

Д. В. Жуков¹, А. Г. Саидов, В. В. Терентьева¹, Е. И. Астахова¹,

¹ Федеральное государственное бюджетное военное образовательное учреждение высшего образования «Военно-космическая академия имени А.Ф. Можайского» Министерства обороны Российской Федерации, Санкт-Петербург, Россия
e-mail: spb.pillgrim@gmail.ru

ОПЫТ ОБУЧЕНИЯ НЕЙРОННОЙ СЕТИ YOLOV8 НА ОГРАНИЧЕННОМ ОБЪЕМЕ ОБУЧАЮЩИХ ДАННЫХ ПРИ РАСПОЗНАВАНИИ АНТРОПОГЕННЫХ ОБЪЕКТОВ С ВЫРАЖЕННОЙ ТЕКСТУРОЙ НА КОСМИЧЕСКИХ ИЗОБРАЖЕНИЯХ

Аннотация. В статье описан успешный опыт обучения нейронной сети YOLOv8 на ограниченном объеме обучающих данных. В качестве решаемой задачи рассматривалась возможность идентификации на цветных трехканальных космических снимках высокого разрешения уникальных по своим геометрическим характеристикам транспортных контейнеров. Всего при обучении нейронной сети было использовано 109 изображений с 1017 меток. Оптимальные для обучения гиперпараметры подбирались с помощью организации итеративного процесса, основанного на генетическом алгоритме. Проблема ошибочного распознавания ложных целей была решена путем добавления в исходный набор данных неразмеченных изображений. В результате проведенных работ была сформирована нейросетевая модель с показателем AP равным 0,86.

Ключевые слова: обучение нейронной сети, ограниченный объем обучающих данных, распознавание антропогенных объектов, подбор значений гиперпараметров, минимизация ошибок второго рода.

Развитие космических систем дистанционного зондирования Земли, общедоступность данных космической съемки, повышенные требования к оперативности обработки материалов съемки обуславливают необходимость автоматизации процесса дешифрирования космических снимков (КС). Одним из основных механизмов решения этой задачи в настоящее время является применение нейросетевых технологий, которые позволяют сравнительно быстро обрабатывать большие объемы данных, и при этом показывают высокую точность при распознавании объектов на изображениях [1], исключая при этом человеческий фактор. Вместе с тем, разработка нейросетевых моделей требуют значительных временных и ресурсных затрат. Связано это прежде всего с необходимостью формирования обучающих наборов данных больших объемов [2]. На практике получить достаточное для полноценного обучения нейронных сетей (НС) количество изображений объектов интереса, которые требуется обнаружить на КС, не всегда представляется возможным, особенно это касается уникальных либо малоизученных элементов сцены. Так, например, при анализе структуры сложносоставного (состоящего из множества отдельных элементов) антропогенного объекта не редко возникает необходимость идентификации отдельных его элементов, которые, в силу развития научно-технического прогресса, могут присутствовать только на

ограниченном количестве обследуемых промышленных предприятий. В ВКА им. А.Ф. Можайского решалась задача, заключающаяся в распознавании на высокодетальных КС складов транспортных контейнеров (СТК). Особенность поставленной задачи заключалась в том, что искомые элементы представляли собой уникальные контейнеры, отличающиеся от повсеместно используемых аналогов прежде всего геометрическими размерами. Анализ текстурных характеристик изображений объекта интереса показал неудовлетворительную разделимость искомых СТК с другими аналогичными элементами – ложными целями (ЛЦ): энтропия Шеннона [3] принимала значение в районе 0,83. Поэтому было принято решение обучить НС для их распознавания. Основная сложность при этом заключалась в том, что достоверная информация о наличии складов с искомыми контейнерами была получена только для 11 промышленных предприятий, что существенно ограничивало объем обучающих данных, дополнительное увеличение которого могло быть осуществлено за счет привлечения разновременных изображений одного и того же объекта и кадрирования больших по площади СТК.

В результате разметки был сформирован набор данных, статистические характеристики которого представлены в таблице 1.

Таблица 1 – Статистические характеристики набора данных для обучения НС при распознавании СТК

Характеристика	Значение
Общее число изображений	109
Общее число классов	1
Общее число меток (объектов интереса)	1017
ЛРМ изображений (м)	0.18...0.68

Обученная на представленном наборе данных НС YOLOv8 [4] имела значение метрики AP равной 0,53.

Повышение показателей качества обучения НС на практике может быть достигнуто за счет подбора оптимальных значений гиперпараметров (ГП).

Настройка значений ГП производилась с помощью генетического алгоритма [5]. Графическое представление этого процесса представлено на рисунке 1.

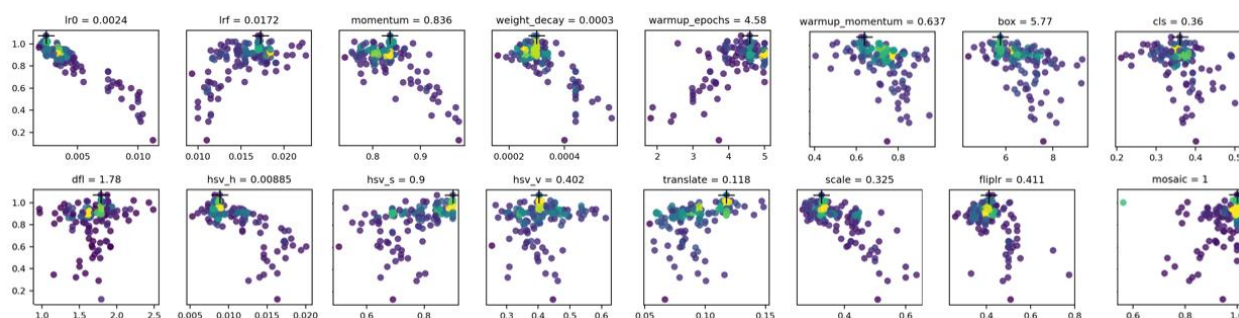


Рисунок 1 – Графическое представление процесса подбора оптимальных значений ГП при обучении YOLOv8 для распознавания СТК

Названия диаграмм, приведенных на рисунке 1, соответствуют названиям варьируемых значений ГП: «lr0», «lrf», «momentum», «warmup_epochs», «warmup_momentum» – блок ГП задающих скорость обучения НС, «weight_decay» – коэффициент, используемый при расчете штрафов в случае переобучения НС, «box», «cls», «dfl» – блок коэффициентов, используемых при расчете функций потерь, «hsv_h», «hsv_s», «hsv_v» – параметры яркостной коррекции при аугментации данных, «translate», «scale», «flipr» – параметры геометрической коррекции при аугментации данных, «mosaic» – параметр, определяющий возможность «склейки» изображения из различных частей при аугментации данных.

Настройка оптимальных значений ГП позволила увеличить метрику AP с 0,53 до 0,86.

Несмотря на сравнительно высокие значения метрики AP проверка работы обученной модели НС на дополнительных данных, представляющих собой территории антропогенных объектов, на которых отсутствуют искомые элементы, выявила значительное количество ЛЦ, которые НС распознает как СТК. К типичным ЛЦ относятся: солнечные панели на крышах зданий, участки дорожной разметки, водная поверхность при волнении, участки железнодорожных путей, открытые склады похожих типов контейнеров, сельскохозяйственные поля. При проверке работы НС на 4570 изображениях, ЛЦ, распознанные как СТК, присутствовали на 348 из них.

Для минимизации ошибок второго рода в исходный набор обучающих данных были добавлены 631 неразмеченных изображений типовых ЛЦ. Обученная на новом наборе данных НС позволила значительно уменьшить количество ошибок второго рода: число изображений с ЛЦ сократилось с 348 до 17. Сравнение графиков, иллюстрирующих качество обучения первоначальной модели НС и модели НС, после доработок (подбора оптимальных значений ГП и включение в состав обучающих данных изображений типовых ЛЦ) приведено на рисунке 2.

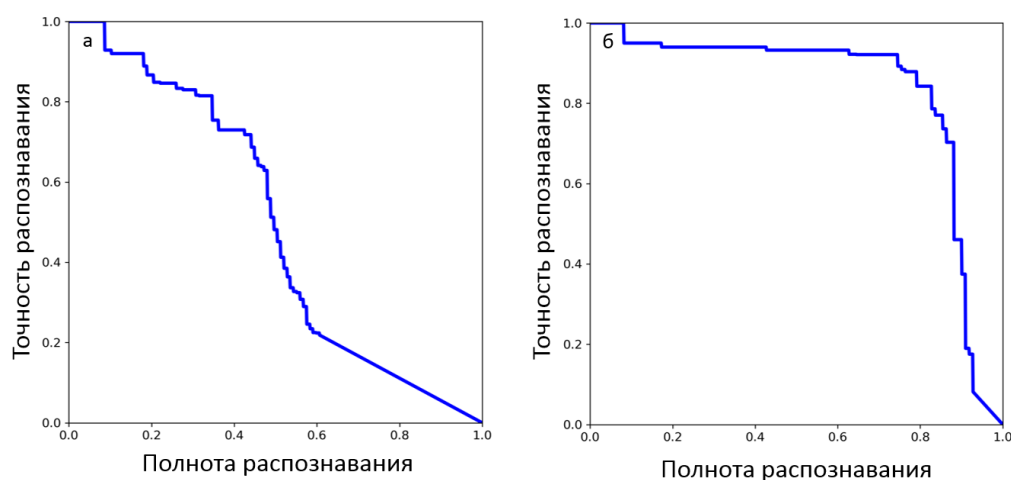


Рисунок 2 – Зависимости точности от полноты распознавания обученных моделей НС YOLOv8 при различных пороговых значениях степени уверенности НС: а) – исходная модель НС, б) – модель НС после доработки.

Таким образом была решена задача обучения модели НС на базе архитектуры YOLOv8 в условиях ограниченного объема обучающих данных. В качестве распознаваемых объектов был выбран только один элемент: СТК, имеющий устойчивые текстурные признаки. Это позволило значительно упростить процесс обучения НС, так как классификатор содержал только один распознаваемый класс и в последствии организовать подбор оптимальных значений ГП с помощью генетического алгоритма, что в других случаях (при разработке модели НС с полноценным классификатором и достаточным объемом обучающих данных) требует значительных временных и ресурсных затрат. Добавление в обучающий набор данных неразмеченных изображений с типовыми ЛЦ минимизировало количество ошибок второго рода.

Список литературы

1. Захватов Д.М., Лукиных А.А. Распознавание изображений с использованием нейронных сетей: проблемы и перспективы // Современные научные исследования и инновации. – 2023. – № 9.
2. Окунев С.В. Рассмотрение способов формирования наборов данных для обучения нейронных сетей // Вестник науки и образования. – 2020. – № 2 (80). Часть 3. – С. 16-19.
3. Абульханов С.Р. Использование информационной энтропии Шеннона для параметризации шероховатой поверхности деталей машиностроения // Известия Самарского научного центра Российской академии наук. – 2022. – т. 24, № 1. – С. 44-54.
4. Mupparaju Sohan, Thotakura Sai Ram, Venkata Rami Reddy Ch. A Review on YOLOv8 and its Advancements // Data Intelligence and Cognitive Informatics. – 2024. – С. 529-545.
5. Давронов Ш.Р. Обзор современных генетических алгоритмов и их применение на практике // Молодой ученый. – 2023. – № 36 (483). – С.15-18.

D.V. Zhukov¹, A.G. Saidov, V.V. Terentyeva¹, E.I. Astakhova¹

¹Military Space Academy of A. F. Mozhaisky, Saint-Petersburg, Russian
e-mail: vka@mil.ru

LIMITED-DATA TRAINING OF YOLOV8 FOR ANTHROPOGENIC OBJECT DETECTION WITH DISTINCT TEXTURAL FEATURES IN REMOTE SENSING IMAGERY

This paper presents a successful case study of training a YOLOv8 neural network with limited training data. The study addresses the challenging task of identifying uniquely-shaped shipping containers in high-resolution RGB satellite imagery. The implemented solution utilized a dataset of 109 images containing 1,017 annotated instances. Optimal training hyperparameters were tuned using an iterative process based on a genetic algorithm. The false positive issue was addressed by augmenting the original dataset with unlabeled images. The resulting model achieved an AP (Average Precision) score of 0.86.

Keywords: instance segmentation, satellite imagery, limited training data, dataset, anthropogenic object, hyperparameter optimization, false positive reduction