

Здравствуйте! Моя тема по ВКР : “Применение нейронных сетей для оценки трафика магазина” 【1】

Я расскажу об этом с пяти сторон. И так начнем. 【2】

Актуальность

Проблема статистики покупательского потока — большая трудность для бизнес-анализа магазинов в течение длительного времени. Можно сказать, что технология оценки потока посетителей является общим требованием всей отрасли. Однако в нашей реальной жизни я заметила такое явление, что большинство магазинов все таки используют традиционные методы подсчета потока пометителей. Традиционные методы включают себе ручные счетчики, автоматические штативы и инфракрасные датчики. Но они часто являются неточным и ненадежным. И самое главное, с помощью этих методов невозможно точно узнать время, когда каждый покупатель заходит в магазин, и нет возможности провести более тщательный бизнес-анализ. Поэтому разработка автоматической системы оценки трафика магазина крайне важна. 【3 4】

Цель, задачи

Целью данной статьи является разработка новой системы оценки трафика магазина на основе технологии нейронных сетей. Для реализации такой цели нужно разрабатывать метод, который сочетает в себе алгоритмы обнаружения, отслеживания и повторной идентификации посетителей. В частности, необходимо реализовать алгоритмы YOLOv5, FastReID и DeepSORT, разработать систему на основе этих алгоритмов и провести экспериментальные оценки всех методов. 【5】

Теперь посмотрим Архитектуру моделей. Буду рассказывать о модули повторной идентификации и модули отслеживания посетителей. 【6】

ReID

Повторная идентификация людей

1. Повторная идентификация людей, также известная как Re-идентификация (или просто ReID) — это новая технология, появившаяся в области интеллектуального анализа видео в последние годы. Re-идентификация людей может быть определена как задача присвоения одного и того же имени или индекса всем образам одного и того же человека, получаемым с пространственно-разнесенных камер, области видимости которых не пересекаются друг с другом, на основе выделения и анализа признаков изображений человека.
2. При выполнении задачи Re-идентификации мы обычно разделяем все изображения людей на две группы: набор запросов и набор галерей. Набор запросов: коллекция изображений, которые мы хотим использовать для поиска подобных изображений в наборе галерей. Набор галерей (Gallery): коллекция всех доступных изображений для поиска. А запрос (probe или query image), это одно изображение или группа связанных изображений для поиска. Цель ReID заключается в том, чтобы найти все изображения людей в наборе галереи, которые похожи или совпадают с

изображениями в наборе запроса, и могут быть точно аутентифицированы или идентифицированы.

3. Будем использовать глубокое обучение для реализации person reid. В частности это метрическое обучение. В исходном пространстве R^F определим функцию отображения $f(x)$, которая преобразует изображение человека x в признаковые представления в пространстве признаков. Далее еще определим такую функцию расстояния D для измерения расстояния между парами признаков представлений. Цель метрического обучения, это найти такие оптимальные отображение $f(x)$ и функцию расстояния D , чтобы расстояние между двумя изображениями одного и того же пешехода было как можно меньше, а расстояние между двумя изображениями разных пешеходов — как можно больше. И это отображение $f(x)$ является обученной глубокой сверточной сетью. Имеются разнообразные модели повторной идентификации. Среди них мы выбрали алгоритм FastReID. 【7】

FastReID

1. Модель FastReID основана на методах глубокого обучения и предназначена для решения задачи Re-идентификации людей.
2. FastReID использует сверточные нейронные сети для извлечения признаков из изображений. Сеть состоит из двух основных компонентов: экстрактора признаков и классификатора. Экстрактор признаков отвечает за извлечение признаков из изображений, которые могут быть использованы для идентификации объектов. Для этого используются предобученные архитектуры нейронных сетей, такие как ResNet или MobileNet. Классификатор отвечает за вычисления расстояния между признаками, ранжирование среди расстояний и присвоение идентификаторов объектам на основе ранжирования.
3. Итак, как обучить такую сеть? На основе этих 4 принципов? Первое: предварительная обработка изображения. Произвольно применять к изображениям преобразование флип, стирание. Второе: использовать предобученную сеть ResNet50 для извлечения функций изображения. Третье: минимизировать потери Triplet Loss и потери Cross Entropy Loss для обучения сети. Четвертое: использовать оптимизатор Адам.
4. Это мой процесс обучения на моем компьютере, и обучение на наборе данных заняло почти 8 часов. 【8】

MOT

Нам нужно сформировать треки для каждого посетителя. Это требует решить задачу отслеживания множества объектов, то есть задача MOT. Задача MOT заключается в сопоставлении детекция объекта на последовательности кадров видеопотока треку объектов. Как показано на рисунке, алгоритм MOT это обнаружение + Ассоциация данных. Пусть мы уже имеем треки в предыдущим кадрах. Теперь в текущем кадре после обнаружения мы получим местоположения всех людей. Нужно ассоциации обнаруженных и треков, чтобы создать новые треки. Повторим обнаружение *~~~*и ассоциацию до конца видеопотока. Таким образом, мы получим треки каждого обнаруженного объектов. В проекте я реализовала трекер на основе алгоритма DeepSORT. 【9】

DeepSORT

1. Полное название DeepSORT — Простое онлайн-отслеживание в реальном времени с метрикой глубокой ассоциации. В настоящее время это популярный алгоритм онлайн-отслеживания в отрасли.
2. На картинке желтая часть — это задача, выполненная DeepSORT. DeepSORT в основном разделен на три части: предсказание Калмана, ассоциативные метрики и венгерское назначение. Непрерывно вводим кадры и результаты детекции в DeepSORT, в качестве выхода созданы траектории каждого объекта в видео. Система оценки трафика магазина
3. На основе YOLOv5 и упомянутого выше FastReID, DeepSORT я спроектировала систему оценки трафика магазина. 【10】

Система оценки трафика магазина

Наша система состоит из нескольких камер и сервера, на котором работает модель нейронных сетей. Устанавливаем камеры в интересующих отделах магазина – например вход, выход. Запускаем все камеры одновременно, и эти камеры будут записывать видеопоследовательности. На нашей модели проходить обработки всех видеопоследовательностей. В результате будет создан отчет, где записано в какое время, какой покупатель проходить какие отделы магазина. 【11 12】

Блок схема системы

Далее мы опишем во временном порядке поведение системы, когда покупатель входит в магазин, проходит через несколько отделов магазина и уходит. Преобразуем записанные камерами видео в кадры для обработки. Когда в кадре появляются люди, модель YOLO обнаружит посетителей и вернет местоположение всех посетителей в кадре. Получили кадр, где все посетители с рамками. Затем модель DeepSORT генерирует треки для каждого посетителя. С этими траекториями будем определять, входил ли пешеход в магазин. Сохраним векторы признаков всех покупателей, входящих в магазин, которые позже будут использоваться для сравнения и запроса. Теперь покупатели вошли в магазин и прошлись по разным отделам. 【13】

Демонстрация системы - Сцена 1: человек входит в магазин Демонстрация системы - Сцена 2: человек выходит из магазина 【14 15】

Проведенные тестирования

Я провела экспериментальные оценки всех используемых моделей, FastReID, DeepSORT и систему. 【16】

FastReID result

1. Я обучил свою собственную модель FastReID на наборе данных MSMT17. На наборах данных MARKET-1501 и DukeMTMC протестировала производительность модели и получила следующие результаты.
2. Измерение проводилось с использованием двух показателей точности – Rank-1 и mAP. Rank-1 это точность первого совпадения, а mAP это средняя точность на целом наборе данных.

3. Результаты экспериментальной оценки приведены в таблице. По результатам экспериментальной оценки видно, что модель FastReID показывает очень хорошие Rank-1 и mAP на обоих наборах. За исключением того, что FastReID немного ниже Circle Loss на ранг-1 на MARKET1501, все показатели у FastReID превосходят аналоги. 【17】

DeepSORT

Мы используем набор MOT16, MOT17 в качестве тестовых наборов для проверки производительности модели DeepSORT. Для оценки мы использовали общие показатели оценки в области многоцелевого отслеживания: MOTA, MOTP, MT, ML, IDs и IDF1. Самые важные из них - MOTA и MOTP. MOT-16 : Результаты анализа на наборе MOT-16: Мы улучшили DeepSORT с помощью FastReID и YOLOv5. По сравнению с исходным алгоритмом DeepSORT видно, что MOTA, MOTP существенно превышает на 4.8% и 1.7%. Количество изменений ID уменьшится на 21 раз. Это скажет, что была повышена возможность сохранения идентификатора пешехода модели. Кроме того, замечено, что производительность в реальном времени немного хуже, чем у алгоритма JDE, потому что JDE является одноэтапным алгоритмом отслеживания, а отслеживание в реальном времени относительно высоко, но переключение идентификаторов относительно частые, что обусловлено взаимным перекрытием мишеней. В розничной среде часто бывают сцены с большим количеством плотных пешеходов, поэтому наш алгоритм является подходящим выбором в это время. MOT-17 : Результаты анализа на наборе MOT-17: На этом наборе наша модель имеет наилучшие MOTA, MT и ML. Метрики IDF1 и IDs чуть хуже у LPC_MOT, но разности не большие. 【18】

Система

Для проверки работоспособности различных функций в этой системе тест проведен в крупном супермаркете ОКЕИ. Мы развернули 3 фиксированные камеры по траектории посетителей, входящих в супермаркет, для получения видео. План этажа супермаркета и расположение трех камер показаны на рисунке. Продолжительность видеозаписи около 2 минут. В общей сложности 20 человек прошли под три камерами. Из них 17 были полностью отслежены, что означает, что они были правильно идентифицированы и отслежены под трех камерами. Два человека не были полностью отслежены из-за пропущенных обнаружений, поэтому они не учитывались в полном отслеживании. Один человек был повторно распознан, но из-за недостаточной степени сходства не удалось восстановить его первоначальный ID, поэтому он также не учитывался в полном отслеживании. Таким образом, всего 20 человек, из которых 17 были успешно полностью отслежены. Мы можем вычислить точность отслеживания системы - это 85%. 【19】

ЗАКЛЮЧЕНИЕ

В данной статье был рассмотрен метод оценки трафика магазина на основе нейронных сетей. Были разработаны алгоритмы отслеживания, повторной идентификации посетителей и подсчета количества посетителей магазина. Результаты показали, что система имеет высокую точность 85%. Однако, при большом скоплении людей точность отслеживания может снижаться. При этом в дальнейшем планируется улучшать алгоритм отслеживания, например, использовать обнаружение центральной точки вместо традиционного обнаружения ограничивающей рамки. 【20】

Спасибо

Результаты работы были представлены на конференции. Доклад отмечен дипломом 1 степени. Всем спасибо за внимание! 【21】