

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 №1**

по дисциплине: «Анализ защищенности систем искусственного интеллекта»

Группа:

ББМО-01-22

Выполнил:

Гребенник Г.С

Проверил:

к.т.н. Спирин А.А.

Москва, 2023

**Содержание**

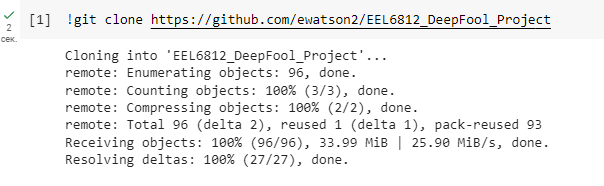
[**Цель работы:** 3](#_Toc157423867)

# **Цель работы:**

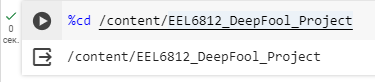
Целью данной работы является отразить отличия для fgsm\_eps=(0.001, 0.02, 0.5, 0.9, 10) и выявить закономерность/обнаружить отсутсвие влияние параметра eps для сетей FC LeNet на датасете MNIST, NiN LeNEt на датасете CIFAR.

# **Ход работы:**

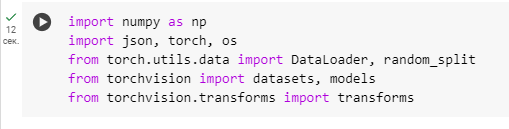
1. Скопировать проект по ссылке в локальную среду выполнения Jupyter (Google Colab): <https://github.com/ewatson2/EEL6812_DeepFool_Project>



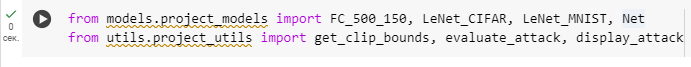
1. Сменим директорию исполнения на вновь созданную папку "EEL6812\_DeepFool\_Project" проекта:



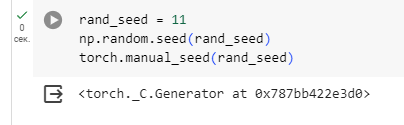
1. Импортируем библиотеки:



1. Импортируем вспомогательные библиотеки из локальных файлов проекта:



1. Установим случайное значение в виде переменной rand\_seed=(порядковый номер в списке гугл-таблицы 11) и установим его для np.random.seed и torch.manual\_seed:



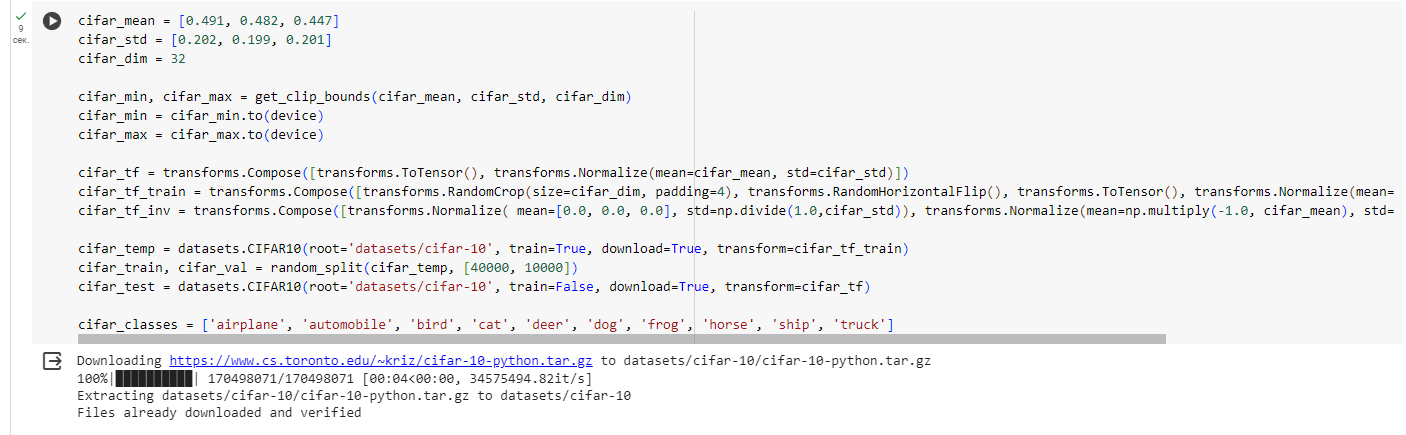
1. Используем в качестсве устройства видеокарту (Среды выполнения--> Сменить среду выполнения --> T4 GPU)



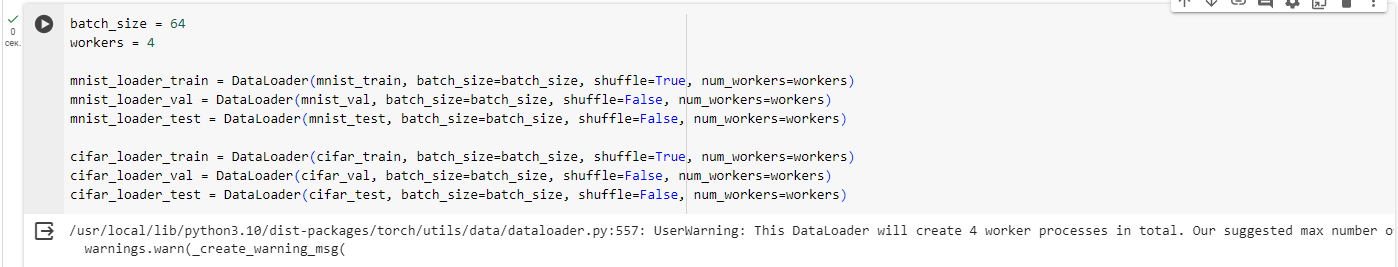
1. Загрузим датасет MNIST c параметрами:



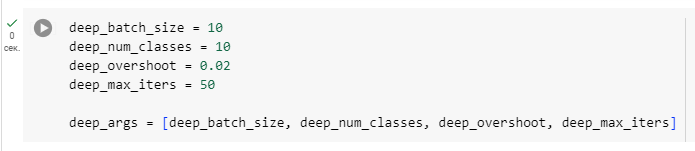
1. Загрузим датасет CIFAR-10 с параметрами:



1. Выполним настройку гиперпараметров и загрузку DataLoader:



1. Зададим параметры deep\_args:



1. Загрузим и оценим стойкость модели Network-In-Network Model к FGSM и DeepFool атакам на основе датасета CIFAR-10:

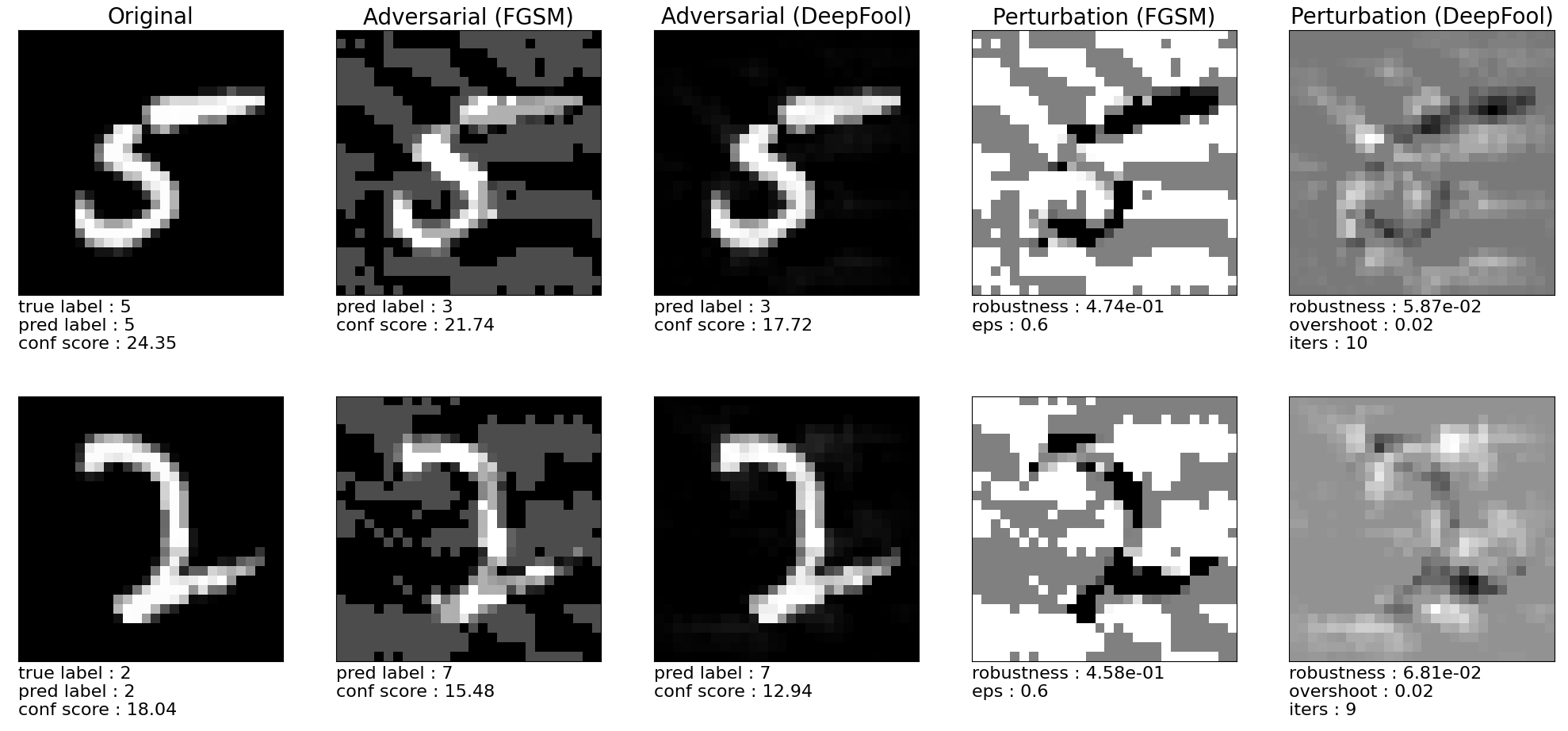


1. Загрузим и оценим стойкость модели LeNet к FGSM и DeepFool атакам на основе датасета CIFAR-10:

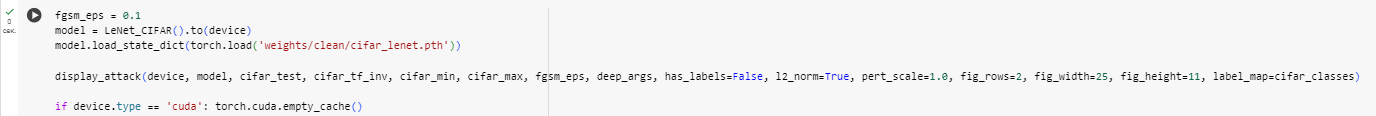


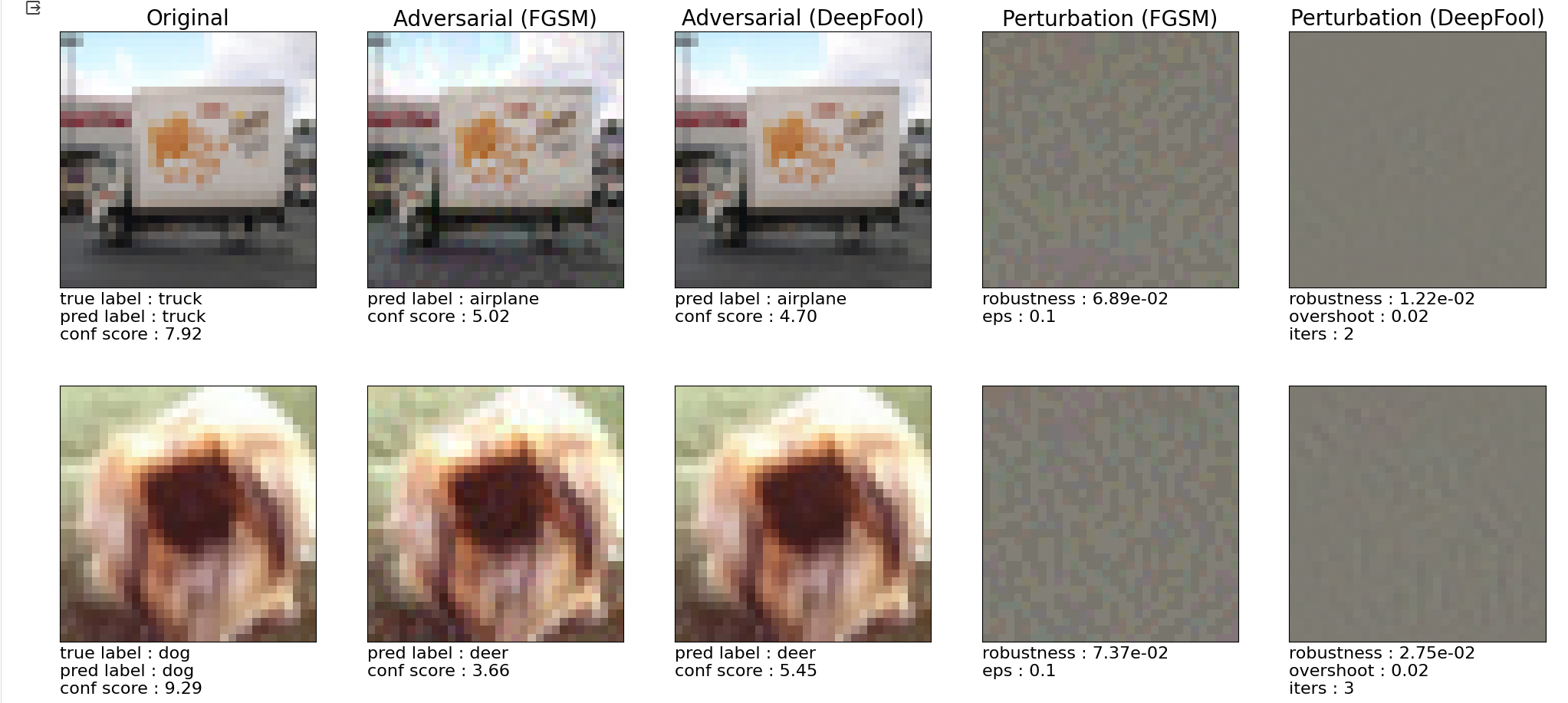
1. Выполним оценку атакующих примеров для сетей:
2. LeNet на датасет MNIST:



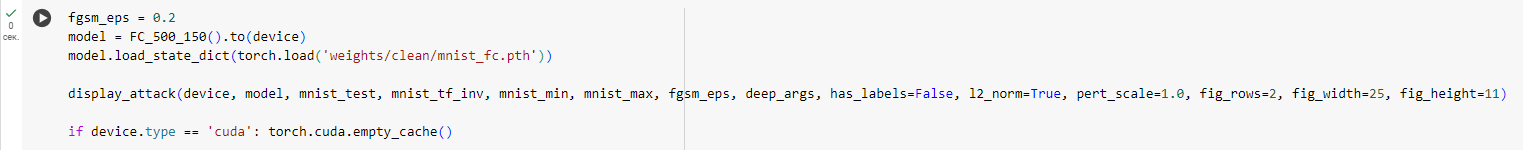


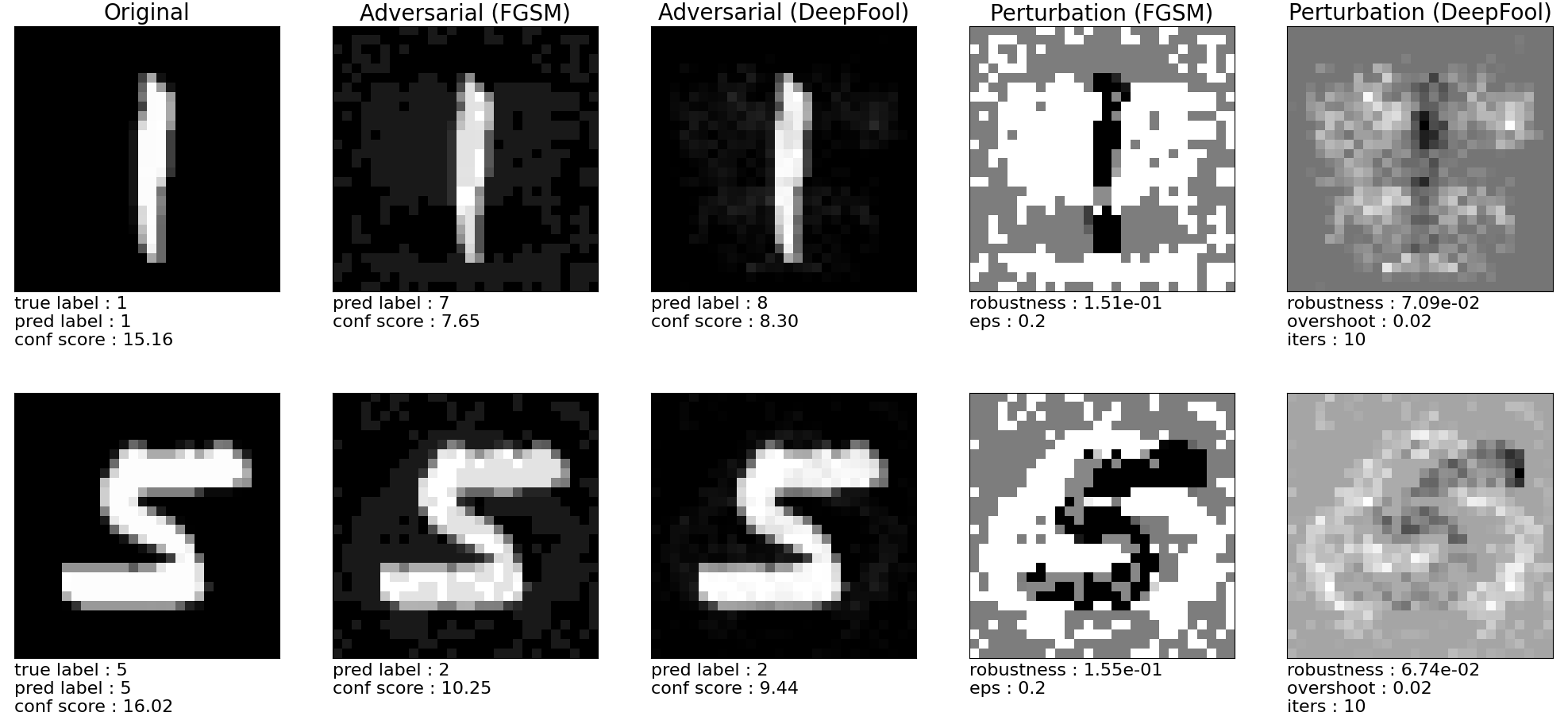
1. LeNet на датасете CIFAR-10:





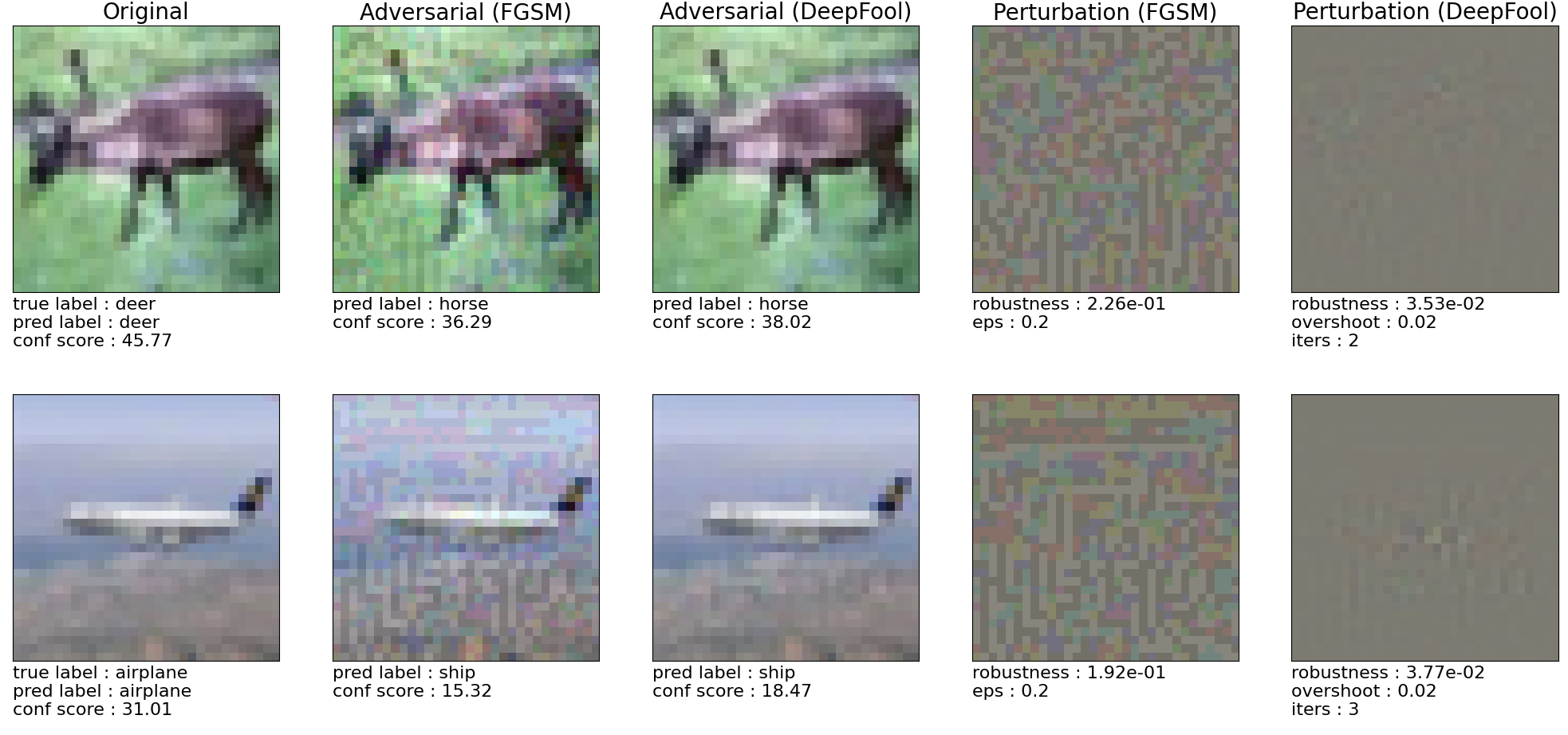
1. FCNet на датасете MNIST:





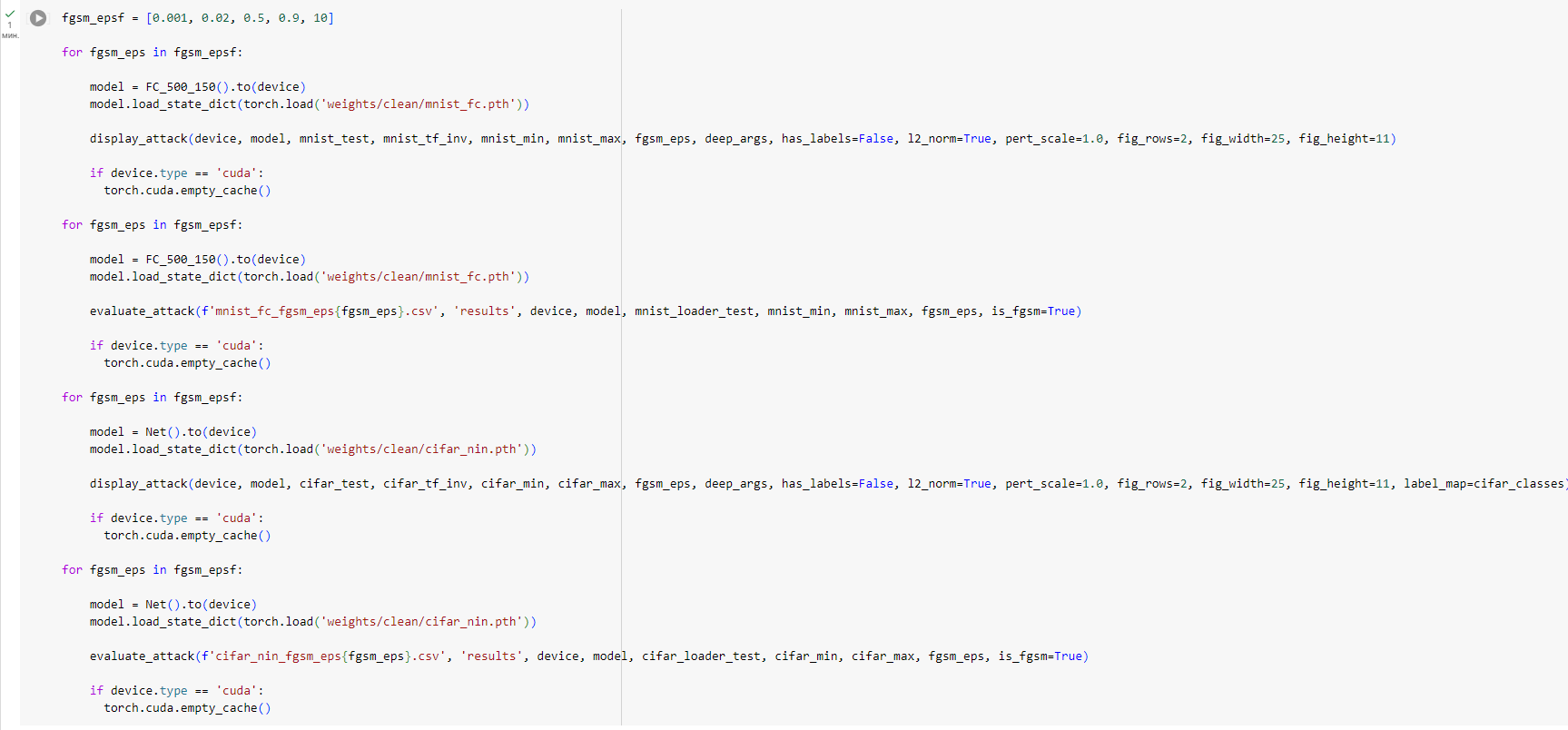
1. Network-in-Network на датасете CIFAR:



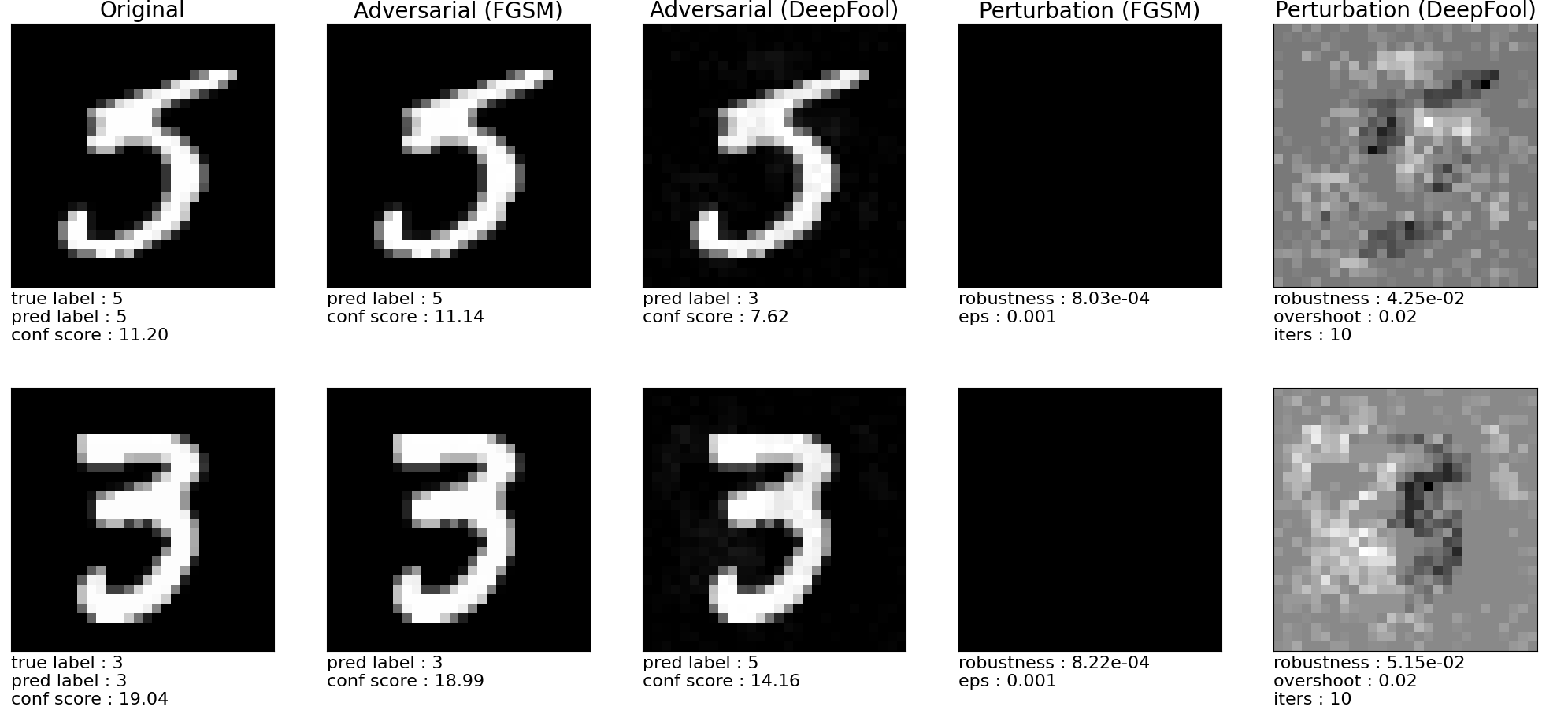


1. Отразим отличия для fgsm\_eps=(0.001, 0.02, 0.5, 0.9, 10) и выявим закономерности/обнаружим отсутсвие влияние параметра eps для сетей FC LeNet на датасете MNIST, NiN LeNEt на датасете CIFAR:

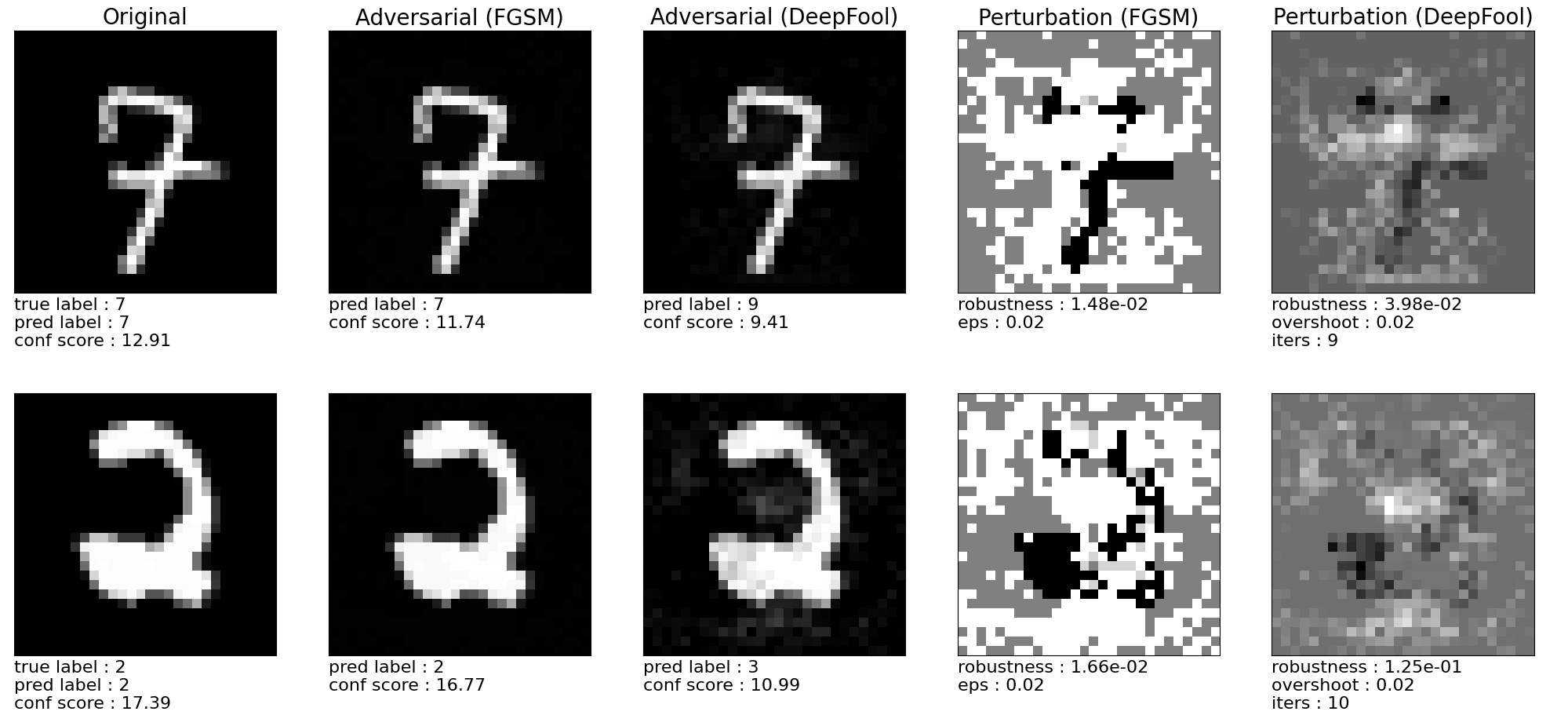
Для выполнения данной задачи реализуем просто цикл который будет перебирать уже известный нам массив значений:



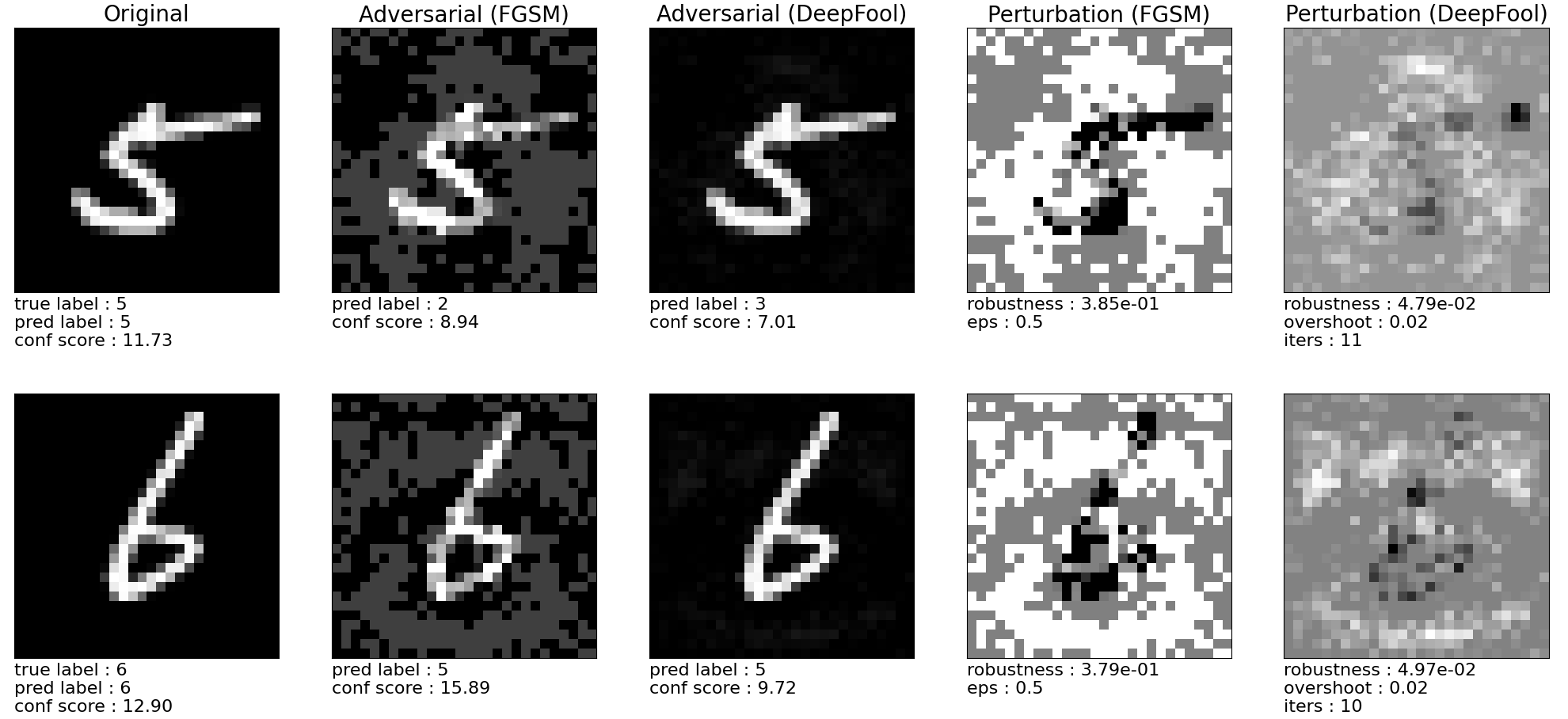
1. FC на MNIST при Esp 0.001



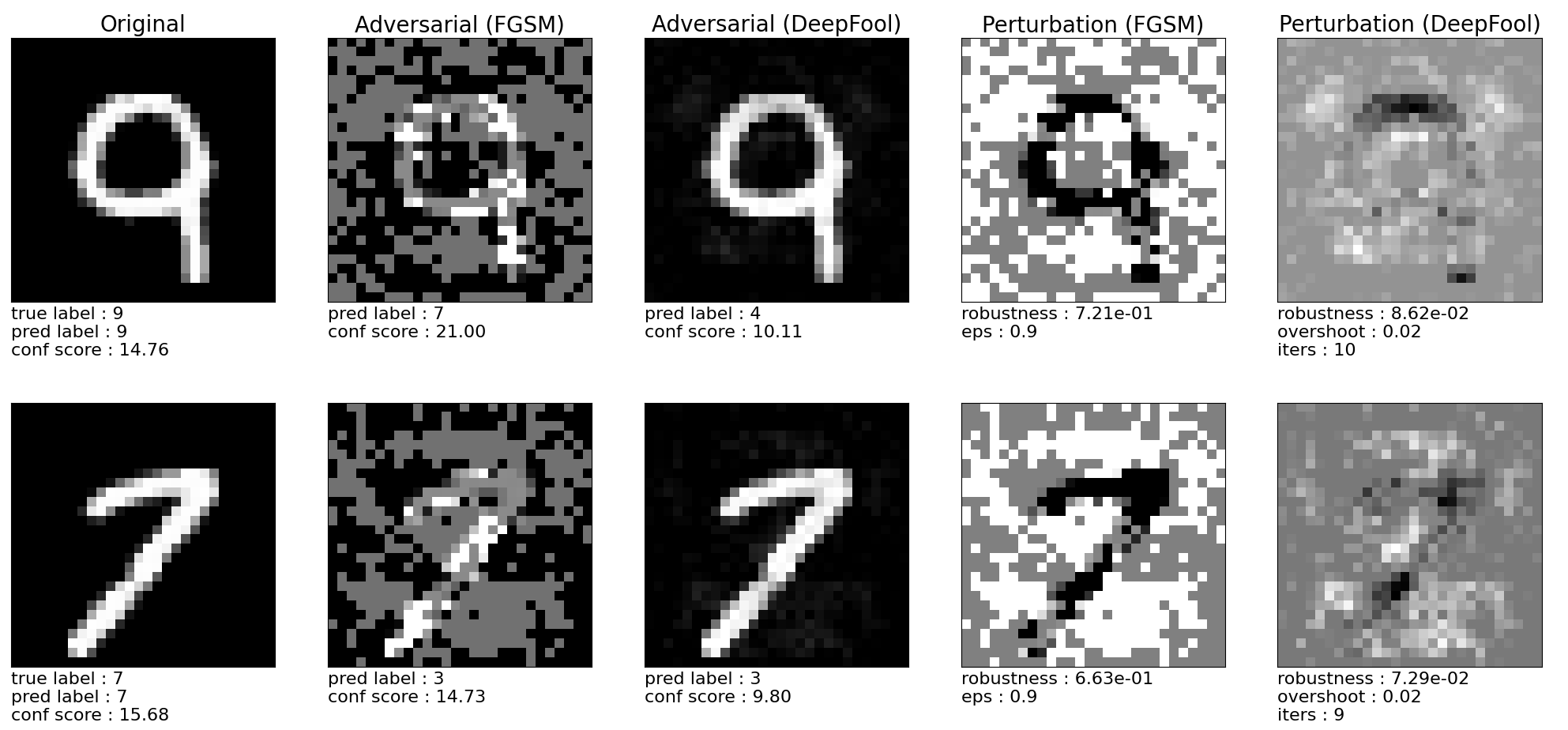
1. FC на MNIST при Esp 0.02



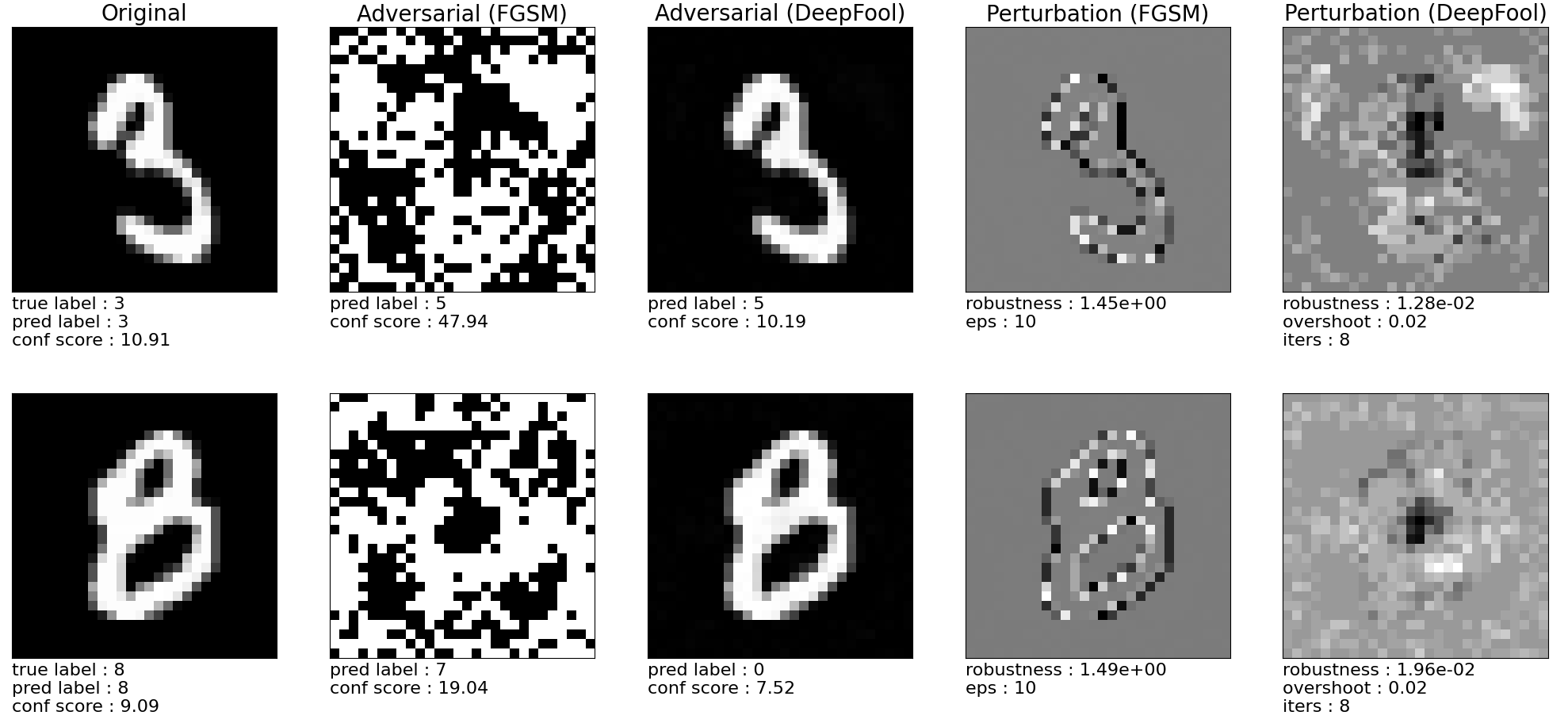
1. FC на MNIST при Esp 0.5

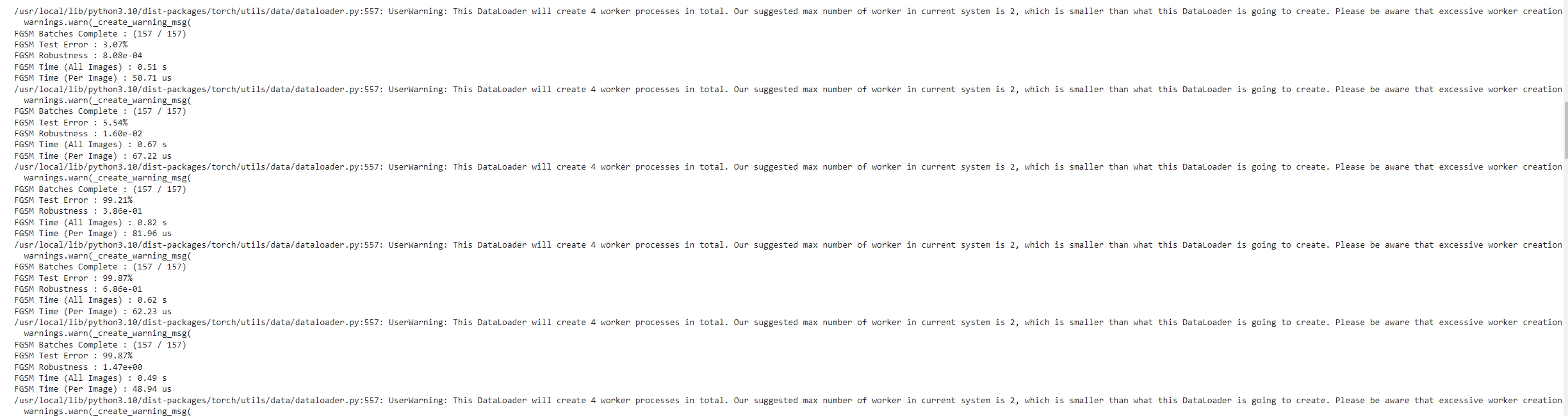


1. FC на MNIST при Esp 0.9



1. FC на MNIST при Esp 10

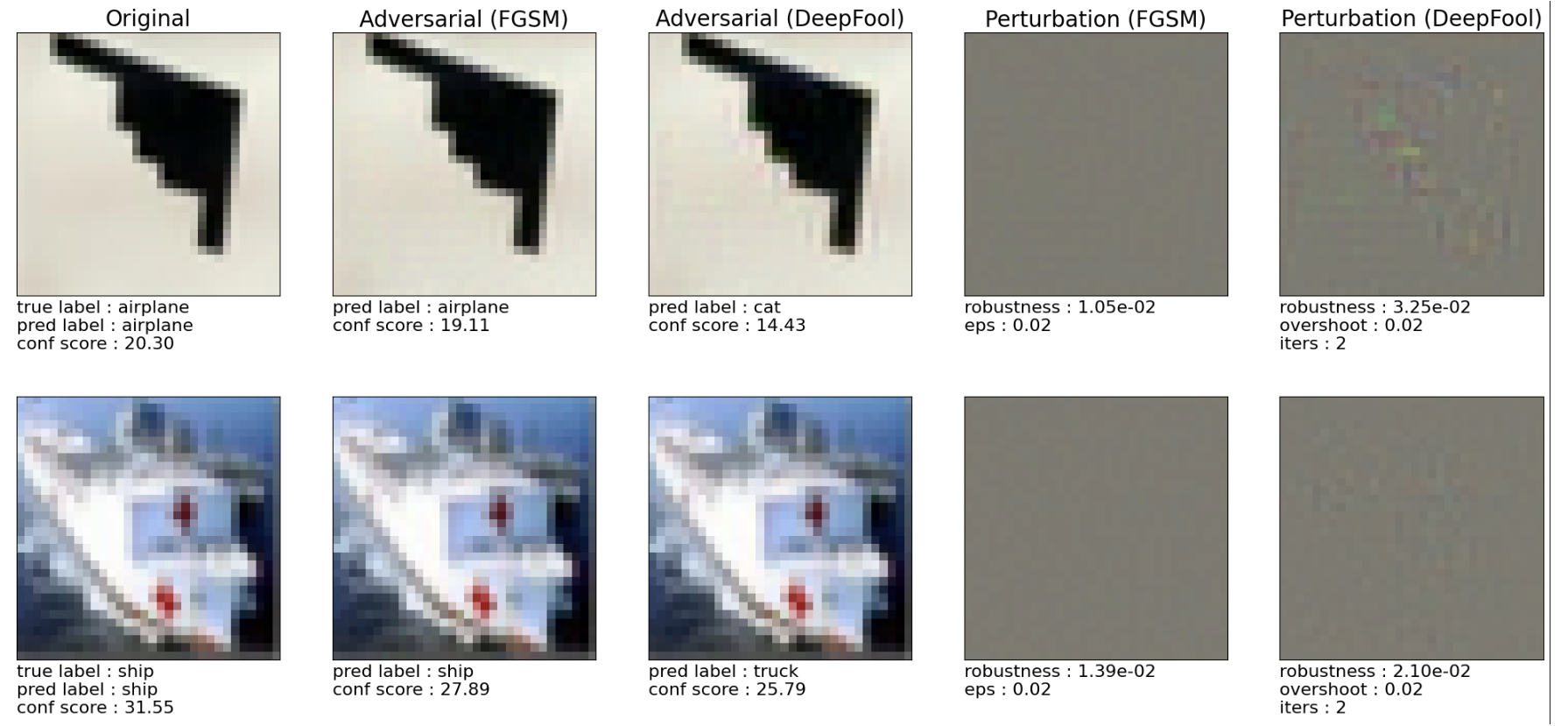




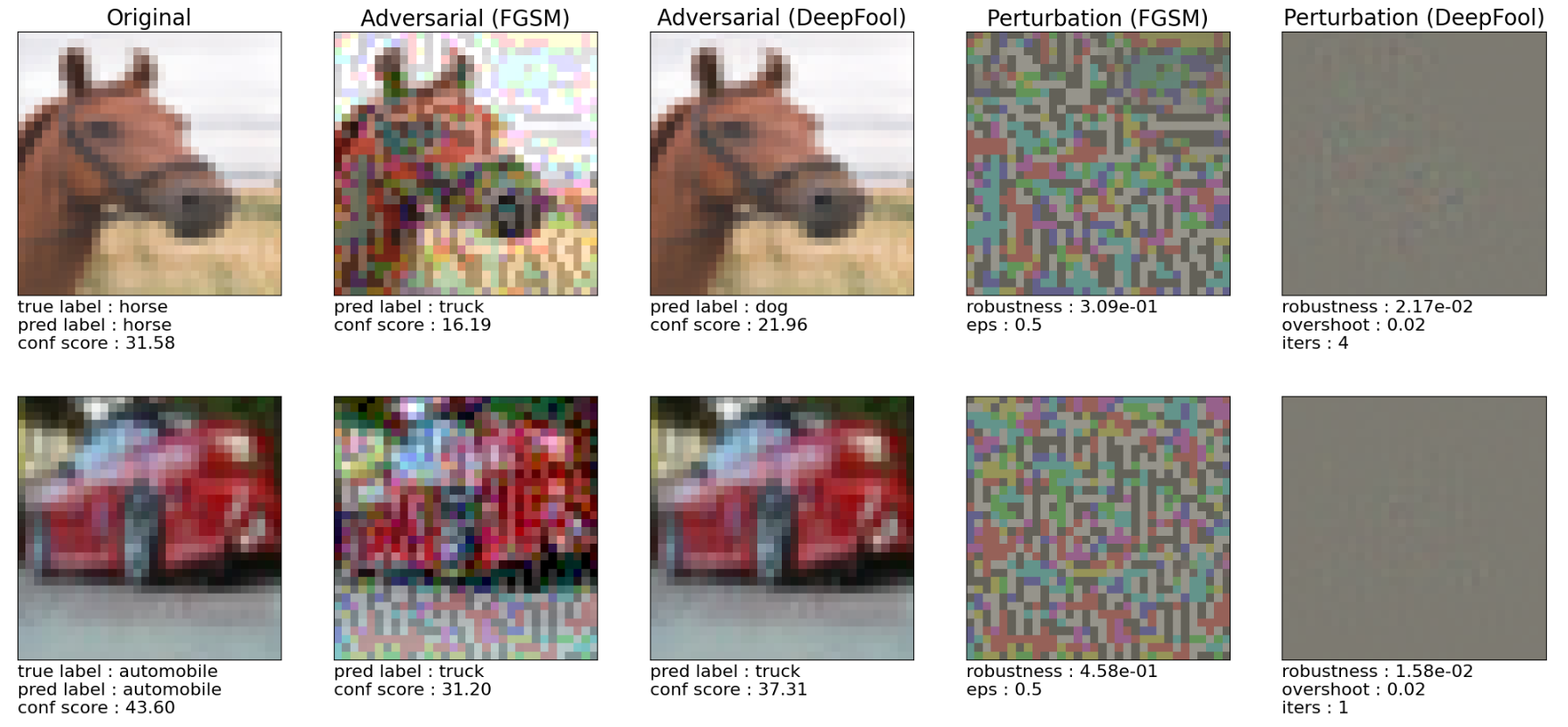
1. Esp 0.001



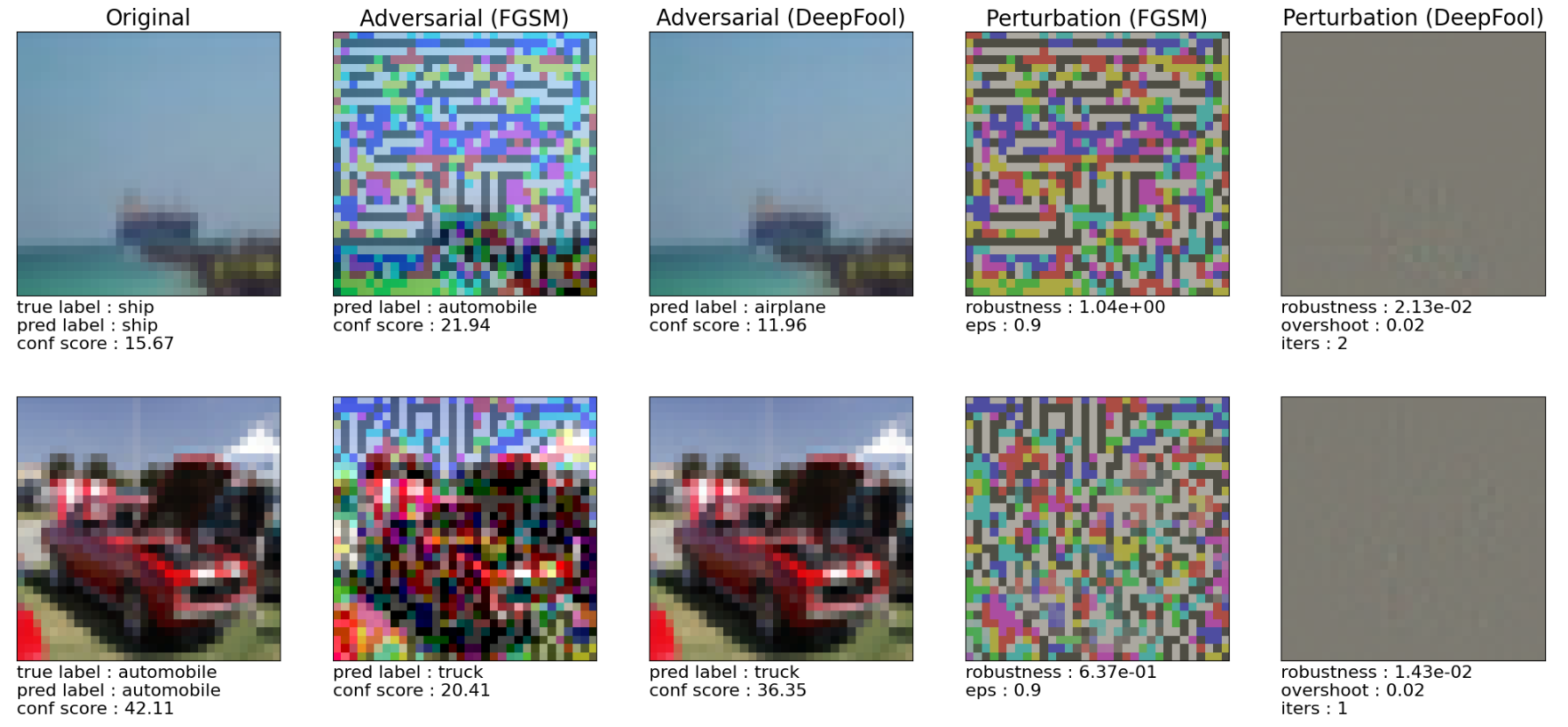
1. Esp 0.02



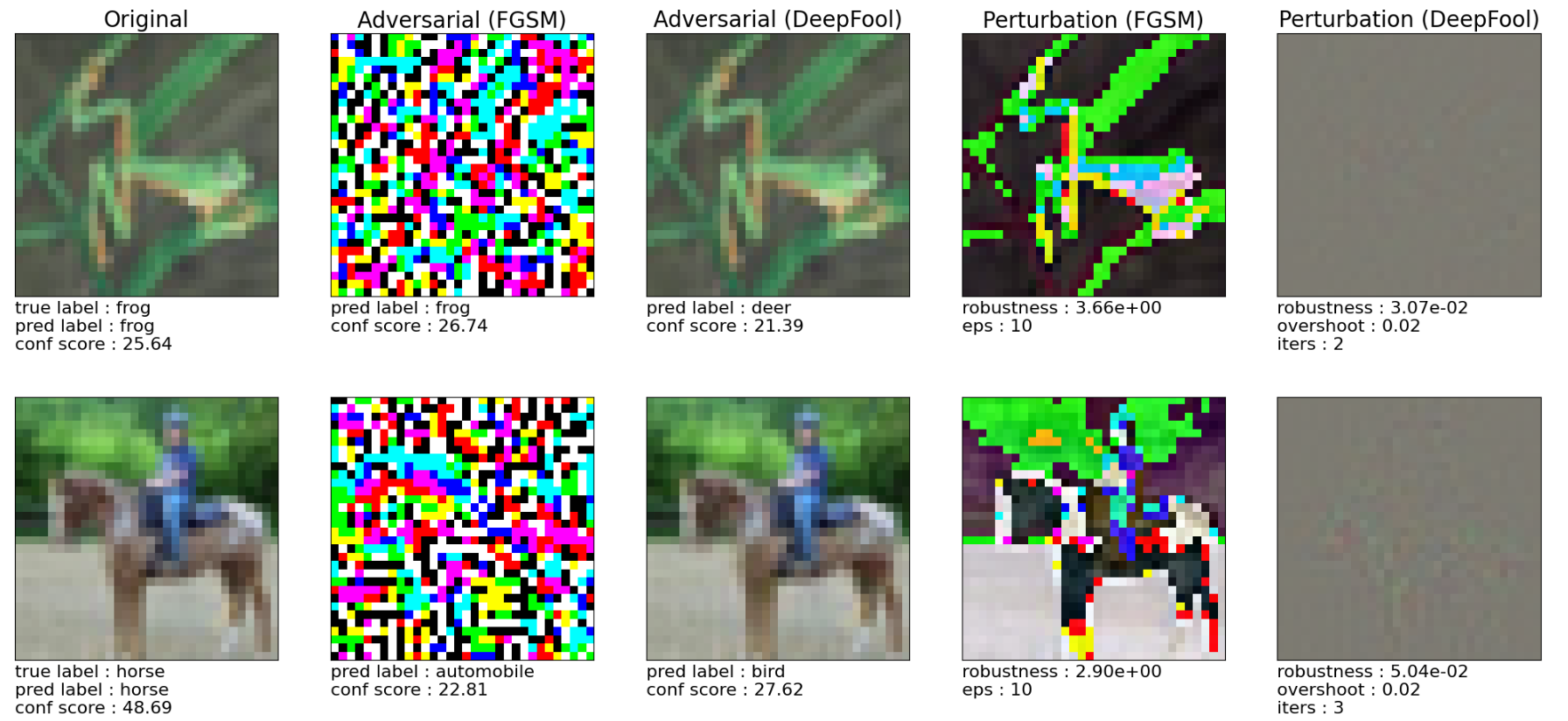
1. Esp 0.5

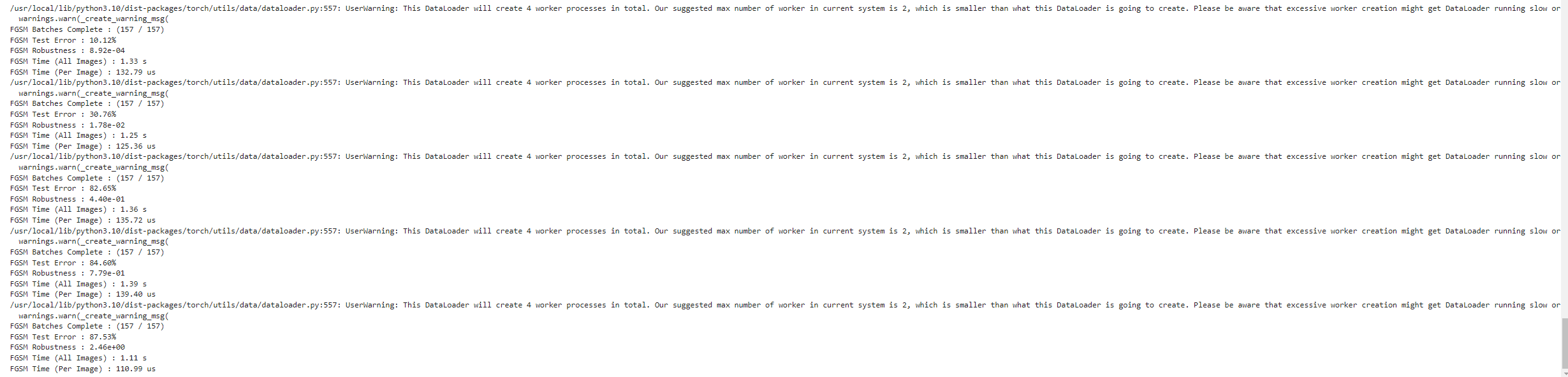


1. Esp 0.9



1. Esp 10





# **Заключение**

При рассмотрении результатов эксперимента, проведенного в рамках данной лабораторий работы, была выявлена закономерность, обозначающая, что при увеличении значения eps, сети становятся более уязвимыми и допускают больше ошибок классификации, нежели при низком значении eps.