

# Гибридный алгоритм автоматического подсчета растений кукурузы по RGB-изображениям с БПЛА на основе комбинации методов компьютерного зрения и машинного обучения

*Пестунов И.А.<sup>1</sup>, Калашников Р.А.<sup>1</sup>, Овчарова Н.В.<sup>2</sup>*

<sup>1</sup>Федеральный исследовательский центр информационных и вычислительных технологий (Новосибирск)

<sup>2</sup>Алтайский государственный университет (Барнаул)



# Актуальность

Современное сельское хозяйство активно внедряет цифровые технологии, включая методы компьютерного зрения и машинного обучения, для повышения эффективности производства.

Густота стояния растений – ключевой агротехнический параметр, напрямую влияющий на урожайность и эффективность технологий точного земледелия.

Традиционные методы оценки (визуальный подсчет на пробных площадках) обладают существенными недостатками: высокая трудоемкость, субъективность оценки и низкая репрезентативность данных.

Анализ RGB-изображений с БПЛА позволяет автоматизировать подсчет растений, снизить трудозатраты, повысить точность и воспроизводимость результатов, масштабировать оценку на большие площади.

Существующие подходы имеют ограничения: классические алгоритмы компьютерного зрения чувствительны к условиям съемки и засоренности полей, традиционные модели ML требуют ручного проектирования признаков, а нейросетевые методы нуждаются в больших размеченных выборках и ресурсах.

Гибридные алгоритмы, сочетающие сильные стороны разных подходов, обеспечивают высокую точность даже при засоренности полей и перекрытии растений, устойчивость к вариативным условиям съемки и практическую значимость для цифрового и точного земледелия.

# Объекты исследования

Исследования проводились в 2022–2025 гг. на сельскохозяйственных землях Алейского и Поспелихинского районов Алтайского края. Объектами служили два опытных поля площадью по 100 га каждое, возделываемые по двум разным агротехнологиям: традиционной плоскорезной обработке и системе «No-Till».



(1)ООО “Золотая осень”  
(2)СПК “Знамя Родины”

# Исходные данные

Исходными данными для обработки послужили 3000 RGB-изображений размером 5472×3648 пикселей с пространственным разрешением 0.3 см/пиксель, полученные с БПЛА.



Агроценоз кукурузы по агротехнике «Классическая плоскорезная обработка» (Алейский район)



Агроценоз кукурузы по агротехнике «No-Till» (Поспелихинский район)

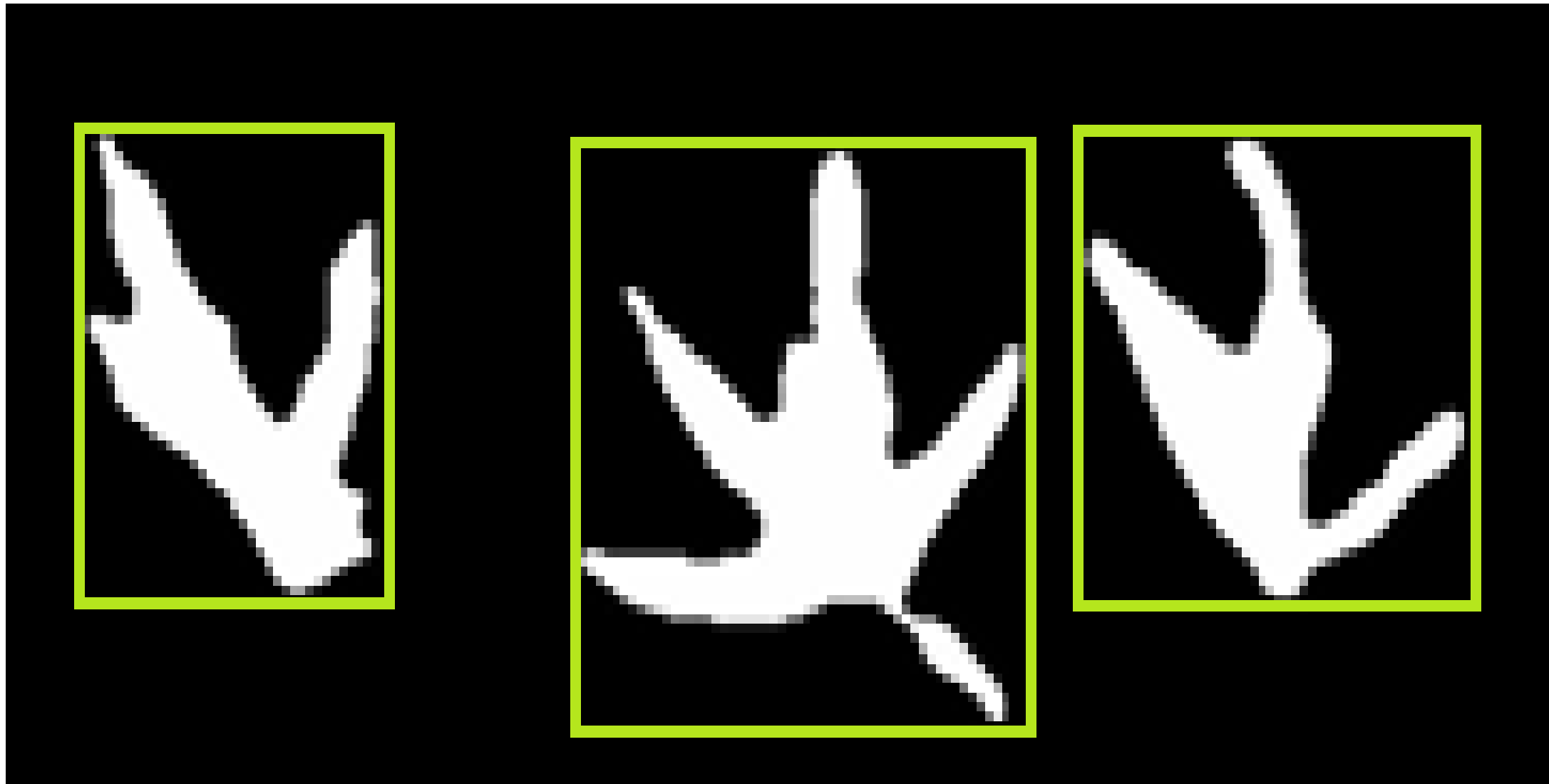
# Примеры фрагментов исходных изображений



# Примеры фрагментов исходных изображений



# Проблематика подсчета количества растений



Несколько одиночных сегментов кукурузы

# Проблематика подсчета количества растений



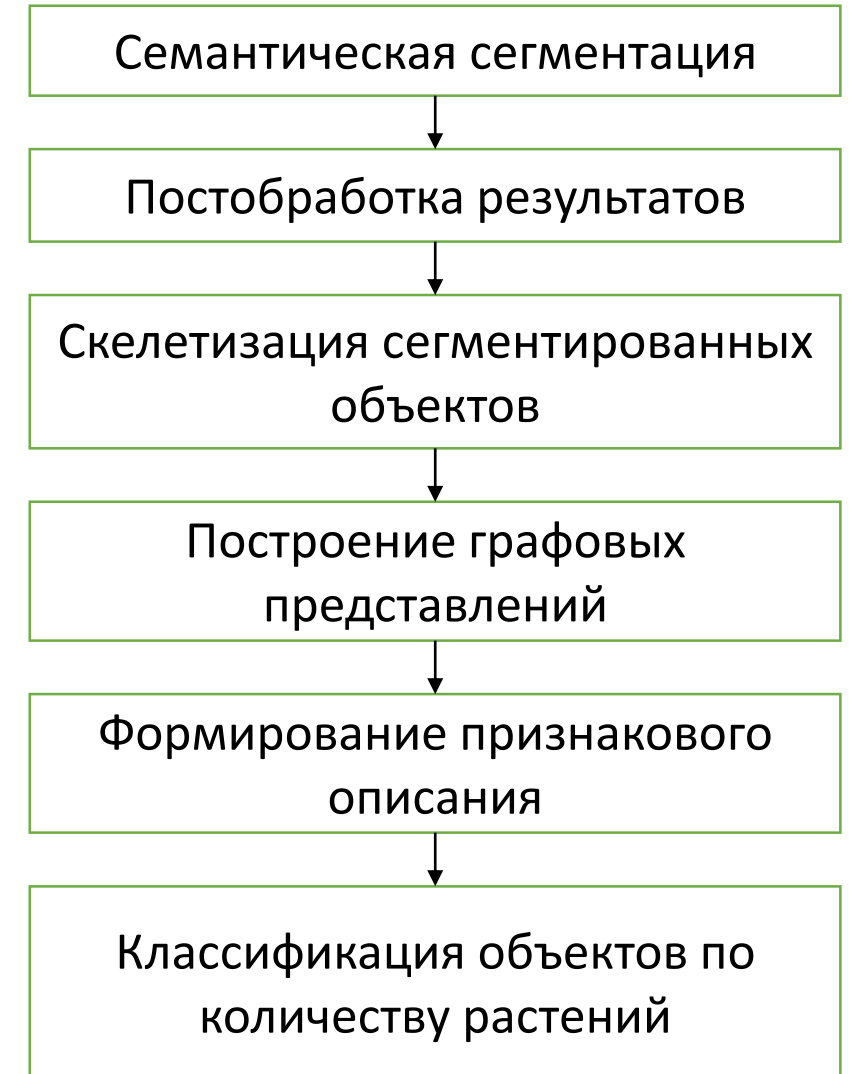
Несколько «слипшихся» сегментов кукурузы

# Возможные подходы к решению задачи

Подход к решению задачи	Краткое описание методов
Алгоритмы компьютерного зрения	Пороговая сегментация и бинаризация, расчёт вегетационных индексов (на RGB), морфологические операции (эрозия/наращивание, открытие/заккрытие), анализ формы и структуры объекта, скелетизация, поиск контуров и ключевых точек. Подход хорошо работает при простых условиях съёмки и высоком контрасте «растение–фон».
Классическое обучение	Ручное извлечение признаков (геометрия масок, контуры, текстуры, признаки скелетных графов, метрики плотности/вытянутости), последующая классификация/регрессия алгоритмами SVM, Random Forest, градиентный бустинг, логистическая/линейная регрессия. Подход обеспечивает интерпретируемость и гибкость, но сильно зависит от качества признаков и плохо переносится на новые условия съёмки.
Глубокое обучение	Сверточные нейросети для сегментации и детекции. Автоматически извлекают многоуровневые признаки, обеспечивая высокую точность на сложных сценах (засорённость, перекрытия), но требуют большого объема размеченных данных и вычислительных ресурсов.
Гибридный подход	Комбинация сильных сторон всех подходов: нейросетевая семантическая сегментация → постобработка (фильтрация шумов, морфология) → скелетизация и построение графовых представлений → извлечение дескрипторов (длины рёбер, ветвления, отношения «скелет/контур», плотность, геометрия bounding boxes и др.) → классификация объектов по количеству растений (SVM/RF/бустинг). Обеспечивает устойчивость к засорённости полей и перекрытиям контуров растений при высокой точности подсчёта.

# Общие этапы гибридного подхода подсчетов экземпляров кукурузы

- 1. Семантическая сегментация RGB-изображений** для выделения масок всходов кукурузы с применением сверточных нейронных сетей, а также постобработка результатов (удаление шумовых объектов, фильтрация по контуру).
- 2. Скелетизация сегментированных объектов и построение графовых представлений**, где каждому скелету соответствует индивидуальный или сгруппированный всход.
- 3. Формирование признакового описания:** вычисление геометрических и топологических характеристик объектов на основе масок и их скелетных графов (длины рёбер, ветвления, соотношения «скелет/контур», плотность, параметры ограничивающих рамок и др.).
- 4. Классификация объектов по количеству растений** с использованием методов машинного обучения (SVM, Random Forest, градиентный бустинг) на основе сформированных признаков.



# Разметка обучающей выборки

Для обучения  
нейронных сетей  
производилась  
полуавтоматическая  
разметка исходных  
изображений



Пример фрагмента размеченного  
изображения

# Формирование выборки для обучения сверточных нейронных сетей

После разметки исходные изображения нарезались на фрагменты 512x512 пикселей. В результате было получено:

- 2854 обучающих фрагментов,
- 500 проверочных фрагментов,
- 400 тестовых фрагментов.



Пример фрагментов изображений и их масок

# Семантическая сегментация

Используется для выделения точных границ всходов кукурузы на фоне почвы.

Тестировалось 3 архитектуры для семантической сегментации: **U-Net**, **DeepLabV3** и **DeepLabV3+**.

Мы также исследовали более сложные подходы: **instance segmentation** и **object detection**, включая современные архитектуры на основе трансформеров:

**Mask R-CNN** – классическая модель для инстанс-сегментации;

**YOLOv8 & YOLOv11** – современные детекторы;

**SegFormer** – трансформерная архитектура для семантической сегментации;

**Mask2Former** – универсальная трансформерная архитектура для различных задач сегментации.

Однако в наших условиях (высокая засорённость, перекрытия растений, низкий контраст с фоном, немногочисленная обучающая выборка) качество распознавания оказалось недостаточным.



Фрагменты исходного изображения

Результат семантической сегментации

# Результаты работы моделей

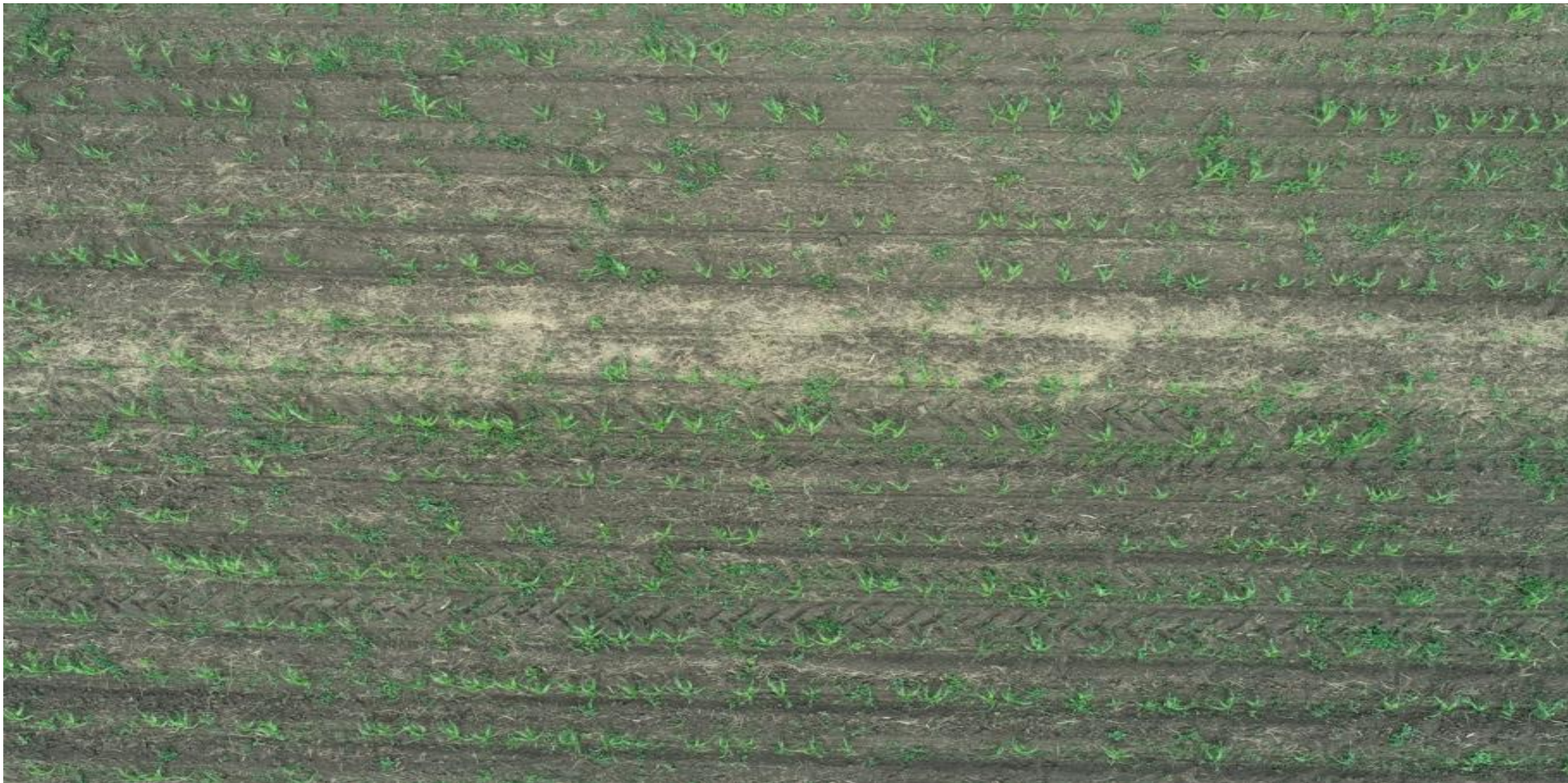
## Значения метрик на обучающей выборке

Модель	Accuracy	Precision	Recall	F1-score	IoU
U-Net	0.9888	0.7999	<b>0.9367</b>	0.8629	0.7589
DeepLabV3	0.9856	0.7692	0.8803	0.8210	0.6964
DeepLabV3+	<b>0.9902</b>	<b>0.8337</b>	0.9228	<b>0.8760</b>	<b>0.7794</b>

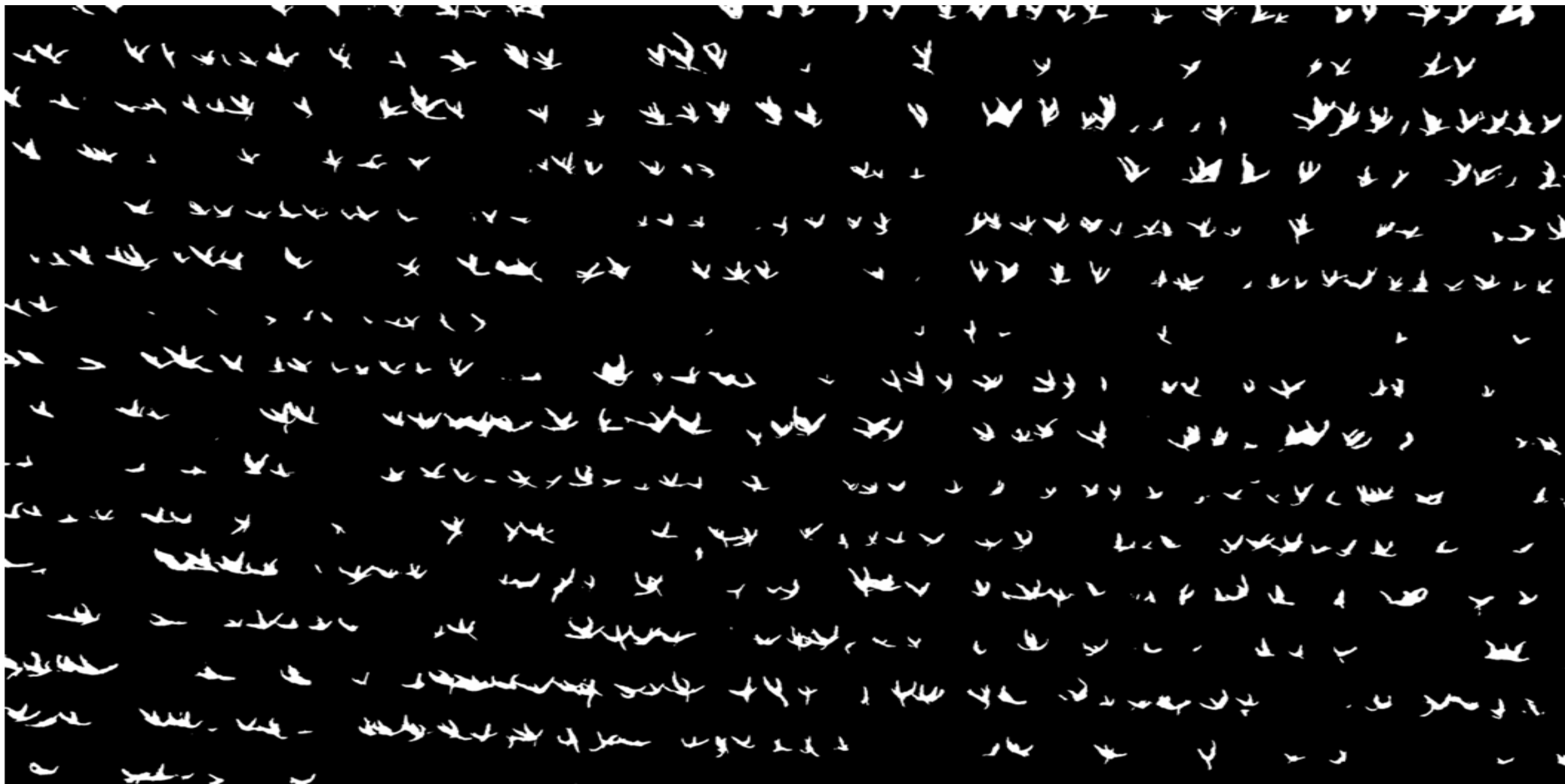
## Значения метрик на тестовой выборке

Модель	Accuracy	Precision	Recall	F1-score	IoU
U-Net	0.9903	0.7356	0.9001	0.8096	0.6801
DeepLabV3	0.9871	0.6846	0.8110	0.7425	0.5904
DeepLabV3+	<b>0.9911</b>	<b>0.7658</b>	<b>0.8828</b>	<b>0.8201</b>	<b>0.6951</b>

# Исходное изображение



# Результат семантической сегментации (DeepLabV3+)



# Предсказание моделей СНС на фрагментах изображений



**Фрагмент  
изображения**



**Маска, полученная  
экспертом**



**U-Net**



**DeepLabV3**



**DeepLabV3+**

# Графовое представление результатов семантической сегментации

Для получения информативных топологических признаков после этапа семантической сегментации применяются **алгоритмы скелетизации Чжана и Ли**, формирующие тонкие линии (скелеты), топологически эквивалентные исходным объектам.

Каждый скелет растения кукурузы можно представить в виде графа:

$$G_{\text{skeleton}} = \langle V, E \rangle, \quad V = V_0 \cup V_1, \quad E = E_0 \cup E_1 \cup E_2 \cup E_3,$$

где  $V_0$  — вершины графа, соответствующие точкам пересечения скелета,

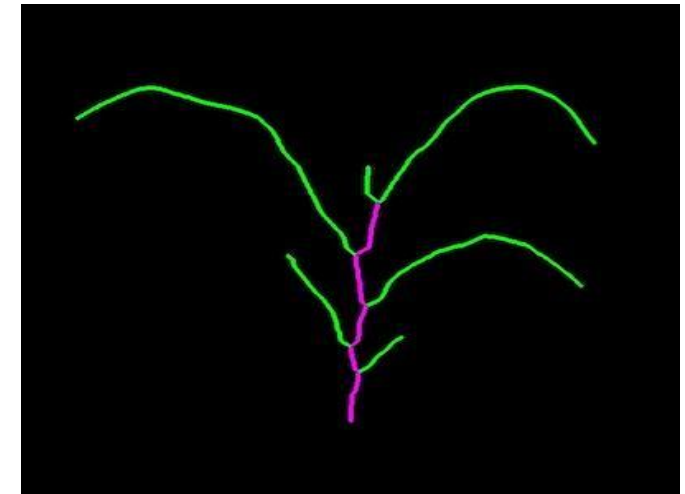
$V_1$  — вершины графа, соответствующие концевым точкам скелета,

$E_0$  — множество изолированных ребер,

$E_1$  — множество концевых ребер,

$E_2$  — ребра, соединяющие точки пересечения скелета (ребра-перекрестки),

$E_3$  — множество изолированных ребер (циклы).



Скелет всхода кукурузы:  
зеленым цветом выделено множество концевых ребер ( $E_1$ ),  
фиолетовым — множество ребер-перекрестков ( $E_2$ )

# Примеры построения скелета



**Исходное  
изображение**



**Метод Чжана**



**Метод Ли**

1. Lee T.-C., Kashyap R.L., Chu C.-N. Building skeleton models via 3-D medial surface/axis thinning algorithms // Computer Vision, Graphics, and Image Processing, 1994, 56(6), 462-478.

2. Zhang T. Y., Suen C. Y. A fast parallel algorithm for thinning digital patterns // Communications of the ACM, 1984, 27(3), 236-239.

# Формирование геометрических признаков

Признаки	Описание
Ширина	Расстояние по горизонтали ограничивающего сегмента
Высота	Расстояние по вертикали ограничивающего сегмента
Площадь объекта	Количество пикселей внутри контура объекта
Периметр контура	Длина внешней границы объекта
Отношение длины скелета к длине контура	Показывает, насколько объект структурирован
Отношение площади объекта к площади сегмента	Характеризует плотность заполнения сегмента объектом
Соотношение сторон сегмента	<b>MIN</b> ( $w/h$ , $h/w$ ) - показывает вытянутость сегмента

# Визуализация геометрических признаков



Ширина



Высота



Площадь объекта



Периметр контура



Отношение площади объекта к площади сегмента



Минимальное отношение между длиной и высотой

# Формирование топологических признаков на основе графовых представлений

Признаки	Описание
Branch_distance	Длина рёбер основного ствола кукурузы
Branch_distance_1	Длина рёбер-перекрёстков
Branch_distance_2	Длина рёбер - концевых
Euclidean_distance	Евклидово расстояние от концов ребра скелета
N_edges (degree = 1)	Количество концевых рёбер
N_edges (degree $\geq 3$ )	Количество рёбер-перекрёстков
Sum_edges	Сумма всех длин рёбер
AVG, MIN, MAX	Статистики от признаков выше

# Отбор наиболее информативных признаков

Избыточные или скоррелированные признаки ухудшают обобщающую способность моделей, увеличивают сложность и время обучения.

**Цель отбора:** сократить размерность признакового пространства, удалить шум и выделить признаки, реально влияющие на предсказания.

Признаковое пространство представлено различными геометрическими и топологическими признаками. **Всего 44 признака.**

## **Методы отбора:**

- Прямой отбор признаков – начиная с пустого множества, признаки добавляются по одному, пока метрика качества (RMSE) улучшается.
- Обратный отбор признаков – начиная со всех признаков, поочередно удаляются наименее значимые, пока качество модели не начнёт падать.
- Дополнительно были учтены признаки, которые показали высокую значимость и устойчивость в наших предыдущих исследованиях.

# Результаты отбора признаков и классификации

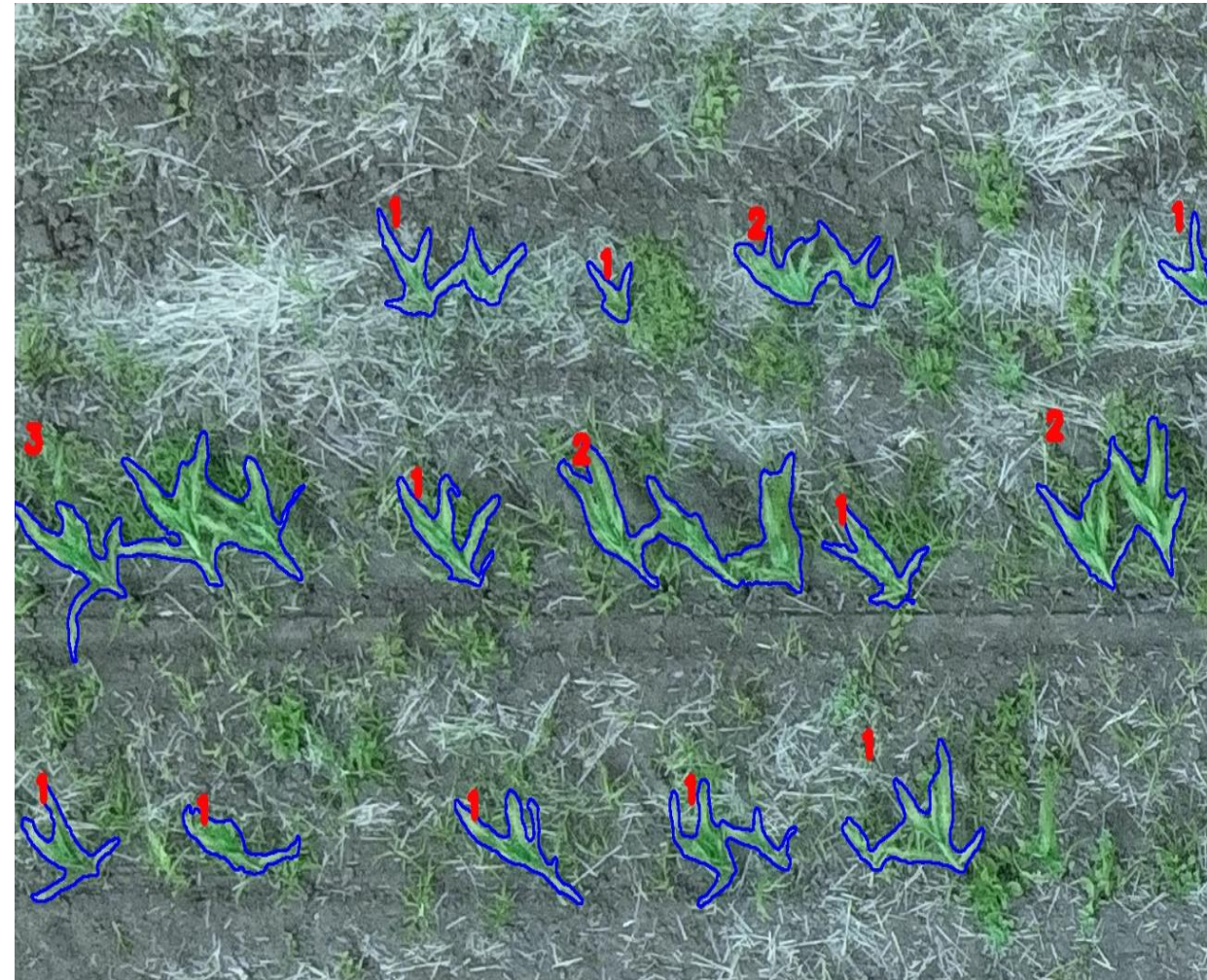
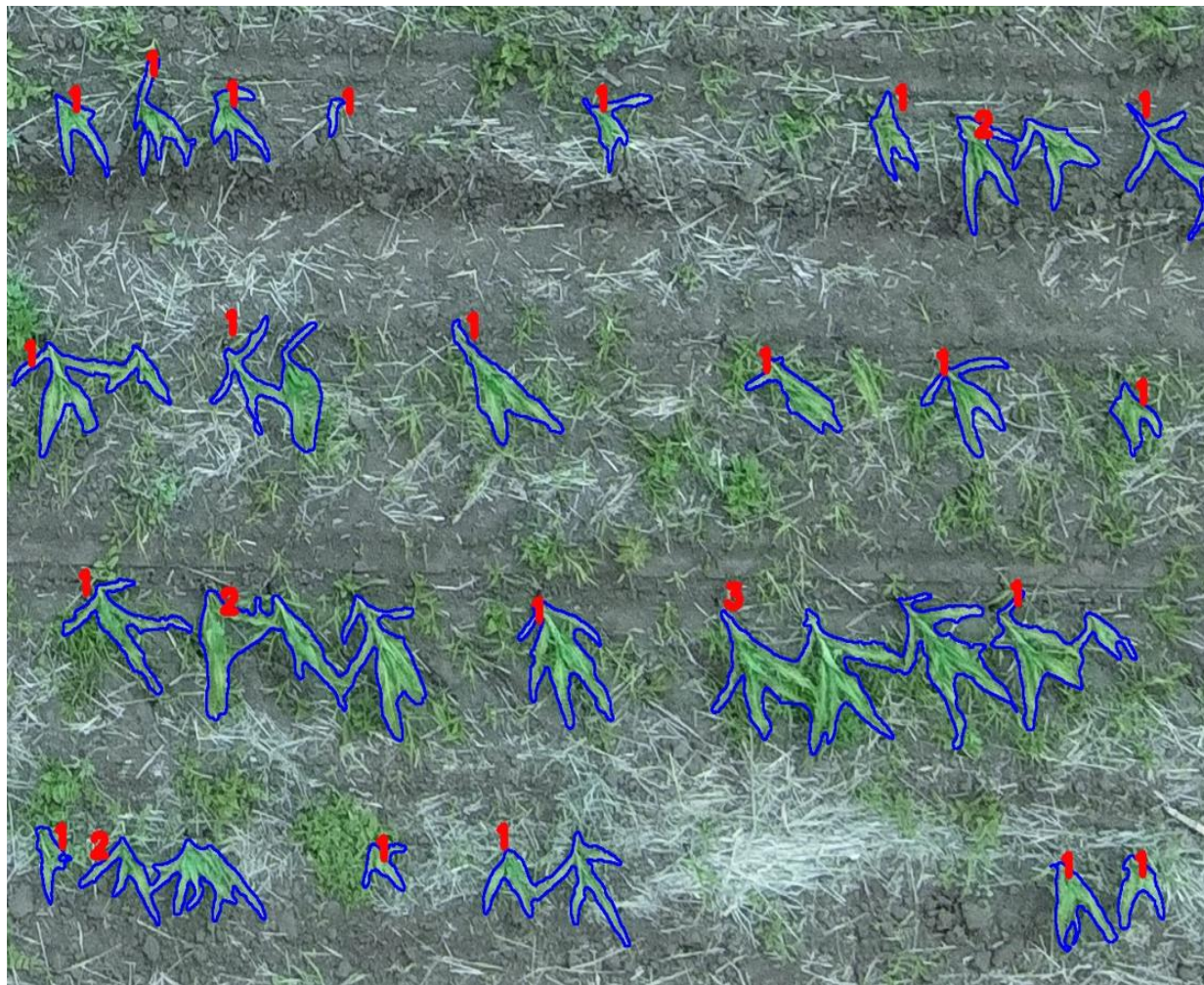
Название признака (в порядке уменьшения значимости)	Значение признака
Length	Длина контура сегмента
Branch_distance_max	Максимальная длина ребер основного ствола кукурузы
Junction_to_endpoint	Длина пути от точки ветвления (junction) до ближайшей конечной точки (endpoint)
Sum_edges	Сумма длин всех рёбер
Branch_distance_sum_1	Сумма длин всех рёбер-перекрестков
Branch_distance_sum	Сумма длин ребер основного ствола кукурузы
Euclidean_distance_max_1	Максимальное евклидово расстояние от концов ребра скелета
Branch_distance_sum_2	Сумма длин всех концевых рёбер
Branch_distance_max_2	Максимальная длина концевых рёбер
Euclidean_distance_sum_1	Максимальное евклидово расстояние от концов ребра скелета

На заключительном этапе были сравнены три модели машинного обучения, обученные на признаках, извлечённых из разных методик отбора признаков.

Модель	Точность
SVM	97,144%
Random Forest	97,128%
Linear Regression	96,601%
AdaBoost	95,668%

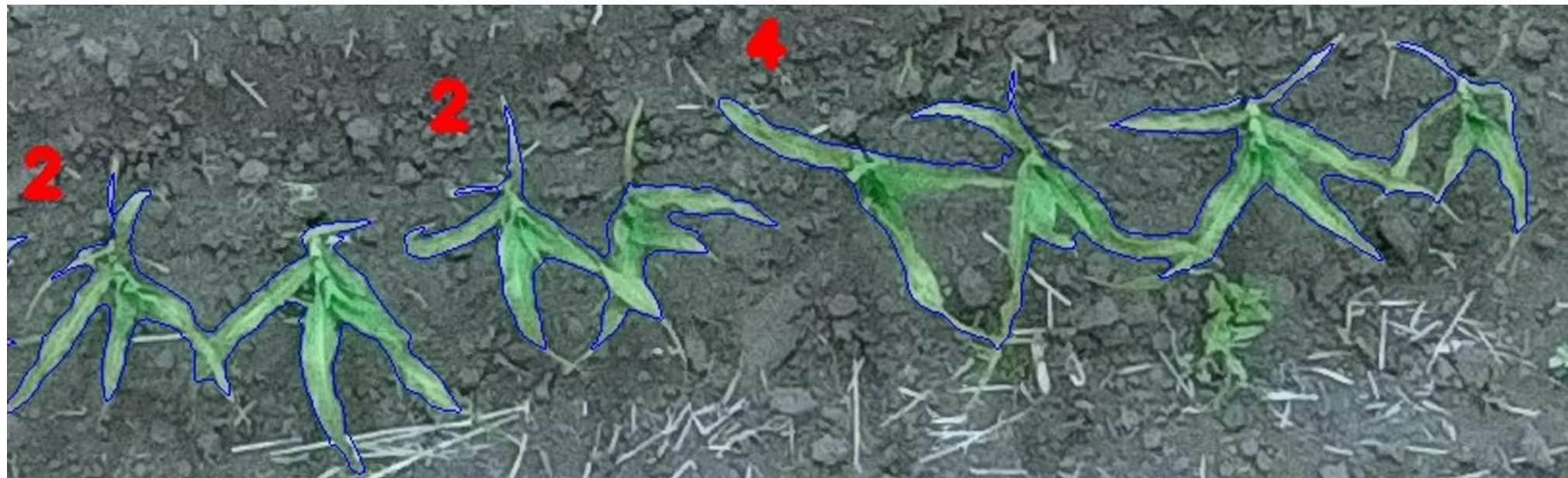
Результаты оценки работы алгоритмов на тестовой выборке

# Примеры результатов метода подсчета всходов



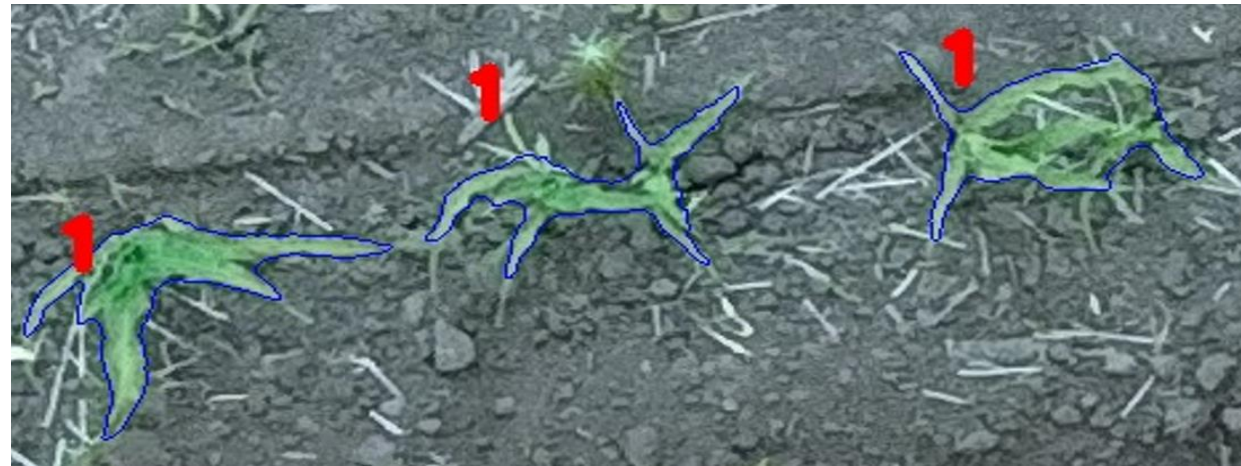
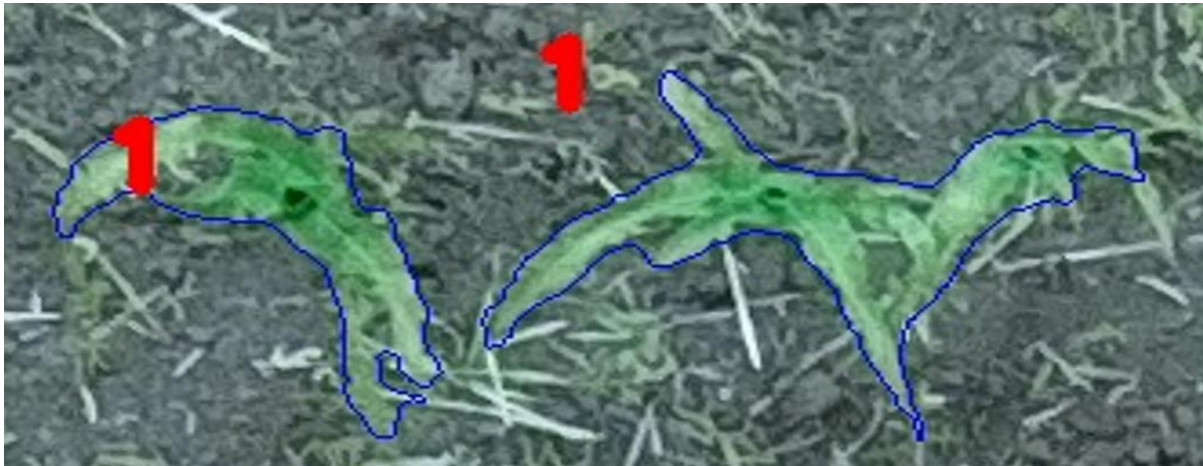
Результаты подсчета всходов растений кукурузы

# Примеры результатов метода подсчета всходов



Результаты подсчета всходов растений кукурузы (в приближении)

# Примеры результатов метода подсчета всходов



Результаты подсчета всходов растений кукурузы (в приближении)

# Заключение

Разработан гибридный алгоритм автоматического подсчёта растений кукурузы, основанный на интеграции методов компьютерного зрения и машинного обучения.

Для семантической сегментации использованы архитектуры U-Net, DeepLabV3 и DeepLabV3+; наилучшие результаты продемонстрировала модель DeepLabV3+.

Сформирован набор геометрических и топологических (графовых) признаков (в общей сложности 44 признака), проведён отбор наиболее значимых.

Наибольшую точность классификации (свыше 97%) показали алгоритмы SVM и Random Forest.

Предложенный метод сохраняет высокую устойчивость к засорённости и частичному перекрытию контуров растений, что подтверждает его практическую применимость в задачах точного земледелия.

Спасибо за внимание!