

Домашнее задание. Нейронные сети

Датасет

В рамках этой домашней задачи мы разработаем модель, которая определит, изображен ли на фотографии динозавр или дракон. Для этого мы будем использовать датасет "Dino or Dragon?", который можно скачать с [Google Drive](#).

На лекциях мы рассмотрели, как использовать предварительно обученную нейронную сеть. В домашнем задании мы обучим гораздо меньшую модель с нуля.

Примечание: Для выполнения этой домашней работы вам потребуется среда с GPU. Рекомендую использовать Google Colab. Вы также можете использовать компьютер без GPU (например, ваш ноутбук), но это будет медленнее.

Подготовка данных

Датасет включает в себя примерно по 1900 изображений динозавров и драконов. Внутри датасета предусмотрены отдельные папки для обучающих и тестовых выборок.

Модель

В рамках данной домашней работы мы применим сверточную нейронную сеть (CNN). Как и во время лекций, для реализации мы будем использовать библиотеку Keras (Tensorflow).

Необходимо разработать модель со следующей структурой:

- Входной слой размером (150, 150, 3)
- Создайте сверточный слой (Conv2D):
 - 32 фильтра
 - Размер ядра (kernel) (3, 3) - размерность фильтра
 - Активация ReLu
- Уменьшите размер карты признаков с помощью max pooling (MaxPooling2D) размером (2, 2)
- Преобразуйте многомерный результат в векторы с помощью слоя Flatten
- Добавьте плотный слой (Dense) с 64 нейронами и активацией ReLu
- Создайте выходной плотный слой с 1 нейроном
- Для выходного слоя используйте активацию, подходящую для бинарной классификации
- В качестве оптимизатора используйте SGD со следующими параметрами:
 - SGD(lr=0.002, momentum=0.8)

Вопрос 1

Так как перед нами стоит задача бинарной классификации, какую функцию потерь лучше всего применить в нашем случае?

- бинарная кросс-энтропия (binary crossentropy)
- фокусные потери (focal loss)
- среднеквадратичная ошибка (mean squared error)
- категориальная кросс-энтропия (categorical crossentropy)

Примечание: Если вы используете активацию для выходного слоя, не требуется устанавливать `from_logits=True`.

Вопрос 2

Определите общее количество параметров в модели. Для этого примените метод `summary`.

- 9215873
- 11215873
- 14215873
- 19215873

Генераторы и обучение

Для следующих двух вопросов используйте следующий генератор данных для обучающих (train) и тестовых (test) наборов:

```
ImageDataGenerator(rescale=1./255)
```

Примечание: Дополнительная предобработка изображений не требуется. При загрузке данных из каталогов обучения/тестирования убедитесь, что параметр `class_mode` установлен правильно для задачи бинарной классификации. Рекомендуемые параметры: `batch_size=20` и `shuffle=True`.

Для обучения примените метод `.fit()` со следующими параметрами:

```
model.fit( train_generator, epochs=10, validation_data=test_generator)
```

Вопрос 3

Какова медиана точности обучения по всем эпохам?

- 0.40
- 0.60
- 0.90
- 0.20

Вопрос 4

Каково стандартное отклонение потерь в процессе обучения по всем эпохам?

- 0.11
- 0.66
- 0.99
- 0.33

Аугментация данных

Для следующего этапа вам потребуется генерировать больше данных с помощью аугментаций.

Добавьте следующие аугментации к генератору обучающих данных:

```
rotation_range=40,  
width_shift_range=0.2,  
height_shift_range=0.2,  
shear_range=0.2,  
zoom_range=0.2,  
horizontal_flip=True,  
fill_mode='nearest'
```

Вопрос 5

Обучите модель еще на 10 эпох с использованием указанного выше кода. Не создавайте модель с нуля; продолжите обучение существующей.

Каково среднее значение потерь на тестовом наборе данных по всем эпохам после аугментации?


- 0.15
- 0.77
- 0.37
- 0.97


Вопрос 6


Каково среднее значение точности на тестовом наборе данных за последние 5 эпох (с 6 по 10) после аугментации?

- 0.84
- 0.54
- 0.44
- 0.24

Отправка результатов

 Формат отправки: Результаты должны быть представлены в виде ссылки на ваш Git репозиторий.

 Ответы на вопросы: Ответы на вопросы должны быть оформлены в одном из следующих форматов: .txt, .pdf или .doc.

 **Важное примечание:** Пожалуйста, не добавляйте сами данные в ваш Git репозиторий. Вместо этого убедитесь, что ваш репозиторий содержит только необходимые файлы с кодом, документацией и ответами на вопросы.

Перед отправкой убедитесь, что вы правильно оформили все материалы и следовали указанным инструкциям. Удачи!

Сроки