|  |  |
| --- | --- |
| Gerb-BMSTU_01 | **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)** |

ФАКУЛЬТЕТ МЕЖДУНАРОДНЫХ ОБРАЗОВАТЕЛЬНЫХ ПРОГРАММ

КАФЕДРА СИСТЕМЫ ОБРАБОТКИ ИНФОРМАЦИИ УПРАВЛЕНИЯ

**РАСЧЕТНО-ПОЯСНИТЕЛЬНАЯ ЗАПИСКА**

***К НАУЧНО-ИССЛЕДОВАТЕЛЬСКОЙ РАБОТЕ***

***НА ТЕМУ:***

***Исследование по обнаружению опасных грузов камерами наблюдения на основе нейронных сетей***

Студент \_\_ ИУ5И-34М\_\_  05. 12. 2024\_\_**\_ \_\_\_\_\_\_\_** Лю Бовэнь **\_\_\_\_\_\_\_**

(Группа) (Подпись, дата) (И. О. Фамилия)

Руководитель НИР  . Ю. Н. Кротов

(Подпись, дата ) (И. О. Фамилия)

*Москва 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 \_\_ИУ-5\_\_

(Индекс)

\_\_\_\_\_\_\_\_\_\_\_\_\_ В.И.Терехов

(И. О. Фамилия)

« \_\_\_\_\_ » \_\_\_\_\_\_\_\_\_\_\_\_\_\_ 2024 г.

**ЗАДАНИЕ**

**на выполнение научно-исследовательской работы**

по теме *Исследование по обнаружению опасных грузов камерами наблюдения на основе нейронных сетей*

Студент группы ИУ5И-34М

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ Лю Бовэнь \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

(Фамилия, имя, отчество)

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

\_\_\_ исследовательская \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

Источник тематики (кафедра, предприятие, НИР) \_\_\_учебная тематика\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

График выполнения НИР: 25% к 13 нед. , 50% к 14 нед. , 75% к 15 нед. , 100% к 16 нед.

Техническое задание Изучить и выбрать технологии по теме работы, выявить возможные источники данных по выбранной тематике, сделать критический анализ выбранных технологий, обосновать недостатки и возможные пути исправления или компенсации недостатков, обучить 3-4 модели \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

***Оформление научно-исследовательской работы:***

Расчетно-пояснительная записка на 19 листах формата А4.

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

Слайды презентации 7-8 шт. \_

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

Дата выдачи задания « 27 » октября 2023 г.

**Руководитель НИР**  15. 10. 2024 г. Ю. Н. Кротов

(Подпись, дата И. О. Фамилия)

**Студент**  15. 10. 2024 г. Лю Бовень

(Подпись, дата И. О. Фамилия)

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

кафедре.

Содержание

[Введение 4](#_Toc12676)

[1. Структура сверточной нейронной сети для распознавания изображений 5](#_Toc15602)

[1.1 Понятие сверточной нейронной сети 5](#_Toc30836)

[1.2 Структура сверточной нейронной сети 6](#_Toc954)

[2. Модель распознавания опасных грузов на основе сверточной нейронной сети 7](#_Toc25332)

[2.1 MobileNetV2 8](#_Toc21201)

[2.2 ROI-CRNN 9](#_Toc26763)

[2.3 Faster R-CNN 11](#_Toc26900)

[2.4 YOLO 15](#_Toc28814)

[3. Построение модели YOLOv8 и анализ результатов эксперимента 23](#_Toc16581)

[3.1 Построение модели YOLOv8 23](#_Toc6989)

[3.2 Показатели оценки производительности 23](#_Toc25917)

3.3 Результаты тренировочного эксперимента модели YOLOv8 25

[3.4 Обнаружение модели YOLOv5x 26](#_Toc4456)

[Заключение 27](#_Toc10515)

[Список использованных источников 29](#_Toc10556)

# Введение

Наличие опасных грузов в общественных местах является серьезной социальной проблемой.К таким опасным грузам относятся легковоспламеняющиеся, взрывоопасные, токсичные, вредные и другие вещества. При неправильном использовании или обращении с ними могут произойти серьезные несчастные случаи, такие как пожар, взрыв, отравление и т.д., что представляет серьезную угрозу для жизни людей и сохранности имущества.В ответ на эту проблему был разработан проект по обнаружению опасных грузов в общественных местах.Этот проект направлен на использование передовых технических средств для быстрого и точного обнаружения опасных грузов в общественных местах, чтобы избежать несчастных случаев и обеспечить безопасность населения.

## 1. Структура сверточной нейронной сети для распознавания изображений

### 1.1 Понятие сверточной нейронной сети

При традиционном распознавании изображений для извлечения признаков обычно используются искусственно разработанные алгоритмы, а затем извлеченные признаки вводятся в обучаемый классификатор для классификации и распознавания.Самым большим преимуществом сверточных нейронных сетей является то, что нет необходимости разрабатывать и изучать алгоритмы искусственного извлечения признаков. Они напрямую используют исходное 2D-изображение в качестве входных данных, скрывая процесс извлечения признаков в сети, и реализуют автоматическое обучение входных данных.Это значительно сокращает ручное участие и повышает эффективность алгоритма.Сверточная нейронная сеть - это особый вид структуры глубокой сети. Ее особенность отражается в двух аспектах. Во-первых, связь между нейронами в каждом слое не полностью связана, а локально связана; во-вторых, веса соединений между нейронами в одном слое являются общими. Эта структура локальной связи и распределения веса сокращается количество подключений в сети, значительно уменьшается количество весов, а также снижается сложность. Три характеристики локального подключения, распределения веса и понижающей дискретизации сверточных нейронных сетей не только уменьшают масштаб сети и объем вычислений, но и обеспечивают устойчивость к смещению, масштабированию или другим формам изменений входного изображения, выполняя извлечение признаков слой за слоем.

Локальное подключение: Сверточные нейронные сети являются продуктом глубокого развития искусственных нейронных сетей. Узлы между слоями традиционных искусственных нейронных сетей полностью связаны (например, многослойные персептроны), что приводит к большому количеству соединений в сети и очень большому объему вычислений при решении.[1] Сверточная нейронная сеть обеспечивает локальное соединение нейронов между двумя соседними слоями. Узлы сверточного слоя подключены только к некоторым узлам предыдущего слоя. Подключенная область называется локальным сенсорным полем. Этот метод разреженного соединения, очевидно, имеет меньше соединений, чем полное соединение, поэтому он является вызывается локальное соединение.

Распределение веса: Сверточные нейронные сети используют метод локального подключения, при котором локальное сенсорное поле соответствует локальной подобласти изображения, а веса локального сенсорного поля (называемые сверточными ядрами или фильтрами объектов) могут использоваться для извлечения таких объектов, как цвет, направленные края, текстуры, углы и т.д. на изображении.В задачах распознавания изображений и классификации классификатору недостаточно использовать один признак для правильной классификации. Сверточные нейронные сети обычно используют несколько наборов весов. Используя этот вес для деконволюции и фильтрации входного изображения, можно получить карту признаков, то есть каждый набор весов соответствует ядро свертки или фильтр. Из-за различных ядер свертки функции извлеченного изображения также различаются.

### 1.2 Структура сверточной нейронной сети

В более типичной структуре сверточной нейронной сети первые несколько слоев представляют собой чередование сверточного слоя и объединяющего слоя, а последние несколько слоев (близких к выходному слою) являются полностью связанными слоями. Роль полностью связанного слоя заключается в соединении карты объектов в длинную-размерный вектор, а затем используйте классификатор для классификации и оценки входного вектора и, наконец, получите результат классификации.

Способность сверточного слоя извлекать признаки в основном зависит от взаимодействия между сверточным ядром и входными данными, то есть каждое сверточное ядро можно рассматривать как средство извлечения каждого признака.Сверточные нейронные сети обычно используют комбинацию из нескольких сверточных ядер, главным образом для улучшения способности извлечения признаков из изображений. Ядро свертки перемещается по всему изображению. При перемещении оно сначала перемещается вправо. Когда строка обрабатывается, оно возвращается к крайнему левому концу и перемещается вниз в вертикальном направлении в соответствии с размером шага и снова перемещается вправо до конца. После подключения сверточного слоя выполняется объединяющий слой, который также называется понижающей дискретизацией.[3] Операцию объединения можно рассматривать как преобразование изображения с более высоким разрешением в изображение с более низким разрешением. С помощью слоя объединения количество узлов в конечном полностью подключенном слое может быть дополнительно уменьшено, чтобы достичь цели уменьшения параметров в нейронной сети. Сверточный слой и объединяющий слой повторяются поочередно, а затем результаты извлеченных объектов соединяются в вектор-столбец и выходной слой через полностью подключенный слой и, наконец, оцениваются классификатором.Это полная структура сверточной нейронной сети.

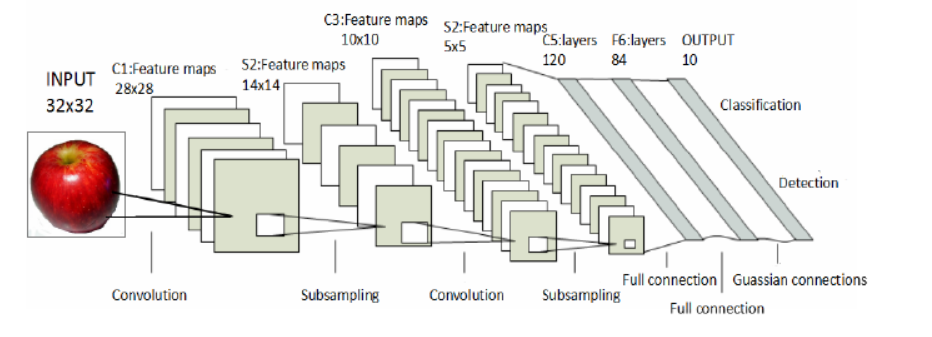


Рисунок 1 Структура сверточной нейронной сети

## 2. Модель распознавания опасных грузов на основе сверточной нейронной сети

Традиционные сверточные нейронные сети, очевидно, были реализованы для уменьшения размерности входных данных, что имеет смысл, но для некоторых специальных сценариев применения скорость обучения нейронной сети все еще не соответствует требованиям, поэтому были созданы некоторые специальные структурированные модели сверточных нейронных сетей, ниже в основном представлены три сложной модели сверточной нейронной сети.

### 2.1 MobileNetV2

MobileNet - это модель, основанная на свертке с разделением по глубине, которая представляет собой разложение стандартной свертки на свертку по глубине и свертку 1\*1, т.е. свертку по точкам. Для Mobilenet глубокая свертка применяет один фильтр к каждому отдельному входному каналу для фильтрации, а затем свертка по точкам применяет операцию свертки 1\*1 для объединения выходов всех глубоких сверток. Стандартная свертка, с другой стороны, объединяет все входы за один шаг, чтобы получить новый набор выходов. Глубоко разделяемая свертка разделяет его на два этапа: фильтрация для каждого отдельного слоя, а затем объединение на следующем этапе. Такая декомпозиция позволяет значительно сократить вычислительные усилия и размер модели.

Теперь мы подробно опишем структуру модели. Как упоминалось ранее, основными строительными блоками являются разделяемые конволюционные блоки. mobileNetV2 содержит начальный полностью связанный слой из 32 конволюционных ядер, за которым следуют 19 остаточных узких слоев. mobileNetV2 использует ReLU6 в качестве нелинейной функции активации, которая является более надежной при использовании для вычислений низкой точности. В модели всегда используется конволюционное ядро размером 3\*3, а во время обучения применяются отсев и нормализация по батч-норме.[6]

Архитектура MobileNetv2 основана на инвертированной остаточной структуре, где оригинальная остаточная структура имеет три свертки в главной ветви и две свертки по точкам с большим количеством каналов, а инвертированная остаточная структура наоборот, с большим количеством сверток в середине (по-прежнему используя разделенную по глубине сверточную структуру) и меньшим в боковых частях.

MobileNetV2 использует постоянную скорость масштабирования во всей сети, за исключением первого уровня. По мере уменьшения размера сети немного меньшая скорость масштабирования работает лучше, в то время как большие сети с большей скоростью масштабирования работают лучше.

表格

描述已自动生成

Рисунок 2.1 Структура сетевой модели MobiletV2.

В целом, MobilenetV2 был создан для приложений компьютерного зрения конечных устройств и характеризуется настраиваемым количеством параметров, низкой задержкой, малым потреблением памяти, эффективной работой и не намного меньшей точностью, а также является облегченноймоделью.

### 2.2 ROI-CRNN

Классификатор обучается путем объединения комбинированных признаков различных каналов, изученных сверточными нейронными сетями и рекуррентными нейронными сетями. Для обучения сверточного ядра предлагается метод выбора блоков изображения на основе ROI интересующей области и последующего выполнения обучения без контроля с помощью K-means, чтобы избежать случайного выбора из всего исходного изображения, тем самым уменьшая избыточную информацию и увеличивая текстуру, цветовую информацию и другие характерные выражения исходного изображения в сверточном ядре. ядро. В вышеупомянутой сверточной нейронной сети признаки, извлеченные из определенных локальных блоков изображений в ROI, используются в качестве детекторов, и различные признаки в разных местоположениях могут быть получены путем сверточных операций с исходным изображением. Последующая сеть использует метод объединения векторов признаков нескольких масштабов и интеграции высокоуровневых признаков для усиления дифференциации признаков, тем самым повышая скорость распознавания изображений.

В основном он разделен на двухуровневые сети: из ROI выбирается большое количество блоков изображений для неконтролируемой кластеризации K-means, и для операции свертки первого уровня получается определенное количество ядер свертки, а затем изображение каналов RGB и оттенков серого, которые были просто преобразованы. предварительно обработанные данные сначала обрабатываются сверточной нейронной сетью первого уровня (CNN) для получения соответствующей карты объектов, и в то же время карта объектов корректируется и локальный контраст нормализуется; затем карта объектов сети первого уровня вводится в соответствующую сверточную нейронную сеть и рекуррентную нейронную сеть сеть второго уровня, рекуррентную нейронную сеть можно рассматривать как комбинацию сверточного слоя и слоя выборки, на котором могут быть изучены объединенные признаки низкоуровневых признаков изображения, и, наконец, объединенные векторы признаков свертки и рекурсивной сети обучаются на классификаторе Softmax, так что что касается получения хорошей модели классификатора.[2]

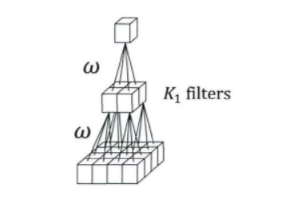
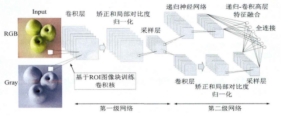


Рисунок 2.2 Схема структуры сети CRNN

Рекуррентные нейронные сети используют набор весов сверточного ядра на разных слоях. Его функция аналогична комбинации операций свертки и выборки. Это уменьшает размер объекта и сокращает объем вычислений за счет рекурсивного использования весов и может эффективно извлекать соответствующие комбинированные объекты из нижних слоев изображения.Установите размер воспринимающего поля фильтра равным L\*L, а размер входной карты объектов равным m2\*n2. [8] Размер шага равен размеру фильтра, то есть, когда принимающая область не перекрывается, размеру каждой карты объектов после прохождения через слой рекуррентной нейронной сети. сеть равна (m2/L) x (n2/L), что эквивалентно уменьшению размера объектов каждого слоя в 1\*1 раз за счет многоуровневого уменьшения размерности, пока размер карты объектов не станет равным 1\*1, то есть она не превратится в пиксель.

### 2.3 Faster R-CNN

При обычных обстоятельствах, после того как CNN извлекает элементы изображения, он классифицирует и оценивает все изображение целиком.В случае, когда на изображении несколько целей, CNN необходимо заранее разделить изображение, но процесс сегментации изображения очень чувствителен к помехам со стороны фоновых факторов.Предлагается серия алгоритмов региональной сверточной нейронной сети(Region with Convolutional Neural Network, R-CNN) для решения ситуации, когда на изображении имеется несколько целей, без необходимости в каких-либо сложных процессах сегментации изображения на ранней стадии.Конкретный процесс обнаружения R-CNN показан на рис.2.3. Сначала Input image извлекает около 2000 полей-кандидатов из исходного изображения с помощью метода выборочного поиска(Selective Search) и нормализует их, а затем использует CNN для извлечения объектов изображения для каждой области-кандидата и, наконец, загрузите объекты в классификатор машины опорных векторов каждой категории, чтобы определить, принадлежат ли они к этой категории, и используйте регрессор для точной настройки положения поля-кандидата для завершения задачи классификации изображений.

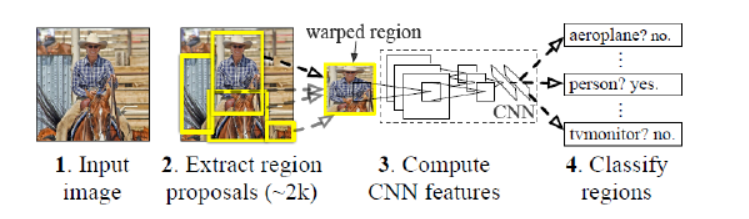


Рисунок 2.3 Процесс проверки RCNN

Хотя R-CNN добилась хороших результатов, есть и некоторые очевидные недостатки. То есть существует около двух тысяч блоков-кандидатов, сгенерированных R-CNN. Каждый блок-кандидат должен выполнять извлечение признаков изображения CNN, и большая часть изображений в этих блоках-кандидатах перекрывают друг друга, поэтому существует есть много мест, которые многократно вычисляются.Чтобы решить эту проблему, была предложена улучшенная версия R-CNN, Fast R-CNN.Как показано на рисунке 2.4, Fast R-CNN сопоставляет все области-кандидаты с последним сверточным слоем сети, а затем использует объединяющий слой интересующей области(Region of Interest, ROI) для унификации размера различных областей-кандидатов. Таким образом, каждое изображение необходимо извлечь только один раз. Извлечение объектов CNN, без необходимости повторного извлечения каждой области-кандидата, что значительно сокращает объем вычислений.

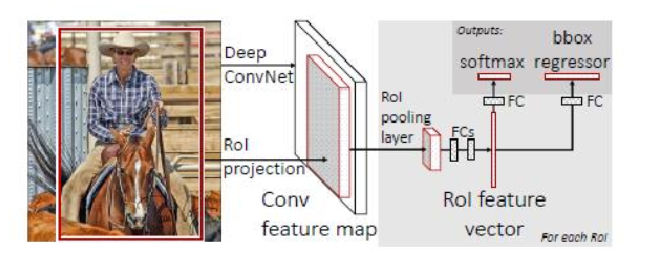


Рисунок 2.4 Процесс проверки Fast-RCNN

Как показано на рис.2.5, более быстрый алгоритм R-CNN в основном состоит из трех частей: первая часть - базовая сеть CNN, ее основная задача - завершить извлечение признаков изображения; вторая часть - сеть RPN, которая в основном использует сверточные нейронные сети для непосредственной генерации региональных целей, определите, существует ли цель, и выполните координатную регрессию; последняя часть - это сеть распознавания, сеть распознавания предназначена для классификации и координации регрессионной коррекции возможных целевых областей, выделенных сетью RPN, и получения окончательных результатов тестирования с помощью алгоритма немаксимального подавления.[5]

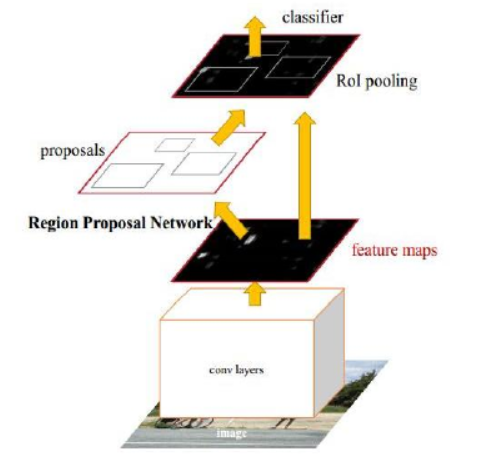


Рисунок 2.5 Система контроля Fast-RCNN

В первой части базовой сети CNN в качестве примера используется модель ZF.Сетевая модель ZF-Net точно настроена на основе сети Alex Net. Сеть ZF-Net содержит 5 сверточных слоев, 3 слоя максимального объединения и 2 полностью подключенных слоя. Конкретная структура показана на рисунке 4.4.

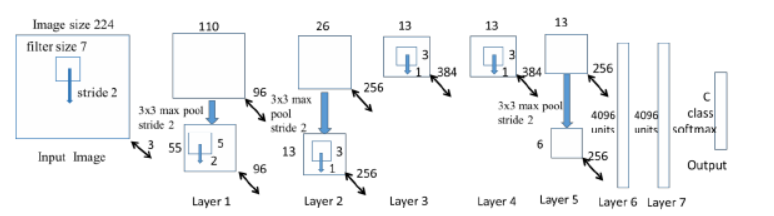


Рисунок 2.6 Схема структуры сети ZF

Конкретная структура второй части сети RPN показана на рис. 2.6.Основная идея состоит в том, чтобы напрямую сгенерировать окно с предложением области в сверточной нейронной сети. По сути, оно использует скользящее окно для поиска целей, но выбирает 9 окон разных размеров и с разными соотношениями сторон, центрированных на каждом пикселе на карте объектов последнего слоя сверточной нейронной сети.За RPN следуют два полностью связанных слоя, которые используются для определения наличия целевой и координатной регрессии, а именно Cls\_score и Bbox\_pred.[7] Среди них Cls\_score содержит 2 элемента, которые представляют вероятность существования целевой и нецелевой. Bbox\_pred используется для регрессии положения координат о цели.Затем, в соответствии с оценкой предложения по региону, отфильтруйте первые 300 предложений по региону и войдите в последующую сеть распознавания для классификации и вынесения решения.

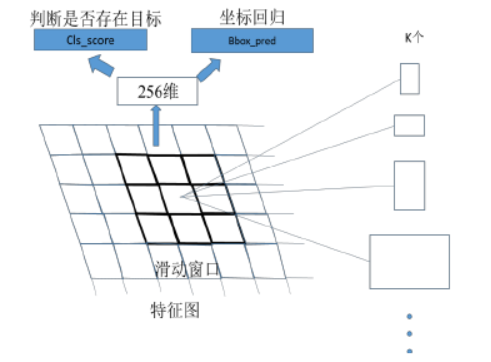


Рисунок 2.7 Схема структуры сети RPN

Третья часть - это сеть распознавания. Она предназначена для классификации и оценки целей, извлеченных сетью RPN. То есть feature maps и proposals, собранные с помощью Roi Pooling, используются через полностью подключенный уровень и уровень Softmax для вычисления вероятности того, что каждое предложение относится к каждой категории, а затем используются bounding box regression для повторного получения информации о смещении позиции proposal bbox\_pred, чтобы получить более точную информацию о поле обнаружения цели.

Несмотря на то, что Faster-RCNN значительно улучшил модель обучения, большинство применений таких алгоритмов обнаружения целей на основе блоков-кандидатов трудно удовлетворить требованиям короткого времени обнаружения цели.

### 2.4 YOLO

Чтобы решить проблему длительного времени обнаружения целевой задачи, появился алгоритм обнаружения цели, основанный на регрессии.Основанный на регрессии алгоритм обнаружения целей использует одновременное позиционирование цели и прогнозируемую классификацию целей непосредственно в сети, чтобы обеспечить точность обнаружения при одновременном повышении скорости обнаружения. Алгоритм серии YOLO (You Only Look Once) и алгоритм SSD (Single Shot Detector) - это два типичных алгоритма обнаружения целей на основе регрессии.

Основанный на регрессии YOLOv1, алгоритм SSD предлагает механизм предварительного выбора блоков, который использует карты объектов разного масштаба для прогнозирования целей разного размера с помощью структуры пирамиды объектов, тем самым улучшая способность многомасштабного обнаружения.Твердотельный накопитель также может повысить точность обнаружения и обеспечить скорость обнаружения. Точность обнаружения SSD относительно высока по сравнению с YOLOv1, но текущий алгоритм серии YOLO теперь обновлен и повторен до YOLOv5, конечно, его производительность давно превзошла или даже превзошла алгоритм SSD.Таким образом, в данной статье алгоритм серии YOLO используется в качестве основы для улучшения модели и оптимизации для идентификации изображений категорий фруктов и овощей. Ниже приводится краткое введение в алгоритм серии YOLO.

Алгоритм серии YOLO использует регрессионные задачи для определения цели объекта, объединяя два этапа определения области-кандидата и распознавания цели в один. Это очень креативно по сравнению с серией R-CNN, в которой сначала выделяются области-кандидаты, а затем идентифицируются цели в областях-кандидатах.Производительность алгоритма претерпела обновления версий и итерации, и она постоянно улучшается.Ниже приводится сравнительное описание алгоритмов серии YOLO прошлых поколений и анализируются преимущества и недостатки каждой версии алгоритма.

1) YOLO v1

YOLOv1 использует CNN в качестве магистральной сети. Сверточный уровень и объединяющий уровень одновременно выполняют две операции: выборку входного изображения и извлечение скрытых объектов изображения, а затем полностью подключенный уровень выводит окончательный результат прогнозирования. На рисунке 2.8 показана сетевая структура YOLOv1. Входное изображение делится сетью на S \*S сетки. Каждая сетка определяет цели в сетке, и ограничивающие рамки, уровни достоверности и все категории векторов вероятности целей, содержащихся во всех сетках, прогнозируются одновременно.Алгоритм YOLO значительно повышает скорость обнаружения, поскольку он использует стратегию создания сетки областей, чтобы избежать дублирования вычислений.Алгоритм YOLO позволяет более эффективно использовать пространственную информацию изображения с помощью задач регрессии и снижает частоту обнаружения фоновых ошибок.Из обзора литературы стало известно, что алгоритм YOLOv1 получил 63,4% и 57,9% значений mAP в тестовых экспериментах с наборами данных VOC2007 и VOC2012 соответственно.В то же время, при сравнении алгоритма YOLOv1 и двухэтапного алгоритма обнаружения целей по-прежнему имелись недостатки, и его точность была низкой при обнаружении небольших целей с аномальным размером или плотностью.

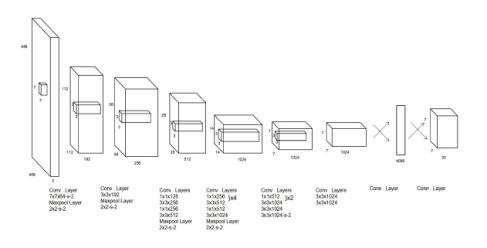


Рисунок 2.8 Схема структуры сети YOLOv1

2）YOLOv2

YOLOv2 добавляет в алгоритм такие стратегии, как пакетная регуляризация, классификаторы высокого разрешения, механизмы привязки кадров, новые магистральные сети и априорное масштабирование, чтобы улучшить производительность обнаружения всей сетевой модели. Как показано на рисунке 2.9, значения карты, полученные YOLOv2 на наборах данных VOC2007 и VOC2012, были изменены. значительно улучшено по сравнению с YOLOv1. В YOLOv2 применяется сеть извлечения объектов Darknet-19, в которой количество каналов объектов увеличено на уровне объединения и используются объемы 3×3 и 1×1 в сети Darknet-19. Ядро используется для извлечения объектов и уменьшения размерности изображения, и окончательный результат прогнозирования выводится на уровне объединения глобальных средних значений.Используя пакетную нормализацию в model trainer, YOLOv2 ускоряет сходимость целевой функции и повышает стабильность обучения модели.На рисунке 2.10 показана структура сети YOLOv2.



Рисунок 2.9 Сравнение YOLOv2 и YOLOv1 по наборам данных о ЛОС

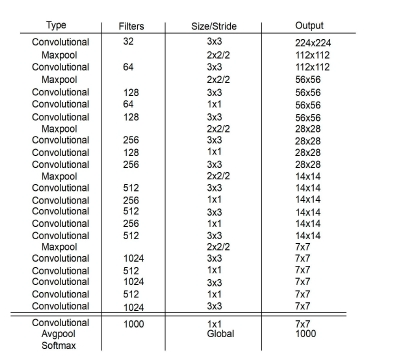


Рисунок 2.10 Схема структуры сети YOLOv2

3）YOLOv3

YOLOv3 использует сеть Darknet-53 для извлечения признаков изображения, увеличивает количество сверточных слоев, что позволяет лучше извлекать характеристические признаки объектов на изображениях. На рисунке 2.11 показана структура сети YOLOv3. Стратегия использования якорных рамок в YOLOv3 заключается в настройке трех групп по три разных размера якорных рамок в соответствии с размером карт признаков, определенных сетью. Количество предсказаний кандидатов в ограничивающие рамки также значительно увеличивается за счет этой стратегии, и вероятность фильтрации положительных примеров в данных снижается, что также немного увеличивает отзывчивость модели. Структура YOLOv3 относительно проста и рафинирована, этот продуманный подход также предоставляет новое направление для последующих одностадийных алгоритмов обнаружения объектов, и хотя его производительность в обнаружении небольших объектов недостаточна по сравнению с двухэтапными алгоритмами, YOLOv3 все же сохраняет высокую скорость обнаружения, даже увеличив количество сверточных слоев. Возможность легкой деплоймента YOLOv3 на мобильных устройствах особенно важна для промышленных задач обнаружения объектов.

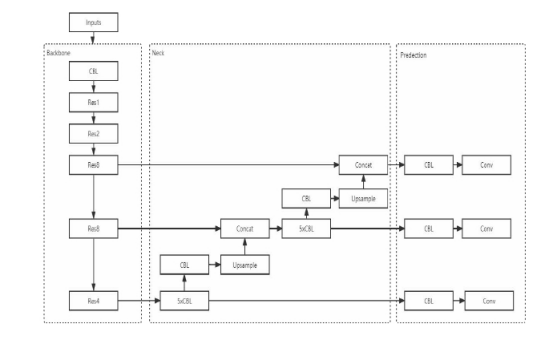


Рисунок 2.11 Схема структуры сети YOLOv3

4）YOLOv4

В сравнении с предыдущей версией YOLOv3, были внесены улучшения в архитектуру YOLOv4, которые включают введение усиления данных Mosaic на входе, улучшение основной сети путем внедрения CSP-модулей в Darknet53. С помощью структур Feature Pyramid Networks (FPН) и Path Aggregation Network (PAN) осуществляется передача и интеграция входной информации, активационной функцией сети становится функция Mish, отбор предсказаний рамок изменен с традиционного NMS на DIoU\_NMS. В YOLOv4 также добавлен SPP-пуллинг, который эффективно предотвращает проблемы обрезки и искажения форм объектов изображения, типичные для операций алгоритма R-CNN, а также решает проблему повторного извлечения признаков изображения, ускоряя производство ограничивающих рамок и уменьшая вычислительные затраты сети. Рисунок 2.12 показывает структуру сети YOLOv4.

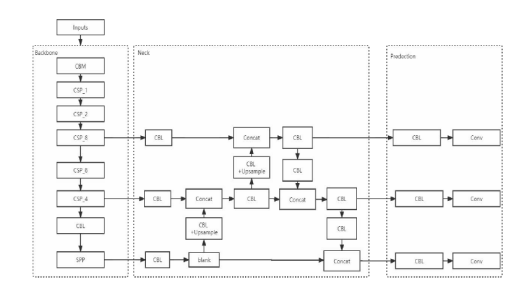


Рисунок 2.12 Схема структуры сети YOLOv4

5）YOLOv5

Алгоритм YOLOv5 является пятым выпуском в серии YOLO и имеет множество нововведений и оптимизаций на основе предыдущих версий алгоритмов YOLO, что делает его одним из лидеров в области одностадийного обнаружения объектов, обладающим относительно высокой точностью и более высокой скоростью детектирования модели в целом.

YOLOv5s, YOLOv5m, YOLOv5l, YOLOv5x и YOLOv5n - это пять различных вариантов сетевой структуры YOLOv5, среди которых YOLOv5n является более легковесной моделью в серии алгоритмов YOLOv5, позволяя при определенной точности распознавания значительно сократить количество параметров модели и вычислительных затрат по сравнению с другими алгоритмами.

Традиционный алгоритм YOLOv5n состоит из четырех частей: ввода Input, основной сети Backbone, сети Neck и блока предсказаний Prediction. На входной стороне модели выполняются три основные операции: усиление данных Mosaic, адаптивный расчет якорных рамок и адаптивное масштабирование изображения. Усиление данных Mosaic улучшает обнаружение мелких объектов и подходит для распознавания изображений фруктов и овощей, в соответствии с данным исследованием. Адаптивное масштабирование изображений приводит изображения различных размеров к одному фиксированному размеру для ввода. Основная сеть YOLOv5n использует модули Conv и CSP, а также модуль SPPF и структуру FPN+PAN для этапа слияния признаков модели. Так же, как и в предыдущих версиях YOLO, на выходном предикторе одновременно выводятся предсказания трех различных размеров, соответствующие обнаружению объектов малого, среднего и крупного размера. Рисунок 2.13 показывает структуру сети YOLOv5n.

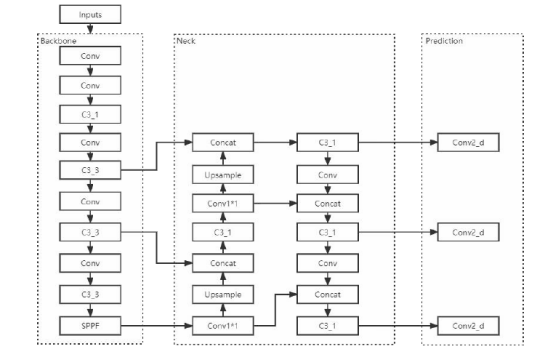


Рисунок 2.13 Схема структуры сети YOLOv5n

Благодаря выдающимся характеристикам YOLOv5 во многих аспектах, включая скорость детекции, точность и способность к обнаружению целей различных размеров, он был широко проверен и принят на практике. Этот алгоритм в процессе реализации конечного обнаружения объектов поддерживает высокую вычислительную производительность, что особенно критично для приложений, требующих обработки в реальном времени. Поэтому я намерен использовать модель YOLOv5 как основу и произвести тщательную настройку и оптимизацию модели YOLOv5, чтобы удовлетворить конкретные потребности этого исследования.

5）YOLOv8

Магистральная сеть YOLOv8 заменила структуру C3 в YOLOv5 на структуру C2F, увеличив градиентный поток и переход от слоя к слою; в шейной части была удалена часть слоя convolution connection и скорректировано количество каналов в моделях различного масштаба; головная часть использует структуру развязывающей головки и отличается от анкерной.- основание должно быть свободным от якоря.

Функция потерь использует стратегию распределения положительной выборки TaskAlignedAssigner и использует потери в фокусе распределения для повышения точности определения модели.

Последняя 10-я ячейка в YOLOX используется для завершения операции mosaic enhanced, которая значительно повышает точность модели.

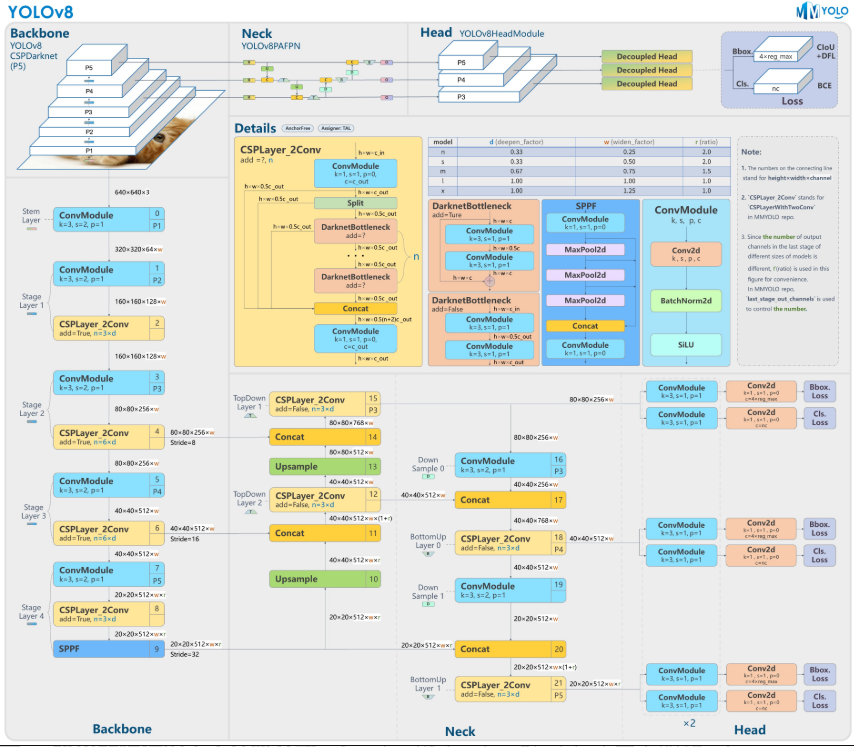


Рисунок 2.14 Схема структуры сети YOLOv8

## 3. Построение модели YOLOv8 и анализ результатов эксперимента

### 3.1 Построение модели YOLOv8

В этой статье описывается процесс создания и тестирования нейронной сети YOLOv8 для распознавания пищевых продуктов.В статье упоминается использование классического набора данных об опасном оружии в общественных местах и дополнительных данных, собранных и подготовленных автором этой статьи, а также характеристики разработки и тестирования модели на основе технической поддержки и стека программного обеспечения.Ниже приводится перевод статьи на русский язык：

В нашем исследовании мы использовали набор данных, основанный на опасном оружии в общественных местах, чтобы повысить универсальность модели.В нашей работе мы использовали шесть различных категорий товаров для формирования набора данных.

Чтобы построить и усовершенствовать модель, мы применили фреймворк глубокого обучения Pytorch, а также обучили и протестировали модель в системе Pycharm.

### 3.2 Показатели оценки производительности

Сравнивая полученные численные значения различных алгоритмов, мы можем получить основания для суждения о хорошей или плохой производительности алгоритмов. Кроме того, на основе этих числовых значений, мы также можем делать целенаправленные улучшения или оптимизации определенных аспектов алгоритма. Таким образом, показатели оценки производительности модели также очень важны.

Для начала представим конкретные показатели оценки следующих алгоритмов обнаружения целей:

(1)Точность (Precision): это показатель, который обозначает долю истинно положительных результатов среди всех результатов, определенных как положительные.

(2) Полнота (Recall): дает оценку того, были ли обнаружены все цели, которые необходимо обнаружить на изображении.

(3) F1-мера (F-Measure): это гармоническое среднее точности и полноты, когда значение F1 в результате выше, это означает, что модель показывает лучшую эффективность.

В вышеупомянутой формуле TP обозначает количество истинно положительных примеров; FP обозначает количество ложноположительных примеров; FN обозначает количество ложноотрицательных примеров; N обозначает количество категорий в выборке.

(4)Кривая P-R: Кривая P-R это график точности и полноты, где P - это значение по вертикальной оси, а R - по горизонтальной оси, отображающий характеристику изображения. В итогах, чем выше точность и полнота, тем лучше производительность модели.

В итоге, в данной статье были выбраны точность (P), полнота (R) в качестве показателей для оценки результатов эксперимента модели.

### 3.3 Результаты тренировочного эксперимента модели YOLOv8

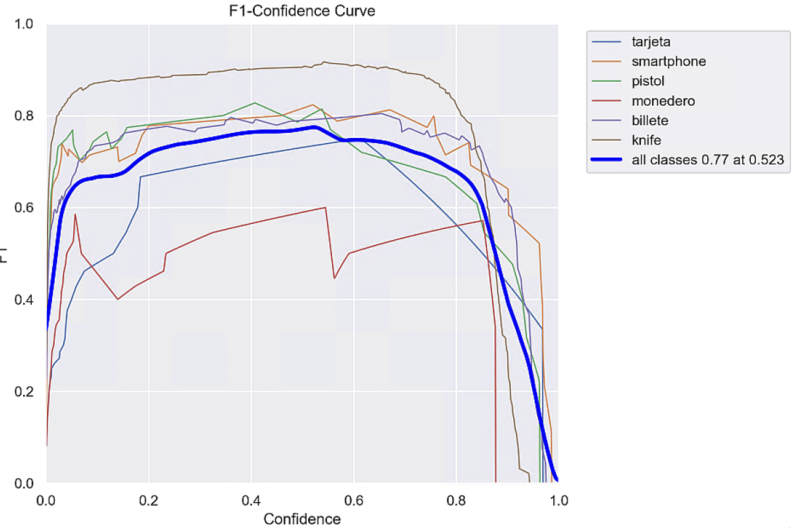


Рисунок 3.1 Схема параметра F1



Рисунок 3.2 Схема параметра PR

### 3.4 Обнаружение модели YOLOv5x

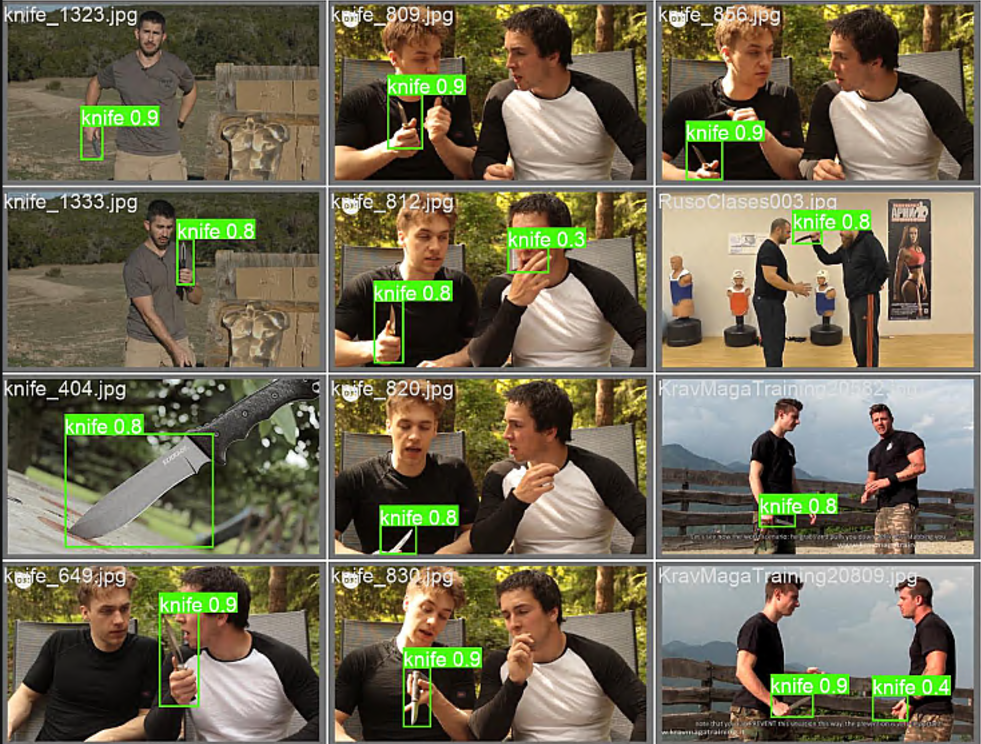


Рисунок 3.3 Результат прогнозирования

# Заключение

YOLOv8 - это итеративный продукт серии "Посмотри только один раз" (YOLO). По сравнению с предыдущей версией, он значительно улучшил точность и скорость работы. Он очень подходит для оборудования с ограниченной вычислительной мощностью, например, для обнаружения опасных грузов.Следует подчеркнуть, что эти устройства требуют высокой точности в режиме реального времени для правильного выполнения таких функций, как мониторинг объектов, раннее оповещение и так далее. Отличительной особенностью YOLOv8 является то, что он обладает высокой производительностью, но для обучения и рассуждения требуется относительно немного вычислительных ресурсов, поэтому его можно интегрировать в продукты бытовой электроники.Это особенно важно для обнаружения опасных грузов, поскольку камеры наблюдения должны обрабатывать изображения с высоким разрешением при сохранении низкой системной задержки. Преимущество модели YOLOv8 заключается в ее гибкости, индивидуализации и хорошей адаптивности, которые могут удовлетворить конкретные потребности пользователей.Например, для обнаружения опасных грузов модель может быть сконфигурирована таким образом, чтобы идентифицировать различные предметы и идентифицировать их в различных условиях окружающей среды, чтобы обеспечить точное обнаружение предметов на месте происшествия. Проверка предметов на наличие опасных грузов в общественных местах имеет большое значение для обеспечения общественной безопасности и поддержания социальной стабильности.Благодаря реализации этого проекта можно эффективно предотвращать несчастные случаи, снижать количество жертв и материальный ущерб, а также повышать уровень безопасности в общественных местах.Поэтому очень важно изучить методы классификации продукции, разработать достаточно точные алгоритмы и интегрировать их в системы обнаружения опасных грузов.В данной статье на основе YOLOv8 оптимизируется модель, создается новый алгоритм классификации товаров YOLOv8 и проводятся тесты его производительности.

# Список использованных источников

1. Du, Cheng-Jin, and Da-Wen Sun. "Recent developments in the applications of image processing techniques for food quality evaluation." Trends in food science & technology 15.5 (2004): 230-249.

2. Zeng Weiliang.Research on image recognition based on deep learning and its application in the classification of fruits and vegetables in refrigerators.2017. Fuzhou University, MA thesis.

3. Zhan Yuan."Research on the design of human-computer interaction in smart refrigerators."Industrial design.12(2021):122-123.

4. Qin Weizhu, Shao Huairong, Luo Shouzhong, Chen Yongsha. Research on the identification of food by refrigerator cameras [C]//.Proceedings of the 2017 China National Electrical Appliance Technology Conference.,2017:230-234.

5. Huang Y, Kangas L J, Rasco B A. Applications of artificial neural networks (ANNs) in food science[J]. Critical reviews in food science and nutrition, 2007, 47(2): 113-126.

6. Zhao Qiuling, Mou Haichun, Dong Lei, Yang Chuang Chuang.Design of intelligent detection and identification system for refrigerator energy efficiency identification [J].Automation application,2014(07):54-56+58.

7. He K, Zhang X, Ren S, et al. Deep residual learning for image recognition[C]//Proceedings of the IEEE conference on computer vision and pattern recognition. 2016: 770-778.