

## Специализированный датасет мультиспектральных аэрофотоснимков для решения задач точного земледелия с применением методов искусственного интеллекта

О. А. Митрофанова<sup>1</sup>, Е. П. Митрофанов<sup>1</sup>, И. С. Блеканов<sup>2</sup>, А. Е. Молин<sup>2</sup>

<sup>1</sup> *Агрофизический научно-исследовательский институт  
Санкт-Петербург, 195220, Россия  
E-mails: omitrofa@gmail.com, mjeka89@gmail.com*

<sup>2</sup> *Санкт-Петербургский государственный университет  
Санкт-Петербург, 199034, Россия*

Благодаря стремительному развитию информационных технологий всё более актуальной становится задача создания наборов больших качественных данных дистанционного зондирования. В Агрофизическом научно-исследовательском институте (АФИ) методы искусственного интеллекта, анализа изображений используются на протяжении около 20 лет. За этот период накоплен большой объём информации для решения задач точного земледелия. Объектом представленного исследования является опытный биополигон АФИ, расположенный в Ленинградской обл., который состоит из 29 полей. Для сбора применялся беспилотный летательный аппарат, разработанный в АФИ, а также беспилотная авиационная система «Геоскан-401». Съёмка осуществлялась в пяти спектрах: красном, зелёном, синем, инфракрасном и дальнем красном, средняя высота полётов составила 80 м, пространственное разрешение снимков — 1–10 см/пиксель. В работе подробно рассмотрены сформированные алгоритмы сбора данных, а также предварительной обработки информации. В исследовании в качестве демонстрации применимости созданного размеченного датасета был проведён эксперимент по анализу ортофотоплана одного из полей биополигона за период 2019–2021 гг., составленного из снимков мультиспектральной камеры Micasense RedEdge MX. В качестве методов анализа изображений использовались адаптированный под задачу метод классического алгоритма машинного обучения Random Forest и метод глубокого обучения на основе архитектуры U-Net. Результаты проведённого эксперимента продемонстрировали преимущество метода глубокого обучения при решении задачи определения азотного режима посевов для дифференцированного внесения удобрений.

**Ключевые слова:** датасет, анализ изображений, мультиспектральные аэрофотоснимки, точное земледелие, машинное обучение, глубокое обучение

Одобрена к печати: 06.10.2023  
DOI: 10.21046/2070-7401-2023-20-6-67-79

### Введение

В современном мире усиливаются негативные факторы, оказывающие существенное влияние на продовольственную безопасность: глобальные изменения климата, экономическая обстановка, прогнозируемое увеличение населения мира на два миллиарда в ближайшие 30 лет (The future..., 2017; Wheeler, Von Braun, 2013). Эти аспекты ставят перед нами беспрецедентные вызовы: необходимо наращивать объёмы производства растениеводческой продукции, при этом снизив негативное воздействие на водные, почвенные ресурсы, замедлив деградацию плодородных земель. Решение таких задач не представляется возможным без перехода от интенсивного сельского хозяйства к устойчивому (точному земледелию — ТЗ).

В последние десятилетия в научных исследованиях достаточно широко представлено использование данных дистанционного зондирования для землепользования и мониторинга состояния культурных растений (Блохин и др., 2020; Якушев и др., 2020; Weiss et al., 2020). При этом следует отметить, что аэрофотосъёмка при решении определённого круга задач точного земледелия имеет преимущества перед спутниковой съёмкой, так как космоснимки, как правило, ограничены временным и пространственным разрешением и не позволяют

оперативно получить высокодетализированные данные для сельскохозяйственной территории (Zhang, Kovacs, 2012).

Применение беспилотных летательных аппаратов (БЛА) и беспилотных авиационных систем (БАС) становится всё более актуальным и перспективным в задачах ТЗ благодаря высокому пространственному разрешению, скорости получения и обработки данных, а также автоматизации процессов сбора информации (Tsouros et al., 2019). Неоднократно свою эффективность продемонстрировало использование методов интеллектуального анализа данных, полученных с изображений сельскохозяйственной территории (Belcore et al., 2021; Lu et al., 2021). Снимки, собранные с помощью БЛА или БАС, применяются во многих задачах ТЗ: в мониторинге состояния культурных растений (Jung et al., 2021), оценке биомассы (Sinde-Gonzalez et al., 2021), обнаружении посевов (Zheng et al., 2020), выявлении заболеваний и засорённости (Ye et al., 2020) и др. Всё более актуальной в растениеводстве становится мультиспектральная съёмка (Kumar et al., 2020).

Таблица 1. Примеры существующих датасетов со снимками, полученными в полевых условиях

Наименование датасета	Используемые каналы съёмки	Платформа для съёмки	Период съёмок	Культура	Количество исходных снимков	Применение
weedNet (Sa et al., 2017)	Red, Green, Blue, NIR	БЛА DJI Mavic	1 сезон (2016)	Сахарная свёкла	465	Мониторинг засорённости
Sugar beet (Chebroly et al., 2017)	Red, Green, Blue, NIR	Наземный робот BoniRob	1 сезон, весенний период (2016)	Сахарная свёкла	>10 000	Картирование, обнаружение посевов, сорняков
WeedMap (Sa et al., 2018)	Red, Green, Blue, NIR, Red Edge	БЛА DJI Mavic pro, Inspire2	1 сезон (2017)	Сахарная свёкла	>10 000	Мониторинг засорённости
Maize, drone set (Wiesner-Hanks et al., 2018)	Red, Green, Blue	БЛА DJI Matrice 600	1 сезон (2017)	Кукуруза	7669	Мониторинг болезней посевов
Weed mapping (Pereira et al., 2019)	Red, Green, Blue	БЛА Horus Aeronaves	1 сезон (2018)	Сахарный тростник	1	Обнаружение сорняков
Deep Seedling (Jiang et al., 2019)	Red, Green, Blue	Наземная съёмка	1 сезон (2015)	Хлопок	5743	Подсчёт посевов
Oil Radish Growth (Mortensen et al., 2019)	Red, Green, Blue	Подвес на трактор	1 сезон (2015)	Редька	5416	Семантическая сегментация и оценка урожайности
Crop/Weed Field Image (Haug, Ostermann, 2014)	Red, Green, Blue, NIR	Наземный робот BoniRob	1 сезон (2013)	Морковь	60	Обнаружение сорняков
DeepWeeds (Olsen et al., 2019)	Red, Green, Blue	Наземный робот AutoWeed, а также инструмент WeedLogger	Июнь 2017 г. – март 2018 г.	8 видов сорняков	17 509	Обнаружение сорняков
Open Plant Phenotype Database (Madsen et al., 2020)	Red, Green, Blue	Наземная установка	4 сезона (2017–2019)	46 видов сорняков	7590	Обнаружение сорняков
Carrots 2017, Onions 2017 (Bosilj et al., 2020)	Red, Green, Blue, NIR	Наземная съёмка	1 сезон (2017)	Морковь, лук	40	Сегментация, мониторинг засорённости

Окончание табл. 1

Наименование датасета	Используемые каналы съёмки	Платформа для съёмки	Период съёмки	Культура	Количество исходных снимков	Применение
Agriculture-Vision (Chiu et al., 2020)	Red, Green, Blue, NIR	БЛА	3 сезона (2017–2019)	Кукуруза, соя	3432 (>90 000 итоговых)	Сегментация аномалий — 9 классов (двойной посев, дефицит удобрений, переувлажнение и др.)
RiceSeedlingDataset (Yang et al., 2021)	Red, Green, Blue	БЛА DJI Phantom 4 Pro	3 сезона (2018–2020)	Рис	>5000	Обнаружение, подсчёт посевов
Wuhan UAV-borne hyperspectral image (WHU-Hi) (Zhong et al., 2020)	Гиперспектр (270 каналов в диапазоне 400–1000 нм))	БЛА DJI Matrice 600 Pro	3 дня в разные годы (2016–2018)	Кукуруза, хлопок, кунжут, соя, рис, клубника, шпинат, арбуз, зелень и др.	3 ортофото (по 270 слоёв)	Классификация культур
Sunflower lodging (Song et al., 2020)	Red, Green, Blue, Red Edge, NIR 1, NIR 2	БЛА DJI Matrice 600	1 сезон (2018)	Подсолнечник	865	Мониторинг полегания посевов
UAV-based Multispectral and Thermal dataset (Kallimani et al., 2020)	Red, Green, Blue, NIR, Red Edge, Thermal Infrared	БЛА eBee, DJI X3, S950	1 день (2017)	Картофель, пшеница, ячмень	>18 000	Изучение суточных вариаций, радиометрической и геометрической точности данных БЛА для ТЗ

Примечание: Red — красный (R); Green — зелёный (G); Blue — синий (B); Red Edge — красный край; NIR — *англ.* near infrared, ближний инфракрасный; Thermal Infrared — тепловой инфракрасный.

В качестве отдельного и важного этапа таких исследований выделяется подготовка набора исходных данных дистанционного зондирования. В табл. 1 представлена выборка существующих датасетов (*англ.* dataset, наборы информации, содержащие тестовые, размеченные экспертами данные) со снимками, полученными в полевых условиях, большинство из которых являются открытыми для исследователей. Следует отметить, что существенная часть изображений получена только в видимом диапазоне с очень ограниченным временным периодом. Кроме того, представленные наборы информации охватывают лишь малую часть культур, выращиваемых в северо-западном регионе России. Соответственно, задача создания наборов больших данных на основе аэрофотосъёмки сельскохозяйственных полей представляется актуальной и перспективной.

В Агрофизическом научно-исследовательском институте (АФИ) методы искусственного интеллекта, анализа изображений используются на протяжении около 20 лет. За этот период накоплен большой объём данных высокого качества для информационной поддержки задач ТЗ.

## Объекты и методы

Объектом исследования является опытный сельскохозяйственный комплекс филиала АФИ, расположенный в дер. Меньково (Гатчинский р-н, Ленинградская обл.). Биополигон состоит из 29 полей (рис. 1, см. с. 71), общая площадь земельных угодий составляет

538,56 га, продолжительность вегетационного периода — 117 дней, основные полевые культуры: пшеница, ячмень, овёс, картофель, многолетние травы.

Таблица 2. Общие сведения об используемых БЛА, БАС и сенсорах

<i>Беспилотные платформы</i>					
Технические характеристики	БЛА, разработанный в АФИ		БАС «Геоскан-401»		
Время полёта	30–45 мин		до 60 мин		
Радиус действия радиолинии	до 1,5 км		15 км		
Рабочая высота полёта	200–800 м		30–200 м		
Время развёртывания комплекса	10–15 мин		5–10 мин		
Площадка для взлёта и посадки	10×30 м		5×5 м		
Тип двигателей	Двигатель внутреннего сгорания		Электрический		
Площадь съёмки за 1 полёт	до 50 га (с разрешением 6 см/пиксель)		до 50 га (с разрешением 2 см/пиксель)		
Диапазон рабочей температуры	–20...+40 °С		–20...+40 °С		
Максимальная скорость ветра у земли	15 м/с		10 м/с		
Полезная нагрузка	Фотокамеры Olympus E-510 и Canon EOS Rebel T1i		Фотокамеры Sony A6000 и Sony RX1, мультиспектральная камера Micasense Rededge-MX		
<i>Сенсоры</i>					
Технические характеристики	Olympus E-510	Canon EOS Rebel T1i	Sony A6000	Sony RX1	Micasense Rededge-MX
Спектры съёмки	Видимый (Red, Green, Blue)	Ближний инфракрасный (NIR)	Ближний инфракрасный (NIR)	Видимый (Red, Green, Blue)	Видимый (Red, Green, Blue), ближний инфракрасный (NIR), красный край (Red Edge)
Эффективное разрешение, МП	10	15,1	24,3	24,3	1,2
Максимальный размер снимка, пиксели	3648×2736	4752×3168	6000×4000	6000×4000	1280×960

Сбор и обработка данных аэрофотосъёмки в АФИ осуществляется с 2003 г. (в датасете сохранены данные с 2006 г.). До 2020 г. использовался беспилотный радиоуправляемый самолёт, разработанный нашим коллегой инженером Айвазовым Г.С., в качестве полезной нагрузки применялись две цифровые камеры, позволяющие получать изображения в видимом и ближнем инфракрасном спектрах. С 2015 г. съёмка осуществляется с помощью БАС «Геоскан-401», которая включает в себя не только БЛА (квадрокоптер), но и встроенную бортовую систему (GPS-приёмник (*англ.* Global Positioning System, система глобального позиционирования), автопилот, стабилизация и т. п.), наземную станцию управления (защищённый ноутбук, программное обеспечение, цифровой канал связи управления и телеметрии, модем, походный стол, генератор и т. п.), а также полезную нагрузку (фотокамеры). Общие сведения об используемых беспилотных платформах и сенсорах представлены в *табл. 2*.

Благодаря переходу к использованию БАС «Геоскан-401» существенно улучшилось качество собираемой информации, повысилась точность геопространственной привязки, а также уменьшилось время предобработки снимков.

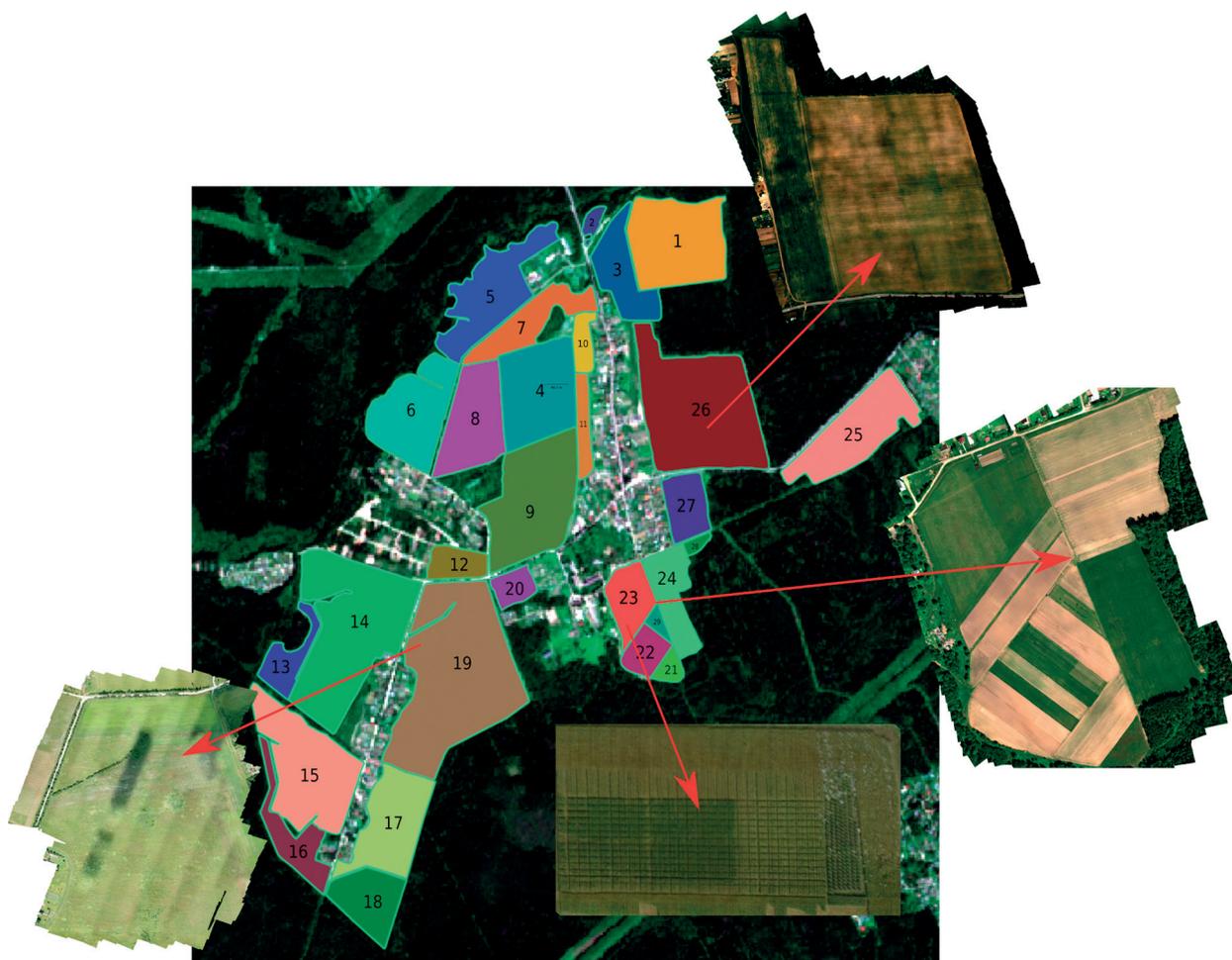


Рис. 1. Биополигон АФИ (дер. Меньково, Гатчинский р-н, Ленинградская обл.)

## Результаты и их обсуждение

Для оптимизации процесса сбора данных аэрофотосъёмки на основе БАС была сформирована общая схема необходимых действий:

*Шаг 1.* Сбор заявок сотрудников АФИ за несколько месяцев до проведения опытов на биополигоне.

*Шаг 2.* Обработка полученных заявок (изучение рельефа обследуемой местности, определение препятствий, выбор участков для съёмки).

*Шаг 3.* Мониторинг прогноза погоды в период проведения опытов, назначение ближайшей даты полёта в соответствии с метеоусловиями (осадки, скорость и направление ветра).

*Шаг 4.* Подготовка оборудования к выезду (заправка генератора, зарядка аккумуляторов комплекса RTK (*англ.* Real Time Kinematic), дронов, проверка сенсоров, погрузка).

*Шаг 5.* Выезд полётной группы на биополигон.

*Шаг 6.* Создание высокоточного контура обследуемого участка (не постоянный шаг, выполняется для каждого участка один раз в сезон).

*Шаг 7.* Развёртывание БАС для полёта, выбор точек посадки и взлёта, калибровка мультиспектральной камеры (при её использовании).

*Шаг 8.* Построение полётного задания в программе GeoScan Planner (нанесение границ обследуемого участка на основе карты Google, расчёт основных параметров: высоты полёта, количества точек на маршруте, скорости БЛА, пространственного разрешения съёмки и т. п.).

*Шаг 9.* Запуск БЛА, сопровождение полёта оператором наземной станции управления.

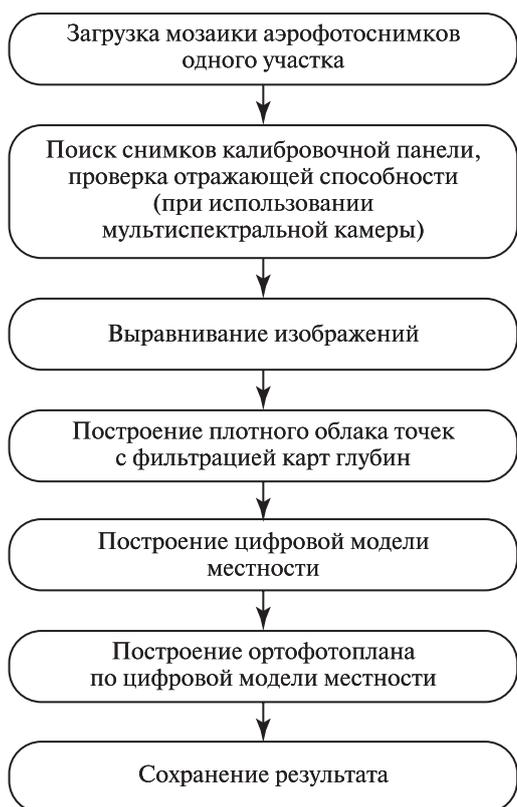


Рис. 2. Алгоритм предобработки данных аэрофотосъёмки для одного обследуемого участка в программе Agisoft Metashape

**Шаг 10.** Посадка БЛА, калибровка мультиспектральной камеры (при её использовании), сохранение данных аэрофотосъёмки, создание файлов геопространственной привязки.

**Шаг 11.** Сбор оборудования, отъезд полётной группы.

**Шаг 12.** Разгрузка, перевод оборудования в режим «хранение».

**Шаг 13.** Обработка полученных результатов в программе Agisoft Metashape (ООО «ЖивойСофт», Санкт-Петербург, <https://www.agisoft.com/>) (общий алгоритм для обработки данных одного участка представлен на рис. 2).

В табл. 3 приведена общая информация о количестве исходных снимков, полётах и сенсорах в период с 2006 до 2021 г. Средняя высота полётов составляет 80 м, пространственное разрешение варьируется в пределах 1–10 см/пиксель. Соответственно, за обозначенный период получено:

- с помощью радиоуправляемого самолёта АФИ — более 50 000 исходных снимков;
- с помощью БАС «Геоскан-401» — около 100 000 исходных снимков;
- более 700 обработанных данных: ортофотопланы, карты распределения индекса NDVI, цифровые модели рельефа местности.

Таблица 3. Общие сведения об аэрофотосъёмочных работах на биополигоне АФИ

Период	Количество дней съёмки за период	Количество исходных снимков за период	Сенсоры	Номера обследуемых полей	Платформа съёмки
2006–2016 гг.	120	23 556	Canon EOS Rebel T1i, Canon EOS M (NIR)	21–29	БЛА АФИ
		25 402	Olympus E-510, Canon EOS M (RGB)		
	18	7 007	Sony Nex-5T (модифицированный)	4, 7–9, 11, 14, 17–19, 21–26, 29	БАС «Геоскан»
		6 394	Sony Nex-5T		
2017 г.	2	734	Canon EOS M (модифицированный)	21–29	БЛА АФИ
		729	Canon EOS M		
	3	2 511	Sony Nex-5T (модифицированный)	14, 19, 21–24, 29	БАС «Геоскан»
		2 073	Sony Nex-5T		
2018 г.	4	552	Canon EOS M (модифицированный)	21–29	БЛА АФИ
		667	Canon EOS M		
	7	4 384	Sony A6000	14–19, 21–24, 26, 29	БАС «Геоскан»
		4 384	Sony RX1		

Период	Количество дней съёмки за период	Количество исходных снимков за период	Сенсоры	Номера обследуемых полей	Платформа съёмки
2019 г.	2	224	Canon EOS M (модифицированный)	21–29	БЛА АФИ
		414	Canon EOS M		
	8	5 531	Sony A6000	1, 3, 9, 15–19, 21–24, 26, 29	БАС «Геоскан»
		5 531	Sony RX1		
2020 г.	1	479	Canon EOS M (модифицированный)	21–29	БЛА АФИ
		239	Canon EOS M		
	5	2 906	Sony A6000	3, 9, 19, 21–24, 29	БАС «Геоскан»
		2 906	Sony RX1		
	10	54 300	Micasense RedEdge MX	3, 4, 9, 11, 21–24, 26, 29	
	2021 г.	2	819	Sony A6000	14, 19
819			Sony RX1		
12		57 125	Micasense RedEdge MX	9, 19, 21–26, 28, 29	

### Прикладные примеры использования датасета в задачах точного земледелия

Созданный многолетний датасет активно используется исследователями АФИ в различных научных проектах, таких как:

- оценка пространственного распространения сорняков (Шпанев, 2022);
- задачи, связанные с размещением и обработкой агрофизических и метеоданных, получаемых с использованием сенсорных узлов, технологии интернета вещей в растениеводстве (Блохин и др., 2019, 2020);
- задачи, связанные с дифференцированным внесением удобрений (Matvejenko et al., 2020), и др.

Одна из прикладных задач в сфере точного земледелия — определение обеспеченности посевов азотом. Так, в работе (Блеканов и др., 2021) при анализе изображений сельскохозяйственного поля, засеянного яровой пшеницей, было выявлено, что модель глубокого обучения U-Net (Ronneberger et al., 2015) показала более точный результат по метрикам качества, чем метод случайного леса (*англ.* Random Forest) (Breiman, 2001). В работе (Blekanov et al., 2023) построенный авторами датасет использовался для обучения различных нейросетевых архитектур в задаче сегментации мультиспектральных снимков для определения необходимых доз удобрения для дифференцированного внесения на сельскохозяйственном поле. В данных исследованиях авторы показали, что подходы на основе глубокого обучения с высокой точностью решают поставленную задачу. Так, например, метод на основе Attention R2U-Net имеет показатель точности более 99 % (Blekanov et al., 2023). Кроме того, в этой же работе исследовалось влияние мультиспектральных характеристик аэрофотоснимков (RGB, NIR, NDVI (*англ.* Normalized Difference Vegetation Index — нормализованный разностный вегетационный индекс) и др.) на качество анализа.

В настоящей работе в качестве демонстрации применимости созданного датасета был проведён эксперимент по анализу ортофотоплана одного из полей биополигона за период 2019–2021 гг. (см. табл. 3), составленного из снимков мультиспектральной камеры Micasense

RedEdge MX. В качестве методов анализа изображений использовались адаптированный под задачу метод классического алгоритма машинного обучения (*англ.* machine learning, ML) Random Forest и метод глубокого обучения (*англ.* deep learning, DL) на основе архитектуры U-Net. В *табл. 4* представлены результаты обучения и оценки качества рассмотренных подходов, которые показывают преимущество метода глубокого обучения.

Таблица 4. Точность анализа снимков

Тип метода	Название	Точность
ML	Random Forest	0,44301
DL	U-Net	0,99814

Общая схема реализации задачи дифференцированного внесения азотсодержащих удобрений на основе созданного датасета представлена на *рис. 3*.

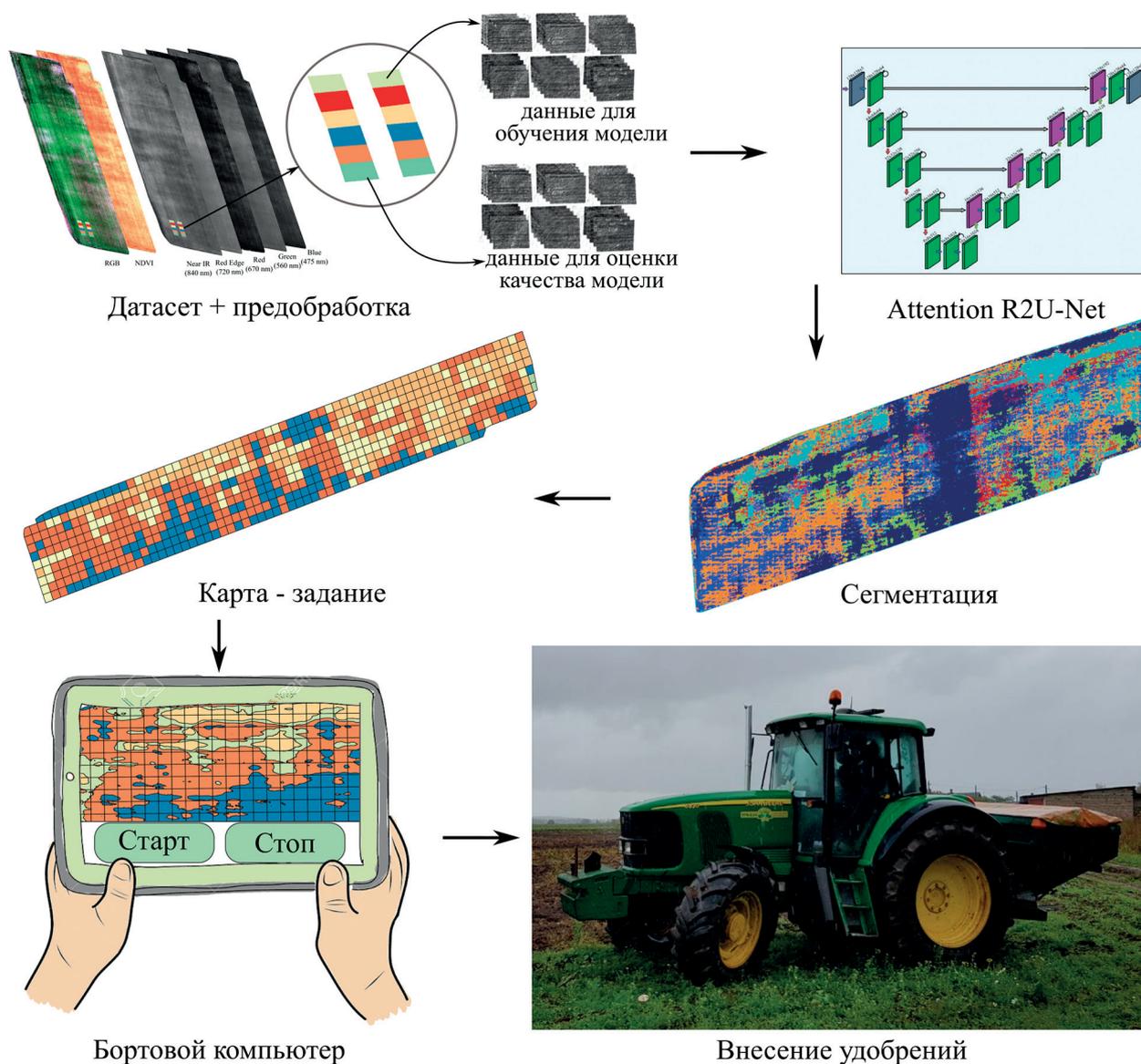


Рис. 3. Общая схема одной из задач применения созданного датасета

На первом этапе из набора данных извлекается информация о тестовых площадках (небольшие участки с известной дозой азота — 0, 40, 60, 80, 120, 200 кг действующего вещества на 1 га) и дополнительно обрабатывается: снимки каждой площадки разрезаются на изображения размером 128×128 пикселей, при этом из первой линии формируется набор обучающей выборки, а из второй — набор тестовой выборки. Следует отметить, что этот шаг необходим только при обучении модели, при дальнейшем использовании на вход подаются исходные ортофотопланы. С помощью обученной нейросетевой модели изображение сегментируется по азотному режиму посевов, на основе этого результата строится специализированная карта-задание, где каждому элементарному участку назначается необходимая доза удобрения. Полученная карта загружается в бортовой компьютер сельскохозяйственной техники, которая выезжает на поле и в соответствии с заданием дифференцированно вносит агрохимикат.

## Заключение

В работе представлен специализированный набор больших данных на основе аэрофотосъемки опытных сельскохозяйственных полей, расположенных в Ленинградской обл. Выделяются следующие основные преимущества созданного датасета:

- большой период охвата (более 15 лет);
- мультиспектральная съемка;
- высокое качество данных (пространственное разрешение 1–10 см/пиксель; построены ортофотопланы, цифровые модели рельефа местности, карты NDVI и т. п.);
- экспертная разметка данных;
- съемка на основных культурах Северо-Западного региона (зерновые, картофель и др.).

Датасет активно используется исследователями в различных задачах точного земледелия. Представленный набор данных закрыт для общего использования, для доступа к нему необходимо обращаться в АФИ с запросом. В работе в качестве демонстрационного примера кратко рассмотрена задача определения азотного режима посевов на основе анализа аэрофотоснимков методами машинного и глубокого обучения. По результатам исследования прослеживается преимущество применения нейросетевых подходов в рассмотренной задаче.

## Литература

1. Блеканов И. С., Молин А. Е., Разумилов Е. С., Ершов Е. А. Мониторинг азотного питания зерновых культур на основе нейросетевых технологий // Применение средств дистанционного зондирования Земли в сельском хозяйстве: материалы 3-й Всерос. науч. конф. с международ. участием. СПб.: ФГБНУ АФИ, 2021. С. 141–145.
2. Блохин Ю. И., Белов А. В., Блохина С. Ю. Комплексная система контроля влажности почвы и локальных метеоусловий для интерпретации данных дистанционного зондирования // Современные проблемы дистанционного зондирования Земли из космоса. 2019. Т. 16. № 3. С. 87–95. DOI: 10.21046/2070-7401-2019-16-3-87-95.
3. Блохин Ю. И., Якушев В. В., Блохина С. Ю. и др. Современные решения для формирования опорной информации с целью повышения точности определения агрофизических свойств почвы по спутниковым данным // Современные проблемы дистанционного зондирования Земли из космоса. 2020. Т. 17. № 4. С. 164–178. DOI: 10.21046/2070-7401-2020-17-4-164-178.
4. Шпанев А. М. Влияние минеральных удобрений на пространственное размещение сорных растений в посевах ярового ячменя // Плодородие. 2022. № 2. С. 8–12. DOI: 10.25680/S19948603.2022.125.02.
5. Якушев В. П., Буре В. М., Митрофанова О. А. и др. Оценка внутривополевой изменчивости посевов с помощью вариограммного анализа спутниковых данных для точного земледелия // Современные проблемы дистанционного зондирования Земли из космоса. 2020. Т. 17. № 2. С. 114–122. DOI: 10.21046/2070-7401-2020-17-2-114-122.

6. *Belcore E., Angeli S., Colucci E. et al.* Precision agriculture workflow, from data collection to data management using FOSS tools: An application in Northern Italy vineyard // Intern. J. Geo-Information. 2021. V. 10. No. 4. Article 236. DOI: 10.3390/ijgi10040236.
7. *Blekanov I., Molin A., Zhang D. et al.* Monitoring of grain crops nitrogen status from UAV multispectral images coupled with deep learning approaches // Computers and Electronics in Agriculture. 2023. V. 212. Article 108047. DOI: 10.1016/j.compag.2023.108047.
8. *Bosilj P., Aptoula E., Duckett T., Cielniak G.* Transfer learning between crop types for semantic segmentation of crops versus weeds in precision agriculture // J. Field Robotics. 2020. V. 37. P. 7–19. DOI: 10.1002/rob.21869.
9. *Breiman L.* Random Forests // Machine Learning. 2001. V. 45. P. 5–32. DOI: 10.1023/A:1010933404324.
10. *Chebrolu N., Lottes P., Schaefer A. et al.* Agricultural robot dataset for plant classification, localization and mapping on sugar beet fields // Intern. J. Robotics Research. 2017. V. 36. No. 10. P. 1045–1052. DOI: 10.1177/0278364917720510.
11. *Chiu M. T., Xu X., Wei Y., Huang Z., Schwing A., Brunner R., Khachatryan H., Karapetyan H., Dozier I., Rose G., Wilson D., Tudor A., Hovakimyan N., Huang T. S., Shi H.* Agriculture-vision: A large aerial image database for agricultural pattern analysis // arXiv:2001.01306 [cs.CV]. 2020.
12. *Haug S., Ostermann J.* A Crop/Weed Field Image dataset for the evaluation of computer vision based precision agriculture tasks // ECCV 2014: Computer Vision — ECCV 2014 Workshops. Lecture Notes in Computer Science. 2014. V. 8928. P. 105–116. DOI: 10.1007/978-3-319-16220-1\_8.
13. *Jiang Y., Li C., Paterson A. H., Robertson J. S.* DeepSeedling: deep convolutional network and Kalman filter for plant seedling detection and counting in the field // Plant Methods. 2019. V. 15. Article 141. DOI: 10.1186/s13007-019-0528-3.
14. *Jung J., Maeda M., Chang A. et al.* The potential of remote sensing and artificial intelligence as tools to improve the resilience of agriculture production systems // Current Opinion in Biotechnology. 2021. V. 70. P. 15–22. DOI: 10.1016/j.copbio.2020.09.003.
15. *Kallimani C., Dehkordi R. H., van Evert F. K. et al.* UAV-based Multispectral and Thermal dataset for exploring the diurnal variability, radiometric and geometric accuracy for precision agriculture // Open Data J. Agricultural Research. 2020. V. 6. 7 p. DOI: 10.18174/odjar.v6i0.16317.
16. *Kumar A., Taparia M., Rajalakshmi P. et al.* UAV based remote sensing for tassel detection and growth stage estimation of maize crop using multispectral images // IGARSS 2020 — 2020 IEEE Intern. Geoscience and Remote Sensing Symp. 2020. P. 1588–1591. DOI: 10.1109/IGARSS39084.2020.9323266.
17. *Lu J., Tan L., Jiang H.* Review on convolutional neural network (CNN) applied to plant leaf disease classification // Agriculture. 2021. V. 11. Article 707. DOI: 10.3390/agriculture11080707.
18. *Madsen S. L., Mathiassen S. K., Dyrmann M. et al.* Open plant phenotype database of common weeds in Denmark // Remote Sensing. 2020. V. 12. No. 8. DOI: 10.3390/RS12081246.
19. *Matvejenko D., Komarov A., Lekomtsev P., Nugis E.* Optimization of options for differential application of nitrogen fertilizers in the precision farming system // Agraarteadus: J. Agricultural Science. 2020. V. 31. No. 1. P. 41–46. DOI: 10.15159/jas.20.06.
20. *Mortensen A. K., Skovsen S., Karstoft H., Gislum R.* The Oil Radish Growth dataset for semantic segmentation and yield estimation // 2019 IEEE/CVF Conf. Computer Vision and Pattern Recognition Workshop. Long Beach, California, US, 2019. 8 p. DOI: 10.1109/CVPRW.2019.00328.
21. *Olsen A., Konovalov D. A., Philippa B. et al.* DeepWeeds: A multiclass weed species image dataset for deep learning // Scientific Reports. 2019. V. 9. No. 1. Article 2058. DOI: 10.1038/s41598-018-38343-3.
22. *Pereira P. C. J., Monteiro A., von Wangenheim A.* Weed Mapping on Aerial Images — A Systematic Literature Review: Technical Report INCoD/LAPIX.01.2019.E / Federal University of Santa Catarina. 2019. DOI: 10.13140/RG.2.2.34979.71204.
23. *Ronneberger O., Fischer P., Brox T.* U-Net: convolutional networks for biomedical image segmentation // Medical Image Computing and Computer-Assisted Intervention — MICCAI 2015. Lecture Notes in Computer Science. 2015. V. 9351. DOI: 10.1007/978-3-319-24574-4\_28.
24. *Sa I., Chen Z., Popovic M. et al.* weedNet: Dense semantic weed classification using multispectral images and MAV for smart farming // IEEE Robotics and Automation Letters. 2017. V. 3. No. 1. P. 588–595. DOI: 10.1109/LRA.2017.2774979.
25. *Sa I., Popovic M., Khanna R. et al.* WeedMap: A large-scale semantic weed mapping framework using aerial multispectral imaging and deep neural network for precision farming // Remote Sensing. 2018. V. 10. No. 9. Article 1423. DOI: 10.3390/rs10091423.
26. *Sinde-González I., Gil-Docampo M., Arza-García M. et al.* Biomass estimation of pasture plots with multi-temporal UAV-based photogrammetric surveys // Intern. J. Applied Earth Observations and Geoinformation. 2021. V. 101. Article 102355. DOI: 10.1016/j.jag.2021.102355.
27. *Song Z., Zhang Z., Yang S. et al.* Identifying sunflower lodging based on image fusion and deep semantic segmentation with UAV remote sensing imaging // Computers and Electronics in Agriculture. 2020. V. 179. Article 105812. DOI: 10.1016/j.compag.2020.105812.

28. The future of food and agriculture — Trends and challenges / Food and Agriculture Organization of the United Nations. Rome: FAO, 2017. 163 p.
29. Tsouros D. C., Bibi S., Sarigiannidis P. G. A review on UAV-based applications for precision agriculture // Information. 2019. V. 10. No. 11. Article 349. DOI: 10.3390/info10110349.
30. Weiss M., Jacob F., Duveiller G. Remote sensing for agricultural applications: A meta-review // Remote Sensing of Environment. 2020. V. 236. Article 111402. DOI: 10.1016/j.rse.2019.111402.
31. Wheeler T., Von Braun J. Climate change impacts on global food security // Science. 2013. V. 341. No. 6145. P. 508–513. DOI: 10.1126/science.1239402.
32. Wiesner-Hanks T., Stewart E. L., Kaczmar N. et al. Image set for deep learning: field images of maize annotated with disease symptoms // BMC Research Notes. 2018. V. 11. Article 440. DOI: 10.1186/s13104-018-3548-6.
33. Yang M. D., Tseng H. H., Hsu Y. C. et al. A UAV open dataset of rice paddies for deep learning practice // Remote Sensing. 2021. V. 13. Article 1358. DOI: 10.3390/rs13071358.
34. Ye H., Huang W., Huang S. et al. Recognition of banana fusarium wilt based on UAV remote sensing // Remote Sensing. 2020. V. 12. No. 6. Article 938. DOI: 10.3390/rs12060938.
35. Zhang C., Kovacs J. M. The application of small unmanned aerial systems for precision agriculture: a review // Precision Agriculture. 2012. V. 13. P. 693–712. DOI: 10.1007/s11119-012-9274-5.
36. Zheng H., Zhou X., He J. et al. Early season detection of rice plants using RGB, NIR-G-B and multi-spectral images from unmanned aerial vehicle (UAV) // Computers and Electronics in Agriculture. 2020. V. 169. Article 105223. DOI: 10.1016/j.compag.2020.105223.
37. Zhong Y., Hu X., Luo C. et al. WHU-Hi: UAV-borne hyperspectral with high spatial resolution ( $H^2$ ) benchmark datasets and classifier for precise crop identification based on deep convolutional neural network with CRF // Remote Sensing of Environment. 2020. V. 250. Article 112012. DOI: 10.1016/j.rse.2020.112012.

## Specialized dataset of multispectral aerophotos for solving precision farming problems using artificial intelligence methods

O.A. Mitrofanova<sup>1</sup>, E. P. Mitrofanov<sup>1</sup>, I. S. Blekanov<sup>2</sup>, A. E. Molin<sup>2</sup>

<sup>1</sup> Agrophysical Research Institute, Saint Petersburg 195220, Russia  
E-mails: omitrofa@gmail.com, mjeka89@gmail.com

<sup>2</sup> Saint Petersburg State University, Saint Petersburg 199034, Russia

Due to fast evolution of information technology, the task of creating large quality remote sensing datasets is becoming increasingly important. At the Agrophysical Research Institute (AFI), artificial intelligence methods and image analysis have been used for about 20 years. During this period, a large amount of information has been collected to solve the problems of precision farming. The object of the presented study is an experimental AFI bio-polygon, located in Leningrad Region, which consists of 29 fields. For collection, an unmanned aerial vehicle developed by the AFI, as well as an unmanned aerial system Geoscan-401, was used. The shooting was carried out in five spectra: red, green, blue, infrared and red edge, the average flight height was 80 meters, the spatial resolution of the images was 1–10 cm/pixel. The paper considers in detail the generated algorithms of data collection and preprocessing. In the study, as a demonstration of the applicability of the created marked-up dataset, an experiment was conducted to analyze the orthophotomap of one of the fields of the biopolygon for the period 2019–2021, compiled from images of the Micasense RedEdge MX multispectral camera. The method of the classical Random Forest machine learning algorithm adapted to the task and the deep learning method based on the U-Net architecture were used as image analysis methods. The results of the experiment demonstrated the advantage of the deep learning method in solving the problem of determining the nitrogen regime of crops for differentiated fertilization.

**Keywords:** dataset, image analysis, multispectral aerial photography, precision farming, machine learning, deep learning

Accepted: 06.10.2023

DOI: 10.21046/2070-7401-2023-20-6-67-79

## References

1. Blekanov I. S., Molin A. E., Razumilov E. S., Ershov E. A., Monitoring of nitrogen nutrition of grain crops based on neural network technologies, *Primenenie sredstv distantsionnogo zondirovaniya Zemli v sel'skom khozyaistve* (Application of Earth Remote Sensing Means in Agriculture), Proc. 3<sup>rd</sup> All-Russia Science Conf. with Intern. Participation, Saint Petersburg: FGBNU AFI, 2021, pp. 141–145 (in Russian).
2. Blokhin Yu. I., Belov A. V., Blokhina S. Yu., Integrated system for control of soil moisture and local weather conditions for remote sensing data interpretation, *Sovremennye problemy distantsionnogo zondirovaniya Zemli iz kosmosa*, 2019, Vol. 16, No. 3, pp. 87–95 (in Russian), DOI: 10.21046/2070-7401-2019-16-3-87-95.
3. Blokhin Yu. I., Yakushev V. V., Blokhina S. Yu. et al., New solutions for the reference data formation to improve the accuracy of the agrophysical soil properties determination from satellite data, *Sovremennye problemy distantsionnogo zondirovaniya Zemli iz kosmosa*, 2020, Vol. 17, No. 4, pp. 164–178 (in Russian), DOI: 10.21046/2070-7401-2020-17-4-164-178.
4. Shpanev A. M., Effect of mineral fertilizers on the spatial distribution of weeds in spring barley crops, *Plodorodie*, 2022, No. 2, pp. 8–12 (in Russian), DOI: 10.25680/S19948603.2022.125.02.
5. Yakushev V. P., Bure V. M., Mitrofanova O. A. et al., Within-field variability estimation based on variogram analysis of satellite data for precision agriculture, *Sovremennye problemy distantsionnogo zondirovaniya Zemli iz kosmosa*, 2020, Vol. 17, No. 2, pp. 114–122 (in Russian), DOI: 10.21046/2070-7401-2020-17-2-114-122.
6. Belcore E., Angeli S., Colucci E. et al., Precision agriculture workflow, from data collection to data management using FOSS tools: An application in Northern Italy vineyard, *Intern. J. Geo-Information*, 2021, Vol. 10, No. 4, Article 236, DOI: 10.3390/ijgi10040236.
7. Blekanov I., Molin A., Zhang D. et al., Monitoring of grain crops nitrogen status from UAV multispectral images coupled with deep learning approaches, *Computers and Electronics in Agriculture*, 2023, Vol. 212, Article 108047, DOI: 10.1016/j.compag.2023.108047.
8. Bosilj P., Aptoula E., Duckett T., Cielniak G., Transfer learning between crop types for semantic segmentation of crops versus weeds in precision agriculture, *J. Field Robotics*, 2020, Vol. 37, pp. 7–19, DOI: 10.1002/rob.21869.
9. Breiman L., Random Forests, *Machine Learning*, 2001, Vol. 45, pp. 5–32, DOI: 10.1023/A:1010933404324.
10. Chebrolu N., Lottes P., Schaefer A. et al., Agricultural robot dataset for plant classification, localization and mapping on sugar beet fields, *Intern. J. Robotics Research*, 2017, Vol. 36, No. 10, pp. 1045–1052, DOI: 10.1177/0278364917720510.
11. Chiu M. T., Xu X., Wei Y., Huang Z., Schwing A., Brunner R., Khachatryan H., Karapetyan H., Dozier I., Rose G., Wilson D., Tudor A., Hovakimyan N., Huang T. S., Shi H., Agriculture-vision: A large aerial image database for agricultural pattern analysis, *arXiv:2001.01306 [cs.CV]*, 2020.
12. Haug S., Ostermann J., A Crop/Weed Field Image dataset for the evaluation of computer vision based precision agriculture tasks, *ECCV 2014: Computer Vision — ECCV 2014 Workshops, Lecture Notes in Computer Science*, 2014, Vol. 8928, pp. 105–116, DOI: 10.1007/978-3-319-16220-1\_8.
13. Jiang Y., Li C., Paterson A. H., Robertson J. S., DeepSeedling: deep convolutional network and Kalman filter for plant seedling detection and counting in the field, *Plant Methods*, 2019, Vol. 15, Article 141, DOI: 10.1186/s13007-019-0528-3.
14. Jung J., Maeda M., Chang A. et al., The potential of remote sensing and artificial intelligence as tools to improve the resilience of agriculture production systems, *Current Opinion in Biotechnology*, 2021, Vol. 70, pp. 15–22, DOI: 10.1016/j.copbio.2020.09.003.
15. Kallimani C., Dehkordi R. H., van Evert F. K. et al., UAV-based Multispectral and Thermal dataset for exploring the diurnal variability, radiometric and geometric accuracy for precision agriculture, *Open Data J. Agricultural Research*, 2020, Vol. 6, 7 p., DOI: 10.18174/odjar.v6i0.16317.
16. Kumar A., Taparia M., Rajalakshmi P. et al., UAV based remote sensing for tassel detection and growth stage estimation of maize crop using multispectral images, *IGARSS 2020 — 2020 IEEE Intern. Geoscience and Remote Sensing Symp.*, 2020, pp. 1588–1591, DOI: 10.1109/IGARSS39084.2020.9323266.
17. Lu J., Tan L., Jiang H., Review on convolutional neural network (CNN) applied to plant leaf disease classification, *Agriculture*, 2021, Vol. 11, Article 707, DOI: 10.3390/agriculture11080707.
18. Madsen S. L., Mathiassen S. K., Dyrmann M. et al., Open plant phenotype database of common weeds in Denmark, *Remote Sensing*, 2020, Vol. 12, No. 8, DOI: 10.3390/RS12081246.
19. Matvejenko D., Komarov A., Lekomtsev P., Nugis E., Optimization of options for differential application of nitrogen fertilizers in the precision farming system, *Agraarteadus: J. Agricultural Science*, 2020, Vol. 31, No. 1, pp. 41–46, DOI: 10.15159/jas.20.06.
20. Mortensen A. K., Skovsen S., Karstoft H., Gislum R., The Oil Radish Growth dataset for semantic segmentation and yield estimation, *2019 IEEE/CVF Conf. Computer Vision and Pattern Recognition Workshop*, Long Beach, California, US, 2019, 8 p., DOI: 10.1109/CVPRW.2019.00328.

21. Olsen A., Konovalov D. A., Philippa B. et al., DeepWeeds: A multiclass weed species image dataset for deep learning, *Scientific Reports*, 2019, Vol. 9, No. 1, Article 2058, DOI: 10.1038/s41598-018-38343-3.
22. Pereira P. C. J., Monteiro A., von Wangenheim A., *Weed Mapping on Aerial Images — A Systematic Literature Review: Technical Report INCoD/LAPIX.01.2019.E*, Federal University of Santa Catarina, 2019, DOI: 10.13140/RG.2.2.34979.71204.
23. Ronneberger O., Fischer P., Brox T., U-Net: convolutional networks for biomedical image segmentation, *Medical Image Computing and Computer-Assisted Intervention — MICCAI 2015, Lecture Notes in Computer Science*, 2015, Vol. 9351, DOI: 10.1007/978-3-319-24574-4\_28.
24. Sa I., Chen Z., Popovic M. et al., weedNet: Dense semantic weed classification using multispectral images and MAV for smart farming, *IEEE Robotics and Automation Letters*, 2017, Vol. 3, No. 1, pp. 588–595, DOI: 10.1109/LRA.2017.2774979.
25. Sa I., Popovic M., Khanna R. et al., WeedMap: A large-scale semantic weed mapping framework using aerial multispectral imaging and deep neural network for precision farming, *Remote Sensing*, 2018, Vol. 10, No. 9, Article 1423, DOI: 10.3390/rs10091423.
26. Sinde-Gonzalez I., Gil-Docampo M., Arza-Garcia M. et al., Biomass estimation of pasture plots with multitemporal UAV-based photogrammetric surveys, *International J. Applied Earth Observations and Geoinformation*, 2021, Vol. 101, Article 102355, DOI: 10.1016/j.jag.2021.102355.
27. Song Z., Zhang Z., Yang S. et al., Identifying sunflower lodging based on image fusion and deep semantic segmentation with UAV remote sensing imaging, *Computers and Electronics in Agriculture*, 2020, Vol. 179, Article 105812, DOI: 10.1016/j.compag.2020.105812.
28. *The future of food and agriculture — Trends and challenges*, Food and Agriculture Organization of the United Nations, Rome: FAO, 2017, 163 p.
29. Tsouros D. C., Bibi S., Sarigiannidis P. G., A review on UAV-based applications for precision agriculture, *Information*, 2019, Vol. 10, No. 11, Article 349, DOI: 10.3390/info10110349.
30. Weiss M., Jacob F., Duveiller G., Remote sensing for agricultural applications: A meta-review, *Remote Sensing of Environment*, 2020, Vol. 236, Article 111402, DOI: 10.1016/j.rse.2019.111402.
31. Wheeler T., Von Braun J., Climate change impacts on global food security, *Science*, 2013, Vol. 341, No. 6145, pp. 508–513, DOI: 10.1126/science.1239402.
32. Wiesner-Hanks T., Stewart E. L., Kaczmar N. et al., Image set for deep learning: field images of maize annotated with disease symptoms, *BMC Research Notes*, 2018, Vol. 11, Article 440, DOI: 10.1186/s13104-018-3548-6.
33. Yang M. D., Tseng H. H., Hsu Y. C. et al., A UAV open dataset of rice paddies for deep learning practice, *Remote sensing*, 2021, Vol. 13, Article 1358, DOI: 10.3390/rs13071358.
34. Ye H., Huang W., Huang S. et al., Recognition of banana fusarium wilt based on UAV remote sensing, *Remote Sensing*, 2020, Vol. 12, No. 6, Article 938, DOI: 10.3390/rs12060938.
35. Zhang C., Kovacs J. M., The application of small unmanned aerial systems for precision agriculture: a review, *Precision Agriculture*, 2012, Vol. 13, pp. 693–712, DOI: 10.1007/s11119-012-9274-5.
36. Zheng H., Zhou X., He J. et al., Early season detection of rice plants using RGB, NIR-G-B and multi-spectral images from unmanned aerial vehicle (UAV), *Computers and Electronics in Agriculture*, 2020, Vol. 169, Article 105223, DOI: 10.1016/j.compag.2020.105223.
37. Zhong Y., Hu X., Luo C. et al., WHU-Hi: UAV-borne hyperspectral with high spatial resolution ( $H^2$ ) benchmark datasets and classifier for precise crop identification based on deep convolutional neural network with CRF, *Remote Sensing of Environment*, 2020, Vol. 250, Article 112012, DOI: 10.1016/j.rse.2020.112012.