальной функции размытия точки с помощью неконтролируемого фильтра Винера, выраженного

из теоремы Байеса. Для осуществления такого решения использованы непосредственно искажённые изображения и сгенерированные из них с помощью генеративно-состязательной сети псевдо-эталонные изображения. Полученные результаты показывают применимость предложенной схемы к задаче восстановления функции рассеяния точки оптической системы, а также к задаче слепой деконволюции.

Вторая схема восстановления неизвестной функции размытия точки реализует итеративный алгоритм попеременной оценки PSF и искомого изображения, основанный на методе Ричардсона-Люси. Отличительной особенностью предложенного алгоритма является использование метода обучения с подкреплением для поиска оптимальных параметров математической модели, аппроксимирующей функцию рассеяния точки. Предложенная идея основана на предположении о том, что иерархическая схема обучения агента позволит повысить эффективность поиска неизвестных параметров по сравнению с последовательным поиском значения каждого параметра. Полученные экспериментальные результаты показывают, что при использовании иерархической схемы обучения, а также при оптимальном сочетании вектора состояний и функции награды, передаваемых агентам, этот подход позволяет получить многообещающие результаты для одновременной оценки PSF и изображения. Разработанный алгоритм позволяет осуществлять восстановление различных изображений и форм функции рассеяния точки. Дальнейшие исследования в этой области планируется направить на сравнение предложенного подхода с использованием обучения с подкреплением с классическими методами глубокого обучения, а также их возможное объединение. Кроме того, в дальнейшем предлагается увеличение количества параметров математической модели, вычисляемых с помощью тренируемых агентов.

Глава 3. Восстановление гиперспектральных изображений, полученных с помощью однопиксельной визуализации

3.1 Однопиксельная визуализация

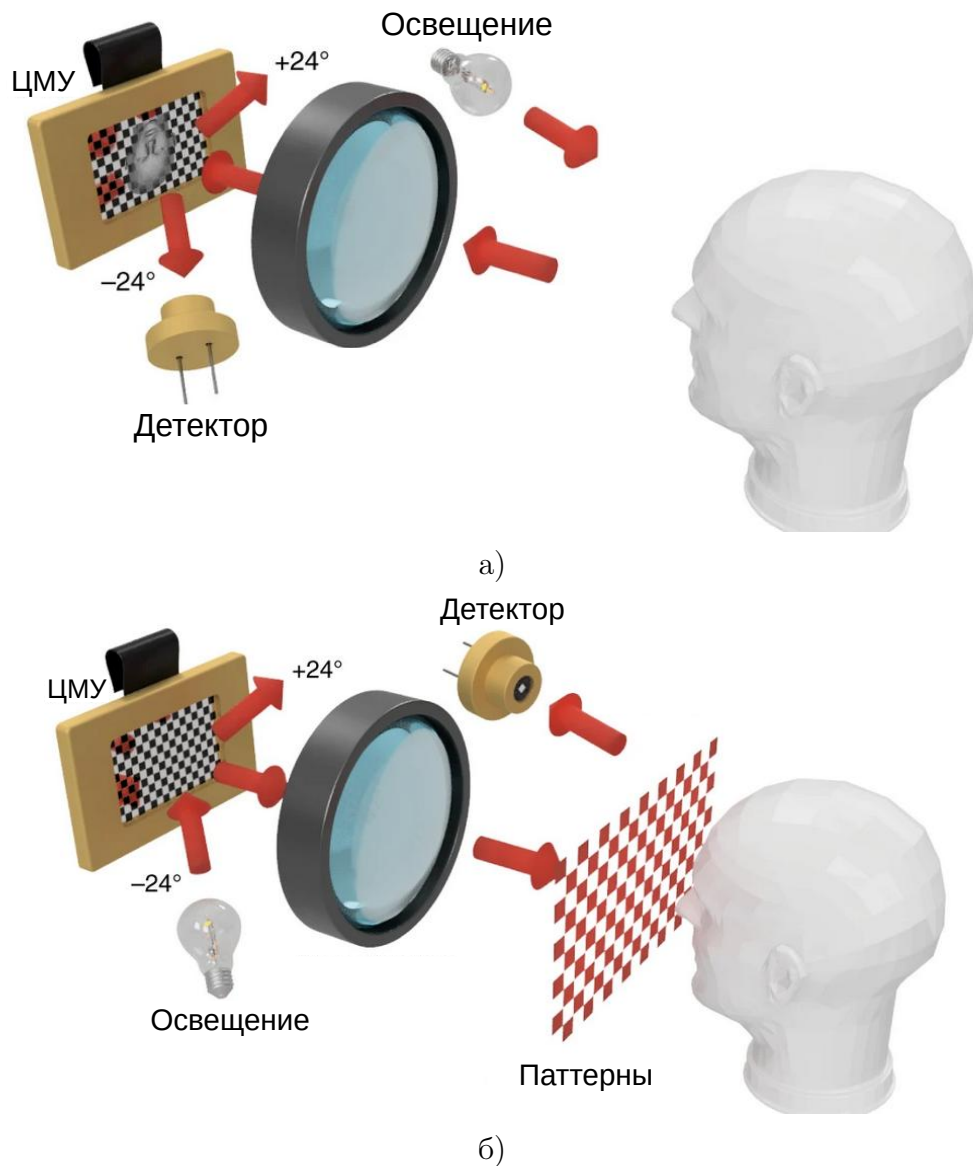
Однопиксельная визуализация – это метод, позволяющий восстанавливать изображение из сигнала, полученного с помощью *однопиксельного датчика*. Концепция однопиксельной визуализации была впервые описана в 2008 году в работе [112] и представляла собой развитие идеи *compressive sensing* или опознания по сжатию [113]. Формирование этой идеи произошло вследствие того, что с развитием технологий число пикселей в датчиках камер увеличивалось, что в свою очередь привело к необходимости располагать значительными вычислительными мощностями для хранения и обработки данных. Достоинством однопиксельной визуализации является возможность восстанавливать изображение, используя меньшее, чем число неизвестных пикселей, количество измерений [8].

Наиболее простым способом получения однопиксельного измерения является последовательная регистрация интенсивности каждого пикселя искомого изображения. Недостатком такого метода является его длительность, а также нерациональность использования освещения [8]. Более рациональным и чаще используемым подходом является применение последовательности пространственно разрешённых паттернов и измерение результата корреляции этих паттернов и искомого изображения объекта. Существует два основных способа получения однопиксельного измерения: *structured detection* (структурированная регистрация) and *structural illumination* (структурированное освещение) [8]. В случае структурированной регистрации устройство, модулирующее световой поток и осуществляющее формирование паттернов, расположено в плоскости формирования изображения. Источник света освещает объект, далее световой поток проходит через линзу, формирующую изображение, и попадает на модулирующее устройство. Затем полученный в результате корреляции паттернов с изображением световой поток фокусируется промежуточной линзой и регистрируется однопиксельным датчиком. Схема этого процесса отображена на Рисунке 3.1а.

В случае структурированного освещения применяется модуляция светового потока, попадающего на объект. Результат корреляции паттернов и объекта затем также фокусируется промежуточной линзой и регистрируется однопиксельным датчиком. Схема этого процесса представлена на Рисунке 3.1б. В обоих случаях восстановление изображения из измерений осуществляется путём перемножения каждого паттерна из набора с соответствующей ему величиной зарегистрированного с помощью однопиксельного датчика измерения. В результате этой операции получается набор взвешенных паттернов, которые затем суммируются для получения восстановленного изображения [8]. Для корректной реконструкции изображения $\mathbf{x} \in \mathbb{R}^N$ необходимо, чтобы количество используемых паттернов было равно общему количеству пикселей в изображении, то есть $P = N$. Однако, в случае, если используемые паттерны не являются ортогональными или в измерениях присутствует шум, потребуется значительно большее количество паттернов $P \gg N$ [8].

Для модуляции светового потока используются различные технологии, включая псевдотермальное излучение [115], жидкокристаллические пространственные модуляторы света [8; 116; 117], светодиоды [118], а также цифровые микрозеркальные устройства [119]. Последние представляют собой матрицу микроскопических зеркал на полупроводниковом чипе. Эти микрозеркала имеют индивидуальную ориентацию $\pm 12^\circ$ относительно плоскости чипа, соответствующую положению зеркала “ВКЛ” и “ВЫКЛ” соответственно. Вследствие этого световой поток, попадающий на чип, перенаправлен в двух направлениях, одно из которых, соответствующее положению “ВКЛ”, принимается за значение бинарного паттерна $+1$, а другое -0 . Использование светового потока может быть оптимизировано путём регистрации потока в обоих направлениях ориентации микрозеркал. В таком случае целесообразно использование двух однопиксельных детекторов с последующим измерением разности между зарегистрированными с их помощью сигналами. Однако чаще всего используется один детектор, и в таком случае вслед за текущими паттернами проецируется их контрастная инверсия [8].

Паттерны, используемые для модуляции светового потока, могут быть представлены случайными значениями в шкале оттенков серого цвета [117]. Другой подход предлагает использование Фурье-кодирования паттернов: с помощью микрозеркал осуществляется формирование синусоидальных паттернов



На Рисунке 3.1а изображена схема получения однопиксельного измерения с помощью структурированной регистрации, когда изображение объекта проецируется на цифровое микрзеркальное устройство (ЦМУ). На Рисунке 3.1б изображена схема получения однопиксельного измерения с помощью структурированного освещения, когда паттерны проецируются на объект.

Рисунок 3.1 — Схемы получения однопиксельного измерения [114].

с фазовым сдвигом относительно друг друга. Затем зарегистрированные результаты корреляции паттернов и изображения подвергаются преобразованию Фурье [8; 120].

В работе [112] было продемонстрировано использование бинарных паттернов Адамара для модуляции светового потока в однопиксельной визуализации. Паттерны Адамара являются ортогональными и имеют значения ± 1 в отли-

чие от описанных ранее значений $+1$ и 0 , которые могут быть реализованы на цифровом микрозеркальном устройстве. Таким образом, для регистрации измерения с использованием паттернов Адамара обычно применяется разностный подход. Однако такой метод требует вдвое большего количества паттернов, реализуемых с помощью микрозеркал [8].

3.2 Формирование однопиксельных измерений

Математическая модель, соответствующая процессу формирования однопиксельного измерения соответствует модели формирования измерений в задачах визуализации, описанной во введении (Формула 1), и может быть выражена как

$$\mathbf{m}^\alpha = \alpha \mathbf{H}_1 \mathbf{x}, \quad (3.1)$$

где $\mathbf{m}^\alpha \in \mathbb{R}^P$ – вектор однопиксельного измерения, $\mathbf{H}_1 \in \mathbb{R}^{P \times N}$, – матрица, состоящая из паттернов, сформированных с помощью модулирующего световой поток устройства, $\mathbf{x} \in \mathbb{R}^N$ – искомое изображение, α – неизвестная интенсивность изображения, выраженная в количестве фотонов, пропорциональном времени освещения объекта. Отметим, что в рамках настоящей работы рассмотрено $P < N$ паттернов, что обеспечивает более высокую скорость сбора данных и снижает требования к памяти. Бинарные паттерны, формирующие \mathbf{H}_1 , выбраны из ортогонального базиса $\mathbf{H} \in \mathbb{R}^{N \times N}$, и в данной работе рассмотрены паттерны Адамара. Они могут быть сгенерированы как заранее, так и непосредственно в процессе формирования измерения [121; 122].

На практике полученное однопиксельное измерение искажено пуассоновским и гауссовским шумами, первый из которых зависит от интенсивности сигнала и обусловлен дискретной природой носителей электрического заряда, а второй является аддитивным и возникает из-за флуктуаций в электрической сети и статистически не зависит от сигнала [39; 123]. Бинарные паттерны Адамара являются ортогональными и содержат значения $+1$ и -1 . Поскольку устройства, модулирующие световой поток, такие как цифровые микрозеркальные устройства, не способны воспроизводить отрицательные значения, возникает необходимость использования разностного подхода, описанного в Разделе 3.1.

С этой целью паттерны \mathbf{H} подвергаются разложению на два набора, содержащие только положительные и только отрицательные значения [124], как

$$\begin{aligned}\mathbf{H}^+ &= \max((\mathbf{H})_{n_x, n_y}, 0), \quad 1 \leq n_x, n_y \leq N \\ \mathbf{H}^- &= \max(-(\mathbf{H})_{n_x, n_y}, 0), \quad 1 \leq n_x, n_y \leq N.\end{aligned}\quad (3.2)$$

С учётом всего вышеизложенного вектор однопиксельных измерений описывается выражением

$$\mathbf{m}_{+,-}^\alpha \sim K\mathcal{P}(\alpha\mathbf{H}_1^{+,-}\mathbf{x}) + \mathcal{N}(\mu_{\text{dark}}, \sigma_{\text{dark}}^2). \quad (3.3)$$

Здесь символ \mathcal{P} обозначает распределение Пуассона, \mathbf{H}_1^+ и \mathbf{H}_1^- представляют собой положительную и отрицательную части \mathbf{H}_1 , выбранные из \mathbf{H}^+ и \mathbf{H}^- соответственно, K – калибровочный коэффициент передачи системы, $\mu_{\text{dark}}, \sigma_{\text{dark}}^2$ – математическое ожидание и дисперсия аддитивного гауссовского шума, появившегося вследствие воздействия темнового тока. В результате применения разностного подхода окончательный вектор однопиксельного измерения выражен как

$$\mathbf{m}^\alpha = \frac{\mathbf{m}_+^\alpha - \mathbf{m}_-^\alpha}{K\alpha}, \quad (3.4)$$

где нормировочный коэффициент $K\alpha$ использован для нормализации изображений в диапазоне значений $[0;1]$.

3.3 Гиперспектральные изображения

В случае применения спектрального детектора, способного регистрировать сигнал в диапазоне излучаемого света и разделять его на каналы, соответствующие поддиапазонам, концепция однопиксельной визуализации расширяется для применения в многоканальной – или гиперспектральной – визуализации. В частности гиперспектральная однопиксельная визуализация была продемонстрирована с помощью компактного оптоволоконного спектрометра [125; 126] или интерферометра Майкельсона [127].

Восстановление однопиксельных измерений, полученных для гиперспектральных изображений, предложено в работе [128]. Подход состоит из двух

этапов: на первом шаге гиперспектральные изображения реконструированы поканально из измерений с помощью ISTA-Net, в то время как на втором шаге результаты реконструкции уточняются с помощью глубокой нейронной сети. Также нейронные сети показали себя надёжным инструментом в задачах обесшумливания гиперспектральных изображений и нашли применение в задачах классификации [129]. Учитывая появление третьего измерения в гиперспектральных данных, для решения задач реконструкции многие исследовательские группы естественно обратились к трёхмерным свёрточным сетям и ядрам [130–133]. Тем не менее, в то время как другие технологии гиперспектральной визуализации были исследованы с точки зрения применимости к ним 3D сетей, задача реконструкции в однопиксельной гиперспектральной визуализации по-прежнему ограничена 2D методами.

В данной главе рассмотрена применимость глубоких свёрточных нейронных сетей для восстановления гиперспектральных изображений из однопиксельных измерений. Основной идеей предложенного подхода является интерпретируемость слоёв нейронной сети: первый слой представляет собой дополнение недостающих измерений и восстановление изображения с помощью минимизации критерия среднеквадратичной ошибки и регуляризации по Тихонову, а последующие слои – дополнительную регуляризацию решения. Предложенная схема в том числе обобщает решение, описанное в работе [134], для случая многоканальных изображений.

3.4 Восстановление изображений из однопиксельных измерений

Наличие шума в зарегистрированных однопиксельных сигналах, а также использование меньшего количества бинарных паттернов для сокращения времени регистрации приводят к формированию проблемы восстановления искомого изображения из однопиксельного измерения как *некорректно поставленной обратной задачи*. Классический подход к решению такой задачи основан на методах compressive sensing [112]. Успех методов глубокого обучения в решении задач восстановления изображений привёл к их применению и в задачах однопиксельной визуализации. Например, в работе [135] нейронная сеть была

использована для улучшения изображения, полученного с помощью так называемой технологии ghost imaging – метода, создающего изображение объекта путем объединения информации от двух детекторов: однопиксельного детектора, регистрирующего изображение объекта, и обычного многопиксельного детектора, не получающего информацию о непосредственно объекте [136]. Затем в [137] была предложена реконструкция однопиксельного измерения, полученного с использованием Фурье-базиса, с помощью свёрточного автоэнкодера. В работе [138] свёрточный автоэнкодер показал лучшие результаты реконструкции изображения по сравнению с классическими методами compressive sensing. В работе [139] с помощью рекуррентной нейронной сети была учтена корреляция между последовательно зарегистрированными с помощью различных паттернов однопиксельными измерениями. Для восстановления изображений с сильно пониженным разрешением в работе [140] была применена генеративно-состязательная сеть. В работе [141] была предложена нейронная сеть, первый слой которой осуществляет поиск недостающих измерений и их отображение в область изображений при наличии не содержащих шум измерений. В работе [134] эта идея была обобщена для случая зашумлённых данных.

В настоящей работе для восстановления изображений из однопиксельных измерений применён подход, предложенный в работе [134]. Как описано в Разделе 3.4, в процессе формирования однопиксельное измерение подвергается искажению вследствие воздействия двух факторов. Во-первых, из-за наличия флуктуаций в электрической сети, а также из-за дискретной природы носителей электрического заряда, измерение искажено гауссовским и пуассоновским шумами. Во-вторых, с целью сокращения времени, затраченного на получение вектора измерений, для его формирования использовано не $P = N$, а $P < N$ паттернов Адамара. Таким образом, полученный сигнал является зашумлённым и требует дополнения недостающих измерений.

Предполагая, что матрица \mathbf{H} является ортогональной, решение задачи восстановления искомого изображения из полученного однопиксельного измерения имеет вид

$$\hat{\mathbf{x}} = \mathbf{H}^T \hat{\mathbf{y}}, \quad (3.5a)$$

$$\hat{\mathbf{y}}(\mathbf{m}^\alpha) = [\mathbf{y}_1^T, \mathbf{y}_2^T]^T. \quad (3.5b)$$

В Формуле 3.5b $\hat{\mathbf{y}}(\mathbf{m}^\alpha)$ представляет собой полный вектор измерений, состоящий из вектора искажённых шумом зарегистрированных измерений \mathbf{y}_1 и вектора недостающих вследствие использования $P < N$ паттернов Адамара измерений \mathbf{y}_2 . Таким образом, задача восстановления изображения из однопиксельного измерения состоит в устранении шума, искажающего измерение, и дополнении восстановленного изображения недостающей информацией. Авторами работы [134] использована теорема Байеса, с точки зрения которой решение задачи восстановления полного вектора измерений задано апостериорным распределением. Одним из способов вычисления решения является его выражение через условное математическое ожидание,

$$\hat{\mathbf{y}}(\mathbf{m}^\alpha) = \mathbb{E}(\mathbf{y}_1, \mathbf{y}_2 | \mathbf{m}^\alpha = \mathbf{m}^\alpha) = \hat{\mathcal{G}}(\mathbf{m}^\alpha), \quad (3.6)$$

где \mathbf{y}_1 и \mathbf{y}_2 – случайные величины, относящиеся к полученным и недостающим измерениям соответственно, а \mathbf{m}^α – полученный вектор измерений, являющийся выборкой из случайной величины \mathbf{m}^α . Оценка условного математического ожидания может быть также задана оператором $\hat{\mathcal{G}}$.

В работе [134] показано, что если случайная величина \mathbf{y} , соответствующая полному вектору измерений, подчиняется нормальному распределению с математическим ожиданием $\boldsymbol{\mu} \in \mathbb{R}^N$ и матрицей ковариаций $\boldsymbol{\Sigma} \in \mathbb{R}^{N \times N}$ и вектор измерений искажён аддитивным гауссовским шумом, решение для выражения 3.5a может быть получено с помощью минимизации критерия среднеквадратичной ошибки и регуляризации по Тихонову как

$$\mathbf{y}_1 = \boldsymbol{\mu}_1 + \boldsymbol{\Sigma}_1 [\boldsymbol{\Sigma}_1 + \boldsymbol{\Sigma}_\alpha]^{-1} (\mathbf{m}^\alpha - \boldsymbol{\mu}_1), \quad (3.7a)$$

$$\mathbf{y}_2 = \boldsymbol{\mu}_2 + \boldsymbol{\Sigma}_{21} \boldsymbol{\Sigma}_1^{-1} [\mathbf{y}_1 - \boldsymbol{\mu}_1]. \quad (3.7b)$$

Здесь $\boldsymbol{\mu}_1 \in \mathbb{R}^P$ и $\boldsymbol{\Sigma}_1 \in \mathbb{R}^{P \times P}$ равны математическому ожиданию и матрице ковариаций измеренных значений, $\boldsymbol{\Sigma}_\alpha \in \mathbb{R}^{P \times P}$ – матрица ковариаций шума, $\boldsymbol{\mu}_2 \in \mathbb{R}^{N-P}$ – математическое ожидание недостающих значений вектора однопиксельных измерений, а $\boldsymbol{\Sigma}_{21} \in \mathbb{R}^{(N-P) \times P}$ – матрица ковариаций между полученными и недостающими значениями измерений. Выражения для вычисления указанных величин представлены в следующих разделах. Таким образом, на первом этапе решения, соответствующем Формуле 3.7a, выполняется удаление

ние шума в полученных измерениях. На втором этапе, описанном Формулой 3.7b, осуществляется дополнение недостающих измерений.

В работе [142] показано, что оператор $\hat{\mathcal{G}}$ в Формуле 3.6 может быть найден как

$$\hat{\mathcal{G}} \in \operatorname{argmin}_{\mathcal{G}} \mathbb{E}(\|\mathcal{G}(\mathbf{m}^\alpha) - \mathbf{y}\|_2^2). \quad (3.8)$$

Как было указано, если \mathbf{y} подчиняется нормальному распределению, оператор $\hat{\mathcal{G}}$ осуществляет линейное отображение, и решение для $\hat{\mathbf{y}}$ может быть найдено в соответствии с Формулами 3.7. Однако в общем случае оператор $\hat{\mathcal{G}}$ является нелинейным, вследствие чего выражение решения в замкнутой форме не представляется возможным [142]. По этой причине авторами работы [134] предложено вычисление $\hat{\mathcal{G}}$ через семейство нелинейных отображений \mathcal{G}_θ . Для этого предложено заменить математическое ожидание в Формуле 3.8 на эмпирическое среднее значение по набору данных из S тренировочных пар измерение-эталонное изображение

$$\hat{\mathcal{G}} \in \operatorname{argmin}_{\theta} \frac{1}{S} \sum_{s=1}^S \left\| \mathcal{G}_\theta(\mathbf{m}^{\alpha(s)}) - \mathbf{x}^{(s)} \right\|_2^2, \quad (3.9)$$

где \mathcal{G}_θ обозначает нейронную сеть с параметрами θ . Обучение сети, то есть обновление её параметров, происходит путём минимизации функции вида

$$L = \|\mathbf{x} - \hat{\mathbf{x}}\|_2^2, \quad (3.10)$$

где $\hat{\mathbf{x}} = \mathcal{G}_\theta(\mathbf{m}^\alpha)$. Традиционно нейронные сети представляют собой композицию каскадных слоёв

$$\mathcal{G}_\theta = \mathcal{G}_\theta^L \circ \dots \circ \mathcal{G}_\theta^1, \quad (3.11)$$

где \mathcal{G}_θ^l , $1 \leq l \leq L$ является l -м нелинейным слоем нейросети, а символ \circ обозначает композицию функций.

Авторами работ [141] и [134] предложено решение Формулы 3.5a и отображение из области измерений в область изображений с помощью полносвязного слоя нейронной сети с фиксированными параметрами

$$\hat{\mathbf{x}} = \mathcal{G}^1(\mathbf{m}^\alpha). \quad (3.12)$$

Здесь \mathcal{G}^1 осуществляет процедуру дополнения обесшумленных с помощью Формулы 3.7a данных, и параметры полносвязаного слоя вычислены согласно Формуле 3.7b. Схема сети для обесшумливания и дополнения измерений представлена на Рисунке 3.2.

3.4.1 Обобщение задачи восстановления для случая гиперспектральных изображений

Процесс поканального формирования однопиксельного измерения для гиперспектрального изображения может быть представлен как

$$\mathbf{m}_\lambda^\alpha = \alpha \mathbf{H}_1 \mathbf{x}_\lambda, \quad 1 \leq \lambda \leq \Lambda, \quad (3.13)$$

где \mathbf{m}_λ и \mathbf{x}_λ являются λ -м каналом вектора измерений и искомого изображения соответственно. Вектор измерений и гиперспектральное изображение соответственно обозначены как $\mathbf{m} = [\mathbf{m}_1^\top \dots \mathbf{m}_\Lambda^\top]^\top$ и $\mathbf{x} = [\mathbf{x}_1^\top \dots \mathbf{x}_\Lambda^\top]^\top$.

Далее описаны два основных подхода к восстановлению гиперспектрального изображения из однопиксельного измерения, реализующих решение с использованием двух- и трёхмерной свёрточной нейронной сети.

3.4.2 Решение с помощью двумерной свёрточной нейронной сети

На первом этапе работы рассмотрен подход, согласно которому нейронная сеть осуществляет восстановление изображения полностью поканально, не учитывая информацию из соседних спектральных каналов. Для реализации этого подхода выбрана двумерная свёрточная нейронная сеть, применяемая к данным как

$$\tilde{\mathbf{x}}_\lambda = \mathcal{G}^1(\mathbf{m}_\lambda^\alpha), \quad 1 \leq \lambda \leq \Lambda, \quad (3.14a)$$

$$\hat{\mathbf{x}}_\lambda = (\mathcal{G}_\theta^L \circ \dots \circ \mathcal{G}_\theta^2)(\tilde{\mathbf{x}}_\lambda), \quad 1 \leq \lambda \leq \Lambda. \quad (3.14b)$$

Здесь $\tilde{\mathbf{x}}$ – промежуточное решение, полученное согласно Формулам 3.7. Сначала происходит устранение шума в полученном векторе измерений \mathbf{m}^α . Затем оператор \mathcal{G}^1 реализует дополнение недостающих измерений и их отображение в область изображений. Оператор \mathcal{G}^1 представлен полносвязным слоем нейронной сети с фиксированными весами [134]. Затем уточнение решения выполнена с помощью свёрточной нейронной сети, а именно UNet [53], слои которой обозначены $(\mathcal{G}_\theta^L \circ \dots \circ \mathcal{G}_\theta^2)$. Эти слои основаны на двумерных свёрточных фильтрах, действующих на входное изображение $\tilde{\mathbf{x}}_\lambda$ исключительно в пространственной области. Параметры сети обучены с помощью чёрно-белых изображений. Схема этой сети представлена на Рисунке 3.3а. Финальная модель названа DC-UNet-2 (Denoised Completion, сеть для обесшумливания и дополнения), и схема её функционирования представлена на Рисунке 3.2.

3.4.3 Решение с помощью трёхмерной свёрточной нейронной сети

На следующем этапе работы для исследования взаимного влияния каналов гиперспектрального изображения предложено использовать трёхмерную свёрточную нейронную сеть. Алгоритм поканального восстановления изображений аналогичен описанному в Разделе 3.4.2. После восстановления отдельных изображений для каждого канала, уточнение полученных результатов реконструкции выполнена с помощью трёхмерной свёрточной нейронной сети,

$$\tilde{\mathbf{x}}_\lambda = \mathcal{G}^1(\mathbf{m}_\lambda^\alpha), \quad 1 \leq \lambda \leq \Lambda, \quad (3.15a)$$

$$\hat{\mathbf{x}} = \mathcal{G}_\theta(\tilde{\mathbf{x}}). \quad (3.15b)$$

В Формуле 3.15 слои свёрточной нейронной сети \mathcal{G}_θ основаны на трёхмерных свёрточных фильтрах, которые взаимодействуют с входным гиперспектральным изображением как в пространственной, так и в спектральной областях, в отличие от двумерной сети. В качестве сети выбрана трёхмерная UNet, впервые описанная в работе [143]. Схема этой сети представлена на Рисунке 3.3б. Финальная модель названа DC-UNet-3, и схема её функционирования представлена на Рисунке 3.2.

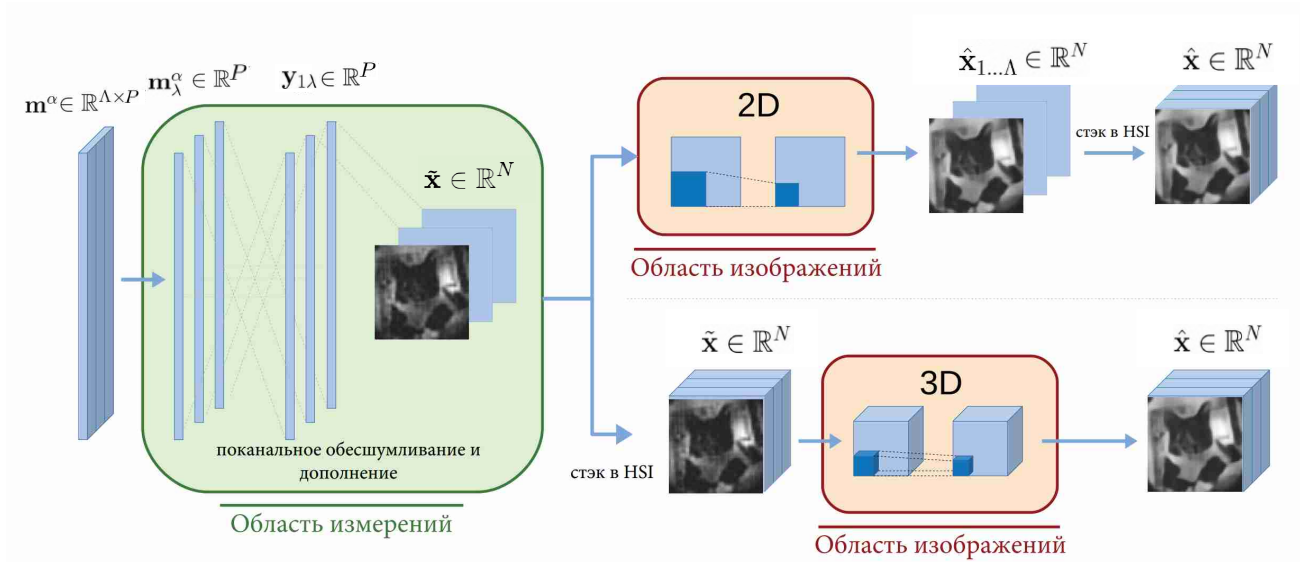


Рисунок 3.2 представляет схему работы двух- и трёхмерной свёрточных сетей, осуществляющих восстановление гиперспектральных изображений из вектора однопиксельных измерений. Обесшумливание и дополнение данных происходит в области измерений для каждого канала в отдельности. Левая часть Рисунка 3.2 соответствует выражению 3.7. В правой части Рисунка 3.2 вверху представлена схема работы DC-UNet-2. В правой части Рисунка 3.2 внизу представлена схема работы DC-UNet-3.

Рисунок 3.2 — Схема двух вариантов модели для восстановления изображений из гиперспектральных однопиксельных измерений.

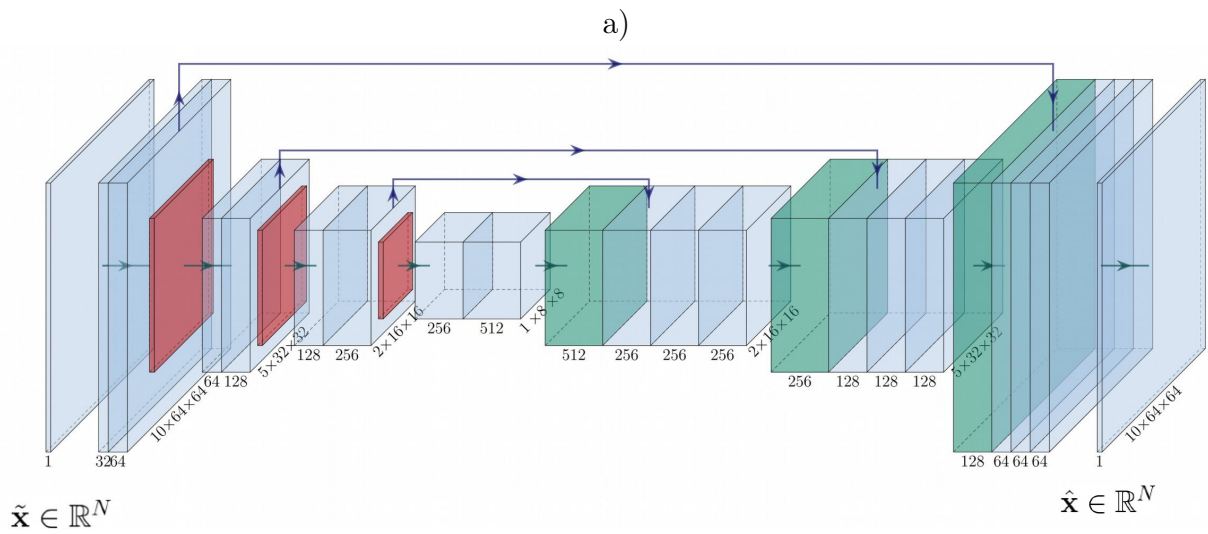
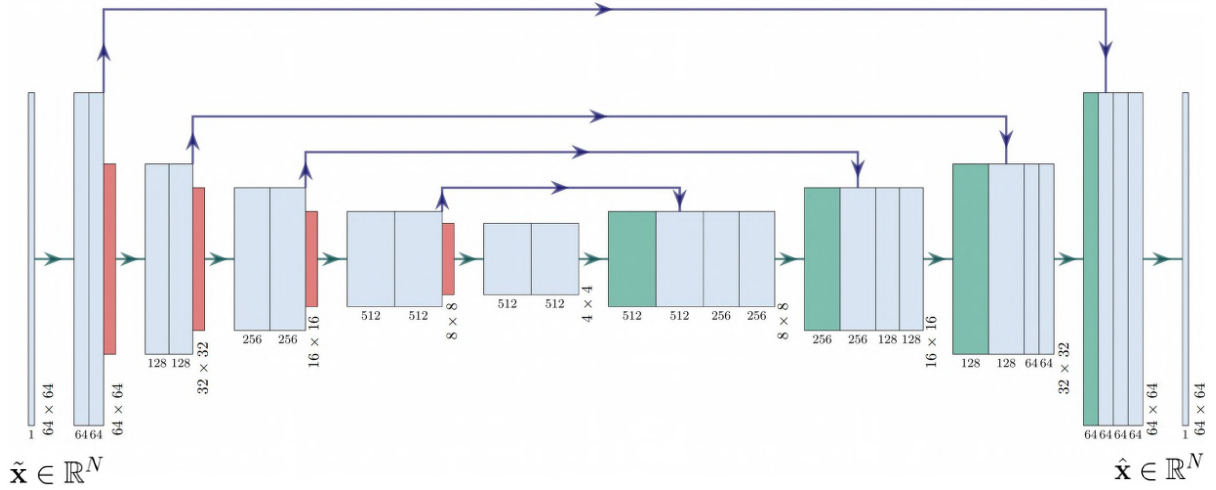
3.4.4 Вычисление математического ожидания и матриц ковариаций

Для восстановления изображения с использованием DC-UNet-2 и DC-UNet-3, необходимо вычислить математическое ожидание μ и ковариацию Σ данных, а также ковариацию шума Σ_α , используемые в Формуле 3.7.

Математическое ожидание μ и ковариация Σ заранее вычислены для тренировочного набора данных, описанного далее. Вычисление осуществлено со-

гласно Формулам

$$\begin{aligned} \boldsymbol{\mu} &= \begin{bmatrix} \boldsymbol{\mu}_1 \\ \boldsymbol{\mu}_2 \end{bmatrix} = \frac{1}{S} \sum_{i=1}^S \mathbf{H}\mathbf{x}_i, \\ \boldsymbol{\Sigma} &= \begin{bmatrix} \boldsymbol{\Sigma}_1 & \boldsymbol{\Sigma}_{21}^\top \\ \boldsymbol{\Sigma}_{21} & \boldsymbol{\Sigma}_2 \end{bmatrix} = \frac{1}{S-1} \sum_{i=1}^S (\mathbf{H}\mathbf{x}_i - \boldsymbol{\mu})(\mathbf{H}\mathbf{x}_i - \boldsymbol{\mu})^\top. \end{aligned} \quad (3.16)$$



б)

На Рисунке 3.3а представлена схема двумерной UNet. На Рисунке 3.3б представлена схема трёхмерной UNet.

Рисунок 3.3 — Схемы двух вариантов свёрточных сетей для уточнения решения задачи восстановления изображения из гиперспектральных однопиксельных измерений.

Важно отметить, что обе величины вычислены один раз и используются во время тренировки и оценки алгоритмов без необходимости дополнительных вычислений.

Ковариация шума в полученных измерениях Σ_α оценивается в соответствии с формулой

$$\Sigma_{\alpha_\lambda} = \text{diag}(\sigma_{\alpha_\lambda}^2) = \frac{1}{K\alpha_\lambda^2} \text{diag}(\mathbf{m}_{+\lambda}^\alpha + \mathbf{m}_{-\lambda}^\alpha) - \rho \frac{2\mu_{\text{dark}}}{\alpha_\lambda^2 K} \mathbf{I} + \rho \frac{2\sigma_{\text{dark}}^2}{\alpha_\lambda^2 K^2} \mathbf{I}, \quad (3.17)$$

где $\text{diag}(\sigma_{\alpha_\lambda}^2)$ является диагональной матрицей, содержащей значения дисперсии шума $\sigma_{\alpha_\lambda}^2$ на главной диагонали, а \mathbf{I} обозначает единичную матрицу. Значения $\sigma_{\alpha_\lambda}^2$ зависят от спектрального канала изображения, поскольку каналы содержат различное количество фотонов. Математическое ожидание μ_{dark} , дисперсия шума σ_{dark} и калибровочный коэффициент K предварительно оценены экспериментально согласно описанному в работе [123] методу. Параметр ρ определяет количество каналов исходного однопиксельного измерения, объединённых для создания одного нового спектрального канала λ из Λ спектральных каналов. Более подробно этот параметр описан в Разделе 3.8.

3.5 Архитектуры свёрточных нейронных сетей

В настоящей работе для восстановления гиперспектральных изображений из однопиксельных измерений предложены следующие алгоритмы: трёхмерная нейронная сеть DC-UNet-3, двумерная нейронная сеть DC-UNet-2 и двумерный алгоритм решения, полученного в соответствии с выражениями 3.7 и не применяющий дополнительную регуляризацию с помощью нейронной сети. Здесь и далее это решение обозначено DC – **D**enoised **C**ompletion, то есть обесшумливание и дополнение измерений. Также исследована способность двумерной нейронной сети восстанавливать многоканальное изображение из многоканальных входных данных путём применения многоканальных свёрточных ядер. Согласно такому подходу, двумерная свёрточная нейронная сеть, в данном случае UNet, получает на входе не двумерное изображение каждого отдельного канала, а трёхмерное гиперспектральное изображение. Фильтры сети осуществляют

свёртку не по одному каналу изображения, а по нескольким. С помощью такого подхода, обозначенного здесь и далее как DC-UNet-2ch, регуляризация также осуществляется в пространственной и спектральной областях, но посредством не трёхмерных свёрточных фильтров, как в случае DC-UNet-3, а набора двумерных свёрточных фильтров. Ввиду отсутствия иных методов, реализующих восстановление гиперспектральных изображений из однопиксельных измерений, для сравнения полученных результатов использован метод, предложенный в работе [137]. Данный метод, названный авторами DL-FSPI, был разработан для однопиксельных измерений, полученных с помощью Фурье-кодированных паттернов. В настоящей работе восстановление изображений из измерений проведено с помощью алгоритма DC согласно Формулам 3.7. Затем уточнение полученного решения осуществлена с помощью свёрточной нейронной сети, предложенной авторами [137], – полученная модель здесь и далее обозначена DC-DL-FSPI. Таким образом решена двойная задача: разработанный в настоящем исследовании метод подвергается сравнению с другим методом, а также исследуется влияние различных архитектур нейронных сетей для этапа дополнительной регуляризации решения на результат восстановления изображения.

В двумерных моделях, обозначенных как DC-UNet-2 и DC-UNet-2ch, в качестве нейронной сети использована архитектура 2D UNet с 13,394,177 обучаемыми параметрами, схема которой представлена на Рисунке 3.3а. Для реализации алгоритма DC-UNet-3 использована 3D UNet с 16,320,257 обучаемыми параметрами, представленная на Рисунке 3.3б. Количество обучаемых параметров в нейронной сети, использованной в моделях DL-FSPI и DC-DL-FSPI, равно 317,121.

В деталях архитектура 2D UNet состоит из сужающегося пути и расширяющегося пути. Начиная с первого свёрточного слоя с 64 каналами или картами признаков с каждой процедурой подвыборки с определением максимального значения это число увеличивается в 2 раза, достигая значения 512 на последнем свёрточном слое сужающегося пути. Размер карт признаков, соответствующих входному изображению на первом слое сети, уменьшается. На расширяющемся пути количество каналов соответственно уменьшается. Этапы расширяющегося пути 2D UNet состоят из билинейной повышающей подвыборки карт признаков, объединении с картами признаков из сужающегося пути и двух свёрточных слоев. Каждый свёрточный слой, кроме последнего, использует свёрточные филь-

тры размером 3×3 , за которыми следует нормализация по набору и слой ReLU. Последний слой, производящий выходные данные сети, состоит из одного сверточного слоя с размером фильтра 1×1 .

Применённая архитектура 3D UNet практически повторяет архитектуру 2D UNet, но каждый слой реализован с помощью трёхмерных свёрточных фильтров размером $3 \times 3 \times 3$. Последний слой сети, производящий на выходе восстановленное изображение, состоит из одного свёрточного слоя, реализуемого с помощью фильтра размером $1 \times 1 \times 1$. Таким образом, начиная с первого свёрточного слоя с количеством каналов, равным 32, это число удваивается с каждым свёрточным слоем, достигая значения 512 в конце сужающегося пути. В данной архитектуре процесс повышающей подвыборки осуществляется с использованием интерполяции.

Свёрточная сеть, описанная в работе [137] и обозначенная DL-FSPI, состоит из сужающегося пути и расширяющегося пути. Первый свёрточный слой состоит из 32 каналов и реализует свёртки с помощью фильтров размером 5×5 . На следующем слое используется 64 канала с фильтрами размером 3×3 , а затем 128 каналов с фильтрами размером 5×5 . На расширяющемся пути количество каналов и размер фильтров меняются, соответственно, как $64 - 32$ и $3 - 3$. В этой сети также использована билинейная повышающая подвыборка карт признаков, подвыборка с определением максимального значения, нормализация по набору и нелинейные слои ReLU. Последний слой, производящий выходные данные сети, состоит из одного сверточного слоя с размером фильтра 1×1 .

3.6 Моделирование тренировочных гиперспектральных данных из RGB-изображений

Для обучения всех моделей использован датасет STL-10 [144], состоящий из 105,000 цветных изображений, разделённых на наборы, зарезервированные для обучения и тестирования. Исходные изображения размером 96×96 преобразованы в 64×64 с помощью бикубической интерполяции. RGB изображения представлены в виде $[\mathbf{x}^r, \mathbf{x}^g, \mathbf{x}^b] \in \mathbb{R}^{N \times 3}$, где каждый цветовой канал обозначен символом \mathbf{x} с соответствующим индексом. Для создания набора данных с необ-

ходимым количеством спектральных каналов предложен метод преобразования RGB изображений в гиперспектральные с использованием кривых спектральной чувствительности человеческого глаза $\eta^r \in \mathbb{R}$, $\eta^g \in \mathbb{R}$, и $\eta^b \in \mathbb{R}$, соответствующих восприятию красного, зелёного и синего цветов соответственно [145; 146]. Для создания Λ спектральных каналов кривые спектральной чувствительности разбиваются на Λ интервалов. Затем каждый канал гиперспектрального изображения сгенерирован как сумма перемноженных значений цветного канала RGB-изображения и значения кривой спектральной чувствительности для выбранного интервала, то есть

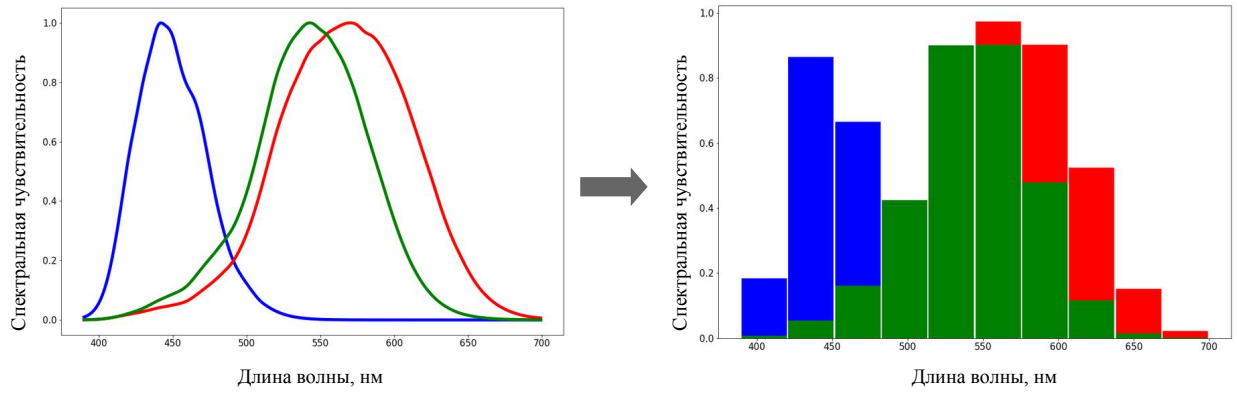
$$\mathbf{x}_\lambda = \mathbf{x}^r \eta_\lambda^r + \mathbf{x}^g \eta_\lambda^g + \mathbf{x}^b \eta_\lambda^b. \quad (3.18)$$

Разбиение кривых спектральной чувствительности на интервалы показано на Рисунке 3.4а. Схема процедуры создания гиперспектрального изображения из RGB представлена на Рисунке 3.4б. Таким образом, в данной работе сгенерированы гиперспектральные изображения, содержащие $\Lambda = 10$ каналов.

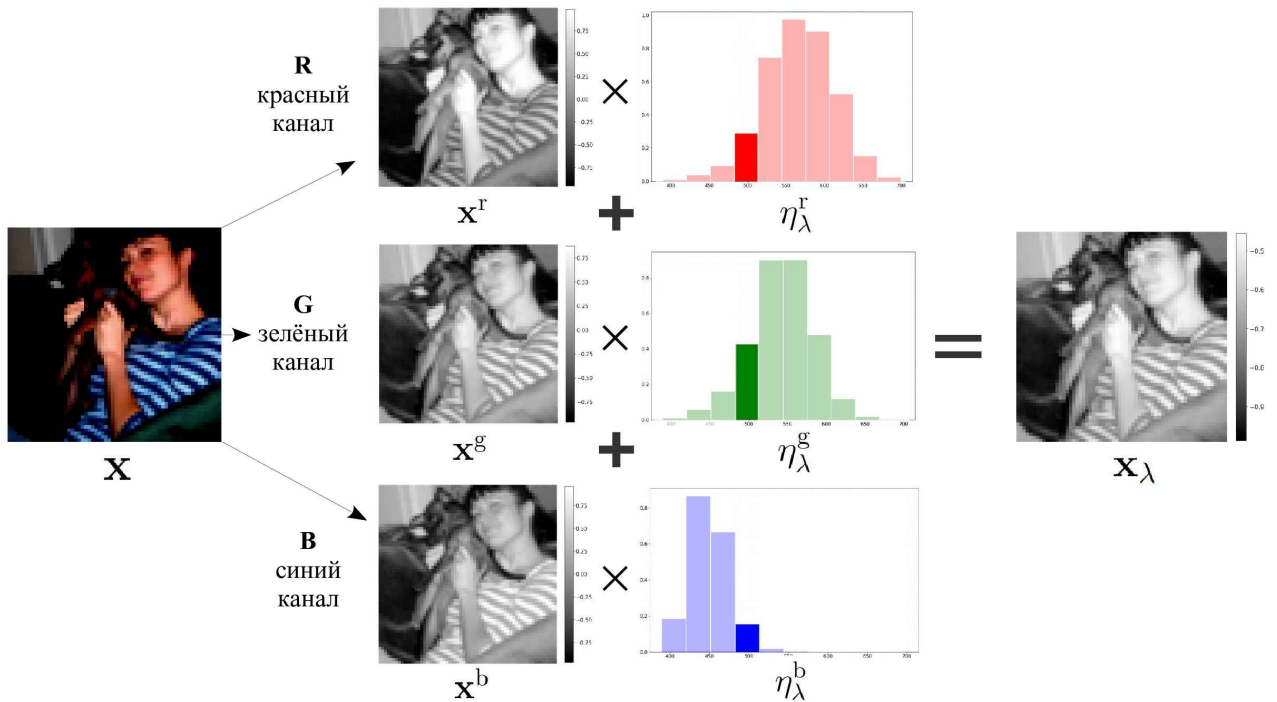
В процессе обучения моделей тренировочный набор однопиксельных измерений сгенерирован из тренировочного набора гиперспектральных изображений в соответствии с Формулой 3.3, где для простоты вычислений калибровочный коэффициент принят равным единице, а шум системы игнорируется. Тем не менее, при восстановлении экспериментальных изображений реальный калибровочный коэффициент и уровень шума системы учтены с помощью Формулы 3.17.

3.7 Обучение моделей для восстановления изображений

Все модели обучены с помощью библиотеки PyTorch [147] с использованием алгоритма оптимизации Adam [63] с изначальным значением скорости обучения, равным 10^{-3} , и уменьшаемым в 5 раз каждые 20 эпох тренировки. Все модели обучены в течение 100 эпох. Предложенный 3D алгоритм восстановления изображений обучен с помощью четырёх видеокарт NVIDIA Quadro RTX 6000 на протяжении 28 часов 51 минуты. Двумерные модели DC-UNet-2, DC-UNet-2ch и DC-DL-FSPI обучены на двух видеокартах NVIDIA GeForce GTX



а)



б)

Рисунок 3.4а представляет схему разбиения кривых спектральной чувствительности человеческого глаза [145; 146] на отдельные интервалы для 10-канального гиперспектрального изображений. На Рисунке 3.4б на примере одного гиперспектрального канала представлена схема преобразования RGB изображения в гиперспектральное с помощью кривых спектральной чувствительности.

Рисунок 3.4 — Схема преобразования RGB изображения в гиперспектральное изображение.

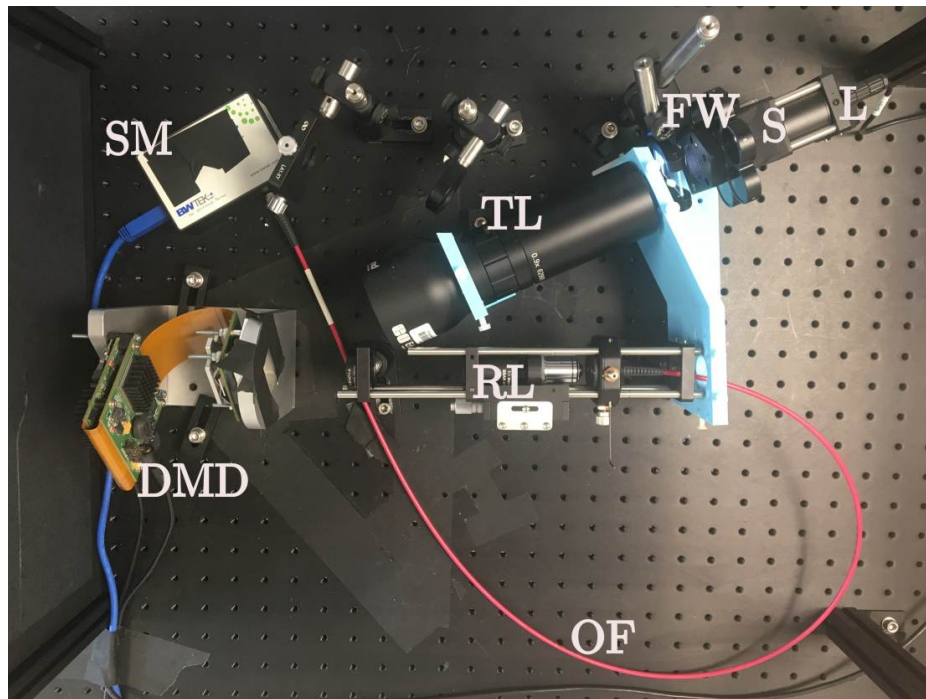
1080Ti на протяжении 10 часов 35 минут для первых двух алгоритмов и 7 часов 50 минут для последнего. Для ускорения и стабилизации процесса обучения, все изображения, как чёрно-белые, так и гиперспектральные, нормированы к $[-1;1]$.

3.8 Экспериментально полученные гиперспектральные однопиксельные измерения

Помимо гиперспектральных изображений, сгенерированных на основе датасета STL-10, предложенные модели также применены к экспериментально полученным однопиксельным гиперспектральным измерениям. Эти измерения получены с помощью экспериментальной установки, изображённой на Рисунке 3.5, и представлены в датасете SPINIM [126; 148]. Набор данных содержит измерения, полученные с использованием 4096 паттернов Адамара размером $N = 64 \times 64$ для 2048 спектральных каналов в диапазоне [317, 1064]. Измерения получены с использованием нейтральных фильтров с различной плотностью, помещённых между источником освещения и объектом, в результате чего получены измерения с различной интенсивностью шума.

Экспериментальная установка для генерирования гиперспектральных однопиксельных измерений представлена на Рисунке 3.5. Изображение освещаемого светодиодной лампой (Thorlabs LIUCWHA/M00441662) объекта формируется с помощью телецентрического объектива (Edmund Optics 62901) и проецируется на цифровое микрозеркальное устройство (vialux V-7001). С помощью цифрового микрозеркального устройства происходит отражение и модуляция светового потока. Затем этот световой поток фокусируется промежуточной линзой и проецируется в оптическое волокно (Thorlabs FT1500UMT 0.39NA). Оптическое волокно, в свою очередь, подключено к спектрометру (BWTeK exemplar BRC115P-V-ST1). Таким образом, при визуализации объекта с использованием спектрометра регистрируется световой поток, полученный в результате корреляции между изображением объекта и бинарными паттернами, разделённый на отдельные каналы, соответствующие диапазонам спектра.

В данной работе для создания однопиксельных гиперспектральных измерений использованы четыре чёрно-белых изображения: изображение кошки из набора данных STL-10, звезда Сименса [149], смещённая звезда Сименса и изображение самого источника света. Путём ограничения светового потока с помощью нейтрального фильтра для каждого изображения получено зашумлённое измерение. Для изображения кошки использован фильтр с плотностью равной 0.6, а для всех остальных изображений – 1.3. Также время получения изме-



Установка состоит из светодиодной лампы (lamp, L), освещающей объект (sample, S), фильтров (filter wheel, FW), телецентрической линзы (telecentric lens, TL), микрозеркального устройства (digital micromirror device, DMD), промежуточной линзы (relay lens, RL), оптического волокна (optic fiber, OF) и спектрометра (spectrometer, SM).

Рисунок 3.5 — Экспериментальная установка для регистрации гиперспектральных однопиксельных измерений [148].

рения, то есть освещения искомого изображения, ограничено 8 мс для звезды Сименса и 4 мс для остальных изображений. Кроме чёрно-белых изображений также получено измерение для цветного изображения звезды Сименса.

Из полного вектора полученных измерений, состоящего из 4096 значений, произведена выборка из 1024 измерений. Из всего диапазона спектральных каналов выбраны те, которые находятся в диапазоне видимого света, то есть [390; 700] нм. Затем эти каналы объединены в 10 интервалов, содержащих $\rho = 77$ спектральных каналов каждый. Псевдо-эталонные изображения для последующей оценки результатов восстановления получены из полного вектора измерений, полученных с плотностью фильтра равной 0.3 или 0.

3.9 Результаты восстановления гиперспектральных изображений из однопиксельных измерений

3.9.1 Результаты для смоделированных однопиксельных измерений

Для оценки качества восстановления изображений с помощью всех алгоритмов использованы следующие метрики: отношение пикового сигнала к шуму (PSNR), индекс структурного сходства (SSIM), реализованные в библиотеке `scikit-image` [67], и метрика спектрального сходства (SAM, spectral angle mapper) [150], реализованная в библиотеке `image-quality-assessment` [151]. Последняя метрика – SAM – определяет сходство между двумя спектрами, рассматривая их как векторы в пространстве и вычисляя угол между ними, и является распространённой метрикой в задачах восстановления гиперспектральных изображений [152]. Для экспериментальных измерений эталонные изображения получены как $\mathbf{x} = \mathbf{H}^T \mathbf{y}_{gt}$, где \mathbf{y}_{gt} – полный вектор измерений, то есть $P = N$, с высоким отношением сигнала к шуму. Перед тем, как вычислить метрики, эталонные изображения нормированы к $[-1;1]$. Для вычисления метрики SAM все изображения нормированы к $[0; 255]$. Значения PSNR и SSIM для каждого изображения вычислены путём усреднения значений соответствующих метрик для каждого канала. Значения метрики SAM вычислены для гиперспектрального изображения целиком.

Все алгоритмы, а именно DC, DC-UNet-2, DC-UNet-2ch, DC-DL-FSPI и DC-UNet-3, применены к тестовому набору сгенерированных гиперспектральных изображений для трёх различных интенсивностей изображения α (150, 200 и 250), соответствующих различным уровням шума. Важно отметить, что в случае DC-UNet-2 и DC-DL-FSPI на вход свёрточной нейронной сети, осуществляющей регуляризацию решения, поступают отдельные каналы изображения размером $1 \times 64 \times 64$, в то время как на вход нейронной сети в моделях DC-UNet-2ch и DC-UNet-3 поступает гиперспектральное изображение размером $10 \times 64 \times 64$. Первое значение соответствует количеству спектральных каналов, а два последних – пространственным размерам изображения. В Таблице

5 представлены значения PSNR, SSIM и SAM, полученные для результатов реконструкции изображений с различным уровнем зашумлённости. Изображения сгенерированы из набора данных STL-10 в соответствии с алгоритмом, описанным в Разделе 3.6. На Рисунке 3.6 представлены эталонные изображения, а также результаты восстановления изображений первого и последнего спектрального канала сгенерированных гиперспектральных изображений с интенсивностью $\alpha = 150$, соответствующей самому высокому уровню шума.

Из результатов, представленных в Таблице 5 и на Рисунке 3.6 можно сделать вывод о том, что методы, учитывающие регуляризацию в спектральной области помимо области изображения (DC-UNet-2ch and DC-UNet-3), достигают лучших результатов реконструкции изображений по сравнению с методами, использующими исключительно пространственную информацию (DC, DC-DL-FSPI, and DC-UNet-2). Позволяя достичь наиболее высоких метрик восстановления, DC-UNet-3 опережает DC-UNet-2 и DC-DL-FSPI по меньшей мере на 4.6 дБ, а DC-UNet-2ch – по меньшей мере на 0.2 дБ. Из результатов, представленных на Рисунке 3.6, видно, что, в то время как реконструкция с помощью DC, DC-DL-FSPI и DC-UNet-2 приводит к недостаточно качественному восстановлению информации, содержащейся в первом и последнем спектральных каналах, включение информации о межканальной корреляции с помощью DC-UNet-2ch и DC-UNet-3 помогает восстановить потерянные данные. Восстановленные изображения большего количества каналов представлены в Приложении Г.

3.9.2 Результаты для экспериментальных однопиксельных измерений

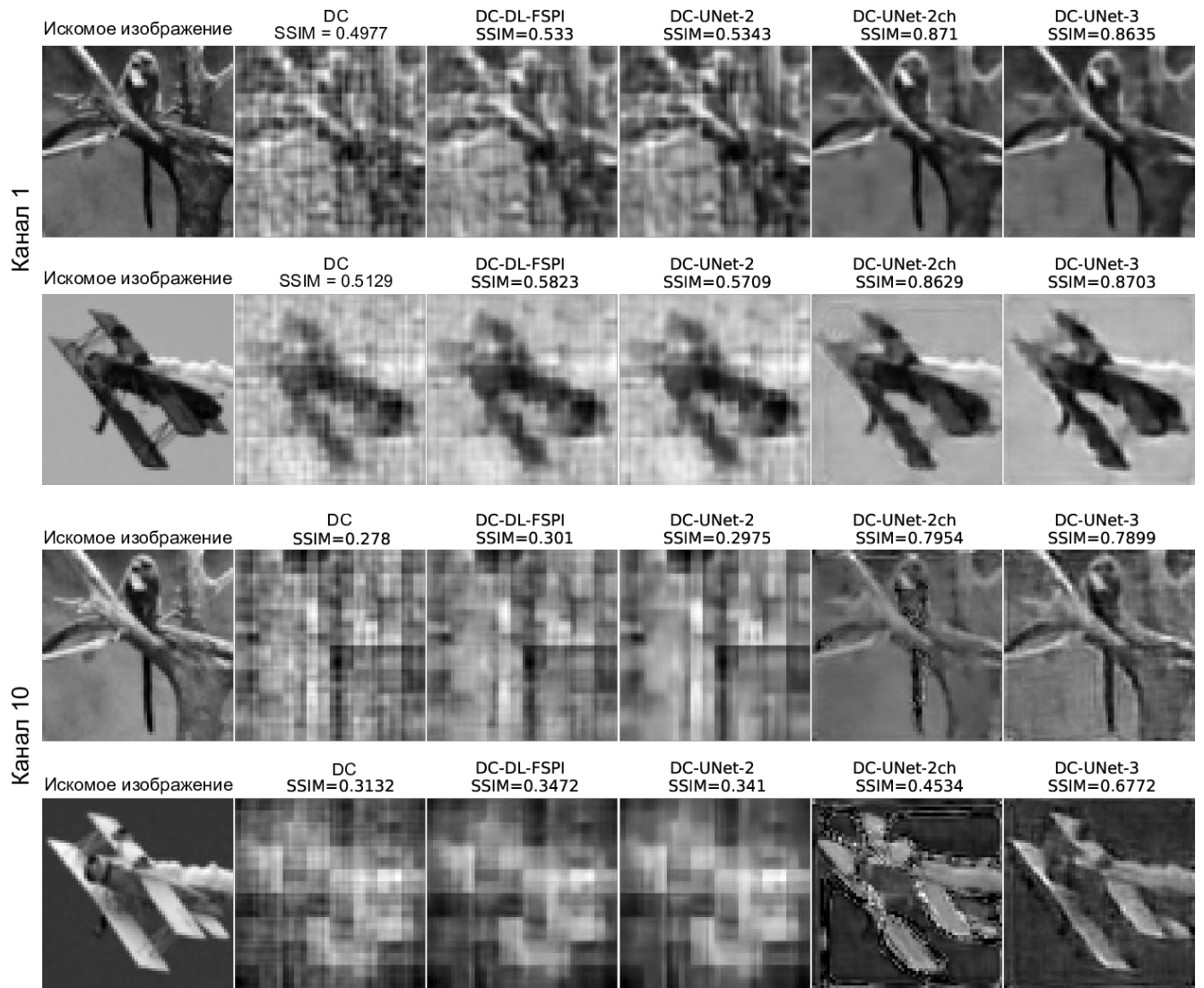
На Рисунках 3.7 и 3.8 представлены результаты восстановления четырёх экспериментальных изображений, описанных в Разделе 3.8. Восстановленные изображения большего количества каналов представлены в Приложении Г. На Рисунке 3.9 отображены графики профилей изменения интенсивности пикселей в небольших областях каждого изображения. Из представленных результатов видно, что включение свёрточной нейронной сети, действующей в пространственной области, приводит к восстановлению менее зашумлённых изображе-

Таблица 5 — Сравнительная таблица метрик качества восстановления смоделированных изображений из тестового набора данных STL-10, соответствующих трём уровням интенсивности (для всего гиперспектрального изображения $\alpha = 150, 200, 500$). Наилучший результат выделен жирным шрифтом

α		DC	DC-DL-FSPI	DC-UNet-2	DC-UNet-2ch	DC-UNet-3
150	PSNR	21.91	23.06	22.90	27.49	27.70
	SSIM	.6699	.7201	.7311	.8861	.8908
	SAM	.0896	.0799	.0685	.0228	.0244
200	PSNR	22.37	23.50	23.33	28.14	28.68
	SSIM	.6947	.7409	.7504	.9001	.9111
	SAM	.0833	.0748	.0643	.0221	.0206
500	PSNR	23.66	24.72	24.58	28.22	29.31
	SSIM	.7639	.7608	.8065	.9008	.9239
	SAM	.0641	.0605	.0488	.0201	.0186

ний, чем при применении алгоритма DC. Таким образом обе пространственно ориентированные модели, DC-UNet-2 и DC-DL-FSPI, восстанавливают изображения для каждого канала лучше, чем алгоритм DC. В то же время, несмотря на наличие некоторых артефактов, DC-UNet-2ch и DC-UNet-3 позволяют осуществить более чёткое восстановление деталей в изображении каждого спектрального канала. Особенно этот эффект заметен в “граничных” спектральных каналах – например, изображение морды кошки из датасета STL-10, изображение букв в смещённой звезде Сименса, а также лучи звезды Сименса наиболее точно соответствуют эталонным изображениям при реконструкции с помощью DC-UNet-3. Определённо, этот эффект достигнут благодаря осуществлению регуляризации в пространственной и спектральной областях, а также использованию трёхмерной свёртки.

В результате анализа графиков изменения интенсивности пикселей, представленных на Рисунке 3.9, можно сделать вывод о том, что, в то время как методы, осуществляющие поканальное восстановление изображений, более точно отображают значения пикселей в “граничных” каналах, DC-UNet-3 более точно восстанавливает значения пикселей в центральных каналах. Отметим,



Две верхние строки представляют результаты восстановления первого спектрального канала изображения, две нижние строки представляют результаты восстановления последнего (десятого) спектрального канала изображения.

Рисунок 3.6 — Результаты восстановления вектора измерений, полученного с интенсивностью, соответствующей $\alpha = 150$, для двух смоделированных изображений.

что методы, не учитывающие спектральную корреляцию, имеют идентичную форму профилей. Из изображений, представленных в первой строке, втором столбце и третьей строке, втором столбце Рисунка 3.9 видно, что восстановление с помощью DC-UNet-3 приводит к более точному воспроизведению формы профиля изменения интенсивности пикселей эталонного изображения. В свою очередь, восстановление с помощью DC-UNet-2 зачастую далеко от эталонного



Рисунок 3.7 — Результаты восстановления экспериментальных измерений двух объектов: изображения кошки из STL-10 и смещённой звезды Сименса.

изображения (Рисунок 3.9, первая строка, вторая строка, третья строка, первый столбец).

Реконструкция изображений с помощью DC-UNet-3 приводит к наилучшему устранению шума в результирующих изображениях и к лучшему восстановлению деталей изображения по сравнению с DC-UNet-2ch. Использование информации о корреляции между соседними спектральными каналами помогает восстанавливать информацию в каналах, которые изначально содержат

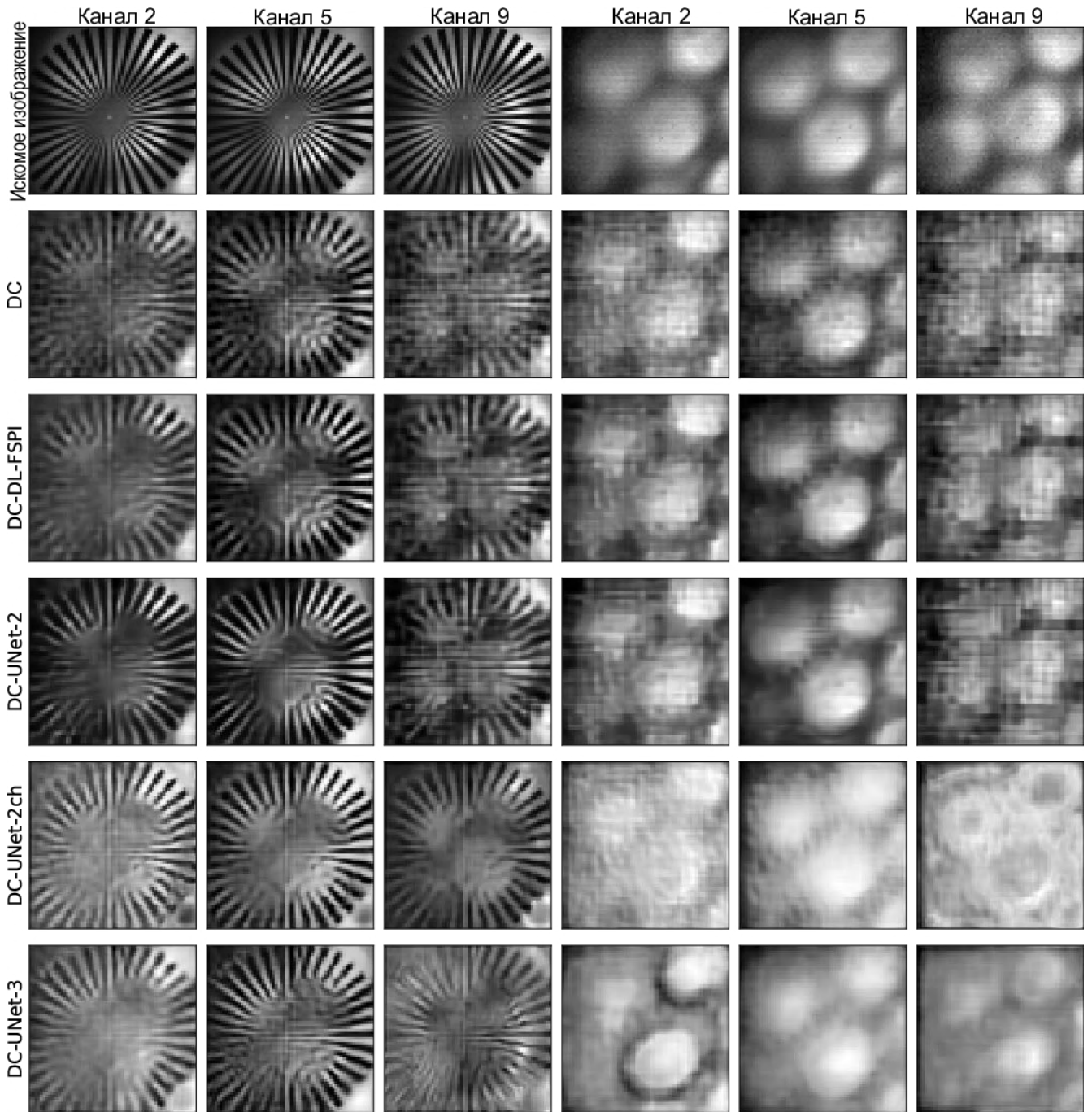


Рисунок 3.8 — Результаты восстановления экспериментальных измерений двух объектов: звезды Сименса и излучающего светодиода (LED).

меньше полезной информации. Это видно, например, на изображении смещённой звезды Сименса, на котором надпись представлена на всех спектральных каналах, в отличие от изображений полученных с помощью DC, DC-UNet-2 и DC-DL-FSPI.

На Рисунке 3.10 представлены результаты восстановления цветного экспериментального изображения звезды Сименса с помощью алгоритмов DC, DC-UNet-2, DC-DL-FSPI, DC-UNet-2ch и DC-UNet-3. Восстановленные изображе-

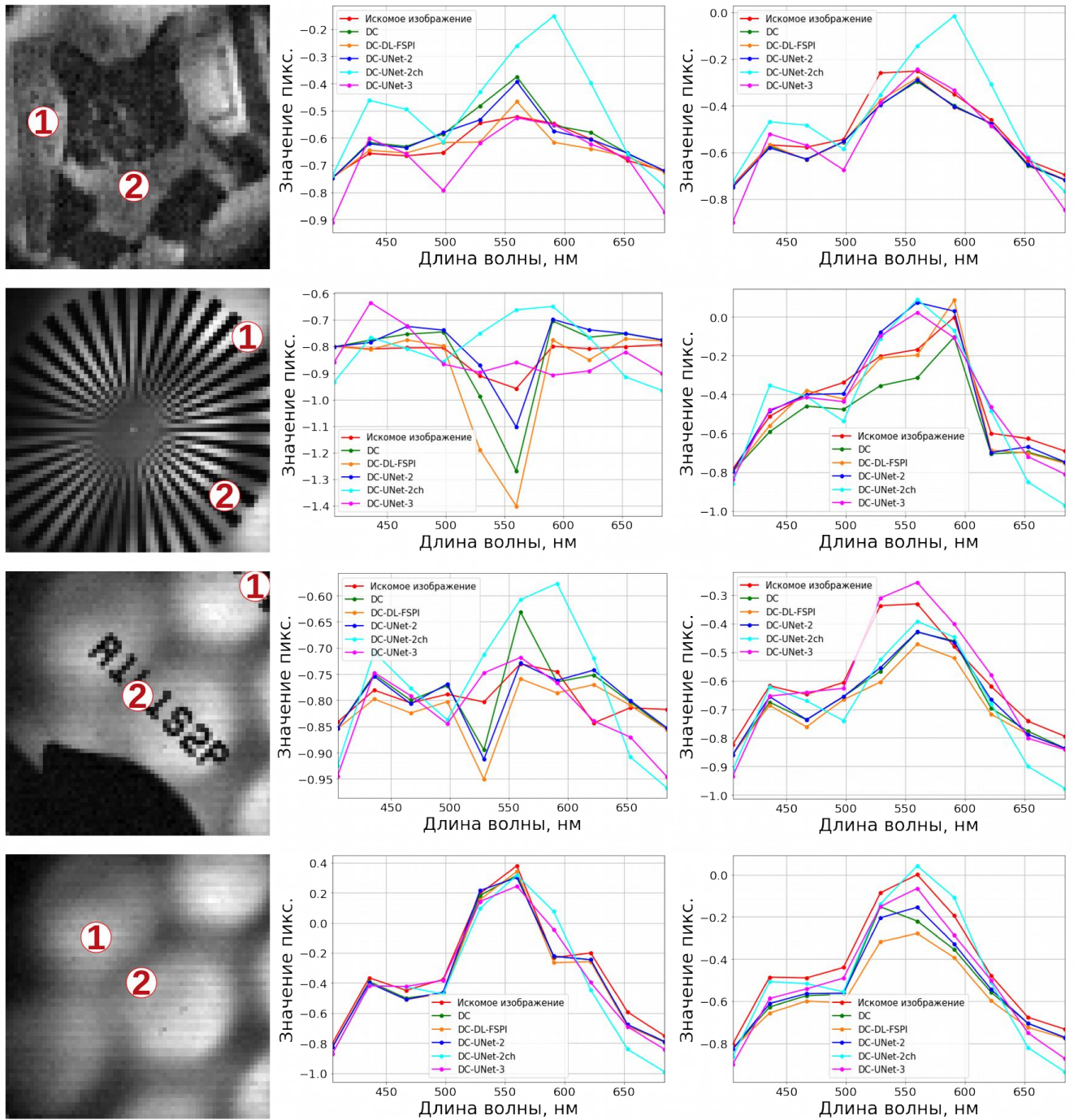
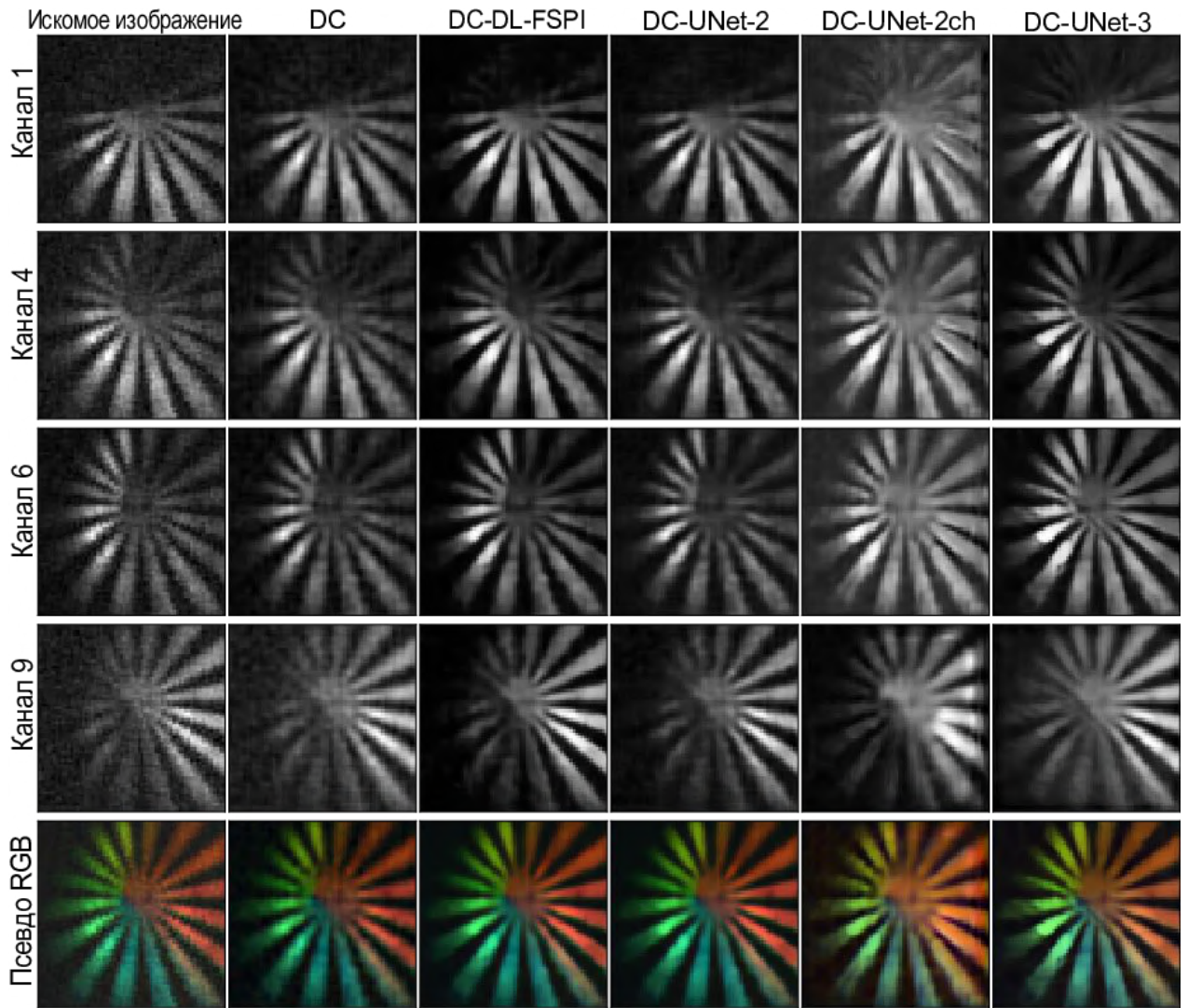


Рисунок 3.9 — Профили изменения интенсивностей пикселей по каналам в восстановленных экспериментальных изображениях.

ния большего количества каналов представлены в Приложении Г. Алгоритм DC приводит к возникновению артефактов квадратной формы на изображении, появляющихся вследствие использования недостаточного количества паттернов Адамара, а также к возникновению зернистых артефактов из-за наличия пуассоновского шума. В то же время все остальные методы, использующие регуляризацию решения, приводят к более гладким восстановленным изображениям. Интересно, что в результате применения DC-DL-FSPI изображения получаются



Нижняя строка показывает псевдо-RGB изображения, полученные из гиперспектральных изображений для каждого алгоритма в соответствии с методом, описанным в [153].

Рисунок 3.10 — Результаты восстановления экспериментального цветного изображения звезды Сименса.

менее шумными, чем при использовании DC-UNet-2. По сравнению с DC-UNet-2 и DC-DL-FSPI, DC-UNet-2ch и DC-UNet-3 позволяют сохранять некоторые детали изображения, утерянные в процессе восстановления первыми двумя алгоритмами. Например, в изображении 4-го канала и 6-го канала, где верхние лучи звезды стерты при реконструкции с использованием методов, не учитывающих спектральную регуляризацию, они остаются практически неповрежденными после восстановления методами, учитывающими гиперспектральное изображение целиком. Кроме того, в 9-м канале восстановление с помощью DC-DL-FSPI и DC-UNet-2 приводит к практически полному исчезновению лучей звезды сле-

ва внизу, в отличие от результатов, полученных с помощью DC-UNet-3. В то же время, применение DC-UNet-2ch приводит к появлению нежелательных артефактов во всех каналах изображения. Всё вышеперечисленное показывает, что DC-UNet-3 является лучшей моделью из предложенных за счёт использования избыточности информации в спектральных каналах. Для визуального сравнения полученных результатов путём преобразования гиперспектрального изображения в изображение из цветового пространства XYZ, а затем из XYZ в RGB, также вычислено псевдо-эталонное цветное изображение [154]. Сравнивая все изображения между собой, можно сделать вывод о том, что все пять методов справились с задачей восстановления цвета. Тем не менее, все алгоритмы осуществили небольшой сдвиг в цветовом спектре. В частности, красные лучи, корректно восстановленные с помощью DC, DC-UNet-2 и DC-DL-FSPI, оказываются смещёнными в оранжевую область при реконструкции с помощью DC-UNet-2ch и DC-UNet-3. Аналогичная ситуация наблюдается с лучами голубого цвета, которые при восстановлении алгоритмом DC, DC-UNet-2 и DC-DL-FSPI отображаются как более синие, а при восстановлении с помощью DC-UNet-2ch и DC-UNet-3 смещены ближе к зелёной области. Однако с точки зрения пространственного разрешения лучи, восстановленные с помощью DC-UNet-3, имеют самую высокую чёткость, особенно в верхней левой части изображения.

3.10 Заключение главы

В настоящей главе рассмотрена задача восстановления гиперспектральных изображений из однопиксельных измерений. Исследованы методы восстановления изображения с помощью минимизации критерия среднеквадратичной ошибки и регуляризации по Тихонову, а также методы, объединяющие эту классическую с точки зрения алгоритмов реконструкции изображений схему с современными методами глубокого обучения для дополнительной регуляризации полученного решения. В данной главе предложено несколько алгоритмов, выполняющих регуляризацию решения для гиперспектральных изображений как исключительно в пространственной области, так и с использованием информации из спектральной области.

Разработанные методы включают в себя две основные схемы – DC-UNet-2 и DC-UNet-3. Обе схемы направлены на решение обратной задачи восстановления изображения. Особенностью полученного однопиксельного измерения является то, что оно получено с использованием меньшего, чем число пикселей искомого изображения, количества паттернов Адамара, а также искажено шумом. Первая схема осуществляет устранение шума и дополнение недостающих однопиксельных измерений и отображение в область изображений с помощью фиксированного полносвязного слоя нейронной сети с дополнительной регуляризацией решения исключительно в пространственной области с помощью двумерной сети UNet. Вторая схема реализует те же самые шаги, но регуляризация в её случае происходит как в пространственной, так и в спектральной областях с помощью трёхмерной сети UNet.

Как итог, эксперименты показывают, что использование трёхмерной UNet улучшает качество восстановления изображений по сравнению с двумерной UNet, действующей только в пространственной области, и двумерной UNet, действующей в пространственной и спектральной областях. Предлагаемая трёхмерная архитектура демонстрирует превосходные результаты как с точки зрения метрик восстановления, так и визуально. Таким образом можно сделать вывод о том, что использование межканальной корреляции с помощью регуляризации в спектральной области позволяет достичь наилучших результатов восстановления гиперспектральных изображений из однопиксельных измерений.

Глава 4. Разделение материалов в изображениях спектральной компьютерной томографии с помощью методов глубокого обучения

4.1 Спектральная компьютерная томография

Рентгеновские лучи, открытые в 1895 году Вильгельмом К. Рентгеном, играют важную роль в медицинской визуализации. Они позволяют обнаруживать и анализировать форму и функциональность внутренних структур тела по проекционным изображениям. Появившаяся в 1970 г. рентгеновская компьютерная томография (КТ) усовершенствовала технологию, предоставляя трехмерные изображения распределения линейного коэффициента ослабления внутри пациента, точно очерчивая органы и ткани [155]. Дальнейшее усовершенствование этой технологии путём внедрения двух источников с различными энергиями излучения позволило выделение информации о различных материалах [156; 157].

Новые поколения сканеров для КТ используют детекторы, осуществляющие подсчёт числа фотонов и определяющие их энергию в результате взаимодействия с различными веществами [155; 158]. Технология КТ, применяющая такие детекторы, называется *спектральной компьютерной томографией*. Дополнительная информация об энергиях фотонов позволяет выделение отдельных изображений для различных материалов [157; 159]. Разделение или декомпозиция материалов применяется, например, для количественного определения содержания кальция в костной ткани, состава камней в почках, концентрации железа в печени, состава костного мозга, визуализации брюшной полости [160], оценки риска рака молочной железы [161], а также определения количества контрастного вещества.

Изображения в компьютерной томографии представляют собой результат ослабления рентгеновского излучения, зарегистрированный с помощью детекторов при прохождении через объект. При использовании детекторов с функцией подсчёта фотонов такие изображения получаются для каждого выбранного диапазона энергий фотонов.

Существуют различные методы восстановления спектральных КТ-изображений. Некоторые из них основаны на методике compressive sensing, или опознании по сжатию, алгоритмах полной вариации функции (total variation, TV) [162—164], низкоранговых методах [165—167] и др. [168—170]. В целом методы восстановления отдельных изображений для каждого материала принято разделять на три группы: методы, действующие в области проекций или синограмм [171; 172], методы, действующие в области изображений [173; 174], и одношаговые методы инверсии [175; 176]. Областью проекций принято называть набор полученных в результате исследования проекционных изображений для каждого угла визуализации. Такие изображения содержат информацию обо всех материалах. В случае спектральной компьютерной томографии эти изображения получаются для каждого энергетического диапазона. Синограмма представляет собой двумерную картину распределения одномерных проекций слоя объекта как функцию от угла проекции [177]. Область изображений составляют томографические реконструкции, полученные из проекций, описанных выше. Такие трёхмерные реконструкции объединяют информацию из всех КТ-проекций и материалов, но являются отдельными для каждого энергетического диапазона. Методы, действующие в области изображений, сначала восстанавливают томографические изображения для каждого энергетического диапазона. Затем полученные изображения подвергаются разделению на отдельные для каждого материала трёхмерные изображения. Основным недостатком такого подхода является то, что он делает существенные предположения относительно прямой модели, описывающей процесс формирования изображения. Неточность формулировки прямой модели может привести к возникновению артефактов в результирующих изображениях [171]. Одношаговые методы инверсии, как следует из названия, восстанавливают изображения каждого материала из измеренных данных за один шаг, что делает их вычислительно крайне затратными методами [172; 178]. Методы, действующие в области проекций, преобразовывают общие для всех материалов изображения КТ-проекций в отдельные проекционные изображения материалов. Полученные изображения затем подвергаются томографической реконструкции. Преимуществом такого подхода, как и в случае с одношаговыми методами инверсии является то, что они учитывают физический процесс формирования измерений. Ещё одним достоинством такого

подхода является возможность независимой декомпозиции каждой проекции, что позволяет распараллеливание и ускорение процесса [179].

Самыми современными методами решения задачи разделения материалов на КТ изображениях являются методы оптимизации [175; 179–182]. Для реализации декомпозиции в области проекций был предложен алгоритм Гаусса-Ньютона с регуляризацией [179]. В [182] авторы проанализировали схему регуляризации, основанную на итеративном методе Брегмана. Тем не менее, оптимизационные методы имеют ряд недостатков. Во-первых, процесс вычисления может быть медленным из-за итеративного характера алгоритмов. Во-вторых, эффективность алгоритмов зависит от наличия информации об энергетической характеристике сканера и от выбора функционала регуляризации. Энергетическая характеристика сканера в свою очередь может быть ухудшена из-за процессов перераспределения заряда, накопления импульса и потери энергии из-за выхода за пределы границ краевого поглощения [155]. Предположение об искаженной энергетической характеристике детектора может отрицательно сказаться на качестве восстановленного изображения [33]. С другой стороны, выбор специфического функционала регуляризации для декомпозиции материалов, как показали исследования, улучшает качество изображения по сравнению с нерегуляризованными подходами [179].

Методы глубокого обучения и в частности свёрточные нейронные сети нашли применение в задачах обработки изображений компьютерной томографии [183; 184]. Некоторыми авторами было предложено применение методов глубокого обучения для выделения изображений различных материалов в спектральной КТ в области изображений. Например, авторы работы [185] обучили сеть UNet на синтетических данных, имитирующих различные комбинации материалов, и протестировали полученную модель на реальных экспериментальных данных. В работе [164] сеть VGG-16 была протестирована на фантоме Шеппа-Логана и цилиндрических фантомах.

В настоящей работе исследована проблема выделения отдельных томографических изображений для различных материалов из изображений спектральной компьютерной томографии, а также решение этой задачи с помощью методов глубокого обучения. Процесс декомпозиции материалов осуществлён как в области КТ-проекций, так и в области восстановленных томографиче-

ских изображений, а результаты сравнены с классическим оптимизационным алгоритмом – методом Гаусса-Ньютона.

4.2 Формирование изображений в спектральной компьютерной томографии

4.2.1 Прямая модель формирования измерений в спектральной компьютерной томографии

Рассмотрим двумерный датчик с N элементами, соответствующими пикселям формируемого изображения при визуализации объекта, состоящего из M материалов и V вокселей. Воксель – это элемент в объёмном изображении, аналог двумерного пикселя в трёхмерном пространстве. Количество проекций, используемых для визуализации объекта в компьютерной томографии, обозначено Φ , а количество диапазонов энергий фотонов – B . Прямая модель формирования вектора измерений \mathbf{y} соответствует

$$\mathbf{y} = \mathcal{K}(\mathbf{x}), \quad (4.1)$$

где $\mathbf{y} = (y_{1,1}^1, \dots, y_{b,n}^\phi, \dots, y_{B,N}^\Phi)^\top$ и $y_{b,n}^\phi$ соответствует количеству фотонов с энергией в диапазоне b , зарегистрированному в пикселе n при визуализации КТ-проекции ϕ . Трёхмерный визуализируемый объект может быть представлен как вектор $\mathbf{x} = (x_{1,1}, \dots, x_{m,v}, \dots, x_{M,V})^\top$, содержащий значения массовой плотности, где $x_{m,v}$ – массовая плотность вещества m в вокселе v . В Формуле 4.1 \mathcal{K} представляет собой совокупность линейного рентгеновского преобразования и нелинейного оператора спектрального смешивания. Рентгеновское преобразование, обозначенное как \mathcal{H} , применяется к каждому материалу по отдельности, то есть

$$\mathbf{a}_m = \mathcal{H}(\mathbf{x}_m), \quad 1 \leq m \leq M. \quad (4.2)$$

Здесь $\mathbf{a}_m = (a_{m,1}^1, \dots, a_{m,n}^\phi, \dots, a_{m,N}^\Phi)^\top$ – вектор значений массовой плотности вещества m , полученный в результате визуализации всех проекций. Это зна-

чение также называется спроецированной массовой плотностью. Нелинейное спектральное смешивание, обозначенное как \mathcal{G} , применяется к каждому виду ϕ по отдельности, то есть

$$\mathbf{y}^\phi = \mathcal{G}(\mathbf{a}^\phi), \quad 1 \leq \phi \leq \Phi, \quad (4.3)$$

где $\mathbf{a}^\phi = (a_{1,1}^\phi, \dots, a_{m,n}^\phi, \dots, a_{M,N}^\phi)^\top$ и $\mathbf{y}^\phi = (y_{1,1}^\phi, \dots, y_{i,n}^\phi, \dots, y_{I,N}^\phi)^\top$. Операция нелинейного спектрального смешивания может быть выражена согласно выражению [180]

$$y_{b,n}^\phi = \int_{E \in \mathcal{E}} s_0(E) d_{b,n}(E) \exp \left[- \sum_{m=1}^M a_{m,n}^\phi \tau_m(E) \right] dE, \quad (4.4)$$

в котором \mathcal{E} – диапазон энергий рентгеновского излучения, $s_0(E)$ – энергетический спектр источника рентгеновского излучения, $d_{b,n}(E)$ – функция отклика детектора в пикселе p для энергетического диапазона b , а $\tau_m(E)$ – зависящий от энергии массовый коэффициент ослабления материала m .

Модель формирования вектора измерений спектральной КТ (Формула 4.1) не является непосредственным вариантом Формулы 1, так как состоит из двух преобразований. Тем не менее, рентгеновское преобразование (Формула 4.2), входящее в состав Формулы 4.1, является линейным, и проблема восстановления изображений в спектральной КТ была рассмотрена в рамках задач восстановления изображений, полученных согласно Формуле 1.

4.2.2 Виртуальная модель формирования измерений спектральной компьютерной томографии

Для решения задачи разделения материалов в области изображений необходимо рассмотреть иную “прямую” модель формирования вектора измерений \mathbf{u} спектральной КТ. В данном контексте модель не является истинно прямой, так как применяется для формулировки этапов решения обратной задачи в области изображений, но не для моделирования измерений.

Первый этап виртуальной модели формирования измерений состоит в формировании набора томографических изображений $\boldsymbol{\mu}$ из вектора значений массовой плотности всех материалов, то есть искомого объекта \mathbf{x} , как

$$\boldsymbol{\mu} = \tilde{\mathcal{G}}(\mathbf{x}). \quad (4.5)$$

Здесь вектор $\boldsymbol{\mu} = (\boldsymbol{\mu}_1^\top, \dots, \boldsymbol{\mu}_b^\top, \dots, \boldsymbol{\mu}_B^\top)^\top$ состоит из томографических изображений для каждого энергетического диапазона b , а $\tilde{\mathcal{G}}$ – оператор спектрального смешивания в области изображений. Каждое отдельное томографическое изображение $\boldsymbol{\mu}_b$ несёт в себе информацию обо всех составных материалах. Отметим, что оператор $\tilde{\mathcal{G}}$ не может быть сформулирован в аналитической форме и необходим только для формулирования обратной задачи.

На втором этапе изображение $\boldsymbol{\mu}_b$ с помощью рентгеновского преобразования трансформируется в набор КТ-проекций $\tilde{\mathbf{y}}_b$, соответствующих энергетическому диапазону b , то есть

$$\tilde{\mathbf{y}}_b = \mathcal{H}(\boldsymbol{\mu}_b). \quad (4.6)$$

Общая формула, соответствующая виртуальной “прямой” модели формирования измерений для работы в области изображений, может быть выражена композицией линейного рентгеновского преобразования \mathcal{H} и нелинейного спектрального смешивания в области изображения $\tilde{\mathcal{G}}$. В Формуле 4.6 вектор измерений $\tilde{\mathbf{y}}_b$ подвергается нормализации по проекциям в соответствии с $\tilde{\mathbf{y}}^\phi = \ln(\mathbf{y}^0/\mathbf{y}^\phi)$, где $\mathbf{y}^0 = \mathcal{G}(0)$ – вектор измерения при отсутствующем объекте. Такая нормализация применена для удобства восстановления томографических изображений с помощью нейронной сети – этот процесс описан далее по тексту.

4.3 Восстановление томографических изображений для отдельных материалов

4.3.1 Разделение материалов в области проекций

Задача разделения материалов в области проекций и получения томографических изображений для каждого вещества состоит из двух этапов. На первом этапе полученные для каждого диапазона энергий b и вида ϕ проекционные изображения спектральной КТ преобразованы в изображения спроецированной массовой плотности для каждого отдельного материала m и вида ϕ . Затем с помощью томографической реконструкции полученные изображения трансформированы в трёхмерные изображения для каждого отдельного материала.

Классический оптимизационный подход. Классический подход к решению задачи декомпозиции материалов и получения проекционных изображений для каждого вещества по отдельности состоит в минимизации функции вида

$$L(\mathbf{a}^\phi) = \|\mathbf{y}^\phi - \mathcal{G}(\mathbf{a}^\phi)\|_{\mathbf{W}^\phi}^2 + \gamma \sum_m r_m(\mathbf{a}^\phi). \quad (4.7)$$

Здесь первый член является взвешенным функционалом, сформулированным по методу наименьших квадратов, $\mathbf{W}^\phi = \text{diag}(1/\sqrt{\mathbf{y}^\phi})$ – матрица весов, diag обозначает диагональную матрицу, на главной диагонали которой находятся элементы $1/\sqrt{\mathbf{y}^\phi}$, γ – регуляризационный коэффициент, контролирующий степень включения регуляризационного члена $r_m(\mathbf{a}^\phi)$, специфического для каждого вещества m . Решение для выражения 4.7 описано в работе [179] и найдено с помощью алгоритма Гаусса-Ньютона (RGN, **R**egularized **G**auss-**N**ewton), реализованного в программе SPRAY [186].

Метод глубокого обучения. Восстановление проекционных КТ изображений для каждого отдельного материала с помощью нейронной сети представляет собой задачу отображения $\tilde{\mathbf{y}}^\phi \mapsto \tilde{\mathbf{a}}^\phi$. Здесь $\tilde{\mathbf{y}}^\phi = \ln(\mathbf{y}^0/\mathbf{y}^\phi)$ и $\tilde{a}_{m,n}^{\phi,s} = a_{m,n}^{\phi,s} / \max_{n,\phi,s} a_{m,n}^{\phi,s}$, $1 \leq s \leq S$, S – количество элементов в тренировочном наборе. Тренировка нейронной сети, осуществляющей отображение, происходит путём минимизации разницы между нормализованным эталонным изображением $\tilde{\mathbf{a}}_m$ и полученным на выходе сети изображением $\hat{\mathbf{a}}_m$. Минимизируемая разница,

или функция потерь, определена как

$$L = \sum_{m=1}^M \|\hat{\mathbf{a}}_m - \tilde{\mathbf{a}}_m\|_2^2. \quad (4.8)$$

На вход сети поступают изображения размера $N_x \times N_y \times B$, а на выходе формируются изображения размера $N_x \times N_y \times M$. Здесь $N_x \times N_y$ – пространственный размер изображения, B – количество энергетических диапазонов и M – количество материалов. Минимизация функции 4.8 осуществлена с помощью алгоритма Adam [63] с использованием скорости обучения, равной 10^{-4} . Размер тренировочного набора равен 16.

Архитектура сети UNet. Архитектура описываемой в данной главе 2D UNet повторяет архитектуру, описанную в Разделе 3.5 Главы 3. В случае отображения изображений спектральной компьютерной томографии в изображения разделённых материалов на вход нейронной сети поступают изображения с пространственным размером $N_x \times N_y$ и количеством признаков карт B , соответствующим количеству энергетических диапазонов. На выходе сети формируются изображения такого же пространственного размера, как и на входе, содержащие M признаков карт, число которых соответствует числу материалов. Общее число пар тренировочных данных соответствует общему количеству КТ-проекций во всех изображениях. Полученная модель обозначена UNet-P (UNet-Projection domain).

В результате первого этапа разделения материалов с помощью метода RGN или с помощью UNet получены изображения спроецированной массовой плотности для каждого вещества m и для каждого вида ϕ . На следующем шаге эти изображения подвергаются томографической реконструкции, в ходе которой все виды для каждого материала объединены в трёхмерное томографическое изображение, соответствующее вектору значений массовой плотности \mathbf{x}_m . Томографическая реконструкция выполнена с помощью метода обратного проецирования с фильтрацией (filtered back projection, FBP), представляющего собой стабилизированную и дискретизированную версию обратного преобразования Радона [187] и реализованного в программе SPRAY [186].

Схема процесса восстановления томографических изображений для каждого материала из проекционных изображений для отдельных диапазонов энергии, осуществлённого в области проекций, представлена на Рисунке 4.1а.

4.3.2 Разделение материалов в области изображений

Как описано в Разделе 4.2.2, для разделения материалов в области изображений использована виртуальная модель формирования вектора КТ-измерений. Для того, чтобы восстановить томографические изображения для каждого материала, соответствующие \mathbf{x}_m , необходимо так же, как и в предыдущем подходе, решить двухэтапную задачу. На первом шаге необходимо восстановить томографические изображения $\boldsymbol{\mu}_b$, а затем полученные изображения перевести в томографические изображения \mathbf{x}_m . Первый шаг осуществлён образом, описанным в Разделе 4.3.1, то есть с помощью метода обратного проецирования с фильтрацией, применённого к изображениям для каждого энергетического диапазона. Второй этап, представляющий собой преобразование набора томографических изображений $\boldsymbol{\mu}$ в томографические изображения для материала, не может быть выражен аналитически с помощью оператора $\tilde{\mathcal{G}}$ из Формулы 4.5. Вместо этого для отображения $\boldsymbol{\mu}^j \mapsto \mathbf{x}^j$ использована свёрточная нейронная сеть. Здесь $j, 1 \leq j \leq J$, обозначает номер среза в трёхмерном томографическом изображении.

Аналогичным описанному в Разделе 4.3.1 образом, тренировка нейронной сети происходит путём минимизации разницы между эталонным изображением \mathbf{x}_m и полученным на выходе сети изображением $\hat{\mathbf{x}}_m$. Минимизируемая разница, или функция потерь, определена как

$$L = \sum_{m=1}^M \|\hat{\mathbf{x}}_m - \mathbf{x}_m\|_2^2. \quad (4.9)$$

Архитектура сети UNet. В случае отображения томографических изображений спектральной КТ в томографические изображения разделённых материалов на вход нейронной сети поступают изображения томографических срезов с пространственным размером $Q_x \times Q_y$ и количеством признаков карт B , соответствующим количеству энергетических диапазонов. На выходе сети формируются изображения такого же пространственного размера, как и на входе, содержащие M признаков карт, число которых соответствует числу материалов. Общее число пар тренировочных данных соответствует общему количеству томографических срезов во всех изображениях. Таким образом, во время

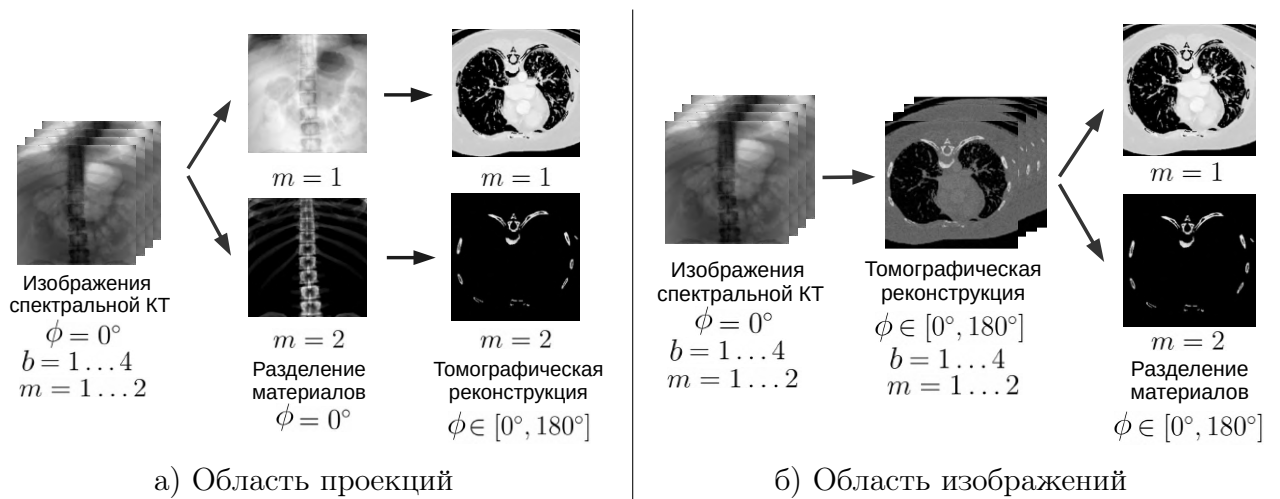


Рисунок 4.1 — Схема реконструкции томографических изображений для энергетических диапазонов и последующей декомпозиции материалов в области изображений. В примере использовано два материала – мягкие ткани (1) и костные ткани (2) – и четыре энергетических диапазона.

обучения на вход сети поступает набор из B двумерных томографических срезов, по одному из каждого томографического изображения, полученного для каждого энергетического диапазона. Тренировка сети осуществлена аналогичным описанному в Разделе 4.3.1 образом. Полученная модель обозначена UNet-I (UNet-Image domain).

Схема процесса восстановления томографических изображений для каждого материала из проекционных изображений для отдельных диапазонов энергии, осуществлённого в области изображений, представлена на Рисунке 4.1б.

4.4 Моделирование данных спектральной компьютерной томографии

На первом этапе работы рассмотрена задача разделения материалов в области проекций с помощью модели UNet-P и сравнение с результатами, полученными с помощью метода RGN. Для обучения нейронной сети использованы томографические изображения из базы данных 3D-IRCADb [188] и анатомического атласа, содержащего сегментированные томографические изображения [189]. Спроецированные массовые плотности для трёх материалов – мягких тканей, костной ткани и контрастного вещества гадолиния – получены из томо-

графических изображений с помощью рентгеновского преобразования согласно Формуле 4.2. Затем изображения спектральной компьютерной томографии для четырёх энергетических диапазонов получены согласно Формуле 4.4. Функция отклика детектора смоделирована согласно параметрической модели, описанной в [180], а функция, описывающая энергетический спектр источника рентгеновского излучения смоделирована с помощью программы SpekCalc [190]. Полученные данные искажены пуассоновским шумом, исходя из предположения, что максимальное число фотонов, попадающих на центральный элемент детектора, составляет 10^8 . Всего смоделировано 360 проекций, из которых для работы выбрано 180. Из этих 180 проекций для тренировки UNet выбрано 135, а 45 отложено для тестирования. Проекция, используемые для обучения модели, подвергнуты симметричному отражению для увеличения количества изображений в тренировочном датасете до 540. Моделирование проекций произведено с помощью программы SPRAY [186]. Обучение модели произведено описанным в Разделе 4.3.1 образом.

Для реализации дополнительного сравнения методов разделения материалов в области проекции, а также для обучения модели UNet-I и сравнения результатов декомпозиции материалов в области томографических изображений использован большой датасет. Набор измерений спектральной компьютерной томографии смоделирован с использованием трёхмерных КТ-изображений, представленных в наборе данных для сегментации опухолей в почках KiTS19 [191]. Для тренировки всех моделей использовано 54 трёхмерных изображения и 9 изображений отложено для тестирования. Изображения подвергнуты сегментации с помощью комбинации порогового метода и морфологических операций для выделения объёмов, содержащих информацию только о мягких тканях и только о костной ткани. Затем спроецированные массовые плотности для двух материалов получены из томографических изображений аналогичным описанному выше образом. Полученные данные искажены пуассоновским шумом исходя из предположения, что максимальное число фотонов, попадающих на центральный элемент детектора, составляет $6 \cdot 10^5$. Таким образом смоделировано 360 КТ-проекций.

Набор данных для обучения UNet-P, действующей в области проекций, состоит из 19440 проекций (360 проекций для 54 трёхмерных томографических изображений). Набор данных для обучения UNet-I, действующей в области

изображений, состоит из 5400 изображений (100 срезов для 54 трёхмерных томографических изображений). В обоих случаях с целью увеличения количества тренировочных данных используются техники приращения данных, основанные на кадрировании изображений. В случае работы в области изображений применены различные степени включения костной ткани в изображение [34].

4.5 Результаты восстановления томографических изображений для отдельных материалов

4.5.1 Метрики оценки результатов разделения материалов

Данная работы была выполнена в сотрудничестве с исследовательской группой из лаборатории исследования медицинских изображений CREATIS в Лионе. Традиционно при решении задач восстановления томографических изображений этой группой для оценки качества восстановления изображений применяются метрики нормализованного Евклидова расстояния и среднеквадратической ошибки (mean squared error, MSE). По этой причине в настоящей главе для оценки качества разделения материалов использованы указанные метрики. Во второй части работы коллегами был дополнительно посчитан индекс структурного сходства (structural similarity index measure, SSIM). Результаты представлены в виде диаграмм размаха.

4.5.2 Результаты восстановления изображений из небольшого набора данных

На Рисунке 4.2 представлены результаты декомпозиции трёх материалов из изображений спектральной компьютерной томографии, полученных для четырёх энергетических диапазонов. Для разделения материалов использованы метод RGN с коэффициентом регуляризации $\gamma = 0.6$ и нейронная сеть UNet,

обученная отображению данных спектральной компьютерной томографии \mathbf{y}^b в изображения спроецированной массовой плотности $\hat{\mathbf{a}}_m$ для трёх материалов. Из результатов, представленных на Рисунке 4.2, видно, что в то время, как качество изображений мягких тканей и контрастного вещества примерно одинаково для двух методов, UNet-P значительно превосходит RGN в восстановлении изображения костной ткани. Результаты декомпозиции материалов также оценены посредством вычисления нормализованного Евклидова расстояния, или ошибки $\frac{\|\hat{\mathbf{a}}_m - \hat{\mathbf{a}}_m\|_2}{\|\hat{\mathbf{a}}_m\|_2}$, между эталонным и восстановленным изображениями, значения которого представлены в Таблице 6. Из результатов, представленных в Таблице 6, Таблица 6 — Сравнительная таблица метрик качества восстановления изображений спроецированной массовой плотности. Наилучший результат выделен жирным шрифтом.

	Ошибка RGN, · 100 %	Ошибка UNet-P, · 100 %
Мягкие ткани	1.21	0.46
Костные ткани	24.60	2.26
Контрастное вещество	8.45	1.00

видно, что использование UNet-P в задаче декомпозиции материалов привело к снижению относительной ошибки для изображений мягких тканей на 62%, костных тканей – на 90%, контрастного вещества – на 88%.

Полученные 180 проекций спроецированной массовой плотности трёх материалов затем подвергнуты томографической реконструкции с помощью метода обратного проецирования с фильтрацией. Предполагается, что первые 180 проекций симметричны вторым 180 проекциям, поэтому для реконструкции использовано суммарно 360 проекций. Один скадрированный срез трёхмерного томографического изображения каждого материала представлен на Рисунке 4.3. Качество томографической реконструкции оценено с помощью стандартной среднеквадратической ошибки (MSE), значения которой представлены в Таблице 7. Несмотря на то, что наилучшие изображения спроецированной массовой плотности получены с помощью UNet-P, томографическая реконструкция этих изображений привела к появлению артефактов в виде полос. По всей видимости, такие артефакты возникают в результате ошибки нейронной сети при декомпозиции материалов в проекциях, которые не были использованы в про-

цессе тренировки. Одним из способов преодоления этого недостатка является использование большего количества тренировочных данных.

Таблица 7 — Сравнительная таблица метрик качества восстановления томографических изображений. Наилучший результат выделен жирным шрифтом.

	MSE RGN	MSE UNet-P
Мягкие ткани, $\cdot 10^3$	0.47	1.06
Костные ткани, $\cdot 10^4$	6.36	4.62
Контрастное вещество, $\cdot 10^6$	1.58	6.25

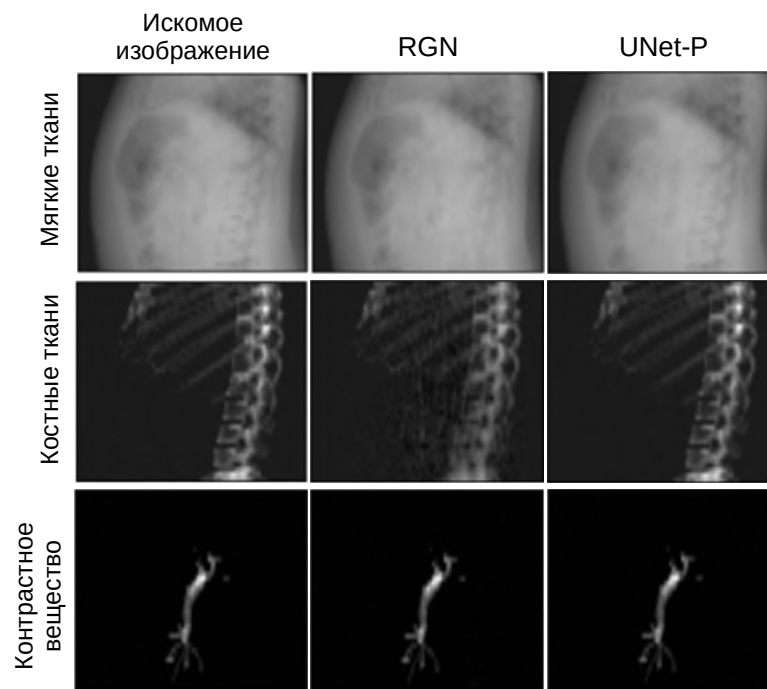


Рисунок 4.2 — Результаты декомпозиции материалов в области проекций для трёх материалов – мягких тканей, костных тканей и контрастного вещества.

Изображения соответствуют одной КТ-проекции.

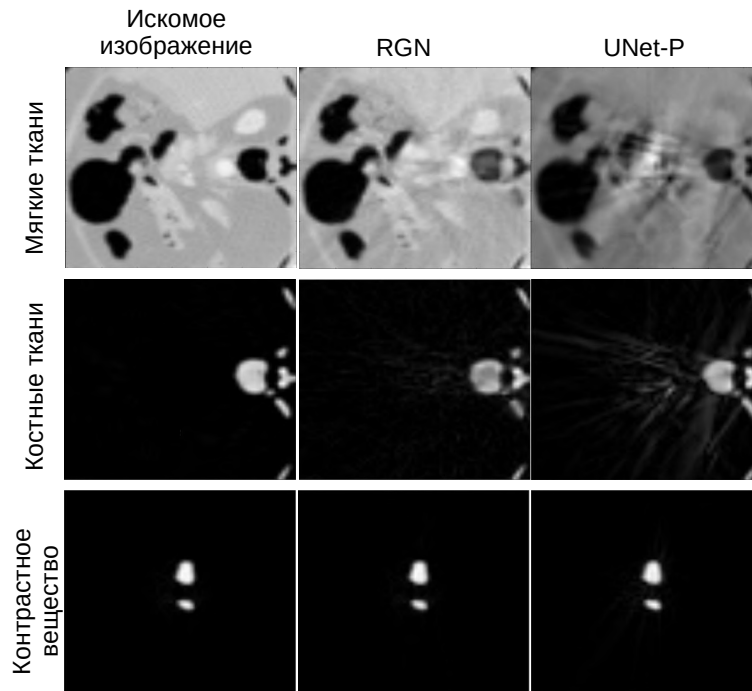


Рисунок 4.3 — Результаты реконструкции томографических изображений для трёх материалов – мягких тканей, костных тканей и контрастного вещества.

4.5.3 Результаты восстановления изображений из большого набора данных

Сравнение результатов декомпозиции материалов с помощью RGN и UNet-P, обученной на большом наборе данных, представлено на Рисунке 4.4. В частности, для алгоритма Гаусса-Ньютона применено два уровня регуляризации – слабая с $\gamma = 0.1$ и средняя с $\gamma = 0.6$. Несмотря на то, что в случае слабой регуляризации RGN позволяет достаточно точное восстановление структур и деталей, изображения спроецированной массовой плотности получаются зашумлёнными. Увеличение уровня регуляризации закономерно способствует устранению шума, но делает изображения более размытыми. В то же время UNet-P позволяет сохранять баланс между точностью восстановления деталей и устранением шума в изображениях.

На Рисунке 4.5 представлены результаты томографической реконструкции изображений каждого материала с помощью трёх методов. В отличие от алгоритма RGN и UNet-P, алгоритм UNet-I восстанавливает непосредственно томографические изображения для каждого материала. Поэтому данный алгоритм не представлен на Рисунке 4.4. На Рисунках 4.7а, 4.7б представлены

графики разбросов значений нормализованной MSE и SSIM для изображений разделённых материалов в области проекций (Рисунок 4.7а) и для изображений томографической реконструкции (Рисунок 4.7б). Отметим, что значения нормализованной MSE для костной ткани выше, чем для мягких тканей, поскольку большая часть изображения костной ткани представлена фоном. В целом применение UNet-P в области проекций снижает значение MSE и увеличивает значение SSIM по сравнению с методом RGN. Реконструкция с помощью UNet-I показывает наименьшее значение MSE и наибольшее значение SSIM. В соответствии с результатами, представленными на Рисунке 4.4, томографические изображения, полученные с использованием алгоритма RGN со слабой регуляризацией, оказываются зашумлёнными. Увеличение значения регуляризационного параметра приводит к устранению шума, но размытию деталей в изображении. Изображения, полученные с применением UNet-P и большего количества тренировочных данных, будучи лучшими визуально, тем не менее имеют артефакты в виде полос. Это происходит из-за того, что ошибки в восстановленных в области проекций изображениях усиливаются при томографической реконструкции изображений. Результаты, полученные с помощью UNet-I, превосходят все остальные томографические реконструкции и позволяют наилучшим образом воспроизвести детали искомым изображениям. Причиной этому, по всей видимости, является наличие этих деталей в томографических изображениях для каждого энергетического диапазона.

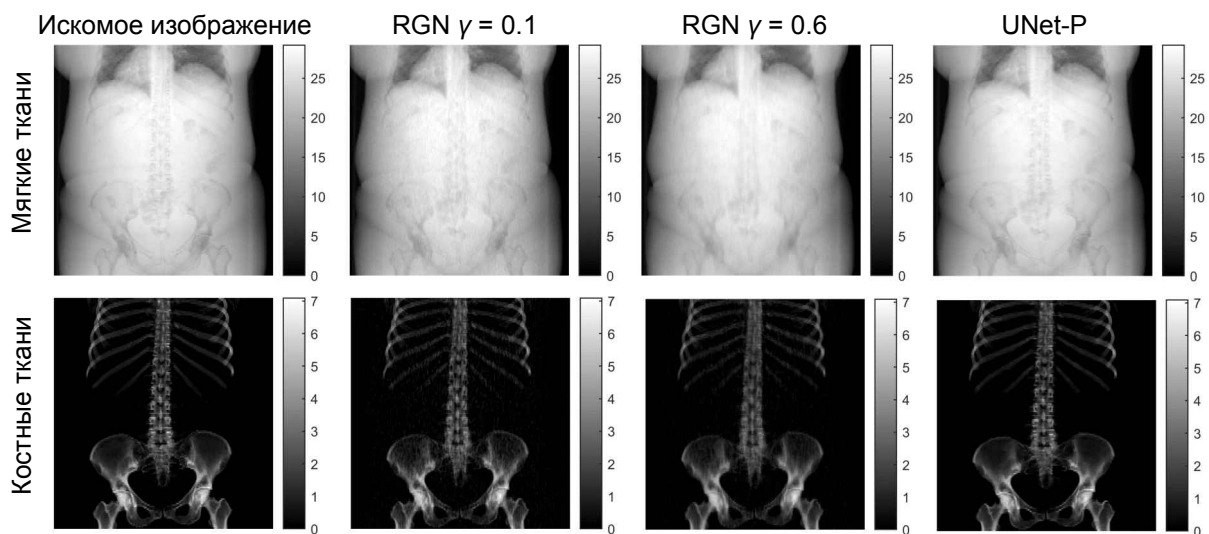


Рисунок 4.4 — Результаты декомпозиции в области проекций для двух материалов – мягких и костных тканей. Изображения соответствуют одной КТ-проекции. Результаты опубликованы в [34].

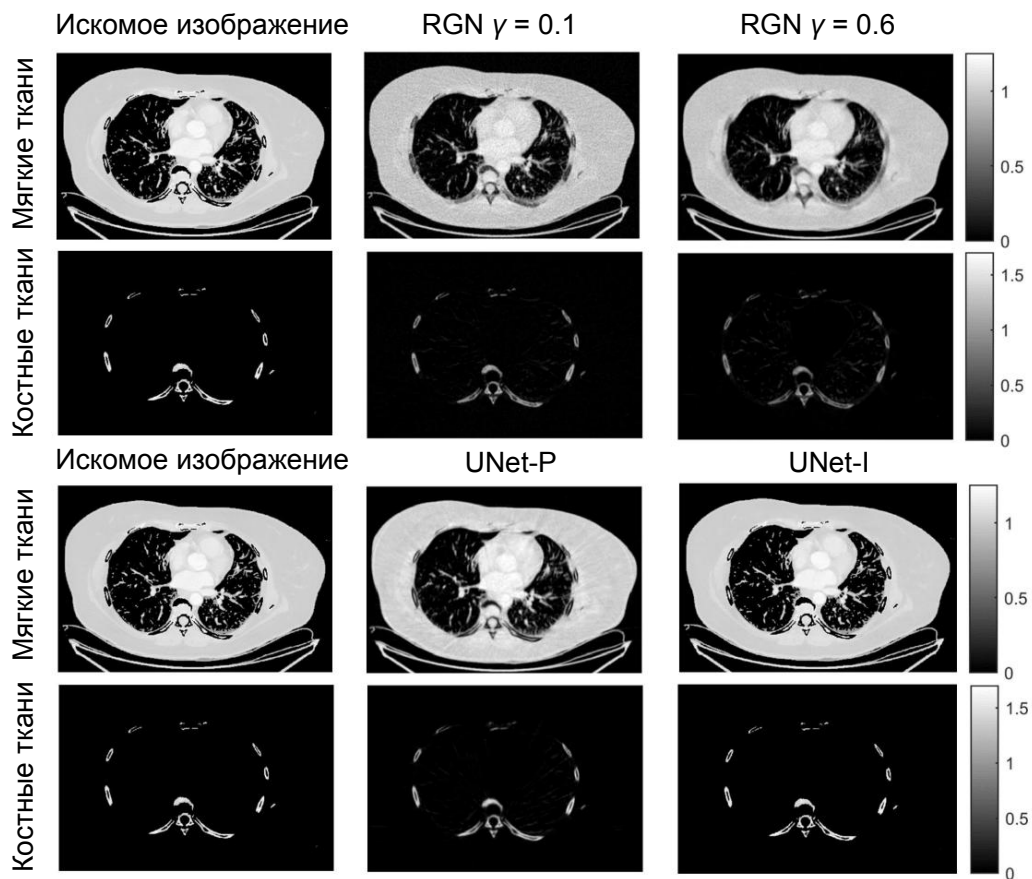


Рисунок 4.5 — Результаты реконструкции томографических изображений для двух материалов – мягких и костных тканей. Результаты опубликованы в [34].

На Рисунке 4.6 представлены результаты томографической реконструкции изображений мягких тканей и костной ткани с помощью трёх методов.

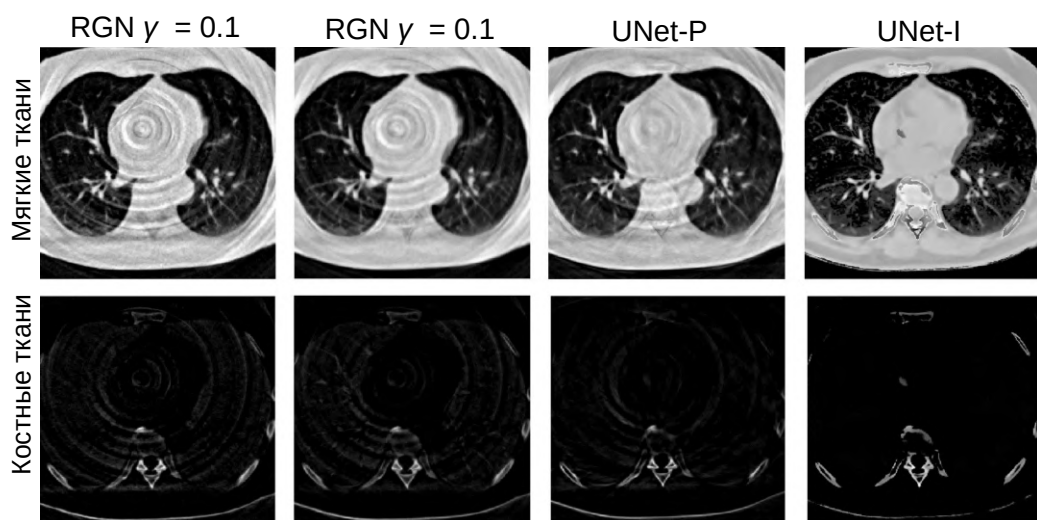
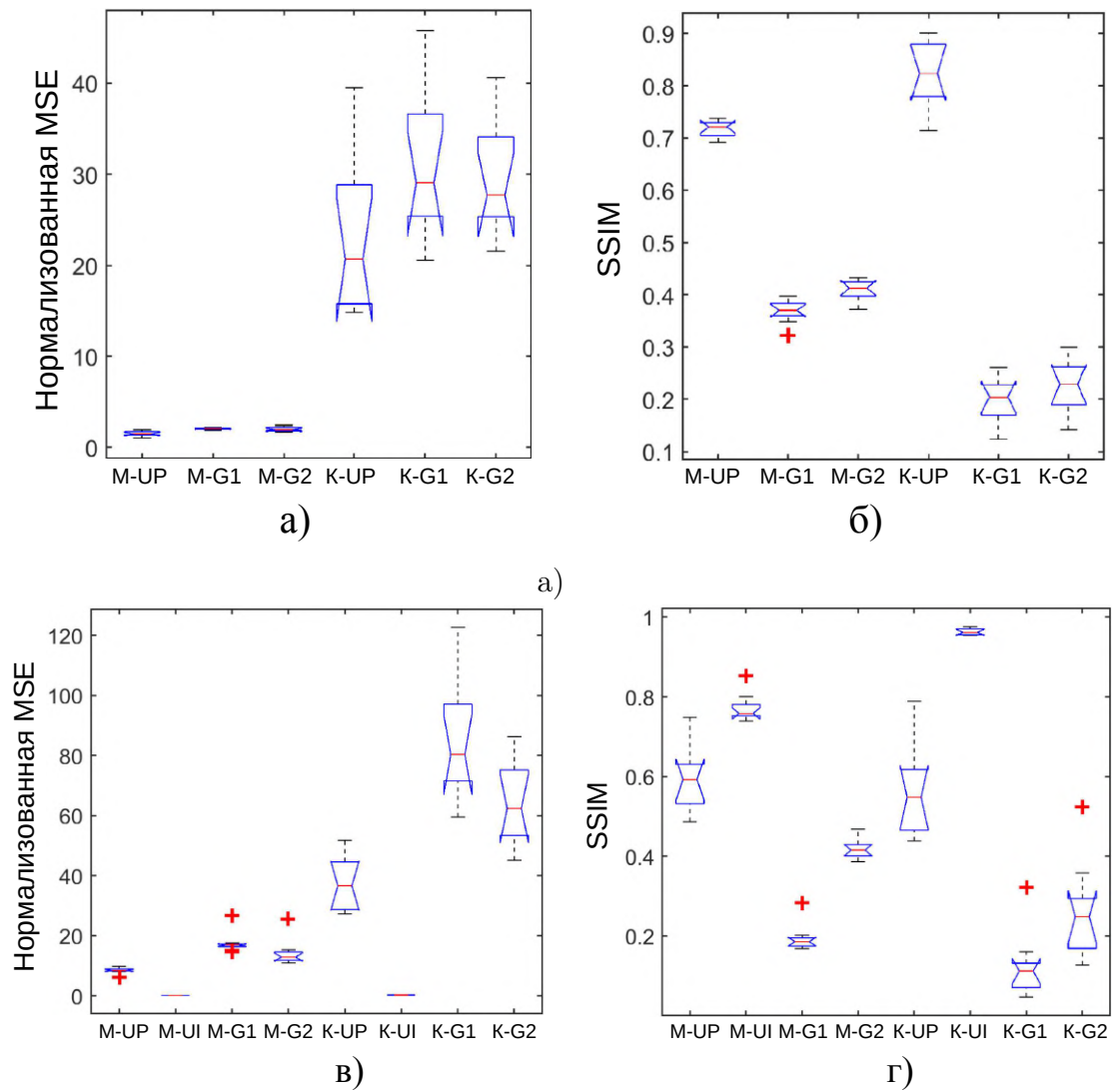


Рисунок 4.6 — Результаты реконструкции томографических изображений реального человеческого торса для двух материалов – мягких и костных тканей. Результаты опубликованы в [34].



а) б)
 RGN с $\gamma = 0.1$ – G1, RGN с $\gamma = 0.6$ – G2, UNet-P – UP, UNet-I – UI; мягкие
 ткани – М, костные ткани – К.

Рисунок 4.7 — Диаграмма размаха нормализованной MSE и SSIM декомпозиции материалов (а, б) и томографической реконструкции (в, г). Диаграммы показывают значение медианы, 25-го и 75-го перцентиля, а также минимальные и максимальные значения, исключая выбросы, которые отображаются красными крестиками. Результаты опубликованы в [34].

Представленные результаты получены с использованием спектральных КТ-изображений реального человеческого торса, полученных с помощью экспериментального спектрального компьютерного томографа [192]. Из результатов, представленных на Рисунке 4.6, видно, что среди методов, работающих в области проекций, RGN реализует изображения, сильнее искажённые шумом и артефактами ложного оконтуривания по сравнению с UNet-P. UNet-I обеспечи-

вает более высокое качество изображения, чем проекционные методы, с лучшим восстановлением деталей как мягких тканей, так и изображений костей.

4.6 Заключение главы

В настоящей главе рассмотрена задача разделения данных спектральной компьютерной томографии на томографические изображения для каждого материала, входящего в состав объекта визуализации. Рассмотрен процесс декомпозиции материалов в двух областях. Первая, так называемая область проекций, представляет собой набор изображений, полученных в процессе КТ-исследования, для различных углов визуализации и диапазонов энергии. В таких изображениях присутствуют данные спектральной КТ для всех материалов. В этом случае осуществляется выделение отдельных для каждого материала проекционных изображений, которые затем подвергаются томографической реконструкции. В результате этого получают отдельные трёхмерные томографические изображения, содержащие информацию о каждом веществе. Процесс декомпозиции материалов в этом случае осуществлён с помощью классического оптимизационного алгоритма Гаусса-Ньютона и с помощью тренируемой нейронной сети UNet-P. Применение нейронной сети показало лучшее качество разделённых в области проекций изображений по сравнению с алгоритмом Гаусса-Ньютона, а именно позволило снизить шум и сохранить детали в изображениях.

Вторая область, в которой осуществлена декомпозиция материалов, – область изображений, а именно восстановленные трёхмерные томографические изображения, объединяющие информацию обо всех материалах в каждом энергетическом диапазоне. Разделение материалов в этой области выполнено с помощью нейронной сети UNet-I. Выделенные трёхмерные томографические изображения для каждого отдельного вещества качественно превосходят изображения, полученные после декомпозиции материалов в области проекций.

Как итог, методы глубокого обучения демонстрируют большой потенциал для спектральной компьютерной томографии, превосходя результаты, полученные с помощью классических оптимизационных методов.

Заключение

В настоящей работе рассмотрены различные задачи восстановления изображений, искажённых в процессе регистрации визуализирующей системой. В ходе работы исследованы классические оптимизационные методы и современные нейросетевые подходы к восстановлению изображений, а также их комбинации. Основные результаты работы заключаются в следующем:

1. Исследование задачи восстановления размытых и зашумлённых изображений, полученных с помощью микроскопа, показало, что объединение оптимизационного алгоритма восстановления изображений и нейросетевого подхода для аппроксимации регуляризационного члена позволяет получать более высокие метрики реконструкции, чем при использовании указанных подходов по отдельности.
2. Разработан набор из последовательно усложнённых схем аппроксимации регуляризационного члена для решения задачи восстановления изображений, искажённых пуассоновским и гауссовским шумами. Некоторые из разработанных алгоритмов превосходят наилучшие существующие на данный момент подходы к решению аналогичных задач. Соответствующие обученные нейросетевые модели опубликованы в открытом доступе.
3. Исследовано и предложено два подхода к решению задачи слепой деконволюции, то есть деконволюции изображений при отсутствии информации о функции размытия точки визуализирующей системы. Согласно первому подходу, оптимизационная задача решена относительно неизвестной функции рассеяния точки. Вторым подходом реализуется итеративную схему попеременного поиска изображения, полученного с помощью микроскопа, и функции рассеяния точки микроскопа. Вычисление параметров последней выполнено с помощью тренируемых агентов. Полученные результаты для обоих методов показывают перспективу их применения к задачам слепой деконволюции.
4. Предложенный алгоритм регуляризации решения в спектральной и пространственной областях с помощью нейронной сети позволяет учитывать межканальную корреляцию при восстановлении гиперспектраль-

- ных изображений из однопиксельного измерения. Совместно с решением с помощью минимизации критерия среднеквадратичной ошибки и регуляризации по Тихонову для обесшумливания и дополнения недостающих измерений, этот подход позволяет получение наилучших визуально и с точки зрения метрик результатов реконструкции.
5. Рассмотрены решения задачи разделения материалов в изображениях спектральной компьютерной томографии, реализуемые в области КТ-проекций и области томографических изображений. Показано, что нейросетевые подходы позволяют не только разделить изображения материалов, но и практически полностью устранить шум в итоговых томографических изображениях в отличие от алгоритма Гаусса-Ньютона с регуляризацией. Кроме того, изображения, полученные вследствие разделения материалов в области изображений, наиболее точно передают детали искомым изображений.

Список сокращений и условных обозначений

- КТ** – компьютерная томография
- CNN** – convolutional neural network, свёрточная нейронная сеть
- DC** – Denoised completion, обесшумливание и дополнение
- DC-Net** – Denoised completion network, сеть для обесшумливания и дополнения
- DC-UNet** – Denoised completion UNet, сеть UNet для обесшумливания и дополнения
- DQN** – deep Q-learning, глубокое Q-обучение
- ESRGAN** – enhanced superresolution GAN
- FBP** – filtered back projection, метод обратного проецирования с фильтрацией
- FFT** – fast Fourier transform, быстрое преобразование Фурье, БПФ
- GAN** – generative adversarial network, генеративно-сопоставительная сеть
- KPN** – нейронная сеть, предсказывающая регуляризационные фильтры
- MSE** – mean squared error, средняя квадратическая ошибка
- PPO** – proximal policy optimization, проксимальная оптимизация политики
- PSF** – point spread function, функция рассеяния точки
- PSNR** – peak signal-to-noise ratio, отношение пикового сигнала к шуму
- ReLU** – rectified linear unit, активационная функция в нейронных сетях
- RL** – reinforcement learning, обучение с подкреплением
- RL+TV** – алгоритм Ричардсона-Люси с регуляризацией total variation
- RGB** – red, green, blue, красный, зелёный, синий, цветовая модель кодирования цвета
- RGN** – regularized Gauss-Newton, алгоритм Гаусса-Ньютона с регуляризацией
- SAM** – spectral angle mapper, метрика спектрального сходства
- SNR** – signal-to-noise ratio, отношение сигнала к шуму
- SOTA** – state-of-the-art, передовые методы в области
- SRGAN** – superresolution GAN
- SSIM** – structural similarity index measure, индекс структурного сходства
- TV** – total variation
- TRLK** – trainable Richardson-Lucy with learnable kernels, алгоритм Ричардсона-Люси с обучаемыми фильтрами

TRL-KPN – trainable Richardson-Lucy with kernel prediction network, алгоритм Ричардсона-Люси с фильтрами, вычисляемыми с помощью нейронной сети

TRL-UNet – trainable Richardson-Lucy with UNet, алгоритм Ричардсона-Люси с регуляризационным параметром, вычисляемым с помощью нейронной сети

UNet-I – UNet-Image domain, сеть UNet, реализованная в области изображений

UNet-P – UNet-Projection domain, сеть UNet, реализованная в области проекций

WFK – Wiener filter with learnable kernels, фильтр Винера с обучаемыми фильтрами

WF-KPN – Wiener filter with kernel prediction network, фильтр Винера с фильтрами, вычисляемыми с помощью нейронной сети

WF-KPN-SA – Wiener filter with kernel prediction network spatially adaptive, фильтр Винера с пространственно адаптируемыми фильтрами, вычисляемыми с помощью нейронной сети

WFT – Wiener filter with Tikhonov regularization, фильтр Винера с регуляризацией по Тихонову

WF-UNet – Wiener filter with UNet, фильтр Винера с регуляризационным параметром, вычисляемым с помощью нейронной сети

Обозначения

Общие обозначения

β	Скорость обучения алгоритма
\cdot^*	Обозначение комплексно-сопряжённой матрицы
\cdot^T	Обозначение транспонированной матрицы
\cdot^H	Обозначение эрмитово-сопряжённой матрицы
γ	Регуляризационный коэффициент
λ	Длина волны излучаемого света
\mathbb{C}^N	Область комплексных чисел размерности N
\mathbb{E}	Математическое ожидание
\mathbb{R}^N	Область вещественных чисел размерности N
$\mathbf{1}$	Вектор, состоящий из единиц
\mathbf{D}	Диагональная матрица
\mathbf{F}	Матрица преобразования Фурье
\mathbf{x}	Искомое изображение или объект
\mathbf{y}	Зарегистрированное при визуализации объекта измерение
\mathcal{G}	Оператор, описывающий нелинейное отображение
$\mathcal{N}(\cdot, \cdot)$	Нормальное распределение
$\mathcal{P}(\cdot)$	Распределение Пуассона
$\Omega[\cdot]$	Стабилизирующий функционал
σ	Среднеквадратическое отклонение
I, i	Количество и номер итерации алгоритма Ричардсона-Люси

K	Размер двумерной функции рассеяния точки
$L(\cdot)$	Целевая функция, подлежащая оптимизации
N, n	Количество и номер пикселя в изображении

Обозначения во введении и главе 1

α	Степень регуляризационного коэффициента
\mathbf{G}	Матрица регуляризационного фильтра
\mathbf{g}	Регуляризационный фильтр
\mathbf{H}	Матрица функции рассеяния точки микроскопа
\mathbf{h}	Функция рассеяния точки микроскопа
\mathbf{n}	Аддитивный фоновый шум системы
\mathbf{P}	Матрица, осуществляющая операцию дополнения нулями
\mathbf{S}	Матрица, осуществляющая операцию сдвига
\mathcal{F}	Преобразование Фурье
\mathbf{H}	Матрица-элемент блочной матрицы функции рассеяния точки
D, d	Количество и номер регуляризационного фильтра
J, j	Количество и номер шага алгоритма градиентного спуска
L	Размер двумерных регуляризационных фильтров
$r(\cdot)$	Функция, определяющая вид регуляризационного члена

Обозначения в главе 2

δz	Степень отклонения плоскости изображения от фокусной плоскости
Δ_x	Разрешающая способность микроскопа в поперечном направлении
Δ_z	Разрешающая способность микроскопа в продольном направлении

$\hat{A}_t(\cdot)$	Функция преимущества выполнения определённого действия по сравнению с другими
\mathbf{H}	Матрица функции рассеяния точки микроскопа
\mathbf{h}	Функция рассеяния точки микроскопа
ν	Обесценивающий коэффициент для будущих наград
NA	Числовая апертура оптической линзы
$\pi_\theta(\cdot)$	Функция, определяющая политику агента
ψ	Коэффициент масштабирования награды
ρ	Нормализованный радиус объектива оптической системы
ε	Коэффициент, ограничивающий степень отличия параметров старой политики от новой
a_t	Действие, выполняемое агентом на t шаге эпизода
C	Максимальная интенсивность в центральной части узора Эйри
d	Номер среза функции рассеяния точки
$h(r)$	Функция рассеяния точки от радиуса в плоскости изображения
$J_0(\cdot)$	Функция Бесселя первого рода нулевого порядка
k_0	Волновое число для длины волны λ
n_i	Иммерсионный показатель преломления
Q	Функция полезности
r	Радиус в плоскости изображения
r_t	Награда, получаемая агентом на t шаге эпизода
s_t	Состояние среды, наблюдаемое агентом на t шаге эпизода
T, t	Количество и номер шага в эпизоде и количество и номер итерации алгоритма слепой деконволюции

z_d	Расстояние от объектива оптической системы до плоскости изображения
z_f	Фокусное расстояние оптической линзы

Обозначения в главе 3

α	Интенсивность изображения, выраженная числом фотонов
η	Значение кривой спектральной чувствительности человеческого глаза
Λ, λ	Количество и номер гиперспектрального канала
\mathbf{H}	Базис бинарных паттернов Адамара
\mathbf{H}_1	Набор бинарных паттернов Адамара, выбранных из базиса \mathbf{H}
\mathbf{m}^α	Однопиксельное измерение, соответствующее количеству фотонов α
μ	Математическое ожидание однопиксельных измерений
Σ	Матрица ковариаций однопиксельных измерений
Σ_α	Матрица ковариаций шума
θ	Параметры нейронной сети
K	Калибровочный коэффициент передачи оптической системы
L, l	Количество и номер слоя нейронной сети
P	Количество использованных паттернов
S, s	Количество и номер элемента в тренировочном наборе

Обозначения в главе 4

\mathbf{a}	Значения массовой плотности, полученные в результате визуализации КТ-проекций
\mathbf{W}	Матрица весов в алгоритме Гаусса-Ньютона

\mathcal{E}	Диапазон энергий рентгеновского излучения
\mathcal{H}	Оператор, осуществляющий линейное рентгеновское преобразование
\mathcal{K}	Оператор-совокупность рентгеновского преобразования и спектрального смешивания
Φ, ϕ	Количество и номер проекции, используемой для визуализации объекта в компьютерной томографии
μ	Томографическое изображение, полученное для различных энергетических диапазонов
τ	Массовый коэффициент ослабления
B, b	Количество и номер диапазона энергий фотонов, рассматриваемых в спектральной КТ
$d(\cdot)$	Функция отклика детектора
E	Энергия рентгеновского излучения
J, j	Количество и номер среза в томографическом изображении
M, m	Количество и номер визуализируемого материала в объекте
$N_{x,y}$	Пространственный размер изображений, соответствующих значениям массовой плотности
$Q_{x,y}$	Пространственный размер изображений, соответствующих срезам томографических изображений для энергетических диапазонов
$s_0(\cdot)$	Функция энергетического спектра источника излучения
V, v	Количество и номер вокселя в трёхмерном объекте

Список литературы

1. Цифровая обработка изображений в информационных системах. Уч. пособие / И. Грузман [и др.]. — Новосибирск: Издательство НГТУ, 2002.
2. *Gonzales R., Woods R.* Цифровая обработка изображений. — Техносфера, 2012.
3. Learning Deep CNN Denoiser Prior for Image Restoration / К. Zhang [и др.] // 2017 IEEE Conference on Computer Vision and Pattern Recognition. — 2017. — с. 2808—2817.
4. *Banham M. R., Katsaggelos A. K.* Digital image restoration // IEEE Signal Processing Magazine. — 1997. — т. 14. — с. 24—41.
5. *Gao Q., Roth S.* How Well Do Filter-Based MRFs Model Natural Images? // Pattern Recognition / под ред. А. Pinz [и др.]. — Berlin, Heidelberg : Springer Berlin Heidelberg, 2012. — с. 62—72.
6. *Kim B.* Numerical Optimization Methods for Image Restoration : дис. ... канд. / Kim Byunggyoo. — Stanford University, 2003.
7. *Born M., Wolf E.* Principles of Optics. — 7-е изд. — Cambridge University Press, 2003.
8. *Gibson G. M., Johnson S. D., Padgett M. J.* Single-pixel imaging 12 years on: a review // Opt. Express. — 2020. — сент. — т. 28, № 19. — с. 28190—28208.
9. Deep learning for undersampled MRI reconstruction / С. М. Hyun [и др.] // Physics in Medicine & Biology. — 2018. — т. 63.
10. *Василенко Г., Тараторин А.* Восстановление изображений. — Радио и связь, 1986.
11. *Hadamard J.* Sur les Problemes Aux Derivees Partielles et Leur Signification Physique // Princeton university bulletin. — 1902. — с. 49—52.
12. *Тихонов А. Н., Арсенин В. Я.* Методы решения некорректных задач. Издание 2-е. — Наука, 1979.
13. 3D Microscopy Deconvolution using Richardson-Lucy Algorithm with Total Variation Regularization / N. Dey [и др.] //. — 2004.

14. *Ying L., Xu D., Liang Z.-P.* On Tikhonov regularization for image reconstruction in parallel MRI // The 26th Annual International Conference of the IEEE Engineering in Medicine and Biology Society. — 2004. — т. 1. — с. 1056—1059.
15. *Шемплинер В.* Восстановление дефокусированных изображений методом двумерного преобразования Фурье и регуляризации Тихонова // Научно-технический вестник информационных технологий, механики и оптики. — 2008. — с. 60—70.
16. *Серезникова Т.* Устойчивые методы восстановления зашумлённых изображений // Вестник ЮУрГУ. — 2011. — т. 25, № 242. — с. 32—42.
17. *Xie J., Xu L., Chen E.* Image Denoising and Inpainting with Deep Neural Networks // Advances in Neural Information Processing Systems. т. 25 / под ред. F. Pereira [и др.]. — Curran Associates, Inc., 2012.
18. Burst Denoising with Kernel Prediction Networks / B. Mildenhall [и др.] //. — 2018. — с. 2502—2510.
19. Joint Demosaicing and Denoising via Learned Nonparametric Random Fields / D. Khashabi [и др.] // IEEE Transactions on Image Processing. — 2014. — т. 23. — с. 4968—4981.
20. *Kokkinos F., Lefkimmiatis S.* Iterative Joint Image Demosaicking and Denoising Using a Residual Denoising Network // IEEE Transactions on Image Processing. — 2019. — авг. — т. 28, № 8. — с. 4177—4188.
21. Detail-Revealing Deep Video Super-Resolution / X. Tao [и др.] // 2017 IEEE International Conference on Computer Vision. — 2017. — с. 4482—4490.
22. Photo-Realistic Single Image Super-Resolution Using a Generative Adversarial Network / C. Ledig [и др.] // 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). — 2017. — с. 105—114.
23. ESRGAN: Enhanced Super-Resolution Generative Adversarial Networks / X. Wang [и др.] // ArXiv. — 2018. — т. abs/1809.00219.
24. Deep Convolutional Neural Network for Image Deconvolution / L. Xu [и др.] // Advances in Neural Information Processing Systems. т. 27 / под ред. Z. Ghahramani [и др.]. — Curran Associates, Inc.,

2014. — URL: <https://proceedings.neurips.cc/paper/2014/file/1c1d4df596d01da60385f0bb17a4a9e0-Paper.pdf>.
25. Learning to Deblur / C. J. Schuler [и др.] // IEEE Transactions on Pattern Analysis and Machine Intelligence. — 2016. — т. 38. — с. 1439—1451.
 26. *Schmidt U., Roth S.* Shrinkage Fields for Effective Image Restoration // 2014 IEEE Conference on Computer Vision and Pattern Recognition. — 2014. — с. 2774—2781.
 27. *Moskalev A.* Trainable regularization for Wiener deconvolution : дис. ... маг. / Moskalev Artem. — Moscow, Russia : Skolkovo Institute of Science, Technology, 2019.
 28. *Koshelev I.* A CNN-based approach for motion deblurring : дис. ... маг. / Koshelev Iaroslav. — Moscow, Russia : Skolkovo Institute of Science, Technology, 2019.
 29. Learning Fully Convolutional Networks for Iterative Non-blind Deconvolution / J. Zhang [и др.] // 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). — 2017. — с. 6969—6977.
 30. Deep Mean-Shift Priors for Image Restoration / S. Arjomand Bigdeli [и др.] // Advances in Neural Information Processing Systems. т. 30 / под ред. I. Guyon [и др.]. — Curran Associates, Inc., 2017.
 31. Microscopy Image Restoration with Deep Wiener-Kolmogorov filters / V. Pronina [et al.] // Computer Vision – ECCV 2020. — Cham : Springer International Publishing, 2020. — P. 185–201.
 32. 3D denoised completion network for deep single-pixel reconstruction of hyperspectral images / V. Pronina [и др.] // Optics Express. — 2021. — нояб. — т. 29, № 24. — с. 39559—39573.
 33. Material Decomposition Problem in Spectral CT: A Transfer Deep Learning Approach / J. F. Abascal [и др.] // 2020 IEEE 17th International Symposium on Biomedical Imaging Workshops. — 2020.
 34. Material Decomposition in Spectral CT Using Deep Learning: A Sim2Real Transfer Approach / J. F. P. J. Abascal [и др.] // IEEE Access. — 2021. — т. 9. — с. 25632—25647.

35. *Aguet F.* Super-Resolution Fluorescence Microscopy Based on Physical Models : дис. . . . канд. / Aguet Francois. — École Polytechnique Fédérale de Lausanne, 2009.
36. *Pawley J. B.* Sources of noise in three-dimensional microscopy data sets // Three-Dimensional Confocal Microscopy: Volume Investigation of Biological Systems (Cell Biology) / под ред. J. K. Stevens [и др.]. — Academic Press, 1994. — гл. 3. с. 47—94.
37. Real-time image denoising of mixed Poisson–Gaussian noise in fluorescence microscopy images using ImageJ / V. Mannam [и др.] // Optica. — 2022. — апр. — т. 9, № 4. — с. 335—345.
38. *Conchello J.-A., Lichtman J. W.* Fluorescence microscopy // Nature Methods. — 2005. — т. 2, № 12. — с. 910—919.
39. Practical Poissonian-Gaussian Noise Modeling and Fitting for Single-Image Raw-Data / A. Foi [и др.] // IEEE Transactions on Image Processing. — 2008. — т. 17. — с. 1737—1754.
40. *Shaw P. J., Rawlins D. J.* The point-spread function of a confocal microscope: its measurement and use in deconvolution of 3-D data // Journal of Microscopy. — 1991. — т. 163, № 2. — с. 151—165.
41. *Пустыльников Л.* Тёплицевы и ганкелевы матрицы и их применения // Успехи математических наук. — 1984. — т. 39, № 4. — с. 53—84.
42. *Воробьёв Н. Н.* Теория игр для экономистов-кибернетиков. — Наука, 1985.
43. Решение обратной задачи восстановления сигнала электронного микроскопа в режиме отражённых электронов на множестве функций ограниченной вариации / Н. А. Кошев [и др.] // Вычислительные методы и программирование. — 2011. — т. 12, № 3. — с. 362—367.
44. *Леонов А. С.* Применение функций нескольких переменных с ограниченными вариациями для численного решения двумерных некорректных задач // Сибирский журнал вычислительной математики. — 1999. — т. 2, № 3. — с. 257—271.
45. *Lefkimmiatis S.* Non-local Color Image Denoising with Convolutional Neural Networks // 2017 IEEE Conference on Computer Vision and Pattern Recognition. — 2017. — с. 5882—5891.

46. *Lefkimiatis S.* Universal Denoising Networks: A Novel CNN Architecture for Image Denoising // 2018 IEEE Conference on Computer Vision and Pattern Recognition. — 2018. — с. 3204–3213.
47. *Kruse J., Rother C., Schmidt U.* Learning to Push the Limits of Efficient FFT-Based Image Deconvolution // 2017 IEEE International Conference on Computer Vision (ICCV). — 2017. — с. 4596–4604.
48. *Kokkinos F., Lefkimiatis S.* Deep Image Demosaicking Using a Cascade of Convolutional Residual Denoising Networks // Computer Vision – ECCV 2020. — Springer International Publishing, 2018. — с. 317–333.
49. *Richardson W. H.* Bayesian-Based Iterative Method of Image Restoration // Journal of the Optical Society of America. — 1972. — т. 62. — с. 55–59.
50. *Lucy L. B.* An iterative technique for the rectification of observed distributions // The Astronomical Journal. — 1974. — т. 79. — с. 745–754.
51. Image Enhancement Based on Discrete Cosine Transforms (DCT) and Discrete Wavelet Transform (DWT): A Review / W. A. Mustafa [и др.] // IOP Conference Series: Materials Science and Engineering. — 2019. — т. 557.
52. *Bertsekas D. P.* Nonlinear Programming: Second Edition. — Athena Scientific, 1999.
53. *Ronneberger O., Fischer P., Brox T.* U-Net: Convolutional Networks for Biomedical Image Segmentation // Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015 / под ред. N. Navab [и др.]. — Springer International Publishing, 2015. — с. 234–241.
54. A Poisson-Gaussian Denoising Dataset With Real Fluorescence Microscopy Images / Y. Zhang [и др.] // 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). — 2019. — с. 11702–11710.
55. A deep learning-based algorithm for 2-D cell segmentation in microscopy images / Y. Al-Kofahi [и др.] // BMC Bioinformatics. — 2018.
56. Diffraction PSF 3D. — Accessed: 2019-05-30. — URL: <https://www.optinav.info/Diffraction-PSF-3D.htm>.
57. *Kirshner H., Sage D.* PSF Generator. — Accessed: 2020-02-25. — URL: <http://bigwww.epfl.ch/algorithms/psfgenerator/#ref>.

58. Fiji: an open-source platform for biological-image analysis / J. E. Schindelin [и др.] // Nature Methods. — 2012. — т. 9. — с. 676—682.
59. *Chen D.-Q.* Regularized Generalized Inverse Accelerating Linearized Alternating Minimization Algorithm for Frame-Based Poissonian Image Deblurring // SIAM J. Imaging Sciences. — 2014. — т. 7. — с. 716—739.
60. *Lefkimmatis S., Unser M.* Poisson Image Reconstruction With Hessian Schatten-Norm Regularization // IEEE Transactions on Image Processing. — 2013. — т. 22. — с. 4314—4327.
61. *Rudin L. I., Osher S., Fatemi E.* Nonlinear total variation based noise removal algorithms // Physica D: nonlinear phenomena. — 1992. — т. 60, № 1—4. — с. 259—268.
62. Loss Functions for Image Restoration With Neural Networks / H. Zhao [и др.] // IEEE Transactions on Computational Imaging. — 2017. — т. 3. — с. 47—57.
63. *Kingma D. P., Ba J.* Adam: A Method for Stochastic Optimization // CoRR. — 2014. — т. abs/1412.6980.
64. *Wiener N.* The Extrapolation, Interpolation and Smoothing of Stationary Time Series, with Engineering Applications. — The MIT Press, 1949.
65. *Marks R. J.* Handbook of Fourier Analysis & Its Applications. — Oxford University Press, 2009.
66. *Shewchuk J. R.* An Introduction to the Conjugate Gradient Method Without the Agonizing Pain : тех. отч. — Pittsburgh, PA, USA, 1994.
67. scikit-image: image processing in Python / S. van der Walt [и др.] // PeerJ. — 2014. — июнь. — т. 2. — e453. — URL: <https://doi.org/10.7717/peerj.453>.
68. *Mäkitalo M., Foi A.* Optimal Inversion of the Generalized Anscombe Transformation for Poisson-Gaussian Noise // IEEE Transactions on Image Processing. — 2013. — т. 22. — с. 91—103.
69. 3D distance-weighted Wiener filter for Poisson noise reduction in sinogram space for SPECT imaging / H. Lu [и др.] // Medical Imaging 2001: Physics of Medical Imaging. т. 4320 / под ред. L. E. Antonuk, M. J. Yaffe. — International Society for Optics, Photonics. SPIE, 2001. — с. 905—913.

70. *Anscombe F. J.* The Transformation of Poisson, Binomial and Negative-Binomial Data // *Biometrika*. — 1948. — т. 35. — с. 246—254.
71. *Mäkitalo M., Foi A.* A Closed-Form Approximation of the Exact Unbiased Inverse of the Anscombe Variance-Stabilizing Transformation // *IEEE Transactions on Image Processing*. — 2011. — т. 20. — с. 2697—2698.
72. *Li J., Luisier F., Blu T.* PURE-LET Image Deconvolution // *IEEE Transactions on Image Processing*. — 2018. — т. 27. — с. 92—105.
73. *Corle T. R., Kino G. S.* Confocal Scanning Optical Microscopy and Related Imaging Systems. — Academic press, 1996.
74. *Пальчикова И., Омелянчук Л., Смирнов Е.* О влиянии дифракции на результаты количественной цитофотометрии // *Автометрия*. — 2012. — т. 48, № 6. — с. 92—101.
75. On-the-fly estimation of a microscopy point spread function. / J. Li [и др.] // *Optics Express*. — 2018. — т. 26 20. — с. 26120—26133.
76. *Richards B., Wolf E.* Electromagnetic diffraction in optical systems, II. Structure of the image field in an aplanatic system // *Proceedings of the Royal Society of London. Series A. Mathematical and Physical Sciences*. — 1959. — т. 253. — с. 358—379.
77. *Gibson S. F., Lanni F.* Diffraction by a circular aperture as a model for three-dimensional optical microscopy. // *Journal of the Optical Society of America. A, Optics and image science*. — 1989. — т. 6 9. — с. 1357—67.
78. Three-dimensional imaging by deconvolution microscopy. / J. G. McNally [и др.] // *Methods*. — 1999. — т. 19 3. — с. 373—85.
79. Blind deconvolution of 3D data in wide field fluorescence microscopy / F. Soulez [и др.] // *International Symposium on Biomedical Imaging*. — Barcelone, Spain, 2012. — CDROM.
80. Image deblurring with blur kernel estimation in RGB channels / X.-H. Xu [и др.] // *2016 IEEE International Conference on Digital Signal Processing (DSP)*. — 2016. — с. 681—684.
81. Blind Deconvolution Microscopy Using Cycle Consistent CNN with Explicit PSF Layer / S. Lim [и др.] // *ArXiv*. — 2019. — т. abs/1904.02910.

82. Unpaired Image-to-Image Translation using Cycle-Consistent Adversarial Networks / J.-Y. Zhu [и др.] // 2017 IEEE International Conference on Computer Vision. — 2017.
83. *Kiryanov R.* Optimization of Imager PSF with GAN : дис. ... маг. / Kiryanov Roman. — Moscow, Russia : Skolkovo Institute of Science, Technology, 2020.
84. *Orieux F., Giovannelli J.-F., Rodet T.* Bayesian estimation of regularization and point spread function parameters for Wiener–Hunt deconvolution // J. Opt. Soc. Am. A. — 2010. — июль. — т. 27, № 7. — с. 1593–1607.
85. *Reznik A., Soloviev A., Torgov A.* Improving the spatial resolution of digital images and video sequences using subpixel scanning // Spatial Data Processing for Monitoring of Natural and Anthropogenic Processes 2021. — 2021.
86. *Williams J.* The electronics revolution: Inventing the future, first edition. — Springer, 2017.
87. *Ian Goodfellow Jean Pouget-Abadie M. M.* Generative adversarial nets. // Advances in neural information processing systems. — 2014. — т. 8. — с. 2672–2680.
88. *Airy G.* On the Diffraction of an Object-glass with Circular Aperture // Transactions of the Cambridge Philosophical Society. — 1835. — т. 5, № 3. — с. 283–291.
89. *Ayers G. R., Dainty J. C.* Iterative blind deconvolution method and its applications // Optics Letters. — 1988. — т. 13. — с. 547–549.
90. *Davey B. L. K., Lane R. G., Bates R. H. T.* Blind deconvolution of noisy complex-valued image // Optics Communications. — 1989. — т. 69. — с. 353–356.
91. *Tsumuraya F., Miura N., Baba N.* Iterative blind deconvolution method using Lucy’s algorithm // Astronomy and Astrophysics. — 1994. — т. 282. — с. 699–708.
92. Blind deconvolution by means of the Richardson–Lucy algorithm / D. A. Fish [и др.] // J. Opt. Soc. Am. A. — 1995. — янв. — т. 12, № 1. — с. 58–65.

93. *Vermorel J., Mohri M.* Multi-armed Bandit Algorithms and Empirical Evaluation // Machine Learning: ECML 2005 / под ред. J. Gama [и др.]. — Berlin, Heidelberg : Springer Berlin Heidelberg, 2005. — с. 437—448.
94. Recurrent Models of Visual Attention / V. Mnih [и др.] // Advances in Neural Information Processing Systems. т. 27 / под ред. Z. Ghahramani [и др.]. — Curran Associates, Inc., 2014.
95. *Ba J., Mnih V., Kavukcuoglu K.* Multiple Object Recognition with Visual Attention // CoRR. — 2015. — т. abs/1412.7755.
96. Attention-Aware Face Hallucination via Deep Reinforcement Learning / Q. Cao [и др.] // 2017 IEEE Conference on Computer Vision and Pattern Recognition. — 2017. — с. 1656—1664.
97. Distort-and-Recover: Color Enhancement Using Deep Reinforcement Learning / J. Park [и др.] // 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. — 2018. — с. 5928—5936.
98. DeepExposure: Learning to Expose Photos with Asynchronously Reinforced Adversarial Learning / R. Yu [и др.] // Advances in Neural Information Processing Systems. т. 31 / под ред. S. Bengio [и др.]. — Curran Associates, Inc., 2018.
99. Crafting a Toolchain for Image Restoration by Deep Reinforcement Learning / K. Yu [и др.] // 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. — 2018. — с. 2443—2452.
100. *Furuta R., Inoue N., Yamasaki T.* Fully Convolutional Network with Multi-Step Reinforcement Learning for Image Processing // Proceedings of the Thirty-Third AAAI Conference on Artificial Intelligence and Thirty-First Innovative Applications of Artificial Intelligence Conference and Ninth AAAI Symposium on Educational Advances in Artificial Intelligence. — Honolulu, Hawaii, USA : AAAI Press, 2019. — (AAAI'19/IAAI'19/EAAI'19). — ISBN 978-1-57735-809-1.
101. *Singhal J., Narang P.* DeblurRL: Image Deblurring with Deep Reinforcement Learning // Computer Vision and Image Processing / под ред. S. K. Singh [и др.]. — Singapore : Springer Singapore, 2021. — с. 445—454.

102. Boosting Denoisers with Reinforcement Learning for Image Restoration / J. Zhang [и др.] // *Soft Comput.* — Berlin, Heidelberg, 2022. — апр. — т. 26, № 7. — с. 3261—3272.
103. *Sutton R. S., Barto A. G.* Reinforcement Learning: An Introduction. — Second. — The MIT Press, 2018.
104. Playing Atari with Deep Reinforcement Learning / V. Mnih [и др.] // *ArXiv.* — 2013. — т. abs/1312.5602.
105. Proximal Policy Optimization Algorithms / J. Schulman [и др.] // *ArXiv.* — 2017. — т. abs/1707.06347.
106. *Latychevskaia T.* Lateral and axial resolution criteria in incoherent and coherent optics and holography, near- and far-field regimes // *Applied Optics.* — 2019. — май. — т. 58, № 13. — с. 3597—3603.
107. *Nyquist.* Scientific Volume Imaging. Deconvolution - Visualization - Analysis NyquistRate. — Accessed: 2022-06-03. <https://svi.nl/NyquistRate>.
108. *Sheppard C. J. R.* Three-dimensional phase imaging with the intensity transport equation // *Appl. Opt.* — 2002. — окт. — т. 41, № 28. — с. 5951—5955.
109. *Sheppard C., Choudhury A., Gannaway J.* Electromagnetic field near the focus of wide-angular lens and mirror system // *Microwaves, Optics and Acoustics, IEE Journal on.* — 1977. — авг. — т. 1. — с. 129—132. — DOI: 10.1049/ij-moa:19770015.
110. *Wilson T., Tan J. B.* Three dimensional image reconstruction in conventional and confocal microscopy // *Bioimaging.* — 1993. — т. 1. — с. 176—184.
111. ГОСТ 13739-78. Масло иммерсионное для микроскопии. Технические требования. Методы испытаний. — Введён 01.01.1980. — М.: Государственный комитет стандартов Совета министров СССР: Издательство стандартов, 1978.
112. Single-Pixel Imaging via Compressive Sampling / M. F. Duarte [и др.] // *IEEE Signal Processing Magazine.* — 2008. — т. 25, № 2. — с. 83—91.
113. *Граничин О., Павленко Д.* Рандомизация данных и ℓ_1 -оптимизация // *Компьютерные инструменты в образовании.* — 2010. — № 1. — с. 4—13.

114. *Edgar M., Gibson G., Padgett M.* Principles and prospects for single-pixel imaging // *Nature Photonics*. — 2018. — т. 13. — с. 13–20.
115. Two-Photon Imaging with Thermal Light / A. Valencia [и др.] // *Phys. Rev. Lett.* — 2005. — февр. — т. 94, вып. 6. — с. 063601.
116. *Shapiro J. H.* Computational ghost imaging // *Phys. Rev. A*. — 2008. — дек. — т. 78, вып. 6. — с. 061802.
117. *Katz O., Bromberg Y., Silberberg Y. R.* Compressive ghost imaging // *Applied Physics Letters*. — 2009. — т. 95. — с. 131110.
118. 1000 fps computational ghost imaging using LED-based structured illumination. / Z.-H. Xu [и др.] // *Optics express*. — 2018. — т. 26 3. — с. 2427–2434.
119. *Sampsel J. B.* Digital micromirror device and its application to projection displays // *Journal of Vacuum Science & Technology B*. — 1994. — т. 12. — с. 3242–3246.
120. *Zhang Z., Ma X., Zhong J.* Single-pixel imaging by means of Fourier spectrum acquisition. // *Nature communications*. — 2015. — т. 6. — с. 6225.
121. Learning-Based Compressive Subsampling / L. Baldassarre [и др.] // *IEEE Journal of Selected Topics in Signal Processing*. — 2016. — т. 10. — с. 809–822.
122. Adaptive Basis Scan by Wavelet Prediction for Single-Pixel Imaging / F. Rousset [и др.] // *IEEE Transactions on Computational Imaging*. — 2017. — т. 3, № 1. — с. 36–46.
123. EMVA 1288 camera characterisation and the influences of radiometric camera characteristics on geometric measurements / M. Rosenberger [и др.] // *ACTA IMEKO*. — 2016. — т. 5. — с. 81.
124. Handling negative patterns for fast single-pixel lifetime imaging / A. L. Mur [и др.] // *Molecular-Guided Surgery: Molecules, Devices, and Applications V. т. 10862* / под ред. B. W. Pogue, S. Gioux. — International Society for Optics, Photonics. SPIE, 2019. — 108620A.
125. *Peller J., Farahi F., Trammell S.* Hyperspectral imaging system based on a single-pixel camera design for detecting differences in tissue properties. // *Applied optics*. — 2018. — т. 57, № 27. — с. 7651–7658.

126. Deep neural networks for single-pixel compressive video reconstruction / A. L. Mur [и др.] // *Unconventional Optical Imaging II*. — 2020.
127. Hyperspectral imaging using the single-pixel Fourier transform technique / S. Jin [и др.] // *Scientific Reports*. — 2017. — т. 7. — с. 45209.
128. *Arias F., Sierra H., Arzuaga E.* A Framework For An Artificial Neural Network Enabled Single Pixel Hyperspectral Imager // *Proceedings of 10th Workshop on Hyperspectral Imaging and Signal Processing: Evolution in Remote Sensing*. — WHISPERS, 2019. — с. 1–5.
129. *Kozik V. I., Nezhevenko E. S.* Selection of features system and network parameters for hyperspectral images classification using convolutional neural networks // *Spatial Data Processing for Monitoring of Natural and Anthropogenic Processes 2021*. — 2021.
130. *Sidorov O., Hardeberg J. Y.* Deep Hyperspectral Prior: Single-Image Denoising, Inpainting, Super-Resolution // *Proceedings of the IEEE/CVF International Conference on Computer Vision Workshops*. — 10.2019. — с. 3844–3851.
131. Hyperspectral Image Denoising Using a 3-D Attention Denoising Network / Q. Shi [и др.] // *IEEE Transactions on Geoscience and Remote Sensing*. — 2021. — с. 1–16.
132. *Wei K., Fu Y., Huang H.* 3-D Quasi-Recurrent Neural Network for Hyperspectral Image Denoising // *IEEE Transactions on Neural Networks and Learning Systems*. — 2021. — т. 32, № 1. — с. 363–375.
133. *Борзов С., Потатуркин О.* Спектрально-пространственные методы классификации гиперспектральных изображений. Обзор // *Автометрия*. — 2018. — т. 54, № 6. — с. 64–86.
134. Single-pixel image reconstruction from experimental data using neural networks / A. Lorente Mur [и др.] // *Optics Express*. — 2021. — май. — т. 29, № 11. — с. 17097–17110.
135. Deep-learning-based ghost imaging / M. Lyu [и др.] // *Scientific Reports*. — 2017. — т. 7. — с. 17865.

136. *Erkmen B. I., Shapiro J. H.* Ghost imaging: from quantum to classical to computational // *Advances in Optics and Photonics*. — 2010. — дек. — т. 2, № 4. — с. 405—450.
137. Improving Imaging Quality of Real-time Fourier Single-pixel Imaging via Deep Learning / S. Rizvi [и др.] // *Sensors*. — 2019. — т. 19, № 19. — с. 4190.
138. Deep learning for real-time single-pixel video / C. F. Higham [и др.] // *Scientific Reports*. — 2018. — т. 8. — с. 2369.
139. Single-pixel imaging using a recurrent neural network combined with convolutional layers / I. Hoshi [и др.] // *Optics Express*. — 2020. — нояб. — т. 28, № 23. — с. 34069—34078.
140. Single-pixel compressive optical image hiding based on conditional generative adversarial network / J. Li [и др.] // *Optics Express*. — 2020. — июль. — т. 28, № 15. — с. 22992—23002.
141. *Ducros N., Mur A. L., Peyrin F.* A Completion Network for Reconstruction from Compressed Acquisition // *Proceedings of IEEE 17th International Symposium on Biomedical Imaging*. — 2020. — с. 619—623.
142. *Mur A. L.* Single-pixel imaging : Compressed video acquisition and reconstruction using deep learning : дис. ... канд. / Mur Antonio Lorente. — Institut National des Sciences Appliquées, INSA - Lyon, 2022.
143. 3D U-Net: Learning Dense Volumetric Segmentation from Sparse Annotation / Ö. Çiçek [и др.] // *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2016* / под ред. S. Ourselin [и др.]. — Cham : Springer International Publishing, 2016. — с. 424—432. — ISBN 978-3-319-46723-8.
144. *Coates A., Ng A., Lee H.* An Analysis of Single-Layer Networks in Unsupervised Feature Learning // *Proceedings of the Fourteenth International Conference on Artificial Intelligence and Statistics*. т. 15 / под ред. G. Gordon, D. Dunson, M. Dudík. — Fort Lauderdale, FL, USA : PMLR, 11–13 Apr.2011. — с. 215—223. — (Proceedings of Machine Learning Research).

145. *Stockman A., Sharpe L.* Cone spectral sensitivities and color matching // Color vision: From Genes to Perception / под ред. К. Gegenfurtner, L. T. Sharpe. — Cambridge University, 1999. — гл. 2. с. 53—87.
146. *Foster D.* Chromatic Function of the Cone // Encyclopedia of the Eye / под ред. D. A. Dartt [и др.]. — Academic, 2010. — с. 266—274. — ISBN 978-0-12-374203-2.
147. PyTorch: An Imperative Style, High-Performance Deep Learning Library / A. Paszke [и др.] // Advances in Neural Information Processing Systems 32. — Curran Associates, Inc., 2019. — с. 8024—8035.
148. *Lorente Mur A., Ducros N.* Single-Pixel Hyperspectral Imaging dataset Version 1.0. — 2020. — <https://gitlab.in2p3.fr/nicolas.ducros/spihim>.
149. Photography — Electronic still picture imaging — Resolution and spatial frequency responses / International Organization for Standardization. — Geneva, CH, 2019. — <https://www.iso.org/standard/71696.html>.
150. *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 4th JPL Airborne Earth Science Workshop. — JPL Publication, 1992. — с. 147—149.
151. Super-resolution of multispectral satellite images using convolutional neural networks / M. U. Müller [и др.] // ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences. — ISPRS, 2020. — с. 33—40.
152. *Rashmi S., Addamani S., Ravikiran A.* Spectral Angle Mapper Algorithm for Remote Sensing Image Classification //. — 2014.
153. Creating RGB Images from Hyperspectral Images Using a Color Matching Function / M. Magnusson [и др.] // Proceedings of 2020 IEEE International Geoscience and Remote Sensing Symposium. — IGARSS, 2020. — с. 2045—2048.
154. *Foster D. H., Amano K.* Hyperspectral imaging in color vision research: tutorial // J. Opt. Soc. Am. A. — 2019. — апр. — т. 36, № 4. — с. 606—627.

155. *Taguchi K., Iwanczyk J. S.* Vision 20/20: Single photon counting x-ray detectors in medical imaging. // *Medical physics*. — 2013. — т. 40 10. — с. 100901.
156. *Alvarez R. E., Macovski A.* Energy-selective reconstructions in X-ray computerized tomography. // *Physics in medicine and biology*. — 1976. — т. 21 5. — с. 733—44.
157. *Dual- and Multi-Energy CT: Principles, Technical Approaches, and Clinical Applications.* / C. H. McCollough [и др.] // *Radiology*. — 2015. — т. 276 3. — с. 637—53.
158. *Roessl E., Proksa R.* K-edge imaging in x-ray computed tomography using multi-bin photon counting detectors. // *Physics in medicine and biology*. — 2007. — т. 52 15. — с. 4679—96.
159. *Atherosclerotic plaque composition: analysis with multicolor CT and targeted gold nanoparticles.* / D. P. Cormode [и др.] // *Radiology*. — 2010. — т. 256 3. — с. 774—82.
160. *Improved Peritoneal Cavity and Abdominal Organ Imaging Using a Biphasic Contrast Agent Protocol and Spectral Photon Counting Computed Tomography K-Edge Imaging* / S. A. Si-Mohamed [и др.] // *Investigative Radiology*. — 2018. — т. 53. — с. 629—639.
161. *Breast tissue decomposition with spectral distortion correction: a postmortem study.* / H. Ding [и др.] // *Medical physics*. — 2014. — т. 41 10. — с. 101901.
162. *Sparsity-regularized image reconstruction of decomposed K-edge data in spectral CT.* / Q. Xu [и др.] // *Physics in medicine and biology*. — 2014. — т. 59 10. — с. N65—79.
163. *Rigie D., Riviere P. J. L.* Joint reconstruction of multi-channel, spectral CT data via constrained total nuclear variation minimization. // *Physics in medicine and biology*. — 2015. — т. 60 5. — с. 1741—62.
164. *Chen Z., Li L.* Robust multimaterial decomposition of spectral CT using convolutional neural networks // *Optical Engineering*. — 2019. — т. 58. — с. 013104—013104.
165. *Multi-energy CT based on a prior rank, intensity and sparsity model (PRISM).* / H. Gao [и др.] // *Inverse problems*. — 2011. — т. 27 11.

166. Sparse-View Spectral CT Reconstruction Using Spectral Patch-Based Low-Rank Penalty / K. Kim [и др.] // IEEE Transactions on Medical Imaging. — 2015. — т. 34. — с. 748—760.
167. Tensor-Based Formulation and Nuclear Norm Regularization for Multienergy Computed Tomography / O. Semerci [и др.] // IEEE Transactions on Image Processing. — 2014. — т. 23. — с. 1678—1693.
168. Tight-frame based iterative image reconstruction for spectral breast CT. / B. Zhao [и др.] // Medical physics. — 2013. — т. 40 3. — с. 031905.
169. Image-Domain Material Decomposition for Spectral CT Using a Generalized Dictionary Learning / W. Wu [и др.] // IEEE Transactions on Radiation and Plasma Medical Sciences. — 2021. — т. 5. — с. 537—547.
170. United Iterative Reconstruction for Spectral Computed Tomography / Y. Xi [и др.] // IEEE Transactions on Medical Imaging. — 2015. — т. 34. — с. 769—778.
171. *Huy Q. L., Molloy S. Y.* Least squares parameter estimation methods for material decomposition with energy discriminating detectors. // Medical physics. — 2011. — т. 38 1. — с. 245—55.
172. *Long Y., Fessler J. A.* Multi-Material Decomposition Using Statistical Image Reconstruction for Spectral CT // IEEE Transactions on Medical Imaging. — 2014. — т. 33. — с. 1614—1626.
173. *Clark D. P., Badea C. T.* Spectral diffusion: an algorithm for robust material decomposition of spectral CT data. // Physics in medicine and biology. — 2014. — т. 59 21. — с. 6445—66.
174. Energy-resolved CT imaging with a photon-counting silicon-strip detector. / M. Persson [и др.] // Physics in medicine and biology. — 2014. — т. 59 22. — с. 6709—27.
175. Comparison of five one-step reconstruction algorithms for spectral CT. / C. Mory [и др.] // Physics in medicine and biology. — 2018. — т. 63 23. — с. 235001.
176. *Rit S., Mory C., Noël P. B.* Image Formation in Spectral Computed Tomography. — CRC Press, 2020. — с. 355—372.

177. ГОСТ 61948-2-2008. Оборудование для радионуклидной диагностики. Эксплуатационные испытания. Часть 2. Сцинтилляционные камеры и однофотонные компьютерные томографы. – Введён 01.01.2009. – Национальный стандарт Российской Федерации: Стандартинформ, 2008.
178. An algorithm for constrained one-step inversion of spectral CT data / R. F. Barber [и др.] // *Physics in Medicine & Biology*. – 2016. – т. 61. – с. 3784–3818.
179. Regularization of nonlinear decomposition of spectral x-ray projection images / N. Ducros [и др.] // *Medical Physics*. – 2017. – т. 44. – e174–e187.
180. Experimental feasibility of multi-energy photon-counting K-edge imaging in pre-clinical computed tomography. / J.-P. Schlomka [и др.] // *Physics in medicine and biology*. – 2008. – т. 53 15. – с. 4031–47.
181. Estimation of Basis Line-Integrals in a Spectral Distortion-Modeled Photon Counting Detector Using Low-Rank Approximation-Based X-Ray Transmittance Modeling: K-Edge Imaging Application / O. Lee [и др.] // *IEEE Transactions on Medical Imaging*. – 2017. – т. 36. – с. 2389–2403.
182. *Abascal J. F. P. J., Ducros N., Peyrin F.* Nonlinear material decomposition using a regularized iterative scheme based on the Bregman distance // *Inverse Problems*. – 2018.
183. Deep Convolutional Neural Network for Inverse Problems in Imaging / K. H. Jin [и др.] // *IEEE Transactions on Image Processing*. – 2017. – т. 26. – с. 4509–4522.
184. Machine-learning-based nonlinear decomposition of CT images for metal artifact reduction / H. S. Park [и др.] // *ArXiv: Medical Physics*. – 2017.
185. *Clark D. P., Holbrook M. D., Badea C. T.* Multi-energy CT decomposition using convolutional neural networks // *Medical Imaging*. – 2018.
186. *Ducros N.* SPRAY – Spectral X-Ray Image Reconstruction Toolbox. – Accessed: 2022-05-31. <https://github.com/nducros/SPRAY>.
187. *Dudgeon D. E., Mersereau R. M.* Multidimensional Digital Signal Processing. – Prentice Hall, 1983.
188. *IRCAD.* ircad France. – Accessed: 2018-08-19. <https://www.ircad.fr/>.

189. Shortest-Path Constraints for 3D Multiobject Semiautomatic Segmentation Via Clustering and Graph Cut / R. Kéchéchian [и др.] // IEEE Transactions on Image Processing. — 2013. — т. 22. — с. 4224—4236.
190. SpekCalc: a program to calculate photon spectra from tungsten anode x-ray tubes / G. G. Poludniowski [и др.] // Physics in Medicine & Biology. — 2009. — т. 54. — N433—N438.
191. The KiTS19 Challenge Data: 300 Kidney Tumor Cases with Clinical Context, CT Semantic Segmentations, and Surgical Outcomes / N. Heller [и др.] // ArXiv. — 2019. — т. abs/1904.00445.
192. Review of an initial experience with an experimental spectral photon-counting computed tomography system / S. Si-Mohamed [и др.] // Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment. — 2017. — т. 873. — с. 27—35. — Imaging 2016.

Список рисунков

1.1	Схемы трёх алгоритмов на основе схемы Ричардсона-Люси, реализующих аппроксимацию регуляризационного члена тренируемыми моделями.	25
1.2	Результаты восстановления размытых изображений, искажённых пуассоновским шумом различной интенсивности. . .	28
1.3	Схемы четырёх алгоритмов на основе фильтра Винера, реализующих аппроксимацию регуляризационного члена тренируемыми моделями.	35
1.4	Результаты восстановления размытых изображений, искажённых гауссовским шумом различной интенсивности.	39
1.5	Примеры функций размытия точки и результатов её свёртки с изображениями.	43
1.6	Результаты восстановления размытых изображений, искажённых пуассоновским шумом различной интенсивности. . .	45
2.1	Схема восстановления функции рассеяния точки с помощью искажённых и псевдо-эталонных изображений.	52
2.2	Послойное сравнение искомой трёхмерной функции рассеяния точки и её восстановленной с помощью предложенного подхода версией.	53
2.3	Сравнение искомого изображения и функции рассеяния точки с их восстановленными версиями.	54
2.4	Пример трёхмерной функции рассеяния точки, полученной с помощью модели Борна и Вольфа. Для наглядности визуализации концентрических колец максимальное значение функции ограничено.	56
2.5	Итеративная схема алгоритма слепой деконволюции, объединяющего метод обучения с подкреплением для поиска параметров PSF и алгоритм Ричардсона-Люси для оценки искомого изображения.	61

2.6	Изменение восстановленной PSF (нижняя строка) и соответствующего ей изображения (верхняя строка) в зависимости от шага итеративного алгоритма слепой деконволюции. Одиночный агент осуществляет поиск параметра σ .	63
2.7	Результаты восстановления искомого изображения и PSF с помощью алгоритма слепой деконволюции и иерархического алгоритма обучения с подкреплением для поиска параметра σ .	64
2.8	Результаты восстановления искомого изображения и PSF с помощью алгоритма слепой деконволюции и иерархического алгоритма обучения с подкреплением для поиска параметров d и NA. Вектор состояний $s_t = \{\hat{\mathbf{x}}^t, \hat{\mathbf{h}}^t, \text{NA}_t, d_t\}$.	68
2.9	Результаты восстановления искомого изображения и PSF с помощью алгоритма слепой деконволюции и иерархического алгоритма обучения с подкреплением для поиска параметров d и NA. Вектор состояний $s_t = \{[a_{t-p+1}, \dots, a_t], [r_{t-p+1}, \dots, r_t]\}$.	70
3.1	Схемы получения однопиксельного измерения [114].	74
3.2	Схема двух вариантов модели для восстановления изображений из гиперспектральных однопиксельных измерений.	83
3.3	Схемы двух вариантов свёрточных сетей для уточнения решения задачи восстановления изображения из гиперспектральных однопиксельных измерений.	84
3.4	Схема преобразования RGB изображения в гиперспектральное изображение.	89
3.5	Экспериментальная установка для регистрации гиперспектральных однопиксельных измерений [148].	91
3.6	Результаты восстановления вектора измерений, полученного с интенсивностью, соответствующей $\alpha = 150$, для двух смоделированных изображений.	95
3.7	Результаты восстановления экспериментальных измерений двух объектов: изображения кошки из STL-10 и смещённой звезды Сименса.	96
3.8	Результаты восстановления экспериментальных измерений двух объектов: звезды Сименса и излучающего светодиода (LED).	97

3.9	Профили изменения интенсивностей пикселей по каналам в восстановленных экспериментальных изображениях.	98
3.10	Результаты восстановления экспериментального цветного изображения звезды Сименса.	99
4.1	Схема реконструкции томографических изображений для энергетических диапазонов и последующей декомпозиции материалов в области изображений. В примере использовано два материала – мягкие ткани (1) и костные ткани (2) – и четыре энергетических диапазона.	111
4.2	Результаты декомпозиции материалов в области проекций для трёх материалов – мягких тканей, костных тканей и контрастного вещества. Изображения соответствуют одной КТ-проекции.	115
4.3	Результаты реконструкции томографических изображений для трёх материалов – мягких тканей, костных тканей и контрастного вещества.	116
4.4	Результаты декомпозиции в области проекций для двух материалов – мягких и костных тканей. Изображения соответствуют одной КТ-проекции. Результаты опубликованы в [34].	117
4.5	Результаты реконструкции томографических изображений для двух материалов – мягких и костных тканей. Результаты опубликованы в [34].	118
4.6	Результаты реконструкции томографических изображений реального человеческого торса для двух материалов – мягких и костных тканей. Результаты опубликованы в [34].	118
4.7	Диаграмма размаха нормализованной MSE и SSIM декомпозиции материалов (а, б) и томографической реконструкции (в, г). Диаграммы показывают значение медианы, 25-го и 75-го перцентиля, а также минимальные и максимальные значения, исключая выбросы, которые отображаются красными крестиками. Результаты опубликованы в [34].	119

Б.1	Пример данных, использованных для обучения ESRGAN [83]. . .	159
В.1	Сравнение регуляризационных фильтров, вычисленных с помощью алгоритмов WF-K и WF-KPN. Результаты восстановления размытых изображений, искажённых гауссовским шумом, с использованием соответствующих фильтров.	162
Г.1	Результаты восстановления экспериментального цветного изображения звезды Сименса.	163
Г.2	Результаты восстановления вектора измерений, соответствующего $\alpha = 150$, для смоделированного изображения. Сверху вниз: каналы 2 – 9.	164
Г.3	Результаты восстановления вектора измерений, соответствующего $\alpha = 500$, для смоделированного изображения. Сверху вниз: каналы 2 – 9.	165
Г.4	Результаты восстановления экспериментальных измерений изображения кошки из STL-10.	166
Г.5	Результаты восстановления экспериментальных измерений смещённой звезды Сименса.	167

Список таблиц

1	Сравнительная таблица метрик качества восстановления изображений, искажённых размытием и пуассоновским шумом четырёх уровней интенсивности. Наилучший результат выделен жирным шрифтом.	28
2	Сравнительная таблица метрик качества восстановления изображений, искажённых размытием и гауссовским шумом пяти уровней интенсивности. Наилучший результат выделен жирным шрифтом.	40
3	Сравнительная таблица метрик качества восстановления изображений, искажённых размытием и гауссовским шумом пяти уровней интенсивности для различных размеров и количества регуляризационных фильтров.	41
4	Сравнительная таблица метрик качества восстановления изображений, искажённых размытием и пуассоновским шумом шести уровней интенсивности. Наилучший результат выделен жирным шрифтом.	46
5	Сравнительная таблица метрик качества восстановления смоделированных изображений из тестового набора данных STL-10, соответствующих трём уровням интенсивности (для всего гиперспектрального изображения $\alpha = 150, 200, 500$). Наилучший результат выделен жирным шрифтом	94
6	Сравнительная таблица метрик качества восстановления изображений спроецированной массовой плотности. Наилучший результат выделен жирным шрифтом.	114
7	Сравнительная таблица метрик качества восстановления томографических изображений. Наилучший результат выделен жирным шрифтом.	115
8	Сравнительная таблица метрик качества восстановления изображений, искажённых размытием и гауссовским шумом пяти уровней интенсивности.	160

- 9 Сравнительная таблица метрик качества восстановления изображений, искажённых размытием и пуассоновским шумом шести уровней интенсивности. 161
- 10 Сравнительная таблица производительности алгоритмом для обработки изображения размером 256×256 . Значения получены с использованием кодов из открытого доступа путём усреднения 10 запусков. 161

Приложение А

Вывод выражений для алгоритма обратного распространения ошибки

Матрицы функции размытия точки \mathbf{H} и регуляризационных фильтров \mathbf{G}_d принято рассматривать как циркулянтные матрицы вещественных чисел, которые могут быть диагонализированы в области Фурье. Благодаря этому решение для предложенных в Главе 1 моделей WF-K и WF-KPN может быть выражено в замкнутой форме как

$$\hat{\mathbf{x}} = \mathbf{F}^H \left(\frac{\mathbf{D}_{\mathbf{H}}^* \mathbf{F} \mathbf{y}}{|\mathbf{D}_{\mathbf{H}}|^2 + e^\alpha \sum_{d=1}^D |\mathbf{D}_{\mathbf{G}_d}|^2} \right), \quad (\text{A.1})$$

где процедура деления осуществлена поэлементно.

Регуляризационные фильтры \mathbf{g}_d и степень регуляризационного коэффициента α , входящие в состав выражения А.1, обновляются в ходе обучения модели с помощью алгоритма обратного распространения ошибки. В то время как функция потерь, минимизируемая в процессе тренировки моделей, вещественнозначная, решение, полученное в соответствии с формулой А.1, включает в себя комплекснозначные величины. В связи с этим невозможно применение готовых алгоритмов вычисления обратного распространения ошибки, реализованных в существующих библиотеках алгоритмов глубокого обучения. По этой причине вычисление градиентов функции потерь относительно обучаемых параметров проведено вручную.

Ниже представлены результаты вычисления градиентов функции потерь относительно обучаемых параметров. Если не указано иное, вычисления произведены с использованием компоновки знаменателя, то есть производная вектора по вектору посчитана как

$$\frac{\partial \mathbf{y}}{\partial \mathbf{x}} = \begin{bmatrix} \frac{\partial y_1}{\partial x_1} & \frac{\partial y_2}{\partial x_1} & \dots & \frac{\partial y_m}{\partial x_1} \\ \frac{\partial y_1}{\partial x_2} & \frac{\partial y_2}{\partial x_2} & \dots & \frac{\partial y_m}{\partial x_2} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial y_1}{\partial x_n} & \frac{\partial y_2}{\partial x_n} & \dots & \frac{\partial y_m}{\partial x_n} \end{bmatrix}, \quad (\text{A.2})$$

а не

$$\frac{\partial \mathbf{y}}{\partial \mathbf{x}} = \begin{bmatrix} \frac{\partial y_1}{\partial x_1} & \frac{\partial y_1}{\partial x_2} & \dots & \frac{\partial y_1}{\partial x_n} \\ \frac{\partial y_2}{\partial x_1} & \frac{\partial y_2}{\partial x_2} & \dots & \frac{\partial y_2}{\partial x_n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial y_m}{\partial x_1} & \frac{\partial y_m}{\partial x_2} & \dots & \frac{\partial y_m}{\partial x_n} \end{bmatrix} \quad (\text{A.3})$$

A.0.1 Производная по α

Обозначим решение выражения A.1 как функцию, заданную параметрами,

$$f(\alpha, \mathbf{y}, \mathbf{g}_d) = \mathbf{F}^H \left(\frac{\mathbf{D}_{\mathbf{H}}^* \mathbf{F} \mathbf{y}}{|\mathbf{D}_{\mathbf{H}}|^2 + e^\alpha \sum_{d=1}^D |\mathbf{D}_{\mathbf{G}d}|^2} \right). \quad (\text{A.4})$$

Для простоты вычислений обозначим числитель и знаменатель в выражении A.4, соответственно,

$$\mathbf{D}_{\mathbf{H}}^* \mathbf{F} \mathbf{y} = \mathbf{z}, \quad (\text{A.5})$$

$$|\mathbf{D}_{\mathbf{H}}|^2 + e^\alpha \sum_{d=1}^D |\mathbf{D}_{\mathbf{G}d}|^2 = \Omega. \quad (\text{A.6})$$

Используя Формулы A.5 и A.6, выражение A.4 приобретает вид

$$f(\alpha, \mathbf{y}, \mathbf{g}_d) = \mathbf{F}^H \Omega^{-1} \mathbf{z}. \quad (\text{A.7})$$

Заметим, что Ω^{-1} является диагональной матрицей. Используя компоновку числителя, производная $f(\alpha, \mathbf{y}, \mathbf{g}_d)$ по α выражена, как

$$\begin{aligned} \frac{\partial f(\alpha, \mathbf{y}, \mathbf{g}_d)}{\partial \alpha} &= \frac{\partial \mathbf{F}^H \Omega^{-1} \mathbf{z}}{\partial \alpha} = -\mathbf{F}^H \Omega^{-1} \frac{\partial \Omega}{\partial \alpha} \Omega^{-1} \mathbf{z} = \\ &= -\mathbf{F}^H \Omega^{-1} e^\alpha \sum_{d=1}^D |\mathbf{D}_{\mathbf{G}d}|^2 \Omega^{-1} \mathbf{z} = -e^\alpha \mathbf{F}^H \Omega^{-2} \sum_{d=1}^D |\mathbf{D}_{\mathbf{G}d}|^2 \mathbf{D}_{\mathbf{H}}^* \mathbf{F} \mathbf{y}. \end{aligned} \quad (\text{A.8})$$

A.0.2 Производная по \mathbf{g}_d

Перепишем выражение A.4 как $f(\alpha, \mathbf{y}, \mathbf{g}_d) = \mathbf{F}^H \mathbf{h}(\mathbf{g})$, где

$$\mathbf{h}(\mathbf{g}) = \frac{\boldsymbol{\lambda}_{\mathbf{H}}^* \odot \mathbf{F}\mathbf{y}}{|\boldsymbol{\lambda}_{\mathbf{H}}|^2 + e^\alpha \sum_{d=1}^D |\mathbf{T}_d \mathbf{g}_d|^2}. \quad (\text{A.9})$$

Таким образом, можно выразить производную функции $f(\alpha, \mathbf{y}, \mathbf{g}_d)$ от \mathbf{g}_d как

$$\frac{\partial f(\alpha, \mathbf{y}, \mathbf{g}_d)}{\partial \mathbf{g}_d} = \frac{\partial \mathbf{F}^H \mathbf{h}(\mathbf{g})}{\partial \mathbf{g}_d} = \frac{\partial \mathbf{h}(\mathbf{g})}{\partial \mathbf{g}_d} \mathbf{F}^*. \quad (\text{A.10})$$

Подчеркнём, что матрицы $\mathbf{D}_{\mathbf{H}}$ и $\mathbf{D}_{\mathbf{G}_d}$ являются диагональными,

$$\begin{aligned} \boldsymbol{\lambda}_{\mathbf{H}} &= \text{vec}(\mathbf{D}_{\mathbf{H}}) \\ \boldsymbol{\lambda}_{\mathbf{G}_d} &= \text{vec}(\mathbf{D}_{\mathbf{G}_d}) = \mathbf{T}_d \mathbf{g}_d \\ \mathbf{T}_d &= \mathbf{F} \mathbf{S}_{\mathbf{G}_d} \mathbf{P}_{\mathbf{G}_d} \in \mathbb{C}^{N \times L^2}. \end{aligned} \quad (\text{A.11})$$

Здесь $\text{vec}(\mathbf{D})$ представляет собой вектор, расположенный на главной диагонали матрицы \mathbf{D} . Символ \odot обозначает операцию поэлементного перемножения, процедура деления в формуле A.9 также осуществлена поэлементно. Здесь $\mathbf{S}_{\mathbf{H}}, \mathbf{S}_{\mathbf{G}_d} \in \mathbb{R}^{N \times N}$ – матрицы, осуществляющие линейную операцию сдвига векторов \mathbf{h} и \mathbf{g}_d , $\mathbf{P}_{\mathbf{H}} \in \mathbb{R}^{N \times N}$, $\mathbf{P}_{\mathbf{G}_d} \in \mathbb{R}^{N \times L^2}$ – матрицы, осуществляющие линейную операцию заполнения нулями векторов \mathbf{h} и \mathbf{g}_d .

Используя выражение A.9 и обозначения в формуле A.11, $\mathbf{h}(\mathbf{g})$ можно переписать как

$$\mathbf{h}(\mathbf{g}) = \begin{bmatrix} \frac{\mathbf{M}_1 \boldsymbol{\lambda}_{\mathbf{H}}^* \odot \mathbf{F}\mathbf{y}}{|\mathbf{M}_1 \boldsymbol{\lambda}_{\mathbf{H}}|^2 + e^\alpha \sum_{d=1}^D |\mathbf{M}_1 \mathbf{T}_d \mathbf{g}_d|^2} \\ \dots \\ \frac{\mathbf{M}_N \boldsymbol{\lambda}_{\mathbf{H}}^* \odot \mathbf{F}\mathbf{y}}{|\mathbf{M}_N \boldsymbol{\lambda}_{\mathbf{H}}|^2 + e^\alpha \sum_{d=1}^D |\mathbf{M}_N \mathbf{T}_d \mathbf{g}_d|^2} \end{bmatrix}, \quad (\text{A.12})$$

где $\mathbf{M}_i \in \mathbb{R}^{1 \times N}$ представляет собой вектор, каждый элемент которого, за исключением i , равен нулю, а i -й элемент равен 1. Таким образом, i -й элемент вектора $\mathbf{h}(\mathbf{g})$ равен

$$\mathbf{h}_i(\mathbf{g}) = \frac{\mathbf{M}_i \boldsymbol{\lambda}_{\mathbf{H}}^* \odot \mathbf{F}\mathbf{y}}{|\mathbf{M}_i \boldsymbol{\lambda}_{\mathbf{H}}|^2 + e^\alpha \sum_{d=1}^D |\mathbf{M}_i \mathbf{T}_d \mathbf{g}_d|^2} = \frac{\mathbf{a}_i}{\mathbf{b}_i + e^\alpha \mathbf{u}_i(\mathbf{g})}. \quad (\text{A.13})$$

Обозначим

$$\begin{aligned}
\mathbf{a}_i &= \mathbf{M}_i \boldsymbol{\lambda}_{\mathbf{H}}^* \odot \mathbf{F} \mathbf{y}, \\
\mathbf{b}_i &= |\mathbf{M}_i \boldsymbol{\lambda}_{\mathbf{H}}|^2 \\
\mathbf{u}_i &= \sum_{d=1}^D |\mathbf{M}_i \mathbf{T}_d \mathbf{g}_d|^2 = \sum_{d=1}^D \mathbf{g}_d^\top \mathbf{T}_d^{*\top} \mathbf{M}_i^\top \mathbf{M}_i \mathbf{T}_d \mathbf{g}_d.
\end{aligned} \tag{A.14}$$

Для вычисления производной $\mathbf{h}_i(\mathbf{g})$ от j -го регуляризационного фильтра \mathbf{g}_j необходимо произвести промежуточные вычисления, начиная с вычисления производной $\mathbf{u}_i(\mathbf{g})$ от \mathbf{g}_j ,

$$\begin{aligned}
\frac{\partial \mathbf{u}_i(\mathbf{g})}{\partial \mathbf{g}_j} &= \sum_{d=1}^D \left(\frac{\partial \mathbf{g}_d}{\partial \mathbf{g}_j} \mathbf{T}_d^{*\top} \mathbf{M}_i^\top \mathbf{M}_i \mathbf{T}_d \mathbf{g}_d + \frac{\partial \mathbf{g}_d}{\partial \mathbf{g}_j} \mathbf{T}_d^\top \mathbf{M}_i^\top \mathbf{M}_i \mathbf{T}_d^* \mathbf{g}_d \right) = \\
&= (\mathbf{T}_j^{*\top} \mathbf{M}_i^\top \mathbf{M}_i \mathbf{T}_j + \mathbf{T}_j^\top \mathbf{M}_i^\top \mathbf{M}_i \mathbf{T}_j^*) \mathbf{g}_j.
\end{aligned} \tag{A.15}$$

Подставляя результат, полученный в выражении A.15 в $\frac{\partial \mathbf{h}_i(\mathbf{g})}{\partial \mathbf{g}_j}$, получаем

$$\begin{aligned}
\frac{\partial \mathbf{h}_i(\mathbf{g})}{\partial \mathbf{g}_j} &= \mathbf{a}_i \frac{\partial (\mathbf{b}_i + e^\alpha \mathbf{u}_i(\mathbf{g}))^{-1}}{\partial \mathbf{g}_j} = -\mathbf{a}_i \frac{e^\alpha}{(\mathbf{b}_i + e^\alpha \mathbf{u}_i(\mathbf{g}))^2} \frac{\partial \mathbf{u}_i(\mathbf{g})}{\partial \mathbf{g}_j} = \\
&= -e^\alpha \frac{\mathbf{M}_i \boldsymbol{\lambda}_{\mathbf{H}}^* \odot \mathbf{F} \mathbf{y}}{(|\mathbf{M}_i \boldsymbol{\lambda}_{\mathbf{H}}|^2 + e^\alpha \sum_{d=1}^D |\mathbf{M}_i \mathbf{T}_d \mathbf{g}_d|^2)^2} \cdot (\mathbf{R}_{ij} + \overline{\mathbf{R}_{ij}}) \mathbf{g}_j,
\end{aligned} \tag{A.16}$$

где $\mathbf{R}_{ij} = \mathbf{T}_j^H \mathbf{M}_i^\top \mathbf{M}_i \mathbf{T}_j$ и $\overline{\mathbf{R}_{ij}}$ – сопряжённая к \mathbf{R}_{ij} матрица.

Производная $\frac{\partial \mathbf{h}(\mathbf{g})}{\partial \mathbf{g}_j}$ может быть записана как

$$\frac{\partial \mathbf{h}(\mathbf{g})}{\partial \mathbf{g}_j} = \left[\frac{\partial \mathbf{h}_1(\mathbf{g})}{\partial \mathbf{g}_j} \quad \frac{\partial \mathbf{h}_2(\mathbf{g})}{\partial \mathbf{g}_j} \quad \dots \quad \frac{\partial \mathbf{h}_N(\mathbf{g})}{\partial \mathbf{g}_j} \right] \in \mathbb{R}^{L^2 \times N} \tag{A.17}$$

Используя A.17 и применяя формулу сложной производной, можно вывести $\frac{\partial \mathbf{h}(\mathbf{g})}{\partial \mathbf{g}_j} \mathbf{F}^* \mathbf{q}$ из формулы A.10, где $\mathbf{q} \in \mathbb{R}^{N \times 1}$ – вектор вещественнозначных чисел, полученный с помощью алгоритма обратного распространения ошибки из последующих слоёв. Таким образом можно записать

$$\frac{\partial \mathbf{h}(\mathbf{g})}{\partial \mathbf{g}_j} \mathbf{F}^* \mathbf{q} = \sum_{i=1}^D \frac{\partial \mathbf{h}_i(\mathbf{g})}{\partial \mathbf{g}_j} \mathbf{Q}_i, \quad \mathbf{Q}_i = \mathbf{M}_i \mathbf{F}^* \mathbf{q}. \tag{A.18}$$

Выведем выражение в сумме в формуле A.18 как

$$\begin{aligned} \frac{\partial \mathbf{h}_i(\mathbf{g})}{\partial \mathbf{g}_j} \mathbf{Q}_i &= -e^\alpha \frac{\mathbf{M}_i \boldsymbol{\lambda}_{\mathbf{H}}^* \odot \mathbf{F} \mathbf{y}}{(|\mathbf{M}_i \boldsymbol{\lambda}_{\mathbf{H}}|^2 + e^\alpha \sum_{d=1}^D |\mathbf{M}_i \mathbf{T}_d \mathbf{g}_d|^2)^2} \cdot \\ &\cdot (\mathbf{R}_{ij} + \overline{\mathbf{R}_{ij}}) \mathbf{g}_j \mathbf{M}_i \mathbf{F}^* \mathbf{q} = -e^\alpha (\mathbf{R}_{ij} + \overline{\mathbf{R}_{ij}}) \mathbf{g}_j \mathbf{M}_i \mathbf{z}. \end{aligned} \quad (\text{A.19})$$

Здесь использовано обозначение $\mathbf{D}_{\mathbf{H}} = \text{diag}(\boldsymbol{\lambda}_{\mathbf{H}})$, $\mathbf{D}_{\mathbf{G}_d} = \text{diag}(\boldsymbol{\lambda}_{\mathbf{G}_d})$, и следовательно $\mathbf{z} = \frac{\mathbf{D}_{\mathbf{H}}^* \mathbf{F} \mathbf{y}}{(|\mathbf{D}_{\mathbf{H}}|^2 + e^\alpha \sum_{d=1}^D |\mathbf{D}_{\mathbf{G}_d}|^2)} \odot \mathbf{F}^* \mathbf{q}$. Процедура деления здесь выполнена с помощью поэлементного деления вектора в числителе на соответствующие элементы диагональной матрицы в знаменателе, расположенные на главной диагонали. Здесь $\text{diag}(\boldsymbol{\lambda})$ обозначает прямоугольную матрицу с вектором $\boldsymbol{\lambda}$ на главной диагонали и нулями на остальных позициях.

Подставляя выражение A.19 в формулу A.18 и принимая $\mathbf{Q} = \mathbf{F}^* \mathbf{q}$, можно записать

$$\begin{aligned} \frac{\partial \mathbf{h}(\mathbf{g})}{\partial \mathbf{g}_j} \mathbf{Q} &= -e^\alpha \sum_{i=1}^D (\mathbf{R}_{ij} + \overline{\mathbf{R}_{ij}}) \mathbf{g}_j \mathbf{M}_i \mathbf{z} = -e^\alpha \sum_{i=1}^D (\mathbf{T}_j^H \mathbf{M}_i^\top \mathbf{M}_i \mathbf{T}_j + \\ &+ \overline{\mathbf{T}_j^H \mathbf{M}_i^\top \mathbf{M}_i \mathbf{T}_j}) \mathbf{g}_j \mathbf{z}_i = -e^\alpha \mathbf{P}_{\mathbf{G}_j}^\top \mathbf{S}_{\mathbf{G}_j}^\top \left(\mathbf{F}^H \sum_{i=1}^D \mathbf{M}_i^\top \mathbf{M}_i \boldsymbol{\lambda}_{\mathbf{G}_j} \mathbf{z}_i + \right. \\ &+ \left. \mathbf{F}^\top \sum_{i=1}^D \mathbf{M}_i^\top \mathbf{M}_i \boldsymbol{\lambda}_{\mathbf{G}_j}^* \mathbf{z}_i \right) = -e^\alpha \mathbf{P}_{\mathbf{G}_j}^\top \mathbf{S}_{\mathbf{G}_j}^\top (\mathbf{F}^H(\boldsymbol{\lambda}_{\mathbf{G}_j} \odot \mathbf{z}) + \\ &+ \mathbf{F}^\top(\boldsymbol{\lambda}_{\mathbf{G}_j}^* \odot \mathbf{z})) = -e^\alpha \mathbf{P}_{\mathbf{G}_j}^\top \mathbf{S}_{\mathbf{G}_j}^\top (\mathbf{F}^H(\boldsymbol{\lambda}_{\mathbf{G}_j} \odot \mathbf{z}) + \overline{\mathbf{F}^H(\boldsymbol{\lambda}_{\mathbf{G}_j} \odot \mathbf{z}^*)}). \end{aligned} \quad (\text{A.20})$$

Заметим, что вектор $\mathbf{F}^H(\boldsymbol{\lambda}_{\mathbf{G}_j} \odot \mathbf{z}^*)$ является вещественнозначным, поэтому $\overline{\mathbf{F}^H(\boldsymbol{\lambda}_{\mathbf{G}_j} \odot \mathbf{z}^*)} = \mathbf{F}^H(\boldsymbol{\lambda}_{\mathbf{G}_j} \odot \mathbf{z}^*)$. Таким образом, выражение A.20 может быть переписано как

$$\begin{aligned} \frac{\partial \mathbf{h}(\mathbf{g})}{\partial \mathbf{g}_j} \mathbf{Q} &= -e^\alpha \mathbf{P}_{\mathbf{G}_j}^\top \mathbf{S}_{\mathbf{G}_j}^\top (\mathbf{F}^H(\boldsymbol{\lambda}_{\mathbf{G}_j} \odot \mathbf{z}) + \mathbf{F}^H(\boldsymbol{\lambda}_{\mathbf{G}_j} \odot \mathbf{z}^*)) = \\ &= -e^\alpha \mathbf{P}_{\mathbf{G}_j}^\top \mathbf{S}_{\mathbf{G}_j}^\top \mathbf{F}^H(\boldsymbol{\lambda}_{\mathbf{G}_j} \odot (\mathbf{z} + \mathbf{z}^*)) = \\ &= -2e^\alpha \mathbf{P}_{\mathbf{G}_j}^\top \mathbf{S}_{\mathbf{G}_j}^\top \mathbf{F}^H(\boldsymbol{\lambda}_{\mathbf{G}_j} \odot \text{Re}(\mathbf{z})). \end{aligned} \quad (\text{A.21})$$

Таким образом производная функции $f(\alpha, \mathbf{y}, \mathbf{g}_d)$ по регуляризационным фильтрам \mathbf{g}_d

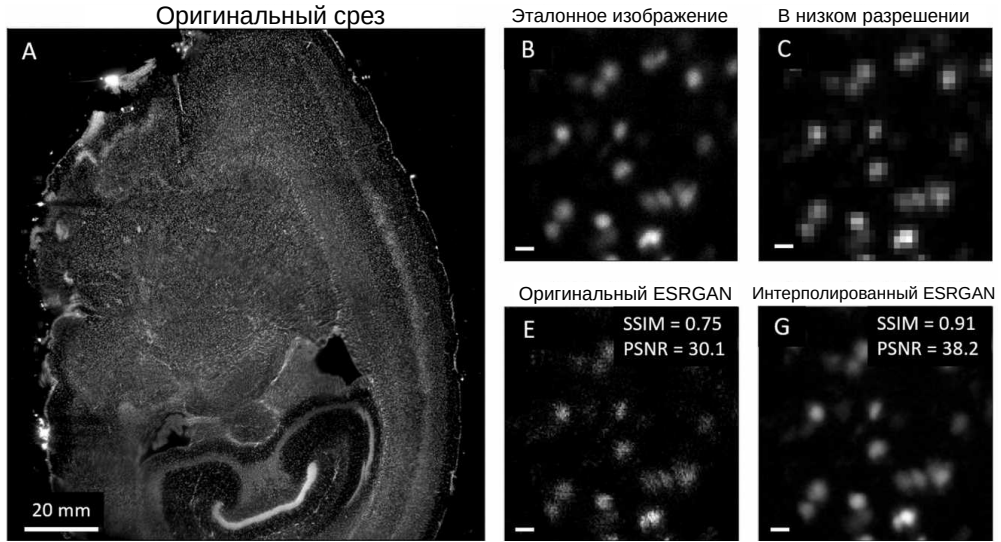
$$\frac{\partial f(\alpha, \mathbf{y}, \mathbf{g}_d)}{\partial \mathbf{g}_d} \mathbf{q} = -2e^\alpha \mathbf{P}_{\mathbf{G}_d}^\top \mathbf{S}_{\mathbf{G}_d}^\top \mathbf{F}^H \cdot \left[\lambda_{\mathbf{G}_d} \odot \operatorname{Re} \left(\frac{\mathbf{D}_H^* \mathbf{F} \mathbf{y}}{(|\mathbf{D}_H|^2 + e^\alpha \sum_{d=1}^D |\mathbf{D}_{\mathbf{G}_d}|^2)^2} \odot \mathbf{F}^* \mathbf{q} \right) \right]. \quad (\text{A.22})$$

Процедура деления в выражении A.22 осуществлена поэлементно.

Приложение Б

Пример данных для обучения ESRGAN

На Рисунке Б.1 представлен пример одного среза трёхмерного изображения из набора данных, использованного для обучения модели ESRGAN. Данные представляют собой трёхмерные изображения мозга мыши, полученные с помощью светлопольного микроскопа.



На Рисунке Б.1 представлен пример одного среза трёхмерного изображения (слева), а также скадрированный участок этого изображения и его сгенерированная в низком разрешении версия – верхняя строка. На нижней строке представлены результаты восстановления искомого скадрированного участка с помощью оригинальной и интерполированной моделей ESRGAN. Рисунок Б.1 — Пример данных, использованных для обучения ESRGAN [83].

Приложение В

Дополнительные результаты для фильтра Винера с обучаемыми параметрами регуляризации

На Рисунке В.1 показаны регуляризационные фильтры, полученные с помощью моделей WF-K и WF-KPN с $D = 8$, $L = 3$. На Рисунке В.1 также представлены изображения, восстановленные с помощью соответствующих моделей. Из результатов видно, что модель WF-K вычисляет идентичные для всех изображений фильтры, в то время как WF-KPN предсказывает индивидуальные для каждого изображения фильтры. Результаты восстановления с использованием обеих моделей демонстрируют, что WF-KPN позволяет восстанавливать визуально лучшие изображения, чем WF-K.

В Таблицах 8 и 9 представлены дополнительные результаты восстановления изображений, искажённых гауссовским и пуассоновским шумами соответственно. Предложенные в Главе 1 модели были обучены на протяжении большего количества эпох. В то время как все модели показывают более высокие значения метрик с увеличением количества эпох, модель WF-KPN-SA для гауссовского шума показала некоторое ухудшение метрик. Это может свидетельствовать о процессе переобучения модели на имеющихся данных и уровнях шума.

Таблица 8 — Сравнительная таблица метрик качества восстановления изображений, искажённых размытием и гауссовским шумом пяти уровней интенсивности.

	Значение стандартного отклонения									
	0.001		0.005		0.01		0.05		0.1	
	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM
Искажённое изображение	36.23	.8955	35.37	.8791	33.93	.8339	26.03	.3858	21.14	.1718
UNet	37.96	.9141	37.56	.9116	36.95	.9057	34.42	.8704	32.79	.8410
WF-K	35.63	.8848	35.58	.8839	35.44	.8814	33.30	.8263	30.70	.7270
WF-KPN	40.04	.9376	38.35	.9215	37.09	.9069	33.13	.8472	31.33	.8030
WF-KPN-SA	38.56	.9270	37.95	.9221	37.11	.9131	34.31	.8657	32.44	.8263
WF-UNet	37.62	.9136	37.41	.9121	37.00	.9078	34.53	.8703	32.71	.8360

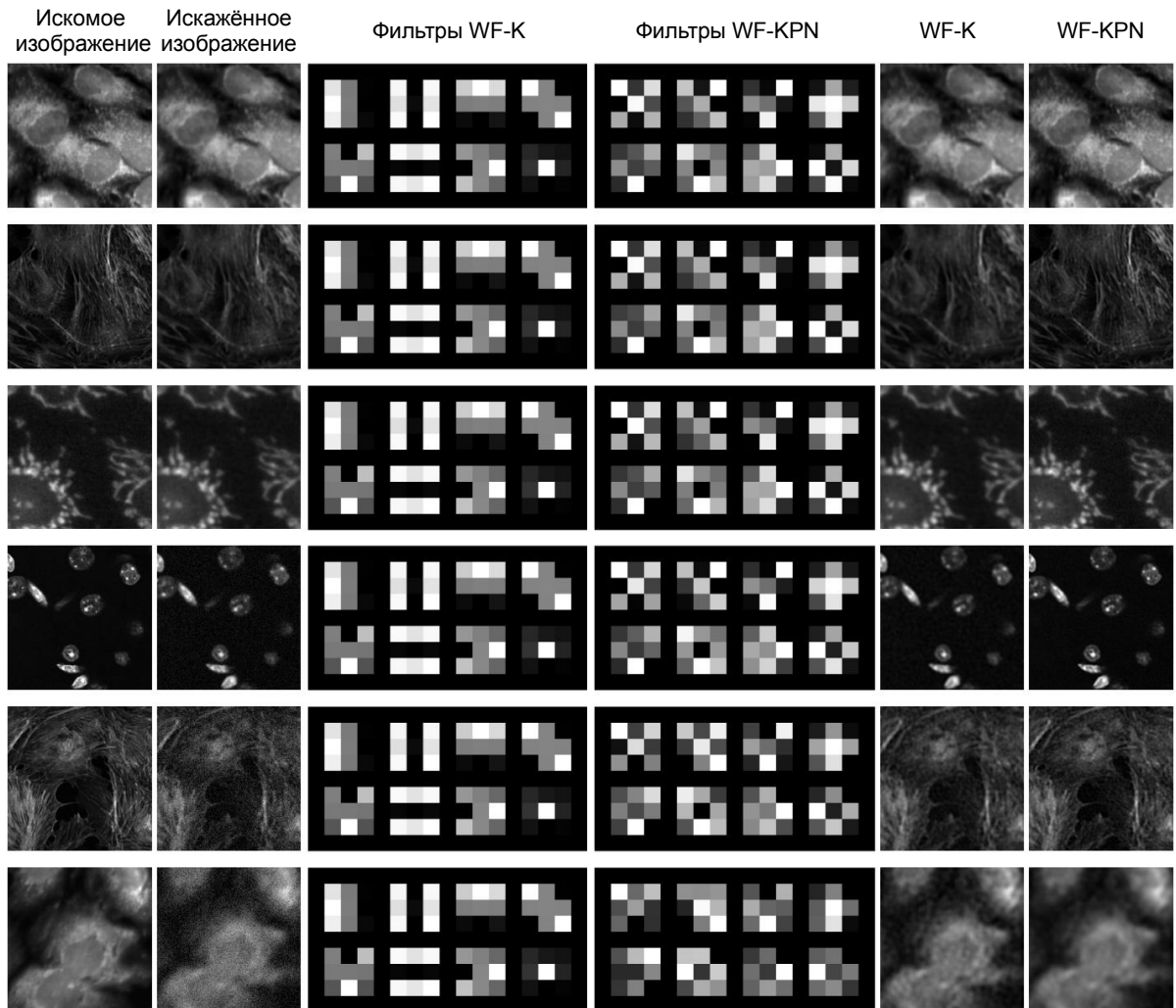
Таблица 9 — Сравнительная таблица метрик качества восстановления изображений, искажённых размытием и пуассоновским шумом шести уровней интенсивности.

	Максимальное значение пикселя в изображении											
	1		2		5		10		25		50	
	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM
Искажённое изображение	10.93	.0981	13.22	.1105	16.67	.1639	19.45	.2425	23.21	.3930	26.00	.5247
UNet	28.77	.7953	30.00	.8177	31.62	.8460	32.84	.8657	34.29	.8887	35.25	.9024
WF-K	25.99	.6071	28.01	.7104	30.10	.7960	31.30	.8325	32.25	.8571	32.64	.8659
WF-KPN	27.35	.7466	28.48	.7760	30.18	.8145	31.59	.8426	33.37	.8739	34.61	.8914
WF-KPN-SA	28.85	.7929	30.11	.8184	31.77	.8483	33.05	.8690	34.61	.8922	35.69	.9066
WF-UNet	29.05	.7977	30.29	.8223	31.90	.8517	33.14	.8723	34.69	.8957	35.83	.9101

В Таблице 10 представлено сравнение количества времени, затраченного на выполнение алгоритмов, представленных во Главе 1. Тренировка всех алгоритмов проводилась с использованием Intel Core i7-8750H CPU и NVIDIA GeForce GTX 1080Ti GPU.

Таблица 10 — Сравнительная таблица производительности алгоритмов для обработки изображения размером 256×256 . Значения получены с использованием кодов из открытого доступа путём усреднения 10 запусков.

	CPU, мс	GPU, мс		CPU, мс	GPU, мс
GILAM	4155.9	–	DMSP	79656.0	19707.7
HSPiRAL	5828.8	–	IRCNN	7890.7	–
PURE-LET	211.3	–	UNet	32.4	2.8
FDN	34542.1	36.6	WF-K	5.6	4.6
WF-KPN	45.4	7.7	WF-KPN-SA	1382.9	187.2
WF-UNet	354.1	39.8	WFT	8.5	–
TRL-KPN	117.6	34.4	TRL-UNet	156.7	22.8
TRLK	73.5	28.9	RL TV	10.0	10.2

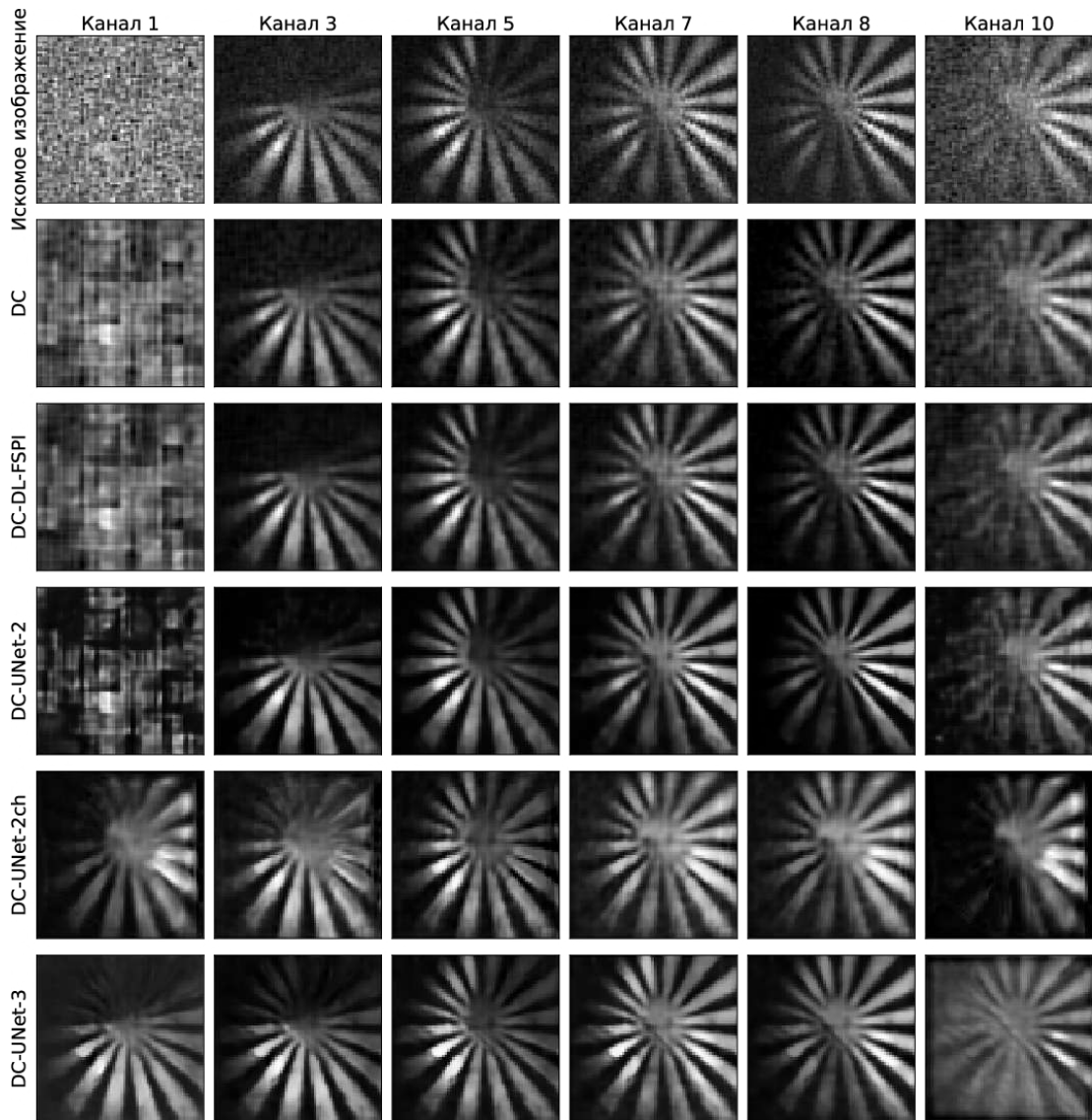


На представленных изображениях видно, что фильтры, вычисленные с помощью WF-K, являются идентичными для всех изображений. Наборы фильтров, вычисленных с помощью WF-KPN, являются индивидуальными для каждого изображения.

Рисунок В.1 — Сравнение регуляризационных фильтров, вычисленных с помощью алгоритмов WF-K и WF-KPN. Результаты восстановления размытых изображений, искажённых гауссовским шумом, с использованием соответствующих фильтров.

Приложение Г

Дополнительные результаты реконструкции гиперспектральных изображений из однопиксельных измерений



Из представленных результатов видно, что использование межканальной корреляции (DC-UNet-2ch, DC-UNet-3) и трёхмерных свёрток (DC-UNet-3) позволяет восстановить информацию даже в каналах, содержащих минимальное количество полезного сигнала (Канал 1).

Рисунок Г.1 — Результаты восстановления экспериментального цветного изображения звезды Сименса.

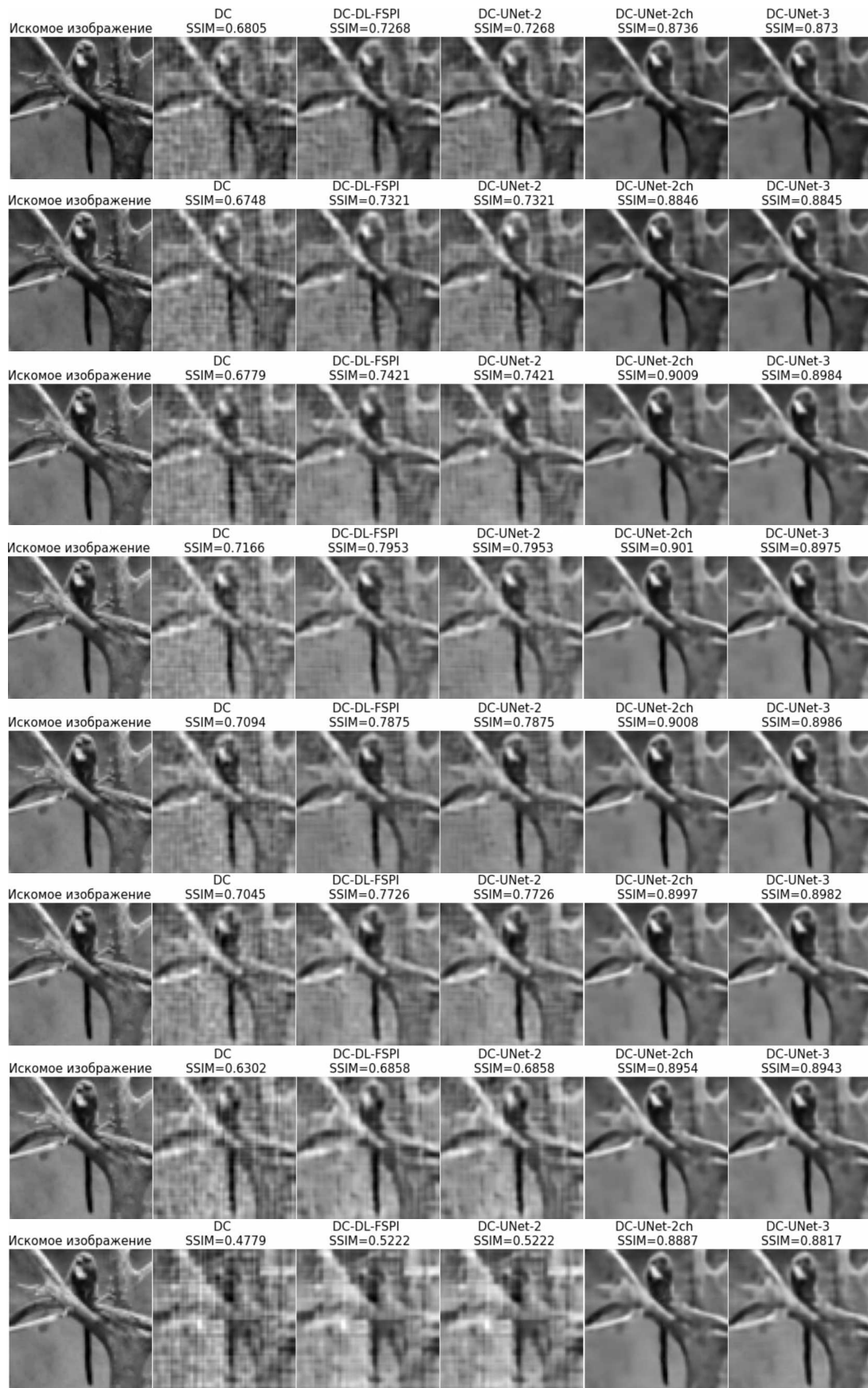


Рисунок Г.2 — Результаты восстановления вектора измерений, соответствующего $\alpha = 150$, для смоделированного изображения. Сверху вниз: каналы 2 – 9.

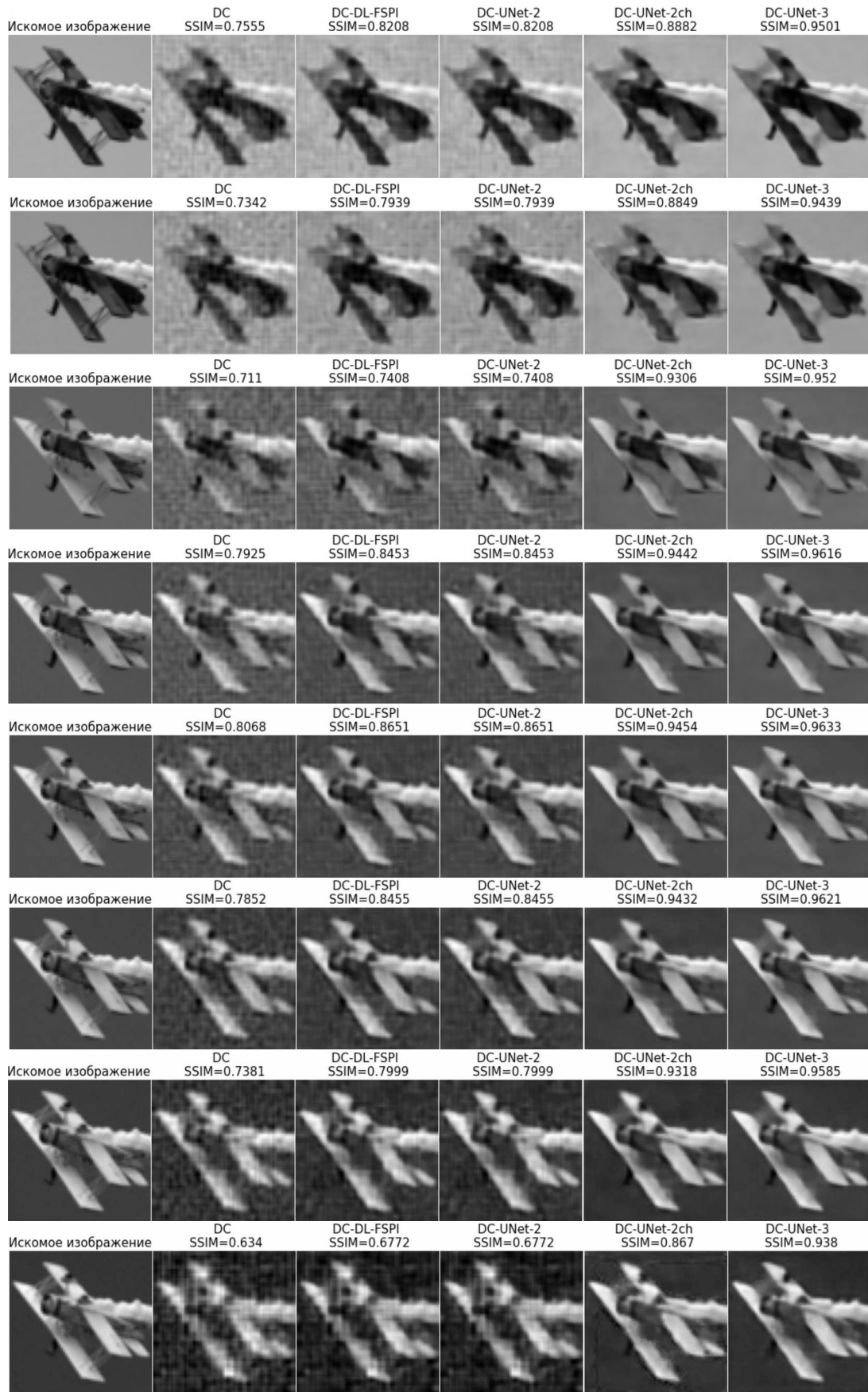


Рисунок Г.3 — Результаты восстановления вектора измерений, соответствующего $\alpha = 500$, для смоделированного изображения. Сверху вниз: каналы 2 – 9.

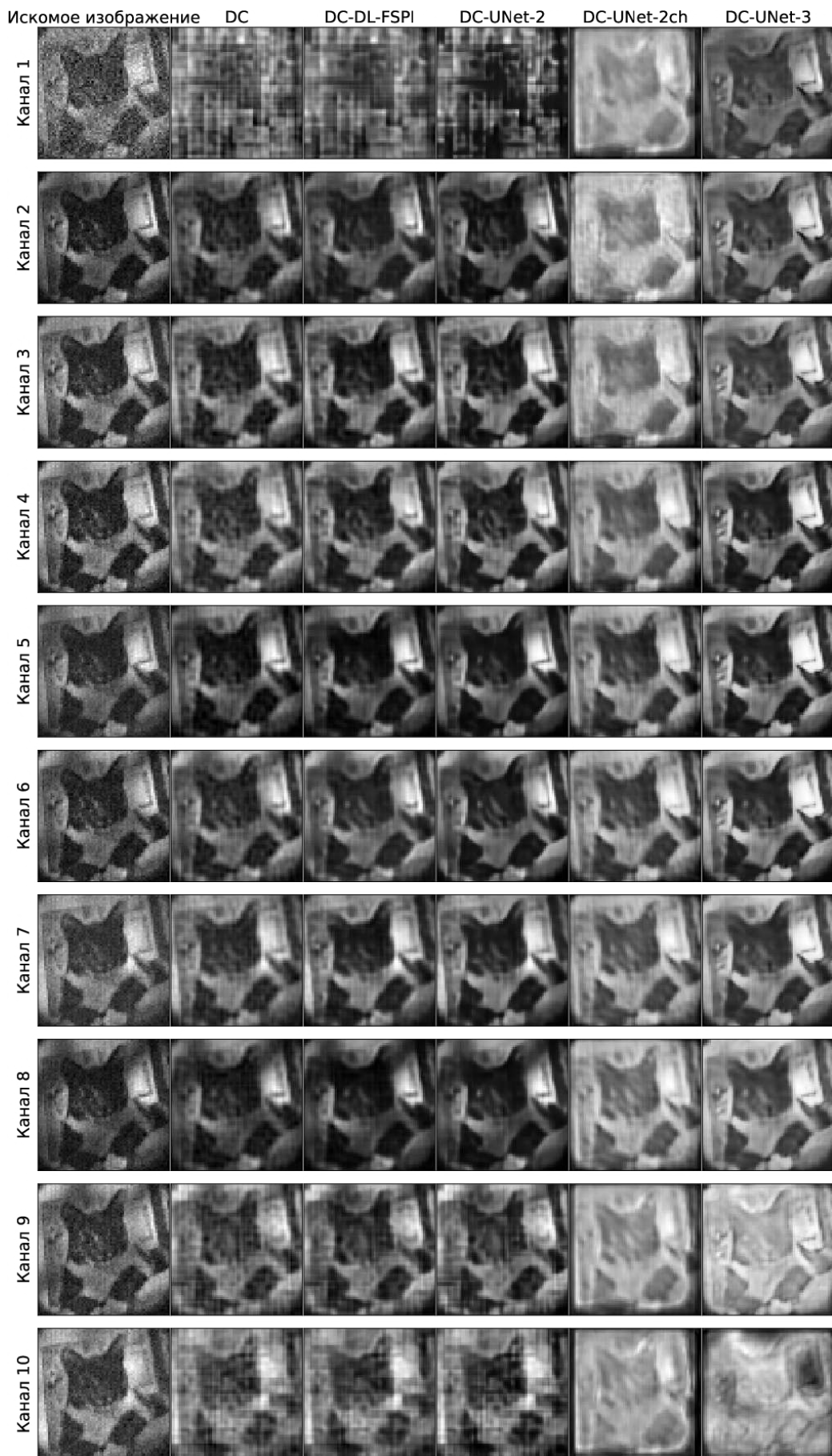


Рисунок Г.4 — Результаты восстановления экспериментальных измерений изображения кошки из STL-10.

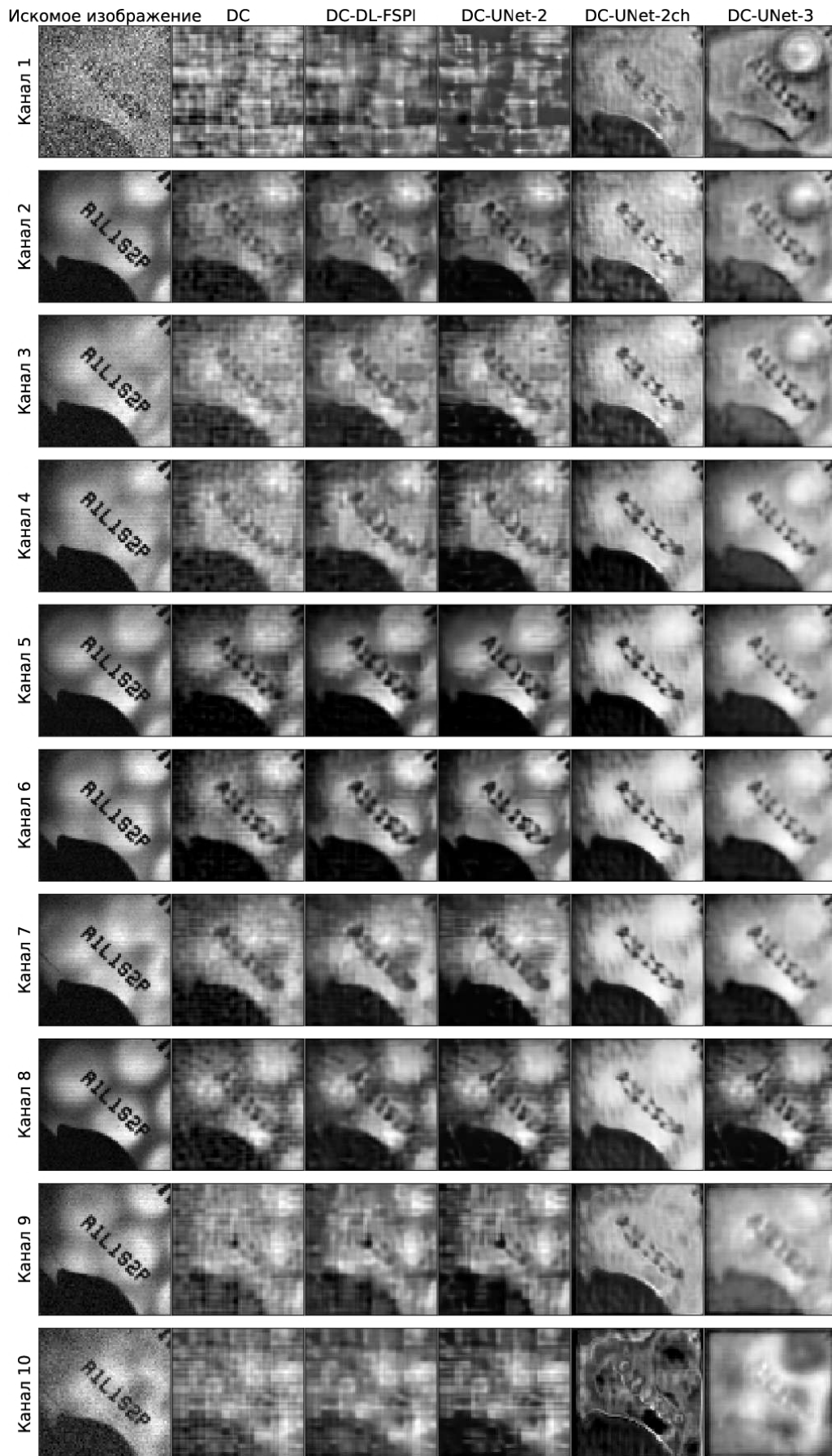


Рисунок Г.5 — Результаты восстановления экспериментальных измерений смещённой звезды Сименса.