|  |  |  |
| --- | --- | --- |
|  |  |  |
| МИНОБРНАУКИ РОССИИ | | |
| Федеральное государственное бюджетное образовательное учреждение  высшего образования  **«МИРЭА– Российский технологический университет»**  **РТУ МИРЭА** | | |
| Институт кибербезопасности и цифровых технологий  Кафедра КБ-4 «Интеллектуальные системы информационной безопасности» | | |

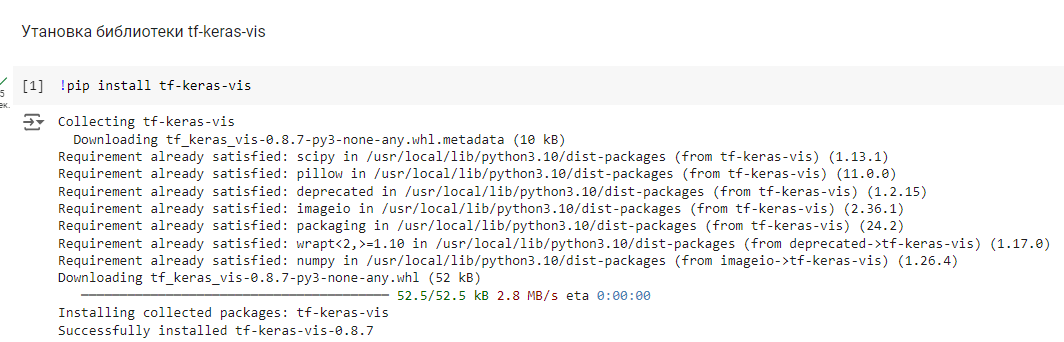
Лабораторная работа 3

по дисциплине «Анализ защищенности систем искусственного интеллекта»

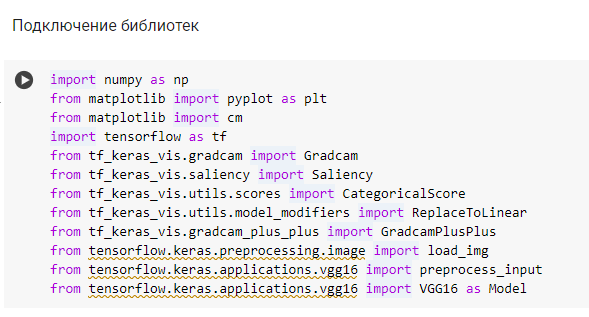
Выполнил

Журавлев Александр Дмитриевич

Устанавливаем библиотеку tf-keras-vls



Подключаем остальные необходимые библиотеки

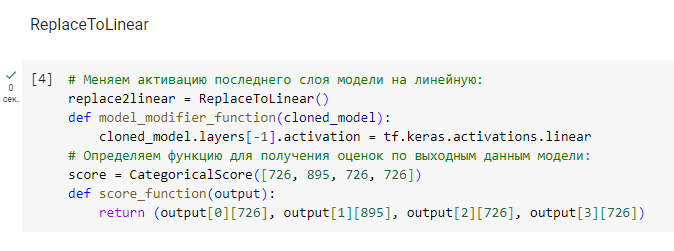


Загружаем четыре изображения для получения карты значимости различных областей изображения

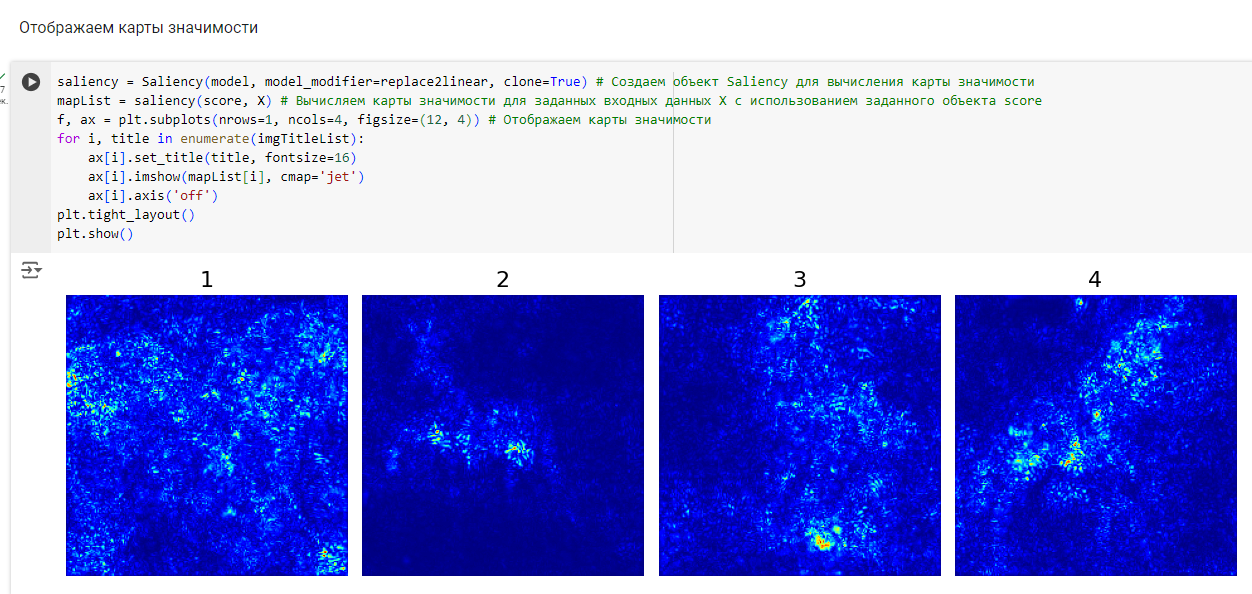




Поменяем функцию активации последнего слоя на линейную, зададим метки классов выбранным изображениям

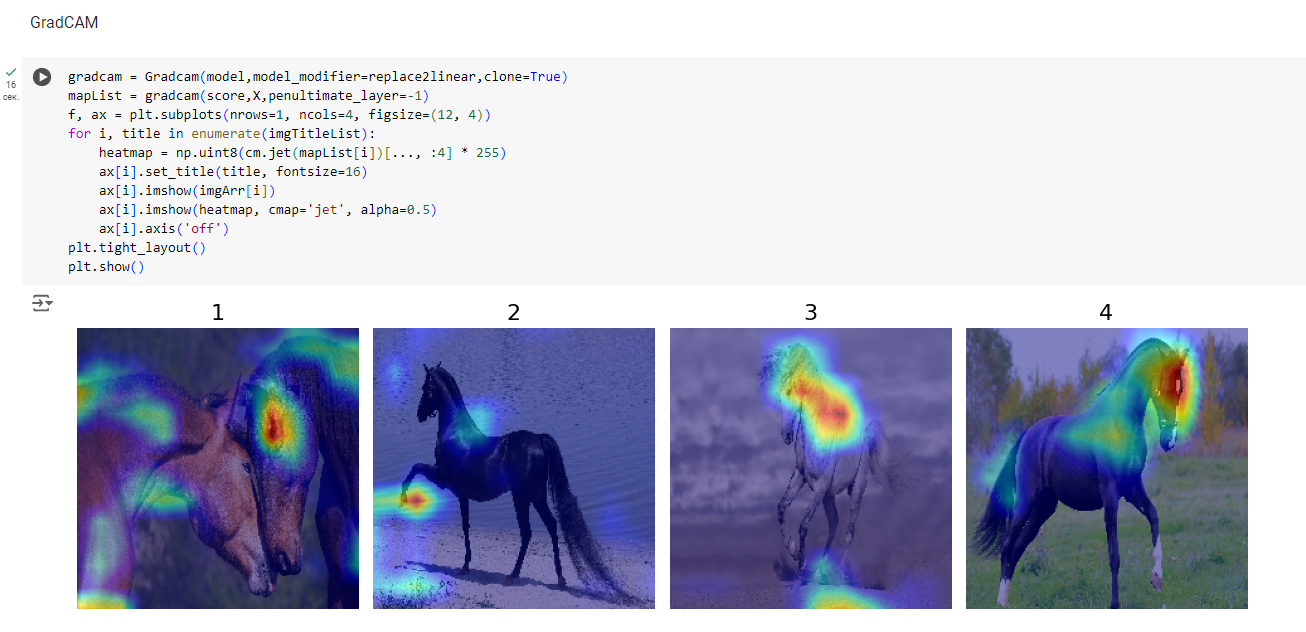


Отобразим карты значимости



Это позволяет вам визуализировать, какие части изображений влияют на вывод модели, тем самым предоставляя более глубокое понимание её работы.

Используем GrandCAM



Это позволяет визуализировать влияние различных областей изображений на предсказания модели

Используем GrandCAM++



Это позволяет понять какие области входных данных были наиболее значимыми для предсказаний модели

Данные методы играют ключевую роль в интерпретации и визуализации работы нейронных сетей.