|  |
| --- |
|  |
| МИНОБРНаУКИ РОССИИ |
| Федеральное государственное бюджетное образовательное учреждение  высшего образования  **«МИРЭА – Российский технологический университет»**  **РТУ МИРЭА** |

Институт комплексной безопасности и специального приборостроения

**Отчет** по лабораторной работе №3

**по дисциплине:** «Анализ защищенности систем искусственного интеллекта»

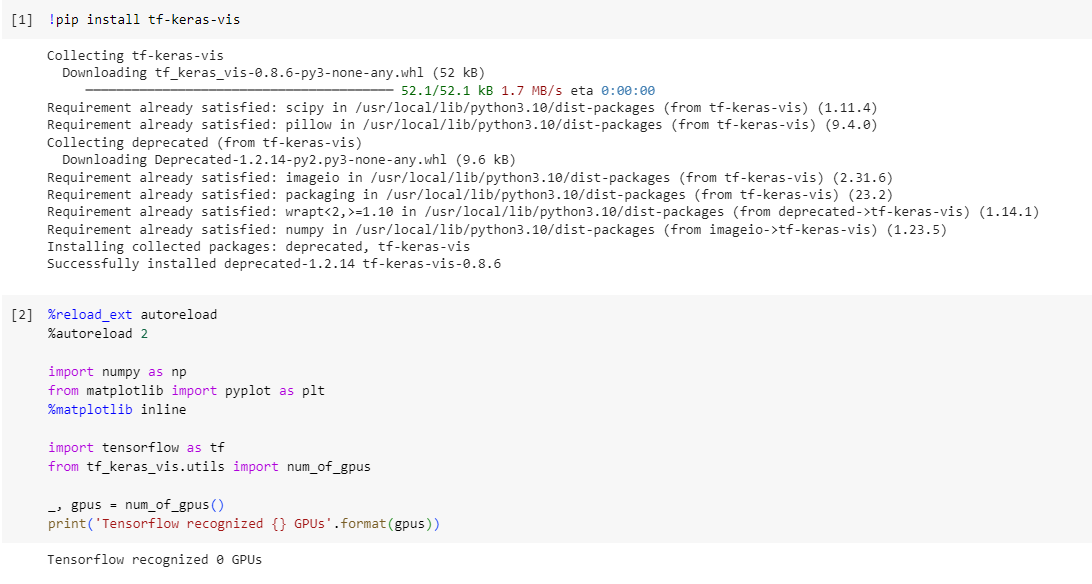
**Выполнил:**

Студент группы ББМО-01-22

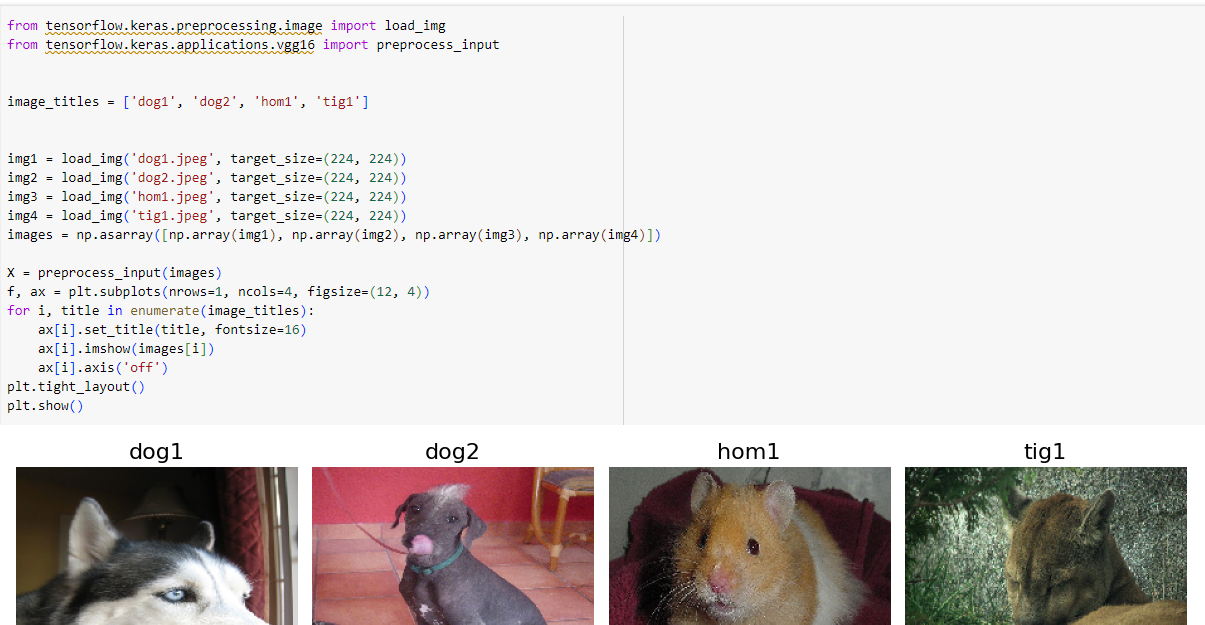
ФИО: Карпова Н.Ю.

**Москва 2023**

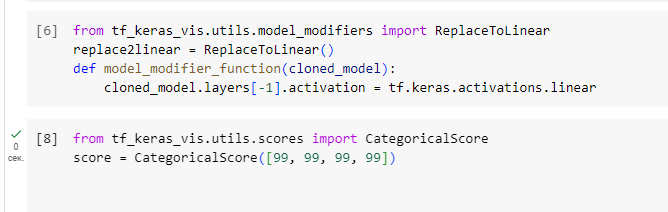
1. Установим tf-keras-vis.



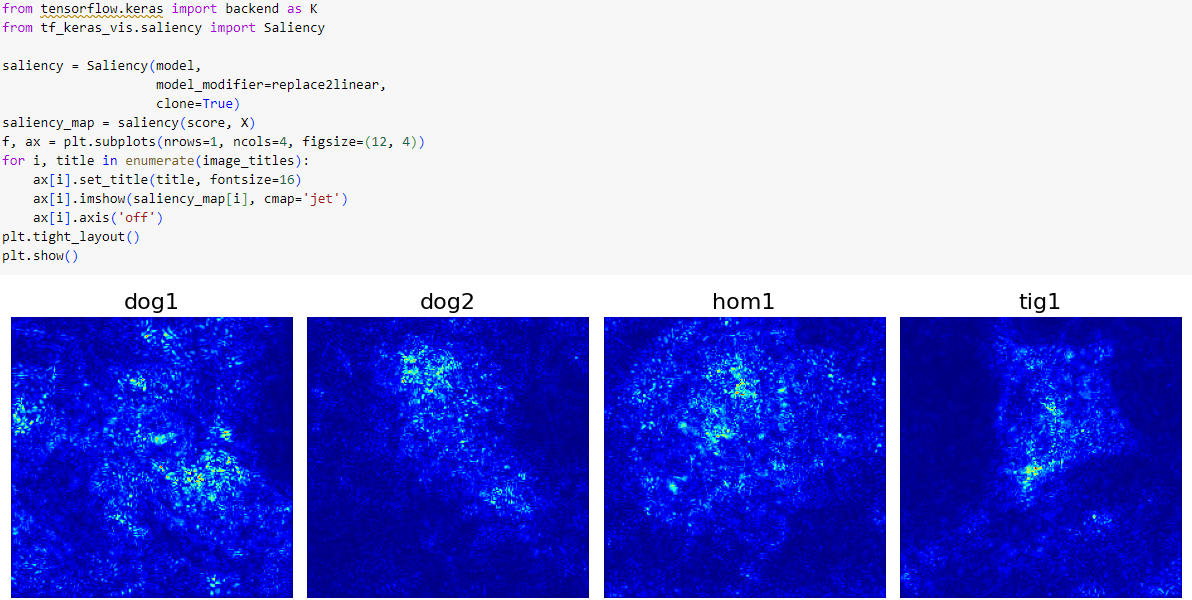
1. Загрузим 4 изображения из датасета ImageNet и отобразим их.



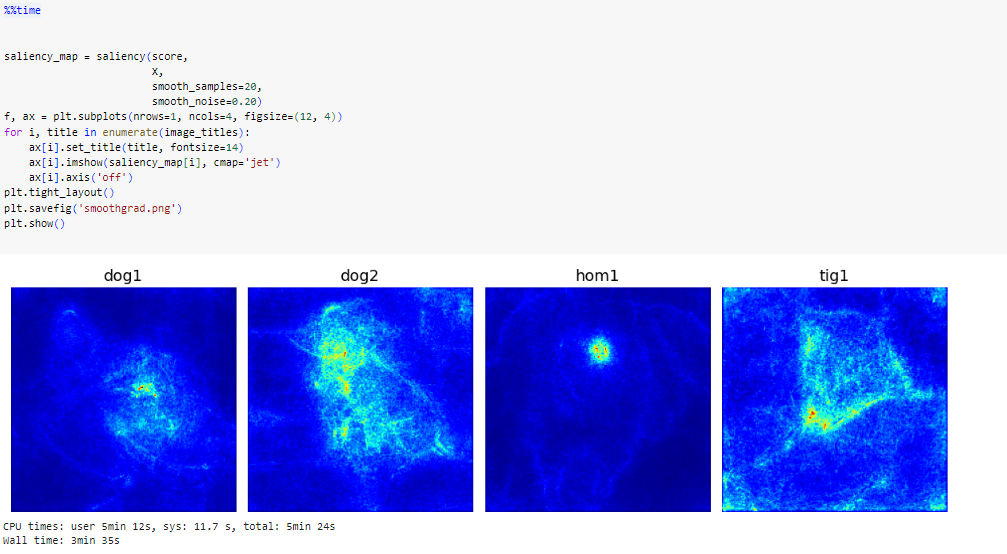
1. Сделаем класс изображения, для примера сделаем все 99.



1. Благодаря Saliency сгенерируем карту внимания, подсвечивая входного изображения.



1. Удалим шум благодаря методу SmoothGrad



1. Попробуем способ генерации внимания GradCAM, которое проявляется в тех областях входного изображения, которые в наибольшей степени способствуют выходному значению



1. Попробуем также GradCAM++, который может предоставить прогнозы модели CNN

