**Курсовой проект**

«Распознавание морских животных на изображениях»

по дисциплине

«Глубокое обучение»

|  |  |
| --- | --- |
| Выполнили студенты гр. 5130904/00101 | Кривенко М.Д.  Прудников Д. А. |
| Руководитель  ст. преподаватель | Малеев О.Г. |

Оглавление

[Описание задачи 3](#_Toc162184529)

[Ход работы 4](#_Toc162184530)

[Сверточная нейронная сеть 4](#_Toc162184531)

[Предобученная модель 10](#_Toc162184532)

[Вывод 13](#_Toc162184533)

# **Описание задачи**

Задача заключается в создании, обучении и тестировании нейронной сети, которая должна быть способна распознавать морских животных на изображениях.

В работе использован датасет морских животных https://www.kaggle.com/datasets/vencerlanz09/sea-animals-image-dataste. Примеры изображений из сета:

A spotted stingray swimming in the ocean

Description automatically generated

Изображение выглядит как млекопитающее, морская выдра, выдра, Канадская выдра

Автоматически созданное описание

# Ход работы

## Сверточная нейронная сеть

Первым шагом является загрузка датасета и разделение датасета на train, val и test датасеты:

A black background with white text

Description automatically generated

Соотношение между сетами примерно равно 0.8 : 0.1 : 0.1 на train, val и test сеты соответственно:

A computer screen shot of white text

Description automatically generated

Получилось три сета данных суммарно на примерно 3000 изображений:

A black background with white text

Description automatically generated

Для предотвращения переобучения используется аугментация данных при помощи ImageDataGenerator. Изображение случайным образом может быть повернуто в пределах 10 градусов, приближено в пределах 10%, сдвинуто по ширине или высоте в пределах 20%, а также отражено горизонтально:

A screen shot of a computer program

Description automatically generated

Далее создается модель Sequential для создания сети с последовательным соединением слоев друг за другом :

A computer screen shot of text

Description automatically generated

Первым слоем идет сверточный слой Conv2D с 32 фильтрами размером 3x3 и функцией активации ReLU. Так как датасеты подготовлены с параметрами (256, 256, 3), для слоя указаны соответствующие параметры входных данных. Значение same параметра padding в данном случае приводит к сохранению размерности данных при проходе через сверточный слой.

Conv2D используется для применения свертки входных данных. Conv2D позволяет извлекать признаки из изображений, что делает их эффективными для обработки изображений и любых других двумерных данных. Conv2D принимает на вход двумерные данные, такие как изображения, и применяет к ним ядро свертки. Это позволяет находить более высокоуровневые признаки, например границы объектов, что способствует распознаванию объектов.

Из активаторов выбран именно ReLU из-за его относительной простоты в вычислении и при этом достаточной эффективности, что будет проверено на этапе тестирования модели.

После Conv2D находится слой MaxPooling2D для уменьшения размера изображения, что с точки зрения сложности вычислений позволяет использовать следующими слои Conv2D с более большим количеством фильтров.

Аналогично описанная выше комбинация слоев дублируется несколько раз с постепенным увеличением количества фильтров в слое Conv2D до 256.

Затем следует слой Dropout. Аргументом является количество случайно отключаемых нейронов. Засчет случайного отключения нейронов данный слой помогает избежать чрезмерной зависимости от конкретных признаков входных данных, таким образом способствуя развитию возможности модели обрабатывать новые данные

Затем добавляется слой Flatten, преобразующий входные данные в одномерный массив для обработки следующими слоями.

Последующие два слоя Dense сначала с 1024 нейронами и функцией активации ReLU, а затем с количеством нейронов, аналогичным количеству распознаваемых классов, и функцией активации softmax используются для создания предсказания о принадлежности входных данных к классу. Softmax используется для интерпретации результата как распределения вероятности о принадлежности изображения к определяемым классам.

В результате получается данная модель:

A screenshot of a computer program

Description automatically generated

За созданием модели следует ее компиляция и обучение с использованием подготовленных ранее датасетов train и val в течение 30 эпох:

A screenshot of a computer

Description automatically generated

A screenshot of a computer program

Description automatically generated

Следует отметить, что точность на тренировочном датасете несильно превышает точность на валидационном датасете в течение всего обучения, а в некоторых эпохах точность на валидационном датасете даже превышает точность на тренировочном датасете, например в 26 эпохе. Это показывает достаточность предпринятых против переобучения мер.

Процесс обучения описывается следующими графиками точности и потерь:

A graph with red and green lines

Description automatically generated

A graph with red and blue lines

Description automatically generated

Финальная оценка точности модели производится с использованием тестового датасета:

A screenshot of a computer

Description automatically generated

В результате полученной cnn моделью достигнута точность 77.5%. Данный результат можно превзойти с помощью предобученных моделей.

## Предобученная модель

Для данной работы выбрана модель ResNet50V2. Предобучение на датасете imagenet из более чем 14 миллионов подписанных изображений должно положительно сказаться на способности нейронной сети классифицировать изображения морских животных.

Рассмотрим процесс создания модели:

A screen shot of a computer screen

Description automatically generated

Сначала загружается предобученная модель ResNet50V2 с указанием весов, полученных при обучении на датасете imagenet. В встроенном внутрь модели слой Dense нет необходимости, так как в дальнейшем мы используем свои слои Dense под специфику поставленной задачи, поэтому аргумент include\_top передается со значением False. Параметры размеры изображения остаются неизменными из процесса создания сверточной сети. Также стоит отметить отключение обучения предобученной модели – в процессе обучения будут обучаться только необученные слои Dense. Пересечение процессов обучения предобученных слоев ResNet50V2 и необученных слоев может привести к ухудшению эффективности работы предобученных слоев.

В итоговой модели после ResNet50V2 следует слой GlobalAveragePooling2D, необходимый для конвертации формата выходных данных из предобученной модели в двумерный.

Затем добавляются слой Dropout для предотвращения переобучения и слои Dense для непосредственно предсказания о пренадлежности изображения к классу. Предназначение данных слоев аналогично уже рассмотренным выше в процессе создания сверточной нейронной сети.

Итоговая модель выглядит следующим образом:

A screenshot of a computer program

Description automatically generated

После создания модель проходит обучение в течение 10 эпох:

A screenshot of a computer

Description automatically generated

Оценим эффективность работы модели на тестовом датасете:

A screenshot of a computer

Description automatically generated

Точность на тестовом датасете составила 93.5%, что значительно превышает результат полученной ранее сверточной нейронной сети без использования предобученной модели.

# Вывод

В данной работе была решена задача классификации морских животных с использованием двух способов. Первым способом является сверточная нейронная сеть без использования предобученной модели. Данный вариант смог достичь точности в 77.5%. Вторым способом является сеть с использованием предобученной модели. Этот способ оказался более эффективным и за меньшее количество эпох смог достичь точности в 93.5%. Также на обучение второй сети понадобилось меньше времени с учетом использования тех же вычислительных ресурсов, что и для первой сети. Таким образом можно сделать вывод, что использование предобученной модели в задаче классификации морских животных является более эффективным подходом.