

Министерство образования Республики Беларусь  
Учреждение образования  
Белорусский государственный университет  
информатики и радиоэлектроники

УДК 004.93

Бочкарев  
Кирилл Юрьевич

Модели и методы распознавания и классификации объектов микологии

**АВТОРЕФЕРАТ**

на соискание академической степени  
магистра технических наук

по специальности 1-40 80 02

Системный анализ, управление и обработка информации

Научный руководитель

Ревотюк М.П.  
к.т.н., доцент

Минск 2019

## ВВЕДЕНИЕ

Развитие науки и технологий привело к тому, что в последние десятилетия во всех сферах нашей жизни используются технические средства, выполняющие автоматическое распознавание образов. Чаще всего его реализация осуществляется путем моделирования операций, выполняемых живыми организмами в процессе взаимодействия и восприятия окружающего мира. Наиболее естественно положить в основу модели распознавания способности человека и его реакции на окружающую действительность. Дополнительным аргументом в пользу такого подхода явилось стремление возложить функции человека на автоматические устройства в тех областях, где условия утомительны или опасны для жизнедеятельности людей.

В процессе распознавания участвует не сам объект, а некоторое его приближение, называемое *образом*. Процедура восприятия образов предшествует процессу распознавания, после чего следует процесс идентификации. Диссертация посвящена распознаванию объектов микологии. Идентификация образов является результатом классификации объектов с последующей их локализацией на основе визуальной информации. Решаемая задача напрямую связана с машинным обучением и компьютерным зрением.

Искусственные нейронные сети (ИНС), особенно сверточные, являются предпочтительной моделью для распознавания. Они способны комбинировать в себе сразу две задачи: классификацию и локализацию образов. Выбор архитектуры нейронной сети является нетривиальной задачей, связанной с анализом большого объема данных. Часто параметры и архитектура сети подбираются экспериментальным путем, что является трудоемким процессом. Это связано с тем, что каждая задача имеет уникальные особенности: данные, ожидаемый результат, необходимую обобщающую способность.

Для автоматизации построения сети и поиска оптимального решения на ней могут применяться генетические алгоритмы. Их цель – определить в

ходе эволюции архитектуру и параметры ИНС, обеспечивающие успешное решение поставленной задачи.

## **ОБЩАЯ ХАРАКТЕРИСТИКА РАБОТЫ**

### **Цель и задачи исследования**

Целью диссертации является разработка моделей и методов для классификации и распознавания заданных образов на изображениях.

Для достижения поставленной цели необходимо решить ряд задач:

- выполнить анализ существующих моделей и методов классификации и распознавания образов;
- разработать модели для классификации и распознавания образов;
- разработать генетический алгоритм для автоматической генерации архитектуры искусственной нейронной сети;
- решить задачу распознавания образов с помощью двух разработанных моделей и провести сравнительный анализ полученных результатов;
- создать мобильное приложение, распознающее объекты микологии;

*Объект исследования* – классификация и распознавание образов.

*Предмет исследования* – модели и методы классификации и распознавания образов, а также способы оптимизации имеющихся моделей.

Основной гипотезой, положенной в основу диссертационной работы, является предположение о возможности применения ИНС и генетических алгоритмов для решения задач классификации и распознавания образов.

### **Личный вклад соискателя**

Результаты, приведенные в диссертации: анализ существующих моделей и методов, разработанные модели и алгоритмы, мобильное программное приложение, получены автором лично.

## **Публикация результатов диссертации**

По теме диссертации было опубликовано две печатные работы: статья в научном журнале и тезисы в сборнике трудов научной конференции. Также результаты работы докладывались на 54-й и 55-й конференциях аспирантов, магистрантов и студентов БГУИР.

## **Структура и объем диссертации**

Диссертация состоит из общей характеристики работы, введения, четырех глав, заключения, списка использованных источников, списка публикаций автора и приложений.

В первой главе выполнен аналитический обзор моделей и методов распознавания и классификации образов, после чего в результате анализа базовых архитектур ИНС был сделан выбор модели в виде сверточной ИНС.

Во второй главе рассмотрены модели, позволяющие решить задачи диссертации двумя способами: на основе базовой архитектуры ИНС и с помощью автоматически сгенерированной ИНС. Для реализации первого подхода проанализировано несколько современных базовых архитектур ИНС, из которых выбрана одна, показавшая лучшие результаты по заданным критериям. Второй подход реализуется на основе модифицированной архитектуры сверточной ИНС, способной решать задачи классификации и распознавания образов.

В третьей главе предложен ряд алгоритмов, обеспечивающих поэтапную реализацию генетического подхода к решению поставленных задач с помощью модифицированной ИНС.

Четвертая глава посвящена разработке мобильного программного приложения, в основу которого положены модели и алгоритмы диссертации. Распознавание заданных образов проводилось с помощью двух моделей, после чего сравнивалось качество полученных результатов. Тестирование разработанных моделей и методов показало их работоспособность и возмож-

ность использования для решения прикладных задач – распознавания объектов микологии.

Общий объем работы составляет 95 страниц, из которых основного текста 76 страниц, 24 рисунка, 3 таблицы, список использованных источников из 31 наименования и приложения на 19 страницах.

## КРАТКОЕ СОДЕРЖАНИЕ РАБОТЫ

Во **введении** определена предметная область, указано основное направление исследования, показана актуальность темы диссертационной работы, дана общая характеристика исследуемых вопросов, обозначена практическая ценность, сформулированы цели и задачи исследования.

Прикладной областью диссертации являются объекты микологии, которые требуется распознавать и распределять на классы: съедобные, несъедобные и условно съедобные. Предусмотрен еще один класс, в который попадают объекты «не грибы».

Задачи, поставленные в диссертации, решаются с помощью моделей распознавания и классификации образов. Поэтому **в первой главе** выполнен аналитический обзор и сравнительный анализ моделей распознавания образов. Рассмотрены модели классификации образов. Особое внимание уделено искусственным нейронным сетям, поскольку их применение к решению задачи распознавания и классификации является достаточно универсальным подходом. Чаще других для глубокого анализа данных применяются сверточные ИНС. К их преимуществам можно отнести инвариантность относительно области применения, хорошие показатели обучаемости и работы на больших данных. В подтверждение этого рассмотрены популярные архитектуры сверточных ИНС: *VGG*, *Inception*, *ResNet*, *MobileNet*, а также программные средства *GoogLeNet Inception* и *ResNet Microsoft*, построенные на основе сверточных ИНС и имеющие достаточно большие функциональные возможности.

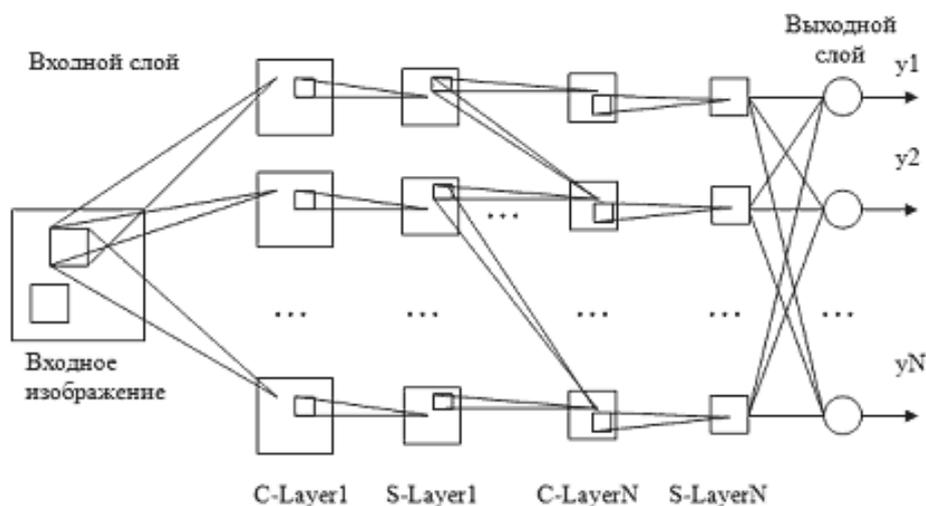
**Во второй главе** на основании анализа архитектур ИНС, рассмотренных ранее, необходимо выбрать архитектуру или ее модификацию, наиболее подходящую для решения задач диссертации. Для проверки работы архитектур ИНС была составлена обучающая выборка, объекты которой сформировали четыре ранее выделенные классы для объектов микологии. Обучающая выборка состояла приблизительно из 36000 изображений, где на каждый класс приходилось около 9000 образов. Обучение проводилось средствами библиотек TensorFlow и Keras при помощи графического ускорителя NVIDIA GeForce RTX 2080ti. Экспериментально было проведено 10 итераций обучения каждой из выбранных моделей сверточных ИНС. С помощью программных средств анализа, предоставляемых библиотекой Keras, модели сравнивались по нескольким важным характеристикам: требуемым ресурсам; потери данных на текущей итерации обучения относительно предыдущей.

В итоге после проведенного обучения, анализа полученных данных и тестирования на случайных изображениях для распознавания и классификации объектов микологи была выбрана модель ИНС на основе архитектуры *ResNet50*. Выбор в ее пользу был сделан по причине оптимального сочетания эффективности решения задачи распознавания и ресурсоёмкости, что является наиболее важными факторами при разработке приложений для мобильных платформ. Появление семейства архитектур *ResNet* явилось результатом исследований возможности создания сверточных ИНС большей глубины.

Кроме вышеперечисленного, сверточные нейронные сети способны комбинировать в себе одновременно решения задач распознавания и классификации образов без дополнительных алгоритмов и средств. Поскольку подбор оптимальных параметров сети является трудоемким процессом из-за ручного перебора и подгона, для автоматизации построения сети, а также для решения проблемы поиска оптимума, могут применяться генетические алгоритмы. Их основная задача – путем эволюции выбрать ту архитектуру и параметры сети, которые успешно обеспечат решение поставленной задачи. Применяя генетический алгоритм, можно получить готовые нейросетевые

решения, зная только обучающую выборку. К достоинствам эволюционных методов относятся: разнообразие получаемых топологий, адаптивность, универсальность.

Для успешной работы генетического алгоритма была выбрана архитектура сверточной ИНС, состоящая из чередующихся сверточных (С-слой) и субдискретизирующих слоев (S-слой), а также одного полносвязного слоя на выходе (F-слой). Пример архитектуры изображен на рисунке 1.



**Рисунок 1 – Архитектура сверточной нейронной сети**

Поскольку задачи классификации и локализации образов тесно связаны между собой, было решено использовать сверточную ИНС, чтобы получать решение этих двух задач в рамках одной модели. Поэтому в ряде сверточных ИНС отсутствует полносвязный слой, замененный слоем усреднения с уменьшением размерности. В результате существенно уменьшается количество параметров сети на последнем слое, что положительно сказывается на скорости и качестве обучения. Слой усреднения позволяет сохранить детектирующие способности сети вплоть до последнего слоя, целенаправленно не обучая сеть задаче обнаружения объектов. Чтобы определить регионы на изображении, в которых расположен объект, необходимо дополнительно построить активационную карту соответствия. В результате удастся построить комбинированную модель классификации и локализации образов.

В третьей главе с целью автоматизации процесса обучения нейронной сети был использован генетический алгоритм, который автоматически подбирает архитектуру сети в ходе естественного отбора. Эффективность применения генетического алгоритма напрямую зависит от способа кодирования архитектуры ИНС, реализации метода кроссовера, мутации и вычисления фитнес-функции. Схематично процесс динамического выбора архитектуры изображен на рисунке 2.



Рисунок 2 – Процесс эволюционного построения сетей

В рамках генетического подхода было реализованы еще несколько необходимых алгоритмов. Алгоритм кроссинговера связан с перемешиванием генов у двух родительских нейронных сетей. Необходимо учитывать, что сеть

состоит из слоев разных типов: одномерных и двумерных. Поэтому случайное перемешивание ее слоев может привести к генерированию сети, которая будет недействительна с точки зрения совместимости слоев между собой.

Алгоритм мутации изменяет структуру сети, добавляя или удаляя слои. При этом необходимо придерживаться правил изменения структуры сети, чтобы после мутации она сохраняла консистентность и совместимость между слоями.

Обучение сети выполняется методом обратного распространения ошибки. Для этого задаются синоптические веса и пороговые значения, так как априорная информация о них отсутствует. Процесс обучения связан с изменением весовых коэффициентов нейронов каждого слоя на основе вычисления градиента. После выполнения прямого и обратного проходов для очередного обучающего примера проверяется, не достигнуто ли условие останова. В случае положительного ответа алгоритм завершается, иначе сеть обучается на очередном образе.

В целях совместимости алгоритма локализации с разработанным генетическим алгоритмом последний слой ИНС не должен содержать полносвязного слоя. Вместо него вводится слой усреднения. Условия определения координат классифицируемого объекта на изображении приведены в разделе 2.5 диссертационной работы.

**В четвертой главе** разработанные модели и алгоритмы были положены в основу мобильного программного приложения, для реализации которого был выбран язык программирования JavaScript.

Сначала с помощью генетического алгоритма генерировались ИНС. Анализ их обучения позволил сделать вывод, что полносвязные слои подходят для небольшой обучающей базы. Для больших обучающих баз важны признаки высокого порядка, предоставляющие высокую обобщающую способность, а это достигается использованием сверточных слоев. В целом можно сделать вывод, что была разработана достаточно эффективная модель построения нейронных сетей.

Для более полной оценки автоматически сгенерированной нейронной сети было выполнено ее сравнение с уже известным аналогом, который также решает задачу классификации. Это ранее построенная модель архитектуры ResNet50, которая обучалась на изображениях из стандартной библиотеки в течение 50 эпох. Обе модели ИНС находились в одинаковых условиях и обучались на одной базе изображений. Результаты приведены в таблице 1.

Таблица 1 – Сравнение точности классификации

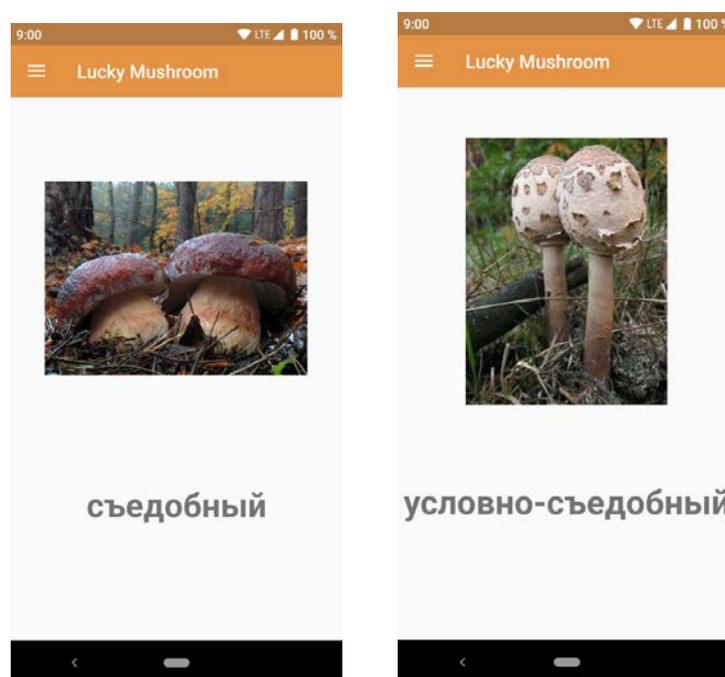
Архитектура	Значение точности распознавания
ResNet50	86,3 %
Автоматически сгенерированная ИНС	73,2%

Получено, что автоматически сгенерированная нейронная сеть уступает реализации ResNet50 на 13% по точности распознавания. Однако, оценивая автоматически сгенерированную ИНС, следует учитывать, что ее построение производилось без участия экспертов предметной области. Поэтому подход автоматического построения сетей имеет место быть и может применяться, например, для реализации прототипов, где на первом месте стоит реализация работающей модели, нежели ее качественные характеристики. В последующем автоматически сгенерированную модель можно подвергать дополнительным изменениям для улучшения конкретных характеристик.

В результате с помощью модели на основе архитектуры ResNet50 (рисунки 3-4) на языке JavaScript было разработано мобильное приложение под операционную систему Android.

Автоматически сгенерированная модель классификации может применяться как основа для решения задачи локализации объектов на изображении. Результаты, полученные с активационной карты, показали, что при использовании сверточных слоев первые из них определяют более специфические признаки анализируемого объекта, а глубокие сверточные слои выяв-

ляют наиболее общие признаки для анализируемого объекта. Именно такие признаки необходимы для успешной локализации объектов.



Рисунки 3-4 – Примеры работы программного средства

## ЗАКЛЮЧЕНИЕ

### Основные научные результаты диссертации

Все задачи диссертации были решены, а цель достигнута. Проведенные исследования показали, что распознавание образов можно выполнять как на основе базовой архитектуры сверточной ИНС, так и с помощью автоматически сгенерированной модели. Приемлемая точность классификации автоматически сгенерированной сети делает возможным использование генетических алгоритмов для решения большого круга задач в машинном обучении. Поскольку все параметры сети определяются автоматически, становится unnecessary подбирать вручную данные для обучения и решения поставленной задачи.

Модель локализации объектов на изображении также показала неплохие результаты. Основным преимуществом разработанного подхода является минимальное изменение модели для классификации.

И хотя предложенная модель требует большого количества вычислительных ресурсов для многократного обучения каждой особи в популяции, но при этом генетические алгоритмы несложно преобразуются для реализации распределенных вычислений. Это позволит в разы уменьшить время работы и обучения всех сетей, так как каждая ИНС будет обучаться независимо от других на отдельном узле.

### **Рекомендации по практическому использованию результатов**

Задачи распознавания образов успешно решены с помощью модели сверточной ИНС на основе базовой архитектуры ResNet 50 и могут использоваться для распознавания и классификации объектов микологии. Кроме того, планируется продолжать исследования по совершенствованию модифицированной модели классификации-локализации образов, а также увеличению эффективности модели за счет использования распределенных вычислений.

### **СПИСОК ОПУБЛИКОВАННЫХ РАБОТ**

1-А Серебряная, Л. В. Модель автоматической классификации и локализации образов/ Л. В. Серебряная, К. Ю. Бочкарев, А. Я. Попитич // Цифровая трансформация. – 2019. – № 1 (6). – С. 43 – 48.

2-А. Бочкарев, К. Ю. / Генетический алгоритм сверточной нейронной сети /К. Ю. Бочкарев // 55-я юбилейная конференция аспирантов, магистрантов и студентов учреждения образования «Белорусский государственный университет информатики и радиоэлектроники», 22-26 апреля 2019 г., БГУ-ИР, Минск, Беларусь: тезисы докладов. – Мн. – 2019. – С. – .