

Алгоритмы компьютерного зрения для распознавания объектов в условиях плохой видимости

С.А. Корчагин, Д.В. Сердечный

Финансовый университет при Правительстве Российской Федерации, Москва

Аннотация: Работа посвящена разработке и анализу алгоритмов компьютерного зрения, предназначенных для распознавания объектов в условиях ограниченной видимости, таких как туман, дождь или недостаточное освещение. В условиях современных требований к безопасности и автоматизации, задача идентификации объектов становится особенно актуальной. Рассмотрены теоретические основы методов компьютерного зрения и их применения в сложных условиях. Проведен анализ алгоритмов обработки изображений, включая методы машинного обучения и глубокого обучения, которые адаптированы для работы в условиях плохой видимости. Представлены результаты экспериментов, демонстрирующие эффективность предложенных подходов, а также сравнение с существующими системами распознавания. Результаты исследования могут быть полезны при разработке автономных транспортных средств и систем видеонаблюдения.

Ключевые слова: компьютерное зрение, математическое моделирование, программный комплекс, методы машинного обучения, автономные транспортные системы.

Введение

Современные технологии компьютерного зрения стремительно развиваются, открывая новые горизонты в сфере автоматизации и интеллектуального анализа данных [1-3]. Одной из наиболее актуальных задач в этой области является распознавание объектов в условиях плохой видимости, что представляет собой значительный вызов для существующих алгоритмов и систем. Плохие погодные условия, низкая освещенность, дым или туман могут существенно затруднить процесс идентификации объектов, что имеет критическое значение для таких секторов, как безопасность [4,5], автомобилизация [6,7] и охрана окружающей среды [8].

Разработка эффективных систем компьютерного зрения, способных адаптироваться к меняющимся условиям, требует комплексного подхода, включающего как совершенствование алгоритмов обработки изображений, так и использование современных технологий машинного обучения. В данной статье рассмотрен ряд инновационных решений, которые направлены на повышение надежности и точности распознавания объектов в сложных

условиях. В работе проанализированы существующие методы, проведено обсуждение их преимуществ и ограничений, а также представлены алгоритмы компьютерного зрения для распознавания объектов в условиях плохой видимости. Актуальность исследований в данной области очевидна, поскольку решение указанной задачи имеет критическое значение для многих сфер, включая безопасность и автономные транспортные средства.

Анализ методов распознавания объектов в условиях плохой видимости

Анализ существующих методов распознавания объектов в условиях плохой видимости включает несколько ключевых направлений, каждое из которых вносит свой вклад в решение этой сложной задачи. Во-первых, широко используются алгоритмы обработки изображений, такие как улучшение контраста, которое позволяет повысить видимость объектов в условиях низкой освещенности или тумана. Методы гистограммного выравнивания и адаптивной фильтрации помогают улучшить качество изображения, что в свою очередь может повысить производительность алгоритмов распознавания. Кроме этого, фильтрация шумов, включая применение средних, медианных и гауссовских фильтров, позволяет уменьшить нежелательный шум и общее качество изображения.

Глубокое обучение, в частности сверточные нейронные сети (CNN), продемонстрировало отличные результаты в задачах распознавания объектов [9]. Архитектуры нейронных сетей, такие как ResNet и DenseNet, показывают хорошие результаты даже в сложных условиях [10]. Использование аугментации данных, которая включает добавление шума, изменение яркости и контрастности, помогает создать более разнообразные наборы данных, что может улучшить обобщающую способность моделей.

В работе [11] приводятся техники фuzziрования – модели нечёткой логики, которые могут эффективно обрабатывать неопределенности и вариации при распознавании объектов. Эти модели играют важную роль в принятии решений на основе нечетких или неполных данных. Кроме того, использование сенсорных технологий, в частности инфракрасных и термальных камер, позволяет значительно улучшить качество распознавания объектов в условиях низкой видимости, так как такие устройства работают на основе теплового излучения. Системы слияния данных, комбинирующие данные с различных сенсоров (например, видимого спектра и инфракрасного), позволяют создать более полное представление о ситуации, что улучшает точность распознавания объектов.

Для учета временных зависимостей в распознавании объектов могут быть использованы рекуррентные нейронные сети (RNN) или их модификации, такие как Long Short-Term Memory (LSTM), что особенно полезно при анализе движущихся объектов в сложных условиях. Улучшенные модели сегментации, такие как U-Net и Mask R-CNN, помогают эффективно разделять объекты на изображениях, даже когда видимость ограничена, если они обучены на соответствующих данных.

Каждый из рассмотренных методов имеет свои преимущества и ограничения, однако комбинирование различных подходов — глубокого обучения, обработки изображений и сенсорных технологий — представляется наилучшей стратегией для повышения точности распознавания объектов в условиях плохой видимости.

Разработка алгоритмов компьютерного зрения

На первом этапе разработки системы был подготовлен набор данных для обучения, на основе фотографий, полученных из открытых источников [12] (рис. 1)

	filename	center_x	center_y	size	label
0	571.jpeg	1746	691	138	1
1	387.jpeg	1132	649	100	0
2	813.jpeg	166	642	108	0
3	1351.jpeg	1693	753	104	1
4	1367.jpeg	1320	647	112	1
...
2451	379.jpeg	634	697	130	0
2452	648.jpeg	508	633	102	1
2453	1700.jpeg	534	636	64	0
2454	691.jpeg	937	685	150	0
2455	1677.jpeg	1885	749	68	1

Рис. 1. – Набор данных для обучения системы в формате .CSV

Одной из ключевых задач было найти оптимальный алгоритм для обнаружения требуемых областей. Мы решили использовать детектор краев Canny. Во-первых, в данном алгоритме используется гауссовский фильтр для сглаживания изображения с целью удаления шума. Уравнение для ядра гауссовского фильтра размером $(2k+1) \times (2k+1)$ задается следующим образом:

$$H_{ij} = \frac{1}{2\pi\sigma^2} \exp\left(-\frac{(i - (k + 1))^2 + (j - (k + 1))^2}{2\sigma^2}\right); 1 \leq i, j \leq (2k + 1)$$

Детектор краев Canny позволяет находить градиенты интенсивности изображения и применять немаксимальное подавление, чтобы избавиться от ложного отклика на обнаружение краев. После этого в алгоритме применяется двойной порог для определения потенциальных краев. Мы использовали 80 для первого порога и 200 для второго. Затем детектор краев Canny завершает обнаружение краев, подавляя все остальные края, которые являются слабыми и не связаны с сильными краями. Основная идея заключается в том, что можно найти края только в области, где изображение

не размыто. Этот алгоритм сработал довольно хорошо на нашем образце изображения.

Далее необходимо было оценить наш алгоритм. Мы решили использовать метрику IoU. IoU определяется, как:

$$IoU = \frac{\text{площадь_перекрытия}}{\text{площадь_объединения}}$$

Для выполнения этой метрики мы преобразовали наши изображения с истинными и предсказанными областями в следующий формат: требуемая область была нарисована белым цветом, а остальная часть изображения была нарисована черным. Однако мы не смогли бы обнаружить что-либо на фотографиях, если его метрика IoU меньше 10, поэтому мы добавили 5,33 % наших фотографий в наш список, где изображения не были обнаружены и отбросили 9,61 % нашего набора данных для обучения, что не критично.

Более того, были обнаружены нетривиальные случаи, когда истинная область могла выходить за границы изображения (поскольку она была квадратной и мы знали ее центр и размер). Чтобы предотвратить это, проверялись границы истинной области на неотрицательность.

Среднее значение метрики IoU составило 78,21. Мы считаем, что этого достаточно, чтобы понять, есть ли на этой области человек или машина. 25 % квантиль для ширины новых изображений составил 71, а для высоты — 66.

В ходе исследования было принято решение масштабировать каждое новое изображение до размера 70x70. После разделения обучающего набора данных на обучающий и проверочный наборы данных была получена задача классификации изображений (рис. 2). В нашем алгоритме автомобилю была присвоена метка 0, а человеку — метку 1. Мы определили следующую структуру CNN (размеры ядер для всех сверток равны 3, а для объединения — 2): Conv2d (от 3 каналов до 4) → MaxPool2d (шаг 2) → Conv2d (от 4 каналов до 6) → преобразование матрицы MxN в вектор размерностью M*N

→ линейное преобразование (от $16 \times 16 \times 6$ до 500) → ReLU → линейное преобразование (от 500 до 50) → ReLU → линейное преобразование (от 50 до 2).

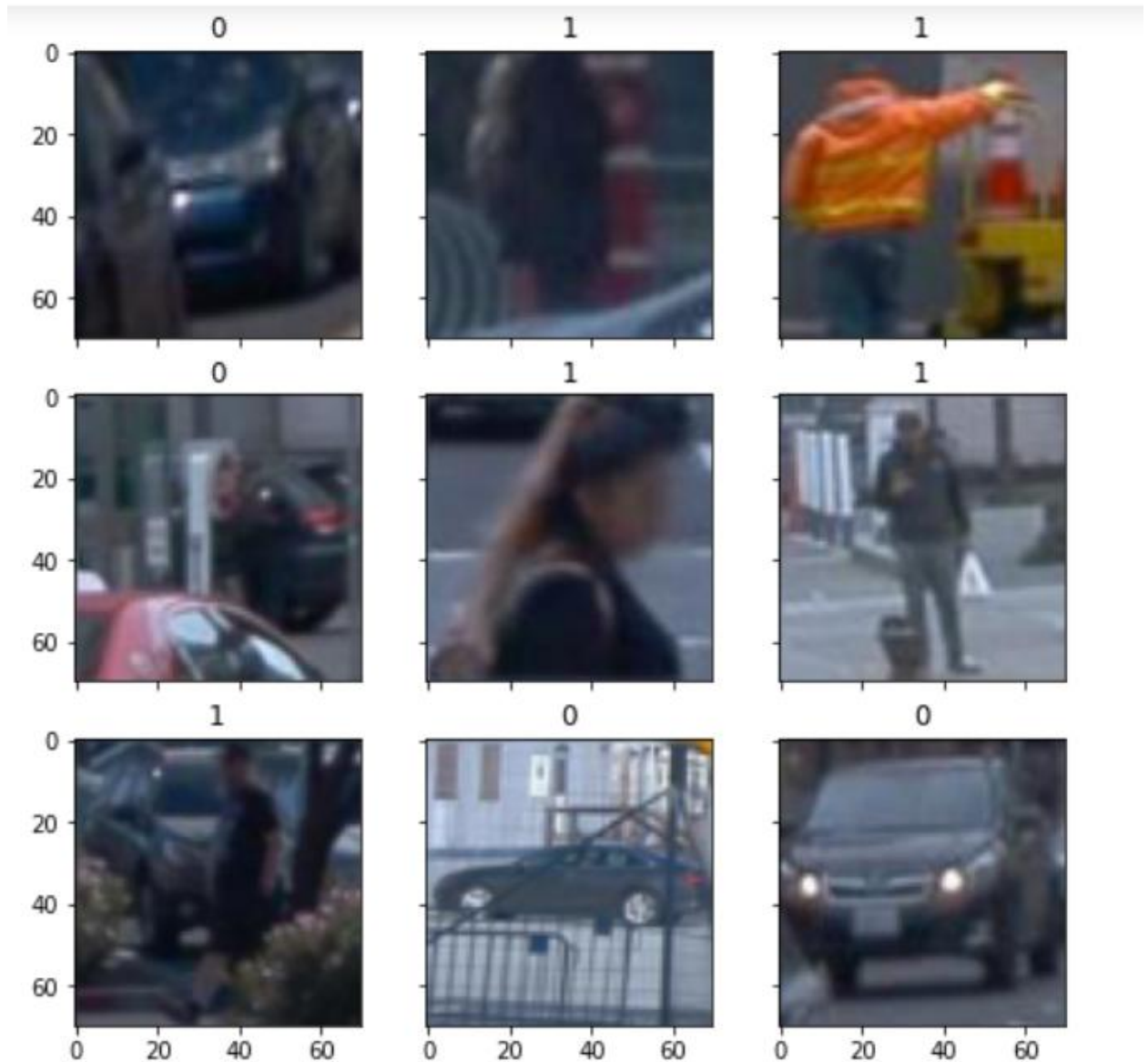


Рис. 2. – Набор данных для решения задач классификации изображений

Модель обучались на графическом процессоре Tesla K80. Мы использовали кросс-энтропийную потерю в качестве функции потерь и оптимизатор Adam со скоростью обучения $1e-4$. Сверточная нейронная сеть CNN обучалась в течение 10 эпох (рис. 3).

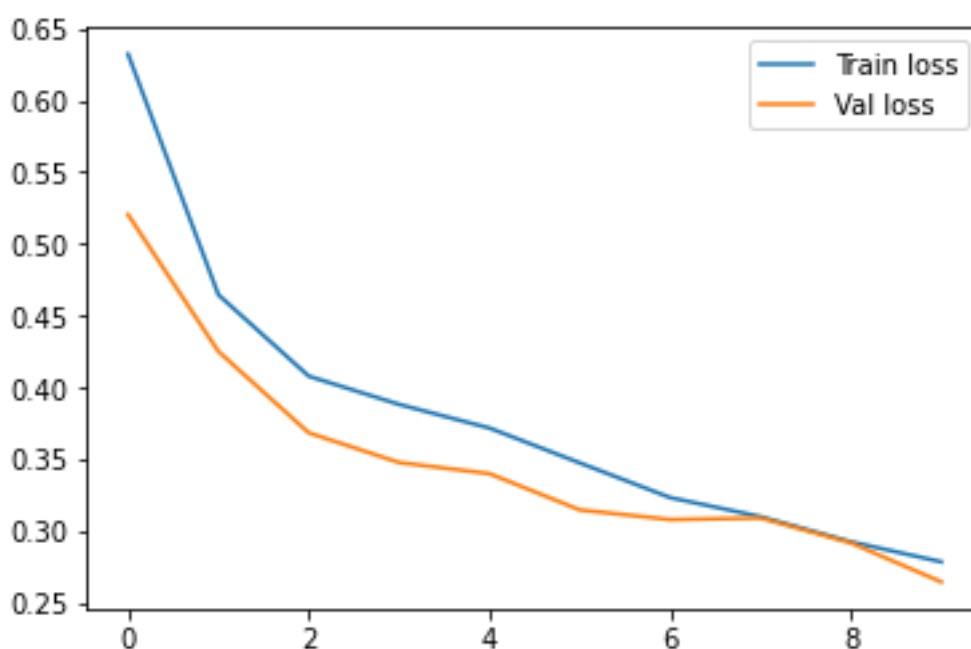


Рис. 3. –Потери сверточной нейронной сети CNN в зависимости от эпох обучения

В результате реализации алгоритма удалось достичь точности 90,625 % на проверочном наборе данных.

Заключение

В нашем исследовании была решена задача компьютерного зрения по распознаванию объектов в сложных погодных условиях. Результаты работы могут быть полезны в беспилотных автомобилях и средствах видеонаблюдения. В плане развития предложенных алгоритмов, актуальными вопросам остаются построение более сложной CNN, изменение ее гиперпараметров и расширения набора данных train новыми маркированными данными. Мы считаем, что эти действия могут повысить точность до более чем 90 %.

Литература

1. Горячкин Б. С., Китов М. А. Компьютерное зрение // E-scio. – 2020. – Т. 9. – № 48. – С. 317-345.

2. Kamyshova G., Kamyshova, G., Osipov, A., Gataullin, S., Korchagin, S., Ignar, S., Gataullin, T., Terekhova N. & Suvorov, S. Artificial neural networks and computer vision's-based phytoindication systems for variable rate irrigation improving //IEEE Access. – 2022. – Т. 10. – pp. 8577-8589.

3. Korchagin, S. A., Gataullin, S. T., Osipov, A. V., Smirnov, M. V., Suvorov, S. V., Serdechnyi, D. V., & Bublikov, K. V. Development of an optimal algorithm for detecting damaged and diseased potato tubers moving along a conveyor belt using computer vision systems //Agronomy. – 2021. – Т. 11. – №. 10. – С. 1980.

4. Панкова М. А., Картавцев Д. В. Применение систем видеонаблюдения для обеспечения пожарной безопасности объектов //Пожарная безопасность: проблемы и перспективы. – 2016. – Т. 2. – №. 1 (7). – С. 56-57.

5. Беспалова Н. В., Корчагин С.А., Сердечный Д.В., Селиверстов В.В. Анализ зарубежного опыта применения интеллектуальных методов в задачах защиты объектов критической информационной инфраструктуры финансового сектора //Инженерный вестник Дона. – 2024. – №. 5. URL: ivdon.ru/ru/magazine/archive/n5y2024/9196

6. Gajjar H., Sanyal S., Shah M. A comprehensive study on lane detecting autonomous car using computer vision //Expert Systems with Applications. – 2023. – Т. 233. – С. 120929.

7. Schneider, S., Taylor, G. W., Kremer, S. C., & Fryxell, J. M. Getting the bugs out of AI: advancing ecological research on arthropods through computer vision //Ecology Letters. – 2023. – Т. 26. – №. 7. – pp. 1247-1258.

8. Koger, B., Deshpande, A., Kerby, J. T., Graving, J. M., Costelloe, B. R., & Couzin, I. D. Quantifying the movement, behaviour and environmental context of group- living animals using drones and computer vision //Journal of Animal Ecology. – 2023. – Т. 92. – №. 7. – pp. 1357-1371.

9. Salehi A. W., Khan S., Gupta G., Alabdullah B. I., Almjally A., Alsolai H., Siddiqui T. & Mellit, A. A study of CNN and transfer learning in medical imaging: Advantages, challenges, future scope //Sustainability. – 2023. – Т. 15. – №. 7. – С. 5930.

10. Авс А., Андриянов Н.А., Соловьев В.И., Соломатин Д. А. Применение глубокого обучения для аугментации и генерации подводного набора данных с промышленными объектами // Вестник Южно-Уральского государственного университета. Серия: Компьютерные технологии, управление, радиоэлектроника. – 2023. – Т. 23. – №. 2. – С. 5-16.

11. Андриянов Н. А., Дементьев В. Е., Ташлинский А. Г. Обнаружение объектов на изображении: от критериев Байеса и Неймана–Пирсона к детекторам на базе нейронных сетей EfficientDet //Компьютерная оптика. – 2022. – Т. 46. – №. 1. – С. 139-159.

12. Sheeny M., De Pellegrin E., Mukherjee S., Ahrabian A., Wang S., & Wallace A. Radiate: A radar dataset for automotive perception in bad weather //2021 IEEE International Conference on Robotics and Automation (ICRA). – IEEE, 2021. – С. 1-7.

References

1. Goryachkin B. S., Kitov M. A. E-scio. – 2020. V. 9. № 48. pp. 317-345.
2. Kamyshova G., Kamyshova, G., Osipov, A., Gataullin, S., Korchagin, S., Ignar, S., Gataullin, T., Terekhova N. & Suvorov, S. IEEE Access. 2022. V. 10. pp. 8577-8589.
3. Korchagin, S. A., Gataullin, S. T., Osipov, A. V., Smirnov, M. V., Suvorov, S. V., Serdechnyi, D. V., & Bublikov, K. V Agronomy. 2021. V. 11(10). P. 1980.
4. Pankova M. A., Kartavtsev D. V. Pozharnaya bezopasnost: problemy i perspektivy. 2016. V. 2(1). pp. 56-57.



5. Bespalova N.V., Korchagin S.A., Serdechny D.V., Seliverstov V.V. Inzhenernyj vestnik Dona. 2024. № 5. URL: ivdon.ru/ru/magazine/archive/n5y2024/9196
6. Gajjar H., Sanyal S., Shah M. Expert Systems with Applications. 2023. V. 233. P. 120929.
7. Schneider, S., Taylor, G. W., Kremer, S. C., & Fryxell, J. M. Ecology Letters. 2023. V. 26(7). pp. 1247-1258.
8. Koger, B., Deshpande, A., Kerby, J. T., Graving, J. M., Costelloe, B. R., & Couzin, I. D. Journal of Animal Ecology. 2023. V. 92(7). pp. 1357-1371.
9. Salehi A. W., Khan S., Gupta G., Alabduallah B. I., Almjally A., Alsolai H., Siddiqui T. & Mellit A. Sustainability. 2023. V. 15(7). p. 5930.
10. Avs A., Andriyanov N.A., Solovyov V.I., Solomatin D.A. Vestnik Yuzhno-Uralskogo gosudarstvennogo universiteta. Seriya: Kompyuternyye tekhnologii. upravleniye. radioelektronika. 2023. V. 23(2). pp. 5-16.
11. Andriyanov N. A., Dementyev V. E., Tashlinskiy A. G. Kompyuternaya optika. 2022. V. 46(1). pp. 139-159.
12. Sheeny M., De Pellegrin E., Mukherjee S., Ahrabian A., Wang S., & Wallace A. Radiate: 2021 IEEE International Conference on Robotics and Automation (ICRA). IEEE. 2021. pp. 1-7.

Дата поступления: 22.09.2024

Дата публикации: 31.10.2024