

## Состав и особенности нового набора данных для решения задачи повторной идентификации человека в системах видеонаблюдения

С. А. Игнатъева, email: ignateva604@gmail.com

Р. П. Богуш, email: r.bogush@psu.by

Полоцкий государственный университет, Беларусь

***Аннотация.** В данной работе предлагается новый набор данных PolReID для обучения и тестирования нейросетевых алгоритмов повторной идентификации человека на изображениях, полученных из видео систем наблюдения. Выполнена и представлена оценка обучающей способности предлагаемого набора данных путем сравнения точности работы алгоритма повторной идентификации, обученного на PolReID и на распространенном наборе данных Market-1501.*

***Ключевые слова:** сверточные нейронные сети, набор данных PolReID, Market-1501, повторная идентификация человека.*

### Введение

Повторная идентификация человека представляет собой поиск интересующего человека на кадрах, полученных с нескольких камер видеонаблюдения. Разные углы обзора и положение человека относительно камеры видеонаблюдения не позволяют выполнять идентификацию по лицу, и приходится учитывать другие дескрипторы, например такие как рост, телосложение, одежда, прическа и пр. Для извлечения признаков наиболее эффективным решением является использование сверточных нейронных сетей (СНС). Для надежной работы алгоритма повторной идентификации СНС необходимо обучить, и чем более разнообразной будет обучающая выборка, тем более эффективные признаки будут извлечены, и тем точнее будет алгоритм повторной идентификации.

Большинство наборов данных для повторной идентификации содержат изображения, где каждый человек показан с разных ракурсов, выборка содержит изображения людей разных полов и возрастов. Однако при этом они формируются на основе видеопоследовательностей, полученных с ограниченного количества статичных камер, что позволяет нейронной сети изучить не только признаки человека на изображении, но также фон и возможные окклюзии, обусловленные ландшафтом или предметами интереса. Это будет иметь положительный эффект, если

применять алгоритм повторной идентификации в тех же условиях, в которых была сформирована выборка для обучения нейронной сети. Алгоритм повторной идентификации, обученный на наборе данных сформированном в открытом пространстве будет работать лучше, если тестовая выборка так же была получена с камер наружного наблюдения. Если в обучающей выборке присутствуют только люди в летней одежде, то обученная система повторной идентификации не сможет показать удовлетворительного результата при тестировании на изображениях, полученных в другое время года. Эта проблема особенно актуальна в условиях умеренного климата, где смена времен года ярко выражена, и в случае применения алгоритма повторной идентификации человека, во избежание необходимости переобучения нейронной сети каждый сезон, для обучения необходимо использовать набор данных, охватывающий различные условия.

Мы предлагаем к рассмотрению набор данных, который был сформирован на основе видеопоследовательностей, полученных в разное время года при различных условиях с 286 различных камер, имеющих разные характеристики и расположенных на различной высоте. Все это позволяет максимально разнообразить состав набора данных, и позволяет повысить обобщающую способность СНС, обученной на такой выборке.

### **1. Подходы к формированию наборов данных для повторной идентификации**

Существующие наборы данных для повторной идентификации, отличаются размером, составом, количеством используемых камер. Для создания набора данных используются кадры, полученные с нескольких камер видеонаблюдения. В качестве изображений для обучения СНС могут использоваться отдельные кадры целиком, как в наборах данных «Large scale person search» [1] и PRW [2], или вырезанные с этих кадров прямоугольные фрагменты, т.н. ограничительные рамки, содержащие только изображение человека. Такие наборы данных могут содержать комплекты ограничительных рамок, полученных с нескольких последовательно идущих кадров, которые называются треклетами (tracklets), как например в наборах данных MARS [3], LPW [4] и др. Также базы данных могут содержать ограничительные рамки, полученные с отдельных кадров, взятых с некоторым интервалом по времени, как Market-1501 [5], CUHN01, CUHN02, CUHN03 [6-8] и др. В зависимости от типа алгоритма повторной идентификации для его обучения используется соответствующий набор данных.

Изображения для наборов данных могут быть получены при различных условиях, съемка может осуществляться на улице (Market-1501, LPW, PRID [9]) или в помещении (QMUL iLIDS, Airport [10]) с

разным количеством камер и способом их расположения. Например, в наборе данных CUHN01 изображения для каждого человека получены с двух камер, области обзора которых не пересекаются. В CUHN02 используется пять таких пар камер, а в CUHN03 изображения получены с шести камер, но для каждого человека предоставляются ограничительные рамки только с двух камер из шести. При формировании LPW использовалось три разных локации, и на первой локации было установлено три камеры, на двух других по четыре камеры. Наборы данных PRW, Market1501 и MARS были получены в одном и том же месте возле супермаркета в университете Циньхуа с одних и тех же шести камер, и отличаются только способом представления данных: кадры целиком, ограничительные рамки с изображением человека, и треклеты.

## **2. Характеристика нового набора данных PolReID**

Сформированный нами набор данных PolReID размещен в [11], включает 31 919 изображений для 271 человека, является расширенной версией набора из [12], и может быть использован в научных целях для обучения и тестирования алгоритмов повторной идентификации людей в распределенных системах видеонаблюдения. От каждого человека, представленного в наборе данных, было получено письменное информированное согласие на использование их изображений в исследовательских целях.

Изображения каждого человека получены с разного количества камер, от двух до десяти, расположенных в разных локациях, и от одной до девяти видеопоследовательности с каждой камеры. Всего использовалось 286 камер на более чем 200 локациях, на улице при разных погодных условиях и временах года (лето, осень и зима) и в помещениях, с естественным и искусственным освещением различной интенсивности. Примеры изображений представлены на рис. 1.

PolReID содержит изображения мужчин и женщин от 18 до 60 лет с нескольких ракурсов в различных локациях в разное время года. Присутствуют примеры людей в масках, с разнообразными предметами в руках (телефон, рюкзак, сумка, папка), окклюзиями. Статистическая информация о составе PolReID приведены в табл. 1.



Рис. 1. Примеры изображений из набора данных PolReID

Таблица 1

Статистические данные о составе PolReID

Характеристика		Количество человек
Пол	Мужской	194
	Женский	77
Возраст	18 – 30 лет	230
	Более 30 лет	41
Время года	Лето	47
	Осень	176
	Зима	48
Условия съемки	В помещении	119
	На улице	112
	На улице и в помещении	40
Наличие маски	В маске	125
	Без маски	137
	В маске и без маски	9

Каждое изображение представляет собой ограничительную рамку, полученную с помощью сверточной нейронной сети YOLOv4 [13] из кадров на видео с разными временными интервалами. Ограничительные рамки не имеют фиксированного размера. Их размер определяется разрешением камеры, с которой было получено изображение и расстоянием между человеком и камерой.

В табл. 2 представлено сравнение PolReID с некоторыми существующими наборами данных.

Таблица 2

*Сравнительная таблица наборов данных для повторной идентификации.*

Набор данных	Количество камер	Количество человек	Количество ограниченных рамок	Размер изображений
Large scale person search	6	8432	99 809	Различный
PRW	6	932	34 304	Различный
MARS	6	1 261	1 191 003	256×128
LPW	3,4,4	2 731	592 438	256×128
Market1501	6	1 501	32 217	128×64
CUHN01	2	971	3 884	160×60
CUHN02	10(5пар)	1 816	7 264	160×60
CUHN03	6	1 360	13 164	Различный
PRID	2	934	24 541	128×64
QMUL iLIDS	2	119	476	Различный
Airport	6	9 651	39 902	128×64
PolReID	286	271	31 919	от 200×70 до 2957×1800

### 3. Результаты экспериментов

При проведении экспериментов используется алгоритм повторной идентификации предложенный в [14], который реализован на фреймворке PyTorch. С помощью СНС извлекаются дескрипторы для каждого изображения человека в галерее тестовой выборки, и формируется таблица признаков. Для каждого изображения-запроса, поступающего на вход алгоритма повторной идентификации так же извлекается дескриптор, и выполняется ранжирование таблицы признаков по максимальному подобию с запросом. В качестве метрики подобию используется косинусное расстояние между дескриптором запроса и дескрипторами всех остальных изображений галереи тестовой выборки. Из ранжированной таблицы признаков исключаются дескрипторы изображений, полученных с той же камеры, что и запрос и определяются

10 наиболее вероятных изображений-кандидатов. На рис. 2 представлен пример работы алгоритма повторной идентификации. Меткой «query» отмечены запросы, «true» – верные предположения, «false» – ошибочные.



Рис. 2. Пример работы алгоритма повторной идентификации при тестировании на PolREID

Для оценки точности повторной идентификации используются метрики Rank1 и mAP. Метрика RankN характеризуется процентом запросов, для которых верное решение было получено среди N первых наиболее вероятных кандидатов, т.о. метрика Rank1 показывает процент запросов, для которых идентификатор первого изображения-кандидата совпадает с идентификатором запроса. mAP (Mean Average Precision) является оценкой точности алгоритма повторной идентификации, отражающей среднее значение средней точности для всех запросов.

При тестировании в качестве экстрактора признаков были рассмотрены три архитектуры СНС, такие как ResNet-50, DenseNet-121 и DarkNet-53 без предварительного обучения и ResNet-50, DenseNet-121 с предварительным обучением. Предварительное обучение подразумевает, что при обучении задаче повторной идентификации начальные веса инициализируются не случайным образом, а загружаются из модели, которая была обучена на ImageNet для задачи классификации изображений 1000 классов. Обучение на наборах PolReID и Market1501 осуществлялось в течение 60 эпох со скоростью 0,03.

PolReID разделен на обучающую и тестовую выборку случайным образом, и обучающая выборка представлена 16770 изображениями для 145 человек, а тестовая – 15149 изображений, из которых 664 изображения являются запросами для 126 человек. В Market1501 для обучения используется 12936 ограничительных рамок для 751 человека, а для тестирования 19732 изображения для 750 человек. В качестве запросов используется 3368 изображений. Обучающие и тестовые выборки не пересекаются. Результаты тестирования представлены в табл. 3.

*Результаты тестирования алгоритма повторной  
идентификации человека на наборах данных Market1501 и  
PolReID с использованием разных СНС.*

СНС	Market1501		PolReID	
	Rank1, %	mAP, %	Rank1, %	mAP, %
<b>ResNet-50*</b>	81,3242	57,6058	83,7349	57,9885
<b>DenseNet-121*</b>	79,1865	56,3257	86,8976	59,8424
<b>DarkNet-53*</b>	80,0772	57,1173	87,5000	61,4378
<b>ResNet-50**</b>	87,7081	70,5364	94,3289	74,0913
<b>DenseNet-121**</b>	89,7821	73,4390	94,5783	76,2445

\* – СНС без предварительного обучения, \*\* – СНС с предварительным обучением на ImageNet.

Анализ табл.3 показал, что точность Rank1 и mAP повторной идентификации для PolReID выше, чем для Market1501 во всех проведенных экспериментах, а лучший показатель точности на наборе данных PolReID с использованием СНС без предварительного обучения был получен на DarkNet-53, и Rank1 выше на 7,42%, mAP на 4,32% по сравнению с точностью на наборе данных Market1501. При использовании предобученных СНС точность повторной идентификации выше при использовании DenseNet-121, а Rank1 и mAP для PolReID выше на 4,8% и 2,81% соответственно, по сравнению с Market1501.

### **Заключение**

Для формирования PolReID использовалось 286 камер, установленных на более чем 200 локациях с различной сложностью фона и в разное время года. Однако PolReID не имеет достаточной сбалансированности классов, что может привести к тому, что точность ре-идентификации для разных групп людей будет различаться. Несмотря на указанный недостаток, разнообразие изображений, разделенных на обучающую и тестовую выборки случайным образом, позволяет получить удовлетворительные результаты при тестировании алгоритма повторной идентификации с СНС обученной на PolReID. Об этом свидетельствуют проведенные эксперименты, показавшие, что точность алгоритма повторной идентификации, обученного с использованием PolReID выше, чем при использовании для обучения набора данных Market1501.

## Список литературы

1. Joint Detection and Identification Feature Learning for Person Search. / T. Xiao [and all] // (IEEE Conference on Computer Vision and Pattern Recognition (CVPR)). – 2017. – P. 3376-3385.
2. Person Re-identification in the Wild / L. Zheng [and all] // IEEE Conference on Computer Vision and Pattern Recognition (CVPR). – 2017. – P. 3346-3355.
3. Mars: A video benchmark for large-scale person re-identification. / L. Zheng [and all] // In European Conference on Computer Vision. Springer International Publishing. – 2016 October. – P. 868-884.
4. Region-based Quality Estimation Network for Large-scale Person Re-identification. / G. Song [and all] // AAAI. [Электронный ресурс]. – Режим доступа : <https://arxiv.org/pdf/1711.08766.pdf>
5. Scalable Person Re-Identification: A Benchmark / L. Zheng [and all] // IEEE International Conference on Computer Vision (ICCV). – 2015. – P. 1116-1124.
6. W. Li. Human Reidentification with Transferred Metric Learning / W. Li, R. Zhao, X. Wang // Proceedings of Asian Conference on Computer Vision (ACCV2012) [Электронный ресурс]. – Режим доступа : [https://www.researchgate.net/publication/283232323\\_People\\_re-identification\\_using\\_two-stage\\_transfer\\_metric\\_learning](https://www.researchgate.net/publication/283232323_People_re-identification_using_two-stage_transfer_metric_learning)
7. W. Li. Locally Aligned Feature Transforms across Views / W. Li., X. Wang // Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR2013) [Электронный ресурс]. – Режим доступа : [https://www.researchgate.net/publication/261276851\\_Locally\\_Aligned\\_Feature\\_Transforms\\_across\\_Views](https://www.researchgate.net/publication/261276851_Locally_Aligned_Feature_Transforms_across_Views)
8. DeepReID: Deep filter pairing neural network for person re-identification. / Li, W. [and all] // Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. – 2014. – P. 152-159.
9. Person re-identification by descriptive and discriminative classification. / Hirzer, M. [and all] // In Scandinavian conference on Image analysis. Springer Berlin Heidelberg. – 2011, May. – P. 91-102.
10. Karanam, S. A Systematic Evaluation and Benchmark for Person Re-Identification: Features, Metrics, and Datasets. / S. Karanam [and all] // [Электронный ресурс]. – Режим доступа : <https://arxiv.org/pdf/1605.09653v3.pdf>
11. PolReID [Электронный ресурс] : База данных. – Режим доступа : <https://github.com/SvetlanaIgn/PolReID>
12. Ihnatsyeva S. Joint Dataset for CNN-based Person Re-identification / S. Ihnatsyeva, R. Bohush, S. Ablameyko // Pattern Recognition and Information Processing (PRIP'2021): Proceedings of the 15th International Conference. – 21–24 Sept. 2021, Minsk, Belarus. – P.33-37
13. Bochkovskiy A. YOLOv4: Optimal Speed and Accuracy of Object Detection. / A. Bochkovskiy, Ch.-Y. Wang, H.-Y. M. Liao// ArXiv, abs/2004.10934. – 2020.
14. Person\_reID\_baseline\_PyTorch. [Электронный ресурс] : PyTorch ReID. – Режим доступа : [https://github.com/layumi/Person\\_reID\\_baseline\\_pytorch](https://github.com/layumi/Person_reID_baseline_pytorch)