

Таким чином була запропонована та описана архітектура, яка дозволить ефективно передавати, обробляти зображення та зберігати тексти у локальній мережі.

Було коротко описано систему OCR та запропоновано для неї архітектурну реалізацію збереження розпізнаних текстів, що зменшить навантаження на пристрій, який виконує задачу розпізнання тексту. Також, якщо використовувати БД в локальній мережі, рівень захисту текстів буде вище. Для цієї БД була запропонована реляційна структура бази даних, яка передбачає користувачів, кожен з яких буде мати доступ до лише розпізнаних ним текстів.

Список використаних джерел:

1. Trier Ø. D., Jain A. K., Taxt T. Feature extraction methods for character recognition – a survey. *Pattern Recognition*. 1996. Vol. 29, № 4. P. 641-622. DOI: [https://doi.org/10.1016/0031-3203\(95\)00118-2](https://doi.org/10.1016/0031-3203(95)00118-2)
2. Pradeep J., Srinivasan E., Himavathi S. Diagonal Based Feature Extraction for Handwritten Alphabets Recognition System Using Neural Network. *International Journal of Computer Science & Information Technology*. 2011. Vol. 3, № 1. P. 27-38. DOI: 10.5121/ijcsit.2011.3103
3. Stones R., Matthew N. *Beginning Databases with PostgreSQL: From Novice to Professional*. New-York: Apress, 2005. 665 p.

Усенко В.В.

студент,

Національний технічний університет України

«Київський політехнічний інститут імені Ігоря Сікорського»

АНАЛІЗ БІБЛОТЕК КОМП'ЮТЕРНОГО ЗОРУ ДЛЯ РОЗРОБКИ ПІДСИСТЕМИ ІДЕНТИФІКАЦІЇ РУХОМИХ ОБ'ЄКТІВ

У сучасному світі потреба в автоматичній ідентифікації об'єктів зростає щорічно. Поступове поліпшення якості зображення та розвиток відеотехнологій дозволяють системам відеоспостереження не тільки

детектувати певний об'єкт, а й надавати можливість ідентифікувати його унікальні характеристики.

Для виявлення та розпізнавання об'єктів на зображенні застосовуються бібліотеки комп'ютерного зору. Власне комп'ютерний зір – це наукова дисципліна, яка займається розробкою та дослідженням методів та алгоритмів обробки цифрових зображень, які можуть виявляти, відслідковувати, а також класифікувати об'єкти і подавати результати в тій формі, зрозумілій для сприймання людиною.

Одним з практичних прикладів використання комп'ютерного зору є підсистема ідентифікації рухомих об'єктів, яка може бути інтегрована з системою контролю надання доступу (СКНД) на певному підприємстві чи режимному об'єкті для контролю доступу на певну територію та для автоматизації прохідного режиму. Основними задачами підсистеми ідентифікації рухомих об'єктів є витяг зображення із відеокамери на сервер, детекція та розпізнавання рухомих об'єктів (в СКНД рухомими об'єктами є зазвичай люди), присутніх на відеофрагментах, збереження результатів аналізу відеофрагментів на сервері для майбутньої обробки користувачем.

Для якісної реалізації підсистеми вкрай важливо провести аналіз доступних бібліотек комп'ютерного зору, які можуть проводити аналіз зображення, а саме детектувати та розпізнавати об'єкти.

VXL

VXL – це бібліотека, що написана на мові програмування C++ і може бути використана на багатьох операційних системах. Основними компонентами бібліотеки є

- *vnl*: Алгоритми для роботи з матрицями, векторами та оптимізатори

- *vil*: Завантаження, збереження та маніпулювання зображеннями в різних форматах.

- *vgl*: Геометрія для точок, кривих та інших геометричних об'єктів в одно-, дво- або трьохвимірному просторі

- *vsl*, *vbl*, *vul*: Різні кросплатформені компоненти для роботи з бінарними потоками та для роботи з файлами

Також існують компоненти, які містять алгоритми та методи відеоманіпуляції, обробки зображень, ймовірнісне моделювання, дизайн графічного інтерфейсу (GUI) та інші алгоритми [1].

Одним з плюсів бібліотеки VXL є незалежність компонентів, тобто кожен з вищезазначених компонентів не пов'язаний один з одним. Також VXL працює з зображеннями високої роздільності.

OpenCV

OpenCV – це бібліотека комп'ютерного зору з відкритим кодом. Бібліотека написана на C та C++ і працює під Linux, Windows та Mac OS X. Також дана бібліотека доступна на інтерфейсах для Python, Java, C#, Ruby, та на інших мовах. OpenCV була розроблена для високої ефективності обробки зображень та з фокусом на настільні додатки.

Однією з цілей OpenCV є створення простої у користуванні інфраструктури комп'ютерного зору, що допомагає людям швидко створювати досить складні програми для обробки зображень та відео. Бібліотека містить понад 2500 функцій, що охоплюють багато областей комп'ютерного зору [2].

Також OpenCV має власний модуль для глибинного навчання DNN (Deep Neural Networks), який може бути використаний для розпізнавання об'єктів на зображенні чи відеофрагменті.

Бібліотека OpenCV розділена на п'ять основних компонентів:

- *MLL* – бібліотека, яка містить алгоритми машинного навчання
- *HighGUI* – функції зберігання та завантаження відео та зображень
- *CvAux* – старі або експериментальні алгоритми обробки зображення
- *CV* – основні алгоритми обробки зображення та комп'ютерного зору
- *CXCORE* – базові структури даних та алгоритми, функції малювання та маніпулювання зображенням [3].

Головною особливістю бібліотеки OpenCV є наявність натренованих класифікаторів для виявлення обличчя людей. OpenCV надає можливість створити свої власні класифікатори для детекції різних об'єктів. Також за допомогою DNN, як вже було згадано, можна імплементувати ідентифікацію об'єктів. Варто зазначити, характерною рисою даної бібліотеки є наявність великої спільноти розробників та наявність російсько- та україномовної документації.

Не зважаючи на переваги, комплексна структура OpenCV може заплутати розробника та сповільнити розробку додатку

AForge.NET

AForge.NET є бібліотекою з відкритим вихідним кодом, створеною на мові C#, яка призначена для розробників і дослідників в області комп'ютерного зору та має широкий спектр можливостей та засобів для розробників та науковців: функції для обробки зображень, модуль для використання нейронних мереж, наявність генетичних алгоритмів. Також доступні алгоритми для роботи з машинним навчанням [4].

В даній документації цієї бібліотеки є велика кількість прикладів, які демонструють її роботу, а також актуальна HTML-документація, яка допомагає початківцям розробникам, які хочуть застосовувати дану бібліотеку в своїх проєктах.

В основному, дана бібліотека інтегрується з фреймворком.NET, що полегшує розробку настільного додатку на мові C#. Головним недоліком є те, що AForge.NET підтримує тільки цю мову програмування.

Аналіз бібліотек комп'ютерного зору

В таблиці 1 розписано основні переваги та недоліки кожної з бібліотек.

Таблиця 1

Перелік переваг та недоліків бібліотек

Бібліотека	Переваги	Недоліки
OpenCV	<ul style="list-style-type: none"> – Наявність україномовної документації – Підтримка широкого набору мов програмування – Наявність засобів для роботи з нейронними мережами та машинним навчанням 	– Велика кількість другорядних функцій
VXL	– Незалежність компонентів бібліотеки	– Підтримка тільки однієї мови програмування: C++
AForge.NET	<ul style="list-style-type: none"> – Інтеграція з фреймворком.NET, що спрощує створення настільних додатків на мові C# – Наявність засобів для роботи з нейронними мережами та машинним навчанням 	– Підтримка тільки однієї мови програмування: C#

Джерело: розроблено автором

Також, особливу увагу треба звернути на швидкодію всіх розглянутих бібліотек, оскільки це може вплинути на якість та швидкість роботи усієї підсистеми розпізнавання об'єктів в цілому. Порівняльна діаграма наведена на рис. 1.

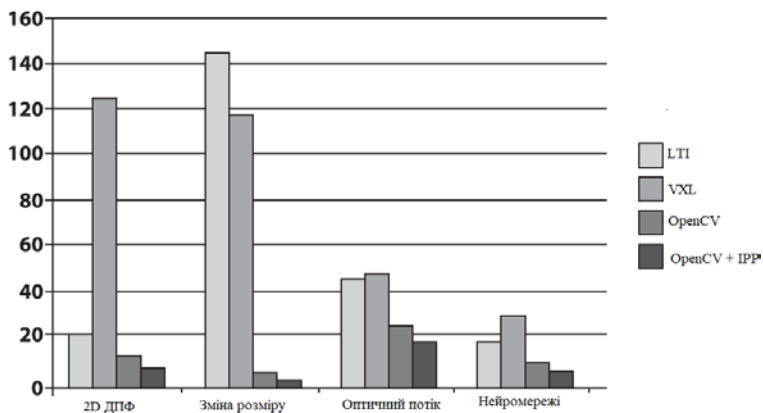


Рис. 1. Порівняння швидкодії бібліотек при виконанні різних операцій

Джерело: розроблено автором [5]

Було детально розглянуто бібліотеки комп'ютерного зору, які є популярними в середовищі Інтернет. Згідно проведеного короткого аналізу можливо помітити, що найбільше переваг має бібліотека OpenCV. До того ж, дана бібліотека є доволі швидкодіюною в порівнянні з іншими програмними засобами. Також, OpenCV, як вже було згадано, містить натреновані класифікатори для детекції облич, що дозволяє легко реалізувати алгоритми детекції облич. При потребі, класифікатори можливо переналаштувати на виявлення інших рухомих об'єктів. Ще однією перевагою на користь OpenCV є підтримка широкого набору мов програмування, а також наявність україномовної документації, що дещо спрощує розуміння бібліотеки та її основних компонентів.

Список використаних джерел:

1. The VXL Homepage : веб-сайт. URL: <https://vxl.github.io/> (дата звернення: 20.05.2020).
2. OpenCV. About : веб-сайт. URL: <https://opencv.org/about/> (дата звернення: 20.05.2020).
3. Pramod Poudel, Mukul Shirvaikar. Optimization of Image Processing Algorithms on Mobile Platforms. *Proceedings of SPIE – The International Society for Optical Engineering* 2011. DOI: 10.1117/12.876520
4. AForge.NET Features : веб-сайт. URL: <http://www.aforge.net/framework/features/> (дата звернення: 20.05.2020).
5. Kaehler A., Bradski G. Learning OpenCV 3 Computer Vision in C++ with the OpenCV Library. O'Reilly 2017. – 967 p.