

ІНФОРМАЦІЙНІ ТЕХНОЛОГІЇ

УДК 004.93

Р. Н. Кветний, Р. В. Маслій, О. М. Кириленко

ДОСЛІДЖЕННЯ АТРИБУТІВ ЛЮДИНИ ДЛЯ ЗАДАЧІ
ПОВТОРНОЇ ІДЕНТИФІКАЦІЇ

Вінницький національний технічний університет, Вінниця

Анотація. В роботі досліджується задача розпізнавання атрибутів людини для покращення повторної ідентифікації. Здійснено аналіз методів повторної ідентифікації та методів розпізнавання атрибутів людини. Розглянуті набори даних атрибутів людини: Market-1501, DukeMTMC-reID, PETA, RAP-v1, RAP-v2 та PA-100k. Проаналізовано проблеми пов'язані як з формування наборів атрибутів людини так і з їх використанням. Надані рекомендації щодо вибору та використання наборів атрибутів людини. Також в роботі наведений огляд метрик оцінювання методів розпізнавання атрибутів людини.

Ключові слова: глибоке навчання, повторна ідентифікація особи, розпізнавання атрибутів людини.

Анотация. В работе исследуется задача распознавания атрибутов человека для улучшения повторной идентификации. Осуществлен анализ методов повторной идентификации и методов распознавания атрибутов человека. Рассмотрены наборы данных атрибутов человека: Market-1501, DukeMTMC-reID, PETA, RAP-v1, RAP-v2 и PA-100k. Проанализированы проблемы связанные как с формированием и наборов атрибутов человека так и с их использованием. Даны рекомендации по выбору и использованию наборов атрибутов человека. Также в работе представлен обзор метрик оценки методов распознавания атрибутов человека.

Ключевые слова: глубокое обучение, повторная идентификация человека, распознавание атрибутов человека.

Abstract. The paper investigates the problem of recognizing human attributes to improve re-identification. The analysis of re-identification methods and methods of recognition of human attributes is carried out. The data sets of human attributes are considered: Market-1501, DukeMTMC-reID, PETA, RAP-v1, RAP-v2 and PA-100k. Problems related to the formation and sets of human attributes and their use are analyzed. Recommendations for the selection and use of human attribute sets are provided. The paper also provides an overview of the metrics for evaluating methods for recognizing human attributes.

Key words: deep learning, person re-identification, human attribute recognition

DOI: <https://doi.org/10.31649/1999-9941-2020-49-3-4-13>.

Вступ

За останні роки зростає кількість мультимедійних даних, доступних в Інтернеті або наданих пристроями відеоспостереження. Саме тому зростає потреба у розробці рішень, які б давали змогу автоматично аналізувати зовнішність, особливості та поведінку людини. Розпізнавання атрибутів людини (Human attribute recognition – HAR) привертає все більше уваги у галузях комп'ютерного зору та розпізнавання образів в першу чергу через його потенційну зручність при використанні у багатьох застосуваннях: аналіз натовпу, пошук людини, виявлення людини, відстеження людини та повторна ідентифікація особи. Розпізнавання атрибутів людини націлене на опис та розуміння особливостей суб'єктів (таких як колір волосся, стиль одягу, стать тощо) або на основі даних про все тіло чи обличчя [1].

Труднощі з розпізнаванням атрибутів людини виникають головним чином через високу мінливість зовнішнього вигляду людини особливо у екземплярах усередині одного класу. Головними критеріями для отримання надійних систем HAR є [1]:

- отримання дискримінаційних ознак з вхідних даних;
- використання внутрішніх кореляційних зв'язків між атрибутами;
- робастність до основних проблем, таких як дані з низькою роздільною здатністю, зміна пози, оклюзії, варіації освітлення та захаращений фон;
- ефективне вирішення проблеми недостатності даних.

Задача повторної ідентифікації особи полягає у ідентифікації однієї і тієї ж людини у відео, знятих з різних камер. Складність задачі полягає у тому що сцени кожної камери, як правило, просторово не перетинаються, перехід людини між сценами кожної камери може значно варіюватися у часі, умови освітлення / пози людини різняться між камерами в реальних ситуаціях.

Проблема просторового зміщення людини відносно камери є однією з ключових проблем при повторній ідентифікації особи. На рис. 1 зображено приклади зображень у яких пішохід знаходиться у різних позах з точок зору різних камер [2].



Рисунок 1 – Приклади різних поз пішохода з точок зору різних камер

Актуальність

Виконуючи повторну ідентифікацію особи, експерти-люди шукають і покладаються на відповідні зовнішній вигляд або функціональні атрибути, які є однозначними в інтерпретації, такі як зачіска, тип взуття чи стиль одягу. Це атрибутивно-орієнтоване подання також використовується, коли опис надається усно (наприклад, очевидцем) оператору. Багато з цих атрибутів можна досить надійно виявити за допомогою сучасних технологій комп'ютерного зору.

Проте, незважаючи на велику кількість досліджень в цій сфері, автоматизована повторна ідентифікація все ще залишається в цілому невіршеною проблемою. Це пов'язано з тим, що більшість ознак все ще є недостатньо дискримінаційними для асоціації об'єктів, особливо для зображень з низькою роздільною здатністю, або ознаки є недостатньо надійними при зміні кута огляду чи освітлення [3].

Мета

Дослідження задачі розпізнавання атрибутів людини для покращення повторної ідентифікації людини у відео, знятих з різних камер.

Задачі

1. Огляд методів повторної ідентифікації особи та методів розпізнавання атрибутів людини.
2. Аналіз існуючих наборів даних для задачі розпізнавання атрибутів людини.
3. Огляд метрик оцінювання методів розпізнавання атрибутів людини.
4. Аналіз проблем пов'язаних з наборами даних атрибутів людини та вибір найкращого підходу та набору даних для подальшого використання у задачі повторної ідентифікації.

Розв'язання задач

Повторна ідентифікація. Сучасні підходи до повторної ідентифікації зазвичай використовують особливості низького рівня, такі як колір, текстура, просторова структура або їх комбінації. Отримавши відповідне подання, алгоритми пошуку найближчого сусіда або алгоритми пошуку відповістей на основі навчання, такі як ранжування, можуть бути використані для повторної ідентифікації. У кожному випадку для вимірювання подібності між двома зразками необхідно вибрати метрику відстані (наприклад, Евклідову чи Бхаттачарія). Також можна дискримінаційно оптимізувати метрику відстані. Інші додаткові аспекти проблеми також досліджувались для покращення продуктивності, такі як підвищення надійності шляхом комбінування ознак у декількох кадрах вздовж треклету, та вивчення топології або кореляції активності мережі камер для зменшення простору відповістей [4].

Атрибути. Моделювання на основі атрибутів нещодавно було використано для ефективного розпізнавання об'єктів та дій. На відміну від функцій низького рівня або класів – високого рівня, атрибути – це середній рівень опису класу чи екземпляра. Існують різні некеровані (наприклад, PCA моделі) або контрольовані (наприклад, нейронні мережі) підходи до моделювання, які створюють на основі даних подання середнього рівня. Ці методи спрямовані на проєкцію даних на базовий набір, визначений припущеннями конкретної моделі (наприклад, максимізація дисперсії, ймовірності чи розрідженості). Навпаки, навчання атрибутів фокусується на представленні екземплярів даних шляхом проєкції на базовий набір, який має семантичне значення для людини. Представлення семантичних атрибутів мають такі переваги:

1. Якщо дані є рідкісними, вони можуть бути більш значимими, ніж ознаки низького рівня.
2. Їх можна використовувати разом із вихідними даними для більшої ефективності.
3. Вони є відповідним поданням для безпосередньої взаємодії людини, отже, дозволяючи пошук визначати або обмежувати атрибутами [5].

Атрибути для повторної ідентифікації особи. Один з поглядів на атрибути – це тип переданого контексту, оскільки вони надають допоміжну інформацію про екземпляр, який допомагає у (повторній) ідентифікації. Тут вони пов'язані з вивченням м'якої біометрії, метою якої є підвищення ефективності біометричної ідентифікації за допомогою допоміжної інформації. В якості альтернативи вони можуть бути використані для семантичної ідентифікації профілю атрибутів, яка в ході ранніх досліджень мала на меті знайти людей, які відповідають словесному опису атрибутів, із мережі камер. Однак цей підхід поки що апробований лише на відносно простих даних із невеликим набором однаково надійних атрибутів обличчя. Однією з центральних проблем використання атрибутів для автоматизованої (повторної) ідентифікації є проблема їх нерівномірної та змінної інформативності та надійності вимірювання на основі вихідних даних [6].

Огляд існуючих методів

Повторна ідентифікація особи на основі згорткових нейронних мереж (ЗНМ). Методи повторної ідентифікації, засновані на ЗНМ, домінують над іншими після успіху глибокого навчання [7, 8, 9]. У низці робіт, що вивчають глибинні метрики [10, 11], пари зображень або триплети подаються в мережу. Зазвичай, просторові обмеження інтегруються у процес навчання подібності [12, 13]. Взагалі кажучи, гли-

бинні метричні методи навчання мають переваги у навчанні на відносно невеликих наборах даних, але їх ефективність з великими наборами даних може бути недостатньою. У роботах [7, 10, 11] автори досліджують глибокі представлення для задачі повторної ідентифікації. У роботах [12, 13] використовується генеративні змагальні мережі (GAN). Нещодавно для вирішення проблеми даних для повторної ідентифікації було запропоновано деякі напівкеровані методи та некеровані методи. Ці методи показують високі результати з невеликою кількістю або взагалі за відсутності анотацій. Інформація про атрибути цим методам теж може бути корисна у напівкерованому режимі.

Атрибути для повторної ідентифікації особи. У деяких ранніх роботах атрибути використовуються як допоміжна інформація для покращення низькорівневих ознак [14, 15, 16]. У роботі [17] дескриптори низького рівня та опорні вектори використовуються для навчання детектору атрибутів. В роботах [14, 15] використовуються як ознаки низького рівня так і атрибути для задачі повторної ідентифікації. Нещодавно запропоновані деякі методи, що базуються на глибокому навчанні. В роботі [18] пропонується метод навчання від «грубого до тонкого», який використовує набір глибоких мереж. Одна мережа навчена здійснювати бінарну класифікацію «людина» / «не людина», інша мережа навчена розпізнавати стать людини, третя мережа навчена здійснювати повторну ідентифікацію людини. У цій роботі мережі навчаються окремо і можуть не помітити взаємодоповнюваність мітки ідентифікатора та мітки атрибута. У роботі [19] мережа спочатку навчається на незалежному наборі даних із мітками атрибутів, а потім здійснюється точніше налаштування мережі на цільовому наборі даних, використовуючи мітки ідентифікації та використовуючи триплет-функцію втрат. Подібним чином, у роботі [20] мережа попередньо навчається на незалежному наборі даних, з розміченими атрибутами, а потім здійснюється «тонке налаштування» на іншому наборі з ідентифікаторами особи. У роботі [21] набір міток атрибутів використовується як запит для отримання зображення людини. Змагальне навчання використовується для генерування зображень-образів певних атрибутів з подальшим отриманням відповідності між зображенням-запитом як на глобальному рівні, так і на рівні семантичного ідентифікатора.

Дослідження анотованих наборів даних

Для дослідження було обрано набори даних з мітками атрибутів: Market-1501 [22], DukeMTMC-reID [23], PETA [24], RAP-v1[25], RAP-v2 [26] та PA-100k [27].

Набір даних Market-1501 Attribute. Market-1501 Attribute – це версія набору даних Market-1501, доповнена анотацією з 27 людськими атрибутами. Початковий набір Market-1501 призначений для дослідження задач повторної ідентифікації осіб. Набір був зібраний на вулиці перед супермаркетом за допомогою 6 камер (5 камер високої роздільної здатності та одна низької роздільної здатності). Атрибути представлені для кожної особи, і в цілому у наборі присутня 1501 анотована особа (751 у тренувальному наборі, 750 у тестувальному наборі). Загалом набір даних містить 32 668 зображень (12937 – тренувальних, 19733 – тестувальних). Атрибути в наборі Market-1501 Attribute, включають демографічну інформацію (стать та вік), інформацію про аксесуари (наявність капелюха, рюкзака, сумки), зовнішній вигляд (довжина волосся) та тип одягу та колір (довжина рукава, довжина одяг нижньої частини тіла, тип одягу нижньої частини тіла, 8 кольорів одягу верхньої частини тіла, 9 кольорів одягу нижньої частини тіла). Позитивні та негативні приклади деяких репрезентативних атрибутів набору даних Market-1501 показано на рис. 2.



Рисунок 2 – Позитивні та негативні приклади атрибутів «короткий рукав», «рюкзак», «сукня», «синій одяг на нижній частині тіла» набору даних Market-1501

Набір даних Duke Attribute. Набір даних DukeMTMC-reid був зібраний в кампусі університету Дюка і містить понад 14 годин відеопослідовностей, зібраних з 8 камер, розташованих таким чином, що фіксувати сцени з натовпом. Головною метою цього набору даних є дослідження задач повторної ідентифікації особи та відстеження особи за допомогою декількох камер. Крім цього підмножина набору даних DukeMTMC-reid - Duke Attribute була анотована людськими атрибутами. Атрибути представлені для кожної особи, і в цілому у наборі присутні 1404 анотовані особи (702 у тренувальному наборі, 702 у тестувальному наборі). Набір даних Duke Attribute містить 36411 зображень (16522 – тренувальних, 19889 – тестувальних). Анотація особи містять 23 атрибути, які стосуються статі (чоловіча, жіноча), аксесуарів: присутність капелюха (так, ні), присутність рюкзака (так, ні), присутність сумочки (так, ні), присутність інших типів сумок (так, ні); стиль одягу: тип взуття (черевики, інше взуття), колір взуття (темний, яскравий), довжина верхнього одягу (довга, коротка), 8 кольорів одягу верхньої частини тіла (чорний, білий, червоний, фіолетовий, сірий, синій, зелений, коричневий) та 7 кольорів одягу нижньої частини тіла (чорний, білий, червоний, сірий, синій, зелений, коричневий). На рис. 3 наведено приклад атрибутів для наборів даних Market-1501 Attribute та Duke Attribute.



Рисунок 3 – Приклад атрибутів для наборів даних (a) Market-1501 та (b) Duke Attribute

Набір даних PETA. Набір даних PEdesTrian Attribute (PETA) містить 19 000 зображень пішоходів, зібраних з 10 загальнодоступних наборів даних; тому зображення мають великі варіації щодо сцени, умов освітлення та роздільної здатності. Роздільна здатність зображень варіюється від 17×39 до 169×365 пікселів. Набір даних надає розширені анотації: зображення вручну позначені 61 двійковим та 4 мультикласовими атрибутами. Двійкові атрибути містять інформацію про демографічні показники (стать: чоловік, жінка; вік: 16–30 років, 31–45 років, 46–60 років, більше 61 року), зовнішній вигляд (довге волосся), одяг (футболка, брюки тощо) та аксесуари (сонцезахисні окуляри, капелюх, рюкзак тощо). Мультикласові атрибути пов'язані з (одинадцятьма основними) кольорами для одягу верхньої та нижньої частини тіла, взуття та волосся суб'єкта. Приклад атрибутів набору даних PETA представлений на рис. 4.

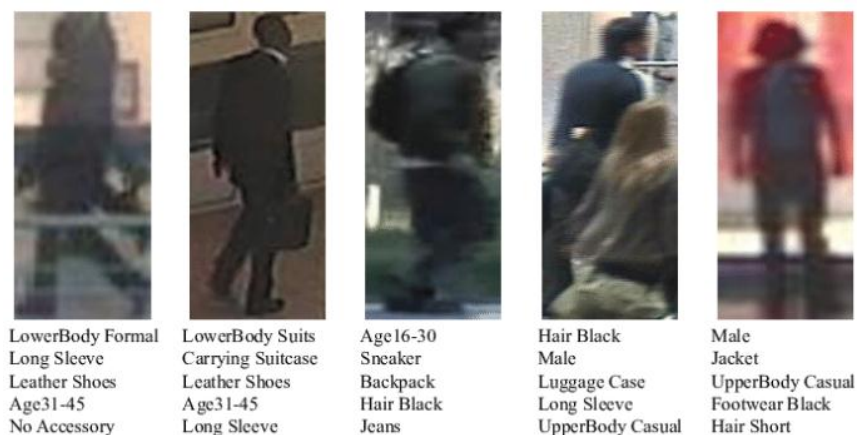


Рисунок 4 – Приклад атрибутів набору даних PETA

Набори даних RAPv1 і RAPv2. В даний час існує дві версії набору даних RAP (Richly Annotated Pedestrian). Перша версія RAPv1 була зібрана з камер спостереження в торгових центрах протягом трьох

місяців; Далі для анотації атрибутів вручну було обрано 17 годин відеозаписів. Загалом набір даних містить 41 585 анотованих зображень. Цей набір має 72 атрибутів, які включають демографічну інформацію (стать та вік), аксесуари (рюкзак, сумка на одне плече, сумочку, поліетиленовий пакет, паперовий пакет тощо), зовнішній вигляд людини (стиль волосся, колір волосся, форма тіла) та інформацію про одяг (стиль одягу, колір одягу, стиль взуття, колір взуття тощо). Крім того, набір даних містить анотації щодо оклюзій, точок зору та інформації про частини тіла.

Друга версія набору даних RAP (RAPv2) призначена як для задачі пошуку людей, так і для розпізнавання атрибутів людини у реальних сценаріях спостереження. Набір даних було зібрано в приміщенні, у торговому центрі, і він містить 84 928 зображень (2589 особи) з 25 різних сцен. Для збору набору даних використовувались камери високої роздільної здатності (1280×720), а роздільна здатність людських силуетів варіюється від 33×81 до 415×583 пікселів. Атрибути версії RAP v2 такі самі, як і у версії RAP v1 (72 атрибути, інформація про оклюзії, точку зору та частини тіла). Приклад атрибутів з наборів даних RAP представлені на рис. 5.



Рисунок 5 – Приклад атрибутів з наборів даних RAP

Набір даних PA-100K. Набір даних PA-100k був розроблений з метою перевершити існуючі набори даних HAR як за кількістю, так і за різноманітністю; набір даних містить більше 100 000 зображень, зроблених у 598 різних сценаріях. Набір даних було зафіксовано камерами зовнішнього спостереження; тому зображення забезпечують велику різницю в роздільній здатності зображення, умовах освітлення та навколишнього середовища. До набору даних додано 26 атрибутів, включаючи демографічні (вік, стать), аксесуари (сумочка, телефон) та інформацію про одяг. Приклад атрибутів набору даних PA-100k наведений на рис. 6.

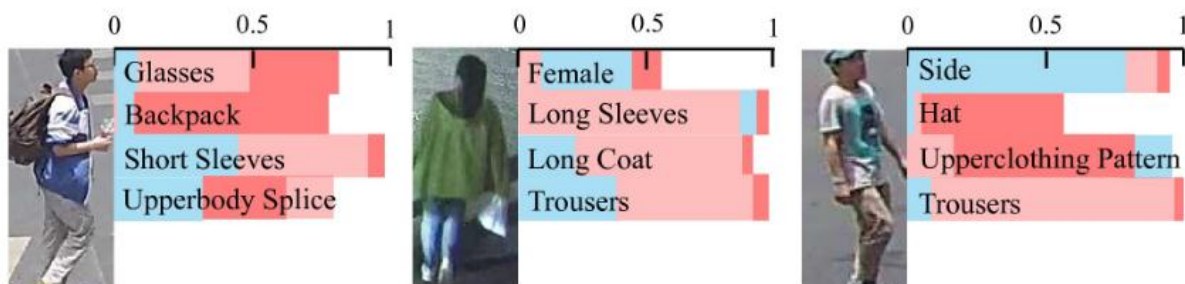


Рисунок 6 – Приклад атрибутів з набору даних PA100k

Метрики оцінювання

Враховуючи, що HAR – це багатокласова проблема класифікації, то найбільш поширеними показниками для вимірювання ефективності цих методів є Accuracy (*Acc*), Precision (*Prec*), Recall (*Rec*) та оцінка *F1*. Загалом ці показники можна обчислити на двох різних рівнях: на рівні міток та на рівні екземпляра.

Оцінка на рівні міток розглядає кожен атрибут незалежно. Наприклад, якщо атрибути статі та зросту відповідають міткам (чоловіча, жіноча) та (низький, середній, високий) відповідно, то оцінка на рівні міток вимірюватиме ефективність кожної комбінації атрибутів та міток. Метрикою, прийнятою в більшості робіт для оцінки рівня міток, є середня точність (*mA*):

$$mA = \frac{1}{2N} \sum_{i=1}^N \left(\frac{TP_i}{P_i} + \frac{TN_i}{N_i} \right) \quad (1)$$

де i відноситься до кожного з N атрибутів. mA визначає середню точність між позитивними та негативними прикладами кожного атрибута.

При оцінці на рівні екземпляра результати вимірюються для кожного атрибута, не враховуючи кількість міток, які він містить. Таким чином, оцінки $Prec$, Rec , Acc та $F1$ для i -го атрибута визначаються:

$$Prec_i = \frac{TP_i}{P_i}, \quad (2)$$

$$Rec_i = \frac{TP_i}{N_i}, \quad (3)$$

$$Acc_i = \frac{TP_i + TN_i}{P_i + N_i}, \quad (4)$$

$$F_i = \frac{2 \cdot Prec \cdot Rec}{Prec + Rec} \quad (5)$$

Використання цих показників є дуже поширеним для порівняльного аналізу різних атрибутів. Загальна продуктивність системи може бути виміряна середнім значенням Acc_i за всіма атрибутами або за допомогою mA . Однак ці показники можуть суттєво розходитися, коли атрибути дуже незбалансовані. mA є кращим, коли автори хочуть оцінити ефект розбалансування даних [28].

Визначення основних проблем пов'язаних з наборами даних для HAR

Першими проблемами, які виникають при розробці нового набору даних для HAR, є знаходження відповіді на такі питання: які атрибути слід анотувати та скільки та які класи потрібно використати для правильного опису атрибута? Очевидно, що обидва ці питання залежать від сфери застосування системи HAR. Як правило, кінцевою метою HAR, незалежно від сфери застосування, буде точний опис зображення з точки зору зрозумілих для людини семантичних міток, наприклад, «шестирічна дівчинка, одягнена у сині джинси, з білою футболкою та з рюкзаком у смужку». Що стосується другого питання, то відповідь знайти просто для деяких атрибутів, таких як стать, але уже вона стає неочевидною та суб'єктивною для інших атрибутів, таких як вік чи інформація про одяг.

Наступною проблемою будь-якого набору даних є їх незбалансованість. Тобто проблема різної кількості екземплярів класу в межах набору даних. Хоча у деяких наборах даних розробники намагалися отримати збалансовані класи, все ж деякі класи залишилися не повністю збалансованими (особливо ті, що стосуються інформації про одяг). І в цілому повністю збалансовані набори даних є не тривіальною задачею. Проблема дисбалансу також існує для демографічних атрибутів. У всіх розглянутих наборах даних HAR, наприклад, клас маленьких дітей представлений погано. Часто набори даних зазвичай збираються в одному регіоні (місті, країні, континенті), дані, як правило, є незбалансованими з точки зору етнічної приналежності. Це важливе питання, оскільки деякі дослідження довели існування ефекту іншої раси – тенденції до більш легкого розпізнавання осіб тієї самої національності – для класифікатора машинного навчання.

Корисним атрибутом в наборі даних може бути анотація поз та оклюзій. Деякі набори даних надають таку інформацію. Ці дані можуть бути корисні дослідникам при оцінюванні системи HAR та дають можливість діагностувати помилки HAR і перевіряти вплив різних факторів.

Ще однією проблемою пов'язаною з анотацією набору даних полягає у врахуванні ситуації коли атрибут для анотації неможливо розрізнити через низьку роздільну здатність, отримання зображення у складних умовах освітлення, оклюзії тощо. Більшість наборів даних, як правило, ігнорують цю проблему та класифікують наявність атрибута або надають багатокласову схему атрибутів. Однак у реальних умовах потрібно враховувати таку ситуацію, оскільки випадки, в яких атрибути важко розпізнати можуть траплятися досить часто.

При роботі з HAR слід ретельно розробити схему розділення набору даних на навчальну, валідаційну та тестувальну вибірки. Загальною помилкою є розбиття кадрів відеопослідовності на навчальну та валідаційну вибірку випадковим чином, незалежно від особи людини. Такий підхід може призвести до несправедливого віднесення суб'єкта до однієї з цих вибірок та спричинення упередженості в процесі навчання. Враховуючи що сучасні методи, як правило, покладаються на архітектури глибоких нейронних мереж, які мають у природі чорну скриньку, не так просто визначити, які особливості зображення ведуть до остаточного результату класифікації. Рішення цієї проблеми полягає у отриманні наборів областей

зображень кожної особи (треклетів) із відеопослідовності або надання анотацій на рівні ідентичності на кожному кадрі. Потім треклет кожної особи можна випадковим чином призначити для однієї з вибірок.

Проблеми конфіденційності. Раніше, оскільки традиційні системи відеоспостереження були простими і включали лише моніторинг людини, конфіденційність не була основною проблемою; проте в наші дні поширеність систем, оснащених передовими технологіями, у громадських місцях (наприклад, торгові центри, приватні та громадські будівлі, автобусні та залізничні вокзали) викликала нові проблеми щодо конфіденційності та безпеки.

Вибір набору даних

Деякі з розглянутих вище проблем потрібно враховувати як при виборі набору даних так і під час створення власного методу розпізнавання атрибутів для врахування можливих недосконалостей існуючих наборів даних.

Крім зазначеного вище при виборі набору даних важливим є вирішення питання про значимість присутніх у наборі атрибутів для цільової області застосування і важливість кожного атрибуту для обрання набору найбільш інформативних атрибутів. Звісно набори даних що мають більшу кількість зображень є більш привабливими для використання оскільки потенційно можуть містити більшу кількість різноманітних прикладів для кожного класу кожного атрибуту. Набори з більшою кількістю атрибутів є теж привабливими оскільки можуть більш повно представляти зображення людини і наблизитися за своєю дискримінаційною здатністю до низькорівневих ознак що використовують глибокі нейронні мережі для задачі повторної ідентифікації.

Крім того при виборі набору даних слід брати до уваги кількість сцен в яких були отримані зображення, оскільки це дозволить отримати різноманітні приклади в межах кожного класу для кожного атрибуту.

Також важливим питанням при виборі набору даних для HAR є умови зйомки: в середині приміщення чи ззовні, оскільки якість освітлення для кожного випадку може варіюватися в різних межах, потрібно врахувати це при цільовому використанні.

Іншим аспектом, який слід враховувати при виборі набору даних HAR, є налаштування камери, що використовується для зйомки зображень або відеопослідовностей. Зазвичай використовуються фіксовані або рухомі камери, і вибір набору знову буде залежати від області застосування програми в яку буде інтегрована система HAR.

Висновки

В результаті аналізу задачі розпізнавання атрибутів людини можна зробити декілька висновків: по-перше, найбільш перспективними для подальшого дослідження задачі HAR є підходи, що базуються на використанні глибоких нейронних мереж; по-друге, можна визначити, що використання атрибутів людини, які представляють собою середньорівневі ознаки зображення, можуть підсилити низькорівневі ознаки, що використовуються, і таким чином покращити достовірність результатів повторної ідентифікації. Але слід також відмітити, що при використанні атрибутів для повторної ідентифікації виникає проблема їх нерівномірної та змінної інформативності та надійності вимірювання на основі вихідних даних. Саме тому ключовим питанням є використання надійного набору даних атрибутів людини для навчання нейромережевого класифікатора.

В роботі на основі аналізу найбільш використовуваних наборів даних атрибутів людини визначено проблеми, які пов'язані як з формуванням власного набору даних атрибутів людини так і з ефективним використанням існуючих наборів даних. Наведено рекомендації щодо вибору наборів даних атрибутів людини. Згідно з цими рекомендаціями найкращими для подальшого використання в дослідженнях автори вважають набори RAPv2 та PA100k.

Список літератури

- [1] Y. Lina, L. Zhengb, Z. Zhenga, Yu Wua, Z. Hua, Ch. Yanc , Yi Yang, «Improving person re-identification by attribute and identity learning», *Pattern Recognition*, Volume 95, 2019, Pages 151-161.
- [2] R. Maslii, O. Kyrylenko, Y. Marushchak, «Analysis of methods of person reidentification in multi camera environment», *Norwegian Journal of development of the International Science*, №47/2020.
- [3] R. Layne, T. Hospedales, S. Gong, «Towards person identification and re-identification with attributes», *School of EECS*, Queen Mary University of London, UK, 2012.
- [4] Z. Zheng, L. Zheng, Y. Yang, «A discriminatively learned cnn embedding for person reidentification», *ACM Transactions on Multimedia Computing, Communications, and Applications* 14 (1) (2017) 13.
- [5] K. He, X. Zhang, S. Ren, J. Sun, «Deep residual learning for image recognition», in *Conference on Computer Vision and Pattern Recognition*, 2016, pp. 770–778.

- [6] H. Fan, L. Zheng, C. Yan, Y. Yang, «Unsupervised person re-identification: Clustering and fine-tuning», *ACM Transactions on Multimedia Computing, Communications, and Applications TOMCCAP* 14 (4).
- [7] R. R. Variator, M. Haloi, G. Wang, «Gated siamese convolutional neural network architecture for human re-identification», in *European Conference on Computer Vision*, 2016, pp. 791–808.
- [8] T. Xiao, H. Li, W. Ouyang, X. Wang, «Learning deep feature representations with domain guided dropout for person re-identification», in *The IEEE Conference on Computer Vision and Pattern Recognition*, 2016, pp. 1249 – 1258.
- [9] S. Zhou, J. Wang, D. Meng, X. Xin, Y. Li, Y. Gong, N. Zheng, «Deep selfpaced learning for person re-identification», *Pattern Recognition* 76 (2018), 739–751.
- [10] L. Zhu, Z. Xu, Y. Yang, A. G. Hauptmann, «Uncovering the temporal context for video question answering», *International Journal of Computer Vision* 124 (3) (2017) 409–421.
- [11] L. Ma, X. Yang, D. Tao, «Person re-identification over camera networks using multi-task distance metric learning», in *The IEEE Transactions on Image Processing* 23 (8) (2014) 3656–3670.
- [12] W. Li, R. Zhao, T. Xiao, X. Wang, «Deepreid: Deep filter pairing neural network for person re-identification», in *The IEEE Conference on Computer Vision and Pattern Recognition*, 2014, pp. 152–159.
- [13] C. Su, S. Zhang, F. Yang, G. Zhang, Q. Tian, W. Gao, L. S. Davis, «Attributes driven tracklet-to-tracklet person re-identification using latent prototypes space mapping», *Pattern Recognition* 66 (2017) 4–15.
- [14] C. Su, F. Yang, S. Zhang, Q. Tian, L. S. Davis, W. Gao, «Multi-task learning with low rank attribute embedding for multi-camera person reidentification», in *The IEEE transactions on pattern analysis and machine intelligence*, 40 (5) (2018) 1167–1181.
- [15] Franco, L. Oliveira, Convolutional covariance features: Conception, integration and performance in person re-identification, *Pattern Recognition* 61 (2017) 593–609.
- [16] Su, S. Zhang, J. Xing, W. Gao, Q. Tian, «Multi-type attributes driven multi-camera person re-identification», *Pattern Recognition* 75 (2018) 77– 89.
- [17] Schumann, R. Stiefelhagen, «Person re-identification by deep learning attribute-complementary information», in *The IEEE Conference on Computer Vision and Pattern Recognition Workshops*, 2017, pp. 1435–1443.
- [18] L. Zheng, L. Shen, L. Tian, S. Wang, J. Wang, Q. Tian, «Scalable person re-identification: A benchmark», in *Proceedings of the IEEE International Conference on Computer Vision*, Santiago, Chile, 7–13 December 2015.
- [19] E. Ristani, F. Solera, R. Zou, R. Cucchiara, C. Tomasi, «Performance Measures and a Data Set for Multi-Target, Multi-Camera Tracking», in *Proceedings of the European Conference on Computer Vision workshop on Benchmarking Multi-Target Tracking*, Amsterdam, The Netherlands, 8–10 October 2016.
- [20] Y. Deng, P. Luo, C.C. Loy, X. Tang, «Pedestrian attribute recognition at far distance», in *Proceedings of the 22nd ACM International Conference on Multimedia*, Orlando, FL, USA, 3–7 November 2014; pp. 789–792.
- [21] Z. Yin, Wei-Shi Zheng, A. Wu, Hong-Xing Yu, Hai Wan, X. Guo, F. Huang, J. Lai, «Adversarial attribute-image person re-identification», *IJCAI*, 2018.
- [22] D. Li, Z. Zhang, X. Chen, H. Ling, K. Huang, «A richly annotated dataset for pedestrian attribute recognition», *arXiv 2016*, arXiv:1603.07054.
- [23] D. Li, Z. Zhang, X. Chen, K. Huang, «A richly annotated pedestrian dataset for person retrieval in real surveillance scenarios», in *IEEE Trans. Image Process.* 2018, 28, 1575–1590.
- [24] X. Liu, H. Zhao, M. Tian, L. Sheng, J. Shao, Yi. Shuai, J. Yan, X. Wang, «Hydraplus-net: Attentive deep features for pedestrian analysis», in *Proceedings of the IEEE International Conference on Computer Vision*, Venice, Italy, 22–29 October 2017, pp. 350–359.
- [25] L. Wu, Y. Wang, J. Gao, X. Li, «Deep adaptive feature embedding with local sample distributions for person re-identification», *Pattern Recognition* 73 (2018) 275–288.
- [26] J. Xu, R. Zhao, F. Zhu, H. Wang, W. Ouyang, «Attention-aware compositional network for person re-identification», in: *The IEEE Conference on Computer Vision and Pattern Recognition*, 2018.
- [27] I. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, Y. Bengio, «Generative adversarial nets», in *Advances in neural information processing systems*, 2014, pp. 2672–2680.
- [28] E. Yaghoubi, D. Borza, SV Aruna Kumar, J. C. Neves, «Human attribute recognition – a comprehensive survey», *Applied Sciences* 10(16):5608, 2020.

Стаття надійшла: 20.11.2020.

References

- [1] Y. Lina, L. Zhengb, Z. Zhenga, Yu Wua, Z. Hua, Ch. Yanc , Yi Yang, «Improving person re-identification by attribute and identity learning», *Pattern Recognition*, Volume 95, 2019, Pages 151-161.
- [2] R. Maslii, O. Kyrylenko, Y. Marushchak, «Analysis of methods of person reidentification in multi camera environment», *Norwegian Journal of development of the International Science*, №47/2020.
- [3] R. Layne, T. Hospedales, S. Gong, «Towards person identification and re-identification with attributes», *School of EECS*, Queen Mary University of London, UK, 2012.
- [4] Z. Zheng, L. Zheng, Y. Yang, «A discriminatively learned cnn embedding for person reidentification», *ACM Transactions on Multimedia Computing, Communications, and Applications* 14 (1) (2017) 13.
- [5] K. He, X. Zhang, S. Ren, J. Sun, «Deep residual learning for image recognition», in *Conference on Computer Vision and Pattern Recognition*, 2016, pp. 770–778.
- [6] H. Fan, L. Zheng, C. Yan, Y. Yang, «Unsupervised person re-identification: Clustering and fine-tuning», *ACM Transactions on Multimedia Computing, Communications, and Applications TOMCCAP* 14 (4).
- [7] R. R. Varior, M. Haloi, G. Wang, «Gated siamese convolutional neural network architecture for human re-identification», in *European Conference on Computer Vision*, 2016, pp. 791–808.
- [8] T. Xiao, H. Li, W. Ouyang, X. Wang, «Learning deep feature representations with domain guided dropout for person re-identification», in *The IEEE Conference on Computer Vision and Pattern Recognition*, 2016, pp. 1249 – 1258.
- [9] S. Zhou, J. Wang, D. Meng, X. Xin, Y. Li, Y. Gong, N. Zheng, «Deep selfpaced learning for person re-identification», *Pattern Recognition* 76 (2018), 739–751.
- [10] L. Zhu, Z. Xu, Y. Yang, A. G. Hauptmann, «Uncovering the temporal context for video question answering», *International Journal of Computer Vision* 124 (3) (2017) 409–421.
- [11] L. Ma, X. Yang, D. Tao, «Person re-identification over camera networks using multi-task distance metric learning», in *The IEEE Transactions on Image Processing* 23 (8) (2014) 3656–3670.
- [12] W. Li, R. Zhao, T. Xiao, X. Wang, «Deepreid: Deep filter pairing neural network for person re-identification», in *The IEEE Conference on Computer Vision and Pattern Recognition*, 2014, pp. 152–159.
- [13] C. Su, S. Zhang, F. Yang, G. Zhang, Q. Tian, W. Gao, L. S. Davis, «Attributes driven tracklet-to-tracklet person re-identification using latent prototypes space mapping», *Pattern Recognition* 66 (2017) 4–15.
- [14] C. Su, F. Yang, S. Zhang, Q. Tian, L. S. Davis, W. Gao, «Multi-task learning with low rank attribute embedding for multi-camera person reidentification», in *The IEEE transactions on pattern analysis and machine intelligence*, 40 (5) (2018) 1167–1181.
- [15] Franco, L. Oliveira, Convolutional covariance features: Conception, integration and performance in person re-identification, *Pattern Recognition* 61 (2017) 593–609.
- [16] Su, S. Zhang, J. Xing, W. Gao, Q. Tian, «Multi-type attributes driven multi-camera person re-identification», *Pattern Recognition* 75 (2018) 77– 89.
- [17] Schumann, R. Stiefelhagen, «Person re-identification by deep learning attribute-complementary information», in *The IEEE Conference on Computer Vision and Pattern Recognition Workshops*, 2017, pp. 1435–1443.
- [18] L. Zheng, L. Shen, L. Tian, S. Wang, J. Wang, Q. Tian, «Scalable person re-identification: A benchmark», in *Proceedings of the IEEE International Conference on Computer Vision*, Santiago, Chile, 7–13 December 2015.
- [19] E. Ristani, F. Solera, R. Zou, R. Cucchiara, C. Tomasi, «Performance Measures and a Data Set for Multi-Target, Multi-Camera Tracking», in *Proceedings of the European Conference on Computer Vision workshop on Benchmarking Multi-Target Tracking*, Amsterdam, The Netherlands, 8–10 October 2016.
- [20] Y. Deng, P. Luo, C.C. Loy, X. Tang, «Pedestrian attribute recognition at far distance», in *Proceedings of the 22nd ACM International Conference on Multimedia*, Orlando, FL, USA, 3–7 November 2014; pp. 789–792.
- [21] Z. Yin, Wei-Shi Zheng, A. Wu, Hong-Xing Yu, Hai Wan, X.Guo, F. Huang, J. Lai, «Adversarial attribute-image person re-identification», *IJCAI*, 2018.
- [22] D. Li, Z. Zhang, X. Chen, H. Ling, K. Huang, «A richly annotated dataset for pedestrian attribute recognition», *arXiv 2016*, arXiv:1603.07054.
- [23] D. Li, Z. Zhang, X. Chen, K. Huang, «A richly annotated pedestrian dataset for person retrieval in real surveillance scenarios», in *IEEE Trans. Image Process.* 2018, 28, 1575–1590.

- [24] X. Liu, H. Zhao, M. Tian, L. Sheng, J. Shao, Yi. Shuai, J. Yan, X. Wang, «Hydraplus-net: Attentive deep features for pedestrian analysis», in *Proceedings of the IEEE International Conference on Computer Vision*, Venice, Italy, 22–29 October 2017, pp. 350–359.
- [25] L. Wu, Y. Wang, J. Gao, X. Li, «Deep adaptive feature embedding with local sample distributions for person re-identification», *Pattern Recognition* 73 (2018) 275–288.
- [26] J. Xu, R. Zhao, F. Zhu, H. Wang, W. Ouyang, «Attention-aware compositional network for person re-identification», in: *The IEEE Conference on Computer Vision and Pattern Recognition*, 2018.
- [27] I. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, Y. Bengio, «Generative adversarial nets», in *Advances in neural information processing systems*, 2014, pp. 2672–2680.
- [28] E. Yaghoubi, D. Borza, SV Aruna Kumar, J. C. Neves, «Human attribute recognition – a comprehensive survey», *Applied Sciences* 10(16):5608, 2020.

Відомості про авторів

Квстний Роман Наумович – доктор технічних наук, професор, зав. кафедри автоматизації та інтелектуальних інформаційних технологій Вінницького національного технічного університету.

Маслій Роман Васильович – кандидат технічних наук, доцент кафедри автоматизації та інтелектуальних інформаційних технологій Вінницького національного технічного університету.

Кириленко Олександр Михайлович – аспірант кафедри автоматизації та інтелектуальних інформаційних технологій Вінницького національного технічного університету.

Г. Н. Кветный, Р. В. Маслий, А. М. Кириленко

ИССЛЕДОВАНИЕ АТРИБУТОВ ЧЕЛОВЕКА ДЛЯ ЗАДАЧИ ПОВТОРНОЙ ИДЕНТИФИКАЦИИ

Винницкий национальный технический университет, Винница

R. N. Kvyetnyy, R. V. Maslii, O. M. Kyrylenko

RESEARCH OF HUMAN ATTRIBUTES FOR THE TASK OF RE-IDENTIFICATION

Vinnitsia National Technical University, Vinnitsia