

РЕИДЕНТИФИКАЦИЯ ЛЮДЕЙ НА ОСНОВЕ КОМБИНИРОВАНИЯ ПРИЗНАКОВ ИЗОБРАЖЕНИЙ ФИГУРЫ И ЛИЦА

Н.А. ТОМАШЕВИЧ, Р.П. БОГУШ, С.А. ИГНАТЬЕВА

Полоцкий государственный университет имени Евфросинии Полоцкой, Республика Беларусь

Поступила в редакцию 4 апреля 2025

Аннотация. В статье предлагается алгоритм повторной идентификации людей в системах, который использует дескриптор для описания особенностей человека на основе нейросетевых признаков его лица и фигуры. Для описания фигуры человека используются глобальный и три локальных нейросетевых дескриптора, которые характеризуют верхнюю, среднюю и нижнюю части изображения. Предложенный подход обеспечивает высокую точность реидентификации при возможности идентификации человека по лицу и позволяет выполнять повторную идентификацию в случае скрытых лиц и частей фигуры человека. Представлены результаты экспериментов по оценке точности реидентификации с использованием предложенного алгоритма.

Ключевые слова: сверточные нейронные сети, идентификация по лицу, дескриптор, ResNet, PolReID

Введение

Непрерывное совершенствование технологий в области компьютерного зрения и машинного обучения позволяет успешно решать все больший объем различных прикладных задач. Среди них реидентификация людей, предполагающая поиск человека по запросу на основе его изображения [1], которая востребована во многих областях включая: мониторинг общественных пространств, например, супермаркетов для анализа предпочтений клиентов и оптимизации обслуживания; розыск пропавших людей даже при частичном закрытии лица или некотором изменении внешнего вида; в управлении транспортными потоками в аэропортах, вокзалах и других транспортных узлах системы реидентификации для анализа перемещения людей с целью предотвращения перегрузок каналов перемещения. Стремительное внедрение систем видеонаблюдения и, соответственно, увеличение объема доступной видеоинформации также свидетельствуют о востребованности и актуальности создания эффективных алгоритмов реидентификации людей. Для этого могут быть использованы подходы, основанные на признаках лица, которые демонстрируют высокую точность в условиях, когда оно не закрыто масками, головными уборами или другими объектами, а человек смотрит в видеокамеру. Однако такие условия не всегда обеспечиваются при мониторинге. Например, в общественных местах люди часто носят очки, шляпы или капюшоны, что значительно снижает эффективность алгоритмов распознавания человека по лицу. Методы реидентификации, основанные на признаках изображения всего тела человека, менее подвержены влиянию скрытия лица. В работе [2] представлены результаты исследований, которые показали, что размытие лица человека незначительно влияет на результат реидентификации, когда она выполняется по признакам изображения всего тела человека. Однако, точность реидентификации при таком подходе может значительно уменьшиться из-за изменчивости внешнего вида человека, например, когда один и тот же человек сменил одежду, аксессуары и в других случаях, когда его общие признаки существенно изменяются, или обнаружены два человека в одинаковой одежде [3]. В настоящее время активно развиваются алгоритмы реидентификации, предложено ряд подходов для повышения точности [4, 5], в том числе и с использованием анализа признаков лиц [6], однако

максимально возможное значение не достигнуто, а в сложных указанных случаях точность низкая. Таким образом, учитывая недостатки каждого из существующих подходов, комбинирование признаков лица и признаков фигуры человека в алгоритме реидентификации людей представляет собой перспективное направление исследований для повышения точности решения указанной задачи.

Формализация алгоритма реидентификации людей на основе признаков фигуры и лица

Разработан алгоритм реидентификации людей, который использует дескриптор для описания особенностей человека на основе признаков его лица и фигуры. Первоначально выполняется обнаружение людей на кадрах видеопоследовательностей и описание области ограничительной рамкой. Затем определяются ключевые точки для изображений каждого человека и формируются глобальные и локальные признаки. В пределах ограничительной рамки выполняется поиск лица, оценивается направление взгляда человека: если он смотрит в видеокамеру, то формируются признаки лица, которые используются для формирования групп людей при реидентификации, а если нет, то реидентификация выполняется на основе признаков фигуры человека на изображении. Таким образом, алгоритм предполагает выполнение следующих шагов обработки для каждого кадра видеопоследовательности:

1. Обнаружение людей на кадре с использованием сверточной нейронной сети (СНС) YOLOv7, формирование ограничивающей рамки локализованной области (A_p) и ключевых точек ($A_1 - A_7$) для каждого человека. Следующие шаги выполняются для каждого обнаруженного человека.

2. Вычисление по ключевым точкам ограничивающих рамок верхней, нижней и средней части человека, и формирование советующих масок для них (M_T, M_M, M_B). В случае, если какая-то из указанных областей скрыта и не может быть локализована, то маска заполняется нулями и вектор признаков такой области не используется в дальнейшем анализе. Список локализованных областей сохраняется в булевы переменные (E_T, E_M, E_B). Далее определяется возможность использования глобального признака на основе выражения

$$E_p = (E_T \wedge E_M \wedge E_B) \quad (1)$$

3. Для извлечения признаков используется сверточная нейронная сеть DenseNet-121, на вход которой подаётся часть изображения, содержащаяся внутри ограничительной рамки (A_p) и масок (M_T, M_M, M_B). На выходе формируется общий вектор признаков F изображения фигуры человека размером 2048 элементов: F_p (глобальный дескриптор), F_T (дескриптор верхней части), F_M (дескриптор средней части), F_B (дескриптор нижней части).

4. Обнаружение лица в локализованной области. Следует отметить, что лицо считается найденным только тогда, когда все указанные ниже условия успешно выполняются.

4.1. На основе ключевых точек A_1 (нос), A_2 (левый глаз), A_3 (правый глаз), A_4 (левое ухо), A_5 (правое ухо) определяется примерная область лица. Данный шаг не позволяет получить точную локализованную область, но приводит к сокращению вычислительных затрат при последующей обработке. При этом уверенность для точек ($A_1 - A_3$) должна быть не ниже 0,5, а для A_4 или A_5 пороговая величина равна 0,2.

4.2. Если уверенность СНС при обнаружении A_4 больше, чем A_5 , то $A_{ear} = A_4$, $A_{eye} = A_2$, иначе $A_{ear} = A_5$, $A_{eye} = A_3$.

4.3. Проверяется, смотрит ли человек в сторону камеры. Для этого вычисляются расстояние D_{ear} между точками A_{ear} и A_{eye} и расстояние D_{nose} между точками A_1 и A_2 . Проверяется условие

$$\frac{D_{nose}}{D_{ear}} > 0,9 \quad , \quad (2)$$

если данное условие выполняется, принимается решение, что человек смотрит в камеру и признаки лица могут быть сформированы и использованы при реидентификации.

4.4. Используется СНС MTCNN для получения более точной ограничивающей рамки лица B_{face} .

5. Для B_{face} применяется СНС buffalo_1 от InsightFace, в результате формируется вектор L из 512 признаков лица.

6. Назначается номер группы вектору признаков следующим образом:

6.1. Составляется массив S из векторов признаков на заданном расстоянии от F_p из хранящихся в древовидной структуре данных T .

6.2. Составляется массив векторов признаков S_{free} из имеющихся в S векторов, которым не назначена группа.

6.3. Подсчитывается количество векторов в S , которое принадлежит каждой группе.

6.4. Выбирается группа S_{group} , которой принадлежит наибольшее количество векторов.

6.5. Если количество принадлежащих массиву S_{free} векторов больше, чем принадлежащих группе S_{group} , то:

6.5.1. Проверяется условие: количество не принадлежащих ни к одной группе векторов больше или равно заданной пороговой величине.

6.5.2. Если условие 6.5.1 выполняется, то создаётся новая группа S_{group} и назначается текущему вектору признаков, а также каждому вектору из S_{free} .

6.5.3. Если условие 6.5.1 не выполняется, то вектор не имеет назначенной группы.

6.6. Если количество векторов в S_{free} меньше, чем в S_{group} , то:

6.6.1. Вектору F_p назначается группа S_{group} .

6.6.2. Находящиеся в S_{free} вектора удаляются из дерева T , чтобы избежать создания новых групп рядом с существующей.

6.7. Новый вектор добавляется в T .

7. Если были сформированы признаки лица L на шаге 5, этому вектору назначается группа, при этом используется тот же алгоритм, что и для признаков всего изображения.

8. Возвращается уникальный идентификатор по признакам лица для текущего изображения человека.

9. Если были сформированы признаки лица на шаге 5 и для них назначена группа на шаге 7, возвращается номер группы, к которой они были добавлены;

10. Иначе, если на шаге 6 назначена группа по признакам всего изображения и, хотя бы одно изображение в данной группе имеет уникальный идентификатор, то возвращается данный идентификатор.

Результаты исследований

Эксперименты выполнены на видеопоследовательностях, полученных с различных камер видеонаблюдения, на которых зафиксированы одни и те же люди, но в разной одежде и в разное время. Сравнение предложенного алгоритма реидентификации, использующего признаки лиц и фигуры человека, с алгоритмом, использующим только признаки фигуры человека выполнено на базе изображений PolReID_990, которая была получена на основе набора из [7], из которой были исключены изображения людей, лица которых в масках. Количество изображений для обучения не изменялось по сравнению PolReID_1077, уменьшено на количество изображений различных людей на 87 для тестирования. В результате в галерее 15020 изображений, а в качестве запросов использовалось 1633 изображения. Следует отметить, что каждая группа может содержать множество изображений одного и того же человека, полученных в разное время, а

также изображения этого человека, полученные в разных местах в такой же одежде. Сложность реидентификации возрастает в случае смены одежды или значительного скрытия части изображения человека другими объектами. На рис. 1 показаны результаты разбиения на группы людей в таких случаях.



Рис. 1. Примеры разделения изображений на группы для сложных случаев

Пример распределения людей по группам на основе признаков изображений лиц и фигур людей показан на рис. 2, анализа которого свидетельствует о повышении эффективности такой процедуры. Например, получено правильное объединение групп «0», «2», «9», «4», «7», «13», «11», «12», «15» из рис. 1.

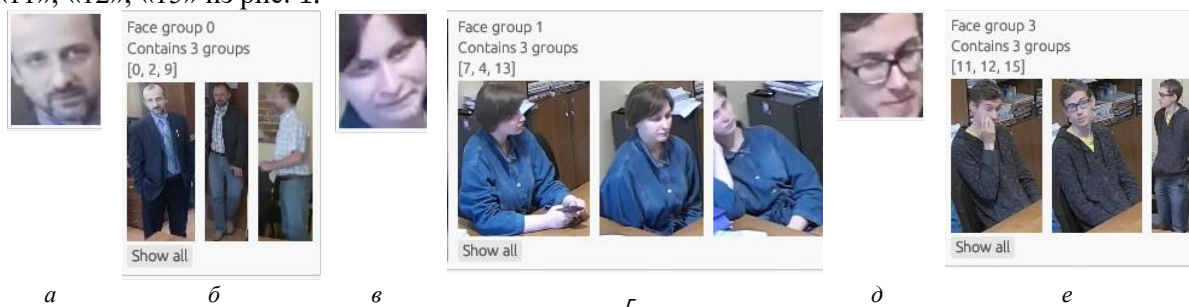


Рис. 2. Примеры объединения людей в группы с учетом признаков лица: а, в, д – изображения лиц идентифицированных людей; б, г, е – объединенные группы

Известно, что оценка эффективности реидентификации людей чаще всего проводится на основе анализа метрик включающая RankN и mAP. Метрики RankN показывают процент от числа запросов, для которых был получен правильный ответ среди первых N результатов. В результате экспериментов для разработанного алгоритма рассчитана метрика Rank1, оценивающая процент запросов, для которых идентификатор первого изображения-кандидата совпадает с идентификатором запроса

$$Rank1 = \frac{|^1Q|}{q}, \quad (3)$$

где $|^1Q|$ – количество запросов, для которых верное решение совпадает с первым выданным результатом; q – общее число запросов.

Результаты экспериментов показали, что на базе данных PolReID алгоритм реидентификации на основе признаков изображения фигуры человека характеризуется метрикой Rank1= 96,6 %, применение разработанного алгоритма позволяет получить значение данной метрики Rank1=97,64 %. Таким образом предложенный алгоритм является перспективным для реидентификации людей в распределённых системах видеонаблюдения.

Заключение

В работе предложен новый алгоритм реидентификации, который требует формирования нейросетевого дескриптора из 2048 значений для изображения фигуры человека, а также получения нейросетевого дескриптора из 512 значений для описания изображения лица. За счет этого достигается увеличение точности при возможности распознавания лиц и обеспечивается реидентификация людей при скрытых лицах на основе анализа признаков изображений их фигур. Выполненные эксперименты на наборе данных PolReID_990 подтвердили эффективность предложенного алгоритма, достигнуто увеличение значения метрики Rank1.

PERSON RE-IDENTIFICATION IN SURVEILLANCE IMAGES BASED ON COMBINATION OF BODY AND FACE FEATURES

N.A. TOMASHEVICH, R.P. BOHUSH, S.A. IHNATSYEVA

Abstract. The paper proposes an algorithm for re-identification of people in systems that uses a descriptor to describe human features based on neural network features of their face and figure. A global and three local neural network descriptors are used to describe a person's figure. The proposed approach ensures high accuracy of re-identification when it is possible to identify a person by face and allows re-identification in the case of hidden faces and parts of a person's figure.

Keywords: convolutional neural networks, face identification, descriptor, ResNet, PolReID

Список литературы

1. Chen H., Ihnatsyeva S. A., Bohush R. P., Ablameyko S.V. // Automation and Remote Control. 2023. Vol. 84, № 5. P. 497–528.
2. Dietlmeier J., Antony J., McGuinness K., O'Connor N.E. // Proc. of ICPR. 2021. P.6 912-6919.
3. Arkushin D., Cohen B., Peleg S., Fried O. // Proc. of. WACVW.2024. P. 143-153.
4. Ye S., Bohush R., Ihnatsyeva S., Ablameyko S.V. // Optical memory and Neural Network. 2023. № 4. P. 233–246.
5. Bohush R., Ihnatsyeva S., Ablameyko S. // Machine Graphics and Vision. 2022. № 31 (1/4). P. 93–109.
6. Singh N.K., Khare M., Bharadwaj H., Jethva H.B. // International Journal of Computer Applications. 2024.Vol.186, №28. P. 21-29.
7. Игнатъева С. А., Богуш Р. П. // Доклады БГУИР. 2023. Т. 21, № 3. С. 87–95.