Министерство науки и высшего образования Российской Федерации Федеральное государственное бюджетное образовательное учреждение высшего образования

ТОМСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ СИСТЕМ УПРАВЛЕНИЯ И РАДИОЭЛЕКТРОНИКИ (ТУСУР)

Кафедра автоматизации обработки информации (АОИ)

Отчёт к Лабораторной работе №6

по дисциплине «Нейронные сети и их применение»

Выполнил:

Студент группы 422–M1

\_\_\_\_Белоус Г.В.

Принял:

К.т.н., Доцент кафедры АОИ

\_\_\_\_Аксёнов С.В.

Оглавление

[1 Введение 3](#_Toc137045847)

[1.1 Цель лабораторной работы 3](#_Toc137045848)

[1.2 Задание на лабораторную работу 3](#_Toc137045849)

[2 Ход выполнения работы 5](#_Toc137045850)

[2.1 Бинарный классификатор изображений 6](#_Toc137045851)

[2.2 Многоклассовый классификатор изображений 8](#_Toc137045852)

[2.3 Итоговые метрика каждой модели 9](#_Toc137045853)

[3 Вывод 12](#_Toc137045854)

# Введение

## 1.1 Цель лабораторной работы

Получить навыки создания и обучения нейронных сетей для классификации изображений, а также переноса обучения.

## **1.2 Задание на лабораторную работу**

Построить нейросетевые модели – бинарный и многоклассовый классификатор, согласно Вашему варианту задачи.

Сначала построить бинарный классификатор для классов 0 и 1.

Выполнить загрузку и предварительную обработку данных из наборов. Разделить каждую выборку на обучающую, тестовую и валидационную. Произвести обучение набора нейросетевых архитектур, отличающихся разным набором параметров: количество карт признаков, распределение слоёв, процедур оптимизации:

- Подобрать архитектуры нейронных сетей, которые с одной стороны позволяют получить модели с лучшими метриками качества работы, с другой стороны не являются избыточными и не переобученными для следующих задач.

Выполнить эксперименты с прореживанием (Dropout).

Привести кривые изменения функции потерь (Loss) и верности (Accuracy) во время обучения на обучающем и валидационном наборах.

Выполнить аугментацию данных и провести повторное обучение на новом расширенном наборе.

Сохранить лучшую архитектуру нейронной сети – бинарного классификатора.

Создать нейронную сеть – многоклассовый классификатор, который классифицирует изображения двух уже имеющихся классов изображений и третьего класса с использованием сверточной основы, полученной ранее. Выполнить дообучение модели. Сохранить лучшую модель. Привести кривые изменения функции потерь (Loss) и верности (Accuracy) во время обучения на обучающем и валидационном наборах из трех классов.

Варианты (выборки): Набор данных: https://www.kaggle.com/datasets/emmarex/plantdisease Класс 0: Potato\_\_Late \_blight, Класс 1: Potato\_\_healthy, Класс 2: Potato\_\_Early\_blight

Содержание отчета.

1. Описание наборов данных.

2. Параметры архитектур и обучения нейронных сетей, использованные для обучения.

3. Графики обучения для архитектур нейронных сетей с лучшими характеристиками эффективности

4. Оценки моделей на тестовых выборках в виде таблиц/ диаграмм, отображающих метрики качества.

5. Программный код с комментариями.

6. Выводы

# Ход выполнения работы

Датасет, данные из которого обрабатывались в данной работе, представляет из себя набор изображений с листьями картофеля, 1000 изображений с листьями подверженных болезни на ранней стадии, 1000 изображений с листьями подверженных болезни на поздней стадии и 152 изображения здоровых листьев.

Обработанные данные были разделены на выборки для обучения и валидации (соответственно 80 и 20%). Предобработка данных осуществлялась с помощью инструментов библиотеки ImageDataGenerator и метода flow\_from\_directory. С помощью этих инструментов изображения приводились к разрешению 150 на 150, а так же над данными осуществлялась аугментация.

Стоит отметить, что для обоих классификаторов использовалась уже предобученная модель. VGG16 (Visual Geometry Group 16) - это глубокая сверточная нейронная сеть, которая была разработана исследовательской группой Visual Geometry Group в Университете Оксфорда. Она получила свое название от количества слоев в сети (16 слоев).

VGG16 была представлена в статье "Very Deep Convolutional Networks for Large-Scale Image Recognition" в 2014 году и стала одной из ведущих архитектур сверточных нейронных сетей для классификации изображений. Она имеет простую и прямолинейную структуру, состоящую из последовательных сверточных слоев и пулинговых слоев, за которыми следуют полносвязные слои.

В данном случае VGG16 используется как базовая модель для извлечения признаков из изображений. Предварительно обученная VGG16 имеет большую вычислительную мощность и способность распознавать широкий спектр объектов на изображениях. Загружая предварительно обученную модель VGG16 и замораживая ее веса, мы можем использовать ее для извлечения высокоуровневых признаков из изображений в качестве входа для последующих слоев сети, которые будут обучаться на новых данных.

Использование предварительно обученной VGG16 позволяет улучшить производительность модели и сократить время обучения, поскольку модель уже обучена на большом наборе данных для классификации изображений.

## Бинарный классификатор изображений

Бинарный классификатор представляет из себя нейронную сеть из четырех слоев:

* Слой добавляющий предварительно обученную модель VGG16 в качестве первого слоя в итоговую модель. Он служит для извлечения высокоуровневых признаков из входных изображений;
* Слой выполняющий операцию "сплющивания" (Flatten), которая преобразует выходные данные предыдущего слоя в одномерный вектор. Это необходимо для передачи данных в последующие полносвязные слои;
* Полносвязный слой с 256 нейронами и функцией активации ReLU (Rectified Linear Unit). Он выполняет нелинейное преобразование данных, добавляя гибкость и сложность модели;
* Выходной слой спредставляет полносвязный слой с одним нейроном и функцией активации сигмоиды (sigmoid). Он используется для бинарной классификации, где выходное значение находится в диапазоне от 0 до 1, представляя вероятность принадлежности к одному из двух классов.

В данном случае был выбран оптимизатор RMSprop с функцией потерь binary\_crossentropy для решения задачи бинарной классификации.

Функция потерь binary\_crossentropy (бинарная перекрестная энтропия) является стандартным выбором для задач бинарной классификации. Она оптимизирует модель, чтобы минимизировать расстояние между предсказанными значениями и фактическими метками классов.

Оптимизатор RMSprop (Root Mean Square Propagation) является адаптивным оптимизатором, который выполняет обновление весов на основе градиентов и скользящего среднего квадратов градиентов. Он позволяет эффективно настраивать скорость обучения для каждого параметра в процессе обучения модели.

Для обучения модели были определены значения эпох (20) и размер выборки (32).

В итоге модель обучалась двадцать эпох, график функции ошибки от количества эпох обучения показан на рисунке 2.2:

Изображение выглядит как текст, снимок экрана, линия, График

Автоматически созданное описание

Рисунок 2.1 – График величины функции ошибки от количества эпох при обучении модели бинарного классификатора

## Многоклассовый классификатор изображений

Бинарный классификатор представляет из себя нейронную сеть из четырех слоев:

* Слой добавляющий предварительно обученную модель VGG16 в качестве первого слоя в итоговую модель. Он служит для извлечения высокоуровневых признаков из входных изображений.
* Слой выполняющий операцию "сплющивания" (Flatten), которая преобразует выходные данные предыдущего слоя в одномерный вектор. Это необходимо для передачи данных в последующие полносвязные слои.
* Полносвязный слой с 256 нейронами и функцией активации ReLU (Rectified Linear Unit). Он выполняет нелинейное преобразование данных, добавляя гибкость и сложность модели.
* Выходной слой представляющий полносвязный слой с количеством нейронов, соответствующим количеству классов, и функцией активации softmax. Он используется для многоклассовой классификации, где выходное значение представляет вероятности принадлежности к каждому классу.

Для данной модели был выбран оптимизатор RMSprop с функцией потерь categorical\_crossentropy для решения задачи многоклассовой классификации.

Функция потерь categorical\_crossentropy (категориальная перекрестная энтропия) является стандартным выбором для задач многоклассовой классификации. Она оптимизирует модель, чтобы минимизировать расстояние между предсказанными вероятностями классов и фактическими метками классов.

Оптимизатор RMSprop (Root Mean Square Propagation) является адаптивным оптимизатором, который выполняет обновление весов на основе градиентов и скользящего среднего квадратов градиентов. Он позволяет эффективно настраивать скорость обучения для каждого параметра в процессе обучения модели.

Для обучения модели были определены значения эпох (20) и размер выборки (32).

В итоге модель обучалась двадцать эпох, график функции ошибки от количества эпох обучения показан на рисунке 2.2:

Изображение выглядит как текст, снимок экрана, График, линия

Автоматически созданное описание

Рисунок 2.2 – График величины функции ошибки от количества эпох при обучении модели многоклассового классификатора

## Итоговые метрика каждой модели

На изображениях 2.3 и 2.4 показаны графики точности моделей от эпох обучения:

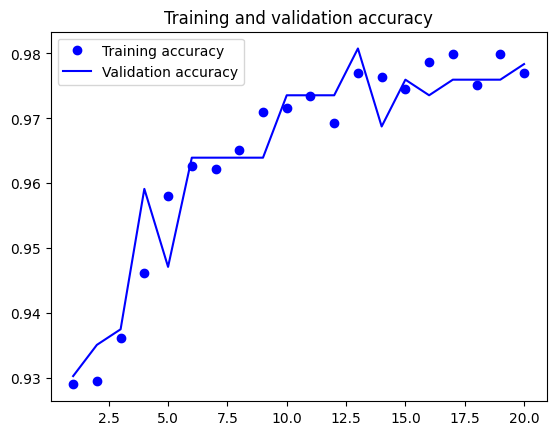


Рисунок 2.3 – График точности модели от количества эпох при обучении модели бинарного классификатора

Изображение выглядит как текст, снимок экрана, линия, График

Автоматически созданное описание

Рисунок 2.4 – График точности модели от количества эпох при обучении модели многоклассового классификатора

В таблице 2.1 собраны метрики всех моделей полученные с помощью библиотеки sklearn.metrics.

Таблица 2.1 – Метрики качества моделей

|  |  |  |
| --- | --- | --- |
| Метрика\Модель | Бинарный классификатор | Многоклассовый классификатор |
| accuracy | 0.9769 | 0.9568 |
| categorical\_crossentropy | 0.0658 | 0.1568 |

# Вывод

В ходе выполнения лабораторной работы были созданы и обучены нейронных сетей для классификации изображений, а также переноса обучения.

Дообучение бинарного и многоклассового классификатора производилось на базе сети VGG16, которая показала высокую точность определения поврежденного болезнью картофеля, при этом судя по метрикам из таблицы 2.1 бинарный классификатор показывает лучшую точность, чем многоклассовый.

Листинг программы находится в соответствующей папке Lab6 по ссылке на репозиторий: https://github.com/peremichka256/NSiIP