ОТЗЫВ ОФИЦИАЛЬНОГО ОППОНЕНТА

на диссертационную работу Бабичевой Маргариты Вадимовны на тему «Автоматизированные системы научных исследований угроз безопасности личности», представленную на соискание ученой степени кандидата технических наук по специальности 2.3.3. «Автоматизация и управление технологическими процессами и производствами» (технические науки).

Актуальность избранной темы.

Возможности применения нейронных сетей в широком спектре современных технологических производств растут. Контроль и прогнозирование сложных процессов ввиду многочисленности аналитических вычислений немыслим без создания автоматизированных систем научных исследований с привлечением элементов искусственного интеллекта. Применение нейронных сетей в стратегии безопасности как производственных процессов, так и государства в целом, позволяет повысить гибкость и быстродействие систем по предотвращению угроз.

Одним из перспективных направлений является создание автоматизированных систем научных исследований угроз безопасности личности, в частности систем аутентификации по биометрическим признакам, систем видеонаблюдения с распознаванием лиц и опасных предметов, автоматизированных систем определения сетевых вторжений и др. Однако трудоемкость и длительность обучения, затратность с точки зрения объема используемых вычислительных ресурсов и математической сложности моделей затрудняет применение нейросетевых алгоритмов в автономных микроконтроллерных автоматизированных системах. При этом важен вопрос уязвимостей, которыми обладают подобные решения.

Решение задачи совершенствования автоматизированных систем научных исследований безопасности личности на основе искусственных нейронных сетей улучшением технологий обработки данных, позволит повысить быстродействие систем и степень достоверности принятия решений, а также снизить уровень уязвимости нейросетевых алгоритмов.

Поэтому, тема «Автоматизированные системы научных исследований угроз безопасности личности», является актуальной. Актуальность темы также подтверждается анализом литературных источников и результатами исследований других авторов.

Степень обоснованности научных положений, выводов и рекомендаций.

Диссертационная работа состоит из введения, четырех разделов с выводами по каждому из них, заключения, перечня условных сокращений, списка литературы из 93 наименований и 6 приложений. Полный объем работы составляет 187 страниц, включая 86 рисунков и 19 таблиц.

Во **введении** диссертантом приведена общая характеристика работы, обоснованы актуальность и степень разработанности темы диссертации, приведены цель, задачи, объект и предмет исследования, сформулированы научная новизна, теоретическая и практическая значимость полученных результатов.

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. Рассмотрены нейронные сети в АСНИ по обработке формализованных данных и обработке изображений, надежность нейросетевых алгоритмов АСНИ. Сделаны выводы о том, что 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.

Краткая характеристика трудностей, возникающих при внедрении ИНС в АСНИ БЛ позволило соискателю определить цель и сформулировать три основные задачи исследования: разработка метода сокращения количества параметров обрезанием проигравших нейронов нейросетевых алгоритмов; программная реализации систем обработки формализованных данных и изображений; снижение уровня уязвимостей АСНИ безопасности личности на основе модифицированных нейросетевых алгоритмов, решение которых необходимо для достижения поставленной цели.

Во **втором разделе** диссертации «Сокращение количества параметров нейронной сети редукцией проигравших нейронов» выдвинута гипотеза о возможности уменьшения параметров сети удалением нейронов, веса которых не меняются в процессе обучения, с дальнейшим дообучением, без потери точности. Разработан алгоритм для однократного и итеративного обрезания нейронной сети, выдвинуто предположение, что результаты не будут зависеть от архитектуры сети и методов обучения.

Выдвинутая диссертантом гипотеза проверена на нейронных сетях различной архитектуры, как полносвязных, так и сверточных. Все исследованные архитектуры нейронных сетей при обрезании, с последующим обучением в среднем показали увеличение точности на 2-3% при уменьшении параметров сети на 70-80%. Результаты сравниваются с результатами, предложенного ранее рандомного удаления весов, которое, как видно из результатов экспериментов, дает меньшее увеличение точности, чем описанное в диссертационной работе.

Диссертантом предложено использовать данную методику в автоматизированных системах, основанных на нейросетевых алгоритмах, для уменьшения ресурсоемкости и увеличения быстродействия без потери точности, описанных в следующем разделе. В итоге в данном разделе решена первая задача диссертационной работы.

В **третьем разделе** диссертации «Разработка АСНИ безопасности личности на основе искусственных нейронных сетей» описаны разработанные автоматизированные системы обработки формализованных данных (анализ файлов логов серверов для обнаружения вторжений и обнаружения радиоканалов утечки информации) и автоматизированные системы обработки изображений (распознавание формы предметов, аутентификации по отпечаткам пальцев, биометрической аутентификации по лицу, видеонаблюдения по распознаванию предметов повышенной опасности) на нейронных сетях.

В разделе подробно описаны методики подготовки данных для обучения, выбор алгоритма обучения и особенностей архитектуры. Показано, что успешность работы таких систем зависит от методики подготовки данных для обучения, алгоритма обучения, выбранной архитектуры или модели и достигает точности 85-97%. Для системы распознавания по отпечаткам пальцев разработан собственный алгоритм бинаризации и применен детектор Харриса для выделения признаков.

Диссертантом Бабичевой М.В. показано, что автоматизированные системы на нейронных сетях, состоящие из 2-х, 3-х скрытых слоев не требуют оптимизации, поскольку не занимают значительных вычислительных ресурсов. Однако в нейронных сетях глубокого обучения, с наличием большого количества сверточных слоев необходимо сократить количество параметров для увеличения быстродействия и уменьшения ресурсоемкости.

Применение разработанного в разделе 2 метода редукции проигравших нейронов позволяет уменьшить количество параметров на 70-80% без потери точности. Таким образом, предложенный в разделе 2, метод был успешно опробован экспериментально на разработанных диссертантом системах.

По итогам рассмотрения раздела можно сделать вывод, что в данном разделе решена вторая задача диссертационной работы.

В **четвертом разделе** диссертации «Угрозы АСНИ БЛ основанных на нейросетевых технологиях» рассмотрены атаки на нейросетевые классификаторы. Диссертантом разработано приложение, позволяющее проводить как атаки на заданный класс, так и атаки на произвольный класс с заданием значения точности распознавания, выбором и контролем оптимизируемой метрики.

В разделе приведены результаты исследования девяти методов генерации состязательных примеров для атак на нейросетевые классификаторы. Сгенерированные состязательные изображения предъявлялись как нейронным классификаторам различной архитектуры, так и поисковым системам Bing, Google и Yandex, работающих на нейросетевых алгоритмах неизвестной архитектуры. Классификаторы известной архитектуры на нейронных сетях оказались уязвимы к атакам как на целевой класс, так и произвольный.

Все сторонние нейронные сети показали устойчивость к атакам на определенный класс, однако алгоритмы Google и Yandex оказались уязвимы к атакам на произвольный класс. На основе анализа полученных результатов

диссертантом предложены методы защиты и сделан вывод о необходимости дополнять автоматизированные системы аутентификации инструментами защиты и поиска уязвимостей.

По итогам рассмотрения раздела можно сделать вывод, что в данном разделе решена третья и последняя из поставленных задач диссертационной работы.

В **заключении** диссертантом приведены основные научные результаты и выводы, полученные при выполнении работы. Сформулированные научные положения и выводы в полной мере отражаются и в достаточной степени обосновываются в тексте диссертации.

Достоверность и новизна научных положений, выводов и рекомендаций.

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

Для подтверждения обоснованности и достоверности своих научных положений диссертант использует значительное количество экспериментов и наблюдений, применяет современные методы исследований, которые соответствуют поставленным в работе цели и задачам.

Сформулированные диссертантом научные положения и выводы, подтверждаются убедительными фактическими данными, результатами экспериментов, наглядно приведенными в таблицах и на графиках.

Достоверность научных положений, помимо этого, подтверждается справками о внедрениях (представлены в Приложении А диссертации) полученных результатов в области практической, научно-исследовательской и преподавательской деятельности, а именно:

1. Внедрение в производственный процесс предприятия ФИРМА «МДЛ» алгоритма распознавания лиц сверточной нейронной сетью оптимизированной разработанным методом редукции нейронов (Акт внедрения от 05.06. 2021 г.)

2. Внедрение в учебный процесс ГОУ ВПО «Донецкий национальный университет» (акты внедрения 01.18/12.1-34 и 02.18/12.1-34 от 16.03.2018 г.).

Кроме того, достоверность и обоснованность полученных научных положений диссертационной работы подтверждается значительным количеством работ, опубликованных по теме диссертации и апробацией на научно-практических конференциях различного уровня. Результаты диссертации опубликованы в 12 печатных работах, в том числе четыре - в рецензируемых научных изданиях, рекомендуемых Высшей Аттестационной Комиссией. Также полученные результаты обсуждались на трех международных научно-практических конференциях и получили одобрение специалистов.

Замечания

1. В третьем разделе приводятся некоторые излишне подробные сведения, особенно по алгоритмам обработки изображений. Объем некоторых подразделов можно было уменьшить, не потеряв при этом информативности.
2. В работе имеются некоторые неточности и опечатки, в частности на рисунке 2.4 указана функция активации Sugma, в то время как правильное название - Sigma.
3. Предложенный во второй главе алгоритм целесообразно было бы представить в виде блок-схемы для большей наглядности.
4. Употребление термина «скорость обучения» в различных значениях, как коэффициента обучения во втором разделе и непосредственно скорости обучения в третьем разделе затрудняет понимание сути исследования.
5. В четвертом разделе сразу не определено, что подразумевается под направленными и ненаправленными атаками.
6. Целесообразно было бы проверить на уязвимости разработанные автоматизированные системы методами, предложенными в четвертой главе.

Заключение.

Несмотря на вышеуказанные замечания, диссертационная работа оценивается положительно и соответствует критериям, установленным пунктом 2.2 «Положения о присуждении ученых степеней», утвержденного Постановлением Совета Министров Донецкой Народной Республики №2-13, от 27.02.2015 г. (с изменениями). Содержание диссертации соответствует областям исследования паспорта научной специальности
2.3.3 – Автоматизация и управление технологическими процессами и производствами (технические науки).

Таким образом, Бабичева Маргарита Вадимовна заслуживает присуждения ученой степени кандидата технических наук по специальности 2.3.3. – Автоматизация и управление технологическими процессами и производствами (технические науки).

Я, Братчун Валерий Иванович, даю согласие на автоматизированную обработку моих персональных данных, приведенных в этом документе.

Официальный оппонент
д.т.н. по специальности 05.23.05 –
строительные материалы и изделия,
профессор, заведующий кафедрой
«Автомобильные дороги и аэродромы»,
Государственного образовательного учреждения
высшего профессионального образования
«Донбасская национальная академия
строительства и архитектуры» _____ В.И. Братчун
(подпись)

Подпись д.т.н., профессора Братчуна В.И. заверяю:
Ученый секретарь
Государственного образовательного учреждения
высшего профессионального образования
«Донбасская национальная академия
строительства и архитектуры»,
к.э.н., доцент _____ М.А. Гракова
(подпись)

Почтовый адрес организации: Государственное образовательное учреждение высшего профессионального образования «Донбасская национальная академия строительства и архитектуры»: 286123, Донецкая Народная Республика, г. Макеевка, ул. Державина, д. 2. Тел.: +7-856-343-7033,
эл. почта: mailbox@donnasa.ru