



ИССЛЕДОВАНИЕ ВОЗМОЖНОСТЕЙ ПРИМЕНЕНИЯ АЛГОРИТМОВ ГЛУБОКОГО ОБУЧЕНИЯ И КОМПЬЮТЕРНОГО ЗРЕНИЯ В ЗАДАЧЕ КЛАССИФИКАЦИИ ГРИБОВ И ЯГОД

Статья содержит эксперименты по исследованию возможностей сверточных нейронных сетей на примере архитектур YOLOv5, YOLOv8, EdgeYOLO применительно к задаче распознавания и классификации грибов и ягод на статических изображениях. Рассмотрен пример обучения моделей компьютерного зрения на наборе данных грибов и ягод. Выполнено сравнение архитектур на отдельном проверочном наборе данных с последующим замером метрик прогресса обучения для сравнения их пригодности к решению поставленной задачи. По результатам экспериментов сделан вывод о возможности применения архитектуры.

Распознавание грибов, распознавание ягод, классификация грибов, классификация ягод.

Нейронные сети в настоящее время широко применяются в тех задачах, где невозможно использовать классические алгоритмы. Существуют задачи, которые очень легко решаются естественным интеллектом человека (при необходимости предварительно обученным и подготовленным), но в то же время очень трудно или практически невозможно разработать формализованный алгоритм, который бы максимально точно моделировал процесс решения этих задач человеком. Одной из этих задач является распознавание и классификация объектов на изображениях. В этом случае на помощь приходят алгоритмы глубокого обучения и компьютерного зрения, переживающие подъем популярности в настоящее время. Они позволяют решить практически любую задачу классификации и распознавания объектов как на статических изображениях, так в видеопотоке данных.

Данная работа посвящена исследованию возможностей применения компьютерного зрения на примере моделей YOLOv5, YOLOv8, EdgeYOLO. Рассмотрим пример обучения этих архитектур распознаванию и классификации грибов и ягод. Данная задача актуальна и имеет прикладной характер. По данным ВЦИОМ, процент походов в лес с 2011 года увеличился с 44 % до 48 %. Сбор грибов по-прежнему популярнее сбора ягод, трав, рыбалки или охоты. В дальнейшем будет выполнена интеграция модели в мобильное приложение для удобства конечных пользователей.

Существуют похожие задачи распознавания и классификации, которые решались с помощью моделей компьютерного зрения YOLOv5, YOLOv8, EdgeYOLO.

В работе «Применение модели YOLOv5 для обнаружения микрообъектов в морской среде» [1] было проведено исследование применимости модели YOLOv5 для распознавания планктона и микропластика в водной среде. Были получены результаты, сравнимые с ручным распознаванием. В частности, точность распознавания (метрика precision) была получена близкая к

100 %. Что говорит о возможности модели YOLOv5 распознавать даже маленькие объекты.

Публикация «Разработка программно-аппаратного комплекса для мониторинга производственной деятельности с использованием нейросети YOLOv8» [2] показывает, что с помощью 8 версии сверточной нейронной сети YOLO можно решать даже такие нетривиальные задачи, как контроль работы цеха с помощью обработки видеопотока с камер на микроконтроллере ESP8266. Это говорит о гибкости архитектуры YOLO, ее способности работать на различных устройствах и применяться для различных задач компьютерного зрения.

В публикации «Детектирование объекта в реальном времени с помощью метода YOLO» [3] авторы отметили высокую скорость работы архитектуры YOLO, а также ее высокую точность, что важно для задач распознавания образов. Также отмечено отсутствие лишних вычислительных затрат, что важно, например, для использования модели на мобильном устройстве, где вычислительные возможности ограничены.

Выбор сети и архитектуры сети для вашей работы. Для решения задачи распознавания и классификации грибов и ягод была выбрана архитектура YOLO в трех версиях: YOLOv5, YOLOv8, а также EdgeYOLO. EdgeYOLO предназначена для работы на устройствах с небольшими вычислительными возможностями и оптимизирована для обработки видеопотока в реальном времени.

Архитектура модели компьютерного зрения (ComputerVision) YOLO использует подход One-Stage Detector. В данной модели сначала предсказываются имеющиеся на входном изображении области, в которых с определенной долей вероятности (показана справа вверху в виде числа до 1, где 0 – 0 %, а 1 – 100 %) осуществляется детекция объекта, затем на последующих этапах обработки осуществляется коррекция найденных областей. Получается многослойная архитектура, характерная для всех моделей глубокого обучения (Deep Learning), представленная на рисунке 1.

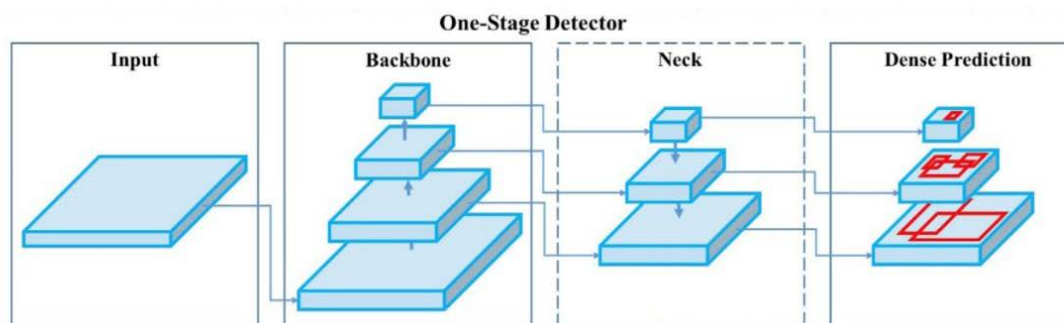


Рис. 1. Архитектура моделей YOLO

Данная модель была выбрана на основании следующих критериев:

1. Высокая популярность и возможность оперативно получить техническую поддержку от сообщества.
2. Открытый исходный код, распространяемый бесплатно под лицензией AGPL-3.0.
3. Наличие популярных инструментов для разметки наборов данных в необходимом для данной модели формате (LabelStudio, Roboflow, CVAT).
4. Простота использования и обучения.
5. Высокая скорость работы архитектуры и наличие в семействе компактных моделей, способных работать на устройствах с ограниченными вычислительными возможностями (смартфоны, планшетные компьютеры, одноплатные компьютеры и т.д.).

Была выбрана самая маленькая по количеству слоев модель YOLO: YOLOv5n6 с входным изображением 1280×1280 px для YOLOv5; YOLOv8n с входным изображением 640×640 px для YOLOv8; EdgeYOLO-Tiny с входным изображением 640×640 px для EdgeYOLO.

Данное решение обусловлено тем, что нейронная сеть будет запускаться на мобильных устройствах, вычислительные возможности которых ограничены.

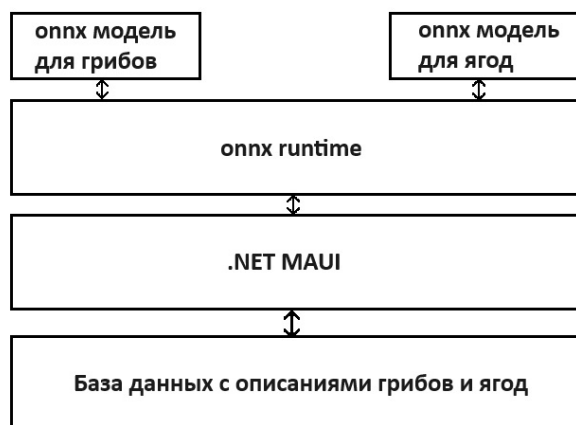


Рис. 2. Архитектура приложения

Архитектура приложения. Место нейронной сети. Для интеграции в приложение модель YOLO была экспортирована в открытый формат ONNX и для запуска ее в приложении используется открытая библиотека-акселератор для запуска моделей ONNX от Microsoft ONNX Runtime. Программа для мобильного устройства написана с использованием фреймворка для создания кросс-платформенных приложе-

ний от Microsoft.NET MAUI на языке C#. Также присутствует база данных с описанием грибов и ягод. Она используется для хранения справочной информации о грибах и ягодах. В приложении присутствуют две отдельные нейронные сети: одна для распознавания и классификации грибов, другая решает аналогичную задачу для ягод. В файле модели хранятся значения параметров (размер входного изображения, количество слоев, веса для входов нейронов и т.д.). Упрощенная архитектура представлена на рисунке 2.

Классы. Модель для классификации грибов обучалась для классов, представленных в таблице 1.

Таблица 1

Классы грибов

0	abotiporus_biennis	6	amanita_regalis
1	agrocycbe_pediades	7	amanita_virosa
2	aleuria_aurantia	8	boletus_edulis
3	alloclavaria_purpurea	9	cyclocybe_erebia
4	amanita_muscaria	10	flammulina_velutipes
5	amanita_phalloides	11	laeticutis_cristata

Модель для классификации ягод была обучена для классов, представленных в таблице 2.

Таблица 2

Классы ягод

0	Black currant	5	Cowberry
1	Blackberry	6	Cranberry
2	Blueberry	7	Raspberries
3	Bog whortleberry	8	Red Ribes
4	Cloudberry	9	Strawberries

Небольшое количество классов было вызвано трудоемкостью ручной разметки данных. Несмотря на их количество, этого достаточно для проведения экспериментов и создания минимального жизнеспособного продукта.

Процесс обучения нейронной сети. Набор данных. Был выполнен поиск готовых размеченных наборов данных для классификации грибов, в частности на портале Kaggle.com. Однако таких наборов данных, в хорошем разрешении обладающих разметкой не было, поэтому разметка осуществлялась вручную в среде label-studio. Набор данных для грибов содержит 219 изображений (включая пустые (null)) для обуче-

ния и 86 (включая пустые) для проверки обучения и вычисления метрик. Набор данных был экспортирован в формат YOLO. Было использовано онлайн-приложение Roboflow. Была выполнена аугментация данных по следующим параметрам: случайный поворот (до 45 градусов), отражение, шум, размытие, перекрытие, за счет этого размер набора данных вырос в 3 раза: 240 изображений для обучения, 30 изображений для тестирования прогресса и вычисления метрик.

Обучение выполнялось в облачном сервисе Google Collab на графическом ускорителе TeslaT4, предобученные модели не использовались, обучение осуществлялось с нуля.

В ходе обучения была получена следующая статистика: для набора данных грибы YOLOv5 показана на рисунке 3; для набора данных грибы YOLOv8 – на рисунке 4; для набора данных ягоды YOLOv8 – на рисунке 5.

Важными метриками являются recall, precision, mAP_0.5, измеряются от 0 до 1, где 0 – 0 %, а 1 – 100 %.

Recall (полнота) – количество всех обнаруженных объектов по отношению ко всем объектам, выделенным разметкой. Учитываются верные и ошибочные обнаружения.

Precision (точность) – отношение верно определенных объектов ко всем найденным объектам.

mAP_0.5 – средняя точность распознавания на основе recall и precision.

Как следует из графика на рисунке 3, средняя точность распознавания смогла превысить 60 %, что является неплохим показателем.

Модель, статистика которой представлена на рисунке 4 показала себя хуже, средняя точность распознавания не смогла пробить порог в 60 %, отчасти это можно объяснить меньшим входным разрешением (матрица 640 против 1280 пикселей).

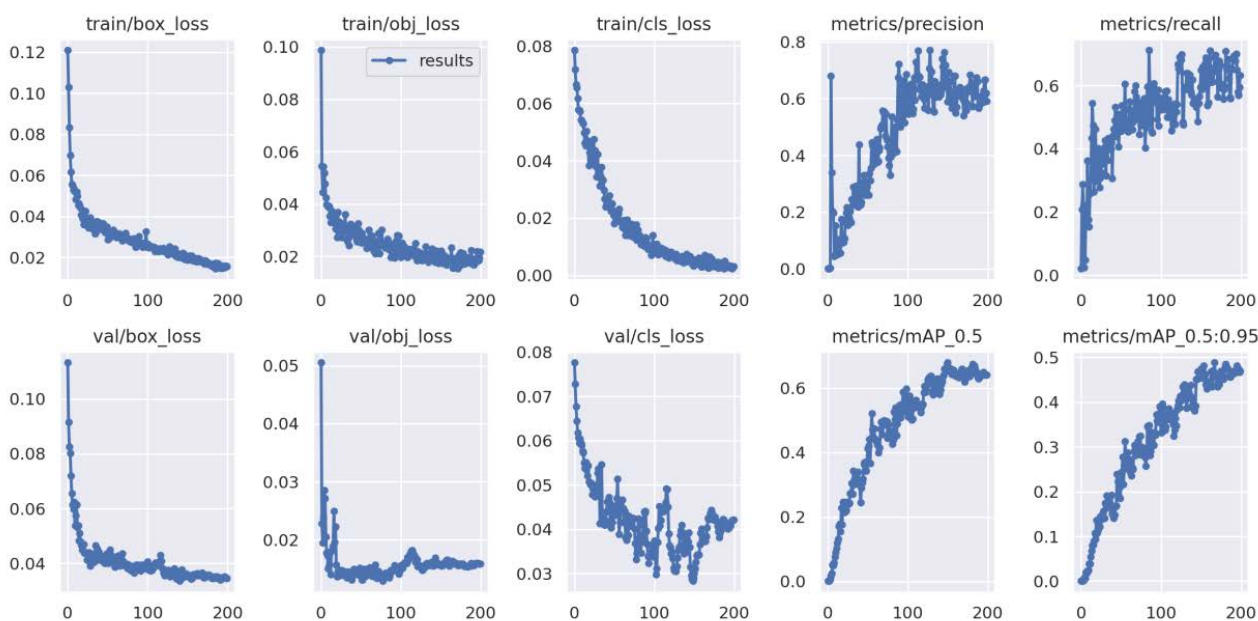


Рис. 3. Статистика поколений YOLOv5

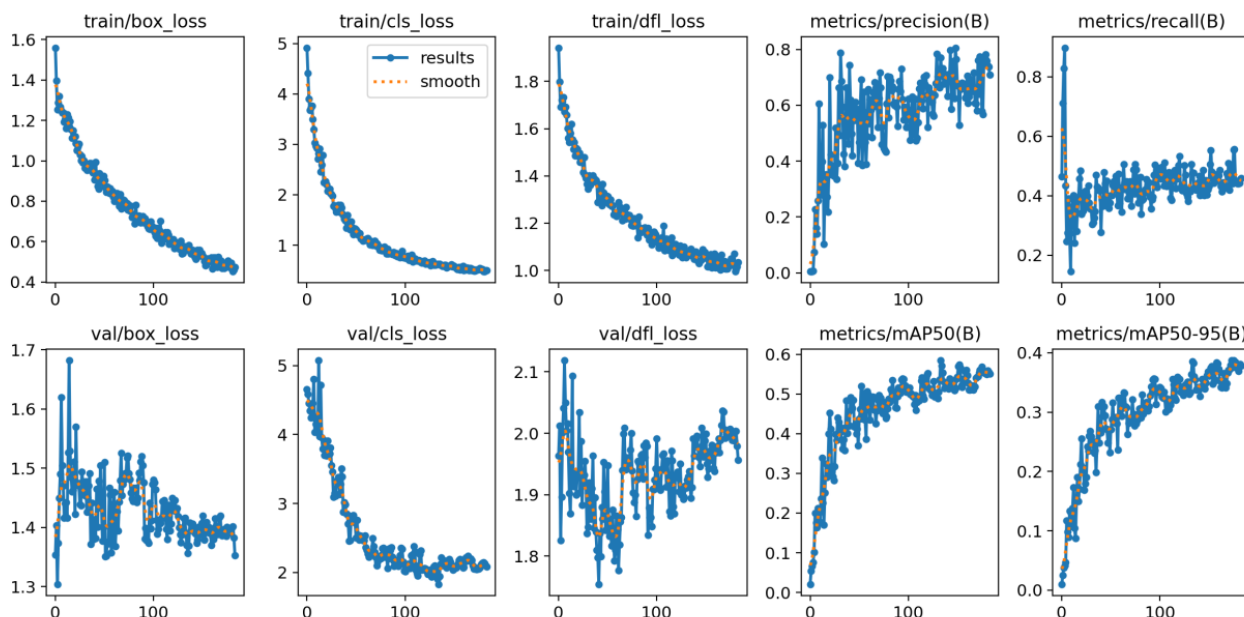


Рис. 4. Статистика поколений YOLOv8

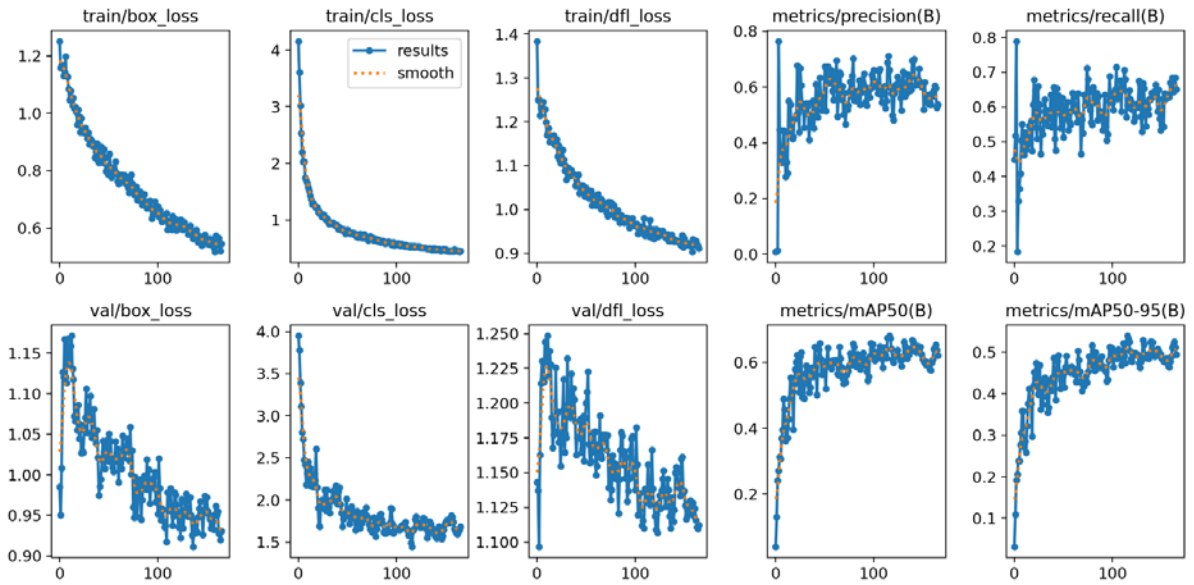


Рис. 5. Статистика поколений YOLOv8 для ягод

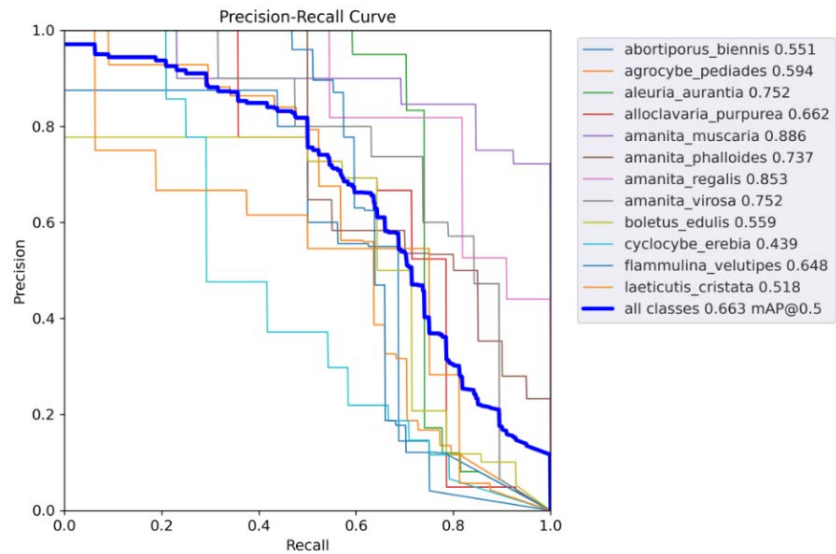


Рис. 6. Precision-Recall для архитектуры YOLOv5

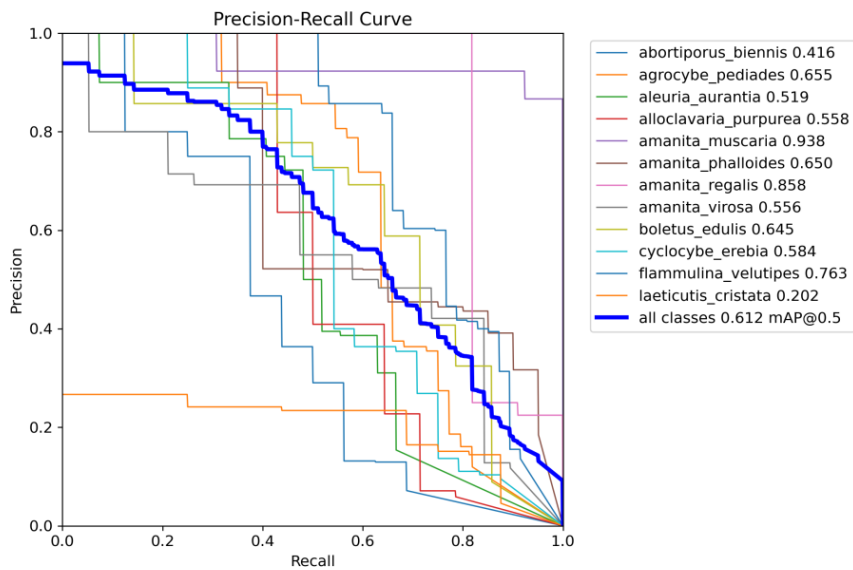


Рис. 7. Precision-Recall для архитектуры YOLOv8

Средняя точность слегка превысила 60 %, что достаточно неплохой показатель для модели, статистика которой представлена на рисунке 5, учитывая высокую скорость работы.

Результаты EdgeYOLO для набора данных «грибы» представлены в таблице 3. В таблице приведены данные по десяти последним поколениям.

Таблица 3

Поколения EdgeYOLO

389:	ap50: 0.6056
ap50_95: 0.4148	390:
ap50: 0.6143	ap50_95: 0.4153
391:	ap50: 0.6133
ap50_95: 0.4162	392:
ap50: 0.6140	ap50_95: 0.4158
393:	ap50: 0.6181
ap50_95: 0.4157	394:
ap50: 0.6234	ap50_95: 0.4231
395:	ap50: 0.6221
ap50_95: 0.4176	396:
ap50: 0.6241	ap50_95: 0.4211
397:	ap50: 0.6237
ap50_95: 0.4198	398:
ap50: 0.6236	ap50_95: 0.4198
399:	ap50: 0.6235
ap50_95: 0.4183	

```

DONE (t=0.65s).
Accumulating evaluation results...
DONE (t=0.28s).
2023-08-14 18:36:32.868 | INFO | __main__:eval_single:158 - ap50 : 0.5918816087360184
2023-08-14 18:36:32.869 | INFO | __main__:eval_single:159 - ap50_95 : 0.39058631769687646
2023-08-14 18:36:32.869 | INFO | __main__:eval_single:160 - Average forward time: 51.92 ms, Average NMS time: 3.85 ms, Average inference time: 55.77 ms
Average Precision (AP) @[ IoU=0.50:0.95 | area= all | maxDets=100 ] = 0.391
Average Precision (AP) @[ IoU=0.50 | area= all | maxDets=100 ] = 0.592
Average Precision (AP) @[ IoU=0.75 | area= all | maxDets=100 ] = 0.419
Average Precision (AP) @[ IoU=0.50:0.95 | area= small | maxDets=100 ] = -1.000
Average Precision (AP) @[ IoU=0.50:0.95 | area=medium | maxDets=100 ] = 0.120
Average Precision (AP) @[ IoU=0.50:0.95 | area= large | maxDets=100 ] = 0.406
Average Recall (AR) @[ IoU=0.50:0.95 | area= all | maxDets= 1 ] = 0.236
Average Recall (AR) @[ IoU=0.50:0.95 | area= all | maxDets= 10 ] = 0.520
Average Recall (AR) @[ IoU=0.50:0.95 | area= all | maxDets=100 ] = 0.567
Average Recall (AR) @[ IoU=0.50:0.95 | area= small | maxDets=100 ] = -1.000
Average Recall (AR) @[ IoU=0.50:0.95 | area=medium | maxDets=100 ] = 0.242
Average Recall (AR) @[ IoU=0.50:0.95 | area= large | maxDets=100 ] = 0.581

```

Рис. 8. mAP0.5 для архитектуры EdgeYOLO



Рис. 9. Классификация гриба

Данная модель показала результат по средней точности распознавания, сопоставимый с YOLOv5. Было выполнено сравнительно тестирование моделей YOLOv5, YOLOv8, EdgeYOLO для набора данных грибы. Для этого был создан отдельный набор данных с теми же классами, модели были запущены на нем в режиме evaluate (проверка и вычисление метрик).

Графики Precision-Recall и метрики mAP0.5 для архитектуры YOLOv5 представлены на рисунке 6.

Графики Precision-Recall и метрики mAP0.5 для архитектуры YOLOv8 представлены на рисунке 7.

Значения mAP0.5 для архитектуры EdgeYOLO представлены на рисунке 8.

Как видим, самый лучший результат показала модель YOLOv5 – 0,663, превзойдя примерно на 5 %, что существенно YOLOv8 с показателем в 0,612. Модель EdgeYOLO показала себя хуже всего с результатом в 0,59.

Для интеграции в приложение были выбраны модели YOLOv5 для классификации грибов и YOLOv8 с аналогичной целью для ягод.

Результат работы сети. На вход нейронной сети подается изображение, на выход мы получаем размеченное изображение (при запуске из консоли, не интегрированной модели в режиме detect). Пример распознавания боровика представлен на рисунке 9.

Пример классификации ягод для интегрированной в мобильное приложение для ОС Android модели с выводом описания представлен на рисунке 10.



Рис. 10. Классификация ягоды

Вывод. Были обучены модели YOLOv5, YOLOv8 и EdgeYOLO для распознавания и классификации грибов. А также YOLOv8 для аналогичной задачи в отношении ягод. Лучше всего показала себя модель YOLOv5 за счет большего входного изображения и со-

ответственно большего количества признаков, которые можно было извлечь из изображения. YOLO-модель подтвердила свой универсальный характер в задачах компьютерного зрения, в частности в задаче классификации. Она может быть применена для обнаружения грибов и ягод на статических изображениях, делает это с достаточно высокой скоростью. Небольшая по размерам модель YOLO Nano может успешно работать на мобильных устройствах, что было проверено путем ее интеграции в мобильное приложение и запуска на смартфоне. Как итог – достаточная точность классификации, высокая скорость работы и возможность запуска отдельных моделей на устройствах с ограниченными вычислительными ресурсами делают ее оптимальным решением применительно к задачам компьютерного зрения.

Литература

1. Применение модели YOLOV5 для обнаружения микрообъектов в морской среде / А. Н. Греков, Ю. Е. Шишкин, С. С. Пелюшенко, А. С. Маврин // Системы контроля окружающей среды. – 2022. – 4 (50). – С. 112–122.
2. Преснецов, А. М. Разработка программно-аппаратного комплекса для мониторинга производственной деятельности с использованием нейросети YOLOv8 / А. М. Преснецов, А. П. Тюрин // Интеллектуальные системы в производстве. – 2023. – Т. 21, № 2. – URL: <https://izdat.istu.ru/index.php/ISM/article/view/5659> (дата обращения: 30.11.2023). – Текст : электронный.
3. Пахомова, О. А. Детектирование объекта в реальном времени с помощью метода YOLO / О. А. Пахомова // Информационные технологии в экономике и управлении : Материалы III Всероссийской научно-практической конференции (Махачкала, 29–30 ноября 2018 г.). – Махачкала : Дагестанский государственный технический университет, 2018. – С. 86–88. – EDN YYQYEX.

G.S. Tikhonovsky, D.V. Kochkin
Vologda State University

RESEARCH ON POSSIBILITIES OF USING DEEP LEARNING AND COMPUTER VISION ALGORITHMS IN CLASSIFYING MUSHROOMS AND BERRIES

This article contains experiments to study the capabilities of convolutional neural networks using the example of the YOLOv5, YOLOv8, EdgeYOLO architectures in relation to the task of recognizing and classifying mushrooms and berries in static images. The example of training computer vision models using a data set of mushrooms and berries is considered. The comparison of these architectures was performed on a separate validation data set followed by measurement of learning progress metrics to compare their suitability for the task. Based on the results of the experiments, it was concluded that this architecture can be used for this task.

YOLOv5, YOLOv8, EdgeYOLO, mushroom recognition, berry recognition, mushroom classification, berry classification.