

И2
Т0
Н2
Т2



СБОРНИК ТРУДОВ

**VIII Международной конференции
и молодежной школы «Информационные
технологии и нанотехнологии»**



**23-27 МАЯ
САМАРА, РОССИЯ**

МИНИСТЕРСТВО НАУКИ И ВЫСШЕГО ОБРАЗОВАНИЯ РОССИЙСКОЙ ФЕДЕРАЦИИ

ФЕДЕРАЛЬНОЕ ГОСУДАРСТВЕННОЕ АВТОНОМНОЕ
ОБРАЗОВАТЕЛЬНОЕ УЧРЕЖДЕНИЕ ВЫСШЕГО ОБРАЗОВАНИЯ
«САМАРСКИЙ НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ
УНИВЕРСИТЕТ ИМЕНИ АКАДЕМИКА С.П. КОРОЛЕВА»
(САМАРСКИЙ УНИВЕРСИТЕТ)

ИНСТИТУТ СИСТЕМ ОБРАБОТКИ ИЗОБРАЖЕНИЙ РАН –
ФИЛИАЛ ФНИЦ "КРИСТАЛЛОГРАФИЯ И ФОТОНИКА" РАН

ИНФОРМАЦИОННЫЕ ТЕХНОЛОГИИ
И НАНОТЕХНОЛОГИИ (ИТНТ-2022)
Том 4. Искусственный интеллект

*Сборник трудов по материалам
VIII Международной конференции и молодежной школы
(г. Самара, 23-27 мая)*

Одобрено редакционно-издательским советом федерального государственного автономного образовательного учреждения высшего образования «Самарский национальный исследовательский университет имени академика С.П. Королева»

© Самарский университет, 2022

ISBN 978-5-7883-1788-5

ISBN 978-5-7883-1792-2 (т. 4)

САМАРА
Издательство Самарского университета
2022

УДК 004.9
ББК 32.973
И741

И741 Информационные технологии и нанотехнологии (ИТНТ-2022): сборник трудов по материалам VIII Международной конференции и молодежной школы (г. Самара, 23-27 мая): в 5 томах / Министерство науки и высшего образования Российской Федерации, Самарский университет, Институт систем обработки изображений РАН – филиал ФНИЦ "Кристаллография и фотоника" РАН. – Самара: Издательство Самарского университета, 2022. – **Том 4. Искусственный интеллект** / под ред. А. В. Никонорова. – 1 CD-ROM (5,6 Мб). – Загл. с титул. экрана. – Текст. Изображение: электронный.

ISBN 978-5-7883-1792-2 (т. 4)
ISBN 978-5-7883-1788-5

Тематика Конференции ИТНТ-2022 охватывает широкий круг областей применения информационных технологий в науке и высокотехнологичных отраслях промышленности. Одним из приоритетных направлений работы Конференции является образовательный аспект, заключающийся в предоставлении студентам и молодым ученым возможности ознакомиться с новейшими научными достижениями по тематике Конференции, а также с уникальным научным оборудованием и лабораторной базой Самарского университета, используемой для реализации современных научных проектов.

УДК 004.9
ББК 32.973

Минимальные системные требования:

PC, процессор Pentium, 160 МГц; оперативная память 32 Мб;
на винчестере 16 Мб; Microsoft Windows XP/Vista/7;
разрешение экрана 1024x768 с глубиной цвета 16 бит;
DVD-ROM2-х и выше, мышь; Adobe Acrobat Reader.

Редактор тома А.В. Никоноров

Выпускающий редактор Д.В. Кудряшов

Подписано для тиражирования 13.07.2022.

Объем издания 5,6 Мб.

Количество носителей 1 диск.

Тираж 11 дисков.

ФЕДЕРАЛЬНОЕ ГОСУДАРСТВЕННОЕ АВТОНОМНОЕ
ОБРАЗОВАТЕЛЬНОЕ УЧРЕЖДЕНИЕ ВЫСШЕГО ОБРАЗОВАНИЯ
«САМАРСКИЙ НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ
УНИВЕРСИТЕТ ИМЕНИ АКАДЕМИКА С.П. КОРОЛЕВА»
(САМАРСКИЙ УНИВЕРСИТЕТ)

443086, САМАРА, МОСКОВСКОЕ ШОССЕ, 34.

Издательство Самарского университета.
443086, Самара, Московское шоссе, 34.

Оглавление

Предисловие	7-10
1. Применение нейросетевых алгоритмов для решения задачи семантической сегментации гиперспектральных данных А.В. Мухин, Д.Н. Грибанов, Р.А. Парингер	040102
2. Адаптация нейросетевой модели распознавания эмоций лиц на основе видеоданных конечного пользователя Е.Н. Чураев, А.В. Савченко	040312
3. Нейросетевой алгоритм для классификации наборов ключевых точек тела человека И.А. Килбас, Д.Н. Грибанов, Р.А. Парингер	040322
4. Исследование методов машинного обучения при анализе добычи и обслуживания нефтяных скважин Е.А. Скачкова, А.Л. Аленин, В.В. Мокшин	040822
5. Полуавтоматическая разметка одноклассовых изображений с применением нейросетевой модели обнаружения объектов В.Н. Гридин, И.А. Новиков, Б.Р. Салем, В.И. Солодовников	040892
6. Нейросетевая классификация сортов кофе на гиперспектральных изображениях А.Р. Макаров, Д.И. Ульянов, А.В. Никоноров, Н.А. Фирсов	040962
7. Использование методов кластеризации текстов на естественном языке в рекомендательных системах Е.Д. Пуговкина, А.А. Белоусов	041022
8. Подход к извлечению и кластеризации библиографической информации А.А. Дырночкин, В.С. Мошкин	041172
9. Гибридный алгоритм классификации кандидатов в термины текста предметной области И.А. Андреев, В.С. Мошкин, Н.Г. Ярушкина	041182
10. Разработка системы гибридного детектирования аномалий временных рядов Д.С. Курило, В.С. Мошкин	041232
11. Использование нейросетей для моделирования человеческой мобильности А.А. Лякишев, А.Ю. Привалов	041272
12. Использование нейросетей для прогнозирования превышения заданного порога длинной очереди сообщений, отправляемых по скоростной магистрали О.Д. Бербасов, А.Ю. Привалов	041353
13. Использование сверточных нейронных сетей для мониторинга безопасности на промышленном объекте А.И. Султанова, Л.М. Шарнин, В.В. Мокшин	041452
14. Сравнение алгоритмов обучения с подкреплением в задаче приобретения навыков передвижения в трёхмерном пространстве Д.А. Козлов	041482
15. Влияние состава наблюдений окружающей среды в задаче приобретения навыков передвижения в трёхмерном пространстве при использовании алгоритмов обучения с подкреплением Д.А. Козлов, В.В. Мясников	041502
16. Морфологический анализ текста с помощью нейронных сетей А.Н. Жданова, А.В. Куприянов, Д.С. Шеренков	041552

17. Исследование и применение свёрточной нейронной сети YOLO для автоматизированного тестирования десктопных и мобильных приложений А.А. Рудь, С.А. Рудь, М.М. Шушкина	041582
18. Автоматическое реферирование текстов В. С. Головизнина	041642
19. Методы предварительной обработки скриншотов десктопных приложений для системы оптического распознавания символов С.А. Рудь, А.А. Рудь, М.М. Шушкина	041752
20. Система управления мышечным тонусом на основе нейронной сети с LIF моделью А.Д. Иванова, Н.В. Харин, Т.В. Балтина, О.А. Саченков	041892
21. Семантическая сегментация рентгенограмм легких на основе нейронной сети класса U-net Н.С. Демин, Н.Ю. Ильясова	041902
22. О компьютерном зрении для мониторинга подводных ограждающих конструкций А.С. Смагин	041932
23. Оценивание координат объектов с использованием сверточных сетей и карт глубины Н.А. Андриянов	041972
24. Мультизадачное обучение интеллектуальных агентов на скрытых представлениях И.Н. Аглоков, К.В. Святков, С.В. Сухов	041982
25. Слияние знаний в импульсных нейронных сетях Д.И. Антонов, С.В. Сухов	042042
26. Hyper-parameter tuning for long short-term memory (LSTM) algorithm to forecast a disease spreading Т. Makarovskikh, М. Abotaleb	042142
27. Automatic classification infectious disease X-ray images based on deep learning algorithms Т. Makarovskikh, А. Salah, А. Badr, Р. Mishra, М. Abotaleb	042302
28. Влияние амплитуды импульса на линейность обновления веса в мемристорах на основе оксида гафния О.О. Пермякова, А.Е. Рогожин, А.В. Мяконьких, Е.А. Смирнова, К.В. Руденко	042312
29. Расчет ошибок коррекции освещения с помощью многомерной гистограммы на основе оригинального стенда А.В. Погадаев, В.А. Васильев, Е.И. Ершов	042322
30. Algorithms for proactive security of industrial systems based on machine learning technologies V. Vasilyev, A. Vulfin, A. Kirillova	042412
31. Стекингвый подход к задаче прогнозирования банкротств Д.В. Полупанов, С.Р. Абдюшева, В.В. Галлямов	042842
32. Метод защиты моделей глубокого обучения цифровыми водяными знаками Ю.Д. Выборнова, Д.И. Ульянов	043062
33. Интерпретация матриц внимания языковых моделей при анализе тональности текстов Д.Э. Пашенко, Е.В. Разова, А.В. Котельникова, С.В. Вычегжанин, Е.В. Котельников	043302
34. Разработка интеллектуальной системы управления ресурсами пищевых предприятий Д.Е. Новичков, С.П. Грачев, Е. Пантелей	043332
35. Разработка прототипа интеллектуальной инсулиновой помпы К.К. Абдулхаликова, П.И. Тутубалин, А.Ю. Александров, С.В. Сотников, И.Н. Урахчинский	043422
36. Обнаружение нарушений целостности поверхностей зданий при помощи нейросетевых алгоритмов Н.А. Фирсов, Л.А. Жердева, Д.А. Жердев, Е.Ю. Минаев	043552

37. Dynamic camera spectral sensitivity estimation
A. Nigmatzyanov, D. Shepelev, V. Vasilev, E. Ershov, M. Tchobanou 043922
38. Распознавание дорожных знаков в интеллектуальной транспортной геоинформационной системе ITSGIS
Н. Клепиков, Т. Михеева 043963
39. Адаптация нейросетевых алгоритмов реконструкции изображений для различных условий экспозиции
В.В. Евдокимова 043602
40. Разработка системы компьютерного зрения для детектирования дефектов на внешней поверхности трубы
Д.В. Шустанов, П.Ю. Якимов 042672

ПРЕДИСЛОВИЕ

Конференция ИТНТ-2022 проводится с целью предоставления возможности научных дискуссий и обсуждения результатов фундаментальных и прикладных исследований в области информационных технологий и нанотехнологий, привлечения молодежи в сферу передовых научных исследований, обмена опытом научно-образовательной деятельности при подготовке ИТНТ-специалистов.

Тематика Конференции ИТНТ-2022 охватывает широкий круг областей применения информационных технологий в науке и высокотехнологичных отраслях промышленности.

Основными направлениями работы Конференции ИТНТ-2022 являются:

Компьютерная оптика и нанофотоника

- дифракционная оптика;
- планарные оптические структуры;
- гиперспектральные системы;
- нанофотоника;
- системы оптической сенсорики, передачи и обработки информации;
- сингулярная оптика.

Информационные технологии дистанционного зондирования Земли

- информационные технологии в проектировании космических аппаратов дистанционного зондирования Земли и полезных нагрузок для них;
- программные и математические решения для управления движением космических аппаратов наблюдения;
- программные и аппаратные средства для получения, обработки и анализа данных, получаемых с космических аппаратов дистанционного зондирования Земли;
- математическое моделирование процессов функционирования космических аппаратов дистанционного зондирования Земли;
- современные проектные решения для создания космических аппаратов мониторинга Земли и околоземного пространства и их группировок, в том числе на базе аппаратов типа CubeSat;
- системы дистанционного зондирования Земли на основе БПЛА.

Распознавание, обработка и анализ изображений

- математические методы цифровой обработки изображений и распознавания образов
- трёхмерное зрение
- биометрические системы на основе изображений
- геоинформационные системы и технологии
- защита и верификация мультимедиа

Искусственный интеллект

- новые подходы, тренды и фундаментальные результаты в сфере искусственного интеллекта и его приложениях к распознаванию образов и анализу изображений, обработке текстов, речевой информации;
- нейросетевые методы и глубокое обучение;
- прикладные технологии искусственного интеллекта в обработке изображений, беспилотном транспорте, производственных и сельскохозяйственных приложениях, медицинских приложениях, экологии, мониторинге окружающей среды и других;
- программные технологии для решения задач искусственного интеллекта – фреймворки, библиотеки, открытые инициативы и сообщества;
- мультидисциплинарные аспекты искусственного интеллекта и машинного обучения.

Науки о данных

- Компьютерные науки:
 - инженерия данных;
 - визуализация данных;
 - математические методы анализа данных;
 - программные платформы и библиотеки для работы с данными;
 - аппаратные средства хранения и обработки данных;
 - высокопроизводительные, параллельные и облачные вычисления, технологии обработки больших данных;
 - базы данных, инструменты и языки для работы с базами данных.
- Прикладные задачи интеллектуального анализа данных:
 - решение актуальных прикладных задач.

Одним из приоритетных направлений работы Конференции ИТНТ-2022 является образовательный аспект, заключающийся в предоставлении студентам и молодым ученым возможности ознакомиться с новейшими научными

достижениями по [тематике](#) Конференции, а также с уникальным научным оборудованием и лабораторной базой [Самарского университета](#), используемой для реализации современных научных проектов.

В рамках Конференции проводится Молодежная школа, где молодые ученые и студенты получают возможность повысить свой профессиональный уровень и [опубликовать](#) свои научные результаты.

В данный сборник трудов вошли материалы по 5 основным направлениям Конференции:

Том 1. Компьютерная оптика и нанофотоника (под редакцией к.ф.-м.н. Е.С. Козловой)

Том 2. Информационные технологии дистанционного зондирования Земли (под редакцией к.т.н. И.С. Ткаченко)

Том 3. Распознавание, обработка и анализ изображений (под редакцией д.ф.-м.н. В.В. Мясникова)

Том 4. Искусственный интеллект (под редакцией д.т.н. А.В. Никонорова)

Том 5. Науки о данных (под редакцией д.т.н. А.В. Куприянова).

Выпускающий редактор томов 1-5: Д.В. Кудряшов.

Официальный сайт Конференции ИТНТ-2022: <http://itnt-conf.org/>

ОРГАНИЗАТОРЫ

- Институт систем обработки изображений РАН (ИСОИ РАН) – филиал ФНИЦ «Кристаллография и фотоника» РАН, г. Самара, Россия;
- Самарский национальный исследовательский университет имени академика С.П. Королева (Самарский университет), г. Самара, Россия.

Организационный комитет

Председатель

Богатырёв В.Д. – д.э.н., профессор, ректор Самарского национального исследовательского университета, г. Самара, Россия

Заместители председателя

Казанский Н.Л. – д.ф.-м.н., профессор, ИСОИ РАН, г. Самара, Россия;

Сергеев В.В. – д.т.н., профессор, Самарский университет, г. Самара, Россия;

Куприянов А.В. – д.т.н., профессор, Самарский университет, г. Самара, Россия.

Ответственный секретарь

Хабибуллин Р.М. – к.т.н., Самарский университет, г. Самара, Россия.

Члены Организационного комитета

Бланк В.А. – к.ф.-м.н., Самарский университет, г. Самара, Россия;

Бояркин Ю.Н. – ИСОИ РАН, г. Самара, Россия;

Востокин С.В. – д.т.н., Самарский университет, г. Самара, Россия;

Гашников М.В. – к.т.н., Самарский университет, г. Самара, Россия;

Гошин Е.В. – к.т.н., Самарский университет, г. Самара, Россия;

Жердев Д.А. – к.т.н., ИСОИ РАН, г. Самара, Россия;

Кадомина Е.А. – Самарский университет, г. Самара, Россия;

Кириш Д.В. – к.т.н., Самарский университет, г. Самара, Россия;

Котов А.П. – к.т.н., ИСОИ РАН, г. Самара, Россия;

Кузнецов А.В. – к.т.н., Самарский университет, г. Самара, Россия;

Куприянов Д.Д. – Самарский университет, г. Самара, Россия;

Максимов А.И. – Самарский университет, г. Самара, Россия;

Мисиевич С.К. – Самарский университет, г. Самара, Россия;

Паренский Н.А. – Самарский университет, г. Самара, Россия;

Подлипов В.В. – ИСОИ РАН, г. Самара, Россия;

Попов С.Б. – д.т.н., проф., Самарский университет, г. Самара, Россия;

Пресняков К.Г. – Департамент информационных технологий и связи Самарской области, г. Самара, Россия;

Рыцарев И.А. – Самарский университет, г. Самара, Россия;

Савельева А.А. – Самарский университет, г. Самара, Россия;

Скиданова Е.Б. – Самарский университет, г. Самара, Россия;

Смагин С.В. – ИСОИ РАН, г. Самара, Россия;

Стафеев С.С. – к.ф.-м.н., ИСОИ РАН, г. Самара, Россия;

Тиц С.Н. – к.т.н., Самарский университет, г. Самара, Россия;

Ткаченко И.С. – к.т.н., доцент, Самарский университет, г. Самара, Россия;

Хнырева Е.С. – Самарский университет, г. Самара, Россия;

Фомченков С.А. – Самарский университет, г. Самара, Россия;

Якуненкова Д.М. – ИСОИ РАН, г. Самара, Россия.

Программный комитет

Председатель

Сойфер В.А. – академик РАН, д.т.н., профессор, президент Самарского университета, Самара, Россия.

Заместитель председателя

Казанский Н.Л. – д.ф.-м.н., профессор, ИСОИ РАН, г. Самара, Россия.

Члены Программного комитета

Korotkova O. – Prof., University of Miami, Coral Gables, USA;
Niemann H. – Prof., Friedrich-Alexander-Universität Erlangen-Nürnberg, Erlangen, Germany;
O’Faolain L. – Prof., Tyndall National Institute, Cork, Ireland;
Sazhin S.- Prof., University of Brighton, Brighton, United Kingdom;
Sobolewski M. – Prof., Polish-Japanese Institute of IT, Warsaw, Poland;
Калачёв Л.В. – профессор, Университет Монтаны, г. Миссула, США;
Коробейников А. – профессор, Педагогический университет Шеньси, г. Сиань, Китай;
Бычков И.В. – академик РАН, д.т.н., профессор, Институт динамики систем и теории управления имени В.М. Матросова Сибирского отделения РАН, г. Иркутск, Россия;
Воеводин В.В. – член-корреспондент РАН, д.ф.-м.н., профессор, Московский государственный университет имени М.В. Ломоносова., г. Москва, Россия;
Головашкин Д.Л. – д.ф.-м.н., профессор, ИСОИ РАН, г. Самара, Россия;
Гуляев Ю.В. – академик РАН, д.ф.-м.н., профессор, Институт радиотехники и электроники им. В.А. Котельникова РАН, г. Москва, Россия;
Желтов С.Ю. – академик РАН, д.т.н., профессор, ГосНИИ авиационных систем, г. Москва, Россия;
Журавлев Ю.И. – академик РАН, д.ф.-м.н., профессор, Вычислительный центр им. А.А. Дородницына РАН, г. Москва, Россия;
Калошин В.А. – д.ф.-м.н., профессор, Институт радиотехники и электроники им. В.А. Котельникова РАН, г. Москва, Россия;
Козлова Е.С. – к.ф.-м.н., ИСОИ РАН, г. Самара, Россия.
Конов В.И. – академик РАН, д.ф.-м.н., профессор, Институт общей физики им. А.М. Прохорова РАН, г. Москва, Россия;
Котляр В.В. – д.ф.-м.н., профессор, ИСОИ РАН, г. Самара, Россия;
Кульчин Ю.Н. – академик РАН, д.ф.-м.н., профессор, Институт автоматизации и процессов управления Дальневосточного отделения РАН, г. Владивосток, Россия;
Куприянов А.В. – д.т.н., профессор, Самарский университет, г. Самара, Россия;
Лабунец В.Г. – д.т.н., профессор, Уральский государственный лесотехнический университет, г. Екатеринбург, Россия;
Мясников В.В. – д.ф.-м.н., профессор, Самарский университет, г. Самара, Россия;
Никитов С.А. – член-корреспондент РАН, д.ф.-м.н., профессор, Институт радиотехники и электроники им. В.А. Котельникова РАН, г. Москва, Россия;
Никоноров А.В. - д.т.н., ИСОИ РАН, г. Самара, Россия;
Новиков Д.А. – чл.-корр. РАН, д.т.н., профессор, Института проблем управления РАН, г. Москва, Россия;
Потатуркин О.И. – д.т.н., профессор, Институт автоматизации и электрометрии Сибирского отделения Российской академии наук, г. Новосибирск, Россия;
Ткаченко И.С. – к.т.н., доцент, Самарский университет, г. Самара, Россия;
Хонина С.Н. – д.ф.-м.н., профессор, ИСОИ РАН, г. Самара, Россия;
Чочиа П.А. – д.т.н., Институт проблем передачи информации им. А.А. Харкевича РАН, г. Москва, Россия.

Применение нейросетевых алгоритмов для решения задачи семантической сегментации гиперспектральных данных

А.В. Мухин
Самарский национальный
исследовательский университет им.
академика С.П. Королева
Самара, Россия
artemmukhinssau@gmail.com

Д.Н. Грибанов
Самарский национальный
исследовательский университет им.
академика С.П. Королева
Самара, Россия
cool.danik01@yandex.ru

Р.А. Парингер
Самарский национальный
исследовательский университет им.
академика С.П. Королева
Институт систем обработки
изображений - филиал ФНИЦ
«Кристаллография и фотоника»
РАН
Самара, Россия
rusparinger@gmail.com

Аннотация—Использование гиперспектральных данных позволяет решать сложные задачи анализа, которые невозможно решить, используя RGB изображения. Гиперспектральные данные часто используются в такой области человеческой жизнедеятельности как сельское хозяйство и агрокультура для наблюдения за растительностью: их степенью влажности и общего здоровья. Успешному применению нейросетевых алгоритмов для анализа гиперспектральных данных препятствует: большая размерность данных, небольшие объемы размеченных данных, влияние условий съемки на качество данных, различие в параметрах и настройках гиперспектральных камер. В данной работе рассматриваются различные подходы к применению нейронных сетей, позволяющих решать задачу семантической сегментации гиперспектральных данных учитывая их специфику. Предложенная в работе архитектура нейронной сети и метод предобработки данных позволил обучить нейронную сеть, превосходящую классические алгоритмы машинного обучения согласно метрике F1.

Ключевые слова— свертка, нейронные сети, сверточная нейронная сеть, гиперспектральные данные.

1. ВВЕДЕНИЕ

Для решения задач анализа гиперспектральных данных обычно применяют либо классические методы машинного обучения, такие как линейные классификаторы, линейные дискриминанты, метод опорных векторов, кластеризация методом k-средних и т.д. Также, для анализа гиперспектральных данных существует отдельный класс методов, основанный на применении различных индексов. Примером такого индекса может быть нормализованный разностный вегетационный индекс NDVI [1]. Данный индекс широко применяется для анализа растений.

Несмотря на все преимущества применения гиперспектральных данных – не все так однозначно. Применение к гиперспектрам классических алгоритмов не всегда позволяет достичь желаемой точности в силу ряда причин: гиперспектральные данные чувствительны к условиям съемки (ветру, освещению); существующие размеченные гиперспектральные наборы данных малочисленны и содержат в себе лишь несколько гиперспектров. Более того, гиперспектральные данные

содержат колоссальное число информации и не все алгоритмы способны обобщать такие данные. Специфичные для гиперспектров индексы также не могут решить все задачи, т.к. для каждой задачи необходим свой индекс [2], нахождение которого нетривиальная задача.

В настоящей работе рассмотрены нейросетевые алгоритмы, как наиболее точные алгоритмы интеллектуального анализа, обладающие высокой способностью к обобщению данных [3].

2. РЕЗУЛЬТАТЫ

В данной работе проведено комплексное исследование влияния архитектур нейронных сетей, методов предобработки данных и конфигурации нейронных сетей на результат семантической сегментации нейронной сети. Модели обучались на наборе из 365 гиперспектров размера 512x512 состоящих из 236 слоев размеченных на 16 классов: яблоня(I), свекла(II), капуста(III), морковь(IV), кукуруза(V), огурец(VI), баклажан(VII), трава(VIII), молочай(IX), овес(X), перец(XI), картофель(XII), щирлица(XIII) (амарант), клубника (XIV), соя (XV) и помидор(XVI).

В эксперименте использовались:

- Нейронная сеть архитектуры Unet [4].
- Нейронная сеть собственной архитектуры (Ours). Данная архитектура была вдохновлена статьёй об архитектуре L2Net [5] и в отличие от Unet имеет существенно меньше обучаемых весов.
- Алгоритм линейной регрессии (LR).
- Случайный лес (RFC).
- Квадратичный дискриминантный анализ (QDA).

Нейронные сети обучались разными способами:

- 1) Обучение нейронной на оригинальных гиперспектрах (Full).
- 2) Обучение нейронной сети на данных, обработанных методом PCA с числом главных компонент равным 17 (PCA).
- 3) Обучение нейронной сети на RGB компонентах гиперспектра (RGB).

Таким образом было исследовано влияние алгоритма уменьшения размерности на результат обучения моделей нейронных сетей.

Исходные гиперспектральные данные были вручную разделены на тренировочную и тестовую выборку. Для обучения нейронных сетей использовалась функция ошибки FocalLoss с параметром $\gamma=5.5$. Количество эпох обучения нейронных сетей 70, используемый оптимизатор – Adam [6]. Шаг обучения контролировался с помощью метода Cosine Annealing With Warm Restart [7] с изначальным значением 0.001 и параметрами $t_0=2$, $t_mult=1$. Размер батча отличался для обеих архитектур в силу ограниченности объема видеопамяти графического процессора. Также при обучении использовалась аугментация данных с помощью поворота на случайный угол и вертикальных\горизонтальных отражений. Нейронные сети были построены и обучались с помощью фреймворка Pytorch.

В таблице 1 представлена сводная таблица результатов проведенных экспериментов в терминах метрики F1.

Таблица I. РЕЗУЛЬТАТЫ ПРОВЕДЕННОГО ЭКСПЕРИМЕНТА В ТЕРМИНАХ МЕТРИКИ F1, УКАЗЫВАЮЩИЕ НА РАЗЛИЧИЯ МЕЖДУ ИССЛЕДУЕМЫМИ АЛГОРИТМАМИ И СПОСОБАМИ ПРЕДОБРАБОТКИ ДАННЫХ

Алгоритмы	Значения метрик F1	
	Среднее	Взвешенное среднее
Unet Full	0,079	0,380
Unet PCA	0,246	0,628
Unet RGB	0,141	0,520
Ours Full	0,027	0,131
Ours PCA	0,584	0,766
Ours RGB	0,083	0,334
LR	0,479	0,582
RFC	0,483	0,589
QDA	0,401	0,441

В таблице 2 представлены результаты эксперимента, где можно сравнить значения метрики F1 классических алгоритмов и “Ours PCA” для каждого класса.

3. ЗАКЛЮЧЕНИЕ

Эффективность применения нейросетевых алгоритмов для решения задач семантической сегментации гиперспектральных данных продемонстрирована в настоящей работе. В ходе проведения эксперимента было выяснено, что нейронные сети способны превзойти по точности классические алгоритмы машинного обучения согласно метрике F1.

Также в данной работе было выяснено, что ключевую роль при применении нейронных сетей играет архитектура сети. Так классическая для задач семантической сегментации архитектура Unet проигрывает в точности классическим алгоритмам. Кроме того, было выяснено, что ключевую роль в подготовке данных для обучения играет метод PCA. Без его применения нейронные сети не могли достичь результатов сопоставимых с классическими методами.

БЛАГОДАРНОСТИ

Результаты исследования были получены при поддержке государственного задания Минобрнауки России Самарскому университету в рамках работ НИЛ-602 "Фотоника для умного дома и умного города" тема 19В-Р001-602 43/21Б.

Таблица II. РЕЗУЛЬТАТЫ ПРОВЕДЕННОГО ЭКСПЕРИМЕНТА В ТЕРМИНАХ МЕТРИКИ F1 ДЛЯ КАЖДОГО КЛАССА, УКАЗЫВАЮЩИЕ НА РАЗЛИЧИЯ МЕЖДУ ИССЛЕДУЕМЫМИ АЛГОРИТМАМИ

Класс	LR	RFC	QDA	NN
Класс нормы (фон)	0,800	0,848	0,396	0,851
Яблоня	0,001	0,001	0,000	0,696
Свекла	0,787	0,709	0,890	0,937
Капуста	0,721	0,641	0,767	0,703
Морковь	0,432	0,381	0,325	0,099
Кукуруза	0,557	0,594	0,559	0,443
Огурец	0,503	0,681	0,393	0,646
Баклажан	0,393	0,280	0,135	0,212
Трава	0,374	0,289	0,130	0,836
Мочай	0,490	0,578	0,662	0,515
Овёс	0,328	0,308	0,177	0,655
Перец	0,267	0,713	0,717	0,894
Картофель	0,752	0,851	0,401	0,828
Ширица	0,860	0,690	0,805	0,907
Клубника	0,557	0,446	0,234	0,703
Соя	0,008	0,000	0,015	0,000
Помидор	0,314	0,196	0,203	0,000
Взвешенное среднее	0,582	0,589	0,441	0,766
Среднее	0,479	0,483	0,401	0,584

ЛИТЕРАТУРА

- [1] Asrar, G.Q. Estimating absorbed photosynthetic radiation and leaf area index from spectral reflectance in wheat 1 / G.Q. Asrar // Agronomy journal. – 1984. – Vol. 76(2). – P. 300-306.
- [2] Парингер, Р.А. Формирование информативного индекса для различения заданных объектов гиперспектральных данных / Р.А. Парингер, А.В. Мухин, А.В. Куприянов // Компьютерная оптика. – 2021. – Т. 45, № 6. – С. 873-878. DOI: 10.18287/2412-6179-CO-930.
- [3] Trajanovski, S. Tongue tumor detection in hyperspectral images using deep learning semantic segmentation / S. Trajanovski // IEEE Transactions on Biomedical Engineering. – 2020. – Vol. 68(4). – P. 1330-1340.
- [4] Ronneberger, O. U-net: Convolutional networks for biomedical image segmentation / O. Ronneberger, P. Fischer, T. Brox // International Conference on Medical image computing and computer-assisted intervention. – Springer, 2015. – P. 234-241.
- [5] Tian, Y. L2-net: Deep learning of discriminative patch descriptor in euclidean space / Y. Tian, B. Fan, F. Wu // Proceedings of the IEEE conference on computer vision and pattern recognition. – 2017. – P. 661-669.
- [6] Kingma, DP. Adam: A method for stochastic optimization / D.P. Kingma, J. Ba // ArXiv, 2014.
- [7] Loshchilov, I. Sgdr: Stochastic gradient descent with warm restarts / I. Loshchilov, F. Hutter // ArXiv, 2016.

Адаптация нейросетевой модели распознавания эмоций лиц на основе видеоданных конечного пользователя

Е.Н. Чураев

Национальный исследовательский университет
Высшая школа экономики
Нижний Новгород, Россия
echuraev@hse.ru

А.В. Савченко

Национальный исследовательский университет
Высшая школа экономики
Нижний Новгород, Россия
avsavchenko@hse.ru

Аннотация—Исследуются способы улучшения качества распознавания эмоций по видео при наличии набора данных с эмоциями конечных пользователей. Используя идею дикторозависимого распознавания речи, предложен новый подход, в котором на первом этапе с использованием набора видео других лиц обучается универсальная нейросетевая модель классификации эмоций, а на втором этапе происходит ее адаптация (дообучение) на основе данных конкретного пользователя. Для систем, нацеленных на работу с большим количеством пользователей, в процессе принятия решения вначале выполняется идентификация лица, после чего эмоции классифицируются с помощью модели, адаптированной под распознанного пользователя. Для набора данных RAVDESS показано, что такой подход позволяет более чем на 20% повысить точность распознавания эмоций.

Ключевые слова— Распознавание эмоций, обработка видео, дообучение нейронных сетей.

1. ВВЕДЕНИЕ

Задача распознавания эмоций заключается в том, что для поступающей на вход последовательности видеок кадров $X(t)$, $t=1,2,\dots,T$, где T – число кадров, требуется поставить в соответствие один из $C>1$ эмоциональных классов (радость, злость и т.п.). В настоящее время технологии распознавания эмоций по видеоизображению лица востребованы во многих областях, например, определение эмоционального состояния водителя для оценки и снижения уровня стресса во время вождения; анализ поведения группы людей в системах видеонаблюдения для предотвращения возможных конфликтных ситуаций; в человеко-машинных интерфейсах для повышения качества понимания состояния пользователя, например, определение реакции покупателя на товар или рекламную компанию. К сожалению, существующие универсальные модели распознают эмоции с точностью 50-70%, что не всегда достаточно для практических приложений. В настоящей работе исследуется возможность повысить точность, если требуется распознавать эмоции лишь ограниченного набора пользователей, при этом для каждого имеется возможность собрать небольшое множество видеоданных с различными эмоциями.

2. ПРЕДЛАГАЕМЫЙ ПОДХОД

Используя идею дикторозависимого распознавания речи, мы предложили новый метод для увеличения точности определения эмоций на видео. На Рис. 1 представлена схема предлагаемого подхода. Алгоритм распознавания эмоций основан на нейросетевом

механизма внимания [1] для дескрипторов лиц, извлеченных из видеок кадров с помощью специальным образом обученных моделей MobileNet и EfficientNet [2]. На этапе обучения базовая дикторонезависимая модель дообучается для каждого пользователя с использованием видеоданных только этого пользователя. В процессе обучения пользователь идентифицируется с помощью известных алгоритмов распознавания лиц, после чего дикторозависимая модель, соответствующая распознанному пользователю, используется для распознавания эмоций. В случае, если пользователь не был найден в базе, то применяется дикторонезависимая модель для определения эмоций на видео.

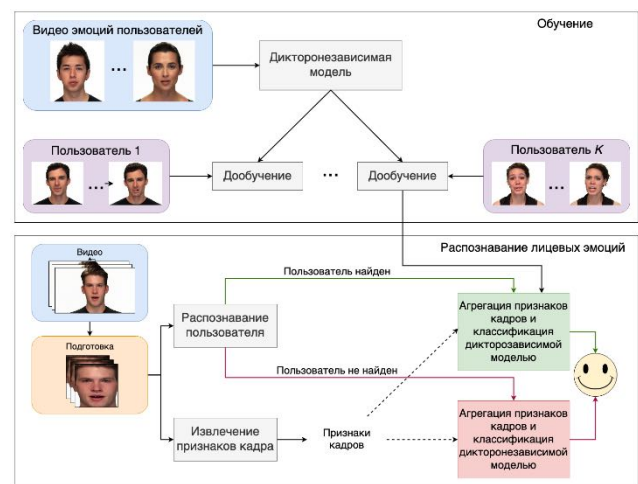


Рис. 1. Предлагаемый алгоритм

В экспериментах мы использовали набор данных RAVDESS, который включает в себя небольшие видео фрагменты, где 24 актёра (12 мужчин и 12 женщин) произносят одну и ту же фразу с разными эмоциями. Для классификации эмоций использовались модель логистической регрессии и три различные реализации механизма внимания. Полученные результаты приведены в Таблице I. Из этих результатов можно сделать вывод, что предложенный подход может применяться для различных архитектур базовой дикторонезависимой модели и позволяет существенно (более чем на 20%) повысить точность распознавания эмоций на видео.

Таблица I. Точность (%) КЛАССИФИКАЦИИ ЭМОЦИЙ ДЛЯ НАБОРА ДАННЫХ RAVDESS

Классификатор	Дикторонезависимая модель	Предложенная адаптация
Logistic regression	72,62	96,61
Single attention	74,17	99,96
Relation attention	73,69	99,76
Self-attention	76,01	99,96

3. ЗАКЛЮЧЕНИЕ

В рамках данной работы на нескольких моделях было показано, что предложенный подход позволяет значительно повысить точность определения эмоций пользователя на видео. Благодаря тому, что для извлечения лицевых признаков, использовались эффективные модели MobileNet/EfficientNet, такой подход может быть реализован на различных энергоэффективных устройствах (в том числе и на

смартфонах) для распознавания эмоций в режиме реального времени. В дальнейшем планируется исследовать возможность повышения точности за счет использования дополнительных модальностей с адаптацией аудиовизуальных нейросетевых моделей.

БЛАГОДАРНОСТИ

Работа выполнено за счет гранта Российского научного фонда (проект № 20-71-10010).

ЛИТЕРАТУРА

- [8] Demochkina, P. Neural network model for video-based facial expression recognition in-the-wild on mobile devices / P. Demochkina, A.V. Savchenko // IEEE International Conference on Information Technology and Nanotechnology (ITNT). – 2021. – P. 1-5.
- [9] Savchenko, A.V. Facial expression and attributes recognition based on multi-task learning of lightweight neural networks / A.V. Savchenko // IEEE International Symposium on Intelligent Systems and Informatics (SISY). – 2021. – P. 119-124.

Нейросетевой алгоритм для классификации наборов ключевых точек тела человека

И.А. Килбас
Самарский национальный
исследовательский университет им.
академика С.П. Королева
Самара, Россия
igor.kilbas.ai@gmail.com

Д.Н. Грибанов
Самарский национальный
исследовательский университет им.
академика С.П. Королева
Самара, Россия
cool.danik01@yandex.ru

Р.А. Парингер
Самарский национальный
исследовательский университет им.
академика С.П. Королева
Институт систем обработки
изображений - филиал ФНИЦ
«Кристаллография и фотоника»
РАН
Самара, Россия
rusparinger@gmail.com

Аннотация—В данной работе представлен нейросетевой алгоритм классификации наборов ключевых точек тела человека. Особенностью данного алгоритма является подход к предобработке данных, позволяющий увеличить обобщающую способность нейронной сети, производящей классификацию. Результаты экспериментов показывают, что предложенный алгоритм обладает высокой точностью и может быть использован в качестве компонента в системах распознавания позы человека.

Ключевые слова— классификация позы, ключевые точки, нейронные сети, распознавание позы.

1. ВВЕДЕНИЕ

Задача понимания человека машинами представляет фундаментальное значение. Благодаря развитию нейросетевых алгоритмов был достигнут существенный прогресс в данном направлении. Особый интерес представляет распознавание позы человека на изображении, поскольку это является неотъемлемой частью понимания состояния человека машинами. Было предложено множество подходов [1-3], большое количество которых применяет нейронные сети.

Задача распознавания позы человека сводится к обнаружению (регрессии) ключевых точек на теле человека и их последующей группировке в отдельные наборы (скелеты). Далее в зависимости от конкретного приложения извлекаются разные типы информации из полученных скелетов: положение человека в кадре, угол наклона конкретной конечности, класс позы. В данной работе рассматривается последняя задача. Представлен нейросетевой алгоритм классификации наборов ключевых точек. Алгоритм разработан таким образом, чтобы быть устойчивым к различным форматам ключевых точек, а также к различного рода трансформациям: сдвиг точек, масштабирование точек.

2. НАБОР ОБУЧАЮЩИХ ДАННЫХ

Для обучения алгоритма был собран набор данных, представляющий собой множество наборов ключевых точек тела человека, размеченные 36 классами («руки на груди», «присед» и т.д.). Пример позы показан на рисунке 1. Классы представляют собой различные упражнения, выполняемые человеком. Наиболее сложными для распознавания классами являются упражнения, включающие касание колен или носков

руками. Далее данные классы будут называться классы позы типа D.

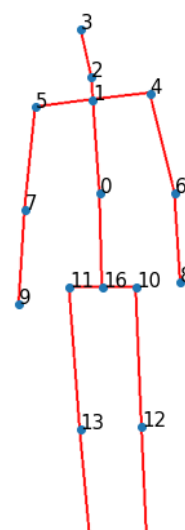


Рис. 1. Пример набора ключевых точек

3. ОПИСАНИЕ АЛГОРИТМА

Алгоритм состоит из двух модулей: модуль предобработки и нейронная сеть. Модуль предобработки отвечает за извлечение признаков из входного набора ключевых точек. Нейронная сеть же производит классификацию вектора извлеченных признаков. Далее будет представлено описание каждого из компонентов.

А. Предобработка точек

Для обеспечения высокого уровня обобщающей способности нейронной сети входные признаки должны содержать только релевантную информацию для исходной задачи. В частности, они должны удовлетворять следующим требованиям:

- 1) Быть инвариантным к масштабу координат ключевых точек.
- 2) Быть инвариантным к сдвигу набора точек.

Для удовлетворения обоих требований на основе набора точек тела человека будет строиться набор нормированных векторов между заданными парами

точек. В качестве таких пар, например, можно взять точки, лежащие на конечностях человека (руки, ноги). Требование 1 выполняется, потому что вектора нормированы (имеют норму, равную единице). Требование 2 выполняется, потому что вектора представляют собой разности значений координат точек, а данная разность не зависит от абсолютного положения точек, делая вектора инвариантными к сдвигу. Помимо выполнения выше озвученных требований, данный набор векторов также инвариантен к пропорциям тела, поскольку не содержит в себе информации о длинах конечностей.

Несмотря на то, что описанный набор векторов удовлетворяет всем требованиям, он не способен запечатлеть наклон торса вперёд, поскольку направления векторов на конечностях не изменяются, меняются лишь их длины до нормировки. Для исправления этой проблемы дополнительно просчитывается относительное расстояние между руками и коленями, а также между руками и стопами. Данное расстояние относительно, поскольку оно будет нормироваться расстоянием между коленом и стопой, дабы не нарушать требование 2.

Б. Нейронная сеть

В качестве нейронной сети используется трёхслойная полносвязная нейронная сеть. Данная нейронная сеть вместе с модулем предобработки данных изображена на рисунке 2.

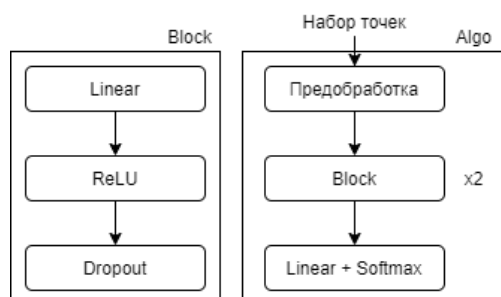


Рис. 2. Схема предложенного алгоритма

Нейронная сеть состоит из двух блоков, каждый блок состоит из слоёв полносвязного слоя (Linear), функции активации и регуляризирующего слоя Dropout. После каждого полносвязного слоя в качестве функции активации используется функция ReLU [4], уже ставшая стандартом во многих приложениях глубокого обучения. Для улучшения обобщающей способности нейронной сети используется Dropout [5], что случайным образом «выключает» нейроны во время обучения и тем самым регуляризирует нейронную сеть, предотвращая переобучение.

После блоков используется еще один полносвязный слой, производящий классификацию, и функция Softmax, что нормирует выходные значения данного слоя, конвертируя их в значения вероятностей принадлежности тому или иному классу.

4. РЕЗУЛЬТАТЫ

Для иллюстрации важности разработанной предобработки наборов ключевых точек было проведено сравнение разных версий алгоритма: исходный алгоритм – версия А, в модуле предобработки не вычисляются относительные расстояния – версия В, вместо модуля предобработки используется нормализация средним и дисперсией – версия С. Также были дополнительно рассчитаны значения точности по наиболее сложным классам поз (позы типа D). Значения точности разных версий алгоритма показаны в таблице 1.

Таблица 1. ЗНАЧЕНИЯ ТОЧНОСТИ КЛАССИФИКАЦИИ ПОЗ РАЗНЫМИ ВЕРСИЯМИ АЛГОРИТМА

Версия алгоритма	Средняя точность по всем классам поз	Средняя точность по классам поз типа D
А	89,8%	78,0%
В	87,2%	63,0%
С	79,9%	51,2%

Применение нейронной сети без модуля предобработки приводит к наименьшей точности, в частности на классах поз типа D. Использование предложенной предобработки, наоборот, позволяет получить наилучшую точность, в том числе более высокую точность на сложных классах поз, что подтверждает её важность.

5. ЗАКЛЮЧЕНИЕ

В данной работе был разработан и исследован нейросетевой алгоритм классификации наборов ключевых точек тела человека. Алгоритм представляет собой многослойную полносвязную нейронную сеть и модуль предобработки входных данных, что делает нейронную сеть инвариантной к ряду разного рода информации во входных данных, тем самым увеличивая её обобщающую способность. Исходя из результатов можно заключить, что предложенный алгоритм классификации наборов ключевых точек может быть использован в качестве компонента в системах распознавания позы человека.

ЛИТЕРАТУРА

- [1] Cao, Z. Realtime multi-person 2d pose estimation using part affinity fields / Z. Cao, T. Simon, S. Wei, Y. Sheikh // Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. – 2017. – P. 7291-7299.
- [2] Artacho, B. OmniPose: A Multi-Scale Framework for Multi-Person Pose Estimation / B. Artacho, A. Savakis // ArXiv, 2021.
- [3] Jin, S. Whole-Body Human Pose Estimation in the Wild / S. Jin // Proceedings of the European Conference on Computer Vision (ECCV), 2020.
- [4] He, K. Delving deep into rectifiers: Surpassing human-level performance on imagenet classification / K. He, X. Zhang, S. Ren, J. Sun // Proceedings of the IEEE International Conference on Computer Vision. – 2015. – P. 1026-1034.
- [5] Srivastava, N. Dropout: a simple way to prevent neural networks from overfitting / N. Srivastava // The journal of machine learning research. – 2014. – Vol. 15(1). – P. 1929-1958.

Исследование методов машинного обучения при анализе добычи и обслуживания нефтяных скважин

Е.А. Скачкова
Казанский национальный
исследовательский технический
университет им А. Н. Туполева -
КАИ
Казань, Россия
elen.rome557@mail.ru

А.Л. Аленин
Казанский национальный
исследовательский технический
университет им А. Н. Туполева -
КАИ
Казань, Россия
freadvice.ru@gmail.com

В.В. Мокшин
Казанский национальный
исследовательский технический
университет им А. Н. Туполева -
КАИ
Казань, Россия
vladimir.mokshin@mail.ru

Аннотация—В данной статье были рассмотрены методы определения значимых признаков, а также методы прогнозирования для оптимизации бурения нефряных скважин, определены их достоинства и недостатки. Для выявления наиболее значимых признаков, влияющих количество добываемой нефти, проводился отбор признаков на основе корреляционного анализа, а также сравнение и идентификация эффективных методов обучения нейронной сети. В результате был выбран метод BRANN (искусственная нейронная сеть с байесовской регуляризацией), так как он показал лучшие результаты при относительно коротком времени обучения. Таким образом, нейронная сеть была обучена прогнозированию количества добываемой нефти на нефтяных скважинах.

Ключевые слова— машинное обучение, корреляционный анализ, искусственная нейронная сеть с байесовской регуляризацией.

1. ВВЕДЕНИЕ

Оптимизация добычи, наряду с управлением резервуаром, является центральной частью стратегии компании по разработке месторождений и добыче полезных ископаемых. Ключевым фактором оптимизации добычи является возможность снижения повреждения пласта при строительстве скважин и проведении регламентных работ. Уменьшение повреждения пласта может быть достигнуто, гарантируя, что эксплуатационные детали будут достигнуты до достижения продуктивной зоны до последних зарегистрированных параметров добычи.

Деятельность по оптимизации производства включает: управление профилем призабойной зоны, газоводяной конус и аппликатура, скважина с максимальным контактом с коллектором и многоствольным заканчиванием, предотвращение отложений органических и неорганических твердых частиц в призабойной зоне/заканчивании/трубопроводе, целостность скважины, предотвращение и устранение разрушения обсадной колонны и цемента, оптимизация производительности механизированной добычи на уровне месторождения и скважины, проектирование наземных сооружений и мощностей по перекачке жидкостей, устранение узких мест производственной системы [1].

Оптимизация производства позволяет повысить производительность существующего поля, и

использование охватывает несколько областей, представляющих интерес. В этом контексте вклад может исходить от применения различных технологий.

2. МЕТОДЫ РЕАЛИЗАЦИИ

Искусственные нейронные сети напоминают нейронную сложность биологического мозга, состоящую из произвольного количества узлов, каждый из которых представляет биологический нейрон. Каждый узел имеет произвольное количество входных и выходных потоков данных, оба из которых соединяют узел с другими узлами в нейронной сети. Каждый узел содержит одну функцию активации, которая вычисляет новое значение для отправки соседним нейронам. Это эквивалентно электрической передаче его биологического аналога [2]. Кроме того, узел также содержит функцию интегрирования, которая преобразует количество входных значений в одно значение. Затем это значение передается в функцию активации.

Два входа подаются в байесовский алгоритм, который представляет собой матрицу $N \times M$ из набора данных. Байесовский классификатор — это простой и эффективный алгоритм классификации, который помогает в разработке быстрых моделей машинного обучения, способных делать быстрые прогнозы. Стандартные сети с обратным распространением менее надежны, чем искусственные нейронные сети с байесовской регуляризацией (BRANN), которые могут свести к минимуму или устранить необходимость длительной перекрестной проверки.

Перед обучением гиперпараметры и веса (w) рассматриваются как случайные величины. Позже это оптимизируется с использованием байесовского подхода к регуляризации. После сбора данных функция плотности для весов может быть изменена с использованием закона Байеса:

$$P(w|D, \alpha, \beta, M) = \frac{P(D|w, \beta, M)P(w|\alpha, M)}{P(D|\alpha, \beta, M)}, \quad (1)$$

где w - вектор весов сети, M - конкретная используемая модель нейронной сети, а D - набор данных.

Поскольку набор для проверки не требуется, итеративная процедура самосогласования должна быть запущена только один раз, чтобы получить «наиболее обобщаемую» модель. Опыт показывает, что меры предосторожности, состоящие в том, чтобы выполнить

всю процедуру пять раз, достаточно, чтобы избежать любого аномального поведения. Это можно противопоставить возможным сотням или тысячам повторных вычислений с нерегулярными искусственными нейронными сетями [5].

Функция правдоподобия весов сети w описывает, насколько вероятно появление данного набора данных при заданном наборе весов:

$$P(w|D, \alpha, \beta, M) = \frac{1}{Z_w(\alpha)Z_D(\beta)} \exp(-(\beta E_D + \alpha E_w)). \quad (2)$$

Метод максимального правдоподобия выбирает веса для максимизации функции правдоподобия, которая в этом гауссовском случае совпадает с минимизацией квадрата ошибки E_D .

Для количественной оценки корреляции между данными в качестве критерия оценки используется коэффициент корреляции в статистике. Вычисляя коэффициент корреляции, можно судить о взаимосвязи между разными выборками и степенью корреляции. Вычисляем коэффициент корреляции Стьюдента и имеем: корреляцию выше 0,3162 и ниже -0,3162 считаем значимой. Таким образом, признаки, имеющие такие значения считаем значимыми [8].

Входными параметрами являются: Номер, Дата, Время, Объект, Товар, Среда, Расход жидкости м3/час, Дозировка ppm, Скорость потока м/с, Частота Гц, Время прохождения жидкости час, Остаточное содержание реагента ppm, Температура оС, pH, O2 (кислород) ppm, Методика определения эффективности, Метод контроля коррозии, Начальная средняя скорость коррозии mpy, Средняя скорость коррозии mpy, Эффективность %.

Выходными параметрами являются: Начальная скорость коррозии mpy, Общая скорость коррозии mpy, Питтинг, Текущая эффективность.

На рисунке (рис. 1) представлены данные, обученные по методу искусственной нейронной сети с байесовской регуляризацией.

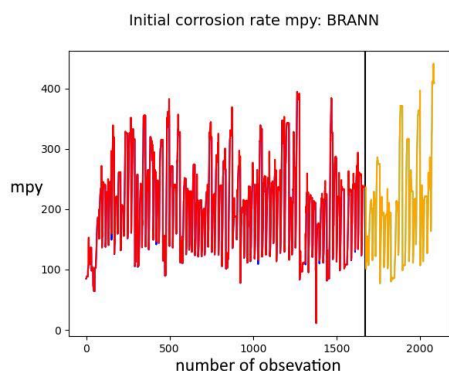


Рис. 1. Начальная скорость коррозии

Берем полученную выборку признаков, чтобы сформировать эффективную модель, основанную на методе искусственной нейронной сети с байесовской регуляризацией. Делим выборку на обучающую (80%) и тестовую (20%), проводим обучение и денормируем данные. Имеем следующие значения точности для каждого из выходных параметров: для параметра Начальная скорость коррозии mpy – 0.95, для параметра Общая скорость коррозии mpy – 0.95, для параметра Питтинг – 0.94, для параметра Текущая эффективность – 0.93. Таким образом, можно говорить о высокой точности выбранного метода.

3. ЗАКЛЮЧЕНИЕ

В работе был проведен анализ существующих методов прогнозирования добычи нефти из нефтяных скважин, предложены различные методы машинного обучения для построения модели прогнозирования показателей результата (характеристик результата) и выбран наиболее оптимальный метод искусственной нейронной сети с байесовской регуляризацией. Проведен анализ методов выделения значимых признаков с использованием метода корреляции. Результаты обучения были получены по каждому методу.

ЛИТЕРАТУРА

- [1] Ticknor, J.L. A Bayesian regularized artificial neural network for stock market forecasting / J.L. Ticknor // *Expert Syst. Appl.* – 2013. – Vol. 14. – P. 5501-5506.
- [2] Xu, L. On convergence properties of the EM algorithm for Gaussian mixtures / L. Xu, M.I. Jordan // *Neural Comput.* – 1996. – Vol. 8. – P. 129-151.
- [3] Peng, H. Feature selection based on mutual information: criteria of max-dependency, max-relevance, and min-redundancy / H. Peng, F. Long, C. Ding // *IEEE Trans Pattern Anal Machine Intell.* – 2005. – Vol. 1. – P. 1226-1238.
- [4] Mokshin, A.V. Adaptive genetic algorithms used to analyze behavior of complex system / A.V. Mokshin, V.V. Mokshin, L.M. Sharnin // *Communications in Nonlinear Science and Numerical Simulation.* – 2019. – Vol. 71. – P. 174-186.
- [5] Yakupov, D.T. Modified spectral clustering method for graphs decomposition. / D.T. Yakupov, V.V. Mokshin // *CEUR Workshop Proceedings.* – 2020. – Vol. 1. – P. 85-90.
- [6] Wali, A.S. Comparative Study of Advance Smart Strain Approximation Method Using Levenberg-Marquardt and Bayesian Regularization Backpropagation Algorithm / A.S. Wali, A. Tyagi // *Materials Today: Proceedings.* – 2020. – Vol. 21. – P. 1380-1395.
- [7] Pan, H. Annealed gradient descent for deep learning / H. Pan // *Neurocomputing.* – 2020. – Vol. 380. – P. 201-211.
- [8] Zhang, T. Adaptive forward-backward greedy algorithm for learning sparse representations / T. Zhang // *IEEE Transactions on Information Theory.* – 2011. – Vol. 57. – P. 4689-4708.
- [9] Mokhtia, M. Feature selection based on regularization of sparsity based regression models by hesitant fuzzy correlation / M. Mokhtia, M. Eftekhari, F. Saberi-Movahed // *Applied Soft Computing.* – 2020. – Vol. 91. – P. 376-381.
- [10] Tutubalin, P.I. The Evaluation of the cryptographic strength of asymmetric encryption algorithms / P.I. Tutubalin, V.V. Mokshin // *Second Russia and Pacific Conference on Computer Technology and Applications.* – 2017. – Vol. 1. – P. 180-183.

Полуавтоматическая разметка одноклассовых изображений с применением нейросетевой модели обнаружения объектов

В.Н. Гридин

Центр информационных технологий
в проектировании РАН
Одинцово, Россия
info@ditc.ras.ru

И.А. Новиков

Центр информационных технологий
в проектировании РАН
Одинцово, Россия
i.novikov@niigb.ru

Б.Р. Салем

Центр информационных технологий
в проектировании РАН
Одинцово, Россия
isub97@gmail.com

В.И. Солодовников

Центр информационных технологий
в проектировании РАН
Одинцово, Россия
vovka_online@mail.ru

Аннотация— В результате резкого роста популярности ресурсозатратных методов искусственного интеллекта возникает серьезная проблема предварительной подготовки данных для эффективного обучения моделей на основе сверточных нейронных сетей. Авторами представлен подход, построенный по принципу итеративного обновления обучающей выборки с применением нейросетевой модели YOLO для обнаружения областей интереса, выделения объектов и автоматизации процесса разметки исходных изображений. Предлагаемый подход был апробирован с различными конфигурациями модели для разметки бактерий на изображениях, полученных с помощью сканирующего электронного микроскопа, и в среднем продемонстрировал показатели точности ~90% в процессе обработки набора данных с увеличением исходной обучающей выборки в 1.75 раз.

Ключевые слова— компьютерное зрение, машинное обучение, обнаружение объектов, разметка данных, предобработка данных.

1. ВВЕДЕНИЕ

Разработки и достижения в области компьютерного зрения и машинного обучения позволили создать основу для апробации и внедрения широкого спектра новых алгоритмических и программных средств для решения прикладных задач в различных предметных областях [1][2][3]. В настоящее время создано множество современных методов нахождения объектов на основе нейронных сетей, что позволяет автоматизировать процесс распознавания образов на изображениях и видео. Однако применение нейросетевого подхода осложняется необходимостью решения задачи предварительной разметки исходных изображений для создания обучающей выборки, причем на дальнейший результат влияет, как объем выборки, так и качество выделения областей интереса. Увеличение числа тренировочных примеров положительно сказывается на точности обнаружения и сегментации объектов [4]. Необходимое количество образцов для обучения является субъективным параметром и варьируется в зависимости от решаемой задачи, специфики исходных данных, характеристик используемой нейросетевой модели. С целью исключения ошибок и улучшения

точности выделения границ объектов процесс первоначальной разметки зачастую выполняется вручную специалистом в исследуемой предметной области и требует значительных временных затрат. В данной работе предложен подход, способный повысить производительность процесса разметки изображений с одним классом объектов при создании обучающей выборки.

2. ИСХОДНЫЕ ДАННЫЕ

Для разметки были выбраны 450 изображений, полученных с помощью сканирующего электронного микроскопа, разрешением 1024x768 с микроорганизмами различного рода. Основной целью являлось выделение областей с бактериями на исходном изображении для дальнейшего использования в обучении нейросетевой модели посредством итеративного пополнения набора тренировочных данных. На первом этапе было решено использовать изображения с наличием одного рода бактерий, что позволило создать первоначальную эталонную базу тренировочных данных. Обучающая выборка исходных изображений была сокращена до 10% от исходной с разделением в пропорции 70/30 на тренировочную и валидационную соответственно.

3. ОПИСАНИЕ ПОДХОДА

В рамках предлагаемого подхода была рассмотрена предварительная разметка изображений с минимальным количеством тренировочных изображений. Основной мотивацией работы являлась минимизация времени разметки тренировочных данных с большим количеством объектов на одиночном изображении. Важной особенностью является возможность промежуточного контроля выделенных областей интереса экспертом в предметной области для корректировки возможных ошибок и улучшения точности выделения границ объектов. В процессе предварительного прототипирования метода автоматизированной разметки все обнаруженные микроорганизмы на изображениях были отмечены как

представители одного домена “бактерии”, тогда как всевозможные небактериологические включения, артефакты, элементы слизистой и т.д. относились к элементам внешнего фона подстилающей поверхности.

На первом этапе на вход модели обнаружения объектов подавались размеченные вручную изображения бактерий. Эти входные данные считались эталонными и составляли обучающую выборку. Тренировка модели осуществлялась на представителях различных родов бактерий, что позволяет составить усредненную карту признаков и в дальнейшем дает возможность находить бактерий иных родов без дополнительного обучения на их представителях. Данный подход схематически изображен на рисунке 1.

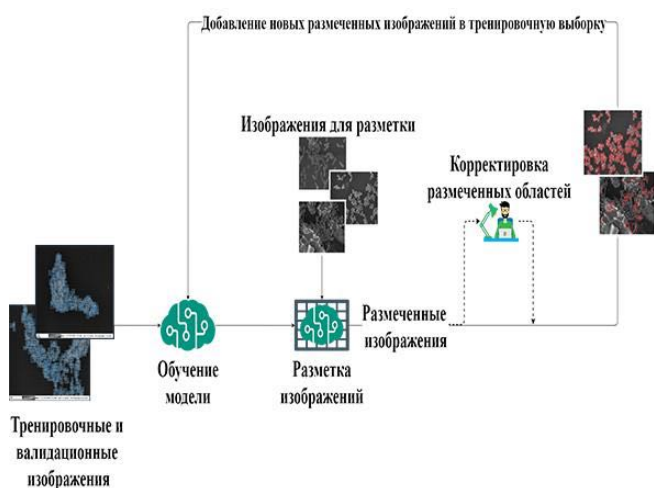


Рис. 1. Общая схема авторазметки одноклассовых изображений

4. ТЕСТИРОВАНИЕ ПОДХОДА

Предлагаемый подход был апробирован с использованием 3 различных сценариев работы модели Yolo:

- модель 1 (M1) обучена на основе YOLOv5l с 50 вручную размеченными изображениями;
- модель 2 (M2) обучена на основе YOLOv5l с 87 изображениями, размеченными, как вручную – 50, так и автоматически моделью M1 – 37,
- модель 3 (M3) дообучена на основе модели M1 с 87 изображениями (50 размечены вручную, 37 размечены предыдущей моделью M1).

5. ЗАКЛЮЧЕНИЕ

Анализ работы предлагаемого подхода автоматизированной разметки и сравнение результатов обучения нейросетевых моделей на вновь полученной расширенной выборке с исходной моделью, обученной на основе вручную размеченных примеров,

продемонстрировал схожие результаты по метрикам оценки качества нахождения объектов, но позволил увеличить объем данных, доступных для последующего использования. В качестве инструментария для обнаружения микробиологических объектов была выбрана модель YOLO v5, однако стоит отметить, что данный подход может быть применен в рамках большинства сверточных нейронных сетей. Осуществлена оценка различных конфигураций выбранной нейросетевой модели, что позволило подобрать наиболее подходящий инструментарий для рассматриваемого подхода. Впоследствии полученные размеченные изображения будут являться входными данными для обучения моделей классификаторов микробиологических объектов. Более того, полученный в ходе исследований подход может быть использован в широком спектре задач, включающих в себя разметку одноклассовых изображений.

Полученные модели были оценены на основе метрик Точность, Полнота, mAP (mean average precision)[4] с областью пересечения 0.5 и [0.5-0.95].

Результаты сравнения трёх вышеуказанных моделей представлены в таблице 1.

Таблица 1. СРАВНЕНИЕ МОДЕЛЕЙ

Метрики качества модели	Модели		
	Модель 1 (M1)	Модель 2 (M2)	Модель 3 (M3)
Точность	0.883	0.909	0.911
Полнота	0.869	0.809	0.817
mAP 0.5	0.896	0.889	0.888
mAP_0.5:0.95	0.542	0.555	0.552

БЛАГОДАРНОСТИ

Работа выполняется в рамках темы № FFSM-2019-0001.

ЛИТЕРАТУРА

- [1] Andriyanov, N.A. Detection of objects in the images: From likelihood relationships towards scalable and efficient neural networks / N.A. Andriyanov, V.E. Dementiev, A.G. Tashlinskiy // Computer Optics. – 2022. – Vol. 46(1). – P. 139-159. DOI: 10.18287/2412-6179-CO-922.
- [2] Kalinina, M.O. Book spine recognition with the use of deep neural networks / M.O. Kalinina, P.L. Nikolaev // Computer Optics. – 2020. – Vol. 44(6). – P. 968-977. DOI: 10.18287/2412-6179-CO-731.
- [3] Gridin, V.N. Predictive data analysis subsystem for patients with subjective and subtle cognitive impairment / V.N. Gridin, B.R. Salem, V.I. Solodovnikov // International Conference on Information Technology and Nanotechnology (ITNT). – 2021. – P. 1-4. DOI: 10.1109/ITNT52450.2021.9649288.
- [4] Liu, L. Mean Average Precision / L. Liu, M.T. Özsu // Encyclopedia of Database Systems. – Boston, MA: Springer US, 2009. – P. 1703-1703.

Нейросетевая классификация сортов кофе на гиперспектральных изображениях

А. Р. Макаров

Самарский национальный исследовательский университет им.
академика С.П. Королева
Самара, Россия
andre.makar1999@gmail.com

Н.А. Фирсов

Самарский национальный исследовательский университет им.
академика С.П. Королева
Самара, Россия
firsov.na98@gmail.com

А. В. Никоноров

Самарский национальный исследовательский университет им.
академика С.П. Королева
Самара, Россия
artniko@gmail.com

Д. И. Ульянов

Самарский национальный исследовательский университет им.
академика С.П. Королева
Самара, Россия
dmitryulyanovhome@gmail.com

Аннотация—В представленной работе предложена модификация существующего метода классификации сортов кофе на основе гиперспектральных изображений в высоком разрешении. В качестве классификатора используется спектрально-пространственная сверточная нейронная сеть с нивелированием влияния условий освещённости сцены на результат классификации. Для компенсации эффекта бликов предложен подход со статистической оценкой формирования обучающей выборки. Показана эффективность предложенного подхода в задаче классификации изображений сортов кофе, полученных сканирующей гиперспектральной камерой.

Ключевые слова— гиперспектральные изображения, сверточные нейронные сети, спектрально-пространственная классификация гиперспектральных изображений.

1. ВВЕДЕНИЕ

Сегодня невозможно отрицать повальную популярность кофе по всему миру. Его сорт является одним из ключевых факторов, влияющих на качество и цену кофе. Идентификация кофейных зерен производится традиционными лабораторными химическими методами на основе реагентов, методами спектроскопии и методами цифровой обработки изображений [1].

Химические методы на основе реагентов требуют много времени и сложны в эксплуатации. Методы спектроскопии и визуализации получили широкое распространение как быстрые, неразрушающие и точные методы. Гиперспектральная визуализация, метод, объединяющий как спектроскопию, так и методы визуализации, привлек повышенное внимание исследователей из разных областей [1].

В отличие от классического RGB-представления изображения, гиперспектральные снимки состоят из гораздо большего числа спектральных компонент – от десятков до нескольких сотен. Как следствие, за счёт большего объёма данных, появляется возможность обнаружить новые признаки, не фиксируемые обычными камерами. Для работы с таким представлением изображений за последние годы разработано множество методов. На волне популярности применения методов глубокого обучения для работы с изображениями, появились методы сегментации и классификации

гиперспектральных изображений (ГСИ) с помощью сверточных нейронных сетей [2,3].

В данной работе представлен подход к решению задачи классификации ГСИ высокого разрешения для определения сортов кофе. В качестве основы для классификатора выступает архитектура [2] с модификациями, нацеленными на повышение устойчивости к изменению условий освещённости фотографируемых объектов.

2. ГИПЕРСПЕКТРАЛЬНЫЕ ДАННЫЕ

Гиперспектральные изображения получены при помощи гиперспектрометра с оптической схемой Оффнера [4, 5]. Отражающая дифракционная решетка с блеском [5] позволяет добиться на матричной сенсоре высокой освещённости, достаточной для получения качественных снимков.

Из-за особенностей поверхности зерен кофе свет от источника может отражаться без каких-либо потерь, создавая блики и засветы на гиперспектральных изображениях. Это создает проблему при обучении сверточных нейронных сетей, внося во все классы искажения со схожими или идентичными спектрами принадлежащими засветам. Для ее решения предлагается использовать пороговую обработку как в работе [2], отталкиваясь от предположения, что спектр бликов близок к максимальной интенсивности по всем длинам волн к источнику освещения, тогда блики можно выделить в отдельный класс по следующему отношению:

$$\begin{cases} X_L = \{HSI_{i,j}\}_{\sum_{k=0}^N HSI_{i,j}^k < LT} \\ X_C = \{HSI_{i,j}\}_{\sum_{k=0}^N HSI_{i,j}^k > LT} \end{cases} \quad (1)$$

где X_L и X_C – множество гиперпикселей классов «блики» и «кофе» соответственно, LT – пороговое значение суммарной интенсивности пиксела. По данным наборам гиперпикселей производится вычисление t-статистики и соответствующему ему значению p-value:

$$t = \left(\frac{\bar{x}_L - \bar{x}_C}{\sqrt{\frac{s_L^2}{n_L} + \frac{s_C^2}{n_C}}} \right) \quad (2)$$

где \bar{X}_L и \bar{X}_C – средние значения гиперпикселей представленных классов; s_L^2 и s_C^2 – оценки дисперсии классов; n_L и n_C – объем выборки.

Для выбора порогового значения интенсивности, на котором будет происходить отсечение бликов, используется наименьший уровень значимости нулевой гипотезы (p-value). График зависимости уровня значимости от выбранного порога приведён на рисунке 1.

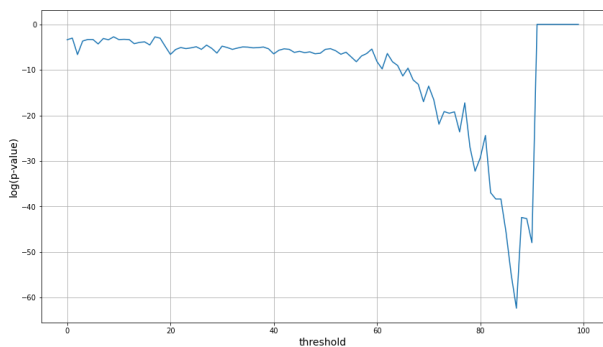


Рис. 1. График уровня значимости нулевой гипотезы

Полученные таким образом гиперпиксели помечались как класс «блики» и накладывались поверх размеченных классов сортов кофе.

3. НЕЙРОСЕТЕВАЯ КЛАССИФИКАЦИЯ

В данной работе используется модификация [2], являющаяся объединением слоёв свёртки для анализа как одномерных, так и двумерных срезов, что делает возможным анализ 3D изображения. Как следствие, архитектура позволяет находить зависимости не только в рамках каждого канала, но и между каналами для каждого гиперпикселя [2].

В качестве модификации архитектуры было предложено не использовать регуляризационный слой dropout. Снижение влияния условий освещённости сцены на снимке достигается за счёт добавления 3d-слоя батч-нормализации [6].

4. РЕЗУЛЬТАТЫ КЛАССИФИКАЦИИ

В данной работе были исследованы два набора гиперспектральных данных – с двумя и тремя сортами кофе на изображении. Показатели accuracy результатов классификации представлены в таблицах 1 и 2 соответственно.

Как видно из представленных данных полученная точность позволяет однозначно определять сорт кофе на гиперспектральном изображении.

Таблица I. РЕЗУЛЬТАТЫ КЛАССИФИКАЦИИ №1 (ACCURACY)

Сорт кофе	Коста-Рика “Тарразу”	Кения “АА Маунт”
Точность классификации	0.829	0.928

Таблица II. РЕЗУЛЬТАТЫ КЛАССИФИКАЦИИ №2 (ACCURACY)

Сорт кофе	Кения “АА Маунт”	Бразилия “Сул-де-Минас”	Коста-Рика “Тарразу”
Точность классификации	0.838	0.772	0.685

5. ЗАКЛЮЧЕНИЕ

В настоящей работе предложена технология классификации сортов кофе на основе гиперспектральной съемки при различных условиях освещенности объектов на снимках. Технология основана на свёрточном классификаторе М3D-DCNN для ГСИ высокого разрешения. Предложенная модификация снижает влияние условий освещённости на качество классификации.

Предложена процедура со статистической оценкой формирования обучающей выборки, позволяющая провести сегментацию бликов на гиперспектральном изображении. Такой подход позволяет увеличить точность распознавания сортов кофе. Результирующая точность классификации составила более 87,85% и 76,5% для двух и трех сортов кофе соответственно.

ЛИТЕРАТУРА

- [1] Zhang, Ch. Identification of coffee bean varieties using hyperspectral imaging influence of preprocessing methods and pixel-wise spectra analysis / Chu Zhang, Fei Liu, Yong He // Scientific Reports. – 2018. – Vol. 8. – P. 2166.
- [2] Firsov, N.A. Neural network-aided classification of hyperspectral vegetation images with a training sample generated using an adaptive vegetation index / N.A. Firsov, V.V. Podlipnov, N.A. Ivliev, P.P. Nikolaev, S.V. Mashkov, P.A. Ishkin, R.V. Skidanov, A.V. Nikonorov // Computer Optics. – 2021. – Vol. 45(6). – P. 887-896. DOI: 10.18287/2412-6179-CO-1038.
- [3] He, M. Multi-scale 3D deep convolutional neural network for hyperspectral image classification / M. He, B. Li, H. Chen // IEEE International Conference on Image Processing (ICIP). – 2017. – P. 3904-3908.
- [4] Podlipnov, V. Experimental determination of soil moisture on hyperspectral images / V. Podlipnov, V. Shchedrin, A. Babichev, S. Vasilyev, V. Blank // Computer Optics. – 2018. – Vol. 42(5). – P. 877-884. DOI: 10.18287/2412-6179-2017-42-5-877-884.
- [5] Karpeev, S. Alignment and study of prototypes of the Offner Hyperspectrometer / S. Karpeev, S. Khonina, A. Murdagulov, M. Petrov // Vestnik of the Samara State Aerospace University. – 2016. – Vol. 15(1).
- [6] Ioffe, S. Batch Normalization: Accelerating Deep Network Training by Reducing Internal Covariate Shift / S. Ioffe, C. Szegedy. – 2015. – URL: <https://arxiv.org/abs/1502.03167>.

Использование методов кластеризации текстов на естественном языке в рекомендательных системах

Е.Д. Пуговкина

Самарский национальный исследовательский университет
им. академика С.П. Королева
Самара, Россия
pugovkinaed@gmail.com

А.А. Белоусов

Самарский национальный исследовательский университет
им. академика С.П. Королева
Самара, Россия
adark@narod.ru

Аннотация—Рекомендательные системы достаточно новое направление в области искусственного интеллекта. Изучение подходов, позволяющих находить предпочтения пользователей, является важной задачей, в частоте, наибольший интерес представляют текстовые данные, генерируемые пользователем или системой. Использование алгоритмов тематического моделирования текстов для поиска схожих пользователей или элементов представляет интерес как с точки зрения обработки естественного языка, так и со стороны построения рекомендательных алгоритмов.

Ключевые слова— рекомендательные алгоритмы, рекомендательные системы, NLP, тематическое моделирование.

1. ВВЕДЕНИЕ

Изучение рекомендательных систем достаточно ново относительно исследования других классических инструментов и методов информационных систем (например, баз данных или поисковых систем). В последние годы интерес к рекомендательным системам возрастает, так как объем данных, генерируемых пользователями, увеличивается, а следовательно, усложняется задача персонализации рекомендательных алгоритмов. Одним из самых популярных генерируемых пользователями контентом является текст, поэтому рассмотрение способов кластеризации текстов на естественном языке как методов поиска схожих пользователей, а следовательно, определение похожих пользователей, является важной задачей. Особый интерес представляют способы кластеризации коротких текстов.

В рамках текущей работы будут рассмотрены способы построения тематической модели корпуса текстов и будет проанализировано их использование в рекомендательных алгоритмах. В статье приводятся как классические модели тематического моделирования текстов, так и более современные подходы, основанные на нейросетевых моделях.

2. АЛГОРИТМЫ ТЕМАТИЧЕСКОГО МОДЕЛИРОВАНИЯ ТЕКСТА

Тематическое моделирование - это метод извлечения скрытых тем из больших объемов текста. На данный момент самым популярным алгоритмом является LDA (Latent Dirichlet Allocation), но существует множество других подходов для разных целей. Например, тематическое моделирование короткого текста - очень сложная задача со специальными алгоритмами. В рамках текущей работы будут рассмотрены следующие

алгоритмы тематического моделирования текста: GSDMM, LDA, LSI, контекстно-нейронные тематические модели.

Кластеризация короткого текста представляет собой сложную проблему из-за его разреженных, многомерных и объемных характеристик. Алгоритм для моделирования коротких текстов, GSDMM, имеет несколько важных свойств: во-первых, GSDMM может автоматически определять количество кластеров, во-вторых, GSDMM предлагает четкий способ сбалансировать полноту и однородность результатов кластеризации, в-третьих, GSDMM быстро сходится, в-четвертых, в отличие от подходов, основанных на модели векторного пространства, GSDMM может справиться с разреженными и многомерными проблемами коротких текстов, в-пятых, подобно тематическим моделям (например, LDA), GSDMM также может получать репрезентативные слова для каждого кластера. Приведённые выше тезисы делают GSDMM хорошим выбором для моделирования тем коротких текстов.

Следующий алгоритм - LDA. Идея латентного размещения Дирихле (LDA) основывается на двух предположениях: человек, который пишет документ, закладывает в текст определенные темы и выбор этой темы, означает целенаправленный подбор слова с определённой вероятностью из некоторого набора слов, относящихся к рассматриваемой теме. В этом случае документ представляется как смесь различных тем. В более общем смысле, LDA-модель помогает объяснить сходство данных посредством группировки свойств этих данных в ненаблюдаемые наборы [1,2].

Третий алгоритм - LSI. LSI является методом уменьшения размерности, который проецирует документы в семантическое пространство более низкой размерности и при этом заставляет документы с аналогичным тематическим содержанием располагаться близко друг к другу в результирующем пространстве. Скрытое пространство создается автоматически на основе совпадения слов в коллекции документов, поэтому степень семантической взаимосвязи между документами в скрытом пространстве будет зависеть от других документов в коллекции [2].

Иное семейство тематических моделей - контекстно-нейронные. В статье [3] авторы объединили контекстуализированные представления с нейронными тематическими моделями. Было обнаружено, что такой подход создает более значимые и последовательные темы, чем традиционные тематические модели набора слов и недавние нейронные модели. Авторы вводят

комбинированную тематическую модель (CombinedTM), чтобы исследовать включение контекстуализированных представлений в тематические модели. Модель строится вокруг двух основных компонентов: нейронной тематической модели ProDLDA [24] и встроенного представления SBERT [4].

В статье [5] авторы предлагают похожую тематическую модель - Zero-Shot Topic Model (ZeroShotTM). Она обучается с использованием представлений входных документов, которые учитывают порядок слов и контекстную информацию, преодолевая одно из основных ограничений моделей BoW. Более того, использование независимого от языка представления документов позволяет выполнять моделирование тем с нулевым выстрелом для невидимых языков. Это свойство важно в условиях нехватки ресурсов, когда мало данных для новых языков.

3. ТЕМАТИЧЕСКОЕ МОДЕЛИРОВАНИЕ ТЕКСТА В РЕКОМЕНДАТЕЛЬНЫХ СИСТЕМАХ

Алгоритмы тематического моделирования текста могут быть использованы при построении как контентных, так и коллаборативных рекомендательных моделей. В зависимости от типа системы персональных рекомендаций появление текстовых данных в системе будет носить различный характер. Построение тематической модели корпуса текстов дает возможность осуществить мягкую кластеризацию пользователей и элементов в системе, а далее, применяя различные метрики текстового сходства, такие как меры семантического сходства с использованием тезаурусов или модели с использованием встраивания слов, можно получать матрицы сходства между пользователями и элементами системы. Также одним из вариантов применения алгоритмов кластеризации текста является ускорение вычислений в системе персональных рекомендаций, так как появляется возможность разбивать пользователей на группы и осуществлять подбор рекомендаций по иным параметрам.

Таблица 1. ЗНАЧЕНИЕ КОГЕРЕНТНОСТИ

Сравн енные алгори тмов	Лучший результат		
	Название модели	Количество тематических топиков	Когерентность
	LSI	14	0,412
	LDA	11	0,406
	GSDMM	30	0,491
	CTM	45	0,663
	ZSTM	30	0,674

4. ВЫЧИСЛИТЕЛЬНЫЙ ЭКСПЕРИМЕНТ

Для проведения вычислительного эксперимента были использованы все модели, рассмотренные в разделе 3 текущей работы. В качестве метрики оценки качества была использована когерентность. Так как выбор оптимального количества тематических топиков во многом отражает качество построенной тематической модели текстов на естественном языке, то для каждого тематического топика была построена модель и оценено ее значение когерентности. Перебор осуществлялся от 10 до 35 тематических топиков. Результаты представлены в таблице 1.

5. ЗАКЛЮЧЕНИЕ

Вопросы, связанные с кластеризацией текста на естественном языке, являются актуальными в связи с колоссальным объемом текстовых данных, генерируемых пользователями в социальных сетях. Подходы и методы, рассмотренные в статье, планируются к апробации над текстовыми данными, полученными из вопросно-ответной системы. В рамках работы произведено сравнение тематических моделей на корпусе текстов и использована мера семантического сходства, представленная языковой моделью RoBERTa, для построения рекомендательной системы с использованием методов тематического моделирования текста на естественном языке.

ЛИТЕРАТУРА

- [1] Вероятностный латентно-семантический анализ // Википедия: свободная энцикл. [Электронный ресурс]. – Режим доступа: https://ru.wikipedia.org/wiki/Вероятностный_латентно-семантический_анализ (19.11.2020).
- [2] Воронцов, К.В. Вероятностные тематические модели / К.В. Воронцов // Профессиональный информационно-аналитический ресурс, посвященный машинному обучению, распознаванию образов и интеллектуальному анализу данных [Электронный ресурс]. — Режим доступа: http://www.machinelearning.ru/wiki/index.php?title=Вероятностные_тематические_модели (21.11.2020).
- [3] Bianchi, F. Pre-training is a Hot Topic: Contextualized Document Embeddings Improve Topic Coherence / F. Bianchi, S. Terragni, D. Hovy // Proceedings of the 59th Annual Meeting of the Association for Computational Linguistics and the 11th International Joint Conference on Natural Language Processing (Volume 2: Short Papers). — 2021. — P. 759-766.
- [4] Reimers, N. SentenceBERT: Sentence embeddings using Siamese BERTnetworks / N. Reimers, I. Gurevych // Proceedings of the Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing (EMNLP-IJCNLP). – 2019. – P. 3982-3992.
- [5] Bianchi, F. Cross-lingual Contextualized Topic Models with Zero-shot Learning / F. Bianchi, S. Terragni, D. Hovy, D. Nozza, E. Fersini // Proceedings of the 16th Conference of the European Chapter of the Association for Computational Linguistics: Main Volume. – 2021. – P. 1676-1683.
- [6] Kadomcev, B.B. Dynamics and the Information / B.B. Kadomcev // Izbrannye trudy: in 6 volumes. – Moscow: “Fizmatlit” Publisher, 2003. – Vol. 2. – P. 508-515.

Подход к извлечению и кластеризации библиографической информации

А.А. Дырочкин
Ульяновский государственный технический университет
Ульяновск, Россия
dyrno4kin@gmail.com

В.С. Мошкин
Ульяновский государственный технический университет
Ульяновск, Россия
postforvadim@ya.ru

Аннотация—В данной статье представлена система извлечения библиографической информации для последующего наукометрического анализа публикаций. Описан алгоритм загрузки и предобработки статей. Предложен подход для формирования научных групп по заданной тематике посредством кластеризации текстов аннотаций статей. Также в работе представлены результаты экспериментов с данными по статьям из научной библиотеки eLibrary.

Ключевые слова— анализ текста, парсинг веб-страниц, векторизация текстов, кластеризация, кластеризация *k-means*.

1. ВВЕДЕНИЕ

Анализ наукометрических данных является важным аспектом при формировании рейтингов научной активности или подборе научных групп по определенной тематике.

Данные о научных публикациях хранятся в цифровых библиографических базах данных. Такие базы помогают отслеживать цитируемость статей, опубликованных в научных изданиях. Также они являются одним из источников получения наукометрических данных, для проведения различных оценочных исследований [1-3].

2. АЛГОРИТМ ФОРМИРОВАНИЯ НАУЧНЫХ ГРУПП ПО ЗАДАННОЙ ТЕМАТИКЕ

В рамках данного исследования был разработан алгоритм извлечения библиографической информации с сайта eLibrary, для последующего наукометрического анализа и формирования научных групп по заданной тематике.

Данный алгоритм включает следующие этапы:

1. Загрузка информации по статьям с сайта eLibrary

Загрузка статей состоит из нескольких этапов. На первом этапе происходит загрузка всех статей по автору в формате: название статьи и уникальный идентификатор. На втором этапе происходит загрузка всех данных по каждой статье (авторы, аннотация, год издания, ключевые слова и др.) [4].

2. Предобработка загруженных статей

Предобработка включает в себя перевод иностранных статей на русский язык, токенизацию и стемминг по методу Портера, который не требует дополнительных словарей, а также фильтрацию по стоп словам. На рисунке 1 представлено описание схемы предобработки статей [5].

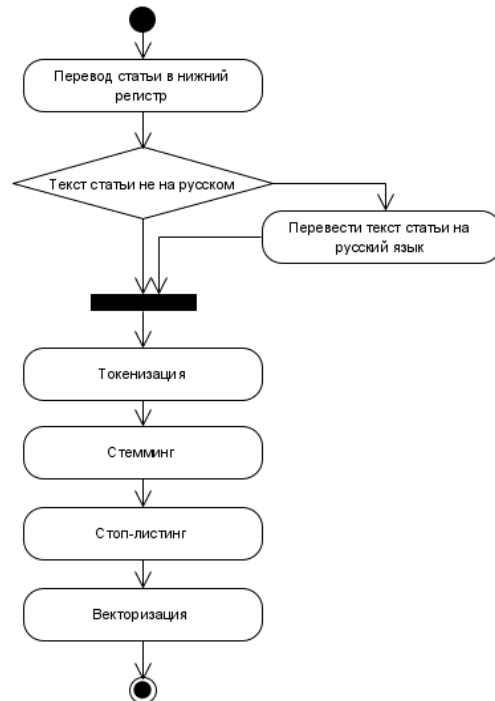


Рис 1. Схема предобработки статей

3. Векторизация предобработанных статей

На данном этапе происходит векторизация предобработанных документов методом TF-IDF.

TF-IDF – мера оценки важности слова в контексте документа, являющегося частью коллекции документов. Вес некоторого слова пропорционален частоте употребления этого слова в документе и обратно пропорционален частоте употребления слова во всех документах коллекции [6].

Частота слова TF рассчитывается по формуле 1.

$$tf(t, d) = \frac{n_i}{\sum_k n_k} \quad (1)$$

где n_i число вхождений слова в документ, а в знаменателе – общее число слов в данном документе.

IDF (обратная частота документа) – инверсия частоты, с которой некоторое слово встречается в документах коллекции. Учёт IDF уменьшает вес часто употребляемых слов [7]. Для каждого слова в пределах одной коллекции документов существует только одно значение IDF. Обратная частота рассчитывается по формуле 2.

$$idf(t, D) = \log \frac{|D|}{|d_i \ni t_i|} \quad (2)$$

где $|D|$ – количество документов в корпусе; $|d_i \supset t_i|$ – количество документов, в которых встречается слово

Таким образом, TF-IDF вычисляется по формуле 3

$$tf - idf(t, d, D) = tf(t, d) \times idf(t, D) \quad (3)$$

4. Кластеризация векторизованных текстов методом k-средних.

Результатом данного этапа является разбиение исходного корпуса текстов, на заданное количество кластеров. Метод k-средних разделяет X документов на k кластеров ($k \leq X$), чтобы минимизировать суммарное квадратичное отклонение точек кластеров от центроидов этих кластеров [8]. Минимальное суммарное отклонение рассчитывается по формуле 4.

$$\min \left[\sum_{i=1}^k \sum_{x(j) \in S_i} \|x^{(j)} - u_i\|^2 \right] \quad (4)$$

где u_i - центроид для кластера S_i

5. Формирование научных групп по заданной тематике.

Для реализации предложенного алгоритма был разработан модуль загрузки статей с сайта eLibrary и модуль предобработки и кластеризации на языке программирования Java.

Общая архитектура системы представлена на рисунке 2.

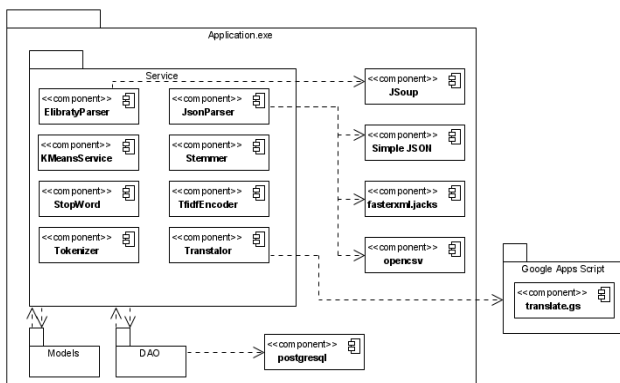


Рис. 2 Общая схема системы

Для проведения кластеризации была загружена библиографическая информация по сотрудникам УЛГТУ, всего было загружено порядка 14 тысяч статей в формате: название статьи и уникальный идентификатор, и в дальнейшем было загружено полное описание 1000 статей для проведения экспериментов.

В результате кластеризации корпус статей был разделен на 12 кластеров. На рисунке 3 представлен результат кластеризации и вывода рекомендаций по формированию научной группы.

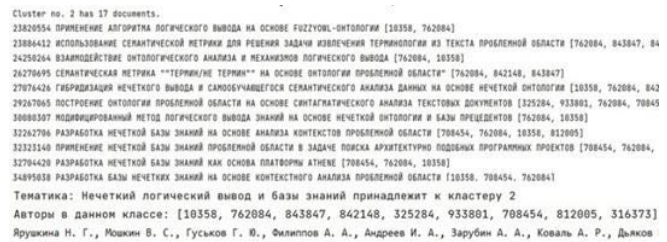


Рис. 3 Результаты кластеризации

3. ЗАКЛЮЧЕНИЕ

Разработанная в рамках исследования система позволяет извлекать библиографическую информацию по авторам из системы eLibrary. Также представлен механизм кластеризации научных статей и формирования рекомендаций при составлении научных групп по заданной тематике.

БЛАГОДАРНОСТИ

Работа выполнена при поддержке Минобрнауки России в рамках проекта № 075-00233-20-05 от 03.11.2020 «Исследование интеллектуального предиктивного мультимодального анализа больших данных и извлечения знаний из различных источников».

ЛИТЕРАТУРА

- [1] Николаев, А.В. Критическая кластеризации научной литературы / А.В. Николаев, В.В. Жуков // Информационно-телекоммуникационные технологии и математическое моделирование высокотехнологичных систем. – 2021. – С. 268-273.
- [2] Низомутдинов, Б.А. Автоматизированный сбор данных для наукометрического анализа / Б.А. Низомутдинов, А.С. Тропников // Научный сервис в сети Интернет. – Федеральный исследовательский центр Институт прикладной математики им. М.В. Келдыша Российской академии наук. – 2019. – Т. 21. – С. 523-531.
- [3] Пархоменко, П.А. Обзор и экспериментальное сравнение методов кластеризации текстов / П.А. Пархоменко, А.А. Григорьев, Н.А. Астраханцев // Труды Института системного программирования РАН. – 2017. – Т. 29, № 2. – С. 161-200.
- [4] Мусаев, А.А. Обзор современных технологий извлечения знаний из текстовых сообщений / А.А. Мусаев // Computer. – 2021. – Т. 13, № 6. – С. 1291-1315.
- [5] Юферев, В.И. Векторизация текстов на основе word-embedding моделей с использованием кластеризации / В.И. Юферев, Н.А. Разин // Моделирование и анализ информационных систем. – 2021. – Т. 28, № 3. – С. 292-311.
- [6] Кравченко, Ю.А. Векторизация текста с использованием методов интеллектуального анализа данных / Ю.А. Кравченко, А.М. Мансур, М.Ж. Хуссайн // Известия Южного федерального университета. Технические науки. – 2021. – № 2(219). – С. 154-167.
- [7] Alam, M. A Review on Clustering of Web Search Result / M. Alam, K. Sadaf // Advances in Computing and Information Technology. Advances in Intelligent Systems and Computing. – Springer, Berlin, Heidelberg, 2013. – Vol. 177.
- [8] Трубников, В.С. Проектирование системы сбора, анализа и визуализации наукометрических данных / В.С. Трубников, К.А. Туральчук // Проблемы современной науки и образования. – 2015. – № 6(36).
- [9] Yarushkina, N. Development of a knowledge base based on context analysis of external information resources / N. Yarushkina, V. Moshkin, A. Filippov // Proceedings of the International conference Information Technology and Nanotechnology. Session Data Science. – Samara, Russia, 2018. – P. 328-337.

Гибридный алгоритм классификации кандидатов в термины текста предметной области

И.А. Андреев
Ульяновский государственный
технический университет
Ульяновск, Россия
ares-ilya@yandex.ru

В.С. Мошкин
Ульяновский государственный
технический университет
Ульяновск, Россия
postforvadim@ya.ru

Н.Г. Ярушкина
Ульяновский государственный
технический университет
Ульяновск, Россия
jng@ulstu.ru

Аннотация—в работе описывается метод классификации кандидатов в термины проблемной области при помощи лингвистических методов с использованием нейронных сетей. Приведен алгоритм работы, представлены результаты экспериментов. По результатам работы достигнут высокий показатель точности.

Ключевые слова— термины, стемминг, нейронные сети, машинное обучение, лингвистика.

1. ВВЕДЕНИЕ

Выделение терминов предметной области – важная задача, которая полезна для широкого спектра задач. В классическом случае составление словаря терминов – это полностью ручной труд эксперта в данной проблемной области. Это занимает очень много времени и высокие материальные затраты, и, кроме того, на итоговый список накладывается отпечаток личного опыта эксперта.

Создание списка терминов без участия эксперта – это задача, решаемая автоматизированной системой классификации кандидатов в термины [1].

2. ГИБРИДНЫЙ МЕТОД КЛАССИФИКАЦИИ КАНДИДАТОВ В ТЕРМИНЫ

Первый этап алгоритма составления списка кандидатов в термины из текста предметной области – получение лингвистических характеристик слов, составляющих данный текст.

Одним из вариантов такой обработки является стемминг. Для разметки текста была выбрана программа *Mystem*, отвечающая необходимым требованиям [2].

Предлагаемый алгоритм автоматизированного составления списка терминов предполагает результат в виде текстового документа, содержащего список терминов, составленных на основе текста при помощи лингвистического метода и отобранных нейронной сетью.

Первое, что делает пользователь – загружает текст на сервер. После загрузки начинается автоматизированная обработка текста. Для начала текст конвертируется в нужную кодировку и обрабатывается программой *Mystem*, которая размечает текст при помощи лингвистических метрик. После разметки загруженный в базу данных текст обрабатывается лингвистическим методом формирования кандидатов в термины [3].

Лингвистический метод извлечения терминологии можно разделить на две части: морфологическую и лексическую.

Морфологический анализ текста:

- определение морфологических признаков и частей речи словоупотреблений;
- определение канонических (начальных, нормализованных) форм слов;
- выделение значимых лексико-грамматических классов [4].

Лексический анализ текста – это извлечение терминов произвольной длины, удовлетворяющих лексико-грамматическим классам [5] [6].

Для возможности выделения терминов из текстов предметной области были разработаны лингвистические шаблоны, с помощью которых удается выделить основные термины. В русском языке синтаксическая структура терминов предметной области более чем в 90 % случаев соответствует следующим пяти шаблонам:

- одиночные сущ.;
- сущ. + сущ.;
- прил. + сущ.;
- прил. + прил. + сущ.;
- сущ. + прил. + сущ.

Вторым этапом является запуск подпрограммы обработки списка терминов при помощи нейронной сети. Нейронная сеть имеет 256 входов. Нейронная сеть представляет собой однослойный перцептрон с прямыми связями, работающий с методом обратного распространения ошибки. Для нормализации входных данных используется функция активации, представленная формулой:

$$R = \frac{1}{1 + \exp(-S)}$$

где R – результат функции активации, S – сумма, получаемая со входов нейронной сети с весами каждого входа [7].

Этот этап запускает пользователь в отдельном приложении. На выходе пользователь получает список прошедших отбор кандидатов в термины. Алгоритм представлен в виде блок-схемы на рисунке 1.

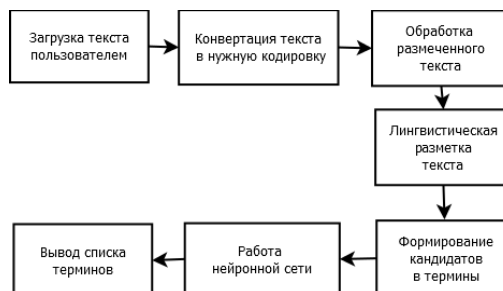


Рис. 1. Алгоритм работы системы

Следующим этапом алгоритма является извлечение терминов на основе обработанных слов текста. Входом для алгоритма является множество существительных, прилагательных, глаголов и служебных частей речи обработанного текста с морфологической информацией о них.

3. РЕЗУЛЬТАТЫ ЭКСПЕРИМЕНТОВ

В качестве входных данных для выполнения эксперимента был выбран текст объемом 40000 слов по теме «Временные ряды». В результате работы системы был получен список терминов, который позволяет более детально изучить полученные результаты экспертом. Перед работой нейронная сеть была обучена на тестовой выборке.

Оцениваются два списка терминов, составленных по одному тексту:

- первый список составлен экспертом;
- второй – программной системой.

При составлении списков использованы одинаковые наборы шаблонов. При нахождении экспертом термина, не подходящего по шаблону, термин игнорируется и не заносится в результирующий список.

Для эксперимента нейронная сеть обучена предварительно на основе списка, содержащего 115 элементов, как являющимися, так и не являющимися терминами.

В результате работы разработанного программного обеспечения и эксперта были сформированы два списка терминов: результат работы нейронной сети, сформированный сервисом, и эталонный – сформированный экспертом.

Общий список терминов, составленный на основе суммы обоих списков терминов, исключая ошибочные, составил 2318 терминов. Характеристики полученных списков представлены в виде таблицы I

Таблица I. ХАРАКТЕРИСТИКИ СПИСОКОВ ТЕРМИНОВ

Источник списка терминов	Количество терминов	Количество ошибок	Необнаруженных терминов
Эксперт	2157	0	52
Разработанное ПО	2286	97	20

На основе полученных данных было рассчитано качество работы программного обеспечения. Процент безошибочных определений терминов составил 100%. Это связано с тем, что, несмотря на некоторое количество ошибок и неопределённых терминов, программное обеспечение определило термины, пропущенные экспертом при вычитывании.

Процент неопределённых терминов, видимых экспертом при вычитывании текста составил 1%. Это позволяет сделать вывод, что разработанная система в

целом может применяться как замена экспертному методу с условием проверки вывода экспертом, т.к. количество ошибочных терминов отлично от 0.

4. ЗАКЛЮЧЕНИЕ

В данной работе рассмотрена проблема автоматизированного формирования списка терминов по тексту предметной области, поднята проблема оценки качества сформированного списка терминов. Был описан процесс автоматизированного составления списка терминов, а также рассмотрен разработанный алгоритм.

Для оценки качества сформированного списка терминов были приглашены эксперты, которые также составили списки терминов по текстам, выбранным для эксперимента. Проведенные эксперименты показали, что результаты работы информационной системы сопоставимы с экспертными результатами, однако уступают им за счёт некоторого количества ошибок, допускаемых системой.

Перспектива исследования может заключаться в интеграции новых лингвостатистических методов и изменении характеристик нейронной сети.

ЛИТЕРАТУРА

- [1] Ковязина, М.А. Извлечение ключевых терминов на базе корпуса текстов о разработке нефтяных и газовых месторождений // Вестник Тюменского государственного университета. Гуманитарные исследования. Humanitates. – 2016. – Т. 2, № 3. – С. 61-69.
- [2] Андреев, И.А. Семантическая метрика терминологичности на основе онтологии предметной области / И.А. Андреев, В.А. Башаев, В.В. Клейн, В.С. Мошкин, Н.Г. Ярушклина // Автоматизация процессов управления. – 2014. – № 4(38). – С. 76-84.
- [3] Kustikova, V.D. A Survey of Deep Learning Methods and Software for Image Classification and Object Detection / V.D. Kustikova, P.N. Druzhkov // Proc. of the 9th Open German-Russian Workshop on Pattern Recognition and Image Understanding, 2014.
- [4] Мошкин, В.С. Семантическая метрика «термин/не термин» на основе онтологии предметной области / В.С. Мошкин, И.А. Андреев, В.А. Башаев, В.В. Клейн // Методы и технологии гибридного и синергетического искусственного интеллекта: материалы I международной Поспеловской летней школы-семинара для студентов, магистрантов и аспирантов. – Калининград: Изд-во БФУ им. И. Канта, 2014. – С. 67-73.
- [5] Лукашевич, Н.В. Использование методов машинного обучения для извлечения слов-терминов / Н.В. Лукашевич, М.Ю. Логачев // Компьютерная лингвистика и семантический Web: по материалам двенадцатой национальной конференции по искусственному интеллекту КИИ, 2010.
- [6] Браславский, П.И. Сравнение четырех методов автоматического извлечения двухсловных терминов из текста / П.И. Браславский, Е.А. Соколов // Компьютерная лингвистика и интеллектуальные технологии. – М.: Изд-во РГТУ, 2006. – С. 88-91.
- [7] Yarushkina, N.G. Hybridization of Fuzzy Inference and Self-learning Fuzzy Ontology-Based Semantic Data Analysis / N.G. Yarushkina, V.S. Moshkin, I.A. Andreev, V.V. Klein, E. Beksaeva // Proceedings of the First International Scientific Conference “Intelligent Information Technologies for Industry” (ITI’16). – Springer International, 2016.

Разработка системы гибридного детектирования аномалий временных рядов

Д.С. Курило

Ульяновский государственный технический университет
Ульяновск, Россия
dimokkurilo@gmail.com

В.С. Мошкин

Ульяновский государственный технический университет
Ульяновск, Россия
postforvadam@ya.ru

Аннотация—В работе описываются результаты разработки программной системы, реализующей метод поиска аномалий с помощью нейронной сети и их последующий анализ, основанный на интеллектуальных алгоритмах. Предложенный подход гибридации методов поиска аномалий позволяет восполнить недостаток объяснимости нейронных сетей данными из онтологии.

Ключевые слова— временной ряд, нейронная сеть, онтология, поиск аномалий.

1. ВВЕДЕНИЕ

Выявление аномалий – область интеллектуального анализа данных, позволяющая находить выделяющиеся из общей массы значения. Эти значения необходимы во многих предметных областях и могут говорить о различных проблемах, сбоях или данных, на которые стоит обратить внимание. Для поиска используются различные технологии, такие как кластерный анализ, скрытые марковские модели, нейронные сети и так далее [1]. Они работают с различной эффективностью, зависящей от затрагиваемой предметной области [2].

2. МОДЕЛЬ, СТРУКТУРА И ОПИСАНИЕ СИСТЕМЫ ПОИСКА АНОМАЛИЙ

Разработанный программный продукт состоит из нескольких модулей. Структура информационной системы показана на диаграмме компонентов (рис 1).

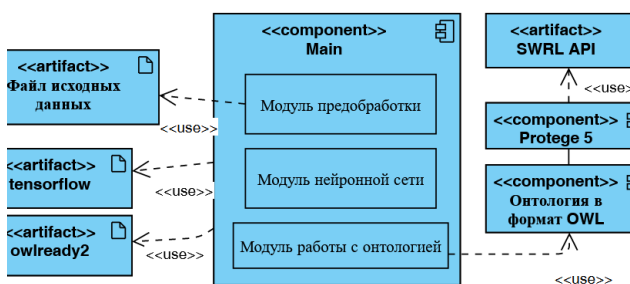


Рис. 1. Диаграмма компонентов

Модули информационной системы используют подключаемые библиотеки:

- owlready - для взаимодействия с онтологией на языке Python;
- tensorflow для создания и управления нейронной сетью;
- SWRL API позволяет взаимодействовать с правилами через интерфейс программы Protégé 5.

Модуль предобработки загружает исходные файлы, получает заголовки столбцов и форматирует данные в DataFrame с необходимыми столбцами. Далее

добавляются временные метки в качестве индексов строк в таблице.

В исходных данных даты замеров отсутствуют, но происходят последовательно, через равные промежутки времени. Принято, что дата проставляется, начиная с текущего дня до окончания строк таблицы с шагом один день [3-5].

В качестве тестовых данных использовались данные производительности нефтяных вышек. Данные разбиваются на несколько независимых друг от друга выборок по имени вышки, на которой происходили замеры. Это позволяет иметь несколько массивов данных, подходящих для последующих экспериментов. Полученные данные при необходимости возможно отобразить на графике. Готовые данные передаются в модуль нейронной сети. Базовый алгоритм анализа состоит из следующих этапов:

- Разбиение исходных данных на тестовую и обучающую выборки. Используется коэффициент разделения 0.5 с обучающей выборкой в первой половине данных.
- Вычисление среднего и стандартного отклонения данных. С помощью этих значений нормализуются исходные данные (рис.2).

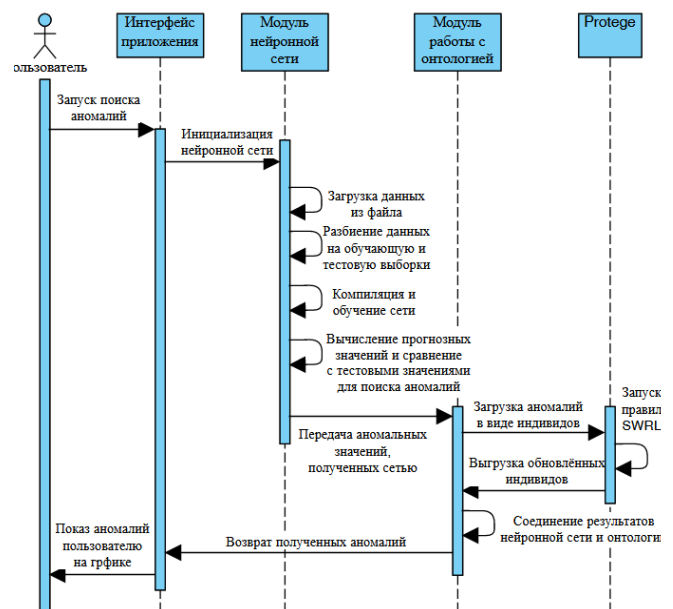


Рис. 2. Диаграмма активностей

- Построение последовательной модели keras. В неё добавляются необходимые слои нейронной сети: входные, свёрточные и рекуррентные [6]. Когда модель построена, задаётся оптимизатор и запускается

обучение модели. Модель обучается на части выборки, полученной ранее.

- С помощью готовой модели генерируется предсказание на диапазон тестовой выборки. Для полученных данных используется оценка модели MAE [7]. Значения, отклоняющиеся от предсказания, являются аномальными. Эти данные передаются в модуль онтологий.

3. ОНТОЛОГИЧЕСКИЙ АНАЛИЗ РЕЗУЛЬТАТОВ

В модуле онтологий содержатся:

- классы измерений;
- свойства измерений;
- правила.

Правила содержат атомы класса и несколько необходимых свойств: аномального значения и дополнительного параметра, по которому подтверждается корректность суждения об аномальности [8].

Цель состоит в проверке достижения параметром константы. Если значение не достигнуто, то замер может считаться аномальным, иначе значение корректно. Проверка осуществляется с помощью встроенных функций библиотеки `sqrlb`. При истинности всех предикатов индивиду в свойство аномальности ставится `true`.

Загрузка и выгрузка индивидов в онтологию в системе происходит с помощью библиотеки `owlready2`. Она позволяет получить из OWL-файла структуру и объекты онтологий и взаимодействовать с ними на языке `python`.

Для загрузки индивидов в онтологию используется класс и свойства, уже описанные в структуре онтологий. После загрузки всех аномалий, происходит пересохранение онтологий и запуск правил. После обработки из файла можно получить все индивиды. В цикле проверяется флаг аномальности и отбираются все значения, принятые онтологией за действительно аномальные (рис.3).

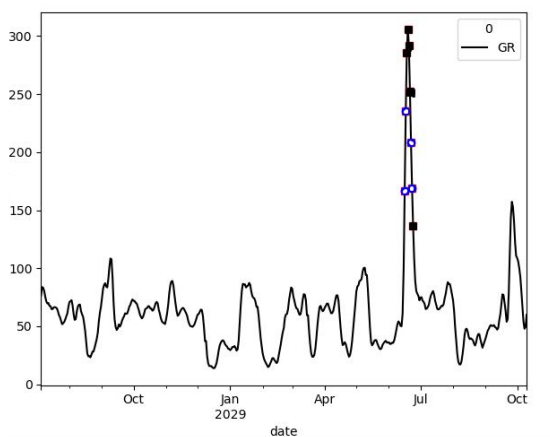


Рис. 3. Результаты работы системы

Чёрные точки - это аномалии, найденные с помощью нейронной сети, но не получившие подтверждения в онтологии. Точки с пустотой в центре - это значения,

подтвержденные правилами онтологии и действительно являющиеся аномалиями.

Для оценки качества модели использовалась метрика средней абсолютной ошибки (MAE). Её значение достигает 80%, что может говорить об обнаружении моделью большинства отклонений и обработкой этих значений онтологией.

4. ЗАКЛЮЧЕНИЕ

Разработанная система обрабатывает исходные данные временного ряда и выдаёт набор аномальных значений. Система использует технологии нейронных сетей и онтологический анализ. Нейронная сеть находит аномалии с помощью построения предсказания, на основании обучающей выборки и сравнения с тестовыми данными. Онтология использует заранее разработанный класс и SWRL правила.

Дальнейшими планами по разработке является усложнение структуры нейронной сети для получения наиболее оптимальных результатов от этого модуля системы. Для онтологии улучшением является добавление классов метрик, которое позволит оперировать переменными, не привязанным к названиям конкретных параметров, а имеющие связи с классами метрик. В следствие этого расширение системы будет заключаться только в добавлении новых классов метрик и написании новых правил.

БЛАГОДАРНОСТИ

Работа выполнена при финансовой поддержке Минобрнауки России в рамках проекта № 075-00233-20-05 от 03.11.2020 «Исследование интеллектуального предиктивного мультимодального анализа больших данных и извлечения знаний из различных источников».

ЛИТЕРАТУРА

- [1] Линдигрин, А.Н. Анализ специфики и проблематики процессов поиска аномалий в сетевых данных / А.Н. Линдигрин // Известия ТулГУ. Технические науки. – 2021. – № 5. – С. 304-309.
- [2] Гасанов, В.И. Выявление аномалий в сетевом трафике на основе нейросетевого моделирования динамики изменения объёмов IP-пакетов / В.И. Гасанов // ММС. – 2018. – № 2.
- [3] Зуев, В.Н. Обнаружение аномалий сетевого трафика методом глубокого обучения / В.Н. Зуев // Программные продукты и системы. – 2021. – № 1. – С. 91-97.
- [4] Wen, T. Time series anomaly detection using convolutional neural networks and transfer learning / T. Wen, R. Keyes // ArXiv preprint: 1905.13628, 2019.
- [5] Zhong, C. Anomaly Detection and Sampling Cost Control via Hierarchical GANs / C. Zhong, M.C. Gursoy, S. Velipasalar // Globecom IEEE Global Communications Conference. – 2020. – P. 1-6.
- [6] Мошкин, В.С. Система онтологического анализа временных рядов / В.С. Мошкин, Н.Г. Ярушклина // Автоматизация процессов управления. – 2014. – № 2(36). – С. 78-85.
- [7] Malhotra, P. Long Short Term Memory Networks for Anomaly Detection in Time Series / P. Malhotra, L. Vig, G.M. Shroff, P. Agarwal. – ESANN, 2015.
- [8] Мошкин, В.С. Особенности интеграции механизмов логического вывода в онтологическую модель представления знаний с помощью SWRL-правил / В.С. Мошкин, Н.Г. Ярушклина // 14 национальная конференция по искусственному интеллекту с международным участием КИИ: труды конференции. – Казань: Изд-во РИЦ «Школа», 2014. – Т. 1. – С. 173-181.

Использование нейросетей для моделирования человеческой мобильности

А.А. Лякишев

Самарский национальный исследовательский университет
им. академика С.П. Королева
Самара, Россия
aqua_phoenix@mail.ru

А.Ю. Привалов

Самарский национальный исследовательский университет
им. академика С.П. Королева
Самара, Россия
privalov1967@gmail.com

Аннотация—В докладе представлен подход к моделированию человеческой мобильности с помощью нейронных сетей.

Ключевые слова—моделирование мобильности, путевые точки, триангуляция Делоне, нейросети.

1. ВВЕДЕНИЕ

Моделирование человеческой мобильности продолжает оставаться одной из актуальных задач, возникающих во многих практических приложениях. Одной из популярных моделей человеческой мобильности является модель SLAW (Self-Similar Least Action Walk) [2], позволяющая учесть сразу несколько ключевых черт реальной мобильности, таких как кластеризация т.н. путевых точек (то есть, точек на маршруте движения, в которых объект проводит много времени), объединяемых в локации, а также статистические характеристики длин и времён переходов между локациями и пребывания в локациях.

Однако, модель SLAW, в которой многие параметры модели описываются эмпирическими формулами, требующими подбора ряда параметров по реальным данным, не обладают желательной точностью моделирования упомянутых выше статистических параметров. В представляемом докладе предлагается использовать обучаемую на реальных трассах мобильности нейросеть, которая в дальнейшем будет использоваться для моделирования трасс мобильности.

2. ОСНОВНАЯ ЧАСТЬ

Кратко опишем основные элементы модели мобильности. Процесс движения объекта описывается, как прямолинейные переходы между путевыми точками и остановки на некоторое время в путевых точках. Путевой точкой является круг радиусом 5 метров, в котором человек проводит более 30 секунд времени. Если построить путевые точки для большого множества объектов, перемещающихся по некоторой местности, то, как правило, получаются скопления путевых точек – локации, т.е. области, включающие близкие путевые точки различных трасс. Локации определяются как транзитивные замыкания точек, находящихся друг от друга на некотором расстоянии. В данной работе это расстояние было равно 30 метрам. Для задания параметров модели необходимо определить границы локаций, которые решено представлять в виде совокупностей треугольников, определяемых с помощью триангуляции Делоне [6].

После предварительной обработки каждая реальная трасса представляется в виде последовательности локаций, в каждой из которых объект проводит некоторое

время, перемещаясь между путевыми точками внутри локации, после чего перемещается в следующую локацию, при этом для каждого перемещения определяются номер следующей локации, длина и скорость. Данные трассы характеризуются с помощью следующих характеристик:

- 1) Распределение длин прыжков между путевыми точками;
- 2) Распределение времени перемещения между путевыми точками;
- 3) Распределение времени пауз в конечной путевой точке прыжков;
- 4) Распределение времени прохождения трасс путевых точек;
- 5) Распределение посещений локации;
- 6) Распределение посещения уникальных локаций.

Исходная выборка делится на две части: обучающую, в которую попадает 80% путевых точек каждой трассы, и проверочную, состоящую из совокупности последних 20% путевых точек каждой трассы.

В работе используется комбинация из нескольких нейронных сетей для генерации этих величин на основе реальных данных:

7) Нейронная сеть для генерации номера очередной локации. На вход сети подаются номера нескольких предыдущих локаций, на выходе сети получается номер следующей локации. В качестве функции активации используется функция softmax , а в качестве функции потерь – категориальная кроссэнтропия [7]. Данное сочетание позволяет быстро кодировать вектор выходных значений, в котором по индексу желаемого выхода находится 1, а остальные значения равны 0 [8]. Сеть состоит из трех слоев:

- входной слой, количество нейронов которого определяется по формуле:

$$M = \text{window} * N, \quad (1)$$

где N – количество локаций в обучающем наборе; window – размер «скользящего окна».

- скрытый слой из 80 нейронов;
- выходной слой, количество нейронов которого равно N .

8) Сеть для генерации параметров времени прыжка и времени паузы. На вход сети подаются параметры нескольких предыдущих прыжков, на выходе сети получаются параметры следующего прыжка. Параметры времени представляют из себя значения функции распределения Леви [9]. За 0 принимается минимальное

допустимое значение, за 0.5 – максимальное. С помощью коэффициента масштабирования значения функции распределения в интервале (0.5; 1) позволяют получить значения времени не только из обучающего набора, но и превосходящие их. Сеть состоит из трех слоев:

- входной слой, количество нейронов которого определяется по формуле:

$$M = \text{window} * N + (\text{window} - 1) * 2 \quad (2)$$

- скрытый слой, состоящий из 80 нейронов;
- выходной слой, состоящий из 2 нейронов.

9) Сеть, определяющая момент остановки движения (окончания трассы). На вход сети подаются номера локаций нескольких предыдущих прыжков, значения высот столбцов гистограмм распределения длин трасс и времени прохождения трасс, а также суммарное время прохождения трассы. На выходе сети получается вероятность остановки генерации. Сеть состоит из 4 слоев:

- входной слой, количество нейронов которого определяется по формуле:

$$M = \text{window} * N + T + P + 1, \quad (3)$$

где T – количество интервалов гистограммы распределения длин трасс; P – количество интервалов гистограммы распределения времени прохождения трасс.

- два скрытых слоя по 40 нейронов в каждом;
- выходной слой с единственным нейроном.

В качестве функции активации выходных слоев сети для генерации параметров прыжка и сети, определяющей момент остановки движения, используется сигмоидальная функция. В качестве функции потерь используется среднеквадратическое отклонение. Для всех сетей в качестве оптимизатора используется оптимизатор Adam [10].

Параметры сгенерированных трасс сравниваются с параметрами проверочной выборки, чтобы убедиться в адекватности полученных данных. На рисунках 1-3 представлены результаты экспериментальных исследований применения нейронных сетей для генерации трасс.

3. ЗАКЛЮЧЕНИЕ

В работе исследована применимость нейронных сетей для работы с моделью человеческой мобильности. В ходе были определены оптимальные структуры таких нейронных сетей и их параметры. Продемонстрированы результаты сравнения характеристик реальных и сгенерированных нейросетями параметров.

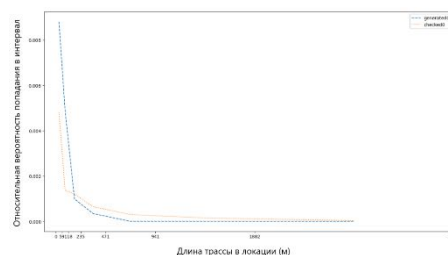


Рис. 1. График распределения длин трасс в локациях

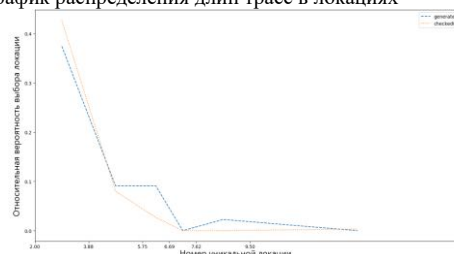


Рис. 2. График распределения уникальных локаций

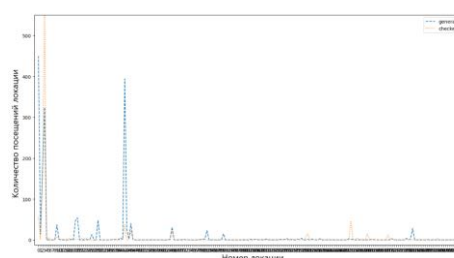


Рис. 3. График количества посещений локаций

ЛИТЕРАТУРА

- [1] Rhee, I. On the Levy-Walk Nature of Human Mobility / I. Rhee, M. Shin, S. Hong, K. Lee, S.J. Kim, S. Chong // IEEE/ACM Transactions On Networking. – 2011. – Vol. 19(3). – P. 630-643.
- [2] Lee, K. SLAW: Self-Similar Least-Action Human Walk / K. Lee, S. Hong, S.J. Kim, I. Rhee, S. Chong // IEEE/ACM Transactions On Networking. – 2012. – Vol. 20(2). – P. 515-529.
- [3] Самуйлов, К.Е. Применение моделей случайного блуждания при моделировании перемещения устройств в беспроводной сети / К.Е. Самуйлов, Ю.В. Гайдамака, С.Я. Шоргин // Информатика и её применение. – 2018. – Т. 12, № 4. – С. 2-8. DOI: 10.14357/19922264180401.
- [4] Mushuang, L. Analysis of the Random Direction Mobility Model with a Sense-and-Avoid Protocol / L. Mushuang, W. Yan, L. Frank // IEEE Globecom Workshops (GC Wkshps). – 2017. – P. 1-6. DOI: 10.1109/GLOCOMW.2017.8269071.
- [5] Barbosa-Filho, H. Human Mobility: Models and Applications / H. Barbosa-Filho, M. Barthelemy, G. Ghoshal, C.R. James, M. Lenormand, T. Louail, R. Menezes, J.J. Ramasco, F. Simini, M. Tomasini // Physics Reports. – 2018. – Vol. 734. – P. 1-74.
- [6] Скворцов, А.В. Триангуляция Делоне и её применение / А.В. Скворцов. – Томск: Изд-во Том. ун-та, 2002. – 128 с.
- [7] Cross-Entropy Loss Function [Electronic resource]. – Access mode: <https://towardsdatascience.com/cross-entropy-loss-function-f38c4ec8643e> (20.01.2022).
- [8] Killer Combo: Softmax and Cross Entropy [Electronic resource]. – Access mode: <https://levelup.gitconnected.com/killer-combo-softmax-and-cross-entropy-5907442f60ba> (20.01.2022).
- [9] Распределение Леви [Электронный ресурс]. – Режим доступа: [https://studme.org/189322/matematika_himiya_fizik/raspredenie_le vi](https://studme.org/189322/matematika_himiya_fizik/raspredenie_le_vi) (27.01.2022).
- [10] Реализуем и сравниваем оптимизаторы моделей в глубоком обучении [Электронный ресурс]. – Режим доступа: <https://habr.com/ru/company/skillfactory/blog/525214/> (27.01.2022).

Использование нейросетей для прогнозирования превышения заданного порога длинной очереди сообщений, отправляемых по скоростной магистрали

О.Д. Бербасов

Самарский национальный исследовательский университет
им. академика С.П. Королева
Самара, Россия
olegoleg.ru@mail.ru

А.Ю. Привалов

Самарский национальный исследовательский университет
им. академика С.П. Королева
Самара, Россия
privalov1967@gmail.com

Аннотация—В работе предложен подход к предсказанию превышения заданного порога значением длины очереди пакетов трафика, отправляемых по высокоскоростной магистрали.

Ключевые слова— прогнозирование, очередь, трафик, нейросети.

1. ВВЕДЕНИЕ

Производительность компьютерной сети является наиболее значимым параметром как для поставщика услуг связи, так и для потребителей их услуг. При проектировании таких сетей важно учитывать множество параметров, одним из которых является нагрузка. Нагрузка это один из параметров, которые влияют на отказоустойчивость и стабильность работы сети. Использование моделирования на этапе проектирования позволяет оценить оптимальную конфигурацию сети и нагрузку на неё в процессе работы, благодаря чему становится проще прогнозирование отказов и оценка её дальнейшей модернизации.

Для создания адекватных моделей необходимо анализировать реальные записи трафика, собранные в различных условиях. При этом, обнаруживается, что реальный трафик обладает особенностями, которые не встречаются у классических моделей теории телетрафика [1,2].

Подобными исследованиями и сбором данных занимаются различные организации по всему миру [2-4], одним из которых является CAIDA [5]. Центр прикладного анализа интернет-данных (CAIDA) проводит сетевое исследование и создаёт исследовательскую инфраструктуру для поддержки крупномасштабного сбора, обработки и распространения данных в научно-исследовательском сообществе. Данные, полученные с этого ресурса, используются для анализа характеристик реальных трасс и для обучения нейронных сетей.

2. ОСНОВНАЯ ЧАСТЬ

В работе используются записи трафика из базы CAID, которые записывались в 2016 году в Чикагском дата-центре Equinix. В процессе обработки сырых данных, отсеиваются все данные, кроме временных меток и значений полезной нагрузки сообщений на уровне IP-пакета, а также время, в течении которого группируются сообщения.

По этим данным были сформированы выборки, которые состоят из временных отметок и суммы объемов полезной нагрузки, переданных за это время сообщений.

Для каждой выборки были рассчитаны выборочное среднее, выборочная средняя дисперсия, корреляционная и автокорреляционные функции. Эти данные в дальнейшем будут использоваться в процессе подготовки обучающих наборов для нейронной сети.

Перед созданием обучающих наборов, требуется преобразовывать записи трафика в запись результатов моделирования ухода сообщений из дискретной очереди. Значимой величиной здесь будет являться объём данных в очереди в каждый момент времени.

Основным параметром для построения модели дискретной очереди будут являться средний уход из очереди и значения входного потока в каждый момент времени. При построении модели дискретной очереди средний уход C будем определять следующим образом:

$$C = MX / k, \quad (1)$$

где C – средний уход, k – коэффициент загрузки, а MX – выборочное среднее входного потока.

Моделируя процесс ухода сообщений из очереди получим значение длины очереди в каждый момент времени. Полученная выборка значений длин очереди будет использоваться для формирования обучающих примеров. В данной работе, значение порога, превышение которого будет предсказывать нейронная сеть, берется равным выборочному среднему этой выборки.

3. ИСПОЛЬЗОВАНИЕ НЕЙРОННОЙ СЕТИ

Для решения задачи, описанной выше, были рассмотрены различные виды нейронных сетей, которые по существующему опыту применения, чаще всего рекомендуются для решения задач прогнозирования.

В работе используется нейронная сеть, состоящая из четырех слоёв:

- первый слой (входной) который, содержит столько нейронов, сколько отсчетов входит в окно прогнозирования;
- второй слой (скрытый) – это LSTM [6] слой, содержащий 64 нейрона, использующих в качестве функции активации гиперболический тангенс;

- третий слой (скрытый) – слой пакетной нормализации [7];
- четвёртый слой (выходной) – слой с единственным нейроном, с линейной активацией.

В качестве метода оптимизации применяется метод стохастического градиентного спуска Adam.

4. ФОРМИРОВАНИЕ ОБУЧАЮЩИХ НАБОРОВ

Для обучения нейронной сети, её необходимо обучить. В данной работе используются методы обучения нейронных сетей с учителем. Для этого используются обучающие примеры, каждый из которых содержит некоторый набор данных.

Содержимым обучающих примеров является набор из входных и выходных данных нейронной сети. Входные данные являются одномерным вектором, который содержит значения длины очереди во временном окне заданной ширины. Выходные данные содержат значения длины очереди, в момент времени, следующий за соответствующим временным окном. Значение ширины окна, промежуток времени до заданного значения, а также параметры нейронной сети – это значения, которые подбираются в процессе анализа прогнозов, полученных от обученной нейронной сети и её переобучении.

5. ОБРАБОТКА РЕЗУЛЬТАТОВ РАБОТЫ НЕЙРОННОЙ СЕТИ

В процессе прогнозирования, обученная нейронная сеть, выдает значения длины очереди в момент времени, следующий за поданным на её вход окном прогнозирования. Таким образом, следующим шагом, является сравнение, полученного от нейронной сети прогноза с некоторым заданным пороговым значением. Результат этого сравнения будет являться результатом прогноза разработанной модели.

6. РАЗРАБОТКА СИСТЕМЫ

На данном этапе, разрабатываемая система прогнозирования превышения длиной очереди сообщений этого порога через заданный промежуток времени, состоит из двух подсистем: подсистема обработки данных и подсистема построения нейронных сетей и обработки результатов их вычислений.

Подсистема обработки данных разрабатывается на языке программирования C#. Она обеспечивает предобработку «сырых» данных файлов *.pcap [6] и дальнейшую работу с этими данными.

В процессе работы этой подсистемы, по исходным данным формируются выборки, которые анализируются и с учетом указанных параметров, преобразовываются в данные для построения модели дискретной очереди.

Данная подсистема так же обеспечивает моделирование процесса ухода сообщений из дискретной очереди по заданным параметрам. Таким образом строится файл данных, содержащий информацию о состоянии очереди и загрузке канала в каждый момент времени.

После формирования файла с данными очередей, подсистема строит обучающие наборы, состоящие из обучающих примеров. Процесс построения обучающих примеров на данном этапе работы часто меняется. Здесь,

по заданным параметрам, формируются обучающие примеры, структуру которых можно задавать через параметры системы.

Подсистема работы с нейронными сетями разрабатывается на языке программирования Python, с использованием технологий tensorflow и keras. Так как структуру обучающих наборов приходится часто менять, был написан специальный адаптер, который строит нейронную сеть используя указанные в подсистеме параметры и данные о размерностях, которые можно извлечь из обучающих наборов.

Результатом работы подсистемы работы с нейронными сетями является обученная нейронная сеть, с помощью которой можно получать прогноз по некоторому окну данных и порогу. Остальные параметры, такие как время до прогнозируемого превышения порога, ширина окна и другие, закладывается в процессе обучения.

Подсистема работы с нейронными сетями поддерживает загрузку уже существующих нейронных сетей для продолжения их обучения, что полезно при получении свежих данных из уже изученных сетей.

7. ЗАКЛЮЧЕНИЕ

В данной работе описана общая структура системы, предсказывающей превышение заданного порога значением длины трафика через заданный интервал времени, а также используемые данные и методы, использующиеся в их обработке.

Ниже приведены результаты исследований на тестовой выборке, извлеченной из набора данных, описанного во 2 части, которая состоит из 1700 временных шагов, следующих друг за другом.

Для получения следующих результатов, нейронная сеть, описанная в части 3 обучалась в течении 600 эпох с размером пакетов 2. Величина окна прогноза была выбрана 4. Обучающий набор содержал 1500 временных отсчетов, а тестовый – 200.

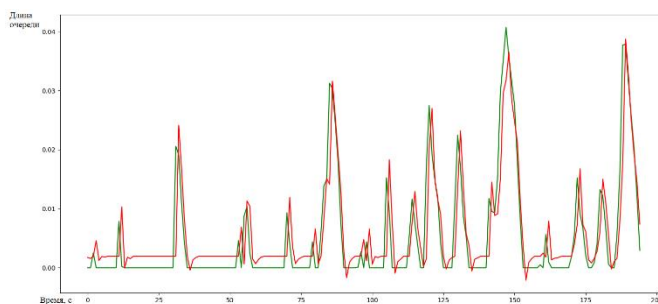


Рис. 1. Результат работы нейронной сети на тестовой выборке

Таким образом, было получено 196 тестовых прогнозов. Среди которых 11 ложно положительных и 4 ложно отрицательных прогнозов переполнения.

На рисунке (1) красной линией соединены значения длины очереди, предсказанные нейронной сетью, а зелёной – действительные значения, взятые из тестовой выборки. Значение по вертикальной оси нормализовано.

ЛИТЕРАТУРА

- [1] Paxson, V. Wide-Area Traffic: The Failure of Poisson Modeling / V. Paxson, S. Floyd // IEEE ACM Transactions on Networking. – 1995. – Vol. 3(3). – P. 226-244.
- [2] Leland, W.E. High time-resolution measurement and analysis of LAN traffic: Implications for LAN interconnection / W. Leland, D.V. Wilson // Proc. IEEE INFOCOM '91. – Bal Harbor, FL, 1991. – P. 1360-1366.
- [3] Mondragón, R.J. Chaotic maps for traffic modelling and queueing performance analysis / R.J. Mondragón, D.K. Arrowsmith, J.M. Pitts // Performance Evaluation. – Netherlands, 2001. – P. 223-240.
- [4] Willinger, W. A bibliographical guide to self-similar traffic and performance modeling for modern high-speed networks / W. Willinger, M.S. Taqqu, A. Erramilli // Clarendon press. – Oxford, 1996. – P. 339-366.
- [5] CAIDA Data – Overview of Datasets, Monitors, and Reports [Электронный ресурс]. – Режим доступа: <https://www.caida.org> (20.01.2022).
- [6] Zhao, Z. LSTM network: a deep learning approach for short-term traffic forecast / Z. Zhao, W. Chen // IET Intelligent Transport Systems. – 2017. – Vol. 11(2). – P. 68-75.
- [7] Слой нормализации [Электронный ресурс]. – Режим доступа: <https://ru.keras.com/norm-layers/> (01.02.2022).
- [8] Libpcap File Format [Электронный ресурс]. – Режим доступа: <https://wiki.wireshark.org/Development/LibpcapFileFormat> (20.01.2022).

Использование сверточных нейронных сетей для мониторинга безопасности на промышленном объекте

А.И. Султанова
Институт компьютерных
технологий и защиты информации
Казанский национальный
исследовательский технический
университет им. А.Н. Туполева -
КАИ
Казань, Россия
sultanovaali@stud.kai.ru

Л.М. Шарнин
Институт компьютерных
технологий и защиты информации
Казанский национальный
исследовательский технический
университет имени А. Н. Туполева -
КАИ
Казань, Россия
lmsharnin@kai.ru

В.В. Мокшин
Институт компьютерных
технологий и защиты информации
Казанский национальный
исследовательский технический
университет им. А.Н. Туполева -
КАИ
Казань, Россия
vladimir.mokshin@mail.ru

Аннотация—В статье рассматривается сверточная нейронная сеть для мониторинга безопасности на промышленном объекте. Описаны этапы обучения и тестирования точности работы нейронной сети. Также представлены график функции ошибки и диаграмма, содержащая среднюю точность по каждому классу и метрику MAP. Полученные результаты могут быть использованы в системах мониторинга безопасности на производственных объектах.

Ключевые слова— распознавание объектов, сверточная нейронная сеть, YOLOv3, безопасность.

1. ВВЕДЕНИЕ

Выявление нарушителей правил техники безопасности на производстве помогает повысить уровень дисциплины сотрудников, минимизировать количество несчастных случаев и улучшить показатели компании. Особое внимание уделяется рискам, связанным с нанесением ущерба здоровью или жизни человека в зонах производственных работ, поскольку уровень безопасности влияет на экономические показатели и статус предприятия. Своевременное обнаружение нарушений техники безопасности может помочь, в предотвращении несчастных случаев на производстве.

На сегодняшний день большую популярность и распространение получили сверточные нейронные сети (CNN). Они прекрасно себя проявили в области компьютерного зрения, как методы детектирования и распознавания объектов на изображениях или в видеопотоке [1-2]. Одним из современных детекторов выступает сверточная нейронная сеть YOLO (You Only Look Once – Ты Смотришь Только Один Раз). Благодаря своей архитектуре, ее достаточно один раз «взглянуть» на изображение, что значительно сокращает время необходимое для обнаружения объекта в кадре, без потери качества распознавания [3-7].

Используя технологию сверточных нейронных сетей проводим работы, связанные с контролем наличия спецодежды (сигнальных жилетов), которая выполняет распознавание и подсчет людей в жилетах / без жилетов. Определяем местоположение грузовых машин (зон повышенной опасности), фиксируем нарушения в виде отчетов и записывает видео с результатами распознавания. Для этого необходимо подготовить набор

данных (изображения с разметкой интересующих объектов), настроить параметры детектора, чтобы обучить нейронную сеть.

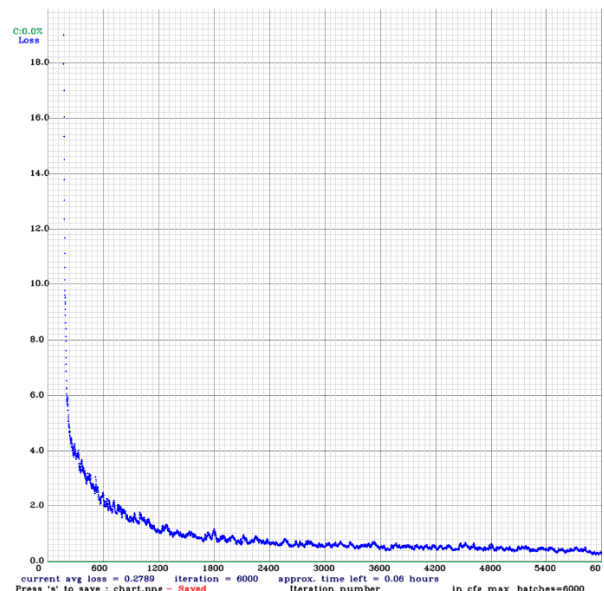


Рис. 1. Общие потери при обучении

На рис. 1 показан график обучения (функция потерь с каждой новой итерацией приближается к 0), среднее значение текущей потери 0.2789 и количество итераций 6000.

2. ЭТАП ТЕСТИРОВАНИЯ

Для оценки точности распознавания обученных моделей, полученных в результате обучения, были рассчитаны средние точности распознавания для каждого класса и метрика mAP. Расчёты проводились на основе тестовой выборки, содержащей 60 изображений и соответствующих аннотаций (рис. 2).

В качестве нейронной сети для дальнейшей работы была выбрана модель под номером 6 (рис. 2), detection_model-ex-0045—loss-0017.756.h5 с метрикой mAP = 0,5041. Данная модель имеет наибольшую точность распознавания объектов на изображении по сравнению с другими сгенерированными во время

обучения моделями. Результаты распознавания защитных касок и сигнальных жилетов представлены на рис. 2.



Рис. 2. Результат работы сверточной нейронной сети YOLOv3 обученной на производственном наборе данных

По рис. 2 (верхний) видно, что два рабочих находятся в зоне повышенной опасности (на грузовой машине – голубой прямоугольник), выполняют погрузочно-разгрузочную работу. В связи с этим оба рабочих обязаны применять сигнальный жилет, однако один рабочий нарушает это правило. Подсчет количества людей показывает, что в кадре находятся 1 рабочий в жилете и 1 рабочий без жилета.

По рис. 2 (нижний) видно, что два рабочих находятся в зоне повышенной опасности (на грузовой машине), выполняют погрузочно-разгрузочную работу. В связи с этим оба рабочих обязаны применять сигнальный жилет, однако один рабочий нарушает это правило. Подсчет количества людей показывает, что в кадре находятся 1 рабочий в жилете и 1 рабочий без жилета.

Программа автоматически фиксирует наличие / отсутствие нарушения в виде отчета в формате .docx с тем же именем, что и выбранное видео, и сохраняет его в отдельную папку «report». Если такой папки нет, программа создаст её автоматически.

Программа фиксирует ситуацию на видео в виде скриншота с надписью: «Нахождение человека без жилета в опасной зоне», если рабочий без жилета находится в зоне повышенной опасности (область грузовой машины). Если по окончании видео нарушений не было выявлено, в отчете прописывается «Нарушений не выявлено».

На некоторых изображениях видны ложные срабатывания и необнаруженные объекты, причинами могут служить недостаточно точная разметка эталонов или недостаточное количество обучающей выборки.

В табл. 1 представлено сравнение методов R-CNN, Fast R-CNN, Faster R-CNN, YOLOv2 и YOLOv3.

По табл. 1 можно сделать вывод, что модель YOLOv3 имеет самую высокую точность распознавания и

быструю скорость обработки входных данных по сравнению с другими алгоритмами.

Таблица 1. СРАВНЕНИЕ МЕТОДОВ РАСПОЗНАВАНИЯ ОБЪЕКТОВ

Алгоритм	Оценки	
	Средняя точность (mAP), %	Скорость обработки, мс
R-CNN	43.56	2991
Fast R-CNN	49.87	2271
Faster R-CNN	58.78	122
YOLOv2 (S x S)	81.64	59.1
YOLOv3 (S x S)	87.42	24.8

3. ЗАКЛЮЧЕНИЕ

В результате научной работы была реализована нейронная сеть YOLOv3 на языке программирования Python для мониторинга безопасности на производственном объекте. Для обучения был сформирован dataset, состоящий из 300 изображений и аннотаций к ним.

Обучение нейронной сети проводилось на 240 изображениях, а для оценки качества было подготовлено 60 изображений. Наилучший результат показала модель detection_model-ex-0045—loss-0017.756.h5. На базе полученных результатов возможно построение более высокоуровневых моделей для применения в исследовательских целях, а так же для создания систем анализа и мониторинга безопасности на производственных объектах.

ЛИТЕРАТУРА

- [1] Мокшин, В.В. Распознавание образов транспортных средств на основе эвристических данных и машинного обучения / В.В. Мокшин // Вестник Технологического университета. – 2016. – Т. 17, № 5. – С. 130-137.
- [2] Стадник, Н.А. Разработка информационной системы организации работ производственного цеха / Н.А. Стадник, А.В. Золотухин, В.В. Мокшин // Информационные технологии и нанотехнологии (ИТНТ-2021): сб. тр. по материалам VII Международной конференции и молодежной школы. – Самара, 2021. – С. 32982.
- [3] YOLO: Real-Time Object Detection [Электронный ресурс]. – Режим доступа : <https://pjreddie.com/darknet/yolo>.
- [4] Redmon, J. You Only Look Once: Unified, Real-Time Object Detection / J. Redmon // Computing Research Repository (CoRR), 2016 [Электронный ресурс]. – Режим доступа: <https://arxiv.org/abs/1506.02640>.
- [5] Redmon, J. Yolo9000: Better, faster, stronger. In Computer Vision and Pattern Recognition (CVPR) / J. Redmon, A. Farhadi // Computing Research Repository (CoRR), 2017 [Электронный ресурс]. – Режим доступа: <https://arxiv.org/abs/1612.08242>.
- [6] Redmon, J. Yolo3: An incremental improvement / J. Redmon, A. Farhadi // Computing Research Repository (CoRR), 2018 [Электронный ресурс]. – Режим доступа: <https://arxiv.org/abs/1804.02767>.
- [7] Kathuria, A. What's new in YOLO v3? / A. Kathuria // Towards data science, 2018 [Электронный ресурс]. – Режим доступа: <https://towardsdatascience.com/yolo-v3-object-detection-53fb7d3bfe6b>.
- [8] Ren, S. Faster r-cnn: Towards real-time object detection with region proposal networks / S. Ren // Computing Research Repository (CoRR), 2015 [Электронный ресурс]. – Режим доступа: <https://arxiv.org/abs/1506.01497>.

Сравнение алгоритмов обучения с подкреплением в задаче приобретения навыков передвижения в трёхмерном пространстве

Д.А. Козлов

Самарский национальный исследовательский университет им. академика С.П. Королева
Самара, Россия
djoade100@gmail.com

Аннотация—В работе выполняется сравнение современных методов обучения с подкреплением на примере решения задачи приобретения агентом навыков передвижения в трёхмерном пространстве. Сравнение производится в симуляторе Unity с использованием пакета ml-agents. В качестве сравниваемых алгоритмов выступают: SAC, PPO, MA-POCA. Они используются для обучения навыкам передвижения нескольких моделей агентов: 3DBall, Crawler, Walker и авторской SimplestBipedal. Результаты экспериментов говорят о преимуществах алгоритма Soft Actor Critic, что делает его более перспективным для использования в реальных средах.

Ключевые слова— обучение с подкреплением, SAC, PPO, MA-POCA, робототехника, Unity ML-Agents, симуляция, MDP, POMDP.

1. ВВЕДЕНИЕ

Проблема передвижения в трёхмерном пространстве автономных устройств – одна из центральных в робототехнике. Она может рассматриваться в различных постановках, начиная с задач самостоятельной ориентации и планирования маршрута (SPLAM-системы, [1]) и заканчивая задачами обучения собственно движению сложных устройств. Настоящая работа посвящена последней проблеме применительно к антропоморфным и животным-подобным устройствам. Выбор объекта исследования обусловлен тем, что для типовых (колесных, гусеничных) беспилотных наземных транспортных средств передвижение по пересеченной местности затруднительно или невозможно. Это дает нетипичным (антропоморфным и многоногим) устройствам преимущество при решении задач в сложных ситуациях, а задачу обучения их передвижению делает актуальной.

В настоящей работе рассматриваются подобные устройства, взаимодействующие с виртуальной симуляцией реального мира в среде игрового движка Unity. Симуляция позволяет ускорить процесс их обучения и тестирования.

Целью работы является сравнение современных эффективных алгоритмов обучения с подкреплением (англ.: Reinforcement Learning, RL) в задаче обучения автономных (антропоморфных и многоногих) мехатронных объектов походке и передвижению. В качестве сравниваемых алгоритмов выступают: Soft Actor-Critic (SAC), Proximal Policy Optimization (PPO) и MultiAgent POsthumous Credit Assignment (MA-POCA). Указанные алгоритмы используются для обучения навыкам передвижения нескольких различных по

структуре моделей агентов: 3DBall, Crawler, Walker и авторской SimplestBipedal.

Наиболее похожее сравнение приводится в работе [2]. Указанная работа отличается от настоящей тем, что в [2] отсутствует сравнение с алгоритмом MA-POCA; кроме того, в настоящей работе мы расширили состав моделей – ввели в рассмотрение новую созданную нами модель SimplestBipedal. Сравнение с MA-POCA было представлено в работе [3], но оно было выполнено авторами алгоритма MA-POCA. Также в работе [3] отсутствует сравнение с алгоритмом SAC, а также для сравнения используется иной набор задач/моделей. Кроме приведенных выше двух работ, авторами не было обнаружено независимых сравнений эффективности использования RL-методов для обучения передвижению разнотипных моделей ни для среды Unity, ни в похожих задачах; а найденные сравнения выполняются для более тривиальных сред/моделей.

Структура работы следующая. В разделе 2 представлены подробности исследуемых алгоритмов, в разделе 3 даны постановка и результаты экспериментов. В конце работы приводится заключение, в котором приведены выводы на основе полученных в экспериментах результатов.

2. МЕТОД И АЛГОРИТМЫ ОБУЧЕНИЯ

Обучение с подкреплением – способ машинного обучения, при котором *агент* (англ.: actor) обучается, взаимодействуя со средой. Для формализации RL-методов используется нотация Марковского процесса принятия решений по причине известного свойства марковости – переходы между состояниями зависят только от последнего состояния, а не от предыдущей истории. А именно: (S, A, R, P, ρ_0) , здесь S — множество допустимых состояний, A — множество допустимых действий, $R: S \times A \times S \rightarrow \mathbb{R}$ — функция вознаграждения, где $r_t = R(s_t, a_t, s_{t+1})$, $P: S \times A \rightarrow \mathcal{P}(S)$ - функция вероятности перехода, то есть $P(s'|s, a)$ - вероятность перехода в состояние s' , если вы находитесь в состоянии s и выполняете действие a , ρ_0 — начальное распределение состояний. Термином (*стохастическая политика*) обозначают стратегию поведения агента, то есть правило или способ выбора конкретного действия в конкретном состоянии: $\pi(a|s)$. Если политика явно не формируется, тогда алгоритм обозначают «off-policy».

A. Алгоритм Soft Actor-Critic

Soft Actor Critic – off-policy алгоритм глубокого RL. В нем агент стремится максимизировать ожидаемое вознаграждение, а также максимизировать энтропию. То

есть преуспеть в задаче, действуя как можно хаотичнее. Сочетая хаотичность со стабильной структурой подхода «astog-critic», метод достигает наилучшей производительности в ряде задач непрерывного управления. Кроме того, в отличие от других off-policy алгоритмов, метод очень стабилен, то есть достигает при разных случайных начальных значениях примерно одинаковой эффективности решения.

Б. Алгоритм Proximal Policy Optimization

Proximal Policy Optimization (PPO)[4] - семейство RL-методов градиента политики, которые чередуют наблюдения (для определения состояния) посредством взаимодействия с окружающей средой и оптимизацию «суррогатной» целевой функции с использованием стохастического градиентного подъема. В отличие от стандартных методов градиента политики, которые выполняют одно обновление градиента для каждого наблюдения, здесь обеспечивается несколько эпох мини-пакетных обновлений.

В. Алгоритм MA-POCA

MA-POCA [3] (MultiAgent POsthumous Credit Assignment), представляет собой многоагентный алгоритм. Он использует искусственную нейронную сеть (ИНС), выступающую в роли критика/оценщика решения, которая действует как «тренер» для целой группы агентов (которые также могут быть реализованы как ИНС).

3. ЭКСПЕРИМЕНТ

А. Среда симуляции

Unity Machine Learning Agents Toolkit [2] (ML-Agents) – проект с открытым исходным кодом, позволяющий использовать возможности игрового движка Unity совместно с реализациями высокопроизводительных алгоритмов машинного обучения, написанных на Python с использованием Pytorch и Tensorflow [5-6]. В качестве задач обучения и агентов выступали: 3DBall, Crawler, Walker и оригинальный SimplestBipedal. На рисунке 1 слева представлен агент 3DBall. Задача этого агента - балансировать шар на платформе. Наблюдение состояния включает координаты шара, скорости вдоль каждой оси и наклон платформы. Награда в этой задаче определяется как время, в течение которого шар находится на платформе. Далее на рисунке 1 представлены агенты Crawler, Walker и SimplestBipedal. Задача этих агентов – переместиться как можно ближе и быстрее к заданной цели в 3D пространстве. Для них награда даётся пропорционально скорости движения агента и направлению движения к цели. Следует отметить, что SimplestBipedal – это самостоятельно созданная задача. Она подразумевает высокую сложность задачи обучения из-за структурных особенностей агента - в верхнем суставе присутствует подвижность в двух плоскостях, а в нижнем – лишь в одной.

Б. Результат экспериментов

В результате проведенных экспериментов были получены усредненные значения суммарной награды, показанные в таблице 1. Из таблицы видно, что SAC превосходит остальные исследуемые алгоритмы.



Рис. 1. Агенты 3DBall, Crawler, Walker, SimplestBipedal

Следует также отметить, что в ходе нашего сравнения были получены результаты, отличающиеся от представленных в работе [2]. Это может быть связано с тем, что указанное исследование проводилось 2 года назад, за которые вносились изменения как в реализации алгоритмов, так и в движок Unity.

Таблица I. РЕЗУЛЬТАТЫ ЭКСПЕРИМЕНТОВ

Агент	3D Ball	Crawler		Walker		Simplest Bipedal	
Шаг (млн.)	0.5	2	5	5	10	2	10
SAC	100	2500	2500	3000	3000	3700	4700
PPO	100	1500	2500	250	500	0	100
MA-POCA	100	1500	2500	250	500	0	100

4. ЗАКЛЮЧЕНИЕ

В результате проведенного исследования было выявлено, что наиболее эффективным методом обучения с подкреплением для решения задачи приобретения навыков передвижения в трёхмерном пространстве является алгоритм Soft Actor Critic. Видно, что SAC в среде Crawler достигает того же уровня награды, что и другие алгоритмы, но за меньшее количество шагов, а в экспериментах Walker и SimplestBipedal SAC значительно превосходит PPO и MA-POCA по значению достигнутой средней суммарной награды. Эти результаты говорят о том, что SAC является наиболее перспективным вариантом для использования в различных робототехнических приложениях. В том числе алгоритм может использоваться для обучения нетипичных устройств в реальной среде.

БЛАГОДАРНОСТИ

Работа выполнена при поддержке Российского научного фонда (проект № 21-11-00321, <https://rscf.ru/en/project/21-11-00321/>).

ЛИТЕРАТУРА

- [1] Kozlov, D. Development of an Autonomous Robotic System Using the Graph-based SPLAM Algorithm / D. Kozlov, V. Myasnikov // International Conference on Information Technology and Nanotechnology (ITNT). – 2021. – P. 1-5. DOI: 10.1109/ITNT52450.2021.9649028.
- [2] Juliani, A. Unity: A General Platform for Intelligent Agents / A. Juliani, V.-P. Berges, E. Teng, A. Cohen, J. Harper, C. Elion, C. Goy, Y. Gao, H. Henry, M. Mattar, D. Lange // ArXiv: 1809.02627, 2020.
- [3] Cohen, A. On the Use and Misuse of Absorbing States in Multi-agent Reinforcement Learning / A. Cohen, E. Teng, V.-P. Berges, R.-P. Dong, H. Henry, M. Mattar, A. Zook, S. Ganguly // ArXiv: 2111.05992, 2021.
- [4] Schulman, J. Proximal Policy Optimization Algorithms / J. Schulman, F. Wolski, P. Dhariwal, A. Radford, O. Klimov // ArXiv: 1707.06347, 2017.
- [5] PyTorch [Electronic resource]. – Mode of access: <https://www.pytorch.org> (19.04.2022).
- [6] TensorFlow [Electronic resource]. – Mode of access: <https://www.tensorflow.org/?hl=ru> (19.04.2022).

Влияние состава наблюдений окружающей среды в задаче приобретения навыков передвижения в трёхмерном пространстве при использовании алгоритмов обучения с подкреплением

Д.А. Козлов

Самарский национальный исследовательский университет
им. академика С.П. Королева
Самара, Россия
djoade100@gmail.com

В.В. Мясников

Самарский национальный исследовательский университет
им. академика С.П. Королева
Самара, Россия
vmyas@geosamara.ru

Аннотация—В работе исследуется влияние состава наблюдений окружающей среды на процесс обучения «двуногого» мехатронного объекта навыкам передвижения в трёхмерном пространстве. Исследования проводятся в среде игрового движка Unity с использованием пакета ML-Agents. В качестве алгоритма обучения был выбран Soft Actor Critic, как один из наиболее эффективных современных алгоритмов обучения с подкреплением (RL), показавший наибольшую эффективность на наборе аналогичных задач. Показано, что состав наблюдений может радикально менять скорость обучения и даже замедлять процесс обучения при наличии «избыточных» данных.

Ключевые слова— SAC, Unity ML-Agents, обучение с подкреплением, симуляция, MDP, POMDP, робототехника.

1. ВВЕДЕНИЕ

Одной из задач робототехники является задача приобретения навыков передвижения (обучение передвижению) в трёхмерном пространстве [1]. Для её решения существует множество подходов, среди которых современным и универсальным является метод обучения с подкреплением (RL-reinforcement learning) [2]. Особый интерес при обучении передвижению представляют человекоподобные механизмы - «двуногие» автономные мехатронные устройства, поскольку они обладают большим потенциалом для движения по пересеченной местности [3]. С точки зрения RL-методов, обучение подобных механизмов интересно большим количеством степеней свободы (по числу подвижных частей) и, как следствие, большим количеством возможных локальных минимумов в решении: локальные минимумы соответствуют таким устойчивым положениям механизма, из которых тяжело выбраться [4].

В настоящей работе исследуется влияние объема и состава наблюдений (данных, информации) об окружающей среде (в терминах RL-алгоритмов – о состоянии) на эффективность решения искомой задачи с использованием RL-метода. В качестве целевого RL-алгоритма выбран Soft Actor Critic, как алгоритм, показавший наибольшую эффективность решения пула аналогичных задач [5]. Исследования проводятся в виртуальной среде игрового движка Unity с использованием пакета ML-Agents.

Структура работы следующая. В разделе 2 кратко описывается используемый алгоритм, в разделе 3 описывается постановка эксперимента и его результаты. В конце работы приведено заключение, обобщающее результаты эксперимента.

2. МЕТОД И АЛГОРИТМ ОБУЧЕНИЯ

Обучение с подкреплением – способ машинного обучения, при котором система обучается, взаимодействуя со средой. В отличие от типовых методов машинного обучения с учителем, требующих заранее определенных наборов данных или указания предопределенных ответов, в RL-методах необходимо указывать лишь награду, которую получает алгоритм обучения (в RL-методах именуется как *агент*) в зависимости от его действий в том или ином состоянии. Состав наблюдений, в свою очередь, описывает *состояние* системы «агент+среда» в каждый момент времени. Задача агента состоит в том, чтобы максимизировать суммарную награду, формируемую в результате всей последовательности его действий.

Задача обучения с подкреплением формализуется как Марковский процесс принятия решений. Марковский процесс принятия решений представляет собой кортеж, (S, A, R, P, ρ_0) , где S — множество всех допустимых состояний, A — множество всех допустимых действий, $R: S \times A \times S \rightarrow \mathbb{R}$ — функция вознаграждения, где $r_t = R(s_t, a_t, s_{t+1})$, $P: S \times A \rightarrow \mathcal{P}(S)$ - функция вероятности перехода, где $P(s'|s, a)$ - вероятность перехода в состояние s' , если вы начинаете в состоянии s и предпринимаете действия a . ρ_0 — начальное распределение состояний. Название «Марковский процесс принятия решений» относится к тому факту, что система подчиняется марковскому свойству: переходы зависят только от самого последнего состояния и действия, а не от предыдущей истории. Под термином (*стохастическая*) *политика* понимают стратегию поведения агента, то есть правило выбора действия в определенном состоянии: $\pi(a|s)$. На практике политику можно реализовать как таблично, так и в виде параметризуемого отображения.

Soft Actor Critic [5] (SAC) — это современный RL-алгоритм, предложенный Google и UC Berkley и позиционированный авторами для использования в

задачах робототехники. Алгоритм агрегирует подходы стохастической оптимизации политики с градиентными методами (англ.: Policy Gradient Algorithms) [6].

Центральной особенностью SAC является регуляризация энтропии. Политика настраивается так, чтобы максимизировать компромисс между ожидаемой доходностью и энтропией, характеризующей меру случайности в политике. Такой подход тесно связан с идеей поиска компромисса между разведкой и эксплуатацией [7], позволяет предотвратить преждевременную сходимость политики к плохому локальному оптимуму.

3. ЭКСПЕРИМЕНТЫ И РЕЗУЛЬТАТЫ

А. Постановка экспериментов

В качестве среды симуляции для экспериментов был использован игровой движок Unity совместно с пакетом ML-Agents [8]. Для проведения экспериментов была создана собственная модель подвижного устройства (агент), названная SimplestBipedal: агент в данной среде имеет две конечности, каждая из которых подвижна в двух суставах. При этом в верхнем суставе присутствует подвижность в двух плоскостях, а в нижнем – лишь в одной. Такая модель подразумевает высокую сложность задачи обучения. Задачей агента являлось достижение «цели» - предопределенного места в 3D среде.

Наблюдения, передаваемые агенту в каждом эксперименте, выбирались в соответствии с таблицей 1. Помимо указанной в ней информации агенту всегда передавалась информация о направлении к цели. *Награда* вычислялась как величина, пропорциональная скорости движения агента и направлению движения к цели.

Таблица 1. СОСТАВ ЭКСПЕРИМЕНТОВ

Передаваемая агенту информация	Номер эксперимента								
	0	1	2	3	4	5	6	7	8
Перемещение частей в глобальных координатах	+	+	+	+	+	+		+	+
Поворот частей в глобальных координатах	+	+	+	+	+		+	+	+
Перемещение частей в локальных координатах			+	+	+	+	+		
Поворот частей в локальных координатах	+		+	+	+		+	+	
Скорость частей в глобальных координатах	+			+	+	+	+		
Скорость частей в локальных координатах				+	+			+	
Положения и углы суставов					+				+
Сила прикладываемая в суставах					+				+

Б. Результат эксперимента

Результаты кратко представлены в таблице 2. В таблице представлены усредненные значения награды к шагу номер 10 млн. Максимальная награда достигнута для пятого эксперимента, она составила 6000.

Из представленных результатов видно, что информация о положении, угле суставов, а также о силе,

приложенной к ним, является (с точки зрения достигаемого результата и скорости обучения) избыточной и оказывающей негативное влияние на качество решения задачи, хотя, очевидно, полезна для решения задачи обучения.

Таблица 2. РЕЗУЛЬТАТЫ ЭКСПЕРИМЕНТОВ

	Номер эксперимента								
	0	1	2	3	4	5	6	7	8
Награда (в тысячах)	5	3	3	5,5	4	6	3	5,7	2

4. ЗАКЛЮЧЕНИЕ

По результатам проведенных исследований становится понятно, что полезные с содержательной точки зрения наблюдения могут оказывать негативное влияние на процесс обучения с использованием RL-методов. Таким образом, необходимым этапом решения задачи обучения с подкреплением в реальной среде (где «стоимость» процесса обучения и самого агента несоизмеримо выше, чем в среде имитационной) становится анализ доступного состава наблюдений и отбор тех, которые целесообразно использовать для RL-обучения.

Остался открытым вопрос, является ли «негативное» влияние наблюдений объективной или субъективной характеристикой, то есть оказывает ли влияние на этот выбор используемый RL-метод.

БЛАГОДАРНОСТИ

Работа выполнена при поддержке Российского научного фонда (проект № 21-11-00321, <https://rscf.ru/en/project/21-11-00321/>).

ЛИТЕРАТУРА

- [1] Peters, J. Towards Motor Skill Learning for Robotics / J. Peters, K. Mülling, J. Kober, D. Nguyen-Tuong, O. Kroemer. – 2009. – P. 469-482.
- [2] Haarnoja, T. Learning to Walk via Deep Reinforcement Learning / T. Haarnoja, S. Ha, A. Zhou, J. Tan, G. Tucker, S. Levine // ArXiv: 1812.11103, 2019.
- [3] Atlas™ // Boston Dynamics [Electronic resource]. – Mode of access: <https://www.bostondynamics.com/atlas> (17.04.2022).
- [4] Kozlov, D. Comparison of Reinforcement Learning Algorithms for Motion Control of an Autonomous Robot in Gazebo Simulator / D. Kozlov // International Conference on Information Technology and Nanotechnology (ITNT). – 2021. – P. 1-5. DOI: 10.1109/ITNT52450.2021.9649145.
- [5] Haarnoja, T. Soft Actor-Critic: Off-Policy Maximum Entropy Deep Reinforcement Learning with a Stochastic Actor / T. Haarnoja, A. Zhou, P. Abbeel, S. Levine // ArXiv: 1801.01290, 2018.
- [6] Silver, D. Deterministic Policy Gradient Algorithms / D. Silver, G. Lever, N. Heess, T. Degris, D. Wierstra, M. Riedmiller. – 2019. – P. 9.
- [7] Rocca, J. The exploration-exploitation trade-off: intuitions and strategies / J. Rocca // Medium, 2021 [Electronic resource]. – Mode of access: <https://towardsdatascience.com/the-exploration-exploitation-dilemma-f5622f8e1e82> (17.04.2022).
- [8] Juliani, A. Unity: A General Platform for Intelligent Agents / A. Juliani, V.-P. Berges, E. Teng, A. Cohen, J. Harper, C. Elion, C. Goy, Y. Gao, H. Henry, M. Mattar, D. Lange // ArXiv: 1809.02627, 2020.

Морфологический анализ текста с помощью нейронных сетей

А.Н. Жданова
Самарский национальный
исследовательский университет им.
академика С.П. Королева
Самара, Россия
zhdan.aleksandra@gmail.com

А.В. Куприянов
Самарский национальный
исследовательский университет им.
академика С.П. Королева
Институт систем обработки
изображений - филиал ФНИЦ
«Кристаллография и фотоника»
РАН
Самара, Россия
akupr@ssau.ru

Д.С. Шеренков
Самарский национальный
исследовательский университет им.
академика С.П. Королева
Самара, Россия
dsherenkov000@gmail.com

Аннотация—Статья посвящена применению технологии нейронных сетей для решения задачи морфологического анализа текста. Для обучения был использован банк размеченного текста, где каждому слову поставлены в соответствие его часть речи и форма слова. Было проведено сравнение точности разметки с доступными сервисами, использующими как нейросетевой подход, так и библиотечный метод морфологического анализа.

Ключевые слова— морфологический анализ, рекуррентная нейронная сеть, автоматическая обработка текста.

1. ВВЕДЕНИЕ

Наука о естественном человеческом языке, лингвистика, имеет множество отраслей и разделов, одним из которых является морфология. Морфологический анализ – одна из базовых задач автоматической обработки текста, относящаяся к типу классификации последовательностей. Такие задачи, как автореферирование текста, определение эмоциональной характеристики, машинный перевод, распознавание сущностей и многие другие, первоочередно включают в себя задачу сопоставления каждого слова, исследуемого текста, с его морфологическими признаками, как один из первых этапов [1]. Небольшие ошибки морфологического анализа могут привести к более серьезным последствиям в дальнейших этапах обработки текста, что обуславливает значимость и актуальность задачи морфологической разметки [2].

2. СРАВНЕНИЕ СО СЛОВАРНЫМИ МЕТОДАМИ

Словарный метод морфологического анализа текста заключается в определении морфологических признаков отдельно взятого слова с помощью нескольких словарей. К таким словарям относятся: словари основ существительных, окончаний существительных, основ прилагательных, основ глаголов, словарь служебных частей речи и т. д. [3]. Это является первым преимуществом в пользу нейросетевого подхода, для которого не нужно использовать настолько объемные и строго структурированные данные. В виду постоянного расширения языков, словари периодически должны дополняться новыми словами. Нейронная сеть же способна обрабатывать не только новые слова, но и несуществующие, опираясь на контекст и общую структуру слова.

Другой проблемой, с которой тяжело справиться словарным методом, являются различные виды омонимии. Обычная омонимия – это совпадение произношения и написания слов, совершенно разных по значению [4]. Частичная лексическая омонимия – совпадение различных форм одного слова [4]. Например, слово «физики» может обозначать как группу ученых, так и являться родительным падежом слова «физика». Также стоит отметить проблему омографии, когда слова имеют одинаковое написание, но разное ударение и смысл [4]. Например, словосочетание «большая часть» имеет разный смысл в зависимости от того на какой слог поставить ударение в слове «большая», на первый или на второй.

Одно из главных преимуществ нейронных сетей – это возможность нахождения скрытых зависимостей между входными и выходными данными, на что не способны методы с использованием словарей, лингвистические методы и т.д.

3. РЕАЛИЗАЦИЯ МОДЕЛИ

Для реализации модели была использована библиотека Keras. С помощью нее можно спроектировать нейронную сеть, задав параметры для ее оптимальной настройки под конкретную задачу. Параметр Dropout был задан 0.2, что означает, что 20% случайно выбранных нейронов игнорируются во время обучения на каждом цикле обновления весов. Это поможет избежать быстрой перенастройки сети. Количество эпох было установлено равным восьми, так как при слишком большом количестве итераций многослойный перцептрон начинает переобучаться. Был задан параметр оптимизации «Adam», реализующих метод стохастического градиентного спуска. Параметр shuffle задан как true, что обеспечивает перетасовку тренировочных данных перед началом каждой эпохи.

4. ОБУЧЕНИЕ НЕЙРОННОЙ СЕТИ

Нейронная сеть обучается на наборе данных, который состоит из 200 000 слов, каждому из которых поставлено в соответствие тэг, определяющий часть речи и форму слова. Пример тренировочного набора представлен на рисунке 1.

```
[('Повязанный', 'ADJ'), ('вокруг', 'ADP'), ('шеи', 'NOUN'),  
( 'шелковый', 'ADJ'), ('платок', 'NOUN'), ('придавал', 'VERB'),  
( 'его', 'DET'), ('довольно', 'ADV'), ('будничному', 'ADJ'),  
( 'костюму', 'NOUN'), ('некоторую', 'DET'), ('элегантность', 'NOUN'),  
( '.', 'PUNCT')]
```

Рис. 1. Часть тренировочного набора данных

Тестовый набор данных имеет ту же структуру для автоматического определения точности полученных результатов.

В результате обучения модели была достигнута точность разметки 93.5%. Результаты проведения экспериментов для русскоязычной модели представлены на рисунке 2.

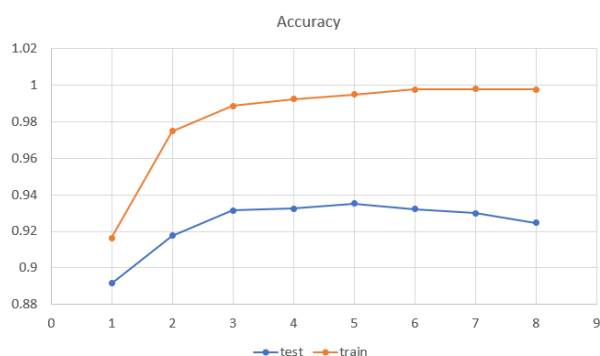


Рис. 2. Графики зависимостей точности разметки текста от количества эпох

Из графиков, представленных на рисунке видно, что максимальная точность разметки для тестового набора данных достигается после пяти эпох обучения и составляет 93.5%. Далее точность разметки для тренировочных наборов продолжает возрастать, а для тестовых начинает убывать, из чего можно сделать вывод о том, что начался процесс переобучения.

По некоторым данным, точность ручной частеречной разметки составляет примерно 98%, что достаточно близко к полученным 93.5%. Можно сказать, что применение модели многослойного перцептрона для решения задачи частеречной разметки оказалось успешным.

5. ЗАКЛЮЧЕНИЕ

В ходе исследования было проведено сравнение нейросетевого и словарного методов для решения задачи морфологической разметки текста. Реализована модель многослойного перцептрона для решения данной задачи.

Достигнутая точность морфологической разметки текста близка к сервисам, использующим словарный метод для решения данной задачи, а в некоторых случаях оказалась выше. Это доказывает перспективность использования нейронных сетей для решения как отдельной задачи морфологической разметки, так и в качестве начального этапа для решения более сложных и практических задач автоматической обработки текста.

ЛИТЕРАТУРА

- [1] Кочконбаева, Б.О. О морфологическом анализе в приложениях автоматической обработки текста / Б.О. Кочконбаева // Бюллетень науки и практики. – 2018 – Т. 4, № 12. – С. 608-612.
- [2] Цитильский, А.М. NLP – Обработка Естественных Языков / А.М. Цитильский, А.В. Иванников, И.С. Рогов // StudNet. – 2020. – № 6. – С. 467-475.
- [3] Бажанова, А.И. Разработка морфологического анализатора для построения понятийного аппарата электронной библиотеки кафедры АСУ / А.И. Бажанова // Информатика и компьютерные технологии. – 2011. – С. 326-330.
- [4] Бочаров, В.В. Прикладная и компьютерная лингвистика / В.В. Бочаров, И.С. Николаев, О.В. Митренина, Т.М. Ландо. – М.: URSS, 2017.

Исследование и применение свёрточной нейронной сети YOLO для автоматизированного тестирования десктопных и мобильных приложений

А.А. Рудь
Самарский национальный
исследовательский университет им.
академика С.П. Королева
Самара, Россия
sasha_96@mail.ru

С.А. Рудь
Самарский национальный
исследовательский университет им.
академика С.П. Королева
Самара, Россия
jin_96@mail.ru

М.М. Шушкина
Поволжский государственный
университет телекоммуникаций и
информатики
Самара, Россия
marija.shushkina@yandex.ru

Аннотация— Настоящая работа направлена на решение задачи обнаружения элементов управления десктопных и мобильных приложений. Представлены проблемы, возникающие при решении данной задачи. В работе показаны способы формирования наборов данных, а также применимость нейросетевого подхода в задачах автоматизированного тестирования программного обеспечения.

Ключевые слова— Свёрточные нейронные сети, генерация наборов данных, YOLO, обнаружение объектов, автоматизированное тестирование.

1. ВВЕДЕНИЕ

В настоящее время имеется множество решений, позволяющих выполнять автоматизированное тестирование программного обеспечения. Однако, существуют приложения, не поддающиеся стандартным методам автоматизации.

В настоящей работе предлагается нейросетевой подход к задаче обнаружения элементов управления десктопных и мобильных приложений. На сколько нам известно, не существует готовых решений на основе нейронных сетей для обнаружения компонентов приложений. Результат данной работы может быть интересен в сферах, где классические подходы уступают человеческому зрению и необходима автоматизация.

2. ОПИСАНИЕ ИСХОДНЫХ ДАННЫХ

Исходные данные глобально можно разделить на две группы: изображения реальных приложений и синтетические изображения. К первым относятся снимки окон существующих приложений, размеченных для обнаружения объектов в формате YOLO. В качестве целевого приложения может выступать любое приложение, имеющее потребность в автоматизированном тестировании.

Синтетические наборы представляют собой генерируемые случайным образом снимки десктопных приложений с соответствующей YOLO-разметкой. Производством синтетических данных занимаются три генератора, написанных на языках программирования: C# (WPF (Windows Presentation Foundation)), Python (PySide), Java (Swing). Невозможно покрыть всё разнообразие дизайна десктопных приложений, однако, существуют

элементы управления, имеющие схожие черты. Синтетический набор данных необходим для:

- 1) повышения обобщающей способности нейронной сети;
- 2) некоторой балансировки классов, путём генерации дефицитных компонентов (см. рисунок 1).

Таким образом, конечный набор данных формируется в результате смещения наборов данных реальных приложений и синтетических.

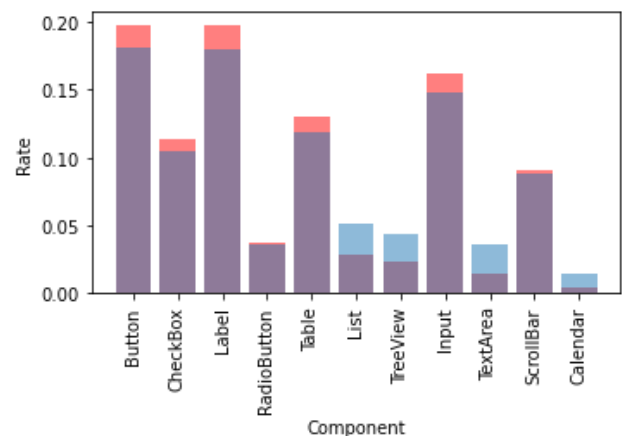


Рис. 1. Небольшая балансировка классов в процентном соотношении. Карсным обозначено уменьшение количества объектов класса, синим – увеличение по отношению к общему числу компонентов

Исходя из требований автоматизированного тестирования, было выделено одиннадцать классов компонентов для десктопных приложений: Button, CheckBox, Input, Label, List, RadioButton, Table, TextArea, TreeView, ScrollBar, Calendar. Как упоминалось ранее в десктопных приложениях присутствует сильный дисбаланс классов. Синтетический набор данных позволяет немного дополнить недостающие классы. Используемый подход не является идеальным решением ввиду значительных отличий в дизайне компонентов.

В качестве альтернативы генераторам возможно использование генеративно-состязательных сетей [1] для воспроизведения максимально схожих компонентов с реальными. Однако, в рамках настоящей работы данное исследование не проводилось.

Объектом исследования мобильных приложений выступило андроид-приложение Mafin [2]. Дизайн мобильных приложений отличается сложностью в задаче классификации компонентов пользовательского интерфейса. Исходя из этого были удалены классы объектов, которые не присутствуют в приложении, либо объединились в единый класс. Таким образом сформировались четыре класса компонентов для мобильного приложения: Button, CheckBox, Input, Label.

3. НЕЙРОСЕТЕВОЙ ПОДХОД

А. Локализация объектов

YOLO (You Only Look Once) [3] – архитектура свёрточной нейронной сети, позиционируемая для обнаружения объектов в реальном времени. Архитектурные особенности менялись от версии к версии. На данный момент последней модификацией сети является версия YOLOv5 [4].

Обучение сети проходило в два этапа:

1) предварительное обучение нейронной сети на синтетическом наборе данных;

2) обучение сети на снимках целевого приложения. Цель первого этапа – охватить максимально возможное количество разнообразий компонентов. На втором этапе происходит сбор набора данных целевого приложения с небольшой долей новых синтетических снимков. Таким образом на валидационной выборке для десктопного приложения mAP (mean Average Precision) [5] $\approx 0,95$; для андроид-приложения mAP $\approx 0,98$.

Б. Классификация атрибутов

Обнаруженные компоненты обладают соответствующими атрибутами. Знание атрибутов позволяет корректно взаимодействовать с элементом управления. Вырезанные из исходного изображения элементы управления подаются на вход простых свёрточных нейронных сетей [6]. Сети состоят из трёх последовательных блоков: слой свёртки и субдискретизации [7]. Последний слой полносвязный. В таблице 1 приведены компоненты и соответствующие им атрибуты, а также показатель F-меры.

Таблица 1. АТРИБУТЫ КОМПОНЕНТОВ

Component	Attributes				F1 score
	<i>enable</i>	<i>has_text</i>	<i>icon</i>	<i>checkable</i>	
Button	+	+	+		0,9859
CheckBox	+			+	0.9809
RadioBox	+			+	0.9918
Input	+	+		+	0.9905
Table	+				0.9874

Ниже приведено описание атрибутов.

1) *enable* – показывает доступность компонента для взаимодействия;

2) *has_text* – показывает наличие текста на компоненте;

3) *icon* – показывает наличие иконки на компоненте;

4) *checkable* – включена флаговая кнопка компонента.

Стоит отметить, что в некоторых задачах автоматизированного тестирования программного обеспечения требуется соотнести некоторые элементы управления другим. Например, элементы CheckBox/RadioButton/Input и Label.

4. ЗАКЛЮЧЕНИЕ

Результатом настоящей работы является разработанное решение, позволяющее обнаруживать элементы управления десктопных и мобильных приложений. Показана актуальность применимости компьютерного зрения для решения задач автоматизированного тестирования. Основными проблемами являются пропуск компонента, что ведёт к неудачному завершению теста, большое разнообразие дизайнов пользовательских интерфейсов, а также задача соотнесения элементов управления.

Кроме того, представлены способы формирования наборов данных. Выделение классов элементов управления для каждого приложения должно носить индивидуальный характер, хотя основные классы компонентов будут идентичны.

БЛАГОДАРНОСТИ

Работа выполнена при поддержке заведующего кафедрой технической кибернетики Самарского национального исследовательского университета имени академика С.П. Королева, доцента, д.т.н. Куприянова Александра Викторовича. Помощь в работе оказывал Шутько Вадим Валерьевич.

ЛИТЕРАТУРА

- [1] Goodfellow, I. Generative Adversarial Nets (NIPS version) / I. Goodfellow, J. Pouget-Abadie, M. Mirza // Advances in Neural Information Processing Systems. – 2014. – P. 27.
- [2] Mafin Product Web Site [Electronic resource]. — Access mode: <https://mafin.ru/> (25.12.2021).
- [3] Fang, W. Tinier-YOLO: A Real-Time Object Detection Method for Constrained Environments / W. Fang, L. Wang, P. Ren // IEEE Access. – 2020. – Vol. 8.
- [4] YOLOv5 Documentation [Electronic resource]. — Access mode: <https://docs.ultralytics.com/> (06.02.2022).
- [5] Revaud, J. Learning with average precision: Training image retrieval with a listwise loss / J. Revaud, J. Almazan, R. Rezende, C. De Souza // Proceedings of the IEEE International Conference on Computer Vision. – 2019. – Vol. 2019.
- [6] LeCun, Y. Backpropagation Applied to Handwritten Zip Code Recognition / Y. LeCun, B. Boser, J.S. Denker // Neural Computation. – 1989. – Vol. 1. – № 4.
- [7] LeCun, Y. Gradient-based learning applied to document recognition / Y. LeCun, L. Bottou, Y. Bengio, P. Haffner // Proceedings of the IEEE. – 1998. – Vol. 86. – № 11.

Автоматическое реферирование текстов

В. С. Головизнина

Вятский государственный университет

Киров, Россия

goloviznina@gmail.com

Аннотация—Автоматическое реферирование текстов – процесс создания краткого изложения текста, содержащего наиболее важную информацию. В настоящей работе исследуется задача создания рефератов русскоязычных текстов с помощью экстрактивных и абстрактивных методов. Для экспериментов был использован корпус новостных статей *Gazeta*. Для оценки качества реферирования использовались метрики ROUGE-N, ROUGE-L и BLEU. Наилучшие результаты показала модель ruT5-large.

Ключевые слова— автоматическое реферирование текстов, экстрактивные и абстрактивные методы реферирования, mBART, ruGPT-3, ruT5.

1. ВВЕДЕНИЕ

Автоматическое реферирование текстов – процесс создания краткого изложения текста, содержащего наиболее важную информацию [1]. Выделяют следующие подходы для реферирования текстов – экстрактивный, абстрактивный и гибридный. При экстрактивном подходе реферат формируется из наиболее важных предложений исходного текста; при абстрактивном содержание реферата генерируется и отличается от предложений исходного текста. Гибридный подход объединяет оба указанных подхода. Методы автоматического реферирования текстов используются в поисковых системах, для резюмирования блогов, научных статей и электронной почты, для генерации заголовков новостных статей, сжатого изложения судебных исков и медицинских текстов [2]. В области автоматического реферирования остаются нерешенные проблемы [2]. Так, в большинстве работ применяется экстрактивный подход, в то время как абстрактивные методы позволяют получить более краткое и близкое к человеческому изложение, отличное от предложений исходного текста. Как и в других областях обработки естественного языка большинство исследований проводится для английского языка [3]. В данной работе сравниваются несколько методов в рамках абстрактивного и экстрактивного подходов на русскоязычном корпусе новостных статей *Gazeta*.

2. ПРЕДЫДУЩИЕ РАБОТЫ

Языковые модели, основанные на архитектуре Transformer [4], стали ключевой технологией для решения задач обработки естественного языка, в том числе и для автоматического реферирования текстов [5]. Для реферирования русскоязычных текстов используют модели mBART [6], ruGPT3 [7] и ruT5 [8]. Так, в работе [9] многоязычная модель mBART обучалась задаче реферирования текстов на русскоязычном наборе данных *Gazeta*. В работе [10] обученная на корпусе *Gazeta* модель ruGPT-3Small используется для реферирования текстов на русском языке. В работе [11] модель mT5 [12] обучалась

реферированию текстов на 44 языках, в том числе и на русском, с использованием корпуса XLSUM.

В данной работе в отличие от [10] вместо модели ruGPT-3Small используется модель ruGPT-3Large. В отличие от [9] помимо mBART обучаются модели ruGPT-3Large и ruT5-large. В отличие от [11] используется не многоязычная модель mT5, а русскоязычная ruT5-large.

3. МЕТОДЫ ИССЛЕДОВАНИЯ

А. Экстрактивные методы

Для экстрактивного реферирования использовались методы TextRank из библиотеки *summa* и LexRank из библиотеки *lexrank* с параметрами по умолчанию.

TextRank [13] – метод, основанный на алгоритме PageRank [14] и применяемый для извлечения ключевых слов и экстрактивного реферирования. LexRank [15] – метод, представляющий текст в виде графа, для вычисления важности текстовых единиц.

Б. Абстрактивные методы

Для абстрактивного реферирования применялись модели mBART, ruGPT-3Large и ruT5-large.

Модель BART – Bidirectional and Auto-Regressive Transformer – основана на архитектуре Transformer и включает в себя двунаправленный кодировщик (как BERT) и авторегрессионный декодировщик (как GPT) [16]. Доступно две версии модели: BART_{BASE} и BART_{LARGE}. Многоязычная версия mBART [6] обучалась на Common Crawl corpus для 25 языков. В работе [9] используется mBART, настроенный для реферирования текстов на наборе данных *Gazeta*.

Модель GPT – Generative Pre-trained Transformer – это 12 слоев декодировщика Transformer [17]. Позднее появились вторая GPT-2 [18] и третья GPT-3 [19] версии модели. Модель ruGPT3 – русскоязычная модель от Сбера [7], основанная на GPT2, доступная в пяти версиях. В работе [20] использовалась модель GPT-2 для реферирования статей о COVID-19 на английском языке. В работе [10] применяли ruGPT-3Small для реферирования статей набора данных *Gazeta*.

Модель T5 – Text-to-Text Transfer Transformer – обучалась на 24 задачах для английского языка [21]. Модель mT5 обучалась для 101 языка, но на одной задаче – заполнения пропусков в тексте [12]. В работе [11] модель mT5 используется для реферирования на корпусе XLSUM. Модель ruT5 – русскоязычная модель T5 от Сбера, доступная в двух версиях: ruT5-base и ruT5-large [8].

4. РЕЗУЛЬТАТЫ ИССЛЕДОВАНИЯ

Для моделей mBART_{LARGE} и ruT5-large длина входа (текста) составляла 1024 токена, длина выходных данных (длина генерируемого реферата) ограничивалась максимальной длиной эталонных рефератов.

Для модели ruGPT-3Large длина выходных данных регулировалась так же, длина входных данных – 2048 токенов. При обучении на вход модели ruGPT-3Large подавались последовательности вида: «Text:text[SEP]Summary:summary», где text – это текст, summary – эталонный реферат для этого текста. При тестировании модель генерировала реферат для текста, подаваемого на вход в виде: «Text:text[SEP]Summary:».

Модель mBART_{LARGE} дообучена на наборе данных Gazeta в работе [9]. Наборы данных для автоматического реферирования – это наборы текстов и рефератов к ним. Набор данных Gazeta – это 63 435 статей из новостного источника Gazeta.ru. Характеристика набора данных приведена в таблице I (длина в токенах указана для токенизатора ruGPT-3Large). В качестве реферата используется описание статьи. Модели ruT5-large и ruGPT-3Large были дообучены на этом же наборе данных.

Для оценки результатов использовались автоматические метрики: ROUGE-N [22], ROUGE-L [22] и BLEU [23]. Результаты экспериментов приведены в таблице II.

Таблица I. ХАРАКТЕРИСТИКА НАБОРА ДАННЫХ GAZETA

Выборка	Размер	Данные	Размер данных в токенах		
			min	max	mean
Обучающая	52 400 (82,6%)	текст	48	2 244	955
		реферат	17	123	64
Валидационная	5 265 (8,3%)	текст	244	1 997	941
		реферат	18	124	69
Тестовая	5 770 (9,1%)	текст	447	2 041	916
		реферат	25	67	127

Таблица II. РЕЗУЛЬТАТЫ ЭКСПЕРИМЕНТОВ

Метод	Значения F ₁ -score			
	R-1	R-2	R-L	BLEU
ruT5 _{LARGE}	32,45	13,97	29,24	10,88
ruGPT3 _{LARGE}	23,45	6,45	20,73	4,93
mBART	31,55	13,54	28,22	11,19
TextRank	21,44	6,27	18,56	3,92
LexRank	23,93	8,00	20,96	5,64

5. ЗАКЛЮЧЕНИЕ

Экстрактивный метод LexRank показал более высокие результаты, чем TextRank. Также стоит отметить, что значения метрик метода LexRank выше, чем значения для модели ruGPT-3Large. В рефератах ruGPT-3Large встречаются фактические ошибки, чего не может быть в экстрактивном методе. Лучшие результаты показали модели ruT5-large и mBART. При этом mBART склонен повторять исходный текст. Рефераты ruT5-large краткие и связные, создают наиболее высокое общее впечатление.

БЛАГОДАРНОСТИ

Результаты статьи получены при поддержке грантовой программы Yandex.Cloud.

ЛИТЕРАТУРА

[1] Fabbri, A.R. SummEval: Re-evaluating Summarization Evaluation / A.R. Fabbri, W. Kryściński, B. McCann, C. Xiong, R. Socher, D. Radev // Transactions of the Association for Computational Linguistics. – 2021. – Vol. 9. – P. 391-409.
[2] El-Kassas, W.S. Automatic text summarization: A comprehensive survey / W.S. El-Kassas, C.R. Salama, A.A. Rafea // Expert Systems with Applications. – 2021. – Vol. 165.

[3] Scialom, T. MLSUM: The Multilingual Summarization Corpus / T. Scialom, P. Dray, S. Lamprier, B. Pivowarski, J. Staiano // Proceedings of the Conference on Empirical Methods in Natural Language Processing (EMNLP). – 2020. – P. 8051-8067.
[4] Vaswani, A. Attention is all you need / A. Vaswani, N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A.N. Gomez // Advances in Neural Information Processing Systems. – 2017. – Vol. 30. – P. 5998-6008.
[5] Liu, Y. Text Summarization with Pretrained Encoders / Y. Liu, M. Lapata // Proceedings of the Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing. – 2019. – P. 3730-3740.
[6] Liu, Y. Multilingual Denoising Pre-training for Neural Machine Translation / Y. Liu, J. Gu, N. Goyal, X. Li, S. Edunov, M. Ghazvininejad, M. Lewis, L. Zettlemoyer // Transactions of the Association for Computational Linguistics. – 2020. – Vol. 8. – P. 726-742.
[7] Sberbank-ai/ruGPTs [Electronic resource]. — Access mode: <https://github.com/sberbank-ai/ru-gpts> (01.02.2022).
[8] Sberbank-ai/model-zoo [Electronic resource]. — Access mode: <https://github.com/sberbank-ai/ru-gpts> (01.02.2022).
[9] Gusev, I. Dataset for Automatic Summarization of Russian / I. Gusev // Artificial Intelligence and Natural Language. – 2020. – P. 122-134.
[10] Nikolich, A. Fine-tuning GPT-3 for Russian Text Summarization / A. Nikolich, I. Oslakova, T. Kudinova, I. Kappusheva, A. Puchkova // Data Science and Intelligent Systems, Proceedings of 5th Computational Methods in Systems and Software. – 2021. – Vol. 2. – P. 748-757.
[11] Hasan, T. XL-Sum: Large-Scale Multilingual Abstractive Summarization for 44 Languages / T. Hasan, A. Bhattacharjee, M. Islam, K. Samin, Y. Li, Y. Kang, M.S. Rahman, R. Shahriyar // Findings of the Association for Computational Linguistics: ACL-IJCNLP. – 2021. – P. 4693-4703.
[12] Xue, L. mT5: A Massively Multilingual Pre-trained Text-to-Text Transformer / L. Xue, N. Constant, A. Roberts, M. Kale, R. Al-Rfou, A. Siddhant, A. Barua, C. Raffel // Proceedings of the Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies. – 2021. – P. 483-498.
[13] Mihalcea, R. TextRank: Bringing Order into Text / R. Mihalcea, P. Tarau // Proceedings of the Conference on Empirical Methods in Natural Language Processing. – 2004. – P. 404-411.
[14] Page, L. The PageRank Citation Ranking: Bringing Order to the Web / L. Page, S. Brin, R. Motwani, T. Winograd. – 1998.
[15] Erkan, G. Graph-based Lexical Centrality as Saliency in Text Summarization / G. Erkan, D.R. Radev // Journal of Artificial Intelligence. – 2004. – Vol. 22. – P. 457-479.
[16] Lewis, M. BART: Denoising Sequence-to-Sequence Pre-training for Natural Language Generation, Translation, and Comprehension / M. Lewis, Y. Liu, N. Goyal, M. Ghazvininejad, A. Mohamed, O. Levy, V. Stoyanov, L. Zettlemoyer // Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics. – 2020. – P. 7871-7880.
[17] Radford, A. Improving Language Understanding by Generative Pre-Training / A. Radford, K. Narasimhan, T. Salimian, I. Sutskever. – 2018.
[18] Radford, A. Language Models are Unsupervised Multitask Learners / A. Radford, J. Wu, R. Child, D. Luan, D. Amodei, I. Sutskever. – 2019.
[19] Brown, T. Language Models are Few-Shot Learners / T. Brown, A. Radford, I. Sutskever. – 2020.
[20] Kieuvoongam, V. Automatic Text Summarization of COVID-19 Medical Research Articles using BERT and GPT-2 / V. Kieuvoongam, B. Tan, Y. Niu. – 2020.
[21] Raffel, C. Exploring the Limits of Transfer Learning with a Unified Text-to-Text Transformer / C. Raffel, N. Shazeer, A. Roberts, K. Lee, S. Narang, M. Matena, Y. Zhou, W. Li, P.J. Liu // Journal of Machine Learning Research. – 2020. – Vol. 21. – P. 1-67.
[22] Lin, C. ROUGE: A Package for Automatic Evaluation of Summaries / C. Lin // Association for Computational Linguistics. – 2004. – P. 74-81.
[23] Papineni, K. BLEU: a Method for Automatic Evaluation of Machine Translation / K. Papineni, S. Roukos, T. Ward, W. Zhu // Proceedings of the 40th Annual Meeting of the Association for Computational Linguistics. – 2002. – P. 311-318.

Методы предварительной обработки скриншотов десктопных приложений для системы оптического распознавания символов

С.А. Рудь
Самарский национальный
исследовательский университет им.
академика С.П. Королева
Самара, Россия
jin_96@mail.ru

А.А. Рудь
Самарский национальный
исследовательский университет им.
академика С.П. Королева
Самара, Россия
sasha_96@mail.ru

М.М. Шушкина
Поволжский государственный
университет телекоммуникаций и
информатики
Самара, Россия
marija.shushkina@yandex.ru

Аннотация—В работе рассматриваются типовые методы предварительной обработки изображений перед подачей их в системы машинного зрения для распознавания текста. Решается задача предобработки данных для систем оптического распознавания символов на примере десктопных приложений.

Ключевые слова— элементы управления приложения, распознавание текста, оптическое распознавание символов, десктопное приложение, виджеты, YOLO, Tesseract, PaddleOCR, CRAFT.

1. ВВЕДЕНИЕ

Задача оптического распознавания символов решается во многих прикладных областях: анализ и распознавание отсканированных документов, распознавание медицинских карт, внедрение электронного документооборота, сбор информации.

Наличие шума, использование различных шрифтов или рукописный текст на изображениях усложняют задачу распознавания текста, что приводит к необходимости разработки неких последовательно применяющихся методов предварительной обработки изображений с целью улучшения их качества.

Вклад авторов в данной работе: собран и размечен набор данных для решения поставленной задачи, решена проблема качественной локализации текстовых блоков.

2. СТАНДАРТНЫЕ МЕТОДЫ ПРЕДОБРАБОТКИ ИЗОБРАЖЕНИЯ

В статьях [1-3] рассматриваются такие методы предварительной обработки изображений для улучшения качества локализации и распознавания текста как: пороговая обработка, операции математической морфологии, удаление шума, увеличение разрешения изображения с помощью свёрточных нейронных сетей.

Применение методов предварительной обработки изображений зависит от области задачи, качества изображения, ориентации текста. Изображения, полученные путем снятия экрана десктоп-приложений, отличаются низким разрешением (100-150 dpi), обширной цветовой палитрой (изменяется как задний фон, так и цвет текста), разнообразием интерфейсов, шрифтов и наличием символов, не входящих в заданный алфавит. Облегчает задачу постоянная горизонтальная ориентация текста.

Некоторые элементы управления содержат иконки, схожие с алфавитом распознаваемого языка, которые трудно удалить. Компоненты интерфейса могут находиться в разных состояниях, например, в фокусе, или недоступен для взаимодействия. В последнем случае оттенок текста кнопок сильно тускнеет, ухудшается контрастность. Существуют и иные случаи, когда часть текста выходит за рамки приведения в нужный вид одной лишь пороговой обработкой. В некоторых случаях частично решается сокращением количества цветочных каналов, приведением изображения в оттенки серого. Поэтому имеет смысл создание неких наборов методов предварительной обработки, применяемых к различным видам компонентов [3].

Использование методов предварительной обработки может улучшить качество как обнаружения текста, так и его распознавания (без учёта ошибок детекции текста).

3. СИСТЕМЫ OCR И ТИПОВОЙ АЛГОРИТМ РАБОТЫ

Tesseract OCR [4] – разработка изначально компании Hewlett-Packard, а в последующем Google, на сегодняшний момент одна из многообещающих реализация технологии оптического распознавания текста. Инструмент является бесплатным, открытым, его исходный код доступен на GitHub [5]. Актуальной, выпущенной на момент написания работы, версией является Tesseract 5.0.

Важными этапами работы системы Tesseract являются: предварительная обработка входящих изображений с помощью библиотеки Leptonika, детектирование и сегментация текста, классификация символов. Для распознавания текста рекомендуется использовать встроенную LSTM-сеть. Tesseract можно интегрировать с внешним механизмом детекции текста. В зависимости от качества изображений авторы рекомендуют использовать собственные методы предварительной обработки. Стоит отметить, что система Tesseract поддерживает множество языков, в том числе и русский.

PaddleOCR – набор инструментов OCR. Основным элементом является система PP-OCR [6], включающая в себя три части: обнаружение текста, коррекция обнаруженного фрейма, распознавание символов с помощью CRNN.

Детекция текста производится в основном с помощью моделей ResNet50 или MobileNetV3. Исходный код с

весами моделей для разных языков доступен на GitHub [7]. Имеется поддержка 80 языков в том числе и русского.

CRAFT [8] – метод обнаружения текста, позволяющий эффективно определять область текста. Состоит из двух компонентов: свёрточная нейронная сеть, оценивающая символьные регионы; связывающая нейронная сеть.

4. ЛОКАЛИЗАЦИЯ ТЕКСТА

Первоначальным этапом большинства технологий OCR является локализация текста. Неточная локализация соответственно приводит к неточному распознаванию и зашумлению результатов.

В контексте решения задачи распознавания текста на изображениях десктоп-данных был собран набор, содержащий 715 вручную размеченных скриншотов, существующих десктопных приложений. Для увеличения набора данных с помощью трёх различных библиотек создания UI-интерфейсов было сгенерировано ещё 548 синтетических изображений. В качестве библиотек UI-интерфейсов были использованы: WPF (C#), Pyside (Python), Java Swing. Использование различных UI-библиотек позволило с наименьшими затратами получить наиболее разнообразные примеры интерфейсов.

Набор данных представлял собой изображения в разрешении 640x640 и файлы разметки в COCO формате. Все элементы в наборе данных представлены одним классом – Text. Разметка ограничивающих рамок производилась на уровне слов. Среднее количество элементов на изображение 32.

Для непредвзятого оценивания качества встроенных методов локализаций Tesseract, PaddleOCR и CRAFT из набора данных была выделена тестовая выборка. Размер выборки составил 190 изображений.

Для улучшения качества локализации была обучена нейронная сеть архитектуры YOLO version 5 [9]. Модель была выбрана из-за высоких значений метрик в решении задачи поиска объектов на общедоступных наборах данных, высокой скорости работы. Обучение производилось на тренировочной части набора данных. Сравнение результатов производилось с помощью метрики IoU (Intersection over Union). Численные значения работы разных методов на тестовой части выборки представлены в таблице 1.

Таблица 1. РЕЗУЛЬТАТЫ ЛОКАЛИЗАЦИИ ТЕКСТА

Метод	IoU
Tesseract	0.31
Paddle (MobileNetv3)	0.50
CRAFT	0.86
YOLO	0.91

На рисунках 1-2 представлены графические примеры разницы детекции текста с помощью Tesseract (рис. 1) и обученной сети YOLO v5 (рис.2).

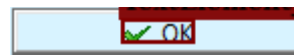


Рис. 1. Локализация текста встроенными методами Tesseract

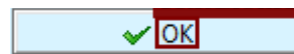


Рис. 2. Локализация текста YOLO

5. ЗАКЛЮЧЕНИЕ

В работе рассмотрены методы классической предобработки изображений для систем оптического распознавания символов. Обращено внимание на наиболее значимые по мнению авторов недостатки бесплатных OCR-систем, продемонстрированы на примере Tesseract, PaddleOCR, CRAFT.

Решена проблема детектирования текстовых блоков на примере скриншотов десктопных приложений с помощью обучения глубокой нейронной сети YOLO.

Разработанные методы имеют потенциал коммерческого применения в такой области как автоматическое тестирование десктопных и мобильных приложений.

БЛАГОДАРНОСТИ

Авторы выражают благодарность заведующему кафедрой технической кибернетики Самарского национального исследовательского университета имени академика С.П. Королева, доценту, д.т.н., Куприянову Александру Викторовичу, за побуждение и мотивацию для написания, Шутько Вадиму Валерьевичу за идею, без которой не было бы этой работы.

Выражаем благодарность ЛАНИТ Экспертиза за предоставленные вычислительные мощности для обучения моделей.

ЛИТЕРАТУРА

- [1] Kshetry, R.L. Image preprocessing and modified adaptive thresholding for improving OCR / R.L. Kshetry // ArXiv, 2021 [Electronic resource]. — Access mode: <https://arxiv.org/abs/2111.14075> (15.04.2022).
- [2] Peng, X. Building super-resolution image generator for OCR accuracy improvement / X. Peng, C. Wang // International Workshop on Document Analysis Systems. – 2020. – P. 145-160.
- [3] Sporici, D. Improving the accuracy of tesseract 4.0 OCR engine using convolution-based preprocessing / D. Sporici, E. Cuşnir, C.A. Boiangiu. – Symmetry. – 2020. – Т. 12, № 5. – P. 715.
- [4] Tesseract User Manual [Electronic resource]. — Access mode: <https://tesseract-ocr.github.io/tessdoc/> (15.04.2022).
- [5] Tesseract Open Source OCR Engine (main repository) [Electronic resource]. — Access mode: <https://github.com/tesseract-ocr/tesseract> (07.02.2022).
- [6] PP-OCR: A Practical Ultra Lightweight OCR System / Y. Du, C. Li, R. Guo // ArXiv, 2020 [Electronic resource]. — Access mode: <https://arxiv.org/abs/2009.09941> (15.04.2022).
- [7] PaddleOCR [Electronic resource]. — Access mode: <https://github.com/PaddlePaddle/PaddleOCR> (15.04.2022).
- [8] Character region awareness for text detection / Y. Baek, B. Lee, D. Han // Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. – 2019. – P. 9365-9374.
- [9] YOLOv5 Documentation [Electronic resource]. — Access mode: <https://docs.ultralytics.com/> (15.04.2022).

Система управления мышечным тонусом на основе нейронной сети с LIF моделью

А.Д. Иванова
Казанский (Приволжский)
федеральный университет
Казань, Россия
ivanovaanastasya0449@gmail.com

Н.В. Харин
Казанский (Приволжский)
федеральный университет
Казань, Россия
NVKharin@kpfu.ru

Т.В. Балтина
Казанский (Приволжский)
федеральный университет
Казань, Россия
tvbaltina@kpfu.ru

О.А. Саченков
Казанский (Приволжский)
федеральный университет
Казань, Россия
OASachenkov@kpfu.ru

Аннотация—В работе описано решение задачи управления, используя метод машинного обучения. В качестве объекта управления была выбрана модель мышечного волокна под действием внешней силы. Задача нейронной сети – найти вид функции управления, чтобы обеспечить сохранение мышц заданного удлинения. Построена математическая модель системы и реализован метод численного решения.

Ключевые слова— спайковая нейронная сеть, управление, математическое моделирование, искусственная нейронная сеть.

1. ВВЕДЕНИЕ

Сегодня все популярнее становится применение машинного обучения, в том числе нейронных сетей, в различных областях. Малоизученным направлением являются импульсные (спайковые) нейронные сети. Цель данной работы: с помощью моделирования механизма сокращения мышечного волокна добиться удержания мышц заданной длины, то есть тонуса мышцы, при условии воздействия на нее внешней растягивающей силы. Сокращение мышцы будет осуществляться при помощи искусственной нейронной сети.

2. МЕТОДЫ

А. LIF модель нейрона

При таком подходе каждый нейрон в каждый момент времени имеет некоторое значение – потенциал, которое сравнивается с пороговым значением нейрона. Если потенциал превышает пороговое значение, то нейрон посылает импульс (спайк) на следующий слой, а его потенциал падает до некоторого уровня. В противном случае происходит накопление потенциала. При выведении из состояния равновесия потенциал нейрона стремится к начальному значению. Динамика функции потенциала V описывается уравнением (1).

$$\tau_{mem} \cdot V' = E_{leak} - V + g_e(E_e - V) + g_i(E_i - V), \quad (1)$$

$$\tau_e \cdot g_e' = -g_e + w_e \cdot h(\Delta L), \quad (2)$$

$$h(\Delta L) = \theta(\Delta L - \Delta L_{THR}), \quad (3)$$

$$\tau_i \cdot g_i' = -g_i + w_i \cdot f(V), \quad (4)$$

$$f(V) = \theta(V), \quad (5)$$

Существует два типа входных сигналов для нейрона: g_e – возбуждающие (положительные) и g_i – тормозящие (отрицательные), динамика которых описывается

уравнениями (2), (4). Именно они в процессе моделирования изменяют значение потенциала. В работе для описания воздействия на входные сигналы использовалась функция Хэвисайда θ (3), (5).

Б. Модель мышцы Хилла

В качестве математической модели мышцы была выбрана трехэлементная модель Хилла (Рис 1.).

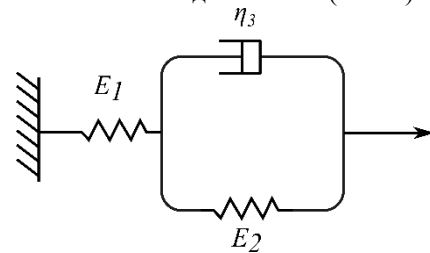


Рис 1. Модель мышцы Хилла, E_1 – коэффициент жесткости последовательного элемента, E_2 – коэффициент жесткости параллельного элемента, η_3 – коэффициент демпфирования

Сокращение мышцы запускает демпфирующий элемент, поэтому управление будет подаваться на него. Изменение длины мышцы инициализирует воздействие на входные сигналы нейрона:

$$\Delta L = \Delta L_1 + \Delta L_2, \Delta L_2 = \Delta L_3, \quad (6)$$

В. Архитектура сети

Архитектура нейронной сети была построена на основе структурных схем биологических систем. В этом случае мотонейрон иницирует активацию сократительного элемента мышцы и при этом посылает тормозящий сигнал (отрицательная обратная связь) на самого себя. Сокращение мышцы регулируется за счет возбуждающего сигнала (положительная обратная связь), посылаемого сенсорным нейроном, который по сути реагирует на удлинение мышцы. Топология описанной сети представлена на Рис.2.

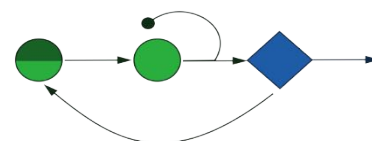


Рис 2. Топология предложенной сети: полужелтый круг – сенсор, зеленый – двигательный нейрон (мотонейрон), синий ромб – мышца; возбуждающие обратные связи обозначены стрелками, тормозящие – кружками

При моделировании была рассмотрена внешняя сила, растягивающая мышцу. При достижении некоторой пороговой величины растяжения сенсорный нейрон генерирует возбуждающие сигналы, которые приходят на мотонейрон, который описывается LIF моделью. Мотонейрон генерирует сигнал активации, вынуждая сокращаться мышечное волокно, и тормозит самого себя (Рис. 3).

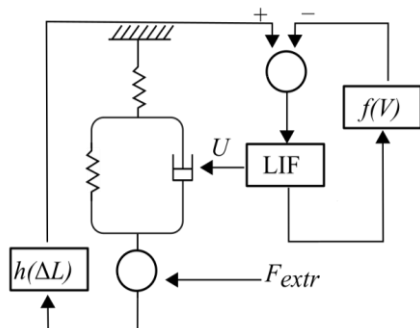


Рис 3. Схема системы управления: F_{extr} – внешняя растягивающая сила, LIF – мотонейрон, U – сила воздействия на сократительный элемент, $h(\Delta L)$ – возбуждающий сигнал мышечной активности, $f(V)$ – тормозящий сигнал нейрона, ΔL – удлинение мышцы

3. РЕЗУЛЬТАТЫ

Для вычисления использовались следующие значения констант: $F_{extr} = 0.2$ Н, $U = \{0; 0.08\}$ Н, заданное растяжение $\Delta L_{prog} = -0.05$ м. В модели Хилла были использованы значения: коэффициенты жесткости упругих элементов $E_1 = 10$ Н/м, $E_2 = 5$ Н/м, коэффициент демпфирования $\eta_3 = 3$ Н с/м.

В работе численно исследовалось влияние синаптического веса положительной обратной связи w_e на поведение динамической системы. В результате расчетов было найдено критическое значение $w_e^* = 6.3$. В случае, когда $w_e = w_e^*$, наблюдается мышечный тонус, и через некоторое время мышца сохраняет заданную длину (Рис. 4). В случае, когда $w_e > w_e^*$, мышечный тонус имеет колебательный характер, то есть мышца не сохраняет постоянную длину. Когда $w_e < w_e^*$, тонус наблюдается, но заданная длина ΔL_{prog} не достигается. Найденное критическое значение суть бифуркационный параметр динамической системы.

4. ЗАКЛЮЧЕНИЕ

В статье исследовалась модель управления мышечным тонусом. Рассмотрена задача мышечного тонуса под действием внешней силы. Архитектура системы управления базировалась на биологических принципах взаимодействия центрального генератора

упорядоченной нейроактивности. LIF модель была использован для моделирования поведения мотонейрона. Возбуждающий сигнал генерировался сенсорным нейроном на основе мышечной активности, тормозящий – мотонейроном. В работе исследовалось влияние величины синаптического веса положительной обратной связи и было найдено его критическое значение.

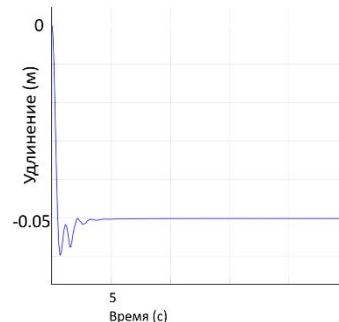


Рис 4. Зависимость удлинения от времени, при воздействии системы управления. Случай мышечного тонуса

БЛАГОДАРНОСТИ

Работа выполнена при финансовой поддержке РФФИ № 20-01-00535.

ЛИТЕРАТУРА

- [1] Дмитричев, А.С. Нелинейные динамические модели нейронов: обзор / А.С. Дмитричев, Д.В. Касаткин, В.В. Клиньшов, С.Ю. Кириллов // Известия высших учебных заведений. Прикладная нелинейная динамика. – 2018. – Т. 26. – С. 5-58.
- [2] Харин, Н.В. Применение нейронной сети для стабилизации маятника / Н.В. Харин, А.Д. Иванова, О.А. Саченков // Труды Математического центра имени Н.И. Лобачевского. – 2020. – Т. 59. – С. 131-134.
- [3] Frigon, A. The neural control of interlimb coordination during mammalian locomotion // Neurophysiol. – 2017. – Vol. 117. – P. 2224-2241.
- [4] Kasabov, N. NeuCube: A spiking neural network architecture for mapping, learning and understanding of spatio-temporal brain data / N.Kasabov // Neural Network. – 2014. – Vol. 52. – P. 467-510.
- [5] Romero, F. A comparison among different Hill-type contraction dynamics formulations for muscle force estimation / F. Romero, F.J. Alonso // Mechanical Sciences. – 2016. – Vol. 7. – P. 19-29.
- [6] McCrea, D.A. Organization of mammalian locomotor rhythm and pattern generation / D.A. McCrea, I.A. Rybak // Brain Research Reviews. – 2008. – Vol. 57. – P. 134-146.
- [7] Schmitt, S. Theoretical Hill-Type Muscle and Stability: Numerical Model and Application / S. Schmitt, M. Günther, T. Rupp, A. Bayer, D. Häufle // Computational and Mathematical Methods in Medicine. – 2013. – P. 570878.
- [8] von Morgenland, J. Hill's Model for Muscle Physiology and Biomechanics / J. von Morgenland, S. Venugopal // Encyclopedia of Computational Neuroscience. – 2020. DOI: 10.1007/978-1-4614-7320-6_100695.

Семантическая сегментация рентгенограмм легких на основе нейронной сети класса U-net

Н.С. Демин

*Институт систем обработки изображений - филиал
ФНИЦ «Кристаллография и фотоника» РАН
Самарский национальный исследовательский университет
им. академика С.П. Королева
Самара, Россия
volfgunus@gmail.com*

Н.Ю. Ильясова

*Институт систем обработки изображений - филиал
ФНИЦ «Кристаллография и фотоника» РАН
Самарский национальный исследовательский университет
им. академика С.П. Королева
Самара, Россия
ilyasova.nata@gmail.com*

Аннотация—В данной работе рассматривается семантическая сегментация рентгенограмм легких с использованием Transfer learning. Используется сеть типа U-net, у которой в качестве энкодера были взяты веса сети EfficientNet, предобученные на наборе ImageNet. В результате дообучения сети в течении 50 эпох была достигнута точность сегментации выше 90% для классов легкие и ребра, и 80% для класса сердце.

Ключевые слова—рентгенограммы легких, сегментация, трансферное обучение.

1. ВВЕДЕНИЕ

В настоящее время несмотря на то, что в основном применяются такие методы лучевой диагностики, как МРТ, имеющие в свою очередь высокую диагностическую информативность, обследования этими методами имеют определенные недостатки: дорогое оборудование и специализированное программное обеспечение. Однако, методы лучевой диагностики закрепились как дешевая и надежная процедура в начальных этапах исследований патологий. Согласно медицинским исследованиям, с помощью рентгенографии грудной клетки, можно выявлять патологические пертурбации анатомических структур, расположенных в грудной полости, а также исследовать изменения костей грудной клетки [1]. Применение методов машинного обучения и нейронных сетей в сфере анализа биомедицинских изображений облегчает раннюю диагностику и лечение заболеваний [2]. Согласно результатам медицинских исследований, на основе отслеживания изменения размеров легких, сердце и параметров костной части грудной клетки как ребра и ключицы, многие заболевания, в том числе хронические, могут быть диагностированы уже на их ранних этапах [3]. Выявлено, что путем наблюдения масса легких неравномерно распределена в органах, а именно с большей плотностью в нижних лепестках. На основе отслеживания изменения размеров легких и параметров ребер многие заболевания могут быть диагностированы уже на их ранних стадиях [4]. Сегментация рентгеновских изображений с помощью нейронных сетей играет большую роль в анализе заболеваний легких [5]. Прорыв в глубоком обучении с использованием сверточных нейронных сетей можно считать разработку AlexNet [6]. Эта сеть, победившая в конкурсе ImageNet, состояла из сверточных, MaxPool и полностью связанных слоев, а также Dropout, который использовался как регуляризатор, помогающий бороться с переобучением. Вслед за этим в статье [7] авторы представили U-Net, архитектуру нейронной сети, основанную на

сужающемся пути для захвата контекста изображения, за которым следует симметричный расширяющийся путь для обеспечения точной локализации. Сети типа U-Net нашли широкое применение в сегментации биомедицинских изображений.

В рамках данной работы предлагается применение Transfer learning для семантической сегментации снимков рентгенограмм.

2. СЕМАНТИЧЕСКАЯ СЕГМЕНТАЦИЯ РЕНТГЕНОГРАММ

Задача семантической сегментации заключается в разбиении исходного растрового изображения или группы изображений на представляющие интереса области или сегменты и последующая одновременная классификация этих сегментов по выбранному классом. В качестве исходных данных был использован набор из 214 размеченных снимков рентгенограмм в формате bmp с размерами 512×512 пикселей. Изображения были размечены на четыре класса: сердце, легкие, ребра и ключицы. Данные были разбиты на обучающую и валидационную выборку в соотношении 200:14. Для увеличения количества данных было проведено их аугментирование. Использовались такие методы, как: HorizontalFlip, OneOf, RandomContrast, RandomGamma, Resize, RandomBrightness, ElasticTransform, GridDistortion, Compose, OpticalDistortion, PadIfNeeded, RandomSizedCrop, Normalize. Что позволило увеличить размеры выборки в 10 раз.

Для получения высоких результатов основных метрик семантической сегментации модели применен подход, который называется трансферным обучением, в обиходе более известен Transfer learning. Этот подход был предложен еще в 1992 [8]. Сама методика трансфер лернинга заключается в передаче знаний нейросети, которая была обучена на данных одной задачи (обычно используется нейронные сети, которые обучались на большом количестве данных) и применение этих знаний для решения другой задачи. В задачах компьютерного зрения данный подход, состоящий из двух этапов, нашел свое широкое применение в различных задачах. Свою оценку исследования трансфер лернинга дала команда из Google AI, где в своей работе [9] оценили вклад этой техники в задачах компьютерного зрения, в частности для задач, связанных с медицинскими изображениями. Результаты исследования показали, что использование данного подхода помогает улучшить, хоть и не значительно, результаты предсказаний нейросети для медицинских задач. Поэтому чаще используется двухэтапный подход трансфер лернинга состоящий из предобучение, когда сеть обучается на большом датасете

с последующей множественной классификацией, и Fine-tuning, когда используется веса из этапа предобучения для целевой нейросети, которая впоследствии дообучается на данных поставленной задачи.

В рамках данной работы в качестве предобученной сверточной нейронной сети был использован класс моделей называемые как EfficientNets, которые в свою очередь, как предполагают ее авторы, порождаются после изменения разрешение изображений в сетке, масштабирование моделей (Скейлинг) и балансирование количество каналов сети [10]. Под операцией скейлинга понимают процесс, когда фиксируются производимые внутри архитектуры сверточной нейронной сети операции и меняются лишь глубина (количество повторений одних и тех же модулей), ширина (количество каналов в свёртках) и разрешение. В работе [11] была предложена улучшенная методология, названная как Noisy-student, для получения более высоких результатов метрик на наборе изображений из датасета ImageNet. Веса данной модели использовались для дообучения целевой задачи семантической сегментации рентгенограмм. Для неё, которая по сути является разновидностью задачи классификации, для выходного слоя была выбрана функция активации Softmax. В качестве метрик оценки работы использовались Dice (Коэффициент Соренсена – Дайса) и IoU (Коэффициент Жаккара).

3. РЕЗУЛЬТАТЫ СЕГМЕНТАЦИИ

С целью сопоставления предсказаний и реальности в области машинного обучения используется матрица ошибок или confusion matrix – это таблица с различными комбинациями прогнозируемых и фактических значений. Прогнозируемые значения описываются как положительные и отрицательные, а фактические – как истинные и ложные. Обычно матрица ошибок используется для оценки точности моделей в задачах классификации. Но задачу семантической сегментации можно рассматривать как частный случай этой проблемы, поэтому confusion matrix актуальна и для измерения точности предсказаний пикселей. Результаты представлены в таблице 1.

Таблица 1. МАТРИЦА ОШИБКИ

		Фактический класс				
		Фон	Легкие	Ребра	Ключицы	Сердце
Предсказанный класс	Фон	0.7560	0.0232	0.0516	0.3680	0.0339
	Легкие	0.0352	0.9290	0.0936	0.0892	0.0064
	Ребра	0.1763	0.0433	0.8410	0.2044	0.0113
	Ключицы	0.0104	0.0015	0.0045	0.3382	0
	Сердце	0.0219	0.0028	0.0091	0	0.9521

В таблице 1 показаны основные классы сегментации, так же указан класс фон – все что не попадает ни под один из классов. Как видно из таблицы проблемным классом являются ключицы, это связано с тем, что это класс, который представлен на изображениях в меньшем размере, чем остальные и часто его пиксели относятся к фону или к ребрам. У остальных классов полученная точность выше 0,8.

Таблица 2. ЗНАЧЕНИЯ МЕТРИК

Метрики	Датасет	Ребра	Ключицы	Легкие	Сердце
IoU	Training	0.8302	0.7169	0.9244	0.7733
	Validation	0.8472	0.7248	0.9401	0.7950
Dice loss	Training	0.1247	0.2072	0.0626	0.1902
	Validation	0.1183	0.2140	0.0558	0.1807

В таблице 2 представлены значения основных метрик.

4. ЗАКЛЮЧЕНИЕ

В данной работе рассматривалось применение transfer learning для задачи семантической сегментации рентгенограмм легких. Для нейронной сети класса U-net, у которой в качестве энкодера использовалась сеть EfficientNets, предобученная на наборе ImageNet, было проведено дообучение для сегментации рентгеновских снимков на 4 класса: сердце, легкие, ребра и ключицы. В результате была получена нейронная сеть, позволяющая выделять классы легкие и ребра с точностью выше 90%.

БЛАГОДАРНОСТИ

Работа выполнена при финансовой поддержке РФФИ (грант № 19-29-01135), Министерства науки и высшего образования РФ, в рамках выполнения работ по государственному заданию ФНИЦ «Кристаллография и фотоника» РАН.

ЛИТЕРАТУРА

- [1] Линденбратен, Л.Д. Медицинская рентгенология / Л.Д. Линденбратен, Ф.М. Лясс. – М.: Рипол Классик, 1986. – 368 с.
- [2] Дюдин, М.В. Нейросетевые модели принятия решений для диагностики заболеваний легких на основе анализа флюорограмм грудной клетки / М.В. Дюдин, Р.А. Томакова, М.В. Томаков // Биомедицинская радиоэлектроника. – 2014. – № 9. – С. 12-16.
- [3] Mathieson, J.R. Chronic diffuse infiltrative lung disease: comparison of diagnostic accuracy of CT and chest radiography / J.R. Mathieson, J.R. Mayo, C.A. Staples, N.L. Müller // Radiology. – 1989. – Vol. 171(1). – P. 111-116.
- [4] Verschakelen, J.A. The role of high-resolution computed tomography in the work-up of interstitial lung disease // Curr Opin Pulm Med. – 2010. – Vol.16(5). – P. 503-510.
- [5] Zhou, B. Lung mass density analysis using deep neural network and lung ultrasound surface wave elastography / B. Zhou, X. Zhang // Ultrasonics. – 2018. – Vol. 89. – P. 173-177.
- [6] Krizhevsky, A. ImageNet classification with deep convolutional neural networks / A. Krizhevsky, I. Sutskever, G.E. Hinton // Commun. ACM. – 2017. – Vol. 60(6). – P. 84-90.
- [7] Ronneberger, O. U-Net: Convolutional networks for biomedical image segmentation / O. Ronneberger, P. Fischer, T. Brox // International Conference on Medical image computing and computer-assisted intervention. – 2015. – P. 234-241.
- [8] Pratt, L.Y. Discriminability-based transfer between neural networks / L.Y. Pratt, S.J. Hanson, J.D. Cowan, C.L. Gile // Advances in Neural Information Processing Systems. – 1992. – Vol. 5. – P. 204-211.
- [9] Raghu, M. Transfusion: Understanding transfer learning for medical imaging / M. Raghu, C. Zhang, J. Kleinberg, S. Bengio // Advances in Neural Information Processing Systems. – 2019. – P. 3342-3352.
- [10] Mingxing, T. EfficientNet: Rethinking Model Scaling for Convolutional Neural Networks / T. Mingxing, Q. Le // Proceedings of the 36th International Conference on Machine Learning. – 2019. – P. 6105-6114.
- [11] Xie, Q. Self-training with noisy student improves imagenet classification / Q. Xie, M.T. Luong, E. Hovy, Q.V. Le // Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. – 2020 – P. 10687-10698.

О компьютерном зрении для мониторинга подводных ограждающих конструкций

А.С. Смагин

Хабаровский федеральный исследовательский центр ДВО РАН
Хабаровск, Россия
smaginkhv@gmail.com

Аннотация—Представлен метод автоматизации мониторинга состояния подводного сетчатого ограждения на примере предприятий аква-культурной направленности. Предложенный алгоритм выделения порывов сетчатых ограждений на основе искусственной нейронной сети показал хорошие результаты на реальных подводных снимках. Результаты, полученные программным комплексом, на основе созданного алгоритма приведены в статье.

Ключевые слова— алгоритм, компьютерное зрение, машинное обучение, программный комплекс.

1. ВВЕДЕНИЕ

Технологии компьютерного зрения и искусственные нейронные сети рассматриваются применительно к мониторингу состояния ограждающих конструкций подводных ферм на предприятиях аква-культурного профиля. Условия эксплуатации конструкций подводных ферм накладывают жесткие ограничения на непосредственное участие человека в их мониторинге и обслуживании. Поэтому разработка алгоритмов для систем компьютерного зрения подводных аппаратов, способных исключить или ограничить работу человека в сложных и опасных условиях подводного мира, является актуальной научно-практической задачей.

2. ПРЕДОБРАБОТКА ИЗОБРАЖЕНИЙ

Одной из основных задач компьютерного зрения является семантическая сегментация изображений [1], связанная с определением формы и положения объекта на изображении. Особую важность решение этой задачи имеет для определения характеристик аномальных областей рассматриваемых сетчатых ограждений, например, их дефектов [2].

На сложность выделения искомым объектов влияют такие факторы как: уровень яркости и его неравномерность, размытость изображения, низкая контрастность, а также посторонние объекты.

В разработанном программном комплексе поиск границ сетчатого ограждения выполняется двухэтапным алгоритмом предобработки изображений.

На первом этапе осуществляется компенсация артефактов, обусловленных освещением, с применением алгоритма Retinex [3, 4]. В его основе лежит идея разделения исходного изображения S на два, так чтобы яркость в каждой точке (x, y) рассчитывалась по формуле

$$S(x,y)=I(x,y)R(x,y).$$

Здесь I (illumination image) и R (reflectance image) - компоненты, отвечающие за освещение и отражение.

Используемая в программном комплексе модификация алгоритма Multi-Scale Retinex with Color Restoration хорошо приспособляется к условиям освещения и справляется с задачей повышения яркости и контрастности изображения, проявляя скрытые из-за слабой освещённости и искажений детали изображения.

На втором этапе выполняется непосредственно семантическая сегментация. Используется модификация сверточной нейронной сети семейства R-CNN (Region Based Convolutional Neural Network), в частности Mask R-CNN [5].

В Mask R-CNN к традиционным для алгоритмов семейства R-CNN метке класса и координатам ограничивающей рамки добавляется также маска объекта — прямоугольная матрица принадлежности пикселя текущему объекту. Ключевым архитектурным отличием является применение RoIAlign вместо RoIPooling. RoIAlign не выполняет округлений сдвигов для пулинга, а сохраняет его значения с плавающей точкой, используя билинейную интерполяцию. Это обеспечивает более точное выделение маски объекта.

3. ОБУЧАЮЩАЯ ВЫБОРКА И ПРОЦЕСС ОБУЧЕНИЯ

Важной частью обучения нейронной сети является подготовка обучающей выборки. Главная проблема её подготовки состоит в сложности поиска нужных изображений, т. к. искомые объекты должны быть похожи на разных изображениях. Желательно, чтобы обучающая выборка состояла из большого числа изображений [6]. В разных источниках рекомендуемый размер обучающей выборки составляет примерно тысячу изображений и более. В нашем случае, в связи с отсутствием такого числа подходящих изображений, выборка состояла из 40 изображений. Разметка производилась с помощью инструмента VGG (Visual Geometry Group) Image Annotator. Из-за сложности структуры объекта разметка одного изображения занимала около часа. В дальнейшем планируется расширение данной выборки за счёт применения различных фильтров, искажающих картину.

В нашем случае, также, определяющее значение имеет, способ разметки. Каждое изображение необходимо поделить на более простые сегменты. Были протестированы 2 варианта разметки: с выделением сторон ячеек и их объединением в линии. Лучшие результаты распознавания были получены при использовании второго варианта разметки.

Обучение нейронной сети производилось на гибридном вычислительном кластере Вычислительного центра ДВО РАН. Использовались 2 10-ти ядерных

процессора IBM POWER8 (3.49 ГГц turbo, 160 потоков); память ECC, 256 ГБ; 2 NVIDIA Tesla P100 GPU.

Обучение сети происходило за 100 эпох, по 200 шагов в каждой и по 100 шагов оценки (validation steps). Длительность обучения ~5200 секунд. Распознавание объектов с помощью весов обученной нейронной сети зависит от размера изображения и не занимает много времени даже на относительно слабом оборудовании. Для изображений с разрешением 1920x1080 время распознавания составляло не более 5 секунд.

4. АЛГОРИТМ ПОИСКА ДЕФЕКТОВ ОГРАЖДАЮЩЕЙ КОНСТРУКЦИИ

После обработки изображения нейронной сетью мы получаем черно-белую маску изображения. Из-за низкого качества изображения или ошибок в работе нейронной сети на полученной маске могут появляться дефекты – порывы ячеек, которых на исходном изображении нет. На рисунке 1, слева внизу на изображении, показаны примеры таких порывов. Они выделены красными контурами.

Алгоритм работает следующим образом.

1. Выделение предполагаемой области дефекта. Предполагаемое место порыва определяется, сравнением числа пикселей изображений внутренних соседних ячеек (площадей ячеек). Пусть N общее число ячеек на изображении, N_i - множество номеров ячеек, соседних с ячейкой с номером i . Зададим числовые параметры $0 < a < b$ и обозначим через θ_{ij} отношение площадей ячеек с номерами i и j . Если $\forall ij, i \in N, j \in N_i$ выполняется условие $\theta_{ij} < a$, то переходим к п.5. Если $\exists ij, i \in N, j \in N_i$, что $\theta_{ij} > b$, то переходим к п.4. В остальных случаях переходим к п.2. Опытным путем установлено что, лучшие результаты достигаются при $a=1.5, b=2.5$.

2. Исключение ложных порывов. Для этого строятся две пары параллельных прямых, продолжающих стороны ячейки в месте порыва. Если угол между парами прямых меньше 20 градусов, то в этом месте может быть ложный порыв и переходим к п.3. Иначе - к п.4.

3. Восстановление маски сетчатого ограждения. Выполняется проверка: если длина предполагаемого дефекта занимает менее 0.2 длины стороны ячейки, то восстанавливаем маску в области дефекта путем закрашивания разрыва. Переходим к п.5. Если длина дефекта занимает более 0.2 от длины стороны ячейки, то делается вывод, что на маске изображен реальный порыв. Переходим к п.4.

4. Помечаем получившуюся область, как область реального порыва сетчатого ограждения. Переходим к анализу следующего изображению.

5. Отмечаем, что на рассмотренном изображении нет дефектов. Переходим к анализу следующему.

Пример области реального порыва представлен на рисунке 1 (верхнее правое изображение).

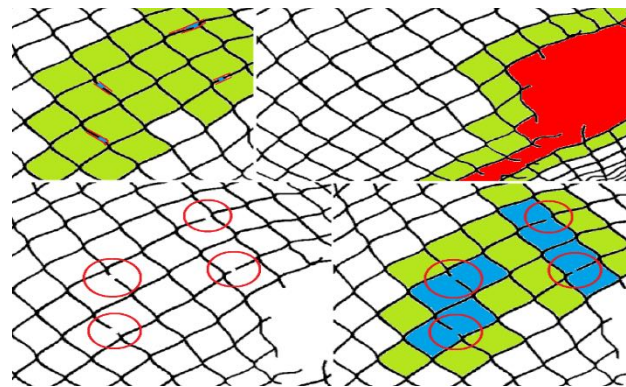


Рис. 1. Пример работы алгоритма

5. ЗАКЛЮЧЕНИЕ

Реализованный в программном комплексе подход к применению технологии компьютерного зрения при поиске порывов сетчатого ограждения, основанный на семантической сегментации, показал достаточно высокую эффективность и может стать основой для решения других сложных задач мониторинга ограждающих конструкций подводных ферм. Точность построения маски по метрике IoU (Intersection over Union) более 64%. Точность определения порывов 96%.

БЛАГОДАРНОСТИ

При выполнении исследования были использованы ресурсы ЦКП «Центр данных ДВО РАН».

ЛИТЕРАТУРА

- [1] Шапиро, Л. Компьютерное зрение (Computer Vision) / Л. Шапиро, Дж. Стокман. – М.: Бинум. Лаборатория знаний, 2006. – 752 с.
- [2] Ronneberger, O. U-Net: Convolutional Networks for Biomedical Image Segmentation / O. Ronneberger, Ph. Fischer, Th. Brox // Medical Image Computing and Computer-Assisted Intervention. – 2015. – Vol. 9351. – P. 234-241.
- [3] Huikai, W. FastFCN: Rethinking Dilated Convolution in the Backbone for Semantic Segmentation / H. Wu, J. Zhang, K. Huang, K. Liang, Y. Yu. – 2019. – P. 1-15. – Access mode: <https://arxiv.org/pdf/1903.11816.pdf>.
- [4] Смагин, А.С. Об алгоритмах компьютерного зрения для поиска порывов в сетчатых ограждающих конструкциях / А.С. Смагин, К.Н. Дубровин // Вычислительные технологии. – 2019. – Т. 24, № 6. – С. 117-124.
- [5] Kaiming, H. Mask R-CNN / H. Kaiming, G. Gkioxari, P. Dollár, R. Girshick. – 2018. – P. 1-12. – Access mode: <https://arxiv.org/pdf/1703.06870.pdf>
- [6] Nordstrom, T. Using and designing massively parallel computers for artificial neural networks / T. Nordstrom, B. Svensson // Journal of Parallel and Distributed Computing. – 1992. – Vol. 14(3). – P. 260-285.

Оценивание координат объектов с использованием сверточных сетей и карт глубины

Н.А. Андриянов
Финансовый университет при Правительстве Российской Федерации
Москва, Россия
naandriyanov@fa.ru

Аннотация—В работе рассмотрена задача оценивания координат объектов, обнаруживаемых по видеопоследовательности изображений с камеры Intel Real Sense. За счет использования сверточной нейронной сети YOLOv5 разработана система обнаружения объектов, для которой реализован алгоритм оценивания их координат в трехмерном пространстве на базе совместной обработки информации с карт глубины и оценок пиксельных координат объектов на изображении. Проведенный сравнительный анализ с методом инструментальных измерений показывал низкую погрешность предложенного алгоритма.

Ключевые слова— компьютерное зрение, искусственный интеллект, стереозрение, обнаружение объектов, оценивание координат.

1. ВВЕДЕНИЕ

Сегодня системы искусственного интеллекта все глубже и шире проникают в различные отрасли экономики. Особый интерес в ряде прикладных задач представляют системы компьютерного зрения, эффективность работы которых значительно повысилась с появлением глубоких сверточных нейронных сетей [1]. Сегодня существует множество популярных архитектур, которые используются в задачах обнаружения и распознавания объектов на изображениях, например, сети архитектуры YOLO [2]. Однако в ряде задач оценки положения объекта в плоскости оптического изображения недостаточно. Например, в робототехнике системы компьютерного зрения должны оценивать положение объектов в трехмерном пространстве [3]. Например, при автоматизации сбора урожая [4], необходимо точно знать, в какое положение направить манипулятор для захвата.

В данной работе рассматривается комплексный алгоритм, который позволяет оценить местоположение в трехмерном пространстве обнаруживаемого на видео объекта по результатам оценки его координат на кадре изображения с помощью сети YOLOv5 и построенной с помощью дальномера камеры Intel Real Sense карты глубины.

2. ДАННЫЕ С КАМЕРЫ И ИХ ОБРАБОТКА

Задача определения дальности до объектов сегодня не составляет большого труда и во многом является решенной. В используемой в работе камере Intel Real Sense Depth Camera D415 (рис. 1) применяется алгоритм измерения стереоглубины. Расстояние рассчитывается по различиям двух принимаемых изображений инфракрасных камер, направленных в одну сторону.



Рис. 1. Вид камеры Intel Real Sense Depth Camera D415

На рис. 1 слева направо: ИК-камера принимающая, ИК-излучатель, ИК-камера принимающая, камера оптического диапазона (непосредственно видеочкамера).

Преимуществом использования такой стереокамеры является то, что карта глубины и цветовое изображение не требуют дополнительного совмещения. Таким образом, мы получаем карту глубины, в которой для каждого пикселя на изображении рассчитано расстояние от камеры и цветовую картинку в каналах яркости RGB. Однако если обратить внимание на рис. 1, то можно заметить, что для корректного позиционирования относительно центра камеры следует учитывать смещение видеочкамеры от центра вдоль горизонтальной оси.

На рис. 2 показан интерфейс программы Real Sense Viewer при построении цветового изображения, а на рис. 3 – соответствующая цветному изображению карта глубины.

Следует отметить, что отдельно взятые кадры, поступающие с камеры, можно рассматривать как статичные изображения. Для этого будет использоваться библиотек OpenCV на Python. Также подключим сверточную нейронную сеть YOLOv5 с уже предобученными весами. В таком формате данная нейронная сеть позволяет производить обнаружение объектов, принадлежащих к стандартным классам датасета COCO [5].

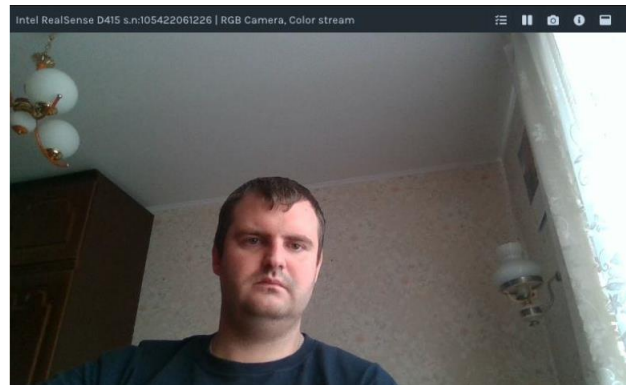


Рис. 2. Изображение в канале RGB

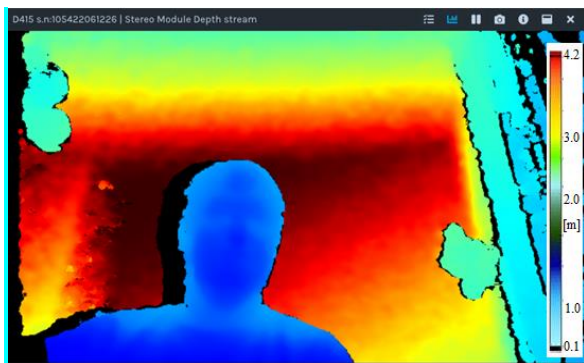


Рис. 3. Карта глубины в Intel RealSense Viewer

На рис. 4 представлен пример совместной обработки информации с цветного и глубинного изображений с использованием YOLOv5. По данным с RGB камеры определяется объект, который обводится прямоугольником, а расстояние до объекта определяется, как значение карты глубины в его центре.

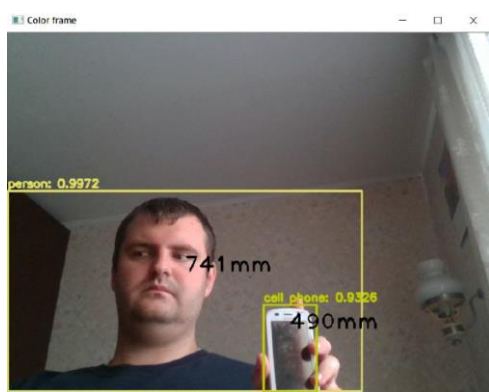


Рис. 4. Определение расстояния до объекта

Анализ рис. 4 показывает, что система компьютерного зрения обнаружила 2 объекта: человека и мобильный телефон на расстоянии 741 мм и 490 мм соответственно. При этом расстояние определяется как расстояние до центральной точки ограничивающего прямоугольника. Зная координаты ограничивающих прямоугольников в пикселях (принимая левую верхнюю точку за начало координат), получим выражение для пересчета относительных координат объектов [6]:

$$\begin{aligned} X[mm] &= d_{x_0,y_0} \frac{(C_x - x_0[px])}{f_x}, \\ Y[mm] &= d_{x_0,y_0} \frac{(C_y - y_0[px])}{f_y}, \end{aligned} \quad (1)$$

где сдвиги относительно камеры по осям X и Y рассчитаны в миллиметрах, d_{x_0,y_0} – значение карты глубины в центре ограничивающего объект прямоугольника, C_x и C_y – координаты центра изображения в пикселях, x_0 и y_0 – координаты центра ограничивающего прямоугольника в пикселях.

Для камеры Intel Real Sense Depth Camera D415 смещение видеофиксатора относительно центра составляет 35 мм, поэтому сделаем поправку:

$$X'[mm] = X[mm] - 35. \quad (2)$$

Понятно, что подставляя $Y[mm]$ из выражения (1), а $X[mm]$ из выражения (2), легко найти $Z[mm]$ по формуле евклидова расстояния, зная значение d_{x_0,y_0} .

Проверка алгоритма выполнялась для 5 произвольных объектов. Сравнение осуществлялось с измерениями с помощью угольников и линеек. В таблице 1 представлены результаты исследования.

Таблица 1. ОЦЕНКА ПОГРЕШНОСТИ ИЗМЕРЕНИЙ

№ объекта	Измерения		
	Пара рассчитанных координат (x,y)	Пара измеренных координат (x,y)	Корень из среднеквадратичной ошибки
1	(186, 25)	(190, 24)	(7.33, 5.27)
2	(320, 85)	(308, 82)	
3	(-76, -64)	(-82, -57)	
4	(-163, 96)	(-155, 100)	
5	(95, 252)	(92, 244)	

По результатам таблицы 1 можно сделать вывод о том, что погрешности в среднем не превышают 1 см.

3. ЗАКЛЮЧЕНИЕ

В работе предложено выполнять оценку координат системами компьютерного зрения с помощью комплексирования карт глубины и оценок на базе глубокого обучения. Исследования предложенного алгоритма показали, что оценки измерений имеют низкую погрешность: в среднем 5-7 мм.

БЛАГОДАРНОСТИ

Исследование выполнено при поддержке РФФИ, Грант №19-29-09048 мк.

ЛИТЕРАТУРА

- [1] Krizhevsky, A. ImageNet Classification with Deep Convolutional Neural Networks / A. Krizhevsky, I. Sutskever, G.E. Hinton // NIPS'12 Proc 25th Int Conf Neural Inf Process Syst. – 2012. – Vol. 1. – P. 1097-1105.
- [2] Nepal, U. Comparing YOLOv3, YOLOv4 and YOLOv5 for Autonomous Landing Spot Detection in Faulty UAVs / U. Nepal, H. Eslamiat // Sensors. – 2022. – Vol. 22. – P. 464. DOI: 10.3390/s22020464.
- [3] Švec, M. Calculation of object position in various reference frames with a robotic simulator / M. Švec, I. Farkas // Proceedings of 36th Annual Conference of the Cognitive Science Society. – 2014. – P. 1-4.
- [4] Andriyanov, N. Intelligent System for Estimation of the Spatial Position of Apples Based on YOLOv3 and Real Sense Depth Camera D415 // N. Andriyanov, I. Khasanshin, D. Utkin, T. Gataullin, S. Ignar, V. Shumaev, V. Soloviev // Symmetry. – 2022. – Vol. 14. – P. 148. DOI: 10.3390/sym14010148.
- [5] COCO [Electronic resource]. – Access mode: <https://cocodataset.org/> (08.02.2022).
- [6] Laganier, R. Robust object pose estimation from feature-based stereo / R. Laganier, S. Gilbert, G. Roth // IEEE Trans. Instrum. Meas. – 2006. – Vol. 55. – P. 1270-1280.

Мультизадачное обучение интеллектуальных агентов на скрытых представлениях

И.Н. Аглюков
Высшая Школа Экономики
Москва, Россия
ildariwe@gmail.com

К.В. Святлов
Ульяновский государственный
технический университет
Ульяновск, Россия
k.svyatov@ulstu.ru

С.В. Сухов
УФирЭ им. В. А. Котельникова РАН
Ульяновск, Россия
ssukhov@ulireran.ru

Аннотация—Рассматривается проблема объединения опыта нескольких агентов при мультизадачном глубоком обучении с подкреплением. Обмен опытом осуществляется путем дистилляции знаний между предобученными агентами-учителями и учеником. Буфер опыта, используемый для хранения накопленных состояний, преобразуется в скрытые представления автокодировщика, что снижает требования к используемым вычислительным ресурсам.

Ключевые слова— обучение с подкреплением, катастрофическое забывание, автокодировщик, дистилляция знаний.

1. ВВЕДЕНИЕ

Показывая поразительные результаты в решении различных частных задач, при попытках мультизадачного обучения, глубокое обучение с подкреплением сильно страдает от проблемы “катастрофической интерференции” [1]. Градиенты, привносимые сразу несколькими задачами, выступают в роли шума, который мешает нормальному обучению агента. При последовательном обучении агента на нескольких задачах в нескольких окружениях предыдущий опыт взаимодействия со средами (окружениями) может храниться в так называемом буфере опыта. При этом возникает другой ряд проблем, связанный с обучением агента без взаимодействия со средой на ранее полученном опыте (при “офлайн обучении”). Здесь очевидной проблемой является ограниченность полученного опыта и отсутствие возможности исследовать среду в поисках областей пространства состояний, дающих большее вознаграждение [2]. Другой не менее важной проблемой является проблема хранения состояний, так как многие подходы обучения с подкреплением основаны на постепенных инкрементальных улучшениях поведения интеллектуального агента, происходящих за счет многократного прохождения по накопленному буферу, а решение проблемы катастрофической интерференции требует хранения максимально возможного количества состояний для процесса передачи знаний.

В нашей работе был разработан метод, способный в значительной мере сократить потери при многозадачном обучении с подкреплением без взаимодействия со средой (Рис. 1). Также был опробован способ сокращения объема необходимой памяти для хранения буфера опыта путем сжатия состояний во внутреннее представление нейронной сети.

2. ОПИСАНИЕ МЕТОДА ПЕРЕДАЧИ ЗНАНИЙ

На первом этапе мы обучаем агентов-учителей с помощью немного модифицированного алгоритма

глубокого Q обучения (Dueling Deep Q learning) [3]. В качестве сред для обучения агентов были выбраны среды эмулирующие игры Atari. Во время обучения мы сохраняли наборы состояний для дальнейшего процесса передачи знаний между учителями и учеником (копией одного из агентов учителей). Сама передача знаний осуществлялась посредством популярного в глубоком обучении метода “Дистилляции” [4], что позволило нам свести проблему обучения с подкреплением к проблеме обучения с учителем. Агентам-учителям подается набор состояний из их собственных буферов опыта; тот же набор подается ученику. Затем значения функции полезности (Q-функции) учителей и ученика пропускаются через функцию softmax, тем самым получая вероятность того или иного действия. Полученные значения подаются в функции потерь для каждой задачи и по сумме результатов функций потерь происходит расчет градиента и оптимизация сети с помощью алгоритма обратного распространения ошибки.

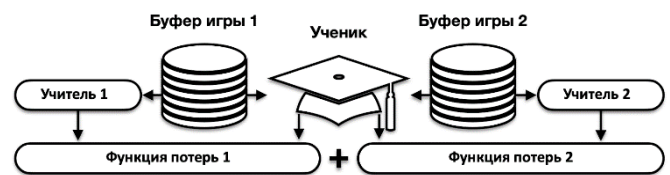


Рис. 1. Диаграмма процесса передачи знаний

Описанный алгоритм апробирован на наборе игр Atari с разным количеством действий (4 действия - Breakout, Atlantis, 6 действий - Pong, Up and Down, DemonAttack) и различными функциями потерь. Наши исследования показали, что агент ученик набирает сопоставимое с учителем количество очков, но немного уступает учителю (Таблица 1).

ТАБЛИЦА 1. РЕЗУЛЬТАТЫ ПЕРЕНОСА ЗНАНИЙ

Учитель 1 Учитель 2	Учитель	Ученик на части буфера	Ученик на скрытых представлен иях
Atlantis	$(2,7 \pm 1,3) \times 10^5$ 212±45	$(2,3 \pm 1,8) \times 10^5$ 200±53	$(1,8 \pm 0,2) \times 10^5$ 8,4±2
Breakout	9251±3301	6169±2599	9284±2987
Demon Attack Up and Down	6607±917	4526±328	4893±471
Pong	15,9±1,43	15,54±0,86	15,43±0,78
Up and Down	6608±918	6336±886	6397±686

Средние набранные очки агентов за тест

Процесс тестирования агента ученика состоит из 10 циклов по 12 500 состояний. Такое количество состояний примерно соответствует тридцати минутам игры в

реальном времени. Ограничение по времени тестирования было продиктовано тем, что агент может заиклиться, и игра будет продолжаться бесконечно, но агент не будет набирать очки. Мы брали среднее количество очков за один цикл, а потом среднее за все циклы.

Одной из проблем «офлайн обучения» является отсутствие возможности поместить весь сохранённый объем знаний учителя в оперативную память. Как следствие, буфер для передачи знаний не покрывает даже весь спектр состояний, которые были доступны учителю при первичном обучении. Одно из решений – использовать для обучения лишь часть буфера (Таблица 1). Другое возможное решение этой задачи – перевод содержимого буфера в сжатое представление. Для сжатия состояний мы воспользовались автокодировщиком [5], который позволил сократить потребляемую память приблизительно в двадцать семь раз путем перевода изображений во внутреннее представление нейронной сети.

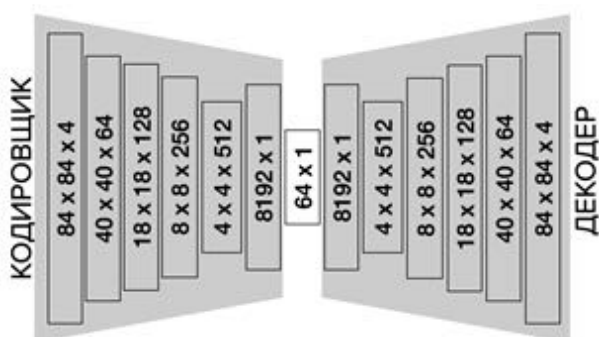


Рис. 2. Схема автокодировщика

Для каждой среды был обучен автокодировщик, сжимающий четыре кадра состояния в 64 цифры с плавающей точкой (Рис. 2). Автокодировщик обучался на всем буфере каждой среды, а затем все сохраненные состояния были закодированы и сохранены на диске. При использовании автокодировщика алгоритм процесса передачи знаний имеет лишь небольшое изменение в части получения состояния, а именно – декодирование из скрытого представления набора состояний для последующей дистилляции знаний. Результаты обучения агента-ученика с помощью скрытых представлений показаны в Таблице 1. Так как известные ограничения

автокодировщика делают изображение мутным, такой способ подходит не ко всем средам и в наших экспериментах в игре Breakout была потеряна часть информации, на которую, по-видимому, опирался агент учитель. Это привело к плохим результатам при передаче знаний в данном конкретном случае.

3. ЗАКЛЮЧЕНИЕ

Мы предложили и протестировали разновидность метода клонирования поведения с помощью дистилляции. Результаты передачи знаний представлены в Таблице 1. Разработанный нами метод успешно справляется с проблемой катастрофического забывания и может использоваться при многозадачном обучении с подкреплением.

Использование набора скрытых признаков позволяет уменьшить объем используемой памяти для хранения буфера опыта, но увеличивает время обучения агента, так как приходится тратить время на декодирование набора состояний для дальнейшей дистилляции знаний. Стоит отметить и тот факт, что такой метод имеет свои ограничения ввиду особенностей автокодировщика и требует дальнейшего совершенствования подхода.

БЛАГОДАРНОСТИ

Исследование выполнено при финансовой поддержке Российского Фонда Фундаментальных Исследований и Правительства Ульяновской области (проект № 18-47-732006).

ЛИТЕРАТУРА

- [1] Teh, Y.W. Distral: Robust multitask reinforcement learning / Y.W. Teh, V. Bapst, W.M. Czarnecki, J. Quan, J. Kirkpatrick, R. Hadsell, N. Heess, R. Pascanu // *Advances in Neural Information Processing Systems*. – 2017. – P. 4496-4506.
- [2] Levine, S. Offline Reinforcement Learning: Tutorial, Review, and Perspectives on Open Problems / S. Levine, A. Kumar, G. Tucker, J. Fu // *ArXiv preprint*: 2005.01643, 2020.
- [3] Wang, Z. Dueling Network Architectures for Deep Reinforcement Learning / Z. Wang, T. Schaul, M. Hessel, H. Hasselt, M. Lanctot, N. Freitas // *PMLR*. – 2016. – Vol.48. – P.1995-2003.
- [4] Hinton, G. Distilling the knowledge in a neural network / G. Hinton, O. Vinyals, J. Dean // *ArXiv preprint*: 1503.02531, 2015.
- [5] Ballard, D. Modular learning in neural networks / D. Ballard // *Proc. of the sixth National conf. on Artificial intelligence*. – 1987. – Vol. 1. – P. 279-284.

Слияние знаний в импульсных нейронных сетях

Д.И. Антонов

Ульяновский государственный технический университет
Ульяновск, Россия
d.antonov@ulireran.ru

С.В. Сухов

УФИРЭ им. В.А. Котельникова РАН
Ульяновск, Россия
ssukhov@ulireran.ru

Аннотация—Показана методика организации слияния знаний (knowledge fusion) импульсной искусственной нейронной сети (ИмНС, Spiking Neural Network) без доступа к первоначальным исходным данным. Разработан метод ранжирования выходных нейронов ИмНС по степени влияния на процесс классификации на основе знания о распределении весов сети. Эксперименты проводились на свободно доступных наборах данных в среде PyTorch.

Ключевые слова— импульсные нейронные сети, спайковые нейронные сети, слияние знаний, ансамбль сетей, прунинг сетей.

1. ВВЕДЕНИЕ

За последнее десятилетие разработчики моделей глубокого обучения (deep learning) добились больших успехов в различных направлениях, в частности в области машинного зрения (computer vision, pattern recognition). Более широкое использование потенциала уже обученных моделей можно реализовать объединением полезной информации нескольких моделей. Один из таких эффективных методов относится к методам ансамблей нейронных сетей – бэггинг (bagging). При бэггинге несколько сетей параллельно обрабатывают тестовые данные и по итогам проводят голосование за консолидированный результат. Но хорошая точность такого способа объединения знаний достигается при условии, что все элементарные классификаторы обучаются и работают параллельно (независимо друг от друга), что может быть довольно громоздким решением.

Работа с громоздкими нейронными сетями стала возможной благодаря широкому распространению графических процессоров (GPU), производящих большое количество параллельных вычислений. Но большую сеть, размещающуюся на сервере, не разместить в небольших цифровых устройствах. Обходят эту проблему, уменьшая размер сети методом обрезки или прунинга (pruning) за счёт сокращения части её параметров.

Попытка прямого дообучения на новых данных обученной ранее сети сталкивается с явлением катастрофического забывания (catastrophic forgetting) предыдущих знаний [1]. В данной работе мы описываем способ избежать катастрофического забывания используя так называемое слияние знаний (knowledge fusion) [2]. Слияние знаний в нашем случае использует идеи прунинга и даёт возможность приближаться к точности бэггинга, избегая его тяжеловесности.

2. ПОСТАНОВКА ЗАДАЧИ

Рассмотрим две импульсные нейронные сети, обученные каждая на своем наборе данных. Будем полагать, что первые слои в обеих сетях одинаковы и выполняют функцию выделения признаков. Последний слой сети выполняет функцию классификатора. Наша цель – определить наименее важные нейроны в выходном слое с целью обрезки соответствующих связей.

Обрезанные связи могут быть заменены связями из другой сети, в результате чего образуется сеть, способная уверенно классифицировать объекты из двух наборов, но обладающая таким же набором параметров, как и исходная сеть. В работе представлено решение поставленной задачи для варианта, который не требует доступа к первоначальному набору обучающих данных.

3. ЭКСПЕРИМЕНТЫ

Для построения ИмНС мы использовали пакет PyTorch, кодирующий информацию временем прихода импульсов. Использованная архитектура ИмНС была взята из работ [3, 4]. Обучение нейронных сетей производилось на выборках из наборов изображений рукописных цифр MNIST [5] и букв EMNIST [6], содержащих по 10 классов каждый и состоящие из 24000 тренировочных и 4000 тестовых изображений.

Архитектура ИмНС состоит из пакета для предобработки поступающих изображений (пакет фильтров разности гауссиан, Difference of Gaussians или DoG-фильтров) и собственно сети (Рис.1). Сеть состоит из последовательно расположенных трех частей, содержащих свёрточный слой (convolution layer) и слой подвыборки или пуллинговый слой (pooling layer). Тренировка первых 2-х свёрточных слоёв производится «без учителя» на основе принципа пластичности временной зависимости импульса (spike-timing-dependent plasticity или STDP), а тренировка 3-го слоя - на базе принципа пластичности временной зависимости импульса с подкреплением (reward-modulated STDP или R-STDP).

Способ слияния знаний, реализованный нами, в существенной мере следовал из двух экспериментов по определению степени важности нейронов выходного слоя для классификации изображений.

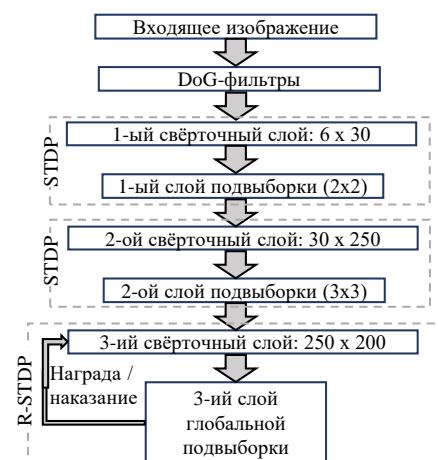


Рис.1. Архитектура нейронной сети

Первый эксперимент заключался в подсчёте количества использований сетью каждого нейрона

выходного слоя в процессе классификации. Было выяснено, что нейроны сильно неравнозначны по степени активности. За время распознавания упомянутого выше тестового набора данных из 4000 изображений частота использования каждого из 200 нейронов выходного слоя колебалась примерно от 0 до 120 раз (Рис.2).

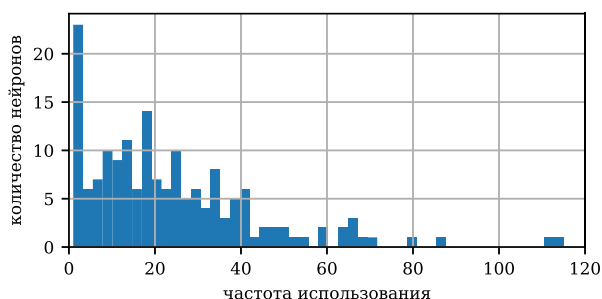


Рис.2 Частота использования нейронов выходного слоя при тестировании набора в 4000 изображений

Второй эксперимент заключался в поиске зависимости активности нейрона от параметров сети. Наиболее высокой оказалась корреляция между степенью активности нейрона и суммой наибольших 10% весов, связанных с каждым нейроном.

Процедура определения степени важности нейронов была использована для создания метода слияния знаний. Предварительно было получено две сети, имеющие идентичную архитектуру, описанную выше. Первая из них была получена тренировкой на выборке из набора изображений цифр MNIST в течение 50 эпох с точностью классификации (accurasy) около 94%. Точность классификации A определялась по следующей формуле:

$$A = D_{\text{correct}} / D_{\text{total}},$$

где D_{correct} – количество правильно классифицированных данных, D_{total} – общее количество данных. Вторая сеть была натренирована на изображениях рукописных букв набора EMNIST. Процесс обучения отличался от обучения первой сети. Первые два слоя второй сети были идентичны первым двум слоям первой сети, уже натренированным на изображениях цифр, и повторно они не тренировались на изображениях букв. Данные первые два слоя выполнили роль универсального выделителя признаков. Третий слой сети был натренирован в течение 50 эпох на наборе из изображений рукописных букв. В итоге натренированная вторая сеть научилась распознавать буквы с точностью около 90%, но её точность классификации цифр была около 8%.

На базе двух сетей методом слияния знаний была создана третья сеть, имеющая такую же архитектуру, как и первые две, и максимально получающая знания от первой и второй сетей. Первый шаг процедуры слияния знаний состоял в прунинге 3-их слоёв первой и второй сети. Третий свёрточный слой данного типа сети содержит веса, относящиеся к 200 нейронам, по 20 нейронов на каждый из 10 классов. В каждом классе нейроны были проранжированы нами по параметру частичной суммы весов. Расчёт параметра частичной суммы весов состоял в выделении из весов, относящихся к каждому нейрону, максимальных 10% весов и суммировании их значений. После ранжирования

нейронов, из каждого класса первой сети было выделено по 10 нейронов с максимальным параметром частичной суммы весов. Такое же выделение нейронов была сделано во второй сети.

Второй шаг процедуры состоял в непосредственном слиянии знаний в создаваемой сети. Первые два слоя без изменений были перенесены из исходных сетей в итоговую сеть. А третий слой был скомпонован из нейронов 3-их слоёв первой и второй сетей таким образом, что каждый класс получил по 10 предварительно выделенных нейронов из первой сети и по 10 нейронов из второй.

В результате проведённого слияния знаний была получена сеть, которая смогла распознавать как цифры со средней точностью 87%, так и буквы со средней точностью 75%. Точность бэггинга на этом же наборе составила 91% на изображениях цифр и 78% на изображениях букв. Таким образом, метод слияния знаний позволил получить сеть вдвое меньшего размера с падением точности классификации на 3-4%.

4. ЗАКЛЮЧЕНИЕ

Полученная методика слияния знаний ИмНС достаточно проста и результативна, но имеет, как минимум, два ограничения:

- тренировочные наборы должны быть достаточно похожими и первые слои сетей должны быть достаточно универсальными, чтобы эти слои смогли бы выделить из наборов общие признаки,
- количество «сливаемых» знаний ограничено сверху количеством выходных нейронов.

Авторы видят дальнейшее развитие своей работы в разработке методики слияния знаний для более общего вида сетей.

БЛАГОДАРНОСТИ

Исследование выполнено при финансовой поддержке РФФИ в рамках научных проектов № 20-07-00974 и № 18-47-732006.

ЛИТЕРАТУРА

- [1] McCloskey, M. Overcoming catastrophic forgetting in neural networks / M. McCloskey, N.J. Cohen // Psychology of Learning and Motivation – Advances in Research and Theory. – 1989. – Vol. 24(C). – P. 109-165.
- [2] Akhlaghi, M.I. Knowledge Fusion in Feedforward Artificial Neural Networks / M.I. Akhlaghi, S.V. Sukhov // Neural Processing Letters. – 2018. – Vol. 48(1). – P. 257-272.
- [3] Mozafari, M. SpykeTorch: Efficient simulation of convolutional spiking neural networks with at most one spike per neuron / M. Mozafari, M. Ganjtabesh, A. Nowzari-Dalini, T. Masquelier // Frontiers in Neuroscience. – 2019. – Vol. 13. – P. 1-12.
- [4] Mozafari, M. Bio-inspired digit recognition using reward-modulated spike-timing-dependent plasticity in deep convolutional networks / M. Mozafari, M. Ganjtabesh, A. Nowzari-Dalini, S.J. Thorpe, T. Masquelier // Pattern Recognition. – 2019. – Vol. 94. – P. 87-95.
- [5] The MNIST database of handwritten digits Web Site [Electronic resource]. — Access mode: <http://yann.lecun.com/exdb/mnist/>.
- [6] The EMNIST Dataset [Electronic resource]. — Access mode: <https://www.nist.gov/itl/products-and-services/emnist-dataset>.

Hyper-parameter tuning for long short-term memory (LSTM) algorithm to forecast a disease spreading

T. Makarovskikh
South Ural State University
Chelyabinsk, Russia
makarovskikh.t.a@susu.ru

M. Abotaleb
South Ural State University
Chelyabinsk, Russia
abotalebmostafa@bk.ru

Abstract—Deep learning, artificial intelligence, and machine learning are ways for technologies to support decision-making in real-time to control the spread of the pandemic, which can help healthcare organizations. This study aims to investigate hyper-parameter tuning for Long Short-Term Memory to forecast Covid-19 infection cases in the Russian Federation by pick the best loss function, activation function, number of epochs, number of neurons in a cell, and optimizer to minimize the error in addition to a good fit for the model where the performance of the model is good on both the training and validation sets. Based on machine learning long short-term memory (LSTM), which has the advantage of analyzing relationships among time series data through its memory function, we propose a forecasting method for daily Covid-19 infection cases based on bidirectional LSTM. In the meanwhile, we use about 10 different forecasting models to forecast the daily Covid-19 infection cases one by one. Moreover, the results of these models are analyzed and compared. Firstly, we concluded that bidirectional long short-term memory to efficiently extract features from the data, which are the items of the previous day. And then, we adopt Bidirectional LSTM to predict the daily infection cases with the extracted feature data. We used BiLSTM to forecast Covid-19 spreading in the Russian Federation for one month. The prediction results were expected. According to the experimental results, the Bidirectional LSTM can provide a reliable daily Covid-19 infection case forecast in Russia with the highest prediction accuracy.

Keywords—Hyper-parameter, LSTM, BiLSTM, ConvLSTMs, LSTMs, Machine learning, time series analysis, forecasting.

1. INTRODUCTION

With the increasing number of Covid-19 infections and the recognition of new Omicron variant, it is necessary to track new infection cases in order to control and monitor meetings, events, and provide assistance in places where the virus is rapidly spreading. Covid-19 has the ability to be the most global epidemic of this decade [1]. For forecasting the global number of Covid-19 infection cases, [2] developed a novel hybrid forecasting model, CNN–LSTM. While the proposed model was compared to 17 baseline models, CNN–LSTM outperformed them all, with the lowest root mean square error (RMSE). There is study for forecasting infection cases in India and Chennai accurately using deep learning models. The authors concluded that stacked LSTM (LSTMs) outperform the ARIMA, LSTM, and Prophet [3]. The LSTM model is the best performer for forecasting cumulative infection cases for 7 and 30 days [4]. A study was held in Egypt, Kuwait, and Saudi Arabia to forecast Covid-19 cases, and authors concluded that LSTM had the best performance in infection cases for the three countries, and GRU had the best performance in death cases for Egypt and Kuwait [5]. For infection and recovered cases, the LSTM-CNN achieved improved performance with an average mean absolute

percentage error (MAPE) among others [6]. Empirical studies conducted show that deep learning-based algorithms such as ANN and LSTM outperform traditional-based algorithms such as the ARIMA model [7]. The aim of this research is to investigate hyper-parameter tuning for long short-term memory (LSTM) models to forecast Covid-19 infection cases in Russia. After we compared the mean absolute error (MAE), RMSE, and R-squared (R^2), we found that the BiLSTM model is the best model for forecasting Covid-19 infection cases in Russia for the current time, as it achieved the least error in MAE and RMSE as well as the highest R^2 .

2. PROPOSED FRAMEWORK

The mechanism of the proposed framework for forecasting of time series $x_1, x_2, x_3, \dots, x_n$ is following.

Step 1. Data preparation. We have some problems with lengthy sequences in the database [8]. The first problem is that it takes a long time for training and needs a lot of memory, and the second problem is getting a poor learned model caused by back-propagating lengthy sequences. So, it is important to prepare and preprocess the data before it is imported to neural networks. Data normalization and standardization are two techniques used in the data preparation phase. We used data standardization, which is realized as a scaling approach to set the mean and standard deviation to 0 and 1, respectively [9].

Step 2. Split dataset into training, validation, and testing sets. Daily Covid-19 infection cases in the Russian Federation Data from the World Health Organization were collected from the January, 3, 2020 to December, 31, 2021. We test our model using 5% of this dataset (37 days). The dataset is further prepared such that the first 95% of the datasets are used for training and the last 5% of the datasets are used for testing purposes. The training set is used to train and improve the models, and we divided the 20% of training data for validation purposes to diagnose overfitting and underfitting, whereas the test set is used to assess the performance of the model. The algorithm and the daily covid 19 infection case data can be obtained from [10].

Step 3. Models' development. In this step, we run our algorithm for LSTM, LSTMs (stacked LSTM), bidirectional LSTM, LSTMs (stacked BiLSTM), convolutional neuron networks (Conv), ConvLSTMs and other forecasting models.

Step 4. Evaluate the forecasting using data from [11]. In order to evaluate the forecasting effect of hyper-parameter tuning for machine learning models, the mean absolute error (MAE), root mean square error (RMSE), and R-square (R^2) are used as the evaluation criteria for the methods (see Table 1).

3. RESULTS

Таблица I. TABLE 1. COMPARISON OF 10 METHODS EVALUATION TESTING COVID 19 DAILY INFECTION CASES IN RUSSIA

Method	MAE	RMSE	R ²
LSTM	621.432	753.053	0.962
Stacked LSTM	1025.523	1273.222	0.893
LSTM 2 window	711.051	857.457	0.949
LSTM 3 window	642.859	776.476	0.957
LSTM 4 window	604.986	741.646	0.960
LSTM 10 window	640.832	742.132	0.949
BILSTM	552.889	658.607	0.971
Stacked BILSTM	965.569	1192.145	0.906
CONV	849.763	921.963	0.578
ConvLSTMs	1189.480	1350.587	0.094

The bold values present the lowest error values of (MAE, RMSE, and R²).

These values show that the proposed approach outperforms the baseline models based on the testing data. After using the processed training set data to train LSTM, Stacked LSTM (LSTMs), LSTM 2 window, LSTM 3 window, LSTM 4 window, LSTM 10 window, BILSTM, Stacked BILSTM, CONV, and ConvLSTMs, respectively, the model completed by training is used to predict the test set data, and the real value is compared with the predicted. From Table 1 the MAE and RMSE of ConvLSTMs are the largest and R² is the smallest, while the MAE and RMSE of bidirectional LSTM (BI-LSTM) are the smallest, R² is the largest, and the closest is 1.

The results show the performance of the BI-LSTM developed model is the best among the other 10 methods. In terms of forecasting accuracy, MAE and RMSE are the smallest among the considered models. Hence, this method has high forecasting accuracy for the investigated dataset. In terms of forecasting performance, the R² of BI-LSTM can be compared with the other methods. Therefore, the BI-LSTM proposed in this paper is superior to the other considered comparative models in terms of fitting degree and error value.

In our research, the role of BiLSTM is crucial, where prediction depends not only on the previous input but also the future input. Because of the nature of daily Covid-19 infection cases in Russia, BiLSTM outperforms the rest of the models. The network needs to increase the amount of information available to the network and does so through two LSTMs that work in two different directions, one taking the input in a forward direction and the other in a backwards direction, so BiLSTM works effectively. We designed the BiLSTM to look at the previous day to forecast the next day with the minimum error.

4. CONCLUSION

Despite that the spread of Covid-19 is recognized as an impossible process for forecasting, the available data sets can be used to study the properties of various predictive models, in particular, for the development of artificial intelligence methods. These methods can be used to forecast not only epidemic data. Developed approaches to improve the accuracy of forecasting can increase the quality of simulation for other processes, such as socio-economic, technological, physical,

etc. In our case, the using more than one time step in BiLSTM lead to increasing of error. Hence, we used only one time step, meaning that the forecast for infection cases for the next day depends on the previous day. We used BiLSTM to forecast Covid-19 spreading in the Russian Federation for one month [11]. The prediction results were expected. In the absence of a new Covid-19 variant, the current wave, which began in the first days of January, will go down in about a month, which is currently observed. If a new variant appears, another wave is possible most probably before the end of summer. Since the situation with virus spreading changes daily, we recommend rerunning our proposed framework every day to get results with lower error. This forecasting method not only provides a new research idea for infection case forecasting but also provides practical experience for scholars. On the other hand, it provides a reference for the health sector in Russia, in particular, and the world health organization WHO, and in general for the health sectors in other countries. In future research, We will develop architecture of the Attention LSTM model and Attention BiLSTM model to minimize the errors and compare the errors with those of this research.

ACKNOWLEDGEMENTS

The work was supported by the Ministry of Science and Higher Education of the Russian Federation (government order FENU-2020-0022).

REFERENCES

- [1] Kannan, S.P. COVID-19 (Novel Coronavirus 2019)-recent trends / S.P. Kannan, P.S. Ali, A. Sheeza, K. Hemalatha // Eur Rev Med Pharmacol Sci. – 2020. – Vol. 24(4). – P. 2006-2011.
- [2] Zuhaira, Z.M. COVID-19 pandemic forecasting using CNN-LSTM: a hybrid approach / Z.M. Zuhaira, N.M. Alturki // Journal of Control Science and Engineering. – 2021. – Vol. 2021.
- [3] Devaraj, J. Forecasting of COVID-19 cases using deep learning models: Is it reliable and practically significant? / J. Devaraj, R.M. Elavarasan, R. Pugazhendhi, G.M. Shafiullah, S. Ganesan, A.K. Jeysree, I.A. Khan, E. Hossain // Results in Physics. – 2021. – Vol. 21. – P. 103817.
- [4] Marzouk, M. Deep learning model for forecasting COVID-19 outbreak in Egypt / M. Marzouk, N. Elshaboury, A. Abdel-Latif, S. Azab // Process Safety and Environmental Protection. – 2021. – Vol. 153. – P. 363-375.
- [5] Omran, N.F. Applying deep learning methods on time-series data for forecasting COVID-19 in Egypt, Kuwait, and Saudi Arabia / N.F. Omran, S.F. Abd-el Ghany, H. Saleh, A. Abdelmgeid, A. Gumaei, M. Al-Rakhami // Complexity. – 2021.- Vol. 2021.
- [6] Dairi, A. Comparative study of machine learning methods for COVID-19 transmission forecasting // A. Dairi, F. Harrou, A. Zeroual, M.M.Hittawe, Y. Sun // Journal of Biomedical Informatics. – 2021. – Vol. 118. – P. 103791.
- [7] Abdollahi, J. Modeling and forecasting Spread of COVID-19 epidemic in Iran until Sep 22, 2021, based on deep learning / J. Abdollahi, A.J. Irani, B. Nouri-Moghaddam // ArXiv preprint: 2103.08178, 2021.
- [8] Khan, S.D. Toward smart lockdown: a novel approach for COVID-19 hotspots prediction using a deep hybrid neural network / S.D. Khan, L. Alarabi, S. Basalamah // Computers. – 2020. – Vol. 9(4). – P. 99.
- [9] Ali, P.J.M. Data normalization and standardization: a technical report / P.J.M. Ali, R.H. Faraj, E. Koya // Mach Learn Tech Rep. – 2014. – Vol. 1(1). – P. 1-6.
- [10] Abotaleb, M. Hyper-parameter Tuning for Long Short-Term Memory (LSTM) Algorithm to Forecast a Disease Spreading / M. Abotaleb, T. Makarovskikh. URL: <https://github.com/abotalebmstafa11/Hyper-parameter-Tuning>.
- [11] Coronavirus: statistics. URL: <https://yandex.ru/covid19/stat>.

Automatic classification infectious disease X-ray images based on deep learning algorithms

T. Makarovskikh
South Ural State University
system of programming department
chelyabinsk, Russia
makarovskikh.t.a@susu.ru

A. Salah
Ain Shams University
Computers and Systems Engineering
Cairo, Egypt
Amir.salah@live.com

A. Badr
New England University
school of Science and Technology
Armidale, NSW, Australia
Amr.mostafa@live.com

P. Mishra
Jawaharlal Nehru Krishi Vishwa
mathematics and Statistics department
Powarkheda, (M.P.), India
pradeepjnkvv@gmail.com

M. Abotaleb
South Ural State University
system of programming department
chelyabinsk, Russia
abotalebmostafa@bk.ru

Abstract—Recent technological advancements allow deep learning to be employed in practically every aspect of life. Because deep learning techniques are so precise, they can be utilized in medicine to classify and detect various diseases. The coronavirus (SARSCoV2) epidemic has recently affected global health systems. SARSCoV2 may be diagnosed via PCR and medical imaging. A chest X-ray is used to diagnose SARSCoV2. This paper proposes a deep learning technique to distinguish SARSCoV2 positive and normal cases. X-rays are the traditional method for diagnosing SARSCoV2, and deep learning models have proven their superior ability to classify medical images, which will be the tool in the future for the classification of any other epidemics that may appear in the future. In this study, not only are the deep learning models fine-tuned, but also the hyperparameters are fine-tuned, which significantly improves the performance of the fine-tuned deep learning models. we developed a system based on deep learning algorithms to classify x-rays. We used five models: Xception, VGG19, ResNet50, DenseNet121, and Inception. We provide deep learning models and algorithms that were trained and evaluated to assist medical efforts and reduce medical staff workload in handling SARSCoV2. Further, the proposed classification model provides good results by leveraging accurate classification of SARSCoV2 disease based on medical images. Besides, the performance of our proposed CNN classification method for medical imaging has been assessed based on different edge-based neural networks. Whenever there is an increasing number of a class in the training network, the accuracy of tertiary classification with CNN will decrease. The proposed model has achieved a 0.9897 accuracy in tertiary classification that includes normal, SARSCoV2 positive, and normal. The proposed algorithm obtains good classification accuracy during the binary classification procedure with the DenseNet121 model by supervised contrastive learning loss function.

Keywords— Convolution Neural Network, Covid 19, Xception, vgg19, ResNet50, DenseNet121, Inception, X-ray Image Classification.

1. INTRODUCTION

The first case of COVID-19 in the Chinese city of Wuhan was misdiagnosed. Whereas specialists initially believed that it was pneumonia, which led to the rapid spread of this epidemic and its spread throughout the country and the world [1-3]. The world health organization called that disease SARS-CoV-2. Covid 19 took thirty days to spread all over China [4]. On January 30, 2020, the world health organization declared the SARS-CoV-2 outbreak [5].

The first imaging tool to be used in the diagnosis of COVID-19 disease is the chest X-ray, which is the most commonly used. A negative chest X-ray is shown on the left in Figure 1, whereas a positive chest X-ray is shown on the right in Figure 1. for the purpose of automatically classifying digitized chest images A number of classical machine learning models have been used [6-7]. Using a Support Vector Machine (SVM) classifier, three statistical characteristics of lung texture were computed and classified as malignant or benign lung cancer [8]. in the case of images as normal or cancerous. With enough annotated images, deep learning approaches have outperformed classical machine learning approaches [9-10]. In medical imaging, the convolution neural network (CNN) architecture is a popular deep learning approach with high accuracy [11]. Deep learning models have great potential in assisting COVID-19 management efforts but require large amounts of training data. When training neural networks for image classification, images from different classes should only differ in the task specific characteristics; it is important, therefore, that all images are taken from the same machines. In this work, we designed the algorithm based on deep learning models to diagnose x-ray images of Covid-19 from Xception, vgg19, ResNet50, DenseNet121, and the Inception model. The results are highly accurate. In this paper, we used the new supervised contrastive learning loss function [14], and in this paper we are the first to use that technique in the classification of the SARS CoV 2 image.



Fig. 1. Normal and covid 19 chest X-ray images

2. DATASET

The datasets were obtained from Kaggle [12], while the code was obtained from GitHub [13]. Chest imaging is commonly used in medicine, and it plays an important role in the detection of SARSCoV2. Through the diagnosis of chest imaging, medical staff can more accurately grasp the imaging modal characteristics of SARSCoV2 cases, such as multiple small patchy shadows and interstitial changes in the early stages, which are obvious outside the lungs. It then develops into multiple ground glass and infiltration shadows in both lungs. In severe cases, lung consolidation and pleural effusion are rare. It has important guiding value for accurately judging the condition and its development, formulating treatment

plans, and evaluating prognoses. There are many SARSCoV2 datasets, but the number of samples is small. The experiment collected 2295 SARSCoV2 chest X rays on Kaggle, consisting of 1811 in a training set and 484 in a validation (test) set [12].

3. PROPOSED METHOD

Our proposed method is based on the well-established deep learning model Xception, vgg19, ResNet50, DenseNet121, Inception models. We prefer to use the DenseNet121 model (see detailed description in Table because , it extracts the features at low-level by using its smaller kernel size, which is appropriate for COVID-19 images with a lower number of layers compared to its others counterpart models. Schema of the algorithm in Figure 2. Our algorithm works on the first input x-ray image, and then it works automatically for all five models: Xception, vgg19, ResNet50, DenseNet121, and Inception. The second step is a comparison of different architectures, and the third step is hyperparameter tuning. The fourth step is a comparison of the results, and finally, the best convolution neural network (CNN) model is selected automatically.

TABLE 2. THE COMPARISON STUDY OF DENSENET121 WITH OTHER EXISTING METHODS

Methods	<i>Sensitivity</i>	<i>Specificity</i>	<i>Accuracy</i>	<i>F-measure</i>
Xception	0.9837	0.9850	0.9793	0.9840
vgg19	0.8992	0.9946	0.9215	0.9430
ResNet50	0.7692	1.0000	0.8099	0.8670
DenseNet121	0.9882	0.9972	0.9897	0.9925
Inception	0.8837	0.7850	0.8793	0.8840

Table 1 provide an elaborate comparative results analysis of the DenseNet121model with other existing techniques. On examining the predictive outcome in terms of sensitivity, the ResNet50 and Inception models have achieved minimal sensitivity values.

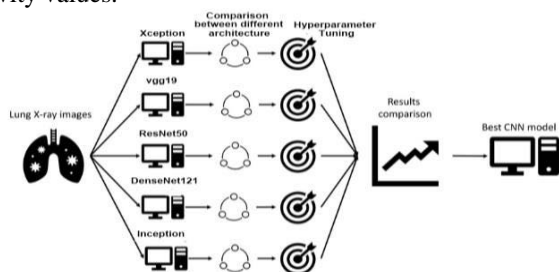


Fig. 2. Proposed method

4. CONCLUSION AND FURTHER RESEARCH

The proposed method has major limitations. The experiment only applies to X-ray images and not CT images, because X-ray images are RGB and CT images are grayscale. This experiment can only be used to classify SARSCoV2 patients and healthy people, and it cannot classify SARSCoV2 and general pneumonia. We will next focus on the classification of SARSCoV2, bacterial pneumonia, and viral pneumonia. This paper has developed a novel algorithm for SARSCoV2 detection and classification. The DenseNet121 model performs well. We evaluated our method on SARSCoV2 image datasets. The evaluation results indicate that our method is not only efficient in terms of classification accuracy but also training parameters. From this result, we can conclude that our proposed method is more appropriate for

SARSCoV2 image classification. However, the performance of our proposed method could be further improved by the following two techniques: First, our method does not utilize offline data augmentation techniques in the experiment. Thus, the use of extensive augmentation techniques such as GAN or Convolution Auto-encoder before training could improve the performance further. This also helps to increase the number of SARSCoV2 images, which results in mitigating the overfitting problem during the training step. Second, the use of other pre-trained deep learning models having a smaller filter size could improve the performance of SARSCoV2 images. This is because a smaller filter size helps extract more discriminating ROIs from SARSCoV2 images

5. ACKNOWLEDGEMENT

The work was supported by Act 211 Government of the Russian Federation, contract No. 02.A03.21.0011. The work was supported by the Ministry of Science and Higher Education of the Russian Federation (government order FENU-2020-0022).

REFERENCES

- [1] Wu, F. A new coronavirus associated with human respiratory disease in China / F. Wu, S. Zhao, B. Yu, Y.M. Chen, W. Wang, Z.G. Song, Y. Hu, Z.W. Tao, J.H. Tian, Y.Y. Pei, M.L. Yuan // Nature. – 2020. – Vol. 579(7798). – P. 265-269.
- [2] Huang, Ch. Clinical features of patients infected with 2019 novel coronavirus in Wuhan, China / Ch. Huang, Y. Wang, X. Li, L. Ren, J. Zhao, Y. Hu, L. Zhang // The lancet 2020. – Vol. 395(10223). – P. 497-506.
- [3] World Health Organization // Pneumonia of unknown cause-China. Emergencies preparedness, 2020.
- [4] Wu, Z. Characteristics of and important lessons from the coronavirus disease 2019 (COVID-19) outbreak in China / Z. Wu, J.M. McGoogan // Chinese Center for Disease Control and Prevention. – 2020.- Vol. 323(13). – P. 1239-1242.
- [5] Timeline-COVID, World Health Organization, 2020. URL: <https://www.who.int/news-room/detail/27-04-2020-who-timeline---covid-19> (2020-04-27).
- [6] Abbas, A. Learning transformations for automated classification of tuberculosis using convolutional neural network / A. Abbas, M.M. Abdelsamea // 13Th international conference on computer engineering and systems (ICCES), IEEE. – 2018. – P. 122-126.
- [7] Kuruvilla, J. Lung cancer classification using neural networks for ct images / J. Kuruvilla, K. Gunavathi // Comput Meth Prog Biomed. – 2014. – Vol. 113(1). – P. 202-209.
- [8] Manikandan, T. Lung cancer detection using fuzzy auto-seed cluster means morphological segmentation and svm classifier / T. Manikandan, N. Bharathi // J Med Syst. – 2016. – Vol. 40(7). – P. 181.
- [9] Anthimopoulos, M. Lung pattern classification for interstitial lung diseases using a deep convolutional neural network / M. Anthimopoulos, S. Christodoulidis, L. Ebner, A. Christe, S. Mougiakakou // IEEE Trans Med Imaging. – 2016. – Vol. 35(5). – P. 1207-1216.
- [10] Sun, W. Computer aided lung cancer diagnosis with deep learning algorithms / W. Sun, B. Zheng, W. Qian // Medical imaging computer-aided diagnosis. – 2016. – Vol. 9785. – P. 97850z.
- [11] LeCun, Y. Deep learning / Y. LeCun, Y. Bengio, G. Hinton // Nature. – 2015. – Vol. 521(7553). – P. 436.
- [12] Abotaleb, M. Classification Infectious disease X-ray images / M. Abotaleb, 2022. URL: <https://www.kaggle.com/mostafaabotaleb/classificationinfectiousdiseasexrayimages>.
- [13] Abotaleb, M. Classification Infectious disease X-ray images / M. Abotaleb, 2022. URL: <https://github.com/abotalebmostafa11/classification-Infectious-disease-X-ray-images>.
- [14] Khosla, P. Supervised contrastive learning / P. Khosla, P. Teterwak, C. Wang, A. Sarna, Y. Tian, P. Isola, D. Krishnan // Advances in Neural Information Processing Systems. – 2020. – Vol. 33. – P. 18661-18673.

Влияние амплитуды импульса на линейность обновления веса в мемристорах на основе оксида гафния

О.О. Пермякова

Московский физико-технический институт
Долгопрудный, Россия
Физико-технологический институт
им. К. А. Валиева РАН
Москва, Россия
o.permyakova@phystech.edu

А.Е. Рогожин

Физико-технологический институт
им. К. А. Валиева РАН
Москва, Россия
rogozhin@ftian.ru

А.В. Мьяконких

Физико-технологический институт
им. К. А. Валиева РАН
Москва, Россия
miakonkikh@ftian.ru

Е.А. Смирнова

Физико-технологический институт
им. К. А. Валиева РАН
Москва, Россия
smirnova@ftian.ru

К.В. Руденко

Физико-технологический институт
им. К. А. Валиева РАН
Москва, Россия
rudenko@ftian.ru

Аннотация—Линейность изменения проводимости мемристора влияет на точность обучения аналоговых нейроморфных систем. Обычно линейности изменения веса добиваются за счет изменения формы сигнала, но это требует включения в аналоговую цепь дополнительной периферии. В работе рассмотрено влияние амплитуды импульса сигнала на потенциацию мемристора на основе оксида гафния. Показано, что при уменьшении модуля амплитуды импульса увеличивается линейность изменения проводимости прибора. Также показано, что существует напряжение при котором проводимость изменяется наилучшим образом.

Ключевые слова— мемристор, резистивное переключение, оксид гафния, потенциация.

1. ВВЕДЕНИЕ

Одним из наиболее перспективных устройств для аппаратной реализации искусственного интеллекта является мемристор. По сути мемристор представляет собой наноразмерный электронный синапс с низким потреблением энергии (пикоджоуль на операцию) [1], сравнимым с потреблением биологического синапса. Обычно мемристор представляет собой планарную структуру металл-диэлектрик-металл. При подаче сигнала напряжения на электроды, с амплитудой выше порогового происходит постепенное (аналоговое) изменение значения сопротивления прибора.

Высокоплотный кроссбар-массив мемристоров может быть использован для нейроморфных вычислений [1, 2]. Такой массив позволяет реализовать не только аналоговые искусственные нейронные сети, но и импульсные нейронные сети.

Одним из основных факторов, влияющих на точность обучения нейроморфных систем на основе мемристоров является линейность и симметрия обновления значения сопротивления (веса) прибора [2]. Почти линейное изменение значения сопротивления мемристора возможно при использовании импульсов с увеличивающейся амплитудой [3]. Но для полностью аналоговой нейроморфной системы на основе мемристоров требуется аппаратная реализация

периферии для настройки таких сигналов в реальном времени. Такая периферия не только занимает дополнительное место на чипе, из-за чего теряется смысл формирования высокоплотного кроссбар-массива, но и потребляет дополнительную мощность.

Увеличении линейности изменения веса мемристора при использовании одинаковой последовательности импульсов не зависимо от состояния мемристора было продемонстрировано для нефиламентарных приборов [4]. Последовательно прикладывая к мемристорам импульсы разной полярности с большей амплитудой переключающих импульсов, можно уменьшить начальный резкий рост проводимости.

На данный момент нет устоявшихся требований к значению амплитуды напряжения для аналоговых нейроморфных систем. Значение амплитуды напряжения для используемых импульсов переключения зависит от структуры мемристора и влияет на мощность потребляемую нейроморфной системой во время обучения. Исследование изменения значения сопротивления мемристоров в зависимости от значения амплитуды напряжения переключающего импульса проводилось для мемристоров на основе NiO_x в диапазоне напряжений от 1,4 В до 3,0 В [5], на основе FeO_x в диапазоне напряжений от 1,3 В до 1,9 В [6] и для мемристоров с химически активным медным электродом Cu/a-Si/Pt в диапазоне напряжений от 1,4 В до 2,4 В [7].

В этой работе рассматривается зависимость линейности обновления сопротивления мемристора во время потенциации одинаковыми импульсами в зависимости от амплитуды напряжения прикладываемого сигнала.

2. ПОДГОТОВКА ОБРАЗЦОВ И МЕТОДИКА ИЗМЕРЕНИЙ

Для изучения синаптических характеристик мемристоров была изготовлена структура с общим электродом $\text{Pt/HfO}_2/\text{HfO}_x\text{N}_y/\text{TiN}$. Активный слой устройства имел толщину 6 нм и был нанесен с помощью плазма-стимулированного атомно-слоевого осаждения. Для слоев HfO_2 , HfO_xN_y использовались следующие прекурсоры ТЕМАН и пары воды, ТЕМАН и плазма

водорода, соответственно. Электроды были нанесены с помощью магнетронного распыления. Верхний электрод Pt был нанесен через теньевую маску и имел площадь $0.0360 \pm 0.0015 \text{ мм}^2$.

Исследование потенциации устройств проводилось с использованием параметрического анализатора Kethly-4200SCS. Импульсы подавались на верхний электрод со встроенной платы генератора импульсов 4205-PG2, при этом нижний электрод прибора был заземлен. Сигнал вышедший с устройства был усилен с помощью трансимпедансного усилителя после чего сигнал попадал на встроенный осциллограф 4200-SCP2HR.

До проведения измерения синаптических характеристик проводилась электроформовка устройств при постоянном напряжении -5 В и последовательно подключенном сопротивлении, со стороны заземленного электрода, 18 кОм .

3. РЕЗУЛЬТАТЫ И ОБСУЖДЕНИЕ

Для исследования линейности обновления значения состояния сопротивления были проведены измерения потенциации структур Pt/HfO₂/HfO_xN_y/TiN. Во время потенциации на верхний электрод подавалась последовательность из ста прямоугольных импульсов амплитуды которых изменялась в диапазоне от $-0,7 \text{ В}$ до $-1,5 \text{ В}$ (Рис. 1), а расстояние между импульсами изменялось в диапазоне от $0,2 \text{ мс}$ до 16 мс (Рис. 2). Длина импульса составляла 5 мкс . Сопротивление изменялось в диапазоне от $0,04 \text{ мСм}$ до $0,3 \text{ мСм}$. Отношение сопротивлений составляло $7,5$.

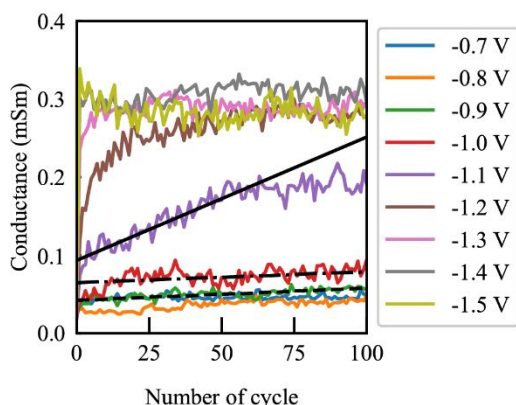


Рис. 1. Потенциация структуры Pt/HfO₂/HfO_xN_y/TiN в течении ста импульсов с расстоянием между импульсами $1,6 \text{ мс}$, в зависимости от амплитуды напряжения сигнала в диапазоне от $-0,7 \text{ В}$ до $-1,5 \text{ В}$

На рис. 1 видно, что при увеличении амплитуды импульса возрастает нелинейность изменения веса в зависимости от числа циклов. При амплитуде сигнала меньше $1,1 \text{ В}$ изменение проводимости было линейным, но при этом незначительным. А значит для изменения веса такого синапса необходимо затратить значительно большее количество импульсов, чем для сигнала с большей амплитудой, что напрямую скажется на скорости работы системы. При амплитуде сигнала больше $1,1 \text{ В}$ изменение проводимости переставало быть линейным. При амплитуде импульса равной $1,1 \text{ В}$

сопротивление возрастало практически линейно в течении первых 65 импульсах. В течении этих импульсов проводимость изменилась на 1 См . После чего скорость изменения сопротивления резко падала. Такое поведение прибора сохранялось при различных расстояниях между импульсами (Рис. 2). Эти данные показывают, что существует выделенное напряжение мемристора, которое является предпочтительным для проведения потенциации устройства.

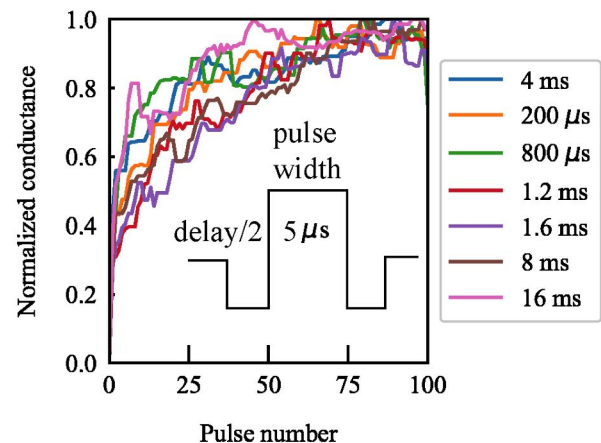


Рис. 2. Потенциация структуры Pt/HfO₂/HfO_xN_y/TiN в течении ста импульсов с амплитудой сигнала $-1,1$ в зависимости от расстояния между импульсами в диапазоне от $0,2 \text{ мс}$ до 16 мс . Вставка: форма импульсов

БЛАГОДАРНОСТИ

Работа выполнена в рамках Государственного задания по теме FFNN-2022-0019 и, частично, в рамках гранта РФФИ проекта № 18-29-27029.

ЛИТЕРАТУРА

- [1] Liu, D. Low-Power Computing with Neuromorphic Engineering / D. Liu, H. Yu, Y. Chai // Adv. Intell. Syst. – 2021. – Vol. 3(2). – P. 2000150. DOI: 10.1002/aisy.202000150.
- [2] Zhang, Y. Brain-inspired computing with memristors: Challenges in devices, circuits, and systems / Y. Zhang, Z. Wang, J. Zhu, Y. Yang, M. Rao, W. Song, Y. Zhuo, X. Zhang, M. Cui, L. Shen, R. Huang // Appl. Phys. Rev. – 2020. – Vol. 7(1). – P. 011308. DOI: 10.1063/1.5124027
- [3] Liu, H. Optimization of non-linear conductance modulation based on metal oxide memristors / H. Liu, M. Wei, Y. Chen // Nanotechnol. Rev. – 2018. – Vol. 7(5). – P. 443-468. DOI: 10.1515/ntrv-2018-0045.
- [4] Wang, I-T. 3D Ta/TaOx/TiO2/Ti synaptic array and linearity tuning of weight update for hardware neural network applications / I-T. Wang, C.-C. Chang, L.-W. Chiu, T. Chou, T.-H. Hou. // Nanotechnol. – 2016. – Vol. 27. – P. 365204. DOI: 10.1088/0957-4484/27/36/365204.
- [5] Hu, S.G. Synaptic long-term potentiation realized in Pavlov's dog model based on a NiOx-based memristor / S.G. Hu, Y. Liu, Z. Liu, T. P. Chen, Q. Yu, L.J. Deng, Y. Yin, S. Hosaka // J. Appl. Phys. – 2014. – Vol. 116. – P. 214502. DOI: 10.1063/1.4902515.
- [6] Wang, C. Investigation and Manipulation of Different Analog Behaviors of Memristor as Electronic Synapse for Neuromorphic Applications / C. Wang, W. He, Y. Tong, R. Zhao // Sci. Rep. – 2016. – Vol. 6. – P. 22970. DOI: 10.1038/srep22970.
- [7] Zhang, X. Emulating Short-Term and Long-Term Plasticity of Bio-Synapse Based on Cu/a-Si/Pt Memristor / X. Zhang, S. Liu, X. Zhao, F. Wu, Q. Wu, W. Wang, R. Cao, Y. Fang, H. Lv, S. Long, Q. Liu, M. Liu // IEEE Electron Device Lett. – 2017. – Vol. 38(9). – P. 1208-1211. DOI: 10.1109/LED.2017.2722.

Расчет ошибок коррекции освещения с помощью многомерной гистограммы на основе оригинального стенда

А.В. Погадаев

Институт проблем передачи
информации им. А. А. Харкевича
Москва, Россия
Московский физико-технический
институт (национальный
исследовательский университет)
Москва, Россия
pogadaev.av@phystech.edu

В.А. Васильев

Институт проблем передачи
информации им. А. А. Харкевича
Москва, Россия
Московский физико-технический
институт (национальный
исследовательский университет)
Москва, Россия
vasilev.va@phystech.edu

Е.И. Ершов

Институт проблем передачи
информации им. А. А. Харкевича
Москва, Россия
Московский физико-технический
институт (национальный
исследовательский университет)
Москва, Россия
ershov@iitp.ru

Аннотация—Неравномерность распределения значений на плоскости цветности является причиной малой информативности расчета ошибок алгоритмов коррекции. В работе предлагается новый подход к оценке алгоритмов коррекции освещения, а именно — хроматической адаптации фон Криса и коррекции на основе спектрозональной модели (СЗМ). Наш подход основан на предположении о равномерности распределения значений цветности, но учитывает для этого дополнительные характеристики. Мы вводим понятия дистанции ошибки, дистанции коррекции и направления коррекции и строим взвешенную гистограмму ошибок работы алгоритмов с учетом распределения этих трех величин. Оценка работы алгоритмов производится на основе оригинального стенда — набора данных и программного обеспечения, принимающего на вход результаты работы алгоритмов. Набор данных состоит из 3 100 230 произведений 38 спектров освещения и 4410 спектров окрасок. Для стандартного наблюдателя CIE 1964 при неравномерной плотности тестовых образцов мы получаем ранжирование алгоритмов коррекции, зависящее от используемого метода взвешивания.

Ключевые слова— коррекция освещения, цветокоррекция, сравнение алгоритмов, стенд.

1. ВВЕДЕНИЕ

Традиционно, основным критерием репрезентативности наборов спектров выступает количество образцов [1], однако этого недостаточно для оценки точности алгоритма коррекции освещения. В нашей работе мы производим эту оценку в предположении, что все возможные цветности в наборе данных распределены равномерно. Это позволяет избежать занижения ошибки в нерепрезентативных случаях, например, для точек в наборе данных, которые находятся на окраине плоскости цветности. Решение, которое мы предлагаем, основано на взвешивании ошибок, полученных в результате работы алгоритмов коррекции.

2. ОСНОВНЫЕ ТЕРМИНЫ И ОБОЗНАЧЕНИЯ

Пусть Λ — видимый диапазон волн, по которому ведется интегрирование, $\Phi(\lambda)$ — спектр отражательной способности (спектр окраски), $\bar{\Psi}(\lambda)$ — функции спектральной чувствительности, а $S_s(\lambda)$ — спектр источника. Тогда эталонным тристимулом назовем:

$$\bar{c}_s = \int_{\Lambda} \Phi(\lambda) S_s(\lambda) \bar{\Psi}(\lambda) d\lambda. \quad (1)$$

Возьмем источник $S_t(\lambda) \neq S_s(\lambda)$. Тогда искомым тристимулом назовем:

$$\bar{c}_t = \int_{\Lambda} \Phi(\lambda) S_t(\lambda) \bar{\Psi}(\lambda) d\lambda. \quad (2)$$

Дистанцией коррекции назовем L_2 - норму между откликами \bar{c}_t и \bar{c}_s . Дистанцией ошибки назовем L_2 - норму между результатом работы алгоритма коррекции и искомым откликом в координатах UV:

$$\Delta E^{uv} = \sqrt{(\Delta u)^2 + (\Delta v)^2}. \quad (3)$$

Мы осуществляем разбиение плоскости UV на квадратные секторы со стороной 0,011. Данный выбор обусловлен минимальным значением JND в данной плоскости [2]. Веса, учитывающие распределение образцов на плоскости цветности, для точки i равны:

$$\omega_i^{density} = \frac{n_{max}^p}{n_i^p}; \omega_i^{distance} = \frac{n_{max}^d}{n_i^d}; \omega_i^{angle} = \frac{n_{max}^\phi}{n_i^\phi}, \quad (4)$$

где n_{max}^p — максимальное количество образцов среди всех секторов, n_i^p — количество образцов в i -м секторе, к которому принадлежит точка, n_{max}^d — максимальное значение гистограммы дистанции коррекции, n_i^d — значение гистограммы, соответствующее i -ой дистанции. Для ω_i^{angle} аналогично. Таким образом, комбинированный вес равен:

$$\omega_i = \omega_i^{density} \cdot \omega_i^{distance} \cdot \omega_i^{angle}. \quad (5)$$

3. ОПИСАНИЕ СТЕНДА

В основу стенда были положены данные спектров отражения из наборов USGS [3] и SFU [4], а также 38 спектров различных источников из стандарта CIE [5]. Трестимулу, заданному определенными источником и спектром окраски, ставились в соответствие трестимулы полученные при том же спектре окраски, но под другими источниками освещения. Все получившиеся трестимулы показаны на Рис. 1.

4. ТЕСТИРОВАНИЕ И РАНЖИРОВАНИЕ АЛГОРИТМОВ

Для тестирования были отобраны следующие алгоритмы: алгоритм фон Криса в XYZ [6], алгоритм фон Криса в LMS [7] и коррекция на основе спектральной модели [8]. Результаты показаны в таблице I.

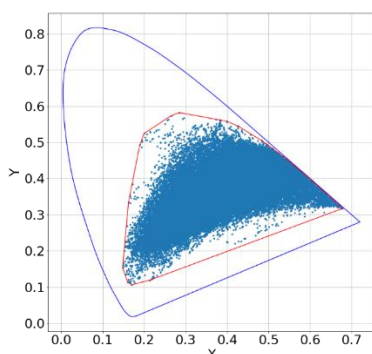


Рис. 1. Синие точки — значения цветности в стенде, красная линия — выпуклая оболочка значений цветности, синяя линия — цветовой треугольник стандартного наблюдателя CIE XYZ 1964

ТАБЛИЦА I. Значения L_2 ошибки алгоритмов с помощью предложенных методов взвешивания

	Коррекция фон Криса в XYZ	Коррекция фон Криса в LMS	Коррекция на основе СЗМ
Среднее без весов	0,0064	0,0068	0,0067
Средневзвешенное с учетом $\omega^{density}$	0,0217	0,0195	0,025
Средневзвешенное с учетом $\omega^{density}$ и $\omega^{distance}$	0,1317	0,0301	0,1646
Средневзвешенное с тремя весами	0,1449	0,0253	0,1821

На Рис. 2 показаны распределения ошибок для коррекции фон Криса в XYZ с разными весами.

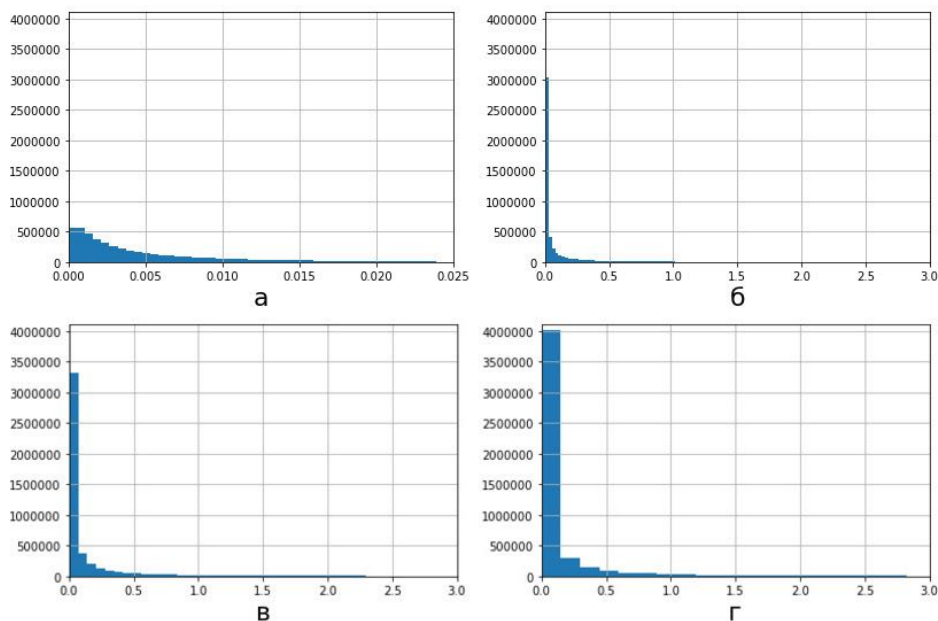


Рис. 2. Распределение ошибок в зависимости от используемых весов. а — без весов, б — с учетом плотности, в — с учетом плотности и дистанции коррекции, г — с учетом всех весов

5. ЗАКЛЮЧЕНИЕ

В работе был предложен стенд для коррекции освещения и осуществлено ранжирование трех алгоритмов по значениям средневзвешенной ошибки: хроматической адаптации фон Криса в цветовых пространствах XYZ и LMS и коррекции на основе спектральной модели.

В результате было продемонстрировано, что ранжирование алгоритмов меняется в зависимости от используемых весов.

Кроме того, удалось достичь равномерного учета ошибки алгоритмов коррекции с помощью весов на основе введенных характеристик.

ЛИТЕРАТУРА

- [1] Li, C. Comparison of real colour gamuts using a new reflectance database / C. Li, M.R. Luo, M.R. Pointer, P. Green // Color Research Application. – 2013. – Vol. 39(5) – P. 442-451. DOI: 10.1002/col.21827.
- [2] MacAdam, D.L. Chromatic Adaptation / D.L. MacAdam // J. Opt. Soc. Am. – 1956. – Vol. 46(7). – P. 500-513.
- [3] Kokaly, R.F. USGS Spectral Library Version 7 / R.F. Kokaly, R.N. Clark, G.A. Swayze // U. S. Geological Survey Data Series. – 2017. – 61. DOI: 10.3133/ds1035.
- [4] Simon Fraser University [Electronic resource]. — Access mode: <https://www2.cs.sfu.ca> (07.02.2022).
CIE 15: Technical Report [Electronic resource]. — Access mode: <https://law.resource.org/pub/us/cfr/ibr/003/cie.15.2004.tables.xls> (07.02.2022).
- [5] Kang, H.R. Computational color technology / H.R. Kang // SPIE. – 2006.
- [6] Kries, A.J. Influence of adaption on the effects produced by luminous stimuli / A.J. Kries // Vieweg, Braunschweig. – 1905. – Vol. 3. – P. 109-282.
- [7] Shepelev, D. Simulation Of Underwater Color Images Using Banded Spectral Model / D. Shepelev, V. Bozhkova, E. Ershov, D. Nikolaev // 34th International ECMS Conference on Modelling and Simulation. – 2020. – Vol. 6. – P. 11-18.

Algorithms for proactive security of industrial systems based on machine learning technologies

V. Vasilyev
Dept. of Computer Engineering and
Information Security
Ufa State Aviation Technical University
Ufa, Russia
vasilyev@ugatu.ac.ru

A. Vulfin
Dept. of Computer Engineering and
Information Security
Ufa State Aviation Technical University
Ufa, Russia
vulfin.alexey@gmail.com

A. Kirillova
Dept. of Computer Engineering and
Information Security
Ufa State Aviation Technical University
Ufa, Russia
kirillova.andm@gmail.com

Abstract—Approaches to improving the efficiency of network attack detection algorithms in heterogeneous industrial networks based on machine learning technologies are considered. An algorithm for analyzing and processing network traffic has been developed in the task of detecting malicious network activity. The Electra dataset is used to train the proposed machine learning models and heterogeneous neural network models.

Keywords—network attacks, machine learning, heterogeneous industrial network, dataset.

1. INTRODUCTION

At the present stage of digital transformation of the industry [1], there is a trend towards the integration of Industrial Internet of Things (IIoT) devices with traditional data collection and control systems and deep penetration of IIoT into critical infrastructure, which has led to an increase in the likelihood and number of potential cyber attacks on industrial facilities.

To detect multi-step network attacks on industrial systems, it is necessary to analyze a significant amount of incoming, outgoing and local network traffic with the ability to compare it with the information security event stream to detect anomalous activity [2].

Attacks using exploits practically do not change the main characteristics of industrial protocol traffic, which makes it very difficult to select signatures for their detection [3]. The use of machine learning (ML) methods makes it possible to identify the features of anomalous traffic and build an appropriate mechanism for their detection [4, 5].

The aim of the work is to improve the efficiency of algorithms for detecting network attacks in a heterogeneous industrial network based on ML technologies.

2. ALGORITHM FOR DETECTING NETWORK ATTACKS IN A HETEROGENEOUS INDUSTRIAL NETWORK

The algorithm for analyzing network traffic parameters in the task of detecting anomalies and malicious network activity based on the use of ML-models is shown in Fig. 1. The main stages of data collection and processing for the construction and use of ML-models are presented.

In order to evaluate the effectiveness of the proposed solution, the Electra dataset [6] was used, generated from the network traffic of a traction electrical substation operating in normal mode and under attack conditions.

To accomplish its task, the electric traction substation consists of 5 PLCs, a SCADA system, a switch and a firewall.

All communications between components are implemented by Modbus and S7comm over TCP/IP.

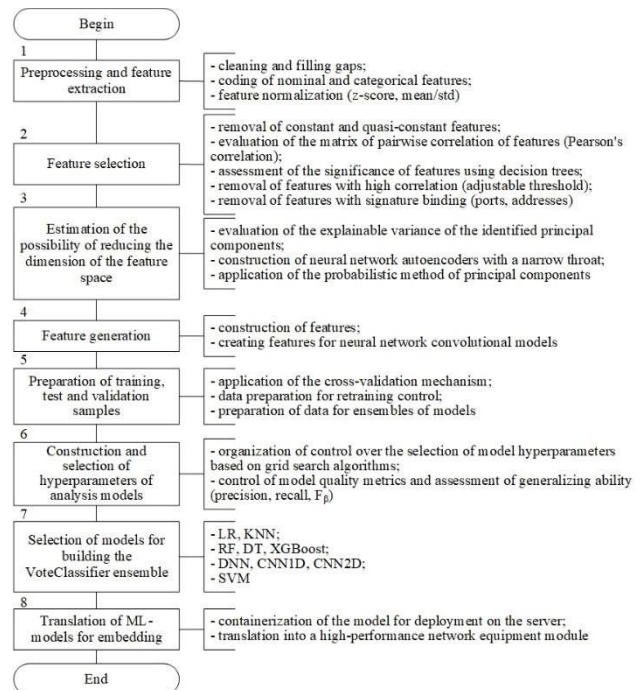


Fig. 1. Generalized network traffic mining algorithm

Electra includes three types of attacks: false data injection, replay, and reconnaissance. Electra is the only dataset available that includes replay attacks. There are two different data sets, one for each Modbus and S7comm communication protocol. The structure of the test bench is shown in Fig. 2.

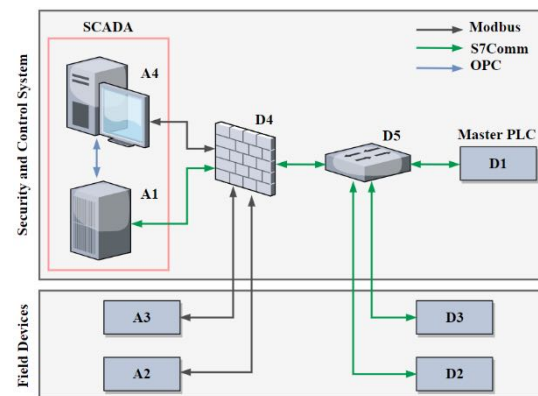


Fig. 2. The structure of the stand for collecting network traffic (A1 (Nanobox) and A4 (HMI) form a SCADA master for Modbus PLC slaves A2 and A3. D1 (PLC) slave for S7Comm A1 slaves and PLCs D2 and D3. D4 – firewall, D5 – switch for connecting devices)

Electra Modbus and Electra S7comm collected network traffic for 12 hours of operation of the stand (Fig. 3), while 94% and 98% of the records correspond to normal operation. The data set contains 387 million records for S7Comm (36.8 GB) and 16 million records for Modbus (1.5 GB). Table 1 shows features of network sessions recorded during traffic collection.

TABLE 1. FEATURES OF NETWORK SESSIONS

Features	Description	Data type
time	Timestamp	string
smac	Source MAC address	string
dmac	Destination MAC address	string
sip	Source IP	string
dip	Destination IP address	string
request	Indicates if the packet is a request (master-to-slave packet)	boolean
fc	Function code	integer
error	Indicates whether there was an error in the read/write operation.	boolean
madd	Memory address to perform a read/write operation	integer
data	Data field	integer
label	Label for attacks and normal patterns	string

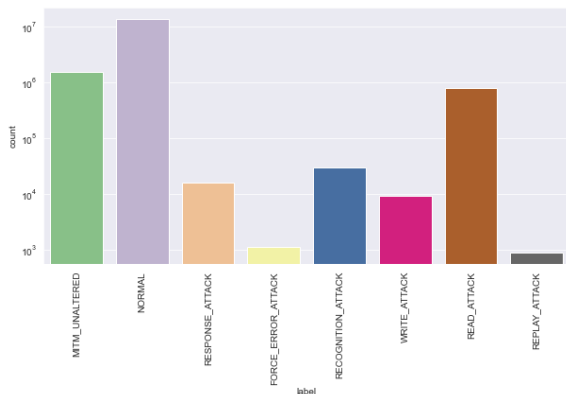


Fig. 3. Distribution of the number of records corresponding to attacks and normal operation

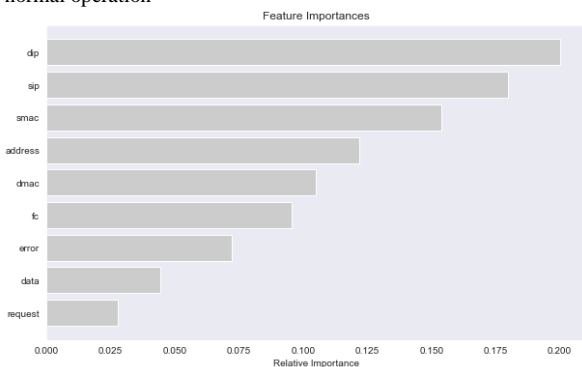


Fig. 4. The histogram of the assessment of the significance of features (y-axis – features), obtained using a classifier based on a committee of decision trees (abscissa – relative units)

3. RESULTS OF THE COMPUTATIONAL EXPERIMENT

Evaluation of the significance of features in the problem of multiclass classification makes it possible to single out those

features that are key to assigning a network session to a particular class. Fig. 4 shows a histogram of the feature significance assessment obtained using a classifier based on the random forest classifier.

An ensemble of classifiers has been implemented, which includes a committee of decision trees (RF), a classifier based on the gradient boosting algorithm on an ensemble of decision trees (XGBClassifier), and an ExtraTreesClassifier. Committee parameters: voting type – “soft” (voting and weighting model predictions for each class); model weights are distributed as {2, 1, 3}. The accuracy metric for the ensemble for the test sample is 0.975, the estimate of the F1- score is 0.964. The assessment of prec recall, F1-score and the number of examples (support) by attack classes in the test sample are shown in Table 2.

Таблица II. EVALUATION OF THE QUALITY OF MULTICLASS CLASSIFICATION FOR A TEST SAMPLE OF AN ENSEMBLE OF CLASSIFIERS

attack classes	precision	recall	F1-score	support
FORCE_ERROR_ATTACK	1.00	1.00	1.00	1043
MITM_UNALTERED	1.00	1.00	1.00	3488
NORMAL	1.00	1.00	1.00	7354
READ_ATTACK	0.84	1.00	0.91	3486
RECOGNITION_ATTACK	1.00	1.00	1.00	3528
REPLAY_ATTACK	1.00	0.03	0.06	686
RESPONSE_ATTACK	1.00	1.00	1.00	3546
WRITE_ATTACK	1.00	1.00	1.00	3469

4. CONCLUSION

Algorithms for intelligent analysis of scattering parameters in the tasks of detecting malicious activity have been developed. An ensemble of classifiers is built based on traditional ML-models. The estimate of the F1- score when working with test samples reaches 96.4%.

ACKNOWLEDGMENT

The reported study was funded by RFBR according to the research project No. 20-38-90078.

REFERENCES

- Threat landscape for industrial automation systems. 2019 year. Kaspersky ICS CERT [Electronic source]. – Available: <https://ics-cert.kaspersky.ru/reports/2020/04/24/threat-landscape-for-industrial-automation-systems-2019-report-at-a-glance/> (31.01.2022).
- Ten, C.W. Cybersecurity for critical infrastructures: Attack and defense modeling / C.W. Ten, G. Manimaran, C.C. Liu // IEEE Transactions on Systems, Man, and Cybernetics-Part A: Systems and Humans. – 2010. – Vol. 40(4). – P. 853-865.
- Cecil, A.A Summary of Network Traffic Monitoring and Analysis Techniques [Electronic source]. – Available: https://www.cse.wustl.edu/~jain/cse567-06/ftp/net_monitoring/index.html (31.01.2022).
- Ten, C.W. Anomaly detection for cybersecurity of the substations / C.W. Ten, J. Hong, C.C. Liu // IEEE Transactions on Smart Grid. – 2011. – Vol. 2(4). – P. 865-873.
- Vulfin, A.M. Network traffic analysis based on machine learning methods / A.M. Vulfin, V.I. Vasilyev, V.E. Gvozdev, K.V. Mironov, O.E. Churkin // Journal of Physics: Conference Series. IOP Publishing. – 2021. – Vol. 2001(1). – P. 012017.
- Gomez, A.L.P. On the generation of anomaly detection datasets in industrial control systems / A.L.P. Gomez, L.F. Maimo, A.H. Celdran, F.J.G. Clemente, C.C. Sarmiento., C.J.D.C. Masa, R.M. Nistal // IEEE Access. – 2019. – Vol. 7. – P. 177460-177473.

Стекинг-подход к задаче прогнозирования банкротств

Д.В. Полупанов
Башкирский Государственный
Университет
Уфа, Россия
demetrious@mail.ru

С.Р. Абдюшева
Башкирский Государственный
Университет
Уфа, Россия
asr_light@mail.ru

В.В. Галиямов
Башкирский Государственный
Университет
Уфа, Россия
tgc38m@gmail.com

Аннотация—В данной работе совершенствуются и расширяются стекинг-методики предсказания банкротства. Основной темой данного исследования является использование моделей разного типа: многослойных перцептронов, решающих деревьев, сетей Кохонена. В различных сочетаниях они отбираются и комбинируются в ансамбли и мета-ансамбли таким образом, чтобы учитывать сильные и слабые стороны каждого типа базовых моделей в контексте задачи.

Ключевые слова— банкротство корпораций, классификация, ансамбли, мета-ансамбли, многослойные перцептроны, решающие деревья, сети Кохонена, стекинг.

1. ВВЕДЕНИЕ

Настоящая работа продолжает исследования в области прогнозирования банкротств при помощи моделей машинного обучения [1]. Актуальность темы исследования с практической точки зрения определяется необходимостью своевременного выявления неблагоприятных тенденций в деятельности предприятий. Моделирование позволяет избежать высоких потерь и затрат вследствие возможной неплатёжеспособности заемщиков. Кроме того, своевременная оценка финансового состояния предприятия, позволит применить регулирующие воздействия, чтобы избежать негативных экономических и социальных последствий. С точки зрения инструментария актуальность обусловлена возрастающей ролью ансамблевых моделей в решении задач машинного обучения, становящихся всё более актуальными с ростом вычислительных мощностей и возможностью распараллеливания. Последние разработки в этой области связаны с нелинейным объединением прогнозов, которые ныне вызывают незатухающий интерес исследователей. Цель данной работы состоит в исследовании методик мета-ансамблирования и стекинга применительно к задаче распознавания банкротства.

2. ПОСТАНОВКА ЗАДАЧИ

Задача прогнозирования банкротств сводится к построению математической модели, способной на основе совокупности финансово-экономических показателей предприятия с высокой точностью предсказывать метки класса («банкрот» / «не банкрот»). Модель строится на основе ретроспективных данных примерно однородной выборки предприятий.

Традиционно, модели классификации могут быть представлены в виде:

$$\hat{y} = F(x, W), x \in X \subset \mathbb{R}^n, y \in Y \subset \mathbb{R} \hat{y} \in \hat{Y} \subset C[\mathbb{R}] \quad (1)$$

где конкретная численная реализация вектора входных факторов ; – конкретная числовая реализация наблюдаемой выходной случайной величины Y ; – объективный эталон – расчетное значение Y ; $\{W\}$ – множество весов, соответствующих базовым моделям; – множество значений вектора входных факторов, – множество декларированных значений выходной величины, – множество расчетных по (1) значений выходной величины, F - оператор ансамблевого отображения.

В нашем случае необходимо решить обратную задачу восстановления скрытой в данных зависимости вероятности банкротства P от вектора экзогенных переменных , которая оценивается с помощью логистической функции [2]:

$$P(t) = 1/[1 + \exp(-\hat{y}(x(t), t))], P \in [0,1] \quad (2)$$

где t – время; показатель экспоненты , играющий роль аргумента в (2), восстанавливается с помощью отображения (1).

3. ИСПОЛЬЗОВАНИЕ АНСАМБЛЕВЫХ И МЕТА-АНСАМБЛЕВЫХ МЕТОДИК

Основная идея ансамбля заключается в том, что модели, имеющие различия, по-разному обрабатывают входные данные и совершают разные ошибки. Объединив прогнозы нескольких таких моделей, мы можем получить зачастую более высокое, чем у лучшей отдельной модели, качество прогнозов [3]. Искусство ансамблирования в этом контексте означает умение создать различия в моделях.

Рассмотрим многоуровневую структуру ансамбля, о котором идёт речь в данной работе. Используются следующие базовые модели: многослойный перцептрон (MLP), решающее дерево (DT) и сеть Кохонена (KCN). На первом уровне ансамбля, который оперирует базовыми моделями, различия достигаются вариацией бэггинга и его разновидностей, кросс-валидации, бустинга (например, XGBoost). На втором уровне ансамблей различия обуславливаются природой базовых моделей. Для каждого типа моделей составляется суперансамбль. Наконец, прогнозы суперансамблей объединяются мета-ансамблем, называемым также ансамблем стекинга.

4. ЗАКЛЮЧЕНИЕ

Предложенные методики прошли проверку на реальных ретроспективных данных предприятий,

размещённых на соревновательной платформе Kaggle. Модель была реализована на языке Python с использованием библиотек машинного обучения Scikit-Learn, Tensor Flow, Keras. Чтобы усилить базовые модели перед непосредственным процессом обучения, производится стандартизация входных данных и оптимизация гиперпараметров при помощи библиотеки Keras Tuner. По итогам вычислительного эксперимента можно сделать вывод, что по большинству метрик многослойный перцептрон лучше других моделей, но ничем не выделяется, дерево решений даёт стабильные, но средние прогнозы. Сеть Кохонена по основным метрикам справляется плохо, зато имеет лучшие показатели точности по не-банкротам (если классификатор назвал фирму не-банкротом, то скорее всего она действительно не-банкрот), и полноты по банкротам (находит высокую долю предприятий-банкротов). На основании этих метрик можно судить, что необходимое разнообразие моделей было достигнуто. Наконец, показатели стекинговой модели почти по всем метрикам превосходят лучшие из показателей

суперансамблей и отдельных моделей, что доказывает возможность успешного применения мета-ансамблирования для решения задач прогнозирования банкротств.

ЛИТЕРАТУРА

- [1] Полупанов, Д.В. Применение нейросетевых ансамблей в оценке банкротства корпораций / Д.В. Полупанов, С.Р. Абдюшева, В.В. Галлямов, С.А. Мустафина. – Вестник технологического университета. – 2021. – Т. 24, № 1. – С. 98-104.
- [2] Гобатков, С.А. Нейросетевые математические модели оценки вероятности банкротства корпораций / С.А. Гобатков, Д.В. Полупанов // Стерлитамак: Стерлитамакский филиал БашГУ. – 2020. – С. 57-74.
- [3] Polupanov, D.V. Improving the neural network mathematical model of corporate bankruptcy / D.V. Polupanov, S.R. Abdiusheva, V.V. Gallyamov // International Conference on Information Technology and Nanotechnology. – 2021. – P. 1-9. DOI: 10.1109/ITNT52450.2021.9649280.
- [4] Goodfellow, I. Deep Learning / I. Goodfellow, Y. Bengio, A. Courville // Adaptive Computation and Machine Learning series. – The MIT Press, 2016.

Метод защиты моделей глубокого обучения цифровыми водяными знаками

Ю.Д. Выборнова

Самарский национальный исследовательский университет
им. академика С.П. Королева
Самара, Россия
vybornovamail@gmail.com

Д.И. Ульянов

Самарский национальный исследовательский университет
им. академика С.П. Королева
Самара, Россия
dmitryulyanovhome@gmail.com

Аннотация—В работе предлагается новый метод защиты авторских прав на глубокие нейронные сети. Основная идея состоит во встраивании цифрового водяного знака в нейросетевую модель путем ее дообучения на уникальном наборе изображений-триггеров. Триггерная выборка формируется путем синтеза псевдоголографических изображений и их встраивания в растровые изображения оригинального датасета, используемого при обучении исходной модели. Результаты экспериментальных исследований подтверждают работоспособность предложенного метода, при этом процесс встраивания цифровых водяных знаков не влияет на точность защищаемой модели.

Ключевые слова— CNN, защита авторских прав, псевдоголографические изображения, цифровые водяные знаки.

1. ВВЕДЕНИЕ

Учитывая растущий интерес к применению искусственного интеллекта, становятся все более актуальными вопросы безопасности при хранении и передаче моделей глубокого обучения. Злоумышленники могут распространять проприетарные модели или незаконно использовать их для предоставления услуг анализа данных. Следовательно, возникает необходимость создания новых методов защиты авторских прав на данный вид интеллектуальной собственности, позволяющих доказать правообладателю факт несанкционированного копирования и распространения предобученных моделей.

Методы встраивания цифровых водяных знаков (ЦВЗ) широко использовались в последние два десятилетия как средство защиты авторских прав на мультимедийные данные (изображения, видео и аудио). Общая идея встраивания ЦВЗ во внедрения в данные защитной информации, незаметной для пользователя, но предоставляющей техническую возможность отследить и доказать факт нарушения авторских прав.

2. ПРЕДЛАГАЕМЫЙ ПОДХОД

Методы встраивания ЦВЗ по принципу Black-Box [1-3] заключаются в дообучении исходной модели на так называемой триггерной выборке таким образом, чтобы модель с ЦВЗ классифицировала изображения-триггеры согласно заданным меткам и чтобы при этом результаты были отличны от прогноза исходной модели. Например, можно дообучить защищаемую модель давать намеренно неверный результат на определенных изображениях (например, классифицировать заданные изображения птиц как самолеты). Процесс проверки авторских прав на модель глубокого обучения заключается в статистической оценке результатов классификации элементов триггерной выборки.

Стоит отметить, что внесение архитектурных изменений в модель (например, добавление в выходной слой нейрона дополнительных нейронов) крайне нежелательно, поскольку в случае предоставления доступа к модели сугубо в качестве сервиса, соответствующие выходы могут блокироваться на уровне программного интерфейса.

Предлагаемый метод защиты авторских прав на глубокие нейронные сети заключается в дообучении модели на уникальном наборе изображений-триггеров, сформированном путем синтеза псевдоголографических изображений (псевдоголограмм). Псевдоголограмма представляет собой двумерный шумоподобный сигнал, который кодирует определенную битовую последовательность длины l .

Для формирования триггерной выборки сгенерированные псевдоголограммы накладываются на растровые изображения исходного набора данных с помощью аддитивной стратегии встраивания ЦВЗ. Предлагается генерировать псевдоголограммы, кодирующие последовательности $S_1, S_2, \dots, S_b, \dots, S_k$, где k – количество классов. Метка класса изображения-триггера определяется не исходной меткой изображения, на которое наложена псевдоголограмма, а индексом i последовательности, которую эта псевдоголограмма кодирует.

3. ФОРМИРОВАНИЕ ТРИГГЕРНОЙ ВЫБОРКИ

Псевдоголографический сигнал формируется на основе синтеза комплексного спектра путем размещения импульсов на двумерной плоскости в спектральной области в зависимости от битов двоичной последовательности S_i . Таким образом, если выполнить обратное дискретное преобразование Фурье, то каждый бит последовательности будет «голографически» отображаться на результирующее изображение.

Правила расположения импульсов могут быть заданы по-разному, но следует отметить, что поскольку псевдоголограмма представляет собой вещественный сигнал, полуплоскости спектра должны быть симметричны. В данной работе импульсы располагаются на двух кольцах радиусов r и $r + \Delta r$, как описано в [4]. Значения радиусов задаются по следующему правилу: $r = 0,36 \times l$, $\Delta r = 6$.

Изображения-триггеры формируются согласно следующему алгоритму: пусть дана псевдоголограмма W размера $N \times N$. Изображение оригинального датасета преобразуется в цветовое пространство YCbCr и приводится к размеру псевдоголограммы $N \times N$. Тогда процесс встраивания сводится к аддитивному наложению

псевдоголограммы на Y-компоненту изображения с последующей нормализацией диапазона яркости:

$$\hat{Y}_{N \times N} = \|Y_{N \times N} + W \times q\|,$$

где q -коэффициент видимости ЦВЗ.

4. НАБОР ДАННЫХ

Для экспериментов был выбран датасет “Cats and Dogs” [5], представляющий собой 24998 изображений кошек и собак. С учётом разделения набора на тренировочную, тестовую и валидационную выборки, каждая из них была разделена в пропорции 9:1 (за исключением тестовой выборки) на два датасета. Большой датасет, состоящий из 15186 тренировочных, 5061 валидационных и 2250 тестовых изображений, использовался для подготовки моделей, в которые впоследствии встраивался ЦВЗ.

Перед процедурой встраивания ЦВЗ было сгенерировано 200 псевдоголограмм на основе двух последовательностей S_1 или S_2 длины $l = 50$ (по 100 для каждого класса). Встраивание ЦВЗ в модель осуществлялось на изображениях меньшего датасета. При этом оригинальным изображениям без наложенных псевдоголограмм соответствовали исходные метки «кошка»/«собака», а тем же изображениям, но с нанесенной псевдоголограммой, метки были назначены в зависимости от того, какая из последовательностей, S_1 или S_2 , закодирована в этой псевдоголограмме. Таким образом, размер датасета составил: тренировочная (1688x2), валидационная (563x2) и верификационная (200). Стоит отметить, что верификационная выборка представляет собой исходные псевдоголограммы (не наложенные на другие изображения) и применяется для анализа доли верно классифицированных ЦВЗ.

Примеры исходных изображений и изображений-триггеров представлены на рисунках 1 и 2 соответственно.



Рис. 1. Оригинальные изображения из датасета «Cats And Dogs»

5. ЭКСПЕРИМЕНТАЛЬНЫЕ ИССЛЕДОВАНИЯ

Для демонстрации возможности встраивания ЦВЗ в нейросети-классификаторы согласно описанному методу, были взяты 3 модели: alexnet, vgg11 и resnet18. Каждая предварительно была обучена на тренировочной части на протяжении 30 эпох.

Параметры обучения: скорость обучения: $lr = 10^{-3}$; коэффициент видимости ЦВЗ $q = 1$; функция ошибки: кросс-энтропия; метод обновления весов: стохастический

градиентный спуск с $momentum = 0,9$; размер батча: 64; число эпох: 30. Полученные результаты представлены в таблице 1. Согласно данным таблицы исходная модель не способна различать псевдоголограммы из верификационного набора, но после процедуры встраивания ЦВЗ в модель точность распознавания псевдоголограмм близка к единице.



Рис. 2. Примеры изображений с ЦВЗ при $q = 1$ и $l = 50$

Таблица I. Точность решения исходной задачи классификации и точность извлечения ЦВЗ

Набор данных	Alexnet	VGG11	Resnet18
Исходная модель, test set	0,968	0,987	0,988
Исходная модель, verification set	0,500	0,545	0,495
Модель с ЦВЗ, test set	0,968	0,988	0,983
Модель с ЦВЗ, verification set	0,995	1,0	0,980

6. ЗАКЛЮЧЕНИЕ

Результаты экспериментальных исследований подтверждают работоспособность предложенного метода, при этом процесс встраивания цифровых водяных знаков не влияет на точность модели.

БЛАГОДАРНОСТИ

Исследование выполнено за счет гранта Российского научного фонда № 21-71-00106, <https://rscf.ru/project/21-71-00106/>.

ЛИТЕРАТУРА

- [1] Rouhani, B.D. Deepsigns: A generic watermarking framework for ip protection of deep learning models / B.D. Rouhani, H. Chen, F. Koushanfar // ArXiv e-prints, 2018.
- [2] Zhang, Y.-Q. DeepTrigger: A Watermarking Scheme of Deep Learning Models Based on Chaotic Automatic Data Annotation / Y.-Q. Zhang, Y.-R. Jia, X. Wang, Q. Niu, N.-D. Chen // IEEE Access. – 2020. – Vol. 8. – P. 213296-213305.
- [3] Adi, Y. Turning your weakness into a strength: Watermarking deep neural networks by backdooring / Y. Adi, C. Baum, M. Cisse, B. Pinkas, J. Keshet // Proceedings of the 27th USENIX Security Symposium (USENIX Security 18). – 2018. – P. 1615-1631.
- [4] Vybornova, Y. Method for Protection of Heterogeneous Data based on Pseudo-Holographic Watermarks / Y. Vybornova // 9th International Symposium on Digital Forensics and Security (ISDFS) . – 2021. – P. 1-5.
- [5] Kaggle: Dogs vs. Cats [Electronic resource]. — Access mode: <https://www.kaggle.com/c/dogs-vs-cats/data> (11.02.2022).

Интерпретация матриц внимания языковых моделей при анализе тональности текстов

Д.Э. Пашенко
Вятский государственный
университет
Киров, Россия
depashchenko@mail.ru

Е.В. Разова
Вятский государственный
университет
Киров, Россия
razova.ev@gmail.com

А.В. Котельникова
Вятский государственный
университет
Киров, Россия
kotelnikova.av@gmail.com

С.В. Вычегжанин
Вятский государственный
университет
Киров, Россия
vychegzhaninsv@gmail.com

Е.В. Котельников
Вятский государственный
университет
Киров, Россия
kotelnikov.ev@gmail.com

Аннотация—Глубокие нейросетевые языковые модели демонстрируют в последнее время впечатляющие результаты при обработке естественного языка, в том числе в области анализа тональности текстов. Однако уверенность пользователей в результатах анализа невысока в связи с плохой интерпретируемостью таких моделей. В работе предлагается метод интерпретации языковых моделей при помощи вероятностно-статистического анализа матриц внимания.

Ключевые слова— анализ тональности, нейронные сети, языковые модели, BERT, интерпретация, внимание, оценочная лексика.

1. ВВЕДЕНИЕ

Стремительное развитие в последние 3–4 года глубоких нейросетевых языковых моделей [1], таких как BERT [2], GPT-3 [3], T5 [4], Switch Transformer [5], ERNIE [6] и других, привело к значительному прогрессу в обработке естественного языка, в частности, при анализе тональности текстов. В этой области современные модели демонстрируют точность выше 97% для задач бинарной классификации (позитив/негатив) и около 60% для задач с пятизначной шкалой [7]. Большинство современных языковых моделей использует в качестве основы архитектуру Transformer [8], в которой ключевая информация содержится в т.н. матрицах внутреннего внимания (self-attention), характеризующих степень влияния различных входных сигналов на результат.

Несмотря на существенные успехи, применение глубоких нейросетевых языковых моделей имеет ряд проблем, в частности, необходимость наличия большого количества размеченных данных для обучения, высокую вычислительную трудоемкость процесса обучения, низкую интерпретируемость.

Альтернативой являются словарные методы (или методы на основе правил) – быстрые, не требующие обучения, хорошо интерпретируемые [9]. Основным лингвистическим ресурсом для них являются словари оценочной лексики, большое количество которых было разработано за последние годы, в том числе для русского языка [10]. Однако словарные методы не обеспечивают высокого качества анализа тональности и значительно уступают в этом плане нейросетевым моделям [11]. Причина такой ситуации остается до конца неясной в связи с плохой интерпретируемостью нейронных сетей.

Таким образом, актуальной является задача улучшения интерпретируемости глубоких нейросетевых языковых моделей с целью обеспечения прозрачности процедур классификации текстов по тональности и для потенциального повышения качества работы словарных методов. В настоящей работе указанная задача решается на основе вероятностно-статистического анализа распределения весов в матрицах внимания, формируемых после дообучения модели типа BERT на корпусе текстов, размеченных по тональности.

2. ОБЗОР ПРЕДЫДУЩИХ РАБОТ

Интерпретация на основе внимания предполагает анализ распределения внимания модели по токенам. Токены с наибольшим вниманием считаются наиболее полезными для предсказания модели. Тан и др. исследовали влияние токенов с наибольшим весом внимания в модели BERT_{BASE} на правильность классификации текстов по тональности [12]. Результаты показали, что первый слой более эффективен для выбора влиятельных токенов, чем последний. Као и др. [13] предложили метод дифференциальной маскировки, позволяющий определить, что «знают» о входных данных разные слои модели и где в разных слоях хранится информация о предсказании. Было установлено, что большинство слоев полагается на строго позитивные или негативные слова. Ву и др. [14] разработали метод послойного отслеживания внимания для анализа структурированных весов внимания в моделях на архитектуре Transformer. В работе была определена доля внимания, уделяемая отдельными головами оценочным словам. Авторы показали, что веса внимания имеют наибольшие значения для сильно позитивных и негативных слов.

3. МЕТОД

Предлагаемый метод интерпретации глубоких нейросетевых языковых моделей с механизмом внутреннего внимания включает четыре этапа.

1) *Построение усредненных матриц внимания для каждого текста.* Значения в исходных матрицах внимания моделей рассредоточены по отдельным входным токенам, которые часто являются не словом целиком, а его частью. Поэтому возникает задача нахождения усредненных весов внимания для слов с учетом лемматизации.

2) *Построение единых матриц внимания по всем текстам анализируемого корпуса.* Единые матрицы внимания формируются на основе усреднения матриц, полученных на предыдущем этапе.

3) *Анализ статистических характеристик сформированных единых матриц.* Анализ осуществляется с учетом различных аспектов: оценочной и нейтральной лексики; эталонной разметки текстов по тональности и разметки, присвоенной модели. В процессе анализа вычисляются и сравниваются средние, медианные, максимальные и минимальные значения весов внимания.

4) *Анализ вероятностных распределений весов внимания сформированных единых матриц.* Расчет дивергенции Кульбака-Лейблера между распределениями весов внимания различных множеств слов – позитивных, негативных и нейтральных.

Предложенный метод позволяет проверять различные гипотезы о степени внимания, уделяемого нейросетевыми языковыми моделями лексике различного типа в разных текстах при разных условиях.

4. РЕЗУЛЬТАТЫ

Для экспериментов с предложенным методом в работе использовалась модель ruRoberta-large [15], продемонстрировавшая лучшие результаты среди моделей типа BERT в рейтинге Russian SuperGLUE [16].

Эксперименты проводились на материале корпуса русскоязычных новостей, включающего 1823 текста, размеченных аннотаторами по трехбалльной шкале: «позитивно», «негативно» и «нейтрально».

Словарь оценочной лексики должен обладать высокой точностью и полнотой. Чтобы создать такой словарь были проанализированы девять общедоступных словарей русскоязычной оценочной лексики [10]. В итоговый словарь включались только те слова, которые входят не менее чем в три исходных словаря. В итоговый словарь вошли 4325 слов, в том числе 1521 позитивных (35,2%) и 2804 негативных (64,8%).

Была выдвинута гипотеза о том, что нейросетевая языковая модель ruRoberta-large уделяет в процессе анализа тональности больше внимания оценочной лексике, чем нейтральной. Применение предложенного метода позволило подтвердить данную гипотезу с помощью критерия суммы рангов Уилкоксона (уровень значимости $p=0,05$) [17]. Также были проанализированы отличия вероятностно-статистических характеристик внимания для правильно определяемых моделью текстов и ошибочных.

5. ЗАКЛЮЧЕНИЕ

Предложенный в исследовании метод интерпретации матриц внимания глубоких нейросетевых языковых моделей позволяет проверять различные гипотезы относительно распределений значений внимания между словами различных множеств при анализе тональности текстов. Метод может быть использован для повышения уверенности пользователей в результатах работы

нейросетевых языковых моделей и для улучшения качества анализа тональности на основе словарных методов.

ЛИТЕРАТУРА

- [1] Han, X. Pre-Trained Models: Past, Present and Future / X. Han, Z. Zhang, N. Ding, Y. Gu, X. Liu // AI Open. – 2021.
- [2] Devlin, J. BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding / J. Devlin, M.-W. Chang, K. Lee, K. Toutanova // Proceedings of 7th Annual Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies (NAACL-HLT). – 2019. – P. 4171-4186.
- [3] Brown, T.B. Language Models are Few-Shot Learners / T.B. Brown, B. Mann, N. Ryder, M. Subbiah, J. Kaplan // Proceedings of the 34th Conference on Neural Information Processing Systems (NeurIPS). – 2020. – P. 1877-1901.
- [4] Raffel, C. Exploring the limits of transfer learning with a unified text-to-text transformer / C. Raffel, N. Shazeer, A. Roberts, K. Lee, S. Narang // Journal of Machine Learning Research. – 2020. – Vol. 21. – P. 1-67.
- [5] Fedus, W. Switch Transformers: Scaling to Trillion Parameter Models with Simple and Efficient Sparsity / W. Fedus, B. Zoph, N. Shazeer // Computing Research Repository. – 2021. – Access mode: <http://arxiv.org/abs/2101.03961> (08.02.2022).
- [6] Sun, Y. ERNIE 3.0: Large-scale Knowledge Enhanced Pre-training for Language Understanding and Generation / Y. Sun, S. Wang, S. Feng, S. Ding, C. Pang // Computing Research Repository. – 2021. – Access mode: <http://arxiv.org/abs/2107.02137> (08.02.2022).
- [7] Papers with code: Sentiment Analysis [Electronic resource]. – Access mode: <https://paperswithcode.com/task/sentiment-analysis> (08.02.2022).
- [8] Vaswani, A. Attention is All you Need / A. Vaswani, N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones // Proceedings of the 31st Conference on Neural Information Processing Systems (NeurIPS). – 2017. – Vol. 30. – P. 6000-6010.
- [9] Taboada, M. Sentiment Analysis: An Overview from Linguistics / M. Taboada // Annual Review of Linguistics. – 2020. – Vol. 2. – P. 325-347.
- [10] Котельников, Е.В. Современные словари оценочной лексики для анализа мнений на русском и английском языках / Е.В. Котельников, Е.В. Разова, А.В. Котельникова, С.В. Вычерганин // Научно-техническая информация. Сер. 2. Информационные процессы и системы. – 2020. – № 12. – С. 16-33.
- [11] Kotelnikova, A.V. Lexicon-based Methods and BERT Model for Sentiment Analysis of Russian Text Corpora / A.V. Kotelnikova, D.E. Paschenko, E.V. Razova // CEUR Workshop Proceedings. – 2021. – Vol. 2922. – P. 73-81.
- [12] Tan, N.O. An explainability analysis of a sentiment prediction task using a transformer-based attention filter / N.O.Tan., J. Bensemann, D. Benavides-Prado, Y. Chen, M. Gahegan // Proceedings of the Ninth Annual Conference on Advances in Cognitive Systems. – 2021. – P. 1-7.
- [13] Cao, N.D. How do Decisions Emerge across Layers in Neural Models? Interpretation with Differentiable Masking / N.D. Cao, M.S. Schlichtkrull, W. Aziz, I. Titov // Proceedings of the Conference on Empirical Methods in Natural Language Processing. – 2020. – P. 3243-3255.
- [14] Wu, Z. Structured Self-Attention Weights Encode Semantics in Sentiment Analysis / Z. Wu, T.-S. Nguyen, D. Ong // Proceedings of the Third BlackboxNLP Workshop on Analyzing and Interpreting Neural Networks for NLP. – 2020. – P. 255-264.
- [15] Sberbank-ai/ruRoberta-large [Electronic resource]. – Access mode: <https://huggingface.co/sberbank-ai/ruRoberta-large> (08.02.2022).
- [16] Russian SuperGLUE: Leaderboard [Electronic resource]. – Access mode: <https://russiansuperglue.com/leaderboard/2> (08.02.2022).
- [17] Wilcoxon, F. Individual comparisons by ranking methods / F. Wilcoxon // Biometrics Bulletin. – 1945. – Vol 6(1). P. 80-83.

Разработка интеллектуальной системы управления ресурсами пищевых предприятий

Д.Е. Новичков
Самарский государственный
технический университет
Самара, Россия
dmitriy.novichkov@gmail.com

С.П. Грачев
Самарский государственный
технический университет
Самара, Россия
sg@kg.ru

Е. Пантелей
Самарский государственный
технический университет
Самара, Россия
mekachiku-san@mail.ru

Аннотация—В работе рассмотрена разработка интеллектуальной системы управления ресурсами пищевых предприятий, основанная на мультиагентных технологиях и сетцентрическом подходе. Разработаны и описаны алгоритмы адаптивного планирования и сформирована онтологическая база знаний. Указаны решаемые системой задачи, приведено описание используемых алгоритмов.

Ключевые слова— мультиагентные технологии, базы знаний, управление ресурсами.

1. ВВЕДЕНИЕ

Современное пищевое предприятие требует решения значительного числа разнородных по содержанию, но связанных между собой задач: логистика доставки кормов, людей и продукции; планирование и управление производством, прогноз закупок и т. д. Решение каждой из этих задач происходит, как правило, в обособленном режиме: разные информационные системы решают задачи независимо друг от друга, после чего происходит согласование планов в ручном режиме. Построенные таким образом решения требуют большого количества доработок и, как следствие, не позволяют в полной мере ИТ-системам выполнять свои функции. Разрабатываемая интеллектуальная система позволяет решать задачи сопряженного планирования логистики и производства, включая элементы прогнозирования, такой подход позволяет повысить эффективность управления ресурсами пищевых предприятий.

Задачу построения расписаний ресурсов при выполнении работ, состоящих из операций, (Job Shop Scheduling Problem, JSP и ее разновидностей, например, Flexible Scheduling Problem и аналогичных) в общем виде можно рассматривать такую задачу можно как задачу программирования в ограничениях. Чаще всего подходы к решению таких задач основаны на традиционных алгоритмах и в основе своей содержат линейное и динамическое программирование. Более современными являются методы генетических алгоритмов и роевой оптимизации, методы искусственного интеллекта с использованием нейросетевых технологий [1].

2. ПОСТАНОВКА ЗАДАЧИ

Необходимо разработать модели, методы и алгоритмы согласования решений из систем планирования и управления ресурсами различной зоны ответственности, обеспечивая при этом адаптивное управление с целью достижения целевых показателей. При этом необходимо учитывать различные прогнозные сценарии продаж и иметь возможность моделировать развитие ситуаций по этим сценариям. Эвристики планирования, полученные от пользователей, необходимо хранить в семантической

базе знаний, что позволит менять логику работы систем без внесения изменений в исходный код системы.

3. ПРЕДЛАГАЕМЫЙ ПОДХОД

В основе решения лежат мультиагентные технологии и сетцентрический принцип: система состоит из набора «маленьких» систем, каждая из которой решает поставленную перед собой задачу в общем информационном пространстве, непрерывно согласуя и корректируя формируемые решения[2]. Каждая такая система представляет собой мультиагентную систему планирования; в этом подходе у каждой сущности реального мира формируется агент – программный объект, представляющий «интересы» в виртуальном мире. Дополнительно вводятся агенты целевых функций системы, технологических требований и требований нормативной документации, что позволяет учесть особенности функционирования в предметной области. Набор основных агентов приведен в табл. 1.

Таблица 1. ПЕРЕЧЕНЬ РЕАЛИЗОВАННЫХ АГЕНТОВ

Тип агента	Цели и поведение
Агент потребности	Формирует запрос на выполнение производственного заказа. Контролирует ход выполнения (соблюдение дат и ограничений).
Агент транспортного средства	Участствует в выполнении логистических операций. Контролирует выполнение нормативных требований (максимальная загрузка, проведение планово-предупредительных ремонтов).
Агент производственной площадки	Консолидирует потребности и ход их выполнения, контролирует выполнение нормативных и технологических требований, обеспечивает минимизацию простоев автотранспорта.
Агент требования	Контролирует выполнение требований в ходе построения плана. Требования могут быть разного типа, в общем случае проверяются сформированные между несколькими агентами отношения и характеристики этих отношений.
Агент технологического процесса	Обеспечивает корректность выполнения технологического процесса в соответствии с семантическим описанием в базе знаний.
Агент сцены	Обеспечивает процесс функционирования агентов, реализуя связь агентов с онтологическим описанием их сущностей.

Схема работы каждого агента выглядит следующим образом: он получает сообщение, исходя из которого выбирает следующее действие. Набор возможных действий зависит от типа агентов, критериев его целевых функций и текущего состояния системы. Часть

функциональности реализована в виде семантической базы знаний, которая является отдельным программным комплексом. В базе знаний представлен инструмент для формирования семантически связанных данных – знаний, которые используются при планировании [3]. К примерам таких данных можно отнести требования нормативной документации и технологические требования, информацию о логистических графах и эвристики функционирования агентов. В общем виде алгоритм работы агента состоит из следующих шагов [4]:

1. Инициализация, в ходе которой агент получает доступ к сущности физического мира, базе знаний и сцене. Сцена представляет собой отображение состояние мира агентов в данный момент времени: их состояния, достигнутые целевые показатели и др.
2. Фаза быстрого планирования, в ходе которой происходит быстрое формирование решений без поиска оптимальных состояний.
3. Фаза проактивности, в которой происходит оптимизация первоначальных планов. В этой фазе агенты могут конкурировать друг с другом и вытеснять друг друга из расписаний, основываясь на функциях бонусов и штрафов.

Пример диаграммы последовательности фазы быстрого планирования в общем виде приведен на рис.1, фаза проактивности отличается наличием рекурсивного вытеснения задач в случае возникновения конфликтов в расписании.

Процесс согласования решений между системами основан на схожем принципе [5]: системы планирования «договариваются» между собой, согласуя планы и моделируя ситуации «а что, если...». При этом имеется

возможность работать как на реальных данных, так и на модельных сценариях, что позволяет учесть различные варианты развития ситуаций.

4. ЗАКЛЮЧЕНИЕ

Разрабатываемая система позволяет решать задачи адаптивного обеспечения кормами производственных площадок для животноводческих предприятий, оптимизации доставки персонала от мест проживания до производственных площадок, оптимизации пребывания автотранспорта на производственных площадках, адаптивного сопряженного планирования производства и продаж. Использование семантической базы знаний предприятия позволяет описать учесть специфические особенности организаций и наработанный в них опыт.

ЛИТЕРАТУРА

- [1] Petcu, A. A class of algorithms for distributed constraint optimizations, *Frontiers in Artificial Intelligence and Applications* / A. Petcu // IOS Press. – 2009. – Vol. 194(1). DOI: 10.5075/epfl-thesis-3942.
- [2] Грачев, С.П. Методы и средства построения интеллектуальных систем для решения сложных задач адаптивного управления ресурсами в реальном времени / С.П. Грачев, А.А. Жилиев, В.Б. Ларюхин // *Автоматика и телемеханика*. – 2021. – № 11. – С. 30-67. DOI 10.31857/S0005231021110039.
- [3] Usman, Z. A Manufacturing Core Concepts Ontology for Product Lifecycle Interoperability / Z. Usman, R.I.M. Young, N. Chungoora, C. Palmer, K. Case, J. Harding // *Proceedings of the International IFIP Working Conference on Enterprise Interoperability IWEI*, Stockholm, Sweden. – Springer: Berlin/Heidelberg, Germany, 2011. – P. 5-18.
- [4] Skobelev, P. Ontology-based open multi-agent systems for adaptive resource management / P. Skobelev, A. Zhilyaev, V.B. Larukhin, S. Grachev, E.V. Simonova // *Proceedings of the 12th International Conference on Agents and Artificial Intelligence (ICAART)*, Valetta, Malta. – SciTePress: Setúbal, Portugal, 2020. – Vol. 1. – P. 127-135.
- [5] Иващенко, А.В. Мультиагентные технологии для разработки сетевых систем управления / А.В. Иващенко, О.В. Карсаев, П.О. Скобелев, А.В. Царев, Р.М. Юсупов // VI Всероссийская научно-практическая конференция "Перспективные системы и задачи управления". – Известия ЮФУ. Технические науки. – 2011. – № 3(116). – С. 11-23.

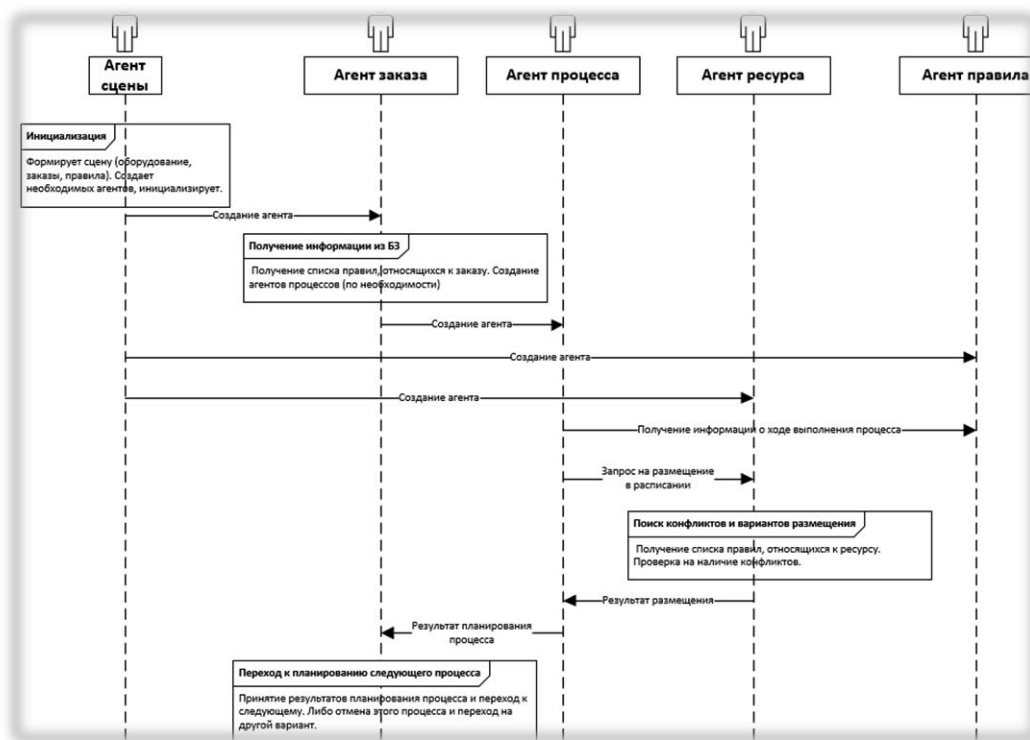


Рис. 1. Диаграмма последовательности фазы быстрого планирования

Разработка прототипа интеллектуальной инсулиновой помпы

К.К. Абдулхаликова
Казанский национальный
исследовательский технический
университет им. А.Н.Туполева - КАИ
Казань, Россия
abdulhalikova.karina@yandex.ru

П.И. Тутубалин
Казанский национальный
исследовательский технический
университет им. А.Н.Туполева - КАИ
Казань, Россия
pitutubalin@kai.ru

А.Ю. Александров
Казанский национальный
исследовательский технический
университет им. А.Н.Туполева - КАИ
Казань, Россия
pitutubalin@kai.ru

С.В. Сотников
Казанский национальный
исследовательский технический
университет им. А.Н.Туполева - КАИ
Казань, Россия
pitutubalin@kai.ru

И.Н. Урахчинский
Казанский национальный
исследовательский технический
университет им. А.Н.Туполева - КАИ
Казань, Россия
pitutubalin@kai.ru

Аннотация—Представлено обоснование энергоэффективного технического решения по разработке прототипа интеллектуальной инсулиновой помпы. Рассмотрен метод, обеспечивающий безопасное регулирование гликемии пациента, больного сахарным диабетом 1 типа. Предложено для реализации метода регулирования гликемии за основу взять устройство, называемое интеллектуальной инсулиновой помпой. Устройство интеллектуальной инсулиновой помпы реализует алгоритмы регулирования гликемии. В качестве основы для алгоритмов регулирования предлагается взять базовый принцип регулирования на основе ПИД-регулятора и перспективный, основанный на использовании нейронных сетей. При этом, среди основных параметров, заложенных в регуляторе, для отмеченных алгоритмов, выступают такие параметры, как: инсулин «на борту» (IOB), углеводы «на борту» (COB), время действия инсулина (DIA), коэффициент чувствительности к инсулину (S) и учитываемые индивидуальные характеристики пациента. На основе предлагаемой модели проведен ряд экспериментов.

Ключевые слова— интеллектуальная инсулиновая помпа, PID-регулятор, нейросеть, лечение сахарного диабета 1 типа.

1. ВВЕДЕНИЕ

Проблема разработки надежной и автономной системы регулирования гликемии для лечения сахарного диабета (СД) 1 типа остается актуальной до сих пор, несмотря на большое количество научных изысканий и продолжающееся увеличение уровня заболеваемости СД. Практически единственным способом лечения СД 1 типа является применение инсулиновой помпы. Однако, применение этого метода сопряжено с рядом проблем, заключающихся в частых ситуациях перерегулирования: гипо- и гипергликемии.

Одним из способов устранения отмеченного недостатка является разработка и внедрение использования интеллектуальной инсулиновой помпы (ИИП). На данный момент такого рода работы находятся на стадии прототипирования готового устройства.

2. УСТРОЙСТВО ПРОТОТИПА ИИП

ИИП – это техническое устройство, предназначенное для помощи пациентам, больным СД 1 типа. Это устройство обеспечивает автоматическое и непрерывное отслеживание и регулирование уровня глюкозы в крови, то есть стремится подменить ряд основных функций поджелудочной железы здорового человека. В качестве прототипа ИИП выступает платформа OpenAPS, являющаяся платформой разработки [1-3]. Вкупе с - инсулиновой помпой, применяемой для лечения больных СД, модули ИИП образуют биоинформационную систему образом схожим с секрецией инсулина у здорового человека.

К разработке такого устройства могут быть применены такие требования, как: безопасность для пациента, высокая надежность компактность, энергоэффективность и простота в эксплуатации. Коротко рассмотрим функциональную схему, представленную на Рис. 1.

В состав ИИП входят:

1. CGM – устройство, позволяющее получать данные о гликемии с интервалом 5 минут;
2. радиомодуль – устройство, предназначенное для обеспечения взаимодействия микрокомпьютера, инсулиновой помпой (ИП);
3. инсулиновая помпа – устройство, которое позволяет изменять свои настройки дистанционно;
4. микрокомпьютер, в состав которого входит микроконтроллер, SD-карта и т.д.

3. ПРИНЦИП РАБОТЫ ПРОТОТИПА ИИП

Существует 2 наиболее распространенных [1] класса алгоритмов регулирования гликемии, реализованных в ИИП и предназначенных для поддержания гликемии в целевом диапазоне:

- PID-регулятор;
- нейросеть.

PID-регулятор представляет собой алгоритм регулирования, учитывающий на каждом шаге регулирования ошибку e посредством обратной связи.

PID-регулятор отличается универсальностью, простотой и быстрой реакцией на возмущающее воздействие. С другой стороны, PID-регулятор не позволяет осуществлять прогнозирование и обеспечить «обучение» и персонализацию ИИП, которые способны обеспечить нейронная сеть. Нейросеть в основе регулятора ИИП позволяет создать более гибкий алгоритм регулирования.

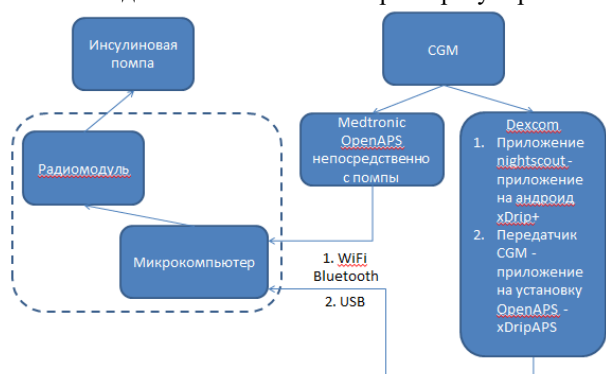


Рис. 1. Схема функциональная ИИП

1. PID-регулятор. На каждом шаге алгоритма корректирует базальную скорость и вводит болюсную дозу (при необходимости) для минимизации ошибки e , являющейся разницей между текущим и целевым значением глюкозы. Ошибка e является результатом текущего значения и целевого значения гликемии. Конечной целью регулирования является ошибка e , которая на каждом шаге равна нулю. Алгоритм регулирования учитывает значение IOB – инсулина, который активно понижает глюкозу в крови/плазме и COB – углеводов, находящихся в организме человека и активно повышающих глюкозу в крови. IOB является суммой введенных единиц инсулина, с учетом времени действия инсулина и времени с момента его введения.

2. Нейросети. В настоящее время актуальным направлением развития биомедицинских технологий является персонализированная медицина. Это подразумевает, что на протяжении всего курса лечения осуществляется мониторинг на основании доступных клинических, а впоследствии и лабораторных критериев с возможной коррекцией путем эмпирических попыток изменения схем лечения. И для такого подхода наиболее подходящими оказываются нейронные сети, обеспечивающие в процессе их обучения приведение модели регулирования к наблюдаемому состоянию пациента. Нейросеть позволяет расширить возможности регулирования и создать более гибкие алгоритмы регулирования [1].

Сравнение двух отмеченных подходов в регулировании гликемии можно провести по следующим критериям:

- время нахождения в целевом диапазоне, т.е. безопасность для пациента;
- быстродействие (время возврата гликемии в целевой диапазон);
- перерегулирование (доля времени гипо- и гипергликемии);
- оценка вычислительной сложности.

4. ЭКСПЕРИМЕНТ

Эксперимент целесообразно провести с регулятором на основе PID-регулятора и на основе нейросети. Это связано с тем, что в основе закона регулирования прототипа лежит PID-регулятор.

Процесс регулирования гликемии был смоделирован в среде *in silico* MATLAB SimBio для PID-регулятора и нейросети. Моделирование осуществлялась с помощью модели *m1* глюкозо-инсулинового ответа [5-6].

На Рис.2 представлены выходные величины, с помощью которых осуществляется регулирование, это *basalDose* и/или *bolusDose*. Значения *bolusDose*, вычисляются для приема пищи, с учетом COB, углеводов и степени энергозатрат пациента больного СД. Значение IOB указывает активный инсулин *basalDose* и/или *bolusDose*, находящийся в организме в момент регулирования.

Дозировка инсулина рассчитывается с учетом:

- 1) коэффициента чувствительности к инсулину;
- 2) расчетного значения суточной дозы инсулина;
- 3) суммарной дозы инсулина, введенной ИИП;
- 4) вида и времени действия инсулина;
- 5) питания, потребленного количества углеводов;
- 6) IOB;
- 7) COB;
- 8) возраста, веса, комплекции.

5. ЗАКЛЮЧЕНИЕ

Регулирование на основе нейросети показало не только эффективность для регулирования гликемии, но и возможность персонализации лечения СД 1 типа. С помощью персонализации алгоритма посредством нейронных сетей можно устранить эту проблему. Однако данное исследование позволяет внедрить платформу в разработку и улучшить качество моделирования алгоритмов регулирования гликемии пациента.

ЛИТЕРАТУРА

- [1] Tutubalin, P. Status of creation of hardware-software complex of automatic control of the insulin delivery / P. Tutubalin // J. Phys.: Conf. Ser. 2019. – Vol. 1368. – P. 042006. DOI:10.1088/1742-6596/1368/4/042006.
- [2] Тутубалин, П.И. О платформе для реализации проекта искусственная поджелудочная железа / П.И. Тутубалин, А.П. Кирпичников, Р.С. Асадов // Вестник технологического университета. – 2019. – Т. 22, № 8. – С. 139-144.
- [3] Тутубалин, П.И. О работе по созданию прототипа образца искусственной поджелудочной железы / П.И. Тутубалин, С.В. Новикова, А.П. Кирпичников. // Вестник технологического университета. – 2019. – Т. 22, № 9. – С. 147-152.
- [4] Карпельев, В.А. Разработка и проверка работы ПИД-регулятора для искусственной поджелудочной железы с интраперитонеальным введением инсулина / В.А. Карпельев, Ю.И. Филиппов, А.В. Аврин // Сахарный диабет. – 2018. – № 21(1). – С. 58-65.
- [5] Dalla Man, C. Meal Simulation Model of the Glucose-Insulin System / C. Dalla Man, R.A. Rizza, C. Cobelli // IEEE Transactions on Biomedical Engineering. – 2007. – Vol. 54(10). – P. 1740-1749.
- [6] Dalla Man, C. A System Model of Oral Glucose Absorption: Validation on Gold Standard Data / C. Dalla Man, M. Camilleri C. Cobelli // IEEE Transactions on Biomedical Engineering. – 2006. – Vol. 53(12). – P. 2472-2478.

Обнаружение нарушений целостности поверхностей зданий при помощи нейросетевых алгоритмов

Н.А. Фирсов

*Самарский национальный исследовательский университет
им. академика С.П. Королева
Самара, Россия*

firsov.na98@gmail.com

Д.А. Жердев

*Институт систем обработки изображений - филиал
ФНИЦ «Кристаллография и фотоника» РАН
Самарский национальный исследовательский университет
им. академика С.П. Королева
Самара, Россия*

denazherdev@gmail.com

Л.А. Жердева

*Самарский национальный исследовательский университет
им. академика С.П. Королева
Самара, Россия*

lara.zherdeva.taskina@gmail.com

Е.Ю. Минаев

*Самарский национальный исследовательский университет
им. академика С.П. Королева
Самара, Россия*

e.minaev@gmail.com

Аннотация—Обнаружение нарушений целостности поверхностей зданий требует привлечения рабочего персонала, при этом зачастую требуются дополнительные приспособления и спецтехника (лестницы, автовышки и пр.) и возникают риски производственных травм. Автоматизация обнаружения нарушений целостности зданий с помощью видеоаналитики данных с БПЛА может ускорить, удешевить и обезопасить данную процедуру. Применение нейросетевых алгоритмов к данной задаче осложнено специфичностью данных, которые требуется собирать на определенных зданиях и площадях инфраструктуры. Для решения этой задачи предлагается использовать синтетические данные, получаемые при помощи виртуальной среды. Синтезированный набор данных позволяет полностью решить поставленную задачу обнаружения трещин фасадов и повреждений кровли.

Ключевые слова— моделирование повреждений, трещины, синтетические данные, сверточные нейронные сети.

1. ВВЕДЕНИЕ

Предприятия с большим количеством инфраструктурных объектов нуждаются в оперативном контроле состояния каждого из объектов для более эффективного управления и предупреждения нежелательных рисков. Вызвано это может быть прежде всего юридическими и нормативными требованиями в отношении устойчивости и безопасности, а также повышением конкурентоспособности организации. На данный момент визуальный осмотр остается основным методом обнаружения дефектов поверхности, таких как трещины, сколы и коррозия. Условия, отмеченные в ходе периодических инспекций, могут потребовать более частых проверок, что заставляет сосредоточиться на повышении эффективности, безопасности и точности визуального контроля. Проблема также заключается в том, что этот тип инспекции по большей части основан на ручном сборе данных невооруженным глазом, а некоторые подходы требуют от инспекторов доступа к различным труднодоступным компонентам инфраструктуры. Таким образом, для обеспечения безопасности, для эффективного и точного осмотра

инфраструктурных объектов требуется применять автоматизированную систему сбора данных [1]. Беспилотная система, уменьшающая роль инспектора, обеспечивает более безопасный и эффективный контроль, т.к. может подлететь ко всем частям большой конструкции для сбора данных с помощью установленных на нём камер и датчиков [2], а автоматизация траектории полета БПЛА приводит к получению пути без столкновений с минимальным перекрытием, максимальным покрытием и минимальным временем полета [3]. Использование нейросетевых алгоритмов позволяет автоматизировать процесс обработки видеоданных, уменьшая нагрузку на персонал и увеличивая эффективность мониторинга повреждений. В связи со сложностью получения набора данных для обучения нейросетей возможно использование технологии моделирования синтетических данных.

2. ОБНАРУЖЕНИЕ ДЕФЕКТОВ

В качестве нейросетевого классификатора для задачи обнаружения дефектов была применена сеть Res-UNet [4]. Данные для обучения собирались с видеоклипов и фотографий, полученных при полевых съемках реальных инфраструктурных объектов. Данные кадры проходили этап ручной разметки под определенную задачу, после чего производилось кадрирование изображения на участки размером 256x256 пикселей. Формировались обучающий и валидационный наборы в соотношении 9 к 1. Обучение проводилось до достижения точности в 95-98% на валидационном наборе. Маски, полученные на тестовых изображениях, проходили этап размывтия и пороговой обработки для отсеивания ложных срабатываний на отдельных пикселях.

В отсутствие возможности съемки сцен реального мира с требуемыми повреждениями, обучающие выборки формировались при помощи моделируемых синтетических данных в среде Unreal Engine (рисунок 1). Таким образом для идентификации механических дефектов кровли и поверхностей было синтезировано порядка тысячи изображений.



Рис. 1. Синтетические данные смоделированные в среде Unreal Engine



Рис. 2. Результаты сегментации нейросетевым алгоритмом

3. ЗАКЛЮЧЕНИЕ

Нейросетевой алгоритм был обучен на полученных синтетических данных и показал 98.6 % точность обнаружения дефектов и нарушений целостности поверхности зданий на изображениях (рисунок 2). Как можно видеть из рисунков, используемый подход позволяет обучить и распознать дефекты с минимальными временными и ресурсными затратами.

БЛАГОДАРНОСТИ

Работа выполнена при поддержке РФФИ грант № МК 19-29-01235.

ЛИТЕРАТУРА

- [1] Liu, P. A review of rotorcraft Unmanned Aerial Vehicle (UAV) developments and applications in civil engineering / P. Liu, A.Y. Chen, Y.N. Huang, J.Y. Han, J.S. Lai, S.C.J. Kang, T.H.R. Wu, M.C. Wen, M.H. Tsai // *Smart Structure and Systems*. – 2014. – Vol. 13(6). – P. 1065-1094.
- [2] Bircher, A. Three-dimensional coverage path planning via viewpoint resampling and tour optimization for aerial robots / A. Bircher, M. Kamel, K. Alexis, M. Burri, P. Oettershagen, S. Omari, T. Mantel, R. Siegwart // *Autonomous Robots*. – 2016. – Vol. 40(1). – P. 1059-1078.
- [3] Hinks, T. Flight optimization algorithms for aerial LiDAR capture for urban infrastructure model generation / T. Hinks, H. Carr, D.F. Laefer, M. Asce // *Journal of Computing in Civil Engineering*. – 2009. – Vol. 23(6). – P. 330-339.
- [4] Diakogiannis, F.I. ResUNet-a: a deep learning framework for semantic segmentation of remotely sensed data / F.I. Diakogiannis, F. Waldner, P. Caccetta, C. Wu // *ISPRS Journal of Photogrammetry and Remote Sensing*. – 2020. – Vol. 16. – P. 94-114.

Dynamic camera spectral sensitivity estimation

A. Nigmatzyanov

*Institute for Information Transmission
Problems (Kharkevich Institute)
Moscow Institute of Physics and
Technology
Moscow, Russia*

D. Shepelev

*Institute for Information Transmission
Problems (Kharkevich Institute)
Moscow, Russia
shepelev@iitp.ru*

V. Vasilev

*Institute for Information Transmission
Problems (Kharkevich Institute)
Moscow Institute of Physics and
Technology
Moscow, Russia
vasilev.va@phystech.edu*

E. Ershov

*Institute for Information Transmission
Problems (Kharkevich Institute)
Moscow, Russia
e.i.ershov@gmail.com*

M. Tchobanou

*Moscow Research Center, Huawei
Technologies Co. Ltd
Moscow, Russia
Mikhail.Tchobanou@huawei.com*

Abstract—In large-scale series production the time for evaluating the camera spectral sensitivity is strongly limited and measured in units of seconds because of production and economic constraints. To estimate variation of spectral sensitivity properties, manufacturers usually precisely measure only a few sensors (the golden set) and use these measurements to perform quick estimation of any other sensor in the released pack. The main drawback of this approach is that the worst color reproduction error cannot be controlled for a particular device: instability of device production process usually causes significantly different sensors, which may not be included in the golden set. In that case the camera will work with low accuracy during the lifetime. To overcome this problem, we consider a new approach to camera spectral sensitivity estimation during its operation. The main idea is based on consistency estimation of images and average scenes spectra. Users receive such a combination of data in practice, for instance modern phone devices have built-in integral spectrometers. Also, the proposed approach can be considered in the scope of classical problem statement of spectral sensitivity estimation with color charts. In the paper we investigated the accuracy of the method of spectral sensitivity estimation based on the basis calculation with singular value decomposition of the sensitivities from the golden set in combination with different types of regularization.

Keywords— *dynamic camera calibration, spectral sensitivity estimation, golden set, quality of color reproduction, color patches.*

1. INTRODUCTION

We consider the problem of evaluating the spectral sensitivity of a sensor in operation, without the use of laboratory equipment and color targets [1]. Such formulation is rare because of the high accuracy requirements for this problem, which is difficult to achieve under uncontrolled conditions. The only example of a similar problem statement is the work [2], where the authors propose to perform the calibration based on averaged sky spectra. They calculated the sky spectra with geo- and meteo data.

In our work the new approach (figure 1) of camera spectral sensitivity calibration was suggested. It based on consistency estimation averaged image color and received spectra data from spectrometer. In this work we imply that FOV of camera and spectrophotometer are coincides. Our work is the first step in creation of the sensor with the real-time auto – spectral calibration system.

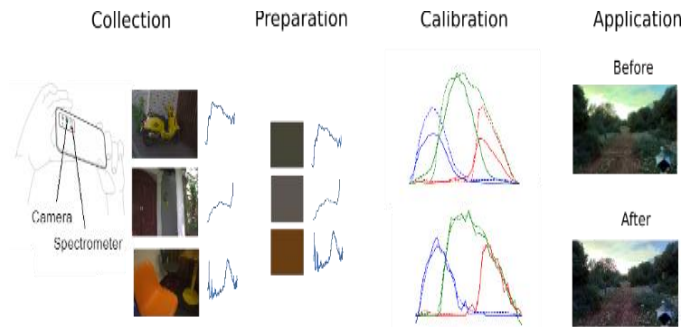


Fig. 1. The illustration of stages of the process of dynamic camera spectral sensitivity correction. Second one is their preparation. Final stage is calibration

2. RECONSTRUCTION PROBLEM

Let the dataset includes m pairs (image, averaged radiance $\mathbf{r} \in \mathbb{R}^n$ over scene), where $n = 431$ – is the wavelengths number. Let's write image linear formation model for all channels in the following form:

$$\mathbf{c} = \mathbf{r}^T X \quad (1)$$

where it is assumed that $\Delta\lambda$ incorporated in \mathbf{r}, \mathbf{c} – is three-dimensional tristimulus vector, $X \in \mathbb{R}^{n \times 3}$ – is the matrix of spectral sensitivities $\chi_i \in \mathbb{R}^n$ in the columns, $i = \overline{1,3}$. For m pairs (1) can be written in matrix form:

$$C = RX, \quad (2)$$

where $C \in \mathbb{R}^{m \times 3}$ – is the tristimulus matrix, $R \in \mathbb{R}^{m \times n}$ – is the reflectance matrix.

In practice, to reduce the number of parameters, χ_i is expressed through a linear combination of a certain basis functions set B_i :

$$\chi_i = B_i^T \mathbf{x}_i, \quad (3)$$

where $\mathbf{x}_i \in \mathbb{R}^{k_i}$ – is the coefficient vector, $B_i \in \mathbb{R}^{k_i \times n}$ is the matrix composed of basis functions (k_i – the number of basis functions for i -th channel). In this case, the tristimulus component \mathbf{c}_i for all images:

$$\mathbf{c}_i = R B_i^T \mathbf{x}_i, \quad (4)$$

where $\mathbf{c}_i \in \mathbb{R}^m$.

According to introduced notation and paper [1] spectral sensitivity estimation of i -th channel can be written in the following form:

$$\hat{\mathbf{x}}_i = \underset{\mathbf{x}_i}{\operatorname{argmin}} (||\mathbf{c}_i - R B_i^T \mathbf{x}_i ||_2^2 + \gamma_i \cdot r(\mathbf{x}_i)), \quad (5)$$

where γ_i – regularization parameter; $r(x_i)$ – regularization term.

For example, in case of Tikhonov regularization:

$$r_i(x_i) = \|L_i x_i\|_2^2, \quad (6)$$

In a similar manner to [1], we applied the singular value decomposition (SVD) method to extract the basis functions from the golden set of known sensitivities. We used best performed optimization methods from scikit-learn to solve (5).

For shape comparison of estimated spectral sensitivity with ground truth (GT) sensitivity, we introduced normalized spectral recovery error:

$$NSE = \frac{1}{3} \sum_{i=1}^3 \frac{\|x_i / \max(X) - \hat{x}_i / \max(\hat{X})\|_2}{\|x_i / \max(X)\|_2} \cdot 100\% \quad (7)$$

where \hat{X} is the vector matrix of estimated sensitivities, $\max(X)$ and $\max(\hat{X})$ are the highest sensitivity values in these vectors over the entire wavelengths grid. This error gives the general quality measure of sensitivity recovery for all channels on an equal scale.

3. EXPERIMENTS AND RESULTS

Here we present the results of a series of experiments performed with singular value decomposition (SVD) on synthetic data of images and spectra for sensitivity estimation according to (5). Synthetic data were calculated based on the 450 public hyperspectral images which were collected as part of the NTIRE2020 workshops [3]. They guarantee the alignment of FOV and Canon noise parameters (assuming the low noise of the hyperspectral data). The spectral sensitivities of 9 cameras Canon EOS 600D, Canon 500D, Canon 300D, Canon 60D, Canon 50D, Canon 40D, Canon 20D, Canon 5D Mark II, Canon 1D Mark III were used as a golden set [5]. In the first experiment the golden set includes the camera that spectral sensitivity will be evaluated. In the second one we exclude estimated sensitivity from the set. According to (5), we use different regularization types of $r(x_i)$ and vary their parameters γ to find the best ones. Both experiments showed that Tikhonov Regularization and Tikhonov Regularization based on Derivatives [1] do not increase accuracy using SVD. For this reason, it was proposed to use Tikhonov Regularization without the first component and L1 regularization (table 1), which give a significant increase in accuracy (1.5 times).

Таблица I. AVERAGED NSE (%) ERRORS FOR CANON CAMERAS BASED ON SYNTHETIC NOISED DATA WHEN THE GOLDEN SET INCLUDES ESTIMATED SENSITIVITY

Basis	The number of basis functions		
	1	2	3
SVD	9.52± 3.57	10.94± 3.42	13.54± 1.26
SVD + L2 regularization without first	9.52± 3.57	8.66± 3.88	8.27±2
SVD + L1 regularization	9.78± 3.76	9.51± 3.19	8.79± 1.27

Tikhonov Regularization without the first component means that matrix L_i in (6) has zero first column.

When estimated sensitivity is not in the golden set, SVD shows lower accuracy. Averaged NSE (%) errors of recovery are given in table 2.

Таблица II. AVERAGED NSE (%) ERRORS FOR CANON CAMERAS BASED ON SYNTHETIC NOISED DATA WHEN THE GOLDEN SET DOES NOT INCLUDE ESTIMATED SENSITIVITY

Basis	The number of basis functions		
	1	2	3
SVD	10.25± 4.25	14.5± 6.52	15.22± 3.78
SVD + L2 regularization without first	10.25± 4.25	10.0± 5.5	9.87± 4.24
SVD + L1 regularization	10.46± 4.47	10.44± 5.09	9.62± 4.7

In the third experiment for each camera from the golden set we found the closest one among others in terms of NSE error, namely, chose for every i -th golden set's sensitivity χ_i such another sensitivity χ_j that $NSE(\chi_i, \chi_j)$ is minimum. We name this operation as *selection of the closest from the set (SCS)*. The average NSE (%) error of SCS over all cameras from the golden set is 10.3 % NSE. The averaged recovery error with SVD from the previous experiment is 9.62 % NSE. Thus, we can conclude that the considered method not only allows to increase the estimation accuracy working with noisy data, but it is better to use a linear mixture of the set's sensitivities.

4. CONCLUSION

The paper considers a new approach to spectral sensitivity estimation of device camera in operation. The results of the work show a high potential for the applicability of this approach. Our research has several restrictions: despite the added noise we use only synthetic data, and it is assumed that camera and spectrometer are aligned by field of view. Nevertheless, this is a necessary first step in the development of the dynamic calibration approach.

ACKNOWLEDGMENT

This research was supported by Huawei Technologies Co. Ltd. Russia.

REFERENCES

- [1] Darrodi, M.M. Reference data set for camera spectral sensitivity estimation / M.M. Darrodi, G. Finlayson, T. Goodman, M. Mackiewicz // J. Opt. Soc. Am. A. – 2015. – Vol. 32(3). – P. 381-391.
- [2] Rei, K. Camera Spectral Sensitivity and White Balance Estimation from Sky Images / K. Rei, H. Zhao, R. Tan, K. Ikeuchi // International Journal of Computer Vision. – 2013. – P. 105.
- [3] Arad, B. “NTIRE 2020 Challenge on Spectral Reconstruction from an RGB Image” Conference / B. Arad, R. Timofte, O. Ben-Shahar, Y. Lin // IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW). – 2020.
- [4] Jahne, B. Digital Image Processing / B. Jahne. – Berlin: Springer-Verlag, 2005. – P. 607.
- [5] Jiang, J. What is the space of spectral sensitivity functions for digital color cameras? / J. Jiang, D. Liu, J. Gu, S. Süsstrunk // IEEE Workshop on Applications of Computer Vision (WACV). – 2013. – P. 168-179.

Распознавание дорожных знаков в интеллектуальной транспортной геоинформационной системе ITSGIS

Н. Клепиков

*Самарский национальный исследовательский университет им.
академика С.П. Королева*
Самара, Россия
klepikov-nikita@mail.ru

Т. Михеева

*Самарский национальный исследовательский университет им.
академика С.П. Королева*
Самара, Россия
Mikheevati@gmail.com

Аннотация—В данной статье описано распознавание дорожных знаков с помощью сверточных нейронных сетей, а также реализация плагина «Дорожные знаки» в интеллектуальной транспортной геоинформационной системе ITSGIS. Плагин "Дорожные знаки" облегчит работу пользователям вышеуказанной разработки. Цель работы – разработать эффективный метод распознавания дорожных знаков, обладающий высокой степенью инвариантности к искажениям. Экспериментальные исследования показали, что с помощью сверточных нейронных сетей можно достичь высокой точности распознавания изображений.

Ключевые слова— дорожные знаки, обработка изображений, системы компьютерного зрения, интеллектуальная транспортная геоинформационная система ITSGIS, сверточные нейронные сети.

1. ВВЕДЕНИЕ

В современном мире дорожные знаки являются обязательным атрибутом при организации дорожного движения. Они информируют водителей об опасных участках дороги, указывают направление движения, не разрешают или дают право проезда, обязывают снижать скорость и выполняют множество других полезных задач.

Плагин «Дорожные знаки» необходим для оптимизации работы отдела управления дорожным движением группы компаний ИнтелТранС, основной задачей которого является размещение дорожных знаков на карте интеллектуальной транспортной геоинформационной системы ITSGIS.

В ходе исследовательской работы будет продемонстрировано полное описание работоспособности плагина «Дорожные знаки», а также рассказано о реализации данной технологии в интеллектуальной транспортной геоинформационной системе ITSGIS.

2. РАЗРАБОТКА ПРОГРАММНОГО КОМПЛЕКСА И ЭКСПЕРИМЕНТАЛЬНЫЕ ИССЛЕДОВАНИЯ

A. Подготовка выборки для нейронной сети

В качестве данных для исследований и разработок использовалась выборка из 51883 изображений дорожных знаков. Изначально изображения имели разное разрешение, но при предобработке полутоновых изображений данные были приведены к единому размеру 50×50 пикселей [1].

Выборка состояла из изображений, разделенных на 43 класса по названию дорожного знака.

Так как изображения дорожных знаков в выборке также имеют разные цветовые переходы и получены с разных устройств, необходимо сделать предобработку для повышения точности алгоритма [2].

Изображение 50×50 из видеопотока извлекает кадр и сегментирует таким образом, чтобы дорожный знак максимально занимал изображение данной размерности. Данная процедура называется обнаружением. Обнаруженный признак масштабируется, сегментируется и подается на вход нейронной сети.

Чтобы изменить цветовое пространство каждого изображения в образце, необходимо применить метод компьютерного зрения для преобразования цветного изображения в градации серого.

B. Распознавание дорожных знаков с помощью сверточной нейронной сети

Разработана сверточная нейронная сеть, состоящая из пяти сверточных, пяти субдискретизирующих и двух полносвязных слоёв.

Для обучения сверточной нейронной сети было использовано 80% изображений от всей выборки, то есть 41506 изображений. Для контроля и оценки эффективности сверточной нейронной сети были взяты оставшиеся 20% изображений от всей выборки, то есть 10377 изображений [3].

Нейронная сеть выводит вероятность того, что изображение принадлежит каждому из 43 классов, после чего вычисляется максимальная вероятность из трех вероятностей. В результате изображение относится к классу с наибольшей вероятностью. Для реализованной сверточной нейронной сети значение средней точности ранжирования меток составило 98.21%.

B. Описание реализованной программы

В разрабатываемой программе на первом этапе использовались библиотеки OpenCV и PIL. Данные библиотеки необходимы для работы с цветовым пространством изображения, с его размером, а также позволяют читать, отображать и записывать изображения. Затем с помощью библиотек глубокого обучения Keras и Tensorflow была создана и обучена модель сверточной нейронной сети [4].

При разработке программы также использовалась библиотека NumPy, которая помогала обрабатывать

многомерные массивы и создавать высокоуровневые математические функции.

Библиотека Matplotlib позволяет визуализировать зависимость точности распознавания с увеличением эпохи, а также график потерь с увеличением эпохи.

Г. Экспериментальные исследования

Для каждого тестового изображения решаем задачу классификации и находим точный процент правильной классификации. Средний показатель качества распознавания на синтетической тестовой выборке составляет 98.21%, что является достоверным показателем [5].

Необходимо использовать алгоритм сегментации и экспериментально оценить качество работы на реальном видеопотоке.

После прохождения алгоритма сегментации изображения дорожных знаков масштабируются и подаются на вход обученной нейронной сети [6].

Средний показатель качества распознавания для изображений, вырезанных из реального видеопотока, составляет 97.64%, что является достоверным показателем.

На основании проведенных экспериментов можно сделать вывод, что процент классификации изображений в реальном времени немного меньше, чем у тестовых изображений. Это различие объясняется тем, что изображения, обученные в нейронной сети, имеют то же происхождение, что и тестовые синтезированные изображения. В свою очередь, реальные изображения не похожи на изображения в обучающей выборке.

3. ИНТЕГРАЦИЯ В ITSGIS

А. Интеллектуальная транспортная геоинформационная система ITSGIS

«ITSGIS» — интеллектуальная транспортная геоинформационная система с многослойной электронной картой города [7].

«ITSGIS» позволяет: визуализировать карты распространенных форматов, редактировать карту с помощью основных графических функций, гибко настраивать пользовательский интерфейс, разрабатывать различные плагины.

Плагины расширяют функциональные возможности и позволяют работать со специализированными геообъектами - точечной, линейной и полигональной геометрией на электронной карте с присоединенной семантикой.

Б. Плагин «Дорожные знаки»

«ITSGIS. Дорожные знаки» — плагин, включающий в себя электронную карту и автоматизированную информационную систему, обеспечивающую работу с геообъектами — дорожными знаками.

На рисунке 1 показан наглядный пример распознавания дорожных знаков в интеллектуальной транспортной геоинформационной системе ITSGIS.

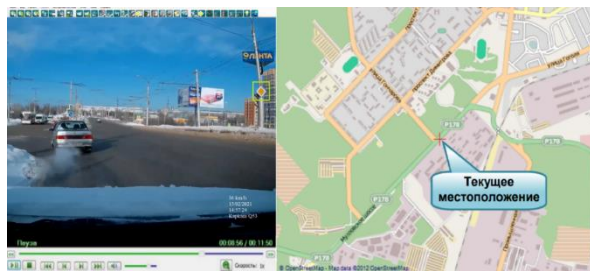


Рис. 1. Демонстрация обнаружения дорожных знаков в ITSGIS

Данная система решает следующие задачи: размещение дорожных знаков на электронной карте города с проверкой разрешения на установку геообъекта, редактирование объекта, изменение направления установленного объекта и его местоположения, проверка наличия установленных объектов на дорожной сети, создание сводных листов и отчетов по установленным объектам.

4. ЗАКЛЮЧЕНИЕ

В результате исследовательской работы была изучена литература по тематике, связанной с нейронными сетями и обнаружением дорожных знаков. В ходе исследования были выявлены наиболее эффективные методы, на основе которых сформирован перечень трудностей и проблем обнаружения образов на изображении.

Данная технология реализована в плагине «Дорожные знаки» в среде интеллектуальной транспортной геоинформационной системы «ITSGIS». В журнале IT&Transport опубликовано 2 статьи по теме исследования, а также результаты научной деятельности были представлены на международной научно-технической конференции PIT 2021. В этих работах исследовались различные методы распознавания изображений, где оказались сверточные нейронные сети быть лучшим.

На внедряемый плагин получено свидетельство о государственной регистрации программы для ЭВМ.

В работе подробно описаны этапы, позволяющие добиться высокой производительности при обнаружении дорожных знаков на максимальной скорости, не требуя при этом больших вычислительных мощностей. Реализованная программа способна распознавать 43 различных дорожных знака с вероятностью 97.64%.

Эту технологию можно использовать даже на недорогих устройствах, часто встречающихся в сельской местности.

БЛАГОДАРНОСТЬ

Исследование выполнено при поддержке группы компаний ИнтелТранС в рамках совместной разработки системы видеораспознавания дорожных знаков.

ЛИТЕРАТУРА

- [1] Haikin, S. Neural networks: full course / S. Haikin. – М.: Williams Publishing House, 2006. – 1104 p.
- [2] Klepikov, N. Road sign recognition methods in the ITSGIS environment / N. Klepikov, T. Mikheeva, S. Mikheev // Advanced Information Technologies (PIT). – Samara, 2021. – P. 470-476.
- [3] Vasyugova, S.A. Driving in emergencies with use of system of the help to the driver / S.A. Vasyugova, A.B. Nikolaev // International journal of arts & sciences. – 2016. – № 4. – P. 90-101.

- [4] Klepikov, N. Research of traffic sign recognition algorithms / N. Klepikov, T. Mikheeva // IT & Transport. – 2021. – № 14. – P. 46-57.
- [5] Zhang, Z. CNN Optimization and its application in traffic signs recognition based on GRA / J. Residuals // Science Technology. – 2016. – № 3. – P. 23-35.
- [6] Demir, H. Quantum dot integrated LEDs using photonic and excitonic color conversion / H. Demir, S. Nizamoglu, T. Erdem // Nano today. – 2011. – № 6. – P. 632-647.
- [7] Hum, Y.C. Multiobjective bi-histogram equalization for image contrast enhancement / Y.C. Hum, K.W. Lai, M. Salim, I. Maheza // Complexity. – 2014. – № 20. – P. 22-36.

Адаптация нейросетевых алгоритмов реконструкции изображений для различных условий экспозиции

В.В. Евдокимова

Институт систем обработки изображений – филиал ФНИЦ «Кристаллография и фотоника» РАН
Самарский национальный исследовательский университет им. академика С.П. Королева
Самара, Россия
vkutikova94@gmail.com

Аннотация—В данной работе проведено исследование влияния экспозиционного числа изображения на качество нейросетевой реконструкции. В рамках экспериментальных исследований показано, что при повышении значения экспозиции ухудшается визуальное качество изображения при нейросетевой реконструкции. Для борьбы с искажениями предлагается использовать предварительную обработку обучающей выборки, заключающуюся в программном изменении значения экспозиции изображений.

Ключевые слова— нейросетевая реконструкция, дифракционная оптика, дифракционные линзы, обработка изображений.

1. ВВЕДЕНИЕ

В настоящее время все больше исследований проводится в области реконструкции изображений, полученных на основе плоских оптических элементов [1 - 5]. В работах [1-4] предлагается нейросетевой подход к реконструкции изображений, который обеспечивает высокий уровень PSNR на изображениях из тестовой выборки. Обучающая и тестовая выборки в данных работах формируются путем съемки изображений с экрана монитора. Однако при обработке изображений реальных сцен появляются артефакты. На Рис. 1 представлен пример реконструкции реального изображения. На Рис. 1б заметны искажения в виде желтых полос и мозаики. В данной работе предполагается, что причиной артефактов может быть съемка обучающего набора данных с фиксированной экспозицией.

Целью данной работы является исследование влияния значения экспозиции изображения на визуальное качество изображения при нейросетевой реконструкции.

Для проверки влияния уровня экспозиции на визуальное качество реконструкции, в том числе на появление артефактов, была проведена съемка одной сцены изображения с разным значением выдержки. Проведенный эксперимент показал, что чем больше выдержка, соответственно больше и значение экспозиции, тем больше артефактов появляется на изображении. Таким образом, в работе предлагается программно имитировать изменение экспозиционного числа обучающей выборки и исследовать качество реконструкции на основе обученной нейронной сети UNet, предложенной в работе [1].

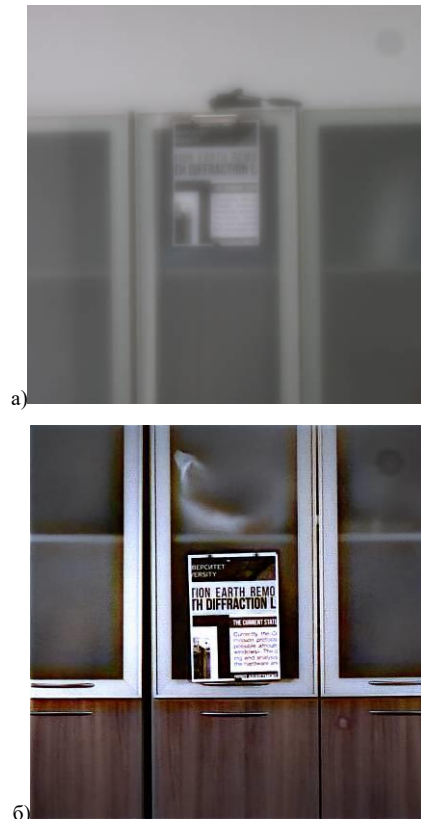


Рис. 1. Пример появления артефактов при нейросетевой реконструкции изображения: а) изображение, полученное на основе МДЛ; б) восстановленное изображение

2. МЕТОДЫ РЕКОНСТРУКЦИИ И ПРЕДВАРИТЕЛЬНОЙ ОБРАБОТКИ ИЗОБРАЖЕНИЙ

В данной работе используется модификация классической архитектуры U-Net [6], предложенная в работе [1]. Для повышения и понижения значения экспозиции обучающей выборки в данной работе используется алгоритм улучшения визуального качества изображений, снятых в условиях слабой освещенности, предложенный в работе [7]. На Рис. 2 представлен пример, демонстрирующий уменьшение количества артефактов на восстановленном изображении при программном уменьшении экспозиционного числа изображения с помощью алгоритма, предложенного в [7]. Обработка изображения проводилась сетью, обученной на наборе данных, сформированном в рамках работы [1]. Съемка набора данных проводилась с фиксированным значением экспозиции. Предварительная обработка

обучающего набора не проводилась. Видно, что программное уменьшение значения экспозиции (Рис. 1б) способствует уменьшению количества артефактов при нейросетевой реконструкции (Рис. 1г).

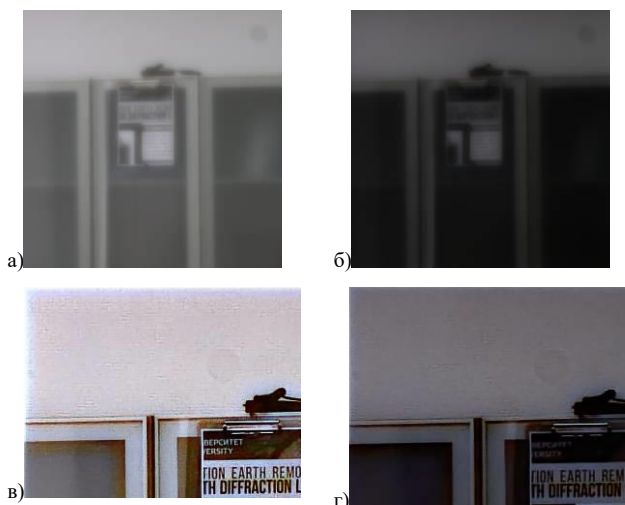


Рис. 2. Пример уменьшения артефактов при программном уменьшении значения экспозиции: а) изображение, полученное на основе МДЛ; б) результат программного уменьшения значения экспозиции; в) результат нейросетевой реконструкции исходного изображения; г) результат нейросетевой реконструкции изображения после уменьшения значения экспозиции

3. РЕЗУЛЬТАТЫ ЭКСПЕРИМЕНТАЛЬНОГО ИССЛЕДОВАНИЯ

Для проведения исследования сформированы 3 обучающих набора данных на основе набора данных, предложенного в работе [1]. Первый набор сформирован путем программного повышения значения экспозиции исходного набора данных, второй – путем понижения значения экспозиции, а третий смешанный набор включает по 1/3 исходных данных, а также данных с повышенной и пониженной экспозицией. На Рис. 3 представлены фрагменты результатов нейросетевой обработки изображения, указанного на Рис. 1а. Обработка проводилась моделями, обученными на трех сформированных и на исходном наборах данных.

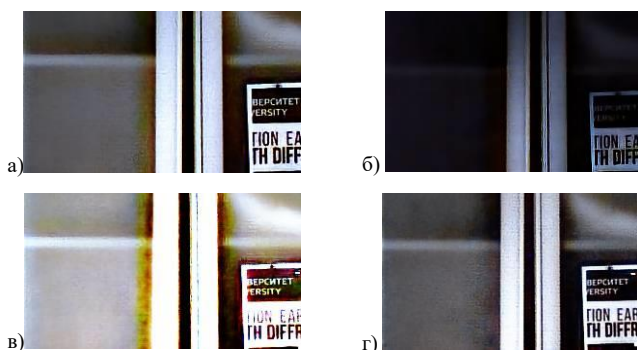


Рис. 3. Результаты нейросетевой обработки при программном изменении экспозиции обучающей выборки: а) обучение сети на исходных данных, фрагмент изображения на Рис. 1б; б) обучение на наборе с программным повышением экспозиции; в) обучение на наборе с программным понижением экспозиции; г) обучение на смешанном наборе данных

4. ЗАКЛЮЧЕНИЕ

В рамках проведенного исследования сделаны следующие выводы: экспозиционное число входного изображения влияет на визуальное качество реконструкции (цвета, количество артефактов), чем выше экспозиция, тем больше артефактов при реконструкции появляется; программное изменение значения экспозиции обучающей выборки позволяет улучшить визуальное качество изображения, а именно уменьшить искажение цветов изображения (желтые полосы на изображении исчезают); обучающая выборка, включающая изображения с разными значениями экспозиции способствует повышению визуального качества изображения. В рамках продолжения описанного исследования планируется сформировать новый обучающий набор, снятый в режиме автоэкспозиции.

БЛАГОДАРНОСТИ

Работа выполнена при поддержке Российского научного фонда № 22-19-00364.

ЛИТЕРАТУРА

- [1] Евдокимова, В.В. Нейросетевая реконструкция видеопотока в дифракционных оптических системах массового производства / В.В. Евдокимова, М.В. Петров, М.А. Клоева, Е.Ю. Зыбин, В.В. Косьянчук, И.Б. Мищенко, В.М. Новиков, Н.И. Сельвесюк, Е.И. Ершов, Н.А. Ивлиев, Р.В. Скиданов, Н.Л. Казанский, А.В. Никоноров // Компьютерная оптика. – 2021. – Т. 45, № 1. – С. 130-141. DOI: 10.18287/2412-6179-CO-834.
- [2] Nikonorov, A. Deep learning-based imaging using single-lens and multi-aperture diffractive optical systems / A. Nikonorov, V. Evdokimova, M. Petrov, P. Yakimov, S. Bibikov, Y. Yuzifovich, R. Skidanov, N. Kazanskiy // Learning for computational imaging. ICCV, 2019.
- [3] Nikonorov, A. Fresnel lens imaging with post-capture image processing / A. Nikonorov, R. Skidanov, V. Fursov, M. Petrov, S. Bibikov, Y. Yuzifovich // Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW). – 2015. – P. 33-41.
- [4] Peng, Y. Learned Large Field-of-View Imaging with Thin-Plate Optics / Y. Peng, Q. Sun, X. Dun, G. Wetzstein, F. Heide // ACM Trans. Graph. – 2019. – Vol. 38. – P. 219.
- [5] Peng, Y. Computational imaging using lightweight diffractive-refractive optics / Y. Peng, Q. Fu, H. Amata, S. Su, F. Heide, W. Heidrich // Opt. Express. – 2015. – Vol. 23 – P. 31393-31407.
- [6] Ronneberger, O. U-Net: Convolutional Networks for Biomedical Image Segmentation / O. Ronneberger, P. Fischer, T. Brox // Medical Image Computing and Computer-Assisted Intervention. Lecture Notes in Computer Science. – 2015. – Vol. 9351. – P. 234-241. DOI: 10.1007/978-3-319-24574-4_28.
- [7] Ying, Z. A new low-light image enhancement algorithm using camera response model / Z. Ying, G. Li, Y. Ren, R. Wang, W. Wang // ICCV. – 2017. – P. 3015-3022.

Разработка системы компьютерного зрения для детектирования дефектов на внешней поверхности трубы

Д.В. Шустанов

Самарский национальный исследовательский университет
им. академика С.П. Королева
Самара, Россия
den.shustanov@gmail.com

П.Ю. Якимов

Самарский национальный исследовательский университет
им. академика С.П. Королева
Самара, Россия
pavel.y.yakimov@gmail.com

Аннотация—В настоящей работе описывается разработка автоматизированной системы для контроля качества внешней поверхности трубы. Описывается метод обработки данных лазерного сканирования: оценка диаметра трубы, измерение отклонения точек поверхности от идеального цилиндра, разворачивание трёхмерной поверхности, сшивка в единую сканограмму трубы, детектирование и определение геометрических параметров дефектов.

Ключевые слова— лазерное сканирование, облако точек, метод наименьших квадратов, сшивание изображений, кластеризация, DBSCAN.

1. ВВЕДЕНИЕ

В настоящее время контроль качества труб осуществляется посредством ручного визуально-измерительного контроля. Вследствие этого результаты оценки могут быть искажены человеческим фактором.

Перспективы выявления дефектов связываются с автоматизацией и стандартизацией процесса контроля качества. Поэтому, автоматизация визуально-измерительного контроля внешних поверхностей труб является актуальной задачей.

Целью данной работы является разработка автоматизированной системы, которая позволит по данным полученным с лазерных профилометров, движущихся вдоль трубы, составить развёртку внешней поверхности трубы, определить наличие на ней дефектов и рассчитать их геометрические параметры.

2. СОСТАВЛЕНИЕ СКАНОГРАММЫ ВНЕШНЕЙ ПОВЕРХНОСТИ ТРУБЫ

Структура программной части системы приведена на Рис. 1 в виде модулей, работающих последовательно.

После захвата поток профилей поступает на модуль интерпретации данных, который производит сшивку профилей в 3D-модель части поверхности трубы. Сектора, поступают в модуль предварительной обработки, в результате чего формируется карта глубин сектора, которая затем поступает на вход модуля сшивки сканограммы. Далее производится детектирование и оценка параметров дефектов по полученной сканограмме.

Синхронизация съёмки профилей происходит с помощью импульсов энкодера, определяющего положение профилометра относительно трубы. По номеру синхроимпульса рассчитывается положение профиля. Все полученные профили размещаются в одном координатном пространстве. Методом наименьших квадратов оценивается окружность для каждого профиля.

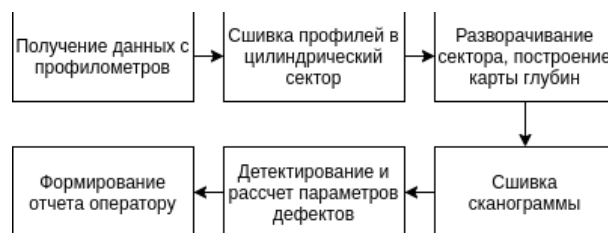


Рис. 1. Структура программной части системы

Далее производится разворачивание профиля — переход от системы координат облака точек в систему координат относительно поверхности сектора. Карты глубин для каждого сектора сшиваются в общую карту поверхности трубы путем сдвига каждой точки сектора вдоль оси на значение энкодера, определяющего поворот трубы. Полученная карта глубин преобразуется в сканограмму - rgb-изображение, цвет каждой точки которого рассчитывается по её глубине с использованием цветовой схемы (Рис. 2).

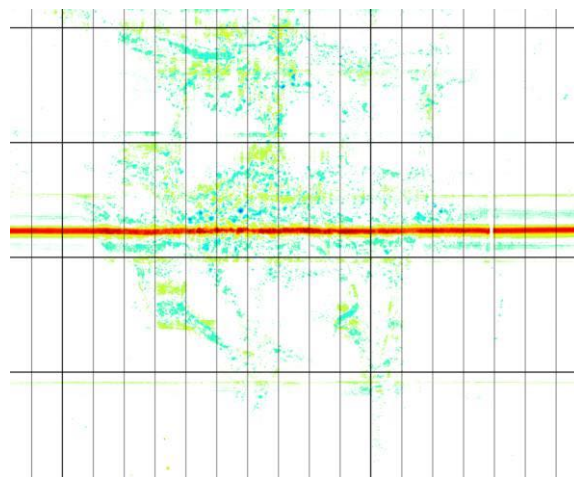


Рис. 2. Часть сканограммы наружной поверхности трубы

3. ДЕТЕКТИРОВАНИЕ ДЕФЕКТОВ

Детектирование дефектов производится путём сегментации изображения и кластеризации мелких дефектов для получения крупных дефектных областей. По карте глубин определяются геометрические параметры областей — максимальная глубина и расстояния от продольного шва и торца трубы.

4. ЭКСПЕРИМЕНТАЛЬНЫЕ ИССЛЕДОВАНИЯ

На Таблице I представлены параметры дефектов трубы, измеренные методами визуально-измерительного контроля и разработанной системой. Погрешность

измерений системой составили 0.58 мм для диаметра и 0.38 мм для глубины дефекта.

ТАБЛИЦА I. ПАРАМЕТРЫ ДЕФЕКТОВ, ИЗМЕРЕННЫЕ ВИК И СИСТЕМОЙ

Дефект	Диаметр			Глубина		
	Измеренный ВИК, мм	Измеренный системой, мм	Погрешность, мм	Измеренная ВИК, мм	Измеренная системой, мм	Погрешность, мм
1	5.6	6.32	0,76	2,6	2,69	0,09
2	5.4	5.32	0,08	2,7	3,32	0,62
3	5.6	5.69	0,09	2,8	2,39	0,41
4	6.0	6.0	0	1,8	1,79	0,01
5	6.1	5.24	0,86	3,0	3,07	0,07

5. ЗАКЛЮЧЕНИЕ

В результате проделанной работы была описана программная часть разработанной системы. Реализованы и протестированы программные модули для шивки и развёртки облака точек, детектирования и определения параметров дефектов.

БЛАГОДАРНОСТИ

Выражаем благодарность ООО "Газпром трансгаз Екатеринбург" за финансирование проекта, ООО НПЦ «Самара» за проектирование и создание прототипа установки для лазерного сканирования трубы.

ЛИТЕРАТУРА

- [1] Zhang, D. Automatic pavement defect detection using 3D laser profiling technology / D. Zhang, Q. Zou, H. Lin, X. Xu, L. He, R. Gui, Q. Li // Automation in Construction. – 2018. – Vol. 96. – P. 350-365.
- [2] Valero, E. Automated defect detection and classification in ashlar masonry walls using machine learning / E. Valero, A. Forsterb, F. Boschéa, E. Hyslopс // Automation in Construction. – 2019. – Vol. 106.
- [3] Park, H.-S. Development of inspection system for defect detection in precast parts using laser scanning data / H.-S. Park, T.U. Mani // Procedia Engineering. – 2014. – Vol. 69. – P. 931-936.
- [4] Jovancevic, I. 3D point cloud analysis for detection and characterization of defects on airplane exterior surface / I. Jovancevic, H.-H. Pham, J.-J. Orteu // Journal of Nondestructive Evaluation. – 2017. – Vol. 36. – P. 74.



Министерство образования
и науки Самарской области



САМАРСКИЙ УНИВЕРСИТЕТ
SAMARA UNIVERSITY

Самарский университет

ИСОИ  **ИПСИ**

Институт систем обработки изображений РАН



Institute of Electrical
and Electronics Engineers



Photonics



ФОТОНИКА



Компьютерная оптика