

УДК 528.9:004.92

<https://doi.org/10.33764/2411-1759-2026-31-1-83-92>

Особенности формирования наборов данных для определения местоположения по фото- и видеоматериалам

Р. В. Гришин¹✉

¹ Сибирский государственный университет геосистем и технологий, г. Новосибирск, Российская Федерация

e-mail: r.grishin54@gmail.com

Аннотация. В статье рассматриваются особенности процесса сбора и разметки данных для обучения моделей компьютерного зрения, определяющих местоположение по фото- и видеоматериалам. Обеспечение высокой точности и разнообразия данных критически влияет на качество обучения моделей, что является основной проблемой при составлении наборов данных. Целью исследования является выявление ключевых факторов, которые влияют на качество материалов, используемых для составления наборов данных, таких как аннотация данных, разнообразие изображений, контекстуальная информация и этические аспекты. В статье анализируются существующие практики и подходы к сбору и обработке данных, а также обсуждаются практические рекомендации по улучшению качества полученных материалов. В результате исследования выявлено, что разнообразие данных, высокое качество изображений и тщательная аннотация значительно повышают точность моделей; на основе этих данных сформированы критерии оценки качества материалов. Создание качественных наборов данных основывается на комплексном подходе от сборки точных и достоверных данных до их валидации и пространственной нормализации.

Ключевые слова: набор данных, определение местоположения, компьютерное зрение, геопространственные данные, классификация

Для цитирования:

Гришин Р. В. Особенности формирования наборов данных для определения местоположения по фото- и видеоматериалам. *Вестник СГУГиТ*. 2026. Т. 31, № 1. С. 83–92. <https://doi.org/10.33764/2411-1759-2026-31-1-83-92>

Features of dataset formation for geolocating images and video materials

R. V. Grishin¹✉

¹ Siberian State University of Geosystems and Technologies, Novosibirsk, Russian Federation

e-mail: r.grishin54@gmail.com

Abstract. The article investigates the key aspects of data collection and annotation for training computer vision models capable of geolocating images and video materials. Dataset accuracy and diversity are pivotal to model efficacy, constituting the central challenge in their compilation. The study delineates critical factors influencing data quality, encompassing annotation precision, image diversity, contextual metadata, and ethical considerations. Prevailing practices in data acquisition and processing are evaluated, accompanied by actionable recommendations to enhance material quality. Di-

verse datasets, high-resolution imagery, and rigorous annotation markedly improve model accuracy, yielding formalized quality assessment criteria. Optimal dataset development demands an integrated methodology, spanning precise data curation, validation, and spatial normalization.

Keywords: datasets, location detection, computer vision, geospatial data, classification

For citation:

Grishin R. V. (2026). Features of dataset formation for geolocating images and video materials. *Vestnik SSUGiT [Vestnik SSUGT]* Vol. 31, No. 1. pp. 83–92. <https://doi.org/10.33764/2411-1759-2026-31-1-83-92>

Введение

Современные технологии геолокации на основе фото- и видеоматериалов становятся ключевым инструментом в таких областях, как навигация, дополненная реальность и урбанистика, однако их эффективность напрямую зависит от качества и репрезентативности обучающих наборов данных [1], которые должны учитывать географическое разнообразие и временные изменения. Несмотря на прогресс в методах компьютерного зрения, создание таких наборов данных остается сложной междисциплинарной задачей, требующей интеграции технологий сбора, аннотирования и валидации данных [2], а также соблюдения норм приватности и лицензирования.

Актуальность исследования обусловлена растущим спросом на точные алгоритмы определения местоположения, которые сталкиваются с проблемой географического и временного уклона. Например, существующие наборы данных, такие как IM2GPS [3] или YFCC100M [4], часто фокусируются на урбанизированных регионах, игнорируя сельские или малодоступные локации, что приводит к снижению точности моделей в реальных условиях. Кроме того, динамические изменения ландшафта (сезонность, погодные условия) требуют постоянного обновления данных, что усложняет их актуальность [5].

Важным аспектом является методология сбора данных. Краудсорсинг обеспечивает масштабируемость, но порождает проблемы с качеством метаданных, такие как неточности GPS или отсутствие семантических аннотаций. Альтернативой выступают синтетические данные, генерируемые с помощью GAN или 3D-моделирования, которые позволяют контролировать вариативность условий съем-

ки, но требуют значительных вычислительных ресурсов [6–8]. Например, методы фотограмметрии, используемые для создания 3D-моделей местности, демонстрируют потенциал в генерации реалистичных сцен, однако их применение ограничено длительностью обработки (до 24 часов для проектов среднего масштаба) [9].

Целью данной статьи является систематизация подходов к созданию наборов данных для задач геолокации, объединив технические и методологические аспекты. На примере анализа существующих решений (ImageNet, Google Landmarks, StreetLearn) и новых методов, таких как гибридные краудсорсингово-синтетические процессы, будут предложены рекомендации по минимизации искажений, оптимизации разметки. Результаты работы направлены на формирование стандартов, которые повысят надежность алгоритмов компьютерного зрения в реальных сценариях.

Методы и материалы

В эпоху цифровых технологий данные стали основой для развития искусственного интеллекта, анализа социальных процессов и принятия решений. Однако сами по себе разрозненные данные, будь то миллионы фотографий из соцсетей или показания датчиков умных устройств, не несут ценности без структурирования и интерпретации. Именно здесь на первый план выходит понятие набора данных – организованной коллекции данных, объединенной общей задачей и дополненной метаданными, которые превращают хаотичную информацию в мощный инструмент для обучения алгоритмов, научных исследований и бизнес-аналитики. Их каче-

ство и репрезентативность напрямую определяют, насколько точными и справедливыми будут решения, принимаемые алгоритмами.

Набор данных может включать текстовую информацию, изображения, видео, аудиозаписи, сенсорные показания или комбинации различных форматов. Каждый элемент набора данных сопровождается метаданными, например географическими координатами, временными отметками или категориальными метками, что позволяет алгоритмам машинного обучения выявлять закономерности и делать прогнозы.

Основная роль наборов данных заключается в обеспечении воспроизводимости и объективности исследований. В машинном обучении они служат источником данных для обучения моделей: нейросети учатся распознавать объекты на изображениях, предсказывать временные ряды или классифицировать тексты именно благодаря разнообразным и репрезентативным данным.

Ключевое требование к наборам данных является репрезентативность, т. е. способность отражать реальные условия, в которых будет применяться модель. Например, набор данных для распознавания дорожных знаков должен включать изображения, сделанные при разном освещении, ракурсах и погодных условиях. Несбалансированные данные (например, преобладание изображений городских пейзажей над сельскими) приводят к смещению (bias) моделей и снижают их точность в реальных сценариях.

Кроме обучения алгоритмов, наборы данных используются:

- для тестирования и сравнения моделей;
- анализа тенденций;
- создания синтетических сред.

Современные наборы данных часто создаются с использованием краудсорсинга (сбор данных от пользователей), автоматического извлечения информации из веб-ресурсов или синтетической генерации (GAN, 3D-моделирование) [10–12]. Однако их подготовка требует решения этических и правовых вопросов: анонимизации персональных данных (GDPR), лицензирования контента и минимизации экологических последствий (энергосза-

траты на обработку). Таким образом, наборы данных – это не просто коллекции данных, а сложные продукты, сочетающие технические, методологические и социальные аспекты.

Формирование наборов данных для задач автоматической геолокации представляет собой многоэтапный процесс, объединяющий методы сбора данных, их аннотирования, валидации и аугментации. Первоначальный этап – сбор данных, который реализуется через три основных подхода. Краудсорсинг, основанный на использовании платформ с пользовательским контентом, предоставляет масштабные массивы изображений с геотегами, однако сопряжен с проблемой неточностей GPS-меток и отсутствия семантической разметки. Альтернативой выступает синтетическая генерация данных, где технологии GAN (Generative Adversarial Networks) и 3D-моделирование (на базе Unity Engine) позволяют имитировать разнообразные условия съемки, включая сезонные изменения и вариации освещения, что, однако, требует значительных вычислительных ресурсов [13–15]. Третий подход предполагает использование открытых репозиторий, таких как Google Landmarks и Open Images, которые предлагают предварительно аннотированные данные, но ограниченные географическим охватом, фокусируясь преимущественно на урбанизированных регионах [16]. Аннотирование данных осуществляется комбинацией ручных и полуавтоматических методов. Ручная разметка с применением инструментов вроде Label Studio обеспечивает высокую точность идентификации объектов и локаций, однако отличается высокой трудоемкостью. Полуавтоматические подходы, такие как кластеризация GPS-меток алгоритмом DBSCAN, позволяют сократить временные затраты, но требуют последующей коррекции для устранения шумов и противоречий в метаданных. Валидация данных включает фильтрацию артефактов через хэширование (pHash) и нейросетевые модели (ResNet-50), а также расширение набора данных, направленное на повышение разнообразия коллекции. Для наглядности этапы создания набора данных представлены на блок-схеме (рис. 1).

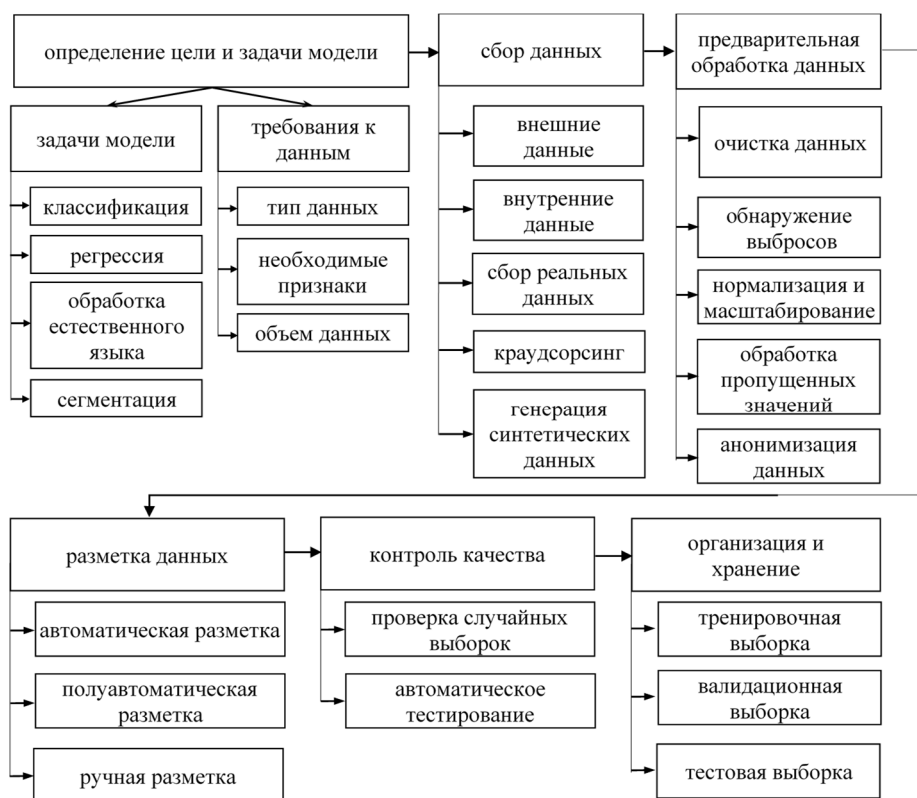


Рис. 1. Этапы создания обучающих выборок

В процессе создания наборов данных, используемых в задачах геолокации по изображениям, возникает ряд уникальных методологических и технических сложностей. В классических наборах данных метки часто представляют собой категории («кошка» или «собака»), но для задач геолокации ключевым элементом являются географические координаты, которые должны точно соответствовать визуальному содержанию снимка [17], что требует сбора изображений с привязкой к GPS-данным, а также тщательной проверки их корреляции [18]. Например, некорректный геотег фотографии Эйфелевой башни, указывающий на Нью-Йорк, сделает набор данных бесполезным для обучения модели. Для решения этой проблемы применяется комбинация подходов: автоматический сбор метаданных (EXIF), ручная верификация экспертами и алгоритмы кластеризации, такие как DBSCAN, для устранения противоречий.

В задачах геопривязки изображений традиционно используются абсолютные географические координаты (широта и долгота) в

качестве меток. Такой подход имеет ряд ограничений: во-первых, он чувствителен к незначительным смещениям координат, во-вторых, приводит к проблемам при кластеризации, агрегации и сравнении данных в разных масштабах. Для решения этих проблем все более широкое распространение получают методы, основанные на дискретизации пространства с помощью регулярных геосеток, таких как H3 (Hexagonal Hierarchical Spatial Index), разработанной Uber.

Система H3 представляет поверхность Земли в виде иерархической структуры шестиугольных ячеек разного разрешения. Каждая ячейка имеет уникальный индекс, позволяющий эффективно выполнять пространственные операции: агрегацию изображений по регионам, определение соседних областей, фильтрацию по плотности и построение графов пространственной связанности. Преимуществом H3 является его масштабируемость: разрешение можно легко адаптировать под нужды конкретной задачи, выбирая более крупные регионы или, наоборот, мелкие локальные области (таблица).

Влияния разрешения НЗ на точность модели

НЗ-разрешение	Средняя площадь ячейки, км ²	Примерное количество ячеек на земле	Применение
res = 5	~2500	~42000	Континенты, страны
res = 7	~300	~588000	Города, регионы
res = 9	~5	37000000	Районы, улицы
res = 10	~0,6	>150000000	Точные координаты

Использование НЗ в наборах данных визуальной геолокации позволяет унифицировать представление пространственных меток и упростить обучение моделей, особенно в задачах классификации. Вместо непрерывных координат модель может предсказывать дискретный индекс НЗ-ячейки, что снижает чувствительность к шуму и позволяет формировать более сбалансированные выборки. Кроме того, пространственное индексирование на основе НЗ облегчает процесс анонимизации, поскольку нет необходимости хранить точные координаты: достаточно идентификатора ячейки, не раскрывающего конкретное местоположение.

На практике использование НЗ может быть комбинировано с другими геопространственными признаками, такими как высота над уровнем моря, плотность застройки или близость к объектам инфраструктуры, что позволяет моделям обучаться на пространственно обобщенных признаках без ущерба для точности. Таким образом, НЗ не только повышает приватность и структурированность набора данных, но и открывает новые возможности для более устойчивого и интерпретируемого обучения в задачах визуальной локализации.

При создании набора данных необходимо учесть вариативность условий съемки. Один и тот же объект может выглядеть по-разному в зависимости от времени суток, сезона, погоды или ракурса. Чтобы обеспечить репрезентативность, набор данных должен включать изображения, сделанные в разнообразных контекстах. Например, зимние снимки альпийских ландшафтов дополняются летними, а городские панорамы – ночными видами с подсветкой. Для расширения разнообразия данных активно используются библиотеки, имитирующие дождь, туман или изменение освещения, что повышает устойчивость моделей к реальным условиям. Однако даже такие методы не заменяют необходимо-

сти сбора «сырых» данных из разных источников – от краудсорсинга до синтетической генерации через 3D-модели.

Критически важным аспектом является минимизация географического и временного смещения. Многие существующие наборы данных (IM2GPS, YFCC100M) неравномерно представлены в сторону развитых регионов (США, Европа), что снижает точность моделей в Азии, Африке или сельской местности [19]. Для устранения этого дисбаланса необходимо целенаправленно включать данные из малоизвестных локаций, например, используя материалы местных фотографов или спутниковые снимки. Кроме того, динамические изменения ландшафта (новые здания, вырубка лесов) требуют регулярного обновления наборов данных, что усложняет их поддержку.

Анонимизация визуальных данных становится ключевым этапом при подготовке наборов данных, включающих изображения с признаками географической привязки. С учетом требований законодательства о защите персональных данных, таких как федеральный закон № 152-ФЗ «О персональных данных» и положение об обработке биометрических данных (особо защищенных категорий), необходимо исключать или модифицировать элементы, способные идентифицировать конкретных людей или объекты. На практике это означает автоматическое размывание лиц, государственных регистрационных знаков автомобилей, уличной рекламы и иных уникальных признаков [20]. Современные методы анонимизации включают в себя применение нейросетевых детекторов (например MTCNN, YOLOv5, Detectron2) для выявления таких объектов с последующим редактированием изображения (например, заменой пикселей или генеративным замещением).

Однако полная анонимизация не всегда гарантирует невозможность восстановления ис-

ходной информации, особенно в случае с привязкой к локации, где контекст сам по себе может быть идентифицирующим (например, уникальные архитектурные объекты). Это порождает необходимость дополнительного анализа рисков повторной идентификации при публикации открытых наборов данных.

Проблематика лицензирования также остается актуальной: многие существующие наборы данных, применимые к задачам визуальной геолокации (например StreetLearn, GeoPlaces, Mapillary Vistas), публикуются под ограничительными лицензиями (Creative Commons BY-NC, исследовательские соглашения), запрещающими их использование в коммерческих целях или распространение модифицированных копий. Это ограничивает применение таких данных в промышленных продуктах и требует либо самостоятельного сбора наборов данных, либо генерации синтетических данных. Последние, хотя и не подпадают под юридические ограничения, все еще уступают реальным по степени визуальной достоверности и разнообразию сцен, что может негативно сказаться на обобщающей способности обучаемых моделей.

Создание специализированных наборов данных для задач компьютерного зрения требует четко организованного процесса, охватывающего все стадии от сбора данных до валидации и экспорта. В рамках данной практики реализован комплексный подход, включающий использование специализированных инструментов на каждом этапе обработки данных [21].

На первом этапе осуществляется сбор данных, для чего используются два ключевых источника: реальные и синтетические данные. Для автоматизированного извлечения информации из открытых веб-источников применяется фреймворк Scrapy, обеспечивающий масштабируемый сбор изображений и метаданных [22, 23]. Дополнительно используется симулятор CARLA, предназначенный для генерации синтетических сцен с заранее заданными

параметрами и автоматической разметкой объектов. Такое комбинирование позволяет расширить охват возможных сценариев и повысить вариативность обучающих данных.

Вторым этапом является аннотирование, реализуемое с помощью платформы CVAT (Computer Vision Annotation Tool). CVAT предоставляет удобный web-интерфейс для ручной и полуавтоматической разметки изображений, включая поддержку форматов bounding box, instance segmentation и keypoints. Для повышения эффективности процесса аннотирования используется стратегия активного обучения, позволяющая автоматически отбирать наиболее информативные и неопределенные примеры для приоритетной разметки [24, 25]. Модель, обученная на предыдущих итерациях, участвует в формировании таких выборок, тем самым формируя замкнутый цикл взаимодействия между аннотированием и валидацией.

На третьем этапе выполняется валидация данных, где применяются модули библиотеки PyTorch, в частности компоненты Torchvision, позволяющие оценивать качество аннотации и выявлять потенциальные артефакты или аномалии. Предобученные модели могут использоваться для сверки полученных меток с предсказаниями, что способствует автоматической диагностике некорректных или неполных разметок. Результаты анализа могут быть визуализированы и направлены обратно в CVAT, обеспечивая итеративное улучшение качества разметки.

Заключительный этап включает экспорт и хранение данных. После валидации аннотированный набор данных сохраняется в стандартизированных форматах с последующей интеграцией в локальные или облачные хранилища. Такая структура процесса (рис. 2) обеспечивает воспроизводимость, масштабируемость и устойчивость всего процесса формирования обучающих выборок.

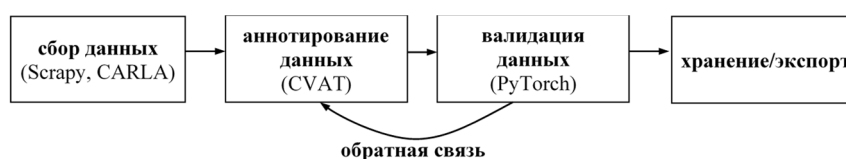


Рис. 2. Схема взаимодействия компонентов при построении набора данных

Результаты

Отсутствие формализованных критериев оценки качества данных может приводить к субъективному отбору изображений, что негативно сказывается на обобщающей способности моделей и воспроизводимости исследований. Разработка и внедрение метрик направлены на стандартизацию подходов и повышение прозрачности формирования обучающих выборок.

Для оценки фото- и видеоматериалов в задачах определения местоположения предлагается использовать следующие ключевые метрики:

- пространственная точность – это расстояние между заявленными координатами (например, из EXIF-метаданных) и фактическим местоположением объекта на фото, которое может быть установлено с помощью ручной верификации или алгоритмов сопоставления;

- оптическое качество изображения, которое определяется резкостью (через анализ градиентов, например, методом вариации Лапласиана), уровнем освещенности и наличием шумов;

- контекстная информативность изображения, отражающая наличие значимых геопривязанных объектов (архитектурных, природных, инфраструктурных), оцененная с помощью детекторов объектов или по плотности признаков на изображении;

- разнообразие условий съемки, включающее временные (день/ночь, сезон) и погодные факторы, фиксируемые либо по метаданным, либо по автоматическому анализу сцены.

Для оценки составленного набора данных предлагается использовать метрики, характеризующие его структуру и полноту. Одной из ключевых является метрика пространственного охвата, основанная на количестве уникальных ячеек НЗ определенного разрешения, которые представлены в наборе данных. Чем выше покрытие, тем шире географическое разнообразие данных. Дополнительно учитывается равномерность распределения изображений по этим ячейкам, которая оценивается через энтропию распределения. Важным показателем является информационная полнота – доля изображений с валидными, непротиворечивыми и проверенными аннотациями, включая координаты, катего-

рии сцены и временные метки. Также вводится индекс визуального разнообразия, отражающий степень различия сцен внутри набора данных. Он вычисляется на основе кластеризации изображений по признакам стиля, цветовой гаммы и композиции, извлекаемым через сверточные нейросети.

Интеграция этих метрик в процесс формирования и контроля качества набора данных позволяет выявить ряд характерных проблем, таких как пространственный и тематический перекос, избыточность однотипных сцен и недоаннотированные изображения. Коррекция набора данных с учетом этих показателей приведет к повышению точности и устойчивости обученных моделей визуальной геолокации, особенно при работе с новыми и географически удаленными локациями.

Заключение

Качество наборов данных для задач визуальной геолокации напрямую зависит от глубины проработки всех этапов их формирования – от сбора и аннотации до валидации и пространственной нормализации. В современных условиях недостаточно просто агрегировать изображения с геометками; необходимо обеспечить репрезентативность, разнообразие условий съемки, точность координатной привязки и юридическую чистоту данных. Использование комбинации реальных и синтетических данных обеспечивает высокую вариативность визуальных сцен и повышает устойчивость обученных моделей к изменяющимся условиям.

Применение пространственных индексов, в частности сетки НЗ, демонстрирует существенные преимущества в сравнении с традиционными методами обработки координатных данных, упрощая процедуры агрегации, анонимизации и анализа геопространственной информации. Для повышения качества итогового набора данных целесообразно применять активное обучение, которое в процессе аннотирования и итеративной валидации повышает точность и согласованность разметки.

В числе приоритетных направлений для дальнейшего развития – генерация синтетических изображений, совершенствование

методов анонимизации и стандартизация обработки метаданных. Эти усилия позволят создать более надежные, масштабируемые и юридически устойчивые наборы дан-

ных, что в свою очередь расширит возможности применения систем визуальной локализации в навигации, безопасности, урбанистике и других прикладных задачах.

БИБЛИОГРАФИЧЕСКИЙ СПИСОК

1. Workman S., Souvenir R., Jacobs N. Wide-Area Image Geo-Localization with Aerial Reference Imagery. Proceedings of the IEEE International Conference on Computer Vision (ICCV). 2015. P. 5007–5015. DOI 10.1109/ICCV.2015.451.
2. Hays J., Efros A. F. im2GPS: estimating geographic information from a single image. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR). 2008. P. 1–8.
3. Lin T.-Y., Cui Y., Belongie S. Hays J. Learning deep representations for ground-to-aerial geolocalization. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR). 2015. P. 5007–5015. DOI 10.1109/CVPR.2015.7299155.
4. Thomee B., Shamma D. A., Friedland G., Elizalde B., Ni K., Poland D., Borth D., Li-Jia L., YFCC100M: The New Data in Multimedia Research. Communications of the ACM. 2016. Vol. 59, No. 2. P. 64–73. DOI 10.1145/2812802.
5. Zheng Z., Wei Y., Yang Y. University-1652: A Multi-view Multi-source Benchmark for Drone-based Geo-localization. Proceedings of the 28th ACM International Conference on Multimedia. 2020. P. 1395–1403. DOI 10.1145/3394171.3413896.
6. Astruc G., Dufour N., Siglidis I. et al. OpenStreetView-5M: The Many Roads to Global Visual Geolocation. Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). 2024. P. 14521–14531. DOI 10.1109/CVPR52729.2024.01411.
7. Zamir A. R., Shah M. Image geo-localization based on multiple nearest neighbor feature matching using generalized graphs. IEEE Transactions on Pattern Analysis and Machine Intelligence. 2014. Vol. 36, № 8. P. 1546–1558. DOI 10.1109/TPAMI.2014.2300476.
8. Neuhold G., Ollmann T. Rota Bulò S., Kotschieder. P. The Mapillary Vistas Dataset for Semantic Understanding of Street Scenes. Proceedings of the IEEE International Conference on Computer Vision (ICCV). 2017. P. 4990–4999. DOI 10.1109/ICCV.2017.534.
9. Clark B., Kerrigan A., Kulkarni P. P., Cepeda V. V., Shah M. Where we are and what we're looking at: query-based worldwide image geo-localization using hierarchies and scenes. Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). 2023. P. 12578–12588. DOI 10.1109/CVPR52729.2023.01236.
10. Ye J., Lv Z., Li W., Yu J., Yang H., Zhong H., He C. Cross-View Image Geo-localization with Panorama-BEV Co-Retrieval Network. Proceedings of the European Conference on Computer Vision (ECCV). 2024. LNCS 14670. P. 532–548. DOI 10.1007/978-3-031-72657-0_32.
11. Yao Y., Sun C. Wang T., Yang J., Zheng E. UAV Geo-Localization Dataset and Method Based on Cross-View Matching. Sensors. 2024. Vol. 24, № 21. P. 6905. DOI 10.3390/s24216905.
12. Zhang X., Sultani W., Wshah S. Cross-View Image Sequence Geo-localization. Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision (WACV). 2023. P. 1980–1989. DOI 10.1109/WACV56688.2023.00206.
13. Khatib O., Lee S., Lee B. S., Yang J. A photo-realistic synthetic dataset for analyzing the effects of moving objects on visual localization algorithms for drones. Proceedings of the International Symposium on Robotics Research (ISRR). 2022. Springer Proceedings in Advanced Robotics. Vol. 20. P. 311–326. DOI 10.1007/978-3-030-95459-8_20.
14. Li B., Xue Y., Shi M. Synthetic data generation based on local-foreground generative adversarial networks for surface defect detection. Journal of Electronic Imaging. 2020. Vol. 29, № 1. P. 013016. DOI 10.1117/1.JEI.29.1.013016.
15. Mununi A., Mununi F., Gerrar N. K. A survey of synthetic data augmentation methods in computer vision. Machine Intelligence Research. 2024. P. 1–39. DOI 10.1007/s11633-022-1411-7.
16. Shorten C., Khoshgoftaar T. M. A survey on image data augmentation for deep learning. Journal of Big Data. 2019. Vol. 6, № 60. P. 1–48. DOI 10.1186/s40537-019-0197-0.
17. Пикалев Я. С. Обнаружение ключевых объектов и перекрестная геолокация: анализ наборов данных и методологические перспективы. Вестник компьютерных и информационных технологий. 2024. № 8. С. 53–60.

18. Павленко Б. В., Пикалев Я. С. Методика создания набора аэрофотоснимков для задачи перекрестной геолокации. Сборник научных трудов. 2024. С. 115–122. DOI 10.24412/2413-7383-2024-4-101-112.
19. Zhang X., Wang L., Su Y. Visual place recognition: A survey from a deep learning perspective. *Pattern Recognition*. 2021. Vol. 113. P. 107760. DOI 10.1016/j.patcog.2020.107760.
20. Klomp S. R., van Rijn M., Wijnhoven R. G. J., Snoek C. G. M., de With P. H. N. Safe Fakes: Evaluating face anonymizers for face detectors. arXiv preprint. 2021. arXiv:2104.11721. DOI 10.48550/arXiv.2104.11721.
21. Прасич А. В., Прасич В. А., Прасич И. В. Формирование обучающей выборки в задачах компьютерного зрения: практики и рекомендации. *Информатика и управление системами*. 2021. Т. 24, № 2. С. 34–43. DOI 10.31799/1684-8853-2021-4-61-70.
22. Гельтман А. И., Горюнов М. Н., Мацкевич А. Г., Рыболовлев Д. А. Методика сбора обучающего набора данных для задач машинного обучения. *Proceedings of the Institute for System Programming of the RAS (ISPRAS)*. 2021. Т. 33, № 5. С. 83–92. DOI 10.15514/ISPRAS-2021-33(5)-5.
23. Пантюшин В. А. Создание элементов инфраструктуры пространственных данных по локальным участкам цифровых изображений. *Землеустройство, кадастр и мониторинг земель*. 2023. № 7. С. 45–51. DOI 10.33920/sel-04-2307-06.
24. Хроль Е. В., Шаронова К. С. Распознавание изображений с помощью искусственного интеллекта. *Вестник науки и образования*. 2023. № 10. С. 99–104. DOI 10.47813/2782-2818-2023-3-4-0311-0321.
25. Хлебникова Т. А., Арбузов А. С., Лисицкий Д. В., Оприцова О. А. Использование материалов БВС для выявления фактов нарушения земельного законодательства на территории г. Новосибирска. *Вестник СГУГиТ*. 2023. Т. 28, № 5. С. 33–40.

REFERENCES

1. Workman, S., Souvenir, R., & Jacobs, N. (2015) Wide-Area Image Geo-Localization with Aerial Reference Imagery. *In Proceedings of the IEEE International Conference on Computer Vision (ICCV)*, 5007-5015, DOI 10.1109/ICCV.2015.451.
2. Hays, J., & Efros, A. F. (2008) im2GPS: estimating geographic information from a single image. *In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 1-8.
3. Lin, T.-Y., Cui, Y., Belongie, S., & Hays, J. (2015) Learning deep representations for ground-to-aerial geolocation. *In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 5007-5015, DOI: 10.1109/CVPR.2015.7299155.
4. Thomee, B., Shamma, D. A., Friedland, G., Elizalde, B., Ni, K., Poland, D., Borth, D., & Li-Jia, L. (2016) YFCC100M: The New Data in Multimedia Research. *Communications of the ACM*, 59(2), 64–73, DOI 10.1145/2812802.
5. Zheng, Z., Wei, Y., & Yang, Y. (2020) University-1652: A Multi-view Multi-source Benchmark for Drone-based Geo-localization. *In Proceedings of the 28th ACM International Conference on Multimedia*, 1395–1403, DOI 10.1145/3394171.3413896.
6. Astruc, G., Dufour, N., Siglidis, I., et al. (2024) OpenStreetView-5M: The Many Roads to Global Visual Geolocation. *In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 14521–14531, DOI 10.1109/CVPR52729.2024.01411.
7. Zamir, A. R., & Shah, M. (2014) Image geo-localization based on multiple nearest neighbor feature matching using generalized graphs. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 36(8), 1546–1558, DOI 10.1109/TPAMI.2014.2300476.
8. Neuhold, G., Ollmann, T., Rota Bulò, S., & Kotschieder, P. (2017) The Mapillary Vistas Dataset for Semantic Understanding of Street Scenes. *In Proceedings of the IEEE International Conference on Computer Vision (ICCV)*, 4990–4999, DOI 10.1109/ICCV.2017.534.
9. Clark, B., Kerrigan, A., Kulkarni, P. P., Cepeda, V. V., & Shah, M. (2023) Where we are and what we're looking at: query-based worldwide image geo-localization using hierarchies and scenes. *In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 12578–12588, DOI 10.1109/CVPR52729.2023.01236.
10. Ye, J., Lv, Z., Li, W., Yu, J., Yang, H., Zhong, H., & He, C. (2024) Cross-View Image Geo-localization with Panorama-BEV Co-Retrieval Network. *In Proceedings of the European Conference on Computer Vision (ECCV)*, LNCS 14670, 532–548, DOI 10.1007/978-3-031-72657-0_32.

11. Yao, Y., Sun, C., Wang, T., Yang, J., & Zheng, E. (2024) UAV Geo-Localization Dataset and Method Based on Cross-View Matching. *Sensors*, 24(21), 6905, DOI 10.3390/s24216905.
12. Zhang, X., Sultani, W., & Wshah, S. (2023) Cross-View Image Sequence Geo-localization. In *Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision (WACV)*, 1980–1989, DOI 10.1109/WACV56688.2023.00206.
13. Khatib, O., Lee, S., Lee, B. S., & Yang, J. (2022) A photo-realistic synthetic dataset for analyzing the effects of moving objects on visual localization algorithms for drones. In *Proceedings of the International Symposium on Robotics Research (ISRR)*, *Springer Proceedings in Advanced Robotics*, 20, 311–326, DOI 10.1007/978-3-030-95459-8_20.
14. Li, B., Xue, Y., & Shi, M. (2020) Synthetic data generation based on local-foreground generative adversarial networks for surface defect detection. *Journal of Electronic Imaging*, 29(1), 013016, DOI 10.1117/1.JEI.29.1.013016.
15. Mununi, A., Mununi, F., & Gerrar, N. K. (2024) A survey of synthetic data augmentation methods in computer vision. *Machine Intelligence Research*, 1–39, DOI 10.1007/s11633-022-1411-7.
16. Shorten, C., & Khoshgoftaar, T. M. (2019) A survey on image data augmentation for deep learning. *Journal of Big Data*, 6(60), 1–48, DOI 10.1186/s40537-019-0197-0.
17. Pikalev, Ya. S. (2024) Object detection and cross-view geolocation: dataset analysis and methodological perspectives. *Vestnik komp'yuternyykh i informatsionnykh tekhnologiy [Bulletin of Computer and Information Technologies]*, 8, 53–60 [in Russian].
18. Pavlenko, B. V., & Pikalev, Ya. S. (2024) Methodology for creating an aerial image dataset for cross-view geolocation tasks. *Collection of Scientific Works*, 115–122, DOI 10.24412/2413-7383-2024-4-101-112 [in Russian].
19. Zhang, X., Wang, L., & Su, Y. (2021) Visual place recognition: A survey from a deep learning perspective. *Pattern Recognition*, 113, 107760, DOI 10.1016/j.patcog.2020.107760.
20. Klomp, S. R., van Rijn, M., Wijnhoven, R. G. J., Snoek, C. G. M., & de With, P. H. N. (2021) Safe Fakes: Evaluating face anonymizers for face detectors. arXiv preprint, arXiv:2104.11721, DOI 10.48550/arXiv.2104.11721.
21. Prasich, A. V., Prasich, V. A., & Prasich, I. V. (2021) Formation of training datasets in computer vision tasks: practices and recommendations. *Informatika i upravlenie sistemami [Informatics and Systems Control]*, 24(2), 34–43, DOI 10.31799/1684-8853-2021-4-61-70 [in Russian].
22. Geltman, A. I., Goryunov, M. N., Matskevich, A. G., & Rybolovlev, D. A. (2021) Methodology for collecting training datasets for machine learning tasks. *Proceedings of the Institute for System Programming of the RAS (ISPRAS)*, 33(5), 83–92, DOI 10.15514/ISPRAS-2021-33(5)-5 [in Russian].
23. Pantyushin, V. A. (2023) Creating spatial data infrastructure elements from local areas of digital images. *Zemleustroystvo, Kadastr i Monitoring Zemel [Land Management, Cadastre and Land Monitoring]*, 7, 45–51, DOI 10.33920/sel-04-2307-06 [in Russian].
24. Khrol, E. V., & Sharonova, K. S. (2023) Image recognition using artificial intelligence. *Vestnik nauki i obrazovaniya [Bulletin of Science and Education]*, 10, 99–104, DOI 10.47813/2782-2818-2023-3-4-0311-0321 [in Russian].
25. Khlebnikova, T. A., Arbuzov, A. S., Lisitsky, D. V., & Opritova, O. A. (2023). Using UAV materials to identify violations of land legislation in the territory of Novosibirsk. *Vestnik SGUGiT, [Vestnik SSUGT]* 28(5), 33–40.

Об авторах

Руслан Владимирович Гришин – аспирант кафедры картографии и геоинформатики.

Author details

Ruslan V. Grishin – PhD Student, Department of Cartography and Geoinformatics.

Получено / Received 02.10.2025

Поступила после рецензирования / Revised 16.01.2026

Принята к публикации / Accepted 30.01.2026