УДК 338.45
DOI

Батырканов Ж.И.,
Доктор технических наук, профессор
bjenish@mail.ru[1]

Бийбосунов Б.И.,
Доктор технических наук, профессор
bbolotbek@mail.ru[1]

Лян Чжаньхао,
соискатель**,**
liangzhanhao@kstu.kg [1]

Кыргызский государственный университет
имени И. Арабаева,
720026,Бишкек,ул. Раззакова,51а[1]

МЕТОДОЛОГИЯ ОБНАРУЖЕНИЯ ЦЕЛЕЙ НА ОСНОВЕ ГЛУБОКОГО ОБУЧЕНИЯ

Аннотация. *В статье представлены методы обнаружения целей на основе применения технологии глубинного обучения. Рассмотрены алгоритмы обнаружения на основе признаков Хаара и признаков (HOG).*

Отмечается, что методы обнаружения целей достигли значительного прогресса в последние годы с появлением моделей глубокого обучения, таких как конволюционные нейронные сети (CNN). Показаны преимущества метода обнаружения целей на основе глубокого обучения перед традиционными сенсорными методами.

Ключевые слова: *методы обнаружения целей, алгоритмы, традиционные методы, глубинное обучение, признаки, компьютерное зрение.*

Введение

Обнаружение целей - это основной и ключевой вопрос в области компьютерного зрения, направленная на идентификацию и определение местоположения целевых объектов на изображениях или видео. Ее основные задачи включают два аспекта: первый - определение категории целевого объекта, а второй - определение положения целевого объекта на изображении, которое обычно представлено ограничивающей рамкой. С быстрым развитием технологии глубокого обучения методы обнаружения целей эволюционировали и получили широкое применение во многих областях, таких как охранное наблюдение, автоматическое вождение, медицинская визуализация, "Умный дом" и т.д.

Методы обнаружения целей можно разделить на традиционные методы, основанные на извлечении признаков, и современные методы, основанные на глубоком обучении. Несмотря на то, что традиционные методы были успешны в определенных сценариях применения, из-за их зависимости от вручную созданных признаков и плохой адаптации к сложным сценариям, все больше исследований переходит к методам глубокого обучения, которые демонстрируют значительные преимущества, особенно при работе с динамически изменяющимися и многомасштабными целями. В этом разделе мы представим обзор традиционных методов обнаружения целей и методов на основе глубокого обучения, уделяя особое внимание ключевым технологиям и истории развития методов глубокого обучения.

Методология

До широкого применения методов глубокого обучения исследования в области обнаружения целей были сосредоточены на традиционных методах компьютерного зрения. Большинство этих методов опирались на алгоритмы ручного извлечения признаков для обнаружения целей путем сопоставления признаков в сочетании с классификаторами.

Метод обнаружения на основе признаков Хаара (Haar)

Метод признаков Хаара - один из самых ранних методов обнаружения целей на основе признаков, а наиболее известное его применение - система обнаружения лиц, предложенная Виолой и Джонсом [1]. Признаки Хаара извлекаются путем вычисления разности яркости в локальной области изображения, и эти признаки классифицируются каскадными классификаторами. Несмотря на то что этот метод позволил добиться значительных результатов в раннем обнаружении лиц, он имеет серьезные ограничения при работе со сложными ситуациями, такими как изменение освещения, различные позы и окклюзии.

Метод обнаружения на основе признаков (HOG)

HOG-функции - еще один распространенный традиционный метод обнаружения целей, который описывает характеристики краев и формы цели путем подсчета направления градиента в локальной области. HOG-методы широко используются в задачах обнаружения пешеходов, особенно при обнаружении пешеходов в сочетании с машинами опорных векторов (SVM) для достижения хороших результатов. Однако метод HOG менее эффективен при работе с крупномасштабными вариациями и сложными фонами, а вычисление признаков занимает много времени, что затрудняет его использование в задачах обнаружения в реальном времени [2].

Хотя традиционные методы успешно справляются с некоторыми специфическими задачами, они в основном опираются на разработанные вручную признаки и хуже адаптируются к меняющимся целям и сложным сценам. С развитием глубокого обучения методы обнаружения целей, основанные на глубоком обучении, постепенно становятся основным направлением исследований и демонстрируют более высокую точность и устойчивость в практических приложениях.

Результаты

Исходя из проблем традиционных методов обнаружения, методы обнаружения целей на основе глубокого обучения достигли значительного прогресса в последние годы с появлением моделей глубокого обучения, таких как конволюционные нейронные сети (CNN). Методы глубокого обучения могут автоматически изучать особенности изображения на основе большого количества данных, что позволяет избежать утомительной ручной разработки особенностей, а также лучше адаптироваться к сложным сценам и изменчивым целям. Благодаря постоянному развитию технологии глубокого обучения классификаторы, основанные на методах глубокого обучения, постепенно демонстрируют высокую производительность. Как показано на рисунке 2.1., с 2012 года AlexNet [3]и далее методы, основанные на глубоком обучении, начали появляться в больших масштабах, однако AlexNet, Overfeat[4] и R-CNN [5] все еще придерживаются подхода конволюционной нейронной сети к извлечению признаков, который, по сути, является стратегией скользящего окна, что все еще влияет на вычисления и производительность алгоритмов в реальном времени. До появления Fast-RCNN [6], который стал первым методом, заменившим традиционный метод скользящего окна на метод ящика-кандидата области, а также ознаменовал официальное появление двухэтапного алгоритма обнаружения целей. Сразу после появления YOLO [7,8,9] и других серий алгоритмов, эта серия алгоритмов использует одноэтапный подход прямой регрессии для замены оригинальной стратегии извлечения кадров-кандидатов регионов, что знаменует собой создание одноэтапных алгоритмов обнаружения.

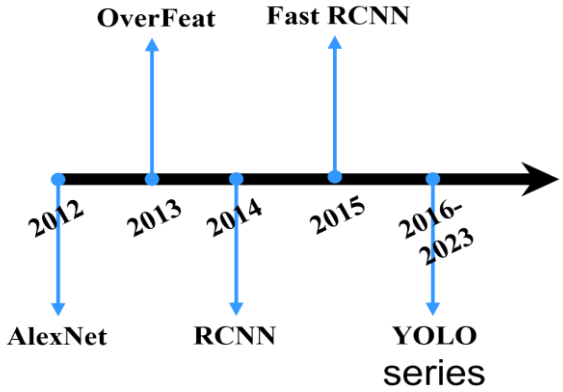

Рисунок 1 - Разработка алгаритмов обнаружения целей на основе глубокого обучения

Двухэтапный метод обнаружения

Первый этап двухэтапного метода обнаружения цели заключается в точном определении положения цели и поиске лучших кадров-кандидатов на основе обеспечения высокой запоминаемости и точности; второй этап классифицирует кадры-кандидаты, полученные на первом этапе, для более точной локализации цели. Двухфазный подход, естественно, требует большого количества вычислений, что приводит к низкой скорости обнаружения, но жертвуя скоростью, можно значительно повысить точность обнаружения. Основными репрезентативными алгоритмами являются R-

CNN, Fast-RCNN. Faster-RCNN [10]. Структура показана на рисунке 2, сначала входное изображение проходит операцию изменения размера для получения$H \times W$ Размер; подается в общую опорную сеть конволюционного слоя для извлечения признаков; используется сеть RPN (Region Proposal Networks) для генерации боксов-кандидатов, и после получения матрицы признаков путем сопоставления исходного изображения сгенерированного бокса-кандидата с изображением признаков, подается в softmax и bbxreg (bounding box regression) соответственно для получения более точного бокса-кандидата. Слой объединения ROI (Region of Interst) понижает выборку матрицы признаков в карту признаков 7×7, объединяет контекстную информацию для извлечения боксов-кандидатов и карт признаков, а затем отправляет их обратно в полностью связанный слой для классификации; наконец, он вычисляет вероятность классифицированных боксов в полностью связанном слое, а затем получает более точную информацию о местоположении и результаты классификации с помощью регрессии ограничивающих боксов.

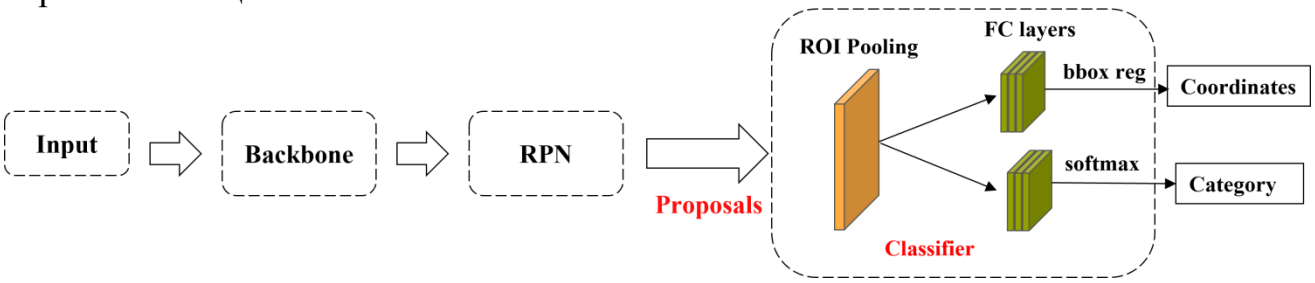

Рисунок 2 - Структура Faster-RCNN

Одноэтапные методы обнаружения

Одноэтапные методы отбрасывают этап поиска лучшего ящика-кандидата, одноразовая прямая регрессия местоположения и класса целевого объекта, через обнаружение можно напрямую вернуться к результатам, такие методы улучшают скорость обнаружения, жертвуя точностью обнаружения, подходят для широкого спектра промышленных развертываний приложений, и поэтому их ищут многие исследователи, основные представительные алгоритмы для YOLO серия алгоритмов [11].

Рисунок 3 - Одноэтапная сеть обнаружения целей

Одноэтапная сеть обнаружения целей, которая отказывается от процесса извлечения кадров-кандидатов в двухэтапном подходе и напрямую регрессирует позиции категории и кадров-кандидатов сразу, жертвует точностью обнаружения, но имеет более высокую скорость обнаружения, чем в литературе [12]. Она улучшается почти в десять раз и открывает новую эру сквозного обнаружения целей в реальном времени. Поток обнаружения этой

сети показан на рисунке 3. Сначала на вход поступает изображение, затем запускается одна CNN для извлечения признаков, и, наконец, выводится уровень доверия с помощью порогового значения, чтобы отбросить результаты с высокой ошибкой. Алгоритм использует глобальный вывод для повышения эффективности вычислений путем объединения изображений в фиксированную сетку.

Заключение

Несмотря на то, что подходы ИИ, а именно глубокого обучения, продемонстрировали значительные преимущества в обнаружении целей, все еще существуют некоторые проблемы:

Точность и эффективность обнаружения: хотя методы обнаружения на основе ИИ значительно превосходят традиционные методы, точность и эффективность обнаружения целей в некоторых сложных сценах (например, в многопользовательских сценах, в условиях недостаточной освещенности и т. д.) все еще сталкиваются с определенными проблемами.

Требования к вычислительным ресурсам: алгоритмы глубокого обучения и компьютерного зрения обычно требуют высокой вычислительной мощности, особенно в приложениях реального времени. Оптимизация алгоритмов для эффективной работы на маломощных устройствах - одно из направлений будущих исследований.

Метод обнаружения целей на основе глубокого обучения имеет следующие преимущества перед традиционными сенсорными методами:

Бесконтактность и высокая точность: алгоритмы искусственного интеллекта обеспечивают бесконтактный мониторинг с помощью видеоустройств без необходимости надевать оборудование, что позволяет избежать неудобств, связанных с ношением или повреждением оборудования.

Обработка данных в реальном времени и интеллектуальное принятие решений: алгоритмы искусственного интеллекта способны обрабатывать данные в реальном времени и автоматически корректировать решения в зависимости от изменений окружающей среды, что делает систему адаптивной в динамичной среде. Благодаря постоянному обучению и оптимизации способность системы к обнаружению будет улучшаться по мере использования.

REFERENCES

1. Xiaojuan, W. Yu. и др. // Laser & Optoelectronics Progress. – 2020. – Vol.57(12). – P.120005.

2. Choudhury, S. Vehicle detection and counting using haar feature-based classifier [Текст] / S. Choudhury, S.P. Chattopadhyay, T.K. Hazra // 2017 8th annual industrial automation and electromechanical engineering conference (IEMECON). – 2017. – P.106–109.

3. Zheng, C. и др. Fatigue driving detection based on Haar feature and extreme learning machine. // The Journal of China Universities of Posts and Telecommunications. – 2016. – Vol.23(4). – P.91–100.

4. Histogram of oriented gradients for human detection in video [Текст]

/ T. Surasak, I. Takahiro, C.H. Cheng и др. // 2018 5th International conference on business and industrial research (ICBIR). – 2018. – P.172–176.

5. Histogram of oriented gradients feature extraction from raw bayer pattern images [Текст] / W. Zhou, S. Gao, L. Zhang, X. Lou // IEEE Transactions on Circuits and Systems II: Express Briefs. – 2020. – Vol.67, №5. – P.946–950.

6. Krizhevsky, A. ImageNet Classification with Deep Convolutional Neural Networks [Текст] / A. Krizhevsky, I. Sutskever, G.E. Hinton // Advances in Neural Information Processing Systems (NeurIPS). – 2012. – Vol.25.

7. OverFeat: Integrated Recognition, Localization and Detection using Convolutional Networks [Текст] / P. Sermanet, D. Eigen, X. Zhang и др. // International Conference on Learning Representations (ICLR). – 2014.

8. Rich feature hierarchies for accurate object detection and semantic segmentation [Текст] / R. Girshick, J. Donahue, T. Darrell, J. Malik // IEEE Conference on Computer Vision and Pattern Recognition (CVPR). – 2014.

9. Girshick, R. Fast R-CNN [Текст] / R. Girshick // IEEE International Conference on Computer Vision (ICCV). – 2015.

10. You Only Look Once: Unified, Real-Time Object Detection [Текст] / J. Redmon, S. Divvala, R. Girshick, A. Farhadi // CVPR. – 2016.

11. Redmon, J. YOLO9000: Better, Faster, Stronger [Текст] / J. Redmon, A. Farhadi // CVPR. – 2017.

12. Redmon, J. YOLOv3: An Incremental Improvement [Текст] / J. Redmon, A. Farhadi // arXiv:1804.02767. – 2018.

Батырканов Ж.И.,
техника ғылымдары докторы, профессор,
bjenish@mail.ru[1]

Бийбосунов Б.И.,
техника ғылымдары докторы, профессор,
bbolotbek@mail.ru[1]

Лян Чжаньхао,
өтініш беруші,
liangzhanhao@kstu.kg [1]

И. Арабаев атындағы Қырғыз мемлекеттік университеті,
720026,Бишкек, Раззаков к-сі,51а[1]

ТЕРЕҢ ОҚЫТУҒА НЕГІЗГЕН МАҚСАТТЫ АНЫҚТАУ ӘДІСІ

Андатпа. *Мақалада терең оқыту технологиясын қолдануға негізделген мақсатты анықтау әдістері берілген. Haar және Gamma-Guided Object (HOG) мүмкіндіктерге негізделген анықтау алгоритмдері қарастырылады. Мақсатты анықтау әдістері соңғы жылдары конволюционды нейрондық*

желілер (CNN) сияқты терең оқыту үлгілерінің пайда болуымен айтарлықтай прогреске қол жеткізгені атап өтілді. Терең оқытуға негізделген мақсатты анықтау әдісінің дәстүрлі сенсорлық әдістерге қарағанда артықшылықтары көрсетілген.

Түйінді сөздер: *мақсатты анықтау әдістері, алгоритмдер, дәстүрлі әдістер, терең оқыту, мүмкіндіктер, компьютерлік көру.*

Batyrkanov Zh.I.,
Doctor of Technical Sciences, Professor, bjenish@mail.ru [1]

Biybosunov B.I.,
Doctor of Technical Sciences, Professor, bbolotbek@mail.ru[1]

Liang Zhanhao,
PhD student , liangzhanhao@kstu.kg [1]

Kyrgyz State University named after. I. Arabaeva,
720026, Bishkek, Razzakova street, 51a[1]

TARGET DETECTION METHODOLOGY BASED ON DEEP LEARNING

Abstract*. The article presents target detection methods based on the application of deep learning technology. Detection algorithms based on Haar features and features (HOG) are considered. It is noted that target detection methods have made significant progress in recent years with the advent of deep learning models such as convolutional neural networks (CNN). The advantages of the deep learning-based target detection method over traditional sensory methods are shown.*

Keywords*: target detection methods, algorithms, traditional methods, deep learning, features, computer vision.*