

# Exploring Patents Visually: An Interactive Search System for Multimodal Patent Image Search and Interpretation

Sushil Awale<sup>1,\*</sup>, Eric Müller-Budack<sup>1</sup>, Rahim Delaviz<sup>2</sup> and Ralph Ewerth<sup>1,3,4</sup>

<sup>1</sup>TIB – Leibniz Information Centre for Science and Technology, Hannover, Germany

<sup>2</sup>European Patent Office, The Hague, The Netherlands

<sup>3</sup>University of Marburg and hessian.AI – Hessian Center for Artificial Intelligence, Marburg, Germany

<sup>4</sup>L3S Research Center, Leibniz University Hannover, Hannover, Germany

## Abstract

Most patent retrieval systems are text-based, which under utilizes the multimodal nature of patent documents. Although a few multimodal patent retrieval systems exist, they fall short in providing efficient and informative visualizations, and facilitating the interpretation of retrieval results. To address these shortcomings, this paper presents iPATENT, a novel web-based multimodal patent image retrieval system. Unlike previous solutions, iPATENT integrates *state-of-the-art* deep learning models for fine-grained unimodal, cross-modal, and multimodal patent image retrieval. Additionally, it employs both traditional machine learning techniques and modern generative methods for interactive visual exploration and insightful interpretation of retrieval results. iPATENT leverages modern web technologies to provide an interactive interface that enables users to explore large patent databases efficiently and in a visually informative way. Source code and demo are publicly available at: <https://service.tib.eu/ipatent/>.

## Keywords

Multimodal patent retrieval, interactive visualization systems, generative AI

## 1. Introduction

Patents protect intellectual property by combining detailed textual descriptions with visual representations. However, most patent search and analysis tools primarily depend on text-based methods to identify similar patents [1]. This approach can be limiting, as visual elements-such as technical drawings and illustrations-often convey crucial supplementary information that text alone may not fully capture. These visual elements can also bridge linguistic and domain-specific gaps that are prevalent in patent documents. Integrating patent images could enhance efficiency in the patent search process [2], which is especially valuable as the volume of patent applications continue to rise<sup>1</sup>. By leveraging both textual and visual information, search systems can identify a broader range of similar patents, thereby improving recall and supporting more comprehensive decision-making for both applicants and examiners [3].

Several systems have been developed to address the unique challenges of patent image retrieval. Early patent image retrieval systems relied on low-level visual features for content-based image retrieval (e.g., PATMEDIA [3], PATSEEK [4]). However, they struggled with scalability and semantic search. In recent years, vision-language models (VLMs) such as CLIP (Contrastive Language-Image Pre-training, [5]) have enabled semantic search in patent image retrieval [6, 7, 8] by mapping both images and text into a shared embedding space. For example, VisPat<sup>2</sup> leverages CLIP to support semantic and cross-modal retrieval. However, existing systems lack support for fine-grained query formulation, such as *text-conditioned image-to-image* retrieval or *sub-region* (component-level) search. Furthermore, these systems lack mechanisms for organizing the retrieved images into meaningful groups or providing

6th Workshop on Patent Text Mining and Semantic Technologies (PatentSemTech) 2025

\*Corresponding author.

✉ [sushil.awale@tib.eu](mailto:sushil.awale@tib.eu) (S. Awale)

ORCID 0000-0003-2575-0134 (S. Awale); 0000-0002-6802-1241 (E. Müller-Budack); 0000-0003-0918-6297 (R. Ewerth)



© 2025 Copyright for this paper by its authors. Use permitted under Creative Commons License Attribution 4.0 International (CC BY 4.0).

<sup>1</sup>[https://report-archive.epo.org/files/babylon/epo\\_patent\\_index\\_2022\\_infographic\\_en.pdf](https://report-archive.epo.org/files/babylon/epo_patent_index_2022_infographic_en.pdf)

<sup>2</sup><https://service.tib.eu/vispat>

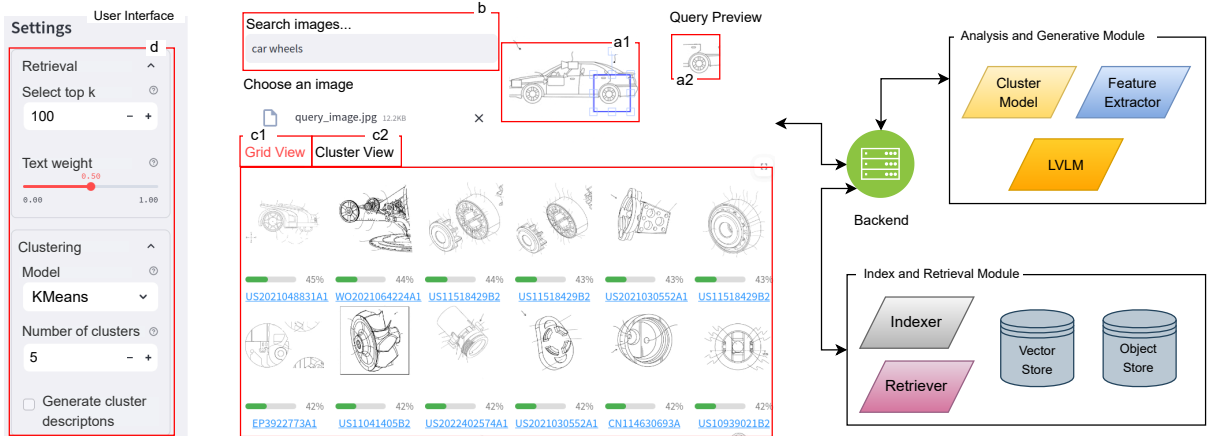
insightful keywords or descriptions to the results, which help in efficient exploration, and interpretation of retrieval results in large databases.

In this paper, we present iPATENT, an open-source, interactive web-based search system for multimodal patent image retrieval and analysis. The system enables semantic patent image search through *image-to-image*, *text-to-image*, and *text-conditioned image-to-image* retrieval approaches. Users can also search for specific sub-regions or components within a query image, supporting fine-grained exploration. Additionally, the system organizes retrieval results into visually and semantically coherent clusters, each accompanied by synthetic descriptions, enhancing exploration and interpretability to facilitate further analysis of the results.

The remainder of this paper is organized as follows. Section 2 describes the system architecture and technical implementation details of iPATENT. The multimodal image retrieval, image clustering, and interpretation approaches are described in Section 3. Section 4 presents practical use cases of iPATENT, and the current functionalities and user interface. In Section 5, we provide quantitative evaluation of CLIP’s performance on patent retrieval task. Section 6 summaries the paper and outlines future work.

## 2. System Architecture of iPATENT

iPATENT facilitates patent image retrieval and analysis through a web-based application system. As illustrated in Figure 1, iPATENT comprises of three primary modules: the user interface and backend module, the index and retrieval module, and the analysis and generative module. The user interface and backend module is implemented using *Streamlit* [9] and *Python*. The backend sub-module orchestrates the data flow among all the other modules. The index and retrieval module implements the indexing and retrieval mechanism using *Qdrant* [10] vector database, and stores the raw image bytes in *MinIO* [11] object store database. The analysis and generative module provides clustering service implemented in *scikit-learn* [12], feature extraction service utilizing a VLM, and generation service using a large vision-language model (LVLM) deployed with *Ollama* [13].



**Figure 1:** System architecture of iPATENT showing the underlying modules and the user interface. The user interface displays the query image of a vehicle (a1) from the patent US2021101616A1, a cropped region depicting the wheels of the vehicle (a2), a conditional text input ‘car wheels’ (b), top-12 retrieved images in a grid view with corresponding patent IDs and similarity scores (c1), and user settings section allowing users to change parameters (d). The cluster view (c2) is shown in Figure 2.

## 3. Multimodal Image Retrieval

iPATENT operates through a sequence of steps that combine deep feature extraction, flexible and fine-grained retrieval, and post-retrieval visualizations. Offline, a large collection of patent images are

processed using a deep learning model to extract high-dimensional feature representations (Section 3.1) which are stored in a vector database. The vector-based retrieval process (Section 3.2) then allows unimodal, cross-modal, and multimodal retrieval of patent images for exploration and discovery. After retrieval, the system further organizes the results through clustering, grouping similar images based on their feature representations (Section 3.3). This organization enhances result navigation and exploration, and together with synthetic descriptions (Section 3.4) enable interpretation and analysis.

### 3.1. Feature Extraction and Indexing

Effective patent image search requires retrieving not only visually similar images but also those that are semantically related. To achieve this, we employ CLIP [5] embeddings, which map both images and text into a shared semantic space through a language-supervised training paradigm. The CLIP model variant employed is ViT-B-16 trained on *laion400m\_e32*, implemented via the *open\_clip* library [14]. CLIP ensures that semantically similar image-text pairs are positioned close together in the embedding space. We encode a large collection of utility patent images using CLIP and index them offline in *Qdrant* vector database along with their metadata such as terms associated with the components depicted in the images.

### 3.2. Multimodal Image Retrieval

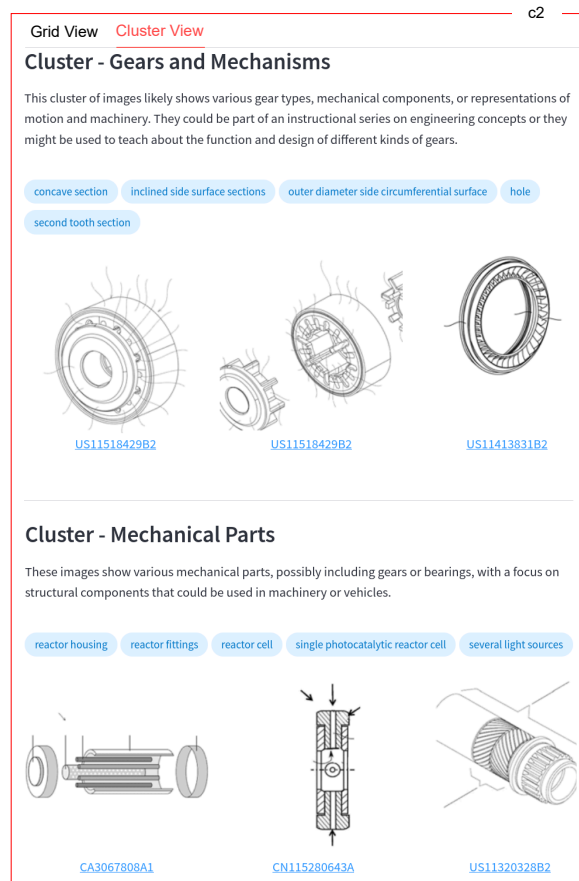
The core image retrieval mechanism in the system is based on a  $k$ -nearest neighbor search within a high-dimensional semantic space, enabling both visual and semantic similarity matching. Central to its efficient and scalable performance is the use of the HNSW (Hierarchical Navigable Small World) [15] algorithm for indexing, which organizes vectors in a multi-layered graph structure. Users can submit queries in the form of images, text, or a combination of both. The query is then passed through the feature extractor (discussed in Section 3.1), which encodes it into the semantic space. The system then performs a vector search with *Qdrant* using cosine distance to efficiently retrieve the most similar images from the index.

iPATENT implements three image retrieval approaches by leveraging the shared semantic space of CLIP: (1) *image-to-image* (unimodal), (2) *text-to-image* (cross-modal), and (3) *text-conditioned image-to-image* (multimodal).

The *image-to-image* retrieval approach allows transcending domain and language barriers during patent search. A user, for example, can upload a technical drawing of a folding chair mechanism to search for images that depict visually and structurally similar folding mechanisms, regardless of differences in drawing style, and without requiring domain knowledge. In this approach, iPATENT takes a single patent image as query, which is then encoded using CLIP, and then performs vector search to retrieve similar images.

The *text-to-image* retrieval approach facilitates searching for relevant images using descriptive language, which bridges the gap between textual and visual information. For example, a user can enter a descriptive textual query such as "portable solar panel with integrated battery storage" encapsulating concepts that sometimes images alone can not. In this cross-modal retrieval approach, iPATENT supports text-based natural language query, which is also embedded into the same semantic space as the images using CLIP. The system, similar to *image-to-image* approach, performs vector search to identify images whose embeddings are most similar to the text query.

The *text-conditioned image-to-image* retrieval is a multimodal approach that leverages the complementary strengths of both modalities, enabling more nuanced and precise queries. For example, users can upload an image of a bicycle frame and refine their search with textual attributes or modifications such as "with built-in suspension system" to find images with additional functional or structural requirements. Here, iPATENT encodes both the query image and the query text into the same semantic space using CLIP, and performs *image-to-image* and *text-to-image* retrieval. The retrieved results are then fused together using weighted averaging, i.e., late fusion. The modality weights are also configurable providing the user further fine-grained query formulation features.



**Figure 2:** Two clusters of the retrieved images in iPATENT along with associated keywords and synthetically generated cluster titles and descriptions for the multimodal query shown in Figure 1.

### 3.3. Image Clustering

Most retrieval systems present results as an ordered list ranked by similarity scores, typically in a list or grid layout. However, such linear layouts often overlook duplicate or near-duplicate results, do not reveal relationships among retrieved items, and often fail to provide meaningful grouping or clustering similar results. Moreover, fixed grid layouts can also become overwhelming when displaying large volume of results, potentially decreasing efficiency in locating specific images of interest.

Displaying retrieved images in visually and semantically coherent cluster groups significantly reduces information overload and enhances users' ability to efficiently navigate and explore the results. To achieve this, we apply the *k-means* [16] clustering algorithm to the CLIP embeddings of the retrieved images. By leveraging CLIP's high-dimensional semantic representations, the clustering process groups together images that share visual and conceptual similarities, making it easier for users to identify patterns and thematic groupings within the search results. Both the number of clusters and the subset of retrieved images to be clustered are user-configurable through the frontend. Figure 2 shows the retrieval results grouped into two clusters paired with keywords and synthetic titles and descriptions.

### 3.4. Cluster Interpretation

Providing rich context for each cluster group—both in relation to its member images and the original query—greatly enhances the interpretability and navigability of retrieval results. To achieve this, each cluster is paired with a set of descriptive keywords as well as synthetic titles and summaries. The keywords consist of component terms associated with the patent images in the cluster, which are extracted directly from the corresponding patent texts. This gives users immediate insight into the technical content of each group.

Beyond keywords, each cluster is further provided with a synthetic title and a description, both generated by an LVLM. For this purpose, we use the LLaVA-1.6 [17] model with Vicuna [18] as the language backbone, quantized to 4-bit precision for efficient inference and resource utilization. To mitigate the significant computational cost associated with LVLM inference, we limit the input to a random selection of 5 images for each cluster. The LVLM then synthesizes human-readable titles and context-aware descriptions that capture the main characteristics and relevance of each cluster. Figure 2 shows a cluster view in iPATENT with two clusters containing three images each, along with synthetic titles and descriptions, and associated keywords.

## 4. Use Cases and Demonstration

iPATENT has a range of impactful use cases across the patent lifecycle. We discuss some use cases of iPATENT in Section 4.1, and highlight some of the user interface and features of iPATENT in Section 4.2.

### 4.1. Use Cases

The multimodal image retrieval capabilities of iPATENT allows for various use cases across the patent lifecycle:

**Prior Art Search:** iPATENT’s multiple search approaches allow for quick and focused searches helping patent examiners and applicants efficiently identify existing inventions that may affect the novelty or non-obviousness of a new patent application.

**Cross-Domain and Multimodal Search:** The cross-modal and multimodal search in iPATENT allow for searching cross-domain where users lack precise technical vocabulary or when inventions are best described visually. This feature is especially helpful in fields where technical drawings and schematics are central to the invention, such as mechanical engineering, and design domains.

**Infringement Detection and Freedom-to-Operate Analysis:** The multimodal search in iPATENT can potentially identify a broader range of relevant patents, enhancing the recall of the search process. This allows companies and legal teams to reduce potential infringement or ensure freedom to operate reducing the risk of litigation post-publication.

### 4.2. Demonstrator

iPATENT is available at <https://service.tib.eu/ipatent/>. Users can perform three distinct types of image retrieval: *image-to-image* retrieval (Figure 1; a), *text-to-image* retrieval (Figure 1; b), and *text-conditioned image-to-image* retrieval (Figure 1; a and b). In *image-to-image* retrieval, users can also crop a specific region of the image, to perform fine-grained searches such as searching for similar sub-figures or components. By adjusting the dimensions of the blue square (shown in Figure 1; a1), the users can acquire a cropped query image (shown in Figure 1; a2).

For each query, the top  $k$  (adjustable by the user) most relevant images are displayed in a grid view (Figure 1; c1). The layout is designed so that the most relevant image appears in the top-left corner, with images arranged from left to right across each row. Relevance decreases progressively rightward along the rows and downward through the grid, making the bottom-right image the least relevant among the displayed results.

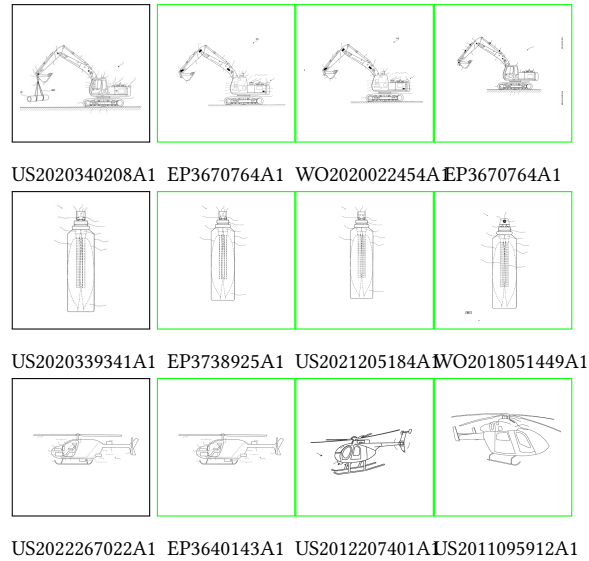
Users are also provided with a cluster view (Figure 2; c2). Here, the retrieved images are organized into groups, each representing a cluster of visually and semantically similar images. Each cluster is displayed as a horizontally scrollable row, which reduces visual information overload by helping users focus on conceptually coherent subsets. The users also have the option to select the number of desired clusters and the number of retrieved images to cluster (Figure 1; d). To further enhance interpretability and efficient navigation, each cluster is also paired with a set of keywords, providing quick insights about the cluster. Beyond the keywords, each cluster is also enriched with synthetic title and description, which are generated using an LVLM.



## 5. Evaluation

For evaluation of the CLIP model [5], we run the *image-to-image* retrieval task on the publicly available benchmark DEEPATENT [19]. This benchmark focuses on image-to-image retrieval task and uses Re-ID (Re-identification) to judge relevance, i.e., two patent images belonging to the same patent are considered to be relevant. The dataset consists of 13,133 patent drawing images as query and more than 38,000 as index images (from 6,927) patents. The CLIP model demonstrates strong performance, achieving an mAP (mean Average Precision) of 0.57 at rank  $k = 1$  without any fine-tuning or domain adaption. This result highlights CLIP’s robust generalization for image retrieval, though further optimization or adaption could yield even higher accuracy [20, 21]. However, relevance based solely on Re-ID does not fully capture the complexity of real-world patent search scenarios. One major challenge is evaluating patent retrieval systems is the lack of reliable, high-quality relevance data, especially for image-based, multimodal and cross-modal search. To circumvent the labor-intensive process of annotating large datasets, pseudo ground-truth labels such as patent citations [22] are used, which are mainly based on text-based retrieval results. As a result, there occurs a misalignment between the ground truth and the multimodal retrieval tasks, often labeling relevant patents as non-relevant or without any relevance judgment.

In addition to the quantitative metrics, we present qualitative results to demonstrate the model’s capability to retrieve visually and semantically similar patent images. In Figure 3, for each query image, the top-three retrieved images exhibit strong visual similarity and consistent semantic content. For example, the query image from patent US2020340208A1 retrieves images of similar machinery with comparable structural details. Similarly, the queries of patents US2020339341A1 and US2022267022A1, yield closely related designs and visualizations. These qualitative examples highlight the effectiveness of CLIP in capturing both fine-grained visual features and relevant semantics, confirming its practical utility in multimodal patent image retrieval.



**Figure 3:** Qualitative examples of *image-to-image* retrieval results using CLIP. Each row corresponds to a different query image (black), accompanied by its top-three retrieved images (green) from the index. The patent IDs are shown below each image.

## 6. Conclusions and Future Work

In this paper, we have introduced iPATENT, a novel web-based platform for patent image retrieval and analysis. Unlike existing systems, iPATENT leverages *state-of-the-art* models for unimodal, cross-modal, and multimodal retrieval of patent images with semantic search capability. The platform also supports

fine-grained searching for sub-figures or components through as interactive query image cropping feature. For organizing search results, iPATENT uses *k-means* clustering, while LVLMS automatically generate informative cluster titles and descriptions. Leveraging modern web technologies, iPATENT provides an interactive and visually informative interface, supporting efficient exploration of large patent databases.

In future work, we aim to enhance iPATENT with additional ranking and interpretation features such as LVLMS-based re-ranking and explanation of ranking results. We plan to improve cluster visualizations with interactive 3D projections for more intuitive exploration. Furthermore, we plan to incorporate faceted search options such as filtering by patent class or figure type. Additionally, LVLMS may introduce biases in the content they generate. A systematic study of the synthetic content and its potential biases is planned as an import direction for future research.

## Acknowledgments

This article has been funded by the Academic Research Programme of the European Patent Office (project "ViP@Scale: Visual and multimodal patent search at scale"). We would like to thank Wolfgang Gritz and Matthias Springstein (both TIB - Leibniz Information Centre for Science and Technology) for their feedback and help in developing iPATENT.

## Declaration on Generative AI

During the preparation of this work, the author(s) used Grammarly in order to: Grammar and spelling check, Paraphrase and reword. After using these tool(s)/service(s), the author(s) reviewed and edited the content as needed and take(s) full responsibility for the publication's content.

## References

- [1] R. Krestel, R. Chikkamath, C. Hewel, J. Risch, A survey on deep learning for patent analysis, *World Patent Information* 65 (2021) 102035. doi:10.1016/j.wpi.2021.102035.
- [2] J. Zellmer, S. Elbeshausen, C. Womser-Hacker, Elicitation of requirements for innovative visual patent retrieval based on interviews with experts, in: *The Information Behaviour Conference, ISIC 2022*, Berlin, Germany, September 26-29, 2022, 2022, p. isic2234. doi:10.47989/irisic2234.
- [3] P. Sidiropoulos, S. Vrochidis, I. Kompatsiaris, Content-based binary image retrieval using the adaptive hierarchical density histogram, *Pattern Recognition* 44 (2011) 739–750. doi:10.1016/J.PATCOG.2010.09.014.
- [4] A. Tiwari, V. Bansal, PATSEEK: content based image retrieval system for patent database, in: *International Conference on Electronic Business, ICEB 2004*, Beijing, China, December 5-9, 2004, Academic Publishers/World Publishing Corporation, 2004, pp. 1167–1171. URL: <https://aisel.aisnet.org/iceb2004/199>.
- [5] A. Radford, J. W. Kim, C. Hallacy, A. Ramesh, G. Goh, S. Agarwal, G. Sastry, A. Askell, P. Mishkin, J. Clark, G. Krueger, I. Sutskever, Learning transferable visual models from natural language supervision, in: *International Conference on Machine Learning, ICML 2021*, Virtual Event, July 18-24, 2021, PMLR, 2021, pp. 8748–8763. URL: <http://proceedings.mlr.press/v139/radford21a.html>.
- [6] K. Pustu-Iren, G. Bruns, R. Ewerth, A multimodal approach for semantic patent image retrieval, in: *Patent Text Mining and Semantic Technologies co-located with the ACM SIGIR Conference on Research and Development in Information Retrieval, PatentSemTech@SIGIR 2021*, Aachen, Germany, July 15, 2021, CEUR-WS.org, 2021, pp. 45–49. URL: <https://ceur-ws.org/Vol-2909/paper6.pdf>.
- [7] H. Lo, J. Chu, J. Hsiang, C. Cho, Large language model informed patent image retrieval, in: *Patent Text Mining and Semantic Technologies co-located with the ACM SIGIR Conference on Research*

- and Development in Information Retrieval, PatentSemTech@SIGIR 2024, Washington D.C., USA, July 28, 2024, CEUR-WS.org, 2024, pp. 51–60. URL: <https://ceur-ws.org/Vol-3775/paper11.pdf>.
- [8] H. H. Shomee, Z. Wang, S. Medya, S. N. Ravi, IMPACT: A Large-scale Integrated Multimodal Patent Analysis and Creation Dataset for Design Patents, in: Conference on Neural Information Processing Systems, NeurIPS 2024, Vancouver, BC, Canada, December 10-15, 2024, Curran Associates, Inc., 2024, pp. 125520–125546. URL: [https://proceedings.neurips.cc/paper\\_files/paper/2024/file/e3301977b92f28e32639ec99eb08f4a1-Paper-Datasets\\_and\\_Benchmarks\\_Track.pdf](https://proceedings.neurips.cc/paper_files/paper/2024/file/e3301977b92f28e32639ec99eb08f4a1-Paper-Datasets_and_Benchmarks_Track.pdf).
  - [9] Streamlit, Streamlit, <https://streamlit.io/>, 2025. Accessed: 2025-04-22.
  - [10] Qdrant, Qdrant documentation, <https://qdrant.tech/documentation/concepts/search/>, 2025. Accessed: 2025-04-22.
  - [11] MinIO, Minio documentation, <https://min.io/docs/minio/linux/index.html>, 2025. Accessed: 2025-04-22.
  - [12] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, E. Duchesnay, Scikit-learn: Machine learning in Python, *Journal of Machine Learning Research* 12 (2011) 2825–2830.
  - [13] Ollama, Ollama documentation, <https://docs.llamaindex.ai/en/stable/examples/llm/ollama/>, 2025. Accessed: 2025-04-22.
  - [14] G. Ilharco, M. Wortsman, R. Wightman, C. Gordon, N. Carlini, R. Taori, A. Dave, V. Shankar, H. Namkoong, J. Miller, H. Hajishirzi, A. Farhadi, L. Schmidt, Openclip, 2021. doi:10.5281/zenodo.5143773.
  - [15] Y. A. Malkov, D. A. Yashunin, Efficient and robust approximate nearest neighbor search using hierarchical navigable small world graphs, *IEEE Transactions on Pattern Analysis and Machine Intelligence* 42 (2020) 824–836. doi:10.1109/TPAMI.2018.2889473.
  - [16] S. P. Lloyd, Least squares quantization in pcm, *IEEE Transactions on Information Theory* 28 (1982) 129–137. doi:10.1109/TIT.1982.1056489.
  - [17] H. Liu, C. Li, Q. Wu, Y. J. Lee, Visual instruction tuning, in: Conference on Neural Information Processing Systems, NeurIPS 2023, New Orleans, LA, USA, December 10-16, 2023, Curran Associates, Inc., 2023, pp. 34892–34916. URL: [https://proceedings.neurips.cc/paper\\_files/paper/2023/file/6dcf277ea32ce3288914faf369fe6de0-Paper-Conference.pdf](https://proceedings.neurips.cc/paper_files/paper/2023/file/6dcf277ea32ce3288914faf369fe6de0-Paper-Conference.pdf).
  - [18] W.-L. Chiang, Z. Li, Z. Lin, Y. Sheng, Z. Wu, H. Zhang, L. Zheng, S. Zhuang, Y. Zhuang, J. E. Gonzalez, I. Stoica, E. P. Xing, Vicuna: An Open-Source Chatbot Impressing GPT-4 with 90%\* ChatGPT Quality, 2023. URL: <https://lmsys.org/blog/2023-03-30-vicuna/>.
  - [19] M. Kucer, D. Oyen, J. Castorena, J. Wu, Deeppatent: Large scale patent drawing recognition and retrieval, in: IEEE/CVF Winter Conference on Applications of Computer Vision, WACV 2022, Waikoloa, HI, USA, January 3-8, 2022, IEEE, 2022, pp. 557–566. doi:10.1109/WACV51458.2022.00063.
  - [20] K. Higuchi, K. Yanai, Patent image retrieval using transformer-based deep metric learning, *World Patent Information* 74 (2023) 102217. doi:<https://doi.org/10.1016/j.wpi.2023.102217>.
  - [21] Y.-H. Lin, M.-C. Hung, C.-F. Lee, Density-refine: Patent image retrieval by density-based region extraction and feature fusion, *Journal of Mechanical Design* 147 (2025) 081703. doi:10.1115/1.4067749.
  - [22] F. Piroi, M. Lupu, A. Hanbury, V. Zenz, CLEF-IP 2011: Retrieval in the intellectual property domain, in: CLEF-IP Workshop co-located with the Conference and Labs of the Evaluation Forum, CLEF-IP@CLEF 2011, Amsterdam, The Netherlands, September 19-22, 2011, CEUR-WS.org, 2011. URL: <https://ceur-ws.org/Vol-1177/CLEF2011wn-CLEF-IP-PiroiEt2011.pdf>.