

Министерство науки и высшего образования Российской Федерации
Санкт-петербургский политехнический университет Петра Великого
Институт электроники и телекоммуникаций
Высшая школа прикладной физики и космических технологий

Работа допущена к защите
директор ВШПФиКТ

_____ А.Л. Гельгор
« _____ » июня 2023 г.

**ВЫПУСКНАЯ КВАЛИФИКАЦИОННАЯ РАБОТА
РАБОТА БАКАЛАВРА**

**ПОВЫШЕНИЕ КАЧЕСТВА РАДИОЛОКАЦИОННЫХ ИЗОБРАЖЕНИЙ
ЗА СЧЁТ ФИЛЬТРАЦИИ МУЛЬТИПЛИКАТИВНОГО ШУМА С
ПОМОЩЬЮ МЕТОДОВ ГЛУБОКОГО ОБУЧЕНИЯ**

по направлению подготовки 11.03.01 – «Радиотехника»
Профиль 11.03.01_01 – «Космические и наземные радиотехнические
системы»

Выполнил
студент гр. 4931101/90102

В.В. Баташев

Научный руководитель
ассистент ВШПФиКТ, к.т.н

В.А. Павлов

Научный руководитель
профессор ВШПФиКТ, д.т.н

С.Б. Макаров

Консультант
по нормоконтролю

Р.И. Зудов

Санкт-Петербург
2023

САНКТ-ПЕТЕРБУРГСКИЙ ПОЛИТЕХНИЧЕСКИЙ УНИВЕРСИТЕТ ПЕТРА ВЕЛИКОГО
Институт электроники и телекоммуникаций

УТВЕРЖДАЮ

Директор высшей школы прикладной
физики и космических технологий

_____ / А.Л. Гельгор

« 28 » _____ апреля 2023 г.

ЗАДАНИЕ

по выполнению выпускной квалификационной работы

студенту _____ Баташеву Вадиму Владимировичу _____ гр. 4931101/90102

фамилия, имя, отчество (при наличии), номер группы

1. Тема работы: Повышение качества радиолокационных изображений за счёт
фильтрации мультипликативного шума с помощью методов глубокого обучения

2. Срок сдачи студентом законченной работы: 01 июня 2023

3. Исходные данные по работе:

Набор данных

<https://ieeexplore.ieee.org/Xplore/home.jsp>

3.1. Использовать открытые образовательные ресурсы и программы поиска и анализа информации.

3.2. Использовать средства автоматизации (автоматизированной) разработки: Matlab,
python

3.3. Применить (протестировать) программное обеспечение Matlab, python

4. Содержание работы (перечень подлежащих разработке вопросов):

1) Обзор подходов для фильтрации мультипликативного шума на радиолокационных
изображениях

2) Разработка и обучение фильтра на основе нейронной сети для фильтрации
мультипликативного шума.

3) Исследование влияния разработанного подхода на объекты различных форм.

4) Сравнение разработанного подхода с классическими и нейросетевыми методами
фильтрации мультипликативного шума.

5. Перечень графического материала (с указанием обязательных чертежей): нет

6. Консультанты по работе (если есть): нет

7. Дата выдачи задания 24.03.2023

Руководитель ВКР _____
(подпись)

В. А. Павлов
инициалы, фамилия

Руководитель ВКР _____
(подпись)

С. Б. Макаров
инициалы, фамилия

Задание принял к исполнению «24» марта 2023г.

Студент _____
(подпись)

В.В. Баташев
инициалы, фамилия

РЕФЕРАТ

На 43 с., 15 рисунков, 6 таблиц, 0 приложений.

КЛЮЧЕВЫЕ СЛОВА: РАДИОЛОКАЦИОННОЕ ИЗОБРАЖЕНИЕ, РАДИОЛОКАЦИОННОЕ СИНТЕЗИРОВАНИЕ АПЕРТУРЫ, МУЛЬТИПЛИКАТИВНЫЙ ШУМ, ФИЛЬТРАЦИЯ СПЕКЛ-ШУМА, ПАРАМЕТРЫ ФИЛЬТРОВ, НЕЙРОННЫЕ СЕТИ, ОПТИЧЕСКОЕ ИЗОБРАЖЕНИЕ, ГЛУБОКОЕ ОБУЧЕНИЕ

Объект исследования – радиолокационные изображения.

Цель работы – разработка и применение алгоритмов фильтрации мультипликативного спекл-шума на основе методов глубокого обучения для повышения качества радиолокационных изображений.

Для фильтрации мультипликативного спекл-шума на радиолокационных изображениях спроектированы различные архитектуры искусственных нейронных сетей. Оценка работы полученных моделей проводилась при помощи специальных метрик качества на разработанном наборе данных. Также проведено сравнение нейросетевого подхода с классическими методами.

Использовались открытые образовательные ресурсы и программы поиска и анализа информации. Использовались средства автоматизации разработки: Python. Применено программное обеспечение: MATLAB.

Результатом является обученная модель нейронной сети, которая позволяет эффективно удалять мультипликативный шум с радиолокационных изображений, тем самым повышая их качество.

ABSTRACT

43 pages, 15 pictures, 6 tables, 0 appendixes

KEYWORDS: RADAR IMAGE, SYNTHETIC APERTURE RADAR, MULTIPLICATIVE NOISE, SPECKLE NOISE FILTERING, FILTER PARAMETERS, NEURAL NETWORKS, OPTICAL IMAGE, DEEP LEARNING

The subject of the graduate qualification work is “radar images”.

The given work is devoted to developing and apply algorithms for filtering multiplicative speckle noise based on deep learning methods to improve the quality of radar images.

Various artificial neural networks have been designed to filter multiplicative speckle noise on radar images. The evaluation of the work of the obtained models was carried out using special quality metrics on the developed data set. The neural network approach is also compared with classical methods.

Open educational resources and information search and analysis programs were used. Development automation tools were used: Python. Software used: MATLAB.

The result is a trained neural network model that effectively removes multiplicative noise from radar images, thereby improving their quality.

СОДЕРЖАНИЕ

ОПРЕДЕЛЕНИЯ, ОБОЗНАЧЕНИЯ И СОКРАЩЕНИЯ.....	7
ВВЕДЕНИЕ.....	8
1 Радиолокационное синтезирование апертуры и фильтрация мультипликативного спекл-шума.....	10
1.1 Принцип действия радиолокационного синтезирования апертуры	10
1.2 Особенности РСА	11
1.3 Спекл-шум на РЛИ	13
1.4 Фильтрация спекл-шума	14
1.5 Выводы.....	17
2 Создание алгоритма на базе ИНС для фильтрации спекл-шума	18
2.1 Генерация набора данных для обучения	18
2.2 Описание архитектуры	21
2.3 Обучение модели	23
2.4 Фильтрация изображения.....	24
2.5 Метрики оценки качества	25
2.6 Выводы.....	28
3 Сравнение нейросетевого подхода к фильтрации с классическими методами	29
3.1 Фильтрация РЛИ при помощи анизотропной диффузии.....	29
3.2 Сравнение рассматриваемых фильтров.....	30
3.3 Сравнение срезов изображений.....	34
3.4 Выводы.....	38
ЗАКЛЮЧЕНИЕ	39
СПИСОК ИСПОЛЬЗОВАННЫХ ИСТОЧНИКОВ	41

ОПРЕДЕЛЕНИЯ, ОБОЗНАЧЕНИЯ И СОКРАЩЕНИЯ

Применяются следующие определения, обозначения и сокращения.

ГНС – глубокая нейронная сеть

ГСС – генеративно-состязательная сеть

ДЗЗ – дистанционное зондирование земли

ИНС – искусственная нейронная сеть

ЛА – летательный аппарат

РЛИ – радиолокационное изображение

РСА – радиолокационное синтезирование апертуры

СКО – среднеквадратическая ошибка

СНС – свёрточная нейронная сеть

GMS – gradient magnitude similarity (степень градиентного сходства)

GMSD – gradient magnitude similarity deviation (отклонение подобия величины градиента)

GMSM – gradient magnitude similarity mean (среднее значение подобия величины градиента)

PSNR – peak signal-to-noise ratio (пиковое отношение сигнал/шум)

SSIM – structural similarity index (индекс структурного сходства)

ВВЕДЕНИЕ

Радиолокационные изображения (РЛИ) — это изображения, получаемые с помощью радара, который отправляет электромагнитные волны и принимает их отражённые от объекта копии. Как правило, необходимая аппаратура устанавливается на спутник, проводящий дистанционное зондирование Земли (ДЗЗ).

РЛИ могут быть получены применением различных технологий, включая многочастотную мультистатическую радиоголограмму, обратные рассеяния, радиолокаторы с синтезированной апертурой (РСА) и другие. Данные, которые удаётся собрать при помощи перечисленных методов, несут в себе информацию о форме изучаемой поверхности, расстоянию до неё и материале. Это позволяет, используя цифровую обработку сигналов, получать двумерное изображение или трёхмерную модель.

Однако, из-за принципа формирования РЛИ, на итоговом изображении неизбежно возникает спекл-шум. Спекл-шум — это особый вид шума, который характеризуется присутствием случайных колебаний яркости на радиолокационных изображениях. Он имеет характерную зернистость, напоминающую небольшие светлые точки на изображении. Спекл-шум возникает из-за интерференции отраженных сигналов, которые подвергаются изменению формы и фазы в зависимости от геометрических и физических свойств, условий передачи, приема и преодолеваемых поверхностей. Когда множество отраженных сигналов формируют радиолокационное изображение, взаимодействие между ними может создать эффект спекла.

Спекл-шум на радиолокационных изображениях оказывает влияние на их качество и точность, что делает его присутствие негативным эффектом для анализа автоматическими системами и интерпретации изображений. Спекл-шум может создавать ложные контуры и детали на изображениях, которых не существуют в реальности, приводить к снижению контрастности, затруднять различение объектов на фоне окружающей среды, ухудшать точность

измерений параметров объектов и становится причиной множества других проблем.

Из-за вышеперечисленных причин очень важна предобработка РЛИ, направленная на удаление спекл-шума. На данный момент можно выделить несколько типов подходов к решению данной задачи.

1. Классические методы, которые извлекают статистическую информацию из изображения, на основании чего происходит фильтрация.

2. Использование различных преобразований для извлечения информации иного рода из изображений. Например, использование вейвлет-преобразования.

3. Подходы, связанные с применением методов глубокого обучения на основе искусственных нейронных сетей (ИНН), которые в процессе обучения автоматически выявляют наиболее значимые признаки для формирования изображения без шума.

Целью данной работы является проектирование и разработка фильтра мультипликативного спекл-шума на основе применения искусственных нейронных сетей.

Для достижения поставленной цели необходимо решить следующие задачи:

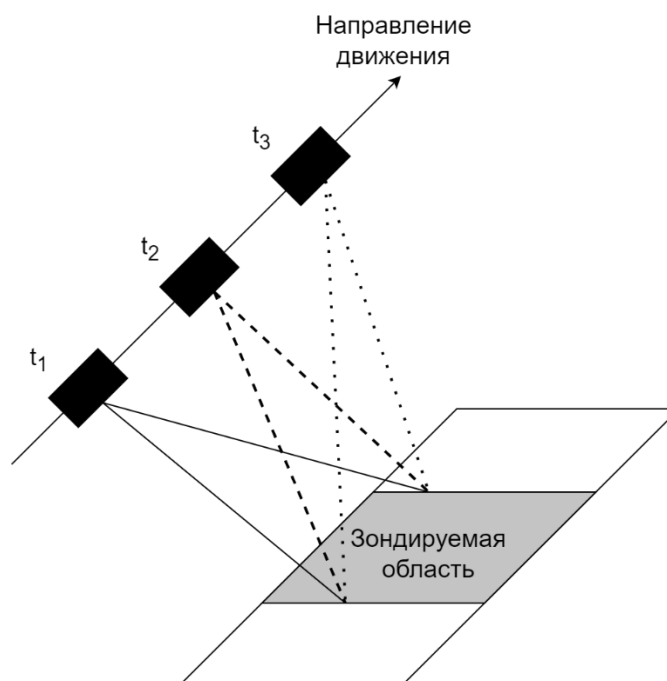
1. Разработка архитектур нейронных сетей.
2. Обучение нейронных сетей на наборе данных.
3. Оценка качества полученного фильтра при помощи метрик.
4. Сравнение различных методов.

1 Радиолокационное синтезирование апертуры и фильтрация мультипликативного спекл-шума

1.1 Принцип действия радиолокационного синтезирования апертуры

Радиолокационное синтезирование апертуры – один из способов дистанционного зондирования поверхностей, основанный на отправлении сигналов и принятии их отражённых копий. РСА имеет множество практических применений в области изучения поверхности планет, например: оценка последствий природных катастроф, слежение за вулканической активностью, изучение влияния таяние ледников, наблюдение за местоположением объектов и их поиск.

Чаще всего аппаратура, выполняющая функции РСА, устанавливается на корпус движущихся по орбите Земли спутников для обеспечения достаточно большого угла обзора. В процессе зондирования устройство отправляет когерентные, поляризованные определённым способом сигналы, после чего принимает их отражённые от изучаемой поверхности копии и сохраняет полученные данные. Из-за достаточно большого расстояния до изучаемой области, изображение, которое формируется в процессе цифровой обработки, имеет низкое пространственное разрешение, вследствие чего даже достаточно большие объекты могут стать неразличимы. Реальное увеличение антенны, которая позволит улучшить качество изображений, влечёт за собой рост стоимости конструкции и понижение её надёжности, поэтому для повышения качества формируемых изображений используется метод синтезированной апертуры. Зондирование одной и той же поверхности происходит в разные моменты времени из разных точек пространства – Рисунок 1.1, тем самым искусственно увеличивая размеры виртуальной антенны, получая намного больше информации о характере исследуемой области, что позволяет повысить пространственное разрешение в несколько раз.



t_1, t_2, t_3 – различные моменты времени зондирования.

Рисунок 1.1 – Схема зондирования поверхности РСА

1.2 Особенности РСА

Радиолокационные изображения, в отличие от оптических, имеют ряд преимуществ. Например, независимость от времени суток. Так как для зондирования используются только отправляемые сигналы и их отражённые копии, отсутствует необходимость в наличии освещения исследуемой поверхности.

Также радиолокационные изображения обладают невосприимчивостью к погодным условиям. Сигналы, отправляемые передатчиком, способны проникать сквозь облака, туман, снегопад, дождь и прочие метеорологические помехи, беспрепятственно достигая поверхности Земли и возвращаясь. Данный эффект работает и на уровне изучения поверхности: открывается возможность исследовать области, которые невозможно увидеть со спутника при помощи оптических изображений. Например, почву в лесу, где кроны деревьев перекрывают обзор.

Подобные возможности появляются благодаря выбору конкретной полосы частот, сигналы на которых способны проникать сквозь некоторые

объекты. В таблице 1 приведены значения частот и то, к каким классам они относятся. От принадлежности к определённому классу зависит то, через какие поверхности смогут проникать сигналы.

Таблица 1.1 Соответствие полос частот их обозначениям для РСА

Название полосы частот	Значения частот, ГГц
Ka	40,0 - 26,5
K	26,5 - 18,0
Ku	18,0 - 12,5
X	12,5 - 8,0
C	8,0 - 4,0
S	4,0 - 2,0
L	2,0 - 1,0
P	1,0 - 0,3

Важным параметром РЛИ является «пространственное разрешение», характеризующее минимальные геометрические размеры, которыми должен обладать объект на поверхности, чтобы его можно было различить на полученном РЛИ. Для технологии РСА это один из показателей эффективности всей системы. Следовательно, для успешного решения задач необходимо стараться максимизировать данный показатель. Есть несколько способов это сделать. Можно выделить три основных подхода повышения пространственного разрешения изображения, но они влекут за собой уменьшение площади изучаемой поверхности: ScanSAR, Stripmap и Spotlight. За одинаковое количество итераций зондирования наилучшее качество изображения покажет метод Spotlight, после него Stripmap и далее ScanSAR. В обратном порядке увеличивается исследуемая площадь. На Рисунке 1.2 схематично изображены процессы исследования поверхности для рассматриваемых трёх подходов.

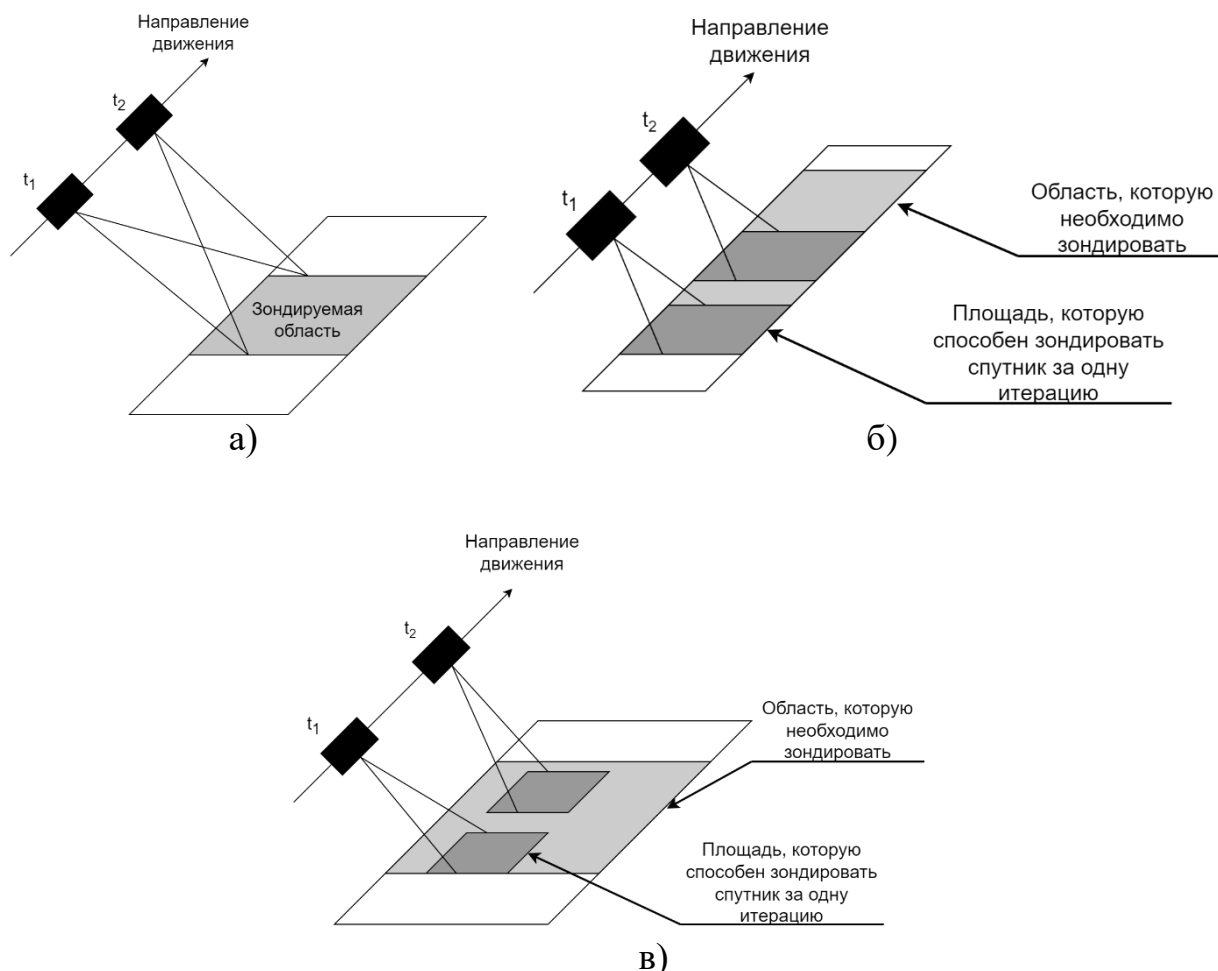


Рисунок 1.2 – Зондирование поверхности тремя способами: а) Spotlight;
б) StripMap; в) ScanSAR.

1.3 Спекл-шум на РЛИ

Спекл-шум представляет собой мультипликативный паттерн случайных ярких и темных точек на изображении, которые могут исказить искомый объект и делать его неразличимым.

В отличие от оптических изображений, где световые волны не меняются при отражении от неоднородных объектов, радиолокационные волны, излучаемые когерентными источниками, могут испытывать случайные изменения фазы и амплитуды при отражении от объектов в окружающей среде. Интерференция, в результате пересечения этих волн, приводит к возникновению спекл-шума на изображении.

Для некоторых типов задач анализ распределения спекл-шума и его изменение во времени позволяет делать сложные выводы о структуре объекта, таким образом, даже данный тип помех способен нести в себе полезную информацию, например: для изучения двойных звёзд, их массы и расстояния до них. Но в случае зондирования поверхности земли, спекл-шум затрудняет визуальное анализирование радиолокационных изображений и понижает качество автоматической обработки, а именно сегментацию объектов, классификацию, детектирование и другое. Поэтому, чтобы получить более точную информацию об исследуемой поверхности, необходимо проводить предобработку данных: фильтрацию спекл-шума.

1.4 Фильтрация спекл-шума

На сегодняшний день существуют десятки различных фильтров радиолокационных изображений, которые можно условно разделить на несколько категорий.

Первые подходы к фильтрации РЛИ были представлены ещё с момента появления самих радиолокационных изображений. Изначально разработанные методы опирались на использование локальных статистических данных в исследуемой квадратной области для извлечения и удаления спекл-шума. Наиболее известными оказались фильтр Ли [1], фильтр Фроста [2], фильтр Куана [3].

Другой подход к фильтрации радиолокационных изображений заключается в применении вейвлет-преобразований [4, 5, 6]. Если стандартное представление сигнала во временной области не даёт информации о частотной составляющей, а использование Фурье-преобразования, наоборот, сохраняет только информацию о частоте сигнала, не оставляя информации о времени, то вейвлет-преобразование является обобщением спектрального анализа и заключается в извлечении частотных признаков из сигналов, но при этом с сохранением временных параметров. Способы фильтрации радиолокационных изображений, основанные на использовании вейвлет-преобразований показывают более хорошие

результаты, чем подходы, связанные с использованием локальных статистических данных в одной области. Применение вейвлет-преобразований стало широко распространенным в области обработки сигналов и изображений, включая такие сферы как компьютерное зрение, медицинская диагностика, видеообработка и другие.

Отдельной группой являются подходы, использующие для фильтрации и генерации нового пикселя без шума информацию обо всём изображении – нелокальные методы [7, 8]. Один из первых таких фильтров – нелокальный фильтр средних [9], в основе которого лежит вычисление евклидова или другого расстояния от каждого пикселя до каждого и введения весовых коэффициентов, которые зависят от полученного значения расстояния. Таким образом, наиболее отдалённые пиксели будут оказывать меньший вклад, чем находящиеся рядом.

С ростом производительности и возможностью параллельных вычислений начали активно развиваться подходы к обработке изображений, использующие глубокие нейронные сети (ГНС). Сложность применения ГНС для фильтрации радиолокационных изображений состоит в том, что для создания модели, способной удалить шум с изображения, необходимо эту модель обучить: подавать на вход зашумлённые изображения, и изображения без шума. Так как на радиолокационных изображениях всегда присутствует спекл-шум, то не существует исходных примеров без шума. Это ограничение можно обойти, искусственно накладывая спекл-шум на чистые оптические изображения.

Как показала практика, применение архитектуры на основе свёрточных нейронных сетей (СНС) в задачах обработки изображений даёт хорошие результаты за счёт использования информации о пространственной структуре. Во время обучения нейронная сеть настраивается на извлечение признаков из РЛИ и на удаление шума, чтобы получить очищенную от шума версию изображения. В статье [10] описывается механизм избавления от аддитивного белого Гауссовского шума с фиксированным уровнем,

основанный на вычитании извлечённого при помощи СНС шума из изображения. Данный подход был развит и в [11] авторы статьи обобщили его на любой уровень шума. Конкретно задача удаления спекл-шума хорошо решается при помощи архитектуры нейронной сети, основанной на свёрточных слоях и остаточных соединениях [12].

Архитектура на основе автоэнкодера позволяет обучить нейронную сеть сжимать исходное изображение до более компактного представления с меньшей размерностью и оставлять наиболее значимые компоненты, после чего расширить полученное представление обратно в изображение исходного размера, но без шума. На основе данной идеи авторы статьи [13] предложили одновременное использование нейронной сети на базе свёрточной архитектуры для увеличения поля восприимчивости с применением механизма автокодировщика для извлечения важных признаков.

Более сложным является применение архитектуры на основе генеративно-сопоставительной сети (ГСС), идея которой базируется на двух нейронных сетях: генератор получает на вход зашумлённое изображение и старается его отфильтровать. дискриминатор получает на вход либо изображение с выхода генератора, либо изображение без шума, после чего определяет, является ли это изображение изначально без шума, или получено после фильтрования генератором. Таким образом, две нейронные сети состязаются и обучаются вместе. Главная проблема заключается в сложности выбора функции потерь для обучения и длительности процесса. На основе такого подхода разработана система [14], задача которой заключается в повышении разрешения радиолокационных изображений.

Отдельно можно выделить обучение нейронной сети на основе трансформеров. Впервые такой тип архитектуры был представлен в 2017 году, её отличительной особенностью стало появление механизма внимания [15] для извлечения зависимостей в последовательностях. Это позволило проектировать модели, способные справляться с решением задач обработки естественного языка и распознавания звука [16] намного лучше

альтернативных методов, в частности, рекуррентных нейронных сетей. Адаптацией данного метода для обработки изображений стал проект Визуальный Трансформер [17]. Использование механизмов внимания в комбинации с другими методами применяется во различных областях. На основе данного подхода были решены многие задачи, в частности и фильтрация радиолокационных изображений. Например, применение вейвлет-преобразования и глубокой нейронной сети на базе Трансформеров [18]. Ключевая идея заключается в извлечении высокочастотной и низкочастотной информации с последующим нахождением зависимостей между ними при помощи механизма внимания.

1.5 Выводы

Исследования в области проектирования продолжают и до сих пор, что говорит об актуальности проблемы и необходимости поиска более оптимальных и менее вычислительно сложных решений, которые можно будет выполнять в режиме реального времени на спутниках, проводящих зондирование земли при помощи РСА.

Целью данной работы является проектирование алгоритмом фильтрации мультипликативного спекл-шума с применением искусственных нейронных сетей, направленные на повышение качества радиолокационных изображений.

Дальнейшие задачи, которые необходимы для достижения поставленной цели:

1. Проектирование различных методов фильтрации РЛИ с использованием методов глубокого обучения
2. Выбор оценки метрики качества после процесса фильтрации
3. Обучение спроектированных архитектур нейронных сетей на заготовленном наборе данных
4. Оценка результата по сравнению с классическими подходами.

2 Создание алгоритма на базе ИНС для фильтрации спекл-шума

2.1 Генерация набора данных для обучения

В качестве данных для обучения нейронной сети необходимо иметь набор из двух пар: изображение, которое нужно отфильтровать и исходное изображение без шума. Но из-за того, что любые РЛИ обладают спекл-шумом, возникает необходимость синтезировать данные для обучения. Процесс генерации новых данных заключается в наложении на обычные оптические изображения поверхности Земли распределения Рэля с параметром масштаба 0,27 [19], которое описывает спекл-шум на РЛИ. На Рисунке 2.1 приведена плотность вероятности используемого распределения шума.

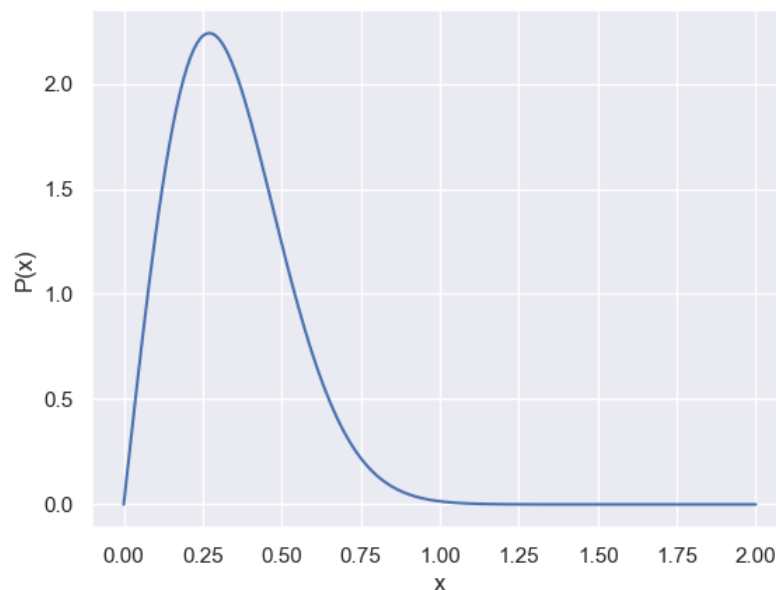


Рисунок 2.1 – Плотность вероятности распределения Рэля с параметром масштаба 0,27.

Добавление шума на изображение соответствует формуле (2.1)

$$\hat{I} = I + I \cdot M_{noise}, \quad (2.1)$$

где \hat{I} – итоговое зашумлённое изображение,

M_{noise} – шум,

I – исходное изображение.

На Рисунке 2.2 показано оптическое изображение без шума и его версия с наложением описанного распределением шума.

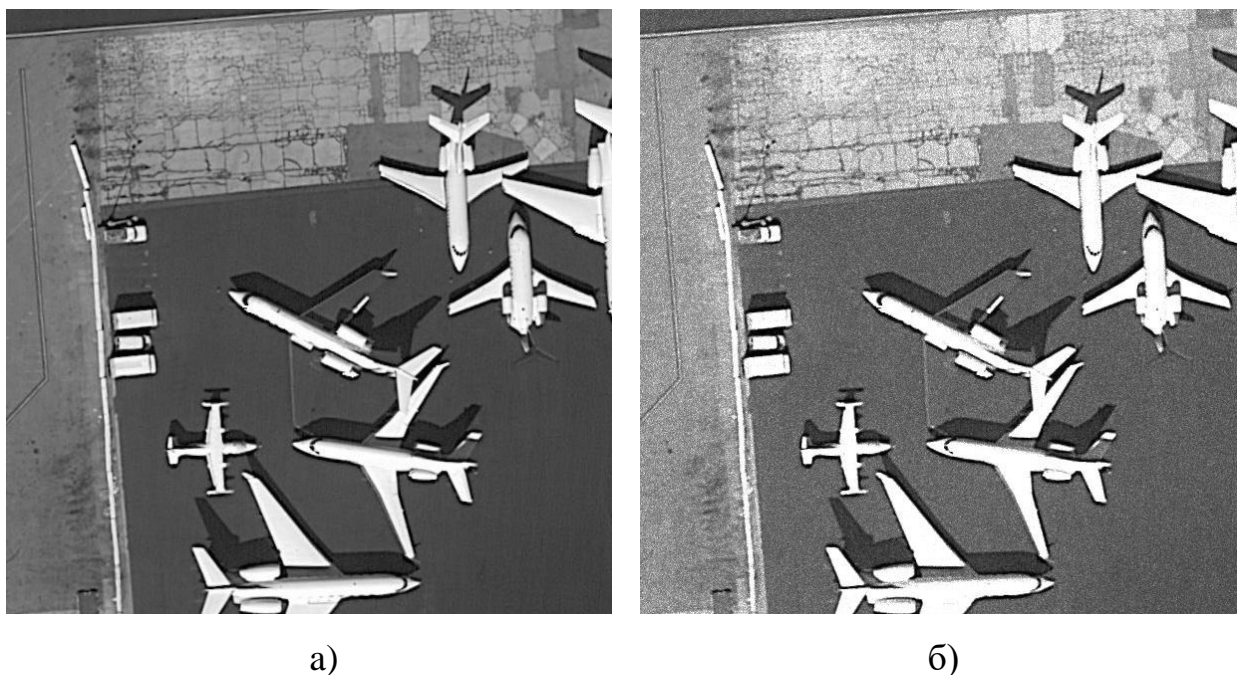


Рисунок 2.2 – а) оригинальное оптическое изображение; б) Оптическое изображение с наложенным спекл-шумом.

Получение набора данных заключается в генерации пар: данные зашумлённого изображения на входе и ожидаемые, незашумлённые, на выходе. В качестве входных данных выбирается квадратное окно из зашумлённого изображения, сторона которого равна N , где N – нечётное. Ожидаемое значение на выходе – это пиксель в центре окна на исходном изображении без шума. Таким образом, окно скользит вдоль всего изображения с заданным шагом. На Рисунке 2.3 графически показан данный процесс на изображении размера 3×3 , к которому добавили границы, равные размеру окна $(N-1)/2$, где ширина окна составляет $N = 3$ и шаг окна равняется 1.

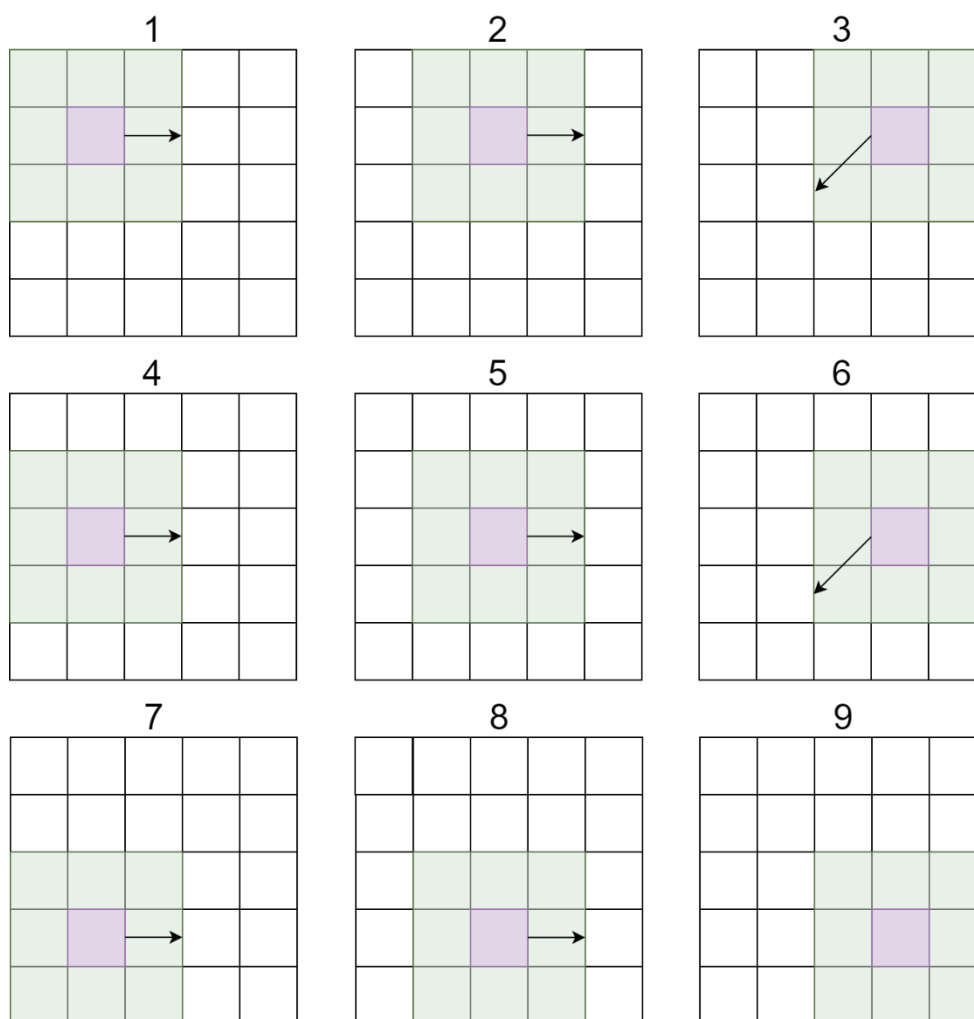


Рисунок 2.3 – Процесс «скольжения» окна с шириной окна $N = 3$ по изображению.

Полученные данные преобразуются в одномерный массив и записываются в файл, либо переменную для дальнейшего использования. Рисунок 2.4 иллюстрирует, каким образом двумерные данные конвертируются в одномерный вектор.

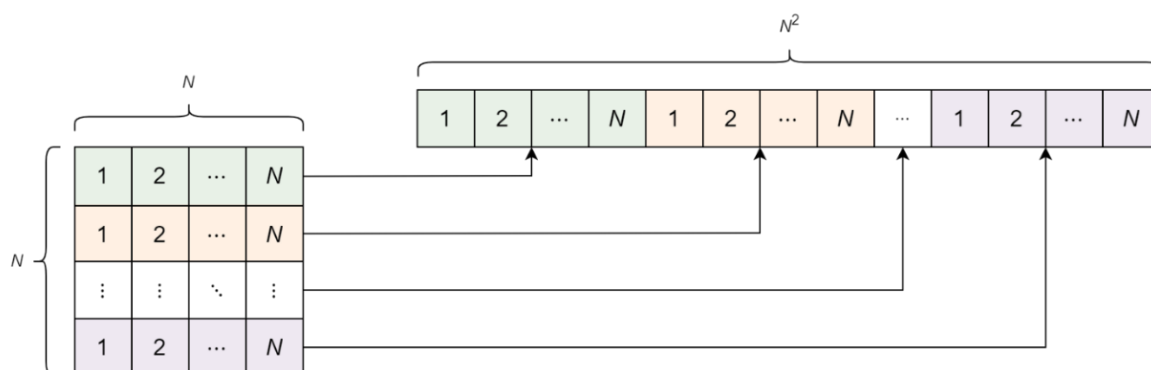


Рисунок 2.4 – Процесс преобразование матрицы в вектор в общем виде. N – количество элементов в строке и количество строк

Изложенный алгоритм действий применяется последовательно ко всем изображениям.

2.2 Описание архитектуры

Архитектура нейронной сети состоит из полносвязных слоёв: входной слой, выходной и 4 скрытых. Для повышения качества работы модели используется пакетная нормализация данных между слоями, в качестве функции активации выбрана ReLU, т. к. она подходит лучше других для задач обработки изображений и является вычислительно простой [20]. Количество нейронов входного слоя равняется квадрату ширины окна N^2 . В задаче регрессии на выходе только один нейрон. Если задача классификации, то на выходе должен быть вектор длиной 256. Общая структура нейронной сети для задачи регрессии и классификации приведены на Рисунках 2.5 и 2.6 соответственно. Количество нейронов для каждого слоя указано в скобках.

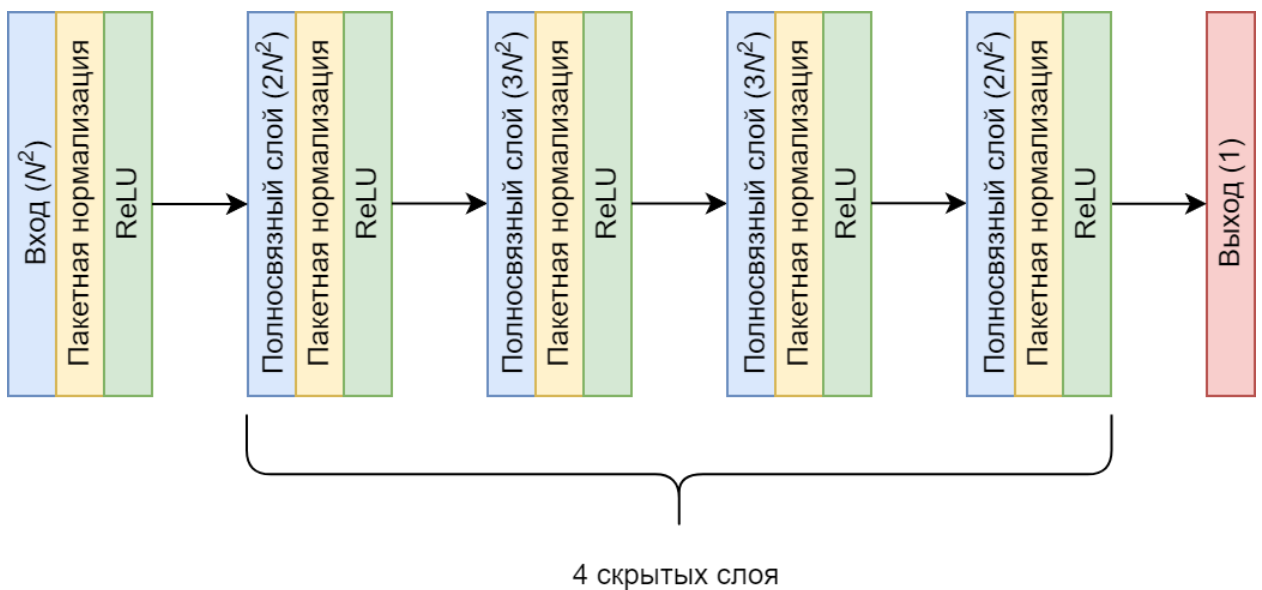


Рисунок 2.5 – Архитектура полносвязной нейронной сети для решения задачи регрессии

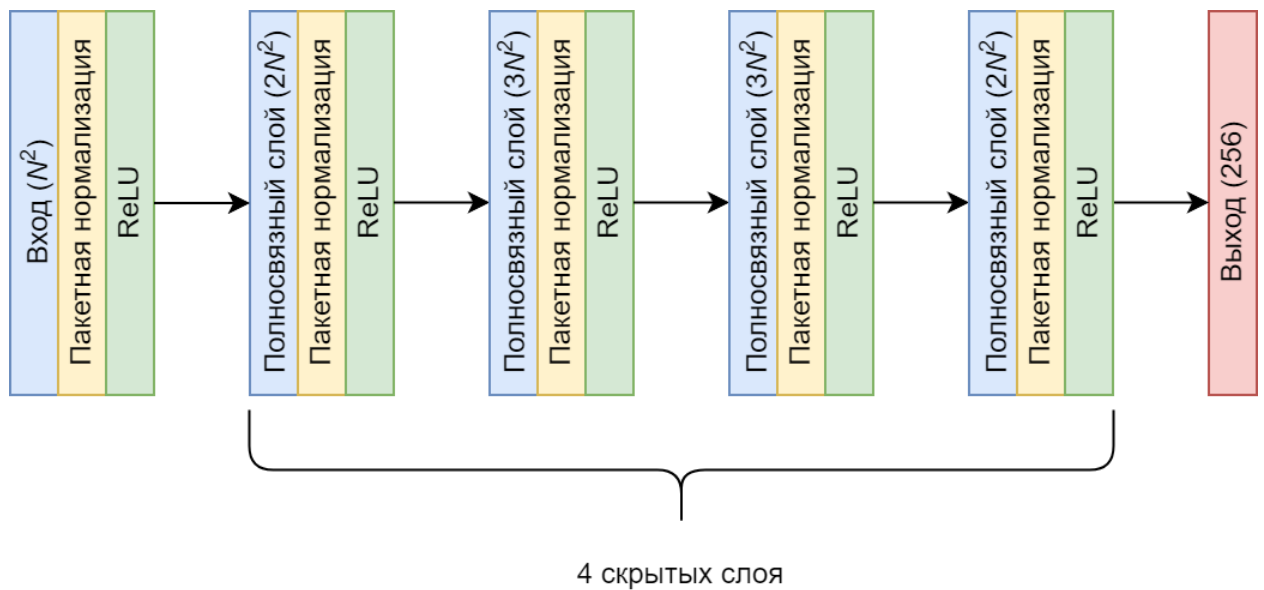


Рисунок 2.6 – Архитектура полносвязной нейронной сети для решения задачи классификации

Архитектура нейронной сети для решения задачи регрессии выдаёт на выходе модели предсказание одного числа исходя из входных данных. В рассматриваемом случае предсказываемым числом является яркость одного пикселя.

Для задачи классификации модель на выходе предсказывает, к какому классу относится объект на входе, где индекс класса эквивалентен яркости. Индекс с максимальным значением в выходном векторе соответствует яркости пикселя. Следовательно, в решаемой задаче классов должно быть столько, сколько значений может принимать яркость пикселя. В нашем случае это отрезок дискретных значений от 0 до 255. Для получения одного значения из массива данных на выходе применяется следующая формула (2.2):

$$p = \arg \max\{out\}, \quad (2.2)$$

где p – яркость предсказываемого пикселя,
 out – данные на выходе модели.

2.3 Обучение модели

Обучение модели происходит на синтезированных данных. На вход подаётся случайная часть изображения – окно.

Функцией потерь для задачи регрессии выбрана среднеквадратическая ошибка (СКО). СКО вычисляет среднее квадрата разности между действительным значением признака и предсказываемым, согласно выражению (2.3). В данном контексте в качестве признака рассматривается отдельный пиксель изображения. Действительное значение признака – пиксель незашумлённого изображения, предсказываемый – пиксель на выходе модели.

$$L_{\text{СКО}} = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2, \quad (2.3)$$

где n – количество признаков,

y_i – исходное значение признака,

\hat{y}_i – предсказываемое моделью значение признака.

Так как в задаче классификации модель возвращает не одно значение, а вектор, то необходимо реальное значение пикселя предварительно закодировать в вектор длины 256, состоящий из нулей и одной единицы, индекс которой в массиве характеризует яркость пикселя. Для решения задачи классификации в качестве функции потерь выбрана кросс-энтропия, выражение 2.4.

$$L_{\text{КЭ}} = - \sum_{i=1}^n t_i \log(p_i), \quad (2.4)$$

где $L_{\text{КЭ}}$ – функция потерь кросс-энтропии,

t_i – номер истинного класса,

p_i – вероятность предсказанного класса.

Вектор вероятностей предсказанных классов вычисляется при помощи логистической функций, формула 2.5

$$P(y)_i = \frac{\exp(y_i)}{\sum_{j=1}^n \exp(y_j)}, \quad (2.5)$$

где $P(y)_i$ – логистическая функция,

y_i, y_j – значение класса на выходе модели,

n – количество классов.

Оптимизатором выбран алгоритм Адам [21], который проводит поиск минимума функции за счёт минимизации функции потерь.

Количество эпох выбрано равным 20. Шаг обучения в процессе оптимизации функции потерь уменьшается в несколько раз, позволяя получить более точную и быструю сходимость. В таблице 2.1 приведены номера эпох, начиная с которых шаг обучение принимает новое значение.

Таблица 2.1 – номера эпох и соответствующие им значения шага обучения

Номер эпохи	Шаг обучения
1	0,1
5	0,01
10	0,001
15	0,0001

2.4 Фильтрация изображения

Фильтрация изображения происходит следующим образом: к РЛИ добавляются границы, равные половине ширины окна с округлением вниз, после чего на вход модели подаётся одно окно пикселей с фиксированной шириной. Так как результат модели – предсказание одного незашумлённого

пикселя, описанный алгоритм применяется столько раз, сколько пикселей в исходном изображении, после чего полученные данные формируют, согласно координатам центра окна, в итоговое отфильтрованное изображение.

2.5 Метрики оценки качества

Для оценивания эффективности модели необходимо задать метрику, которая будет сравнивать отфильтрованное изображение с исходным. В данной работе оценка производилось при помощи следующих алгоритмов: индекс структурного сходства (SSIM – Structural Similarity Index) [22] и отклонение подобия величины градиента (GMSD – gradient magnitude similarity deviation) [23].

SSIM позволяет оценить сходство между двумя изображениями, основываясь на их структурных характеристиках. Значение метрики SSIM формируется на основе трех факторов: яркости, контрастности и структуры.

Параметр яркости оценивает сходство между средними яркостями с помощью коэффициента яркости l , который вычисляется по формуле 2.6.

$$l = \frac{2 \cdot \mu_1 \cdot \mu_2 + c_1}{\mu_1^2 + \mu_2^2 + c_1}, \quad (2.6)$$

где μ_1 и μ_2 – средние яркости эталонного и сравниваемого изображений соответственно,

c_1 – константа для предотвращения деления на ноль.

Контрастность: этот фактор оценивает сходство между стандартными отклонениями пикселей изображений. Оценивается с помощью коэффициента контрастности c , который вычисляется по формуле 2.7.

$$c = \frac{2 \cdot \sigma_1 \cdot \sigma_2 + c_2}{\sigma_1^2 + \sigma_2^2 + c_2}, \quad (2.7)$$

где σ_1 и σ_2 – стандартные отклонения пикселей изображений соответственно,

c_2 – константа для предотвращения деления на ноль.

Структура: данный коэффициент оценивает сходство между корреляцией пикселей изображений. Оценивается с помощью коэффициента структуры s , который вычисляется по формуле 2.8.

$$s = \frac{\sigma_{12} + c_3}{\sigma_1 \cdot \sigma_2 + c_3}, \quad (2.8)$$

где σ_{12} – ковариация между пикселями изображений,

c_3 – константа для предотвращения деления на ноль.

Общий коэффициент SSIM вычисляется как произведение трех вышеперечисленных коэффициентов по формуле 2.9.

$$SSIM = l^\alpha \cdot c^\beta \cdot s^\gamma, \quad (2.9)$$

где α , β и γ – весовые коэффициенты, которые задают важность каждого из факторов.

Обычно значения α , β и γ принимают равными 1, также сделано и в текущей работе.

Значение метрики SSIM лежит в пределах $[-1, 1]$. Чем ближе к 1, тем сравниваемые изображения считаются более схожими по своей структуре. Если значение метрики достигает 1 – изображения идентичны.

GMSD — это метод оценки качества изображения, который измеряет отклонение между градиентами эталонного и сравниваемого изображений, что делает ее более чувствительной к искажениям, связанным с высокочастотными деталями и текстурой, чем другие метрики, такие как пиковое отношение сигнал/шум (PSNR – Peak Signal-to-Noise Ratio) или SSIM.

Градиент вычисляется обычной свёрткой изображения с линейным фильтром, например: Собеля, Щарра, Прюитта. Последний используется для

поставленной задачи в силу своей простоты, его матрицы h_x и h_y для осей x и y соответственно представлены в выражении 2.10.

$$h_x = \begin{bmatrix} \frac{1}{3} & 0 & -\frac{1}{3} \\ \frac{1}{3} & 0 & -\frac{1}{3} \\ \frac{1}{3} & 0 & -\frac{1}{3} \end{bmatrix}, h_y = \begin{bmatrix} \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \\ 0 & 0 & 0 \\ -\frac{1}{3} & -\frac{1}{3} & -\frac{1}{3} \end{bmatrix} \quad (2.10)$$

Значения градиентов m_r и m_d вычисляются для каждой области i исходного и сравниваемого изображений по формулам 2.11 и 2.12 соответственно.

$$m_r(i) = \sqrt{(r \otimes h_x)^2(i) + (r \otimes h_y)^2(i)}, \quad (2.11)$$

$$m_d(i) = \sqrt{(d \otimes h_x)^2(i) + (d \otimes h_y)^2(i)}, \quad (2.12)$$

где h_x, h_y – матрицы линейного фильтра для соответствующих осей,

r – исходное изображение,

d – сравниваемое,

\otimes – оператор свёртки.

Для упрощения введём два промежуточных значения: степень градиентного сходства (GMS – gradient magnitude similarity), формула (2.13) и среднее значение подобия величины градиента (GMSM – gradient magnitude similarity mean), формула (2.14). Они позволяют пересчитать значение искомой метрики. Первое выражение характеризует карту градиентов, второе характеризует усреднение карт градиентов.

$$GMS(i) = \frac{2m_r(i) \cdot m_d(i) + c}{m_r^2(i) + m_d^2(i) + c}, \quad (2.13)$$

$$GMSM = \frac{1}{N} \sum_{i=1}^N GMS(i), \quad (2.14)$$

где c – константа для предотвращения деления на ноль.

Итоговая метрика GMSD вычисляется по формуле (2.15). Чем ближе получаемое значение к нулю, тем более похожими считаются изображения. Если значение метрики строго равно нулю, значит изображения на входе идентичны.

$$GMSD = \sqrt{\frac{1}{N} \sum_{i=1}^N (GMS(i) - GMSM)^2} \quad (2.15)$$

2.6 Выводы

Нейронные сети зачастую решают задачи обработки изображений лучше, чем классические алгоритмы за счёт итеративного обучения на наборе данных. Это позволяет вести автоматический поиск оптимальных параметров – весовых коэффициентов модели. Чтобы данный процесс ускорить и сделать более стабильным, применяются различные дополнительные техники, например, пакетная нормализация между слоями, адаптивные оптимизаторы целевой функции, изменение шага обучения в зависимости от количества пройденных эпох и другие. С повышением числа параметров нейронной сети её обобщающая способность увеличивается, но при этом также растёт время, необходимое на обучение большего числа весовых коэффициентов и уменьшается стабильность.

Для оценки эффективности работы полученной модели используются метрики качества. В задачах обработки изображений таковыми метриками могут являться алгоритмы SSIM и GMSD, которые способны численно оценить степень сходства между двумя изображениями.

3 Сравнение нейросетевого подхода к фильтрации с классическими методами

3.1 Фильтрация РЛИ при помощи анизотропной диффузии

В качестве классического метода выбран фильтр анизотропной диффузии, так как он справляется с поставленной задачей лучше других классических подходов [24].

Анизотропная диффузия – один из классических методов обработки изображений, цель которого заключается в устранении шумовых составляющих, при этом оставляя нетронутыми важные части изображения, такие как границы, линии и другие детали, несущие в себе основную информацию для интерпретации. Идея фильтра заключается в том, чтобы различать направление основных градиентов в изображении. Он выполняет итеративные операции, в которых учитывается интенсивность пикселей и их изменения в разных направлениях. Таким образом, позволяет сохранить резкие границы и детали изображения, минимизируя размытие. Процесс фильтрации анизотропной диффузии основывается на уравнении диффузии, которое определяет, как интенсивность пикселя изменяется со временем. Это уравнение (3.1) имеет следующий вид.

$$\frac{\partial I}{\partial t} = \text{div}(c(x, t)\nabla I), \quad (3.1)$$

где I – матрица интенсивностей пикселей изображения,

t – номер итерации,

$c(x, t)$ – коэффициент диффузии,

∇I – градиент яркости пикселей изображения,

div – оператор дивергенции.

Коэффициент $c(x, t)$ определяет, насколько быстро диффузия будет происходить в каждой точке изображения. В анизотропной диффузии этот параметр настраивается в зависимости от разности интенсивностей пикселей

и их градиентов. Если разность интенсивностей большая, то коэффициент диффузии будет небольшим, что помогает сохранить границы и детали изображения. Если разность интенсивностей мала, то коэффициент диффузии будет большим, что способствует сглаживанию шума.

Фильтр анизотропной диффузии принимает на вход 4 аргумента: само изображение, количество итераций t , шаг по времени Δt и параметр k . Оптимальные значения для последних трёх переменных при решении задачи фильтрации спекл-шума были исследованы в статье [24]. Данные применяются в текущей работе и приведены в таблице 3.1.

Таблица 3.1 – Оптимальные параметры для фильтра анизотропной диффузии квадратичной функции

Параметр	Оптимальное значение
t	13
Δt	0,25
k	0,05

3.2 Сравнение рассматриваемых фильтров

На Рисунке 3.1 приведены два оптических изображения, а также их копии с наложением мультипликативного спекл-шума. Так как данные изображения несут в себе информацию, схожую с той, которая содержится на радиолокационных изображениях: поверхность Земли, то имеет смысл численно проводить оценку работы разработанных алгоритмов на них.



а)



б)



в)



г)

Рисунок 3.1 – а), б) исходные изображения; в), г) зашумлённые изображения

Рисунок 3.2 иллюстрирует результат работы фильтра анизотропной диффузии с оптимальными параметрами над вышеуказанными изображениями с наложенным шумом.



а)



б)

Рисунок 3.2 – результат фильтра анизотропной диффузии

Применение нейросетевого фильтра с задачей регрессии показано на Рисунке 3.3.



а)



б)



в)



г)

Рисунок 3.3 – изображения, отфильтрованные нейронной сетью, с задачей регрессии а), б) с применением нормализации входных данных; в), г) без применения нормализации

Применение нейросетевого фильтра с задачей классификации показано на Рисунке 3.4.



а)



б)



в)



г)

Рисунок 3.4 – изображения, отфильтрованные нейронной сетью, с задачей классификации а), б) с применением нормализации входных данных; в), г) без применения нормализации

Визуально затруднительно объективно определить то, какой из представленных фильтров мультипликативного спекл-шума работает лучше. Для повышения точности использовался набор данных с 1000 оптических изображений земной поверхности, на каждое из которых накладывался шум, а после применялся каждый из рассматриваемых фильтров. Оценка работы различных фильтров проводилась при помощи вышеописанных метрик сходства изображений SSIM и GMSD. Результаты представлены в таблице 3.2.

Таблица 3.2 – численные значения метрик для исходных изображений и отфильтрованных различными способами.

Тип фильтра	Размер окна N , пикселей	Количество скрытых слоёв	Использование нормализации	SSIM	GMSD
ИНС Регрессия	7×7	5	Нет	0,877	0,056
ИНС Регрессия	7×7	5	Да	0,846	0,064
ИНС Классификация	7×7	5	Нет	0,864	0,062
ИНС Классификация	7×7	5	Да	0,863	0,065
Фильтр анизотропной диффузии	-	-	-	0,827	0,078

Лучший результат показал фильтр на основе нейронной сети, решающий задачу регрессии без нормализации.

3.3 Сравнение срезов изображений

Для более визуально понятного сравнения рассматриваемых фильтров воспользуемся следующим способом. На Рисунке 3.5 приведено оптическое изображение, в центре которого взят срез пикселей. Координата среза по оси y составляет $y = 350$. По оси x взяты все пиксели.

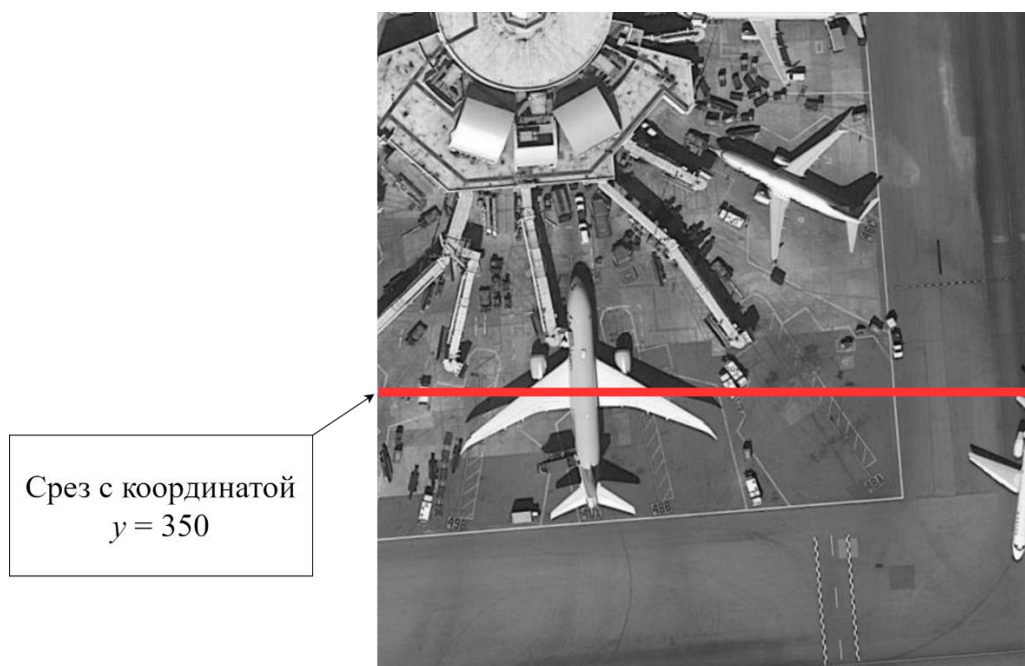


Рисунок 3.5 – срез пикселей на оптическом изображении

Далее проведём операцию взятия среза для такого же изображения, но с наложенным шумом, а также после фильтрации. Графики полученных срезов приведены на Рисунке 3.6. На оси x отображены координаты пикселей, соответствующие координатам на исходном изображении. По оси y показана интенсивность пикселей. В идеале результирующий график должен повторять график исходного изображения, что означает отсутствие отличий между оригинальным изображением и отфильтрованным.

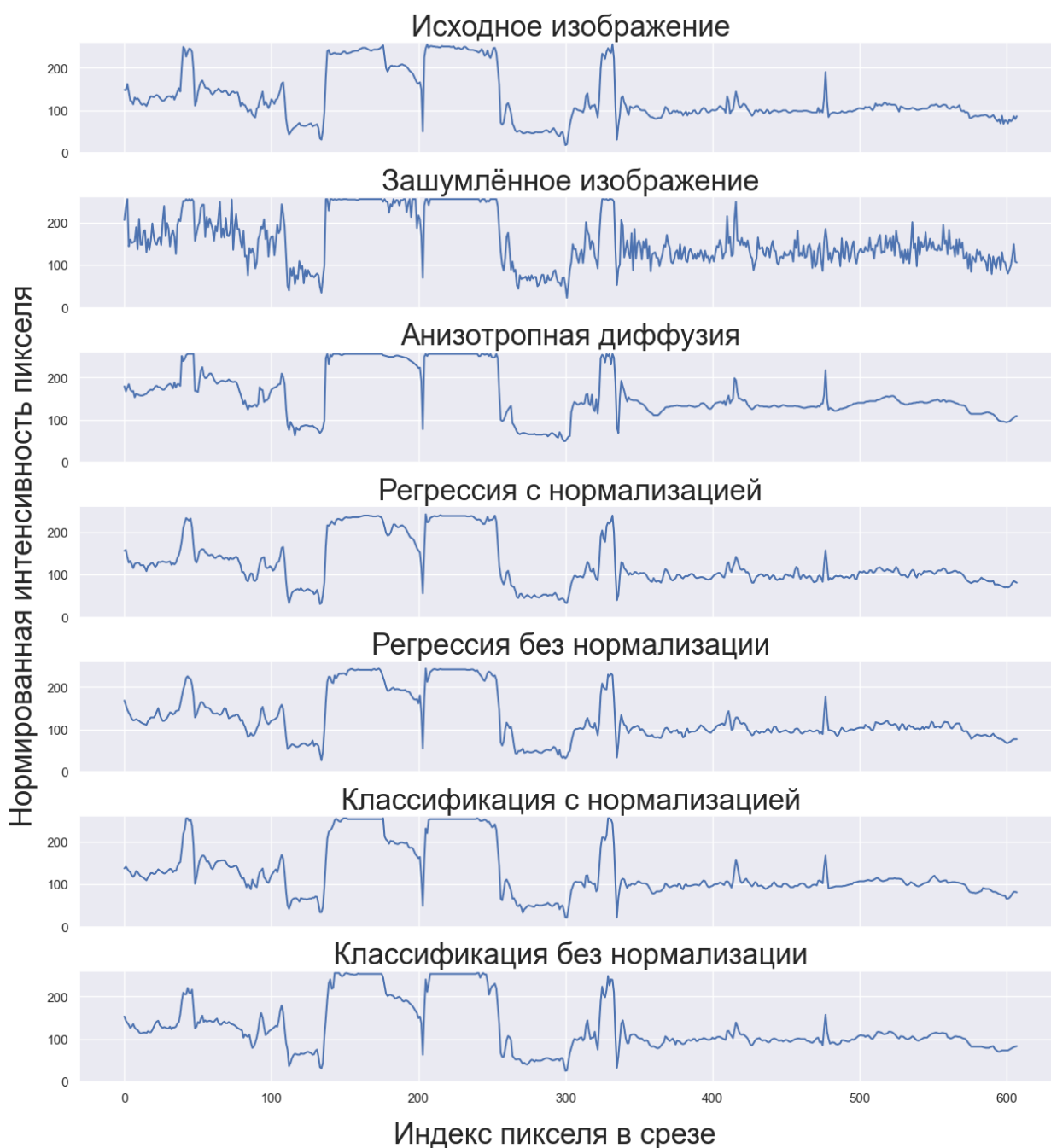


Рисунок 3.6 – интенсивность пикселей на срезе различных изображений

Из рисунка 3.6 видно, что фильтр анизотропной диффузии хорошо справляется с задачей сглаживания и размытия, но вследствие этого теряются явные границы и контуры на изображении, которые способны нести в себе важную информацию. Фильтры на базе нейронных сетей меньше подвержены размывать изображения, что позволяет сохранить после обработки исходное пространственное разрешение.

Рисунок 3.7 аналогичен предыдущему, но на графиках отображена разность между преобразованным срезом и исходным. Это позволяет наглядно оценить, на сколько в среднем отклоняется яркость в данном срезе, изменение которой произошло в процессе фильтрации зашумлённого изображения различными методами.

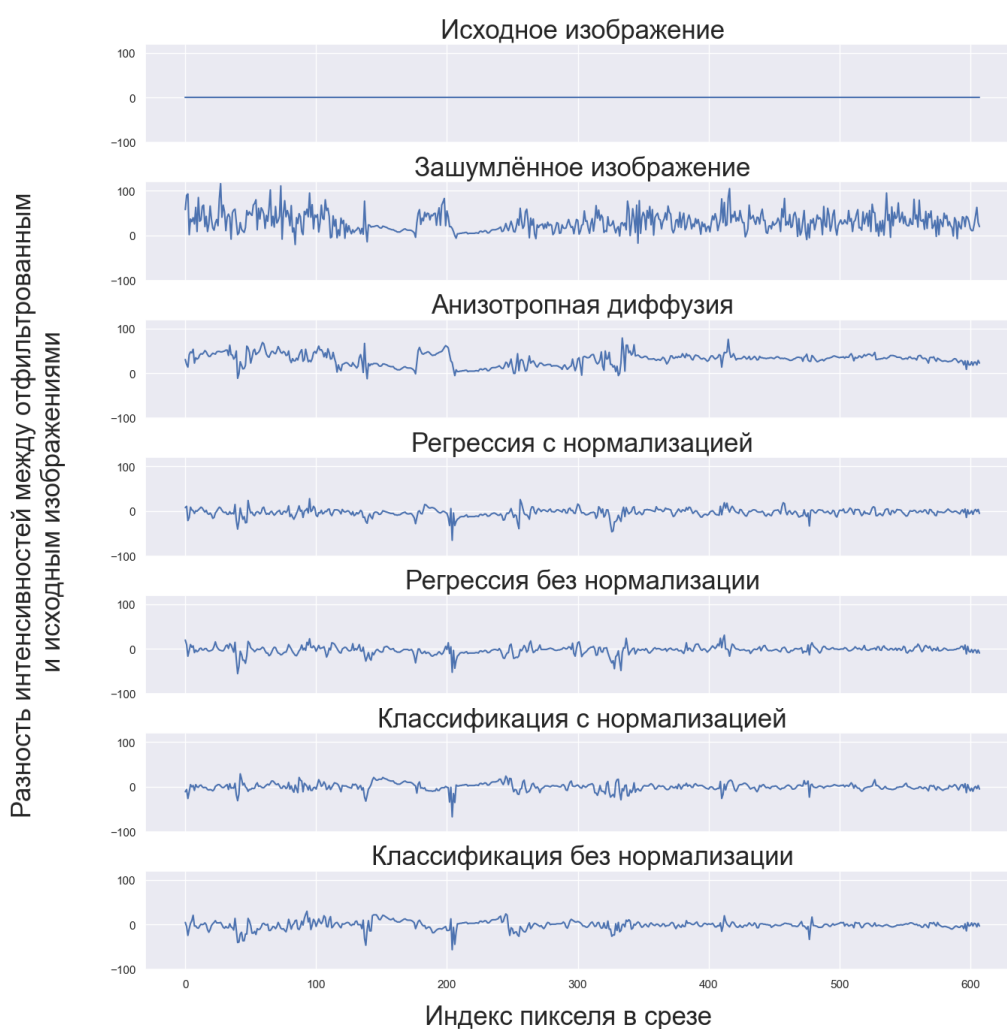


Рисунок 3.7 – разность между интенсивностью пикселей на срезе оригинального изображения и рассматриваемого

В таблице 3.3 приведены значения стандартного отклонения σ и среднего μ для каждой разности среза между обработанным изображением и исходным.

Таблица 3.3 – значения стандартного отклонения и среднего для разности срезов между рассматриваемым обработанным изображением и оригинальным.

Название среза	μ	σ
Исходное изображение	0	0
Зашумлённое изображение	30,06	21,4
Анизотропная диффузия	30,92	14,07
Регрессия с нормализацией	-3,06	8,79
Регрессия без нормализации	-2,39	8,94
Классификация с нормализацией	0,22	8,6
Классификация без нормализации	-0,93	9,37

Исходя из данных в таблице 3.3 видно, что больший эффект на снижение дисперсии оказывает применение именно нейросетевых фильтров. Для данного среза при $x = 350$ наименьшее отклонение достигается путём применения нейросетевого фильтра с решением задачи классификации.

Значение среднего для зашумлённого изображения и отфильтрованного фильтром анизотропной диффузии слабо отличаются. Из этого следует, что классический метод не способен подавлять избыточную яркость, которая возникает из-за спекл-шума. Нейросетевые подходы, в свою очередь, справляются с этой задачей, оставляя лишь незначительные отклонения от нуля.

Для полноценного сравнения в таблице 3.4 представлены значения стандартного отклонения и среднего для разности обработанного изображения и исходного.

Таблица 3.4 значения стандартного отклонения и среднего для разности между рассматриваемым обработанным изображением и оригинальным.

Название среза	μ	σ
Исходное изображение	0	0
Зашумлённое изображение	34,34	21,99
Анизотропная диффузия	34,98	13,14
Регрессия с нормализацией	-1,66	9,64
Регрессия без нормализации	-1,97	9,29
Классификация с нормализацией	-1,7	9,46
Классификация без нормализации	-1,47	9,33

Из полученных данных следует, что среди разработанных фильтров наименьшее отклонение отфильтрованного изображения от исходного получается путём применения нейросетевого фильтра на базе решения задачи регрессии без нормализации к зашумлённому. Зашумлённое РЛИ обладает повышенным уровнем яркости, но фильтр анизотропной диффузии не подавляет её к исходному значению, что сказывается на увеличении среднего значения разности между отфильтрованным и исходным изображениями.

3.4 Выводы

Разработанные алгоритмы показывают более высокую эффективность по сравнению с классическим методом: анизотропной диффузией несмотря на то, что данный подход считается одним из самых новых и качественных.

Результаты сравнения эффективности работы различных методов фильтрации мультипликативного спекл-шума на РЛИ показали, что оптимальным является использование нейросетевого подхода с задачей регрессии без нормализации. Такой способ позволяет максимизировать значение метрики SSIM и минимизировать значение GMSD, а значит данный подход работает лучше остальных среди рассматриваемых и справляется с задачей сохранения структурной составляющей изображения.

ЗАКЛЮЧЕНИЕ

Радиолокационные изображения применяются во многих областях за счёт невосприимчивости к погодным условиям. Но на каждом РЛИ присутствует спекл-шум, который может затруднить обработку изображений или вовсе привести к неверным выводам. Для решения этой задачи разрабатываются различные фильтры, применение которых позволяет избавиться от помех.

В ходе работы были разработаны различные алгоритмы фильтрации мультипликативного спекл-шума при помощи искусственных нейронных сетей с различными подходами и параметрами. Архитектура нейронных сетей включает в себе 5 скрытых слоёв с пакетной нормализацией и функцией активации ReLU. Для обучения ИНС применялся метод обучения с учителем на синтезированном наборе оптических изображений земной поверхности с наложенным специальным образом мультипликативным спекл-шумом, полученным при помощи распределения Рэлея с параметром масштаба 0,27. Генерация обучающего набора данных является вынужденной мерой, так как нет способа создать радиолокационные изображения без спекл-шума. Размер входного вектора ИНС выбирался равным, квадрату ширины «скользящего окна», которое итеративно проходит по всему изображению.

Для проведения оценки работы полученных алгоритмов фильтрации использовались специальные метрики сходства изображений SSIM и GMSD. Они способны учитывать структурные особенности изображений, несущую в себе важную информацию для дальнейшей интерпретации.

Цель работы достигнута, поставленные задачи были выполнены. Результаты сравнения метрик показали, что наиболее успешно с задачей фильтрации мультипликативного спек-шума на радиолокационных изображениях справляется алгоритм на основе нейронной сети с задачей регрессии без предварительной нормализации входных данных. Значения метрик SSIM и GMSD при таком подходе достигают значений 0,877 и 0,056

соответственно, что на 0,05 и 0,022 лучше, чем при применении классического фильтра анизотропной диффузии.

Также были проанализированы срезы изображений для наглядного сравнения работы фильтров. Выявлено, что нейросетевые фильтры способны подавлять избыточную яркость, которая возникает вследствие появления спекл шума, в то время как фильтр анизотропной диффузии данную задачу не решает.

СПИСОК ИСПОЛЬЗОВАННЫХ ИСТОЧНИКОВ

1. J.-S. Lee, "Speckle analysis and smoothing of synthetic aperture radar images," *Computer Graphics and Image Processing*, vol. 17, no. 1, pp. 24 – 32, 1981.
2. V. S. Frost, J. A. Stiles, K. S. Shanmugan, J. C. Holtzman, "A model for radar images and its application to adaptive digital filtering of multiplicative noise," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. PAMI-4, no. 2, pp. 157–166, March 1982.
3. D. Kuan, A. Sawchuk, T. Strand, P. Chavel, "Adaptive restoration of images with speckle," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. 35, no. 3, pp. 373–383, March 1987.
4. F. Argenti, A. Alparone, "Speckle removal from SAR images in the undecimated wavelet domain," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 40, pp. 2363–2374, Nov. 2002.
5. F. Argenti, T. Bianchi, A. Alparone, "Multiresolution MAP despeckling of SAR images based on locally adaptive generalized gaussian pdf modeling," *IEEE Transactions on Image Processing*, vol. 15, no. 11, pp. 3385–3399, Nov. 2006.
6. M. Dai, C. Peng, A. K. Chan, D. Loguinov, "Bayesian wavelet shrinkage with edge detection for SAR image despeckling," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 42, no. 8, pp. 1642– 1648, Aug. 2004.
7. C. Deledalle, L. Denis, F. Tupin, "Iterative weighted maximum likelihood denoising with probabilistic patchbased weights," *IEEE Transactions on Image Processing*, vol. 18, no. 12, pp. 2661–2672, Dec. 2009.
8. Zhong H., Xu J., Jiao L. Classification based nonlocal means despeckling for SAR image //MIPPR 2009: Automatic Target Recognition and Image Analysis. – SPIE, 2009. – T. 7495. – C. 231-238.
9. P. Coupe, P. Hellier, C. Kervrann, C. Barillot, "Bayesian non local means-based speckle filtering," 2008 5th IEEE International Symposium on Biomedical Imaging: From Nano to Macro, Paris, France, 2008, pp. 1291-1294, doi: 10.1109/ISBI.2008.4541240.

10. Antoni Buades, Bartomeu Coll, Jean-Michel Morel. A review of image denoising algorithms, with a new one. *Multiscale Modeling and Simulation: A SIAM Interdisciplinary Journal*, 2005, 4 (2), pp.490-530. fffhal-00271141.
11. G. Chierchia, D. Cozzolino, G. Poggi, L. Verdoliva, “SAR image despeckling through convolutional neural networks,” in *2017 IEEE International Geoscience and Remote Sensing Symposium (IGARSS)*, July 2017, pp. 5438–5441.
12. P. Wang, H. Zhang, V. M. Patel, “SAR Image Despeckling Using a Convolutional Neural Network,” *IEEE Signal Processing Letters*, vol. 24, no. 12, pp. 1763–1767, Dec 2017.
13. Qianqian Zhang, Ruizhi Sun “SAR Image Despeckling Based on Convolutional Denoising Autoencoder” – URL: <https://arxiv.org/abs/2011.14627>
14. Ao, Dongyang, Corneliu Octavian Dumitru, Gottfried Schwarz, Mihai Datcu. 2018. "Dialectical GAN for SAR Image Translation: From Sentinel-1 to TerraSAR-X" *Remote Sensing* 10, no. 10: 1597. – URL: <https://doi.org/10.3390/rs10101597>
15. Vaswani, Ashish, Noam M. Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N. Gomez, Lukasz Kaiser, Illia Polosukhin. “Attention is All you Need.” *NIPS* (2017). – URL: <https://arxiv.org/abs/1706.03762>
16. Radford, Alec, Jong Wook Kim, Tao Xu, Greg Brockman, Christine McLeavey, Ilya Sutskever. “Robust Speech Recognition via Large-Scale Weak Supervision.” – URL: <https://arxiv.org/abs/2212.04356>.
17. Dosovitskiy A., Beyer L., Kolesnikov A., Weissenborn D., Zhai X., Unterthiner T., Houlsby N. (2020). “An image is worth 16x16 words: Transformers for image recognition at scale.” – URL: <https://arxiv.org/abs/2010.11929>
18. Li, H., Yang, Z., Hong, X., Zhao, Z., Chen, J., Shi, Y., Pan, J. (2022). “DnSwin: Toward real-world denoising via a continuous Wavelet Sliding” – URL: <https://arxiv.org/abs/2207.13861>.
19. Belov A. A., Pavlov V. A., Tuzova A. A. A Method of Finding Optimal Parameters of Speckle Noise Reduction Filters. 2020 *Internet of Things, Smart*

Spaces and Next Generation Networks and Systems, Springer Intern. Publishing, 2020, pp. 133–141. doi: 10.1007/978-3-030- 65729-1_12.

20. Ian Goodfellow, Yoshua Bengio, Aaron Courville (2016). Deep Learning. MIT Press – C. 166–270.

21. Kingma, Diederik P., Jimmy Ba. "Adam: A method for stochastic optimization." arXiv preprint arXiv:1412.6980 (2014) – URL: <https://arxiv.org/abs/1412.6980>

22. Wang Z., Bovik A.C., Sheikh H.R., Simoncelli E.P. Image quality assessment: from error visibility to structural similarity. IEEE Trans Image Process. 2004 Apr;13(4):600-12. doi: 10.1109/tip.2003.819861. PMID: 15376593.

23. B. Zhang, P. V. Sander, A. Bermak, "Gradient magnitude similarity deviation on multiple scales for color image quality assessment," 2017 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), New Orleans, LA, USA, 2017, pp. 1253-1257, doi: 10.1109/ICASSP.2017.7952357.

24. Тузова А.А., Павлов В.А., Белов А.А. Подавление мультипликативного шума на радиолокационных изображениях. Известия высших учебных заведений России. Радиоэлектроника. 2021;24(4):6-18. <https://doi.org/10.32603/1993-8985-2021-24-4-6-18>