



Министерство науки и высшего образования Российской Федерации
Мытищинский филиал
Федерального государственного бюджетного образовательного
учреждения
высшего образования
«Московский государственный технический университет
имени Н.Э. Баумана
(национальный исследовательский университет)»
(МФ МГТУ им. Н.Э. Баумана)

ФАКУЛЬТЕТ _____ИНФОРМАТИКА И СИСТЕМЫ УПРАВЛЕНИЯ_____

КАФЕДРА _____СИСТЕМЫ ОБРАБОТКИ ИНФОРМАЦИИ И УПРАВЛЕНИЯ_____

РАСЧЕТНО-ПОЯСНИТЕЛЬНАЯ ЗАПИСКА К НАУЧНО-ИССЛЕДОВАТЕЛЬСКОЙ РАБОТЕ НА ТЕМУ:

Применение каскада обучающихся моделей для распознавания и региональной классификации китайских автомобильных номеров

Студент ИУ5И-31М
(Группа)

(Подпись, дата)

Ли Лююй
(И.О.Фамилия)

Руководитель курсовой работы

(Подпись, дата)

Ю.Е. Гапанюк
(И.О.Фамилия)

2024 г.

Министерство науки и высшего образования Российской Федерации
Федеральное государственное бюджетное образовательное учреждение
высшего образования
«Московский государственный технический университет имени Н.Э. Баумана
(национальный исследовательский университет)»
(МГТУ им. Н.Э. Баумана)

УТВЕРЖДАЮ
Заведующий кафедрой ИУ5
(Индекс)
В.И. Терехов
(И.О.Фамилия)
« 04 » декабря 2024 г.

ЗАДАНИЕ
на выполнение научно-исследовательской работы

по теме Применение каскада обучающихся моделей для распознавания и региональной классификации китайских автомобильных номеров

Студент группы ИУ5И-31М

Ли Лююй

(Фамилия, имя, отчество)

Направленность НИР (учебная, исследовательская, практическая, производственная, др.)

ПРАКТИЧЕСКАЯ

Источник тематики (кафедра, предприятие, НИР) КАФЕДРА

График выполнения НИР: 25% к ____ нед., 50% к ____ нед., 75% к ____ нед., 100% к ____ нед.

Техническое задание:

Применение технологии компьютерного зрения в области безопасности дорожного движения

Оформление научно-исследовательской работы:

Расчетно-пояснительная записка на 40 листах формата А4.

Перечень графического (иллюстративного) материала (чертежи, плакаты, слайды и т.п.)

Дата выдачи задания « 04 » декабря 2024 г.

Руководитель НИР

(Подпись, дата)

Ю.Е. Гапанюк

(И.О.Фамилия)

Студент

(Подпись, дата)

Ли Лююй

(И.О.Фамилия)

Примечание: Задание оформляется в двух экземплярах: один выдается студенту, второй хранится на кафедре.

каталог

1. Тема дипломной работы	4
2. Обзор литературы.....	4
2.1 Текущий уровень технологий распознавания номеров	4
2.2 Применение машинного и глубокого обучения в распознавании номеров.....	6
2.3 Обзор каскадной модели обучения	8
3. Предварительные предложения по реализации	9
3.1 Подготовка и предобработка данных	9
3.2 Разработка каскадной модели обучения.....	10
3.3 Архитектура и интеграция системы.....	11
3.4 Оптимизация модели	11
4. Задачи, которые необходимо решить.....	12
4.1 Проблемы с набором данных.....	12
4.2 Разработка и обучение каскадной модели.....	13
4.3 Реализация региональной классификации	13
5. Ожидаемые результаты	15
5.1 Высокая точность распознавания номеров	15
5.2 Эффективная региональная классификация	15
5.3 Устойчивость и расширяемость системы.....	15
5.4 Реализация в реальном времени и практическое применение	15
Рекомендации	15

1. Тема дипломной работы

Название темы: Применение каскадной модели обучения для распознавания и региональной классификации автомобильных номеров Китая.

Описание темы: В данной работе будет рассмотрено, как использование каскадной модели обучения может повысить точность распознавания и эффективность региональной классификации автомобильных номеров Китая. Пошаговый механизм обработки каскадной модели обучения особенно подходит для решения сложных задач распознавания номеров, начиная с начального этапа обнаружения и заканчивая точным распознаванием символов, обеспечивая систематизированную классификацию. С помощью данной модели можно эффективно учитывать различия автомобильных номеров разных провинций Китая, такие как форма и расположение символов, что способствует повышению уровня интеллектуализации и эффективности системы управления дорожным движением.

2. Обзор литературы

2.1 Текущий уровень технологий распознавания номеров

Распознавание номерных знаков (License Plate Recognition, LPR) как важная составляющая интеллектуальных транспортных систем (ITS) имеет многолетнюю историю развития. Оно широко применяется в таких областях, как мониторинг дорожного движения, управление парковками, автоматические платежные системы, правоохранительные органы и предотвращение преступности. Основная цель технологии распознавания номерных знаков заключается в захвате изображений номерных знаков автомобилей и их преобразовании в читаемую последовательность символов для дальнейшей обработки. Этот процесс обычно включает три основных этапа: обнаружение номерного знака, сегментация символов и распознавание символов.

На ранних стадиях технологии распознавания номерных знаков в основном полагались на традиционные методы обработки изображений. Эти методы часто базировались на предварительно определённых алгоритмах обнаружения границ или технике сегментации с использованием пороговых значений яркости для распознавания областей номерных знаков. Например, оператор Собеля и метод обнаружения границ Канни были широко используемыми методами раннего обнаружения номерных знаков. Эти методы показывали хорошие результаты в идеальных условиях, но в сложных ситуациях (например, при неравномерном освещении, изменении углов обзора, повреждении или частичном закрытии номерного знака) они демонстрировали слабые результаты, так как сильно зависели от качества предварительной обработки изображения и строго определённых пороговых значений характеристик. На этапе сегментации и распознавания символов традиционные методы в основном полагались на технику сопоставления шаблонов (Template Matching), то есть производили поочерёдное сравнение символов с предварительно определёнными шаблонами символов для их распознавания. Однако этот подход часто оказывался неэффективным при изменении шрифтов, деформации символов или размытых изображениях ¹¹.

С развитием технологий компьютерного зрения методы распознавания номерных знаков, основанные на машинном обучении, постепенно заменили традиционные методы обработки изображений. Введение методов машинного обучения значительно повысило устойчивость и способность к обобщению системы распознавания. Например, такие классификационные алгоритмы, как метод опорных векторов (Support Vector Machine, SVM) и метод k-ближайших соседей (K-Nearest Neighbors, KNN), постепенно начали применяться для распознавания символов. Эти алгоритмы опирались на разработанные вручную признаки изображений (такие как преобразование Хафа, гистограмма ориентированных градиентов (HOG) и другие) для

извлечения признаков, а затем использовали классификаторы для классификации символов. По сравнению с традиционным сопоставлением шаблонов методы машинного обучения обладают большей способностью к обобщению, особенно при сложных фонах и изменяющемся освещении, что позволяет им сохранять высокую точность распознавания.

Тем не менее, методы машинного обучения также сталкиваются с рядом вызовов. Во-первых, этап извлечения признаков имеет решающее значение для конечной производительности алгоритма, а разработка признаков часто требует экспертных знаний и может не подходить для различных сценариев. Во-вторых, модели машинного обучения сильно зависят от большого объёма данных, особенно при различиях в стандартах номерных знаков и шрифтах в разных странах или регионах. Для обучения модели требуется большое количество размеченных данных, что ограничивает применение методов машинного обучения на практике.

С появлением эпохи глубокого обучения свёрточные нейронные сети (Convolutional Neural Network, CNN) кардинально изменили подход к распознаванию номерных знаков. CNN с помощью многоуровневых свёрток способны автоматически извлекать глубокие признаки из исходных изображений без необходимости в ручной разработке признаков. Особенно заметны улучшения производительности моделей глубокого обучения на всех этапах — от обнаружения номерных знаков до сегментации и распознавания символов. Например, модели на основе глубокого обучения YOLO (You Only Look Once) и Faster R-CNN широко используются для задач обнаружения номерных знаков. Эти модели позволяют быстро и точно локализовать области с номерными знаками даже в сложных условиях фона. Кроме того, рекуррентные нейронные сети (Recurrent Neural Network, RNN) и сети долгой краткосрочной памяти (Long Short-Term Memory, LSTM) также были введены для распознавания символов, решая проблему последовательной зависимости при распознавании символов^[1].

Методы глубокого обучения не только значительно повысили точность распознавания номерных знаков, но и улучшили устойчивость моделей. При работе в условиях неравномерного освещения, наклонных углов и поврежденных номерных знаков модели глубокого обучения могут автоматически подстраивать стратегии распознавания, чтобы эффективно справляться с такими ситуациями. Например, в некоторых последних исследованиях использовались генеративные состязательные сети (Generative Adversarial Networks, GANs) для генерации данных номерных знаков с шумом и частичным закрытием, что повысило способность моделей к обобщению в реальных условиях^[1]. Кроме того, использование переноса обучения значительно сократило время обучения моделей, позволяя предварительно обученным моделям быстро адаптироваться к различным задачам распознавания номерных знаков.

Несмотря на значительный прогресс, достигнутый методами глубокого обучения в области распознавания номерных знаков, остаются проблемы при их применении на практике. Во-первых, оперативность системы распознавания номерных знаков является серьёзным узким местом. В системах мониторинга дорожного движения и автоматических платежных системах система распознавания должна завершить обнаружение и распознавание каждого номерного знака в пределах миллисекунд, чтобы обеспечить плавную работу системы. Поэтому задача повышения производительности модели при сохранении высокой точности распознавания является важным вопросом, требующим решения. Во-вторых, не менее значимой является проблема региональной классификации номерных знаков, особенно в таких странах, как Китай, где существует множество провинций с различными форматами номерных знаков. Обеспечение того, чтобы модель могла эффективно распознавать и различать номерные знаки различных регионов, является одним из ключевых направлений оптимизации системы^[1].

Подводя итог, можно сказать, что технология распознавания номерных знаков прошла путь от традиционных методов обработки изображений до машинного и глубокого обучения.

Широкое применение глубокого обучения в задачах обнаружения и распознавания номерных знаков значительно повысило точность и устойчивость системы, особенно в сложных условиях. Однако, несмотря на высокий уровень точности, достигнутый в экспериментальных условиях, в реальных приложениях всё ещё существуют вызовы, такие как оперативность и точность региональной классификации. В будущем, с введением более передовых алгоритмов и расширением доступных наборов данных, системы распознавания номерных знаков станут ещё более автоматизированными и интеллектуальными.

2.2 Применение машинного и глубокого обучения в распознавании номеров

С распространением интеллектуальных транспортных систем (ITS) технологии машинного и глубокого обучения постепенно стали основными методами в области распознавания автомобильных номеров (LPR). Традиционные методы обработки изображений уже не могут справиться с реальными сложными условиями, а модели, основанные на машинном и глубоком обучении, способны значительно повысить точность и устойчивость распознавания номеров, особенно при изменении углов обзора, неравномерном освещении и повреждениях номерных знаков.

Применение машинного обучения в распознавании номеров началось с оптимизации традиционных методов, в частности, в области распознавания символов. Метод опорных векторов (Support Vector Machine, SVM) является одним из первых классических методов машинного обучения, применённых в распознавании автомобильных номеров. SVM выполняет задачу классификации, находя гиперплоскость, которая разделяет различные классы в пространстве признаков. В процессе распознавания символов SVM используется для классификации изображений символов, извлечённых из номерных знаков. Благодаря своей высокой способности к обобщению, SVM особенно хорошо подходит для распознавания символов номерных знаков на небольших наборах данных. В некоторых исследованиях учёные успешно реализовали точное распознавание символов номеров, комбинируя гистограмму ориентированных градиентов (HOG) с SVM ^[2].

Ещё один популярный метод машинного обучения — это метод k-ближайших соседей (K-Nearest Neighbors, KNN). KNN — это метод обучения на основе примеров, который классифицирует образцы на основе вычисления расстояний между ними и образцами из обучающей выборки. В распознавании автомобильных номеров KNN часто используется на этапах сегментации и распознавания символов. Несмотря на простоту реализации, KNN может иметь ограничения при работе с признаками высокой размерности, и по мере увеличения объёма данных его эффективность значительно снижается. Поэтому, хотя KNN применялся на ранних стадиях исследования распознавания номеров, с появлением более продвинутых технологий глубокого обучения его использование сократилось.

Введение глубокого обучения полностью изменило подход к распознаванию автомобильных номеров. Свёрточные нейронные сети (Convolutional Neural Network, CNN), являясь классической архитектурой в глубоком обучении, широко применяются для классификации изображений и задач обнаружения объектов. CNN автоматически извлекают важные признаки из изображений через свёрточные операции, что делает их особенно эффективными в задачах обнаружения номерных знаков и распознавания символов. В отличие от традиционных методов машинного обучения, CNN не зависят от вручную разработанных методов извлечения признаков, а обучаются на нескольких свёрточных слоях, что позволяет автоматически извлекать как низкоуровневые, так и высокоуровневые признаки изображений, тем самым повышая устойчивость и способность модели к обобщению. Например, модели на основе CNN достигли высокой точности распознавания автомобильных номеров в

многочисленных исследованиях, особенно при сложных фонах и изменении углов обзора, обеспечивая точное обнаружение и распознавание номеров ^[31].

В задачах обнаружения номеров модели глубокого обучения, такие как YOLO (You Only Look Once) и Faster R-CNN, демонстрируют выдающиеся результаты. YOLO — это одноэтапная архитектура для обнаружения объектов, которая выполняет задачу обнаружения за одно пропускание изображения через модель, что обеспечивает высокую скорость работы. Эта особенность делает YOLO широко используемой в системах мониторинга трафика и системах автоматических платежей в реальном времени, поскольку она способна за очень короткое время завершить обнаружение и распознавание номеров. В отличие от неё, Faster R-CNN использует двухэтапную архитектуру, которая сначала генерирует кандидаты на области интереса, а затем классифицирует их. Хотя скорость обнаружения Faster R-CNN несколько ниже, чем у YOLO, её точность обнаружения выше, что делает её особенно подходящей для задач распознавания автомобильных номеров, где требуется высокая точность ^[41].

Помимо обнаружения номеров, глубокое обучение также добилось значительных успехов в распознавании символов. Модели, основанные на рекуррентных нейронных сетях (Recurrent Neural Network, RNN) и сетях долгой краткосрочной памяти (Long Short-Term Memory, LSTM), используются для последовательного распознавания символов на номерных знаках. LSTM, благодаря своей способности эффективно обрабатывать долгосрочные зависимости, широко используется в задачах распознавания рукописного текста и речи. В области распознавания автомобильных номеров LSTM решает проблему последовательной зависимости символов при их распознавании по одному, особенно в случаях, когда символы на номерных знаках прерываются повреждениями или частично закрыты. Способность LSTM к последовательному обучению помогает значительно повысить точность распознавания ^[31].

Ещё одно важное достижение глубокого обучения — это применение генеративных состязательных сетей (Generative Adversarial Networks, GANs). GANs создают реалистичные синтетические изображения через противоборство между генерирующей моделью и дискриминирующей моделью. В задачах распознавания номерных знаков исследователи используют GANs для генерации изображений номеров с различными условиями освещения, изменёнными углами обзора и повреждениями, чтобы повысить устойчивость моделей. Последние исследования показывают, что использование GANs для увеличения данных и синтеза изображений позволяет улучшить способность моделей глубокого обучения распознавать номера в сложных условиях ^[41].

Несмотря на огромный прогресс, достигнутый благодаря глубокому обучению в области распознавания автомобильных номеров, остаются нерешённые проблемы при его применении на практике. Во-первых, модели глубокого обучения требуют больших объёмов размеченных данных для обучения, что может быть проблемой для задач распознавания номеров в некоторых странах или регионах. Во-вторых, модели глубокого обучения часто требуют больших вычислительных ресурсов. Вопрос о том, как повысить скорость работы системы при сохранении высокой точности, остаётся ключевым направлением исследований. Кроме того, необходимо решить проблему адаптации моделей к различным форматам номеров в разных странах и регионах ^{[21] [41]}.

В заключение, применение машинного и глубокого обучения в распознавании автомобильных номеров значительно улучшило точность, устойчивость и скорость работы систем. С постоянной оптимизацией алгоритмов и повышением производительности аппаратного обеспечения технологии распознавания номеров будут всё шире применяться в интеллектуальных транспортных системах.

2.3 Обзор каскадной модели обучения

Каскадная модель обучения (Cascade Learning Model) — это структура модели, которая использует поэтапную обработку нескольких классификаторов и широко применяется в различных задачах обнаружения изображений. В отличие от традиционной одноуровневой модели, каскадная модель обучения опирается на несколько каскадных классификаторов, где каждый этап постепенно отфильтровывает отрицательные образцы, а положительные образцы обрабатываются более тщательно. Этот метод особенно подходит для решения сложных задач распознавания образов, таких как обнаружение лиц или автомобильных номеров. Стратегия постепенной оптимизации каскадного обучения не только повышает точность модели, но и значительно снижает вычислительные затраты, что делает её высокоэффективной в системах реального времени^[5].

Основные принципы каскадного обучения восходят к классическому алгоритму AdaBoost. AdaBoost итеративно обучает группу слабых классификаторов и объединяет их в сильный классификатор. Каскадная модель обучения располагает несколькими классификаторами по порядку и поэтапно применяет их к входным данным, причём каждый классификатор отвечает за распознавание определённого признака или образа. Для задач распознавания автомобильных номеров этот процесс обычно делится на несколько этапов: сначала обнаруживается область номерного знака, затем выполняется сегментация символов, и, наконец, происходит их распознавание. Такая многоэтапная обработка обеспечивает более тщательную обработку на каждом этапе и позволяет адаптировать сложность модели под задачи разного уровня сложности.

Каскадное обучение нашло важное применение в области распознавания автомобильных номеров. В задаче обнаружения автомобильных номеров определение области номерного знака обычно является первым шагом процесса распознавания. Каскадная модель обучения через несколько каскадных детекторов постепенно сужает область поиска возможных номеров. Например, каскадный классификатор, предложенный Виолой и Джонсом, изначально был разработан для задач обнаружения лиц, но позже был адаптирован для обнаружения номерных знаков. Этот метод постепенно применяет слабые классификаторы к различным областям изображения, фильтруя участки без номерных знаков и оставляя только те области, которые могут содержать номерной знак. Благодаря такому поэтапному фильтрованию система эффективно идентифицирует области с номерами среди большого количества фоновой информации, существенно снижая вычислительные затраты^[6].

В распознавании символов на номерах каскадная модель обучения также демонстрирует свои сильные стороны. Традиционные методы распознавания символов обычно полагались на отдельные классификаторы для каждой задачи классификации символов, в то время как каскадная модель обучения может поэтапно улучшать классификацию символов. Например, исследователи могут сначала использовать простой классификатор для первоначального распознавания кода провинции на номере автомобиля. Этот шаг относительно прост, так как коды провинций в Китае имеют фиксированный формат и ограниченное количество символов. Далее более сложные классификаторы могут сосредоточиться на распознавании букв и цифр, что требует большей точности. Этот метод постепенной обработки не только снижает вероятность ошибок распознавания, но и позволяет эффективно справляться с проблемами, такими как повреждения символов или изменения углов обзора^[6].

В отличие от других моделей глубокого обучения, каскадная модель обучения имеет явные преимущества в условиях ограниченных ресурсов. Поскольку каждый классификатор обрабатывает только небольшое количество возможных областей или образцов, вычислительная нагрузка каскадной модели относительно невелика, что делает её весьма применимой в

системах реального времени. Например, в автоматических платёжных системах или системах управления парковками предъявляются высокие требования к оперативности, и система должна быстро распознавать номерные знаки каждого проезжающего автомобиля. Механизм быстрого фильтрации каскадной модели позволяет завершить обнаружение и распознавание символов за миллисекунды. Таким образом, каскадные модели широко применяются в тех областях, где требуются высокие показатели времени отклика^[5].

Ещё одним важным преимуществом каскадной модели обучения является её масштабируемость. Поскольку структура модели состоит из нескольких каскадных классификаторов, её можно адаптировать в зависимости от требований конкретной задачи, изменяя глубину каскада и сложность каждого классификатора. Например, в задаче распознавания автомобильных номеров исследователи могут проектировать разные каскадные структуры для разных целей: для распознавания кода провинции можно использовать простой классификатор, а для сложных комбинаций букв и цифр — более сложные классификаторы для обеспечения точности. Модульная архитектура каскадной модели позволяет гибко настраивать её структуру и производительность в зависимости от конкретных условий использования^[6].

Ограничения каскадного обучения в основном связаны с эффективностью фильтрации отрицательных образцов, которая зависит от качества работы каждого классификатора. В сложных условиях, если начальный классификатор неэффективно отфильтровывает отрицательные образцы, последующие классификаторы будут нести большую вычислительную нагрузку, что может привести к снижению общей производительности системы. Кроме того, процесс проектирования и обучения каскадной модели довольно сложен, особенно при работе с большими объемами данных. Как найти баланс между производительностью каждого классификатора и эффективностью всей модели — это вопрос, требующий дальнейшей оптимизации^[5] ^[6].

Тем не менее, с появлением гибридных моделей, которые объединяют каскадное обучение с методами глубокого обучения, таких как модели глубокого каскадного обучения, появляются новые перспективы. Например, некоторые исследователи внедряют свёрточные нейронные сети (CNN) в структуру каскадного обучения, используя их мощные возможности для извлечения признаков и комбинируя их с поэтапным механизмом фильтрации каскадного обучения для улучшения общей производительности системы. В будущем оптимизация каскадных моделей обучения и их интеграция с глубоким обучением продолжат стимулировать развитие систем распознавания автомобильных номеров^[6].

3. Предварительные предложения по реализации

3.1 Подготовка и предобработка данных

Производительность системы распознавания номерных знаков в значительной степени зависит от используемого набора данных. Поэтому в данном исследовании подготовка набора данных является ключевым первым шагом. Для того чтобы обучение и тестирование каскадной модели обучения были эффективными, набор данных должен включать изображения номерных знаков в различных реальных условиях.

Источники данных: Источники данных могут быть как открытыми наборами данных с номерными знаками, так и созданными собственными наборами. Например, набор данных с китайскими номерными знаками можно получить через дорожные службы или системы интеллектуального транспортного мониторинга. Если невозможно собрать достаточно большой

набор данных, можно использовать технологии увеличения данных, чтобы сгенерировать разнообразные изображения номерных знаков, включающие коды различных провинций, условия освещения, повреждения и изменения углов наклона номерных знаков.

Разметка данных: Каждое изображение номерного знака должно быть точно размечено. Разметка включает в себя информацию о местоположении номерного знака (для его обнаружения) и каждый символ на номере (для распознавания символов). Кроме того, разметка кодов провинций необходима для проведения региональной классификации.

Предварительная обработка данных: Предварительная обработка изображений является важным шагом для обеспечения качества входных данных для модели. Основные шаги включают:

Обрезка изображений: Обрезка области номерного знака на изображении для удаления ненужного фона.

Преобразование в градации серого: Преобразование цветных изображений в черно-белые для уменьшения объема вычислений.

Удаление шума: Использование медианной или гауссовой фильтрации для удаления шума с изображений и повышения чёткости символов на номерном знаке.

Нормализация: Нормализация значений пикселей изображения для устранения влияния изменений освещённости.

Цель предварительной обработки данных заключается в создании чётких и чистых изображений номерных знаков, которые будут использоваться для последующего обучения и тестирования модели.

3.2 Разработка каскадной модели обучения

Разработка каскадной модели обучения является ключевым аспектом данного исследования. Путём разделения задач обнаружения номерных знаков, распознавания символов и региональной классификации на несколько каскадных этапов можно постепенно повышать точность и эффективность модели.

Этап начального обнаружения: На этом этапе основная задача модели — быстро обнаружить область номерного знака на изображении. Свёрточные нейронные сети (CNN), как одна из ведущих моделей для обнаружения объектов, способны эффективно выполнять эту задачу. Можно использовать лёгкие архитектуры, такие как YOLO или Faster R-CNN, для обеспечения работы в режиме реального времени при обнаружении номерных знаков. Цель данного этапа — сформировать кандидаты областей номерных знаков для последующей точной обработки.

Этап распознавания символов: После завершения обнаружения номерного знака следующая задача — распознавание символов. Преимущество каскадного обучения заключается в возможности поэтапного распознавания символов на номере. Вначале модель может распознать код провинции на номере, затем последовательно распознать буквы и цифры. Распознавание кода провинции можно свести к задаче многоклассовой классификации, так как количество кодов провинций в Китае ограничено. На этом этапе можно использовать модели на основе свёрточных нейронных сетей, а также комбинировать их с рекуррентными нейронными сетями (RNN) или сетями с длинной краткосрочной памятью (LSTM) для предсказания последовательности символов.

Этап региональной классификации: После успешного распознавания кода провинции на этапе распознавания символов система может выполнить региональную классификацию на основе этого кода. Этот этап относительно прост и может быть выполнен с использованием

таблицы поиска (Lookup Table), которая связывает распознанный код провинции с конкретным регионом или провинцией.

3.3 Архитектура и интеграция системы

Проектирование эффективной архитектуры системы критически важно для того, чтобы все модули могли работать согласованно. Вся система распознавания номерных знаков включает в себя полный цикл от получения изображения до обнаружения номерных знаков, распознавания символов и региональной классификации.

Проектирование системы: Система распознавания номерных знаков будет состоять из трёх основных модулей: модуль обнаружения номерных знаков, модуль распознавания символов и модуль региональной классификации. Эти модули будут работать последовательно, чтобы обеспечить быструю реакцию системы распознавания. Для повышения расширяемости системы модули должны иметь унифицированные интерфейсы для данных, что позволит легко заменять или обновлять отдельные модули в будущем.

Рабочий процесс:

Получение изображений: Изображения номерных знаков поступают с камеры или из существующего набора данных.

Обнаружение номерного знака: Модуль обнаружения номерных знаков использует модель CNN для определения местоположения номера.

Распознавание символов: Модуль распознавания символов поочередно определяет код провинции, буквы и цифры на номерном знаке.

Региональная классификация: На основе результатов распознавания символов система классифицирует номерной знак в соответствующий регион или провинцию.

Эта система может быть дополнительно интегрирована в интеллектуальные транспортные системы (ITS) для выполнения распознавания и классификации номеров в режиме реального времени.

3.4 Оптимизация модели

Для повышения производительности модели могут быть применены различные техники оптимизации на этапах проектирования и обучения каскадной модели обучения.

Применение переноса обучения: Модели глубокого обучения обычно требуют большого объема данных и вычислительных ресурсов для обучения. Перенос обучения помогает значительно сократить время обучения модели и снизить зависимость от данных. Используя предварительно обученные модели для других задач обнаружения объектов (например, предварительно обученные модели на ImageNet), можно значительно повысить эффективность распознавания номерных знаков и символов. Эти предварительно обученные модели могут быть использованы как начальные, затем может быть выполнена их донастройка (fine-tuning) на наборе данных с номерными знаками для адаптации к специфике задачи распознавания.

Оптимизация гиперпараметров: Выбор гиперпараметров (таких как скорость обучения, размер пакета, инициализация весов и другие) в процессе обучения напрямую влияет на скорость сходимости модели и конечную точность. Можно использовать такие методы, как случайный поиск (Random Search) или байесовская оптимизация (Bayesian Optimization), для автоматической настройки гиперпараметров и нахождения их оптимального сочетания. Кроме того, можно применять кросс-валидацию (Cross Validation), чтобы гарантировать хорошую обобщающую способность модели.

Увеличение данных: Создание изображений с различными углами поворота, масштабированием, сдвигами и добавлением шума с помощью техник увеличения данных позволяет существенно повысить точность модели в сложных условиях. Также использование

генеративных состязательных сетей (GANs) для генерации синтетических изображений номерных знаков может дополнительно расширить набор данных и повысить устойчивость модели.

4. Задачи, которые необходимо решить

В процессе разработки и реализации каскадной модели обучения возникает множество технических и практических задач. Начиная с обработки набора данных и заканчивая обучением модели, оптимизацией и интеграцией системы, каждый этап требует преодоления конкретных проблем. В следующей части подробно рассматриваются ключевые вопросы, которые необходимо решить в процессе реализации, и предлагаются возможные стратегии их решения.

4.1 Проблемы с набором данных

Задача 1: Как обрабатывать данные с изображениями номерных знаков в различных условиях (изменение освещённости, углов обзора, повреждение номеров и т.д.), чтобы гарантировать устойчивость модели

Одной из основных проблем системы распознавания номерных знаков является обработка изображений в разных сложных условиях. Изменения освещения, углов съёмки, повреждения или закрытие номера могут влиять на результаты распознавания. Для решения этих проблем модель должна обладать хорошей устойчивостью и способностью адаптироваться к различным ситуациям.

Изменение освещения: Изображения номерных знаков, снятые при разных условиях освещённости (днём, ночью или при неравномерном освещении), могут иметь изменённое качество. Для решения этой проблемы можно использовать технологии предварительной обработки изображений, такие как адаптивное гистограммное выравнивание (Adaptive Histogram Equalization), чтобы улучшить контраст изображения и уменьшить влияние неравномерного освещения.

Изменение углов обзора: Изменение углов, под которыми номера снимаются камерой, является распространённой проблемой. Для повышения способности модели распознавать номера под различными углами можно увеличить количество изображений номеров с разными углами в наборе данных или использовать технологии увеличения данных (например, поворот и перспективные преобразования изображений) для создания многоугольных изображений, которые помогут обучить модель распознавать номера с разных углов обзора.

Повреждение номеров: Повреждения или закрытие номера могут ухудшить видимость и целостность символов. Для улучшения адаптивности модели можно использовать генеративные состязательные сети (GANs) для создания изображений повреждённых номеров. Также архитектура модели, включающая свёрточные нейронные сети (CNN) и сети долгой краткосрочной памяти (LSTM), поможет модели лучше обрабатывать неполные или размытые символы.

Задача 2: Как очистить и расширить набор данных, чтобы гарантировать его разнообразие и репрезентативность для обучения модели

Очистка и расширение набора данных напрямую влияет на качество обучения модели и её способность к обобщению. Для того чтобы модель могла адаптироваться к различным условиям распознавания номеров, крайне важно обеспечить разнообразие набора данных.

Очистка данных: Необходимо фильтровать изображения с низким качеством (например, сильно размытые, неполные или серьёзно повреждённые номера), чтобы гарантировать качество входных данных для модели. Кроме того, важно проверять точность разметки данных. Ошибки в разметке, такие как неправильные границы символов или метки символов, следует исправлять вручную.

Расширение данных: Чтобы увеличить репрезентативность набора данных, можно использовать технологии увеличения данных для создания более разнообразных изображений номеров. Это могут быть изменения яркости и контраста изображений, добавление гауссовского шума или моделирование различных погодных условий (например, дождь, туман). Также можно использовать GANs для генерации новых образцов номеров, что расширит набор данных и улучшит способность модели к обобщению.

4.2 Разработка и обучение каскадной модели

Задача 3: Как спроектировать и обучить каскадную модель, подходящую для распознавания номерных знаков, и как разработать этапы каскада для обеспечения поэтапной оптимизации от обнаружения номеров до распознавания символов

Каскадная модель обучения обрабатывает задачи поэтапно, что позволяет значительно повысить точность и эффективность распознавания. Однако важным вопросом остаётся правильное проектирование каждого этапа каскадного обучения, так как оно напрямую влияет на производительность модели.

Многоступенчатая структура обнаружения: Каскадная модель должна на первом этапе быстро обнаружить область номера на изображении, а на следующих этапах выполнить точное распознавание символов. В конструкции модели можно использовать лёгкие модели для обнаружения объектов, такие как YOLO или Faster R-CNN, на первом этапе обнаружения номера. Далее можно использовать более точные классификаторы для распознавания символов. Каждый классификатор на каждом этапе должен поэтапно фильтровать нерелевантные области, чтобы обеспечить более точную обработку данных на следующем этапе.

Пошаговое распознавание символов: На этапе распознавания символов можно использовать многоступенчатую классификацию для поэтапного распознавания различных категорий символов на номере. Например, сначала можно использовать простой классификатор для распознавания кода провинции, а затем более сложные классификаторы для распознавания букв и цифр. Структура каскада должна быть спроектирована таким образом, чтобы для простых задач использовались простые классификаторы, а для более сложных — более сложные классификаторы, что позволит повысить общую вычислительную эффективность и точность.

Задача 4: Как оптимизировать процесс обучения для повышения точности распознавания и уменьшения числа ошибок распознавания

Для того чтобы каскадная модель обучения точно распознавала символы на номерах и минимизировала ошибки, требуется множество оптимизаций на этапе обучения.

Оптимизация функции потерь: В процессе обучения выбор правильной функции потерь имеет решающее значение для производительности модели. Для задач распознавания символов стандартным выбором является использование перекрёстной энтропии (Cross-Entropy Loss), однако для случаев с размытыми или неполными символами можно применить фокусную функцию потерь (Focal Loss) для более сложных образцов, что уменьшит внимание модели к легко классифицируемым образцам.

Регуляризация модели: Чтобы избежать переобучения модели, в процессе её обучения можно применить такие методы регуляризации, как L2-регуляризация, слои Dropout или нормализация пакетов (Batch Normalization), что повысит способность модели к обобщению.

4.3 Реализация региональной классификации

Задача 5: Как реализовать точную региональную классификацию на основе распознавания кода провинции и как обеспечить точное распознавание номеров из разных провинций на большом объёме данных

Формат и расположение символов на номерах в различных провинциях Китая может немного отличаться, поэтому задача региональной классификации требует, чтобы модель точно распознавала код провинции и классифицировала номера в правильный регион.

Извлечение признаков символов: Для распознавания кода провинции можно использовать CNN для извлечения признаков из первых двух символов (кода провинции) на номере. Так как код провинции состоит из ограниченного набора символов и имеет фиксированный формат, задача его распознавания относительно проста. Разработка специального классификатора для кода провинции обеспечит высокую точность на этом этапе.

Таблица соответствий регионов: После распознавания кода провинции можно использовать таблицу соответствий (Lookup Table), чтобы сопоставить распознанный код конкретной провинции или региону. Таблица должна включать коды всех провинций и автономных регионов Китая.

Задача 6: Как справляться с незначительными различиями в форматах номеров между провинциями

Хотя общая структура номеров в Китае едина, между номерами разных провинций могут быть небольшие различия, особенно в расположении символов и в стилях шрифтов.

Сопоставление шаблонов и адаптация модели: Чтобы учесть эти различия, можно использовать алгоритмы сопоставления шаблонов, чтобы на этапе распознавания символов учесть особенности форматирования номеров в различных провинциях. Это улучшит способность модели к распознаванию. Также можно выполнить дальнейшую тонкую настройку модели, чтобы она адаптировалась к различным особенностям номеров в каждой провинции.

4.4 Интеграция и применение системы

задача 7: Как интегрировать модель в систему реального времени, обеспечив возможность распознавания и классификации автомобильных номеров в режиме реального времени в системах мониторинга трафика

В реальных приложениях важнейшей характеристикой системы распознавания номерных знаков является её способность работать в режиме реального времени, особенно в таких сценариях, как мониторинг дорожного движения или автоматические системы оплаты. Система должна выполнять задачи распознавания и классификации номеров за минимально короткое время.

Оптимизация системы: Использование лёгких архитектур моделей (например, YOLO и других моделей для детекции в реальном времени) и эффективных фреймворков для выполнения вычислений (таких как TensorRT, ONNX и другие) может значительно ускорить процесс вывода модели и обеспечить её работу в режиме реального времени. Также параллельная обработка потоков данных (например, с использованием GPU-ускорения) может повысить скорость реакции системы.

Интеграция в встроенные системы: В таких приложениях, как мониторинг дорожного движения или управление парковками, где требуется локальное развертывание, можно интегрировать оптимизированную модель в устройства на базе встроенных систем. Использование граничных вычислений (edge computing) позволяет уменьшить зависимость от облачных технологий, что повышает скорость отклика системы и её стабильность.

Задача 8: Как преодолеть узкие места в производительности модели при её реальном применении, включая скорость отклика системы и стабильность

При масштабном применении системы распознавания номерных знаков требуется обрабатывать большие объёмы данных, что предъявляет высокие требования к скорости отклика и стабильности системы.

Механизмы балансировки нагрузки и кэширования: В условиях высокой нагрузки и большого количества одновременно поступающих данных можно применить механизмы балансировки нагрузки и кэширования для обеспечения стабильности системы. Использование

кэширования для хранения результатов распознавания часто встречающихся номерных знаков позволяет сократить время повторных вычислений и повысить скорость отклика системы.

Механизмы отказоустойчивости: Чтобы избежать ошибок распознавания и неправильных выводов модели в реальных условиях, можно разработать систему отказоустойчивости, которая будет выполнять повторное распознавание в случае ошибки или использовать несколько моделей для голосования, чтобы повысить точность распознавания.

5. Ожидаемые результаты

5.1 Высокая точность распознавания номеров

Использование каскадной модели обучения позволяет значительно повысить точность распознавания китайских автомобильных номеров в сложных условиях, особенно при наличии размытости, изменении углов обзора и других сложных ситуациях.

5.2 Эффективная региональная классификация

Система может автоматически распознавать код провинции на номере и выполнять точную региональную классификацию, обеспечивая возможность классификации автомобильных номеров Китая на уровне провинций или городов.

5.3 Устойчивость и расширяемость системы

Система обладает хорошей устойчивостью и способна адаптироваться к различным условиям (освещение, углы обзора, повреждения и т.д.) для эффективного распознавания номеров.

Ожидается, что система будет обладать возможностью расширения, что позволит адаптировать её под другие региональные кодовые системы и применить для распознавания автомобильных номеров в других странах или регионах.

5.4 Реализация в реальном времени и практическое применение

Система способна обрабатывать данные в реальном времени, что делает её подходящей для практического применения в таких областях, как мониторинг дорожного движения, управление парковками и платные проезды.

Предполагается, что модель может быть реализована на встроенных системах или облачных платформах, что обеспечит возможность обработки большого объёма данных по распознаванию номеров с высокой практической ценностью.

Рекомендации

1. Юань И., Се Ф., Чжан Ж. Достижения в распознавании автомобильных номеров с использованием методов машинного и глубокого обучения: обзор // IEEE Transactions on Intelligent Transportation Systems. 2020. Т. 21, № 6. С. 2764-2775.

2. Чжоу В., Ван Ч., Ян С. Методы машинного обучения для распознавания автомобильных номеров: обзор // Journal of Intelligent Transportation Systems. 2019. Т. 23, № 5. С. 408-419.

3. Ли Х., Ли Ч., Ли С. Сравнительное исследование алгоритмов глубокого обучения для распознавания автомобильных номеров // IEEE Access. 2020. Т. 8. С. 41718-41726.

4. Хуан Ч., Ван Ч., Ду С. Применение GANs в системах распознавания автомобильных номеров // Pattern Recognition Letters. 2021. Т. 144. С. 62-70.

5. Виола П., Джонс М. Быстрое обнаружение объектов с использованием каскада простых признаков // Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR). 2001. Т. 1. С. 511-518.

6. Чжоу Л., Лю Ц., Ван С. Каскадное обучение для обнаружения и распознавания автомобильных номеров в сложных условиях // Pattern Recognition Letters. 2020. Т. 131. С. 234-241.