БЕЛОРУССКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ

Факультет прикладной математики и информатики

Кафедра дискретной математики и алгоритмики

Аннотация к магистерской диссертации

«Проблема устойчивости глубоких нейронных сетей на примере задач анализа биомедицинских изображений»

Войнов Дмитрий Михайлович

Научный руководитель – кандидат технических наук, заведующий лабораторией анализа биомедицинских изображений ОИПИ НАН Республики Беларусь Ковалев В. А.

Минск, 2021

Реферат

Магистерская диссертация, 50 страниц, 10 рисунков, 1 таблица, 14 источников.

Ключевые слова: ГЛУБОКИЕ НЕЙРОННЫЕ СЕТИ, СОСТЯЗАТЕЛЬНЫЕ АТАКИ, АЛГОРИТМЫ ГЕНЕРАЦИИ АТАКУЮЩИХ ИЗОБРАЖЕНИЙ, АНАЛИЗ БИОМЕДИЦИНСКИХ ИЗОБРАЖЕНИЙ.

Объектом исследования являются глубокие классификационные нейронные сети.

Предметом исследования является устойчивость глубоких нейронных сетей при их применении в задачах анализа биомедицинских изображений.

Целью работы было постановлено исследовать влияние состязательных атак на нейронные сети при решении различных задач классификации биомедицинских изображений; определить зависимость эффективности проводимых атак от режима их проведения, выбранного алгоритма генерации атакующих изображений, значений контрольных параметров.

В ходе работы была разработана методика экспериментального исследования, показывающего необходимые характеристики состязательных атак. В результате проведения этого исследования были определены искомые зависимости и построены соответствующие графики.

Полученный результат можно использовать в разработке безопасных систем автоматического и полуавтоматического диагностирования заболеваний, основанных на анализе изображений.

Abstract

Master thesis, 50 pages, 10 figures, 1 table, 14 resources.

Keywords: DEEP NEURAL NETWORKS, ADVERSARIAL ATTACKS, ALGORITHMS OF GENERATING ADVERSARIAL EXAMPLES, BIOMEDICAL IMAGE ANALYSIS.

The object of research is deep neural networks.

The subject of study is robustness of deep neural networks in the biomedical image analysis domain.

The aim of this work is to investigate the influence of adversarial attacks on neural networks in scope of different biomedical image classification problems, and to determine dependence of attacks effectiveness on its mode, algorithm of generating adversarial examples, values of control parameters.

The methodology of experimental research showing necessary characteristics of adversarial attacks was developed during the study. As a result of this research, the desired dependencies were determined and related plots were made.

The result can be applied to development of secure and safe computerized systems of automatic or semi-automatic disease diagnostics based on image analysis.