

МИНОБРНАУКИ РОССИИ

Федеральное государственное бюджетное образовательное учреждение высшего образования

**«МИРЭА – Российский технологический университет»**

# РТУ МИРЭА



ИКБ направление «Киберразведка и противодействие угрозам с применением технологий искусственного интеллекта» 10.04.01

Кафедра КБ-4 «Интеллектуальные системы информационной

безопасности»

# Лабораторная работа №3

## По дисциплине

“Анализ защищенности системы информационной

## безопасности”

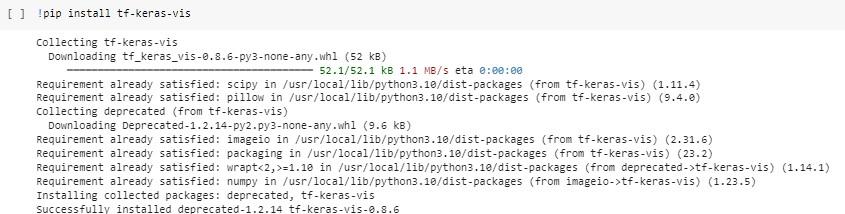
Группа: ББМО-01-22

Выполнил: Челышев Д.И.

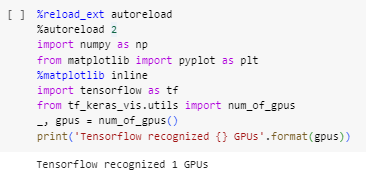
Проверил: Спирин А.А.

Москва 2023

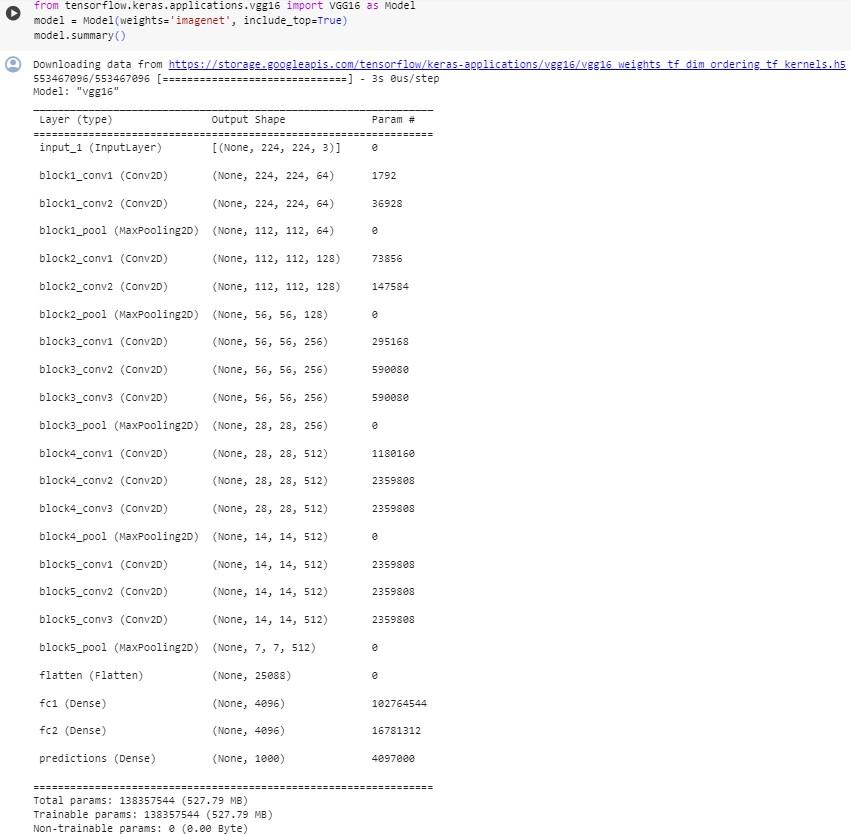
Для начала установим инструмент для визуализации для TensorFlow Keras.



Далее подключим необходимые библиотеки, а также активируем перезагрузку при изменении кода, чтобы не нужно было явно выполнять перезагрузку



Загрузим предварительно обученную модель VGG16, на ImageNet датасете. После чего отобразим сводку по модели.

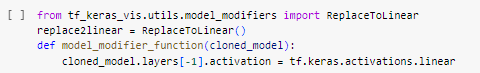


Загрузим несколько изображений датасета ImageNet и выполним их предварительную обработку перед использованием.

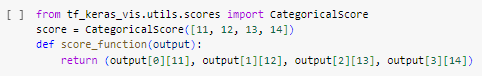
Отобразим на одном графическом представлении все наши изображения.



Заменим функцию активации на линейную



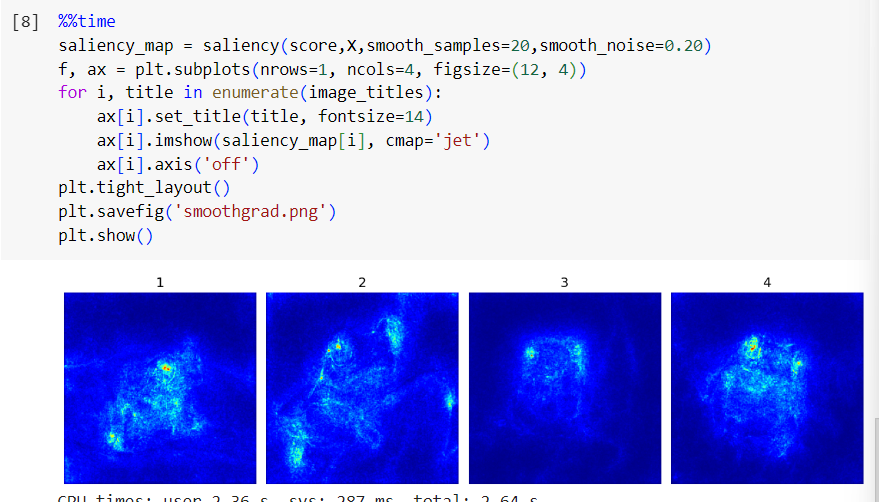
Создадим функцию очков соответствия каждому изображению



Создадим карту внимания (vanilla)



Уменьшим шум для карт влияния



Используем GrandCAM



Используем GradCAM++



**Вывод**

Таким образом, использование методов визуализации Grad-CAM, Grad- CAM++, Saliency, SmoothGrad и подобных, может быть полезным для понимания того, какие части изображений были наиболее важными при принятии решений моделью машинного обучения. Методы Grad-CAM и Grad- CAM++ позволяют визуализировать активации в различных частях изображений, позволяя понять, где модель фокусируется при определении

классов. Карты активаций могут помочь исследовать, какие объекты или части изображения были ключевыми для принятия решения моделью. Использование метода SmoothGrad может помочь снизить шум и сделать карты сайленси более интерпретируемыми