

НЕОБХІДНІСТЬ РОЗРОБКИ МОБІЛЬНОГО ДОДАТКУ ДЛЯ РОЗПІЗНАВАННЯ РУКОПИСНИХ ДІАГРАМ ТА СХЕМ З ВИКОРИСТАННЯМ ТЕХНОЛОГІЙ КОМП'ЮТЕРНОГО ЗОРУ

Для покращення сприйняття інформації та пояснення різних процесів та процедур використовують схематичні зображення, що складаються з геометричних фігур, коротких текстових написів та інших символів, наприклад, таких як стрілки з різними закінченнями. До прикладів таких схематичних зображень можна віднести блок-схеми, діаграми класів, мапи, макети інтерфейсів та їхні окремих елементів. Вони дозволяють наочно пояснювати взаємодію між різними елементами, послідовності дій та зображати прототипи майбутніх інтерфейсів. Зазвичай, для створення згаданих вище схематичних зображень використовують програмні продукти, що розраховані на роботу з персональним комп'ютером. Використовуючи клавіатуру та мишу, можна доволі швидко створити потрібну діаграму чи макет без великих зусиль. Проте, коли під рукою лише мобільний пристрій з відносно невеликим розміром екрану з сенсорним керуванням, цей процес стає доволі кропітким. На мобільному пристрої важко проводити маніпуляції з фігурами, стрілками та написами через невеликі розміри екрану.

Тому для створення візуальних схем на мобільних пристроях можна розробити спеціальний додаток. Він міститиме різний функціонал для створення та редагування різноманітних схем та макетів. Він дозволить експортувати створені візуальні елементи в різноманітних форматах для подальшого використання або покращення. Також він дозволить ділитися створеними візуалізаціями з іншими за допомогою популярних месенджерів, соціальних мереж та інших популярних способів комунікації.

Для вирішення проблеми незручності створення візуальних схем з мобільних пристроїв, пропонується спростити процес додавання елементів на робоче полотно шляхом сканування знімку рукописної схеми. Користувач матиме змогу зробити замальовку схеми на аркуші паперу або на дошці, зробити знімок за допомогою камери мобільного пристрою і отримати схему в цифровому форматі, доступному для редагування. На папері простіше зображати такого роду схеми, так як є відчуття простору і фізичного контролю. Реалізувати процес перетворення рукописної схеми в цифровий вигляд можна за допомогою технологій комп'ютерного зору.

Комп'ютерний зір (Computer Vision) – це технологія, яка дозволяє створювати цифрові системи, які здатні сприймати візуальну інформацію як людська система зору. Комп'ютерний зір можна розглядати з двох поглядів. З погляду біології, комп'ютерний зір спрямований на створення обчислювальних моделей зорової системи людини. З інженерного погляду, комп'ютерний зір спрямований на створення автономних систем для виконання завдань, які може виконувати зорова система людини [1].

Для реалізації комп'ютерного зору використовуються такі технології як глибоке навчання (deep learning) та згорткові нейронні мережі (Convolutional Neural Network або CNN).

Глибоке навчання – це техніка машинного навчання, яка вчить комп'ютери робити те, що є природним для людини: вчитися на прикладі. В глибокому навчанні комп'ютерна модель вчиться виконувати класифікаційні завдання напряму з зображень, тексту або звуку. Моделі глибокого навчання можуть бути дуже високоточними і навіть перевершувати ефективність людини в виконанні таких завдань. Моделі натреновані з використанням великих наборів даних і архітектури нейронних мереж з багатьма шарами.

Більшість глибоких методів навчання використовують архітектуру нейронних мереж, які часто називають глибокими нейронними мережами. На відміну від традиційних нейронних мереж, які зазвичай містять 2-3 прихованих шари, глибокі нейронні мережі можуть містити до 150 прихованих шарів [2].

Для реалізації комп'ютерного зору застосовують згорткові нейронні мережі. Вони містять згорткові шари, які застосовують до входу операцію згортки (математична операція над двома функціями, результат якої третя функція), передають результат до наступного шару і мережа робить прогнози щодо того, що вона «бачить». Нейронна мережа запускає згортки та перевіряє точність своїх прогнозів у серії ітерацій, доки прогнози не почнуть збуватися.

Комп'ютерний зір вирішує наступні завдання:

- *класифікація* – визначає, що саме знаходиться на певному зображенні (людина, яблуко, собака);
- *виявлення об'єктів* – використовує класифікацію для ідентифікації певного класу зображень, а потім відстежує їхню появу на зображенні чи відео;
- *відстеження об'єкта* – стежить за об'єктом після його виявлення. Це завдання часто виконується за допомогою послідовних зображень або відео в реальному часі;
- *пошук зображень на основі вмісту* – використовує комп'ютерний зір для перегляду, пошуку та отримання зображень із великих сховищ даних на основі вмісту зображень, а не пов'язаних із ними тегів метаданих [3].

Комп'ютерний зір застосовується в різноманітних цілях. Зокрема для розпізнавання та перекладу текстів в додатку Google Translate. Користувачу варто лише навести камеру на текст, який потрібно перекласти і комп'ютерний зір розпізнає його та відправляє на переклад. Також комп'ютерний зір активно застосовується при розробці автопілотів для різноманітних транспортних засобів. З його допомогою програма керування

автомобілем здатна розрізняти дорожні знаки, пішоходів, розмітку та інших учасників дорожнього руху. Комп'ютерний зір полегшує відбір необхідних зображень чи відео, наприклад, вирізняючи важливі моменти в спортивних трансляціях.

У підсумку, незважаючи на те, що галузь комп'ютерного зору має велику кількість напрацювань, в ній залишається ще велика кількість питань, які слід дослідити. Для того аби вирішити проблему розпізнавання рукописних візуальних схем, слід правильно підібрати модель навчання та коректний набір даних.

Список використаних джерел

1. Huang, T. (1999) Computer vision, CERN Document Server. CERN. Available at: <https://cds.cern.ch/record/400313> (Accessed: November 6, 2022).
2. What is deep learning?: How it works, techniques & applications (no date) How It Works, Techniques & Applications - MATLAB & Simulink. Available at: <https://www.mathworks.com/discovery/deep-learning.html> (Accessed: November 6, 2022).
3. What is Computer Vision? (no date) IBM. Available at: <https://www.ibm.com/topics/computer-vision> (Accessed: November 6, 2022).