

БЕЛОРУССКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ

Факультет прикладной математики и информатики

Кафедра биомедицинской информатики

Аннотация к дипломной работе

**«Исследование взаимосвязей между параметрами
состязательных атак и ошибок классификации
изображений нейронными сетями»**

Войнов Дмитрий Михайлович

Научный руководитель – кандидат технических наук, доцент Ковалев В. А.

Минск, 2019

Реферат

Дипломная работа, 42 страницы, 17 рисунков, 1 таблица, 16 формул, 19 источников.

ГИСТОЛОГИЧЕСКИЕ ИЗОБРАЖЕНИЯ, РЕНТГЕНОВСКИЕ ИЗОБРАЖЕНИЯ, ГЛУБОКИЕ НЕЙРОННЫЕ СЕТИ, СОСТЯЗАТЕЛЬНЫЕ АТАКИ, СОСТЯЗАТЕЛЬНЫЕ ПРИМЕРЫ.

Объект исследования – состязательные атаки, алгоритмы генерации состязательных примеров.

Цель работы – исследование зависимости влияния состязательных атак на глубокие нейронные сети от параметров алгоритмов генерации состязательных примеров и типов медицинских изображений.

В ходе работы изучено явление состязательных атак, рассмотрены существующие алгоритмы генерации состязательных примеров. Постановлено и проведено экспериментальное исследование, выявляющее ключевые зависимости устойчивости глубоких нейронных сетей от главных параметров алгоритмов генерации. Обнаружены зависимости от исходной вероятности изображения и модальности медицинских изображений.

Областью применения является обучение устойчивых и безопасных моделей глубоких нейронных сетей.

Abstract

Diploma thesis, 42 pages, 17 figures, 1 tables, 16 formulas, 19 sources.

HISTOLOGY IMAGES, X-RAY IMAGES, DEEP NEURAL NETWORKS, ADVERSARIAL ATTACKS, ADVERSARIAL EXAMPLES.

Object of research – adversarial attacks, algorithms of crafting adversarial examples.

Objective – studying the dependence of impact of adversarial attacks to deep neural networks on parameters of algorithms of crafting adversarial examples and biomedical image type.

The phenomenon of adversarial attacks was studied as well as existing algorithms of crafting of adversarial examples. The experimental investigation that detects key patterns of vulnerability of deep neural networks on the main parameters of crafting algorithms was built and carried out. Some dependencies on original image probability and biomedical image type were recognized.

The scope is training of robust and secure models of deep neural networks.