**МИНИСТЕРСТВО ОБРАЗОВАНИЯ И НАУКИ
ДОНЕЦКОЙ НАРОДНОЙ РЕСПУБЛИКИ
ГОСУДАРСТВЕННОЕ ОБРАЗОВАТЕЛЬНОЕ УЧРЕЖДЕНИЕ
ВЫСШЕГО ПРОФЕССИОНАЛЬНОГО ОБРАЗОВАНИЯ
«ДОНЕЦКИЙ НАЦИОНАЛЬНЫЙ УНИВЕРСИТЕТ»**

На правах рукописи

Бабичева Маргарита Вадимовна

АВТОМАТИЗИРОВАННЫЕ СИСТЕМЫ НАУЧНЫХ ИССЛЕДОВАНИЙ УГРОЗ БЕЗОПАСНОСТИ ЛИЧНОСТИ

Специальность 2.3.3. «Автоматизация и управление технологическими процессами и производствами» (технические науки)

АВТОРЕФЕРАТ
диссертации на соискание ученой степени
кандидата технических наук

Донецк – 2023

Работа выполнена в ГОУ ВПО «ДОННУ» Министерства образования и науки Донецкой Народной Республики

Научный руководитель: доктор технических наук, профессор **Данилов Владимир Васильевич** ГОУ ВПО «ДОННУ» (г. Донецк), заведующий кафедрой радиофизики и инфокоммуникационных технологий

Официальные оппоненты: **Братчун Валерий Иванович** доктор технических наук, профессор ГОУ ВПО «ДОНБАССКАЯ НАЦИОНАЛЬНАЯ АКАДЕМИЯ СТРОИТЕЛЬСТВА И АРХИТЕКТУРЫ» (ДОННАСА) (г. Макеевка), заведующий кафедрой «Автомобильные дороги и аэродромы»

Маренич Ольга Константиновна кандидат технических наук ГБУ «Донгипрошахт» (г. Донецк), ведущий инженер отдела электромеханики, автоматизации и связи

Ведущая организация: **Государственное бюджетное образовательное учреждение высшего образования «Донецкий институт железнодорожного транспорта»** (г. Донецк)

Защита диссертации состоится «30» мая 2023 года в 10:00 часов на заседании диссертационного совета 02.2.006.02 при ГОУВПО «ДОННТУ» и ГОУ ВПО «ДОННУ, по адресу: 283001, г. Донецк, ул. Артема, 58, корп. 1, ауд. 203. Тел./факс: +7(856)304-30-55, e-mail uchensovet@donntu.ru

С диссертацией можно ознакомиться в библиотеке ГОУВПО «ДОННТУ» по адресу: 283001, г. Донецк, ул. Артема, 58, корп. 2. Адрес сайта университета: http://donntu.ru

Автореферат разослан «___»_____2023 г.

Ученый секретарь диссертационного совета 02.2.006.02 кандидат технических наук, доцент Т.В. Завадская

ОБЩАЯ ХАРАКТЕРИСТИКА РАБОТЫ

Актуальность темы исследований. Контроль и прогнозирование сложных процессов, например, стратегии нацбезопасности, «…включающей оборону страны, государственную, общественную, информационную, экологическую, экономическую, транспортную, энергетическую и безопасность личности», по многочисленности аналитических вычислений немыслим без создания автоматизированных систем научных исследований (АСНИ) с привлечением элементов исскуственного интеллекта, в частности, искусственных нейронных сетей (ИНС).

Известен круг задач, решаемых ИНС: распознавание образов, классификация, принятие решений и управление, кластеризация, прогнозирование, аппроксимация, сжатие данных и ассоциативная память. Основными преимуществами нейронных сетей перед традиционными вычислительными методами являются: решение задач в условиях неопределенности, устойчивость к шумам во входных данных, гибкость структуры нейронных сетей, высокое быстродействие, адаптация к изменениям окружающей среды, отказоустойчивость. Известны и недостатки искусственных нейронных сетей, основные из которых: трудоемкость и длительность обучения; неспособность принятия решений в несколько этапов; трудозатратность с точки зрения объема используемых вычислительных ресурсов и математической сложности моделей, внедрение продвинутых нейросетевых систем возможно только по лицензии от производителя и т.д.

Тем не менее, анализ существующих научных работ показал, что создание автоматизированных систем научных исследований безопасности личности (АСНИ БЛ) на основе искусственных нейронных сетей является актуальной научно-технической задачей, имеющей практическое значение, решение которой позволит повысить их быстродействие, степень достоверности принятия решений, а также снизить уровень уязвимости нейросетевых алгоритмов.

Степень разработанности темы исследования. Вопросам применения нейросетевых алгоритмов для решения технологических задач посвящены исследования Т.В. Филатовой, А.И. Павлова, О.Ю. Лончакова, А.В. Жвакина, в которых рассмотрены преимущества нейросетевого подхода в решении задач автоматизированной обработки формализованных данных перед традиционными численными статистическими методами. В работах В.Е. Сорокина, А.А. Ежова, А.С. Новикова, И. В. Крысова, И. Л. Чулкова, А. В. Брагина исследованы нейросетевые технологии обработки изображений и применение их в автоматизированных системах машинного зрения, разработки технической документации, проектирования, обработки результатов научных исследований. Однако остается нерешенным вопрос о возможности применения в автономных автоматизированных системах малозатратных нейросетевых решений, а также оптимизации глубоких сверточных сетей для использования в микроконтроллерных системах управления. При этом актуален вопрос

уязвимостей, которыми обладают подобные системы. В работах Н. Папернота, А. Фази, Н. Карлини, Д. Вагнера и др. рассмотрены уязвимости нейросетевых алгоритмов, однако нет исследований обобщающих, сравнивающих и тестирующих известные уязвимсти нейронных сетей и степень их опасности.

Связь с научными направлениями, планами, темами. Диссертационное исследование выполнено в соответствии с планами научно-исследовательских работ ГОУ ВПО «ДОННУ», выполненных на кафедре радиофизики и инфокоммуникационных технологий (РФ и ИКТ) в рамках НИР №0118D000013 Г-18/39 «Моделирование защищенных инфокоммуникационных систем» и госбюджетной темы №0119Д000024 19-1/вв41 «Диагностика и контроль структуры конструкционных материалов по данным анализа ихпараметров методами акустической спектроскопии» в части исследований методов применения нейронных сетей для обработки экспериментальных данных.

Цель и задачи исследований. Целью работы является создание АСНИ безопасности личности на основе искусственных нейронных сетей путем совершенствования технологий обработки данных, что позволит повысить быстродействие систем и степень достоверности принятия решений, а также снизить уровень уязвимости нейросетевых алгоритмов.

Для достижения цели в работе поставлены и решены следующие задачи:

- разработка метода сокращения количества параметров обрезанием проигравших нейронов нейросетевых алгоритмов;
- программная реализация систем обработки формализованных данных и изображений, включая архитектурные решения;
- снижение уровня уязвимостей АСНИ безопасности личности на основе модифицированных нейросетевых алгоритмов.

Объект исследования. Автоматизированные системы научных исследований безопасности личности на основе искусственных нейронных сетей.

Предмет исследования. Нейросетевые модели представления и обработки данных.

Научная новизна полученных результатов заключается в следующем.

1. Впервые разработан метод сокращения количества параметров нейронной сети обрезанием проигравших нейронов, позволяющий уменьшить ресурсоемкость и увеличить быстродействие без потери точности нейросетевых алгоритмов АСНИ безопасности личности. Показано, что результаты сокращения количества параметров не зависят от архитектуры нейронной сети и методов обучения;

2. Дальнейшее развитие получили:

- метод обработки файлов логов серверов для обнаружения угроз;
- метод распознавания формы предметов на основе нейронной сети LVQ;
- процедура аутентификация с распознаванием лиц на основе сверточной нейронной сети;

- процедура видеонаблюдения по распознаванию предметов повышенной опасности.

3. Впервые обосновано применение метода Харриса для выделения признаков, поступающих на нейронную сеть, а также, разработан собственный итерационный алгоритм бинаризации, для автоматизированных систем доступа по отпечаткам пальцев.

4. Впервые предложены 9 методов генерации состязательных примеров для ненаправленных и направленных угроз на нейросетевые классификаторы и системы распознавания лиц, в том числе Bing, Google, Yandex.

Теоретическая и практическая значимость работы.

Теоретическая значимость результатов работы заключается в обосновании принципов применения нейросетевых технологий в автоматизированных системах научных исследований безопасности личности и определении ограничений, которые накладывают нейросетевые алгоритмы на такие системы.

Практическое значение результатов исследований.

1. Разработан метод, позволяющий на 70% сокращать количество параметров нейронных сетей, с сохранением и даже увеличением точности и скорости работы нейросетевых алгоритмов.

2. Разработаны принципы построения компактных нейросетевых решений, которые можно использовать для прошивки микроконтроллерных устройств в автономных автоматизированных системах.

3. Разработаны алгоритмы для обработки изображений в автоматизированных системах аутентификации и классификации.

4. Предложены способы защиты АСНИ безопасности личности на нейросетевых алгоритмах от атак генерацией состязательных примеров.

Практическая ценность работы подтверждается следующим:

а) Алгоритм распознавания лиц сверточной нейронной сетью оптимизированной разработанным методом редукции нейронов внедрен на предприятии ФИРМА «МДЛ» в систему аутентификации доступа на базе монитора видеодомофона М-480М с системой на кристалле Allwiner A-13 (Акт внедрения от 05.06. 2021 г.)

б) Методика аутентификации пользователей и программное приложение для распознавания лиц на основе нейронной сверточной сети, методика проверки цифровых документов на подлинность, компьютерное приложение для защиты цифровых документов от редактирования внедрены в учебный процесс кафедры радиофизики и инфокоммуникационных технологий путем использования в лабораторном практикуме по дисциплине «Основы информационной безопасности» в Донецком национальном универитете (акты внедрения 01.18/12.1-34 и 02.18/12.1-34 от 16.03.2018 г.).

Методология и методы исследований. Для решения поставленных задач использовались следующие научные методы: математическое моделирование (для представления данных), многомерный статистический анализ (для

обработки формализованных данных), цифровая обработка сигналов и теория распознавания образов (для обработки изображений), прогнозирование и оптимизация (для разработки архитектур систем), современные методы программирования (для практической реализации автоматизированных систем).

Научные положения, выносимые на защиту:

1. Предложенный метод сокращения количества параметров нейронных сетей обрезанием проигравших нейронов позволяет на 70% снизить количество параметров нейронных сетей, с сохранением точности и повышением скорости работы нейросетевых алгоритмов.

2. Внедрение обоснованного теоретически метода построения компактных нейросетевых АСНИ широкого назначения, в том числе безопасности личности и методов обработки изображений для АСНИ по распознаванию образов позволяет повысить их эффективность и надежность.

Степень достоверности и апробация результатов. Обоснованность и достоверность научных положений, выводов и рекомендаций подтверждается результатами математического моделирования, компьютерных экспериментов и тестирования разработанных прототипов автоматизированных систем научных исследований безопасности личности.

По направлению исследований, содержанию научных положений и выводов, существу полученных результатов диссертационная работа соответствует паспорту специальности 2.3.3. «Автоматизация и управление технологическими процессами и производствами» в частности: п.3 «Методология, научные основы, средства и технологии построения автоматизированных систем управления технологическими процессами (АСУТП) и производствами (АСУП), а также технической подготовкой производства (АСТПП) и т. д.», п. 16 «Средства и методы проектирования технического, математического, лингвистического и других видов обеспечения АСУ», п. 18 «Разработка автоматизированных систем научных исследований».

Основные положения диссертационной работы апробированы на научно-технических конференциях: IV Международная научная конференция «Донецкие чтения 2019: образование, наука, инновации, культура и вызовы современности» «Нейронные сети в системах для научных исследований», г. Донецк, 2919 г., XVIII Международная научно-практическая конференция, г. Пенза, «Атаки на нейросетевые автоматизированные системы распознавания опасных предметов» 2021 г., II Международная научно-практическая конференция «Программная инженерия: методы и технологии разработки информационно-вычислительных систем (ПИИВС-2018)» «The authenticity of digital documents checking and protection by own steganography algorithm», г. Донецк, 2018 г.

Личный вклад автора. Все результаты и положения, составляющие основное содержание диссертации, вынесенные на защиту, получены автором самостоятельно. Личный вклад соискателя заключается в обосновании идеи

работы и ее реализации, в выполнении теоретических и экспериментальных исследований, разработке программного обеспечения и прототипов АСНИ.

Вклад соискателя в работы, опубликованные в соавторстве, конкретизирован в списке работ, опубликованных по теме диссертации.

Публикации. Основные научные результаты диссертации опубликованы автором самостоятельно и в соавторстве в 12 научных изданиях, из них 4 – в изданиях, включенных в перечень ВАК ДНР, 4 – в иных научных изданиях, 4 – в материалах и тезисах конференций.

Структура и объём диссертации. Диссертационная работа изложена на 188 страницах, состоит из введения, четырех разделов, заключения, перечня условных сокращений, списка литературы из 93 наименований и 6 приложений.

ОСНОВНОЕ СОДЕРЖАНИЕ РАБОТЫ

В первом разделе диссертационной работы «Автоматизированные системы научных исследований безопасности личности (АСНИ БЛ) на основе искусственных нейронных сетей» выделен класс задач, которые решаются при помощи нейросетевых технологий. Сформулированы проблемы, возникающие при внедрении нейросетевых алгоритмов в АСНИ БЛ: затратность с точки зрения машинных ресурсов, необходимость лицензирования существующих решений и уязвимость к спецефическим нейросетевым атакам.

Рассмотрены

- нейронные сети в АСНИ по обработке формализованных данных;
- нейронные сети в АСНИ по обработке изображений;
- надежность нейросетевых алгоритмов АСНИ.

Сделаны следующие выводы по разделу. Задачи аппроксимации, интерполяции, экстраполяции и классификации решаются с помощью сетей с радиально-базисными функциями, вероятностных сетей, обобщенно-регрессионных сетей, многослойного персептрона, сетей с конкурирующими слоями, сверточных сетей. Основное внимание необходимо уделить архитектуре сети, методу обучения, выделению и подготовке данных для обучения и входных данных.

Для решения задач по обработке изображений используются специальные библиотеки, которые предоставляют возможность создавать собственную архитектуру или пользоваться готовыми моделями, которые можно модифицировать и обучать, однако такого рода решения затратны по ресурсам.

Большинство исследований сосредоточено на архитектуре нейронных сетей, алгоритмах обучения и принципиальной возможности решения задачи распознавания для того или иного вида данных, однако не столь широко исследованы такие характеристики как быстродействие, время обучения, объем использованной оперативной и постоянной памяти, точность распознавания, что является ключевыми характеристиками в случае применения нейронных сетей в автономных автоматизированных системах на нейронных сетях.

Автоматизированные системы на нейросетевых алгоритмах могут быть введены в заблуждение, состязательнми примерами, созданными той же, либо сторонней нейронной сетью, на сегодняшний день успешность таких атак невелика и в среднем ниже 60 %.

Состязательные примеры, генерируемые методом градиентного спуска, более эффективны в случае, если сгенерированы нейронной сетью той же архитектуры, что и целевая, если же архитектура целевой сети неизвестна, то применяют генетические алгоритмы.

На сегодняшний день не существует универсального метода защиты от такого рода атак, для каждого вида атаки необходимо подбирать соответствующий метод. Наиболее успешны атаки на АСНИ, работающие без вмешательства человека и с ростом числа подобных систем, проблема уязвимости нейросетевых алгоритмов становится актуальной.

Остается нерешенным вопрос о возможности применения в АСНИ малозатратных нейросетевых решений, а также снижения количества параметров глубоких сверточных сетей для использования в микроконтроллерных системах управления. При этом актуален вопрос уязвимостей, которыми обладают подобные системы, поскольку при получении доступа к управлению технологическим процессом или к системе аутентификации злоумышленник может нанести ощутимый вред, однако исследований, обобщающих, сравнивающих и тестирующих известные атаки на нейронные сети на сегодняшний день крайне мало.

Краткая характеристика трудностей, возникающих при внедрении ИНС в АСНИ БЛ позволила сформулировать следующие задачи исследования: разработка метода сокращения количества параметров обрезанием проигравших нейронов нейросетевых алгоритмов, программная реализации систем обработки формализованных данных и изображений, включая архитектурные решения, снижение уровня уязвимостей АСНИ безопасности личности на основе модифицированных нейросетевых алгоритмов.

Второй раздел диссертации «Сокращение количества параметров нейронной сети редукцией проигравших нейронов», предлагает решение первой задачи исследования, а именно, разработке метода сокращения количества параметров обрезанием проигравших нейронов нейросетевых алгоритмов.

Выдвинута гипотеза о том, что если приравнять к нулю веса нейронов обученной нейронной сети, которые не изменяются в процессе обучения (обрезать сеть), а затем снова начать обучать такую обрезанную сеть, то можно добиться увеличения точности, при уменьшении количества параметров сети, разработан алгорим обрезания.

Выдвинутая гипотеза проверена экспериментально на полносвязной сети LeNet и сверточных сетях Conv-2, Conv-4, Conv-6 с различным количеством сверточных слоев, а затем на глубоких нейронных сетях VGG-19 и ResNet-18 с различными особенностями архитектуры и обучения. Получено увеличение точности распознавания на 2-3 % после обрезания более, чем на 30 % и

уменьшения параметров сети на 70 %. На рисунке 1 представлен характерный график зависимости точности классификации предметов сверточными нейронными сетями различной архитектуры с обучающим датасетом CIFAR10.

Выигрыш в точности распознавания зависит от алгоритма сокращения количества параметров, скорости обучения, набора данных для обучения и архитектуры сети.

Экспериментально подтверждено, что описанное в ранее опубликованных работах рандомное обрезание дает падение точности распознавания в среднем 21 % за итерацию, больше, чем падение точности при обрезке неуспешных нейронов (в среднем 2,9 % за итерацию).

Для сравнения проводилось рандомномное обрезание нейронов, предложенное в ранее опубликованных работах. При последующем обучении обрезанных нейронных сетей сети для предложенного обрезания проигравших нейронов точность распознавания увеличивалась в среднем на 7 % по сравнению с рандомным обрезанием.

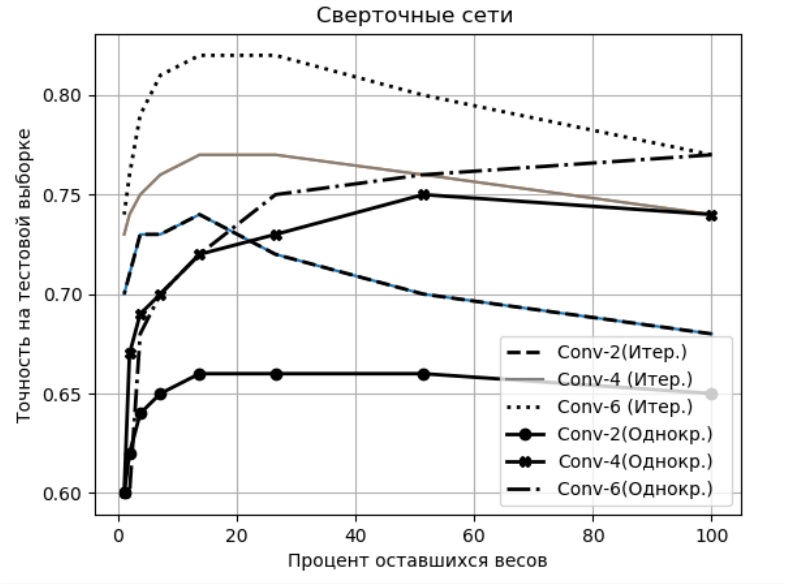

Рисунок 1 – Точность до ранней остановки обрезания весов с учетом выигравших нейронов при использовании однократной и итеративной обрезки для сверточных сетей с разным количеством сверточных слоев

На рисунке 2 представлены результаты сравнения рандомного обрезания с обрезанием по предложенному автором алгоритму для сети Lenet-300-100, обучаемой на наборе MNIST, распознающей рукописные цифры от 0 до 9.

Предложенную методику целесообразно использовать в автоматизированных системах, основанных на нейросетевых алгоритмах, для уменьшения ресурсоемкости и увеличения быстродействия без потери точности.

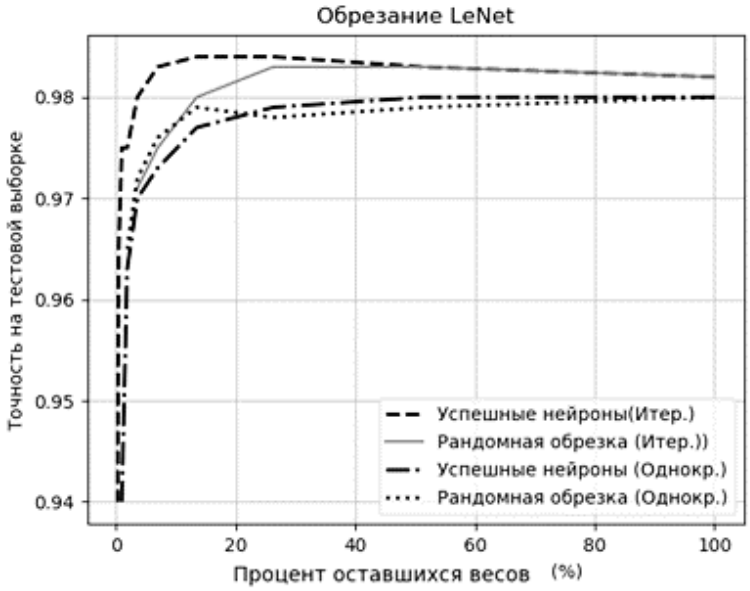

Рисунок 2 – Точность до ранней остановки для рандомного обрезания и обрезания весов с учетом выигравших нейронов при использовании однократной и итеративной обрезки

В **третьем разделе «**Разработка АСНИ БЛ на основе искусственных нейронных сетей» рассматривается решение второй задачи исследований, в частности, программная реализация систем обработки формализованных данных и изображений, включая архитектурные решения.

Рассмотрены:

- автоматизированные системы научных исследований по обработке формализованных экспериментальных данных, в частности, автоматизированная система анализа файлов логов на основе нейронной сети и логистической регрессии и АСНИ по обнаружению радиоканалов утечки информации;

- АСНИ по обработке изображений, в частности, автоматизированная система распознавания формы предметов, автоматизированная система аутентификации по отпечаткам пальцев, автоматизированная система биометрической аутентификации по лицу и атоматизированная система видеонаблюдения по распознаванию предметов повышенной опасности.

По разделу сделаны следующие выводы.

Предложены описания этапов разработки двух автоматизированных систем обработки формализованных данных (анализа файлов логов и определения каналов утечки информации) на нейронных сетях типа двухслойный персептрон и рекуррентной сети Элмана и четырех автоматизированных системы обработки изображений (распознавания формы предметов на основе сети LVQ, аутентификации пользователей по отпечаткам пальцев на основе сети RBF и по лицу на основе сверточной сети с пулингом,

распознавания опасных предметов в режиме реального времени на основе сверточной сети глубокого обучения).

Показано, что успешность работы таких систем зависит от методики подготовки данных для обучения, алгоритма обучения, выбранной архитектуры и достигает точности 85-97 %.

Экспериментально доказано, что для систем, работающих с изображениями, большое значение, имеют алгоритмы предварительной обработки изображений, поэтому был предложен собственный алгоритм бинаризации и впервые применен детектор Харриса для выделения признаков.

Показано, что автоматизированные системы на нейронных сетях, состоящие из 2-х, 3-х скрытых слоев не требуют сокращения количества параметров, поскольку не занимают значительных вычислительных ресурсов и могут быть применены в микроконтроллерных системах или одноплатных мини-компьютерных системах. Однако для систем глубокого обучения для распознавания изображений, с наличием большого количества сверточных слоев требуется сокращение количества параметров и предложенный метод редукции проигравших нейронов позволяет уменьшить количество параметров на 70-80% без потери точности.

Результаты исследований внедрены в систему аутентификации и доступа на базе монитора видеодомофона М-480М с системой на кристалле Allwiner A-13 (Рисунок 3).

Рисунок 3 – Процесс распознавания лица на мониторе видеодомофона М-480М

Достигнута точность распознавания лиц 89 % и скорость распознавания 0,5 с/лицо при записи 5-6 лиц в базу данных, 4,5 с/лицо при записи 150 лиц в базу данных.

В **четвертом разделе** диссертации «Угрозы АСНИ БЛ основанных на нейросетевых технологиях», предложено решение третьей задачи пункта исследований, в частности, снижение уровня уязвимостей АСНИ безопасности личности на основе модифицированных нейросетевых алгоритмов.

Рассмотренны:

- угрозы для автоматизированных нейросетевых классификаторов;
- уязвимости АСНИ распознавания лиц;
- методика компьютерных экспериментов;
- результаты тестирования на поисковых АСНИ БЛ;
- методы защиты от угроз для нейросетевых алгоритмов.

Из проведенных исследований сделаны следующие выводы.

У использованных нейросетевых алгоритмов исследованы виды угроз. Разработано приложение с пользовательским интерфейсом, позволяющее создавать состязательные примеры эксплуатирующие уязвимости нейросетевых алгоритмов с возможностью задания параметров угроз. С помощью данного приложения проведено исследование 9-ти методов генерации состязательных примеров для ненаправленных и направленных атак. Для ненаправленных атак лучшие результаты показали GaussianBlurAttack, GradientAttack и Single Pixel Attack, для направленных атак - Carlini and Wagner атака. Показано, что нейронные сети поисковых систем Bing, Google и Yandex устойчивы к направленным атакам типа «черный ящик». Алгоритм google и Яндекс оказались уязвимы к ненаправленным атакам. На рисунке 4 представлены результаты поиска по состязательным примерам Additive Gaussian Noise Attack.

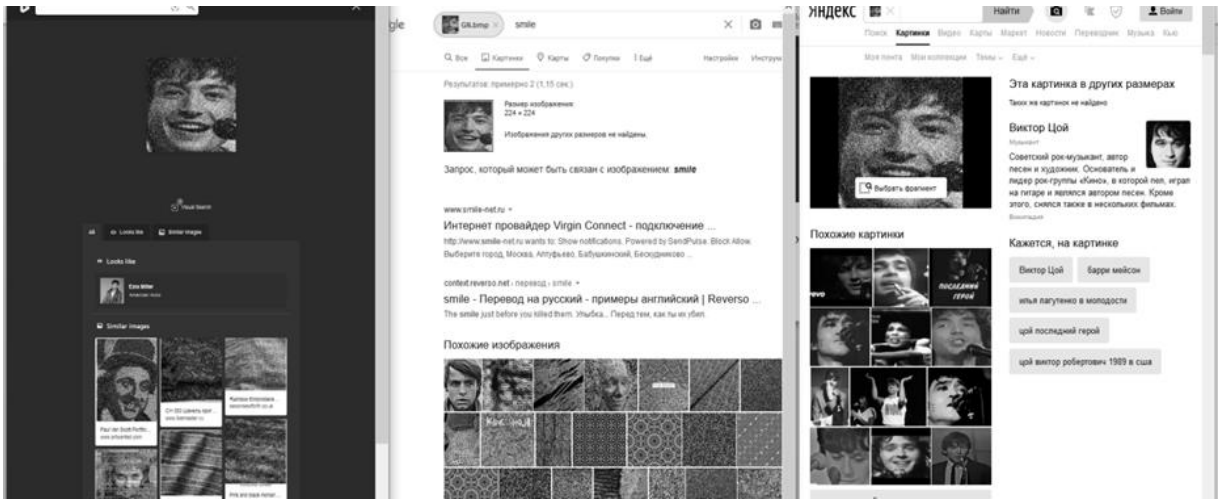

Рисунок 4 Поисковая система Yandex распознала модифицированное алгоритмом Additive Gaussian Noise изображение «Эрза Миллер», как «Виктор Цой».

Предложены методы защиты от атак на нейросетевые алгоритмы, на основе анализа полученных результатов по каждой атаке, в частности подсчет четных и нечетных значений яркости пикселей, обучение на состязательных примерах, повышение контрастности и фильтрация, уменьшение обратной связи с атакующим,

Сделан вывод, что целесообразно дополнять автоматизированные системы аутентификации инструментами поиска подобных уязвимостей и предусматривать защитные меры.

ЗАКЛЮЧЕНИЕ

В диссертационной работе дано решение актуальной научно-технической задачи совершенствования АСНИ безопасности личности на основе искусственных нейронных сетей путем улучшения технологий обработки данных, что позволяет повысить быстродействие систем и степень достоверности принятия решений, а также снизить уровень уязвимости нейросетевых алгоритмов. Основные результаты работы состоят в следующем:

1. Разработан метод сокращения количества параметров нейросетевых алгоритмов редукцией проигравших нейронов, который позволяет снизить количество параметров нейронных сетей до 30 %, повышает скорость работы и снижает ресурсоемкость программных решений.

2. Разработаны архитектурные решения и программная реализация двух систем обработки формализованных данных (для анализа файлов логов сервера и обнаружения радиоканалов утечки информации) и четырех систем обработки изображений (определение формы предмета, распознавание по отпечатку пальца, по лицу, распознавание опасных предметов).

3. Исследованы уязвимости АСНИ БЛ на основе нейросетевых алгоритмов путем создания состязательных примеров и проведения 9-ти типов направленных и ненаправленных атак на классификаторы изображений, в том числе поисковики Bing, Google, Yandex. Предложены методы защиты от подобных атак.

4. Анализ полученных экспериментальных данных подтверждает возможность внедрения нейросетевых технологий в автономные автоматизированные системы другого назначения, например, промышленного и широкого назначения за счет уменьшения ресурсоемкости разработанных решений без понижения эффективности и качества.

5. Методика аутентификации пользователей и программное приложение для распознавания лиц на основе нейронной сверточной сети, методика проверки цифровых документов на подлинность, компьютерное приложение для защиты цифровых документов от редактирования внедрены в учебный процесс кафедры радиофизики и инфокоммуникационных технологий путем использования в лекционных курсах и лабораторных практикумах по дисциплинам «Пакеты прикладных программ для обработки изображений», «Нейронные сети», «Основы информационной безопасности».

6. Результаты диссертационных исследований внедрены в систему аутентификации и доступа на базе монитора видеодомофона М-480М с системой на кристалле Allwiner A-13. Достигнута точность распознавания 89 % и скорость распознавания 0,5 с/лицо. при записи 5-6 лиц в базу данных, 4,5 с/лицо при записи 150 лиц в базу данных, что позволило увеличить скорость работы и уменьшить ресурсоемкость системы.

СПИСОК РАБОТ, ОПУБЛИКОВАННЫХ АВТОРОМ ПО ТЕМЕ ДИССЕРТАЦИИ

− в изданиях из перечня рецензируемых научных изданий ВАК ДНР:

1. Бабичева М.В., Автоматизация средств обнаружения радиоканалов утечки информации / М.В. Бабичева, А.С. Попов, А. В. Яновский // Вестник Донецкого национального университета. Серия Г: Технические науки. – 2020. – № 1. – С. 9-13.

2. Бабичева М.В., Атаки на автоматизированные системы аутентификации на нейросетевых классификаторах/ М.В. Бабичева, Д. В. Василенко // Вестник Донецкого национального университета. Серия Г: Технические науки. – 2020. – № 1. – С. 23-30.

3. Бабичева М.В. Автоматизированная система видеонаблюдения по распознаванию предметов повышенной опасности / М.В. Бабичева, В. В. Данилов «Сборник научных трудов Донецкий институт железнодорожного транспорта». – 2020. – № 56. – С. 20-25

4. Бабичева М.В. Вероятностный подход к оптимизации нейронных сетей случайной редукцией нейронов / М. В. Бабичева, В. В. Данилов, С. В. Борщевский // Вестник Донецкого национального университета. Серия Г: Технические науки. – 2021. – № 1. – С.63-71.

− в иных научных изданиях:

5. Бабичева М.В. Атаки на нейросетевые классификаторы М.В. Бабичева // «Вестник Донецкого национального университета. Серия Г: Технические науки». – 2019. – № 2. – С. 51-55

6. Бабичева, М.В. Выделение особых точек отпечатков пальцев детекторами углов / М.В. Бабичева, А.С. Юрченко // Вестник Донецкого национального университета Серия Г. Технические науки. – 2019. – № 2. – С. 10- 15.

7. Бабичева, М. В. Итеративный алгоритм пороговой бинаризации для обработки отпечатков пальцев в биометрических системах доступа / М.В. Бабичева, А. С. Юрченко // Вестник Донецкого национального университета Серия Г. Технические науки, 2018. – № 3. – С. 41-46.

8. Бабичева, М. В. Распознавание лиц в режиме реального времени сверточной нейронной сетью / М.В. Бабичева, А.С. Шевченко // Вестник Донецкого национального университета. Серия Г: Технические науки. – 2018. – № 2. – С. 67 - 71.

− в материалах конференций:

9. Babicheva, M.V. The authenticity of digital documents checking and protection by own steganography algorithm / M.V. Babicheva // Сборник материалов II Международной научно-практической конференции Программная инженерия: методы и технологии разработки информационно-вычислительных систем (ПИИВС-2018). – Том. 1., 14-18 ноября 2018 г. – С. 96-101.

10. Бабичева, М.В. Нейронные сети в системах для научных исследований / М.В. Бабичева, В. В. Данилов, А.С. Юрченко // Материалы конференции Материалы IV Международной научной конференции «Донецкие чтения 2019». – Том 1, Часть 2. – С. 164-165.

11. Бабичева, М.В. Оптимизация нейронной сети редукцией проигравших нейронов// Материалы конференции Материалы V Международной научной конференции «Донецкие чтения 2020». – Том 1, Часть 2. – С. 156-158.

12. Бабичева, М.В. Атаки на нейросетевые автоматизированные системы распознавания опасных предметов/ М. В. Бабичева, П. А. Майоров // Современные научные исследования: актуальные вопросы, достижения и инновации: сборник статей XVIII Международной научнопрактической конференции. – Пенза: МЦНС «Наука и Просвещение». – 2021. – С. 63-69.

Личный вклад автора в публикациях: [1] – сформулирована идея применения нейросетевых алгоритмов в системах обнаружения несанкционированного излучения; [2] – предложены сценарии эксплуатации уязвимостей на 12 ресурсов для распознавания и классификации изображений нейронными сетями; [3] – предложены улучшения архитектуры нейронной сети Yoolo, позволяющие повысить точность распознавания предметов повышенной опасности; [4] – предложена вероятностная модель редукции нейронов скрытого слоя; [5] – предложена собственная классификация современных методов генерирования вредоносных данных для нейросетевых систем; [6] – предложено использование угловых детекторов Харриса для выделения особых точек локальных признаков отпечатков пальцев; [7] – предложен метод итеративной пороговой бинаризации изображений отпечатков пальцев; [8] – выбран алгоритм обнаружения лица и разработаны архитектура нейронной сети и методика тестирования; [9] – предложена авторская методика определения подлинности электронных документов и собственный алгоритм стеганографии; [10] – выбраны архитектры нейронных сетей и определены исследуемые параметры; [11] – предложен метод оптимизации нейронных сетей редукцией проигравших нейронов; [12] – предложена математическая модель угроз, алгоритм создания патча, методика проведения экспериментальных атак.

АННОТАЦИЯ

Бабичева М. В. **Автоматизированные системы научных исследований угроз безопасности личности.** - На правах рукописи.

Диссертация на соискание ученой степени кандидата технических наук по специальности 2.3.3. «Автоматизация и управление технологическими процессами и производствами» – ГОУВПО «ДОНЕЦКИЙ НАЦИОНАЛЬНЫЙ ТЕХНИЧЕСКИЙ УНИВЕРСИТЕТ», Донецк, 2023.

В диссертационной работе дано решение актуальной научно-технической задачи создания АСНИ безопасности личности на основе искусственных нейронных сетей, путем совершенствования технологий обработки данных, что позволяет для систем повысить их быстродействие и степень достоверности

принятия решений, а также снизить уровень уязвимости нейросетевых алгоритмов.

Разработан метод сокращения количества параметров нейросетевых алгоритмов редукцией проигравших нейронов, который позволяет снизить количество параметров нейронных сетей до 20 %, не понижая точности и повышает скорость работы и понижает ресурсоемкость программных решений.

Разработаны архитектурные решения и программная реализация двух систем обработки формализованных данных (для анализа файлов логов сервера и обнаружения радиоканалов утечки информации) и четырех систем обработки изображений (определение формы предмета, распознавание по отпечатку пальца, по лицу, распознавание опасных предметов).

Исследованы уязвимости автоматизированных систем на основе нейросетевых алгоритмов путем создания состязательных примеров и проведения 9-ти типов направленных и ненаправленных атак на классификаторы изображений, в том числе поисковики Bing, Google, Yandex. Предложены методы защиты от подобных атак.

Анализ полученных экспериментальных данных подтверждает возможность внедрения нейросетевых технологий в автономные автоматизированные системы для промышленного и широкого назначения за счет уменьшения ресурсоемкости разработанных решений без понижения эффективности и качества.

Результаты диссертационных исследований внедрены в систему аутентификации и доступа на базе монитора видеодомофона М-480М. Достигнута точность распознавания 89 % и скорость распознавания 0,5 с/ лицо. при записи 5-6 лиц в базу данных, 4,5 с/лицо при записи 150 лиц в базу данных, а также в учебный процесс кафедры радиофизики и инфокоммуникационных технологий путем использования в лекционных курсах и лабораторных практикумах по дисциплинам «Пакеты прикладных программ для обработки изображений», «Нейронные сети», «Основы информационной безопасности».

Ключевые слова: автоматизированные системы научных исследований, угрозы безопасности личности, нейронные сети.

ABSTRACT

Babicheva M. V. Automated scientific systems research of threats for personal security. - Manuscript.

Ph.D. (Candidate’s) Thesis in Engineering Science by specialty 2.3.3. «Automation and control of technological processes and industries» – State Higher Educational Establishment «Donetsk National Technical University», Donetsk, 2023.

The dissertation work provides a solution to the urgent scientific and technical problem of creating an automated scientific systems research of threats for personal security based on artificial neural networks, by improving data processing technologies, which allows systems to increase their speed and the degree of reliability of decision-making, as well as reduce the level of vulnerability of neural network algorithms.

A method for reducing the number of parameters of neural network algorithms by reducing losing neurons has been developed, which allows reducing the number of parameters of neural networks to 20% without lowering the accuracy and increases the speed and resource intensity of software solutions.

Architectural solutions and software implementation of two systems for processing formalized data (for analyzing server log files and detecting radio channels of information leakage) and four image processing systems (determining the shape of an object, recognition by a fingerprint, by a face, recognition of dangerous objects) have been developed.

Automated systems based on neural network algorithms vulnerabilities were investigated by creating adversarial examples and conducting nine types of directed and undirected attacks on image classifiers, including search engines Bing, Google, Yandex. Protection methods against such attacks are proposed.

The analysis of the obtained experimental data confirms the possibility of introducing neural network technologies into autonomous automated systems for industrial and general purpose by reducing the resource intensity of the developed solutions without reducing efficiency and quality.

The results of the dissertation research were used in the authentication and access system based on the M-480M video intercom monitor with the Allwiner A-13 chip system. Recognition accuracy of 89 % and recognition speed of 0.5 c/face are achieved when 5-6 persons are recorded in the database and 4.5 c/person when 150 persons are recorded in the database, as in the educational process of the Department of Radiophysics and Infocommunication Technologies by using them in lecture courses and laboratory workshops on the disciplines "Application packages for image processing", "Neural networks", "Fundamentals of information security".

Keywords: automated research systems, personal security threats, neural networks.