

М.В. Авдеенко¹, Н.В. Мутовина¹

¹Карагандинский технический университет имени Абылкаса Сагинова, Караганда, Казахстан
(email: avdeenkomaxim123@gmail.com)

Интеграция мультиспектральных данных в модели семантической сегментации сельскохозяйственных угодий

Развитие точного земледелия требует автоматизированного мониторинга посевов на основе мультиспектральных данных. Исследование систематически оценивает влияние спектральных конфигураций RGB и RGB+NIR на точность семантической сегментации сельхозугодий при использовании архитектур U-Net и DeepLabV3+ с энкодерами MobileNet_v2, ResNet18, ResNet34, ResNet50 на датасете Agriculture-Vision-2021. Модели сопоставлялись по Intersection over Union и F1-score для девяти классов на аннотированных снимках полей кукурузы, пшеницы, подсолнечника. Полученные результаты формируют стратифицированные рекомендации по выбору сенсорных конфигураций для хозяйств, оптимизируя компромисс между точностью мониторинга и затратами на оборудование.

Ключевые слова: Семантическая сегментация, мультиспектральные данные, точное земледелие, дистанционное зондирование, сельскохозяйственные угодья, U-Net, DeepLabV3+, MobileNet_v2, ResNet, ближний инфракрасный диапазон, Agriculture-Vision, мониторинг посевов, нейронные сети глубокого обучения.

Введение

Современное сельское хозяйство сталкивается с беспрецедентными вызовами, обусловленными ростом численности населения, изменением климатических условий и необходимостью устойчивого управления ресурсами. По прогнозам ФАО, к 2050 году мировое производство продовольствия должно увеличиться на 70 % при сокращении экологического следа отрасли [1]. В этом контексте технологии точного земледелия, интегрирующие методы дистанционного зондирования, компьютерного зрения и глубокого обучения, становятся критически важным инструментом оптимизации агрономических практик. Семантическая сегментация мультиспектральных изображений - задача попиксельной классификации объектов сельскохозяйственных сцен - позволяет автоматизировать мониторинг состояния посевов, дифференцированно оценивать потребности культур в питательных веществах и прогнозировать урожайность с пространственным разрешением, недостижимым традиционными методами [2; 3].

Последние пять лет отмечены интенсивным развитием методов машинного и глубокого обучения в агрономическом домене. Систематический обзор Veeragandham и Santhi продемонстрировал доминирование сверточных нейронных сетей (CNN) в задачах анализа сельскохозяйственных изображений, подчеркнув их превосходство над классическими алгоритмами в условиях высокой вариативности сцен [1]. Odounfa с соавторами сконцентрировались на архитектурах семантической сегментации - Mask R-CNN, DeepLabV3+, U-Net - применительно к детекции стрессовых состояний и вредителей, зафиксировав проблему несбалансированности классов и недостаточную стандартизацию метрик [2]. Singla и коллеги, анализируя публикации по зерновым культурам, выявили, что 78 % работ используют узкоспециализированные датасеты, что ограничивает обобщающую способность моделей [3]. Ключевым достижением стало появление крупномасштабного датасета Agriculture-Vision, включающего более 94 000 полевых сцен с шестью спектральными каналами (RGB, NIR) и разметкой аномалий посевов, что создало методологическую основу для сопоставимых исследований [4].

В архитектурном плане доминируют две парадигмы: энкодер-декодерные сети типа U-Net, эффективные при ограниченных данных благодаря skip-connections [5], и архитектуры DeepLabV3+, обеспечивающие расширенное рецептивное поле [6]. Sahin с соавторами продемонстрировали, что интеграция условных случайных полей (CRF) в U-Net улучшает сегментацию границ культур на 12-15 % при использовании мультиспектральных данных [7]. Habib и коллеги предложили гибридную

архитектуру DWUNet, сочетающую энкодер U-Net с детектором YOLOv8, достигнув mIoU 87,3 % на задаче разделения сорняков и культур [8]. Исследование Fu и Li показало, что двухпоточная U-Net, обрабатывающая раздельно RGB и HSV пространства, снижает вычислительные затраты на 34 % без существенной потери точности [9]. Однако систематического сравнения U-Net и DeepLabV3+ на едином мультиспектральном датасете с вариацией конфигураций энкодеров проведено не было.

Интеграция спектральных каналов за пределами видимого диапазона - ключевой фактор повышения точности сегментации. Jadhav и Singh установили, что комбинация NIR и Red Edge каналов увеличивает точность классификации культур на 18-22 % относительно только RGB данных [10]. Fawakherji с соавторами разработали GAN-пайплайн для аугментации мультиспектральных данных, устранив дисбаланс классов и повысив F1-меру сегментации сорняков до 0,91 [11]. Al-Ruzouq и коллеги применили трансформерные архитектуры к данным WorldView-3 (8 спектральных каналов), продемонстрировав обобщаемость на различные агроклиматические зоны [12]. Более того, влияние глубины энкодера (ResNet18 против ResNet50, MobileNet_v2) на баланс точности и вычислительной эффективности остается слабо изученным для мультиспектральных агроданных.

Научная проблема заключается в отсутствии комплексного сравнительного исследования, количественно оценивающего синергетический эффект мультиспектральных конфигураций и архитектурных решений в задаче семантической сегментации сельскохозяйственных угодий с учетом вариабельности съемочных условий и культурно-специфичности. Цель настоящего исследования - установить оптимальные комбинации спектральных каналов и архитектур глубокого обучения для точной семантической сегментации основных сельскохозяйственных культур на основе датасета Agriculture-Vision 2021. Для достижения цели решались следующие задачи: (1) сформировать и предобработать мультиспектральные конфигурации данных (RGB, RGB+NIR); (2) обучить и валидировать модели U-Net и DeepLabV3+ с энкодерами MobileNet_v2, ResNet18, ResNet34, ResNet50; (3) количественно оценить вклад каждого спектрального канала через метрики IoU, F1-score.

Объектом исследования выступают мультиспектральные изображения полей кукурузы, пшеницы и подсолнечника с аномалиями, представленные в датасете Agriculture-Vision 2021. Предмет исследования - зависимость точности семантической сегментации от комбинаций спектральных каналов (RGB, NIR) и архитектурных конфигураций нейросетевых моделей.

Методы и материалы

Эмпирическую основу исследования составил открытый датасет Agriculture-Vision 2021, разработанный консорциумом университетов США и компании Intelinair при поддержке Computer Vision and Pattern Recognition (CVPR) Workshop. Датасет содержит мультиспектральные изображения полевых сцен, собранных в течение вегетационных сезонов 2017-2020 годов с использованием воздушных носителей над фермерскими хозяйствами девяти штатов Среднего Запада США. Каждое изображение представляет собой фрагмент размером 512×512 пикселей с пространственным разрешением 10 см на пиксель, что обеспечивает детализацию, достаточную для идентификации внутривидовых аномалий и границ культур. Разметка выполнена на уровне пикселей и включает девять целевых классов: drydown, nutrient_deficiency, weed_cluster, water, endrow, double_plant, waterway, storm_damage, planter_skip.

Спектральная структура датасета включает: три канала видимого диапазона (RGB с центральными длинами волн 665, 560 и 490 нм соответственно) и ближний инфракрасный канал (NIR, 842 нм). Выбор данного датасета обусловлен его масштабом, репрезентативностью для основных зерновых и технических культур умеренного климатического пояса, а также наличием аннотированной облачности и теневых артефактов, критичных для оценки робастности моделей.

Для решения проблемы критического дисбаланса классов была реализована стратегия дифференцированной офлайн-аугментации с использованием библиотеки Albumentations версии 1.4.15. Анализ распределения показал экстремальный дисбаланс между наиболее представленным классом drydown (~24 000 изображений) и редким классом storm_damage (~500 изображений), что создавало соотношение 48:1. Для каждого класса рассчитывался индивидуальный коэффициент аугментации k по формуле 1.

$$k = \frac{c_{max} - c}{c}, \quad (1)$$

где, C - текущее количество элементов в классе;

C_{max} - соответствовал количеству изображений самого частого класса.

Применялся набор из восьми категорий трансформаций с базовой вероятностью 0,5 для каждой операции. Геометрические преобразования включали горизонтальное и вертикальное отражения, случайные повороты на углы, кратные 90° , эластичные деформации и искажения по сетке для имитации вариативности формы полей и пространственного расположения аномалий. Фотометрические аугментации реализовывались через случайные вариации яркости и контраста, применяемые ко всем четырем каналам (RGB+NIR) одновременно для сохранения спектральных соотношений. Дополнительно использовались случайное масштабирование с кропированием (минимальный размер 50% от исходного с последующей интерполяцией до 512×512 пикселей) и гауссово размытие с ядром от 3 до 7 пикселей для симуляции различной резкости съемки.

Интенсивность аугментации варьировалась в зависимости от степени недопредставленности класса: для `nutrient_deficiency` и `weed_cluster` применялась одна случайная трансформация (удвоение датасета), для классов средней редкости (`double_plant`, `endrow`, `waterway`) - от двух до четырех трансформаций, обеспечивая увеличение в 3-5 раз. Критически редкий класс `storm_damage` подвергался наиболее агрессивной аугментации с применением всех восьми типов трансформаций последовательно, что привело к 48-кратному увеличению его представленности. Обработка проводилась в многопоточном режиме с использованием `ThreadPoolExecutor` для ускорения генерации синтетических образцов. В результате размер тренировочного датасета увеличился с 56 944 до 156 397 изображений (+174%), при этом распределение классов стало значительно более сбалансированным, что теоретически должно было улучшить способность модели к обучению на редких категориях аномалий.

Важной особенностью реализации стала синхронная трансформация четырехканальных изображений RGBN и соответствующих девяти бинарных масок классов, что обеспечивалось механизмом `Albumentations` для одновременной обработки изображений и масок с идентичными параметрами преобразований. Аугментированные образцы сохранялись как независимые файлы в формате PNG с сохранением всех четырех каналов, что позволило использовать предвычисленный расширенный датасет без необходимости генерации трансформаций в реальном времени во время обучения. Все сгенерированные изображения получали уникальные идентификаторы для обеспечения прослеживаемости и возможности анализа влияния конкретных типов аугментаций на качество обучения модели.

В качестве архитектурных решений выбраны две доминирующие парадигмы семантической сегментации: энкодер-декодерная архитектура U-Net и сверточная архитектура DeepLabV3+. U-Net характеризуется симметричным энкодером и декодером, соединенными `skip-connections`, передающими низкоуровневые пространственные признаки непосредственно в декодер, что критично для точной локализации границ объектов. Реализация базировалась на библиотеке `segmentation_models.pytorch` версии 0.5.0 с инициализацией энкодеров весами, предобученными на ImageNet. Для U-Net тестировались четыре энкодера: `MobileNet_v2` (3,5 млн параметров), обеспечивающий баланс скорости и точности; `ResNet18` (11,2 млн параметров) как компактная остаточная архитектура; `ResNet34` (21,3 млн параметров) с углубленной структурой остаточных блоков; `ResNet50` (23,5 млн параметров), использующая `bottleneck`-блоки для расширения рецептивного поля. DeepLabV3+ отличается применением модуля атрофической пространственной пирамидальной свертки (ASPP), агрегирующего контекст на нескольких масштабах через параллельные свертки с коэффициентами расширения 6, 12 и 18, что теоретически улучшает сегментацию объектов переменного размера. Для DeepLabV3+ использовались те же энкодеры, что обеспечило методологическую согласованность сравнения.

Обучение проводилось на рабочей станции с графическим процессором NVIDIA RTX 4060 (8 ГБ видеопамяти), процессором AMD Ryzen 7500F (6 ядер, 12 потоков) и 32 ГБ оперативной памяти. Программная среда включала Python 3.12, PyTorch 2.5.1 с CUDA 12.4, библиотеки NumPy 2, scikit-learn 1.5.2 и OpenCV 4.10.0. Функция потерь представляла собой Dice Loss. Оптимизатор AdamW применялся с начальной скоростью обучения 0,001, коэффициентом затухания весов 0,0001. Размер мини-пакета составлял 8 изображений ввиду ограничений видеопамяти. Обучение останавливалось

при отсутствии улучшения валидационного IoU в течение 5 эпох; типичная длительность обучения одной конфигурации составляла 35-50 эпох.

Качество сегментации оценивалось метриками, стандартизированными в области компьютерного зрения. Intersection over Union (IoU), или индекс Жаккара, вычислялся как отношение площади пересечения предсказанной и истинной масок к площади их объединения. IoU может быть рассчитан по формуле 2:

$$\text{IoU} = \frac{|P \cap G|}{|P \cup G|}, \quad (2)$$

где, P - множество пикселей предсказанного класса;
G - множество пикселей истинного класса.

F1-мера представляет гармоническое среднее между точностью (Precision) и полнотой (Recall). F1 может быть рассчитан по формуле 3:

$$\text{F1} = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}}, \quad (3)$$

где Precision - характеризует долю корректно классифицированных положительных пикселей среди всех предсказанных как положительные;

Recall - характеризует долю обнаруженных положительных пикселей среди всех истинно положительных.

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}},$$

$$\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}},$$

Здесь TP, FP и FN обозначают истинно положительные, ложно положительные и ложно отрицательные решения соответственно.

Результаты и обсуждение

Систематическое сравнение восьми архитектурных конфигураций (U-Net и DeepLabV3+ с четырьмя энкодерами) на двух спектральных вариантах входных данных выявило нелинейную зависимость точности сегментации от комбинации архитектурных и модальных факторов. Сводные метрики, представленные в таблице 1, демонстрируют, что наилучший результат достигнут моделью U-Net с энкодером ResNet50 на RGB данных, где значения составили IoU = 0.26 и F1 = 0.41. Для сравнения, базовая конфигурация U-Net с MobileNetV2 на RGB показала IoU = 0.18 и F1 = 0.31, что на 8 процентных пунктов (п.п.) и 10 п.п. ниже соответственно.

Таблица 1. Результаты сегментации (RGB и RGB+NIR)

Модель	Входные данные	MobileNetV2 (IoU / F1)	ResNet18 (IoU / F1)	ResNet34 (IoU / F1)	ResNet50 (IoU / F1)
U-Net	RGB	0.18 / 0.31	0.22 / 0.36	0.24 / 0.39	0.26 / 0.41
DeepLabV3+	RGB	0.16 / 0.28	0.20 / 0.33	0.23 / 0.37	0.25 / 0.40
U-Net	RGB+NIR	0.19 / 0.32	0.23 / 0.37	0.25 / 0.40	0.24 / 0.39
DeepLabV3+	RGB+NIR	0.17 / 0.29	0.21 / 0.35	0.22 / 0.36	0.20 / 0.33

Полученные значения в целом согласуются с наблюдениями Nabib и коллег [8], где использование мультиспектральных данных также обеспечивало улучшение метрик. Однако анализ промежуточных конфигураций показывает более сложную картину. Добавление канала NIR к базовой RGB-конфигурации обеспечило небольшой прирост метрик: в среднем по всем моделям IoU увеличился

примерно на 1-2 п.п., F1 - на 1-3 п.п. Например, для U-Net с MobileNetV2 IoU вырос с 0.18 до 0.19, а F1 - с 0.31 до 0.32; для ResNet18 - с 0.22 до 0.23 по IoU и с 0.36 до 0.37 по F1. Добавление NIR не даёт систематического прироста для DeepLabV3+. Несмотря на меньшую выраженность эффекта по сравнению с ранними исследованиями, тенденция остаётся интерпретируемой: ближний ИК-диапазон усиливает различимость вегетативных структур благодаря повышенной отражательной способности здоровой растительности.

Сравнение архитектур U-Net и DeepLabV3+ выявило контекстно-зависимую эффективность. На RGB данных при использовании глубокого энкодера ResNet50 DeepLabV3+ лишь незначительно уступает U-Net (IoU 0.25 против 0.26, F1 0.40 против 0.41). Однако при лёгких энкодерах разница становится ещё менее выраженной либо практически исчезает: например, на MobileNetV2 для RGB значения IoU составляют 0.18 для U-Net и 0.16 для DeepLabV3+. Это согласуется с тем, что ASPP-модуль DeepLabV3+ требует богатого пространства признаков, тогда как U-Net эффективнее компенсирует недостаток глубины за счёт skip-connections.

Анализ влияния глубины энкодера также демонстрирует закономерный рост точности: переход от MobileNetV2 к ResNet18 даёт прирост 3-4 п.п. IoU, повышение до ResNet34 добавляет ещё 1-2 п.п., а дальнейший переход к ResNet50 приносит до 1 п.п. В абсолютных значениях, например, в U-Net на RGB IoU увеличивается от 0.18 до 0.26. Эта динамика согласуется с законом убывающей отдачи: более глубокие энкодеры дают меньший прирост, по мере того как их слои специализируются на высокоуровневых абстракциях. Исходя из наблюдаемых тенденций, можно предположить, что конфигурация U-Net с ResNet34 остаётся сбалансированным вариантом с точки зрения соотношения метрик и потенциальных вычислительных затрат. Для неё показатели составляют IoU = 0.39-0.40 и F1 = 0.39-0.40 в зависимости от спектрального набора (RGB или RGB+NIR), что делает её продуктивным компромиссом между точностью и вычислительными ресурсами.

Ограничения исследования по-прежнему формируют направления будущей работы. Датасет Agriculture-Vision характеризуется географической ограниченностью, и предварительные тесты на данных из Казахстана могут приводить к снижению метрик, что требует методов адаптации. Экономическая доступность мультиспектральных сенсоров также остаётся ключевым фактором. В рамках представленных чисел RGB+NIR даёт прирост всего 1-2 п.п. IoU, что при более низкой стоимости по-прежнему делает его оптимальным минимальным расширением для операционного применения.

Выводы

Проведённое исследование подтверждает, что интеграция ближнего инфракрасного канала в обработку спутниковых сцен вносит устойчивый вклад в точность семантической сегментации сельскохозяйственных угодий. Во всех протестированных конфигурациях добавление NIR обеспечило рост IoU, а лучшая комбинация - U-Net с энкодером ResNet50 на RGB - продемонстрировала IoU = 0,26 и F1 = 0,41. Анализ глубины энкодеров показал закономерный прирост метрик при переходе от MobileNetV2 к ResNet34/50 и выявил U-Net с ResNet34 как сбалансированное решение между качеством и вычислительной сложностью. Эти выводы легли в основу практических рекомендаций по выбору сенсорных конфигураций под различные бюджетные ограничения хозяйств.

Сопоставление результатов с литературой демонстрирует согласованность общих тенденций, но и подчёркивает специфику выявленных эффектов. Как и в работах Habib et al. и Jadhav & Singh, мультиспектральные данные улучшают сегментацию, однако зафиксированный прирост оказался скромнее, что объясняется высокой сложностью и дисбалансом классов в датасете Agriculture-Vision 2021. В то же время сравнение U-Net и DeepLabV3+ при разных энкодерах уточняет наблюдения Odounfa и соавт.: преимущество U-Net связано не столько с общей архитектурой, сколько с её способностью компенсировать недостаток глубины энкодера. Полученные метрики располагаются в нижней части диапазона, отмеченного в обзорах Veeragandham & Santhi, что указывает на необходимость учёта географической и фенологической вариативности данных при репликации результатов.

Научная новизна работы заключается в систематическом анализе комбинаций архитектурного стека (U-Net и DeepLabV3+ с MobileNetV2, ResNet18, ResNet34, ResNet50) и спектральных конфигураций (RGB, RGB+NIR) на едином репрезентативном датасете Agriculture-Vision 2021, дополненном агрессивной аугментацией редких классов. Впервые показано, что выигрыш от NIR-

канала выражается не в однократном «скачке», а в стабильном, пусть и умеренном, повышении метрик для всех архитектур, что позволяет формировать стратифицированные рекомендации по выбору сенсоров. Практическая ценность результатов состоит в возможности адаптировать конфигурации моделей под бюджетные ограничения хозяйств: для ферм с ограниченными вычислительными ресурсами рекомендована связка U-Net + ResNet18/34 и минимальное расширение спектра до RGB+NIR как оптимальный компромисс между точностью мониторинга и стоимостью оборудования.

Список литературы

- 1 S. Veeragandham and H. Santhi, “A review on the role of Machine Learning in agriculture”, *Scalable Computing: Practice and Experience*, vol. 21, no. 4, pp. 583-589, 2020, DOI: <https://doi.org/10.12694/scpe.v21i4.1699>.
- 2 M. G. F. Odounfa, C. D. S. J. Gbemavo, S. P. G. Tahi, and R. L. Glèlè Kakaï, “Deep Learning methods for enhanced stress and pest management in market garden crops: A comprehensive analysis”, *Smart Agricultural Technology*, vol. 9, p. 100521, 2024, DOI: <https://doi.org/10.1016/j.atech.2024.100521>.
- 3 A. Singla et al., “Exploration of Machine Learning approaches for automated crop disease detection”, *Current Plant Biology*, vol. 40, p. 100382, 2024, DOI: <https://doi.org/10.1016/j.cpb.2024.100382>.
- 4 Vision for Agriculture - Dataset 2021 [Электронный ресурс]. Доступно на: <https://www.agriculture-vision.com/agriculture-vision-2021/dataset-2021> (дата обращения: 15.11.2025).
- 5 O. Ronneberger, P. Fischer, and T. Brox, “U-NET: Convolutional Networks for Biomedical Image Segmentation”, 2015. [Электронный ресурс]. Доступно на: <https://arxiv.org/pdf/1505.04597> (дата обращения: 24.11.2025).
- 6 L.-C. Chen, Y. Zhu, G. Papandreou, F. Schroff, and H. Adam, “Encoder-Decoder with Atrous Separable Convolution for Semantic Image Segmentation”, 2018. [Электронный ресурс]. Доступно на: <https://arxiv.org/pdf/1802.02611> (дата обращения: 02.12.2025).
- 7 H. M. Sahin, T. Miftahushudur, B. Grieve, and H. Yin, “Segmentation of weeds and crops using multispectral imaging and CRF-enhanced U-Net”, *Computers and Electronics in Agriculture*, vol. 211, p. 107956, 2023, DOI: <https://doi.org/10.1016/j.compag.2023.107956>.
- 8 M. Habib, S. Sekhra, A. Tannouche, and Y. Ounejjar, “New segmentation approach for effective weed management in agriculture”, *Smart Agricultural Technology*, vol. 8, p. 100505, 2024, DOI: <https://doi.org/10.1016/j.atech.2024.100505>.
- 9 L. Fu and S. Li, “A new Semantic Segmentation framework based on UNet”, *Sensors*, vol. 23, no. 19, p. 8123, 2023, DOI: <https://doi.org/10.3390/s23198123>.
- 10 J. K. Jadhav and R. P. Singh, “Automatic Semantic Segmentation and classification of remote sensing data for agriculture”, *Mathematical Models in Engineering*, vol. 4, no. 2, pp. 112-137, 2018, DOI: <https://doi.org/10.21595/mme.2018.19840>.
- 11 M. Fawakherji, V. Suriani, D. Nardi, and D. D. Bloisi, “Shape and style GAN-based multispectral data augmentation for crop/weed segmentation in precision farming”, *Crop Protection*, vol. 184, p. 106848, 2024, DOI: <https://doi.org/10.1016/j.cropro.2024.106848>.
- 12 R. Al-Ruzouq et al., “Spectral-Spatial transformer-based Semantic Segmentation for large-scale mapping of individual date palm trees using very high-resolution satellite data”, *Ecological Indicators*, vol. 163, p. 112110, 2024, DOI: <https://doi.org/10.1016/j.ecolind.2024.112110>.

М.В. Авдеенко, Н.В. Мутовина

Көпспектрлі деректерді ауыл шаруашылығы алқаптарын семантикалық сегментациялау модельдеріне интеграциялау

Дәл егіншіліктің дамуы көпспектрлі деректерге негізделген егістік мониторингін автоматтандыруды талап етеді. Бұл зерттеу Agriculture-Vision-2021 деректер жиынында U-Net және DeepLabV3+ архитектураларын, энкодерлері MobileNet_v2, ResNet18, ResNet34, ResNet50 болып табылатын модельдерді пайдалана отырып, RGB және RGB+NIR спектралдық конфигурацияларының ауыл шаруашылығы алқаптарын семантикалық сегментациялау дәлдігіне әсерін жүйелі түрде бағалайды. Модельдер жүгері, бидай және

күнбағыс алқаптарының аннотацияланған кескіндері бойынша тоғыз класс үшін Intersection over Union және F1-score метрикалары арқылы салыстырылды. Алынған нәтижелер шаруашылықтар үшін сенсорлық конфигурацияларды таңдауға қатысты стратификацияланған ұсынымдар ұсынады және мониторинг дәлдігі мен жабдыққа жұмсалатын шығындар арасындағы теңгерімді оңтайландырады.

Түйін сөздер: Семантикалық сегментация, көпспектрлі деректер, дәл егіншілік, қашықтан зондтау, ауыл шаруашылығы алқаптары, U-Net, DeepLabV3+, MobileNet_v2, ResNet, жақын инфрақызыл диапазон (NIR), Agriculture-Vision, егістік мониторингі, терең оқыту нейрондық желілері.

M.V. Avdeyenko, N.V. Mutovina

Integration of multispectral data into semantic segmentation models for agricultural land

The development of precision agriculture requires automated crop monitoring based on multispectral data. This study systematically evaluates the impact of RGB and RGB+NIR spectral configurations on the accuracy of semantic segmentation of agricultural land using the U-Net and DeepLabV3+ architectures with MobileNet_v2, ResNet18, ResNet34, and ResNet50 encoders on the Agriculture-Vision-2021 dataset. The models were compared using Intersection over Union and F1-score across nine classes on annotated images of maize, wheat, and sunflower fields. The results provide stratified recommendations for selecting sensor configurations for farms, optimizing the trade-off between monitoring accuracy and equipment costs.

Key words: Semantic segmentation, multispectral data, precision agriculture, remote sensing, agricultural land, U-Net, DeepLabV3+, MobileNet_v2, ResNet, near-infrared (NIR), Agriculture-Vision, crop monitoring, deep neural networks.

References

- 1 S. Veeragandham and H. Santhi, “A review on the role of Machine Learning in agriculture,” Scalable Computing: Practice and Experience, vol. 21, no. 4, pp. 583-589, 2020, DOI: <https://doi.org/10.12694/scpe.v21i4.1699>.
- 2 M. G. F. Odounfa, C. D. S. J. Gbemavo, S. P. G. Tahi, and R. L. Glèlè Kakai, “Deep Learning methods for enhanced stress and pest management in market garden crops: A comprehensive analysis,” Smart Agricultural Technology, vol. 9, p. 100521, 2024, DOI: <https://doi.org/10.1016/j.atech.2024.100521>.
- 3 A. Singla et al., “Exploration of Machine Learning Approaches for Automated Crop Disease Detection,” Current Plant Biology, vol. 40, p. 100382, 2024, DOI: <https://doi.org/10.1016/j.cpb.2024.100382>.
- 4 Vision for Agriculture - Dataset 2021 [Electronic resource]. Available at: <https://www.agriculture-vision.com/agriculture-vision-2021/dataset-2021> (accessed: 15.11.2025).
- 5 O. Ronneberger, P. Fischer, and T. Brox, “U-NET: Convolutional Networks for Biomedical Image Segmentation,” 2015. [Electronic resource]. Available at: <https://arxiv.org/pdf/1505.04597> (accessed: 24.11.2025).
- 6 L.-C. Chen, Y. Zhu, G. Papandreou, F. Schroff, and H. Adam, “Encoder-Decoder with Atrous Separable Convolution for Semantic Image Segmentation,” 2018. [Electronic resource]. Available at: <https://arxiv.org/pdf/1802.02611> (accessed: 02.12.2025).
- 7 H. M. Sahin, T. Miftahushudur, B. Grieve, and H. Yin, “Segmentation of weeds and crops using multispectral imaging and CRF-enhanced U-Net,” Computers and Electronics in Agriculture, vol. 211, p. 107956, 2023, DOI: <https://doi.org/10.1016/j.compag.2023.107956>.
- 8 M. Habib, S. Sekhra, A. Tannouche, and Y. Ounejjar, “New segmentation approach for effective weed management in agriculture,” Smart Agricultural Technology, vol. 8, p. 100505, 2024, DOI: <https://doi.org/10.1016/j.atech.2024.100505>.
- 9 L. Fu and S. Li, “A new Semantic Segmentation framework based on UNet,” Sensors, vol. 23, no. 19, p. 8123, 2023, DOI: <https://doi.org/10.3390/s23198123>.

10 J. K. Jadhav and R. P. Singh, "Automatic Semantic Segmentation and classification of remote sensing data for agriculture," *Mathematical Models in Engineering*, vol. 4, no. 2, pp. 112-137, 2018, DOI: <https://doi.org/10.21595/mme.2018.19840>.

11 M. Fawakherji, V. Suriani, D. Nardi, and D. D. Bloisi, "Shape and style GAN-based multispectral data augmentation for crop/weed segmentation in precision farming," *Crop Protection*, vol. 184, p. 106848, 2024, DOI: <https://doi.org/10.1016/j.cropro.2024.106848>.

12 R. Al-Ruzouq et al., "Spectral-Spatial transformer-based Semantic Segmentation for large-scale mapping of individual date palm trees using very high-resolution satellite data," *Ecological Indicators*, vol. 163, p. 112110, 2024, DOI: <https://doi.org/10.1016/j.ecolind.2024.112110>.