

Разработка алгоритма автоматического определения числа пропусков растений картофеля по снимкам посевов с БПЛА с применением машинного обучения

Д.А. Полещенко¹, И.С. Михайлов¹

¹ Старооскольский технологический институт им. А.А. Угарова (филиал) ННТУ «МИСИС», 309516, Россия, г. Старый Оскол, мкр. Макаренко, д. 42

Аннотация

В работе представлен разработанный подход к автоматической идентификации рядов картофеля и обнаружению пропусков по аэрофотоснимкам с БПЛА. Метод сочетает сегментацию растений с использованием моделей YOLOv11, адаптивную морфологическую фильтрацию с учётом пространственного разрешения и автоматическую коррекцию ориентации рядов. Экспериментальная апробация на полевых данных показала высокую точность выделения рядов (97,4 %) и эффективность выявления пропусков (F1-score 0,91). Полученные результаты свидетельствуют о корректности идентификации структуры посевов, что обосновывает применимость подхода для поддержки решений в цифровом сельском хозяйстве.

Ключевые слова: мониторинг посевов, картофель, сегментация растений, ортофотоплан, YOLOv11, Ground Sample Distance, дистанционное зондирование, БПЛА.

Цитирование: Полещенко, Д.А. Разработка алгоритма автоматического определения числа пропусков растений картофеля по снимкам посевов с БПЛА с применением машинного обучения / Д.А. Полещенко, И.С. Михайлов // Компьютерная оптика. – 2026. – Т. 50, № 4. – 1875. – doi: 10.18287/COJ1875.

Citation: Poleshchenko DA, Mikhailov IS. Development of a machine-learning algorithm for the automatic determination of the number of missing potato plants from UAV imagery. Computer Optics 2026; 50(4): 1875. doi: 10.18287/COJ1875.

Введение

Точное земледелие ориентировано на повышение эффективности производства за счёт цифровых технологий и автоматизированного анализа данных. Одной из ключевых задач является автоматическое определение числа пропусков сельскохозяйственных культур по изображениям с беспилотных летательных аппаратов (БПЛА; англ. UAV).

Нарушения в структуре посадочных рядов – пропуски и пространственная неравномерность – возникают из-за низкой всхожести, повреждений вредителями, стрессовых погодных условий и ошибок посева, статистически значимо снижая урожайность. Их своевременное автоматизированное выявление уменьшает влияние человеческого фактора и упрощает интеграцию результатов в цифровые системы поддержки принятия решений.

Современные алгоритмы обработки изображений, включая методы машинного обучения и глубокие нейронные сети, обеспечивают детальный анализ структуры посевов. Свёрточные архитектуры класса U-Net и семейства YOLO демонстрируют высокую точность в задачах сегментации и обнаружения объектов на сельскохозяйственных изображениях с БПЛА [1].

Перспективность компактных моделей для полевых условий подтверждают работы по автоматизированной сегментации рядов культур: облегчённая модель VC – U-Net для масличного рапса достигает точности порядка 94 % при скорости, достаточной для систем реального времени [2], что важно при ограниченных вычислительных ресурсах и энергопотреблении.

Одновременно в литературе представлены исследования, ориентированные на сравнительный анализ спектральных каналов и методических подходов. В работе на материалах виноградников, персиков и томатов показано, что алгоритмы пороговой сегментации, классификации и байесовской сегментации позволяют выделять растительность с точностью более 90 % по данным БПЛА, а прирост от использования ближнего инфракрасного диапазона (NIR) по сравнению с RGB-снимками при распознавании рядов оказывается ограниченным. Это, в совокупности с высоким пространственным разрешением и частотой съёмки, показывает, что RGB-данные с сантиметровым Ground Sample Distance (GSD) достаточны для детектирования растений и рядов, а также оперативного контроля однородности всходов и наличия пропусков [3].

Для более сложных сцен с плотным переплетением листовой массы и вариативным фоном эффективность демонстрируют специализированные модификации детекторов: архитектура SugarRow – YOLO, включающая модули внимания и усовершенствованные функции потерь, достигает точности выделения рядов до 96,6 % [4], что подтверждает значимость адаптации базовых алгоритмов под конкретные агросценарии.

Наряду с глубокими моделями, сохраняют актуальность классические методы, основанные на геометрической обработке изображения и разделении классов «растительность/фон». В исследовании [5] используются

коррекция перспективы, цветовая сегментация, кластеризация и устойчивая подгонка прямых линий; алгоритм достигает IoU порядка 0,73 и точности обнаружения рядов более 90 % даже при сильном засорении сорняками. Сравнительные работы показывают, что при анализе ортомозаик классический алгоритм Хафа нередко уступает альтернативным подходам. В частности, в [6] метод наименьших квадратов при вычислении метрики CRDA (Crop Row Detection Accuracy) обеспечивает значения 0,99 – 1,00 для хлопка и 0,66 – 0,82 для пшеницы против 0,31 – 0,53 для преобразования Хафа.

Современные исследования активно используют детекторы объектов и сегментационные модели для растениеориентированного анализа. В работе [7] детектор Faster R–CNN для распознавания кустов картофеля на RGB-снимках БПЛА показал среднее значение AP около 0,78 и высокую корреляцию карт плотности с фактической урожайностью ($R^2 \approx 0,80$), что подчёркивает практическую значимость подхода. В [8] облегчённая модификация Maize-YOLOv8n для детекции пропавших всходов кукурузы достигла mAP $\approx 97,4\%$, точности 94,3 % и полноты 93,1 % при времени инференса 12,8 мс на изображение, демонстрируя устойчивость к вариациям освещения и засорённости фона.

Пропуски растений внутри рядов снижают продуктивность и часто требуют досева. Распространённые подходы основываются на детекции отдельных растений и анализе расстояний между ними с адаптивными порогами: для посевов сои ансамбль моделей (SVM и многослойный перцептрон) по данным БПЛА обеспечил AUC около 0,98 и точность порядка 95 % [9]. В работе [10] геометрический метод для рядов кукурузы по RGB-снимкам (выделение рядов и анализ расстояний между растениями) достиг точности 93 % для пропусков свыше 75 см и 91 % для диапазона 45 – 75 см, демонстрируя устойчивость к неоднородности всходов и потенциальную переносимость на другие культуры, включая картофель. Для плантаций сахарного тростника подход, комбинирующий классификацию ортомозаик методом К-ближайших соседей, оценку ориентации рядов по RGB-градиенту и морфологическую обработку, позволил измерять пропуски с малой погрешностью [11].

В условиях цифровизации агросектора всё более актуальной становится автоматизация объективной оценки структуры и качества посадок картофеля по аэрофотоснимкам с БПЛА. Настоящая работа посвящена разработке интегрированного алгоритма методов глубокого обучения для решения задач автоматического выделения посадочных рядов и процедур морфологической фильтрации и коррекции ориентации рядов для обнаружения пропусков.

1. Материалы и методы

1.1. Исходные данные и аэрофотосъёмка

Эксперимент проводился на картофельном поле общей площадью 8,5 га в Городецком районе Нижегородской области (56°26'59.44" с.ш., 43°54'17.14" в.д.). Почва – дерново-подзолистая, средней влажности и хорошей воздухопроницаемости, характерная для данной агроклиматической зоны.

Аэрофотосъёмка выполнена 25 июня 2024 года в утренние часы с использованием БПЛА *DJI Phantom 4 Pro v2.0* с камерой *DJI FC6310S* (20 Мп, сенсор 1" CMOS, фокусное расстояние 8,8 мм). Съёмка велась с высоты 15 м с продольным перекрытием кадров 80 % и поперечным 70 %.

Получены изображения с пространственным разрешением 0,36 см / пиксель (GSD), что обеспечило детальную визуализацию растений и междурядий. Геопривязка выполнялась штатным GNSS-приёмником БПЛА (точность $\pm 1,5$ м). Относительная точность ортофотоплана составила 0,5 – 0,8 см, что достаточно для корректных измерений в пределах поля.

Фенологическая стадия развития посадок определялась по международной шкале BBCH на основе полевого осмотра и соответствовала фазе 35 – 39 (активный рост ботвы, частичное смыкание рядов). Полученные снимки общим объемом более 1500 кадров были обработаны в программном обеспечении семейства *PIX4D* с формированием единого ортофотоплана, использованного на последующих этапах анализа.

1.2. Подготовка данных и обучение моделей

В рамках настоящего исследования для автоматической идентификации и оценки качества посадочных рядов картофеля применялись методы сегментации изображений, полученных с беспилотного летательного аппарата. Ключевым этапом подготовки данных стала ручная разметка исходных аэрофотоснимков с использованием платформы *Roboflow* [12]. Всего было размечено 84 полноразмерных изображения (5472 × 3648 пикселей), обеспечивающих репрезентативное покрытие исследуемого участка.

Для повышения эффективности обучения и оценки обобщающей способности моделей использовался следующий пайплайн подготовки данных. На первом этапе исходный массив из 84 полноразмерных ортофотоснимков был разделен на три независимые подвыборки: обучающую (70 %), валидационную (15 %) и тестовую (15 %). Разделение выполнялось строго на уровне исходных кадров до процедуры нарезки. Это критически важно для исключения утечки данных, при которой перекрывающиеся фрагменты одного участка поля могли бы оказаться одновременно в обучающей и тестовой частях.

Далее изображения каждой подвыборки были нарезаны на тайлы размером 1024 × 1024 пикселя с перекрытием 30 % по горизонтали и вертикали. Такой подход позволил сохранить контекст на границах фрагментов. Для тайлов обучающей выборки дополнительно применялась процедура аугментации, что

позволило кратно увеличить объем тренировочных данных. Валидационная и тестовая выборки использовались без аугментаций. В итоговый набор данных вошло 2870 изображений: обучающая выборка – 2500 тайлов, валидационная – 185 тайлов, тестовая – 185 тайлов.

Комплекс аугментаций для обучающей выборки включал: геометрические трансформации (горизонтальные и вертикальные отражения, случайные повороты), фотометрические искажения (корректировка яркости, контраста, оттенка, насыщенности, добавление гауссовского шума и размытия), а также метод *Cutout* (случайное удаление прямоугольных областей до 25 % площади кадра).

Для обучения использовались три модификации *YOLOv11-Seg* (nano, small, medium) [13]. Дополнительно, в целях сравнительного анализа с классическими методами семантической сегментации, применяемыми в задачах агромониторинга [14], были реализованы архитектуры *U-Net* и *DeepLabV3* с энкодером *ResNet-34*. Основные гиперпараметры и применённые аугментации приведены в табл. 1. При этом особое внимание уделялось подбору числа эпох, параметру *freeze* и стратегии аугментации, поскольку именно они наиболее существенно влияют на качество сегментации и устойчивость моделей.

Табл. 1. Основные гиперпараметры и обоснование выбора

Гиперпараметр	Значения	Обоснование выбора
epochs	80	Баланс между сходимостью и риском переобучения
patience	20	Ранняя остановка при отсутствии улучшения метрик
batch size	24 (nano), 16 (small), 8 (medium)	Оптимальное использование GPU-памяти
freeze	0, 5, 10	Анализ влияния заморозки слоёв на сегментацию
augmentations	Mosaic, MixUp (<i>close mosaic</i> = 30, <i>scale</i> = 0.5)	Повышение устойчивости к вариациям данных

В исследовании применён подход увеличения обобщающей способности модели *Mosaic*, который комбинирует четыре случайных изображения в один батч, искусственно увеличивая разнообразие и сложность объектов на тайлах. Это способствует лучшей генерализации модели, особенно на малых выборках или в условиях ограниченного разнообразия сцен.

Параметр *close_mosaic* = 30 автоматически отключает *Mosaic*-аугментацию на последних 30 % эпох обучения, что позволяет модели на финальном этапе более точно подстроиться к реальным распределениям входных данных и снизить вероятность появления артефактов, связанных с синтетическими композициями.

Обучение моделей проводилось на рабочей станции с графическим ускорителем *NVIDIA TITAN RTX* (24 ГБ VRAM) и процессором *Intel Core i9-10920X*, что обеспечивало возможность параллельной работы с крупными батчами и ускоряло перебор гиперпараметров.

Экспериментально было установлено, что конфигурации *nano* и *small* демонстрируют высокую скорость инференса (3 – 5 мс на тайл), тогда как модификация *medium* достигает наилучших результатов по показателю точности (*mAP@0.5* до 0,94 – 0,96). Среднее время обработки одного тайла (1024 × 1024 пикселя) для модели *medium* составило 27–30 мс (FPS ≈ 35). Для сравнения, время инференса моделей *U-Net* и *DeepLabV3* превышает 60 – 70 мс на тайл. Хотя задача картирования пропусков формально не относится к системам жёсткого реального времени, фактор производительности становится критическим при масштабировании. Полный ортофотоплан поля площадью 8,5 га содержит около 13 000 тайлов, обработка которых моделью *medium* занимает 6 – 7 минут. Однако типичные площади полей в картофелеводческих хозяйствах РФ варьируются от 50 до 300 га. При экстраполяции на участок в 100 га объём данных возрастает до порядка 150 000 тайлов. В этом случае использование *YOLOv11m* потребует порядка 1,2 часа, тогда как применение более тяжёлых архитектур увеличит время анализа до 3 – 4 часов, что существенно ограничивает возможности оперативного мониторинга и принятия решений в день облёта.

Критерием оптимальности модели служили метрики *mAP@0.5*, *mAP@0.5:0.95* и F1-score, вычислявшиеся на тестовой выборке.

1.3. Обнаружение рядов на сегментированных масках

Входными данными для данного этапа является бинарная маска сегментации M_{segm} , полученная на выходе нейронной сети. На маске значения пикселей, равные 1 (белый цвет), соответствуют классу «растение» (картофель), а значения 0 (чёрный цвет) – классу «фон» (почва). Обнаружение рядов осуществляется с целью дальнейшего количественного анализа структуры посадок и поиска пропусков. Этап включает последовательную обработку результатов сегментации: коррекцию ориентации рядов, морфологическую фильтрацию для удаления артефактов и выделение их геометрии.

1.3.1. Оценка и коррекция ориентации рядов

Первым шагом обработки является приведение изображения к положению, при котором ряды располагаются строго вертикально. Это упрощает последующий анализ структуры посадок и повышает точность выделения отдельных рядов. Для определения доминирующего направления рядов были исследованы два подхода: метод главных компонент (Principal Component Analysis, PCA) [16], применённый к координатам центроидов связанных

компонент сегментированной маски, и rotation-based алгоритм, аналогичный описанному в [17]. Оба метода обеспечили сопоставимое качество оценки ориентации, однако PCA продемонстрировал кратно более высокую скорость работы (примерно в 10 раз быстрее на тестовых данных), поэтому в дальнейшем в рабочем пайплайне использовался именно этот подход.

Пусть (x_i, y_i) – координаты центров масс компонент. Рассчитывается ковариационная матрица согласно формуле (1):

$$\mathbf{C} = \frac{1}{N} \sum_{i=1}^N \begin{pmatrix} x_i - \bar{x} \\ y_i - \bar{y} \end{pmatrix} \begin{pmatrix} x_i - \bar{x} & y_i - \bar{y} \end{pmatrix}, \quad (1)$$

где $(\bar{x}, \bar{y})$ – средние значения по координатам; N – число компонент (центроидов).

Вычисляются собственные значения и собственные векторы матрицы $\mathbf{C}$. Главная ось (направление рядов) соответствует собственному вектору с максимальным собственным значением:

$$\mathbf{v} = \begin{pmatrix} v_x \\ v_y \end{pmatrix}, \quad (2)$$

где v_x и v_y – компоненты по осям абсцисс и ординат соответственно. Угол ориентации определяется как

$$\theta = \arctan2(v_y, v_x). \quad (3)$$

Далее маска поворачивается на угол $90^\circ - \theta$, чтобы ряды стали параллельны оси ординат.

В результате коррекции ориентации дальнейшие вычисления (поиск осей рядов, анализ пропусков) выполняются в единой координатной системе.

1.3.2. Морфологическая очистка и улучшение геометрии маски

После получения сегментированной маски и выравнивания изображения так, чтобы ряды ориентировались вертикально, выполняется адаптивная морфологическая очистка, направленная на удаление поперечных перемычек (см. рис. 1а) и мелких шумовых включений, а также на приведение геометрии рядов к согласованному виду.

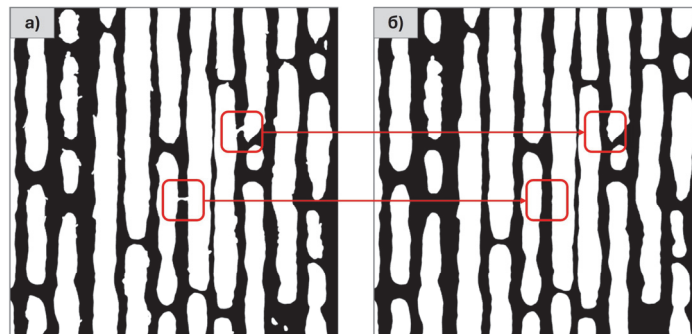


Рис. 1. Морфологическая очистка маски: а) исходная маска с шумом и горизонтальными перемычками; б) результат адаптивного морфологического открытия

Размер структурного элемента подбирается автоматически на основе GSD (см/пиксель) и максимально допустимой физической ширины удаляемых перемычек w_{bridge} (см). Требуемый размер ядра в пикселях определяется согласно (4):

$$k = \begin{cases} \left\lceil \frac{w_{\text{bridge}}}{GSD} \right\rceil + 1, & \left\lceil \frac{w_{\text{bridge}}}{GSD} \right\rceil = 0 \pmod{2}, \\ \left\lceil \frac{w_{\text{bridge}}}{GSD} \right\rceil, & \left\lceil \frac{w_{\text{bridge}}}{GSD} \right\rceil \neq 0 \pmod{2}, \end{cases} \quad (4)$$

где k – нечётный линейный размер ядра (в пикселях), обеспечивающий центральную опорную точку. Параметр w_{bridge} имеет физический смысл – максимальная ширина смыкания крон между соседними рядами, которую алгоритм должен интерпретировать как ложную перемычку. В данной работе для фазы активного роста ботвы экспериментально выбрано значение $w_{\text{bridge}} = 20$ см.

Поскольку ряды вертикальные, а перемычки имеют горизонтальную ориентацию, применяется вытянутый по вертикали структурный элемент (5):

$$K = \mathbf{1}_{k \times w_{\text{thin}}}, \quad (5)$$

где K – прямоугольный структурный элемент, заданный матрицей из единиц $\mathbf{1}_{a \times b}$ с центральной опорной точкой. «Длинная» сторона равна k , определяемая по (4). «Тонкая» сторона w_{thin} (в данной работе принята равной 7 пикселям) выбрана исходя из разрешения снимка $\approx 0,36$ см/пикс так, чтобы соответствовать масштабу порядка

нескольких сантиметров (2–3 см). Это позволяет гарантированно удалять мелкий шум и тонкие листовые связи, не повреждая основной стебель растения. В общем случае параметр может быть пересчитан для других условий съёмки по соотношению между желаемым физическим размером на местности и значением GSD.

Очистка маски выполняется морфологическим открытием со структурным элементом K :

$$M_{\text{cleaned}} = (M_{\text{segm}} \ominus K) \oplus K, \quad (6)$$

где M_{segm} – исходная бинарная маска сегментации; M_{cleaned} – результат фильтрации; $\ominus$ – операция эрозии (уменьшения) бинарной маски с использованием структурного элемента K , удаляет мелкие объекты и шумы по границе сегментации; $\oplus$ – операция дилатации (расширения) бинарной маски с тем же структурным элементом K , восстанавливающая форму крупных объектов после эрозии. После применения (6) малые связные компоненты исключаются по порогу площади, в результате чего формируется геометрически согласованная маска с чётко разделёнными протяжёнными рядами, пригодная для извлечения их осей и последующего анализа разрывов. Эффект фильтрации иллюстрируется на рис. 1: на рис. 1а заметны шум и горизонтальные перемычки, на рис. 1б – после применения фильтра по (6) перемычки устранены, контуры рядов упорядочены.

1.3.3. Поиск связных компонент и кластеризация рядов

На повернутой и бинаризованной маске (см. рис. 2а) выделяются все связные компоненты.

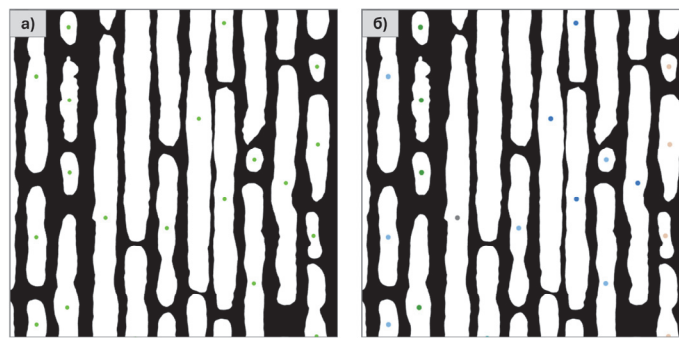


Рис. 2. Этап кластеризации центроидов: а) центры масс связных компонент на повернутой и очищенной маске (зелёные точки); б) результат кластеризации: точки раскрашены по принадлежности к кластерам

Перед вычислением координат центров масс выполняется предварительная фильтрация шума: объекты, площадь которых не превышает порогового значения $S_{\min}$ (в работе принято $S_{\min} = 50$ пикселей), исключаются из анализа как артефакты сегментации. Для каждой оставшейся (значимой) компоненты вычисляется центр масс (x_i, y_i) . В дальнейшем для вертикально ориентированных рядов используется абсцисса центроида x_i . Сформированное множество координат центров масс имеет вид:

$$X = \{x_i\}_{i=1}^N, \quad (7)$$

где N – число компонент, прошедших фильтрацию по площади.

Далее выполняется одномерная кластеризация вдоль оси x алгоритмом DBSCAN [17]. Радиус окрестности задаётся в пикселях как

$$\varepsilon_{\text{px}} = k \cdot \frac{S}{\text{GSD}}, \quad (8)$$

где S – расстояние между рядами (см), $k \in (0; 1)$ – эмпирический коэффициент (в настоящей работе $k = 0,5$). Определение кластеров реализуется как:

$$\{C_j\} = \text{DBSCAN}(X; \varepsilon_{\text{px}}), \quad (9)$$

где $\{C_j\}$ – набор кластеров.

Положение оси ряда для каждого кластера C_j определяется медианой абсцисс его точек:

$$X_j^{\text{row}} = \text{median}\{x_i: x_i \in C_j\}. \quad (10)$$

В силу особенностей сегментации и разрывов возможна регистрация нескольких близко расположенных осей, фактически соответствующих одному и тому же ряду (см. рис. 3а).

На финальном шаге выполняется объединение таких осей: если для пары $(X_p^{\text{row}}, X_q^{\text{row}})$ выполняется

$$|X_p^{\text{row}} - X_q^{\text{row}}| < \tau_{\text{merge}}, \quad \tau_{\text{merge}} \equiv \varepsilon_{\text{px}} \text{ из (8)}, \quad (11)$$

то такие оси объединяются в класс эквивалентности. Для каждого класса координата результирующей оси вычисляется как медиана координат элементов класса:

$$X_j^{\text{final}} = \text{median}\{X_i^{\text{row}} | X_i^{\text{row}} \in [X_j^{\text{row}}]_{\sim}\}, \quad (12)$$

где $\sim$ – отношение эквивалентности по правилу (11), $[X_j^{\text{row}}]_{\sim}$ – класс эквивалентности. Тем самым получается конечный набор уникальных осей:

$$\{X_i^{\text{row}}\}_{i=1}^N / \sim = \{X_j^{\text{final}}\}_{j=1}^M, \quad M \leq N, \quad (13)$$

где M – итоговое число рядов на участке. В качестве примера на рис. 3а отмечены две близкие оси X_i^{row} и X_{i+1}^{row} , которые после объединения переходят в одну ось X_j^{final} (рис. 3б).

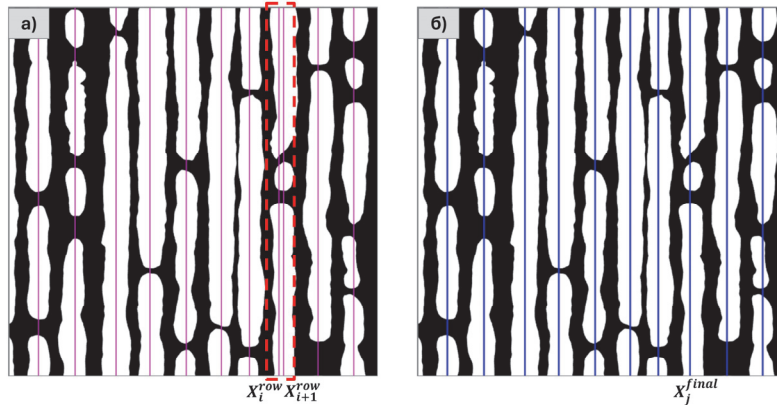


Рис. 3. Устранение дублирования осей: а) оси, выделенные до объединения (фиолетовые линии); пунктирной красной рамкой выделена зона с двумя близкими осями X_i^{row} и X_{i+1}^{row} ; б) финальный набор осей после объединения по порогу τ_{merge} из (11) (синие линии), где $X_i^{\text{row}}, X_{i+1}^{\text{row}} \rightarrow X_j^{\text{final}}$

1.3.4. Анализ пропусков в рядах

После построения координат осей рядов производится автоматический анализ непрерывности вдоль каждой оси. Для выбранного ряда формируется бинарная последовательность

$$S = [s_1, s_2, \dots, s_L], \quad s_i \in \{0, 1\}, \quad (14)$$

где $s_i = 1$ соответствует наличию сегментированного растения, а $s_i = 0$ – отсутствию растений (пропуску).

Критический порог для классификации разрыва как пропуска определяется на основе агротехнологических норм. Пусть L_{plant} – номинальный шаг посадки (расстояние между кустами в см), установленный при высадке. Для учёта естественной вариабельности развития растений вводится коэффициент допуска $\alpha \geq 1$. Пороговое значение длины пропуска в пикселях вычисляется как:

$$\tau_{\text{gap}} = \left\lceil \frac{L_{\text{plant}} \cdot \alpha}{\text{GSD}} \right\rceil. \quad (15)$$

В данной работе принято $L_{\text{plant}} = 30$ см, $\alpha = 1,5$, что определяет порог $\tau_{\text{gap}} \approx 310$ пикселей при текущем разрешении.

Пропуск определяется как *максимальный* интервал последовательных нулей $[a_i, b_i] \subset \{1, \dots, L\}$, где L – длина рассматриваемой линии ряда (в пикселях), для которого

$$s_i = 0 \quad \forall i \in [a_i, b_i], \quad (b_i - a_i + 1) > \tau_{\text{gap}}, \quad (16)$$

а «максимальность» задаётся краевыми условиями

$$(a_i = 1 \text{ или } s_{a_i-1} = 1), \quad (b_i = L \text{ или } s_{b_i+1} = 1). \quad (17)$$

Множество обнаруженных пропусков записывается как набор троек (a_i, b_i, ℓ_i) , где $\ell_i = b_i - a_i + 1$ – длина пропуска в пикселях:

$$\text{Gaps} = \{(a_i, b_i, \ell_i) : \ell_i > \tau_{\text{gap}}\}. \quad (18)$$

Для отчётности границы переводятся в физические координаты:

$$y_{\text{start},i} = a_i \cdot \text{GSD}, \quad y_{\text{end},i} = b_i \cdot \text{GSD}, \quad \ell_{\text{cm},i} = \ell_i \cdot \text{GSD} \text{ (см)}.$$

Предложенная процедура линейна по длине ряда L (анализ по формулам (16)–(18) реализуется через один проход) и обеспечивает фиксацию только тех разрывов, чья длина строго превышает порог, рассчитанный с учётом агротехнологических требований.

2. Экспериментальные исследования и анализ результатов

2.1. Оценка качества сегментации

В рамках исследования была проведена серия экспериментов (1 – 6), охватывающих различные конфигурации моделей *YOLOv11* (nano, small, medium) для оценки качества сегментации посадочных рядов картофеля на тестовой выборке. Для каждой архитектуры сравнивались два режима обучения: базовый (без дополнительных аугментаций) и расширенный – с использованием *Mosaic* и *MixUp*.

Сравнительный анализ результатов (см. табл. 2) показал стабильное улучшение метрик сегментации при применении расширенных аугментаций во всех вариантах моделей. Наиболее заметный прирост был достигнут для конфигурации *YOLOv11m*, где максимальные значения *mAP@0.5* составили 0,944, а F1-score достиг 0,93. Полученные результаты подтверждают, что использование сложных аугментаций повышает адаптивность и устойчивость моделей к разнообразию условий съёмки, включая частичные перекрытия растений, изменения освещённости и наличие теней.

Табл. 2. Сравнение конфигураций и качества моделей для сегментации рядов

№ эксп.	модели	mAP @0.5	mAP @0.5:0.95	Precision	Recall	F1
1	YOLOv11n (no augm., freeze=5)	0,932	0,719	0,948	0,880	0,913
2	YOLOv11n (Mosaic+MixUp, freeze=5)	0,943	0,734	0,957	0,899	0,927
3	YOLOv11s (no augm., freeze=5)	0,939	0,728	0,979	0,882	0,928
4	YOLOv11s (Mosaic+MixUp, freeze=5)	0,943	0,739	0,971	0,896	0,932
5	YOLOv11m (no augm., freeze=5)	0,936	0,716	0,981	0,877	0,926
6	YOLOv11m (Mosaic+MixUp, freeze=5)	0,944	0,740	0,964	0,899	0,930
7	YOLOv11m (MixUp, freeze=5)	0,941	0,734	0,956	0,899	0,927
8	YOLOv11m (Mosaic, freeze=5)	0,943	0,715	0,979	0,892	0,933
9	YOLOv11m (Mosaic+MixUp, freeze=0)	0,939	0,736	0,974	0,887	0,928
10	YOLOv11m (Mosaic+MixUp, freeze=10)	0,936	0,704	0,984	0,877	0,927
11	DeepLabV3+ (resnet34)	0,928	0,701	0,971	0,865	0,913
12	U-Net (resnet34)	0,921	0,689	0,958	0,851	0,901

Для детального анализа выбранной архитектуры *YOLOv11m* была проведена дополнительная серия экспериментов (7 – 10), направленная на исследование влияния глубины «заморозки» начальных слоёв (freeze=0, 5, 10) и отдельных видов аугментаций на итоговое качество сегментации. Полученные данные показали, что оптимальным является значение freeze=5, а комбинация *Mosaic* и *MixUp* обеспечивает наилучший баланс между точностью, полнотой и вычислительными затратами.

Для сравнительного анализа были также протестированы архитектуры сетей сегментации – *DeepLabV3+* и *U-Net* с энкодером ResNet34 (эксперименты 11 – 12). Несмотря на хорошие показатели, обе модели продемонстрировали качество сегментации несколько ниже, чем оптимальная конфигурация *YOLOv11m*.

В целом, представленные результаты демонстрируют высокую воспроизводимость и стабильность работы алгоритма. Оценка, выполненная на тестовой части датасета, подтверждает способность модели обобщать выученные признаки на новые изображения картофеля, не встречавшиеся при обучении. Все конфигурации *YOLOv11* показывают *mAP@0.5* выше 0,93 и F1-score выше 0,91, что соответствует уровню современных исследований в данной области.

На рис. 4 представлен пример результатов сегментации аэрофотоснимка участка поля. Модель демонстрирует высокую точность пространственной локализации растений, однако на границах изображения возможны частично усечённые объекты и неполные сегментированные области. В интегрированном пайплайне данный эффект нивелируется посредством скользящей обработки тайлов с перекрытием и последующего объединения предсказанных масок. Такой механизм обеспечивает устранение краевых артефактов и формирование единой согласованной маски для всего ортофотоплана.

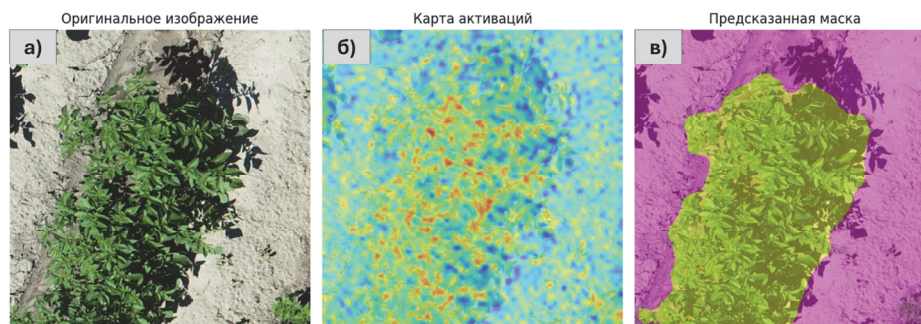


Рис. 4. Визуализация этапов сегментации: а) исходное изображение, б) карта активаций нейросети, в) предсказанная маска (зелёный – растения, фиолетовый – фон/почва)

Полученные сегментированные данные используются в качестве входной информации для последующих процедур анализа, включая автоматическое определение осей рядов и выявление пропусков, что подтверждает эффективность предложенного метода для мониторинга состояния посадок с применением аэрофотосъемки.

2.2. Оценка качества выделения рядов

Для проверки корректности автоматического выделения осей посадочных рядов была проведена количественная и визуальная валидация (см. рис. 5).

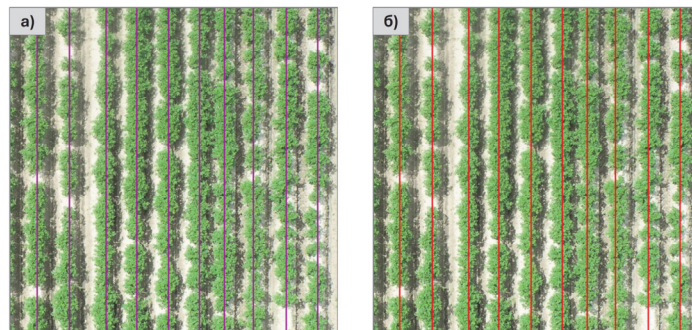


Рис. 5. Сравнение выделения осей рядов: а) результаты ручной разметки (фиолетовые линии); б) результаты автоматической разметки (красные линии)

В качестве основных метрик использовались: среднее отклонение угла между автоматически определённой и эталонной (ручной) осью ряда, среднее смещение оси (в пикселях и сантиметрах), а также доля корректно распознанных рядов.

На тестовой выборке с ручной разметкой получены следующие значения: среднее угловое отклонение между автоматически найденной и эталонной осью ряда составило $1,8^\circ$, ошибка положения оси – 6,2 пикселя (0,9 см), а доля совпадений рядов достигла 97,4 %.

Сравнение с экспертной разметкой показало, что предложенный алгоритм точно определяет положения рядов даже при сложной структуре посевов и локальных дефектах. В частности, среднее абсолютное отклонение координат рядов на автоматической маске относительно ручной оценки составило всего 1,8 пикселя, что демонстрирует высокую надёжность метода.

2.3. Оценка качества определения пропусков

Для оценки точности автоматического выявления пропусков вдоль осей рядов была проведена количественная и визуальная валидация на независимой тестовой выборке с ручной экспертной разметкой. Качество работы метода измерялось по метрикам Precision, Recall и F1-score. Итоговые значения составили: Precision – 0,93, Recall – 0,89 и F1-score – 0,91.

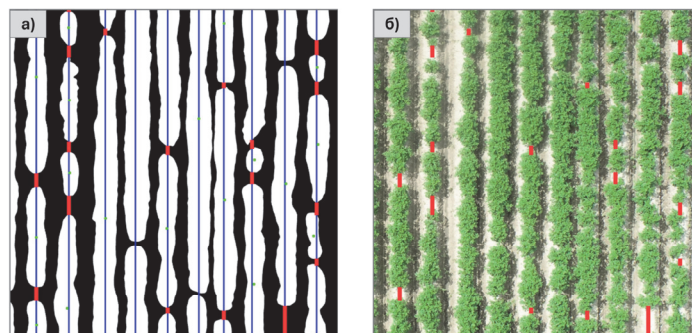


Рис. 6. Визуализация обнаружения пропусков в рядах картофеля: а) автоматическое выделение пропусков (красные прямоугольники) на бинарной маске сегментированных рядов; б) наложение найденных пропусков на исходный аэрофотоснимок участка

Полученные результаты демонстрируют высокую эффективность подхода на основной площади посевов. Однако анализ ошибок показал, что снижение метрик (в частности, Recall) коррелирует с наличием на снимках участков, к которым относятся:

- 1) зоны локального затенения (собственные тени растений и тени от соседних кустов при смыкании рядов), где снижается контрастность границ объектов;
- 2) технологические колеи и края поля с нарушенной геометрией посадки;
- 3) участки с высокой засоренностью, где сорняки могут ошибочно интерпретироваться как культурные растения.

Примеры автоматического определения пропусков вдоль осей рядов приведены на рис. 6: пропуски отображаются на бинарной маске и исходном изображении цветowymi метками, что иллюстрирует корректную работу алгоритма на типичных участках поля.

Основным направлением дальнейшего развития метода является расширение обучающей выборки сценами с разнообразным агрофоном (вариативное освещение, засорённость, нарушения геометрии рядов) для повышения устойчивости к указанным типам сложных участков.

3. Схема функционирования метода

Предлагаемый подход к автоматизации идентификации рядов картофеля и обнаружению пропусков по данным аэрофотосъёмки с беспилотных летательных аппаратов включает последовательность этапов обработки, представленных на рис. 7. Каждый этап направлен на поэтапное преобразование исходных изображений в аналитически интерпретируемые результаты, обеспечивающие количественную оценку структуры посадок.



Рис. 7. Структурная схема алгоритма автоматической идентификации рядов картофеля и выявления пропусков

На первом этапе выполняется генерация ортомозаики, включающая съёмку исследуемого участка с БПЛА и построение ортофотоплана с геопривязкой, обеспечивающего высокую точность пространственных измерений. Далее реализуется разбиение ортофотоплана на фрагменты с частичным перекрытием с использованием подхода SAHI (Sliced Aided Hyper Inference) [18], позволяющего эффективно обрабатывать крупноформатные изображения за счёт их деления на перекрывающиеся тайлы и последующего объединения результатов инференса.

На следующем этапе выполняется сегментация растительного покрова и объединение масок: для каждого тайла формируется бинарная маска с использованием модели *YOLOv11-Seg*, после чего маски интегрируются в единую карту поля. По совмещённой маске определяется доминирующее направление рядов методом главных компонент, и изображение поворачивается для их выравнивания вдоль оси ординат. В повернутой системе координат выполняется адаптивная морфологическая фильтрация с прямоугольным структурным элементом, ориентированным поперёк рядов, что обеспечивает разрыв перемычек и удаление мелких артефактов при сохранении линейной структуры посадок.

Затем выполняется извлечение и кластеризация осей рядов: по очищенной маске определяются центры, которые объединяются алгоритмом DBSCAN в непрерывные линии, соответствующие реальным посадкам. На заключительном этапе производится детекция пропусков и расчёт метрик качества рядов.

Таким образом, алгоритм обеспечивает полный цикл анализа – от сегментации аэрофотоснимков и морфологической очистки до выделения осей рядов и определения пропусков в посадках.

Заключение

Результаты экспериментов показали, что применение современных моделей сегментации в сочетании с адаптивными фильтрами позволяет достичь высокой точности выделения осей рядов (97,4 %) и обнаружения пропусков (F1-score до 0,91).

Сравнительный анализ показал, что модель *YOLOv11m* обеспечивает оптимальный баланс производительности: достигая метрики *mAP@0.5* на уровне 0,944, она существенно превосходит архитектуры *DeepLabV3+* и *U-Net* по скорости обработки. Ограничения метода в основном проявляются на локальных участках с нарушенной структурой изображения, где возможны ошибки детекции. Предложенный подход может быть адаптирован для анализа других рядовых культур (кукуруза, свёкла, соя) при минимальной дополнительной настройке.

Практическая значимость работы заключается в применимости предложенного алгоритма в составе цифровых систем агромониторинга. Высокая скорость инференса модели *YOLOv11m* (в эксперименте обработка

массива тайлов для поля 8,5 га заняла 6–7 минут) обеспечивает возможность оперативного анализа производственных площадей в день облёта. Масштабируемость и универсальность подхода делают его эффективным инструментом точного земледелия.

References

- [1] Zuolkernan I, Abuhani DA, Hussain MH, Khan J, ElMohandes M. Machine learning for precision agriculture using imagery from unmanned aerial vehicles (UAVs): a survey. *Drones* 2023; 7(6): 382. doi:10.3390/drones7060382.
- [2] Li G, Le F, Si S, Cui L, Xue X. Image segmentation-based oilseed rape row detection for infield navigation of agri-robot. *Agronomy* 2024; 14(9): 1886. doi:10.3390/agronomy14091886.
- [3] Ronchetti G, Mayer A, Facchi A, Ortuani B, Sona G. Crop row detection through UAV surveys to optimize on-farm irrigation management. *Remote Sens* 2020; 12(12): 1967. doi:10.3390/rs12121967.
- [4] Deng G, Zhou F, Dong H, Xu Z, Li Y. Accurate sugarcane detection and row fitting using SugarRow-YOLO and clustering-based spline methods for autonomous agricultural operations. *Appl Sci* 2025; 15(14): 7789. doi:10.3390/app15147789.
- [5] Khan MN, Rahi A, Rajendran VP, Al Hasan M, Anwar S. Real-time crop row detection using computer vision: application in agricultural robots. *Front Artif Intell* 2024; 7: 1435686. doi:10.3389/frai.2024.1435686.
- [6] Ferreira R, Lima AM, da Silva DL, Arias-Garcia J, Yudi J, Hernandez AC. Data labelling and heuristics to estimate sugarcane rows orientation for UAV-aided precision agriculture. *Smart Agric Technol* 2025; 10: 101425. doi:10.1016/j.atech.2025.101425.
- [7] Chen P, Ma X, Wang F, Li J. New method for crop row detection using unmanned aerial vehicle images. *Remote Sens* 2021; 13(17): 3526. doi:10.3390/rs13173526.
- [8] Mhango JK, Harris EW, Green R, Monaghan JM. Mapping potato plant density variation using aerial imagery and deep learning techniques for precision agriculture. *Remote Sens* 2021; 13(14): 2705. doi:10.3390/rs13142705.
- [9] Gao J, Tan F, Cui J, Hou Z. Unmanned aerial vehicle image detection of maize-YOLOv8n seedling leakage. *Front Plant Sci* 2025; 16: 1569229. doi:10.3389/fpls.2025.1569229. [web:354]
- [10] de Souza FLP, Dias MA, Setiyono TD, Campos S, Shiratsuchi LS, Tao H. Identification of soybean planting gaps using machine learning. *Smart Agric Technol* 2025; 10: 100779. doi:10.1016/j.atech.2025.100779.
- [11] Rocha BM, da Fonseca AU, Pedrini H, Soares F. Automatic detection and evaluation of sugarcane planting rows in aerial images. *Inf Process Agric* 2023; 10(3): 400-415. doi:10.1016/j.inpa.2022.04.003.
- [12] Roboflow – platform for computer vision datasets and models. Available at: <https://roboflow.com/>.
- [13] Khanam R, Hussain M. YOLOv11: an overview of the key architectural enhancements. *arXiv* 2024; arXiv:2410.17725. Available at: <https://arxiv.org/abs/2410.17725>.
- [14] Pestunov IA, Kalashnikov RA, Ovcharova NV, Bayramov RR. Otsenka polevoy vskhozhesti kukuruzy po RGB-izobrazheniyam s BPLA s ispolzovaniem metodov mashinnogo obucheniya [Field emergence assessment of maize using UAV RGB imagery and machine learning methods, in Russian]. In: *Regionalnye problemy dstantsionnogo zondirovaniya Zemli: materialy XI Mezhdunarodnoy nauchnoy konferentsii*; 2024; Krasnoyarsk, Russia. Krasnoyarsk: Sibirsky federalny universitet; 2024. pp. 334-339.
- [15] Abdi H, Williams LJ. Principal component analysis. *Wiley Interdiscip Rev Comput Stat* 2010; 2(4): 433-459. doi:10.1002/wics.101.
- [16] Alt VV, Pestunov IA, Melnikov PV, Yolkin OV. Avtomatizirovannoe obnaruzhenie sornyakov i otsenka kachestva vskhodov sel'skokhozyaystvennykh kultur po RGB-izobrazheniyam [Automated weed detection and assessment of crop emergence quality from RGB images, in Russian]. *Sib Vestn Selskokhozyaystvennoy Nauki* 2019; 48(5): 52-60. doi:10.26898/0370-8799-2018-5-7.
- [17] Schubert E, Sander J, Ester M, Kriegel HP, Xu X. DBSCAN revisited, revisited: why and how you should (still) use DBSCAN. *ACM Trans Database Syst* 2017; 42(3): 19. doi:10.1145/3068335.
- [18] Akyon FC, Altinuc SO, Temizel A. Slicing aided hyper inference and fine-tuning for small object detection. In: *Proc 2022 IEEE Int Conf Image Process (ICIP)*; 2022. pp. 966-970. doi:10.1109/ICIP46576.2022.9897990.

Сведения об авторах

Полещенко Дмитрий Александрович, 1980 года рождения, кандидат технических наук, заведующий кафедрой автоматизации и информационных технологий им. Ю.И. Еременко СТИ НИТУ «МИСИС». Сфера научных интересов: машинное зрение, нейронные сети, анализ виброакустических сигналов. E-mail: po-dima@yandex.ru ORCID: 0000-0003-3544-4296

Михайлов Илья Сергеевич, 2001 года рождения, аспирант кафедры автоматизации и информационных технологий им. Ю.И. Еременко СТИ НИТУ «МИСИС». Область научных интересов: компьютерное зрение, машинное обучение. E-mail: mikhaylov.is@yandex.ru ORCID: 0000-0002-1292-4871

Поступила в редакцию 27 октября 2025 г.



Окончательный вариант – 15 декабря 2025 г.

Development of a machine-learning algorithm for the automatic determination of the number of missing potato plants from UAV imagery

D.A. Poleshchenko¹, I.S. Mikhailov¹

¹ Sary Oskol Technological Institute n.a. A.A. Ugarov (branch)

National University of Science and Technology MISIS, 309516, Sary Oskol, Russia, Makarenko Microdistrict 42

Abstract

This paper presents a method for automatic identification of potato rows and detection of missing plants from UAV aerial imagery. The approach combines plant segmentation using YOLOv11 models, adaptive morphological filtering that accounts for spatial resolution, and automatic correction of row orientation. Experimental evaluation on field data demonstrated high row extraction accuracy (97.4 %) and effective gap detection (F1-score 0.91). The obtained results confirm the correct identification of crop structure, justifying the applicability of the approach for decision support in digital agriculture.

Keywords: crop monitoring, potato, plant segmentation, orthomosaic, YOLOv11, Ground Sample Distance, remote sensing, UAV.

Citation: Poleshchenko DA, Mikhailov IS. Development of a machine-learning algorithm for the automatic determination of the number of missing potato plants from UAV imagery. *Computer Optics* 2026; 50(4): 1875. doi: 10.18287/COJ1875.

Author's information

Dmitry Alexandrovich Poleshchenko, (b. 1980), Ph.D. (Eng.), Head of the Department of Automation and Information Technologies named after Yu.I. Eremenko, STI NUST "MISIS". Research interests: computer vision, neural networks, vibroacoustic signal analysis. E-mail: po-dima@yandex.ru ORCID: 0000-0003-3544-4296

Ilya Sergeevich Mikhailov, (b. 2001), Ph.D. student at the Department of Automation and Information Technologies named after Yu.I. Eremenko, STI NUST "MISIS". Research interests: computer vision, machine learning. E-mail: mikhaylov.is@yandex.ru ORCID: 0000-0002-1292-4871

Received October 27, 2025. The final version – December 15, 2025.
