

doi: 10.17586/2226-1494-2026-26-1-104-115

УДК 004.825

## Разработка алгоритма формирования рекомендаций выбора моделей детектирования объектов на основе мета-признаков данных и базы знаний экспериментов

Артём Алексеевич Сметанин<sup>1</sup>✉, Алексей Валентинович Духанов<sup>2</sup>,  
Михаил Юрьевич Герасимчук<sup>3</sup>, Евгений Валерьевич Гудов<sup>4</sup>

<sup>1,4</sup> ООО «Планетарий 1», Санкт-Петербург, 196084, Российская Федерация

<sup>1,2,3</sup> Университет ИТМО, Санкт-Петербург, 197101, Российская Федерация

<sup>3</sup> ООО МТС ДИДЖИТАЛ, Санкт-Петербург, 197046, Российская Федерация

<sup>4</sup> Академия труда и социальных отношений, Москва, 119454, Российская Федерация

<sup>1</sup> [artem\\_smetanin@niuitmo.ru](mailto:artem_smetanin@niuitmo.ru)✉, <https://orcid.org/0000-0002-8158-7450>

<sup>2</sup> [dukhanov@itmo.ru](mailto:dukhanov@itmo.ru), <https://orcid.org/0000-0002-1011-9932>

<sup>3</sup> [mishagerasimchuk@itmo.ru](mailto:mishagerasimchuk@itmo.ru), <https://orcid.org/0009-0008-8005-9550>

<sup>4</sup> [evgenes@gmail.com](mailto:evgenes@gmail.com), <https://orcid.org/0009-0006-4534-0262>

### Аннотация

**Введение.** Детектирование объектов в системах видеонаблюдения, медицинской диагностики, в автономных транспортных средствах и иных технических системах с компонентами компьютерного зрения связано с выбором лучшей модели для конкретного набора данных. Эта задача до сих пор остается трудоемкой и требует значительных вычислительных и временных ресурсов. Предлагается метод автоматизированного формирования рекомендаций по выбору модели детектирования объектов на основе анализа мета-признаков изображений и базы знаний результатов выполненных ранее научных исследований. **Метод.** Предложенный метод предполагает формирование векторов мета-признаков на основе набора данных (изображения и аннотации) и поиске похожих наборов данных с использованием алгоритма CatBoostClassifier (F1-score = 0,82) в собранной базе знаний. После определения аналогичных наборов данных осуществляется ранжирование моделей детектирования объектов по точности (mean Average Precision) и производительности (Frames Per Second) на основе накопленных экспериментальных данных, что позволяет рекомендовать наиболее подходящие архитектуры. Метод реализован в открытом рекомендательном фреймворке Object Detection Recommendation System. **Основные результаты.** Экспериментально подтверждено, что применение предложенного алгоритма сокращает время выбора модели на 40 % по сравнению с традиционными подходами, требующими ручного тестирования. Результаты проверки на 12 тестовых наборах данных продемонстрировали высокую эффективность метода в задачах с ограниченными вычислительными ресурсами и разнообразными характеристиками изображений. **Обсуждение.** Благодаря интеграции в Object Detection Recommendation System метод доступен специалистам, не обладающим глубокими знаниями в области машинного обучения/компьютерного зрения. Метод особенно полезен в условиях ограниченных временных и вычислительных ресурсов, когда необходимо быстро получить обоснованное начальное приближение к выбору модели. В будущем планируется расширение базы знаний, а также адаптация метода для задач семантической сегментации и трекинга объектов.

### Ключевые слова

детектирование объектов, мета-признаки, автоматизация выбора моделей, компьютерное зрение, база знаний экспериментов

**Ссылка для цитирования:** Сметанин А.А., Духанов А.В., Герасимчук М.Ю., Гудов Е.В. Разработка алгоритма формирования рекомендаций выбора моделей детектирования объектов на основе мета-признаков данных и базы знаний экспериментов // Научно-технический вестник информационных технологий, механики и оптики. 2026. Т. 26, № 1. С. 104–115. doi: 10.17586/2226-1494-2026-26-1-104-115

## Development of an algorithm for formulating recommendations for the selection of object detection models based on data meta-features and an experimental knowledge base

Artem A. Smetanin<sup>1</sup>✉, Alexey V. Dukhanov<sup>2</sup>, Mikhail Yu. Gerasimchuk<sup>3</sup>, Evgeny V. Gudov<sup>4</sup>

<sup>1,4</sup> Planetarium one, Saint Petersburg, 196084, Russian Federation

<sup>1,2,3</sup> ITMO University, Saint Petersburg, 197101, Russian Federation

<sup>3</sup> MTS DIGITAL LLC, Saint Petersburg, 197046, Russian Federation

<sup>4</sup> Academy of Labour and Social Relations, Moscow, 119454, Russian Federation

<sup>1</sup> artem\_smetanin@niuitmo.ru✉, <https://orcid.org/0000-0002-8158-7450>

<sup>2</sup> dukhanov@itmo.ru, <https://orcid.org/0000-0002-1011-9932>

<sup>3</sup> mishagerasimchuk@itmo.ru, <https://orcid.org/0009-0008-8005-9550>

<sup>4</sup> evgenes@gmail.com, <https://orcid.org/0009-0006-4534-0262>

### Abstract

Object detection in video surveillance systems, medical diagnostics, autonomous vehicles, and other technical systems incorporating computer vision components is closely tied to selecting the optimal model for a specific dataset. This task remains labor-intensive and requires substantial computational and time resources. In this work, the authors propose a method for automated generation of recommendations for selecting an object detection model based on the analysis of image meta-features and a knowledge base of results from previous studies. The proposed method involves forming meta-feature vectors from a dataset (images and annotations) and searching for similar datasets within a collected knowledge base using the CatBoostClassifier algorithm (F1-score = 0.82). After identifying analogous datasets, object detection models are ranked by accuracy (mAP50) and performance (FPS) based on previously accumulated experimental data, enabling the recommendation of the most suitable architectures. The method is implemented in the open-source recommendation framework Object Detection Recommendation System (ODRS). Experimental evaluation confirmed that applying the proposed algorithm reduces model selection time by 40 % compared to traditional approaches that require manual testing. Validation on 12 test datasets demonstrated high efficiency of the method in tasks with limited computational resources and diverse image characteristics. Due to its integration into ODRS, the method is accessible to specialists without deep expertise in machine learning or computer vision. It is particularly useful in scenarios with constrained time and computational resources, where a quick, well-founded initial approximation for model selection is needed. Future work includes expanding the knowledge base and adapting the method for semantic segmentation and object tracking tasks.

### Keywords

meta-algorithms, object detection, recommendation systems, meta-features, model selection automation, computer vision, algorithm optimization, experimental knowledge base

**For citation:** Smetanin A.A., Dukhanov A.V., Gerasimchuk M.Yu., Gudov E.V. Development of an algorithm for formulating recommendations for the selection of object detection models based on data meta-features and an experimental knowledge base. *Scientific and Technical Journal of Information Technologies, Mechanics and Optics*, 2026, vol. 26, no. 1, pp. 104–115 (in Russian). doi: 10.17586/2226-1494-2026-26-1-104-115

### Введение

В последние годы наблюдается значительный рост интереса к разработке и применению моделей детектирования объектов в таких областях, как автономные транспортные средства, системы видеонаблюдения [1], сортировка мусора [2] и многих других. Одной из ключевых задач в исследованиях остается выбор модели для конкретного набора данных, требующий значительных временных и вычислительных ресурсов. Современные подходы, основанные на эмпирическом подборе, часто оказываются неэффективными с точки зрения временных и человеческих ресурсов: ручной подбор моделей требует длительной работы экспертов и ограничивает число анализируемых исследований, что тормозит прогресс в области компьютерного зрения и, в частности, детектирования объектов [3, 4].

Анализ существующих решений показывает, что методы, опирающиеся на метапризнаки данных (числовые характеристики набора данных), такие как число примеров, количество классов и базовые статистические распределения — обладают потенциалом для ускорения процесса выбора модели [5, 6]. Однако большинство

таких подходов ограничены узкими предметными областями или недостаточной адаптивностью к разнородным наборам данных. При этом чаще всего используются четыре-пять мета-признаков (например, число экземпляров, число классов, среднее и дисперсия по признакам), что ограничивает полноту характеристики данных и снижает качество рекомендаций [7]. Кроме того, существующие подходы, как правило, оперируют отдельными группами признаков и не используют объединенные векторы мета-признаков, которые одновременно описывают визуальные характеристики изображений, структуру аннотаций и предпочтения пользователя. Вместе с тем не удалось найти решения задачи создания адаптивного и автоматизированного метода выбора моделей детектирования объектов, который может учитывать совокупность признаков различных уровней и позволяет получать обоснованные рекомендации с учетом специфики данных и пользовательских требований.

В настоящей работе предложен метод автоматизированного подбора моделей машинного обучения для детектирования объектов, основанный на анализе метапризнаков изображений, аннотаций и предпочте-

ний пользователя, с применением классификационных алгоритмов (CatBoost, Random Forest) и методов кластеризации ( $k$ -means, Agglomerative Clustering) на базе знаний результатов выполненных ранее экспериментов. В отличие от существующих решений, данный подход использует объединенный вектор метапризнаков (включающий статистики распределения, характеристики аннотаций и пользовательские требования), что позволяет повысить точность рекомендаций (Precision) с 85 % до 92 %. Такая точность считается высокой для задач автоматического подбора моделей и демонстрирует надежность рекомендаций [8]. В рамках исследования разработана и сформирована база знаний на основе результатов обучения 20 моделей глубокого обучения, включая You Only Look Once (YOLO), Single Shot MultiBox Detector (SSD) и Faster Region-based Convolutional Neural Network (Faster R-CNN), на 12 разнообразных наборах данных; проведено сравнение алгоритмов машинного обучения и предложен алгоритм ранжирования моделей с учетом требований пользователя. Экспериментальные результаты показывают, что предложенный метод позволяет сократить время выбора модели на 40 % по сравнению с традиционными подходами.

#### **Обзор методов машинного обучения для автоматизированного выбора моделей детектирования объектов**

Современные методы выбора моделей машинного обучения основаны на анализе характеристик данных и применении алгоритмов, позволяющих выявлять взаимосвязи между мета-признаками наборов и качеством работы моделей. Наиболее часто используются два класса подходов — кластеризация и классификация, обеспечивающие автоматизированный поиск оптимальных решений [9, 10].

Алгоритмы кластеризации, такие как  $k$ -means [11] и Agglomerative Clustering [12], применяются для группировки схожих наборов данных без необходимости предварительной разметки. Они просты в реализации и позволяют выделять скрытые структуры, однако чувствительны к выбросам и требуют предварительного выбора числа кластеров. Методы классификации, напротив, ориентированы на размеченные данные и обеспечивают более высокую точность рекомендаций. Среди них наиболее эффективны Random Forest, CatBoost [13, 14] и AdaBoost [15], демонстрирующие высокие значения метрик Precision, Recall и F1-score, отражающих способность модели корректно определять и различать классы. Нейросетевые подходы, такие как Multi-Layer Perceptron (MLP) [16], хорошо обрабатывают сложные нелинейные зависимости, но подвержены переобучению и требуют тщательной настройки гиперпараметров.

Для задач детектирования объектов наибольшее распространение получили архитектуры глубокого обучения YOLO, SSD и Faster R-CNN [17, 18]. Их производительность и точность оцениваются с помощью метрик mean Average Precision (mAP50), Precision, Recall и Frames Per Second (FPS). Показатель mAP50

при 50-процентном перекрытии, Intersection over Union) характеризует среднюю точность локализации и классификации объектов. Метрики Precision и Recall описывают соответственно корректность и полноту детектирования, а FPS отражает скорость обработки изображений на центральном и графическом процессорах.

Сравнительные результаты на эталонных наборах данных показывают, что YOLOv8 достигает mAP50 около 0,62 при 155 FPS, SSD демонстрирует mAP50 — 0,77 при 46 FPS, а Faster R-CNN обеспечивает наибольшую точность (mAP50 — 0,87), но работает медленнее, порядка 7 FPS. Таким образом, выбор архитектуры определяется балансом между точностью и производительностью, что делает особенно актуальным использование мета-признаков данных для предварительного анализа.

Мета-признаки представляют собой совокупность статистических, визуальных и структурных характеристик изображений и аннотаций, таких как энтропия, яркость, контраст, плотность объектов, дисбаланс классов, размеры и пропорции ограничивающих рамок [19]. Они позволяют количественно описать сложность набора данных и подобрать оптимальную модель без полного перебора всех архитектур.

#### **Метод автоматизированного выбора архитектуры на основе мета-признаков**

Комплексный анализ мета-признаков лежит в основе формирования базы знаний и разработки системы рекомендаций по выбору моделей детектирования объектов. База знаний представляет собой накопленный опыт применения архитектур глубокого обучения к разнообразным наборам данных, описанным через мета-признаки изображений и аннотаций. Основная идея заключается в том, чтобы накапливать информацию о характеристиках наборов данных, примененных моделях и достигнутых результатах, что позволяет проводить поиск аналогий при появлении новых задач.

Выполнена разработка алгоритма подбора моделей на основе базы знаний. Алгоритм опирается на принципы сопоставления входных мета-признаков с существующими записями, ранжирования моделей по их эффективности в схожих условиях и генерации рекомендаций по выбору наиболее подходящей архитектуры для решения конкретной задачи.

Эксперименты по обучению глубоких нейросетевых архитектур на различных аннотированных наборах растровых изображений послужили основой для формирования базы знаний, необходимой для обучения алгоритма рекомендаций. Это позволяет системе эффективно обрабатывать новые данные и формировать релевантные предсказания на основе ранее накопленного опыта. Ключевым этапом при этом является отбор репрезентативных и разнообразных наборов данных, отражающих как технические характеристики изображений (разрешение, контраст, экспозиция), так и предметную специфику задач (количество классов, плотность аннотаций, наличие частичных перекрытий объектов).

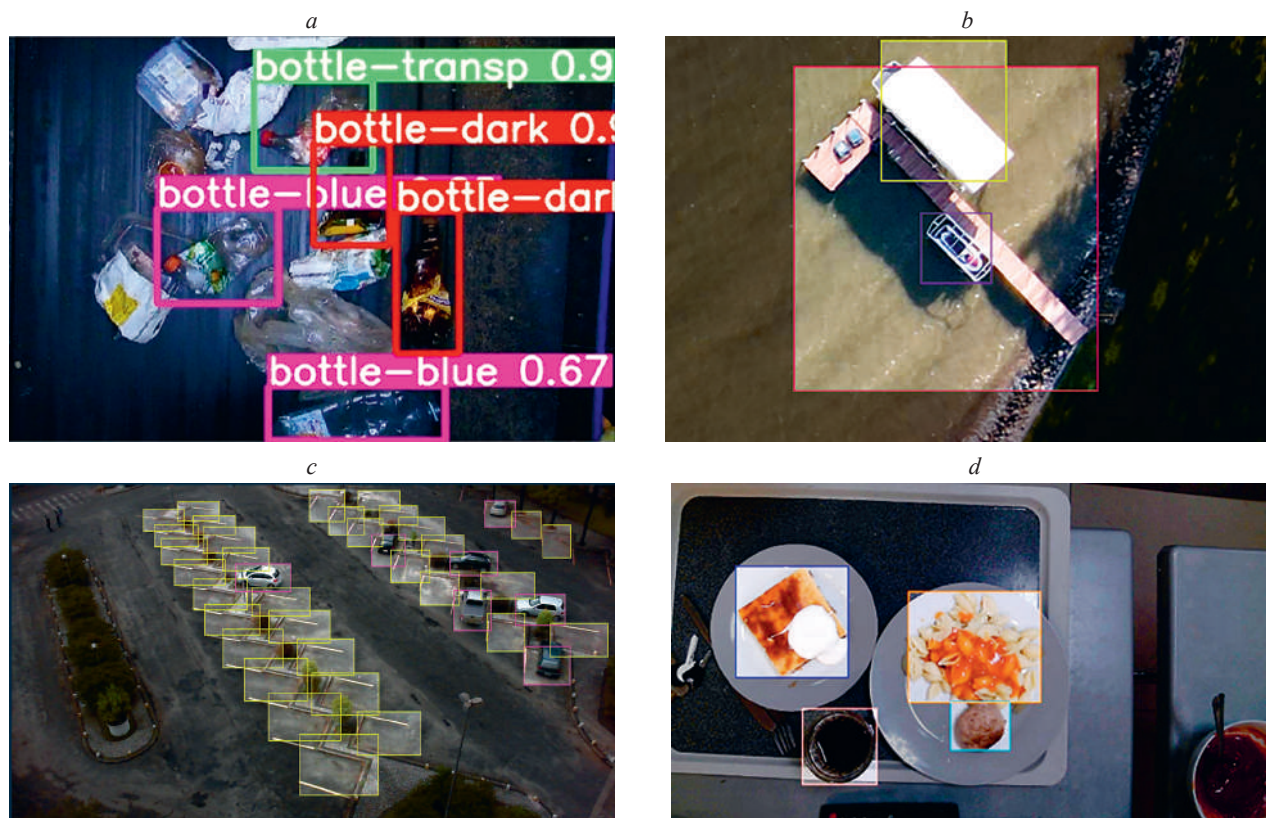


Рис. 1. Примеры аннотированных изображений из наборов данных: Warp (a); Aerial-maritime (b); PKLot (c); Food (d)  
 Fig. 1. Examples of annotated images from datasets: Warp (a); Aerial-maritime (b); PKLot (c); Food (d)

В рамках исследования использовались наборы данных, собранные с открытых платформ, таких как Roboflow, Kaggle и др., предоставляющих аннотированные изображения из различных прикладных областей. Выбранные наборы данных содержат сцены с частично скрытыми объектами, что особенно актуально для реальных условий эксплуатации (рис. 1).

Наборы данных различались по количеству изображений, числу классов и типам сцен, охватывая широкий спектр сценариев и условий съемки (табл. 1), полный список наборов данных доступен по ссылке<sup>1</sup>.

Для проведения экспериментов были выбраны различные модели глубокого обучения, каждая из которых обладает своими уникальными характеристиками и предпочтительными областями применения. В настоящей работе использовались следующие архитектуры: yolov5l, yolov5m, yolov5n, yolov5s, yolov5x, yolov7, yolov7x, yolov7-tiny, yolov8n, yolov8s, yolov8m, yolov8x, а также модели SSD и Faster R-CNN. Все модели были обучены на предварительно собранных наборах аннотированных изображений, что позволило сформировать базу знаний, содержащую примеры решений задач в различных сценариях. Для замера метрик FPS\_CPU использовался процессор Intel Core i9-12900KF (12th Gen). Обучение моделей и сбор данных для теку-

щей версии базы знаний потребовали около 960 ч работы графического процессора NVIDIA Quadro P6000. Обобщенные характеристики базы знаний включают: 20 тематически разнообразных наборов изображений; 307 уникальных классов объектов; 54 756 аннотированных изображений; 39 извлекаемых мета-признаков. Часть результатов обучения примененных моделей приведена в табл. 2. Полный массив результатов и метрик доступен по ссылке<sup>1</sup>.

В разработанной базе знаний отражены три типа сведений: мета-параметры наборов данных (**X**), признаки аннотаций (**Y**) и результаты оценки эффективности архитектур глубокого обучения (**B**).

Алгоритм подбора рекомендаций базируется на предварительно сформированной базе знаний состоит из двух этапов.

На этапе 1 алгоритм выделяет и преобразует мета-признаки (**X** и **Y**) из входного аннотированного набора данных в векторное представление. Затем, используя алгоритм поиска, из базы знаний выбирается наиболее «похожий» на входной набор данных, опираясь на сходство по мета-признакам (рис. 2). В качестве базовой функции поиска «похожего» набора данных в базе знаний использован алгоритм выбора минимального евклидова расстояния между векторами; при этом переменная **max\_sim** обозначает текущее максимальное значение меры сходства, а переменная **best\_id** — идентификатор набора данных (или соответствующей модели) в базе знаний, для которого данное значение достигается.

<sup>1</sup> BTO.ipynb — Colab [Электронный ресурс]. URL: [https://colab.research.google.com/drive/1IAIOdqRVE\\_7G83wzbHW0EoYDE9eWoGQv#scrollTo=Z5gJWdH8ie7b](https://colab.research.google.com/drive/1IAIOdqRVE_7G83wzbHW0EoYDE9eWoGQv#scrollTo=Z5gJWdH8ie7b) (дата обращения: 25.11.2025).

Таблица 1. Характеристики наборов растровых изображений

Table 1. Description of the characteristics of the image datasets

Название набора данных	Описание	Характеристики данных		
		Размер, пиксели	Количество классов	Количество изображений
WARP	Набор данных промышленных изображений твердых коммунальных отходов	1920 × 1080	28	5948
Aerial-maritime	Размеченные снимки доков, лодок, автомобилей и других объектов с беспилотного летательного аппарата	800 × 600	5	1016
Food	Размеченные изображения еды	256 × 256	7	802
PKLot	Изображения с парковок с разметкой мест для автомобилей	640 × 640	2	12 416

Таблица 2. Фрагмент результатов экспериментов по обучению моделей

Table 2. Fragment of the results of experiments on model training

Модель	Название набора данных <sup>1</sup>	Параметры обучения моделей							
		mAP50	mAP95	FPS_GPU, кадр/с	FPS_CPU, кадр/с	Ширина Bounding Box (BBox), отн. ед.	Высота BBox, отн. ед.	Соотношение BBox	Количество объектов на изображении
yolov5l	Warp	0,50	0,39	50	4	0,14	0,23	0,66	3,41
yolov5m	Food	0,74	0,42	61	7	0,17	0,27	0,64	3,11
yolov5n	Plantdoc	0,86	0,58	80	18	0,12	0,21	0,71	2,98
yolov5s	Website-screenshot	0,64	0,51	79	13	0,13	0,22	0,69	3,33
yolov5x	PKLot	0,51	0,40	42	2	0,15	0,25	0,61	3,21
yolov7x	Warp	0,56	0,43	15	4	0,13	0,23	0,67	3,41
yolov7	Food	0,53	0,41	15	6	0,13	0,23	0,67	3,41
yolov7-tiny	Website-screenshot	0,47	0,34	86	34	0,13	0,23	0,67	3,41

На этапе 2, из базы знаний экспериментов, проведенных на этапе 1 подобном наборе данных и с учетом пользовательской конфигурации (выбираются предпочтения по скорости и точности модели, а также аппаратной составляющей). Далее с помощью алгоритма ранжирования подбираются наиболее подходящие модели (рис. 3).

В результате работы алгоритмов (рис. 2 и 3) на выходе получаем список моделей (рис. 4), наиболее подходящий под набор данных и заданные условия.

### Экспериментальное исследование эффективности алгоритмов сопоставления наборов данных на основе мета-признаков

Для выбора наиболее точного алгоритма поиска «похожего» набора данных из базы знаний были проведены экспериментальные исследования, опирающиеся

<sup>1</sup> ВТО.ipynb — Colab [Электронный ресурс]. URL: [https://colab.research.google.com/drive/1IAIOdqRVE\\_7G83wzbHW0EoYDE9eWoGQv#scrollTo=Z5gJWdH8ie7b](https://colab.research.google.com/drive/1IAIOdqRVE_7G83wzbHW0EoYDE9eWoGQv#scrollTo=Z5gJWdH8ie7b) (дата обращения: 25.11.2025).

на методы, рассмотренные в разделе «Обзор методов машинного обучения для автоматизированного выбора моделей детектирования объектов». В качестве базовой модели (Baseline) выбран алгоритм поиска похожих наборов (рис. 2).

Были отобраны 12 тестовых наборов данных, сгруппированных по три в каждой из трех предметных областей (табл. 3), которые представлены в базе знаний. Задачей эксперимента является выбор алгоритма, который с наибольшей точностью и полнотой (метрика F1-score) сможет сопоставить тестовый набор данных с соответствующим набором данных из той же предметной области в базе знаний.

Первым шагом стала визуализация данных из базы знаний, крайне важный этап анализа, позволяющий наглядно представить извлеченную информацию, оценить ее качество и выявить ключевые особенности. Для алгоритма визуализации данных был выбран метод главных компонент (Principal Component Analysis, PCA) с двумя компонентами. На рис. 5 представлена двумерная (2D) проекция пространственного распределения наборов данных. Оси «Главная компонента 1» и «Главная компонента 2» отображают две первые главные ком-

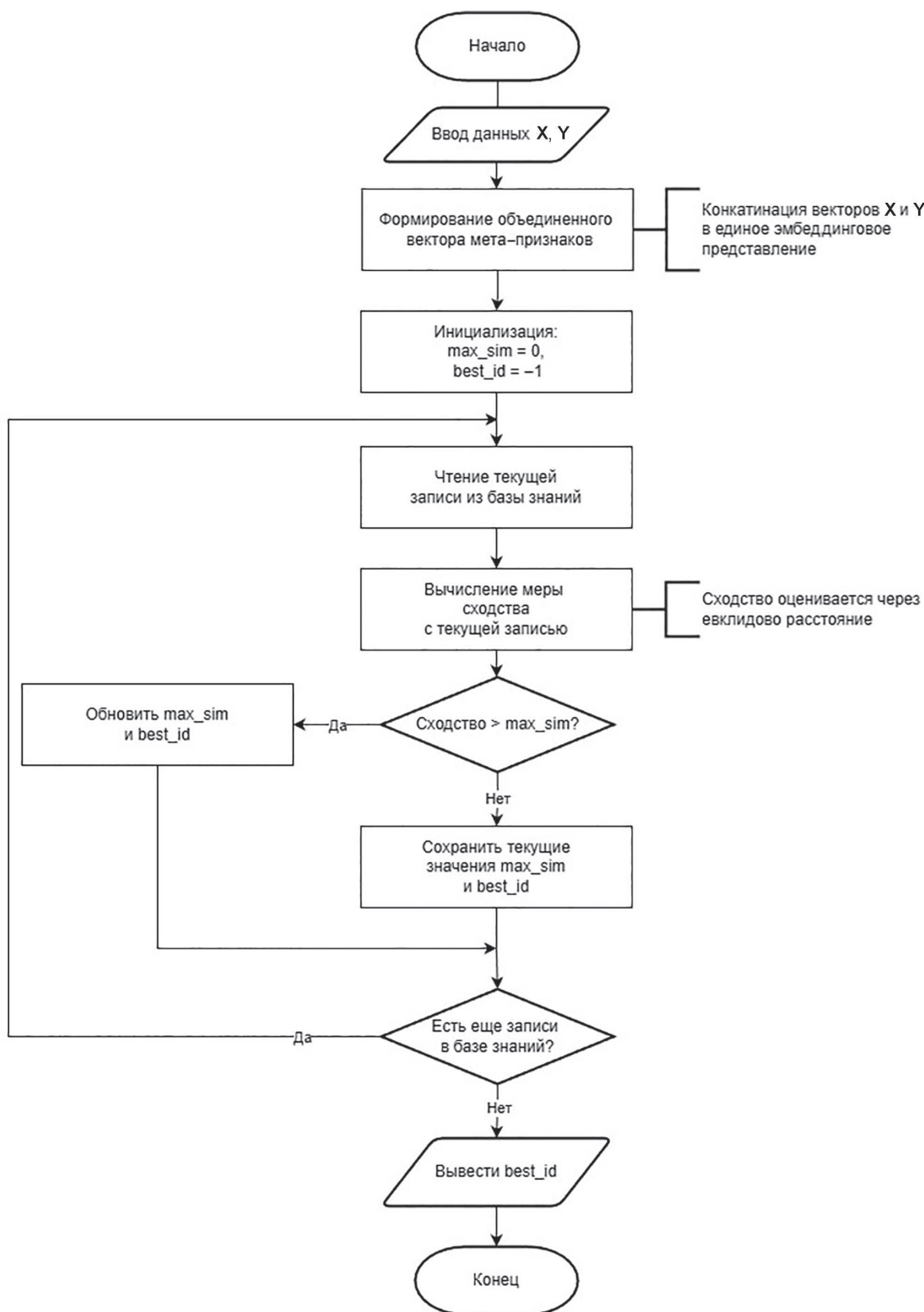


Рис. 2. Блок-схема алгоритма поиска схожих наборов данных  
 Fig. 2. Block diagram of the algorithm for searching similar datasets

поненты, которые объясняют наибольшую долю дисперсии исходных признаков. Каждая точка на графике соответствует одному набору данных, описанному в пространстве мета-признаков базы знаний. Положение точки на плоскости отражает схожесть или различие

наборов данных — чем ближе точки расположены друг к другу, тем более похожи соответствующие наборы данных по своим мета-признакам. Для PCA-графиков оси — это значения главных компонент, и у них нет физических единиц измерения. Главные компоненты

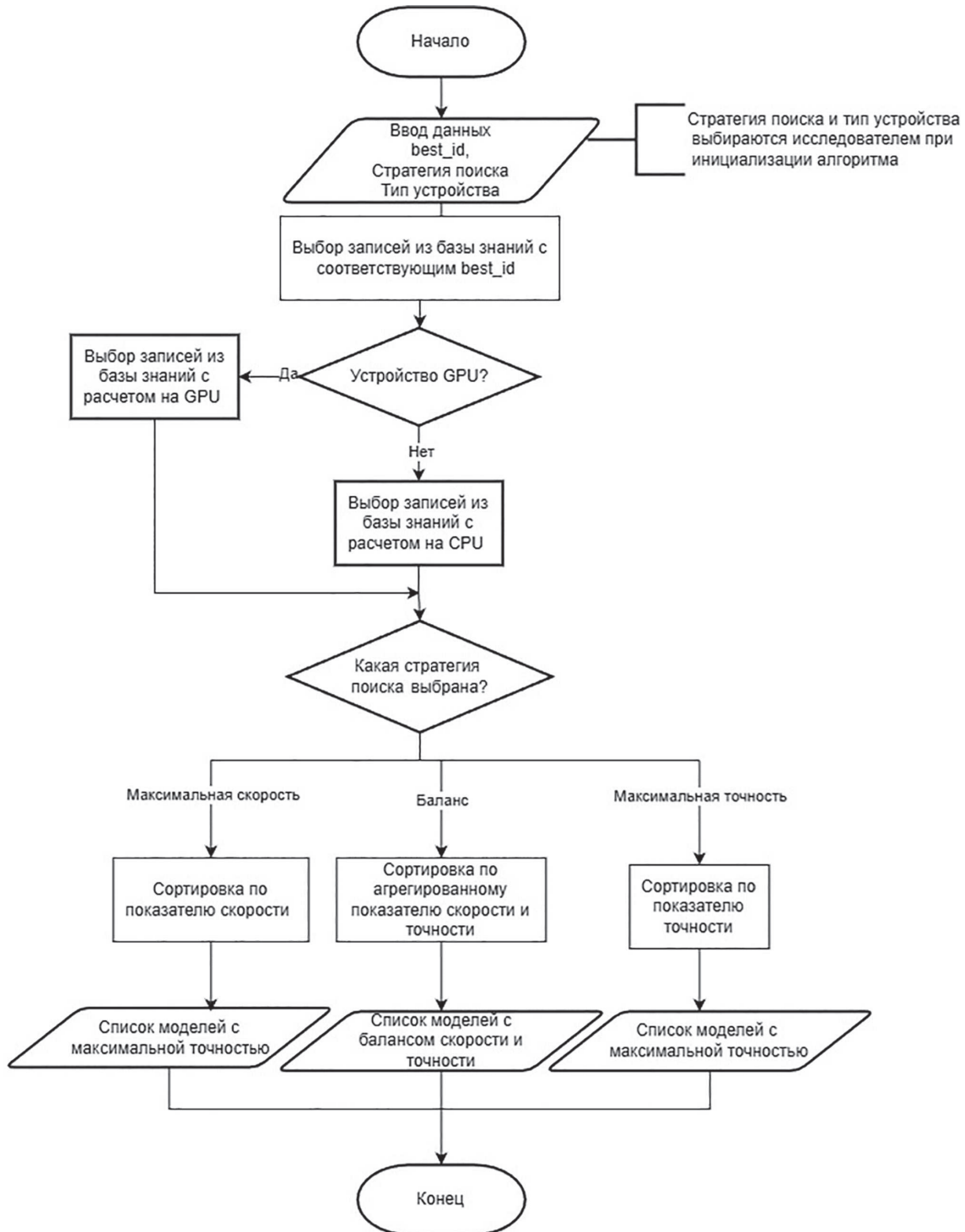


Рис. 3. Блок-схема алгоритма ранжирования моделей.

GPU — графический процессор, CPU — центральный процессор

Fig. 3. Block diagram of the model ranking algorithm where GPU stands for Graphics Processing Unit and CPU stands for Central Processing Unit

линейные комбинации стандартизированных признаков, поэтому значения выражаются в условных (безразмерных) единицах.

Визуализация распределения наборов данных после понижения размерности методом PCA (рис. 5) показала, что проекции мета-признаков формируют кластеры с минимальным перекрытием между классами, что

свидетельствует о частичной линейной разделимости в пространстве признаков. Для количественного подтверждения этого наблюдения были применены методы классификации (CatBoost, Random Forest, MLP), показавшие высокие значения метрики F1-score (до 0,82 — табл. 4), что дополнительно указывает на возможность точного разделения классов по мета-признакам. Это

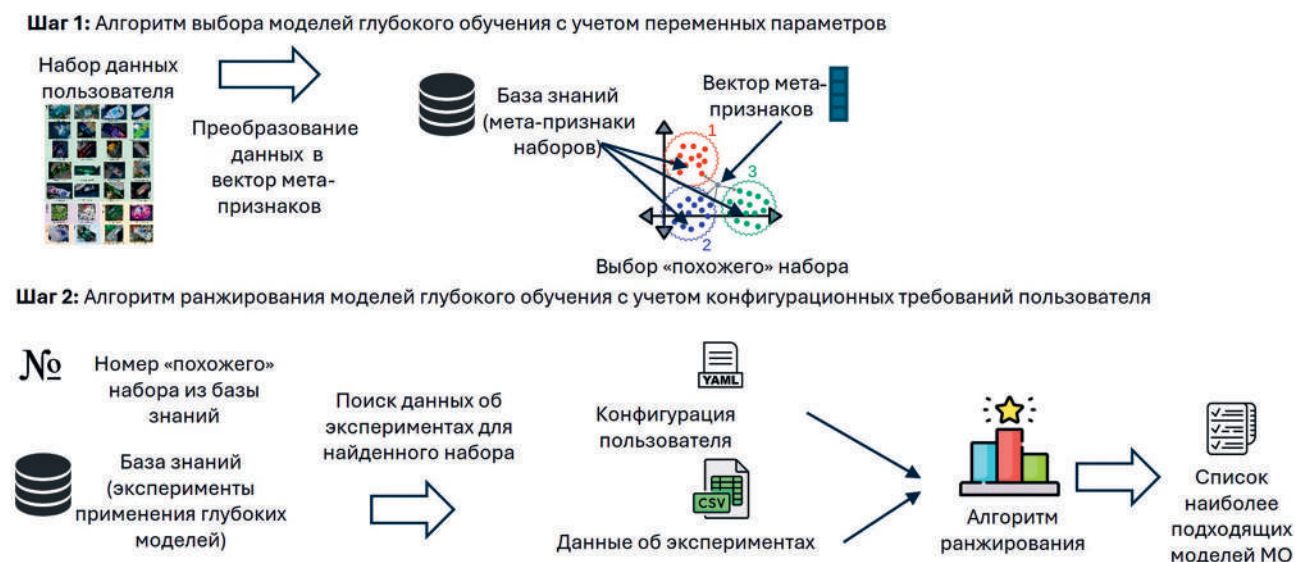


Рис. 4. Общая схема алгоритма подбора модели машинного обучения (МО).

YAML — YAML Ain't Markup Language, CSV — Comma-Separated Values

Fig. 4. The general algorithm scheme for selecting a machine learning model where YAML stands for YAML Ain't Markup Language and CSV stands for Comma-Separated Values

Таблица 3. Описание тестовых наборов данных

Table 3. Description of test datasets

Название набора данных	«Похожий» набор данных из базы знаний	Описание
Card 1	GRZ Dataset	Номерные знаки
Card 2		
Card 3		
Socker player 1	Football players	Игроки на футбольном поле
Socker player 2		
Socker player 3		
Car_plate 1	Card object detection	Различные типы карт
Car_plate 2		
Car_plate 3		
PPE 1	Constuction site safety	Средства индивидуальной защиты
PPE 2		
PPE 3		

позволяет обоснованно использовать классификационные методы для сопоставления новых наборов данных с аналогами в базе знаний. В ходе эксперимента были проведены измерения для предложенных методов, а также измерения для алгоритма поиска похожих наборов (рис. 2) с использованием различных расстояний, а также рассмотрено дополнительное преобразование векторов-признаков методом PCA перед подачей в цикл алгоритма. Для поиска наиболее «близкого» набора для методов кластеризации использовалось косинусное (CD) и евклидово (ED) расстояния. Наиболее эффективными показали себя методы кластеризации и классификации. Результаты экспериментов показаны в табл. 4.

На основе понижения размерности с использованием метода PCA визуализировано распределение на-

боров данных в пространстве мета-признаков. Такая проекция позволяет оценить взаимное расположение объектов в признаковом пространстве и выявить кластеры близких по структуре и содержанию наборов данных. На графике (рис. 6) показан результат работы алгоритма CatBoost классификации.

Учитывая стабильность и высокую точность классификации, CatBoostClassifier был выбран в качестве алгоритма для решения задачи поиска «похожих» наборов данных. Производительность выбранного алгоритма делает его надежным кандидатом для интеграции в автоматизированную систему рекомендаций архитектур.

Для оценки временной эффективности предложенного алгоритма были проведены сравнения с тради-

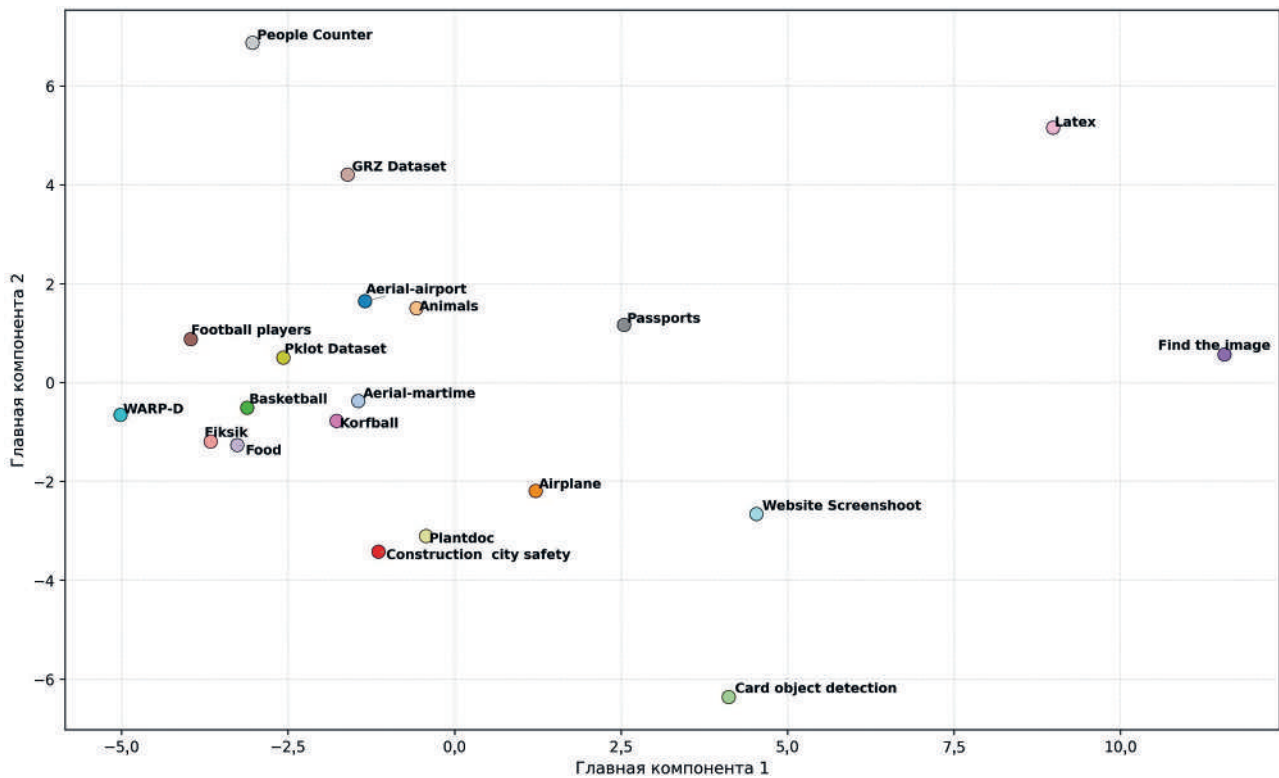


Рис. 5. Распределение наборов из базы знаний в 2D-пространстве после применения PCA

Fig. 5. Distribution of the knowledge base datasets in 2D space after applying PCA

ционным эмпирическим подходом подбора моделей. В типичных сценариях специалист вручную тестирует 3–5 моделей на небольшом подмножестве данных, что в среднем занимает  $25 \pm 3$  мин на задачу, включая выбор

архитектуры, настройку параметров и предварительное обучение.

В случае применения предложенного алгоритма выбор «похожего» набора из базы знаний и ранжиро-

Таблица 4. Полученные метрики для исследуемых методов

Table 4. The obtained metrics for the investigated methods

Название алгоритма	Precision	Recall	F1-score
<b>Baseline и модификации</b>			
GetSimilarDatasetNumber (ED)	0,75	0,33	0,45
GetSimilarDatasetNumber (CD)	0,25	0,17	0,20
PCA + GetSimilarDatasetNumber (ED)	0,88	0,50	0,63
PCA + GetSimilarDatasetNumber (CD)	0,88	0,42	0,55
<b>Методы классификации</b>			
Random Forest Classifier	0,92	0,67	0,74
CatBoost Classifier	0,92	0,75	0,82
AdaBoost Classifier	0,36	0,42	0,35
Multi-Layer Perceptron	0,88	0,42	0,55
<b>Методы кластеризации</b>			
k-means (ED)	0,75	0,60	0,67
Agglomerative clustering (ED)	0,71	0,50	0,59
k-means (CD)	0,78	0,70	0,74
Agglomerative clustering (CD)	0,67	0,40	0,5
k-means (MD)	0,75	0,60	0,67
Agglomerative clustering (MD)	0,71	0,50	0,59

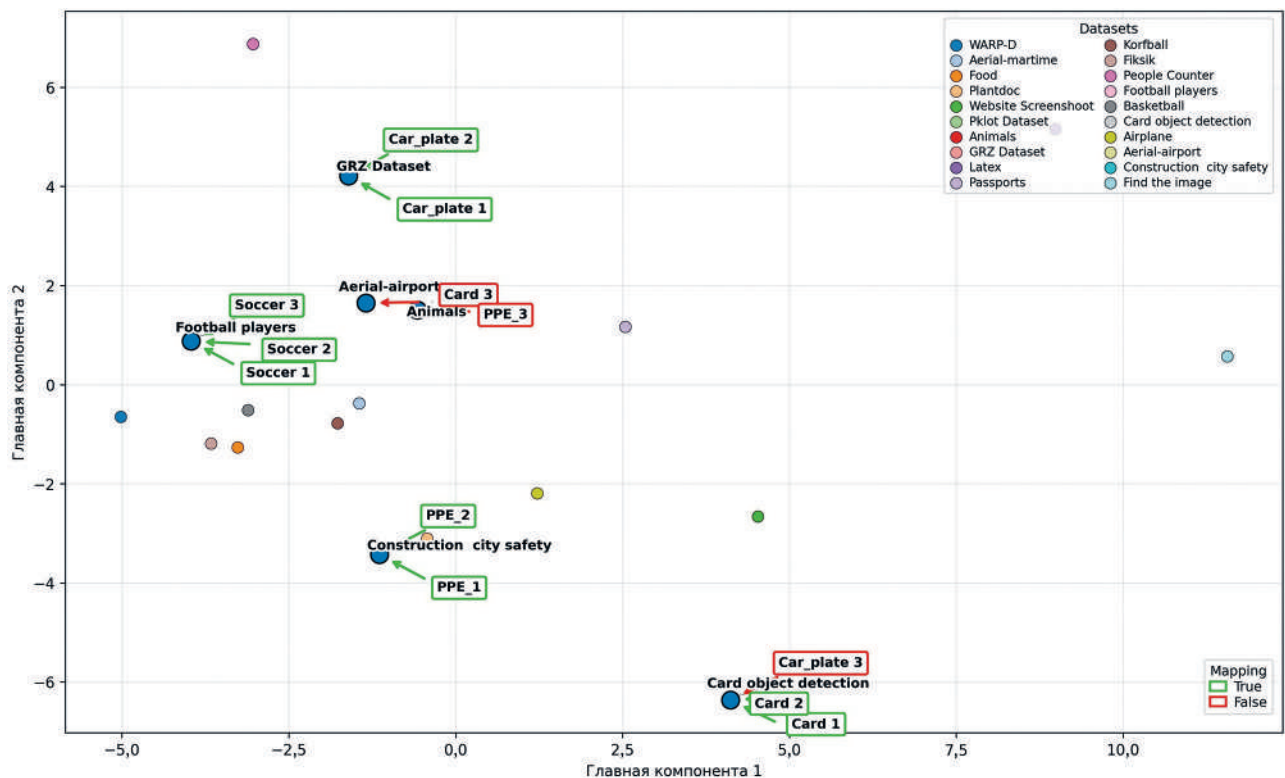


Рис. 6. 2D-распределение PCA классов с указанием принадлежности тестовых наборов данных для метода классификации CatBoostClassifier

Зеленой рамкой отмечены наборы данных, для которых найден набор данных в базе знаний из той же предметной области; красной рамкой — неверно сопоставленные наборы данных по результатам экспериментов

Fig. 6. 2D distribution of PCA classes with indication of the test sets belonging for the CatBoostClassifier. The green box indicates data sets for which a set of data has been found in the knowledge base from the same subject area; the red box indicates incorrectly matched data sets based on experimental results

вание моделей по предпочтениям пользователя выполняется автоматически и занимает около  $10 \pm 2$  мин, включая извлечение мета-признаков, сопоставление и генерацию рекомендаций. Значительная часть ручных операций (экспертная оценка сложности данных, подбор архитектуры, первичное тестирование моделей) не проводится, поскольку информация уже присутствует в базе знаний по аналогичным задачам. Таким образом, общее сокращение времени достигает порядка 40 %, что обусловлено автоматизацией ключевых этапов принятия решения.

### Заключение

Представлен метод рекомендательного выбора моделей детектирования объектов, основанный на анализе мета-признаков изображений и использовании базы знаний выполненных ранее экспериментов. Предложенный алгоритм реализован в открытом фреймворке Object Detection Recommendation System и обеспечивает автоматизацию процесса подбора моделей, что позволило сократить среднее время выбора на 40 % по сравнению с традиционными подходами (с  $25 \pm 3$  до

$10 \pm 2$  мин,  $p < 0,001$ , где  $p$  — уровень статистической значимости различий). Модель CatBoostClassifier, лежащая в основе алгоритма, демонстрирует высокую точность (F1-score = 0,82), превосходя популярные методы кластеризации, такие как  $k$ -means (F1-score = 0,67). Экспериментальные результаты подтверждают устойчивость метода при работе с разнородными наборами данных и в условиях ограниченных вычислительных ресурсов. В частности, при значительном дисбалансе классов ( $D > 0,5$ ), где  $D$  — коэффициент дисбаланса классов в наборе данных, точность рекомендаций сохраняется на уровне 85.

Предложенный метод показал практическую применимость в задачах, где критически важен баланс между точностью и производительностью, включая системы видеонаблюдения, медицинскую диагностику, автономный транспорт и промышленный контроль. В дальнейшем планируется расширение базы знаний за счет интеграции данных с 3D-сенсоров и LiDAR (Light Detection and Ranging), сравнение с AutoML-платформами, оптимизация под edge-устройства и адаптация метода для задач семантической сегментации и трекинга объектов.

## Литература

## References

1. Edozie E., Shuaibu A.N., John U.K., Sadiq B.O. Comprehensive review of recent developments in visual object detection based on deep learning // *Artificial Intelligence Review*. 2025. V. 58. N 9. P. 277. <https://doi.org/10.1007/s10462-025-11284-w>
2. Yudin D., Zakharenko N., Smetanin A., Filonov R., Kichik M., Kuznetsov V., et al. Hierarchical waste detection with weakly supervised segmentation in images from recycling plants // *Engineering Applications of Artificial Intelligence*. 2024. V. 128. P. 128. <https://doi.org/10.1016/j.engappai.2023.107542>
3. Alyari F., Navimipour N.J. Recommender systems: a systematic review of the state of the art literature and suggestions for future research // *Kybernetes*. 2018. V. 47. N 5. P. 985–1017. <https://doi.org/10.1108/K-06-2017-0196>
4. Feuer M., Klein A., Eggersperger K., Springenberg J.T., Blum M., Hutter F. Efficient and robust automated machine learning // *Proc. of the 29th International Conference on Neural Information Processing Systems*. 2015. V. 2. P. 2755–2763.
5. Bhanuse R., Mal S. A Systematic Review: deep learning based E-Learning Recommendation System // *Proc. of the International Conference on Artificial Intelligence and Smart Systems (ICAIS)*. 2021. P. 190–197. <https://doi.org/10.1109/icaais50930.2021.9395835>
6. Samin H., Azim T. Knowledge based recommender system for academia using machine learning: a case study on higher education landscape of Pakistan // *IEEE Access*. 2019. V. 7. P. 67081–67093. <https://doi.org/10.1109/ACCESS.2019.2912012>
7. Rivolli A., Garcia L.P.F., Soares C., Vanschoren J., de Carvalho A.C.P.L.F. Characterizing classification datasets: a study of meta-features for meta-learning // *arXiv*. 2018. arXiv:1808.10406v2. <https://doi.org/10.48550/arXiv.1808.10406>
8. Okezie Ben-John. *Beyond Accuracy: Understanding Precision and Recall in Machine Learning*. [Электронный ресурс]. URL: <https://benjohnokezie.medium.com/beyond-accuracy-understanding-precision-and-recall-in-machine-learning-9a07db9bc46c> (дата обращения: 20.07.2025)
9. Gao X., Luo H., Wang Q., Zhao F., Ye L., Zhang Y. A human activity recognition algorithm based on stacking denoising autoencoder and lightGBM // *Sensors*. 2019. V. 19. N 4. P. 947. <https://doi.org/10.3390/s19040947>
10. Salehin I., Islam M.S., Saha P., Noman S.M., Tunj A., Hasan M.M., Baten M.A. AutoML: A systematic review on automated machine learning with neural architecture search // *Journal of Information and Intelligence*. 2024. V. 2. N 1. P. 52–81. <https://doi.org/10.1016/j.jiixd.2023.10.002>
11. Wu B.K. K-means clustering algorithm and Python implementation // *Proc. of the IEEE International Conference on Computer Science, Artificial Intelligence and Electronic Engineering (CSAIEE)*. 2021. P. 55–59. <https://doi.org/10.1109/csaiee54046.2021.9543260>
12. Manoj T., Sajeev G.P. Conductivity based agglomerative spectral clustering for community detection // *Proc. of the 6th International Conference on Image Information Processing (ICIIP)*. 2021. P. 385–389. <https://doi.org/10.1109/iciip53038.2021.9702554>
13. Dorogush A.V., Ershov V., Gulina A. CatBoost: gradient boosting with categorical features support // *arXiv*. 2018. arXiv:1810.11363v1. <https://doi.org/10.48550/arXiv.1810.11363>
14. Mollas I., Bassiliades N., Tsoumakas G. Conclusive local interpretation rules for random forests // *Data Mining and Knowledge Discovery*. 2022. V. 36. N 4. P. 1521–1574. <https://doi.org/10.1007/s10618-022-00839-y>
15. Chelms C., Qi W. Hierarchical MultiClass AdaBoost // *Proc. of the IEEE International Conference on Big Data*. 2021. P. 5063–5070. <https://doi.org/10.1109/bigdata52589.2021.9671291>
16. Singh G., Sachan M. Multi-layer perceptron (MLP) neural network technique for offline handwritten Gurmukhi character recognition // *Proc. of the IEEE International Conference on Computational Intelligence and Computing Research*. 2014. P. 1–5. <https://doi.org/10.1109/iccic.2014.7238334>
17. Alkentar S.M., Alshawa B., Assalem A., Karakolla D. Practical comparison of the accuracy and speed of YOLO, SSD and Faster RCNN for drone detection // *Journal of Engineering*. 2021. V. 27. N 8. P. 19–31. <https://doi.org/10.31026/j.eng.2021.08.02>
18. Diwan T., Anirudh G., Tembhone J.V. Object detection using YOLO: challenges, architectural successors, datasets and applications // *Multimedia Tools and Applications*. 2023. V. 82. N 6. P. 9243–9275. <https://doi.org/10.1007/s11042-022-13644-y>
1. Edozie E., Shuaibu A.N., John U.K., Sadiq B.O. Comprehensive review of recent developments in visual object detection based on deep learning. *Artificial Intelligence Review*, 2025, vol. 58, no. 9, pp. 277. <https://doi.org/10.1007/s10462-025-11284-w>
2. Yudin D., Zakharenko N., Smetanin A., Filonov R., Kichik M., Kuznetsov V., et al. Hierarchical waste detection with weakly supervised segmentation in images from recycling plants. *Engineering Applications of Artificial Intelligence*, 2024, vol. 128, pp. 128. <https://doi.org/10.1016/j.engappai.2023.107542>
3. Alyari F., Navimipour N.J. Recommender systems: a systematic review of the state of the art literature and suggestions for future research. *Kybernetes*. 2018, vol. 47, no. 5, pp. 985–1017. <https://doi.org/10.1108/K-06-2017-0196>
4. Feuer M., Klein A., Eggersperger K., Springenberg J.T., Blum M., Hutter F. Efficient and robust automated machine learning. *Proc. of the 29th International Conference on Neural Information Processing Systems*, 2015, vol. 2, pp. 2755–2763.
5. Bhanuse R., Mal S. A Systematic Review: deep learning based E-Learning Recommendation System. *Proc. of the International Conference on Artificial Intelligence and Smart Systems (ICAIS)*, 2021, pp. 190–197. <https://doi.org/10.1109/icaais50930.2021.9395835>
6. Samin H., Azim T. Knowledge based recommender system for academia using machine learning: a case study on higher education landscape of Pakistan. *IEEE Access*, 2019, vol. 7, pp. 67081–67093. <https://doi.org/10.1109/ACCESS.2019.2912012>
7. Rivolli A., Garcia L.P.F., Soares C., Vanschoren J., de Carvalho A.C.P.L.F. Characterizing classification datasets: a study of meta-features for meta-learning. *arXiv*, 2018. arXiv:1808.10406v2. <https://doi.org/10.48550/arXiv.1808.10406>
8. Okezie Ben-John. *Beyond Accuracy: Understanding Precision and Recall in Machine Learning*. Available at: <https://benjohnokezie.medium.com/beyond-accuracy-understanding-precision-and-recall-in-machine-learning-9a07db9bc46c> (accessed: 20.07.2025)
9. Gao X., Luo H., Wang Q., Zhao F., Ye L., Zhang Y. A human activity recognition algorithm based on stacking denoising autoencoder and lightGBM. *Sensors*, 2019, vol. 19, no. 4, pp. 947. <https://doi.org/10.3390/s19040947>
10. Salehin I., Islam M.S., Saha P., Noman S.M., Tunj A., Hasan M.M., Baten M.A. AutoML: A systematic review on automated machine learning with neural architecture search. *Journal of Information and Intelligence*, 2024, vol. 2, no. 1, pp. 52–81. <https://doi.org/10.1016/j.jiixd.2023.10.002>
11. Wu B.K. K-means clustering algorithm and Python implementation. *Proc. of the IEEE International Conference on Computer Science, Artificial Intelligence and Electronic Engineering (CSAIEE)*, 2021, pp. 55–59. <https://doi.org/10.1109/csaiee54046.2021.9543260>
12. Manoj T., Sajeev G.P. Conductivity based agglomerative spectral clustering for community detection. *Proc. of the 6th International Conference on Image Information Processing (ICIIP)*, 2021, pp. 385–389. <https://doi.org/10.1109/iciip53038.2021.9702554>
13. Dorogush A.V., Ershov V., Gulina A. CatBoost: gradient boosting with categorical features support. *arXiv*, 2018. arXiv:1810.11363v1. <https://doi.org/10.48550/arXiv.1810.11363>
14. Mollas I., Bassiliades N., Tsoumakas G. Conclusive local interpretation rules for random forests. *Data Mining and Knowledge Discovery*, 2022, vol. 36, no. 4, pp. 1521–1574. <https://doi.org/10.1007/s10618-022-00839-y>
15. Chelms C., Qi W. Hierarchical MultiClass AdaBoost. *Proc. of the IEEE International Conference on Big Data*, 2021, pp. 5063–5070. <https://doi.org/10.1109/bigdata52589.2021.9671291>
16. Singh G., Sachan M. Multi-layer perceptron (MLP) neural network technique for offline handwritten Gurmukhi character recognition. *Proc. of the IEEE International Conference on Computational Intelligence and Computing Research*, 2014, pp. 1–5. <https://doi.org/10.1109/iccic.2014.7238334>
17. Alkentar S.M., Alshawa B., Assalem A., Karakolla D. Practical comparison of the accuracy and speed of YOLO, SSD and Faster RCNN for drone detection. *Journal of Engineering*, 2021, vol. 27, no. 8, pp. 19–31. <https://doi.org/10.31026/j.eng.2021.08.02>
18. Diwan T., Anirudh G., Tembhone J.V. Object detection using YOLO: challenges, architectural successors, datasets and applications. *Multimedia Tools and Applications*, 2023, vol. 82, no. 6, pp. 9243–9275. <https://doi.org/10.1007/s11042-022-13644-y>

19. Chen X.Q., Zhang Q.Y., Lin M.H., Yang G.Y., He C. No-reference color image quality assessment: from entropy to perceptual quality // EURASIP Journal on Image and Video Processing. 2019. V. 2019. N 1. P. 77. <https://doi.org/10.1186/s13640-019-0479-7>

19. Chen X.Q., Zhang Q.Y., Lin M.H., Yang G.Y., He C. No-reference color image quality assessment: from entropy to perceptual quality. EURASIP Journal on Image and Video Processing, 2019, vol. 2019, no. 1, pp. 77. <https://doi.org/10.1186/s13640-019-0479-7>

#### Авторы

**Сметанин Артём Алексеевич** — главный программист, ООО «Планетарий 1», Санкт-Петербург, 196084, Российская Федерация; инженер, Университет ИТМО, Санкт-Петербург, 197101, Российская Федерация, [sc](https://orcid.org/0000-0002-8158-7450) 57221840206, <https://orcid.org/0000-0002-8158-7450>, [artem\\_smetanin@niuitmo.ru](mailto:artem_smetanin@niuitmo.ru)

**Духанов Алексей Валентинович** — доктор технических наук, доцент, доцент, Университет ИТМО, Санкт-Петербург, 197101, Российская Федерация, [sc](https://orcid.org/0000-0002-1011-9932) 56218447300, <https://orcid.org/0000-0002-1011-9932>, [dukhanov@itmo.ru](mailto:dukhanov@itmo.ru)

**Герасимчук Михаил Юрьевич** — ведущий инженер, ООО МТС ДИДЖИТАЛ, Санкт-Петербург, 197046, Российская Федерация, аспирант, Университет ИТМО, Санкт-Петербург, 197101, Российская Федерация, [sc](https://orcid.org/0009-0008-8005-9550) 57919215000, <https://orcid.org/0009-0008-8005-9550>, [mishagerasimchuk@itmo.ru](mailto:mishagerasimchuk@itmo.ru)

**Гудов Евгений Валерьевич** — специалист, аспирант, Академия труда и социальных отношений, Москва, 119454, Российская Федерация; генеральный директор, ООО «Планетарий 1», Санкт-Петербург, 196084, Российская Федерация, <https://orcid.org/0009-0006-4534-0262>, [evgenes@gmail.com](mailto:evgenes@gmail.com)

#### Authors

**Artem A. Smetanin** — Chief Programmer, Planetarium one, Saint Petersburg, 196084, Russian Federation; Engineer, ITMO University, Saint Petersburg, 197101, Russian Federation, [sc](https://orcid.org/0000-0002-8158-7450) 57221840206, <https://orcid.org/0000-0002-8158-7450>, [artem\\_smetanin@niuitmo.ru](mailto:artem_smetanin@niuitmo.ru)

**Alexey V. Dukhanov** — D.Sc., Associate Professor, Associate Professor, ITMO University, Saint Petersburg, 197101, Russian Federation, [sc](https://orcid.org/0000-0002-1011-9932) 56218447300, <https://orcid.org/0000-0002-1011-9932>, [dukhanov@itmo.ru](mailto:dukhanov@itmo.ru)

**Mikhail Yu. Gerasimchuk** — Leading Engineer, MTS DIGITAL LLC, Saint Petersburg, 197046, Russian Federation; PhD Student, ITMO University, Saint Petersburg, 197101, Russian Federation, [sc](https://orcid.org/0009-0008-8005-9550) 57919215000, <https://orcid.org/0009-0008-8005-9550>, [mishagerasimchuk@itmo.ru](mailto:mishagerasimchuk@itmo.ru)

**Evgeny V. Gudov** — Specialist, PhD Student, Academy of Labour and Social Relations, Moscow, 119454, Russian Federation; CEO, Planetarium one, Saint Petersburg, 196084, Russian Federation, <https://orcid.org/0009-0006-4534-0262>, [evgenes@gmail.com](mailto:evgenes@gmail.com)

*Статья поступила в редакцию 14.02.2025  
Одобрена после рецензирования 11.11.2025  
Принята к печати 21.01.2026*

*Received 14.02.2025  
Approved after reviewing 11.11.2025  
Accepted 21.01.2026*



Работа доступна по лицензии  
Creative Commons  
«Attribution-NonCommercial»