

UDC 004.8

Денис Смішний, Олексій Алещенко

ОБРОБКА ВІЗУАЛЬНИХ ДАНИХ ЗА ДОПОМОГОЮ ТЕХНОЛОГІЇ COMPUTER VISION

VISUAL DATA PROCESSING BY COMPUTER VISION TECHNOLOGY

В даній статті розглядається питання розпізнавання об'єктів та образів з використанням технологій комп'ютерного зору. В ході дослідження описуються популярні алгоритми, такі як VGG16, DeepMask. Проводиться короткий огляд існуючих рішень в цій галузі. Визначаються перспективи використання комп'ютерного зору, зокрема на ринку мобільних пристроїв. Для цього проведено огляд можливостей бібліотеки OpenCV, та, як приклад, використання її у промислових продуктах.

Ключові слова: комп'ютерний зір, VGG16, OpenCV, розпізнавання об'єктів та образів.

Рис.: 3. Бібл.: 5.

In this article the question of recognition of objects and images using computer vision technologies is considered. The research describes popular algorithms such as VGG16, DeepMask. A brief overview of existing solutions in this area is conducted. The prospects for using computer vision, in particular in the mobile devices, are determined. To do this, reviewed the capabilities of the OpenCV library, and, for example, use it in industrial products.

Keywords: computer vision, VGG16, OpenCV, object and image recognition.

Fig. 3. Bibl.: 5.

Актуальність теми дослідження. В зв'язку зі зростаючими темпами впровадження інформаційних технологій в повсякденне життя користувачів, технології, що потребують залучення комп'ютерного зору набувають дедалі більшої популярності. Також подібні технології знаходять активне застосування в галузі криміналістики та пошукових операцій. Дані тенденції спонукають до широкого аналізу проблематики, пов'язаної з темою комп'ютерного зору та проведення певних досліджень в цій області.

Постановка проблеми. Недостатня кількість оглядових матеріалів з існуючих імплементацій машинного бачення. Дефіцит аналітичних статей з порівняння різних алгоритмів розпізнавання в контексті реальних прикладних задач. Заплутаність існуючих прикладів реалізації можливостей бібліотеки OpenCV.

Аналіз останніх досліджень та публікацій. Наразі у широкому доступі можна знайти величезну кількість матеріалів, присвячених тематиці комп'ютерного зору. Зокрема, доволі корисними в цьому плані будуть публікації на таких ресурсах як Хабрахабр [1] та Medium [2]. З іншого боку, подібні статті не мають акредитаційного підтвердження і можуть містити помилкову інформацію. Корисною для початківців буде стаття Learning to Refine Object Segments від науковців центру досліджень штучного інтелекту компанії Facebook [3]. Серед існуючих рішень корисним в плані розуміння існуючих алгоритмів можна виділити документацію бібліотеки з відкритим кодом OpenCV [4].

Виділення недосліджених частин загальної проблеми. На даний момент в науковій літературі спостерігається невелика кількість ґрунтовних публікацій з аналізом та порівнянням алгоритмів машинного бачення. У цих публікаціях рідко приділяється увага існуючим бібліотекам реалізацій алгоритмів та написанню прикладних програм з їх використанням.

Постановка завдання. Завданням є огляд основних алгоритмів реалізації технології комп'ютерного зору, наведення прикладу навчання нейронної мережі для визначення об'єктів на картинці та огляд бібліотеки OpenCV.

Викладення основного матеріалу. Комп'ютерним зором називають складну систему визначення та аналізу об'єктів та образів, отриманих системою ззовні. Вивчення можливостей та потенціалу комп'ютерів в галузі машинного бачення розпочалися ще в 70-х роках минулого століття, але активного розвитку зазнали саме час. Зокрема, було створено багато алгоритмів для розпізнавання образів: VGG16, VGG32, VGG29, ResNet, DenseNet, група алгоритмів Inception, тощо. Розглянемо детальніше кілька з них.

VGG16. VGG розшифровується як Visual Geometry Group, тобто група науковців з факультету інженерних наук Оксфордського університету, а 16 - число слоїв нейронної мережі, необхідних для розпізнавання образів. Суть даного алгоритму наступна:

- 1) поділ зображення пополам, доки не дійдемо до глибини зображення 14×14 пікселів (мінімальна можлива глибина - 3×3 пс);
- 2) в процесі поділів отриманих шматків зображення намагаємось отримати маски шуканого образу та коефіцієнти ймовірності знаходження шуканого образу в даній ділянці зображення;
- 3) на основі аналізу зменшених частин зображення будуються вихідні незалежні описи. Потім результат усереднюється для всіх масок.

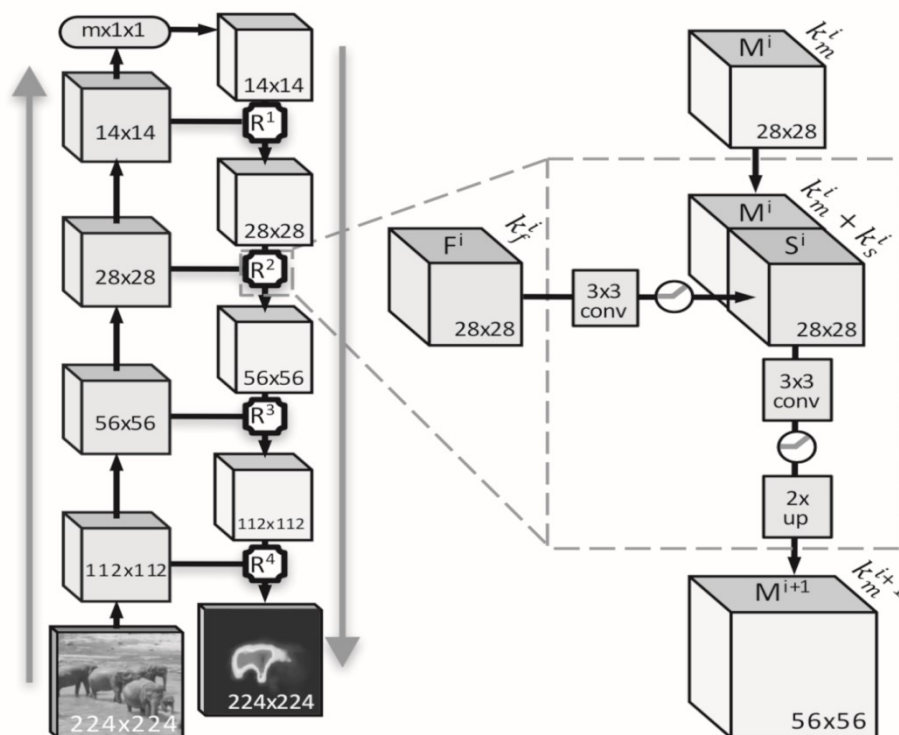


Рис 1. Схема роботи алгоритму

Даний підхід добре працює для знаходження образів, що контрастують з фоном, але не годиться для вилучення одноманітних об'єктів (Наприклад, знайти кількість овець на фото) або знаходження об'єктів, колір яких збігається з кольором фону. Для цього слід використовувати вдосконалений алгоритм згорткової нейронної мережі - DeepMask.

DeepMask. По факту, даний алгоритм є вдосконаленим алгоритмом VGG16. Суть в тому, що на кожному кроці ми додаємо не тільки маску образу, але й використовуємо інформацію, отриману від нижніх шарів, а також дані з розпізнаних верхніх шарів.

Кожний етап зменшення вихідного зображення дозволяє отримати кодовану маску, згенеровану при проходженні від загального зображення до зменшеного. Разом з цим у вдосконаленій згортковій нейронній мережі відбувається рух від меншого зображення до більшого з отриманням характерних функцій (точок). Підсумком є маска, отримана шляхом двонаправленого злиття функцій і опорних точок.



Рис.2. Результат роботи DeepMask

Головна проблема DeepMask в тому, що ця модель використовує просту мережу прямого поширення, яка успішно створює маски з відносно низьким рівнем дискретизації, але не виконує сегментацію з точністю до пікселя [3].

OpenCV. OpenCV (Open Source Computer Vision Library) - це бібліотека програмного забезпечення з відкритим кодом для комп'ютерного зору та машинного навчання. OpenCV була побудована, щоб забезпечити загальну інфраструктуру для програм комп'ютерного зору та прискорити використання машин у комерційних продуктах. OpenCV, що поширюється за ліцензією BSD, дозволяє підприємствам легко використовувати та змінювати код [5].

Бібліотека реалізує понад 2500 оптимізованих алгоритмів. Вони можуть бути використані для виявлення та розпізнавання облич, визначення об'єктів, класифікації людських дій у відео, відстеження руху камери, відстеження рухомих об'єктів, витягування 3D-моделей об'єктів, створення 3D-точкових хмар із стереокамер, зшивання зображень разом для отримання високої роздільної здатності зображення цілісної сцени, пошуку схожих зображень з тими, що знаходяться в базі даних, корекції ефекту "червоних очей", що виникає при використанні спалаху під час фотографування, відстежування рухів очей, визначення декорації та встановлення маркерів для подальшого накладання їх на доповнену реальність тощо.

Дослідження. Використовуючи бібліотеку OpenCV реалізуємо детекцію QR-коду з урахуванням кутових відхилень. Для цього скористаємось теоремою про співвідношення.

Визначення: співвідношення площ N числа двовірних полігонів залишається інваріантним до рівномірного геометричного перетворення:

Якщо a_i - це площа багатокутника i , то:

$$a_1 : a_2 \dots : a_N = R \quad (1)$$

в 2D площині X . Тоді:

$$a'_1 : a'_2 \dots : a'_N = R \quad (2)$$

для будь-якої іншої 2D-площини Y , яка є геометричним перетворенням X :

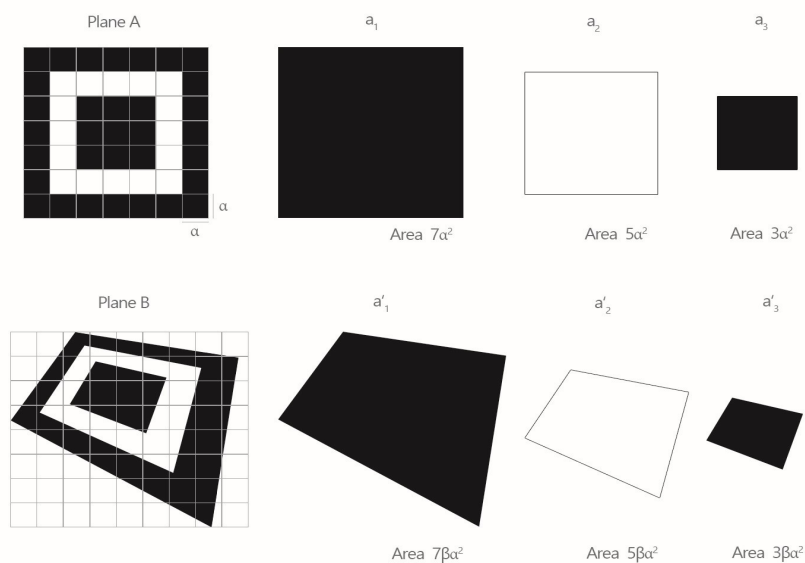


Рис. 3. Співвідношення площ полігонів

Вихідний код проекту доступний за посиланням <https://github.com/denGlob/CppDemo>.

Висновки. В ході проведеної роботи було розглянуто перспективи використання комп'ютерного зору. Оглянуто основні алгоритми виявлення образів та об'єктів на зображеннях. Зазначається, що алгоритм VGG16 добре працює для об'єктів, що контрастують з фоном, але є дуже неефективним в типових практичних задачах, де фон часто зливається з шуканим об'єктом. Для цього краще використовувати модернізації алгоритму такі як DeepMask та SharpMask. Оглянуто бібліотеку OpenCV. На її основі створено тестовий додаток для детекції QR-кодів.

Список використаних джерел

1. ComputerVision и с чем его едят [електронний ресурс] - <https://habr.com/company/funcorp/blog/351638/>
2. Lars Hulstaert. Going deep into object detection [електронний ресурс] - <https://towardsdatascience.com/going-deep-into-object-detection-bed442d92b34>
3. Pedro O. Pinheiro, Tsung-Yi Lin, Ronan Collobert, Piotr Doll'ar. (Jul 2016) Learning to Refine Object Segments (p.15)
4. OpenCV Documentation [електронний ресурс] - <https://docs.opencv.org/>
5. Zheng, S., Jayasumana, S., Romera-Paredes, B., Vineet, B., Su, Z., Du, D., Huang, C., Torr, P. (2015). Conditional random fields as recurrent neural nets. In: ICCV.

ДОВІДКА ПРО АВТОРІВ

Смішний Денис Миколайович - студент IV курсу, група ІО-41, комп'ютерна інженерія, кафедра обчислювальної техніки, Національний технічний університет України «Київський політехнічний інститут імені Ігоря Сікорського».

Smishnyi Denys - student of IV course, group IO-41, computer engineering, Department of Computer Engineering, National Technical University of Ukraine "Kyiv Polytechnic Institute named after Igor Sikorsky".

E-mail: denshiksmile@gmail.com

Алещенко Олексій Вадимович – асистент, кафедра обчислювальної техніки, Національний технічний університет України "Київський політехнічний інститут імені Ігоря Сікорського".

Aleshchenko Oleksii – assistant professor, Department of Computer Engineering, National Technical University of Ukraine "Igor Sikorsky Kyiv Polytechnic Institute".

E-mail: alexey.aleshchenko@gmail.com

Denys Smishnyi, Oleksii Aleshchenko

VISUAL DATA PROCESSING BY COMPUTER VISION TECHNOLOGY

Relevance of research. In connection with the increasing rate of implementation of information technology in the daily life of users, technologies that require the attraction of computer vision are gaining popularity.

Formulation of the problem. Creation of a review of existing implementations of machine vision. Analysis of popular algorithms for the operation of neural networks for finding images and objects. Guide examples of commercial implementation of computer vision technology.

Analysis of recent research and publications. At present, you can find a huge amount of materials on the subject of computer vision in broad access. For beginners, the Learning to Refine Object Segments article is from the research team at Facebook's Artificial Intelligence Research Center. Among the existing solutions useful in understanding the existing algorithms, you can select the library documentation with OpenCV open source code.

Uninvestigated parts of general matters defining. To date, there is a small number of solid scientific articles with analysis and comparison of machine vision algorithms, as well as a review of libraries and their implementations.

Setting objectives. The task is to examine the main algorithms for the implementation of computer vision technology, give an example of studying the neural network to identify objects in the picture and review the library OpenCV.

Presentation of the main material. At the first stage we overview the VGG16 algorithm. Research shows, that this algorithm is good for contrast objects, but rather better to use DeepMask algorithm or SharpMask in partial tasks. Also, we overview in a short OpenCV library. It contains 2500 algorithms for computer vision and bridge code for popular languages such as Java and Python for comfortable coding.

Conclusions. During the work the prospects of using computer vision were considered. The main algorithms for the detection of images and objects in images are reviewed. It is determined that the VGG16 algorithm works well for objects contrasting with the background but is very ineffective in typical practical tasks, where the background often merges with the target object. To do this, it's best to use an upgrade algorithm such as DeepMask and SharpMask. The OpenCV library is under review. On its basis, a test application for the detection of QR-codes was created.

Keywords: computer vision, VGG16, OpenCV, object and image recognition.