

МИНОБРНАУКИ РОССИИ

Федеральное государственное бюджетное образовательное учреждение высшего образования

«МИРЭА– Российский технологический университет»

# РТУ МИРЭА

Институт кибербезопасности и цифровых технологий

КБ-4 «Интеллектуальные системы информационной безопасности»

**Отчет полабораторной работе №3**

по дисциплине: «Анализ защищенности систем искусственного

интеллекта»

Выполнил:

Студент группы ББМО-01-23

Бакин Д.И.

2024

**Установка библиотек**



Рисунок 1 – Импорт библиотек

**Загрузим модель**

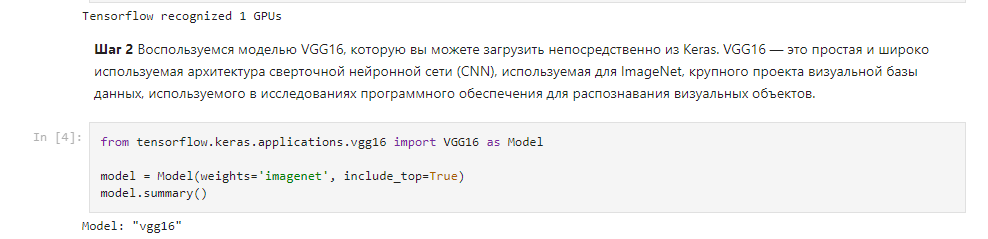


Рисунок 2 – Загрузка модели

**Загрузка и выполнение предобработки 4 изображений**

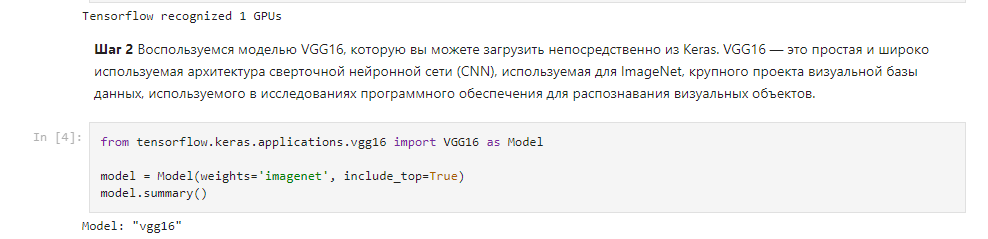
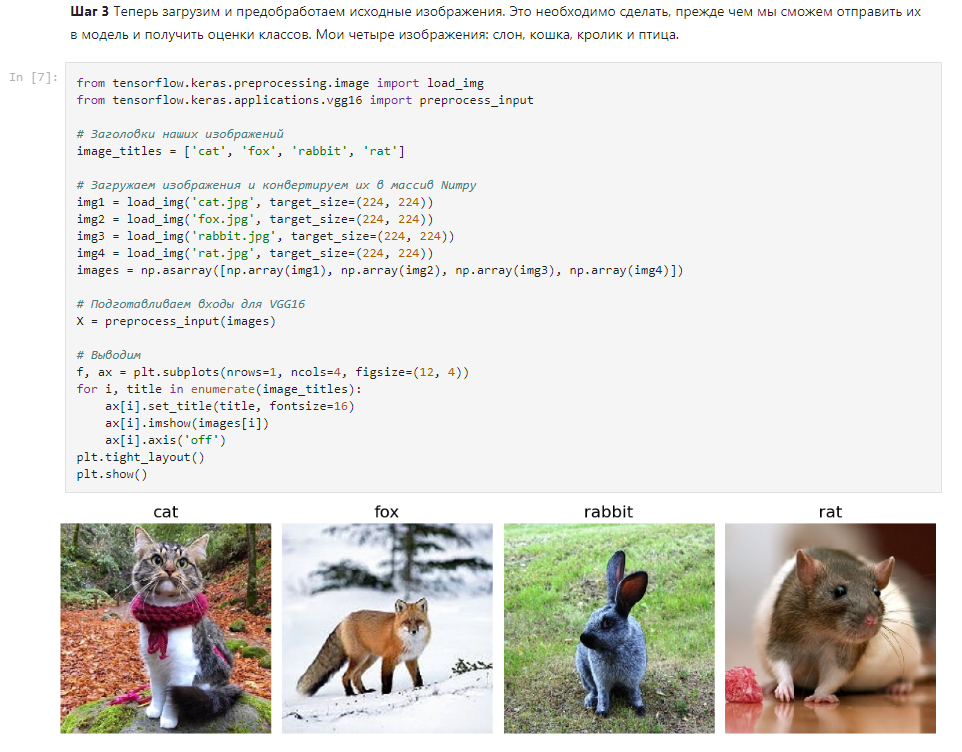


Рисунок 3 –Загрузка и выполнение предобработки 4 изображений

**Заменим функцию на линейную функцию, а также создадим функцию по подсчету очков соответствия каждого изображения определённой группы**



Рисунок 4 –Заменим функцию на линейную функцию, а также создадим функцию по подсчету очков соответствия каждого изображения определённой группы

**Сгенерируем карту внимания и подсветим области наибольшего внимания**



Рисунок 5 –Сгенерируем карту внимания и подсветим области наибольшего внимания

**Уменьшим шум для карт влияния, и выведем результат.**

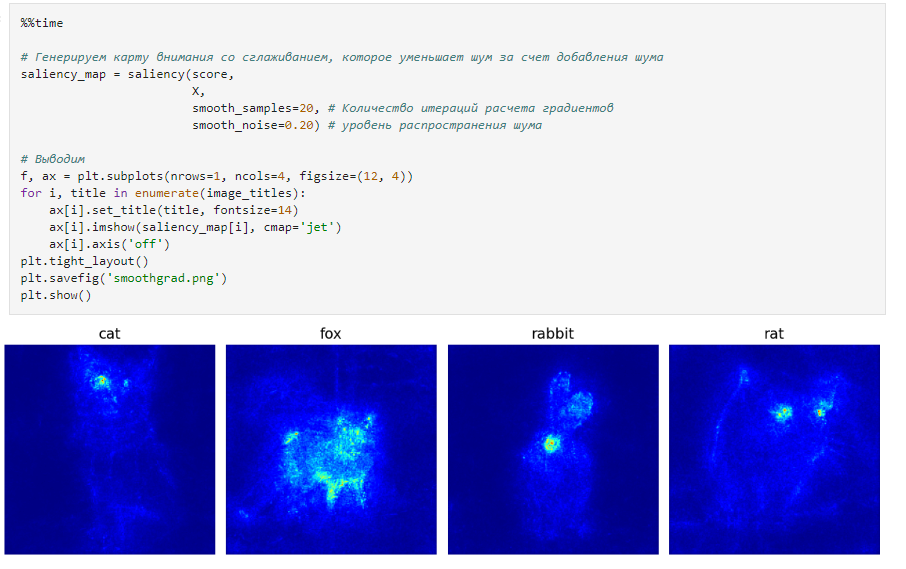


Рисунок 6 –Уменьшим шум для карт влияния, и выведем результат.

**Сравним значения с GradCam**

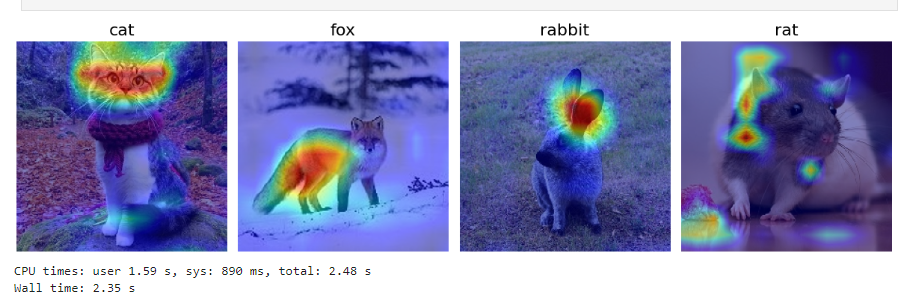


Рисунок 7 –Сравним значения с GradCam

**Сравним значения с GradCam+**



Рисунок 8 –Сравним значения с GradCam

**Вывод**

В результате выполнения работы изучили процесс построения карт

внимания в нейронных сетях для анализа изображений, с помощью которого

можно подробнее понять работу обученной модели по категоризации

изображения.

В ходе работы были выполнены следующие шаги:

 Замена функции активации softmax на линейную для корректного

вычисления градиентов;

 С помощью метода SmoothGrad, уменьшили шумы на карте, что

позволило получить более точные результаты карт внимания;

 Построение карт значимости классов для выбранных

изображений разными методами.

После завершения экспериментов и проведения анализа полученных

результатов было выявлено, что метод GragCAM++ выделяет более полные и

точные контура животных, чем GragCAM.