

## **РАЗРАБОТКА АРХИТЕКТУРЫ СИСТЕМ ИНТЕЛЛЕКТУАЛЬНОГО ПОИСКА ИЗОБРАЖЕНИЙ ПО СОДЕРЖАНИЮ**

**Аннотация.** В данной статье предлагается архитектура системы поиска изображений по содержанию (Content-Based Image Retrieval) на основе модели векторизации, имеющая свойства, не достигнутые в более ранних работах.

**Ключевые слова:** интеллектуальный поиск изображений, компьютерное зрение, обработка изображений, векторная база данных.

**Введение.** В ходе жизненного цикла хранилища изображений количество их постоянно увеличивается, что усложняет поиск изображений вручную [1].

Разрабатываемая архитектура системы поиска изображений нацелена на использование в следующих типах систем:

- 1) фотогалереи;
- 2) соцсети;
- 3) поисковые системы;
- 4) системы аутентификации;
- 5) система распознавания в электронном документообороте.

Есть два способа поиска:

- поиск по содержанию (CBIR);
- поиск по текстовым аннотациям (Description-Based Image Retrieval) [2].

CBIR — поиск изображения в базе данных на основе сравнения низкоуровневых характеристик, в отличие от DBIR не требует дополнительной пометки изображения [3-4].

Для его получения есть методы вычисления вектора признаков:

- вариационный ряд (гистограмма);
- локальные дескрипторы;
- дескрипторы особых точек;

- дескрипторы на основе хэш-функций;
- нейросетевые дескрипторы [5].

Требования, которым должны отвечать системы, построенные с применением разрабатываемой архитектуры, следующие:

- 1) отказоустойчивость;
- 2) высокая скорость добавления новых изображений;
- 3) сохранение высокой скорости поиска изображений с ростом общего количества изображений;
- 4) возможность замены реализаций модулей без больших трудозатрат.

**Проблема исследования.** Рассмотренные архитектуры не обеспечивают одновременное удовлетворение всех выдвигаемых требований (табл. 1).

*Таблица 1*

**Сравнение архитектур системы поиска изображений**

Архитектура	Требование			
	1	2	3	4
Архитектура баз данных для задачи поиска изображений [1]	X	X	+	X
Retrieval Systems Image [3]	X	X	+	+
Multilayer Architecture for Content-based Image [4]	X	X	+	+
Архитектура системы биометрической аутентификации [5]	X	X	+	X

Цель работы — разработать архитектуру системы СВIR, которая позволит обеспечивать удовлетворение вышеописанных требований.

**Материалы и методы.** Чтобы система соответствовала заявленным требованиям, требуется наличие следующих компонентов:

- 1) объектное хранилище;
- 2) индексатор;
- 3) модель для векторизации;
- 4) поисковый сервис;

- 5) векторная СУБД;
- 6) очередь сообщений.

Клиенты взаимодействуют с объектным хранилищем и поисковым сервисом, перед которыми опционально могут располагаться API-шлюзы, обратные прокси, балансировщики нагрузки и прочие сущности, которые не являются значимыми в рамках данной работы.

Объектное хранилище решает проблемы, связанные с хранением и доступом к изображениям, в т.ч. горизонтальное масштабирование и абстракция от нижележащих файловых систем. Для построения индекса изображений требуется, чтобы хранилище предоставляло уведомления о добавлении новых объектов. Реализация протокола AWS S3 не является необходимым требованием, но т. к. он стал де-факто стандартом для объектных хранилищ и поддерживается не только продуктами Amazon, в дальнейшем будем предполагать, что используется именно он, и называть хранилище S3.

Модель для векторизации (далее векторизатор) преобразует изображения (при индексировании) или запросы клиента (при поиске) в векторы одинаковой размерности так, чтобы расстояние между точками, полученными из запроса и изображения было тем меньше, чем они ближе по содержанию.

Индексатор выполняет роль слушателя сообщений о добавлении нового изображения от S3 хранилища (паттерн наблюдатель). При получении сообщения он должен:

- получить изображение из S3;
- векторизовать изображение, используя векторизатор;
- сохранить соответствие изображение-вектор в векторную БД.

Поисковый сервис обрабатывает запросы клиентского приложения следующим образом:

- 1) векторизует запрос клиента, используя векторизатор;
- 2) запрашивает у векторной базы данных список наиболее подходящих изображений;
- 3) возвращает полученную информацию клиенту.

Наиболее подходящими считаются изображения с наименьшим расстоянием до точки, полученной в ходе векторизации запроса.

Векторная база данных (например, Milvus, Weaviate, Qdrant, PostgreSQL + pgvector) хранит соответствия векторных представлений изображений и самих изображений и предоставляет возможность эффективно выполнять поиск по большому количеству векторов. Важно, чтобы база данных поддерживала метрики расстояния, которые требуются для выбранного векторизатора.

Очередь сообщений (например, Apache Kafka, RabbitMQ, Apache ActiveMQ, NATS) обеспечивает отказоустойчивость взаимодействия S3-индексатора и скорости загрузки изображений. Для примера рассмотрим следующую ситуацию: индексатор недоступен. Без очереди при загрузке система вынуждена либо ожидать доступности индексатора, либо не добавлять изображение в индекс. При наличии очереди она гарантирует доставку и обработку сообщения после того, как индексатор станет доступен. Описываемая архитектура предполагает, что быстрая загрузка изображения более важна, чем малая задержка между загрузкой и доступностью при поиске.

При необходимости можно использовать несколько разных векторизаторов.

1. CLIP,
2. UForm,
3. ImageBind.

В таком случае поисковый сервис должен либо выбрать один из результатов, либо агрегировать их.

**Заключение.** В результате данной работы было реализовано следующее:

- рассмотрены четыре готовых архитектуры и выявлены их недостатки;
- выявлены компоненты, необходимые для удовлетворения списка требований к архитектуре;
- установлены возможные реализации модулей системы.

На основе реализованного была разработана архитектура, удовлетворяющая всем описанным требованиям.

В дальнейшем архитектура, разработанная в результате данной работы, станет основой программного продукта, разрабатываемого авторами.

## СПИСОК ЛИТЕРАТУРЫ

1. Десятников И.Е. Алгоритмы поиска изображений в базах видеоданных / Десятников И.Е., Утробин В.А. — Текст: непосредственный // Компьютерная оптика. — 2011. — Т. 35, № 3. — С. 416-422.
2. Десятников И.Е. Построение архитектуры баз данных для задачи поиска изображений / Десятников И.Е., Утробин В.А. — Текст: непосредственный // Вестник Нижегородского государственного университета. — 2012. — № 5(2). — С. 319-325.
3. Grycuk R. Multilayer Architecture for Content-Based Image Retrieval Systems / R. Grycuk, P. Najgebauer, R. Nowicki, R. Scherer. — Direct text // International Conference on Service-Oriented Computing and Applications: 12<sup>th</sup> int. sci.-tech. conf. 12th 19-20 Jan. 2019 y. — Taiwan, 2019. — Pp. 119-126.
4. Rao S. An approach for cbir system through multi-layer neural network / P. S. Rao, E. Vamsidhar, G. Raju, R. Satapati, K. Varma. — Direct text // International Journal of Engineering Science and Technology. — 2010. — Vol. 2, № 4. — Pp. 559-563.
5. Прозоров Д.Е. Применение легковесной сиамской нейросети для формирования вектора признаков в системе васкулярной аутентификации / Д.Е. Прозоров, А.В. Земцов — Текст: непосредственный // Компьютерная оптика. — 2023. — Т. 47, № 3. — С. 433-441.