

## Стан та розвиток технологій повторної ідентифікації людини в інтелектуальних відеосистемах

А. С. Колпіков, Е. М. Забарна, М. В. Лобачев  
Національний університет «Одеська політехніка»

**Анотація.** В статті розглянуто сучасний стан і основні напрями розвитку технологій повторної ідентифікації людини (re-ID) в інтелектуальних системах відеоспостереження. Визначено формальну постановку задачі повторної ідентифікації та охарактеризовано її роль у забезпеченні безперервного автоматичного моніторингу об'єктів між неперетинними камерами. Проаналізовано основні виклики, що ускладнюють вирішення цієї задачі в реальних умовах, серед яких – варіативність зовнішнього вигляду, зміни ракурсу та освітлення, часткове перекриття, перенос знань між доменами (domain shift), обмеженість навчальних даних та апаратних ресурсів. В роботі узагальнено підходи до вилучення ознак, зіставлення зображень та методів глибинного навчання, які використовуються для побудови ефективних re-ID моделей. Особливу увагу приділено сучасним тенденціям у використанні генеративних змагальних мереж (GAN) та attention-механізмів (механізмів уваги), що сприяють покращенню точності ідентифікації. Наведено перспективні напрями подальших досліджень, зокрема у сфері мультимодального навчання, адаптації до нових доменів та етичного впровадження технологій у публічні середовища.

**Ключові слова:** повторна ідентифікація людини, інтелектуальні відеосистеми, комп'ютерний зір, глибоке навчання, ознаковий простір, метричне навчання, генеративні змагальні мережі, attention-механізм (механізм уваги).

### Вступ

У контексті зростаючого попиту на системи забезпечення громадської безпеки, автоматизований моніторинг публічних просторів набуває дедалі більшого значення. Інтелектуальні відеосистеми, оснащені засобами комп'ютерного зору та штучного інтелекту, відкривають нові можливості для автоматичного аналізу поведінки об'єктів у реальному часі. Однією з ключових задач, що постає в рамках таких систем, є повторна ідентифікація людини – процес встановлення відповідності між візуальними зображеннями однієї й тієї ж особи, зафіксованими різними камерами або в різні моменти часу [1].

На відміну від класичного розпізнавання обличчя, повторна ідентифікація передбачає ідентифікацію особи в умовах, коли її зовнішній вигляд може змінюватися, а зображення містять лише загальний вигляд постаті, без чітко видимих біометричних рисунків.

Актуальність досліджень у цій сфері зумовлена поширенням систем відеоспостереження у

містах, транспортних вузлах, торгових центрах, аеропортах та інших об'єктах критичної інфраструктури. У таких умовах необхідність автоматизованого та надійного встановлення ідентичності особи між неперетинними камерами стає ключовим чинником ефективності роботи відеоаналітичних систем. Повторна ідентифікація дозволяє не лише відстежувати переміщення людей у просторі, а й сприяє розширенню функціональності систем безпеки, виявленню аномальної поведінки, аналізу потоків відвідувачів та інтеграції з іншими аналітичними модулями [2].

Разом з тим, розв'язання задачі повторної ідентифікації є складним технічним завданням, що супроводжується низкою суттєвих викликів. До основних проблем належать варіативність зовнішнього вигляду людини внаслідок змін одягу, ракурсу, освітлення, якості зображення, а також складність умов зйомки у різних камерах і наявність шумових факторів у фоновому середовищі.

Крім того, ідентифікація вимагає стійкості моделей до часткових оклюзій, різного масштабу зображень та обмеженої кількості прикладів для навчання (рис. 1) [3]. Особливу складність становить перенесення моделей, навчених на одних

© Колпіков А. С., Забарна Е. М.,  
Лобачев М. В., 2025

даних, у нові середовища, де умови зйомки значно відрізняються.

У відповідь на ці виклики сучасна наукова спільнота активно розробляє нові алгоритмічні рішення, засновані на методах глибокого навчання, зокрема згорткових нейронних мереж, attention-механізмів, трансформерних архітектур та навчання на подібності ознак (metric learning). Важливу роль відіграє також створення великих відкритих датасетів, що дозволяють дослідникам

розробляти та тестувати моделі в умовах, наближених до реальних.

Метою цієї статті є всебічний аналіз сучасного стану технологій повторної ідентифікації людини в системах відеоспостереження.

У роботі розглядаються ключові підходи до розв'язання задачі re-ID, аналізуються основні труднощі її практичної реалізації, порівнюються найбільш поширені методи та моделі, а також окреслюються перспективні напрями подальших досліджень у цій галузі.



Рис.1. Кілька прикладів зображень з навчального набору даних ReID780. Кожен рядок відповідає різним зображенням для однієї особи. В якості прикладу можна навести рядок зображення мають різну орієнтацію та позу відповідної людини, а також різні умови освітлення та фони [3]

## 1. Постановка задачі

Задача повторної ідентифікації людини (*person re-identification*, re-ID) полягає у встановленні того, чи належать два або більше зобра-

жень, отриманих з різних камер спостереження або в різні моменти часу, одній і тій самій особі.

В рамках інтелектуальних систем відеоспостереження ця задача є ключовим компонентом, що забезпечує безперервне відстеження об'єктів

між неперетинними камерами без ручної участі оператора.

Нехай задано множину відеокамер  $\{C_1, C_2, \dots, C_K\}$ , які генерують зображення або відео потоки, в яких виявляються об'єкти (люди). Після етапу детекції з кожного кадру виділяються фрагменти зображень, що містять окремих осіб. Позначимо набір зображень людей як  $\mathcal{I} = \{I_1, I_2, \dots, I_N\}$ , де кожне зображення  $I_i$  містить зону кадру з однією особою.

Кожному зображенню  $I_i$  відповідає вектор ознак  $f_i \in \mathbb{R}^d$ , отриманий за допомогою функції екстракції ознак  $f(\cdot)$ , яка реалізується, як правило, глибокою нейронною мережею:

$$f_i = f(I_i; \theta), \quad (1)$$

де  $\theta$  – параметри мережі, що підлягають навчанню.

Мета повторної ідентифікації полягає в тому, щоб для заданого запиту (query image)  $I_q$  знайти серед галереї зображень  $G = \{I_1, \dots, I_M\}$  всі ті, що відповідають тій же особі, що й  $I_q$ . Задача формулюється як задача пошуку найбільш схожих зображень у ознаковому просторі:

$$\hat{I} = \arg \min_{I_j \in G} d(f_q, f_j), \quad (2)$$

де  $f_q = f(I_q)$ ,  $f_j = f(I_j)$ , а  $d(\cdot, \cdot)$  – функція відстані, наприклад, евклідова відстань:

$$d(f_q, f_j) = \|f_q - f_j\|_2, \quad (3)$$

або косинусна відстань:

$$d_{cos}(f_q, f_j) = 1 - \frac{f_q^T f_j}{\|f_q\| \|f_j\|}. \quad (4)$$

Для навчання моделі використовують або підхід класифікації, де кожному класу відповідає окрема особа, або метричне навчання (metric learning), яке націлене на формування такого ознакового простору, в якому зображення однієї й тієї ж особи знаходяться близько одне до одного, а зображення різних осіб – далеко. Один із найпоширеніших підходів – використання функції триплетних втрат:

$$\mathcal{L}_{triplet} = \max(0, d(f_a, f_p) - d(f_a, f_n) + \alpha), \quad (5)$$

де  $f_a$ ,  $f_p$ ,  $f_n$  – вектори ознак для відповідно анкори (anchor), позитивного прикладу (той самий клас) і негативного прикладу (інший клас), а  $\alpha$  –

пороговий параметр (margin), що визначає мінімальну бажану відстань між класами.

В альтернативних підходах також використовується контрастивна втрата, класифікаційна втрата на softmax-шарі, або комбіновані функції втрат, які водночас оптимізують якість розподілу ознак і здатність класифікувати об'єкти.

У загальному випадку, задача повторної ідентифікації розглядається як задача пошуку за зразком (retrieval), яка вимагає:

- формування дискримінативного ознакового простору;
- використання ефективної метрики схожості;
- забезпечення узагальнюваності моделей на нові дані (переносимість).

Таким чином, з формальної точки зору, задача повторної ідентифікації полягає у побудові функції  $f: \mathcal{I} \rightarrow \mathbb{R}^d$ , яка мінімізує внутрішньокласову відстань:

$$\forall (I_i, I_j) \in \text{SameID}: d(f(I_i), f(I_j)) \rightarrow \min, \quad (6)$$

та максимізує міжкласову:

$$\forall (I_i, I_k) \in \text{DifID}: d(f(I_i), f(I_k)) \rightarrow \max, \quad (7)$$

В реальних системах до цього формулювання додаються додаткові обмеження: реальний час обробки, обмеженість пам'яті, змінність умов зйомки, відсутність маркованих даних у нових середовищах.

Це зумовлює актуальність розробки ефективних і стійких моделей, здатних працювати в умовах невизначеності та складних вхідних даних.

## 2. Виклики та перешкоди задачі повторної ідентифікації людини

Розв'язання задачі повторної ідентифікації людини в умовах реального середовища супроводжується низкою суттєвих технічних і прикладних викликів. Серед основних з них можна виокремити наступні [4]:

- висока варіативність зовнішнього вигляду людини. Одна й та сама особа може мати різний вигляд у різні моменти часу: змінюється одяг, взуття, зачіска, наявність або відсутність аксесуарів (наприклад, сумки, капелюхи, маски). Модель повинна бути здатною фокусуватися на сталих індивідуальних рисах, зберігаючи стійкість до несуттєвих змін зовнішності;

– *зміна умов зйомки в різних камерах.* Камери можуть відрізнятися кутом огляду, якістю зображення, рівнем освітлення, кольоровим балансом і фоновим шумом. Це ускладнює безпосереднє порівняння зображень і знижує якість обчислення схожості ознак;

– *часткове перекриття тіла або неповне входження в кадр.* Людина може бути частково закрита іншими об'єктами або потрапити в кадр лише частково. Це знижує ефективність виділення ознак і потребує застосування методів, здатних працювати з неповною інформацією;

– *схожість зовнішнього вигляду різних осіб.* У публічних місцях багато людей можуть мати подібні силуети або одяг. Це підвищує ризик хибних відповідей між різними особами та потребує побудови більш дискримінативного ознакового простору;

– *проблема переносу знань між різними доменами (domain shift).* Модель, навчена на одному датасеті, часто демонструє погану продуктивність на нових даних із відмінними характеристиками. Необхідне застосування методів доменної адаптації або навчання без учителя;

– *відсутність чітких біометричних ознак.* У більшості випадків обличчя особи або повністю відсутнє в кадрі, або має недостатню роздільну здатність для використання класичних методів біометричної ідентифікації. Це вимагає зосередження на непрямих ознаках – кольорах, формах, ході;

– *високі обчислювальні витрати.* Для обробки відеопотоків з багатьох камер у реальному часі потрібні обчислювальні ефективні алгоритми. Великі моделі можуть бути точними, але непридатними для застосування у ресурсно обмежених системах або на вбудованих пристроях;

– *етичні та правові обмеження.* Застосування технологій повторної ідентифікації пов'язане з ризиками порушення приватності, контролю за громадянами без їхньої згоди та зловживанням аналітичними можливостями. Це викликає необхідність забезпечення прозорості, регулювання та дотримання норм захисту персональних даних.

Таким чином, задача повторної ідентифікації людини виходить за межі суто технічної проблеми. Вона передбачає розв'язання комплексу питань – від інженерних і алгоритмічних до етичних і соціальних.

Успішне подолання вказаних викликів є запорукою ефективною інтеграції re-ID у повноцінні інтелектуальні системи відеоспостереження нового покоління.

### 3. Огляд підходів до вирішення задачі повторної ідентифікації людини

Підходи до повторної ідентифікації особи, як правило, використовують дескриптор зовнішності для представлення особи та функції зіставлення для порівняння цих дескрипторів зовнішності.

Протягом багатьох років робилися спроби покращити як представлення, так і алгоритм зіставлення, щоб підвищити стійкість до варіацій пози, освітлення та фону, притаманних цій проблемі.

Нещодавно, для повторної ідентифікації людини були застосовані підходи глибокого навчання, що дозволило досягти найсучасніших результатів [5].

Підходи глибокого навчання для ідентифікації особи спільно вивчають візуальні представлення ознак і метрику подібності. У літературі розпізнавання осіб здебільшого виконується за зовнішнім виглядом людини на одноколірних зображеннях [6].

Однак, багато підходів використовують інші ознаки для виконання завдання ідентифікації особи, такі як часова інформація, зображення глибини, хода, топологія камери тощо [7].

#### 3.1 Підходи до вилучення ознак

Вилучення ознак – це фундаментальний крок у повторній ідентифікації особи, спрямований на створення надійних і дискримінаційних візуальних зображень людей. У даній статті традиційні методи поділяються на три основні підходи: холістичний, патч-орієнтований і виділення ознак на основі частин тіла [8].

Холістичні підходи часто використовують глобальні кольорові гістограми для опису зовнішності людей, а методи фотометричної нормалізації (наприклад, вирівнювання гістограми або функції передачі яскравості) застосовують для пом'якшення впливу освітлення та варіацій камери [9].

Підходи на основі патчів усувають обмеження цілісних ознак шляхом щільної вибірки локальних областей (наприклад, патчі 10×10) і вилучення дескрипторів, таких як гістограми LAB/HSV, SIFT, LBP або градієнтні ознаки [10]. Ці ознаки часто агрегуються за допомогою таких методів, як Bag-of-Words або локальне кодування координат для формування компактних, але інформативних зображень.

Підходи на основі частин тіла підвищують просторову стійкість, використовуючи структуру

людського тіла. Такі методи, як сегментація на основі симетрії (наприклад, SDALF) або горизонтальний поділ на смуги, спрямовані на вирівнювання та кодування семантично значущих областей, таких як тулуб або ноги [11]. Деякі методи поєднують ознаки на основі смуг і плям для покращення дискримінаційної здатності [12].

Додаткові стратегії включають використання коваріаційних дескрипторів, моделювання регіонів на основі Гаусса та вибір ознак на основі бустінгу [13].

Незважаючи на різноманітність цих методів, всі вони спрямовані на виявлення інваріантних і відмінних патернів у різних ракурсах камери для покращення ефективності повторної ідентифікації.

### 3.2 Підходи на базі зіставлення

Підходи на базі зіставлення при повторній ідентифікації особи зосереджені на вимірюванні схожості між представленнями ознак, витягнутими з різних зображень [14]. Оскільки традиційні метрики відстані, такі як евклідова відстань або відстань Бгаттачар'я, часто є недостатніми для обробки великих внутрішньокласових варіацій (наприклад, через зміну точки зору або освітлення), для підвищення надійності та дискримінаційної здатності застосовуються спеціалізовані методи зіставлення [15].

Важливим напрямком є метричне навчання, метою якого є вивчення функції відстані, пристосованої до завдання повторної ідентифікації людини. Більшість методів використовують махаланобісову відстань, параметризовану позитивною напіввизначеною (PSD) матрицею [16]. Безпосередня оптимізація цієї матриці може бути обчислювальною трудомісткою, тому багато підходів розкладають її на підпростори меншої розмірності або застосовують обмеження для збереження опуклості. Методи метричного навчання можна умовно поділити на два типи [17]:

- оптимізація на основі обмежень на відстані, де цільова функція призначена для зближення схожих пар і розсування несхожих пар. Такі методи, як KISSME (Keep It Simple and Straightforward Metric), забезпечують ефективні рішення закритої форми, засновані на гауссівських припущеннях про простір ознак;

- дискримінативне навчання підпростору, метою якого є вивчення проєкції, де внутрішньокласова дисперсія мінімізується, а міжкласова максимізується. Приклади включають методи на основі LDA, LFDA (локальний дискримінантний аналіз Фішера), XQDA (який інтегрує QDA з KISSME) і NFST (нульове перетворення Фолі-

Саммона), яке проєктує зразки одного класу на спільну точку, максимізуючи при цьому міжкласову відокремлюваність [17].

Крім того, для покращення масштабованості та точності, були застосовані ансамблеві методи. Наприклад, ансамбль RankSVM формулює повторну ідентифікацію як задачу ранжування та агрегує декілька слабких SVM для зменшення обчислювальних витрат [18].

Загалом, підходи зіставлення доповнюють екстракцію ознак, уточнюючи спосіб вимірювання схожості, що відіграє життєво важливу роль у підвищенні точності ідентифікації в складних сценаріях перехресного ракурсу.

### 3.3 Підходи на базі глибинного навчання

Глибоке навчання значно вдосконалило повторну ідентифікацію людини, уможлививши спільне навчання представлень ознак і метрик схожості в наскрізний спосіб. На відміну від традиційних ручних функцій, згорткові нейронні мережі (CNN) автоматично вивчають дискримінаційні ознаки з необроблених даних зображення, що призводить до більш надійної роботи при зміні пози, освітлення та фону.

Підходи, засновані на глибокому навчанні, часто використовують структурні пріоритети, що застосовуються в класичних методах – наприклад, архітектури на основі смуг, патчів або частин тіла [19]. Моделі на основі смуг ділять вхідні зображення на горизонтальні області, припускаючи грубе вирівнювання між зображеннями людей. Наприклад, сіамські мережі зі спільними параметрами можуть порівнювати смуги незалежно один від одного, тоді як механізми вентилявання покращують відповідні характеристики на рівні смуг під час зіставлення [20].

Методи на основі частин тіла використовують інформацію про позу для вилучення та об'єднання ознак із семантично значущих областей (наприклад, голови, тулуба, ніг) [21]. Деякі моделі використовують мережі пропозицій областей тіла або просторові трансформантні мережі (STN) для локалізації частин тіла пішоходів без явних анотацій [22]. Інші використовують механізми уваги для автоматичного навчання дискримінативних областей тіла під час тренування, використовуючи кілька гілок мережі або вагові маски [23].

Щодо функцій втрат, глибокі моделі використовують парні втрати (наприклад, контрастні втрати), потрібні втрати або їхні комбінації (наприклад, квадруплетні втрати) для оптимізації схожості ідентифікації. Ці втрати покращують узагальнення моделі, зменшуючи внутрішньо-

класову дисперсію та збільшуючи міжкласову відокремлюваність у вивченому просторі вбудування [24].

Останні тенденції включають генеративні змагальні мережі (Generative Adversarial Networks, GAN) для доповнення навчальних даних або зменшення розривів між наборами даних, а також механізми уваги для вибіркового виділення важливих областей або кадрів (при повторній ідентифікації на основі відеозаписів) [25]. Ці вдосконалення вирішують такі проблеми, як дефіцит даних і варіації точок зору, ще більше підвищуючи продуктивність.

Таким чином, підходи глибокого навчання революціонізували вилучення та співставлення ознак при повторній ідентифікації особи, пропонуючи масштабовані, адаптивні та високодискримінаційні репрезентації за допомогою наскрізних навчальних фреймворків.

### 3.4 Інші підходи

Хоча більшість методів повторної ідентифікації людини покладаються на зовнішній вигляд RGB-зображення, інші ознаки можуть надати додаткову або більш надійну інформацію, особливо в складних умовах. У цьому пункті обговорюються альтернативні способи та контекстна інформація, які підвищують ефективність повторної ідентифікації.

Часова інформація, яку зазвичай отримують з відеопослідовностей, фіксує патерни руху, такі як хода. Ранні роботи використовували відео для віднімання фону або сегментації частин, тоді як більш сучасні методи агрегують ознаки на рівні кадру за допомогою стратегій об'єднання або вимірюють мінімальну відстань між послідовностями [26]. Детектори на основі ходи, такі як Gait Energy Images (GEI) або Flow Energy Profiles (FEP), використовують динаміку ходи, яка часто є унікальною для кожної людини. Глибокі рекурентні моделі (наприклад, RNN або LSTM) також використовуються для кодування часових залежностей для представлення на рівні послідовностей [27].

Біометричні ознаки – такі як риси обличчя, візерунки райдужної оболонки ока або відбитки пальців – пропонують високодискримінаційну інформацію про особу. Однак їхнє використання обмежене у спостереженні через потребу в зображеннях високої роздільної здатності, спеціалізованих сенсорах і суб'єктах, які співпрацюють. Тим не менш, атрибути обличчя, такі як зачіска або окуляри, досліджувалися як м'які біометричні ознаки для пошуку людей [28].

Методи, засновані на атрибутах, використовують семантичні дескриптори (наприклад, стиль одягу, колір або аксесуари), які є зрозумілими для людини і придатними для пошуку з нуля. Ці м'які біометричні ознаки забезпечують представлення середнього рівня, яке може доповнювати низькорівневі ознаки зовнішності [29].

Інформація про глибину і 3D-зображення, отримана за допомогою таких пристроїв, як бортова камера, дає змогу створювати зображення, що не залежать від точки зору і освітленості. Сигнали глибини сприяють кращій сегментації пішоходів і дозволяють виокремити скелетні та тривимірні ознаки форми тіла. Методи моделюють тіло, використовуючи довжину кінцівок, хмари точок або особливості скелета, іноді доповнені глибоким навчанням [30].

Топологія камер і мережеві обмеження також можуть допомогти в повторній ідентифікації шляхом включення просторово-часового контексту, наприклад, прогнозування ймовірних переходів камер або пішохідних шляхів на основі макета мережі.

Загалом, підходи на основі ознак спрямовані на подолання обмежень, пов'язаних лише з RGB-характеристиками зовнішності, шляхом включення багатшої, часто мультимодальної, контекстної або фізіологічної інформації для підвищення точності ідентифікації в складних середовищах [31].

## 4. Новітні тренди в задачі повторної ідентифікації людини

Нові тенденції у вирішенні задачі повторної ідентифікації цілком базуються на використанні глибокого навчання. Постійно пропонуються нові архітектури ШНМ. Ці нові ідеї дають деякі нові тенденції для застосування ШНМ до задачі повторної ідентифікації людини. Однією з таких тенденцій може бути генеративна змагальна мережа (GAN).

В задачах ідентифікації особи існує багато варіацій. Але через обмежений розмір наборів даних наведені приклади не можуть відображати всі можливі варіації. Деякі підходи розглядають використання генеративних моделей. Це моделі, які можуть навчитися створювати дані, подібні до даних, які ми надаємо. GAN було запропоновано в [32] і швидко застосовано до різних завдань. GAN навчає дві підмережі: генератор і дискримінацію. Дискримінація виявляє, чи є зразок згенерованим або справжнім, тоді як генератор створює зразки, щоб обдурити дискримінацію.

В роботі [33], автори вперше застосували GAN до проблеми повторної ідентифікації. Ідея проста: використання згенерованих зображень для доповнення навчальних даних. Вони використовують глибоку згорткову генеративну змагальну мережу (DCGAN) для генерації зразків і базову CNN для навчання репрезентації. Експерименти показують, що додавання даних, згенерованих GAN, покращує вивчені вбудовування CNN.

В роботі [34], замість того, щоб генерувати довільні зображення пішоходів, автори пропонують генерувати зображення особи в різних позах з різними скелетами. Для накладання ідентичності згенерованих зображень пропонується модуль-провідник. В роботі [35] автори запропонували вирішити проблему розриву між доменами за допомогою GAN.

Ця проблема зазвичай існує між наборами даних, спричинена варіаціями налаштувань камери та умов зйомки. Це призводить до того, що модель, навчена на наявних навчальних даних, не може бути ефективно використана для нових тестових областей.

Щоб зменшити дорогі витрати на анування нових навчальних вибірок, автори виконали переклад зображень за допомогою GAN для подолання розриву між доменами, тобто перенесення зображень з одного домену в інший.

Інша тенденція – це механізм уваги. Моделі глибокого навчання без механізму уваги покладаються на всі зображення або всі кадри в послідовності зображень.

Однак, з одного боку, зміщення пішохода в обмежувальній рамці або зміна його пози може призвести до того, що деяким областям слід приділяти більше уваги, з іншого боку, через затемнення або зміну точок зору, для послідовності зображень деякі кадри є більш важливими.

З цією метою, автори роботи [36] запропонували спільно навчатися багаторівневій м'якій піксельній увазі та жорсткій регіональній увазі з одночасною оптимізацією представлення ознак. Автори роботи [37] об'єднують просторову і часову увагу в моделі просторово-часової уваги для вирішення проблем повторної ідентифікації людини на основі відео. Для просторової уваги використовується термін «пеналізація» для впорядкування множинної надлишкової уваги. Часова увага використовується для присвоєння ваг різним важливим областям на основі кожного кадру, щоб повною мірою використати переваги дискримінативних областей зображення.

## Висновки

У статті було проведено комплексний аналіз сучасного стану та перспектив розвитку технологій повторної ідентифікації людини в інтелектуальних системах відеоспостереження. Встановлено, що задача re-ID є надзвичайно актуальною в контексті стрімкого поширення відеоаналітики для забезпечення громадської безпеки, моніторингу публічних просторів і побудови аналітичних платформ нового покоління.

Детально розглянуто формальну постановку задачі повторної ідентифікації, що ґрунтується на пошуку відповідностей між зображеннями однієї особи у багатовимірному ознаковому просторі. Показано, що центральним викликом є побудова дискримінативного, інваріантного до умов зйомки простору ознак, а також забезпечення здатності моделі до узагальнення на нові домени.

Проаналізовано основні виклики, що ускладнюють розв'язання задачі в реальних умовах, серед яких – варіативність зовнішнього вигляду, зміна освітлення і ракурсу, часткове перекриття, domain shift, відсутність біометричних ознак, а також обмеження обчислювальних ресурсів і вимоги до конфіденційності. Ці фактори зумовлюють потребу в розробці адаптивних, високопродуктивних та етичних систем повторної ідентифікації.

У статті узагальнено провідні підходи до вилучення ознак, включаючи холістичні, патч-орієнтовані та методи на основі частин тіла. Окремо проаналізовано підходи до зіставлення ознак на основі класичних та метричних методів, а також сучасні архітектури глибокого навчання. Показано, що використання CNN, attention-механізмів, трансформерів та генеративних моделей дозволяє суттєво підвищити точність повторної ідентифікації.

Окреслено новітні тенденції у цій галузі, зокрема застосування генеративних змагальних мереж (GAN) для доповнення навчальних даних та зменшення domain gap, а також впровадження механізмів просторової та часової уваги для покращення локалізації та динамічного аналізу.

Загалом, дослідження підтверджує, що сучасні методи, базовані на глибинному навчанні, відкривають широкі можливості для подальшого розвитку re-ID-систем. Водночас перспективними напрямками залишаються: підвищення стійкості моделей до змін середовища, оптимізація ресурсоспоживання, розробка мультимодальних підходів та створення етичних, прозорих і контрольованих систем відеоспостереження нового покоління.

## Список використаної літератури:

1. L. Zheng, Y. Yang, and A. G. Hauptmann, *Person Re-identification: Past, Present and Future*, arXiv preprint arXiv:1610.02984, 2016. [Online]. Available: <https://arxiv.org/abs/1610.02984>
2. A. Angelova, A. Krizhevsky, and V. Vanhoucke, "Pedestrian detection with a Large-Field-Of-View deep network," in *Proc. 2015 IEEE Int. Conf. Robotics and Automation (ICRA)*, Seattle, WA, USA, 2015, pp. 704–711. doi: 10.1109/ICRA.2015.7139256.
3. M. E. Chasmai and T. Banerjee, *Person Re-Identification*, arXiv preprint arXiv:2204.13158, 2022. [Online]. Available: <https://arxiv.org/abs/2204.13158>
4. H. Huang, D. Li, Z. Zhang, X. Chen, and K. Huang, "Adversarially Occluded Samples for Person Re-identification," in *Proc. 2018 IEEE/CVF Conf. on Computer Vision and Pattern Recognition (CVPR)*, Salt Lake City, UT, USA, 2018, pp. 5098–5107. doi: 10.1109/CVPR.2018.00535.
5. L. Ren, J. Lu, J. Feng, and J. Zhou, "Uniform and Variational Deep Learning for RGB-D Object Recognition and Person Re-Identification," *IEEE Transactions on Image Processing*, vol. 28, no. 10, pp. 4970–4983, Oct. 2019. doi: 10.1109/TIP.2019.2915655.
6. J. Lin, L. Ren, J. Lu, J. Feng, and J. Zhou, "Consistent-Aware Deep Learning for Person Re-identification in a Camera Network," in *Proc. 2017 IEEE Conf. on Computer Vision and Pattern Recognition (CVPR)*, Honolulu, HI, USA, 2017, pp. 3396–3405.
7. L. Bazzani, M. Cristani, A. Perina, M. Farenzena, and V. Murino, "Multiple-Shot Person Re-identification by HPE Signature," in *Proc. 2010 20th Int. Conf. on Pattern Recognition (ICPR)*, Istanbul, Turkey, 2010, pp. 1413–1416. doi: 10.1109/ICPR.2010.349.
8. W. Li, X. Zhu, and S. Gong, *Harmonious Attention Network for Person Re-Identification*, arXiv preprint arXiv:1802.08122, 2018. [Online]. Available: <https://arxiv.org/abs/1802.08122>
9. Y. Yang, J. Yang, J. Yan, S. Liao, D. Yi, and S. Z. Li, "Salient Color Names for Person Re-identification," in *Computer Vision – ECCV 2014*, D. Fleet, T. Pajdla, B. Schiele, and T. Tuytelaars, Eds. Cham: Springer, 2014, pp. 536–551. doi: 10.1007/978-3-319-10590-1\_35.
10. Y. Shen, W. Lin, J. Yan, M. Xu, J. Wu, and J. Wang, *Person Re-identification with Correspondence Structure Learning*, arXiv preprint arXiv:1504.06243, 2015. [Online]. Available: <https://arxiv.org/abs/1504.06243>
11. D. S. Cheng, M. Cristani, M. Stoppa, L. Bazzani, and V. Murino, "Custom Pictorial Structures for Re-identification," in *Proc. British Machine Vision Conference (BMVC)*, Dundee, UK, 2011.
12. D. Gray and H. Tao, "Viewpoint Invariant Pedestrian Recognition with an Ensemble of Localized Features," in *Computer Vision – ECCV 2008*, D. Forsyth, P. Torr, and A. Zisserman, Eds. Berlin, Heidelberg: Springer, 2008, pp. 262–275. doi: 10.1007/978-3-540-88682-2\_21.
13. T. Matsukawa, T. Okabe, E. Suzuki, and Y. Sato, "Hierarchical Gaussian Descriptor for Person Re-identification," in *Proc. 2016 IEEE Conf. on Computer Vision and Pattern Recognition (CVPR)*, Las Vegas, NV, USA, 2016, pp. 1363–1372. doi: 10.1109/CVPR.2016.152.
14. W. R. Schwartz and L. S. Davis, "Learning Discriminative Appearance-Based Models Using Partial Least Squares," in *Proc. 2009 XXII Brazilian Symp. on Computer Graphics and Image Processing (SIBGRAPI)*, Rio de Janeiro, Brazil, 2009, pp. 322–329. doi: 10.1109/SIBGRAPI.2009.42.
15. Z. Lin and L. S. Davis, "Learning Pairwise Dissimilarity Profiles for Appearance Recognition in Visual Surveillance," in *Advances in Visual Computing*, G. Bebis et al., Eds. Berlin, Heidelberg: Springer, 2008, pp. 23–34. doi: 10.1007/978-3-540-89639-5\_3.
16. A. Mignon and F. Jurie, "PCCA: A new approach for distance learning from sparse pairwise constraints," in *Proc. 2012 IEEE Conf. on Computer Vision and Pattern Recognition (CVPR)*, Providence, RI, USA, 2012, pp. 2666–2672. doi: 10.1109/CVPR.2012.6247987.
17. S. Pedagadi, J. Orwell, S. Velastin, and B. Boghossian, "Local Fisher Discriminant Analysis for Pedestrian Re-identification," in *Proc. 2013 IEEE Conf. on Computer Vision and Pattern Recognition (CVPR)*, Portland, OR, USA, 2013, pp. 3318–3325. doi: 10.1109/CVPR.2013.426.
18. B. J. Prosser, W. Zheng, S. Gong, and T. Xiang, "Person Re-Identification by Support Vector Ranking," in *Proc. British Machine Vision Conference (BMVC)*, Aberystwyth, UK, 2010.
19. D. Yi, Z. Lei, and S. Z. Li, *Deep Metric Learning for Practical Person Re-Identification*, arXiv preprint arXiv:1407.4979, 2014. [Online]. Available: <https://arxiv.org/abs/1407.4979>
20. R. R. Varior, M. Haloi, and G. Wang, "Gated Siamese Convolutional Neural Network Architecture for Human Re-identification," in *Computer Vision – ECCV 2016*, B. Leibe, J. Matas, N. Sebe, and M. Welling, Eds. Cham: Springer, 2016, pp. 791–808. doi: 10.1007/978-3-319-46484-8\_48.

21. R. R. Varior, B. Shuai, J. Lu, D. Xu, and G. Wang, *A Siamese Long Short-Term Memory Architecture for Human Re-Identification*, arXiv preprint arXiv:1607.08381, 2016. [Online]. Available: <https://arxiv.org/abs/1607.08381>
22. J. You, A. Wu, X. Li, and W.-S. Zheng, "Top-Push Video-Based Person Re-identification," in *Proc. 2016 IEEE Conf. on Computer Vision and Pattern Recognition (CVPR)*, Las Vegas, NV, USA, 2016, pp. 1345–1353. doi: 10.1109/CVPR.2016.150.
23. D. Cheng, Y. Gong, S. Zhou, J. Wang, and N. Zheng, "Person Re-identification by Multi-Channel Parts-Based CNN with Improved Triplet Loss Function," in *Proc. 2016 IEEE Conf. on Computer Vision and Pattern Recognition (CVPR)*, Las Vegas, NV, USA, 2016, pp. 1335–1344. doi: 10.1109/CVPR.2016.149.
24. S. Ding, L. Lin, G. Wang, and H. Chao, *Deep Feature Learning with Relative Distance Comparison for Person Re-identification*, arXiv preprint arXiv:1512.03622, 2015. [Online]. Available: <https://arxiv.org/abs/1512.03622>
25. Z. Zheng, L. Zheng, and Y. Yang, *Unlabeled Samples Generated by GAN Improve the Person Re-identification Baseline in vitro*, arXiv preprint arXiv:1701.07717, 2017. [Online]. Available: <https://arxiv.org/abs/1701.07717>
26. N. Gheissari, T. B. Sebastian, and R. Hartley, "Person Reidentification Using Spatiotemporal Appearance," in *Proc. 2006 IEEE Computer Society Conf. on Computer Vision and Pattern Recognition (CVPR)*, vol. 2, New York, NY, USA, 2006, pp. 1528–1535. doi: 10.1109/CVPR.2006.223.
27. Y. Yan, B. Ni, Z. Song, C. Ma, Y. Yan, and X. Yang, *Person Re-Identification via Recurrent Feature Aggregation*, arXiv preprint arXiv:1701.06351, 2017. [Online]. Available: <https://arxiv.org/abs/1701.06351>
28. D. A. Vaquero, R. S. Feris, D. Tran, L. Brown, A. Hampapur, and M. Turk, "Attribute-based people search in surveillance environments," in *Proc. 2009 Workshop on Applications of Computer Vision (WACV)*, Snowbird, UT, USA, 2009, pp. 1–8. doi: 10.1109/WACV.2009.5403131.
29. M. Munaro, A. Fossati, A. Basso, E. Mene-gatti, and L. Van Gool, "One-Shot Person Re-identification with a Consumer Depth Camera," in *Person Re-Identification*, S. Gong, M. Cristani, S. Yan, and C. C. Loy, Eds. London: Springer, 2014, pp. 161–181. doi: 10.1007/978-1-4471-6296-4\_8.
30. A. Haque, A. Alahi, and L. Fei-Fei, *Recurrent Attention Models for Depth-Based Person Identification*, arXiv preprint arXiv:1611.07212, 2016. [Online]. Available: <https://arxiv.org/abs/1611.07212>
31. O. Javed, K. Rasheed, M. Shafique, and M. Shah, "Tracking across multiple cameras with disjoint views," in *Proc. Ninth IEEE Int. Conf. on Computer Vision (ICCV)*, vol. 2, Nice, France, 2003, pp. 952–957. doi: 10.1109/ICCV.2003.1238451.
32. I. J. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, and Y. Bengio, *Generative Adversarial Networks*, arXiv preprint arXiv:1406.2661, 2014. [Online]. Available: <https://arxiv.org/abs/1406.2661>
33. J. Liu, B. Ni, Y. Yan, P. Zhou, S. Cheng, and J. Hu, "Pose Transferrable Person Re-Identification," in *Proc. IEEE Conf. on Computer Vision and Pattern Recognition (CVPR)*, Salt Lake City, UT, USA, June 2018.
34. L. Wei, S. Zhang, W. Gao, and Q. Tian, *Person Transfer GAN to Bridge Domain Gap for Person Re-Identification*, arXiv preprint arXiv:1711.08565, 2018. [Online]. Available: <https://arxiv.org/abs/1711.08565>
35. S. Liao, Y. Hu, X. Zhu, and S. Z. Li, *Person Re-identification by Local Maximal Occurrence Representation and Metric Learning*, arXiv preprint arXiv:1406.4216, 2015. [Online]. Available: <https://arxiv.org/abs/1406.4216>
36. S. Li, S. Bak, P. Carr, and X. Wang, *Diversity Regularized Spatiotemporal Attention for Video-based Person Re-identification*, arXiv preprint arXiv:1803.09882, 2018. [Online]. Available: <https://arxiv.org/abs/1803.09882>
37. R. Li, X. Yuan, W. Liu, and X. Xu, "Event-based Video Person Re-identification via Cross-Modality and Temporal Collaboration," in *Proc. ICASSP 2025 – IEEE Int. Conf. on Acoustics, Speech and Signal Processing (ICASSP)*, Hyderabad, India, 2025, pp. 1–5. doi: 10.1109/ICASSP49660.2025.10889628

## State and Development of Human Re-Identification Technologies in Intelligent Video Systems

A. Kolpikov, E. Zabarna, M. Lobachev  
Odesa Polytechnic National University

**Abstract.** The article discusses the current state and main directions of development of human re-identification (re-ID) technologies in intelligent video surveillance systems. The formal statement of the re-ID problem is defined and its role in ensuring continuous automatic monitoring of objects between non-overlapping cameras is characterized. The main challenges that complicate the solution of this problem in real conditions are analyzed, including variability of appearance, changes in angle and lighting, partial overlap, domain shift, limited training data and hardware resources. The paper summarizes the approaches to feature extraction, image matching, and deep learning methods used to build effective re-ID models. Particular attention is paid to current trends in the use of generative adversarial networks (GANs) and attention mechanisms that improve identification accuracy. Promising areas for further research, in particular in the field of multimodal learning, adaptation to new domains, and ethical implementation of technologies in public environments, are presented.

**Keywords:** person re-identification, intelligent video systems, computer vision, deep learning, feature space, metric learning, generative adversarial networks, attention mechanism.

Отримано 10.08.2025



**Колпиков Андрій Сергійович**, аспірант кафедри штучного інтелекту та аналізу даних, Національний університет «Одеська Політехніка»; проспект Шевченка, 1. Одеса, 65044, Україна.  
E-mail: kolpikov.andrey@gmail.com

**Kolpikov Andrii**, PhD Student of Artificial Intelligence and Data Analysis Department, Odesa Polytechnic National University, 1, Shevchenko Avenue, Odesa, 65044, Ukraine.  
E-mail: kolpikov.andrey@gmail.com

**ORCID ID:** <https://orcid.org/0009-0006-9116-7501>



**Забарна Елеонора Миколаївна**, доктор економічних наук, професор, зав. каф. інтегрованих технологій управління, Національний університет «Одеська Політехніка»; проспект Шевченка, 1. Одеса, 65044, Україна. E-mail: e.m.zabarna@op.edu.ua

**Zabarna Eleonora**, Doctor of Economics, Professor, Head of the Department of Integrated Management Technologies, Odesa Polytechnic National University; 1, Shevchenko Avenue, Odesa, 65044, Ukraine. E-mail: e.m.zabarna@op.edu.ua

**ORCID ID:** <https://orcid.org/0000-0002-2659-5909>



**Лобачев Михайло Вікторович**, кандидат технічних наук, професор, директор Інституту штучного інтелекту та робототехніки, Національний університет «Одеська Політехніка»; проспект Шевченка, 1. Одеса, 65044, Україна.  
E-mail: lobachev@op.edu.ua

**Lobachev Mykhaylo**, Ph.D., Professor, Head of the Institute of Artificial Intelligence and Robotics, Odesa Polytechnic National University; 1, Shevchenko Avenue, Odesa, 65044, Ukraine.  
E-mail: lobachev@op.edu.ua

**ORCID ID:** <https://orcid.org/0000-0002-4859-304X>