SEMANTIC SEARCH OF RELEVANT IMAGES USING VECTOR DATABASES

Ostapchenko Nazar
Bachelor student
Tyshchenko Olena
Teacher
Denysenko Bohdan
Bachelor student
Mazurets Oleksandr
Ph.D in Engineering Science, Associate Professor
Computer Science Department
Khmelnytskyi National University, Ukraine

Image search is a fundamental element of the modern world, and an essential component of the Internet, providing fast and intuitive access to a vast space of visual content. Thanks to the advancement of artificial intelligence and machine learning technologies, image search continues to evolve, becoming increasingly sophisticated and capable of meeting user needs in increasingly accurate and reliable ways [1].

Not so long ago, the easiest way to search for a relevant image was to use a regular search engine and enter a text query, and add the word "image", but thanks to the general evolution of technology, there are now more efficient ways to search for images, i.e. using appropriate image search engines or using reverse search function.

Now a user can use image search, such as Google, and get search results that will display related images, products, websites that use similar images, as well as results with content related to the image using image classification technology. Visual search features and visual search APIs are being added wherever there is a search feature: Google, Pinterest, Instagram, Amazon, and Microsoft have all implemented visual search features. Visual search is already being used in e-commerce, social media, websites, apps, and even in real life. Planning and optimizing for visual search is more important than ever. It won’t replace text search, but combination of text, voice, and visual search all serve one purpose: to make user’s life easier and more convenient. The use of images in search will continue to grow, and more importantly, new technologies for search will continue to emerge.

The authors’ investigations reveal that modern vector database systems, when combined with powerful embedding models such as transformer-based encoders [2] and convolutional neural network-derived feature extractors [3, 4], excel at automatically retrieving semantically relevant images from vast multimedia archives. By projecting both query inputs and image collections into a shared high-dimensional vector space, these methods enable rapid nearest-neighbor searches that accurately capture visual and contextual similarity, far surpassing traditional keyword-based retrieval in both precision and recall [5, 6]. When deployed in applications ranging from digital asset management to real-time surveillance feed analysis, automated vector search significantly reduces the time and labor required to sift through massive image repositories [7], a capability of paramount importance in domains where swift access to pertinent visual information can drive critical decision-making [8].

The purpose of the work is designing of method for automated selection of relevant images using vector databases.

The method of automated selection of relevant images using vector databases allows, using existing data and semantic search, to find images relevant to the user's query. The method of automated selection of relevant images using vector databases works by converting input data in the form of a text query and documents using vectorization tools into vector representations, which are stored and analyzed to determine semantic similarity and find relevant results.

The step 1 is preprocessing the document collection. It involves vectorizing each document – that is, converting text data into numeric data. At the same stage, vector representations are saved in a vector database that supports fast similarity search.

The step 2 is to prepare the search query. The user's query goes through the same vectorization stages as documents: cleaning, transformation into a vector form using the same model that was used for indexing.

In step 3, the nearest neighbors are searched for – analyzing the similarity between the query vector and the vectors stored in the vector database. For this purpose, Euclidean distance, cosine similarity or other appropriate metric is used [9]. The system finds documents that have the highest level of similarity to the query.

In step 4, semantic search occurs – processing of the results of similarity analysis. The selected documents are displayed in the form of relevant results that are as similar as possible to the query at the content level.

The last step is the formation of the final result, which is returned to the user. This result includes the closest images or objects in terms of content that were found in the vector database based on semantic similarity to the query. The output data is a list of results ordered by similarity that are relevant to the user's query, taking into account the semantic context.

The application created based on the method of automated selection of relevant images based on semantic search using vector databases, search by image by software for selection of relevant images using vector databases is shown in Figure 1. It has the following functionality: image search by image context, image search by text, saving the image to the database. The user can search by text, is shown in Figure2.

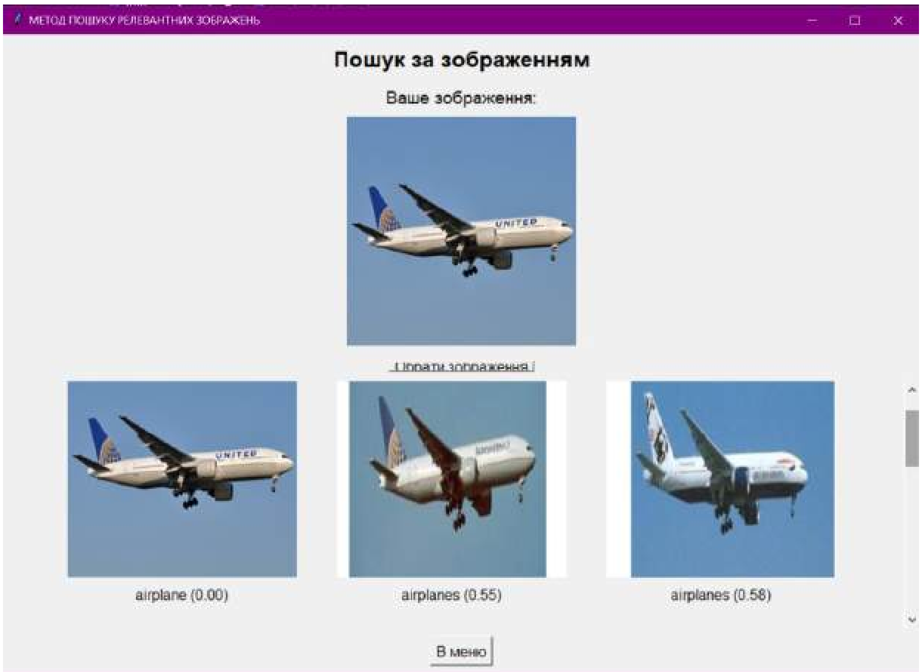

Figure 1. Search by image by software for selection of relevant images using vector databases

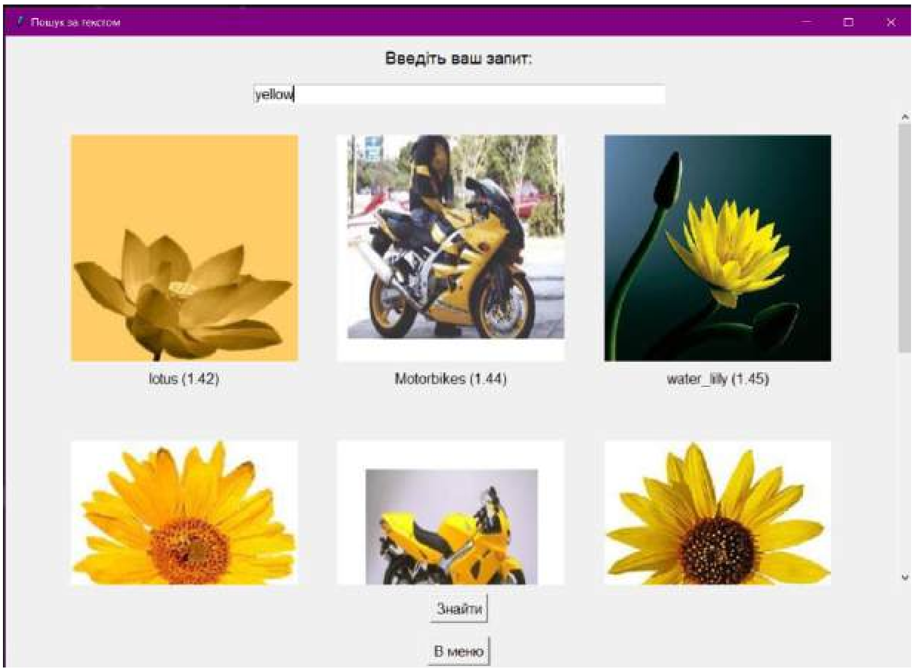

Figure 2. Search by text by software for selection of relevant images using vector databases

To do search by text, click the text search button and enter any search query, for example, "yellow". After starting the search, the system sends this text to the OpenCLIP model, which converts the text into a vector. Then the text vector is compared with all image vectors in the database. Those images whose vectors are closest to the text vector are returned to the user as results.

When the user wants to find similar images based on another image, he opens a search window. In this window, you can select an image from your own device. After selecting an image, the system displays it on the screen, and processing begins in the background. The selected file is transferred to the OpenCLIP model, which creates a vector from it. This vector is a mathematical representation of the main features of the image. Next, a search is performed in the vector database, where all available images already have their own vectors. The system measures the distances between the vector of the selected image and all other vectors in the database. The smaller this distance, the more similar the images are considered.

So, the stages of the semantic search method using data vectorization and vector comparison were described. The method allows you to find relevant images even in cases where the query formulation differs from the document text, thanks to the analysis of the content of connections in a multidimensional vector space.

References

1. AIContentfy. Content analysis is becoming more and more relevant. URL: https://aicontentfy.com/en/blog/future-of-content-analysis-with-ai.
2. Mazurets, O., Zalutska, O., Tyschenko, O., & Bohdanova, A. (2024). An approach to using MobileNet CNN-model for gesture recognition. Proceedings of XXIII International Scientific and Practical Conference "Problems of Science and Technology: the Search for Innovative Solutions", 59–64.
3. Kharysh, I., Sobko, O., & Mazurets, O. (2024). Designing CNN neural network model for detecting fractures of lower extremities by X-ray images. The Impact of Scientific Research on the Development of the Modern World. Proceedings of the XLIV International Scientific and Practical Conference. Dubrovnik, Croatia, 91–96.
4. Zharnovskyi, O., Mazurets, O., & Sobko, O. (2024). Approach to identification of artificial intelligence-generated people images by means of machine learning. Key Aspects of the Development of Scientific Research in Modern Conditions. Proceedings of the XLV International Scientific and Practical Conference, 69–73.
5. Mazurets, O. V., Klimenko, V. I., Molchanova, M. O., & Sultanov, A. V. (2024). Object-oriented intelligent system for neural network detection of sugar crystallization zones. Global Science: Prospects and Innovations. Proceedings of the 10th International Scientific and Practical Conference. Cognum Publishing House, Liverpool, United Kingdom, 198–207.
6. Zalutska, O. O., Hladun, O. V., & Mazurets, O. V. (2025). Method of preventing failures of rotating machines by vibration analysis using machine learning techniques. Radio Electronics, Computer Science, Control, 1, 142–152.
7. Hladun, O., Mazurets, O., Molchanova, M., & Sobko, O. (2024). Real time detection the person emotion state using neural network. Scientific Research: Modern Innovations and Future Perspectives. Montreal, Canada, 119–123.
8. Novak Y., & Mazurets O. (2023). Practical Application of Method of Automated Personal Identification by Fingerprints Using Convolution Neural Networks. Proceedings of V International Scientific and Practical Conference "Modern strategies of global scientific solutions". Stockholm, Sweden, 136–140.

9. Sobko, O. V., Dydo, R. A., & Mazurets, O. V. (2025). Method of neural network detection of defects based on the analysis of rotating machines vibrations. Radio Electronics, Computer Science, Control, 1, 106–115.

ВИЯВЛЕННЯ ДЕФЕКТІВ НА ЗОБРАЖЕННЯХ АВТОМОБІЛЬНИХ ДОРІГ З ВИКОРИСТАННЯМ НЕЙРОННИХ МЕРЕЖ

Величко Віталій
здобувач вищої освіти магістерського рівня
Науковий керівник:
Мельник Р.А.
проф.
Кафедра програмного забезпечення
Національний університет «Львівська Політехніка»

Анотація. Стаття присвячена дослідженню ефективності комбінованого підходу до виявлення дефектів дорожнього покриття, що поєднує сегментацію зображення з подальшим виявленням об’єктів за допомогою нейронної мережі YOLOv8. Проведено порівняння з класичним методом прямої детекції на повному зображенні, виконано оцінку точності та швидкодії обох підходів. Встановлено, що комбінований підхід не забезпечує стабільного покращення результатів і поступається класичному за швидкодією.

Ключові слова: дефекти дорожнього покриття, сегментація зображення, виявлення об’єктів, нейронні мережі.

Введення. Стан дорожнього покриття впливає на безпеку та комфорт учасників дорожнього руху, тому завдання виявлення дефектів є актуальним. Одним із перспективних напрямів його вирішення є застосування методів комп’ютерного зору на основі згорткових нейронних мереж, зокрема YOLO. Попри високу ефективність існуючих підходів, існує припущення, що точність виявлення можна підвищити шляхом обмеження зони аналізу лише ділянкою дорожнього полотна. Як показано в статті[1], сегментація дорожнього полотна перед виконанням детекції сприяє зниженню кількості хибнопозитивних спрацювань і покращенню точності аналізу. Такий ефект можна досягти за допомогою комбінованих підходів, які включають попередню сегментацію зображення для виділення області дороги. Залишається відкритим питання, чи забезпечує такий підхід покращення результатів та наскільки він виправданий, зважаючи на додаткові обчислювальні витрати.

Мета та завдання дослідження. Метою дослідження є оцінка ефективності комбінованого підходу до виявлення дефектів дорожнього покриття, що поєднує сегментацію зображення з подальшою детекцією об'єктів. У рамках роботи