

М. С. ГОЛИКОВ, аспірант, Харківський національний університет імені В.Н. Каразіна, м. Харків, Україна; e-mail: maksym.holikov@karazin.ua; ORCID: <https://orcid.org/0000-0003-4842-7823>

В. В. ДОНЕЦЬ, доктор філософії (PhD), викладач кафедри математичного моделювання та аналізу даних, Харківський національний університет імені В.Н. Каразіна, м. Харків, Україна; e-mail: vol.donets@gmail.com; ORCID: <https://orcid.org/0000-0002-5963-9998>

В. Є. СТІЛЕЦЬ, кандидат технічних наук (PhD), доцент кафедри комп'ютерних систем та робототехніки, Харківський національний університет імені В.Н. Каразіна, м. Харків, Україна; e-mail: viktoria.strilets@karazin.ua; ORCID: <https://orcid.org/0000-0002-2475-1496>

ІНТЕЛЕКТУАЛЬНИЙ АНАЛІЗ ОПТИЧНИХ ЗОБРАЖЕНЬ НА ОСНОВІ ГІБРИДНОГО ПІДХОДУ

У статті розглянуто інтелектуальний підхід до аналізу оптичних зображень у реальному часі, який базується на поєднанні методів розпізнавання обличчя із використанням глибокого навчання та класичних алгоритмів комп'ютерного зору для їх відстеження. Запропоновано систему аналізу відеопотоку з гібридним підходом, яка інтегрує попереднє розпізнавання обличчя на основі векторних ознак (вбудовувань), згенерованих за допомогою нейронної мережі Facenet, та трекінг обличчя за допомогою алгоритму CSRT (Channel and Spatial Reliability Tracker), що входить до складу бібліотеки OpenCV. Реалізована система дозволяє розпізнавати та автоматично ідентифікувати користувачів на відеопотоці з камери, зберігати нові обличчя у базі даних, а також ефективно відстежувати ідентифіковані обличчя протягом наступних кадрів. Алгоритм обробки кадрів реалізовано у багатопоточному режимі з використанням черг (queue) та механізмів синхронізації потоків, що дозволяє забезпечити стабільну роботу в реальному часі. Для розпізнавання невідомих осіб передбачено автоматичне створення унікального ID та додавання їхніх ознак до загальної бази вбудовувань. Особливу увагу приділено оцінці просторового перекриття зон виявлення для уникнення дублювання трекерів при одночасній присутності кількох осіб у кадрі. Система аналізу відеопотоку реалізована як вебсервіс на базі Flask, що забезпечує зручну інтеграцію з іншими програмними модулями та можливість віддаленого моніторингу через веб-інтерфейс. Запропонований гібридний підхід поєднує точність сучасних моделей глибокого навчання з гнучкістю класичних алгоритмів трекінгу, що робить систему придатною для використання в системах безпеки, розумних офісах, освітньому середовищі та в інших сферах, де важлива точна ідентифікація осіб у динамічних умовах. У підсумку, дана робота демонструє практичну реалізацію інтелектуальної системи аналізу зображень, яка може бути адаптована до різних сценаріїв використання, зокрема в системах відеонагляду, контролю доступу, управління потоками людей, а також у дослідницьких та освітніх проектах.

Ключові слова: інтелектуальний аналіз зображень, розпізнавання обличчя, OpenCV, гібридна система, відстеження об'єктів, комп'ютерний зір.

Вступ. Сучасний етап розвитку інформаційних технологій характеризується стрімким зростанням обсягів оптичних даних, що генеруються системами відеоспостереження, мобільними пристроями, розумними камерами та іншими джерелами. Одним із ключових напрямків обробки таких даних є автоматизований аналіз оптичних зображень, зокрема, розпізнавання та відстеження об'єктів, що є фундаментальними задачами для побудови систем управління динамічними системами, безпеки, контролю доступу, аналітики поведінки, розумного міського середовища тощо.

Особливої актуальності набуває задача розпізнавання та трекінгу обличчя у відеопотоці в реальному часі. Дана функціональність забезпечує основу для створення інтелектуальних систем моніторингу, які здатні не лише ідентифікувати користувачів, а й забезпечити їх динамічне відстеження у просторі без потреби в повторному розпізнаванні. Однак побудова надійної системи, здатної працювати в умовах змінного освітлення, часткових перекриттів, змін ракурсу та присутності кількох осіб одночасно, залишається складною прикладною задачею.

Існуючі підходи до розпізнавання обличчя, що базуються на глибоких нейронних мережах (наприклад, FaceNet, VGGFace, ArcFace [1]), забезпечують високу точність ідентифікації. Проте їх інтеграція в реальний відеопотік вимагає значних обчислювальних ресурсів, що ускладнює забезпечення безперервної обробки кад-

рів. З іншого боку, класичні алгоритми трекінгу, такі як CSRT, KCF або MOSSE [2], мають високу швидкість, але не забезпечують ідентифікації об'єктів, а лише відстежують рух вже знайдених. Це пояснює необхідність розробки гібридних систем, які б поєднували сильні сторони обох підходів.

У зв'язку з цим постає науково-практичне завдання, яке полягає у розробці ефективного інтелектуального підходу до аналізу відео з камер спостереження, який би поєднував точність сучасних моделей розпізнавання обличчя із ефективністю та швидкістю традиційних методів відстеження. Такий підхід повинен враховувати обмеженість ресурсів, необхідність роботи в реальному часі, динамічність сцени, а також можливість автоматичної реєстрації нових користувачів.

У роботі запропонований гібридний підхід до побудови системи аналізу відеопотоку, яка поєднує глибоку модель для генерації вбудовувань обличчя (FaceNet) з алгоритмом трекінгу CSRT. Система реалізована у вигляді веб-сервісу з використанням бібліотек OpenCV [3], DeepFace [4] та Flask [5], що забезпечує інтерактивну обробку відеопотоку у реальному часі. Гібридний підхід дозволяє досягти балансу між точністю розпізнавання та швидкістю обробки, роблячи систему придатною для широкого спектра застосувань у сфері інтелектуального відеоспостереження.

© Голіков М. С., Донець В. В., Стрілець В. Є., 2025



Дослідницька стаття: Цю статтю опубліковано видавництвом *НТУ «ХПІ»* у збірнику «Вісник Національного технічного університету «ХПІ». Серія: Системний аналіз, управління та інформаційні технології». Ця стаття поширюється за міжнародною ліцензією [Creative Commons Attribution \(CC BY 4.0\)](https://creativecommons.org/licenses/by/4.0/). **Конфлікт інтересів:** Автор/и заявив/или про відсутність конфлікту.



Опис існуючих рішень і формулювання проблеми для розв'язання. Відомі алгоритми аналізу оптичних зображень поєднують класичні методи комп'ютерного зору з глибоким навчанням, а також використовують мультимодальні та багаторівневі архітектури. Класичні методи (наприклад, фільтри Хаара, детектори особливостей SIFT/ORB, оптичний потік, алгоритми RANSAC, Kalman-фільтри тощо) надають швидкі і прозорі рішення для простих завдань, але поступаються у складності та масштабованості. Глибокі нейронні мережі (наприклад, CNN, трансформери ViT) можуть навчатися складним ознакам без ручного конструювання, але потребують великих обчислювальних ресурсів і даних. Гібридні підходи намагаються поєднати сильні сторони обох: наприклад, cascaded архітектури, де згорткові мережі поєднуються з трансформерами, дозволяють одночасно захоплювати локальні деталі та глобальні залежності [6].

Іншим прикладом є використання пірамідальних мереж (Feature Pyramid Network) для багатомасштабного аналізу. FPN-структура дозволяє визначати об'єкти різних розмірів, побудувавши топ-даун піраміду ознак всередині CNN [7].

Мультимодальні системи поєднують оптичні зображення з даними інших сенсорів. Наприклад, події камери (event-камери) разом зі звичайними кадрами забезпечують дуже високу тимчасову роздільну здатність (частоту оновлення), високу динамічну чутливість і відсутність розмиття руху [8]. Багаторівневе (каскадне) оброблення часто включає часові компоненти: наприклад, у відеоаналітиці використовують рекурентні мережі (LSTM/GRU) разом зі згортковими для врахування послідовностей кадрів. Загалом, гібридні підходи можуть включати: класичні трекари (Kalman-фільтр, Lucas-Kanade оптичний потік, ковзні вікна) разом із CNN-детекторами, каскадні архітектури «швидкий детектор + точний детектор» (наприклад, швидкий YOLO плюс повільний Faster R-CNN), а також багаторівневі багатшаблонні моделі, що зберігають різні рівні абстракції зображення [6, 9].

У роботі [10] для задач трекінгу пішоходів автори поєднали класичний алгоритм оптичного потоку з архітектурою глибокої мережі. Завдяки цьому вийшов добрий компроміс між точністю та швидкістю: модель показала MOTA ≈ 0.608 (проти 0.549 у чисто DL-системі) та скоротила час обробки приблизно вдвічі. Високу ефективність досягають за рахунок того, що потік обробляється простим алгоритмом у місцях, де його результатів достатньо, а глибока мережа запускається рідше. Таким чином, затрати обчислень суттєво знижуються, а точність залишається майже на рівні найкращих нейромережних рішень.

Алгоритм гібридного трекінгу об'єктів з подіями використовує кадри звичайної камери та дані подій з датчика DAVIS [8]. Система спочатку визначає об'єкт на зображенні, а потім використовує подійні маски для відстеження між кадрами. За рахунок подійної інформації досягається надзвичайно висока частота трекінгу до 500 Hz при звичайних 24 FPS на вході. Такий підхід використовує переваги обох типів датчиків: рух за кадром обробляється подіями (без розмиття), що дозволяє дуже швидко реагувати на динамічні зміни.

Система багатокласового трекінгу [11], що розширює простий алгоритм SORT комбінуванням класичної схеми Kalman+Hungarian із глибокими метриками подібності, отриманими через CNN. Вона навчає зовнішні ознаки (appearance) для призначення треків, що суттєво зменшує «заміни ідентичності» об'єктів у треці – на 45 % менше, ніж без глибокого компонента. DeepSORT добре показує себе в реальному часі з високими FPS, оскільки більша частина важкого навчання відбувається офлайн, а онлайн-алгоритм – відносно легкий.

Гібридні підходи і рішення об'єднують сильні сторони класичних алгоритмів і глибоких мереж, що дає низку переваг, але водночас накладає свої обмеження.

Однією з головних переваг гібридних рішень є висока точність, якої вдається досягти завдяки використанню глибоких нейронних мереж. Ці моделі здатні вивчати й розпізнавати складні ознаки, що робить їх надзвичайно ефективними для задач класифікації, сегментації чи відстеження об'єктів. Водночас, щоб зменшити обчислювальні витрати, гібридні підходи часто поєднують легкі класичні етапи попередньої обробки з менш частим запуском обчислювально важких мереж. Це дозволяє досягти обробки відео в реальному часі без значної втрати якості.

Класичні методи, зокрема фільтрація, виявлення контурів чи оцінка оптичного потоку, мають ще одну важливу перевагу: стійкість до окремих типів шуму або несприятливих умов, наприклад, слабкого освітлення. Саме тому їх використання у системах на основі гібридних підходів дозволяє доповнити слабкі сторони глибоких моделей. До того ж гібридні підходи легко адаптуються до різних типів сенсорів. Наприклад, поєднання із тепловізорами RGB-камер чи LiDAR дає змогу значно підвищити якість роботи систем у нічних або складних погодних умовах.

Втім, попри всі переваги, системи з гібридними підходами стикаються з низкою серйозних викликів. Одним з основних є потреба глибоких компонентів у великій кількості анотованих даних для навчання. Це особливо актуально для нових типів сенсорів, як-от камери чи тепловізори, для яких публічно доступних датасетів досить мало. У таких умовах системи можуть погано узагальнювати нові об'єкти або сцени, що призводить до помилок.

Ще одним важливим викликом є складність інтеграції. Об'єднання різних алгоритмів потребує тонкого налаштування великої кількості гіперпараметрів: порогів, ваг, частоти оновлення компонентів тощо. Це суттєво ускладнює розробку, тестування та підтримку системи, особливо в умовах, де потрібна стабільна робота в реальному часі.

Не менш важливою є й проблема обчислювальної складності. Навіть якщо класичні методи є легкими, глибокі моделі з десятками мільйонів параметрів потребують потужних обчислювальних ресурсів. У випадках, коли система має працювати на вбудованому або енергообмеженому обладнанні, доводиться спрощувати архітектуру, що часто веде до компромісу між точністю та швидкістю.

Особливості деяких сенсорів також можуть створювати труднощі. Наприклад, подієві камери не фіксують абсолютну яскравість сцени, що ускладнює класифікацію або відстеження об'єктів. Крім того, системи з гібридними підходами можуть ставати вразливими до паразитних явищ – спалахів світла або змін форми об'єктів, які не були представлені у навчальних даних.

Загалом, розробка ефективної гібридної системи є завжди мистецтвом балансу. Додавання швидкої попередньої обробки, такої як оптичний потік, може пришвидшити роботу, але дещо знизити точність. Аналогічно, спрощення глибинної моделі для зменшення обчислень може призвести до втрати важливих деталей. Тому налаштування таких систем часто здійснюється емпірично, з урахуванням конкретного завдання та апаратних обмежень.

Мета та задачі дослідження.

Метою даного дослідження є вдосконалення гібридного підходу до інтелектуального аналізу оптичних зображень, що забезпечує точне розпізнавання та стабільне відстеження об'єктів (зокрема обличч) у відеопотоці в режимі реального часу при обмежених обчислювальних ресурсах. Запропонований підхід має поєднувати переваги глибоких нейронних мереж (висока точність і здатність до узагальнення) із легкістю та швидкодією класичних трекінг-алгоритмів, забезпечуючи адаптацію до змін середовища, зміни ракурсів, часткового перекриття і динамічних сцен.

Для досягнення цієї мети були сформульовані такі основні задачі дослідження:

1) вдосконалити систему аналізу відеопотоку із застосуванням гібридного підходу, що дозволить використовувати розпізнавання лише при появі нових об'єктів або втраті треків, тоді як основне відстеження виконуватиметься класичним трекером;

2) реалізувати модель системи, що обробляє відеопотік, виконує розпізнавання обличч, асоціює вбудовування з ID-треками та зберігає інформацію про об'єкти.

Гібридний підхід до розпізнавання та відстеження обличч. Запропонований підхід базується на інтеграції сучасних технологій глибокого навчання для розпізнавання обличч з класичними алгоритмами трекінгу для відстеження, які реалізовані за допомогою мови програмування Python [12]. Такий гібридний підхід дозволяє мінімізувати обчислювальні витрати шляхом розділення задачі на два основних компоненти: ідентифікацію об'єкта (обличчя), яка виконується періодично або за необхідності, та відстеження вже розпізнаного об'єкта у відеопотоці за допомогою швидких класичних трекерів.

У якості механізму розпізнавання використовується глибока нейронна мережа FaceNet, яка генерує векторні представлення обличч, а саме вбудовування. Мережа перетворює зображення обличчя у простір ознак, де близькі за візуальними характеристиками обличчя мають близькі вбудовування, що дозволяє застосовувати евклідову або косинусну метрику для зіставлення обличч з базою вже відомих представлень. Це забезпечує високу точність ідентифікації навіть у складних умовах, а саме при зміні освітлення, ракурсу чи наявності часткового перекриття.

Для трекінгу обличч між кадрами використовується класичний алгоритм відстеження об'єктів, а саме CSRT (Channel and Spatial Reliability Tracking). Він базується на кореляційних фільтрах, які ефективно оновлюють модель об'єкта при кожному кадрі. CSRT добре справляється зі зміною масштабу, фоновими перешкодами та частковим перекриттям. Найбільшою перевагою CSRT є швидкодія: на сучасних середовищах обробки відео він здатен відстежувати кілька об'єктів із частотою понад 30 кадрів за секунду.

Гібридний підхід працює за таким принципом: при початковому виявленні обличчя у кадрі запускається процес розпізнавання через модель FaceNet. Створюється унікальний ID-трек обличчя з прив'язкою до його вбудовування. Після цього активується трекер CSRT, який веде об'єкт протягом наступних кадрів без потреби у повторному розпізнаванні. Якщо трекер втрачає об'єкт (наприклад, через вихід за межі кадру або сильне перекриття), ініціюється повторне розпізнавання через появу нового об'єкта, або при поверненні вже відомого.

Додатково реалізовано механізм асоціації треків: при появі нового обличчя система порівнює вбудовування з наявними у базі, і якщо знаходить подібний, асоціює об'єкт з існуючим ID. Цей процес зображено на рис. 1.

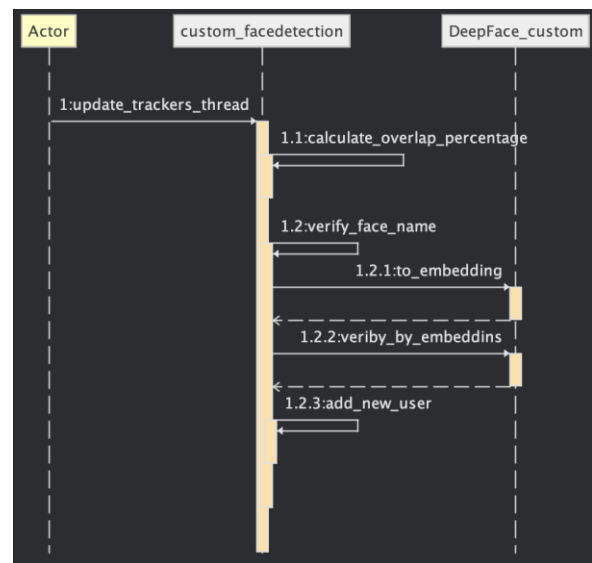


Рис. 1. Діаграма послідовності процесу оновлення та верифікації трекерів

Механізм асоціації треків дозволяє уникнути дублювання особистостей у трекінгу та зберігати цілісну історію переміщення. Для зберігання вбудовувань і треків використовується in-memo (внутрішня) база, що дозволяє швидко звертатися до ознак обличч.

Таким чином, інтелектуальний аналіз здійснюється лише в критичні моменти, а точніше при першій появі об'єкта або втраті треку. Але більшу частину часу система працює в режимі швидкого трекінгу. Це дозволяє досягти обробки відеопотоку в реальному часі, знижуючи навантаження на обчислювальні ресурси, зберігаючи при цьому високу точність та стійкість до змін середовища.

На рис. 2 представлено покроковий процес розпізнавання та відстеження обличчя у відеопотоці. Він починається з надання вхідного кадру, продовжується виявленням обличчя, їх розпізнаванням за допомогою FaceNet, ініціалізацією трекера CSRT та оновленням відстеження.

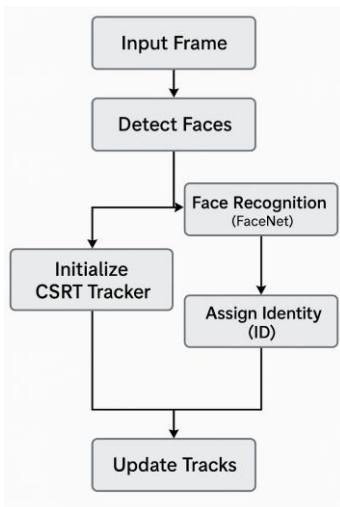


Рис. 2. Покроковий процес виявлення та відстеження обличчя

Впроваджений гібридний підхід легко масштабувати: кількість одночасно відстежуваних обличчя обмежена лише апаратними ресурсами. Водночас система зберігає адаптивність: наприклад, можна змінювати пороги для повторного запуску розпізнавання або комбінувати кілька трекерів для різних умов (CSRT, KCF, MOSSE). У перспективі можливе розширення функціоналу, а саме визначення емоцій, віку, статі, а також аналітики групової поведінки на основі ідентифікованих треків.

Висновки. У цьому дослідженні було розглянуто та реалізовано гібридний підхід до інтелектуального аналізу оптичних зображень, зокрема для задачі розпізнавання та відстеження обличчя у відеопотоці. Поєднання методів глибокого навчання з класичними алгоритмами комп'ютерного зору дозволило досягти балансу між точністю, продуктивністю та обчислювальною ефективністю, що є критично важливим у системах реального часу з обмеженими ресурсами.

Розроблена система аналізу відеопотоку успішно використовує попереднє розпізнавання обличчя за допомогою моделі FaceNet лише в моменти появи нових об'єктів або втрати треків, а для основного трекінгу застосовує класичний алгоритм CSRT. Така структура системи дозволяє уникнути надмірного навантаження на систему, зберігаючи при цьому високу точність асоціації об'єктів у кадрах.

Гібридний підхід показав високу гнучкість та здатність до масштабування. Його можна адаптувати для інших типів об'єктів або використовувати в комбінації з додатковими сенсорами (наприклад, інфрачервоними або подієвими камерами), що розширює сферу застосування: від систем відеоспостереження до управління динамічними системами.

Разом з тим, інтеграція глибоких моделей та класичних алгоритмів вимагає тонкого налаштування та

ретельного тестування. Складність налаштувань, обмежена доступність анотованих даних і апаратні обмеження залишаються відкритими викликами для практичного впровадження таких рішень.

Список використаної літератури

1. Jiachen Yan, Guang Yang, Weihong Li et al. A Credibility Scoring Algorithm to match surveillance video target and UWB tag. *Research Square*. 2024. DOI: <https://doi.org/10.21203/rs.3.rs-4015518/v1>.
2. Muhammad Indra Ardiawan, Gede Putra Kusuma Negarara. Comparative Analysis of FaceNet, VGGFace, and GhostFaceNets Face Recognition Algorithms For Potential Criminal Suspect Identification. *J. Appl. Artif. Intell.* 2024, vol. 5, no. 2, pp. 34–49. DOI: 10.48185/jaai.v5i2.1237.
3. Kumar S. B., Raju V. S., Maheswari U. V. OpenCV libraries for computer vision. *Computer Vision: Applications of Visual AI and Image Processing*. Boston: De Gruyter. 2023, pp. 1–22.
4. Poorni R., Charulatha S., Amritha B., Bhavyashree P. Real-time face detection and recognition using deepface convolutional neural network, *The Electrochemical Society*. 2022, vol. 107 (5091).
5. Ashley D. *Foundation Dynamic Web Pages with Python*. Apress: Austin, TX, USA, 2020, p. 213.
6. Goutam Yelluru Gopal, Maria A Amer. Separable self and mixed attention transformers for efficient object tracking. *Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision*. 2024, pp. 6708–6717.
7. Xie J., Pang Y. W., Nie, J., Cao, J., Han, J. G. Latent Feature Pyramid Network for Object Detection. *IEEE Trans. Multimed.* 2023, vol. 25, pp. 2153–2163.
8. Zaid El Shair, Samir A. Rawashdeh. High Speed Hybrid Object Tracking Algorithm using Image and Event Data. *TechRxiv*, 2021.
9. Lin T.-Y., Dollár P., Girshick R., He K., Hariharan B., Belongie S. Feature pyramid networks for object detection. *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2017, pp. 936–944.
10. Scarrica V. M., Panariello C., Ferone A., Staiano A. A hybrid approach to Real-Time Multi-Target Tracking. *ArXiv. Computer Science*. 2023. DOI: 10.48550/arXiv.2308.01248.
11. Wojke N., Bewley A., Paulus D. Simple online and realtime tracking with a deep association metric. *ICIP. IEEE*, 2017, pp. 3645–3649.
12. Sharma V. K., Kumar V., Sharma S., et al. *Python Programming: A Practical Approach*. CRC Press. 2021 :213–5.

References (transliterated)

1. Jiachen Yan, Guang Yang, Weihong Li et al. A Credibility Scoring Algorithm to match surveillance video target and UWB tag. *Research Square*. 2024. DOI: <https://doi.org/10.21203/rs.3.rs-4015518/v1>.
2. Muhammad Indra Ardiawan, Gede Putra Kusuma Negarara. Comparative Analysis of FaceNet, VGGFace, and GhostFaceNets Face Recognition Algorithms For Potential Criminal Suspect Identification. *J. Appl. Artif. Intell.* 2024, vol. 5, no. 2, pp. 34–49. DOI: 10.48185/jaai.v5i2.1237.
3. Kumar S. B., Raju V. S., Maheswari U. V. OpenCV libraries for computer vision. *Computer Vision: Applications of Visual AI and Image Processing*. Boston: De Gruyter. 2023, pp. 1–22.
4. Poorni R., Charulatha S., Amritha B., Bhavyashree P. Real-time face detection and recognition using deepface convolutional neural network, *The Electrochemical Society*. 2022, vol. 107 (5091).
5. Ashley D. *Foundation Dynamic Web Pages with Python*. Apress: Austin, TX, USA, 2020, p. 213.
6. Goutam Yelluru Gopal, Maria A Amer. Separable self and mixed attention transformers for efficient object tracking. *Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision*. 2024, pp. 6708–6717.
7. Xie J., Pang Y. W., Nie, J., Cao, J., Han, J. G. Latent Feature Pyramid Network for Object Detection. *IEEE Trans. Multimed.* 2023, vol. 25, pp. 2153–2163.
8. Zaid El Shair, Samir A. Rawashdeh. High Speed Hybrid Object Tracking Algorithm using Image and Event Data. *TechRxiv*, 2021.
9. Lin T.-Y., Dollár P., Girshick R., He K., Hariharan B., Belongie S. Feature pyramid networks for object detection. *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2017, pp. 936–944.

10. Scarrica V. M., Panariello C., Ferone A., Staiano A. A hybrid approach to Real-Time Multi-Target Tracking. *ArXiv. Computer Science*. 2023. DOI: 10.48550/arXiv.2308.01248.
11. Wojke N., Bewley A., Paulus D. Simple online and realtime tracking with a deep association metric. *ICIP. IEEE*, 2017, pp. 3645–3649.
12. Sharma V. K., Kumar V., Sharma S., et al. *Python Programming: A Practical Approach*. CRC Press. 2021 :213–5.

Надійшла (received) 05.05.2025

UDC 004.8

M. S. HOLIKOV, Postgraduate student, V. N. Karazin Kharkiv National University, Kharkiv, Ukraine; e-mail: maksym.holikov@karazin.ua; ORCID: <https://orcid.org/0000-0003-4842-7823>

V. V. DONETS, Doctor of Philosophy (PhD), V. N. Karazin Kharkiv National University, Lecturer at the Department of Mathematical Modelling and Data Analysis, Kharkiv, Ukraine; e mail: vol.donets@gmail.com; ORCID: <https://orcid.org/0000-0002-5963-9998>

V. Y. STRILETS, Candidate of Technical Sciences (PhD), V. N. Karazin Kharkiv National University, Associate Professor at the Department of Computer Systems and Robotics, Kharkiv, Ukraine; e mail: viktoria.strilets@karazin.ua; ORCID: <https://orcid.org/0000-0002-2475-1496>

INTELLIGENT ANALYSIS OF OPTICAL IMAGES BASED ON A HYBRID APPROACH

The article considers an intelligent approach to real-time analysis of optical images based on a combination of face recognition methods using deep learning and classical computer vision algorithms for tracking them. A system with hybrid approach is proposed that integrates preliminary face recognition based on vector features (embeddings) generated by the FaceNet neural network and face tracking using the CSRT (Channel and Spatial Reliability Tracker) algorithm, which is part of the OpenCV library. The implemented system allows to recognise and automatically identify users in a video stream from a webcam, store new faces in the database, and effectively track identified faces over subsequent frames. The frame processing algorithm is implemented in a multi-threaded mode using queues and thread synchronisation mechanisms to ensure stable operation in real time. To recognise unknown persons, a unique ID is automatically created and their features are added to the common database of emblems. Particular attention is paid to assessing the spatial overlap of detection zones to avoid duplication of trackers when several people are present in the frame at the same time. In addition, the system is implemented as a web service based on Flask, which provides convenient integration with other software modules and the possibility of remote monitoring via a web interface. The proposed hybrid approach combines the accuracy of modern deep learning models with the flexibility of classical tracking algorithms, making the system suitable for use in security systems, smart offices, educational environments, and other areas where accurate face identification in dynamic environments is important. In summary, this paper demonstrates the practical implementation of an intelligent image analysis system that can be adapted to various use cases, including video surveillance, access control, and crowd management systems, as well as research and educational projects.

Keywords: intelligent image analysis, face recognition, OpenCV, hybrid system, object tracking, computer vision.

Повні імена авторів / Author's full names

Автор 1 / Author 1: Голиков Максим Сергійович / Holikov Maksym Serhiiovych

Автор 2 / Author 2: Донець Володимир Віталійович / Donets Volodymyr Vitaliiovych

Автор 3 / Author 3: Стрілець Вікторія Євгенівна / Strilets Viktoriia Yevhenivna