<https://doi.org/10.38013/2542-0542-2021-1-93-102>

УДК 621.384.3:004.032.26

Некоторые результаты экспериментального исследования прототипа автономной инфракрасной системы распознавания наземных объектов

А. И. Мальцев¹, Д. Г. Откупман^{1,2}, В. К. Осташенкова¹, М. В. Останин¹

¹ Публичное акционерное общество «Научно-производственное предприятие «Импульс», Москва, Российская Федерация

² Федеральное государственное бюджетное образовательное учреждение высшего образования «Московский государственный университет геодезии и картографии» (МИИГАиК), Москва, Российская Федерация

Представлены результаты экспериментов с прототипом автономной инфракрасной системы распознавания наземных объектов на базе отечественных физических компонентов и открытой архитектуры сверточной нейронной сети YOLOv3. Объект распознавания – легковой автофургон. Нейронная сеть обучена на наборе изображений, снятых в видимом диапазоне. Анализируются инфракрасные видеокдры неидеального качества, снятые на движущемся и вибрирующем воздушном носителе – октокоптере.

Ключевые слова: сверточная нейронная сеть, беспилотный летательный аппарат, дальность обнаружения, распознавание

Для цитирования: Мальцев А. И., Откупман Д. Г., Осташенкова В. К., Останин М. В. Некоторые результаты экспериментального исследования прототипа автономной инфракрасной системы распознавания наземных объектов // Вестник Концерна ВКО «Алмаз – Антей». 2021. № 1. С. 93–102. <https://doi.org/10.38013/2542-0542-2021-1-93-102>

For citation: Maltsev A. I., Otkupman D. G., Ostashenkova V. K., Ostanin M. V. Experimental study of a prototype for an autonomous infrared system for ground object recognition// Vestnik Koncerna VKO "Almaz – Antey". 2021. No. 1. P. 93–102. <https://doi.org/10.38013/2542-0542-2021-1-93-102>

Поступила 25.11.2020 Отрецензирована 16.12.2020 Одобрена 18.01.2021 Опубликована 17.02.2021

Введение

Развитие теории и практики сверточных нейронных сетей (*Convolutional Neural Network*) обеспечило значительный прогресс алгоритмов автоматического распознавания объектов для различных гражданских и военных приложений. Типичный процесс построения системы распознавания с алгоритмом на базе нейронной сети предполагает наличие базы (*dataset*) изображений с маркированными (*labeled*) типовыми объектами в различных ракурсах и условиях. Алгоритм нейронной сети, реализованный в виде программы для ЭВМ, настраивается для обнаружения и распознавания заданного типа объектов путем обучения

(*training*) на большой базе изображений объектов, распознавание которых должна обеспечить система. Предполагается, что изображения должны быть сняты в том же диапазоне, в котором будет работать приемная часть системы распознавания. Однако для инфракрасной (ИК) области спектра подходящие наборы изображений, особенно для объектов, представляющих интерес в военных приложениях, малодоступны или имеют ограниченный объем и часто сняты в неподходящем ракурсе. Поэтому представляет интерес исследование возможностей распознавания при использовании базы изображений другого диапазона.

Применение различных архитектур сверточных нейронных сетей в системах, работающих в ИК диапазоне, было исследовано рядом авторов. В том числе в [1] было рассмотрено

применение различных архитектур, обученных на наборе данных видимого диапазона и на наборе данных, предоставленных компанией *FLIR*. Однако в статье не рассматривалась возможность использования архитектуры *YOLO*.

Для того чтобы заранее рассмотреть все возможности архитектуры *YOLOv3*, авторами данной статьи были изучены материалы, представленные в [2, 3]. Однако, несмотря на схожесть исследований с нашим исследованием в части обучения известной архитектуры *YOLOv3* на наборе изображений видимой области электромагнитного спектра для применения в ИК системах, в большинстве зарубежных публикаций основное внимание уделяется распознаванию людей, что не представляло интереса для нас, к тому же в известных публикациях не уделено внимание анализу результатов обработки видеопотока, снятого с колеблющегося и вибрирующего носителя.

В доступных отечественных публикациях представленный в настоящей работе подход также не отражен. Рассматриваются либо автоматические тепловизионные системы распознавания с другими архитектурами, либо иные средства распознавания [4, 5].

Основная часть

Проведено экспериментальное исследование прототипа автономной системы распознавания

наземных объектов со стороны верхней полусферы, построенной на основе приемной части (объектив, матрица) длинноволновой части ИК диапазона и алгоритма распознавания на базе нейронной сети. Для практической реализации приемной части выбраны отечественные компоненты компании «ОКБ «Астрон» [6] – микроболометрический модуль на оксиде ванадия с разрешением 640×480 , шагом 17 мкм и беззатворной калибровкой, а также тепловизионный германиевый объектив с пассивной атермализацией, фокусным расстоянием 100 мм, относительным отверстием 1:1,4 и качеством изображения близким к дифракционному пределу. Приемная часть с аппаратурой передачи данных размещалась на беспилотном летательном аппарате (БЛА) – октокоптере. Упрощенная схема эксперимента представлена на рисунке 1. Общий вид полетной части экспериментального оборудования представлен на рисунке 2.

Для разработки алгоритма распознавания выбрана описанная в открытых источниках архитектура сверточной нейронной сети *YOLO* (*You Only Look Once*), обеспечивающая распознавание множественных объектов на изображениях и работающая с многочисленными классификациями объектов. По сравнению с другими известными архитектурами *YOLOv3* является одной из наиболее точных и быстрых действующих [7].

Рис. 1. Упрощенная схема эксперимента

Рис. 2. Трехмерная модель беспилотного носителя с аппаратурой

Если говорить о теоретических основах работы алгоритма *YOLOv3*, то стоит упомянуть, что классификатор любой сверточной нейронной сети обычно при каждой интерпретации делает предположение того, какой тип объекта находится в окне. Для каждого изображения приходится выполнять большое количество прогнозов, характеризуемое четырехзначными числами. В связи с этим работа алгоритма проходит довольно медленно. Именно эту отрицательную черту помогает устранить архитектура *YOLO*, преимущество которой в том числе отражено в названии – «*You Only Look Once*».

Сверточная нейронная сеть *YOLO* использует все изображение для предопределения каждой ограничивающей рамки (*bounding box*). Она также прогнозирует все рамки для каждого класса изображений одновременно. Это означает, что *YOLO* за один раз анализирует все изображение и все объекты, присутствующие на нем. Этот факт становится ключевым для продолжительности обработки кадра. Подтверждением преимуществ варианта архитектуры *YOLOv3* является график, представленный на рисунке 3, где показано

быстродействие *YOLOv3* в сравнении с другими известными архитектурами. Как следует из данных, приведенных на рисунке 3, при обучении на одном и том же наборе данных *COCO dataset* и при сравнимой средней точности обнаружения объекта (*mAP* – *mean Average Precision*) *YOLOv3* превосходит остальные рассмотренные сети по скорости нахождения объекта.

Для создания специализированной программы, реализующей нейронную сеть, выбран высокоуровневый язык программирования *Python* с дополнениями в виде открытых библиотек (модулей) для научных вычислений, глубокого обучения и компьютерного зрения (*NumPy*, *TensorFlow*, *Keras*, *OpenCV*).

В качестве объекта наблюдения был выбран легковой автофургон (4150×1960×1820 мм).

Для выбора оптимальной дальности распознавания объекта определенного размера можно воспользоваться упрощенным критерием, рассмотренным в [8]:

$$L(N) = \frac{v \cdot f' \cdot h_{кр}}{N},$$

Рис. 3. График быстродействия на примере различных архитектур, обученных на наборе данных *COCO* (*Common Objects in Context*) [7]

— YOLOv3, — RetinaNet-50, — RetinaNet-101

Рис. 4. Теоретическая зависимость дальности обнаружения от заданного критерием обнаружения числа пикселей разрешения на изображении объекта при критическом размере объекта 2,75 м. Зависимость получена для выбранной конфигурации приемной системы в геометрическом приближении

где L – дальность распознавания (дистанция до объекта), м;
 $\nu = 58,82$ 1/мм – пространственная частота;
 $f' = 100$ мм – фокусное расстояние;
 $h_{кр} = 2,7$ м – критический размер объекта;
 N – количество активных элементов (пикселей), шт.

На рисунке 4 приведена определенная в геометрическом приближении зависимость дальности обнаружения нашего объекта от числа пикселей разрешения, заданного критерием обнаружения.

Исходя из допущения о разрешении, заведомо достаточном для обнаружения рассматриваемого объекта, а также исходя из удобства работы на площадке, выбранной для экспериментов, съемку объекта наблюдения было

решено проводить с борта БЛА на дальности около 250 м. Для того чтобы оценить вероятности распознавания объекта на других дальностях, полученные исходные изображения обрабатывались путем объединения пикселей разрешения, кратного изменению дальности.

Представленные ниже результаты съемок получены в дневное время на фоне подстилающей поверхности типа «неоднородный луг – грунтовая дорога». Объект наблюдения (автофургон) в процессе съемки разворачивался, соответственно менялся ракурс наблюдения объекта. Начальный и конечный кадры необработанной (без цифровых улучшений) видеозаписи наблюдаемого сюжета в длинноволновой части ИК диапазона представлен на рисунке 5.

а

б

Рис. 5. Начальный и конечный кадры съемки сюжета с объектом наблюдения

В связи с отсутствием доступных достаточно объемных баз изображений объектов в длинноволновой части ИК диапазона обучение нейросети проводилось по открытой базе изображений видимого диапазона *COCO dataset* [9] с маркированными объектами типов «автомобиль», «автобус», «грузовой автомобиль». Отметим, что создатели *YOLO* предоставляют открытый исходный код и подробное описание для обучения собственной модели на различных наборах данных, в том числе на наборе данных *COCO dataset* [10]. Таким образом, при обучении использовались изображения видимого диапазона форматом 416×416 , количество которых составляло более 10 000, точность распознавания при обучающей выборке составляла более 0,8 (80 %).

Обученной таким образом программе, реализующей нейросеть, предъявлялись кадры видеосюжета с объектом, разворачивающимся от положения «сверху под углом – в лоб» до положения «сверху под углом – в борт». Программа обнаруживала объект на кадре и распознавала его, определяя тип объекта из набора «автомобиль», «автобус», «грузовой автомобиль» и оценивая вероятность распознавания типа объекта (рис. 6).

Соответственно, для типа «автомобиль» это будет вероятность правильного распознавания, а применительно к двум другим типам – вероятность принятия действительного объекта за объект другого типа, то есть вероятность ложного распознавания.

Для иллюстрации того, как меняются точечные оценки вероятностей от кадра к кадру при развороте наблюдаемого объекта, был построен график, на котором ось абсцисс – номер кадра, ось ординат – вероятность. Пример, поясняющий структуру графика, приведен на рисунке 7.

Следует отметить сильный разброс оценок вероятности от кадра к кадру при достаточно устойчивом характере аппроксимирующих оценок, что свидетельствует о целесообразности накопления оценок по нескольким кадрам для повышения вероятности правильного принятия решения алгоритмом распознавания, реализуемом в конечном изделии.

Рис. 6. Визуализация процесса распознавания для одного из кадров. Над рамкой, которой выделен обнаруженный объект, показан тип объекта, распознанного на кадре с максимальной вероятностью, и значение вероятности

Рис. 7. Графическое представление процесса анализа последовательности кадров при развороте объекта наблюдения. Абсцисса – номер кадра, ордината – вероятность. Вероятность правильного распознавания типа «автомобиль» – красные точки, вероятность ложного распознавания типа «автобус», «грузовой автомобиль» – треугольник синего и ромб зеленого цвета. Точки, попавшие на ось абсцисс, соответствуют ситуациям, когда оценка программой вероятности распознавания составляет менее 0,1. Сплошные цветные линии – линейная аппроксимация значений вероятностей

— Автомобиль, — Автобус, — Грузовой автомобиль

Рис. 8. Результаты покадровой оценки вероятности распознавания в обнаруженном объекте типа «автомобиль» (правильное), «автобус», «грузовой автомобиль» (ложное) для последовательности кадров от положения легкового автофургона «в лоб» до положения «в борт». Линейная аппроксимация показывает, как менялась усредненная вероятность в зависимости от ракурса наблюдения объекта. Точки на оси абсцисс соответствуют ситуациям, когда оценка вероятности распознавания составляет менее 0,1

— Автомобиль, — Автобус, — Грузовой автомобиль

Результаты обработки последовательных кадров от положения «в лоб» до положения «в борт» представлены на рисунке 8.

Для оценки вероятности распознавания объекта на дальностях 500 и 1000 м проводилась цифровая обработка полученного видеофайла с целью имитации увеличения дальности наблюдения (объединение 2×2 и 4×4 соседних пикселей изображения в кадре

с усреднением яркости). Результаты анализа для обработанных последовательностей кадров представлены на рисунке 9.

Сравнение результатов, представленных на рисунках 8 и 9, показывает, что при увеличении дальности (ухудшении разрешения) происходит существенное снижение вероятности правильного распознавания объекта при неблагоприятном ракурсе наблюдения.

Рис. 9. Результаты оценки вероятности распознавания в обнаруженном объекте типа «автомобиль», «автобус», «грузовой автомобиль»: а) при имитации дальности наблюдения 500 м (разрешение 320×240); б) при имитации дальности наблюдения 1000 м (разрешение 160×120) объект распознается уверенно как «автомобиль» только при развороте в положение «в борт». Линейная аппроксимация показывает, как менялась усредненная вероятность в зависимости от ракурса наблюдения объекта. Точки на оси абсцисс соответствуют ситуациям, когда оценка вероятности распознавания составляет менее 0,1

— Автомобиль, — Автобус, — Грузовой автомобиль

Для качественной иллюстрации этого в таблице для разных дальностей приведены результаты линейной аппроксимации оценки вероятности в начале анализируемой записи при расположении «сверху под углом – в лоб»

и для конца записи при расположении «сверху под углом – в борт».

Для проверки работоспособности обученной программы распознавания в сложной фоновой обстановке при наличии дрожаний

Таблица

Результаты линейной аппроксимации вероятности правильного распознавания для разных дальностей и ракурсов наблюдения

Дальность	Линейно аппроксимированная вероятность	
	«сверху под углом – в лоб»	«сверху под углом – в борт»
250 м	~ 0,38	~ 0,7
500 м	~ 0,26	~ 0,7
1000 м	0	~ 0,6

а

б

Рис. 10. Фрагмент видеозаписи сюжета «одно-двухэтажные дома – заборы – дороги с покрытием – легковые автомобили»: а) исходная видеозапись; б) видеозапись с распознаванием.

камеры и «смаза» изображения была проведена видеосъемка сюжета пригородной застройки «одно-двухэтажные дома – заборы – дороги с покрытием – легковые автомобили» при произвольных движениях БЛА – носителя аппаратуры. Фрагмент полученной видеозаписи в исходном виде и с обнаруженными программой распознавания объектами приведен на рисунке 10.

Реализованный в программе распознавания алгоритм сети продемонстрировал в условиях смаза и дрожания уверенное распознавание визуально хорошо различимого на фоне объекта. Но даже более важным представляется то, что, во-первых, была получена ненулевая оценка вероятности правильного распознавания в сложных фоновых условиях визуально плохо различимого объекта (ближе к левому верхнему углу картинки), и, во-вторых, ни разу комбинация контрастных фоновых участков не была принята за искомый объект.

Выводы

1. Разработан прототип автономной системы распознавания наземных объектов на основе отечественного матричного приемника длинноволновой части ИК диапазона, отечественного объектива и программы, реализующей алгоритм сверточной нейронной сети с архитектурой типа *YOLOv3*, взятой из открытых источников.

2. Экспериментальное исследование разработанного прототипа видеосъемкой с воздушного носителя и последующей наземной

обработкой кадров показало его работоспособность, т.е. способность, при размещении на движущемся воздушном носителе, выдавать ИК изображения с разрешением, достаточным для обнаружения и распознавания заданного объекта, а также способность программы распознавания обеспечивать распознавание типа объекта.

3. Нейронная сеть, обученная на наборе изображений типовых объектов в видимом диапазоне, показала достаточно высокую вероятность распознавания объектов на изображениях, полученных в длинноволновой части ИК диапазона, что упрощает подготовку баз изображений при разработках реальных прикладных устройств.

4. Обработка видеопотока, снятого с воздушного носителя, подверженного колебаниям и вибрациям, показала наличие существенного разброса оценок вероятности распознавания объекта для соседних кадров видеосъемки, поэтому при разработке алгоритма функционирования реального прикладного устройства целесообразно проводить совместную обработку серии последовательных кадров с накоплением оценки вероятности.

5. Созданный прототип системы распознавания в составе приемной части, выполненной на отечественной элементной базе, и открытого алгоритма сверточной нейронной сети *YOLOv3* может служить основой для решения прикладных задач, в том числе для разработок координаторов, способных автономно обнаруживать заданные объекты.

Список литературы

1. Devaguptapu Ch., Akolekar N., Sharma M.M., Balasubramanian V.N. Chaitanya Devaguptapu Borrow from Anywhere: Pseudo Multi-modal Object Detection in Thermal Imagery [Electronic resource] // IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshop (CVPRW). 2019. 10 p. URL: <https://www.semanticscholar.org/paper/Borrow-From-Anywhere%3A-Pseudo-Multi-Modal-Object-in-Devaguptapu-Akolekar/b95ce8e6249302e9f4c88248ab587f43760d88b1> (date of request: 01.10.2020). DOI: 10.1109/CVPRW.2019.00135
2. Kristo M., Ivasic-Kos M., Pobar M. Thermal Object Detection in Difficult Weather Conditions Using YOLO // IEEE Access. 2020. Vol. 8. P. 125459–125476. DOI: 10.1109/ACCESS.2020.3007481
3. Ivašić-Kos M., Kristo M., Pobar M. Human detection in thermal imaging using YOLO // Conference Paper. April, 2019. 5 p. URL: <https://www.researchgate.net/publication/333360405> (date of request: 01.10.2020). DOI: 10.1145/3323933.3324076
4. Мингалев А.В., Белов А.В., Габдуллин И.М., Агафонова Р.Р., Шушарин С.Н. Распознавание тест-объектов на тепловизионных изображениях // «Компьютерная оптика». 2019. Т. 43, № 3. С. 402–411. DOI: 10.18287/2412-6179-2019-43-3-402-411
5. Фомичева О.А., Стреляев С.И. Методы распознавания ИК-изображения // Известия ТулГУ. Технические науки. 2018. Вып. 11. С. 207–213.
6. Оптико-механическое конструкторское бюро «Астрон»: сайт. – МО, г. Лыткарино, 2020. URL: <https://astrohn.ru> (дата обращения: 01.09.2020).
7. Redmon J., Farhadi A. YOLOv3: An Incremental Improvement. 8 April, 2018. arXiv:1804.02767v1 [cs.CV]. URL: <https://arxiv.org/pdf/1804.02767.pdf> (date of request: 01.09.2020).
8. Якушенков Ю.Г., Тарасов В.В. Инфракрасные системы «смотрящего» типа. М.: Лорос, 2004. 444 с.
9. COCO – Common Objects in Context: website. 2015. URL: <https://cocodataset.org> (date of request: 01.09.2020).
10. Survival Strategies for the Robot Rebellion: website. 2015. URL: <https://pjreddie.com/> (date of request: 01.09.2020).

Об авторах

Мальцев Андрей Иванович – кандидат технических наук, старший научный сотрудник, главный научный сотрудник – заместитель главного конструктора Публичного акционерного общества «Научно-производственное предприятие «Импульс», Москва, Российская Федерация.

Область научных интересов: высокоточное вооружение, системы самонаведения.

Откупман Дмитрий Григорьевич – старший инженер-исследователь Публичного акционерного общества «Научно-производственное предприятие «Импульс», Москва, Российская Федерация; преподаватель кафедры оптико-электронных приборов Федерального государственного бюджетного образовательного учреждения высшего образования «Московский государственный университет геодезии и картографии» (МИИГАиК), Москва, Российская Федерация.

Область научных интересов: синтез оптических систем, тепловидение, компьютерное зрение, фотоника, лазерная техника.

Осташенкова Виктория Константиновна – инженер-исследователь Публичного акционерного общества «Научно-производственное предприятие «Импульс», Москва, Российская Федерация.

Область научных интересов: инфракрасная техника, искусственные интеллектуальные системы.

Останин Михаил Васильевич – начальник отдела Публичного акционерного общества «Научно-производственное предприятие «Импульс», Москва, Российская Федерация.

Область научных интересов: инфракрасная техника, оптико-электронные системы, лазерная дальнометрия.

Experimental study of a prototype for an autonomous infrared system for ground object recognition

Maltsev A. I.¹, Otkupman D. G.^{1,2}, Ostashenkova V. K.¹, Ostanin M. V.¹

¹ JSC Research and Production Enterprise "Impulse", Moscow, Russian Federation

² Moscow State University of Geodesy and Cartography (MIIGAiK), Moscow, Russian Federation

The results of experiments with a prototype of an autonomous infrared system for recognition of ground objects based on domestic physical components and open architecture of the YOLOv3 convolutional neural network are presented. The object of recognition is a car van. The neural network is trained on a set of images taken in the visible range. Infrared video footage of imperfect quality recorded by a moving and vibrating air carrier – octocopter – is analysed.

Keywords: convolutional neural network, unmanned aerial vehicle, detection range, recognition

Information about the authors

Maltsev Andrey Ivanovich – Cand. Sci. (Engineering), Senior Researcher, Deputy Chief Designer, JSC Research and Production Enterprise "Impulse", Moscow, Russian Federation.

Research interests: precision weapons, homing systems.

Otkupman Dmitriy Grigoryevich – Senior Research Engineer, JSC Research and Production Company "Impulse", Moscow, Russian Federation; Lecturer, Department of Optical Electronic Devices, Moscow State University of Geodesy and Cartography (MIIGAiK), Moscow, Russian Federation.

Research interests: synthesis of optical systems, thermal imaging, computer vision, photonics, laser technology.

Ostashenkova Viktorya Konstantinovna – Research Engineer, JSC Research and Production Company "Impulse", Moscow, Russian Federation.

Research interests: infrared technology, artificial intelligent systems.

Ostanin Mikhail Vasilievich – Departmental Head, JSC Research and Production Company "Impulse", Moscow, Russian Federation.

Research interest: infrared technology, optic-electronic systems, laser rangefinder.