

МИНОБРНАУКИ РОССИИ

Федеральное государственное бюджетное образовательное учреждение
высшего образования

«САРАТОВСКИЙ НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ
ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ ИМЕНИ Н.Г.ЧЕРНЫШЕВСКОГО»

Кафедра информатики и программирования

**ПРИМЕНЕНИЕ ГЕНЕРАТИВНО-СОСТЯЗАТЕЛЬНЫХ СЕТЕЙ ДЛЯ
ВАРИАТИВНОЙ АУГМЕНТАЦИИ НАБОРОВ ДАННЫХ,
ИСПОЛЬЗУЮЩИХСЯ В ЗАДАЧАХ РАСПОЗНАВАНИЯ
ЗАБОЛЕВАНИЙ РАСТЕНИЙ**

АВТОРЕФЕРАТ БАКАЛАВРСКОЙ РАБОТЫ

студентки 4 курса 441 группы

направления 02.03.03 Математическое обеспечение и администрирование
информационных систем

факультета компьютерных наук и информационных технологий

Роговой Наталии Викторовны

Научный руководитель:

старший преподаватель

А.А. Казачкова

подпись, дата

Зав. кафедрой:

к.ф.-м.н., доцент

М.В. Огнева

подпись, дата

Саратов 2021

ВВЕДЕНИЕ

Актуальность темы. Нейронные сети являются одним из наилучших способов решения задачи распознавания образов. Наборы данных, созданные для их обучения, содержат множество различных изображений и помогают добиваться высоких показателей распознавания. Для каждой из задач в рассматриваемой области может требоваться свое определенное количество изображений для обучающего набора, поэтому не существует их универсального минимального или максимального предела. Обычно считается, что чем больше имеется изображений, тем выше шанс получить точный результат. На практике же собрать достаточно обширную обучающую выборку не является тривиальной задачей, особенно если решаемая проблема довольно специфичная. Также важным критерием является соблюдение сбалансированности и репрезентативности такого набора, поэтому недостаточно добиться только большого количества изображений в выборке, а необходимо следить и за качеством собираемых данных.

Обычно для накопления большого набора реалистичных изображений некоторого объекта необходимо прикладывать немалые усилия. Поэтому очень часто на практике встречается проблема нехватки данных для расширения существующей базы новыми примерами. Ведь для того, чтобы нейронная сеть была полезной, ее ошибка на проверочных данных должна продолжать уменьшаться вместе с ошибкой на обучающем наборе, но дефицит изображений, даже если они сами максимально качественны, не позволяет этого добиться.

Увеличение обучающей выборки — очень мощный способ решения данной проблемы. Расширенные данные будут намного полезнее при обучении и позволят минимизировать расстояние между обучающим и проверочным набором, а также любыми новыми данными для тестирования модели. Один из способов расширения обучающей выборки представляет собой процедуру, называемой аугментацией данных, которая может быть

произведена множеством различных способов. Аугментация данных оказывает очень полезное влияние на процесс обучения нейронных сетей и во много раз повышает их качество. В данной работе рассматривается способ проведения аугментации с помощью генеративно-сопоставительных сетей.

Цель бакалаврской работы – применение генеративно-сопоставительных нейронных сетей различных архитектур для проведения вариативной аугментации наборов данных, состоящих из изображений растений с последующим расширением обучающей выборки, сравнением полученных результатов и их визуализацией в Android-приложении.

Поставленная цель определила **следующие задачи**:

1. проанализировать проблему нехватки данных для обучения нейронных сетей при решении задачи распознавания образов, а также способов ее решения;
2. изучить процесс обучения и работы генеративно-сопоставительной нейронной сети, а также ее вида — CycleGAN;
3. сравнить архитектуры двух рассматриваемых в работе нейронных сетей CycleGAN;
4. аугментировать визуальные данные, содержащие изображения здоровых и больных растений, при помощи двух нейронных сетей CycleGAN;
5. провести оценку полученных результатов с помощью сверточной нейронной сети;
6. собрать новый обучающий набор, который содержит изображения другого растения, и провести их аугментацию с помощью CycleGAN для расширения возможности распознавания заболеваний растений;
7. реализовать приложение на платформе Android, которое позволяет создавать изображения с помощью камеры устройства и визуализировать их аугментированное представление.

Методологические основы применения генеративно-состязательных нейронных сетей для аугментации данных представлены в работах Ф. Шолле, С. О. Емельянова, А. Мюллера, Ч. Аггарвала, О. Жерона, Д. Фостера, С. Николенко, А. Кадурина, Е. Архангельской.

Практическая значимость бакалаврской работы.

Созданные посредством генерации выборки изображений здоровых и больных растений могут дополнить существующие наборы данных, содержащие примеры подобные сгенерированным, для обучения моделей машинного обучения в целях повышения качества распознавания этих моделей.

Реализованное в рамках работы приложение для проведения аугментации изображений предоставляет возможность ознакомления с сгенерированными данными в удобном для пользователя формате.

Структура и объём работы. Бакалаврская работа состоит из введения, 4 разделов, заключения, списка использованных источников и 7 приложений. Общий объём работы – 95 страниц, включая 57 рисунков, список использованных источников информации – 32 наименований.

КРАТКОЕ СОДЕРЖАНИЕ РАБОТЫ

Первый раздел «Обзор предметной области» посвящен рассмотрению области компьютерного зрения. Анализируется проблема переобучения нейронных сетей и способ ее решения – аугментация данных, а также проведение аугментации с помощью генеративно-сопоставительных сетей (GAN).

Компьютерное зрение – область искусственного интеллекта, которая применяется для анализа изображений и видео. В настоящий момент существует ряд проблем, решение которых с помощью инструментов компьютерного зрения позволяет улучшить жизнь многих людей. Сейчас в странах Африки остро стоит проблема, заключающаяся в поражении важной продовольственной культуры – маниоки. Существуют болезни, которые поражают маниоку и делают ее корни несъедобными, что является основным источником низких урожаев. Один из способов борьбы с заболеваниями маниоки основан на применении методов компьютерного зрения, для предотвращения бедственных ситуаций. В целях профилактики и предупреждения ее заболеваний необходим инструмент, позволяющий определять по фотографии здорово ли растение, а если же оно оказывается больным, то предсказывал наиболее схожий по симптомам вид заболевания. В качестве такого инструмента может выступать нейронная сеть, но какой бы инструмент компьютерного зрения не был бы выбран, нужно учитывать, что любая модель должна обладать максимальной обобщающей способностью, а случай, когда она этого не умеет называется переобучением. Возникновению переобучения может способствовать малое количество обучающих данных и предотвратить такую ситуацию позволяет сбор как можно большего количества разнообразных изображений.

Аугментация данных – это методика создания дополнительных обучающих данных из уже имеющихся, то есть через модификацию существующих объектов. Наиболее часто используются следующие модификации: отражения вдоль осей, смена цветов, изменение

насыщенности и другие. Альтернативный способ аугментации заключается в порождении одного изображения на основе многих имеющихся, которое почти невозможно отличить от реальной фотографии с помощью GAN.

GAN включает в себя две составляющие: сеть-генератор и сеть-дискриминатор, которые постепенно обучают друг друга. Сеть-генератор ожидает на входе случайный вектор и при получении декодирует его к виду искусственного изображения, похожего на пример из обучающей выборки. Сеть-дискриминатор ожидает на входе изображение, которое может оказаться как настоящим, так и искусственным, а затем определяет к какой категории отнести его: обучающей выборке или поддельным изображениям.

Тип генеративно-состязательной сети – CycleGAN предназначен для трансляции изображений, состоящей в извлечении из изображений, представляющих образцы стиля, стилистических деталей и их внедрении в исходное изображение. В архитектуре такой сети присутствуют четыре модели: два генератора и два дискриминатора. Первый генератор преобразует изображения из первого пространства во второе. Вторым генератором преобразуют изображения из второго пространства в первое. Оба дискриминатора оценивают убедительность изображений, созданных генераторами.

Второй раздел «Постановка задачи» посвящен обобщению предыдущего теоретического материала и обоснованию актуальности аугментации изображений растения маниока с помощью сетей CycleGAN.

Третий раздел «Сравнительный анализ архитектур CycleGAN» посвящен рассмотрению архитектур двух используемых в работе нейронных сетей CycleGAN, с последующим анализом процессов их обучения.

Аугментация обучающих выборок проводилась с помощью двух CycleGAN сетей. Обе сети написаны на языке программирования Python с использованием библиотеки Keras.

Генератор первой сети основывается на сверточной сети U-Net, состоящей из двух частей, где одна из половин отвечает за понижение

разрешения, а другая за повышение. В качестве дискриминатора используется предварительно обученная на наборе ImageNet модель сверточной сети VGG16, для распознавания объектов на изображениях.

Генератор второй сети основан на классическом подходе к построению архитектуры CycleGAN и использует три блока: кодировщик, трансформер и декодер. Кодировщик понижает разрешение, трансформер состоит из остаточных сетевых блоков, представляющих сеть ResNet, а декодер повышает разрешение. Для дискриминатора используются веса обученной сети MobileNetV2, предназначенной для распознавания изображений.

Для обучения первой сети было выбрано 13 эпох, где во время первых 3 дискриминаторы необучаемы, чтобы не нарушать в первые эпохи загруженные и предварительно натренированные веса сети VGG16. Для обучения второй сети количество эпох было выбрано равным 30 и все модели были обучаемы с самого начала. Размер пакета, означающий количество выборок на обновление градиента, для первой сети равен 1, а для второй 2. Количество шагов в эпоху для первой сети равно $2000//\text{batch_size}$, а для второй $2577//\text{batch_size}$.

Четвертый раздел «Проведение аугментации данных с помощью CycleGAN» посвящен применению CycleGAN для проведения аугментации обучающего набора из изображений маниоки, также описывается процедура предобработки данных. Далее полученные изображения сравниваются как визуально, так и с помощью сверточной нейронной сети (CNN). В конце раздела описывается расширение обучающей выборки изображениями других растений и проведение аугментации на них.

Предобработка данных включает их считывание в формате TFRecord и приведение в более удобный вид для последующей работы. Затем проводится их аугментация, которая включает усечение изображений, повороты, изменение яркости и контрастности, регулировку RGB оттенка. После все данные повторяются в нескольких эпохах и перемешиваются. Заключительный этап – нормализация и изменение размера изображения.

Для проведения аугментации изображений больной маниоки из 4 возможных видов ее заболевания были выбраны 2 — Cassava Bacterial Blight (CBB) и Cassava Brown Streak Disease (CBSD).

Первая сеть справилась со своей задачей и смогла сгенерировать довольно реалистичные изображения, при этом не ухудшив качество оригинала, а лишь наложив стилевые характеристики. Результаты второй сети на наборе, содержащем изображение маниоки больной CBB немного менее качественны, чем у первой, а на наборе данных с заболеванием CBSD оказались достаточно реалистичны в обе стороны.

Для оценки качества сгенерированных изображений сверточной нейросетью с помощью обеих CycleGAN были сгенерированы 2576 изображений здоровой маниоки из изображений больной обеими рассматриваемыми болезнями, а также 1086 больной бактериальным ожогом и 2189 больной вирусом коричневой полосы из здоровой. В качестве тестовых данных ей поочередно были отданы все вышеперечисленные выборки и затем получена классификация сгенерированных изображений. Далее приведены результаты для первой сети CycleGAN.

Из синтетических изображений маниоки больной бактериальным ожогом, большую часть изображений (1702 образцов) сверточная сеть распознала как здоровые, при этом среди классов изображений больных растений большая часть была верно отнесена к классу CBB (447 образцов).

Больше половины сгенерированных изображений маниоки больной вирусом коричневой полосы были распознаны правильно (1681 образцов).

Из синтетических изображений здоровой маниоки классифицированы 590 изображений как здоровые листья, а следующая по величине группа (331 образцов) была ошибочно определена как листья больные бактериальным ожогом из которых и генерировались здоровые.

Большая часть сгенерированных изображений здоровой маниоки из изображений маниоки больной вирусом коричневой полосы классифицирована правильно (1683 образцов), а из классов, относящихся к

больным растениям больше всего, занимают изображения, определенные как CBSD из которой генерировалась здоровая маниока.

Далее приведены результаты для второй сети CycleGAN.

Правильно классифицированы 422 синтетических изображения (СВВ) и неверно определены 1472 образца как здоровые из которых происходила генерация.

Правильно классифицированы 1604 синтетических изображения (CBSD) и неверно определены 742 образца как здоровые.

Правильно классифицированы 707 синтетических изображений здоровой маниоки, а 158 как СВВ из которой генерировалась здоровая.

Правильно классифицированы 1092 синтетических изображения здоровой маниоки, а 867 как CBSD из которой генерировалась здоровая.

Таким образом в генерации из изображений здоровой маниоки в изображения больной бактериальным ожогом, вирусом коричневой полосы и из изображений маниоки больной вирусом коричневой полосы в изображения здоровой лучше себя показала первая сеть CycleGAN. У второй сети получилось лучше сгенерировать изображения здоровой из больной бактериальным ожогом чем у первой.

Полезно проводить процедуру аугментации и для других растений. Поэтому в рамках данной работы было собрано два обучающих набора данных, один из которых содержит изображения листьев хлопчатника, а другой листьев кукурузы. Было собрано 282 изображения здоровых листьев хлопчатника и 279 изображений листьев с солнечным ожогом. Для второго обучающего набора было собрано 1624 изображения здоровых листьев кукурузы и 1426 изображений листьев кукурузы, поврежденных кукурузной листовой совкой.

Для проведения аугментации на новых наборах данных была выбрана первая сеть CycleGAN, так как она показала более высокие результаты. Три из четырех случаев генерации изображений растений на новых наборах

данных оказались успешными и в дальнейшем эти сгенерированные изображения можно использовать для аугментации имеющихся данных.

Пятый раздел «Разработка приложения на платформе Android» посвящен реализации Android-приложения, которое позволяет наглядно продемонстрировать результат генерации, а также выбрать способ создания изображения, подлежащего преобразованию и модель, способную выполнять это преобразование.

В данной работе был выбран наиболее часто используемый в разработке Android-приложений язык Java. В качестве среды разработки использовалась IDE Android Studio. Вместо эмулятора приложение запускалось и проверялось на работоспособность на реальном устройстве Samsung Galaxy A50.

Шесть обученных моделей CycleGAN были сохранены в формате .h5 и переведены в формат библиотеки TensorFlow Lite с разрешением .tflite. Для загрузки моделей в приложение, импортируется и используется интерпретатор Tflite. Таким образом объект Interpreter будет инициализирован конкретной моделью, используя файл .tflite, и содержать в себе рассматриваемую модель для генерации новых изображений.

В файле activity_main.xml был добавлен элемент-контейнер ImageView и две кнопки одна из которых отвечает за открытие системного приложения «Камера», а другая любого хранилища фотографий. Были добавлены две кнопки для генерации изображения и выбора модели, а также второй элемент ImageView для помещения в него нового изображения.

В результате нажатия кнопки Generate в нижнем элементе будет получено новое изображение, в соответствии с выбранной моделью.

ЗАКЛЮЧЕНИЕ

В ходе работы были выполнены все поставленные задачи:

- изучено устройство генеративно-состязательных сетей, в частности CycleGAN;
- проведена аугментация базы изображений маниоки с помощью двух сетей CycleGAN;
- получена оценка качества аугментированных изображений путем их классификации сверточной нейронной сетью;
- расширен обучающий набор данных, состоящий из маниоки, добавлением к нему изображений хлопчатника и кукурузы;
- проведена аугментация расширенного набора изображений;
- создано приложение на платформе Android для визуализации полученных результатов.

Таким образом, поставленная цель, заключающаяся в проведении аугментации посредством обучения CycleGAN, выполнена.

Полученные аугментированные изображения достаточно качественные, но не всегда удавалось добиться совершенного результата, поэтому в дальнейшем есть возможность подбора наиболее подходящих архитектур нейронных сетей и параметров для их обучения. При этом проведение аугментации доступно на ограниченном количестве видов растений и их заболеваний, поэтому также возможно более широкое расширение обучающей выборки изображениями других растений и их различных болезней.

В продолжение разработки можно усовершенствовать Android-приложение путем добавления в него большего функционала, например такого как классификация здоровых и больных растений и получение информации о заболевании из приложения, отправка сгенерированного изображения напрямую из приложения, а также создание более удобного и гармоничного интерфейса.

Основные источники информации:

1. Шолле, Ф. Глубокое обучение на Python / Ф. Шолле ; пер. А. Киселев. — СПб. : Питер, 2019. — 400 с.
2. Емельянов, С. О. Методы аугментации обучающих выборок в задачах классификации изображений / С. О. Емельянов, А. А. Иванова, Е. А. Швец, Д. П. Николаев // Сенсорные системы. — 2018. — Т. 32, № 3. — С. 236-245.
3. Мюллер, А. Введение в машинное обучение Python / А. Мюллер, С. Гвидо ; пер. А. В. Груздева. — СПб. : ООО «Альфа-книга», 2017. — 480 с.
4. Аггарвал, Ч. Нейронные сети и глубокое обучение. Учебный курс. / Ч. Аггарвал ; пер. А. Г. Гузикевич. — СПб. : ООО «Диалектика», 2020. — 752 с.
5. Жерон, О. Прикладное машинное обучение с помощью Scikit-Learn и TensorFlow: концепции, инструменты и техники для создания интеллектуальных систем / О. Жерон ; пер. Ю. Н. Артеменко. — СПб. : ООО «Альфа-книга», 2018. — 688 с.
6. Фостер, Д. Генеративное глубокое обучение. Творческий потенциал нейронных сетей / Д. Фостер ; пер. А. Киселева. — СПб. : Питер, 2020. — 336 с.
7. Николенко, С. Глубокое обучение / С. Николенко, А. Кадури, Е. Архангельская. — СПб. : Питер, 2018. — 480 с.