

DSPA:

Вопросы применения цифровой обработки сигналов

№4

2022

СОДЕРЖАНИЕ

Ерохин С.Д., Петухов А.Н., Пилюгин П.Л. ПРОЦЕССЫ АДАПТАЦИИ В АСИМПТОТИЧЕСКОМ УПРАВЛЕНИИ БЕЗОПАСНОСТЬЮ	4
Медведев Е.Ю., Воронова Л.И. НЕЙРОСЕТЕВОЕ РАСПОЗНАВАНИЕ "COVID-19" НА ОСНОВЕ РЕНТГЕНОВСКИХ ИЗОБРАЖЕНИЙ ГРУДНОЙ КЛЕТКИ С ПОМОЩЬЮ АЛГОРИТМА МАШИННОГО ОБУЧЕНИЯ VGG-16	9
Скрябин В.И., Белов Н.В. ПРИМЕНЕНИЕ НЕЙРОННЫХ СЕТЕЙ В ОБЛАСТИ РАСПОЗНАВАНИЕ ЖЕСТОВ	17
Харшиладзе Д.Г. НЕЙРОСЕТЕВОЕ РАСПОЗНАВАНИЕ ПОЛА И ВОЗРАСТА ПО ФОТОГРАФИИ ЛИЦА	24
Дергачёв Д.А., Тимагин Д.А., Боликов С.А. ИСПОЛЬЗОВАНИЕ ФОТОГРАММЕТРИИ ДЛЯ СКАНИРОВАНИЯ ОБЪЕКТОВ И ПОСЛЕДУЮЩЕГО ИСПОЛЬЗОВАНИЯ В САПР	30
Ашурбеков С.А., Матвиенко Д.А., Семкина Е.В., Суфиянов Р.Ш. ИНФРАКРАСНЫЙ ЛАЗЕР НА ГРАНАТЕ С НАКАЧКОЙ ТРЕМЯ ЛАЗЕРНЫМИ ДИОДАМИ ДЛЯ МАРКИРОВКИ ИЗДЕЛИЙ АВТОМОБИЛЬНОЙ ПРОМЫШЛЕННОСТИ	33
Ашурбеков С.А., Чубова-Белошникова М.М., Воеводин Д.Ю., Суфиянов Р.Ш. ОДНОЧАСТОТНЫЙ ЛАЗЕР НА КРИСТАЛЛЕ ВАНАДАТА С НАКАЧКОЙ ЛАЗЕРНЫМ ДИОДОМ НА ГАРМОНИКАХ ДЛЯ СКАНИРОВАНИЯ АВТОМОБИЛЬНЫХ ДОРОГ И ЗДАНИЙ ПРИ ПРОЕКТИРОВАНИИ И РЕМОНТЕ	36

ПРОЦЕССЫ АДАПТАЦИИ В АСИМПТОТИЧЕСКОМ УПРАВЛЕНИИ БЕЗОПАСНОСТЬЮ

Ерохин Сергей Дмитриевич,

Московский технический университет связи и информатики, ректор, к.т.н., Москва, Россия
esd@mtusi.ru

Петухов Андрей Николаевич,

*Московский технический университет связи и информатики, начальник отдела, к.т.н.,
Москва, Россия*
anpetukhov@yandex.ru

Пилюгин Павел Львович

*Московский технический университет связи и информатики, старший научный сотрудник, к.т.н.,
Москва, Россия*
ppl@mail.ru

Аннотация

Определяются функции адаптивного контура архитектуры асимптотического управления информационной безопасностью. Выделяются направления адаптации, обеспечивающие полнофункциональную реакцию на события безопасности, компьютерную атаку или инцидент, устанавливается роль ретроспективного и эталонного потока событий безопасности, показаны свойства алгоритмов детектирования атак, обеспечивающих адекватность и монотонность транзакций адаптации. Предлагается использование концепции «Общих критериев» для асимптотического управления политиками безопасности.

Ключевые слова

Информационная безопасность, асимптотическое управление, политики безопасности, адаптация, признаковое пространство, детектирование атак.

Введение

В архитектуре асимптотического управления [1,2] источником знаний о факторах опасности является мониторинг потока событий безопасности. Анализ этого потока наряду с использованием сведений о новых угрозах или уязвимостях от органов централизованного управления выполняется как для индикации инцидента, так и для детектирования компьютерной атаки.

Инцидент – это специальное состояние объекта защиты, эти состояния определены и составляют множество критичности. Компьютерная атака – это активность субъекта, которая может привести к возникновению инцидента. Сигналом для инициализации процессов адаптации является детектирование атаки или индикация/прогноз инцидента.

Возможны следующие варианты:

- был зафиксирован инцидент, но атака не детектирована – корректируются процедуры детектирования, и изменяются настройки средств защиты;
- была детектирована атака, но прогнозирование инцидента дало отрицательный результат - коррекции могут подвергаться процедуры детектирования;
- была детектирована атака, и прогнозирование инцидента дало положительный результат - изменяются настройки средств защиты и могут корректироваться процедуры детектирования.

Адаптация признакового пространства

Корректировка процедур детектирования наряду с алгоритмическими изменениями включает в себя изменения признакового пространства. Таким образом, необходимыми являются, три направления адаптации:

- изменение признакового пространства мониторинга;
- изменение алгоритмов детектирования на основании анализа потоков событий безопасности

(в случае пропуска атаки) и на основании оценки эффективности отдельных процедур (в случае детектирования атаки);

- изменение политик безопасности и соответствующих настроек средств защиты.

Условием выполнения любой транзакции адаптации является развитие защищенности, т.е. невозможность повторения причины данной транзакции (адекватность изменений) при сохранении ранее накопленного защитного потенциала (монотонность изменений).

В контексте асимптотического управления задача адаптации признакового пространства имеет особенности. Целью является внесение изменений в действующий набор так, чтобы обеспечить развитие защищенности. Кроме того, события безопасности обладают внутренней структурой – схемой агрегирования [3], которая может подвергаться адаптации наравне с номенклатурой атрибутов мониторинга. Результаты, полезные для адаптации признакового пространства, получены в отношении сигнатурных методов детектирования атак в [4], где решается задача последовательного выявления различий защищенности в зависимости от состава сигнатур.

Основанием выбора признаков из некоторого множества могут быть корреляция признаков или интенсивность ошибок детектирования. Для оценки могут использоваться парные коэффициенты корреляции. Если атака детектируется композицией из нескольких событий безопасности, то показателем связности может служить коэффициент множественной корреляции [4].

Использование интенсивности ошибок детектирования может использовать фактор интенсивности и фактор опасности [5]. В случае совпадения стратегий оптимизации каждого из этих факторов, следование любой из стратегий обеспечивает рост защищенности на каждом шаге адаптации.

Адаптации признакового пространства инициируется идентификацией атаки и выбирает признаки из предопределенного списка событий безопасности. Процедура идентификация включает анализ причинно-следственных переходов объекта защиты, завершившихся инцидентом.

Оценка статистических характеристик является результатом анализа отличий ретроспективного потока событий от эталонного потока, который формируется в условиях, исключающих вредоносные воздействия. Ретроспективный поток включает события, непосредственно предшествующие инциденту и представляет собой образ причин, вызвавших инцидент.

Даже в самых обширных наборах признаков возможности выбора ограничены. Можно расширить диапазон потенциальных признаков, если привлечь процедуру генерации. Хотя состав терминов для описания структуры признаков ограничен, он позволяет сформировать больше структур, чем используется в заданном наборе. Структуры признаков формируются из терминов в соответствии с правилами, которые составляют грамматику, порождающую язык признаков. В [6] анализу подвергается набор признаков для обучающих корпусов данных KDD 99 [7]. Для формирования грамматики описан и применен генетический алгоритм грамматической эволюции. В результате формируются продукционные правила, обеспечивающие наряду с предопределенными признаками вывод целого ряда дополнительных.

Адаптация процедур детектирования компьютерных атак

Исходными данными для адаптации процедур детектирования является идентификация атаки в процессе расследования причин инцидента. Этот результат может быть представлен в эксплицитном виде сценария проведенной атаки или в имплицитной форме проявлений этой атаки. События безопасности имплицитной формы могут быть выражены как на базе действующего признакового пространства, так и за его пределами. Новые признаки вводятся автоматически, либо привлекаются решения по адаптации признакового пространства.

Для идентификации атаки в имплицитной форме может быть использован эталонный поток событий безопасности. Сравнение этого потока с потоком реальных событий позволяет выявить отклонения, которые являются проявлением атаки и дают основание для формирования новых сигнатур или внесения изменений в настройки выявления аномалий.

Если для сигнатурных алгоритмов детектирования атак возможно монотонное развитие защищенности при пошаговом расширении состава сигнатур (экстенсивное решение) [4], то для алгоритмов выявления аномалий такая схема адаптации неприменима. Это вызвано следующими причинами:

- алгоритмы выявления аномалий имеют уникальные схемы управления [8], и эти схемы не поддаются структурной унификации;

- изменения алгоритмов, направленные на детектирование новой атаки, может непредсказуемым образом изменить реагирование на те атаки, которые успешно детектировались ранее;
- Стремление преодолеть такие ограничения получило развитие в следующих направлениях:
- вариативная организация комплекса детектирования на базе нескольких алгоритмов;
 - мультиагентная организация комплекса детектирования с использованием агенто-интерфейсов, поддерживающих адаптацию;
 - использование машинного обучения и тестирования в процессе адаптации;
 - реализация моделей, предусматривающих самостоятельное накопление сведений о предыдущих транзакциях адаптации, например иммунных моделей.

Вариативная организация комплекса детектирования наряду с расширением диапазона детектируемых атак, обеспечивает поддержку процесса адаптации. Для произвольного набора действующих алгоритмов невозможно установить и гарантировать, что внесение изменений не нарушит детектирование ранее установленных атак. Поэтому множество действующих алгоритмов дополняется новым элементом. Если новый алгоритм изменяет признаковое пространство, то настройки уже действующих алгоритмов требуют проверки своей эффективности, особенно это важно при использовании частотных и иных статистических признаков.

Вариативная организация не ограничивает порядок выполнения отдельных алгоритмов (последовательный, параллельный, смешанный), но в любом случае предполагает управление этим порядком и формированием общего результата на базе реакций отдельных алгоритмов.

При реализации мультиагентной организации адаптации [9,10] каждый алгоритм детектирования снабжается специальным агенто-интерфейсом, который поддерживает изменение соответствующего алгоритма, интерпретируя унифицированное описание идентификации атаки.

Преимущество агенто-интерфейсов в том, что это универсальный инструмент для поддержки адаптации. Распространение этого подхода ограничено потому, что для сигнатурных методов интерпретация тривиальна, а для развитых алгоритмов выявления аномалий (вейвлет-анализ, фрактальный анализ и др.) она очень сложна. Кроме того, следует исключить алгоритмы без последующей идентификации, которые не требуют корректировки для детектирования новой атаки, а также, алгоритмы, предназначенные для детектирования только определенных атак и не подлежащие изменению. Таким образом, остается небольшой спектр решений, для которых применение мультиагентной организации адаптации эффективно.

Вторым ограничением является эффект «насыщения», который проявляется после выполнения нескольких транзакций адаптации. Начиная с какого-то момента агент-интерфейс показывает, что вносимые изменения снижают эффективность реагирования на ранее детектируемые атаки, и корректировка нецелесообразна.

Алгоритмы с машинным обучением [11] упрощают процедуру адаптации. При этом возрастает роль ретроспективного потока событий безопасности, который становится основным источником имплицитного знания о пропущенной атаке. Кроме того, снижается предсказуемость результата адаптации с точки зрения влияния на ранее детектируемые атаки. Для развития защищенности достаточно переобучить алгоритм, используя поток событий безопасности, непосредственно предшествующий инциденту. Ретроспективный поток содержит признаки необнаруженной атаки, и после такого обучения алгоритм приобретет способность детектировать атаки такого типа.

В случае эксплицитной формы идентификации атаки для обучения может быть использован метод моделирования признаков атаки [2]. В эталонный поток событий безопасности вносятся изменения, моделирующие признаки атаки, и формируется «зараженный» поток, который используется для обучения алгоритмов. Перед использованием моделирующих признаков, необходимо их локализовать для условий объекта защиты.

Модели иммунной системы [12, 13] являются развитием решений, включающих в себя компоненты поддержки процессов адаптации и использующих машинное обучение. Иммунная система представляет собой распределенный механизм защиты от атак, который использует структуры данных, имитирующие биологические объекты, и обладает свойствами не только адаптивности, но и иммунной памяти. Это позволяет сохранять возможность детектирования ранее идентифицированных атак.

Адаптация политик безопасности и настроек средств защиты

Каждый защитных механизм имеет специфические средства представления реализуемых ограничений, но общей формой является политика безопасности. Для обеспечения адекватности изменений необходимо выполнить отображение характеристик идентифицированной атаки на пространство ограничений. Дополнительным условием при определении ограничений является их бесконфликтность с функциональностью объекта защиты, т.е. вносимые изменения должны сохранять возможность выполнения объектом защиты своих функций с прежней эффективностью (бесконфликтность изменений).

В матрице техник и тактик атак MITRE [14] для многих техник указываются защитные механизмы, препятствующие реализации атаки с соответствующей техникой. Эти указания являются общими рекомендациями, требующими локализации к условиям объекта защиты. Кроме того, однозначная идентификация атаки в терминах тактик и техник MITRE в ряде случаев затруднена.

Среди нотаций политик безопасности можно выделить функциональные и информационные ограничения. В первом случае политикой специфицируются функции безопасности, а во втором – определяются легитимные и запрещенные информационные потоки. Одной из наиболее развитых платформ для описания и управления функциональными ограничениями является концепция «Общих критериев» [15] и ее базовая структура – профиль защиты, который содержит следующие согласованные компоненты:

- функциональные требования, которые специфицируют содержание функций безопасности.
- угрозы, которым противостоит комплекс защитных мер
- политики профиля – аксиоматические положения, определяющие внешние свойства функций безопасности
- предположения безопасности, устанавливающие сферу применения комплекса защитных мер

Функциональные требования соответствуют ограничениям, которым подвергается объект защиты. Это нотации его политики безопасности, они подвергаются изменениям в процессе адаптации. Угрозы соответствуют элементам проведения идентифицированной атаки. Политики профиля определяют условия бесконфликтности изменений, а предположения задают условия применения профиля.

Функциональные требования представлены в обобщенном виде структур-шаблонов, которые требуют заполнения и установления параметров. Анализируя идентифицированную атаку и выделив ее компоненты, можно сопоставить их с требованиями профиля и сформулировать необходимые дополнения в (настроить требование). Если подходящего требования в составе профиля нет, то привлекается дополнительное из predetermined множества [16] с последующей настройкой этого требования. Бесконфликтность оценивается путем согласования с политиками профиля каждого изменения. Весь комплекс изменений сопоставляется с предположениями для оценки и исключения дублирования.

Методически профиль защиты снабжен механизмами согласования функциональных требований с угрозами, политиками профиля и предположениями, что позволяет утверждать адекватность и бесконфликтность изменений в процессе адаптации профиля защиты. Для достижения монотонности изменений необходимо дополнить механизмы корректировки правилами, изолирующими влияние вносимых изменений на действие ранее сформулированных требований и добавить проверку условия монотонности при согласовании. Это требует сохранения сведений о непосредственно предшествующих коррекциях.

Политики безопасности с информационными ограничениями определяют фильтрацию и маршрутизацию передачи фрагментов данных, границы и формы распространения потоков информации. Поскольку атрибуты коррекции и идентифицированной атаки могут располагаться на разных уровнях модели OSI, политика должна быть многоуровневой и способной к трансляции между уровнями [17]. Могут быть решения, обеспечивающие формулировку и выполнение корректировок в терминах одного уровня, например, для межсетевого экрана это могут быть предикаты условий прохождения или отбрасывания пакетов [18]. В отношении управления политиками с информационными ограничениями трудно оценить условия монотонности и бесконфликтности.

Заключение

Асимптотическое управление безопасностью предусматривает три направления адаптации: признакового пространства, процедур детектирования атак и самих средств защиты. Условиями асимптотического управления являются адекватность и монотонность транзакций адаптации. Анализ отличий ретроспективного потока событий безопасности от эталонного образца позволяет выполнять адаптацию признакового пространства и процедур детектирования атак. Дополнение процедур детектирования функциями интерпретации идентифицированной атаки и внутренней памяти предыдущих транзакций обеспечивает адекватность и монотонность адаптации. Внесение изменений в политики безопасности дает возможность развивать защищенность, предупреждая возникновение и развитие инцидентов.

Литература

1. *Ерохин С.Д., Пилюгин П.Л., Петухов А.Н.* Принципы и задачи асимптотического управления безопасностью критических информационных инфраструктур // Т-Comm: Телекоммуникации и Транспорт, 2019, № 12, стр. 29-35.
2. *Ерохин С.Д., Петухов А.Н.* Архитектура асимптотического управления безопасностью критических информационных инфраструктур // DSPA: Вопросы применения цифровой обработки сигналов, №1, 2022. С. 18-31.
3. *Ерохин С.Д., Петухов А.Н.* Сбор и агрегирование событий безопасности // Технологии информационного общества. Сборник трудов XIV Международной отраслевой научно-технической конференции, Москва, 2020. С. 239-244.
4. *Sergey Erokhin, Andrey Petukhov, Pavel Pilyugin.* Comparison of information security systems for asymptotic information security management of critical information infrastructures (CII) // 28th FRUCT Conference. 23 января 2021.
5. *Ерохин С.Д., Петухов А.Н., Пилюгин П.Л.* Эффективность активного мониторинга событий сетевой безопасности // Электросвязь, 2020, № 2. С. 46-51.
6. *Dominic Wilson, Devindar Kaur.* Using Grammatical Evolution for Evolving Intrusion Detection Rules // Proceedings of the 5th WSEAS Int. Conference on Information Security and Privacy, Venice, Italy, November 20-22, 2006.
7. *Protic D.* Review of KDD // Cup '99, NLS-KDD and Kyoto 2006+ datasets. Military Nehical Courier. Vol. 66, issue 3. 20, pp. 580-596.
8. *Шелухин О.И.* Сетевые аномалии. Обнаружение, локализация, прогнозирование. М.: Горячая линия – Телеком, 2020.
9. *Никишова А.В.* Принципы функционирования многоагентной системы обнаружения атак // Известия ЮФУ. Технические науки. Тематический выпуск. «Информационная безопасность». Таганрог: Изд-во ТТИ ЮФУ, 2012, №12 (137). С. 28-33.
10. *Ляпустин А.Е.* Безопасность мультиагентной платформы // Программные системы и вычислительные методы. 2017. № 3. С. 16-24.
11. *Шелухин О.И., Ерохин С.И., Полковников М.В.* Технологии машинного обучения в сетевой безопасности. М.: Горячая линия – Телеком, 2021.
12. *Kim J., Bentley P.* An Artificial Immune Model for Network Intrusion Detection // <http://neuro.bstu.by/our/immune3.pdf>, дата доступа: 15 мая 2022 г.
13. *Васильев В.И., Котов В.Д.* Система обнаружения сетевых вторжений на основе механизмов иммунной модели // Известия ЮФУ. Технические науки. № 12(125). 2011. С. 180-189.
14. <https://attack.mitre.org/matrices/18/>
15. ГОСТ Р ИСО/МЭК 15408-1-2012 Информационная технология. Методы и средства обеспечения безопасности. Критерии оценки безопасности информационных технологий. Часть 1. Введение и общая модель
16. ГОСТ Р ИСО/МЭК 15408-2-2012 Информационная технология. Методы и средства обеспечения безопасности. Критерии оценки безопасности информационных технологий. Часть 2. Функциональные компоненты безопасности
17. *Ерохин С.Д., Пилюгин П.Л.* Многоуровневая политика безопасности сети передачи данных //DSPA: Вопросы применения цифровой обработки сигналов. №2. 2022. С. 13-23.
18. *Мьо Т.Т.* Формальная модель анализа ошибок конфигурирования межсетевых экранов // Наука и Образование. МГТУ им. Н.Э. Баумана, Электрон. журн. 2015. № 6. С. 305-328.

НЕЙРОСЕТЕВОЕ РАСПОЗНАВАНИЕ "COVID-19" НА ОСНОВЕ РЕНТГЕНОВСКИХ ИЗОБРАЖЕНИЙ ГРУДНОЙ КЛЕТКИ С ПОМОЩЬЮ АЛГОРИТМА МАШИННОГО ОБУЧЕНИЯ VGG-16

Медведев Евгений Юрьевич,

Московский Технический Университет Связи и Информатики, Москва, Россия
gn.medvedev@gmail.com

Воронова Лилия Ивановна,

Московский Технический Университет связи и Информатики, д.ф.-м.н., профессор, Москва, Россия
voronova.lilia@yandex.ru

Аннотация

В статье описывается процесс проектирования и реализации нейронной сети на основе сети VGG16, для нейросетевой классификации рентгеновских снимков грудной клетки здоровых пациентов и больных COVID-19. Рассматриваются популярные модели сверточных нейронных сетей, приведено их краткое математическое описание. Описан процесс подготовки, обучения и тестирования набора данных. На основе анализа полученных результатов сделано заключение о возможности дальнейшего применения разработанного авторами ресурса в сфере медицинской диагностики для проверки наличия повреждений в лёгких пациентов, вызванных SARS-CoV-2 с высокой точностью.

Ключевые слова

Компьютерное зрение, классификация, нейронная сеть, VGG-16, машинное обучение, COVID-19, рентгеновские снимки.

Введение

Быстрое распространение по всему миру коронавируса SARS-CoV-2 в 2020 году [1, 2], в значительной степени связано с отсутствием технических возможностей для проведения массовой, быстрой и функциональной диагностики потенциальных пациентов. Обладающие на данный момент, наибольшей достоверностью методы тестирования на основе полимеразной цепной реакции (ПЦР) требуют, как правило, несколько часов на получение корректного результата, с учётом требований к профессионализму медперсонала и соблюдения им всей технологии анализа. В связи с этим, разработка надёжных, быстрых и легко масштабируемых систем обнаружения COVID-19 необходима для сокращения пандемии этого заболевания.

Внедрение методов интеллектуального анализа данных в сферу первичного анализа и интерпретации результатов рентгенологической диагностики, является в сложившейся в мире ситуации весьма перспективной идеей [3]. Первичный анализ рентгеновского снимка проводится вручную, что не может не сказаться на трудоёмкости и точности полученных результатов. На мой взгляд, именно поэтому необходимо переложить эту задачу на программное обеспечение, использующее алгоритмы машинного обучения. Исследования в этом направлении ведутся и на кафедре ИСУиА МГУСИ [4, 5, 6].

В данной статье описывается процесс разработки классификатора рентгеновских изображений, на базе нейросети, построенной с помощью предварительно обученной модели VGG16 и осуществляющей классификацию набора данных, состоящего из небольшого числа РИ.

Теоретическое описание свёрточной нейронной сети. Популярные модели СНС

1. Проект ImageNet и его влияние на развитие машинного обучения

В настоящее время, в мире одним из основных проектов, занимающихся развитием систем машинного обучения, является ImageNet. Данный проект занимается созданием и поддержкой массивной базы аннотированных изображений, которые предназначены для отработки и тестирования на ней методов распознавания образов и машинного зрения [7].

По состоянию на декабрь 2021 года, в ней было записано около пятнадцати миллионов URL адресов с изображениями, которые прошли через ручную модерацию и аннотацию. В аннотации перечисляются объекты, попавшие на изображение и их координаты. Также с 2010 года ведется проект ImageNet Large Scale Visual Recognition Challenge (Кампания по широкомасштабному распознаванию образов в ImageNet), в рамках которого проводятся соревнования в классификации и распознавании объектов и сцен в базе данных ImageNet [8].

Большинство самых популярных моделей свёрточных нейросетей являются победителями данного соревнования разных годов.

2. Теоретическое описание свёрточной нейронной сети

Свёрточная нейронная сеть (СНС) – это одна из вариаций многослойного перцептрона. Такие сети состояются из свёрточных слоёв, слоёв дискретизации и полносвязных слоёв. СНС чаще всего используется в компьютерном зрении.

Получая на входе серию изображений или видео из реального мира, используя приёмы свёртки, система ИИ учится автоматически извлекать свойства из этих входящих данных для выполнения конкретной задачи, например аутентификации лица, семантической сегментации изображений, или классификации изображений. Механизм свертки представлен на рисунке 1 [9].

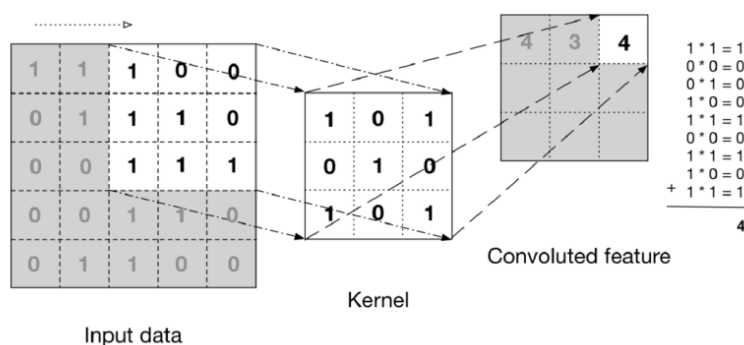


Рис. 1. Механизм операции свёртки

Свёртка описывается следующей формулой

$$y_n^l = f_l \left(\sum_{m \in v_n^{l-1}} \{w_{m,n}^{l-1}\} \otimes \{z_n^{l-1}(i,j)\} + b_n^l \right),$$

где l – номер текущего слоя ($l = 1, 3, \dots, 2a+1$), y_n^l – n -ая карта признаков на l -ом слое, f_l функция активации l -го слоя, v_n^{l-1} – список всех уровней слоя ($l-1$), которые соединяются с картой признаков n слоя l , $\{w_{m,n}^{l-1}(i,j)\}$ – свертка, применяемая к карте признаков m слоя ($l-1$), на слое l с картой признаков n , b_n^l – пороговые значения, присоединяемые к карте признаков n на слое l .

Информация со свёрточных слоев поступает на слои субдискретизации или Pooling-слои, на которых происходит уменьшение размера матрицы и увеличение инвариантности к масштабу входного изображения. При выполнении данной операции определяется наличие признака на ограниченном пространстве. Субдискретизация выполняется по следующей формуле

$$y_n^l = f_l \left(\{z_n^{l-1}(i,j)\} \otimes \{w_{m,n}^l(i,j)\} + b_n^l \right),$$

где l – номер текущего слоя ($l = 2, 4, \dots, 2a$), $\{z_n^{l-1}(i,j)\}$ – сумма пикселей по блоку размером 2×2 .

Свёрточный и субдискретизирующий слои чередуются, пока карты признаков не преобразуются в вектор или скаляр. Далее сигнал отправляется на полносвязный слой, где происходит обработка данных и формируется результат [10].

3. Краткий обзор современных моделей свёрточных нейронных сетей

Используя разнообразные свёрточные фильтры, модели свёрточной нейронной сети могут фиксировать высокоуровневое представление входных данных, что делает их наиболее популярными для решения задач, связанных с компьютерным зрением – классификацией изображений и обнаружением объектов. Одними из самых известных моделей, для классификации изображений являются: ResNet, GoogleNet, AlexNet, и VGG16. Ниже кратко описаны три из четырёх самых популярных на данный момент моделей свёрточных нейросетей. Модель VGG16, на основе которой, будет создана и обучена нейросеть в данной статье, описана в параграфе 4.

ResNet

ResNet – модель свёрточной сети, использующая технику т.н. «ярлыков» для достижения точности на уровне человека с ошибкой ниже 5%, путём пропуска соединений или прыжков между слоями. Также «ярлыки» используются для решения проблемы исчезновения градиента в процессе обучения, позволяя обучать глубокие нейронные сети модели с более глубокой структурой [11].

AlexNet

AlexNet – модель свёрточной сети, созданная для классификации изображений и первая сеть, выигравшая конкурс ImageNet Challenge в 2012 году. Данная модель состоит из пяти свёрточных слоёв и трёх полносвязных слоёв. Таким образом AlexNet требуется 61 миллион весов и провести 724 миллиона операций совмещенного умножения-сложения (MAC) для классификации изображения размером 227 на 227 пикселя [12].

GoogLeNet

GoogLeNet - модель свёрточной сети, основанная на архитектуре Inception. Является победителем конкурса ImageNet Challenge в 2014 году. Данная модель состоит в общей сложности из 22 слоёв. Отличительной особенностью данной, СНС, является то, что она основана на архитектуре Inception, суть которой состоит в том, что вместо выбора одного единственного значения размера ядра СНС, используется сразу несколько вариантов с последующей конкатенацией результатов [13].

4. Обзор архитектуры VGG16

VGG16 — это простая и широко применяемая архитектура предварительно обученной свёрточной нейронной сети, которая вошла в топ-5, при тестировании на ImageNet в распознавания различных объектов на изображении. Архитектура VGG16 была разработана и представлена Кареном Симонианом и Эндрю Зиссерманом из Оксфордского университета в их статье «Very Deep Convolutional Networks for Large-Scale Image Recognition». VGG – это аббревиатура от Visual Geometry Group, группы исследователей из Оксфордского университета, разработавшая эту архитектуру, а «16» означает, что эта архитектура имеет 16 слоёв с весами. Эта сеть довольно большая и имеет около 138 миллионов параметров [14].

Основной особенностью этой свёрточной нейросети, стало то, что вместо большого количества гиперпараметров, в ней все операции свёртки выполняются с помощью фильтра размером 3 с шагом 1 и всегда одним и тем же заполнением, а все слои субдискретизации при помощи фильтра размером 2 с шагом 2. Такое расположение слоёв свёртки и субдискретизации последовательно выполняется во всей архитектуре. В конце он имеет два полносвязных слоя, за которыми следует softmax для вывода.

Проектирование и реализация свёрточной нейронной сети на основе модели VGG16

1. Описание набора данных разрабатываемой сети.

Набор данных представляет собой сто, вручную отобранных из большего репозитория-датасета, рентгеновских снимков грудной клетки различных пациентов, выложенного на сайте GitHub [15]. По 50 снимков здоровых людей и 50 снимков людей больных COVID-19. Структура папки данных, передаваемых в ImageDataGenerator представлена ниже на рисунке 2 и выглядит следующим образом

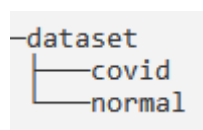


Рис. 2. Структура папки данных

Фактическое содержимое каждой из этих двух папок представляет собой 50 изображений рентгеновских снимков в формате jpeg разного качества, разрешения, чёткости и цвета. На рисунке 3 представлены по одному изображению из каждой папки.

Нейронная сеть построена на основе архитектуры свёрточной нейронной сети VGG16 с предварительно обученными в ImageNet, которая содержит 16 слоёв и без встроенного полносвязного слоя.



Рис. 3. Примеры изображений из набора данных

2. Архитектура разрабатываемой сети

Сеть состоит из двух частей:

Первая часть сети выделяет характерные признаки в изображении. Состоит из чередующихся каскадов свёртки и подвыборки. Сначала идут два каскада свёртка-свёртка-подвыборка, а затем три каскада свёртка-свёртка-свёртка-подвыборка. Размер узла свёртки во всех слоях 3*3. При подвыборке производится выбор максимального значения (MaxPooling) из квадрата размером 2*2.

Вторая часть отвечает за классификацию изображения по выделенным на предыдущем этапе признакам. Эта часть содержит пять полносвязных слоёв. В первых двух слоях 512 нейронов, в следующих двух слоях по 64 нейрона и в последнем слое два нейрона – по количеству классов изображений в наборе данных.

На вход сеть VGG16 получает изображение размером 224*224 пикселя, 3 канала цвета (красный, зеленый и синий). На входе сеть выдаёт вероятности в формате one hot encoding, что на изображении тот или иной класс объекта (болен человек ковид или нет).

Схема архитектуры нейросети из данной статьи представлена ниже на рисунке 4.

Layer (type)	Output Shape	Param #
input_1 (InputLayer)	[(None, 224, 224, 3)]	0
block1_conv1 (Conv2D)	(None, 224, 224, 64)	1792
block1_conv2 (Conv2D)	(None, 224, 224, 64)	36928
block1_pool (MaxPooling2D)	(None, 112, 112, 64)	0
block2_conv1 (Conv2D)	(None, 112, 112, 128)	73856
block2_conv2 (Conv2D)	(None, 112, 112, 128)	147584
block2_pool (MaxPooling2D)	(None, 56, 56, 128)	0
block3_conv1 (Conv2D)	(None, 56, 56, 256)	295168
block3_conv2 (Conv2D)	(None, 56, 56, 256)	590080
block3_conv3 (Conv2D)	(None, 56, 56, 256)	590080
block3_pool (MaxPooling2D)	(None, 28, 28, 256)	0
block4_conv1 (Conv2D)	(None, 28, 28, 512)	1180160
block4_conv2 (Conv2D)	(None, 28, 28, 512)	2359808
block4_conv3 (Conv2D)	(None, 28, 28, 512)	2359808
block4_pool (MaxPooling2D)	(None, 14, 14, 512)	0
block5_conv1 (Conv2D)	(None, 14, 14, 512)	2359808
block5_conv2 (Conv2D)	(None, 14, 14, 512)	2359808
block5_conv3 (Conv2D)	(None, 14, 14, 512)	2359808
block5_pool (MaxPooling2D)	(None, 7, 7, 512)	0
average_pooling2d (AveragePo	(None, 1, 1, 512)	0
flatten (Flatten)	(None, 512)	0
dense (Dense)	(None, 64)	32832
dropout (Dropout)	(None, 64)	0
dense_1 (Dense)	(None, 2)	130

Total params: 14,747,650		
Trainable params: 32,962		
Non-trainable params: 14,714,688		

Рис. 4. Сводное представление структуры нейронной сети

3. Описание процесса обучения и тестирования

Программа работает согласно блок-схеме, приведенной выше на рисунке 5. Сперва с помощью стандартного модуля Python 3 argparse читается папка dataset в которых содержится две подпапки covid и normal (см. рис. 2) в которых, в свою очередь содержится по пятьдесят рентгеновских снимков грудных клеток людей больных covid 19 и, соответственно, здоровых.

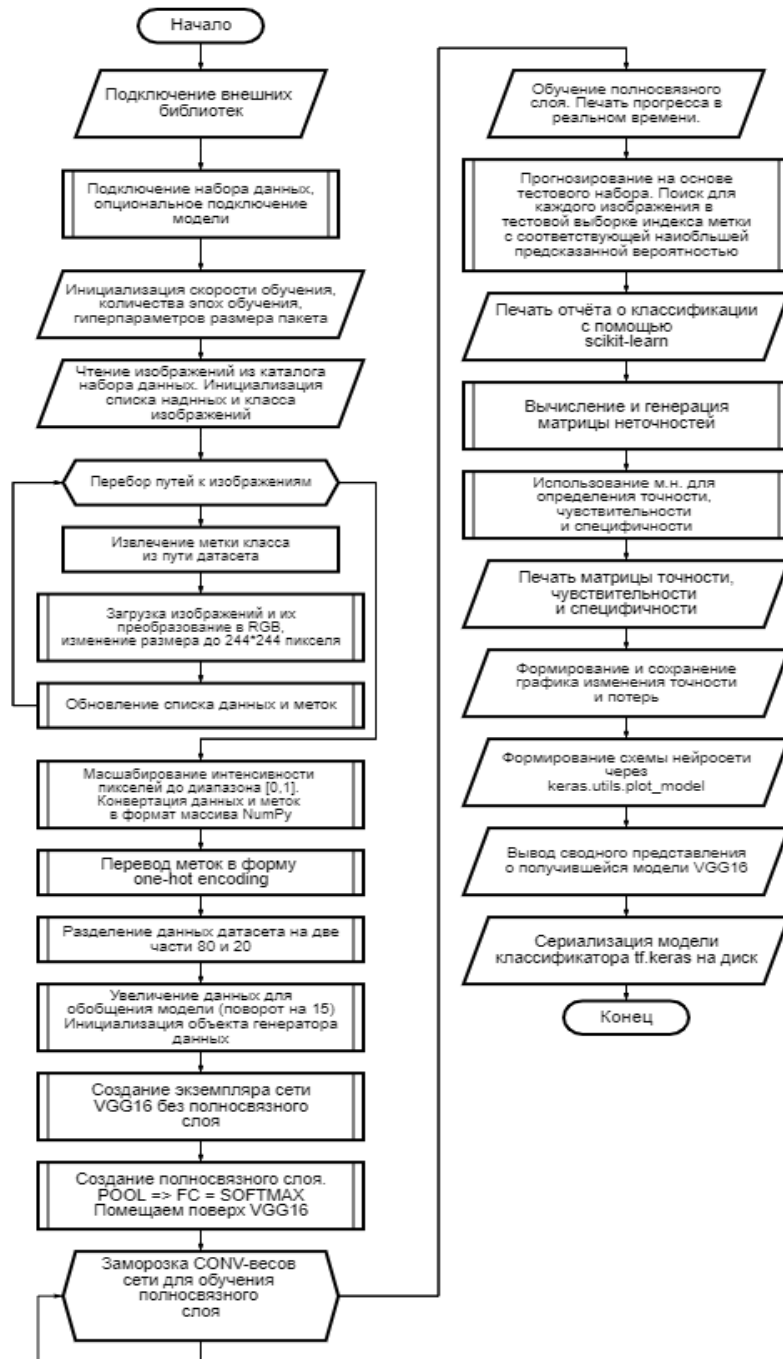


Рис. 5. Блок-схема работы программы

Затем инициализируются переменные для начальной скорости обучения, количества эпох и гиперпараметра размера пакета обучения, на основании значения которого выполняется аппроксимация градиента по всему обучающему множеству.

```
[[[0.00392157 0.00392157 0.00392157]
 [0.00392157 0.00392157 0.00392157]
 [0.00392157 0.00392157 0.00392157]
 ...
 [0.00784314 0.00784314 0.00784314]
 [0.00784314 0.00784314 0.00784314]
 [0.00784314 0.00784314 0.00784314]]

 [[0.00392157 0.00392157 0.00392157]
 [0.00392157 0.00392157 0.00392157]
 [0.00392157 0.00392157 0.00392157]
 ...
 [0.00784314 0.00784314 0.00784314]
 [0.00784314 0.00784314 0.00784314]
 [0.00784314 0.00784314 0.00784314]]

 [[0.00392157 0.00392157 0.00392157]
 [0.00392157 0.00392157 0.00392157]
 [0.00392157 0.00392157 0.00392157]
 ...
 [0.00784314 0.00784314 0.00784314]]
```

Рис. 6. Преобразование изображений в списки вводного слоя НС

После чего начинается процесс подготовки изображений, к дальнейшей обработке их нейросетью. Для начала инициализируется список данных, т. е. изображений и класс изображений. Часть данного списка изображена на рисунке 6 в значениях переменной data.

После чего в цикле происходит перебор и извлечение метки класса (covid или normal) из пути к изображению в датасете, попутно само изображение загружается и преобразовывается из цветового пространства BGR в цветовое пространство RGB, а затем уменьшается до размера 224 на 224 пикселя, тем самым оно становится готовым для обработки свёрточной нейронной сетью. Списки data и labels обновляются соответственно.

Затем происходит масштабирование интенсивности пикселей до диапазона [0, 1] и конвертация списков data и labels в формат массива NumPy. После чего происходит one-hot преобразование списки labels который после преобразования приобретёт следующий формат (рис. 7). Где каждое «закодированное» значение состоит из двухэлементного массива, один из которых является «горячим» (т. е. 1), а другой «нет» (т. е. 0)

```
[1. 0.]
 [1. 0.]
 [1. 0.]
 [1. 0.]
 [1. 0.]
 [1. 0.]
 [1. 0.]
 [1. 0.]
 [0. 1.]
 [0. 1.]
 [0. 1.]
 [0. 1.]
 [0. 1.]
 [0. 1.]
 [0. 1.]
 [0. 1.]
 [0. 1.]
 [0. 1.]
```

Рис. 7. Список labels после обработки его one-hot-encoding

После чего данные датасета разделяются на две части в соотношении 80% данных для обучения и 20% данных для тестирования. Чтобы гарантировать успешное обобщение будущей модели нейросети, выполняется увеличение данных путём поворота каждого из изображений на 15 градусов по часовой или против часовой стрелки используя метод ImageDataGenerator из библиотеки keras.

Теперь инициализируется модель VGGNet и подготавливается к тонкой настройке. Создаётся экземпляр сети VGG16 с весами предварительно обученными в ImageNet, но без использования полносвязного слоя. Это обусловлено особенностью работы с полносвязными слоями, так как конечные полносвязные слои могут принимать только входные данные фиксированного размера, который ра-

нее был определен входной формой и всей обработкой в сверточных слоях. Любое изменение формы ввода изменит форму ввода для полностью связанных слоёв, что сделает веса несовместимыми.

Затем создаётся новый полносвязный слой, состоящий из слоёв pool, fully connected и softmax и помещается наверх VGG16. После чего замораживаются CONV-веса, чтобы обучался только полносвязный слой. После чего компилируется сеть с затуханием скорости обучения и оптимизатором Adam. Учитывая, что это задача двухклассовой классификации, используется loss-функция бинарной перекрёстной энтропии.

Для запуска обучения используется метод keras fit_generator которому передаются рентгеновские снимки посредством объекта расширения данных. После чего выполняется прогнозирование на основе тестового набора и на выходе получаем прогнозные индексы. Затем генерируется и выводится с помощью вспомогательной утилиты scikit-learn отчёт о классификации.

Для будущей статистической оценки происходит вычисление матрицы неточностей путём её генерации и её использования для определения точности, чувствительности и специфичности.

После завершения всех эпох обучения строится и сохраняется в файл plot.png – график истории изменений точности и потерь. Пример такого графика изображён на рисунке 8.

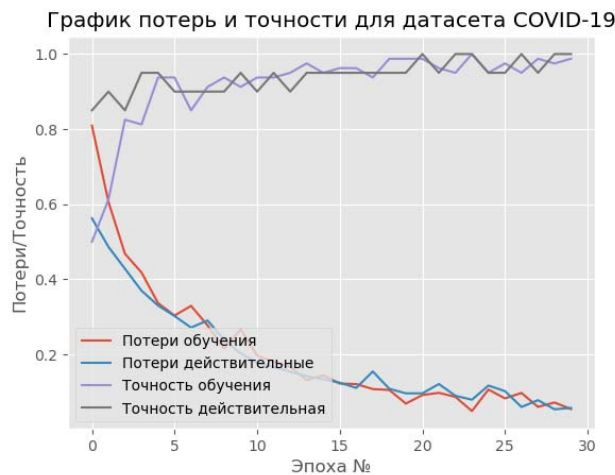


Рис. 8. Процесс изменения потерь и точности для 30 эпох

4. Исследование результатов работы нейросети

После того, как мы написали код программы для нейросети, становится возможным подобрать экспериментальным путём количество эпох и скорость обучения. После нескольких итераций запусков, с различными комбинациями значений входных параметров, по графикам аналогичным тому, что мы можем наблюдать выше, на рисунке 8, можно сделать вывод, что нейросеть работает верно. Анализируя полученный результат, можно утверждать, что имея сравнительно небольшой набор данных, полученной точности хватает для решения поставленной задачи и сеть способна корректно определять наличие COVID19 в ткани лёгких на рентгеновских снимках.

Выводы

В статье приведён пример разработки и реализации нейронной сети по выбранной авторами предметной области, в частности, выполнено проектирование и реализация нейросети для оперативной интеллектуальной классификации рентгеновских снимков здоровых людей и людей больных COVID19. Разработана и представлена архитектура сети, подготовлен набор учебных и тестовых данных для разрабатываемой сети. Исходный набор данных состоит из 100 снимков, обучение производилось на 80 снимках, тестирование на 20 снимках.

Итогом данной работы является спроектированная и корректно работающая нейронная сеть, способная с высокой точностью отличать на основании рентгеновских снимков, пациентов с заболеваниями пневмонии или COVID19. Полученной точности хватает для решения поставленной задачи.

Для увеличения точности и уменьшения ошибки в работе сети, при обучении одной используется метод подбора оптимальных значений параметров количества эпох и скорости обучения.

В рамках проведения данных работ удалось добиться, при достижении значения количества эпох равному 30, скорости обучения НС, равной 0.0025 и размера пакета обучения равному 10, значения ошибки (Точность действительная) равной 0.99. Аналогичным образом этого возможно добиться, проводя более тонкую оптимизацию работы VGG16, либо увеличением набора данных. При изменении архитектуры VGG16, выбран подход с ручным созданием экземпляра сети VGG16 с весами, предварительно обученными в ImageNet, но без использования полносвязного слоя и последующим созданием одного слоя.

Разработанная авторами нейронная сеть полностью оправдывает возможность своего дальнейшего применения в сфере медицинской диагностики для проверки наличия повреждений в лёгких пациентов, вызванных SARS-CoV-2, т.к. показала высокую точность работы.

Литература

1. *Wu F., Zhao S., Yu B. и др.* A new coronavirus associated with human respiratory disease in China // Nature. 2020. № 580(7803). С. 265.
2. Pneumonia of unknown cause – China. World Health Organization. URL: <https://www.who.int/emergencies/disease-outbreak-news/item/2020-DON229>.
3. *Mriganka N., Chandrajit Choudhury.* Automatic Detection of Pneumonia from Chest X-Rays Using Deep Learning // International Conference on Machine Learning, Image Processing, Network Security and Data Sciences. Silchar: MIND 2020, 2020. С. 175-182.
4. *Лешин Е.В., Ретинский В.Н., Воронова Л.И.* Распознавание изображений на основе технологии TensorFlow // Сборник трудов XV Международной отраслевой научно-технической конференции «Технологии информационного общества». 2021. С. 321-323.
5. *Брус В.П., Воронова Л.И.* О разработке системы нейросетевого прогнозирования состояния здоровья человека // Искусственный интеллект и цифровая экономика: взгляд студенчества. Материалы I Всероссийской студенческой научно-практической конференции. М.: Министерство науки и высшего образования Российской Федерации, Государственный университет управления, 2020. С. 171-173.
6. *Баширов А.Н., Воронова Л.И.* Прогнозирование вероятности бронхиальной астмы у детей с применением алгоритма случайного леса // Искусственный Интеллект И Цифровая Экономика: Взгляд Студенчества. материалы I Всероссийской студенческой научно-практической конференции. М.: Министерство науки и высшего образования Российской Федерации, Государственный университет управления, 2020. С. 169-171.
7. ImageNet. Wikipedia.org. URL: <https://ru.wikipedia.org/wiki/ImageNet>.
8. ImageNet Large Scale Visual Recognition Challenge (ILSVRC). ImageNet. URL: <https://imagenet.org/challenges/LSVRC>.
9. Свёрточная нейронная сеть. Wikipedia. URL: https://ru.wikipedia.org/wiki/Свёрточная_нейронная_сеть.
10. *Костецкая Г.Ю., Федяев О.И.* Распознавание изображений человеческих лиц с помощью свёрточной нейронной сети // Сб. трудов Международной студенческой научно-технической конференции «Информационные управляющие системы и компьютерный мониторинг». 2010. С. 216-218.
11. *He, Kaiming, Zhang и др.* Deep Residual Learning for Image Recognition // 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Las Vegas, NV, 2016. С. 770-778.
12. *Krizhevsky Alex, Sutskever и др.* ImageNet classification with deep convolutional neural networks // Communications of the ACM. 2017. С. 84-90.
13. *Richmond Alake.* Deep Learning: GoogLeNet Explained // Towards Data Science. URL: <https://towardsdatascience.com/deep-learning-googlenet-explained-de8861c82765>.
14. *Simonyan K., Zisserman A.* Very deep convolutional networks for large-scale image recognition // arXiv preprint arXiv:1409.1556. 2014.
15. covid-chestxray-dataset // Github.com. URL: <https://github.com/ieee8023/covid-chestxray-dataset>.

ПРИМЕНЕНИЕ НЕЙРОННЫХ СЕТЕЙ В ОБЛАСТИ РАСПОЗНАВАНИЕ ЖЕСТОВ

Скрябин Владислав Игоревич,
МТУСИ, магистрант, Москва, Россия
s.vlad9914@gmail.com

Белов Никита Вадимович,
МТУСИ, ассистент каф. ИСУиА, Москва, Россия
djnikone2777@yandex.ru

Аннотация

В наше время технологии получают всё больше возможностей, упрощая жизнь людей с различными ограничениями. В данной работе исследуется применение нейронных сетей в области распознавание жестов. Приведен подробный разбор проектирования, реализации и тестирования системы распознавания жестов.

Ключевые слова

Жесты, машинное обучение, свёрточная нейронная сеть, google Colab, Jupyter Notebook.

Введение

С развитием технологий и существенным увеличением мощности вычислительных систем, стали актуальными задачи автоматизации процессов и создания искусственного интеллекта (ИИ).

На сегодняшний день востребованной задачей является прогнозирование и распознавание объектов, для которой используются подходы машинного обучения. С машинным обучением (МО) тесно связаны такие области, как искусственный интеллект, интеллектуальный анализ данных, статистика, распознавание образа и базы данных.

Многие исследователи занимаются решением задач в данной области, в том числе, кафедра Интеллектуальные системы в управлении и автоматизации МТУСИ [1, 2, 3].

Анализ технологий для проектирования и реализации

В ходе анализа технологий было выявлено два возможных способа реализации системы:

1. Локальный.
2. Облачный.

В локальном способе стоит выделить два инструмента разработки:

1. Jupyter notebook [4].
2. Pycharm [5].

Рассмотрим каждый из них отдельно. Jupyter notebook это среда разработки, предназначенная для разработки программного обеспечения с открытым исходным кодом, открытых стандартов и сервисов для интерактивных вычислений на многих языках программирования.

У этого инструмента есть свои плюсы и минусы. Его можно использовать в браузере, но вычислительных мощностей не будет достаточно для выполнения расчетов. При запуске «Jupyter» локально с помощью сервиса «anaconda» у данной среды разработки будет недостаточно мощностей для тренировки модели, поскольку все расчеты будут опираться на вычислительный процессор персонально ПК. Поэтому стоит отказаться от этого способа, несмотря на его удобство.

PyCharm – интегрированная среда разработки, которая включает редактор кода и компилятор для написания и компиляции программ на одном или нескольких языках программирования.

Кроме того, IDE поставляется с большим количеством функций, которые облегчают всестороннюю разработку программного обеспечения.

PyCharm локальный вариант, но основная проблема реализации в этой среде снова упирается в вычислительные мощности ПК. В момент тренировки модели все вычислительные процессы переходят

дят на мощности процессора, а этого недостаточно для реализации поставленной задачи. Поэтому стоит рассмотреть облачный вариант создания модели.

Когда речь заходит об облачной реализации, то стоит обратить внимание на бесплатный облачный сервис на основе Jupyter Notebook – Google Colab, который поддерживает бесплатный графический процессор и позволяет выполнять всю работу на серверах Google.

Достоинства:

- Бессерверный режим: Colab позволяет мгновенно создать полностью бессерверный блокнот Jupyter в браузере. Это означает, что не нужно собственное мощное оборудование и установка языка Python
- Нет привязки к операционной системе: это могут быть такие системы, как Windows или MacOS, так и обычный телефон с возможностью выхода в сеть
- Совместное использование: можно предоставить общий доступ к записной книжке, как к документу Google. Таким образом проектом могут воспользоваться и другие участники проектирования
- «Google» бесплатно предоставляет графический процессор GPU и TPU

Недостатки:

- Трудно работать с большими наборами данных, так как они должны загружаться и храниться на диске Google (15 ГБ свободного места с идентификатором gmail)
 - У него есть ограничения, которые не позволяют использовать его в производственной сфере
- Colab подходит для программирования на Python с использованием библиотек глубокого обучения, такими как:
- PyTorch.
 - Keras.
 - TensorFlow.
 - OpenCV.

Появляется возможность создавать, загружать, хранить и делиться записными книжками в Colab, а также монтировать Google Диск и использовать все, что там хранится. Дистрибутив Anaconda Jupyter Notebook уже содержит в себе несколько предустановленных библиотек данных [6]. Google Colab предоставляет еще больше предустановленных библиотек машинного обучения, таких как Keras, TensorFlow и PyTorch.

Метод CNN (convolutional neural network)

Свёрточная нейросеть работает по следующему принципу: берется исходное изображение (стоит помнить, что изображение по своей сути является матрицей), далее оно пропускается через серию свёрточных, нелинейных слоев и полносвязных слоев, после чего генерируется вывод.

Свёрточный слой является первым в нейросети, на вход которого поступает исходное изображение. Он состоит из некоторого фильтра (матрица заданного размера, которая выводится в ходе обучения нейросети), который движется по входному изображению и умножает значения фильтра на исходные значения пикселей изображения (поэлементное умножение), после чего умноженные значения суммируются в одно число. Далее фильтр сдвигается на один шаг вправо и так далее. Принцип работы приведён ниже (рис. 1).

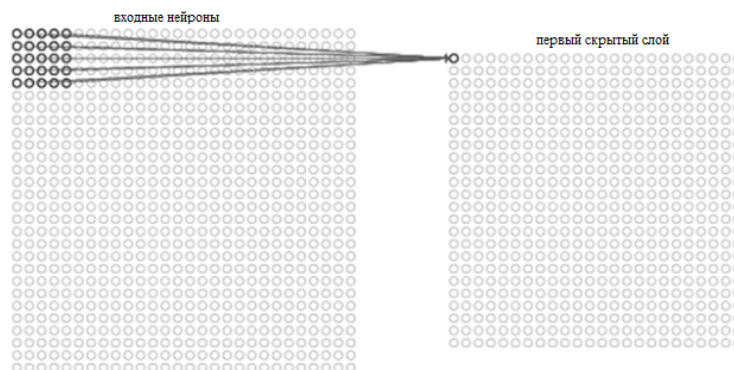


Рис. 1. Свёрточный слой

На данном слое выделяются свойства базового уровня: прямые границы, простые цвета и кривые, причем так, что один фильтр – это одно примитивное свойство. Вывод данного слоя – карта свойств, т.е. отображение областей, в которых больше вероятности наличия кривых. Чем больше фильтров, тем больше глубина карты свойств и больше информации о входной картинке. Рассмотрение данного алгоритма на простом примере приведено ниже:

Имеется фильтр (рис. 2). Слева имеется его матричная форма, а справа – графическое отображение.

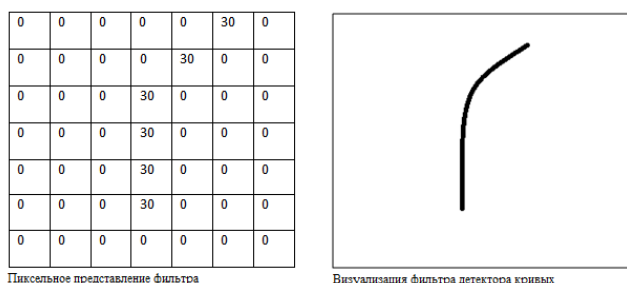


Рис. 2. Фильтр первого свёрточного слоя

Также имеется некоторое изображение, которое будет пройдено фильтром, производя перемножение значений пикселей и суммируя их (рис. 3). На выходе получаем число, причем достаточно большое.

Далее фильтром начинается прохождение по всему изображению, производя эти же операции. Например, ниже проиллюстрирована ситуация, при которой для заданного фильтра нет положительного результата (рис. 4).



Рис. 4. Фильтр не смог определить кривую, на которую не был обучен свёрточный слой

Как было сказано выше, проходя по всему изображению на выходе слоя получается карта свойств, т.е. отображение областей, в которых больше вероятности наличия кривых.

Для обнаружения более сложных деталей изображения (к примеру, руки, либо туловища человека), поверх первого фильтра, выходом которого является карта простейших свойств, можно применить новые фильтры, которые реагируют на свойства более высокого уровня (полукольца, комбинации прямых с изгибом, квадратов). Чем больше свёрточных слоев проходит изображение, тем более сложные характеристики вводятся в карты свойств. Например, в конце сети могут быть фильтры, которые активируются при наличии текста или руки на изображении.

Распознавание жестов в видео будет реализован на основе свёрточной нейросети (CNN) YOLO четвёртой версии (Scaled YOLO v4). Выбор обусловлен в первую очередь тем, что на сегодняшний день это одна из самых лучших нейронных сетей для обнаружения, среди опубликованных нейронных сетей. Кроме того, она является лучшим с точки зрения соотношения скорости и точности распознавания согласно данным компании Microsoft [7].

Выбор технологии для реализации подсистемы

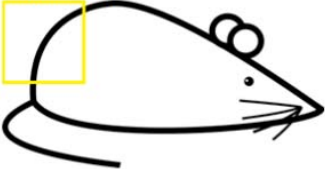
Можно производить обучение модели локально на своей машине. Однако в ходе анализа технологий для проектирования было выбрано другое средство разработки – Google Colab [8]. В его основу также входит «Jupyter Notebook», но Google предоставляет бесплатно виртуальную машину с мощностями, которыми не обладает обычный ПК. Google предоставляет бесплатный графический процессор Tesla K80 объемом около 12 ГБ (рис. 5).

1. Оригинальное изображение:



Исходное изображение

2. Помещаем фильтр в верхний левый угол изображения:



Визуализация фильтра на изображении

3. Теперь производим свёртку: перемножаем значения фильтра со значениями части изображения, которые попали в область фильтрации (желтый квадрат), а потом суммируем их. Получаем некоторое число



Визуализация рецептивного поля

0	0	0	0	0	0	30
0	0	0	0	50	50	50
0	0	0	20	50	0	0
0	0	0	50	50	0	0
0	0	0	50	50	0	0
0	0	0	50	50	0	0
0	0	0	50	50	0	0

Пиксельное представление воспринимающего поля

*

0	0	0	0	0	30	0
0	0	0	0	30	0	0
0	0	0	30	0	0	0
0	0	0	30	0	0	0
0	0	0	30	0	0	0
0	0	0	30	0	0	0
0	0	0	0	0	0	0

Пиксельное представление фильтра

Умножение и суммирование = $(50*30)+(50*30)+(50*30)+(20*30)+(50*30) = 6600$ (A large number!)

Рис. 3. Применение фильтра к части изображения



Рис. 5. Графический процессор Tesla K80

Все вычисления происходят быстрее. Это всё можно реализовать на локальной машине, но для этого нужно соответственно ПК с более высокими системными требованиями. Ниже приведено сравнение с другими графическими процессорами (рис. 6).

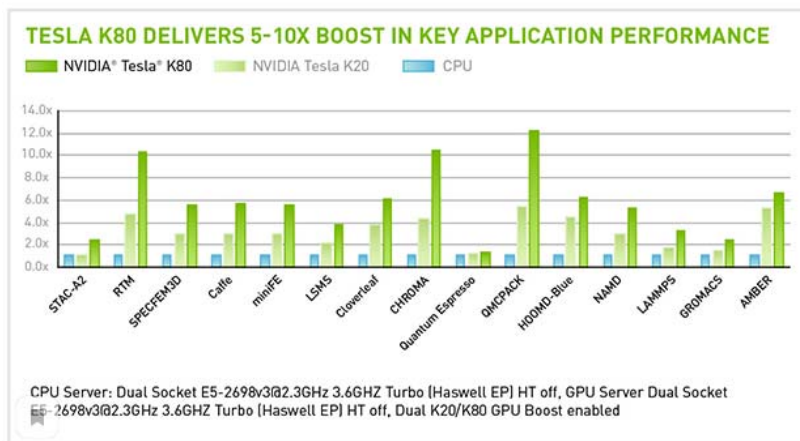


Рис. 6. Сравнение с другими графическими процессорами

Программная реализация интеллектуальной системы распознавания жестов

Для начала в «Google Colab» будет создан документ на «Google-диске» с именем «System1.ipynb» (рис. 7).

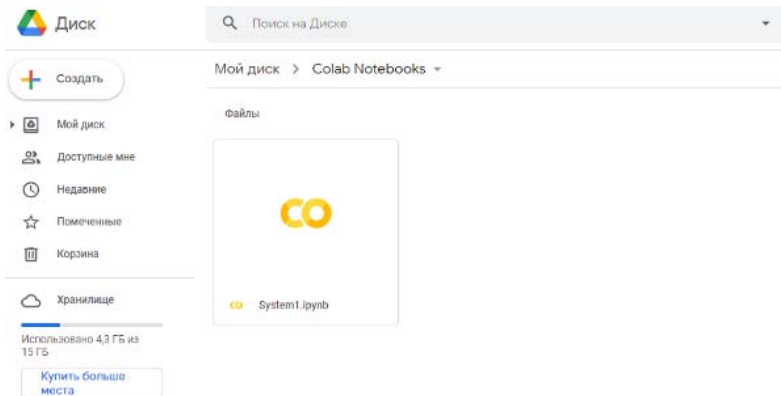


Рис. 7. Создание нового документа

После создания необходимо включить графический процессор. GPU на ноутбуке Colab сможет обрабатывать изображения в разы быстрее, чем процессор (рис. 8).

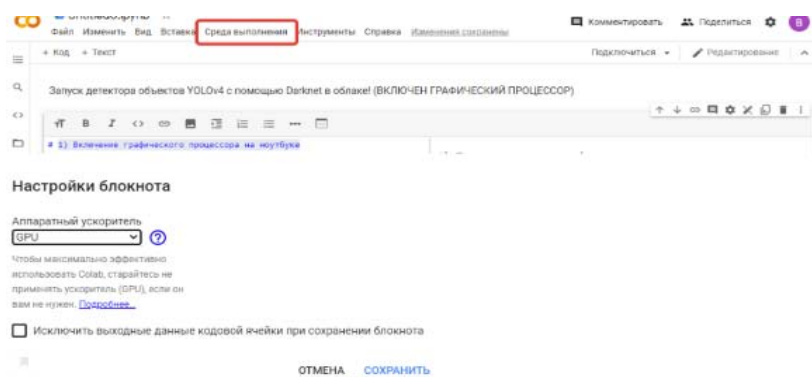


Рис. 8. Включение графического процессора

После включения аппаратного ускорителя в виртуально среду выполнения необходимо импортировать «darknet yolo». «Darknet» – это фреймворк нейронной сети с открытым исходным кодом, написанный на языках C и CUDA. «Darknet» содержит уже обученные веса на базе «dataset COCO». В работе «Darknet» понадобится для тренировки классификатора.

Необходимо также подключить «Google-диск». На диске будут храниться данные, необходимые для работы. Ниже представлен пример того, как выглядит папка, расположенная на диске (рис. 9).

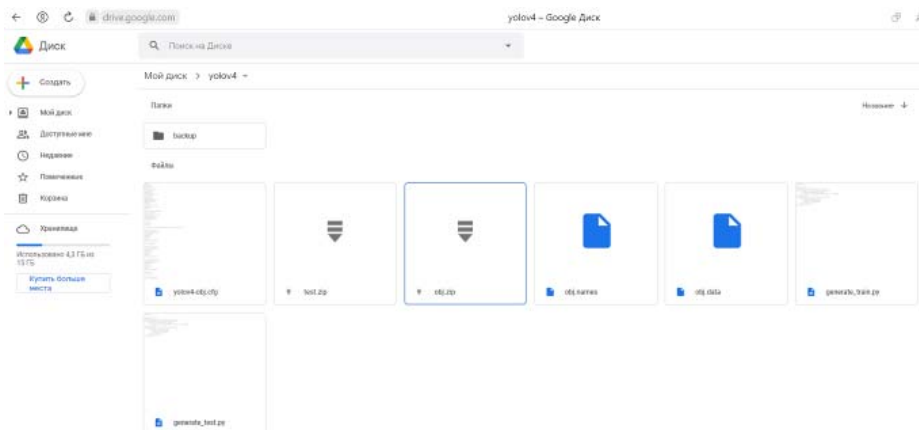


Рис. 9. Рабочая папка на Google-диске

«Obj.zip» содержит набор изображений с уже подготовленными заранее метками. Вторая группа изображений, которая находится в «test.zip» содержит набор для тестирования. Сделано это для того, чтобы в момент тренировки была возможность протестировать модель на наличие потерь с изображениями, которые не используются в тренировочном процессе. Для тестирования используется 20% изображений от общего количества изображений, содержащихся в «Obj.zip».

Файл «yolov4-obj.cfg» необходимо отредактировать на основе детектора объектов. Изменения вносятся на основании того, сколько классов будет использовано для тренировки.

Первая переменная, на которую стоит обратить внимание при работе с файлом – это «max_batches». Рассчитывается она по следующему принципу – (n-классов) * 2000, следовательно max_batches = 52000. Следующая не менее важная переменная – это «steps». Для расчета шагов берётся 80% от «max_batches» и 90% «max_batches», следовательно «steps» = 41600, 46800.

Необходимо настроить ширину и высоту (это может быть любое число кратное 32). Необходимо это для того, чтобы задать размер сети, поэтому каждое изображение будет изменено под размер сети во время обучения. В работе использовались «width» = 416 и «height» = 416. Можно улучшить результаты распознавания, путём увеличения размеров, например 608, но тогда это существенно замедлит тренировочный процесс. Ещё один не мало важный параметр – это «Filters». Рассчитывается он по принципу «n классов + 5 * 3», следовательно, их будет 41 в рамках работы. Также необходимо прописать количество классов, необходимых для тренировочного процесса. Это лишь часть основных параметров, которые подвергаются изменениям, в зависимости от количества классов.

После подготовки всех исходных файлов необходимо выполнить импортирование их в Darknet для дальнейшей тренировки.

После подготовки всех данных необходимо перейти к обучению. Так как будет выполняться 6000 итераций при тренировке, чтобы не потерять уже обученные веса, каждый 100 итераций будут сохраняться на диск. Таким образом, через «n-ое» количество часов тренировочного процесса в папке «backup» будет располагаться обученная модель для 26 классов.

После подготовленной модели был подготовлен скрипт, при помощи которого интегрировалась наша модель, которая способна распознавать жесты в реальном времени с камеры видеослежения. К каждому жесту будет присвоена вероятность того, к какому из классов его отнести. В зависимости от уровня освещения в помещении и качества входного видеопотока будет выдаваться разная вероятность. Точность можно увеличить, путем применения камер большего разрешения.

При визуальном представлении был написан JS-скрипт в HTML-странице в блокноте. В коде была прописана ссылка на диск, где расположена обученная модель. В визуальном представлении рассмотрено три жеста. В первом случае был рассмотрен жест, который обозначает букву «В» (рис. 10).

Во втором случае был рассмотрен жест, который обозначает букву «W». Системе удалось со 100% вероятности определить тот факт, пользователь показывает букву «W» (рис. 11).

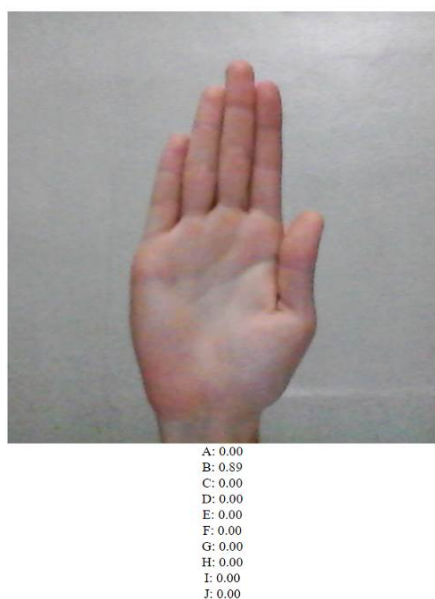


Рис. 10. Распознавание жеста «В»

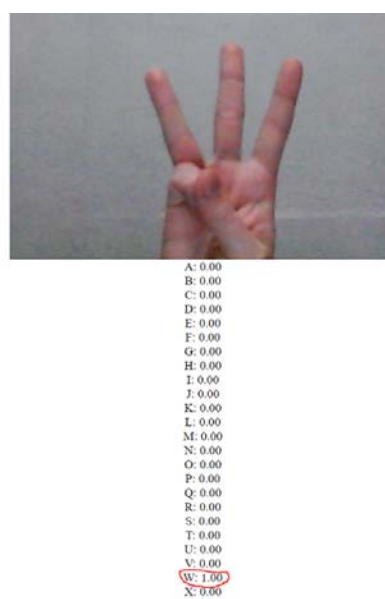


Рис. 11. Распознавание жеста «W»

Третья ситуация фиксирует распознанную букву «К» (рис. 12).



A: 0.00
 B: 0.00
 C: 0.00
 D: 0.00
 E: 0.00
 F: 0.00
 G: 0.00
 H: 0.00
 I: 0.00
 J: 0.00
K: 0.93
 L: 0.00
 M: 0.00
 N: 0.00
 O: 0.00
 P: 0.00
 Q: 0.00
 R: 0.00
 S: 0.00
 T: 0.00
 U: 0.00
 V: 0.04
 W: 0.01

Рис. 12. Распознавание жеста «К»

Заключение

В статье исследована предметная область применения нейронных сетей в области распознавание жестов. Выполнено исследование методов интеллектуального слежения, в результате чего сделан выбор в пользу CNN-метода. Подготовлен набор данных, каждое изображение относилось к определённому классу и имело заданные координаты детектирования.

Произведено тестирование разработанной системы, для этого использован google colab. На вход подается изображение жеста руки, подготовленная модель относит его к одному из классов. Проведено тестирование при помощи камеры в реальном времени.

Построенная система распознает жесты с вероятностью не менее 89%.

Литература

1. Воронов В.И., Воронова Л.И., Вовик А.Г. и др. Разработка программного комплекса биометрической идентификации на основе распознавания лиц для системы безопасности университета с использованием нейросетевых методов и современных программных решений // Отчет о НИР № 15-2/2019-Б от 25.03.2019 (Московский технический университет связи и информатики).
2. Челябин И.А., Воронова Л.И. Применение нейронной сети прямого распространения для распознавания русского жестового языка // Сборник трудов XV Международной отраслевой научно-технической конференции «Технологии информационного общества», 2021. С. 346-348.
3. Мартынов Д.А., Воронова Л.И. Применение перцептрона для распознавания дактилем русского жестового языка // DSPA: Вопросы применения цифровой обработки сигналов. 2020. Т. 10. № 2. С. 37-46.
4. Jupyter.org. 2021. URL: <https://jupyter.org/> (дата обращения: 08.04.2021).
5. PyCharm. jetbrains.com. 2021. URL: <https://www.jetbrains.com/pycharm/> (дата обращения: 08.04.2022).
6. Google Colab. Colab.research.google.com. 2021. URL: <https://Colab.research.google.com/notebooks/intro.ipynb#scrollTo=OwuxHmxllTwN> (дата обращения: 26.04.2021).
7. Microsoft.com. 2021. URL: <https://www.microsoft.com/en-us/research/publication/microsoft-coco-common-objects-in-context/> (дата обращения: 01.05.2021).
8. Cloud.google. 2007. URL: <https://cloud.google.com/why-google-cloud?hl=ru> (дата обращения: 16.04.2021).

НЕЙРОСЕТЕВОЕ РАСПОЗНАВАНИЕ ПОЛА И ВОЗРАСТА ПО ФОТОГРАФИИ ЛИЦА

Харшиладзе Дмитрий Георгиевич,

Московский Технический Университет Связи и Информатики, Москва, Россия,
dimahg2000@mail.ru

Аннотация

Данная статья рассматривает процесс разработки нейронной сети для распознавания пола и возраста человека по фотографии с помощью OpenCV и глубокого обучения. Рассмотрены этапы разработки, модель нейронной сети, принцип её работы, а также набор данных, с помощью которого нейронная сеть будет предсказывать пол и возраст человека по фотографии. Прогнозирование возраста рассматривается как задача регрессии. Результаты исследования показывают, что сеть умеет приблизительно определять возраст и пол человека корректно.

Ключевые слова

OpenCV, Python, Haar Cascade, Leanear SVM, распознавание лица, глубокое обучение, компьютерное зрение, нейронные сети, методы распознавания лиц.

Введение

Возраст и пол человека играют важнейшую роль в социальной жизни индивида и в его взаимодействии с другими людьми. Например, в зависимости от возрастной группы человека во многих языках меняется форма обращения к нему.

Определение пола и возраста является популярной задачей машинного обучения, и в настоящий момент существуют различные подходы к её решению.

Решение данной задачи имеет множество практических приложений. К примеру, появляется возможность быстрого принятия решения о допуске лиц на мероприятие, имеющее возрастной ценз, или на продажу им некоторых товаров. В области маркетинга и анализа статистики сегментирование посетителей по социально-демографическому признаку позволяет лучше понять состояние рынка и предпочтения пользователей. Исследование в этом направлении проводится на кафедре ИСУиА МТУСИ [1, 2, 3].

OpenCV (Open Source Computer Vision Library) – это библиотека программного обеспечения для компьютерного зрения и машинного обучения с открытым исходным кодом.

Библиотека содержит более 2500 оптимизированных алгоритмов, включая полный набор как классических, так и современных алгоритмов компьютерного зрения и машинного обучения.

Эти алгоритмы могут использоваться для обнаружения и распознавания лиц, идентификации объектов, классификации действий человека в видео, отслеживания движений камеры, отслеживания движущихся объектов, извлечения трехмерных моделей объектов, создания трехмерных облаков точек со стереокамер, сшивания изображений вместе для получения высокого разрешения. изображение всей сцены, найти похожие изображения из базы данных изображений, удалить эффект красных глаз с изображений, снятых с помощью вспышки, следить за движениями глаз, распознавать пейзаж и устанавливать маркеры, чтобы наложить его на дополненную реальность и т. д.

Компьютерное зрение включает в себя классификацию групп пикселей в изображении или видеопотоке как фоновых или уникальных. Каждый из следующих примеров демонстрирует различные алгоритмы, которые отделяют уникальные функции от фона с помощью различных методов.

У библиотеки также много других полезных функций, однако нас конкретно интересует алгоритм распознавания лиц. Алгоритм может обнаруживать и другие объекты, кроме лиц, OpenCV как раз также включает инструменты, позволяющие обучить нейронную сеть поиску любого объекта, а не только лица [4].

Результаты исследования

Определение возраста – это процесс автоматического определения возраста человека исключительно по фотографии его лица.

Определение пола – это также процесс автоматического определения пола человека исключительно по фотографии его лица.

Данные задачи можно разделить на двухэтапный процесс:

- Этап №1: обнаружение лиц во входном изображении
- Этап № 2: выделение интересующей области лица и применение алгоритма определения возраста, а затем и пола, для прогнозирования возраста и пола человека.

Для этапа № 1 можно использовать любой детектор лиц, способный создавать ограничивающие рамки для лиц на изображении, включая, помимо прочего, алгоритм Haar cascades (Каскады Хаара), модель HOG + линейная SVM, детекторы одиночного выстрела (SSD) и т. д.

- Алгоритм Haar cascades будет очень быстрым, он способен работать в реальном времени на встроенных устройствах, однако проблема в том, что он менее точен и очень подвержен ложным срабатываниям, когда лицо находится там, где его нет;

- Модели HOG и линейная SVM более точны, чем Haar cascades, но работают медленнее. Они также плохо справляются с ситуацией, когда не все лицо видно;

- Детекторы лиц на основе глубокого обучения являются наиболее надежными и обеспечивают наилучшую точность, но требуют даже больше вычислительных ресурсов, чем Haar cascades и линейную SVM с моделью HOG.

Поэтому стоит, для начала, определиться что важнее – скорость или точность, учитывая то, что некоторые алгоритмы гораздо больше требуют вычислительных ресурсов, нежели другие.

После определения детектором координат, ограничивающих рамки лица в изображении, переходим к этапу №2 – определение возраста и пола человека. Возраст и пол определяются по ограничивающей рамке (x, y) – координатам лица, сначала извлекается прямоугольник лица, остальная часть изображения игнорируется, чтобы программа не фокусировалась и не пыталась определить возраст другого “шума”. Прямоугольник пропускается через алгоритм определения лица, что и даёт фактическое предсказание. Предсказание возраста и пола или же классификация на возрастную группу и гендерную группу как раз будет происходить с помощью нейронных сетей.

Существует несколько алгоритмов определения возраста, но наиболее популярными из них являются детекторы возраста на основе глубокого обучения.

Используемая модель глубокого обучения распознавания пола и возраста по лицу была первоначально реализована и обучена Джил Леви и Толом Хаснером в статье “ Age and Gender Classification using Convolutional Neural Networks” [5].

На рисунке 1 представлена модель глубокого обучения распознавания пола и возраста по лицу.

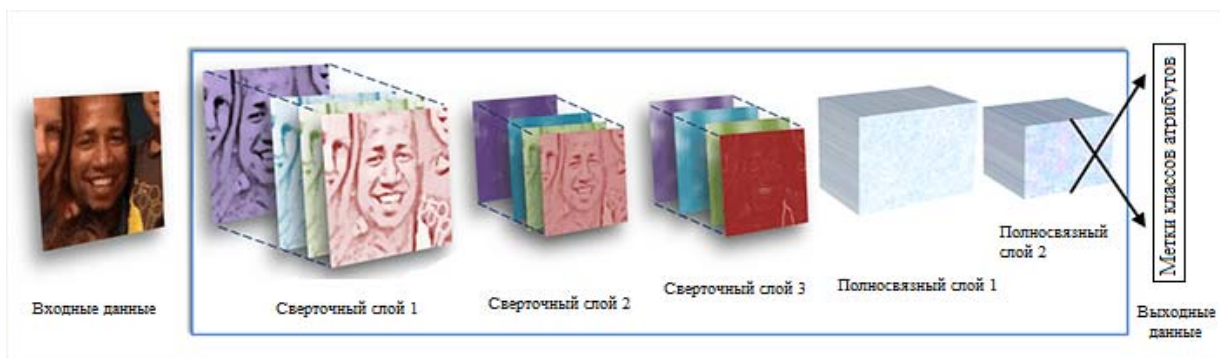


Рис. 1. Модель глубокого обучения распознавания пола и возраста по лицу

Сеть содержит три сверточных слоя, за каждым из которых следует исправленная линейная операция и слой объединения. Первые два слоя также следуют нормализации с использованием локальной нормализации отклика.

Первый Сверточный слой содержит 96 фильтров размером 7×7 пикселей, второй Сверточный слой содержит 256 фильтров размером 5×5 пикселей, третий и последний сверточный слой содержит 384 фильтра размером 3×3 пикселя. Наконец, добавляются два полносвязных слоя, каждый из которых содержит 512 нейронов [2].

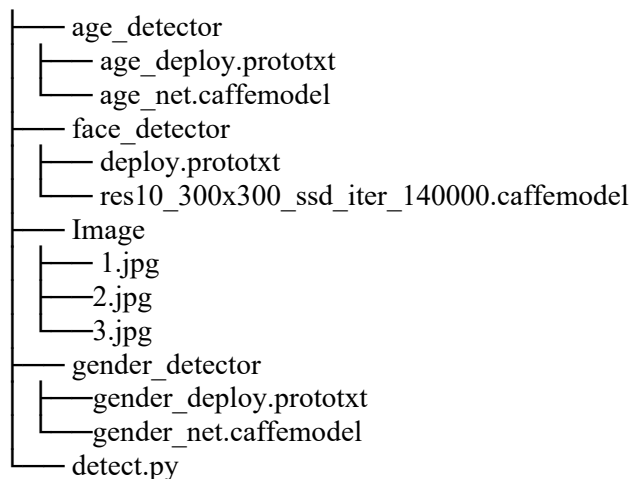
В данном случае задача определения возраста будет рассматриваться как задача классификации, а не регрессии, так как сам вопрос определения внешности строится исключительно на внешности, из-за этого регрессионный метод с большей вероятностью будет выдавать неверный результат, нежели чем через классификационный метод с определенными возрастными группами.

Диапазон будет таковым:

- 0-3
- 4-7
- 8-14
- 15-24
- 25-37
- 38-47
- 48-59
- >60

Набор данных Adience [6], используемый для обучения модели, содержит в себе данные, соответствующие фотографиям, сделанными в реальных условиях, где есть блики, шумы, с разными позами и разным освещением, чтобы можно было фиксировать все изменения во внешнем виде, которые можно ожидать от изображений, снятых без тщательной подготовки или позирования.

Сам проект выглядит следующим образом:



Каталоги проекта состоят из определителя возраста, гендера и поиска лица на фотографии. Каждая из этих моделей глубокого обучения основана на Caffe.

Также внутри есть три тестовых фотографии, которые показывают возможности нейронной сети (по факту их больше).

Detect.py является основной программой для обнаружения лица, после чего выполняется прогноз возраста с помощью OpenCV.

Модель, используемая в данном проекте, как раз была предобучена на ImageNet dataset.

Каждая модель представлена двумя файлами, это значит, что она была обучена с помощью Caffe (фреймворк глубокого машинного обучения, нацеленный на простое использование, высокую скорость и модульность). Файл с расширением prototxt отвечает за архитектуру сети, а файл с расширением caffemodel за веса модели.

Для определения возраста используются age_deploy.prototxt и age_net.caffemodel, для определения пола gender_deploy.prototxt и gender_net.caffemodel.

Для самой программы нам потребуются следующие библиотеки:

- `import argparse` – это модуль для обработки аргументов командной строки, так как саму программу мы будем вызывать из командной строки
- `import os` – это библиотека функций для работы с операционной системой
- `import cv2` – это библиотека компьютерного зрения
- `import numpy as np` – это библиотека языка Python, добавляющая поддержку больших многомерных массивов и матриц, вместе с большой библиотекой высокоуровневых (и очень быстрых) математических функций для операций с этими массивами.
- `import time` – модуль для работы со временем в Python

Наш скрипт требует четыре аргумента командной строки:

- `--image`: Предоставляет путь к входному изображению для определения возраста.
- `--face`: Путь к нашему предварительно обученному каталогу моделей детектора лиц.
- `--age`: Каталог моделей нашего предварительно обученного детектора возраста.
- `--confidence`: минимальный порог вероятности для фильтрации слабых обнаружений.

Детектор возраста и пола в данном случае – это классификатор, который предсказывает возраст и пол человека, используя ограничивающие линии его лица в соответствии с предопределенными сегментами.

В данном случае, задание сегментов будет выглядеть так:

```
AGE_LIST = [ "(0-3)", "(4-7)", "(8-14)", "(15-24)", "(25-37)", "(38-47)", "(48-59)", "(>60)" ]
GENDER_LIST = [ "Male", "Female" ]
```

Наш возраст и пол определяются в сегментах (т. е. метках классов) для нашего предварительно обученного детектора. Мы будем использовать эти списки и связанные с ними индексы, чтобы получить возрастную и гендерную группы и добавить аннотации к выходному изображению.

В процессе разработки идёт загрузка модели детектора лиц с компьютера, загрузка модели детектора возраста с компьютера, загрузка модели детектора пола с компьютера, загрузка изображения и создание входного двоичного объекта для изображения, передача двоичного объекта и получение данных об обнаружении лиц, заикливание обнаружение с помощью цикла в котором происходит фильтрация слабых обнаружений, убедившись, что достоверность больше минимальной достоверности, вычисление (x, y) – координат ограничивающей рамки для объекта, извлечение из фотографии лица только нужных рамок по координатам, прогнозирование по возрасту и нахождение возрастной группы с наибольшей соответствующей вероятностью, прогнозирование по полу и нахождение пола с наибольшей соответствующей вероятностью, отображение прогнозируемого возраста и пола на терминале и рисование ограничивающей рамки лица вместе с соответствующим прогнозируемым возрастом и полом. После выходное изображение сохраняется в отдельный файл.

Оценить результат работы можно только практически, поэтому укажу в качестве пути свою фотографию.



Рис. 2. Моя фотография №1

На момент фотографии я входил в данный диапазон возраста, пол определен верно.

На другой фотографии мой возраст определен некорректно (мне было 19 лет), в данном случае сказывается то, что сам детектор работает исключительно на ограничивающей рамке, он не может определить возраст корректно, так как не видит, что вне рамки. Также сказывается освещение. Пол определен корректно.



Рис. 3. Моя фотография №2

В качестве примера также возьмем фотографию актера Моргана Фримена, которому на момент фотографии было 76 лет. Возраст и пол определены корректно.

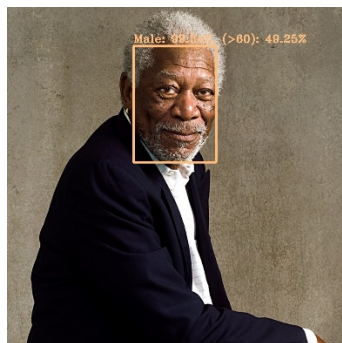


Рис. 4. Фотография Моргана Фримена

Заключение

Целью данного исследования было построить нейросетевую модель, прогнозирующую пол и возраст человека по фотографии лица. На определенном наборе данных удалось выполнить задачу классификации как для определения возраста, так и для определения пола, которое, в общем случае, дало удовлетворительные результаты.

Однако нейронная сеть всё равно делает субъективные результаты, так как она полагается именно на то, как человек выглядит. Чтобы оценить детектор возраста, вы не можете полагаться на фактический возраст человека. Вместо этого вам нужно измерить точность между предсказанным возрастом и предполагаемым возрастом. Касательно пола, нейронной сети гораздо проще классифицировать, так как в определение происходило именно на двух гендерах.

Если брать общемировую ситуацию по отношению к гендерному вопросу, то классифицировать человека, причисляя его к одному из 46-78 гендеров [7], то это может уже обидеть человека, на котором работало предсказание пола, так как определитель пола с высокой вероятностью покажет некорректный результат.

Литература

1. *Воронова Л.И., Воронов В.И.* Machine Learning: регрессионные методы интеллектуального анализа данных: Учебное пособие МТУСИ. М., 2018 (дата обращения – 20.01.2022).
2. *Жаров И.А., Воронов В.И., Воронова Л.И., Быков А.Д., Вовик А.Г., Усачев В.А.* Программа для идентификации изображения в программном комплексе нейросетевого распознавания лица (дата обращения – 20.01.2022).
3. *Воронов В.И., Воронова Л.И., Вовик А.Г., Фадеев М.В., Быков А.В., Жаров И.С.* Разработка программного комплекса биометрической идентификации на основе распознавания лиц для системы безопасности университета с использованием нейросетевых методов и современных программных решений.
4. <https://blog.skillfactory.ru/glossary/opencv/> (дата обращения – 20.01.2022).
5. https://talhassner.github.io/home/projects/cnn_agegender/CVPR2015_CNN_AgeGenderEstimation.pdf (дата обращения – 20.01.2022).
6. <https://talhassner.github.io/home/projects/Adience/Adience-data.html#agegender> (дата обращения – 20.01.2022).
- <https://ria.ru/20171214/1510843156.html> (дата обращения – 20.01.2022).

ИСПОЛЬЗОВАНИЕ ФОТОГРАММЕТРИИ ДЛЯ СКАНИРОВАНИЯ ОБЪЕКТОВ И ПОСЛЕДУЮЩЕГО ИСПОЛЬЗОВАНИЯ В САПР

Дергачёв Денис Андреевич,
МАДИ, Москва, Россия
dergdenand@gmail.com

Тимагин Данила Александрович,
МАДИ, Москва, Россия
Yararaw@mail.ru

Боликов Савелий Андреевич,
МАДИ, Москва, Россия,
kojyxovo2308@gmail.com

Аннотация

В современном машиностроении инженеры нередко сталкиваются с ситуациями, когда на основе реально имеющейся детали необходимо составить техническую документацию, которая зачастую не предоставляется заводом изготовителем или просто утеряна. Авторы статьи провели анализ фотограмметрии как метода 3D-сканирования, выявили его преимущества и недостатки, а также на практике применили данный метод для снятия геометрических размеров заднего кулака.

Ключевые слова

3D-сканирование, фотограмметрия, машиностроение, цифровая модель.

Введение

Существуют некоторые изделия, которые трудно представить в виде цифровой модели по причине сложной геометрии и невозможности снятия размеров обычными средствами измерения. Решить данную проблему представляется возможным, используя различные методы 3D-сканирования. В настоящем исследовании авторы статьи провели анализ фотограмметрии как метода 3D-сканирования, выявили его преимущества и недостатки, а также на практике применили данный метод для снятия геометрических размеров заднего кулака.

Описание процесса сканирования методом фотограмметрии

Часто встречается мнение что 3D-сканирование – достаточно сложный процесс с дорогостоящим профессиональным оборудованием. Была поставлена задача найти более простой и бюджетный метод. В качестве такого метода была выбрана фотограмметрия. Для осуществления сканирования данным методом нужны лишь смартфон и три бесплатных программы.

Первый шаг – снимаем один контрольный размер с детали, которую нам необходимо перевести в цифровую модель для дальнейшего масштабирования полученной модели (рис. 1,а). Второй шаг – фотографируем деталь с разных углов и ракурсов. Для уменьшения вероятности некорректного отображения модели желательно использовать контрастный одноцветный фон с хорошим освещением. Чем больше фотографий, тем точнее передаётся геометрия детали. Далее импортируем полученные фотографии в программу Mushroom, при помощи которой получаем 3D-геометрию детали из её фотографий (рис. 1,б).

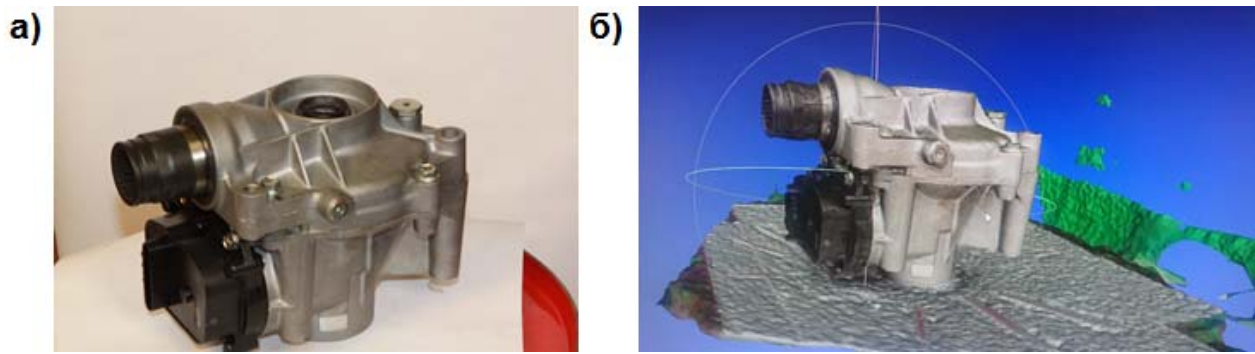


Рис. 1. Фотография изделия (а) и его 3D-модель (б)

После получения 3D-геометрии детали переносим её в программу Meshlab. Данная программа используется для удаления всего лишнего, что не относится к нашей модели. Mushroom, помимо самой детали, добавляет в модель ещё и часть фона. Для удаления ненужных нам областей используем команду «select faces». После удаления лишних элементов сохраняем модель в формате STL.

Теперь осталось только масштабировать почти готовую 3D-модель, так как программа Mushroom составляет по фотографиям только геометрию детали, не зная её реальных размеров. Открываем модель в программе Netfabb, проводим замер того размера, что выбрали в качестве контрольного на построенной модели, и вычисляем коэффициент масштабирования, на который необходимо увеличить модель, чтобы она совпадала по размерам с реальной деталью:

$$\text{коэффициент масштабирования} = \frac{\text{размер на реальной детали}}{\text{размер на модели}}$$

Далее при помощи команды «Масштабирование» увеличиваем модель на коэффициент масштабирования и получаем готовую модель с реальной геометрией и размерами. Если необходимо объединить две плоскости в одну модель, чтобы точнее передать внешнюю и внутреннюю поверхность детали, можно воспользоваться программой Meshlab.

Сравнение фотограмметрии с другими методами 3D сканирования и применение его на практике

Данный метод 3D-сканирования оказался достаточно эффективным, простым, бюджетным и не требующим профессионального оборудования. Возможно, результат получается и не столь точный, как у профессиональных сканеров, но для наших нужд большая точность не требуется, и полученный результат вполне устраивает.

Данный метод успешно применён на практике при сканировании готовых фабричных деталей, которые используются для постройки нашего изделия. Важно иметь модели данных деталей в компьютерной сборке машины для более точной оценки промежуточных построений и инженерных расчётов [1-4].

Недостатки фотограмметрии

В процессе сканирования данным методом были выявлены следующие недостатки:

1. Сильная зависимость от качества фотоаппарата, что отрицательно сказывается на доступность данного метода ввиду высокой стоимости хорошей камеры. Но даже в этом случае фотограмметрия доступнее в экономическом плане, нежели 3D-сканнер.

2. Необходимость хорошего освещения и статичности объекта. При проведении фотосъемки изделия тени, отбрасываемые самой деталью или же фотографом, отрицательно сказываются на итог обработки фотографий программой, что может привести к затруднениям в виде некорректного отображения изделия после обработки.

3. Необходимость масштабирования промежуточного результата и устранения лишних поверхностей. Так как программа не учитывает реальные размеры, то промежуточный результат не будет соответствовать реальному прототипу. Для исправления этого недочёта необходимо выполнить сохранение сканирования в формате STEP и последующее масштабирование результата в любой доступной САПР (например, Inventor).

4. Необходимость «очищать» промежуточный результат от всего лишнего. Проблема в том, что при выполнении фотографий объект находится в статичном положении на плоской поверхности (столе). В таком случае необходимо проводить фотограмметрию в двух положениях. Но данная проблема решаема, если проводить съёмку объекта в подвешенном состоянии, например, привязав объект двумя нитками. В таком случае не придётся проводить повторную фотограмметрию, соответственно, процесс «очистки» и масштабирования сократится.

Заключение

В дальнейшем планируется активно использовать метод фотограмметрии при выполнении инженерных проектов и в основной учебной деятельности.

Литература

1. *Вихров А.В.* Несущие системы транспортных средств специального назначения: учеб. пособие. М.: МАДИ, 2015. 112 с.
2. *Малиновский М.П.* Компьютерная графика в Compas: учеб. пособие. В 2 ч. Ч. 1: Конструкторская документация. М.: МАДИ, 2015. 108 с.
3. *Малиновский М.П.* Компьютерная графика в Compas: учеб. пособие. В 2 ч. Ч. 2: Оформление чертежей. М.: МАДИ, 2015. 124 с.
4. *Малиновский М.П., Малахов Д.Ю.* Компьютерное моделирование транспортных средств специального назначения: Применение библиотек и расчётных приложений КОМПАС: учеб. пособие. М.: МАДИ, 2020. 164 с.

ИНФРАКРАСНЫЙ ЛАЗЕР НА ГРАНАТЕ С НАКАЧКОЙ ТРЕМЯ ЛАЗЕРНЫМИ ДИОДАМИ ДЛЯ МАРКИРОВКИ ИЗДЕЛИЙ АВТОМОБИЛЬНОЙ ПРОМЫШЛЕННОСТИ

Ашурбеков Сефер Ашурбекович,
зав. лаб., к.т.н., доцент, Бронницкий филиал МАДИ,
ashsefer118@gmail.com

Матвиенко Диана Алексеевна,
РТУ МИРЭА, Москва, Россия
dianamatvienko9703@gmail.com

Семкина Елена Викторовна,
Бронницкий филиал МАДИ,
elena.semkina.00@mail.ru

Суфиянов Ракип Шайхиевич,
зав. каф., д.т.н., доцент, Бронницкий филиал МАДИ
mr.surash@mail.ru

Аннотация

Представлены результаты разработки инфракрасного лазера на гранате с примесными ионами неодима с накачкой тремя четырёхваттными лазерными диодами с мощностью выходного излучения 6 Вт для маркировки изделий автомобильной промышленности. Выполнен цифровой анализ профиля лазерного пучка в различных режимах маркировки с применением ПЗС-камеры.

Ключевые слова

Лазерная маркировка, лазера с диодной накачкой, профиль лазерного пучка, цифровой анализ, автомобильные детали.

Введение

Лазерная маркировка представляет способ нанесения изображений на поверхность за счёт лазерного пучка, сфокусированного оптической системой в пятно минимально возможного диаметра. При этом в области воздействия лазерного пучка создаётся высокая температура, достаточная для выжигания и испарения материала [1-5]. Центральный элемент установки для маркировки – мощный лазер. Чаще используются СО₂-лазеры, реже – неодимовые и волоконные, в которых активной средой служит оптоволокно, легированное интербием или эрбием [6-8]. Стоимость волоконных лазеров гораздо выше, чем у традиционных лазеров. Основной недостаток СО₂-лазеров это большие габариты излучателей и необходимость охлаждения системы. В случае выхода из строя газоразрядная трубка лазера может быть отремонтирована только на заводе-изготовителе. Система перемещения управляется компьютером согласно чертежам [9-12].

В нашей работе впервые предложено использовать для маркировки изделий автомобильной промышленности компактные и бурно развивающиеся в последние годы твёрдотельные лазеры с накачкой полупроводниковыми лазерными диодами.

Функциональная схема лазерного излучателя на кристалле V₃Al₅O₁₂: Nd³⁺ с накачкой тремя лазерными диодами приведена на рисунке 1.

Профиль лазерного пучка на выходе излучателя исследовалась путём компьютерной обработки сигнала с выхода ПЗС-камеры по специальной программе.

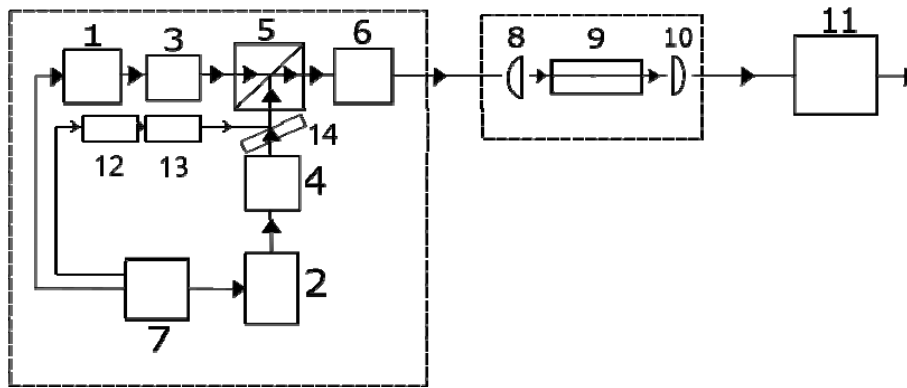


Рис. 1. Функциональная схема инфракрасного лазерного излучателя с диодной накачкой тремя лазерными диодами

Анализ энергетических и геометрических характеристик лазерного излучателя по профилю пучка позволил определить уникальные параметры излучателя.

Режим работы – одночастотный, длина волны 1064 нм. Диаметр пучка на выходе излучателя < 1.5 мм, угловая расходимость < 0.6 миллирадиана, мощность излучателя 6.5 Вт, шумы излучения < 4%.

1, 2, 12 - четырёхлазерные диоды накачки (808 нм); 3, 4, 13 – конденсоры; 5 – призма-куб; 6 – объектив; 7 – стабилизированный источник питания 5V; 8, 10 – зеркала оптического резонатора; 9 - кристалл граната; 11 – телескоп двухкомпонентный для формирования геометрических параметров лазерного пучка.

Полупроводниковые лазерные диоды накачки (1, 2, 13) установлены в корпусе осветителя. Диоды питаются стабилизированным источником питания 5 В и имеют возможность в небольших пределах перемещаться относительно конденсоров (3), (4), (13). Излучение диодов накачки направляется конденсорами на призму-куб (5). Далее лазерный пучок фокусируется на активном элементе (9), который вместе с зеркалами резонатора (8) и (10) установлен в корпусе резонатора. Зеркало (8) с коэффициентом отражения 0,98 для линий основной гармоники 1064 нм, зеркало (10) полупрозрачное с коэффициентом пропускания 40-60%. Корпуса осветителя и резонатора размещены на элементах Пельтье для стабилизации температуры диодов накачки и активного элемента.

Геометрические параметры излучателя формируются с помощью двухкомпонентного телескопа. Каскадная конструкция излучателя позволяет создать многокаскадную накачку активного элемента и довести мощность излучения до 60-80 Вт, что позволит осуществить маркировку и гравировку также металлических деталей автомобилей. Конструкция предложенного излучателя позволяет осуществить также внутррезонаторную акусто-оптическую модуляцию выходного излучения, что позволяет использовать лазеры с диодной накачкой для закалки и упрочнения ответственных металлических деталей автомобилей.

Заключение

1. Разработан компактный и эффективный лазер на гранате с накачкой тремя лазерными диодами с выходной мощностью 6,5 Вт и угловой расходимостью менее 0,5 миллирадиан и отработаны режимы маркировки неметаллических деталей автомобилей

2. Выполнен цифровой анализ лазерного пучка с применением ПЗС-камеры и компьютера.

3. Представляется целесообразным продолжение исследований с целью разработки многокаскадных излучателей для увеличения выходной мощности до 60-80 Вт.

Литература

1. *Вейко В.П., Петров А.А., Самохвалов А.А.* Введение в лазерные технологии. Опорный конспект лекции по курсу «Лазерные технологии» под редакцией Вейко В.П. // СПб: Университет ИТМО. 2018. 161 с.
2. *Вейко В.П.* Опорный конспект лекции «Лазерной микро- и нанотехнологии в микроэлектронике». СПб: НИУ ИТМО, 2011, 141 с.
3. *Вейко В.П., Шахно Е.А.* Лазерные технологии в примерах и задачах; учебное пособие. СПб: Университет ИТМО, 2014, 88 с.
4. *Семашко В.И.* Методы лазерной маркировки // Доклады БГУПР, октябрь-декабрь, № 4, 2004.
5. *Пойзнер Б.И.* Физические основы лазерной техники. Учебное пособие. М.: ИНФРА-М, 2018, 160 с.
6. *Дьюми У.* Лазерная технология и анализ материалов: перевод с английского. М.: МИР, 1986, 504 с.
7. *Голубчик А.М., Филатов С.А.* Документарное оформление экспорта международных транспортно-экспедиторских услуг // Вестник Московского автомобильно-дорожного государственного технического университета (МАДИ). 2019. № 1 (56). С. 3-8.
8. *Ушаков Д.В., Максимов В.А., Солнцев А.А., Поживилов Н.В.* Оценка качества работы сервисного центра при обслуживании и ремонте подвижного состава автобусного АТП // Вестник Московского автомобильно-дорожного государственного технического университета (МАДИ). 2021. № 4 (67). С. 10-17.
9. *Котомчин А.Н., Синельников А.Ф., Корнейчук Н.И.* Использование износостойкого хромирования при восстановлении и упрочнении деталей автомобилей // Вестник Московского автомобильно-дорожного государственного технического университета (МАДИ). 2021. № 1 (64). С. 11-17.
10. *Мороз С.М.* Запас работоспособности – комплексный показатель надежности автотранспортных средств // Вестник Московского автомобильно-дорожного государственного технического университета (МАДИ). 2020. № 3 (62). С. 3-10.
11. *Беляев А.И., Евтюков С.А.* Формирование параметров поверхностного слоя деталей дорожно-строительных машин при восстановлении // Вестник Московского автомобильно-дорожного государственного технического университета (МАДИ). 2020. № 2 (61). С. 31-38.
12. *Арифуллин И.В., Терентьев А.В., Таланова И.Н.* Многокритериальная аналитическая модель определения срока эксплуатации автомобиля // Вестник Московского автомобильно-дорожного государственного технического университета (МАДИ). 2019. № 4 (59). С. 3-8.

ОДНОЧАСТОТНЫЙ ЛАЗЕР НА КРИСТАЛЛЕ ВАНАДАТА С НАКАЧКОЙ ЛАЗЕРНЫМ ДИОДОМ НА ГАРМОНИКАХ ДЛЯ СКАНИРОВАНИЯ АВТОМОБИЛЬНЫХ ДОРОГ И ЗДАНИЙ ПРИ ПРОЕКТИРОВАНИИ И РЕМОНТЕ

Ашурбеков Сефер Ашурбекович,
зав. лаб., к.т.н., доцент, Бронницкий филиал МАДИ,
ashsefer118@gmail.com

Чубова-Белошникова Мария Максимовна,
РТУ МИРЭА, Москва, Россия
dianamatvienko9703@gmail.com

Воеводин Данила Юрьевич,
Бронницкий филиал МАДИ,
danyash13@gmail.com

Суфиянов Ракип Шайхиевич,
зав. каф., д.т.н., доцент, Бронницкий филиал МАДИ
mr.surash@mail.ru

Аннотация

Разработан компактный одночастотный лазер на кристалле $YVO_4: Nd^{3+}$ с накачкой одноаттным лазерным диодом на гармониках 1064 нм и 532 нм с термостабилизацией для сканирования автомобильных дорог и зданий: номинальные параметры лазера – мощность 50 мВт, угловая расходимость $< 0,45$ миллирадиан, диаметр пучка 1,12 мм, эллиптичность 0,98, уровень шумов $< 4\%$. Показана эффективность и надёжность использования ванадатного лазера в нивелирах, дальномерах и при сканировании зданий и автомобильных дорог.

Ключевые слова

Ванадатный лазер, диодная накачка, сканирование дорог и зданий, одночастотный лазер.

Введение

В последние годы обозначился бурный рост твёрдотельных лазеров с диодной накачкой как компактных, малогабаритных лазеров с высоким качеством лазерного пучка и работающих в широкой области оптического спектра [1-6]. В нашей работе реализован лазерный излучатель на одном из самых перспективных активных материалов лазерной техники – на кристалле ванадата $YVO_4: Nd^{3+}$ с накачкой одним лазерным диодом (700-800 нм) с выходной мощностью 1 Вт с нелинейным удвоением частоты, со стабилизацией температуры активного элемента и диода накачки для использования в дальномерах, нивелирах и при сканировании автомобильных дорог и зданий [7-10].

Лазерное сканирование уже набрало широкую популярность за десять лет применения на наших дорогах [11-15]. Сканирование автомобильных дорог – мобильный и высокоточный метод, который может применяться на всех этапах жизненного цикла автомобильных трасс, от инженерных изысканий и территориального планирования до ремонта и реконструкции трассы [16-21].

Функциональная схема излучателя приведена на рисунке 1.

Излучение лазерного диода (3) (700-800 нм) с помощью конденсора (4) и объектива (5) фокусируется на активный элемент (6) в железной оправке для отвода тепла. Активный элемент (6) и зеркала резонатора (8-9), нелинейный удвоитель частоты установлены в корпусе резонатора. Диод накачки питается стабилизированным источником тока 5V (1) через плату управления током диода (2), которая запаивается на основание диода. На выходе излучателя установлен телескоп (13) для формирова-

ния необходимого диаметра пучка и его угловой расходимости. Для стабилизации выходной мощности установлен отводной узел, состоящий из полупрозрачной пластины (11), установленной под углом Брюстера к излучению и фотоприемника (12), сигнал с которого подается в плату управления током (2).

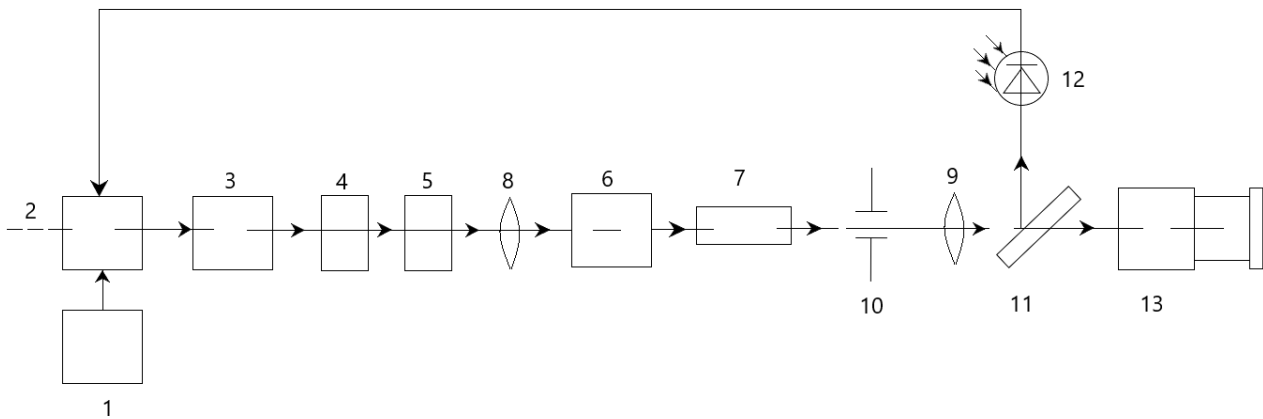


Рис. 1. Функциональная схема лазерного излучателя на кристалле $YVO_4: Nd^{3+}$ с накачкой одним полупроводниковым лазерным диодом и удвоением частоты:

- 1 – стабилизированный источник питания 5V; 2 – плата управления током диода накачки; 3 – лазерный диод накачки; 4 – конденсор; 5 – объектив; 8,9 – зеркала резонатора для второй гармоники 532 нм; 6 – активный элемент; 7 – нелинейный кристалл; 10 – диафрагма; 11 – полупрозрачная пластина; 12- фотоприёмник; 13 – двухкомпонентный телескоп

Отработана методика цифрового анализа профиля лазерного пучка для измерения оптических характеристик с применением ПЗС-камеры и компьютера. В результате получены следующие характеристики лазерного излучателя:

- режим работы – одночастотный;
- длина волны – 1064 нм; 532 нм;
- диаметр пучка 1,12 мм;
- угловая расходимость < 0.45 миллирадиан;
- номинальная мощность излучения в непрерывном режиме < 50 мВт;
- максимальная вариация выходной мощности < 4%.

Исследована возможность сканирования дорог при проектировании и ремонте. Максимальный диаметр лазерного пучка на расстоянии 100 м менее 10 мм.

Заключение

1. Разработан малогабаритный стабильный лазерный источник на кристалле ванадата на гармониках мощностью в непрерывном режиме 50 мВт и с угловой расходимостью меньше 0,45 миллирадиан.
2. Практическая ценность результатов заключается в возможности эффективного использования ванадатного лазера в качестве надежного источника в лазерных дальномерах, нивелирах и для сканирования автомобильных дорог и зданий при проектировании и ремонте.

Литература

1. Вейко В.П., Петров А.А., Самохвалов А.А. Введение в лазерные технологии. Опорный конспект лекций по курсу “Лазерные технологии” под редакцией Вейко В. П. СПб: Университет ИТМО, 2018. 161 с.
2. Вейко В.П., Шагно Е.А. Лазерные технологии в задачах и примерах. СПб: Университет ИТМО, 2014. 88 с.

3. *Комиссаров А.В.* Теория и технология лазерного сканирования для пространственного моделирования территорий. Новосибирск: СГТА, 2015. 103 с.
4. *Пойзнер Б.И.* Физические основы лазерной техники. Учебное пособие. М.: ИНФРО-М, 2018. 160 с.
5. *Звелто О.* Принципы лазеров. Перевод с английского. М.: Мир, 1990. 558 с.
6. *Мотуз В.О., Сарычев Д.С.* Применение лазерного сканирования и 3D-моделей в жизненном цикле автомобильных дорог // САПР и ГИС автомобильных дорог. 2014. №1(2). С. 12-15.
7. *Бойков В.Н., Федотов П.А., Пуркин В.И.* Автоматизированное проектирование автомобильных дорог. М.: МАДИ СГТУ, 2005. 224 с.
8. *Лыгина Л.А., Строчков Д.М., Котов А.А.* Организация и безопасность движения в малых населенных пунктах сельского типа вдоль автомобильных дорог и режимы движения автомобилей // Вестник Московского автомобильно-дорожного государственного технического университета (МАДИ). 2021. № 4 (67). С. 104-113.
9. *Донченко В.В.* Построение безопасных транспортных систем как идеология снижения смертности на автомобильных дорогах // Вестник Московского автомобильно-дорожного государственного технического университета (МАДИ). 2021. № 3 (66). С. 114-121.
10. *Крамаренко Б.А., Болдин А.П.* Возможности статистического прогнозирования изменения общих и частных показателей безопасности дорожного движения при технической эксплуатации транспортных средств на основе использования программы STATISTICA-10-12 // Вестник Московского автомобильно-дорожного государственного технического университета (МАДИ). 2021. № 3 (66). С. 3-12.
11. *Андронов В.Д.* Надежность бетонных покрытий аэродромов после выполнения ремонтных работ // Вестник Московского автомобильно-дорожного государственного технического университета (МАДИ). 2021. № 2 (65). С. 56-60.
12. *Ретин С.В., Шиманова А.А., Лутов Д.А.* Методика повышения эксплуатационной надежности сложного технического объекта посредством анализа его структурной надежности // Вестник Московского автомобильно-дорожного государственного технического университета (МАДИ). 2021. № 1 (64). С. 23-30.
13. *Суладзе М.Д., Федулов В.К., Подопригора В.Г.* Основы управления эксплуатационно-техническим состоянием аэродромных покрытий // Вестник Московского автомобильно-дорожного государственного технического университета (МАДИ). 2020. № 4 (63). С. 64-71.
14. *Ле Д.Л.* Автомобилизация крупных агломераций и развитие транспортной инфраструктуры // Вестник Московского автомобильно-дорожного государственного технического университета (МАДИ). 2020. № 1 (60). С. 3-10.
15. *Корочкин А.В.* Прогнозирование требуемой прочности дорожной одежды автомобильных магистралей // Вестник Московского автомобильно-дорожного государственного технического университета (МАДИ). 2020. № 1 (60). С. 66-71.
16. *Попов В.И., Чан Д.З.* Эффективность применения пенополистирола в насыпях за интегральными устоями мостов // Вестник Московского автомобильно-дорожного государственного технического университета (МАДИ). 2020. № 1 (60). С. 95-100.
17. *Немчинов М.В., Холин А.С.* Принципы проектирования прочных и долговечных дорожных одежд // Вестник Московского автомобильно-дорожного государственного технического университета (МАДИ). 2019. № 4 (59). С. 70-76.
18. *Бураков А.В., Соболев Е.Е.* Развитие технологии санации деформационных швов и трещин жестких аэродромных покрытий // Вестник Московского автомобильно-дорожного государственного технического университета (МАДИ). 2019. № 4 (59). С. 76-81.
19. *Ларина Т.А., Зубарев Н.Р.* Износ асфальтобетонных покрытий под воздействием колес легкового транспорта // Вестник Московского автомобильно-дорожного государственного технического университета (МАДИ). 2019. № 3 (58). С. 55-61.
20. *Попов В.И., Чан Д.З.* Современные способы снижения деформаций насыпей подходов к мостам // Вестник Московского автомобильно-дорожного государственного технического университета (МАДИ). 2019. № 2 (57). С. 96-101.
21. *Ушаков В.В., Горячев М.Г., Жустарёва Е.В.* Конструктивные особенности нежестких дорожных одежд федеральных автомобильных дорог России со слоями из асфальтобетона // Вестник Московского автомобильно-дорожного государственного технического университета (МАДИ). 2019. № 1 (56). С. 90-94.