

<https://doi.org/10.21603/2074-9414-2026-3-2657>
<https://elibrary.ru/ALMTCW>

Оригинальная статья
<https://fptt.ru>

Определение товарного сорта ягод вишни с использованием сверточных нейронных сетей



А. М. Попов^{ORCID}, А. В. Шафрай*^{ORCID}, В. С. Косинов^{ORCID},
Г. И. Подберезен^{ORCID}, Д. В. Сухоруков^{ORCID}, С. С. Комаров^{ORCID}

Кемеровский государственный университет^{ORCID}, Кемерово, Россия

Поступила в редакцию: 15.06.2026

Принята после рецензирования: 24.08.2026

Принята к публикации: 01.09.2026

©

А. М. Попов, А. В. Шафрай, В. С. Косинов, Г. И. Подберезен,
Д. В. Сухоруков, С. С. Комаров, 2026

*e-mail: shafraia@mail.ru



Аннотация.

Ручной контроль качества плодово-ягодного сырья на этапе приемки остается одной из ключевых задач пищевой промышленности. Существующие методы визуальной оценки субъективны, трудоемки и зависят от квалификации оператора, что ограничивает их эффективность в условиях современного производства. Цель исследования – применить модели сверточных нейронных сетей семейства YOLO для классификации ягод вишни по товарным сортам на этапе приемки сырья.

Объект исследования – ягоды вишни культурных сортов (*Prunus subg. Cerasus*), собранные в Кемеровской области в 2024 г. (Россия). Исследование выполнено на базе ООО «НПО Здоровое питание» (Кемерово, Россия). Собраны и размечены изображения четырех категорий качества. Проведено сравнение пяти моделей YOLO (YOLOv8m, YOLOv9m, YOLO11m, YOLO12m, YOLO26m) и трех вариантов архитектуры YOLO12 по размеру. Для выбранной модели YOLO12m оптимизированы гиперпараметры: сравнивались восемь оптимизаторов, четыре значения размера батча и пять значений коэффициента регуляризации Dropout.

Установлено, что наилучшую точность показала модель среднего размера YOLO12. Среди оптимизаторов наивысшие метрики зафиксированы для SGD. Определен оптимальный размер обучающего пакета; изменение параметра Dropout не оказало заметного влияния на результаты. Увеличение числа эпох финального обучения повысило качество распознавания по всем четырем категориям. Работоспособность модели подтверждена на производственной выборке. На основе модели разработан прототип программы с графическим интерфейсом для загрузки изображений, визуализации детекции и формирования отчета. Результаты подтверждают пригодность моделей YOLO для автоматизации входного контроля качества ягод вишни, обеспечивая объективный и воспроизводимый анализ взамен субъективной ручной оценки, и создают основу для внедрения промышленных систем контроля качества.

Ключевые слова. Искусственный интеллект, нейронные сети, YOLO, пищевая промышленность, переработка сырья, ягоды вишни

Финансирование. Исследование выполнено за счет гранта Российского научного фонда № 25-26-00136, <https://rscf.ru/project/25-26-00136/>

Для цитирования: Попов А. М., Шафрай А. В., Косинов В. С., Подберезен Г. И., Сухоруков Д. В. и др. Определение товарного сорта ягод вишни с использованием сверточных нейронных сетей. Техника и технология пищевых производств. 2026. Т. 56. № 3. С. 516–525. <https://doi.org/10.21603/2074-9414-2026-3-2657>

Convolutional Neural Networks in Commercial Grading of Cherry



Anatoly M. Popov^{ID}, Anton V. Shafrai*^{ID}, Vitaly S. Kosinov^{ID},
Grigory I. Podberezen^{ID}, Dmitry V. Sukhorukov^{ID}, Sergey S. Komarov^{ID}

Kemerovo State University^{ROR}, Kemerovo, Russia

Received: 15.06.2026
Revised: 24.08.2026
Accepted: 01.09.2026

*e-mail: shafraia@mail.ru

© A.M. Popov, A.V. Shafrai, V.S. Kosinov, G.I. Podberezen,
D.V. Sukhorukov, S.S. Komarov, 2026



Abstract.

Manual quality control of fruit and berries remains a major challenge in the food industry. Existing visual assessment methods are subjective, labor-intensive, and dependent on operator skills, which limits their effectiveness in modern production environments. However, convolutional neural network models of the YOLO type can be used to classify cherries by commercial grade at the raw material acceptance stage.

The study focused on cultivated cherries (*Prunus* subg. *Cerasus*) harvested in the Kemerovo Region (Russia) in 2024. The research was conducted at the Zdorovye Pitaniye Research and Production Association (Kemerovo, Russia). The comparative analysis evaluated images across four quality categories using five YOLO models (YOLOv8m, YOLOv9m, YOLO11m, YOLO12m, YOLO26m) and three YOLO12 architectural variants of different sizes. Hyperparameter optimization was performed for the selected YOLO12m model, assessing eight optimizers, four batch sizes, and five Dropout regularization coefficient values.

The medium-sized YOLO12 model demonstrated the highest accuracy. Among the optimizers, SGD yielded the best performance metrics. The optimal training batch size was determined, while adjusting the Dropout parameter had no significant effect on the results. Increasing the number of final training epochs improved recognition performance across all four categories. The model's performance was validated on an industrial sample. Based on this model, a prototype software application with a graphical user interface (GUI) was developed for image loading, detection visualization, and report generation.

The results confirm the suitability of YOLO models for automating incoming cherry quality control, providing objective and reproducible analysis instead of subjective manual assessment, thereby paving the way for advanced industrial quality control systems.

Keywords. Artificial intelligence, neural networks, YOLO, food industry, raw material processing, cherry

Funding. The study was supported by the Russian Science Foundation, Grant No. 25-26-00136, <https://rscf.ru/en/project/25-26-00136/>

For citation: Popov AM, Shafrai AV, Kosinov VS, Podberezen GI, Sukhorukov DV, *et al.* Convolutional Neural Networks in Commercial Grading of Cherry. Food Processing: Techniques and Technology. 2026;56(3):516–525. (In Russ.) <https://doi.org/10.21603/2074-9414-2026-3-2657>

Введение

Применение искусственного интеллекта в пищевой инженерии расширяется с каждым годом. Значительная часть исследований посвящена оценке безопасности, качества и срока годности пищевых продуктов с привлечением методов машинного обучения. Среди наиболее распространенных инструментов фигурируют классические алгоритмы машинного обучения и нейронные сети. Из нейросетевых архитектур наиболее востребованы полносвязные, рекуррентные и сверточные нейронные сети, в т. ч. в рамках глубокого обучения. Данные методы охватывают широкий круг задач: от прогнозирования технологических параметров до визуальной оценки продукции.

В пищевой промышленности и сельском хозяйстве прочно вошли в практику механизация и автоматизация

производственных операций. В последние десятилетия наблюдается устойчивый тренд к их роботизации и цифровизации. Даже операции, традиционно выполнявшиеся человеком (например органолептическая оценка сырья), все чаще уступают место цифровым методам с включением искусственного интеллекта. Для этой цели используются электронные носы, электронные языки и системы компьютерного зрения [1, 2].

Е. Yavuzer *et al.* [3] разработали метод определения свежести фаршированных мидий на основе электронного носа (платформа Arduino) в сочетании с нейросетями ResNet и GoogLeNet [3]. D. R. Wijaya *et al.* [4] оценили свежесть морепродуктов с помощью электронного носа и алгоритмов машинного обучения, достигнув точности, близкой к 100 %. I. Zlobin *et al.* [5] предложили одноразовые электронные языки на основе

металлоорганических каркасов для классификации сортов чая с применением нейронных сетей. В другом исследовании электронный нос выступал в качестве инструмента для мониторинга порчи мяса, рыбы и молока [6]. Также предложен метод оценки микробиологических показателей молока с помощью массива газовых сенсоров на основе пьезокварцевых микровесов [7]. Отдельное направление посвящено комбинированию электронного носа и компьютерного зрения для определения степени готовности жареной курицы [8].

Спектроскопические и мультиспектральные методы в сочетании с машинным обучением также заняли значимое место в пищевой промышленности. X. Wu et al. [9] осуществили спектроскопический анализ совместно со сверточными сетями для выявления фальсификации меда с точностью классификации более 97 %. M. G. Lanjewar et al. [10] задействовали методы спектроскопии для обнаружения примесей крахмала в куркуме, где алгоритм Random Forest достиг коэффициента детерминации, равного 0,999. Гиперспектральная визуализация в ближнем инфракрасном диапазоне применялась для количественного определения красителя метанилового желтого в муке из нута с включением сверточных нейронных сетей [11]. R. Belugina et al. [12] предложили метод оценки концентрации фосфолипидов в растительных маслах на основе потенциометрических мультисенсорных систем с помощью метода опорных векторов и наименьших квадратов.

Компьютерное зрение представляет собой эффективный инструмент визуальной оценки качества пищевого сырья, позволяющий объективно анализировать внешний вид продукта независимо от квалификации оператора. C. N. Sánchez et al. [13] провели многомерный анализ изменения цвета говядины с помощью системы компьютерного зрения и методов машинного обучения для определения свежести продукта. В исследовании [14] методы машинного обучения использованы для прогнозирования свежести черники в условиях холодильной цепи по параметрам газовой среды хранения с точностью до 94 %. В обзорных зарубежных публикациях подчеркивается, что сочетание электронных носов с алгоритмами машинного обучения представляет собой мощный неразрушающий инструмент оценки качества пищевых продуктов в таких категориях, как мясо, молочная продукция, рыба, чай и кофе [15].

Вишня – распространенное сырье для быстрорастворимых гранулированных продуктов: морсов, киселей, сухих завтраков. Ее ягоды богаты органическими кислотами, сахарами, антоцианами и витаминами, которые определяют вкусовые и питательные свойства готовой продукции. Качество последней напрямую зависит от кондиции исходного сырья: некондиционные ягоды, удвоенные плоды и экземпляры с отростками негативно влияют на переработку, нарушая гранулирование и снижая однородность продукта.

Традиционный ручной контроль на приемке трудоемок и субъективен. Применение сверточных нейронных сетей позволяет автоматизировать оценку: сети извлекают визуальные признаки из изображений и относят каждую ягоду к соответствующему классу качества. Такой подход обеспечивает объективность и воспроизводимость результата независимо от квалификации оператора. Цель исследования – применить модели сверточных нейронных сетей семейства YOLO для классификации ягод вишни по товарным сортам на этапе приемки сырья.

Для достижения поставленной цели решались следующие задачи:

- подготовить набор данных, содержащий изображения ягод вишни;
- обучить модели семейства YOLO на сформированном наборе данных;
- сравнить различные архитектуры и гиперпараметры;
- разработать прототип программного продукта на основе наиболее точной модели.

Новизна работы заключается в использовании актуальных моделей семейства YOLO для классификации ягод вишни по четырем категориям качества на этапе приемки сырья. В научной литературе задача классификации ягод вишни по товарным сортам в условиях реального производства освещена фрагментарно. Полученные результаты могут служить основой для разработки промышленных систем автоматического контроля качества плодово-ягодного сырья.

Объекты и методы исследования

Объект исследования – ягоды вишни культурных сортов (*Prunus subg. Cerasus*), собранные в Кемеровской области в 2024 г. (Россия). Работа выполнена на базе ООО «НПО Здоровое питание» (Кемерово, Россия). Предмет исследования – методы классификации ягод вишни по товарным сортам на этапе приемки сырья с применением сверточных нейронных сетей семейства YOLO. Алгоритм использования нейронных сетей включал следующие этапы (рис. 1) [1]:

1. Формирование и подготовка датасета. На этом этапе создается размеченный набор изображений, репрезентирующий все товарные сорта вишни и типичные дефекты сырья.

1.1. Сбор и предварительная обработка изображений. Исходные снимки получены с помощью камеры Nikon D5300 kit 18–140 mm (Япония) в варьируемых условиях освещения и при разных ракурсах – для обеспечения инвариантности модели к внешним факторам. Поскольку сверточные архитектуры требуют фиксированного входного размера, все кадры приводятся к единому разрешению, диктуемому выбранной сетью.

1.2. Аннотирование. Каждому изображению присваивается метка класса (сорт или категория качества), которая служит целевой переменной в задаче классификации. Разметка выполнена в открытом инструменте Computer Vision Annotation Tool.

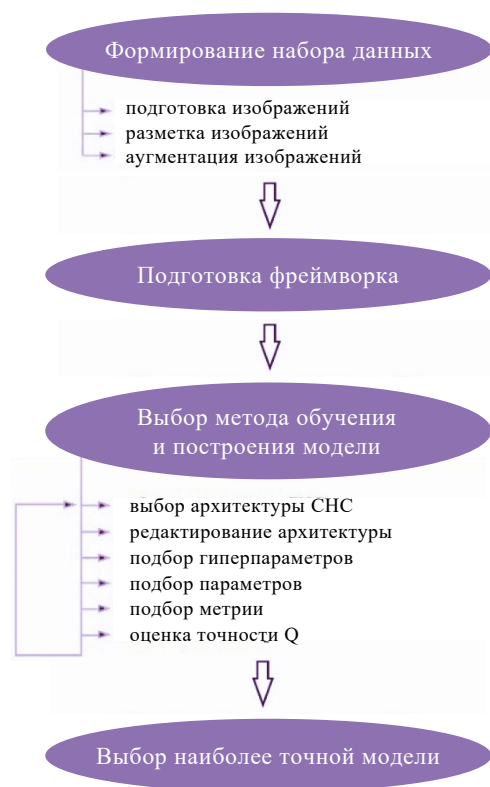


Рисунок 1. Принципиальная схема применения нейронных сетей

Figure 1. Neural network application diagram

1.3. Аугментация. При дефиците данных применяется их искусственное расширение за счет геометрических (повороты, отражения) и фотометрических (изменение яркости, контраста, добавление шума) преобразований. Это повышает разнообразие обучающей выборки и устойчивость сети к вариативности внешнего вида объектов, что в итоге улучшает обобщение.

2. Подключение библиотек. Загружаются программные пакеты для обработки изображений, управления данными и построения нейросетевых моделей.

3. Проектирование и обучение модели. Этап включает итерационный подбор архитектуры, гиперпараметров и стратегии обучения для достижения максимума целевой метрики.

3.1. Выбор базовой архитектуры. Используется подход переноса обучения (transfer learning) [14]: за основу берется сверточная сеть, предобученная на крупном наборе данных (например ImageNet), из семейств ResNet, EfficientNet, MobileNet и др. Внутри одного семейства перебираются варианты разной глубины (ResNet18, ResNet50 и т. п.) с учетом компромисса «точность – вычислительные затраты», критичного при потенциальном внедрении в производственные условия с ограниченными ресурсами.

3.2. Адаптация под задачу. Выходной слой предобученной сети (рассчитанный на 1000 классов ImageNet)

заменяется на полносвязный слой с числом нейронов, равным количеству классифицируемых категорий.

3.3. Настройка гиперпараметров. Определяются ключевые гиперпараметры: функция потерь, оптимизатор (Adam, SGD и др.), методы регуляризации (дропаут, L2-регуляризация), а также вспомогательные техники, такие как batch normalization, способствующие стабильности и скорости обучения [15]. Следует отметить, что эффекты отдельных параметров регуляризации могут различаться в зависимости от решаемой задачи и особенностей реализации в используемых фреймворках.

3.4. Параметры обучения. Фиксируются размер батча, число эпох, пропорция разбиения выборки (обучение/валидация/тест – например 80/10/10), а также возможность применения ранней остановки (early stopping) или динамического изменения скорости обучения (learning rate scheduling).

3.5. Оценка и выбор наилучшей модели. Качество классификации оценивается по метрикам: точность (precision), полнота (recall), F1-мера, матрица ошибок [16]. Поиск оптимальной конфигурации (архитектура + гиперпараметры) выполняется итеративно, в т. ч. с привлечением автоматизированных средств, и завершается при достижении требуемой точности и обобщающей способности.

Качество сырья, поступающего на переработку, регламентируется техническими условиями и государственными стандартами. Для гранулированных продуктов, в состав которых входит вишня (кисели, сухие напитки, морсы, сухие завтраки и др.), действует ГОСТ 33801-2016 «Вишня и черешня свежие. Технические условия». Стандарт устанавливает три товарных сорта: высший, первый и второй. Критерии сортового деления – доля здоровых, свежих ягод с однородной окраской, отсутствие дефектов формы и окраски, повреждений насекомыми-вредителями, а также избыточной внешней влажности.

Сорт партии сырья определяют по соотношению кондиционных (соответствующих требованиям сорта) и некондиционных ягод. На практике такую оценку проводят визуально по малой выборке, что делает процесс субъективным и трудоемким. При этом задача хорошо ложится на методы компьютерного зрения, в первую очередь на сверточные нейронные сети.

Съемка проведена на производственной площадке ООО «НПО Здоровое Питание» (Кемерово, Россия) камерой, установленной на этапе приемки сырья. Изображения получены при варьируемых освещении и ракурсах – для обеспечения инвариантности модели к внешним условиям. Кадры фиксируют как одиночные, так и групповые размещения ягод, отличающихся по размеру, форме и кондиции, что повышает устойчивость сети к вариативности объекта и сложным случаям распознавания. Собранные снимки сформировали структурированный датасет для последующей разметки и обучения модели.

Для получения целевых значений, требуемых для обучения нейронных сетей детектированию ягод, полученные снимки размечали: каждой ягоде на изображении присваивали метку. Разметка выполнена в программе для аннотирования изображений и видео Computer Vision Annotation Tool. В соответствии с технологической пригодностью сформировано четыре класса (рис. 2): класс 0 – кондиционная ягода; класс 1 – удвоенная ягода; класс 2 – ягода с отростком; класс 3 – некондиционная ягода. Разметка сохранена в формате YOLO, поддерживающем большинство нейросетевых фреймворков.

Нейронные сети – это вычислительные модели, состоящие из взаимосвязанных узлов (нейронов) [16], обучающиеся на данных: в процессе обучения веса связей корректируются так, чтобы минимизировать ошибку предсказания, за счет чего сеть последовательно повышает точность решений. Для обработки изображений применяются сверточные нейронные сети (СНС), учитывающие не только значения пикселей, но и их пространственное расположение – именно оно несет смысл визуальной информации. Существует множество архитектур, каждая из которых ориентирована на свой класс задач.

В рамках исследования необходимо обнаружить ягуду вишни на изображении, локализовать ее и классифицировать по категории качества. Это задача детекции объектов, которая требует одновременного прогнозирования координат ограничивающей рамки и принадлежности объекта к классу. Для ее решения чаще всего применяют предобученные СНС, среди которых выделяют семейства ResNet, EfficientNet, MobileNet и др. В настоящей работе выбран подход на основе архитектуры YOLO [17] – одного из наиболее распространенных решений для детекции изображений.

Ключевое отличие YOLO от других детекторов – изображение проходит через нейронную сеть ровно один раз. Принцип работы: изображение разбивается на сетку ячеек, каждая из которых самостоятельно предсказывает ограничивающие рамки, оценки уверенности и вероятности классов. Параллельная обработка ячеек обеспечивает высокое быстродействие, что критично для производственных задач с непрерывным потоком кадров.

Архитектура YOLO включает три последовательных блока: backbone извлекает признаки на разных масштабах; neck агрегирует их для устойчивого распознавания объектов различного размера; head по полученным признакам формирует итоговые предсказания.

За время развития семейство YOLO сменило несколько поколений. YOLOv8 остается одной из самых распространенных версий. В YOLOv9 реализован способ борьбы с потерями информации в глубоких сетях. В YOLO11 пересмотрена структура светочных блоков для улучшения баланса скорости и точности. В YOLO12 внедрены механизмы внимания, учитывающие глобальный контекст изображения и существенно повышающие точность детекции. В актуальной версии YOLO26 архитектура упрощена в пользу скорости обработки.

Обучение модели YOLO начинается с загрузки предобученных весов – этот подход (transfer learning) позволяет использовать уже выделенные визуальные признаки, что сокращает потребность в данных и времени при дообучении на целевом датасете. Далее начинается подбор гиперпараметров. Этот процесс носит итерационный характер и продолжается до тех пор, пока модель не достигает приемлемых значений.

Функция потерь YOLO состоит из трех компонентов: box loss (точность локализации рамок), cls loss (правильность классификации) и distribution focal loss (точность границ рамки через моделирование распределения вероятностей). Итоговая функция потерь – взвешенная сумма указанных компонент.

Аугментация данных – это важный шаг перед началом обучения. Суть метода состоит в применении различных преобразований к изображениям и их аннотациям. Особую роль аугментация играет тогда, когда датасет невелик и собрать дополнительные данные затруднительно. Преобразования искусственно увеличивают число обучающих примеров и повышают их разнообразие, тем самым снижая риск переобучения.

Качество моделей оценивали по стандартным метрикам детекции [18]. Precision – доля верных предсказаний среди всех предсказанных объектов; высокое значение указывает на редкие ложные срабатывания. Recall – доля найденных объектов среди всех реальных; высокое значение говорит о малом числе пропусков.

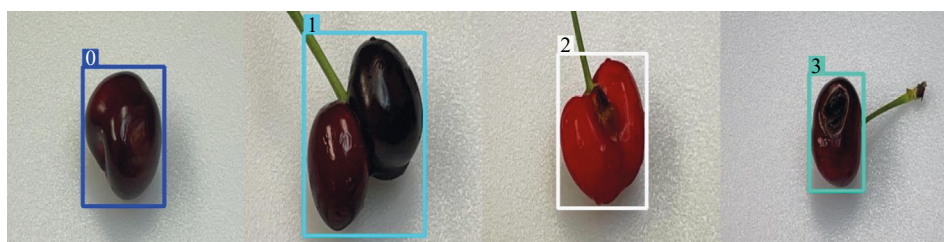


Рисунок 2. Изображение ягод вишни соответствующих классов

Figure 2. Cherries by grade

Обе метрики зависят от IoU – отношения площади пересечения предсказанной и истинной рамок к площади их объединения. mAP50 – средняя точность при пороге $\text{IoU} \geq 0,50$ (предсказание считают верным при перекрытии не менее 50 %). mAP50-95 – среднее значение mAP по порогам IoU от 0,50 до 0,95 с шагом 0,05; эта метрика строже и отражает стабильность детекции при различных требованиях к точности локализации.

Обучение моделей выполнено на следующем программно-аппаратном комплексе: ОС Windows 10 (версия 10.0.19041), интерпретатор Python 3.8 (выбран из-за совместимости с библиотеками машинного обучения), облачная среда Google Colab с доступом к графическому процессору NVIDIA T4, хранилищем данных и поддержкой Jupyter Notebooks. Использование GPU существенно ускоряет матричные вычисления, лежащие в основе обучения нейронных сетей. Выбранный комплекс позволил эффективно реализовать полный цикл обучения и тестирования нейросетевых моделей.

Результаты и их обсуждение

На этапе приемки сырья получено 530 собственных изображений (1024×1024 пикселя), чего оказалось недостаточно для полноценного обучения: малый объем выборки повышает риск переобучения. Для расширения датасета привлечены открытые изображения вишни с платформы Roboflow Universe. Собственные и сторонние изображения перемешаны случайным образом и размечены по целевым классам. Итоговый датасет составил более 1546 изображений с аннотациями (рис. 3).

Датасет разбивали на три выборки: обучающую (80 %), валидационную (10 %) и тестовую (10 %).

Обучающая служит для обновления весов, валидационная – для контроля качества и ранней остановки, тестовая – для финальной оценки на данных, не участвовавших в обучении.

Для повышения устойчивости модели к вариациям освещения, масштаба и ориентации объектов, характерным для производства, применены следующие преобразования:

1. Горизонтальное отражение с вероятностью 0,5.
2. Случайное изменение оттенка на 0,015.
3. Изменение насыщенности на 0,7.
4. Изменение яркости на 0,4.
5. Масштабирование на 0,5.
6. Мозаичная аугментация, при которой четыре изображения объединяются в одно.

Проведено сравнение пяти моделей YOLO (суффикс «m» указывает на средний размер архитектуры) в идентичных условиях: 50 эпох; размер батча 16; начальная скорость обучения 0,01; оптимизатор SGD с инерцией 0,937. Веса функции потерь: box loss – 7,5; cls loss – 0,5; distribution focal loss – 1,5. Итоговые метрики детекции по всем классам для каждой модели на тестовой выборке представлены в таблице 1.

Наилучшие результаты продемонстрировала модель YOLO12m; YOLO26m – наихудшие. Вероятно, это связано с тем, что YOLO26 ориентирована на упрощение архитектуры ради скорости, тогда как в YOLO12 сделан упор на механизмы внимания, повышающие точность.

Далее проведено сравнение версий YOLO12, различающихся количеством параметров и вычислительной сложностью (суффиксы: «n» – нано-, «s» – малый, «m» – средний размер). Обучение выполнено в идентичных условиях: 20 эпох; размер батча 16; оптимизатор SGD;



Рисунок 3. Примеры изображений из датасета

Figure 3. Examples of dataset images

Таблица 1. Сравнение метрик обученных моделей YOLO

Table 1. Metrics of trained YOLO models

Модель	Precision	Recall	mAP50	mAP50-95
YOLOv8m	0,9256	0,8926	0,9543	0,8344
YOLOv9m	0,9211	0,9056	0,9569	0,8395
YOLO11m	0,9116	0,9119	0,9536	0,8390
YOLO12m	0,9141	0,9042	0,9563	0,8430
YOLO26m	0,8946	0,8865	0,9402	0,8253

начальная скорость 0,01; конечная – 0,0001 (косинусный планировщик); параметры аугментации – по умолчанию. Результаты на тестовой выборке приведены в таблице 2.

Увеличение размера модели повышает точность: YOLO12m показала лучшие результаты среди версий, YOLO12n – наихудшие. Рост числа параметров улучшает распознавание, но требует больших вычислительных затрат.

Затем на модели YOLO12m проведено сравнение восьми оптимизаторов в идентичных условиях: 20 эпох; размер батча 16; косинусный планировщик скорости; начальная скорость 0,01 для SGD и MuSGD; 0,001 для остальных; параметры аугментации – по умолчанию. Результаты на тестовой выборке приведены в таблице 3.

Лучшие результаты показал SGD, близкие – MuSGD; наихудшие – AdamW. SGD обеспечивает стабильную сходимость и эффективен для задач детекции.

Влияние размера батча исследовалось на модели YOLO12m при 20 эпохах, оптимизаторе SGD, начальной скорости 0,01, косинусном планировщике и параметрах аугментации по умолчанию. Протестировано четыре варианта; результаты на тестовой выборке приведены в таблице 4.

Лучшие результаты показал батч 16, наихудшие – батч 8. Увеличение размера батча до 16 оптимизирует оценку градиентов; дальнейшее увеличение не дает существенного прироста, поэтому батч 16 обеспечивает баланс точности и скорости.

Влияние коэффициента регуляризации Dropout исследовалось на модели YOLO12m при 20 эпохах,

батче 16, оптимизаторе SGD, начальной скорости 0,01, косинусном планировщике и параметрах аугментации по умолчанию. Протестировано пять значений Dropout; результаты на тестовой выборке приведены в таблице 5.

Изменение Dropout не повлияло на метрики: в фреймворке Ultralytics этот параметр активируется только для классификационных сетей, где применяется к регуляризирующим слоям. В архитектуре YOLO12m такие слои отсутствуют, поэтому переданное значение Dropout игнорируется при обучении и не отражается на результатах. Таким образом, неизменность показателей в таблице 5 обусловлена техническим ограничением, а не эффективностью регуляризации.

Финальное обучение YOLO12m выполнено с увеличенным до 50 числом эпох, батчем 16, оптимизатором SGD, начальной скоростью 0,01, конечной – 0,0001 (косинусный планировщик), отключением мозаичной аугментации за 20 эпох до завершения и ранней остановкой при отсутствии улучшений в течение 10 эпох; остальные параметры аугментации – по умолчанию (табл. 6).

Результаты финального обучения показали высокую точность модели: mAP50-95 достигла 0,8448. Модель корректно классифицирует все четыре категории ягод. Примеры детекции приведены на рисунке 4.

Достигнутые значения метрик, представленные в таблице 6, получены на тестовой выборке, сформированной

Таблица 2. Сравнение метрик обученных версий модели YOLO12

Table 2. Metrics of trained YOLO12 versions

Модель	Precision	Recall	mAP50	mAP50-95
YOLO12n	0,8817	0,8771	0,9372	0,8069
YOLO12s	0,9170	0,8926	0,9537	0,8341
YOLO12m	0,9253	0,9049	0,9600	0,8436

Таблица 3. Сравнение метрик модели YOLO12m с различными оптимизаторами

Table 3. YOLO12m model metrics with different optimizers

Оптимизатор	Precision	Recall	mAP50	mAP50-95
Adam	0,9023	0,8948	0,9493	0,8290
AdamW	0,8991	0,8802	0,9464	0,8268
Adamax	0,9100	0,9069	0,9571	0,8392
MuSGD	0,9251	0,9012	0,9586	0,8418
NAdam	0,9092	0,8970	0,9528	0,8293
RAdam	0,9065	0,9024	0,9530	0,8354
RMSProp	0,9129	0,8930	0,9507	0,8323
SGD	0,9253	0,9049	0,9600	0,8436

Таблица 4. Сравнение метрик модели YOLO12m с разным размером батча

Table 4. YOLO12m model metrics with different batch sizes

Размер батча	Precision	Recall	mAP50	mAP50-95
8	0,9204	0,8995	0,9532	0,8366
12	0,9174	0,9026	0,9596	0,8404
16	0,9253	0,9049	0,9600	0,8436
20	0,9201	0,9027	0,9578	0,8428

Таблица 5. Сравнение метрик модели YOLO12m с разной величиной Dropout

Table 5. YOLO12m model metrics with different Dropout values

Dropout	Precision	Recall	mAP50	mAP50-95
0,10	0,9253	0,9049	0,9600	0,8436
0,15	0,9253	0,9049	0,9600	0,8436
0,20	0,9253	0,9049	0,9600	0,8436
0,25	0,9253	0,9049	0,9600	0,8436
0,30	0,9253	0,9049	0,9600	0,8436

Таблица 6. Значения метрик после финального обучения

Table 6. Metric values after final training

Модель	Precision	Recall	mAP50	mAP50-95
YOLO12m	0,9034	0,9243	0,9562	0,8448



Рисунок 4. Определение ягод вишни на изображении

Figure 4. Visual identification of cherries

Таблица 7. Значения метрик финальной модели YOLO12m на выборке производственных изображений

Table 7. Final YOLO12m model: Metric values on the production image subset

Модель	Precision	Recall	mAP50	mAP50-95
YOLO12m	0,8893	0,9019	0,9214	0,8124

рованной из объединенного датасета, включающего как производственные изображения, так и изображения из открытого источника. Такой состав соответствует общепринятой практике оценки моделей, обученных с расширением данных, и позволяет судить об общей способности модели распознавать ягоды различных категорий при вариативных условиях съемки. Для оценки применимости модели в реальном производстве проведена дополнительная валидация YOLO12m на 50 новых снимках с линии ООО «НПО Здоровое Питание» (Кемерово, Россия), собранных и размеченных по ранее описанной методике. Результаты приведены в таблице 7.

При тестировании модели исключительно на изображениях, полученных в производственных условиях, значения метрик несколько снижаются по сравнению с результатами на смешанной тестовой выборке, однако остаются на уровне, достаточном для практического применения.

На основе обученной модели разработана программа ЭВМ, направленная на регистрацию в Федеральный институт промышленной собственности (Москва, Россия). Программа предоставляет графический интерфейс для загрузки изображения ягод вишни и отображения результатов их классификации. Алгоритм работы представлен на рисунке 5.

Работа с программой начинается с выбора файла изображения. Для этого пользователь нажимает кнопку «Загрузка». Диалоговое окно позволяет выбрать файл с локального устройства. После загрузки изображение

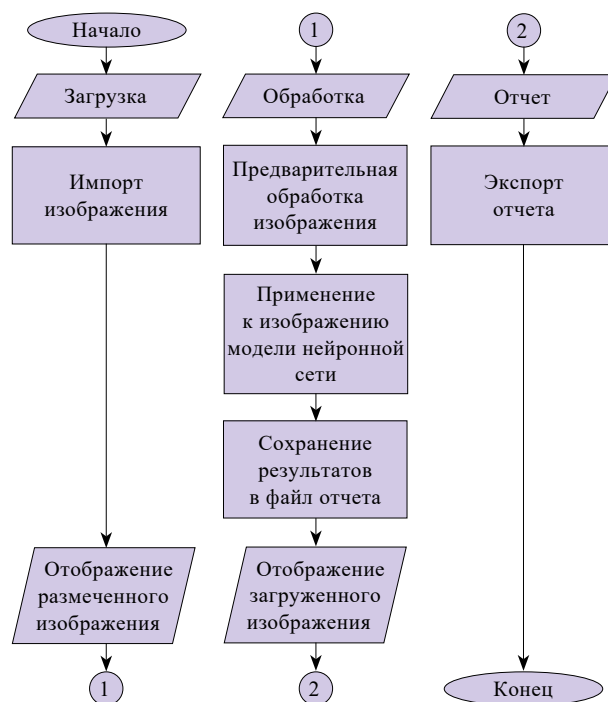


Рисунок 5. Алгоритм работы программы

Figure 5. Operation algorithm

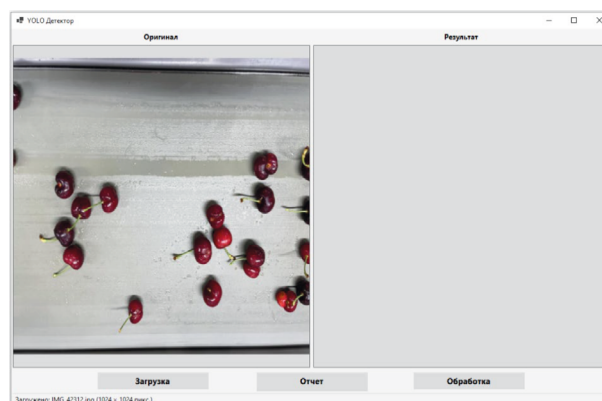


Рисунок 6. Интерфейс программы после загрузки изображения

Figure 6. Program interface after image loading

отображается в левой части окна программы (рис. 6), что позволяет проверить корректность выбранного файла перед анализом.

После нажатия кнопки «Обработка» программа передает изображение в модель YOLO12m, которая выполняет детекцию и возвращает результаты. Обработанное изображение с рамками и метками классов отображается в правой части окна (рис. 7). Результаты автоматически сохраняются в текстовый файл с указанием количества ягод каждого класса. Программа также сравнивает долю кондиционных и некондиционных

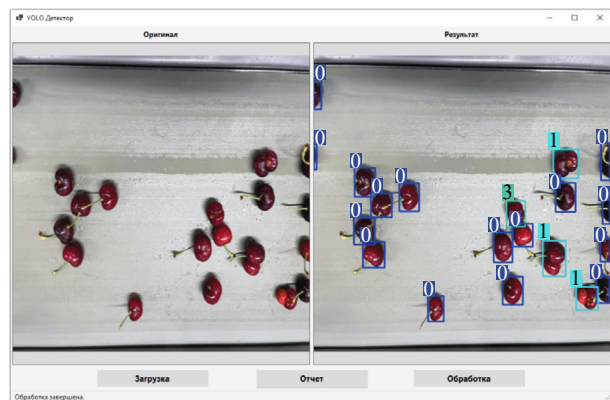


Рисунок 7. Интерфейс программы после обработки изображения

Figure 7. Program interface after image processing

ягод, определяет товарный сорт партии и формирует отчет, доступный для скачивания по кнопке «Отчет».

Программа разработана для цифровизации органолептической оценки внешнего вида сырья на этапе приемки. Ее применение снижает субъективность ручного контроля и сокращает время оценки партии. В перспективе программа может использоваться на непрерывных производствах для контроля соответствия сырья требованиям ГОСТ и технологических инструкций – в этом случае на вход будет подаваться видеопоток вместо отдельных изображений, что потребует адаптации архитектуры ПО. Данная задача относится к автоматизации производственных процессов и выходит за рамки настоящего исследования.

Выводы

В ходе работы собран и подготовлен обучающий набор данных, состоящий из изображений ягод вишни четырех категорий качества, полученных в реальных производственных условиях на этапе приемки сырья. Все изображения размечены вручную в инструменте CVAT – каждой ягоде присвоена метка класса и ограничивающая рамка. Это позволило обучить модель самостоятельному распознаванию категорий без участия оператора.

Для детекции выбраны модели семейства YOLO. Проведено последовательное сравнение пяти версий (YOLOv8m, YOLOv9m, YOLO11m, YOLO12m и YOLO26m) и трех вариантов YOLO12 по размеру (nano, small и medium). Для YOLO12m выполнена

оптимизация гиперпараметров: сравнивались восемь оптимизаторов, четыре значения батча и пять значений параметра Dropout. Все конфигурации обучались в идентичных условиях с аугментацией данных и финальным тестированием на отдельной выборке.

По результатам сравнения пяти моделей преимущество продемонстрировала архитектура YOLO12m, показав наилучшие значения метрики mAP50-95 среди всех рассмотренных архитектур. Из всех оптимизаторов лучший результат дал SGD, оптимальный батч – 16. Проверка Dropout показала его неэффективность: в используемом фреймворке механизм регуляризации реализован только для классификационных задач, что учтено при интерпретации. Финальное обучение YOLO12m с увеличенным числом эпох повысило mAP50-95 до 0,8448; на производственной выборке значение составило 0,8124, что подтверждает применимость модели в реальных условиях.

На основе обученной модели разработана программа для ЭВМ. Оператор загружает снимок, программа передает его в модель, после обработки на экране отображается размеченное изображение с рамками и метками классов. Для работы с программой не требуется специальных знаний.

Исследование показало, что модели семейства YOLO пригодны для решения задачи классификации ягод вишни на этапе приемки сырья. YOLO12m уверенно распознает все четыре категории. Выявление некондиционных ягод до переработки снижает риск ухудшения качества готовой продукции. Предложенный подход обеспечивает стабильный и воспроизводимый результат, а также уменьшает время и трудозатраты на входной контроль.

Критерии авторства

Авторы в равной степени участвовали в написании рукописи.

Конфликт интересов

Авторы заявляют об отсутствии конфликта интересов.

Contribution

The authors contributed equally to the research and bear equal responsibility for the information published in this article.

Conflict of interest

The authors state that there is no conflict of interest.

Список литературы / References

1. Шафрай А. В., Попов А. М., Косинов В. С., Резниченко И. Ю., Бондарчук О. Н. Применение сверточных нейронных сетей для контроля показателей технологического потока производства гранулированных киселей. Техника и технология пищевых производств. 2025. Т. 55. № 4. С. 845–855. [Shafrai AV, Popov AM, Kosinov VS, Reznichenko IYu, Bondarchuk ON. Convolutional neural networks in granulated kissel production. Food Processing: Techniques and Technology. 2025;55(4):845–855. (In Russ.)] <https://doi.org/10.21603/2074-9414-2025-4-2613>

2. Daiaeddine MJ, Badrouss S, El Harti A, Bachaoui EM, Biniz M, *et al.* UAV imagery, advanced deep learning, and YOLOv7 object detection model in enhancing citrus yield estimation. *Foods and Raw Materials*. 2025;13(2):242–253. <https://doi.org/10.21603/2308-4057-2025-2-650>
3. Yavuzer E, Köse M, Uslu H. Determining the quality level of ready to-eat stuffed mussels with Arduino-based electronic nose. *Journal of Food Measurement and Characterization*. 2024;18:5629–5637. <https://doi.org/10.1007/s11694-024-02593-9>
4. Wijaya DR, Syarwan NF, Nugraha MA, Ananda D, Fahrudin T, *et al.* Seafood quality detection using electronic nose and machine learning algorithms with hyperparameter optimization. *IEEE Access*. 2023;11:62484–62495. <https://doi.org/10.1109/ACCESS.2023.3286980>
5. Zlobin I, Sinelnikov A, Toroptsev N, Titov K, Antoshkina E, *et al.* Metal–organic frameworks at the tip of the e-tongue: Machine learning-driven disposable electrochemical sensors. *Journal of Materials Chemistry C*. 2025;13(23):11776–11788. <https://doi.org/10.1039/D5TC00965K>
6. Anisimov DS, Abramov AA, Gaidarzhi VP, Kaplun DS, Agina EV, *et al.* Food freshness measurements and product distinguishing by a portable electronic nose based on organic field-effect transistors. *ACS Omega*. 2023;8(5):4649–4654. <https://doi.org/10.1021/acsomega.2c06386>
7. Shuba A, Umarhanov R, Bogdanova E, Anokhina E, Burakova I. Possibilities of an electronic nose on piezoelectric sensors with polycomposite coatings to investigate the microbiological indicators of milk. *Sensors*. 2024;24(11):3634. <https://doi.org/10.3390/s24113634>
8. Fedorov FS, Yaqin A, Krasnikov DV, Kondrashov VA, Ovchinnikov G, *et al.* Detecting cooking state of grilled chicken by electronic nose and computer vision techniques. *Food Chemistry*. 2021;345:128747. <https://doi.org/10.1016/j.foodchem.2020.128747>
9. Wu X, Xu B, Ma R, Niu Y, Gao S, *et al.* Identification and quantification of adulterated honey by Raman spectroscopy combined with convolutional neural network and chemometrics. *Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy*. 2022;274(1):121133. <https://doi.org/10.1016/j.saa.2022.121133>
10. Lanjewar MG, Asolkar S, Parab JS, Morajkar PP. Detecting starch-adulterated turmeric using Vis-NIR spectroscopy and multispectral imaging with machine learning. *Journal of Food Composition and Analysis*. 2024;136:106700. <https://doi.org/10.1016/j.jfca.2024.106700>
11. Saha D, Senthilkumar T, Singh CHB, Manickavasagan A. Quantitative detection of metanil yellow adulteration in chickpea flour using line-scan near-infrared hyperspectral imaging with partial least square regression and one-dimensional convolutional neural network. *Journal of Food Composition and Analysis*. 2023;120:105290. <https://doi.org/10.1016/j.jfca.2023.105290>
12. Belugina R, Senchikhina A, Volkov S, Fedorov A, Legin A, *et al.* Quantification of phosphatides in sunflower oils using a potentiometric e-tongue. *Analytical Methods*. 2022;14(32):3064–3070. <https://doi.org/10.1039/D2AY00736C>
13. Sánchez CN, Orvañanos-Guerrero MT, Domínguez-Soberanes JU, Álvarez-Cisneros YEM. Analysis of beef quality according to color changes using computer vision and white-box machine learning techniques. *Heliyon*. 2023;9(7):e17976. <https://doi.org/10.1016/j.heliyon.2023.e17976>
14. Huang W, Wang X, Zhang JU, Xia J, Zhang X. Improvement of blueberry freshness prediction based on machine learning and multi-source sensing in the cold chain logistics. *Food Control*. 2023;145:109496. <https://doi.org/10.1016/j.foodcont.2022.109496>
15. Anwar H, Anwar T, Murtaza SH. Review on food quality assessment using machine learning and electronic nose system. *Biosensors and Bioelectronics*. 2023;14:100365. <https://doi.org/10.1016/j.biosx.2023.100365>
16. Ксенофонов В. В. Нейронные сети. Проблемы науки. 2020. № 11. С. 28–29. [Ksenofontov VV. Neural networks. Problems of Science. 2020;(11):28–29. (In Russ.)] <https://elibrary.ru/DTVJNS>
17. Sapkota R, Flores-Calero M, Qureshi R, Badgujar CH, Nepal U. YOLO advances to its genesis: A decadal and comprehensive review of the You Only Look Once (YOLO) series. *Artificial Intelligence Review*. 2025;58(9):274. <https://doi.org/10.1007/s10462-025-11253-3>
18. Mirzaei B, Nezamabadi-Pour H, Raoof A, Derakhshani R. Small object detection and tracking: A comprehensive review. *Sensors*. 2023;23(15):6887. <https://doi.org/10.3390/s23156887>

Дополнительная информация об авторах / Additional information about the authors

Попов Анатолий Михайлович / Anatoly M. Popov ORCID 0000-0003-0728-7211; eLIBRARY SPIN 6259-7555
Шафрай Антон Валерьевич / Anton V. Shafrai ORCID 0000-0003-4512-1933; eLIBRARY SPIN 1840-5242
Косинов Виталий Сергеевич / Vitaly S. Kosinov ORCID 0009-0005-5852-9809; eLIBRARY SPIN 8397-7880
Подберезен Григорий Игоревич / Grigory I. Podberezen ORCID 0000-0003-4499-8426; eLIBRARY SPIN 8477-8514
Сухоруков Дмитрий Викторович / Dmitry V. Sukhorukov ORCID 0000-0001-7995-3813; eLIBRARY SPIN 1595-7666
Комаров Сергей Сергеевич / Sergey S. Komarov ORCID 0000-0002-1517-1596; eLIBRARY SPIN 2719-1509