

ҚАЗАҚСТАН РЕСПУБЛИКАСЫ ҒЫЛЫМ ЖӘНЕ ЖОҒАРЫ БІЛІМ МИНИСТРЛІГІ

«Л.Н. ГУМИЛЕВ АТЫНДАҒЫ ЕУРАЗИЯ ҰЛТТЫҚ УНИВЕРСИТЕТІ» КЕАҚ

**Студенттер мен жас ғалымдардың
«GYLYM JÁNE BILIM - 2024»
XIX Халықаралық ғылыми конференциясының
БАЯНДАМАЛАР ЖИНАҒЫ**

**СБОРНИК МАТЕРИАЛОВ
XIX Международной научной конференции
студентов и молодых ученых
«GYLYM JÁNE BILIM - 2024»**

**PROCEEDINGS
of the XIX International Scientific Conference
for students and young scholars
«GYLYM JÁNE BILIM - 2024»**

**2024
Астана**

УДК 001

ББК 72

G99

«ǴYLYM JÁNE BILIM – 2024» студенттер мен жас ғалымдардың XIX Халықаралық ғылыми конференциясы = XIX Международная научная конференция студентов и молодых ученых «ǴYLYM JÁNE BILIM – 2024» = The XIX International Scientific Conference for students and young scholars «ǴYLYM JÁNE BILIM – 2024». – Астана: – 7478 б. - қазақша, орысша, ағылшынша.

ISBN 978-601-7697-07-5

Жинаққа студенттердің, магистранттардың, докторанттардың және жас ғалымдардың жаратылыстану-техникалық және гуманитарлық ғылымдардың өзекті мәселелері бойынша баяндамалары енгізілген.

The proceedings are the papers of students, undergraduates, doctoral students and young researchers on topical issues of natural and technical sciences and humanities.

В сборник вошли доклады студентов, магистрантов, докторантов и молодых ученых по актуальным вопросам естественно-технических и гуманитарных наук.

УДК 001

ББК 72

G99

ISBN 978-601-7697-07-5

**©Л.Н. Гумилев атындағы Еуразия
ұлттық университеті, 2024**

doi: 10.1109/CVPR52688.2022.00581. keywords: {Reflectivity;Learning systems;Adaptation models;Computer vision;Codes;Noise reduction;Lighting;Low-level vision},

3. Wei, C., Wang, W., Yang, W., & Liu, J. (2018, August 14). Deep RETINEX decomposition for low-light enhancement. arXiv.org. <https://arxiv.org/abs/1808.04560>

4. Цветовые пространства в OpenCV и бинаризация // Robx URL: https://robx.org/wiki/prog/opencv-biblioteka/cvet_opencv/ (дата обращения: 20.03.2024).

5. Щеглов Д.Ю. Исследование алгоритма "гамма-коррекции" для повышения визуального качества изображения // Теория и практика современной науки. 2017. №7 (25). URL: <https://cyberleninka.ru/article/n/issledovanie-algoritma-gamma-korreksii-dlya-povysheniya-vizualnogo-kachestva-izobrazheniya> (дата обращения: 18.03.2024).

УДК 004.85

ИССЛЕДОВАНИЕ АЛГОРИТМОВ ДЕТЕКЦИИ И ОТСЛЕЖИВАНИЯ ПОКРЫТИЯ ДОРОГИ НА ВИДЕО С ИСПОЛЬЗОВАНИЕМ КОМПЬЮТЕРНОГО ЗРЕНИЯ

Райман Евгений Сергеевич¹, Айымбай Сұнғат Жұмабекұлы²,
Керімхан Бекжан Темірханұлы³
e.raiman@bk.ru

¹студент 4-го курса по ОП 6В06104 - «Вычислительная техника и программное обеспечение», ЕНУ им. Л.Н.Гумилева, Астана, Казахстан

²докторант 2-го курса по ОП 8D06104 - «Вычислительная техника и программное обеспечение», ЕНУ им. Л.Н.Гумилева, Астана, Казахстан

³преподаватель кафедры компьютерной и программной инженерии ЕНУ им. Л.Н.Гумилева, Астана, Казахстан

Научный руководитель – А.К. Жумадилаева

Введение. В основе интеллектуальной видеоаналитики лежит компьютерное зрение. Компьютерное зрение – научная область, описывающее создание и обучение специализированных машин для обнаружения, отслеживания и классификации различного рода объектов. Принято считать, что в конце 1970-х годов началось усиленное изучение данной проблемы. Ровно с того момента, когда компьютеры смогли совладать с обработкой больших наборов данных, к числу которых можно отнести и изображения.

В настоящий момент вопрос интеллектуальной видеоаналитики вышел на новый уровень. С развитием технологий алгоритмы получают все более широкое распространение. Количество сфер, в котором применяется интеллектуальная видеоаналитики – колоссальное. Безопасность на производственных объектах, ассистирование врачебной деятельности, помощь на дорогах, распознавание текста – это лишь малая часть. Стоит также отметить, что видеоаналитика с помощью компьютерного зрения является неотъемлемой частью концепции «Умного города».

Сравнивая компьютерное и биологическое зрение можно точно сказать, что компьютеры видят не так, как люди. Люди, с учетом своего жизненного опыта способны с особой точностью и скоростью идентифицировать объекты на изображении. Компьютеры же в свою очередь без должного обучения не способны это делать. Особую роль в этом занимает машинное обучение. Машинное обучение – методика анализа данных, которая в результате процесса решения большого количества сходных задач на дата-сете дает возможность обучения аналитической системе распознавать определенные классы объектов на изображении.

Алгоритмы нахождения объектов на изображении. Во-первых, алгоритм нахождения объектов на изображении зависит от постановки задачи.

Задача сегментации – задача, при которой на вход модели подается изображение, а на выходе для каждого пикселя определяется принадлежность этого пикселя к определенной

категории. Например, в исходном изображении представлена дорога, автомобили и дорожный знак. Искомая категория – дорожные знаки. Для каждого пикселя будет определено, является ли он дорожным знаком или нет.

Задача классификации – задача, в которой определяется категории объектов, присутствующих на изображении. В случае, если речь идет о классификации с локализацией, помимо определения категории отображается еще и рамка, ограничивающая местонахождение объекта на изображении.

Задача детекции – задача, в которой происходит выделение объектов, определение их класса и отображение ограничивающих рамок этих объектов. В данных задачах классы определены заранее и, в отличие от задач классификации, число объектов, имеющих на изображении, изначально неизвестны.

Для определения достоверности расположения ограничивающих рамок используются специальные метрики. В задачах классификации с локализацией и детекции объектов может быть использована метрика IoU (Intersection over Union). Для вычисления данной метрики используется следующая формула:

$$IoU = \frac{S(A \cap B)}{S(A \cup B)},$$

где $S(A \cap B)$ – площадь пересечения ограничивающих рамок и $S(A \cup B)$ – общая площадь ограничивающих рамок.

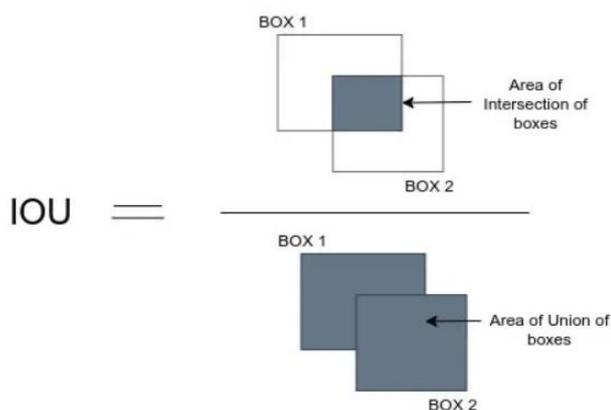


Рисунок 1. Наглядная формула метрики IoU

В задачах детекции объектов на изображении также может быть использована метрика mAP (mean average precision) – усредненная по всем категориям величина средней точности. Метрика имеет следующую формулу:

$$AP = \int_0^1 p(r)dr,$$

где p – точность, r – полнота из предположения, что ограничивающая рамка определена верно, если $IoU \geq 0.5$.

Box Loss (потеря обрамляющего прямоугольника) – функция потерь, которая измеряет разницу между предсказанными и истинным координатами обрамляющего прямоугольника. Для определения Box Loss используются различные метрики, такие как средняя абсолютная ошибка (MAE) или среднеквадратичная ошибка (MSE).

Class Loss (потеря классификации) – функция потерь, которая измеряет разницу между предсказанной и истинной меткой класса объекта, то есть классификационная потеря измеряет, насколько точно модель предсказывает класс объекта.

Object Loss (потеря объекта) – общая потеря, которая объединяет в себе и Box Loss (потерю обрамляющего прямоугольника), и Class Loss (потерю классификации). Цель потери

объекта – обеспечить, чтобы модель одновременно правильно предсказывала координаты ограничивающего прямоугольника и класс объекта.

Имеется два подхода к решению задач детекции объектов – двухэтапный и одноэтапный.

Двухэтапный подход разделен на две задачи. На первом этапе либо с помощью селективного метода, либо с использованием специального слоя нейронной сети выделяются регионы интереса – области, в которых вероятность содержания в себе какого-либо объекта высока. На втором этапе с помощью классификатора определяется принадлежность объекта к исходному классу. Алгоритмы, использующие двухэтапный метод – R-CNN, Fast R-CNN, Faster R-CNN и Mask R-CNN.

Одноэтапный подход. Здесь, в свою очередь, отсутствует этап определения регионов интереса. Алгоритм сразу переходит к предсказыванию координат ограничивающих рамок со следующими характеристиками – степенью уверенности и результатом проведения классификации. При этом постепенно происходит корректировка местоположения ограничивающих рамок. Отличным примером алгоритма с одноэтапным подходом является алгоритм YOLO. Именно о нем речь пойдет далее.

Алгоритм YOLO. YOLO (You Only Look Once) – современный алгоритм глубокого обучения, использующийся для обнаружения объектов на изображениях и видео. Алгоритм YOLO был разработан в 2016 году и впервые представлен на конференции CVPR2016 Джозефом Редмоном и Али Ризой Мохаммади.

Главным отличием YOLO от других алгоритмов сверточной нейронной сети (CNN), используемых для детекции объектов, следующее: изображение проходит через сверточную нейронную сеть всего один раз, то есть происходит внедрение всего изображения сразу. В других же алгоритмах изображение проходит через CNN снова и снова, многократно внедряя изображение, поэтому главное преимущество алгоритма YOLO – высокая скорость его работы и поддержка работы в реальном времени.

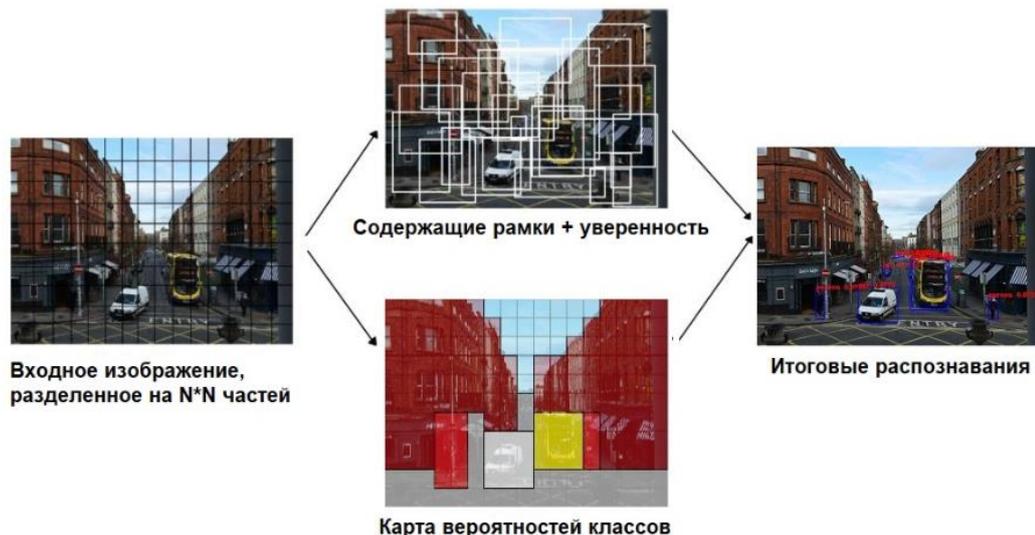


Рисунок 2. Принцип детекции объектов на изображении с помощью алгоритма YOLO

Сбор дата-сета для последующего обучения модели. Для начала необходимо сформировать набор данных – дата-сет. Собрать большое количество изображений и загрузить все эти изображения на сервис разметки изображений, в нашем случае это Roboflow. Был собран дата-сет, состоящий из изображений с дорожными знаками.



Рисунок 3. Примеры изображений, использующихся в обучении модели

После сбора изображений необходимо сформировать набор классов, по которым будут распознаваться объекты и произвести аннотирование изображений. Создание аннотаций на изображении – процесс создания описательных данных, помогающих в идентификации и классификации объектов, людей или сцен на картинке.

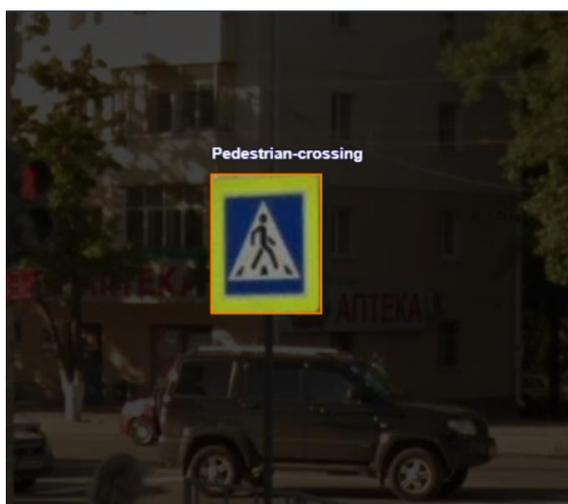


Рисунок 4. Процесс аннотирования изображения

Обучение модели. После создания дата-сета необходимо разбить его на три части: обучающая (train), валидационная (valid) и тестовая (test), в соотношении 70-20-10.

Обучающая выборка (train) – набор, который используется непосредственно для обучения модели. Модель проходит обучение путем подстройки весов и параметров, что в итоге позволяет выбранную функцию потерь свести до минимума. Содержит в себе изображения с разметкой объектов.

Валидационная выборка (valid) – этот набор позволяет проводить оценку производительности модели во время обучения и для настройки гиперпараметров. На валидационной выборке модель не обучается напрямую, но при этом ее производительности измеряется, что позволяет выбирать лучшие значения для параметров. Также с помощью валидационной части обучения модели предотвращается переобучение. Данная выборка также содержит в себе изображения с разметкой объектов.

Тестовая выборка (test) – набор данных, который не был использован ни в обучении, ни в валидации модели. С использованием тестовой выборки измеряются окончательные оценки производительности модели после завершения обучения и настройки гиперпараметров.

За все время обучения модель будет некоторое количество раз проходить по всей обучающей выборке. Для каждого такого прохождения существует термин «эпоха». После каждой эпохи модель обычно оценивается на проверочном (валидационном) наборе данных для оценки ее производительности и контроля за переобучением.

В целом, получилось сформировать дата-сет из 2400 изображений, из которых 1000 были уникальным, а остальные 1400 – результатом различного рода взаимодействия на изображения.

Результат работы обученной модели. Обучение проводилось в течение 4-х часов. Было проведено около 200 эпох. При этом на пике метрики имели следующие значения: mAP – 0.98, Box Loss – 0.386, Class Loss – 0.227 и Object Loss – 0.82. В конце всех вышеперечисленных действий получится модель, способная идентифицировать категорию дорожного знака на изображении и на видео.

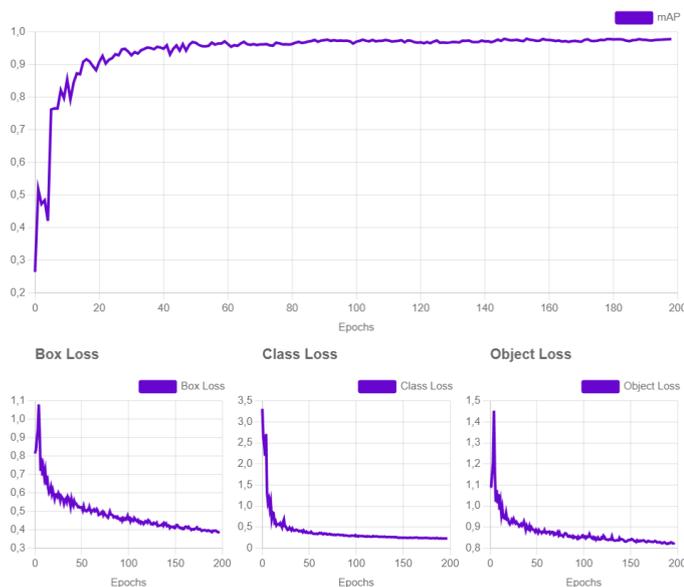


Рисунок 5. Графическое представление процесса обучения модели



Рисунок 6. Пример работы обученной модели

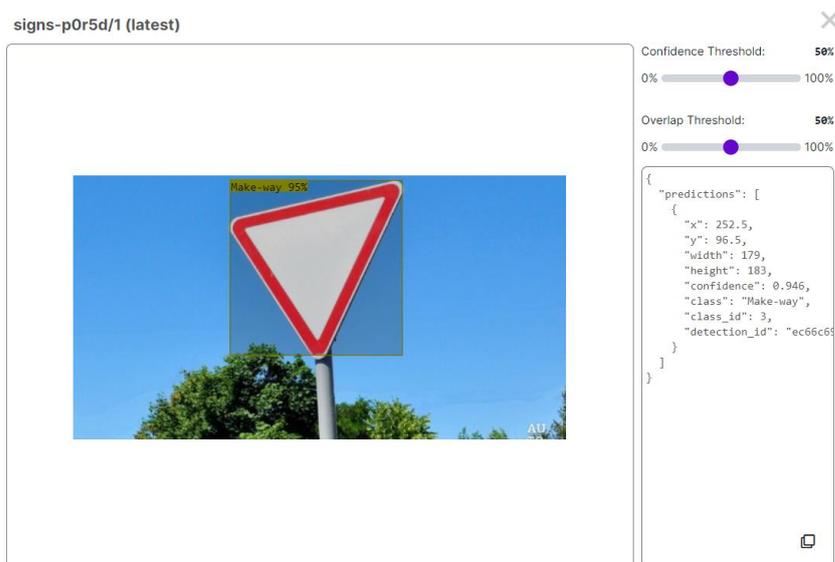


Рисунок 7. Пример работы обученной модели

Заключение. В заключении, создание модели для детекции классов на изображениях с применением алгоритма YOLO (You Only Look Once) представляет собой значимый шаг в развитии компьютерного зрения и распознавания объектов. Результаты данной работы подтверждают эффективность метода YOLO, позволяя точно и быстро обнаруживать различные объекты на изображениях. Использование глубокого обучения и концепции единовременного просмотра для детекции объектов делает этот подход особенно привлекательным для множества приложений, включая автономные транспортные средства, системы безопасности, медицинскую диагностику и многие другие. Продолжение исследований в этой области будет способствовать созданию более эффективных и универсальных систем компьютерного зрения, способных решать широкий круг задач в различных областях применения.

При поддержке: Данное исследование профинансировано Комитетом науки Министерства науки и высшего образования Республики Казахстан (грант №AP19678989 «Интеллектуальная видеоаналитика и отчетность по дорожному покрытию и освещению улиц города»).

Список использованных источников

1. Nelson, J. (2024, March 5). What is Yolo? the ultimate guide [2024]. Roboflow Blog. <https://blog.roboflow.com/guide-to-yolo-models/>
2. Zvornicanin, E. (2024, March 18). What is Yolo Algorithm?. Baeldung on Computer Science. <https://www.baeldung.com/cs/yolo-algorithm>
3. Илья Г. (2020). Распознавание объектов с помощью YOLO v3 на Tensorflow 2.0 // Proglib URL: <https://proglib.io/p/raspoznavanie-obektov-s-pomoshchyu-yolo-v3-na-tensorflow-2-0-2020-11-08> (дата обращения: 12.03.2024).
4. Подробное объяснение алгоритма YOLO // Русские Блоги URL: <https://russianblogs.com/article/9704510420/> (дата обращения: 19.03.2024).
5. Роман (2021). Алгоритм YOLO простым языком // Medium URL: <https://medium.com/nuances-of-programming/%D0%B0%D0%BB%D0%B3%D0%BE%D1%80%D0%B8%D1%82%D0%BC-yolo-%D0%BF%D1%80%D0%BE%D1%81%D1%82%D1%8B%D0%BC-%D1%8F%D0%B7%D1%8B%D0%BA%D0%BE%D0%BC-a4e666be4d41> (дата обращения: 21.03.2024).