

Концепція застосунку, що забезпечує ефективний пошук картин за текстовим описом.

Владислав Концевий, д-р філ., асистент¹ (ORCID:0000-0002-6636-5489), Олександра Юхимович, студент¹ (ORCID: 0009-0003-1024-2118), Євгенія Крутікова, студент¹ (ORCID:0009-0007-4871-2559).

¹ Київський Національний Університет Будівництва і архітектури, Україна

АНОТАЦІЯ

У статті розглядається концепція та архітектура інтелектуальної системи пошуку мистецьких творів за текстовим описом або зображенням. Основна ідея полягає у створенні інструменту, що дає змогу користувачам з мінімальними знаннями швидко знаходити потрібні картини навіть за неповним чи нечітким описом. Запропонована архітектура включає модулі індексації текстів та зображень, пошук за векторними представленнями, фільтрацію результатів за стилем, періодом та автором, а також систему ранжування для підвищення релевантності. Система орієнтована на широке коло користувачів і може бути масштабована з додаванням багатомовної підтримки та рекомендаційних модулів.

Ключові слова: живопис, картини, штучний інтелект, нейронні мережі, концепт застосунку, CNN.

1. ВСТУП

У сучасному інформаційному просторі користувачі дедалі частіше стикаються з проблемою пошуку художніх творів, зокрема живопису, за неповною інформацією. Типова ситуація полягає в тому, що глядач пам'ятає лише загальне враження від картини або окремі деталі композиції, однак назва та автор залишаються невідомими. Традиційні пошукові системи, орієнтовані на ключові слова, у такому випадку виявляються малоефективними: вони повертають велику кількість результатів, серед яких важко віднайти саме потрібний твір. Це створює суттєві перешкоди як для пересічних поціновувачів мистецтва, так і для дослідників, викладачів чи студентів.

Запропонована в даній роботі концепція полягає у створенні застосунку з використанням ШІ, здатного здійснювати пошук картин за їхнім описом у текстовому форматі. На відміну від класичних пошукових механізмів, така система повинна аналізувати навіть мінімальний обсяг вхідної інформації, пропонуючи користувачеві декілька релевантних варіантів із можливістю подальшої фільтрації за стилем, історичним періодом або школою живопису. Очікується, що це дозволить не лише ідентифікувати конкретний твір, а й розширити коло знань користувача шляхом ознайомлення з подібними зразками.

2. МЕТА ТА ЗАВДАННЯ ДОСЛІДЖЕННЯ

Метою роботи є концептуальна розробка прототипу застосунку, що забезпечує ефективний пошук картин за текстовим описом. Основними завданнями є формування зручного інтерфейсу для введення опису, розробка механізму фільтрації та ранжування результатів, а також інтеграція модулю для пошуку за зображенням. Особлива увага приділяється адаптивності системи: вона повинна коректно обробляти як короткі, так і розгорнуті текстові запити, без необхідності у користувача володіти спеціальними знаннями у сфері мистецтва.

3. ПОПЕРЕДНЄ ДОСЛІДЖЕННЯ

Для оцінки актуальності проблеми було проведено експериментальне опитування серед 50 користувачів різних вікових груп. Учасникам пропонувалося пригадати будь-яку картину, яка справила на них враження, і спробувати знайти її за допомогою стандартного пошуку Google. Результати

показали, що у 62% випадків пошук займав понад 10 хвилин, а у 28% випадків картина взагалі не була знайдена. Це свідчить про необхідність створення інструменту, здатного ефективніше інтерпретувати неповні та описові запити.

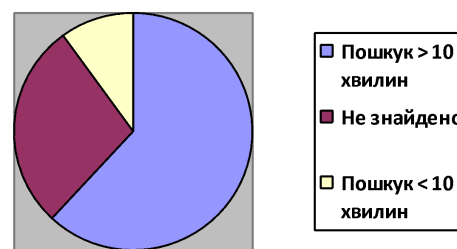


Рисунок 1. Результати опитування

Дослідження [1] представляє модель CLIP, яка поєднує текстові та візуальні ембедінги. Ця робота є фундаментальною для побудови системи пошуку картин за описом, адже дозволяє узгоджувати текстові та зображувальні простори. Стаття [2] описує Faiss — ефективну бібліотеку для пошуку найближчих сусідів у векторному просторі. Це безпосередньо пов'язано з оптимізацією швидкості пошуку у великих базах картин. Робота [3] вводить архітектуру ResNet, яка стала основою для багатьох сучасних CNN-моделей. Для нашої системи це джерело важливе як приклад надійної архітектури для створення візуальних ембедінгів.

4. ВИКЛАД ОСНОВНОГО МАТЕРІАЛУ

У рамках даної роботи передбачається створення концепції застосунку з базовою можливістю пошуку за текстом, додавання фільтрів, системи рекомендацій, багатомовної підтримки та комбінованого пошуку за текстом і зображенням. Реалізація такого підходу забезпечить швидкий і зручний доступ до інформації про художні твори, що має практичне значення для широкого кола користувачів. На рівні клієнта реалізується введення опису користувачем, відображення результатів, використання фільтрів за художником, періодом чи стилем, а також можливість завантаження зображення для пошуку

за візуальними ознаками. Концептуальний дизайн зображений на рисунку 2. Бекенд відповідає за обробку запитів, індексацію, доступ до бази даних і модулі обробки зображень. Для зберігання використовується реляційна база, у якій фіксуються ключові метадані: назва твору, автор, рік створення, опис, URL зображення та теги.



Рисунок 2. Концептуальний дизайн інтерфейсу користувача

Функціональність системи забезпечується низкою ключових модулів. Індикатор відповідає за нормалізацію текстових описів шляхом видалення пунктуації та приведення тексту до єдиного реєстру, створення метаданих та формування векторних представлень. Для кожного запису обчислюється текстовий та візуальний ембедінг, які зберігаються у векторній базі даних. Модуль пошуку за зображенням використовує завантажене користувачем фото, яке нейронна мережа типу CNN або модель класу CLIP перетворює у векторне представлення. Пошук виконується за принципом найближчих сусідів у векторному просторі, після чого система повертає найбільш релевантні результати. Користувацький інтерфейс побудований із застосуванням сучасних UI-компонентів і забезпечує підтримку текстових підказок, автодоповнення, фільтрації та сортування результатів за релевантністю, періодом чи популярністю. Результати відображаються у вигляді карток із мініатюрами картин, зазначенням автора, року та короткого опису з підсвіченими ключовими словами. Для детального ознайомлення передбачене модальне вікно з повним описом та можливістю збереження або поширення інформації.

Алгоритм пошуку в системі реалізуватиметься покроково. Спочатку користувач вводить текстовий опис, на рівні фронтенду здійснюється попередня обробка запиту та передає його на бекенд. У випадку семантичного пошуку бекенд формує векторне представлення опису і виконує пошук найближчих сусідів у векторній базі даних. Альтернативно, може застосовуватись класичний метод TF-IDF. Після цього результати фільтруються та сортуються відповідно до заданих користувачем параметрів. Кінцевим кроком є формування ранжованого списку картин, який повертається до фронтенду для візуалізації.

Особлива увага приділяється аспектам зручності використання. Інтерфейс передбачає відображення кількості знайдених результатів і надання підказок щодо уточнення запиту, наприклад через додавання кольорових характеристик, опис об'єктів чи вказівку настрою. У

текстових результатах застосовується підсвічування знайдених ключових токенів. Система також пропонує варіанти уточнення запиту у форматі «Можливо ви мали на увазі...», використовуючи алгоритми нечіткого співпадіння та аналіз поширених помилок.

Якість роботи системи оцінюється на основі експериментального тестування. Для кількісної оцінки застосовуються метрики Precision, Recall та Mean Reciprocal Rank. Для підвищення продуктивності застосовується CDN для розповсюдження зображень і кешування результатів пошуку. Горизонтальне масштабування векторної бази забезпечує можливість роботи системи з великими наборами даних.

Розвиток проекту планується поетапно. На першому місяці створюється мінімальний прототип із реалізацією базового пошуку за допомогою TF-IDF. Протягом другого місяця додається бекенд, база метаданих і простий індикатор. Третій етап передбачає інтеграцію векторних ембедінгів і покращення користувацького досвіду, а на четвертому — розширення датасету, підтримку пошуку за фотографіями, багатомовність та систему рекомендацій.

Окремим аспектом є безпека та конфіденційність даних. Система передбачає політику обмеженого часу зберігання завантажених фото та можливість їхнього видалення користувачем. При необхідності передбачається шифрування зображень у сховищі. Усі процеси обробки персональних даних повинні відповідати вимогам GDPR та локального законодавства.

5. ВИСНОВКИ

Запропонована система пошуку мистецьких творів за текстовим описом має потенціал стати інструментом у сфері цифрової гуманітаристики та культурної спадщини. Концепція зв'язує демонструє можливість знаходження картин навіть за мінімальними описами, а інтеграція сучасних моделей ембедінгів дозволяє значно підвищити точність результатів порівняно з класичними методами. Подальший розвиток системи передбачає удосконалення алгоритмів ранжування, розширення бази даних та інтеграцію мультимодальних пошукових стратегій. Це відкриває перспективи для створення потужної платформи, здатної не лише допомогти користувачам знаходити картини, а й стимулювати їх інтерес до вивчення мистецтва.

Список літератури

- [1] Radford A., Kim J. W., Hallacy C., Ramesh A., Goh G., Agarwal S., ... Sutskever I. Learning Transferable Visual Models From Natural Language Supervision // *Proceedings of the 38th International Conference on Machine Learning (ICML)*. – 2021. – P. 8748–8763.
- [2] Johnson J., Douze M., Jégou H. Billion-scale similarity search with GPUs // *IEEE Transactions on Big Data*. – 2019. – Vol. 7, No. 3. – P. 535–547.
- [3] He K., Zhang X., Ren S., Sun J. Deep Residual Learning for Image Recognition // *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. – 2016. – P. 770–778.