Генеративные технологии для создания обучающих данных в задачах ДЗЗ*

А.В. Смирнов, И.П. Тищенко

Федеральное государственное бюджетное учреждение науки Институт программных систем им. А.К. Айламазяна Российской академии наук, г. Переславль-Залесский, Россия

Аннотация. В работе представлен метод генерации обучающих наборов данных для задач дистанционного зондирования Земли (ДЗЗ) с использованием генеративных инструментов, таких как Stable Diffusion. В частности, рассматривается задача поиска и распознавания лесных пожаров, незаконных вырубок и несанкционированных свалок на снимках, полученных с БПЛА. В ходе настоящего исследования был сгенерирован обучающий набор данных, состоящий из трех классов, на котором была обучена модель семейства YOLO. При инференсе модели на реальных снимках, найденных в Интернете, были достигнуты значения точности в 72,4 и 65,8% по метрикам F1 и mAP50, соответственно. На основании полученных результатов, сделан вывод о целесообразности применения представленного метода, как минимум, в качестве инструмента аугментации существующих данных.

Ключевые слова: *генерация набора данных, нейросетевая модель, анализ снимков, дистанционное зондирование Земли, беспилотный летательный аппарат, лесные пожары, незаконные вырубки, несанкционированные свалки.*

DOI: 10.14357/20790279260305 **EDN:** QGKCZA

Введение

Сфера применения дистанционного зондирования Земли (ДЗЗ) очень широка и охватывает множество областей, в том числе такие направления, как «Экология и охрана природы», «Коммунальное хозяйство» и «Безопасность и оборона». Одним из наиболее актуальных методов ДЗЗ является анализ снимков, полученных с беспилотных летательных аппаратов (БПЛА), за счет того, что при их использовании можно выполнить аэрофотосъемку труднодоступных и отдаленных районов с необходимой частотой и меньшими затратами по сравнению со спутниковым мониторингом.

Благодаря развитию технологий машинного обучения появилась возможность проводить анализ снимков, полученных с помощью БПЛА, с использованием искусственных нейронных сетей (ИНС). Основными объектами интереса при анализе снимков БПЛА стали экологические ситуации, наиболее значимые для состояния окружающей среды. К таким ситуациям можно отнести лесные пожары, незаконные вырубки деревьев и несанкционированные свалки отходов жизнедеятельности.

Одной из проблем нейросетевого поиска и распознавания лесных пожаров на снимках с БПЛА является отсутствие в необходимом количестве общедоступных наборов данных, особенно с выполненной разметкой целевых областей. В работе [1] предложен набор данных, составленный из изображений, полученных с БПЛА в финских бореальных лесах, с аннотациями, выполненными как человеком, так и с использованием алгоритмов компьютерного зрения. Все изображения были собраны во время нескольких контролируемых выжиганий и были успешно использованы для обучения нейросетевых моделей.

Еще один набор данных UAVs-FFDB для распознавания лесных пожаров был предложен в статье [2]. Он включает в себя 1653 высококачественных RGB-изображения, полученных с использованием стандартного квадрокоптера S500 [3] в сочетании с камерой RaspiCam V2 [4]. Также набор содержит дополненные данные, в результате чего общее количество изображений достигает 15560, что повышает его разнообразие и полноту. Исходные изображения в наборе имеют разрешение от 353x314 до 640x480 пикселей и были сня-

ты с высоты от 5 до 15 метров. Все изображения были аннотированы с использованием платформы Makesense.ai [5], что позволило точно определить границы пожаров.

Задача обнаружения лесных пожаров была рассмотрена в статье [6]. В своем исследовании авторы предлагают новый набор данных FLAME (Fire Luminosity Airborne-based Machine learning Evaluation) [7], который включает покадрово размеченные видеозаписи и тепловые карты, полученные с помощью инфракрасных камер. В статье также рассматриваются задачи бинарной классификации изображений на наличие/отсутствие пожара и обнаружения пожара с использованием методов сегментации для точного определения его границ. Используемый авторами метод бинарной классификации на основе искусственных нейронных сетей (ИНС) достиг точности в 76%, а разработанный метод глубокого обучения на базе U-Net [8] показал 92% по метрике Precision [9] и 84% по Recall [9].

Технико-экономическая оценка применения БПЛА для обнаружения незаконных мусорных свалок представлена в исследовании [10]. Согласно опубликованному исследованию, использование дронов может снизить затраты на поиск и обработку незаконных свалок на 58 миллионов долларов США, а в некоторых случаях свести их к нулю. Это подтверждает первоначальное предположение авторов о том, что БПЛА могут быть эффективным инструментом для идентификации различных типов отходов, размещенных на земле в пределах нелегальных свалок.

Исследование, опубликованное в статье [11], охватывает проблему отсутствия релевантных наборов данных для поиска и обнаружения незаконных свалок. С целью решения данной проблемы авторы представляют новый набор данных AerialWaste, изображения для которого собраны из трех различных источников: ортофотоснимки AGEA [12], снимки WorldView-3 [13] и Google Earth [14]. Из этих изображений 3478 являются положительными примерами, представляющими подозрительные места свалок, а 6956 — отрицательными. Все изображения в наборе данных аннотированы в формате MS COCO [15]. Используемый авторами классификатор на основе ResNet50 при бинарной классификации показал точность в 80,7% по метрике F1 [16].

Так как наличие незаконных свалок мусора может нанести реальный вред экологической обстановке на определенной территории, была создана специализированная платформа экологического мониторинга Raven Scan [17]. Данная платформа использует передовые технологии получения изображений с помощью БПЛА и спутников, а также обеспечивает интеграцию с моделями искусственного интеллекта (ИИ) для разметки изображений, тем самым предоставляя исчерпывающий инструментарий для выявления незаконных свалок.

Вырубка лесов оказывает многогранное воздействие на окружающую среду, биоразнообразие и климат. Мониторинг вырубки лесов является важным шагом на пути к управлению и смягчению ее последствий. Наряду с традиционными методами мониторинга, для отслеживания состояния леса используется аэрофотосъемка, в том числе с применением БПЛА и ИНС. Так, в работе [18] представлена новая система Интернета вещей (IoT) [19] на основе БПЛА для обнаружения звука бензопилы. Система использует недорогие IoT-узлы с микрофонами для обнаружения звуков бензопилы с помощью модели сверточной нейронной сети (СНС) [20]. При обнаружении дрон активирует звуковые и световые сигналы тревоги для предотвращения незаконной вырубки леса. Результаты экспериментов показывают, что узлы IoT могут обнаруживать звуки бензопилы на расстоянии до 100 м, при этом нейросетевая модель достигает точности 99,37 %.

Оценка масштаба незаконной вырубки лесов с помощью картографирования, выполненного с использованием БПЛА, предлагается в статье [21]. Авторы заявляют, что высокое разрешение изображений, полученных с дрона, позволяет создавать подробные трехмерные базовые карты лесов и дает возможность количественно оценить потери лесных массивов. Потенциально данное решение позволит местным неправительственным организациям использовать полученные карты в сочетании с еженедельными оповещениями о вырубке лесов для оперативного реагирования на местах.

Исходя из представленных работ можно сделать вывод о том, что использование БПЛА для решения различных задач ДЗЗ актуально и пользуется популярностью. Применение различных нейросетевых методов и алгоритмов для анализа снимков, полученных с БПЛА, также имеет место и все чаще используется благодаря развитию технологий компьютерного зрения (КЗ) и машинного обучения (МО). Тем не менее, сейчас наблюдается нехватка релевантных обучающих наборов данных для поиска и распознавания различных экологических ситуаций, таких как лесные пожары, незаконные вырубки деревьев и несанкционированные свалки отходов жизнедеятельности. В настоящем исследовании рассмотрено применение генеративных технологий для создания обучающего набора данных. Результаты проведенного исследования

могут быть использованы как для генерации данных, так и в качестве способа их аугментации.

1. Актуальность, цель и задачи исследования

Целью настоящего исследования является оценка практичности применения генеративных нейросетевых технологий для создания обучающего набора данных при выполнении задач ДЗЗ. В частности, рассматривается поиск и распознавание лесных пожаров, незаконных вырубок деревьев и несанкционированных свалок отходов жизнедеятельности с использованием снимков, полученных с БПЛА.

Таким образом, основная задача настоящего исследования заключается в генерации обучающего набора данных, состоящего из трех вышеперечисленных классов изображений с имитацией ракурса съемки характерного для БПЛА. Следующей задачей является разметка полученного набора данных и обучение нейросетевой модели для распознавания целевых объектов (очаги пожаров, места вырубки деревьев и скопления мусора) методом семантической сегментации.

Непосредственная оценка целесообразности применения генеративных технологий для ДЗЗ возможна при анализе результатов инференса модели на реальных снимках. Если достигнутые значения точности по метрикам F1 и mAP50 [22] будут не ниже 70 и 65%, соответственно, то использование генеративных технологий будет оправдано, по крайней мере, как способ расширения существующих обучающих наборов данных.

2. Используемые нейросетевые решения

В качестве инструмента для генерации изображений была использована модель Stable Diffusion. На текущий момент Stable Diffusion является одной из самых продвинутых нейронных сетей для генерации изображений. В основном данная модель используется для генерации детализированных изображений на основе текстовых описаний, хотя ее также можно применять к другим задачам обработки изображений. В августе 2022 модель Stable Diffusion вышла в публичный доступ от группы Stability AI [23]. Исходный код модели также опубликован на платформе GitHub.

Существует два основных варианта использования модели Stable Diffusion: онлайн через сайт или приложение и с помощью специального программного обеспечения (ПО) с графическим интерфейсом [24]. Для настоящего исследования был выбран второй вариант, так как он позволяет развернуть модель на собственном аппаратном обеспечении, а также обладает возможностью гибкой настройки параметров и быстрого просмотра результатов благодаря наличию графического интерфейса.

Несмотря на то, что выход оригинальной Stable Diffusion состоялся в 2022 году, одна из ее первых версий – 1.5, до сих пор остается самой популярной в сообществе. Исходя из этого, именно Stable Diffusion 1.5 была использована для генерации изображений в ходе настоящего исследования.

Следует также отметить основной критерий выбора данной модели в качестве генератора изображений. Им стала возможность бесплатного полноценного развертывания модели Stable Diffusion на собственном аппаратном обеспечении. Среди недостатков такого решения можно отметить относительно высокие системные требования к используемому компьютеру.

В роли нейросетевой модели для проведения экспериментов по распознаванию целевых объектов на сгенерированных изображениях выступала архитектура YOLO. Это широко известная нейросетевая архитектура, которая знаменита большим количеством версий: первая из которых вышла в 2016 году и решала только задачу детекции объектов на изображении, а последняя уже представляет собой целую фундаментальную модель, которую можно использовать для классификации, детекции, сегментации и определения поз.

Сейчас одной из последних итераций данной архитектуры является YOLOv11 [25]. В зависимости от количества гиперпараметров архитектура YOLOv11 подразделяется на несколько вариантов моделей, таких как: n – nano, s – small, m – medium, l – large и x – extra large. В рамках эксперимента для проведения настоящего исследования были выбраны варианты YOLOv11n, как самая быстрая и ресурсосберегающая модель, и YOLOv11m, как наиболее сбалансированная модель по показателям производительности и достигаемой точности.

3. Существующие наборы данных

На сегодняшний день уже существует несколько наборов данных для обучения нейросетевых алгоритмов в задачах анализа снимков ДЗЗ. Наиболее полные наборы посвящены поиску лесных пожаров. Так, в работе [26] представлен набор Wildfire Dataset, состоящий из 2700 снимков, разрешение которых варьируется от 153x206 до 19699x8974 пикселей. Набор разделен на три выборки для обучения, тестирования и валидации. Каждая выборка содержит по два класса изображе-

ний: 0 – есть огонь, 1 – нет огня. Также существуют наборы [27-29], созданные независимыми исследователями и находящиеся в открытом доступе. Количество изображений в таких наборах может варьироваться от нескольких сотен до нескольких тысяч.

Наборы данных с изображениями вырубки лесов встречаются реже и содержат меньше снимков. Тем не менее существует несколько наборов [30-32] с 69, 23 и 57 изображениями соответственно. Каждый из рассмотренных наборов, по мимо самих изображений, содержит аннотационные файлы разметки, что упрощает процесс обучения ИНС. Однако из-за малого количества снимков возникает проблема недостаточного количества обучающих образцов.

Для решения задач по обнаружению свалок с высоты полета БПЛА следует обратить внимание на набор данных Dataset for Waste/Dumpsite Detection using drone imagery [33]. В нем содержится 2115 снимков незаконных свалок, сделанных с помощью БПЛА с разрешением 1280x1280 пикселей. Все изображения в наборе аннотированы в формате COCO.

Несмотря на наличие обучающих наборов данных для задач распознавания и детектирования лесных пожаров, незаконных свалок и нелегальной вырубки деревьев, есть ряд недостатков и проблем, с которыми можно столкнуться при их использовании. К ним можно отнести следующие:

– Отсутствие единого разрешения изображений. Популярные нейросетевые архитектуры обычно имеют фиксированный размер входа. При подаче изображения несоответствующего разрешения его размер будет изменен к размеру входа. Изменение размера оригинального изображения может привести к потере значительного количества деталей.
– Ракурс съемки не всегда соответствует тому, что необходим для анализа ДЗЗ, так как часть изображений снята «с земли». Хоть это и позволяет разнообразить набор данных, ИНС, как правило, дают лучший результат, если обучались на данных, наиболее приближенных к реальной ситуации использования.
– Дисбаланс классов. Судя по статистике рассмотренных наборов данных, при одновременном распознавании лесных пожаров, незаконных свалок и нелегальных вырубок, перекос по количеству снимков будет в сторону изображений лесных пожаров, что может негативно сказаться на итоговой точности нейронной сети.

Обучающий набор данных, созданный с использованием генеративных нейросетевых инструментов, может быть лишен перечисленных недостатков. Во-первых, при генерации можно зафиксировать размер создаваемого изображения в соответствии с размером входа ИНС. Во-вторых, необходимый ракурс съемки достигается путем создания соответствующего запроса (промта). В-третьих, можно отслеживать точное количество изображений, сгенерированных для того или иного класса, что положительно влияет на баланс.

4. Генерация обучающего набор данных

Разрешение генерируемых изображений было установлено на 640x640 пикселей в соответствии с входным размером модели YOLOv11. Для генерации снимков использовались так называемые промты (от англ. prompt) – текстовые запросы, которые подавались на вход Stable Diffusion. Генерация изображений происходила для трех классов:

– forest_fire – изображения лесных пожаров, снятые с высоты полета БПЛА;
– illegal_dumps – изображения с незаконными мусорными свалками, снятые с высоты полета БПЛА;
– illegal_logging – изображения с зонами несанкционированной вырубки деревьев, снятые с высоты полета БПЛА.

При генерации использовался метод семплирования (Sampling method) – DPM++ 2M. Количество шагов семплирования (Sampling steps) варьировалось от 20 до 40. Параметр CFG Scale, отвечающий за то, на сколько точно модель будет следовать запросу, устанавливался в диапазоне от 5 до 10. Эти шаги позволили генерировать более разнообразные изображения, что потенциально положительно повлияло на релевантность итогового обучающего набора данных.

Для генерации изображений класса forest_fire использовались следующие текстовые запросы (prompts):

– A drone photo of a fire in a mixed forest.
– A drone photo of a fire in a forest.
– A drone photo of a fire at the edge of a forest.

В классе illegal_dumps содержались изображения, сгенерированные по таким запросам, как:

– A drone photo of an illegal dump.
– A drone photo of an illegal dump taken from a high altitude.
– A drone photo of an illegal dump in forest.

В свою очередь, для класса illegal_logging использовались следующие запросы:

– A drone photo of illegal logging.
– A drone photo of illegal tree felling.

– A drone photo of illegal logging taken from a high altitude.
– A drone photo of illegal tree felling taken from a high altitude.

В итоге было сгенерировано 150 снимков для каждого класса. На рис. 1 показаны примеры изображений каждого из классов.

Определить высоту и/или поле зрения съемки у сгенерированных изображений не представляется возможным. Оценить, насколько синтетические изображения похожи на реальные снимки можно лишь визуально. Тем не менее, данная ситуация не должна вносить каких-либо систематических смещений в модель, так как достигается некоторый уровень разнообразия обучающей выборки, благодаря чему она учится выделять универсальные закономерности.

Далее все полученные изображения размечались вручную с использованием специализированного ПО – LabelMe [34]. В классе forest_fire разме-

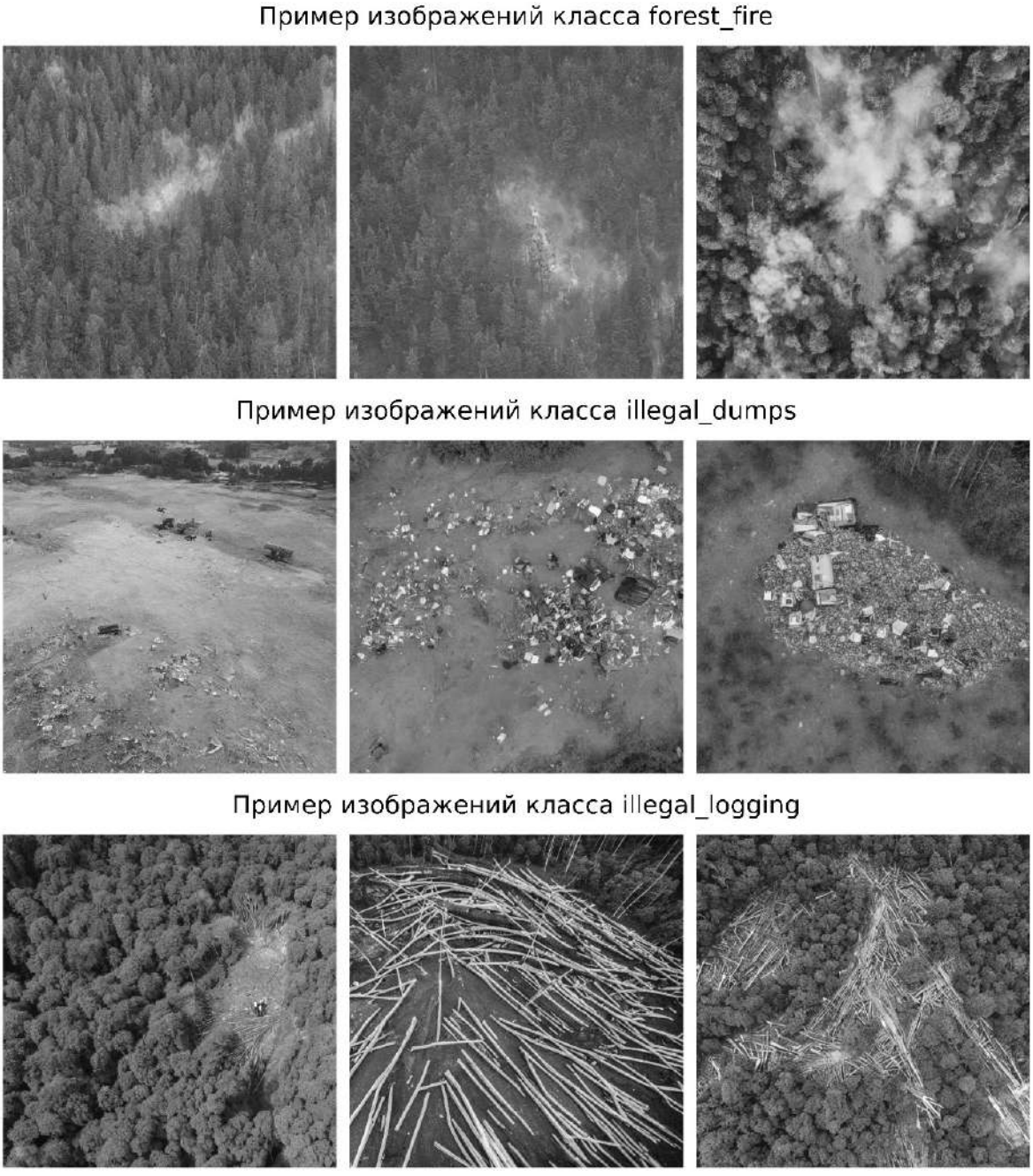


Рис. 1. Примеры сгенерированных изображений для каждого из классов

Табл. 1

Точность модели после валидации по метрикам Precision, Recall и mAP50

Модель	**YOLOv11n**			**YOLOv11m**		
Класс	Precision	Recall	mAP50	Precision	Recall	mAP50
Общее	0.897	0.806	0.869	0.871	0.788	0.847
forest_fire	0.883	0.905	0.918	0.826	0.876	0.881
illegal_dumps	0.831	0.670	0.758	0.857	0.643	0.770
illegal_logging	0.976	0.844	0.931	0.931	0.844	0.889

чались очаги пожара, охваченные огнем деревья, а также отдельные языки пламени. Для класса illegal_dumps отмечались места скопления мусора различного объема, а также некоторые отдельные объекты, внешний вид которых соответствовал мусору. Изображения класса illegal_logging размечались так, чтобы целевая область соответствовала зоне предполагаемой вырубки деревьев, характеризующейся наличием пней, а также поваленных веток и стволов деревьев.

Сгенерированный набор данных подвергся процессу искусственного расширения, в ходе которого к 50 случайным копиям исходных изображений каждого класса были применены процедуры случайного изменения яркости и отражения по горизонтали с сохранением разметки.

Полученный набор данных был преобразован в формат YOLO и разделен на обучающую и валидационную выборки по 450 и 150 изображений, соответственно.

5. Обучение модели

Реализация моделей YOLOv11, а также инструментарий для ее обучения и валидации на языке программирования Python были взяты из официальной библиотеки Ultralytics [35] версии 8.3.255. Обучение моделей происходило на протяжении 500 эпох. Остальные параметры обучения были установлены на значения по умолчанию. Результаты валидации по метрикам Precision, Recall и mAP50 для сегментации представлены в табл.1.

Исходя из полученных результатов валидации можно сделать вывод, что в среднем младшая модель (YOLOv11n) показывает более высокую точность по метрике mAP50. Более «тяжелая» модель YOLOv11m превосходит ее только для класса illegal_dumps. Затраченное время инференса на изображение у модели YOLOv11n также ниже и составляет 1.8 мс против 6.7 мс у модели YOLOv11m. Следовательно, использование модели YOLOv11n в реалистичных сценариях более рационально, так как в данном случае она показывает схожий со старшей моделью результат при кратно меньшем потреблении ресурсов.

6. Инференс модели и анализ результатов

Для выполнения инференса модели YOLOv11n, обученной на сгенерированном наборе данных, была собрана тестовая выборка, содержащая по 10 изображений на класс. Тестовые изображения представляли собой снимки, полученные с БПЛА, которые были в открытом доступе в сети Интернет.

Большинство изображений были прикреплены к соответствующим новостным статьям. Часть изображений располагалась на специальных платформах (фотобанках) без указания их точного происхождения. География съемки затрагивала различные регионы России, США, Турции, Бразилии, Украины и Австралии. Дата съемки используемых изображений варьировалась в диапазоне от конца 1988 г. вплоть до начла 2026 г. Изображения тестовой выборки не имели единого размера. Всего тестовая выборка насчитывала 30 изображений, каждое из которых было вручную размечено для получения эталонных координат целевых объектов (очагов лесных пожаров, мест незаконного складирования мусора и областей нелегальной вырубки деревьев).

Визуализация результатов и подсчет точности инференса по метрикам F1 и mAP50 были выполнены с использованием библиотеки Supervision [36]. Полученные показатели точности модели YOLOv11n при инференсе представлены в табл. 2.

Табл. 2

Результаты инференса модели по метрикам F1 и mAP50

Класс	**F1**	**mAP50**
Общее	0.724	0.658
forest_fire	0.688	0.561
illegal_dumps	0.759	0.691
illegal_logging	0.800	0.723

На рис. 2 показаны примеры детектирования целевых объектов (очагов лесных пожаров, мест незаконного складирования мусора и областей нелегальной вырубки деревьев).

В процессе инференса было достигнуто общее значение точности в 72.4% (0.724) по F1 и 65.8% (0.658) по mAP50. Данные показатели удовлетворяют требованиям настоящего исследования.

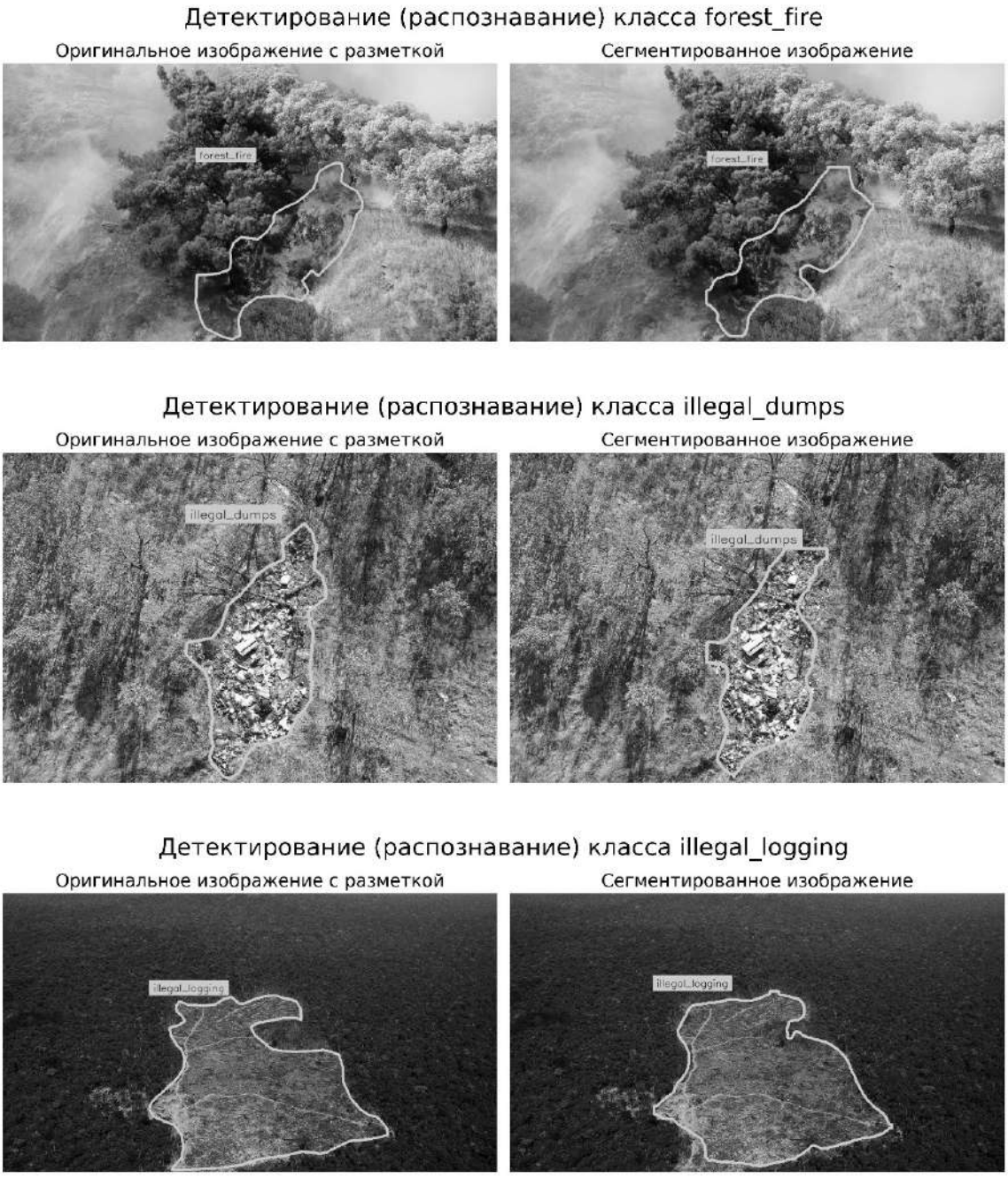


Рис. 2. Примеры детектирования целевых объектов

Стоит отметить, что наименьшая точность была получена для класса forest_fire – 68.8% (0.688) и 56.1% (0.561), соответственно. Вероятно это связано со специфическими особенностями отображения огня и языков пламени, при которых затруднительно определить их точные границы. К тому же, неясным остается вопрос, касающийся дыма и/или обугленных деревьев, а именно стоит ли добавлять их в текущий класс или сформировать новый. Подобные ситуации, хоть и в меньшей степени, также наблюдались в остальных классах, что негативно влияло на релевантность ручной разметки и достигнутую точность.

Чтобы подтвердить эффективность разработанного метода, был проведен сравнительный эксперимент. В качестве исходных данных был выбран рассмотренный ранее набор UAVs-FFDB. Он содержит как «сырые» данные, так и уже аугментированные методами поворота, сдвига и масштабирования. Для эксперимента случайным образом

Табл. 3

Набор	**UAVs-FFDB_aug**			**UAVs-FFDB_gen**		
Класс	Precision	Recall	mAP50	Precision	Recall	mAP50
forest_fire	0.906	0.783	0.783	0.822	0.808	0.875

были отобраны 500 «сырых» снимков, а также 100 аугментированных. Также были отобраны 100 сгенерированных снимков класса forest_fire. Разметка всех отобранных снимков была преобразована в формат YOLO для детекции. В итоге были получены два тестовых набора данных: UAVs-FFDB_aug и UAVs-FFDB_gen, которые помимо «сырых» данных содержали аугментированные и синтетические снимки, соответственно. Оба набора были разделены на обучающую и валидационную выборки для обучения модели YOLOv11n. Результаты тестового обучения модели на 100 эпохах представлены в табл. 3.

В данном случае при использовании синтетических снимков удалось добиться повышения точности по mAP50 примерно на 9.2%, что подтверждает эффективность разработанного метода.

Заключение

В настоящем исследовании представлен метод генерации искусственного обучающего набора данных ДЗЗ с использованием генеративных механизмов, таких как Stable Diffusion. Полученный таким образом набор может быть использован для решения определенного спектра задач ДЗЗ, в данном случае для поиска и распознавания лесных пожаров, незаконных вырубок деревьев и несанкционированных свалок отходов. Достигнутые значения точности при выполнении экспериментального инференса нейросетевой модели удовлетворяют изначальным требованиям настоящего исследования, тем самым подтверждая целесообразность применения представленного метода по меньшей мере в качестве инструмента аугментации существующих данных.

В процессе проведения настоящего исследования были выявлены некоторые ограничения разработанного метода, связанные с процессом генерации искусственных изображений. Для адекватной генерации необходимо выполнить подбор параметров генератора, а также сформировать запрос, отвечающий определенным требованиям. Полученные изображения также необходимо отфильтровать, так как генератор не всегда выдает нужный результат, и некоторые изображения выглядят чрезмерно искусственными и нереалистичными. К тому же, процесс генерации искусственных изображений занимает значительно больше времени, чем использование классических методов аугментации (геометрические трансформации, цветовые преобразования).

Для будущих исследований рационально использовать новую версию Stable Diffusion с применением более структурированного подхода к формированию запросов (prompts) по средством специального синтаксиса. Таким образом, можно добиться более стабильной генерации целевых изображений. В свою очередь, для повышения точности определения целевых объектов/областей планируется кратно увеличить созданный набор данных с возможным переходом от задачи сегментации к задаче детекции целевых объектов.

Литература

1. Pesonen J., Raita-Hakola AM., Joutsalainen J., Hakala T., Akhtar W., Koivumäki N., Markelin L., Suomalainen J. Alves de Oliveira R., Pölönen I., Honkavaara E. Boreal Forest Fire: UAV-collected wildfire detection and smoke segmentation dataset // Scientific Data. 2025. Vol. 12. P. 1419.
2. Mowla M.N., Asadi D., Tekeoğlu K., Masum S., Rabie K. UAVs-FFDB: A high-resolution dataset for advancing forest fire detection and monitoring using unmanned aerial vehicles (UAVs) // Data in brief. 2024. Vol. 55. P. 110706.
3. Holybro S500 - Copter documentation. URL: https://ardupilot.org/copter/docs/reference-frames-holybro-s500.html (дата обращения: 10.03.2026).
4. Raspberry Pi Camera Module 2. URL: https://www.raspberrypi.com/products/camera-module-v2/ (дата обращения: 10.03.2026).
5. Make Sense. URL: https://www.makesense.ai/ (дата обращения: 10.03.2026).
6. Shamsoshoara A., Afghah F., Razi A., Zheng L., Fulé P., Blasch E. Aerial imagery pile burn detection using deep learning: The FLAME dataset // Computer Networks. 2021. Vol. 193. P. 108001.
7. The FLAME dataset: Aerial Imagery Pile burn detection using drones (UAVs). URL: https://ieee-dataport.org/open-access/flame-dataset-aerial-imagery-pile-burn-detection-using-drones-uavs (дата обращения: 10.03.2026).
8. Ronneberger O., Fischer P., Brox T. U-net: Convolutional networks for biomedical image segmentation // International Conference on Medical image computing and computer-assisted intervention. Cham: Springer international publishing. 2015. P. 234-241.

9. Precision & Recall. URL: https://mlu-explain.github.io/precision-recall/ (дата обращения: 10.03.2026).
10. Mager A., Blass V. From illegal waste dumps to beneficial resources using drone technology and advanced data analysis tools: A feasibility study // Remote Sensing. 2022. Vol. 14. P. 3923.
11. Torres R. N., Fraternali P. AerialWaste dataset for landfill discovery in aerial and satellite images // Scientific Data. 2023. Vol. 10. P. 63.
12. Ortofoto AGEA 2020 RGB. URL: https://geoportale.regione.emilia-romagna.it/catalogo/dati-cartografici/cartografia-di-base/immagini/layer-6 (дата обращения: 10.03.2026).
13. WorldView-3 (США). URL: https://www.scanex.ru/data/satellites/worldview-3/ (дата обращения: 10.03.2026).
14. Google Earth. URL: https://www.google.com/earth/about/versions/ (дата обращения: 10.03.2026).
15. The COCO dataset format. URL: https://docs.aws.amazon.com/rekognition/latest/customlabels-dg/md-coco-overview.html (дата обращения: 10.03.2026).
16. F1 Score in Machine Learning. URL: https://www.geeksforgeeks.org/machine-learning/f1-score-in-machine-learning/ (дата обращения: 10.03.2026).
17. Welcome to the Raven Scan platform documentation. URL: https://github.com/IntelligentNetworkSolutions/IllegalDumpSiteDetectionAndLandfillMonitoring (дата обращения: 10.03.2026).
18. Ramadan M., Ali M., Khoo S.Y., Alkhedher M. AI-powered IoT and UAV systems for real-time detection and prevention of illegal logging // Results in Engineering. 2024. Vol. 24. P. 103277.
19. What is the Internet of Things (IoT). URL: https://www.ibm.com/think/topics/internet-of-things (дата обращения: 10.03.2026).
20. Purwono I., Ma'arif A., Rahmaniar W., Fathurrahman H. I. K., Frisky A., Haq Q. M. Understanding of convolutional neural network (cnn): A review // International Journal of Robotics and Control Systems. 2022. Vol. 2. P. 739-748.
21. Williams J. UAV survey mapping of illegal deforestation in Madagascar // Plants, People, Planet. 2024. Vol. 6. P. 1413-1424.
22. Mean Average Precision (mAP) in Object Detection. URL: https://learnopencv.com/mean-average-precision-map-object-detection-model-evaluation-metric/ (дата обращения: 10.03.2026).
23. Stable Diffusion Public Release. URL: https://stability.ai/news-updates/stable-diffusion-public-release (дата обращения: 10.03.2026).
24. Stable Diffusion web UI. URL: https://github.com/AUTOMATIC1111/stable-diffusion-webui (дата обращения: 10.03.2026).
25. Khanam R., Hussain M. Yolov11: An overview of the key architectural enhancements (2024) // arXiv preprint arXiv:2410.17725. 2024.
26. El-Madafri I., Peña M., Olmedo-Torre N. The wildfire dataset: Enhancing deep learning-based forest fire detection with a diverse evolving open-source dataset focused on data representativeness and a novel multi-task learning approach // Forests. 2023. Vol. 14. P. 1697.
27. Forest Fire and Smoke Detection using UAV imaging_3 Computer Vision Dataset. URL: https://universe.roboflow.com/auebmsc-ba-machine-learning/forest-fire-and-smoke-detection-using-uav-imaging_3 (дата обращения: 10.03.2026).
28. FireForest Computer Vision Dataset. URL: https://universe.roboflow.com/deteccin-de-objetos-kksqo/fireforest-qwrdo (дата обращения: 10.03.2026).
29. fireforest Computer Vision Dataset. URL: https://universe.roboflow.com/yolo-sjgkl/fireforest (дата обращения: 10.03.2026).
30. Deforestation Computer Vision Dataset. URL: https://universe.roboflow.com/ahsan-iqbal-669fetbseef19/deforestation-tgch1 (дата обращения: 10.03.2026).
31. Deforestation Computer Vision Dataset. URL: https://universe.roboflow.com/ahsan-iqbal-ex-3wv/deforestation-e1sye (дата обращения: 10.03.2026).
32. Deforestation Computer Vision Dataset. URL: https://universe.roboflow.com/object-detection-lazxt/deforestation-yzbgm (дата обращения: 10.03.2026).
33. Dataset for Waste/Dumpsite Detection using drone imagery. URL: https://universe.roboflow.com/dimash-ub63w/dumpsite-detection-iqzw1 (дата обращения: 10.03.2026).
34. Labelme - Image annotation with Python. URL: https://github.com/wkentaro/labelme (дата обращения: 10.03.2026).
35. Ultralytics YOLO for SOTA object detection, multi-object tracking, instance segmentation, pose estimation and image classification. URL: https://pypi.org/project/ultralytics/ (дата обращения: 10.03.2026).
36. A set of easy-to-use utils that will come in handy in any Computer Vision project. URL: https://pypi.org/project/supervision/ (дата обращения: 10.03.2026).

Смирнов Александр Владимирович. Федеральное государственное бюджетное учреждение науки Институт программных систем им. А.К. Айламазяна Российской академии наук, г. Переславль-Залесский, Россия. Старший научный сотрудник. Область научных интересов: компьютерное зрение; нейронные сети; робототехника; автоматизация и управление. E-mail: asmirnov_1991@mail.ru (Ответственный за переписку).

Тищенко Игорь Петрович. Федеральное государственное бюджетное учреждение науки Институт программных систем им. А.К. Айламазяна Российской академии наук, г. Переславль-Залесский, Россия. Исполняющий обязанности директора. Кандидат технических наук. Область научных интересов: компьютерное зрение; нейронные сети; робототехника; автоматизация и управление. E-mail: igor.p.tishchenko@yandex.ru.

Generative technologies for creating training data for EO tasks

A.V. Smirnov, I.P. Tishchenko
Program Systems Institute of the RAS

Abstract. This paper presents a method for generating training datasets for Earth observation (EO) tasks using generative tools such as Stable Diffusion. Specifically, the problem of detecting and recognizing forest fires, illegal logging, and unauthorized dumpsites in UAV imagery is considered. In this study, a training dataset consisting of three classes was generated and used to train a YOLO model. When inferring the model to real-world images found online, accuracy values of 72.4% and 65.8% were achieved for the F1 and mAP50 metrics, respectively. Based on the obtained results, it is concluded that the presented method is feasible, at least as a tool for augmenting existing data.

Keywords: *dataset generation, neural network model, image analysis, Earth observation (EO), unmanned aerial vehicle, forest fires, illegal logging, illegal dumping*.
DOI: 10.14357/20790279260305 **EDN:** QGKCZA

References

1. Pesonen J., Raita-Hakola AM., Joutsalainen J., Hakala T., Akhtar W., Koivumäki N., et al. Boreal Forest Fire: UAV-collected wildfire detection and smoke segmentation dataset. *Scientific Data*. 2025;(12): 1419. doi: https://doi.org/10.1038/s41597-025-05634-0.
2. Mowla M.N., Asadi D., Tekeoglu K. N., Masum S., Rabie K. UAVs-FFDB: A high-resolution dataset for advancing forest fire detection and monitoring using unmanned aerial vehicles (UAVs). *Data in Brief*. 2024; (55):110706. doi: https://doi.org/10.1016/j.dib.2024.110706.
3. *Holybro S500 - Copter documentation*. Available from: https://ardupilot.org/copter/docs/reference-frames-holybro-s500.html [Accessed 10 March 2026].
4. *Raspberry Pi Camera Module 2*. Available from: https://www.raspberrypi.com/products/camera-module-v2/ [Accessed 10 March 2026].
5. *Make Sense*. Available from: https://www.makesense.ai/ [Accessed 10 March 2026].
6. Shamsoshoara A., Afghah F., Razi A., Zheng L., Fulé P., Blasch E. Aerial imagery pile burn detection using deep learning: The FLAME dataset. *Computer Networks*. 2021;(193):108001. doi: https://doi.org/10.1016/j.comnet.2021.108001.
7. *The FLAME dataset: Aerial Imagery Pile burn detection using drones (UAVs)*. Available from: https://ieee-dataport.org/open-access/flame-dataset-aerial-imagery-pile-burn-detection-using-drones-uavs [Accessed 10 March 2026].
8. Ronneberger O., Fischer P., Brox T. U-net: Convolutional networks for biomedical image segmentation. *In International Conference on Medical image computing and computer-assisted intervention. Cham: Springer international publishing*. 2015:234-241. doi: https://doi.org/10.48550/arXiv.1505.04597.
9. *Precision & Recall*. Available from: https://mlu-explain.github.io/precision-recall/ [Accessed 10 March 2026].

10. Mager A., Blass V. From illegal waste dumps to beneficial resources using drone technology and advanced data analysis tools: A feasibility study. *Remote Sensing*. 2022;(14):3923. doi: https://doi.org/10.3390/rs14163923.
11. Torres R., Fraternali P. AerialWaste dataset for landfill discovery in aerial and satellite images. *Scientific Data*. 2023;(10):63. doi: https://doi.org/10.1038/s41597-023-01976-9.
12. *Ortofoto AGEA 2020 RGB*. Available from: https://geoportale.regione.emilia-romagna.it/catalogo/dati-cartografici/cartografia-di-base/immagini/layer-6 [Accessed 10 March 2026].
13. *WorldView-3 (USA)*. Available from: https://www.scanex.ru/data/satellites/worldview-3/ [Accessed 10 March 2026].
14. *Google Earth*. Available from: https://www.google.com/earth/about/versions/ [Accessed 10 March 2026].
15. *The COCO dataset format*. Available from: https://docs.aws.amazon.com/rekognition/latest/customlabels-dg/md-coco-overview.html [Accessed 10 March 2026].
16. *F1 Score in Machine Learning*. Available from: https://www.geeksforgeeks.org/machine-learning/f1-score-in-machine-learning/ [Accessed 10 March 2026].
17. *Welcome to the Raven Scan platform documentation*. Available from: https://github.com/IntelligentNetworkSolutions/IllegalDumpSiteDetectionAndLandfillMonitoring [Accessed 10 March 2026].
18. Ramadan M., Ali M., Khoo S.Y., Alkhedher M. AI-powered IoT and UAV systems for real-time detection and prevention of illegal logging. *Results in Engineering*. 2024;(24):103277. doi: https://doi.org/10.1016/j.rineng.2024.103277.
19. *What is the Internet of Things (IoT)*. Available from: https://www.ibm.com/think/topics/internet-of-things [Accessed 10 March 2026].
20. Purwono I., Ma'arif A., Rahmaniar W., Fathurrahman H. I. K., Frisky A., Haq Q. M. Understanding of convolutional neural network (cnn): A review. *International Journal of Robotics and Control Systems*. 2022;(2):739-748. doi: https://doi.org/10.31763/ijrcs.v2i4.888.
21. Williams, J. UAV survey mapping of illegal deforestation in Madagascar. *Plants, People, Planet*. 2024;(6):1413-1424. doi: https://doi.org/10.1002/ppp3.10533.
22. *Mean Average Precision (mAP) in Object Detection*. Available from: https://learnopencv.com/mean-average-precision-map-object-detection-model-evaluation-metric/ [Accessed 10 March 2026].
23. *Stable Diffusion Public Release*. Available from: https://stability.ai/news-updates/stable-diffusion-public-release [Accessed 10 March 2026].
24. *Stable Diffusion web UI*. Available from: https://github.com/AUTOMATIC1111/stable-diffusion-webui [Accessed 10 March 2026].
25. Khanam R., Hussain M. Yolov11: An overview of the key architectural enhancements. *arXiv preprint arXiv:2410.17725*. 2024. doi: https://doi.org/10.48550/arXiv.2410.17725.
26. El-Madafri, I., Peña, M., Olmedo-Torre, N. The wildfire dataset: Enhancing deep learning-based forest fire detection with a diverse evolving open-source dataset focused on data representativeness and a novel multi-task learning approach. *Forests*. 2023;(14):1697. doi: https://doi.org/10.3390/f14091697.
27. *Forest Fire and Smoke Detection using UAV imaging_3 Computer Vision Dataset*. Available from: https://universe.roboflow.com/auebmsc-ba-machine-learning/forest-fire-and-smoke-detection-using-uav-imaging_3 [Accessed 10 March 2026].
28. *FireForest Computer Vision Dataset*. Available from: https://universe.roboflow.com/deteccin-de-objetos-kksqo/fireforest-qwrdo [Accessed 10 March 2026].
29. *Fireforest Computer Vision Dataset*. Available from: https://universe.roboflow.com/yolo-sjgkl/fireforest [Accessed 10 March 2026].
30. *Deforestation Computer Vision Dataset*. Available from: https://universe.roboflow.com/ahsan-iqbal-669fetbseef19/deforestation-tgch1 [Accessed 10 March 2026].
31. *Deforestation Computer Vision Dataset*. Available from: https://universe.roboflow.com/ahsan-iqbal-ex-3wv/deforestation-e1sye [Accessed 10 March 2026].
32. *Deforestation Computer Vision Dataset*. Available from: https://universe.roboflow.com/object-detection-lazxt/deforestation-yzbgm [Accessed 10 March 2026].
33. *Dataset for Waste/Dumpsite Detection using drone imagery*. Available from: https://universe.roboflow.com/dimash-ub63w/dumpsite-detection-iqzw1 [Accessed 10 March 2026].

34. *Labelme – Image annotation with Python*. Available from: https://github.com/wkentaro/labelme [Accessed 10 March 2026].
35. *Ultralytics YOLO for SOTA object detection, multi-object tracking, instance segmentation, pose estimation and image classification*. Available from: https://pypi.org/project/ultralytics/ [Accessed 10 March 2026].
36. *A set of easy-to-use utils that will come in handy in any Computer Vision project*. Available from: https://pypi.org/project/supervision/ [Accessed 10 March 2026].

Smirnov Alexander V. Program Systems Institute of the RAS, 4a Petra Pervogo str., Pereslavl-Zalessky, 152021, Russia. E-mail: asmirnov_1991@mail.ru

Tishchenko Igor P. PhD, Program Systems Institute of the RAS, 4a Petra Pervogo str., Pereslavl-Zalessky, 152021, Russia. E-mail: igor.p.tishchenko@yandex.ru