

Министерство образования Республики Беларусь

Учреждение образования  
Белорусский государственный университет  
информатики и радиоэлектроники

УДК 004.855.5

Гусак  
Янина Олеговна

## **Решение прикладных задач методами машинного обучения**

### **АВТОРЕФЕРАТ**

на соискание академической степени  
магистра информатики и вычислительной техники

по специальности 1-40 81 04 – Обработка больших объемов информации

Научный руководитель  
Егорова Н.Г.  
кандидат технических наук,  
доцент

Минск 2019

## КРАТКОЕ ВВЕДЕНИЕ

Потенциал машинного обучения к упрощению и удешевлению многих процессов обуславливает неугасающий интерес к данной области, что в свою очередь обуславливает ее развитие. Распространение автоматизации повышает качество жизни, облегчает и ускоряет производство, сокращает расходы на человеческие ресурсы, экономит время. Даже небольшая деталь, будучи избавленной от ручного труда, на одном из этапов производства может помочь сэкономить значительную сумму. Заинтересованность бизнеса в автоматизации обуславливает и заинтересованность научной среды в развитии области машинного обучения.

Одним из бизнесов, остро нуждающихся в автоматизации процессов, является дистрибуция розничной продукции в странах, где существует иерархия компаний для продвижения некоторых видов продукции. Во многих странах, в частности в США, компания, производящая продукцию, в общем случае не может поставлять ее торговым сетям напрямую. Такое правило обусловлено противодействием образованию монополий. Движения товара к торговым сетям осуществляют компании-дистрибьюторы. В этом секторе также существует конкуренция. В ходе работы компаниям-дистрибьюторам необходимо заключать контракты с поставщиками на закупку продукции и торговыми сетями на реализацию продукции. В схемах дистрибуции продукции существует множество условий: количество свободных складов, количество поставленной продукции, количество торговых точек, в которые необходимо поставить продукцию, какое количество уже поставлено, как часто необходимо возобновлять поставки. Учет этих условий, как правило, ведётся в обычных таблицах, которые постоянно должны обновляться и дополняться специалистами компании-дистрибьютора. Развозом продукции по торговым точкам занимаются торговые представители. Они не только доставляют продукцию к месту назначения, но и расставляют ее на торговых площадках. Для подтверждения выполнения работы торговые представители должны сделать фотографию расставленной продукции и снабдить фото информацией о торговой точке, о торговых марках и количестве продукции. Затем эта фотография отправляется в компанию-дистрибьютор для того, чтобы на основе информации из фото обновить таблицы.

Для облегчения работы этого сектора существует система, которая позволяет компаниям ускорить множество этапов процесса дистрибуции. Для торговых представителей существует специальное мобильное приложение, в котором есть функция, позволяющая сделать фото и снабдить его необходимой информацией. На данный момент фотографии не снабжаются информацией о

ценах на продукцию, так как это может быть затратным по времени для торговых представителей - указывать цены всех видов выставленной продукции. Было принято решение создать алгоритм, который мог бы определять ценники на изображении. Для решения данной проблемы нет общедоступных наборов данных. Попытка решения задачи в условиях недоступности набора размеченных данных была предпринята в предыдущей работе над данной задачей. В данной работе рассматривается решение задачи с разметкой небольшого количества изображений и построением моделей машинного обучения на получившемся наборе.

## **ОБЩАЯ ХАРАКТЕРИСТИКА РАБОТЫ**

### **Цель и задачи исследования**

Целью работы является разработка решения для задачи определения положения объектов на изображении и последующая их классификация в условиях малого количества доступных данных. Результатом должен быть набор компонент, способных локализовать ценники на изображении, выделить их и распознать цену.

Решение должно быть реализовано на языке программирования Python. Исходными данными являются фотографии стендов с товарами и ценниками в формате RGB без какой-либо дополнительной информации. Точность решения является более приоритетной по сравнению со скоростью его работы.

Для достижения поставленной цели необходимо решить следующие задачи: проанализировать современное состояние области машинного обучения и анализа данных, выделить этапы решения задачи и подходы к решению каждого этапа, определить подход к разметке данных и осуществить разметку, провести эксперименты, сравнить результаты и сделать вывод о том, какими методами и моделями машинного обучения можно наилучшим образом (в заданных ограничениях) решить поставленную задачу.

Объектом исследования является выбранная прикладная задача.

Предметом исследования – применение алгоритмов машинного обучения для решения поставленной задачи.

### **Связь работы с приоритетными направлениями научных исследований и запросами реального сектора экономики**

Исследуемые в работе подходы для анализа изображений позволяют автоматизировать процесс выделения данных о ценниках из изображений стендов с товарами, что повлечет за собой сокращение времени на сбор и агрегацию информации, что в свою очередь приведет к экономии средств на оплату труда сотрудников, которые на текущий момент занимаются сбором данной информации вручную.

Работа выполнялась в соответствии с научно-техническим заданием и планом работ кафедры Информатики и компании “Леверекс Интернешнл”. В качестве прикладной задачи была выбрана задача детектирования ценников на изображении и последующего распознавания цифр на ценнике. В рамках исследовательской работы над выбранной задачей была изучена предметная область машинного обучения и компьютерного зрения. Было разработано

решение для локализации ценников и исследованы подходы к распознаванию их содержимого с учетом недоступности размеченных данных.

### **Личный вклад соискателя**

Результаты, приведенные в диссертации, получены соискателем лично. Вклад научного руководителя Н.Г. Егоровой, заключается в формулировке целей и задач исследования.

### **Опубликованность результатов диссертации**

По теме диссертации опубликованы две печатные работы в сборниках трудов и материалов международных научных конференции.

### **Структура и объем диссертации**

Диссертация состоит из введения, общей характеристики работы, трех глав, заключения, списка использованных источников, списка публикаций автора и приложений. В первой главе находится обзор предметной области: компьютерного зрения и машинного обучения. Во второй главе приведена информация об используемых технологиях и их возможностях. В третьей главе дано последовательное описание этапов работы над решением поставленной задачей: разбиение на подзадачи (разметка данных, локализация объектов, классификация) и подходы к решению каждой подзадачи. Приведено также сравнение результатов работы моделей локализации объектов, а также описаны недостатки подходов к дальнейшей классификации содержимого объектов.

Общий объем диссертации – 66 страниц. Работа содержит 10 формул и 38 рисунков, одну таблицу. Библиографический список включает 34 наименования.

## ОСНОВНОЕ СОДЕРЖАНИЕ

Во введении рассмотрено приведено поверхностное описание потребности бизнеса в автоматизированных решениях и в решении, рассматриваемом в работе, в частности. В первой главе рассмотрены история современное состояние области машинного обучения и компьютерного зрения, в частности большое внимание уделяется сверточным нейронным сетям, особенностям их обучения. В третьей главе задача разбивается на 3 подзадачи: выбор подхода к разметке данных (обрамляющие прямоугольники для обучения моделей сегментации или же или же выделение контуров для обучения модели сегментации), выбор модели сегментации (UNet) или детекции (RetinaNet), исследование подходов к распознаванию содержимого ценников после того, как они были локализованы в результате решения предыдущей подзадачи. В качестве решения первой подзадачи была выбрана разметка полигонами, максимально близко к границам ценника так, чтобы по таким полигонам можно было составить маски для решения задачи сегментации и в то же время по точкам этих полигонов можно было бы выделить обрамляющие прямоугольники для подачи на обучение модели детекции. В результате было размечено 835 изображений. При исследовании второй подзадачи было выявлено, что модель сегментации показывает лучшее качество при количестве обучающих данных 635 изображений. В рамках третьей подзадачи были рассмотрены четыре подхода к распознаванию цены на ценнике, однако в условиях отсутствия обучающих данных для этой задачи, ни один из подходов, использующий сторонние данные или инструменты оптического распознавания символов, не показал устойчивого качества. В заключении приведены краткие результаты работы и намечены возможные дальнейшие шаги для решения задачи распознавания содержимого локализованных ценников.

## ЗАКЛЮЧЕНИЕ

В ходе проведенной работы были исследованы подходы к локализации объектов интереса на изображении с использованием моделей нейронных сетей для детектирования и сегментации. Путем проведения экспериментов с использованием зафиксированной валидационной выборки было выяснено, что при наличии небольшой обучающей выборки (635 изображений, размеченных вручную в рамках работы), модель сегментации UNet показывает лучшие результаты даже с учетом того, что разметка изображений проводилась по контуру искомым объектов, а путем обводки полигоном из четырех точек. В результате, лучшая модель показала коэффициент Серенсена-Дайса на валидационной выборке (200 изображений) равный 0.8040. Потенциально, при наращивании размера тренировочной выборки имеет смысл перейти на модель детекции с глубокой архитектурой для извлечения признаков, например, ResNet.

После выделения ценников стоит задача классификации содержимого (цен) с учетом отсутствия размеченного набора данных из этого домена. В рамках данной подзадачи были исследованы четыре различных подхода, ни один из которых не привел к устойчивому решению в силу набора недостатков, связанных, как правило, с недоступностью набора размеченных данных близкому к домену (изображения с ценником с размеченным содержанием) и необходимостью использовать сторонние общедоступные наборы данных (в частности набор с номерами домов).

Однако благодаря успешной реализации подхода к локализации ценников с помощью сегментации, доразметка данных упрощается и, как минимум, один этап внедрен в работу. Следующим шагом в направлении распознавания содержимого ценников может стать ручная разметка цифр для построения модели близкой к домену, которая также может быть упрощена использованием одного из подходов к локализации и распознаванию цифр, исследованных в данной работе.

## СПИСОК ОПУБЛИКОВАННЫХ РАБОТ

1-А. Гусак, Я. О. Использование алгоритмов машинного обучения в анализе текстов естественного языка на примере определения вопросов-дубликатов / Я. О. Гусак, А. А. Шлеменков // Дистанционное обучение – образовательная среда XXI века : материалы X международной научно-методической конференции (Минск, 7 - 8 декабря 2017 года). – Минск : БГУИР, 2017. – С. 205 - 206.

2-А. Гусак, Я. О. Использование многоуровневой модели для эффективного управления дамбой и предсказания наводнений / Я. О. Гусак, А. А. Шлеменков // Информационные технологии и системы 2018 (ИТС 2018) = Information Technologies and Systems 2018 (ITS 2018) : материалы международной научной конференции, Минск, 25 октября 2018 г. / Белорусский государственный университет информатики и радиоэлектроники ; редкол. : Л. Ю. Шилин [и др.]. – Минск, 2018. – С. 260 - 261.