

ІЛАЩУК МИКОЛА

Чернівецький національний університет ім. Ю. Федьковича

<https://orcid.org/0009-0002-7996-6176>e-mail: ilashchuk.mykola.m@chnu.edu.ua

КУШНІР ІРИНА

Чернівецький національний університет ім. Ю. Федьковича

e-mail: i.kushnir@chnu.edu.ua

МЕЛЬНИЧУК СТЕПАН

Чернівецький національний університет ім. Ю. Федьковича

e-mail: s.melnychuk@chnu.edu.ua

РОЗПІЗНАВАННЯ ОБЛИЧ В РЕАЛЬНОМУ ЧАСІ ЗА ДОПОМОГОЮ БІБЛІОТЕКИ OPENCV ТА МОВИ ПРОГРАМУВАННЯ PYTHON

В статті розглядається процес розпізнавання облич, що є однією з ключових задач в області комп'ютерного зору. Основна увага приділяється використанню бібліотеки OpenCV, яка забезпечує інструменти для ефективного та швидкого виявлення облич на зображеннях та відео в реальному часі.

У роботі описується каскад Хаара та алгоритм Віоли-Джонса, який лежить в основі визначення облич, реалізованого в OpenCV. Також розглядаються покрокові етапи реалізації проекту на мові програмування Python, обробка зображень. Наводяться результати експериментів, що демонструють ефективність і точність запропонованого підходу. Стаття буде корисною для розробників, інженерів та дослідників, зацікавлених у впровадженні технологій розпізнавання облич у своїх проектах використовуючи базові алгоритми комп'ютерного зору.

Ключові слова: OpenCV, Python, зображення, Хаар каскад, алгоритм Віоли-Джонса, ефективність.

ILASHCHUK MYKOLA, KUSHNIR IRYNA, MELNYCHUK STEPAN

Yuriy Fedkovych Chernivtsi National University

REAL-TIME FACE RECOGNITION USING OPENCV LIBRARY AND PYTHON PROGRAMMING LANGUAGE

The article discusses the process of face recognition, which is one of the key tasks in the field of computer vision. The main focus is on the use of the OpenCV library, which provides tools for efficient and fast face detection in images and real-time video.

The paper describes the Haar cascade and the Viola-Jones algorithm, which form the basis of the face detection implemented in OpenCV. It also discusses the step-by-step stages of project implementation in the Python programming language, as well as image processing. The results of experiments demonstrating the effectiveness and accuracy of the proposed approach are presented. The article will be useful for developers, engineers, and researchers interested in implementing face recognition technologies in their projects using basic computer vision algorithms.

In addition to the fundamental techniques, the article delves into the nuances of optimizing the face detection process. This includes adjusting parameters such as the scale factor and the minimum number of neighbors, which are crucial for balancing detection accuracy and performance. The importance of preprocessing steps like grayscale conversion and using the correct type's cascade file is also highlighted, as these steps significantly enhance the detection rate by improving image contrast.

The theoretical implementation section provides a step-by-step conception of face detection and recognition using the Haar cascade and Viola-Jones method. Also, need to underline the importance of this open topic about the recognition of human emotions using the research by Robert Plutchyk. He determines the main human emotions that are the basis for combining the recognition of a person and his emotions.

The practical implementation section provides a realization of face detection and recognition using OpenCV and Python: loading image data, applying the Haar cascade classifier, and testing the program solution.

To validate the effectiveness of the approach, several experiments were conducted using diverse image datasets. The results, summarized in tabular format, indicate a high detection rate with minimal input data. The experiments also compare the performance with different counts of input data, highlighting the advantages and disadvantages of the Viola-Jones method.

Keywords: opencv, python, image, haar cascade, viola-jones, efficiency.

Постановка проблеми у загальному вигляді

та її зв'язок із важливими науковими чи практичними завданнями

Розпізнавання облич є однією з найважливіших і найскладніших задач у сфері комп'ютерного зору. Основна проблема полягає в необхідності розробки та впровадження систем, які можуть ефективно та точно розпізнавати обличчя в умовах реального часу та визначити емоційний стан людини за допомогою аналізу виразів обличчя, враховуючи різноманітність умов освітлення, кутів зору, виразів облич та інших факторів, що можуть впливати на якість розпізнавання.

Використання бібліотеки OpenCV та мови програмування Python для вирішення цієї задачі є базовим напрямком завдяки інструментам, які надає OpenCV для обробки зображень та відео, а також простоті та гнучкості Python. Алгоритми, реалізовані в OpenCV, такі як каскад Хаара та алгоритм Віоли-Джонса, дозволяють точно виявляти обличчя на зображеннях. Проте, для ефективного застосування цих алгоритмів у реальних додатках необхідно проаналізувати та оптимізувати їхню роботу під конкретні умови та вимоги.

Аналіз досліджень та публікацій

В роботі [1] наведено дані про емоції з точки зору психології, біології та еволюції, розкриваючи їхню роль у людському житті. Робота підкреслює взаємозв'язок між емоціями та адаптаційними механізмами, які сприяють виживанню.

В статті [2] Матсумото аналізує методологічні аспекти оцінки емоцій у міжкультурних дослідженнях, зокрема коментуючи роботу Ельфенбейн та Амбаді. Він обґрунтовує важливість врахування культурних відмінностей при вивченні розпізнавання емоцій, пропонуючи вдосконалені методи дослідження.

В роботі [3, 6] автори представили революційний підхід до розпізнавання обличчя в реальному часі, використовуючи каскад Хаара та метод Adaboost. Вони показали, що їхній алгоритм забезпечує високу точність і швидкість, що робить його придатним для практичного застосування.

В праці [4] автор досліджує сучасні досягнення в теорії вейвлетів та їх застосування в інженерії, фізиці та технологіях. Також демонструє різноманітні способи застосування вейвлетів для аналізу сигналів і зображень, підкреслюючи їхню ефективність і універсальність.

В роботі [5] наведено дані про аналіз останніх досягнень в області розпізнавання обличчя, зокрема використання інтегрального представлення зображень. Робота висвітлює різні алгоритми та підходи, що сприяють підвищенню точності та швидкості розпізнавання.

Формулювання цілей статті

Метою роботи є розробка та впровадження ефективної системи розпізнавання обличчя в реальному часі, використовуючи бібліотеку OpenCV та мову програмування Python. Також ключовим аспектом являється аналіз алгоритмів, що забезпечують ефективне виявлення обличчя на зображеннях та відео, зокрема каскад Хаара та алгоритм Віоли-Джонса.

Виклад основного матеріалу

Не заперечним є факт, що цивілізація в останні десятиліття здійснила стрибок у розвитку науки і технологій. Разом із тим ця сама цивілізація все частіше стикається з проблемами взаєморозуміння між людьми. Масові соціальні мережі не можуть зменшити ці проблеми, скоріше навпаки – вони їх усугубляють. Живі люди потребують прямого, часто емоційного спілкування. Крім того, часто вміння спілкуватися з людьми стає запорукою успішності професійної діяльності. Нездатність зрозуміти співбесідника, увійти в його положення, приводить до професійної некомпетентності. Ефективна комунікація передбачає адекватну оцінку емоційного стану співбесідника, що часто можна оцінити за допомогою інформативних невербальних проявів, зокрема по виразу його обличчя.

Таким чином аналіз виразу обличчя є засобом розпізнавання емоцій людини іншою людиною. Відповідно за допомогою методів комп'ютерного зору цей процес можна автоматизувати. Комп'ютерний зір – це розділ науки, в рамках якого створюються та аналізуються алгоритми розпізнавання об'єктів. Американський психолог Роберт Плутчик визначив вісім основних емоцій людини: гнів, очікування, відразу, страх, радість, смуток, подив і довіру [1]. Емоції проявляються по-різному і в комбінації однієї емоції з іншою можуть вийти абсолютно нові емоційні стани. Виявляється, що різні люди незалежно від вікових категорій, від статі та інших характеристик однакові у виразі своїх емоцій: вони однаково усміхаються коли радіють або хмурять брови, розсердившись. Ці базові емоції називаються «дискретними».

Розпізнавання емоцій – це частинний випадок більш загального поняття - розпізнавання образів. Розпізнати об'єкт означає за набором ознак даного об'єкта віднести його до одного із заданого набору можливих значень. В даному випадку до однієї із базових емоцій.

В психології розроблено цілий ряд методів для визначення та розпізнавання емоцій за виразом обличчя. Наприклад методика JACFEE дозволяє розпізнавати емоцій за фотографіями, і її результати можуть розглядатися як об'єктивні показники емоційних проявів [2].

На початку XXI століття до задач розпізнавання емоцій почали підключатись технологічні засоби. Зокрема розроблялися алгоритми і програми, спрямовані на розпізнавання та визначення емоцій по фотографіях окремих індивідуумів. Вивчення різних методів розпізнавання емоцій за допомогою ІТ технологій триває і в даний час. Причому, реалізовані методи доповнюються новими, які дозволяють досягати якісніших результатів. Цікавою видається задача розпізнати емоції певної групи людей, використовуючи методи психології та методи, якими оперують ІТ технології та провести порівняльний аналіз отриманих результатів. Це достатньо громіздке завдання, розв'язання якого почнемо із задачі розпізнавання обличчя.

Задача розпізнавання обличчя може бути розв'язана за допомогою використання бібліотеки комп'ютерного зору OpenCV, варіанти реалізації якої є для багатьох існуючих систем. Існує цілий ряд методів розпізнавання об'єктів, серед яких популярним і ефективним є метод Віоли-Джонса [3].

Метод Віоли-Джонса

У сучасній теорії та практиці вивчення як одновимірних так і двовимірних зображень активно використовується опис сигналів за допомогою вейвлетів [4]. В сучасних практиках використовують кілька десятків вейвлетів, кожен із яких має свої особливості. Тому вдалий вибір вейвлета є необхідною умовою розв'язання поставленої задачі. Вейвлети визначаються як функції однієї або двох дійсних змінних в залежності від області визначення, та структури області можливих значень. Розрізняють дискретні та неперервні вейвлети. Вейвлети Хаара задаються на скінчених інтервалах різних змінних та приймають два значення $\{-1; +1\}$.

Віола і Джонс адаптували ідею використання вейвлетів Хаара і розробили так звані ознаки Хаара, які складаються з суміжних прямокутних областей білого та чорного кольорів (Рис.1).

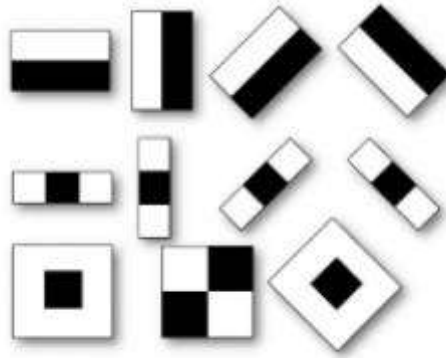


Рис. 1. Візуалізація вейвлетів Хаара

Оскільки на чорно-білій фотографії людського обличчя апіорі присутні темніші (очі, рот) та світліші (лоб, щоки) області, то визначений діапазон значень яскравості (0-255) може бути використаний для створення класифікатора. Такий класифікатор повинен базуватися на відношеннях яскравостей різних зон об'єкта.

У методі Віоли-Джонса використовується інтегральне представлення зображення [5] – матриця, яка за розмірами співпадає із матрицею вихідного зображення. Кожний елемент матриці вихідного зображення задає інтенсивність даного пікселя і є числом в інтервалі (0,255), а кожний елемент матриці інтегрального зображення знаходиться згідно виразу:

$$L(x, y) = \sum_{i=0, j=0}^{i \leq x, j \leq y} I(i, j) \tag{1}$$

де, $I(i, j)$ – яскравість пікселів вихідного зображення. Кожен елемент матриці $L(x,y)$ знаходиться як сума інтенсивностей пікселів прямокутника від (0,0) до (x,y). Тобто значення пікселя (x,y) рівне сумі значень інтенсивностей всіх пікселів, які розташовані зліва від пікселя (x,y) та вище від нього. Суму значень пікселів довільного прямокутника ABCD можна розрахувати через суму і різниці сусідніх прямокутників згідно виразу, як це видно із рис.2

$$S_{ABCD} = L(A) + L(C) - L(B) - L(D) \tag{2}$$

де, для розрахунку значення в конкретній точці із координатами (x,y) використовуємо формулу

$$L(i, j) = I(i, j) + L(i - 1, j) + L(i, j - 1) - L(i - 1, j - 1) \tag{3}$$

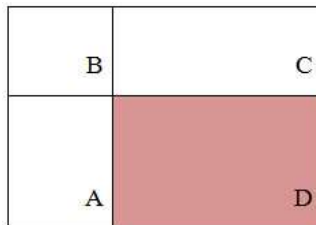


Рис. 2. Математична візуалізація прямокутника ABCD

Таким чином, матриця інтегрального зображення містить інформацію про сам об'єкт зображення. Оскільки на обличчі є темніші і світліші зони, то можна виділити певні ознаки відношенням яскравості цих зон, що слугуватиме певним класифікатором. Тобто класифікатор базується на відношенні яскравості певних зон досліджуваного об'єкта. Як було показано вище, ознаки Хаара складаються з суміжних прямокутних областей, які позиціонуються на зображенні, далі підсумовуються інтенсивності пікселів в чорній і білій областях і обчислюється різниця між сумами. Ця різниця і буде значенням конкретної певної ознаки, яка заданим чином позиціонована на зображенні.

Ці ознаки є досить слабкими, кожна з них окремо не є показовою. Тому було запропоновано об'єднати їх в каскад, що вже може мати задовільну точність, за умови застосування ефективного алгоритму навчання каскаду. Частіше за інших використовується алгоритм AdaBoost [6].

Вхідними даними для даного проекту являється відео тривалістю близько 5 секунд зняте в анфас, де студенти відтворюють «мрійливий характер» (тобто усмішка, радість). Взято за основу саме відео по причині більшої кількості вхідних даних кожної людини для більш точного навчання системи. Тобто чим більше фото конкретної людини різного характеру, тим система має менше схожості з іншими особами.

Також варто зауважити, що на відео повинні бути в полі зору обидва ока, оскільки, в навчанні системи розпізнавання використовується Хаар каскад з фронтальним розпізнаванням обличчя, який є складовою частиною бібліотеки OpenCV [7].

Після розкадрування відео на фото-файли у .jpg форматі, наступним ключовим фактором являється робота з фотографіями, а саме виявлення обличчя та їх ідентифікація.

Для обробки даної операції виконуються наступні кроки:



Важливо, що у даному випадку система працює в реальному часі. Тобто підключається камера, ідентифікується її унікальне ID, зображення презентується на екран та виділяється актуальна частина обличчя. Для аналізу роботи системи було взято групу студентів з 10, 20 та 30 осіб.

Для реалізації було сформовані файли, що містили 10, 20 та 30 облич. Беремо конкретне обличчя (присутнє (фото 1) або відсутнє у файлах (фото 2)), готуємо його до обробки в системі OpenCV згідно кроків приведених вище і здійснюємо пошук в цих трьох файлах.

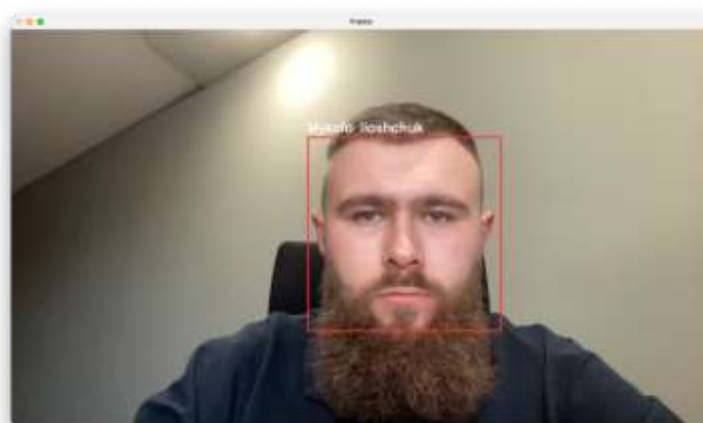


Рис. 1. Людина присутня у файлах

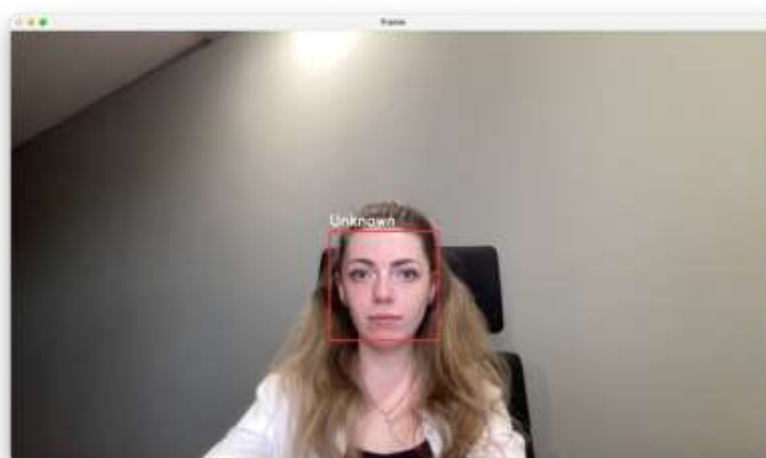


Рис. 2. Людина відсутня у файлах

Результати досліджень розпізнавання облич за допомогою OpenCV

Кількість даних, відео	Якість розпізнавання, %	Кількість помилок розпізнавання	Час підготовки до розпізнавання, хв	Якість візуалізації відео на екрані
10	100	1	1	в нормі
20	90,5	2	4	чіткість візуалізації
30	81	4	7	чіткість візуалізації

Як видно з Таблиці 1, у випадку коли у файлі 10 облич, система розпізнала 10 з 10 студентів, а невідому особу – ідентифікувала як “Unknown”. У випадку 20 об'єктів у файлі, відсоток розпізнавання залишається високим, оскільки 19 з 20 студентів були розпізнані та правильно ідентифікує невідомі особи. У випадку 30 зображень - відсоток розпізнаних облич складав 81% та невідомі особи також визначалися правильно.

Час підготовки відео до розпізнавання залежить від кількості заведених даних. У випадку 10 – до 1 хвилини, 20 – 4 хвилини, 30 – близько 7 хвилин.

Оскільки система порівнює вибране обличчя з даними відео в реальному часі, то погіршується затримка візуалізації відео на екрані. Зі збільшенням кількості облич у файлі, якість погіршується.

**Висновки з даного дослідження
і перспективи подальших розвідок у даному напрямі**

Методика розпізнавання обличчя з використання каскад Хаара та бібліотеки OpenCV не погано працює при малих об'ємах вхідних даних (до 20 осіб). Збільшення вхідних даних потребує використання кращих алгоритмів з використанням нейронних мереж.

References

1. Plutchik, R. Emotions and Life: Perspectives from Psychology, Biology, and Evolution. 2002.
2. Matsumoto D. Methodological Requirements to Test a Possible In-Group Advantage in Judging Emotions Across Cultures: Comments on Elfenbein and Ambady (2002) and Evidence. JACFEE. 2002.
3. Viola, P., Jones, M. J. Robust Real-time Face Detection. International Journal of Computer Vision. 2004. № 57(2). 137–154.
4. Baleanu D. Advances In Wavelet Theory And Their Applications In Engineering, Physics And Technology. 2012.
5. Zhang C., Zhang Z. A Survey of Recent Advances in Face Detection. Microsoft Research Technical Report MSR-TR-2010-66. 2010.
6. Zhu J., Rosset S., Hastie T., Tibshirani R. Multiclass AdaBoost. Statistics and Its Interface. 2009. № 2(3). 349–360.
7. Haar Cascade Classifier for Face Detection. OpenCV. https://github.com/kipr/opencv/blob/master/data/haarcascades/haarcascade_frontalface_default.xml