

Способы распознавания категорий товаров при загрузке на маркетплейс

Александрович Юлия Геннадьевна

Российский экономический университет имени Г.В. Плеханова

Студент

Аннотация

В данной статье описаны способы, которые можно использовать для определения категории товара при загрузке на маркетплейсы. Автором рассмотрены статья других авторов по теме Deep Learning и рассмотрены возможные реализации, а также плюсы и минусы данных подходов. Более детально приведена информация о Глубоком обучении, которое было выбрано для распознавания категорий товаров. Также в данной статье рассмотрены популярные фреймворки, которые могут быть взяты для реализации распознавания категории товара по фотографии.

Ключевые слова: Deep Learning; Фреймворки Deep Learning, Text Mining, bounding box, распознавание, фотографии.

Types of product category recognition when uploading to the marketplace

Alexandrovich Yulia Gennadievna

Plekhanov Russian University of Economics

Student

Abstract

This article describes the methods that can be used to determine the product category when uploading to marketplaces. The author reviewed an article by other authors on the topic of Deep Learning and considered possible implementations, as well as the pros and cons of these approaches. More detailed information is provided about Deep Learning, which was selected for the recognition of product categories. Also, this article discusses popular frameworks that can be used to implement product category recognition based on a photo.

Keywords: Deep Learning; Deep learning frameworks, Intelligent Text analysis, bounding box, recognition, photos.

Введение

Сегодня слово «маркетплейс» у всех на слуху. Во время пандемии люди в полной мере ощутили все возможности, которые может предоставить им сектор e-commerce. Это касается тех, кто приобретает услуги, товары на площадке и тех, кто их предоставляет и продает. Более того, это стало очень популярно и появились различные курсы и специалисты, которые могут

помочь тебе выйти на рынок и торговать товарами. Также перспективы для себя оценивают крупные компании и тоже интегрируются с маркетплейсами.

В случае таких крупных компаний все основные процессы уже настроены. Там есть и 1С, Мой Склад, Внутреннее самописное ПО и другое. И такие компании имеют тысячи позиций товаров. При работе с маркетплейсом в таком случае такие компании используют API, чтобы автоматизировать свои процессы. Основные задачи, которые преимущественно хотелось бы решать без ошибок – загрузка карточек товаров и их успешное прохождение модерации, так как это занимает большое количество времени и влияет на продажи и начало работы. При анализе работы API на маркетплейсе – это является узким горлом работы с API. Исходя их данных проблем сторонние компании разработали скрипты и способы того, как упростить загрузку товаров и предоставляют это в качестве платной услуги. Настолько это проблема кажется явной.

Цель статьи - рассмотреть методы и способы реализации распознавания категории по загруженным фотографиям, актуальные на текущий момент и доступные для того, чтобы упростить работу с API и загрузкой товаров.

Основная часть

Распознавание изображений, текста, видео все чаще интересует крупные компании. Это может помочь исследовать целевую аудиторию и сделать маркетинг инновационным. С помощью распознавания также можно провести анализ пространства магазина и даже дать оценку эффективности рекламы в оффлайне. Распознавание штрих-кодов используется для автоматической идентификации товаров.

Если говорить о распознавании категорий товаров или самих товаров при загрузке для более личного использования, то такие статьи не были найдены. Можно предположить, что глубоким обучением и распознаванием интересуются только довольно крупные компании для обширных проектов, которые могут быть распространены на большие масштабы пользователей, возможно из-за сложности реализации и ограничения ресурсов. Распознавание продуктов относится к использованию технологий, которые в основном основаны на методах компьютерного зрения, поэтому компьютеры могут заменить процесс ручной идентификации и классификации продуктов.

Для автоматизации загрузки и определения категории товаров можно использовать следующие методы: Text Mining, Object Detection.

- 1) Если мы говорим о распознавании категории с помощью Text Mining, то в качестве решения мы можем получить следующую реализацию.

Будет создан скрипт предсказания категории товаров. Он будет вызываться отдельным методом API и иметь несколько уровней точности предсказания. Для получения предсказания пользователю необходимо будет передать наименование товара. То есть определение категории будет

происходить по распознаванию введенного текста. Модель должна будет обучиться на каталоге и, увидев наименование товара, предсказать его категорию.

В чем плюсы данного подхода?

Пользователи переключают часть работы на модель. С их стороны не нужно загружать категорийное дерево и разбираться с тем, как это правильно должно работать. Они грузят только наименования товаров и предпочтительную достоверность распознавания категории.

В чем минусы данного подхода?

Пользователю также необходимо использовать дополнительный метод и далее его переиспользовать. При недостаточной точности предсказания введенной категории может быть ошибка модерации и тогда пользователю придется вернуться в самое начало или использовать стандартный метод получения категорий. Ухудшается пользовательский опыт.

2) Если говорить об использовании Object Detection, то в качестве решения мы сможем получить следующую реализацию.

Пользователю не нужно будет вызывать специальные методы для выгрузки или определения категории. Ему необходимо будет загрузить всю информацию по карточке товара, обязательно приложив фотографии.

На основании фотоизображения должна будет определена категория товара. От корневой до самой детализированной.

В чем плюсы данного подхода?

Упрощается работа клиента с API, с загрузкой товаров. Сокращается время на загрузку товаров на маркетплейсы. Улучшается пользовательский опыт.

В чем минусы данного подхода?

При ошибке модели будут происходить ошибки модерации. Показатель точности модели должен быть на высоком уровне. При появлении совершенно новых уникальных товаров могут происходить ошибки в определении категории.

В своей статье «Deep Learning for Retail Product Recognition: Challenges and Techniques» Yuchen Wei , Son Tran , Shuxiang Xu ,Byeong Kang , and Matthew Springer подробно рассматривают тему Глубокого изучения распознавания продуктов.[1] В ней они представляют всесторонний обзор литературы о недавних исследованиях по распознаванию розничных продуктов на основе глубокого обучения. Данная информация помогла в уточнении некоторых нюансов и решений категоризации розничных товаров.

С помощью статьи были изучены методы компьютерного зрения:

1. Классический (Захват изображения -> Предварительная обработка изображений -> Извлечение признаков -> Классификация признаков -> Результат распознавания)

2. Deep learning

3. CNN (используется для классификации изображений с 1988 года, структура и методы улучшились в связке с датасетом ImageNet)

Также были рассмотрены проблемы, связанные с данными методами и общее представление о том, с какими трудностями сталкиваются исследователи при распознавании категорий розничных товаров.

1. Детекторы. Основная задача компьютерного зрения – научить машину видеть и распознавать объекты. Детекторы – основа многих алгоритмов компьютерного зрения. Если знать, как детектировать объекты – можно решать большое количество задач компьютерного зрения. Современные методы PASCAL VOC содержит 20 классов объектов, а MS COCO содержит фотографии 80 категорий объектов. То есть текущие детекторы объектов не подходят для непосредственного применения к распознаванию продуктов из-за их ограничений категорий.
2. Выбор объектов. Для построения модели необходимо большое количество значений в датасете. Данные объекты возможно необходимо собственноручно размечать следующим образом:

На входе: изображение

На выходе: набор из bounding box, каждый из которых представлен в виде:

(Cx, Cy) - координаты центра;

h – высота;

w – ширина;

class – имя класса;

[confidence – вероятность класса].

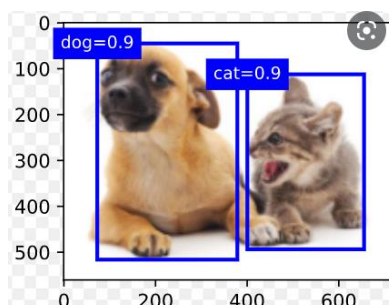


Рисунок 1 – Пример определения класса и вероятность класса

Создание набора данных с достаточным количеством обучающих данных для глубокого обучения требует много времени. Для решения задач необходимо собрать датасет под задачу. Когда мы обучаем нейросеть – она находит общие паттерны в картинках. И далее на основании этих паттернов обучаются классификаторы. Картинки должны быть разных размеров и качества, чтобы нейросеть не привыкала к одному шаблону.

3. Данные. Некоторые готовые наборы данных о розничных продуктах содержат меньше изображений и больше классов. Поэтому необходимо предоставить большой набор данных для обучения модели глубокого обучения, когда необходимо, чтобы эта модель могла распознавать объекты из различных категорий.

За последнее время Deep Learning, в особенности если брать компьютерное зрение, значительно продвинулось и стало основным решением для классификации изображений и обнаружения объектов.

Основное различие между глубоким обучением и традиционными методами распознавания образов заключается в том, что первые могут напрямую изучать признаки из данных изображения, а не использовать созданные вручную признаки. Еще одно отличие глубокого обучения — более глубокие слои, которые могут извлекать более точные признаки, чем традиционные нейронные сети.

Вышеуказанные преимущества позволяют методам Deep Learning предлагать новые идеи для решения некоторых важных проблем компьютерного зрения, таких как сегментация изображений и обнаружение ключевых точек.

В настоящее время применение Deep Learning для распознавания розничных товаров используется для:

1) Классификация изображений: это фундаментальная задача компьютерного зрения, которая стремится разделить разные изображения на разные категории.

2) Обнаружение объектов: это относится к обнаружению объектов с прямоугольниками при категоризации изображений. За последние несколько лет, в связи с непрерывным ростом Deep Learning, многие ученые и разработчики создали и оптимизировали некоторые фреймворки Глубокого обучения, чтобы ускорить процедуры обучения и прогнозирования. Например, TensorFlow, PyTorch, MXNet, Caffe. Это самые распространенные фреймворки и они упрощают работу с Deep Learning.

Как говорилось ранее, для построения модели глубокого обучения можно использовать готовые фреймворки. Особенно важно обратить внимание на следующие моменты:

1. Возможность решать задачи компьютерного зрения и распознавания изображений.
2. Дополнительные/готовые Инструменты
3. Документирование и поддержка сообщества

В таблице 1 рассмотрены фреймворки Deep Learning.

Таблица 1 – Сравнение фреймворков Deep Learning

	Tensorflow[3]	PyTorch[2]	Keras[4]	Caffee[5]
Описание	Используется для решения задач построения и тренировки нейронной сети с целью автоматического нахождения и	Используется для решения различных задач: компьютерное зрение, обработка естественного языка.	Настройка над фреймворком TensorFlow. Обеспечивает взаимодействие с искусственным и нейронными сетями.	Некая среда для глубинного обучения

	Tensorflow[3]	PyTorch[2]	Keras[4]	Caffee[5]
	классификации образов			
Порог вхождения	Средний	Низкий	Низкий	Средний
Язык	Python	Python, C++ / Java	Python	C++, интерфейс Python
Поддержка, сообщества	+	+	+	Ограниченная
Достаточная документация	+	-	+	+
Распределённое обучение	+	+	+	+
Комментарии		Фреймворк имеет много предварительно обученных моделей и готовых модульных частей. Их легко между собой комбинировать	Фреймворк содержит множество инструментов для упрощения работы с изображениями и текстом. Подходит для новичков	Используется в компьютерном зрении, распознавании речи и мультимедиа. Есть отдельная презентация о возможностях Caffee в распознавании изображений с примерами использования

Заключение

Мы рассмотрели методы, которые смогут помочь нам с задачей распознавания категорий. Используя детекцию объектов, наряду с рассмотренными фреймворками Deep Learning мы сможем построить модель распознавания категорий товаров по фотографии. Также мы можем использовать Сверточные нейронные сети (Convolution Neural Nets), они также помогут с распознаванием объектов и их классификацией.

Данные модели должны упростить загрузку товаров и модерацию карточек товаров маркетплейса и улучшить пользовательский опыт пользователей, которые собираются интегрироваться к e-com площадке или уже пользуются API, но с ограниченной функциональностью.

Библиографический список

1. Deep Learning for Retail Product Recognition: Challenges and Techniques
URL: <https://www.hindawi.com/journals/cin/2020/8875910/>
2. FROM RESEARCH TO PRODUCTION. URL: <https://pytorch.org/>
3. Сравнение фреймворков для глубокого обучения: TensorFlow, PyTorch,

Keras, MXNet, Microsoft Cognitive Toolkit, Caffe, etc URL:
<https://habr.com/ru/company/otus/blog/443874/>
4. Caffe URL: <http://postoev.ru/caffe/>