

УДК 004.93`1

DOI <https://doi.org/10.32851/tnv-tech.2022.3.6>

АНАЛІЗ МЕТОДІВ КОМП'ЮТЕРНОГО ЗОРУ В ЗАДАЧАХ ІДЕНТИФІКАЦІЇ ОСІБ У ВІДЕОПОТОЦІ

Мозговенко А. А. – асистент кафедри комп'ютерних наук
Таврійського державного агротехнологічного університету
імені Дмитра Моторного
ORCID ID: 0000-0002-7445-8925

Зінов'єва О. Г. – старший викладач кафедри комп'ютерних наук
Таврійського державного агротехнологічного університету
імені Дмитра Моторного
ORCID ID: 0000-0003-3760-8952

Завдання виявлення обличчя на зображенні часто є першим кроком у процесі розв'язання задачі вищого рівня – розпізнавання особи, деталей особи або її ліміки. Крім того, інформація про присутність та кількість осіб на зображенні може бути корисною в системах автоматичного обліку кількості відвідувачів; системах пропускнуго контролю в установах, аеропортах та метро; автоматичних системах запобігання нещасним випадкам; інтелектуальних інтерфейсах «людина-комп'ютер»; у фототехніці для автоматичного фокусування на обличчі людини, а також для стабілізації зображення особи з метою полегшення розпізнавання емоцій; для розширення зони стереобачення під час створення систем 3D відображення. Те, що для нас є фотографією, для системи розпізнавання образів – лише набір пікселів з різними параметрами кольору. Щоб навчити систему розпізнавати на зображенні окремі об'єкти, необхідно надати їй датасета – набір з тисяч зображень, в яких зазначено, де саме знаходиться потрібний об'єкт. Наприклад, якщо ми хочемо, щоб система навчилася розпізнавати на знімках людей, потрібно показати їй безліч фотографій людей різного віку, в різних позах і одязі, в різних умовах. Після подібного тренування система зможе безпомилково розпізнати людину на фотографіях. Однак напрошується інше запитання: якщо для системи фотографія – це просто сукупність пікселів, то як же нейросеть розуміє, що саме зображено на фото?

Значне зростання обчислювальної потужності комп'ютерів та поява нових математичних моделей алгоритмів останніми роками дозволило домогтися начительного прогресу у сфері комп'ютерного зору. Нові технології, засновані на машинному навчанні, дозволяють створювати надійні рішення для складних задач розпізнавальних образів.

Дана стаття є оглядом основних підходів до вирішення завдань комп'ютерного зору, з акцентом на глибинне машинне навчання та штучні нейронні мережі. У статті розглянуто алгоритми розпізнавання об'єктів на зображенні, проведено аналіз методів, що застосовуються при обробці зображень, а також описано використання засобів машинного навчання в рамках роботи із зображеннями.

Ключові слова: комп'ютерний зір, нейронні мережі, розпізнавання об'єктів, методи.

Mozghovenko A. A., Zinovieva O. G. Analysis of computer vision methods in the problems of personal identification in video flow

The task of detecting a face in an image is often the first step in the process of solving a higher-level problem – facial recognition, facial details or facial expressions. In addition, information about the presence and number of people in the image can be useful in systems of automatic accounting of the number of visitors; access control systems in institutions, airports and the metro; automatic accident prevention systems; intelligent human-computer interfaces; in photography for automatic focusing on a person's face, as well as to stabilize the image of the face in order to facilitate the recognition of emotions; to expand the stereo vision area when creating 3D display systems. What is photography to us, to the image recognition system, is just a set of pixels with different color settings. To teach the system to recognize individual objects in an image, you need to provide it with a dataset – a set of thousands of images that indicate exactly where the desired object is located. For example, if we want the system to learn to recognize people in pictures, we

need to show it a lot of photos of people of different ages, in different poses and clothes, in different conditions. After such training, the system will be able to unmistakably recognize the person in the photos. However, another question arises: if for the system the photo is just a set of pixels, then how does the neural network understand what is depicted in the photo?

Significant growth in the computing power of computers and the emergence of new mathematical models of algorithms in recent years has made significant progress in the field of computer vision. New technologies based on machine learning allow us to create reliable solutions for complex adaptive image recognition.

This article is an overview of the main approaches to solving computer vision problems, with an emphasis on deep machine learning and artificial neural networks. The article considers algorithms for recognizing objects in the image, analyzes the methods used in image processing, and describes the use of machine learning tools in working with images.

Key words: *computer vision, neural networks, object recognition, methods.*

Постановка проблеми. Для розпізнавання об'єктів на зображенні використання об'єктів використовуються різні методи, але одним з найбільш перспективних визнаний метод гістограм орієнтованих градієнтів (HOG). Зображення знебарвлюється, а потім в блоках 16x16 пікселів система знаходить напрямки змін кольору (вектор градієнта), буде по всьому зображенню карту цих векторів, і тим самим «знімок» ознак об'єкта, які не змінюються в залежності від пози/положення і освітлення.

Аналіз останніх досліджень і публікацій. На сьогодні на ринку програмних продуктів існує велика кількість програмних забезпечень по розпізнаванню облич. Серед яких можна виділити наступні:

1. Search Face.
2. Google picture.
3. Пошук по фото.

Програмне забезпечення Search Face виконує розпізнавання осіб по фото в Інтернеті з бази даних. Розумний алгоритм, спираючись на дані вбудованої нейромережі, швидко аналізує образ людини і знаходить образи схожих людей. Сервіс працює онлайн, тобто, нічого завантажувати або встановлювати на пристрій не потрібно. Все, що потрібно – браузер і Інтернет підключення.

До переваг програмного забезпечення відносять:

- висока швидкість аналізу даних;
- висока точність (вихідне фото жодного разу не потрапляло в Інтернет);
- простота інтерфейсу користувача;
- можливість користуватись додатком на мобільного та комп'ютера;
- наявність додатку на Play Markt [8];

До недоліків відносять: необхідність вводу власного номеру телефону;

Сервіс Google.picture. Програмне забезпечення розпізнавання осіб коштує чимало, і воно навряд чи підійде звичайному користувачеві Інтернету. Таке обладнання набувають великі компанії, супермаркети, державні служби.

Ось кілька компаній, які реалізують програми для розпізнавання осіб в відеоспостереження:

- Findface;
- Hikvision;
- Verilook Standard SDK.

Середня ціна за стартовий пакет додатку – від 450 доларів. Професійні продукти навряд чи зацікавлять стандартного користувача.

Один з найбільш популярних веб-сервісів, який відмінно працює безкоштовно – це Google.photo. Програма виконує не тільки розпізнавання осіб по фото, а й авторизує користувача по образу.

До переваг google.picture можна віднести:

- аналізує не тільки базу даних, а й фото с папки;
- більше шансів на авторизацію особи;

До недоліків відносять наступне:

- довге зчитування бази даних;
- у зв'язку з великою кількістю даних має повільну швидкість роботи.

Мобільний додаток «Пошук по фото» Вона виконує пошук і розпізнавання особи з фото вихідника, ґрунтуючись на даних пошукових систем Google. Цей додаток сумісний з операційною системою IOS 9.1 і вище.

До переваг додатку можна віднести:

- програма розпізнає особу вельми непогано і швидко;

До недоліків додатку:

- часто вилітає;
- довго доводиться розбиратися з інтерфейсом, навіть за допомогою інструкції;
- Порівняльний аналіз аналогів розглянуто і зведено у таблиці 1.

Таблиця 1

Порівняння аналогів

Системі розпізнавання облич	Швидкість роботи	Точність	Простота інтерфейсу
Search Face	+	+	+
Яндекс.Зображення	-	+	+
Пошук по фото	+	-	-

Постановка завдання. Мета даного дослідження – розглянути різні методи, що використовуються для цифрової обробки зображень у сучасних системах комп'ютерного зору, виділити галузі застосування таких систем, а також показати нові шляхи та тенденції їх розвитку.

Виклад матеріалу дослідження. Визначимо перелік завдань, що вирішуються системами розпізнавання та ідентифікації осіб по відеофіксації в реальному часі:

1. Верифікація. Система розпізнавання та ідентифікації осіб по відеофіксації в реальному часі може підтверджувати особистість людини шляхом порівняння висунутого зразка з еталонним зразком, записаним в систему раніше. Фактично, виконується порівняння за схемою «один до одного».

2. Ідентифікація на закритому безлічі. Система розпізнавання та ідентифікації осіб по відеофіксації в реальному часі зіставляє отримане з відеопотоку зображення з безліччю записаних раніше шаблонів різних людей з метою встановлення особи, якій належить цей зразок. Дане завдання може бути вирішене з урахуванням того, що персона, якій належить зразок, присутній в базі. У цьому завданні виконується порівняння зразків за схемою «один до багатьох».

По суті, системи розпізнавання є комп'ютерними програмами, які аналізують зображення облич людей з метою їх ідентифікації. Програма бере зображення особи і вимірює такі його характеристики, як відстань між очима, довжина носа, кут щелепи, на основі чого створюється унікальний файл, який називається «шаблон». Використовуючи шаблони, програма порівнює дане зображення з іншими зображеннями в базі, а потім оцінює, наскільки зображення є схожими один на одного. Звичайними джерелами зображень для використання при ідентифікації по обличчю є сигнали від відеокамер або раніше отримані фотографії, на зразок тих, що зберігаються в базі даних водійських посвідчень.

Нейронні мережі не програмуються в звичному сенсі цього слова, вони навчаються. Можливість навчання – одне з головних переваг нейронних мереж перед традиційними алгоритмами. Розвиток штучних нейронних мереж надихається біологією.

Математична модель, а також її програмне або апаратне втілення, побудована за принципом організації та функціонування біологічних нейронних мереж – мереж нервових клітин живого організму. Це поняття виникло при вивченні процесів, що протікають в мозку, і при спробі змоделювати ці процеси.

Після того, як зображення у вигляді цифрових даних з камер передається на комп'ютер – воно обробляється за допомогою спеціального алгоритму, який визначає розташування області особи за її основними рисами (очей, рота, брів, носа і т. д.). Таких методів виявлення осіб існує багато і більшість з них представляють собою комбінацію інших методів. Але всі їхні можна розбити на дві категорії: методи на основі знань, які ґрунтуються на досвіді людини і методи виявлення особи за зовнішніми ознаками (методи при яких необхідно провести етап навчання системи, напрямом обробки тестових зображень

Також варто відразу зазначити, що продуктивність обраних методів істотно залежить від різних чинників, наприклад: освітлення, положення особи, заднього фону і т.д. Тому, для забезпечення коректної роботи системи, необхідно створити сприятливі умови для її використання:

- передне, або близьке до нього положення особи;
- особа, яка не перекривається іншими об'єктами;
- нейтральне вираз обличчя;

Методи, засновані на знаннях, використовують інформацію про особу, його рисах, формі, текстурі або кольорі шкіри. У цих методах виділяється якийсь набір правил (властивостей і особливостей особи), яким повинен відповідати фрагмент кадру, для того щоб вважатися людським обличчям. Визначити такий набір правил досить легко. Всі правила це формалізовані знання, якими керується людина, коли визначає, особа перед ним або не обличчя. Наприклад, основні правила: області очей, носа і рота відрізняються по яскравості щодо іншої частини особи; очі на обличчі завжди розташовуються симетрично відносно один одного. Спираючись на ці та інші схожі властивості, будуються алгоритми, які в ході виконання перевіряють наявність правил на зображенні. До цієї ж групи методів відносять більш загальний метод – метод порівняння з шаблоном. У цьому методі, за допомогою опису властивостей окремих областей особи і їх заданому взаємному розташуванню визначається стандарт особи (шаблон), з яким в подальшому порівнюють вихідне зображення.

Методи на основі знань отримали досить широке поширення і мають непогані показники, однак вони показують хороші результати тільки на зображеннях з хорошим розширенням, без шумів і з нескладним фоном. На кадрах з відеопотоку або камер встановлених в громадських місцях, де можливі різні ракурси і повороти осіб, а також мінливі освітлення і велика кількість заважають об'єктів на задньому плані, існує велика ймовірність виникнення помилок.

Методи виявлення осіб за зовнішніми ознаками підходять до проблеми з іншого боку, вони не намагаються в явному вигляді формалізувати відбуваються в людському мозку процеси, а навпаки намагаються виявити закономірності і властивості зображення особи неявно, застосовуючи методи математичної статистики і машинного навчання.

Методи цієї групи позбавлені вище зазначених недоліків і тому стали частіше застосовуватися в системах відеоспостереження. Виявлення осіб в таких методах здійснюється перебором всіх прямокутних фрагментів зображення з метою визначення, до якого класу належить зображення: до класу містять особа або до класу зображень без обличчя. За рахунок такого великого об'єму роботи методи мають надмірністю і великою обчислювальною складністю.

Щоб зменшити кількість обчислень і прискорити процес відшукування осіб, автори застосовують різні додаткові методи для скорочення кількості розглянутих фрагментів. Кілька найбільш актуальних і заслуговують на увагу методів виявлення осіб розглянуті нижче:

Метод Віоли-Джос (Viola-Jones object detection). Метод був запропонований Паулом Віолою і Майклом Джонсом в 2001 році і став першим методом, демонструє високі результати при обробці зображень в реальному часу. У методу є безліч реалізацій, в тому числі в складі бібліотеки комп'ютерного зору OpenCV (функція `cvHaarDetectObjects()`).

Переваги даного методу:

- висока швидкість роботи (за рахунок використання каскадного класифікатора);
- високий точність виявлення повернутих осіб на кут до 30 градусів (якщо кут більше, ефективність даного методу сильно знижується);
- недоліки:
 - тривалий час навчання. алгоритму необхідно проаналізувати велика кількість тестових зображень;
 - при виявленні, на становище особи можуть бути встановлені обмеження.

Метод гнучкого порівняння на графах (Elastic graph matching). Метод відноситься до 2D моделювання. Його суть полягає в зіставленні графів, які описують особи (особа представляється у вигляді сітки з індивідуальним розташуванням вершин і ребер).

Процедура розпізнавання відбувається наступним чином – еталонний граф, що характеризує основний параметр розпізнавання, залишається без зміни, в той час як інші деформуються під впливом структура особи з прив'язкою до основним антропометричним точкам: відстань між очима, вухами, лінія носа, ширина губ і т.п. Чим більше цих точок використовується, тим точніше буде система розпізнавання, але і істотно збільшиться час на обробку одного об'єкта.

Недоліки методу:

- складність алгоритму розпізнавання призводить до необхідності використання значних обчислювальних потужностей;
- складна процедура введення нових шаблонів в базу даних;
- швидкодія аналітичної системи обернено пропорційно розмірами баз даних.

Приховані Марковські моделі (СММ). Метод заснований на статистичному порівнянні об'єкта з базою шаблонів. Приховані Марковські моделі використовують статистичні властивості сигналів і враховують їх просторові характеристики. Елементи моделі: початкова ймовірність станів, безліч спостережуваних станів, безліч прихованих станів, матриця перехідних ймовірностей. Кожному елементу відповідає своя Марковська модель. Під час розпізнавання людини, перевіряються всі згенеровані Марковські моделі і шукається найбільша з спостережуваних ймовірність того, що послідовність спостережень для об'єкта згенерована відповідної моделлю.

Недоліки:

- низька швидкість спрацьовування;
- низька розрізняє здатність і не оптимальний алгоритм навчання;

– система може оптимізувати тільки час обробки даних і відгуку на власну модель, але не може мінімізувати час перебору інших моделей.

Метод головних компонент (РСА). Головною метою РСА є зменшення простору ознак без значущої втрати інформації і так, щоб воно як найкраще описувало «типові» образи, що належать безлічі осіб. В задачі розпізнавання осіб його використовують головним чином для того, щоб уявити обличчя як вектор малої розмірності, який потім порівнюється з еталонними векторами з бази.

Набір власних векторів, який був отриманий один раз на навчальній вибірці, використовується для кодування інших зображень осіб, які можна уявити зваженої комбінацією власних векторів. При використанні обмеженої кількості власних векторів можна отримати стислу апроксимацію вхідному зображенню особи, яку згодом можна зберігати в БД, як вектор коефіцієнтів, який служить одночасно ключем пошуку в БД. РСА добре зарекомендував себе в додатках. Однак, тоді, коли на зображенні особи є значні зміни у виразі обличчя або освітленості, ефективність методу значно падає. Це відбувається через того, що метод головних компонент вибирає підпростір з метою максимальної апроксимації вхідного набору даних, а не з метою виконання дискримінації між класами осіб.

Метод опорних векторів (Support Vector Machines, SVM) – це набір подібних алгоритмів навчання на основі вчителів, які використовуються для завдань класифікації та регресійного аналізу [4]. Суть методу опорних векторів полягає у знаходженні гіперплощини в просторі ознак, що відокремлює клас зображень осіб від зображень «не-персон». максимум.

Переваги цього методу:

- висока стійкість до перепідготовки;
- висока швидкість в порівнянні з нейронними мережами;
- точність роботи методу поступається багатьом методам.

Нейромережеві методи. Досить поширені методи, які включають в себе близько десятка різних алгоритмів. основною особливістю таких мереж є їх здатність до навчання на наборі готових прикладів заздалегідь занесених в базу даних. Під час навчання нейронних мереж, мережа автоматично витягує ключові ознаки і будує взаємозв'язок між ними. Після цього, для того щоб розпізнати ранне невідомий об'єкт, навчена нейронна мережа застосовує отриманий досвід. Нейромережеві методи показують одні з найкращих результатів в області розпізнавання, але вважаються найбільш складними для реалізації.

Переваги даного методу:

– висока точність виявлення при правильному налаштуванні параметрів мережі.

- висока стійкість до перенавчання;

Недоліки:

Низька швидкість спрацьовування;

– складна процедура внесення змін (будь-яка зміна вимагає перенавчання мережі);

- важко формалізувати архітектуру мережі (кількість нейронів, шарів,
- характер зв'язків)
- висока обчислювальна складність.

Зведемо розглянуті методи до порівняльної таблиці 2. Методи на основі знань до зрівняння долучатись не будуть, оскільки ці методи залежать від багатьох факторів, які впливають на результат.

Таблиця 2

Порівняльна таблиця розглянутих методів

Назва методу	Швидкість роботи	Точність	Складність відтворення
Метод Віоли-Джонса	+	+	+
Метод гнучкого порівняння на графах	–	+	–
Приховані Марковські моделі	–	–	–
Метод головних компонент	+	+	–
Метод опорних векторів	+	+	–
Нейромережеві методи	+	+	–

Висновки. У цій статті було проведено аналіз алгоритмів, які застосовуються при обробці зображень. Алгоритм Віоли-Джонса є найбільш підходящим, оскільки є можливість навчання. Використовуючи цей алгоритм у системах розпізнавання об'єктів, можна змінювати характеристики об'єкта, що шукається, проводячи навчання на конкретних прикладах

СПИСОК ВИКОРИСТАНОЇ ЛІТЕРАТУРИ:

1. Путятін Є. П., Гороховатській В.О., Матат О.О. Методи та алгоритми комп'ютерного зору: Навч. посібник. Х: СМІТ, 2006. 236 с.
2. Довбиш, А.С. Основи теорії розпізнавання образів: навч. посіб.: у 2-х ч. Суми : Сумський державний університет, 2015. Ч.1. 109 с.
3. Тимошук П. В. Штучні нейронні мережі. Навчальний посібник. Львів : Видавництво Львівської політехніки. 2011. 444 с.
4. Вовк С.М., Гнатушенко В.В., Бондаренко М.В. Методи обробки зображень та компютерний зір: Навчальний посібник. Д. : ЛІРА, 2016. 148 с.

REFERENCES:

1. Putiatin, Ye. P., Horokhovatskii, V.O. & Matat, O.O. (2006). *Metody ta alhorytmy kompiuternoho zoru [Methods and algorithms of computer vision]*, Kharkiv : SMIT [in Ukrainian]
2. Dovbysh, A.S. (2015). *Osnovy teorii rozpoznavannia obraziv [Fundamentals of pattern recognition theory]*. (Vols. 1-2). Sumy: Sumskiy derzhavnyi universytet [in Ukrainian]
3. Tymoshchuk, P. V. (2011). *Shtuchni neironni merezhi [Artificial neural networks]*. Lviv : Yydavnytstvo Lvivskoi politekhniky [in Ukrainian]
4. Vovk, S.M., Hnatushenko, V.V. & Bondarenko M.V. (2016). *Metody obrobky zobrazhen ta kompiuternyi zir [Image processing methods and computer vision]*. Dnipro : LIRA [in Ukrainian].