

На правах рукописи



Сафонова Анастасия Николаевна

**МЕТОДЫ МАШИННОГО ОБУЧЕНИЯ ПРИ ОБРАБОТКЕ
ИЗОБРАЖЕНИЙ СВЕРХВЫСОКОГО ПРОСТРАНСТВЕННОГО
РАЗРЕШЕНИЯ НА ПРИМЕРЕ ЗАДАЧ КЛАССИФИКАЦИИ
РАСТИТЕЛЬНОСТИ**

Специальность 05.13.17 – Теоретические основы информатики

АВТОРЕФЕРАТ
диссертации на соискание ученой степени
кандидата технических наук

Красноярск – 2020

Работа выполнена в федеральном государственном автономном образовательном учреждении высшего образования «Сибирский федеральный университет»

Научный руководитель кандидат технических наук, доцент,
Маглинец Юрий Анатольевич

Официальные оппоненты: **Романов Андрей Николаевич**, доктор технических наук, доцент, Федеральное государственное бюджетное учреждение науки Институт водных и экологических проблем Сибирского отделения Российской академии наук, заместитель директора по научной работе, лаборатория физики атмосферно-гидросферных процессов, заведующий лабораторией.

Евстратова Лариса Геннадьевна, кандидат технических наук, доцент, Федеральное государственное бюджетное образовательное учреждение высшего образования «Государственный университет по землеустройству», кафедра картографии, доцент.

Ведущая организация: Федеральное государственное бюджетное учреждение науки Институт вычислительных технологий Сибирского отделения Российской академии наук, г. Новосибирск.

Защита диссертации состоится «27» февраля 2020 года в 14:00 часов на заседании диссертационного совета Д 212.099.22, созданного на базе Сибирского федерального университета по адресу: 660074, г. Красноярск, ул. Киренского, 26, ауд. УЛК 112.

С диссертацией можно ознакомиться в библиотеке и на сайте Сибирского федерального университета по адресу <http://www.sfu-kras.ru>

Автореферат разослан «___» января 2020 г.

Ученый секретарь
диссертационного совета



Покидышева Людмила Ивановна

ОБЩАЯ ХАРАКТЕРИСТИКА РАБОТЫ

Актуальность работы. Искусственные нейронные сети (ИНС) в настоящее время переживают свое второе рождение, что, в первую очередь, обусловлено увеличением вычислительных мощностей современных компьютеров и появлением сверхбольших наборов данных для обучения, присутствующих в глобальных сетях. На основе ИНС разрабатываются решения в области классификации данных, сегментации изображений дистанционного зондирования Земли (ДЗЗ), поддержки принятия решений, сопоставимые по качеству, а зачастую превышающие результаты, полученные на основе классических методов распознавания образов.

Прикладная область диссертационного исследования связана с решением задач сегментации, классификации и категоризации характера повреждения растительности на основе использования сверточных нейронных сетей (СНС). В последнее время активно ведутся исследования в работах ученых США, Китая, Испании, Бразилии, Швейцарии, Германии, Украины, Швеции, Дании, Турции, России и других стран G.V. Bonan, M.C. Hansen, Z. Ma, W. Li, E. Guirado, S. Tabik, R. Baeta, N. Kussul, L.T. Waser, Z. Deli, M. Långkvist, M. Dyrmann, S. Razavi, С.А. Кривца, И.А. Керчева, Э.М. Бисирова, Д.А. Демидко, Н.В. Пашеновой, Ю.Н. Баранчикова, В.М. Петько, С.А. Астапенко, Е.Н. Акулова, А.Н. Горбань, М.Н. Фаворской, Л.Ф. Ноженковой и др., позволяющие классифицировать растительность на основе применения таких СНС, как Inception, ResNet, DenseNet, RCNN, VGG, а также классических методов, включая ОБІА.

Однако, зачастую при решении новых классов задач внутри рассматриваемой прикладной области возникают проблемы, связанные с малыми объемами выборки и недостаточным качеством исходного материала для обучения, что определяет актуальность исследования и разработки новых методов и алгоритмов классификации категорий повреждения растительности на основе машинного обучения, а именно СНС, позволяющих обеспечить лучшее качество распознавания в условиях малых выборок и ограниченного количества спектральных каналов.

Целью работы является разработка методов машинного обучения при обработке изображений сверхвысокого пространственного разрешения в условиях малых выборок по искусственно увеличенным данным на примере задач сегментации, классификации и категоризации характера повреждения растительности.

Решаемые задачи:

1. Анализ современных алгоритмов СНС и их использования при решении задач классификации объектов, представленных на мультиспектральных и трехканальных изображениях сверхвысокого пространственного разрешения.
2. Разработка алгоритма построения и искусственного увеличения размеченного набора данных на изображениях сверхвысокого пространственного разрешения в условиях малых выборок.
3. Разработка новой архитектуры СНС для задач точной классификации размеченного набора данных на трехканальных цветных изображениях сверхвысокого пространственного разрешения.
4. Разработка новой методики формирования контрольной выборки на трехканальных цветных изображениях сверхвысокого пространственного разрешения в условиях малых выборок для проведения тестирования разработанной архитектуры СНС и ее сравнения с современными моделями СНС.
5. Модифицирование алгоритма сегментации изображений сверхвысокого пространственного разрешения в условиях малых выборок на основе СНС ResNet50 и ResNet101 для построения маски регионов объектов.
6. Выполнение вычислений основных параметров метрики оценки разработанной архитектуры СНС с использованием различных наборов данных изображений сверхвысокого пространственного разрешения для оценки результатов точности.

Научная новизна.

1. Разработана новая архитектура СНС, обученная в условиях малых выборок на искусственно увеличенном наборе данных ДЗЗ, отличающаяся структурой и составом слоев нейронной сети, их порядком, позволяющая автоматически классифицировать размеченные категории повреждения растительности на трехканальных цветных изображениях сверхвысокого пространственного разрешения.
2. Разработана новая методика формирования контрольной выборки, отличающаяся более высокой степенью автоматизации, позволяющая локализовать категории повреждения растительности на трехканальных цветных изображениях сверхвысокого пространственного разрешения в условиях малых выборок, с целью ускорения процесса подготовки тестируемого набора данных.
3. Получена модификация алгоритма сегментации растительности на основе СНС ResNet50 и ResNet101, отличающаяся уникальным подбором параметров, включением искусственного увеличения размеченного набора данных, позволяющая строить маски регионов объектов на изображениях сверхвысокого пространственного разрешения в условиях малых выборок.

Теоретическая значимость работы обоснована тем, что применительно к проблематике диссертации результативно использован комплекс существующих методов построения СНС; произведена модернизация архитектуры СНС и алгоритма сегментации на основе СНС ResNet50 и ResNet101. Предложенные в работе модели и алгоритмы могут быть использованы для развития методов построения и использования СНС в условиях малых выборок.

Практическая значимость работы. Разработанные методы машинного обучения позволяют улучшить численные показатели при обработке изображений сверхвысокого пространственного разрешения на примере задач сегментации, классификации и категоризации характера повреждения растительности, а также могут быть использованы в системе мониторинга лесного и сельского хозяйства с применением беспилотных летательных аппаратов (БПЛА).

Методы исследования. Для решения поставленных в работе задач использовались методы предварительной обработки данных ДЗЗ, методы искусственного увеличения объема обучающей выборки, методы машинного обучения, ГИС-технологии, методы статистического анализа.

Достоверность полученных результатов диссертационной работы подтверждена экспериментальными исследованиями, проведенными автором на основе аэросъемки с БПЛА и сопоставлением полученных данных с результатами полевых (наземных) измерений, а также использованием существующих, теоретически обоснованных и проверенных на практике методов (методов математической статистики).

Основные положения, выносимые на защиту.

1. Новая архитектура СНС для задач точной классификации размеченного набора данных на трехканальных цветных изображениях сверхвысокого пространственного разрешения в условиях малых выборок.

2. Новая методика формирования контрольной выборки на трехканальных цветных изображениях сверхвысокого пространственного разрешения в условиях малых выборок для проведения тестирования разработанной архитектуры СНС и ее сравнения с современными моделями СНС.

3. Модифицированный алгоритм сегментации изображений сверхвысокого пространственного разрешения в условиях малых выборок на основе СНС ResNet50 и ResNet101 для построения маски регионов объектов.

Апробация работы. Основные положения и результаты диссертационной работы докладывались и обсуждались на IV, V, VI международных научных

конференциях «Региональные проблемы дистанционного зондирования Земли» (Красноярск, 2017, 2018, 2019 гг.), II Congreso Nacional / IV Jornadas de Investigadores en Formación: Fomentando la Interdisciplinariedad (Университет Гранады, г. Гранада, Испания, 2019 г.), всероссийских конференциях «Обработка пространственных данных в задачах мониторинга природных и антропогенных процессов» (г. Бердск, Новосибирская область, 2017, 2019 гг.).

Внедрение результатов работы. Представленные в работе новая архитектура СНС, методика формирования контрольной выборки и модифицированный алгоритм сегментации реализованы программно, обеспечены методически и встроены в курс практических занятий по дисциплине «Интеллектуальные технологии и представление знаний» в раздел «Модели нейронных сетей» для студентов специальности 27.03.03 «Системный анализ и управление».

Результаты диссертационной работы были использованы при реализации следующих контрактов:

- Грант РФФИ №16-11-00007 в 2016-2018 гг., «Автоматизация распознавания природно-антропогенных объектов на аэрокосмических изображениях высокого спектрального и пространственного разрешения» для проведения фундаментальных научных исследований и поисковых научных исследований отдельными научными группами;
- Грант Европейского союза Erasmus+ Student Credit Mobility в Университете Гранады, Гранада, Испания в 2016-2017 гг., «Методы детектирования категорий повреждения растительных сообществ на основе машинного обучения по данным съемки беспилотного летательного аппарата»;
- Грант Европейского союза Erasmus+ Student Credit Mobility в Университете Гранады, Гранада, Испания в 2017-2018 гг., «Методы детектирования категорий повреждения растительных сообществ на основе машинного обучения по данным съемки беспилотного летательного аппарата»;
- Грант по проекту 5-100 для проведения научно-исследовательской деятельности в Университете Гранады, Гранада, Испания в период с 19.10.2019 по 04.11.2019 г.

Соответствие диссертации паспорту специальности.

Диссертационное исследование выполнено в соответствии с паспортом специальности 05.13.17 – «Теоретические основы информатики» в части:

- пункта 5 «Разработка и исследование моделей и алгоритмов анализа данных, обнаружения закономерностей в данных и их извлечение, разработка и исследование методов и алгоритмов анализа текста, устной речи и изображений»;

- пункта 7 «Разработка методов распознавания образов, фильтрации, распознавания и синтеза изображений, решающих правил. Моделирование формирования эмпирического знания».

Публикации. По результатам исследования опубликовано 13 работ, из них в рецензируемых научных изданиях по перечню ВАК – 2 [1,2], в изданиях, индексированных в Scopus и Web of Science – 3 [3-5].

Структура и объем диссертации. Диссертационная работа состоит из введения, четырех глав, заключения, списка литературы и приложений. Основной текст работы включает 104 страницы. Изложение иллюстрируется 35 рисунками и 10 таблицами (1 таблица в приложении Б). Список литературы содержит 111 наименований.

ОСНОВНОЕ СОДЕРЖАНИЕ РАБОТЫ

Во введении обосновывается актуальность темы диссертационной работы, формулируется цель, задачи исследования, описываются методы исследования, научная новизна, практическая значимость полученных результатов, результаты, выносимые на защиту и краткое содержание работы по главам.

В первой главе представлены результаты анализа развития СНС в задачах классификации объектов на изображениях. Проведен обзор существующих в мире работ в области обработки изображений ДЗЗ с применением современных СНС (VGG, ResNet, Inception, InceptionResNet, Xception и DenseNet). Сформулированы основные проблемы, которые необходимо решить в настоящем исследовании. В первом подразделе рассмотрены методы функционирования современных СНС, способы формирования структуры СНС, влияние настраиваемых параметров сети на каждый составляющий ее элемент. Во втором подразделе представлен аналитический обзор работ с использованием современных СНС при решении задач классификации объектов, представленных на мультиспектральных и трехканальных изображениях сверхвысокого пространственного разрешения. Рассмотрены преимущества и недостатки современных СНС. В третьем подразделе описаны основные проблемы и задачи, которые необходимо решить в настоящем исследовании. В заключении представлены выводы по главе.

В результате проведенного аналитического обзора литературы и рассмотрения современных СНС можно заключить, что СНС являются мощным инструментом, позволяющим эффективно распознавать и классифицировать объекты на изображениях, при этом по точности и скорости работы СНС значительно превосходят классические методы машинного обучения. СНС

широко используются в различных областях, где требуется анализ объектов на изображениях. В частности, в последние годы СНС стали применяться и в задачах обработки изображений ДЗЗ.

В тоже время, обзор показал, что описанные в открытых источниках СНС и методики формирования контрольной выборки на изображениях ДЗЗ зачастую обладают избыточной вычислительной сложностью и требуют адаптации для решения рассматриваемого класса задач, выдвигаемых в диссертационном исследовании. Поэтому целесообразна разработка и модификация алгоритмического и методического обеспечения СНС, направленная на сокращение вычислительных и временных ресурсов при обработке данных. Другим направлением для развития указанных методов является повышение точности сегментации и классификации анализируемых данных.

Вторая глава посвящена разработке алгоритмов и архитектуры СНС при решении задач сегментации, классификации и категоризации характера повреждения растительности в условиях малых выборок по искусственно увеличенным данным ДЗЗ со сверхвысоким пространственным разрешением.

Первый подраздел посвящен описанию алгоритма построения и искусственного увеличения размеченного набора данных на изображениях сверхвысокого пространственного разрешения в условиях малых выборок. Алгоритм, путем применения искусственных геометрических преобразований к исходным объектам, обеспечивает создание новых обучающих примеров, что, в условиях малых выборок, позволяет более качественно обучить нейронную сеть и уменьшить переобучение. В диссертационной работе использовались следующие методы искусственного увеличения набора данных: усиление яркости зеленого цвета, применение фильтра размытия по Гауссу, усреднение пикселей изображения с помощью нормализованного пиксельного фильтра, поворот изображения относительно его центра с разными углами поворота, обрезка изображения и уменьшение размерности в два раза, изменение размера центральной части изображения до размеров 150×200 пикселей согласно требованиям СНС с использованием кубической интерполяции. Реализация алгоритма построения и искусственного увеличения набора данных изображений осуществлялась средствами языка программирования Python с использованием библиотек NumPy, SciPy и др.

Второй подраздел главы описывает разработку новой архитектуры СНС для задач точной классификации размеченного набора данных на трехканальных цветных изображениях сверхвысокого пространственного разрешения в условиях малых выборок. Общая архитектура СНС состоит из шести сверточных блоков (каждый включает в себя один сверточный слой). Первый и третий сверточные

блоки включают в себя слои пулинга (функция уменьшения размерности пространства карт признаков) с функцией максимума. В конце СНС находятся два полносвязных слоя и один выходной слой. В последних четырех сверточных блоках используется функция активации ReLU (1), а в выходном слое - функция активации Softmax (2):

$$f(x) = \max(0, x), \quad (1)$$

$$z = \omega^T x - \theta, \quad (2)$$

где x – вектор-столбец признаков объекта размерности $M \times 1$, ω^T – транспонированная матрица весовых коэффициентов признаков, имеющая размерность $K \times M$, где K – количество классов объектов, а M – количество признаков объектов.

Чтобы контролировать переобучение сети, по результатам проведенных экспериментов было решено использовать метод регуляризации для ИНС Dropout, с помощью которого было достигнуто уменьшение сложности модели, с сохранением количества ее параметров на низком уровне. Также было выбрано оптимальное значение коэффициента регуляризации. По результатам проведенных численных экспериментов было определено, что оптимальными являются значения коэффициента 0,25 после второго, четвертого и пятого слоев, и 0,5 перед выходным слоем. В качестве целевой функции, которая должна быть минимизирована в ходе обучения нейронной сети, была выбрана категориальная кросс-энтропийная потеря между входными данными и фактической классификацией изображений. Данная функция хорошо подходит для вычисления вероятности принадлежности исходного изображения к определенной категории. К настоящему времени предложено достаточно большое количество алгоритмов оптимизации при расчете градиентного спуска в нейронных сетях. В предлагаемой модели использовалась оптимизация ADAM (adaptive moment estimation, адаптивная оценка момента), поскольку для рассматриваемой проблемы это наиболее подходящий оптимизатор, в частности, благодаря возможности начальной калибровки ИНС (Рисунок 1).

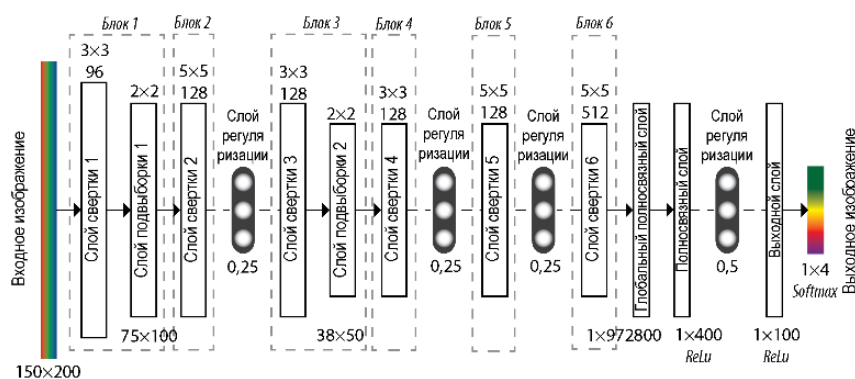


Рисунок 1 – Новая архитектура СНС

Структуру новой СНС можно формально представить в следующем виде:

- Входное изображение учитывает двумерную топологию и состоит из нескольких матриц, где каждая матрица соответствует изображению конкретного цветового канала: красный, синий и зеленый. Входная информация каждого конкретного значения пикселя нормализуются в диапазоне от 0 до 1 по формуле (3):

$$f(p, \min, \max) = \frac{p - \min}{\max - \min}, \quad (3)$$

где f – функция нормализации, p – значение конкретного пикселя от 0 до 255, \min – минимальное значение пикселя – 0, \max – максимальное значение пикселя – 255.

- Во время обучения СНС значения каждой матрицы сверточного слоя равны 0, а значения весов ядер задаются случайным образом в диапазоне от -0,5 до 0,5. Ядро проходит попиксельно по предыдущей матрице и производит операцию свертки по следующей формуле (4):

$$(f * g)[m, n] = \sum_{k, l} f[m - k, n - l] * g[k, l], \quad (4)$$

где f – исходная матрица изображения, g – ядро свертки, m – ширина исходной матрицы изображения, n – высота исходной матрицы изображения, k – ширина ядра, l – высота ядра.

- В зависимости от метода обработки краев исходной матрицы изображения на этапе свертки, результат может быть меньше исходной матрицы, такого же размера или большего размера, что можно представить в виде формулы (5):

$$x^l = f(x^{l-1} * k^l + b^l), \quad (5)$$

где x^l – выход слоя l , $f(\)$ – функция активации, b^l – коэффициент сдвига слоя l , $*$ – операция свертки входа x с ядром k .

- Таким образом, в результате краевых эффектов, исходная матрица на выходе уменьшается и имеет следующий вид (6):

$$x_j^l = f\left(\sum_i x_i^{l-1} * k_j^l + b_j^l\right), \quad (6)$$

где x_j^l – карта признаков j (выход слоя l), k_j^l – ядро свертки j карты, слоя l .

- Слой подвыборки разработанной СНС можно описать следующей формулой (7):

$$x^l = f(a^l * \text{subsample}(x^{l-1}) + b^l), \quad (7)$$

где a^l , b^l – коэффициент сдвига слоя l , $\text{subsample}()$ – операция выборки локальных максимальных значений.

- Последний тип слоя СНС – полносвязный слой или обычный многослойный персептрон. Целью слоя является – классификация, где

происходит моделирование сложной нелинейной функции, при оптимизации которой, улучшается качество распознавания (8):

$$x_j^l = f\left(\sum_i x_i^{l-1} w_{i,j}^{l-1} + b_j^{l-1}\right), \quad (8)$$

где $w_{i,j}^{l-1}$ – матрица весовых коэффициентов слоя l .

Реализация новой архитектуры СНС производилась средствами языка программирования Python с использованием библиотек TensorFlow и Keras.

В третьем подразделе описана новая методика формирования контрольной выборки на трехканальных цветных изображениях сверхвысокого пространственного разрешения с целью ускорения процесса подготовки тестируемого набора данных. Структура предлагаемой методики включает в себя совокупность последовательных элементов:

- Преобразование трехканального цветного изображения в цветовую палитру серых оттенков;
- Размытие изображения в градациях серого с использованием высокочастотного Гауссова фильтра для уменьшения шума на изображениях со следующими параметрами функции: размер ядра 11×11 пикселей со стандартным отклонением, равным 0;
- Создание двоичного изображения (чёрно-белое) из размытых изображений в градациях серого путем применения пороговой функции с оптимальным пороговым значением яркости пикселей входного изображения, равным 100;
- Структурирование контуров элементов изображения путем применения двух последовательных функций (эрозия и расширение) с несколькими итерациями для двоичных изображений с целью различения контуров отдельных крон деревьев и минимизации эффекта их слияния в одном объекте;
- Обнаружение участков изображения, которое было реализовано с использованием функции вычисления площади контура на основе формулы Грина (9):

$$S = \iint_R dx dy = \frac{1}{2} \oint_C x dy - y dx, \quad (9)$$

где S – это площадь области R , ограниченной контуром C . Символ \oint_C – указывает на то, что контур потенциального объекта является замкнутым, и обход при интегрировании вдоль этого контура производится против часовой стрелки. Размер объекта для функции был установлен в диапазоне от 50×50 до 200×200 пикселей в соответствии с оригинальными размерами изображений обучающей выборки.

Использование предложенной методики формирования контрольной выборки на изображениях сверхвысокого пространственного разрешения позволяет производить подготовку тестируемого набора данных за относительно более короткое время, нежели в аналогах в режиме, близком к автоматическому. При этом достигается существенное ускорение процесса подготовки тестируемого набора данных. Конкретные временные характеристики варьируются в зависимости от размера входного изображения и конфигурации компьютера. В условиях проведенных экспериментов, формирование контрольной выборки занимало в среднем около 5 минут. Реализация методики формирования контрольной выборки была осуществлена средствами языка программирования Python с использованием библиотек OpenCV, Pillow и др.

В четвертом подразделе описан модифицированный алгоритм сегментации изображений сверхвысокого пространственного разрешения в условиях малых выборок на основе СНС ResNet50 и ResNet101 для построения маски регионов объектов. Алгоритм был модифицирован за счет использования искусственно увеличенного набора данных и тонкой настройки алгоритма для решения поставленной задачи. В данной СНС реализован двухэтапный подход. На первом этапе сканируется входное изображение и генерируются предложения (области, которые могут содержать объект). На втором этапе выполняется классифицирование предложения и формирование ограничивающих рамок и масок.

Mask R-CNN состоит из следующих структурных компонентов:

1. Основа – стандартная СНС ResNet101, на ранних слоях которой обнаруживаются объекты низкого уровня (края и углы), а на более поздних слоях сети обнаруживаются элементы более высокого уровня (дерево, человек, здание). Проходя через нейронную сеть, изображение преобразуется из матрицы вида $1024 \times 1024 \times 3$ (RGB) в карту объектов формы $32 \times 32 \times 2048$. Эта карта характеристик становится входом для следующих уровней сети. Для улучшения качества извлечения объектов сетью использовался метод Feature Pyramid Network, который берет обнаруженные СНС элементы высокого уровня и передает их в нижние слои. Это позволяет функциям на каждом уровне иметь доступ к функциям как нижнего, так и более высокого уровня.

2. Region Proposal Network – СНС, которая сканирует изображение в режиме скользящего окна и находит области, содержащие объекты. В результате сеть генерирует: класс привязки (передний план и фон, где в первом подразумевается присутствие классифицируемого объекта) и ограничительную рамку с оценкой дельты (процент относительно точности измерения координат, ширины, высоты) для уточнения поля привязки для лучшего соответствия объекту.

3. Классификатор и ограничивающее окно-регрессор. Данный компонент генерирует два выхода для каждого объекта: класс объекта и ограничительная рамка. Данная рамка необходима для дальнейшего уточнения расположения и размера объекта.

Так как классификаторы не всегда справляются с определением размеров объектов, для них обычно требуется объекты фиксированного размера. Для этого использовался дополнительный слой подвыборки (субдискретизации), который позволяет обрезать часть карты объекта до фиксированного размера.

4. Сегментные маски – это СНС, которая генерирует маски выделенных классификатором объектов с предыдущего шага. Сгенерированные маски имеют низкое разрешение: 28×28 пикселей. Во время обучения используются уменьшенные маски объектов до 28×28 пикселей, чтобы вычислить потери, а во время вывода предсказанные маски увеличиваются до размеров ограничивающей рамки объекта, что и является окончательными масками, по одной на объект.

Пятый подраздел главы представляет собой описание метрики оценки эффективности разработанной архитектуры СНС, с такими рассчитываемыми характеристиками, как: доля правильных ответов (Accuracy) (10), точность (Precision) (11), полнота (Recall) (12) и F-мера (F_score) (13):

$$Accuracy = \frac{TP+TN}{TP+TN+FN+FP}, \quad (10)$$

где количество истинно положительных предсказаний (True Positive, TP), истинно отрицательных предсказаний (True Negative, TN), ложноположительных предсказаний (False Positive, FP) и ложноотрицательных предсказаний (False Negative, FN),

$$Precision = \frac{TP}{TP+FP}, \quad (11)$$

$$Recall = \frac{TP}{TP+FN}, \quad (12)$$

$$F_score = 2 \times \frac{Precision \times Recall}{Precision + Recall}. \quad (13)$$

В третьей главе представлены результаты экспериментального применения разработанной архитектуры СНС. Глава включает в себя четыре подраздела.

В первом подразделе описан тестовый участок и материалы ДЗЗ, на которых проводятся эксперименты настоящего диссертационного исследования. Представлен процесс подготовки обучающего набора данных и его искусственное увеличение (дополнение). Проведена подготовка тестового набора данных для проведения независимой проверки новой архитектуры СНС.

Определение категории состояния здоровья сибирских пихт (лат. *Abies sibirica*) проводилось на основе энтомологического подхода, предложенного Кривцом. В данном подходе состояние пихт описывается с помощью шести категорий, согласно уровню проникновения вредителей *P. proximus* в ствол и их влиянию на полог: I - здоровые деревья; II - ослабленные деревья; III - сильно ослабленные деревья; IV - умирающие деревья; V - недавно умершие деревья; VI - старый мертвый лес. Следует отметить, что различия между первой, второй и третьей категориями проявляются только на стволе деревьев, но они не видны в его кроне. Поэтому в данном исследовании классификация выполнялась с помощью четырех категорий (пропуская вторую и третью). Таким образом, объектом исследования являются следующие категории повреждения дерева: *категория 1* - полностью здоровое дерево или недавно подвергнутое нападению жуков, *категория 2* - дерево, колонизированное жуками, *категория 3* - недавно умершее дерево и *категория 4* – сухостой.

Для исследовательских площадок был получен набор изображений RGB со сверхвысоким пространственным разрешением ($\approx 5-10$ см/пиксел) во время нескольких полетов квадрокоптера DJI Phantom 3 Pro (со стандартной камерой) в июле 2016 года (графики А и В) и гексакоптера Yuneec Typhoon H (с камерой CGO3 +) в мае 2016 года (участок Б) и августе 2018 года (участок Г). Изображения для исследовательских участков А и В были получены в облачных погодных условиях на высотах над уровнем моря 670 м (А) и 700 м (В) (высота относительно уровня земли 120-150 м), а участки Б и Г были обследованы в солнечную погоду на высоте 120 м. При получении всех аэрофотоснимков использовались настройки камеры по умолчанию (автоматический баланс белого, ISO 100). Для разметки обучающих данных было подготовлено 50 изображений (патчей) для каждой из четырех категорий повреждения дерева. При этом из подготовленных данных было создано два набора данных. Первый набор – это данные для обучения, валидации и тестирования. Он состоял из 50 выбранных вручную фрагментов изображений отдельных деревьев для каждой категории повреждения дерева, в результате чего было получено 200 фрагментов изображений. Для проведения качественного обучения модели СНС использовалось 80% изображений, а оставшиеся 20% изображений были использованы для внутренней проверки модели (тестирования). Вторым набор – это данные для внешнего тестирования или внешней проверки. Этот набор состоял из 88 патчей, сгенерированных разработанной методикой формирования тестируемого набора данных. Для повышения надежности и точности классификации с помощью разработанной архитектуры СНС, было выполнено дополнение объема исходной выборки с использованием алгоритма построения и

искусственного увеличения размеченного набора данных с 160 до 3520 изображений для обучения сети, с 40 до 880 – для независимой ее валидации.

Подготовка второго набора тестовых изображений для проведения независимой проверки разработанной архитектуры СНС выполнялась по новой методике формирования патчей на новых тестовых участках В и Г. Результатом выполнения этого этапа являются области-кандидаты, обозначенные рамками красного цвета, представленные на рисунке 2.

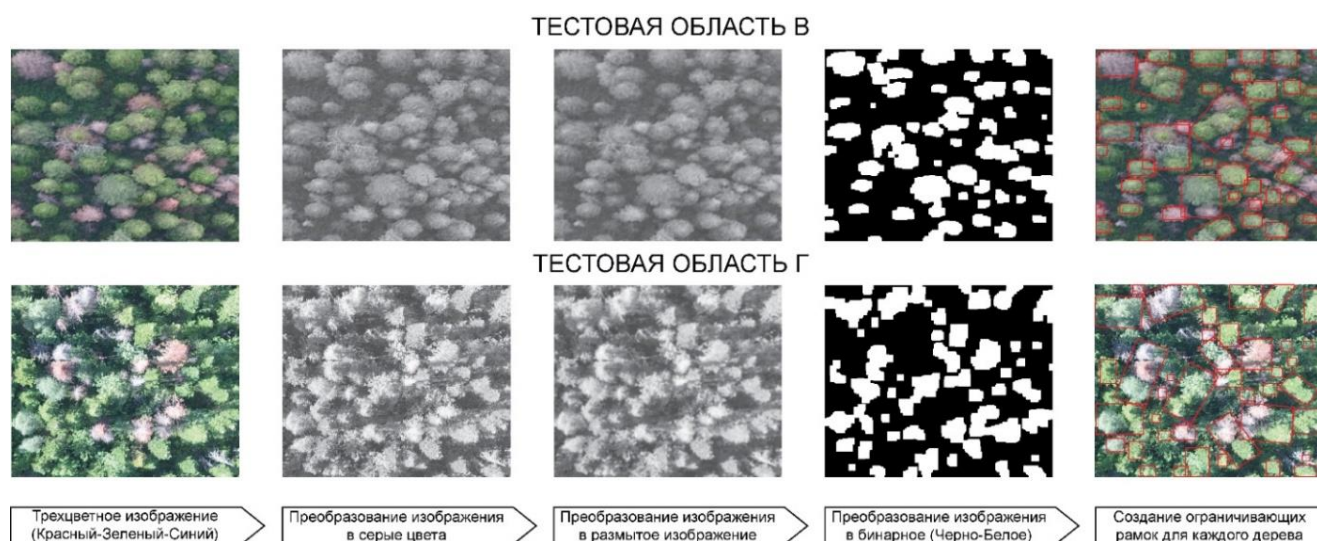


Рисунок 2 – Методика формирования патчей на тестовых областях В и Г

Предварительная обработка состояла из следующих этапов: (1) преобразование трехцветного изображения в изображение в серой палитре (PAN); (2) преобразование изображения в градациях серого в размытое изображение; (3) преобразование размытого изображения в двоичное изображение на основе порога цифрового значения; (4) обнаружение деревьев на изображениях. На крайних справа изображениях красными прямоугольниками отмечены 48 потенциальных (с высокой вероятностью содержащих дерево) патчей, идентифицированных в тестовой области В, и 40 потенциальных патчей, идентифицированных в тестовой области Г.

Второй подраздел содержит описание полученных результатов обучения новой архитектуры СНС и их сравнение с результатами, полученными с использованием типопредставителей современных моделей нейронных сетей. Максимальная производительность новой архитектуры СНС с расширенным набором данных была достигнута на 23-й эпохе обучения, обеспечивая точность внутреннего теста 99,7% и минимальные потери (мера количества ошибок) при обучении ниже 0,01. После 23-й эпохи потери обучения модели были стабилизированы, а разница между потерями при обучении и проверки увеличилась. После обучения созданной модели СНС было проведено её

сравнение с известными моделями нейронных сетей на том же наборе данных. В сравнении участвовали следующие известные и хорошо зарекомендовавшие себя модели: Xception, VGG-16, VGG-19, ResNet-50, Inception-V3, InceptionResNet-V2, DenseNet-121, DenseNet -169 и DenseNet-201. Выходные слои всех участвовавших в сравнении моделей СНС были адаптированы к нашей задаче (четыре категории повреждения пихт).

В третьем подразделе представлена оценка результатов тестирования новой архитектуры СНС. Результаты расчета стандартных показателей оценки, основанные на матрице ошибок классификации, рассчитанной по каждой категории повреждения дерева с увеличением и без увеличения данных, представлены в таблице 1.

Таблица 1 – Производительность новой архитектуры СНС с увеличением и без увеличения на тестовом наборе для каждой категории повреждения дерева. Производительность характеризуется следующими показателями: истинно положительных результатов (TP), истинно отрицательных (TN), ложно положительных (FP), ложно отрицательных (FN), доли правильных ответов (accuracy), точности (precision), полноты (recall) и F-меры (F_score)

Категории повреждения пихт	TP	TN	FP	FN	Accuracy, %	Precision, %	Recall, %	F_Score, %
	Число объектов							
Без увеличения набора данных изображений								
1	18	44	3	14	78,48	85,71	56,25	67,92
2	30	32	15	6	74,7	66,67	83,33	74,07
3	10	52	0	6	91,18	100	65,5	76,92
4	4	58	8	0	88,57	33,33	100	50
С увеличением набора данных изображений								
1	32	48	5	0	94,12	86,49	100	92,75
2	31	49	2	5	91,95	93,94	86,11	89,86
3	13	67	0	3	96,39	100	81,25	89,66
4	4	76	1	0	98,77	80	100	88,89

Как видно из таблицы 1, обучение новой архитектуры СНС на увеличенном наборе данных изображений обеспечивает лучшую точность и F-меру, по сравнению с результатами, полученными при обучении СНС на исходном наборе данных (без увеличения). Доля правильных ответов, точность, и F-мера улучшились на 14,5%, 20,4%, и 34,3% соответственно. Также в диссертационной работе представлено сравнение результатов классификации разработанной модели с результатами работы более сложных моделей. Несмотря на то, что Xception, VGG16, VGG19, ResNet50, InceptionV3, InceptionResNetV2, DenseNet121, DenseNet169 и DenseNet201 являются мощными и вычислительно интенсивными моделями, в данной задаче распознавания категорий повреждения пихт они все же уступают разработанной в настоящем исследовании архитектуре СНС. Модель VGG-16 показала результаты, в среднем наиболее высокие среди других альтернативных моделей, рассмотренных в эксперименте. Она распознала

первую, вторую и четвертую категории с точностью 85,9%, 79,76 и 94,37% соответственно. Распознавание третьей категории с точностью выше 88,89% достигнуто моделью ResNet-50. Наименьшую точность распознавания продемонстрировала сеть InceptionResNetV2. Разработанная в этом исследовании архитектура СНС обеспечивает гораздо лучшие результаты при решении задач классификации категорий повреждения дерева (Пихта Сибирская). Предлагаемая модель, обученная на увеличенном наборе изображений, достигает высокого значения F-меры для категорий повреждения пихт 1, 2, 3, и 4 с 92,75%, 89,86%, 89,66% и 88,89% соответственно. Это можно объяснить тем фактом, что модель правильно различает цвет, форму и текстуру каждой из четырех категорий повреждения дерева. Результаты классификации категорий повреждения дерева новой архитектурой СНС на изображениях испытательных участков В и Г представлены на рисунках 3 и 4.

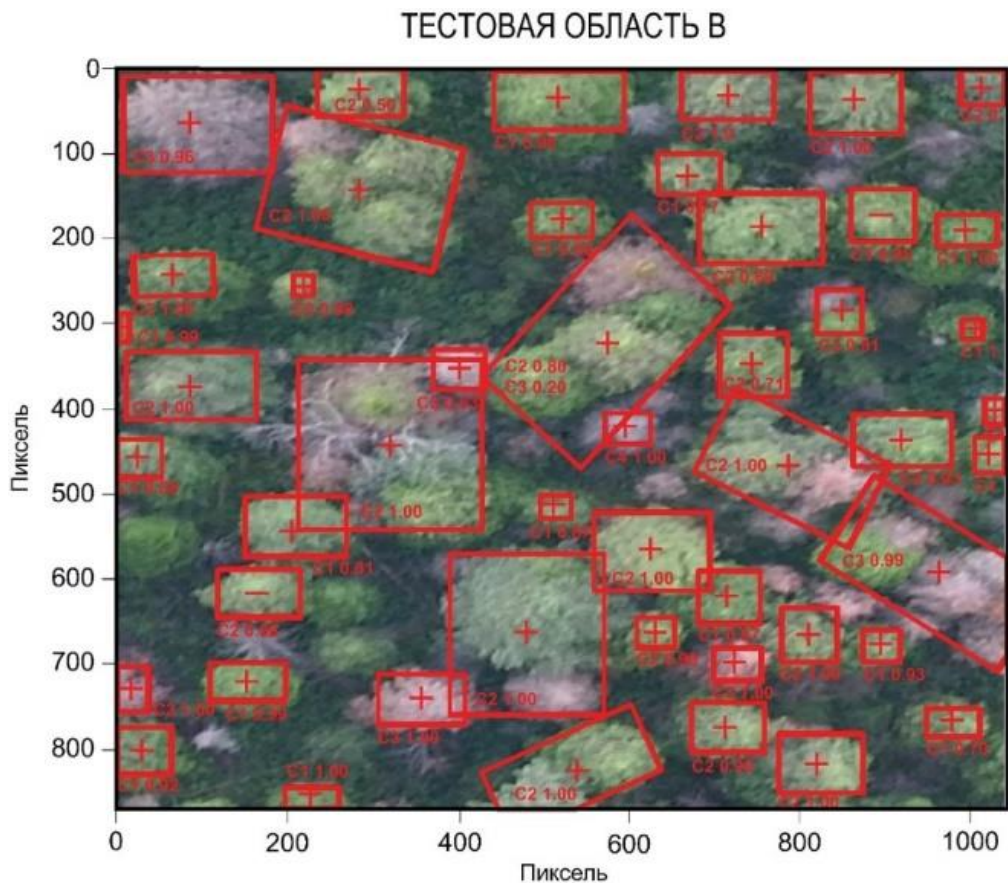


Рисунок 3 – Результаты классификации категорий повреждения дерева в тестовой области В.

C1, C2, C3 и C4 указывают класс, предсказываемый новой архитектурой СНС, вместе с соответствующей вероятностью. Символы «+» и «-» указывают правильную и неправильную оценку класса

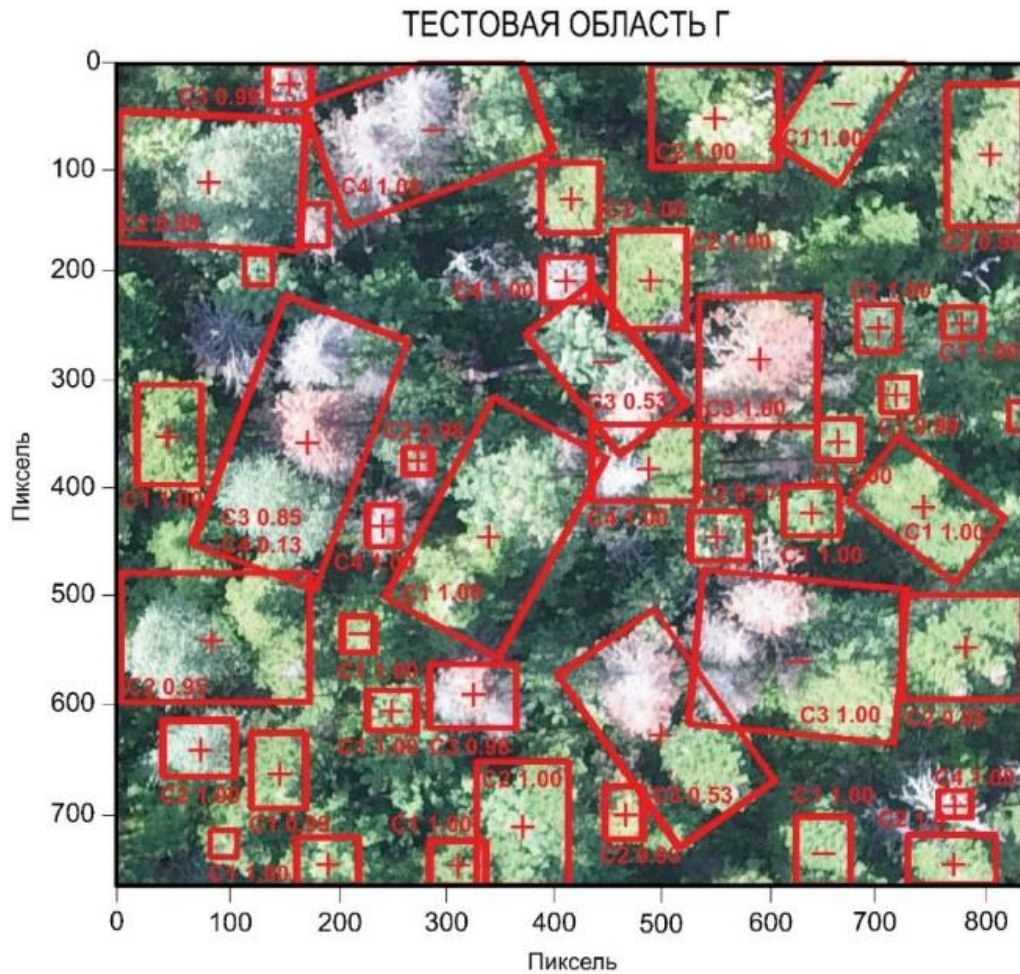


Рисунок 4 – Результаты классификации категорий повреждения дерева (Пихта Сибирская) в тестовой области Г. Обозначения те же, что и на рисунке 3

Красные прямоугольники отмечают выбранные области, полученные с помощью разработанного алгоритма формирования контрольной выборки категорий повреждения растительности (Пихта Сибирская) на изображениях сверхвысокого пространственного разрешения (подраздел 2.3). При этом рядом с каждым прямоугольником имеется два числа – первое показывает категорию повреждения дерева, предсказанную СНС (C1, C2, C3 и C4 соответствуют категориям повреждения дерева 1, 2, 3 и 4), а второе число показывает вероятность принадлежности объекта к соответствующей категории, которая рассчитана классификатором СНС. Отметим, что эта вероятность указывает на уровень достоверности модели.

Исходный код разработанной модели находится в свободном доступе в среде Internet по адресу: <https://github.com/ansaf0/OurCNN>.

В четвертой главе представлены результаты экспериментальных исследований использования модифицированного алгоритма на примере сегментации растительности (Оливковые деревья, лат. *Olea europaea*) на основе СНС ResNet50 и ResNet101 по данным изображений сверхвысокого

пространственного разрешения в условия малых выборок для построения маски регионов объектов. Глава включает в себя три подраздела. В первом подразделе описан тестовый участок и материалы ДЗЗ, на которых проводятся эксперименты настоящего диссертационного исследования. Описан процесс формирования набора данных для обучения модифицированного алгоритма сегментации на основе СНС ResNet50 и ResNet101.

Для проведения эксперимента использовались материалы съемки БПЛА. В исследовании использовались два БПЛА: Parrot Disco-Pro AG с многоспектральным датчиком Parrot Sequoia с четырьмя каналами (зеленый (Green), красный (Red), крайний красный (Red_2), ближний инфракрасный (NIR)) и DL-Phantom 4 Pro с RGB камерой. Четырехканальные изображения сверхвысокого разрешения были получены в феврале 2019 года с разрешением 13 см/пиксель, а цветное изображение было получено в июне 2019 года с разрешением 3 см/пиксель. В обоих случаях полеты БПЛА производились на высоте 120 метров. Эксперименты проводились по следующим комбинациям каналов: RGB изображение, Green-Red-NIR, ВИ NDVI и ВИ GNDVI. На выбранных комбинациях далее велась подготовка набора данных для обучения и тестирования Mask R-CNN. Было подготовлено 150 изображений (патчей) из которых 140 было предназначено для обучения модели и 10 для валидации и тестирования обученной модели. Далее тестовый набор данных был искусственно увеличен до 1500 изображений для каждой комбинации набора данных.

Во втором подразделе представлены результаты экспериментальных исследований модифицированного алгоритма сегментации объектов на основе СНС ResNet50 и ResNet101. Количество эпох (итераций) при обучении СНС варьировалось в пределах от 100 до 180, в зависимости от используемой комбинации каналов, что соответствовало времени расчета от 24 до 96 часов.

Результаты сегментации и построения пиксельных масок для каждого дерева на изображениях дают среднюю точность 100% на каждом из наборов данных (Рисунки 5 и 6). При этом анализ результатов также показал, что точность сегментации модифицированного алгоритма на искусственно увеличенном наборе данных на 2% выше точности для модели, обученной на наборе данных без дополнения.

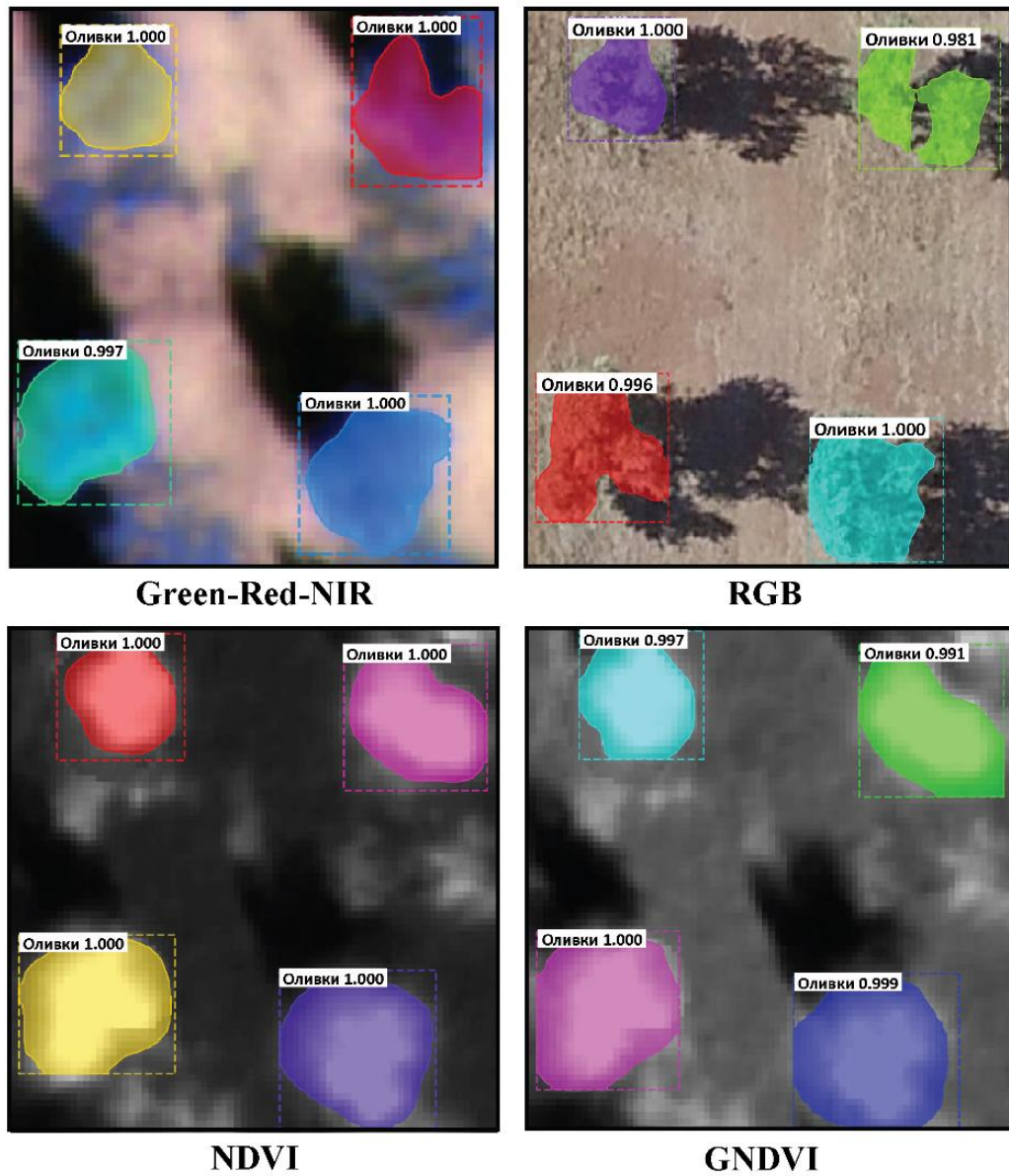


Рисунок 5 – Результат сегментации класса «Оливковые деревья» с использованием Mask R-CNN, обученной на наборе данных без искусственного увеличения

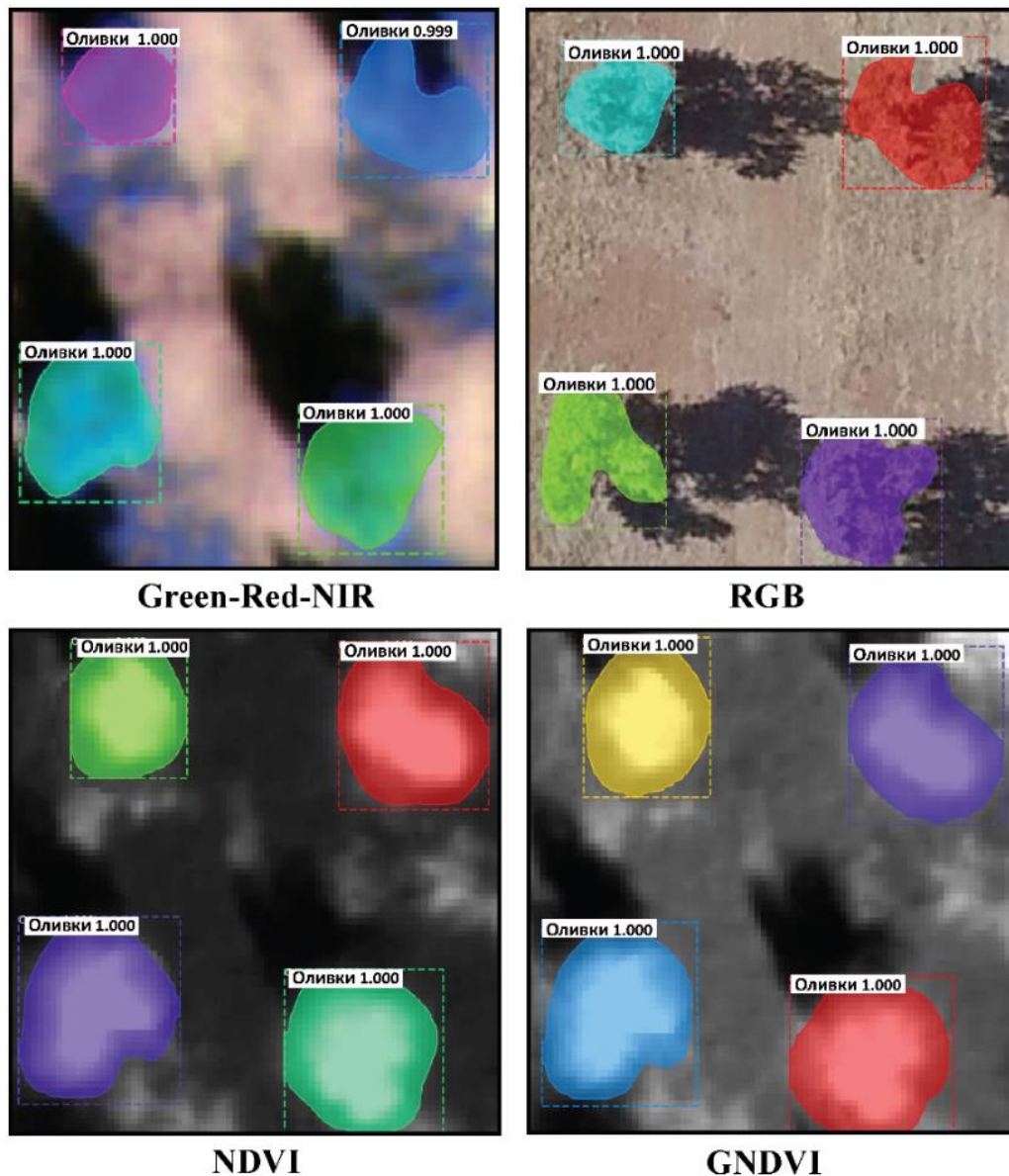


Рисунок 6 – Результат сегментации класса «Оливковые деревья» с использованием Mask R-CNN, обученной на искусственно увеличенном наборе данных

Приводятся дополнительные автоматические расчеты размеров оливковых деревьев для возможного прогнозирования прибыли и урожайности по следующим показателям:

- Попиксельная площадь объекта;
- Приближенная высота дерева (h), определяемая по размеру тени дерева в метрах (l_T) с учетом высоты солнца над горизонтом в зависимости от географического положения (широты и долготы), даты и времени съемки ($tg\gamma$) (14):

$$h = l_T \times tg\gamma; \quad (14)$$

- Приближенный объем кроны дерева, без учета нижней части ствола дерева, на котором нет веток (в среднем до 0,5 метров).

Результаты расчета отличались от полевых (наземных) измерений в среднем на 10%. Таким образом, данный подход имеет перспективы для использования на фермерских хозяйствах, в частности, для прогнозирования урожайности и прибыли.

ЗАКЛЮЧЕНИЕ

Представленная диссертационная работа содержит описание и результаты тестирования разработанных методов машинного обучения при обработке изображений сверхвысокого пространственного разрешения в условиях малых выборок по искусственно увеличенным данным на примере задач сегментации, классификации и категоризации характера повреждения растительности. Основные научные и практические результаты работы заключаются в следующем:

1. Проведен анализ современных алгоритмов СНС и их использования при решении задач классификации объектов, представленных на мультиспектральных и трехканальных изображениях сверхвысокого пространственного разрешения, позволяющий определить основные направления исследования в области разработки эффективных алгоритмов СНС для обработки данных ДЗЗ.

2. Разработан алгоритм построения и искусственного увеличения размеченного набора данных на изображениях сверхвысокого пространственного разрешения в условиях малых выборок.

3. Разработана новая архитектура СНС для задач точной классификации размеченного набора данных на трехканальных цветных изображениях сверхвысокого пространственного разрешения.

4. Разработана новая методика формирования контрольной выборки на трехканальных цветных изображениях сверхвысокого пространственного разрешения в условиях малых выборок для проведения тестирования разработанной архитектуры СНС и ее сравнения с современными моделями СНС.

5. Проведено модифицирование алгоритма сегментации изображений сверхвысокого пространственного разрешения в условиях малых выборок на основе СНС ResNet50 и ResNet101 для построения маски регионов объектов.

6. Выполнены экспериментальные исследования по разработанным архитектурам СНС и алгоритмам на тестовых участках: заповедник «Столбы», расположенный на северо-востоке от города Красноярск в Центральной Сибири Российской Федерации по данным изображений, полученных с БПЛА за 2016, 2018 гг.; плантация оливковых деревьев *Picular*, расположенная на севере города Гранады, Андалусия, Испания по данным изображений БПЛА за 2019 г.

7. Произведено вычисление основных параметров метрики оценки разработанной архитектуры СНС с использованием различных наборов данных изображений сверхвысокого пространственного разрешения. По проведенным расчетам выявлено, что от использования искусственного увеличения тестового набора данных для обучения СНС доля правильных ответов (accuracy), точность (precision) и F-мера (F_score) улучшились на 14,5%, 20,4% и 34,3% соответственно.

СПИСОК ПУБЛИКАЦИЙ ПО ТЕМЕ ДИССЕРТАЦИИ

В изданиях, рекомендованных ВАК:

1. **Safonova, A.** Classification of agricultural crops from middle – resolution satellite images using Gaussian processes based method / A. Safonova, Y. Dmitriev // Journal of Siberian Federal University. Engineering & Technologies. – 2018. – № 11(8). – P. 909-921.

2. Дмитриев, Е.В. Комплексование классификаторов в задаче тематической обработки гиперспектральных аэрокосмических изображений / Е.В. Дмитриев, В.В. Козодеров, А.О. Дементьев, **А.Н. Сафонова** // Автометрия. – 2018. – Том 54, № 3 – С. 3-12.

3. **Safonova, A.** Detection of Fir Trees (*Abies sibirica*) Damaged by the Bark Beetle in Unmanned Aerial Vehicle Images with Deep Learning / A. Safonova, S. Tabik, D. Alcaraz-Segura, A. Rubtsov, Y. Maglinets, F. Herrera // Remote Sensing. – 2019. – № 11. – P. 643. (индексируется в **Scopus** и **Web of Science**)

4. Dmitriev, E.V. Combining classifiers in the problem of thematic processing of hyperspectral aerospace images / E.V. Dmitriev, V.V. Kozoderov, A.O. Demytyev, **A.N. Safonova** // Optoelectronics, Instrumentation and Data Processing. – 2018. Volume 54, Issue 3. – P. 213–221. (индексируется в **Scopus** и **Web of Science**)

5. **Safonova, A.** Classification of agricultural crops of earth remote sensing data using Gaussian based methods / A. Safonova, E. Dmitriev // CEUR Workshop Proceedings. – 2018. – № 2033. – P. 225-230. (индексируется в **Scopus** и **Web of Science**)

В других изданиях и материалах конференций:

6. Дмитриев, Е.В. Классификация и оценка состояния смешанных древостоев по аэроизображениям сверхвысокого пространственного разрешения / Е.В. Дмитриев, В.А. Козуб, П.Г. Мельник, А.А. Соколов, **А.Н. Сафонова** // Известия высших учебных заведений. Лесной журнал. – 2019. – № 5. (индексируется в **Web of Science**)

7. **Сафонова, А.Н.** Сегментация оливковых деревьев с использованием нейронной сети Mask R-CNN / А.Н. Сафонова, С. Табик, Ю.А. Маглинец // Сборник трудов Всероссийской конференции с международным участием «Обработка пространственных данных в задачах мониторинга природных и антропогенных процессов» (SDM-2019). – Новосибирск. – 2019. – С. 456-460.

8. **Safonova, A.** Mask R-CNN for segmentation of olive trees // Материалы VI Международной научной конференции «Региональные проблемы дистанционного зондирования Земли» (РПДЗЗ-2019) / A. Safonova, E. Guirado, S. Tabik, D. Alcaraz-Segura, Y. Maglinets // Красноярск. – 2019. – С. 167-172.

9. **Safonova, A.** Detection of Fir Trees (*Abies Sibirica*) Damaged by the Bark Beetle in Unmanned Aerial Vehicle Images with Deep Learning / A. Safonova, S. Tabik, D. Alcaraz-Segura, A. Rubtsov, Y. Maglinets, F. Herrera // II Congreso Nacional – IV Jornadas de Investigadores en Formación: Fomentando la Interdisciplinariedad. – Granada, España. – 2019.

10. Дмитриев, Е.В. Определение таксационных параметров лесных территорий на основе космических изображений высокого разрешения / Е.В. Дмитриев, В.В. Козодеров, А.А. Соколов, **А.Н. Сафонова** // Материалы V Международной научной конференции «Региональные проблемы дистанционного зондирования Земли» (РПДЗЗ-2018). – Красноярск: 2018. – С. 16-20.

11. **Safonova, A.** Application of convolutional neural networks in detection of forest damage caused by the Polygraphus Proximus beetle / A. Safonova, A. Rubtsov, S. Tabik, D. Alcaraz-Segura // Материалы V Международной научной конференции «Региональные проблемы дистанционного зондирования Земли» (РПДЗЗ-2018) / Красноярск. – 2018. – С. 235-241.

12. **Сафонова, А.Н.** Классификация сельскохозяйственных культур по данным дистанционного зондирования Земли с использованием метода Гаусса / А.Н. Сафонова // Материалы IV Международной научной конференции «Региональные проблемы дистанционного зондирования Земли» (РПДЗЗ-2017). – Красноярск: 2017. – С. 161-166.

13. **Сафонова, А.Н.** Классификация сельскохозяйственных культур по данным дистанционного зондирования Земли с использованием метода Гаусса / А.Н. Сафонова, Р. Молина // Сборник трудов Всероссийской конференции «Обработка пространственных данных в задачах мониторинга природных и антропогенных процессов» (SDM-2017). – Новосибирск: 2017. – С. 225-230.