
Опыт применения нейронной сети YOLOv5 для детектирования растений подсолнечника

Д.А. Полещенко, В.А. Петров, И.С. Михайлов

Старооскольский технологический институт им. А.А. Угарова (филиал) НИТУ «МИСиС». Россия, г. Старый Оскол

Аннотация: В статье описываются результаты исследования возможности детектирования растений подсолнечника по фотографиям, произведенным БПЛА. Решение данной проблемы позволит обеспечить автоматизированный контроль за важным сельскохозяйственным параметром – густотой всходов. Задача усложняется ограниченно малым объемом обучающей выборки и «возмущениями», связанными с засоренностью поля. В результате получено, что нейронная сеть YOLOv5m способна на выборке из 122 картинок качественно детектировать растения с ошибкой обучения 0.077%. Искусственное увеличение выборки до 363 фотографий снижает ошибку обучения до 0.063%. Возмущения снижают эффективность детектирования растений подсолнечника на тестовых изображениях. Повысить эффективность детектирования можно, как за счет добавления оригинальных изображений в обучающую выборку, так и за счет ее искусственного увеличения.

Ключевые слова: детектирование, YOLOv5, подсолнечник, густота всходов, нейронная сеть.

Введение

С развитием цифровизации и бережливого производства, в том числе, в агропромышленном комплексе, появилось такое понятие, как точное земледелие. В данной концепции лежат представления о существовании неоднородностей в пределах одного поля (например, большое количество сорняков, не взрослые и поврежденные культуры).

Важным показателем урожайности является густота растений, который позволяет прогнозировать объем посева к уборке, а, следовательно, оптимизировать и снизить затраты на работу сельскохозяйственной техники [1].

Густота отражает площадь питания посевов. Низкая густота приводит к чрезмерному потреблению питательных веществ посевными культурами, что усложняет уборку урожая. Также происходит увеличение сорной растительности, что требует увеличение затрат на борьбу с ней. При высокой

густоте происходит истощение земли и ухудшается питание растений, что приводит к снижению урожайности. Поэтому оптимизация густоты растений помогает решить данные проблемы.

Традиционные методы оценки раннего насаждения растений основаны на ручных измерениях и визуальном наблюдении, которые требуют много времени, субъективны из-за небольших используемых площадей для отбора проб и неспособны уловить пространственную изменчивость в масштабе поля. С развитием дистанционного зондирования с помощью небольших и легких беспилотных летательных аппаратов (БПЛА) начали появляться эффективные и точные методы оценки урожайности, основанные на методах глубокого обучения. Можно выделить две категории: методы, основанные на сегментации, и методы, основанные на детектировании объектов.

Методы на основе сегментации обычно используют алгоритмы полной сверточной сети (FCN) [2] или U-net [3, 4].

Данные методы возможно применять в решаемой задаче, однако нецелесообразно, по причине того, что при сегментации возможно осуществлять и классификацию растений, но всегда остается вопрос по подсчету числа сегментированных растений на фотографии. Для этого необходимо применять дополнительные алгоритмы, что усложняет задачу и увеличивает время решения.

Более целесообразно при контроле густоты всходов растений использовать методы детектирования объектов и соответствующие структуры нейронных сетей.

Для задач обнаружения растений в работе [5] проведен сравнительный анализ таких алгоритмов, как EfficientDet, Single Shot MultiBox Detector (SSD) и You Only Look Once (YOLOv4), где последний показал наилучшие результаты.

В статье рассмотрено решение задачи детектирования растений на фотографиях поля, произведенных с помощью БПЛА, на примере подсолнечника для последующего их подсчета с целью определения густоты посевов.

При решении подобных задач существует проблема – малая емкость обучающей выборки. Целесообразно использовать модель YOLO, которая ориентируется на обучении сети за один проход данных, что может позволить нивелировать проблему малого объема обучающих данных.

Концепция распознавания объектов YOLO предполагает переход к регрессионной форме задачи [6]. YOLO – одноэтапный алгоритм глубокого обучения, который использует сверточные нейронные сети для обнаружения объектов. На сегодняшний день были разработаны несколько версий алгоритма. Принцип работы YOLOv1 заключается в том, что заданное изображение разделяется на n -ное количество одинаковых фрагментов, в каждом из которых обнаруживается центр ячейки. Каждая ячейка предсказывает фиксированное количество ограничивающих рамок со значением достоверности (confidence score). В свою очередь, ограничивающие рамки базируются на пяти составляющих: координатам по оси абсцисс и ординат, высоте и ширине рамки, а также на значении достоверности. Для выбора наиболее подходящего значения на следующем этапе алгоритм использует IoU — пересечение над объединением. Кроме того, YOLO включает в себя опцию немаксимального подавления (NMS) для удаления лишних рамок. Вторая версия алгоритма дополнена пакетной нормализацией (batch normalization) вместе со сверточными слоями для снижения вероятности переобучения и повышения точности [7]. В YOLOv3, в связи со слабым уровнем обнаружения небольших объектов, была произведена замена магистральной сети извлечения признаков (feature extraction backbone) Darknet19 на Darknet53 [8]. YOLOv4 является точнее и

быстрее предыдущих алгоритмов, поскольку в нее включена новая магистральная сеть извлечения признаков на CSPDarknet53[9]. В пятой версии алгоритма фреймворк PyTorch сменил магистральную сеть Darknet, за счет чего он был значительно облегчен [10]. Добавление нового focus layer, заменившего первые три слоя магистральной сети YOLOv3, в YOLOv5 повысило скорость обнаружения объектов с минимальными потерями в точности [11].

В работе исследуется возможность применения модели YOLOv5 и оценивается эффективность детектирования растений подсолнечника в условиях малой выборки.

Подготовка обучающего набора данных

Для обучения моделей сверточных нейронных сетей требуется большая выборка изображений, чтобы получить достаточно эффективную модель. Для задачи детектирования подсолнечника на начальном этапе в открытом доступе нет подходящих наборов данных, поэтому собран собственный набор.

Над полем с подсолнечником были осуществлены полеты на БПЛА, на высоте 5 метров над землей и сделан ряд фотографий в разрешении 2592x1728. В результате визуального анализа было отобрано 153 фотографии, которые поделили в объеме 122 шт. для обучения, и 31 шт. для валидации. Далее фотографии обрезали до размера 1920x1080, чтобы они могли разместиться в памяти видеокарты при обучении. Также емкость данного набора была искусственно увеличена до 363 фотографий путем аугментации, где использовались такие методы, как поворот на 90 градусов, изменение ориентации, яркости. Примеры изображений представлены далее (см. рис. 1).



Рис. 1. – Изображение с БПЛА

Аннотирование изображений

Так как используемый метод распознавания объектов является разновидностью метода обучения с учителем, необходимо передавать модели информацию о местоположении ограничивающих рамок людей на изображении. Для этого необходимо произвести аннотирование изображений. Для выполнения этой задачи использовали приложение Roboflow, которое позволяет не только размечать изображения, но также дает возможность экспортировать аннотированные наборы данных в различных форматах, что ускоряет процесс разработки. Процесс аннотирования представлен далее (см. рис. 2).

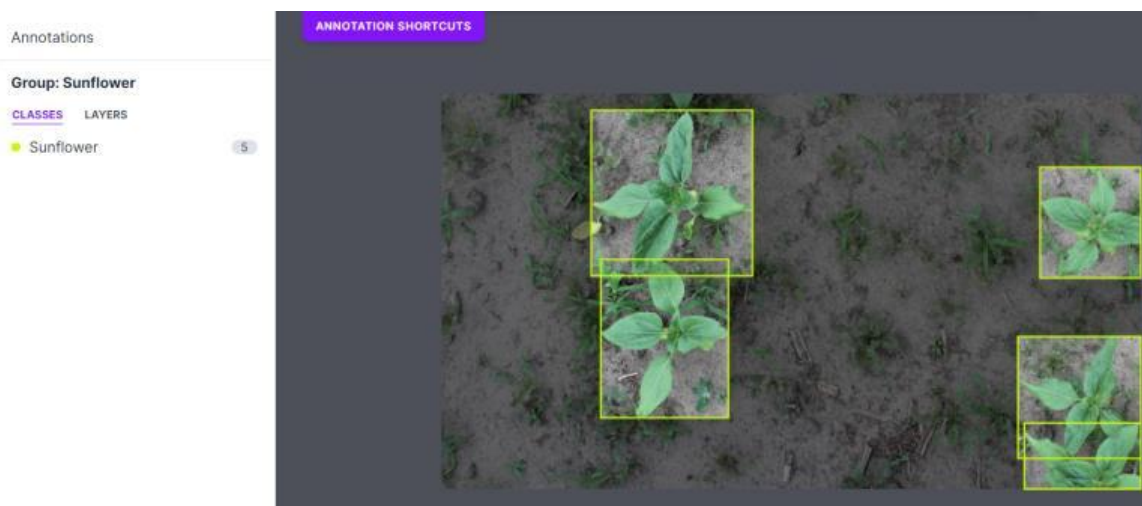


Рис. 2. – Процесс аннотирования изображения

Обучение

Для улучшения результатов обучения были выбраны уже предобученные модели. Библиотекой yolov5 предоставляются 5 различных предобученных на наборе данных COCO моделей, которые отличаются размером, скоростью выполнения и точностью. Были рассмотрены модели YOLOv5x и YOLOv5m, различные по точности и скорости обработки изображений. Данные модели были обучены при различных размерах обучающей выборки при следующих гиперпараметрах:

- Размер пакета – 2.
- Количество эпох – 15 и 30.

Для сравнения результатов обучения применены стандартные метрики, используемые для оценки моделей машинного обучения:

- Точность (Precision) определяет отношение числа корректно обнаруженных объектов ко всему количеству обнаруженных объектов.
- Полнота (Recall) определяет отношение числа корректно обнаруженных объектов ко всем объектам класса.
- Средняя точность (AP) определяется площадью под кривой точности полноты.

Средняя точность является одной из наиболее используемых метрик для определения качества модели обнаружения объектов. При этом средняя точность может включать в себя различные метрики, в зависимости от IoU. Наиболее используемыми является значение AP при IoU не меньше 0.5 и среднее значение AP при пороговых значениях IoU в диапазоне от 0.5 до 0.95 с шагом 0.5, которые обозначаются AP@.5 и AP@.5:.95 соответственно.

Тестирование на валидационном наборе показало следующие результаты (см. таблицу 1, рисунок 3):

Таблица № 1

Результаты валидации моделей на различных наборах данных

Модель	Кол-во эпох	Размер обучающей выборки	Precision, %	Recall, %	AP@.5, %	AP@.5:.95, %
YOLOv5m	15	122	0,944	0,937	0,967	0,76
		363	0,964	0,928	0,976	0,878
	30	122	0,949	0,923	0,961	0,839
		363	0,936	0,937	0,98	0,924
YOLOv5x	15	122	0,927	0,937	0,975	0,798
		363	0,923	0,944	0,963	0,903
	30	122	0,934	0,944	0,971	0,888
		363	0,936	0,914	0,959	0,914

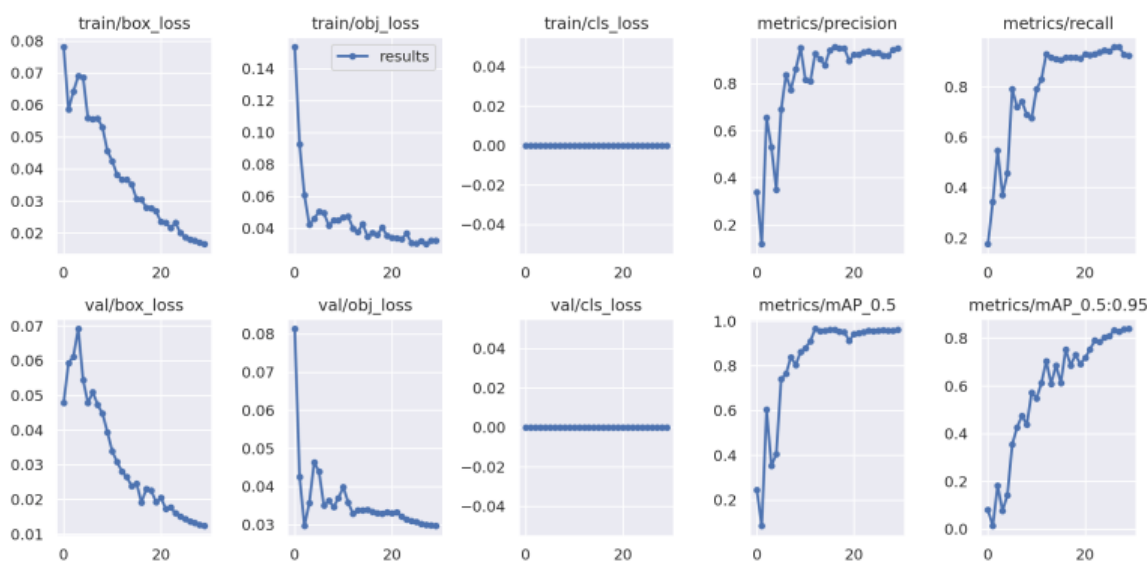


Рис. 3. – Графики обучения модели YOLOv5m на выборке из 363 изображений

Результаты эксперимента

На основании результатов обучения для тестирования была выбрана модель YOLOv5m, обученная при 30 эпохах на наборе данных из 363 картинок, показавшая наилучшие результаты при валидации.

На вход сети была подана картинка размером 1920x1080 из того же набора данных, но не участвовавшие в обучении (см. рисунок 4). Модель показала высокую точность при распознавании, детектировав все растения без ошибок.



Рис. 4. – Результат работы модели YOLOv5m

Для следующего эксперимента были поданы картинки размером 5472x3648 разной фазы роста подсолнечника в условиях с незначительным и большим количеством сорняков (см. рисунки 5-6). На «чистом» участке модель показала следующие показатели качества: precision (0,621%) и recall (0,976%). На засоренном участке увеличивается ошибка детектирования. Поэтому требуется расширять набор данных для обучения изображениями при различных условиях для улучшения качества работы модели.

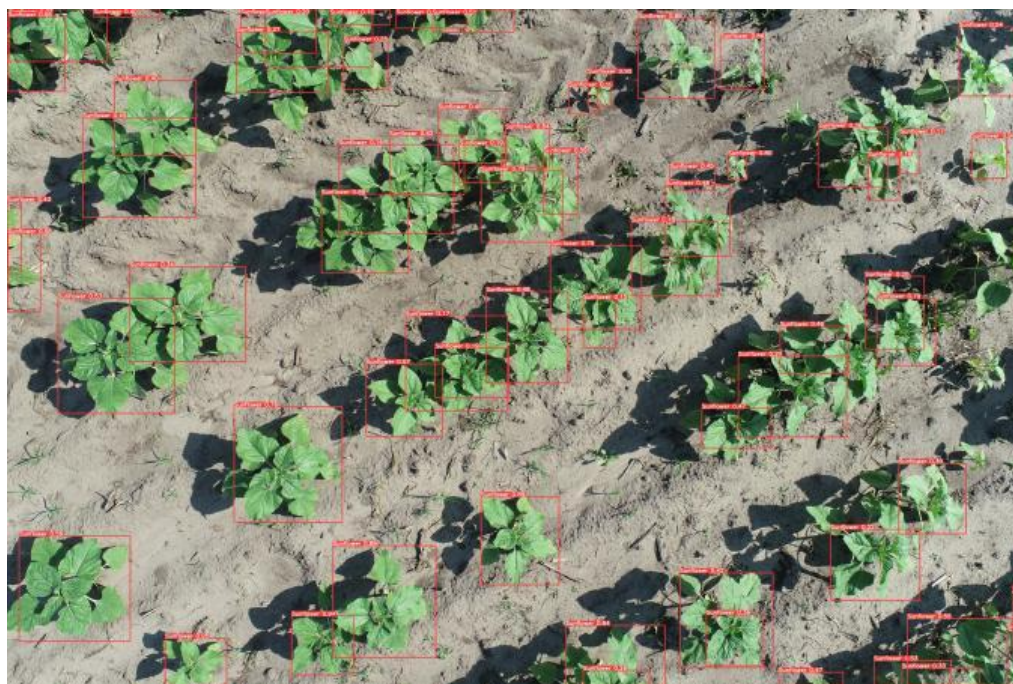


Рис. 5. – Результат работы модели YOLOv5m на более поздней фазе развития растений

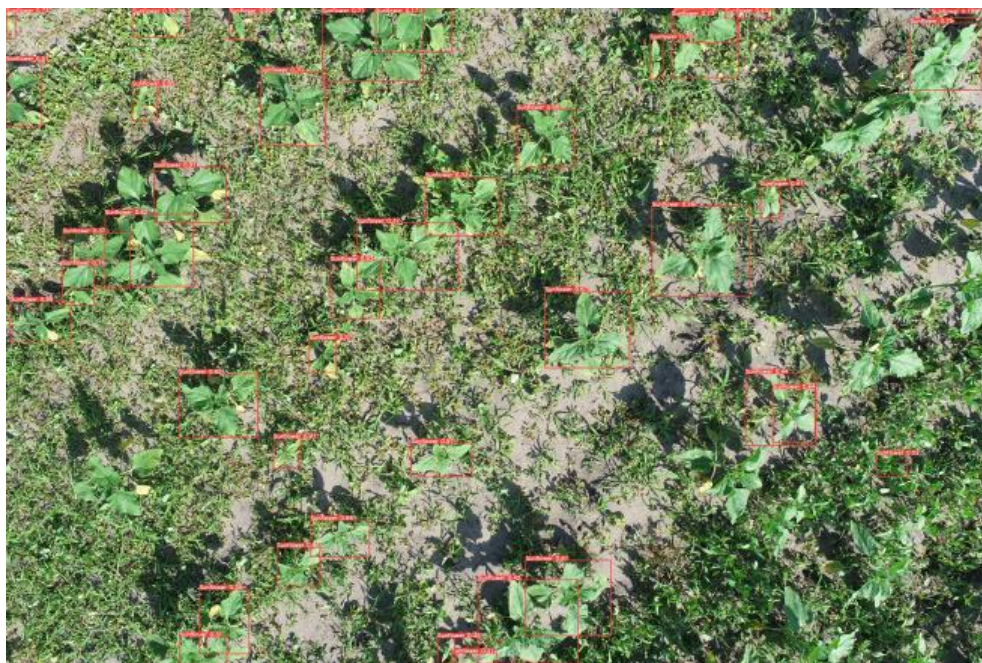


Рис. 6. – Результат работы модели YOLOv5m на зашумленном изображении

Заключение

В результате исследования был собран набор данных с поля подсолнечника и проаннотирован при помощи RoboFlow. Выбранный алгоритм детектирования YOLOv5 был обучен с использованием

подготовленного набора данных. Экспериментальным образом были получены метрики моделей и выбрана наилучшая для последующих опытов. Было выяснено, что YOLOv5 теряет точность на новых наборах данных. Поэтому, расширив набор данных изображениями в различных условиях и фазах роста подсолнечника, можно обучить модель для получения более точных результатов. Таким образом, алгоритм YOLOv5 может быть использован для решения задачи подсчета растений на поле, что требует дальнейшей работы.

Работа выполнена при частичной финансовой поддержке Совета по грантам Президента РФ (проект СП-551.2022.5).

Литература

1. Златопольский А. А., Яровой С. А. Автоматизированный расчёт густоты растений по фотографиям с беспилотного летательного аппарата // Современные проблемы дистанционного зондирования Земли из космоса. – 2020. – Т. 17. – №. 7. – С. 155-166.
2. Shelhamer, E.; Long, J.; Darrell, T. Fully Convolutional Networks for Semantic Segmentation. IEEE Trans. Pattern Anal. Mach. Intell. 2017, 39, Pp. 640–651.
3. Vong, C. N., Conway, L. S., Zhou, J., Kitchen, N. R., & Sudduth, K. A. Early corn stand count of different cropping systems using UAV-imagery and deep learning // Computers and Electronics in Agriculture. – 2021. – Т. 186. – P. 106214.
4. Fawakherji, M., Youssef, A., Bloisi, D. D., Pretto, A., & Nardi, D. Crop and weed classification using pixel-wise segmentation on ground and aerial images //Int. J. Robot. Comput. – 2020. – Т. 2. – №. 1. – Pp. 39-57.
5. Li H., Wang P., Huang C. Comparison of Deep Learning Methods for Detecting and Counting Sorghum Heads in UAV Imagery //Remote Sensing. – 2022. – Т. 14. – №. 13. – Pp. 3143.

6. Redmon, J., Divvala, S., Girshick, R., & Farhadi, A. You Only Look Once: Unified, Real-Time Object Detection // Proceedings of the IEEE conference on computer vision and pattern recognition. – 2016. – P. 10.

7. Redmon, J., Divvala, S., Girshick, R., & Farhadi, A. YOLO9000: Better, Faster, Stronger // Proceedings of the IEEE conference on computer vision and pattern recognition. – 2016. – P. 7263-7271.

8. Redmon, J. and Farhadi, A. YOLOv3: An Incremental Improvement // arxiv.org. – 2018. – URL: arXiv:1804.02767v1 (дата обращения 20.07.2022)

9. Bochkovskiy A., Wang C. Y., Liao H. Y. M. Yolov4: Optimal speed and accuracy of object detection // arxiv.org.– 2020. – URL: arXiv:2004.10934 (дата обращения 20.07.2022)

10. Nelson J., Solawetz J. YOLOv5 is Here: State-of-the-Art Object Detection at 140 FPS // RoboflowBlog. – 2020. – URL: blog.roboflow.com/yolov5-is-here (дата обращения 20.07.2022)

11. Jocher G. YOLOv5 Focus Layer // GitHub. – 2021. – URL: github.com/ultralytics/yolov5/discussions/3181 (дата обращения 20.07.2022)

References

1. Zlatopolsky, A.A. and Yarovoy, S.A. Sovremennye Problemy Distantionnogo Zondirovaniya Zemli iz Kosmosa, 2020. 17(7), pp. 155–166.

2. Shelhamer, E.; Long, J.; Darrell, T. IEEE Trans. Pattern Anal. Mach. Intell. 2017, 39, pp. 640–651.

3. Vong, C. N., Conway, L. S., Zhou, J., Kitchen, N. R., & Sudduth, K. A. Computers and Electronics in Agriculture, 2021. V. 186. 106214 p.

4. Fawakherji, M., Youssef, A., Bloisi, D. D., Pretto, A., & Nardi, D. Int. J. Robot. Comput, 2020. V. 2. №. 1. pp. 39-57

5. Li H., Wang P., Huang C. Remote Sensing, 2022. V. 14. №. 13. 3143 p.

6. Redmon, J., Divvala, S., Girshick, R., & Farhadi, A. Proceedings of the IEEE conference on computer vision and pattern recognition, 2016. 10 p.



7. Redmon, J., Divvala, S., Girshick, R., & Farhadi, A. Proceedings of the IEEE conference on computer vision and pattern recognition, 2016. pp.7263-7271.
8. Redmon, J. and Farhadi, A. arxiv.org, 2018. URL: [arXiv:1804.02767v1](https://arxiv.org/abs/1804.02767v1) (Accessed 20.07.2022).
9. Bochkovskiy A., Wang C. Y., Liao H. Y. M. arxiv.org, 2020. URL: [arXiv:2004.10934](https://arxiv.org/abs/2004.10934) (Accessed 20.07.2022).
10. Nelson J., Solawetz J. RoboflowBlog, 2020. URL: blog.roboflow.com/yolov5-is-here (Accessed 20.07.2022).
11. Jocher G. GitHub, 2021. URL: github.com/ultralytics/yolov5/discussions/3181 (Accessed 20.07.2022).