Journal of Electrical and Computer Engineering Innovations
(JECEI)
Journal homepage: http://www.jecei.sru.ac.ir

SRTTU

JECEI

Research paper

Image Recreating in Improving the Performance of Architectures for Person Re-Identification

*S. H. Zahiri**, R. Iranpoor*

Department of Electrical Engineering, Faculty of Engineering, University of Birjand, Birjand, Iran.

Article Info

Article History:
Received 12 January 2024
Reviewed 27 February 2024
Revised 30 March 2024
Accepted 03 April 2024

Keywords:
Person re-identification
Deep learning
Image processing
Convolutional neural network
Computer vision
Image detection

*Corresponding Author's Email Address:
hzahiri@birjand.ac.ir

Abstract

Background and Objectives: Re-identifying individuals due to its capability to match a person across non-overlapping cameras is a significant application in computer vision. However, it presents a challenging task because of the large number of pedestrians with various poses and appearances appearing at different camera viewpoints. Consequently, various learning approaches have been employed to overcome these challenges. The use of methods that can strike an appropriate balance between speed and accuracy is also a key consideration in this research.

Methods: Since one of the key challenges is reducing computational costs, the initial focus is on evaluating various methods. Subsequently, improvements to these methods have been made by adding components to networks that have low computational costs. The most significant of these modifications is the addition of an Image Re-Retrieval Layer (IRL) to the Backbone network to investigate changes in accuracy.

Results: Given that increasing computational speed is a fundamental goal of this work, the use of MobileNetV2 architecture as the Backbone network has been considered. The IRL block has been designed for minimal impact on computational speed. By examining this component, specifically for the CUHK03 dataset, there was a 5% increase in mAP and a 3% increase in @Rank1. For the Market-1501 dataset, the improvement is partially evident. Comparisons with more complex architectures have shown a significant increase in computational speed in these methods.

Conclusion: Reducing computational costs while increasing relative recognition accuracy are interdependent objectives. Depending on the specific context and priorities, one might emphasize one over the other when selecting an appropriate method. The changes applied in this research can lead to more optimal results in method selection, striking a balance between computational efficiency and recognition accuracy.

Introduction

With the development of deep learning and continuous improvement in computational power, significant progress has been achieved in the field of object-detection, which is one of the crucial issues in the realm of machine vision. By introducing the first solutions for addressing this problem using convolutional neural networks, notable methods gradually emerged to better tackle the object-detection challenge. Person re-identification (ReID), which has recently gained substantial attention due to its extensive applications in various domains, serves as a fundamental and essential function in intelligent surveillance systems. An essential task for a network-based surveillance system consisting of distributed cameras is to associate individuals across camera views at different locations and times. This is referred to as the person ReID problem and forms the basis for many other important application programs. A comprehensive person ReID system comprises three main components: person identification, person tracking, and person retrieval [1]. Current research efforts to solve the ReID problem have mainly focused on two aspects:

1. Feature learning, or in other words, developing feature representations that are distinct for identity while being invariant to viewpoint and constant lighting conditions [2].
2. Metric learning, which practically involves developing machine learning methods to optimize the discriminative parameters of a ReID model [3].

However, achieving automatic ReID remains a significant challenge due to inherent limitations in most visually generated features from individuals' appearances. The pedestrian images captured by non-overlapping cameras usually under an uncontrolled environment and most of the images are low quality, which leads to some conventional biometrics features like gait and face are not feasible to be used for the task. In these circumstances, the appearance features of the pedestrians, which are extracted from their clothes' colors or objects carried with them, seem to be more suitable for the person ReID task. In addition, different pedestrians may share similar appearances, which often leads to the appearances of different pedestrians looked quite similar [4]. In this work, given the importance of feature extraction alongside the need for sufficient speed for real-time applications, faster architectures are employed for feature extraction. As the main component of the work, to enhance performance, a module called the Image Recreation Layer (IRL) is added to the Backbone, injecting input images between the blocks of the Backbone network. This component can be incorporated alongside various architectures. The obtained results demonstrate an improvement in the utilized method.

Related Works

The primary objective of the research conducted by SUN et al. [5] is to acquire superior features using a Part-based Convolutional Neural Network (PCB). In this section, a convolutional descriptor is generated from the input image, capturing features related to various image components. Subsequently, an enhanced method for Part Pooling (RPP) is introduced.

Additionally, Zhong et al. [6] address camera style adaptation (camStyle) as a solution to enhance data diversity and improve the performance differences among cameras. They initially utilize CycleGAN, which not only increases data variety but also introduces a significant level of noise. Zheng and colleagues propose a learning framework that employs the DG-NET network for the pairing of a generator module and a discriminator module for re-identification purposes.

Mohammed et al. [7] introduce the real-time ReID-DeePNet system, where they utilize the fusion of scores from two different deep learning models. In another study, Y. Zhu et al. [8] mimic the human visual perception process, which involves transitioning from coarse to fine-grained observation. The core of this approach involves a multi-scale structure consisting of two key elements: the Global Channel-Aware Attention (GCA) module for capturing global structural information and the Adaptively Spatial Feature Fusion (ASFF) module for highlighting distinctive features. Furthermore, they introduce a Bidirectional Pairwise Metric (BPM) loss function. The topological relationship between global and local features forms a framework that derives a novel feature representation through a graph transformation network. This representation is then trained and tested, based on the work of Wang et al. [9].

Various other approaches have also been presented to address some challenges such as attention to non-human sections, high complexity of the model and inference time, use of false labels for unsupervised methods, etc. For example, In the work of Zhu et al. [10], alignment scheme in transformer architecture is discussed and automatic transformer (AAformer) is introduced for automatic localization of human and non-human parts. Cho et al. [11], proposed a new framework for Part-based Pseudo Label Refinement (PPLR) proposed to decrease label noises with a reliable complementary relationship. In another work, two types of attention maps have been used to inform feature maps about the individual and relevant body parts, for which a holistic attention branch (HAB) and a partial attention branch (PAB) have been proposed [12].

Datasets

Research in the field of deep learning technology requires a substantial amount of data for model training. To exploit robust person ReID models, it is crucial to have the available ReID datasets with the characteristics of cluttered background, occlusions and overlapped bodies, etc [4].

Given that person ReID is a thoroughly investigated problem, numerous datasets have been made available for research purposes. Datasets like VIPeR [13], GRID [14], and CUHK01 [15] not only have a limited number of individuals but also a small number of images per individual. Most of them employ handcrafted labeling methods for person identification.

With the advancement of deep learning, small-scale datasets can no longer satisfy its training requirements. Therefore, large-scale datasets such as CUHK03[16], Market1501 [17], and DukeMTMC-reid[18] have been proposed and accepted.

The Market-1501 dataset is one of the most widely used datasets in the field of person detection and identification in images. This dataset includes 1501 different individuals in an open environment, with a total of approximately 32,217 images collected from 6 surveillance cameras equipped with various sensors. Each individual has around 6 to 20 images in the dataset, captured as full-body images. The dataset is divided into two parts: training and testing (query and gallery). The training part comprises images of the first 751 individuals, each having around 12 images. The testing part involves another set of 750 individuals, with only one query image per individual and approximately 4 to 18 gallery images for each individual [17].

CUHK03 is much larger dataset which includes 13164 images of 1360 pedestrians. Unlike existing datasets, which only provide manually cropped pedestrian images, this dataset provides automatically detected bounding boxes for evaluation close to practical applications [16].

In Fig. 1 an example of images from the two datasets used is presented.

Fig. 1: Samples of images in the datasets, with part (a) being images captured from two separate cameras in the first and second rows from the cuhk03 dataset, and part (b) from the market-1501 dataset.

To evaluate a Re-ID system, Cumulative Matching Characteristics (CMC) [19] and mean Average Precision (mAP) [17] are two widely used measurements. CMC is accurate when only one ground truth exists for each query, since it only considers the first match in evaluation process. However, the gallery set usually contains multiple groundtruths in a large camera network, and CMC cannot completely reflect the discrim- inability of a model across multiple cameras [20].

Backbone Networks

Backbone networks serve as the primary feature extractors for object-detection and person ReID tasks. These networks take images as input and output feature maps corresponding to the input image. Most backbone networks used for object detection are originally designed for classification tasks, with the fully connected layers often removed. Enhanced versions of classification networks are also available. For instance, researchers have designed new backbones tailored to better address the specific challenges of object-detection.

specific needs in tasks like [21] and [22] are addressed by modifying backbone networks, such as adding or removing layers or replacing some layers with other specialized design layers.

Due to various requirements regarding accuracy alongside efficiency, individuals can choose deeper and more compact backbones like ResNet [23], ResNeXt [24], AmoebaNet, or lightweight backbones like MobileNet [25], ShuffleNet [26], SqueezeNet [27], Xception[28], MobileNetV2 [28]. When targeting mobile devices, lightweight backbones can meet the requirements effectively.

The MobileNetV2 architecture is a deep neural network design for image processing with the goal of reducing the number of parameters and computational operations. This architecture is particularly beneficial for applications and devices with limited resources. As input images in many applications are larger than traditional matrices, the use of architectures with a high number of parameters and computational operations can lead to excessive computational resource consumption. This is where the MobileNetV2 architecture employs specific techniques to minimize resource consumption.

The main structure of MobileNetV2 is composed of Depthwise Separable Convolution layers. These layers are divided into two main stages:

Depthwise Convolution: In this stage, each input channel is convolved separately with a smaller kernel filter (usually 3x3 or 5x5). These filters are applied to each input channel and ultimately fulfill the role of a regular convolution.

Pointwise Convolution (1x1 Convolution): In this stage, 1x1 kernel filters are applied to the output of the previous stage. These filters are used to combine and transfer information between channels. Each input channel is transformed into an output channel.

The combination of these two stages allows the network to aggregate information in a concise manner while significantly reducing the number of parameters and computational operations. Therefore, MobileNetV2 is recognized as a lightweight architecture suitable for image processing on mobile devices and constrained resources [28].

Methods

In this section of the research plan, the focus is on reviewing the efforts made in the field of person ReID using feature-based partitioning methods. Feature partitioning is a common approach in person ReID identification, where feature maps are divided into several non-overlapping regions or predefined blocks based on learning local features.

Fig. 2 illustrates the overall workflow of the person ReID process. Initially, the CNN of interest is trained using the dataset to perform feature extraction. The images are divided into training and evaluation sets, with 80% used for training and 20% for evaluation. Subsequently, the network is tested using images from the query and gallery sets. One image is selected as the main image, and based on the extracted features, a decision is made regarding the similarity of individuals in the recognition task.

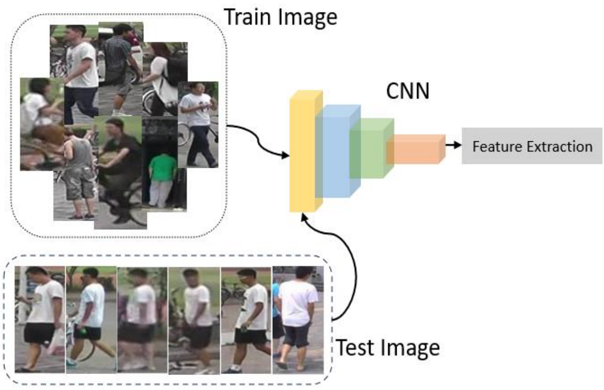

Fig. 2: General approach in person re-identification.

Here, the PCB (Part-based Convolutional Baseline) architecture has been employed to enhance common methods and align features [29]. This architecture takes the whole image as input and divides the resulting feature map into regular convolutional layers to segment it into different regions. Furthermore, the RPP (Random Partitioning Pooling) [5] method has been proposed, aiming to enhance the re-allocation of key points within each local area and strengthen stability to mitigate non-alignment issues.

In all these cases, Backbone networks that have been pre-trained on large datasets like ImageNet have been used. Architectures like ResNet [23] and DenseNet [31] are employed for tasks with larger dimensions and a need for higher computational and time resources. Additionally, lightweight architectures like MobileNetV2 [29] are used to reduce computational costs.

In this section, by modifying certain structures of the Backbone networks, the modified architectures are compared with the pre-trained ones. Due to the requirement for reduced computational costs in real-time applications, one of the main objectives of this comparison is to reduce training time. For this reason, the MobileNetV2 architecture has been implemented with modifications for use as a Backbone network for the purpose of conducting comparisons.

In the design of the MobileNetV2 architecture, key parameters such as kernel sizes and channel numbers are carefully adjusted to strike a balance between accuracy and performance. Moreover, different versions of MobileNetV2 with varying depths and complexities can be used for different tasks.

As depicted in Fig. 3, the MobileNetV2 architecture is utilized for the purpose of result comparison. The reason for choosing this architecture is its faster performance compared to other available architectures, ensuring lower computational costs and thus making it suitable for broader usage. Alongside this section, a block for image re-creation is injected into intermediate layers of the main architecture. In the Backbone network, the input image with dimensions of 256x128 and 3 channels from the color image is transformed through five stages to dimensions of 8x4 with a channel length of 1280. Finally, it is converted to dimensions of 1x1280 through average pooling. Subsequently, the final class is determined using a fully connected layer, PCB, or RPP method.

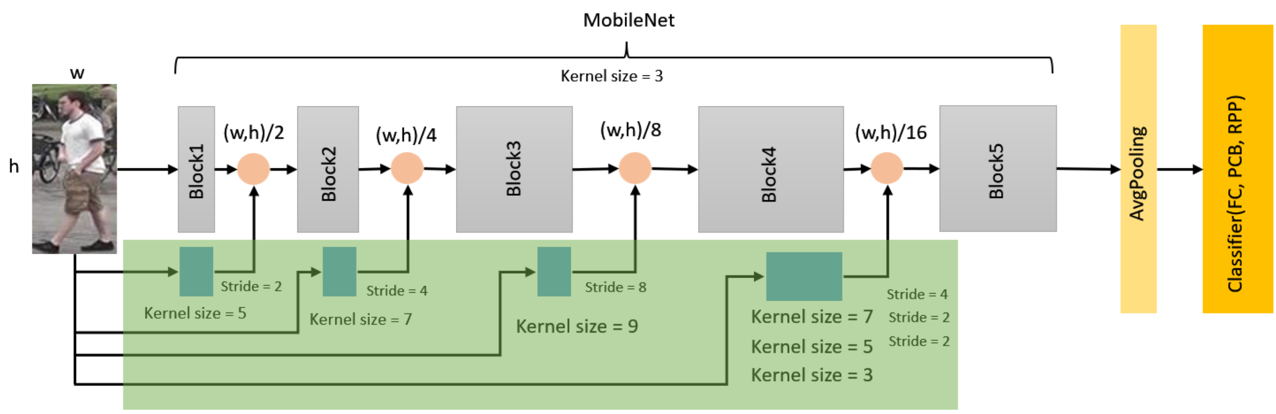

Fig. 3: The MobileNetV2 architecture, coupled with the Image Recreating Layers (IRL) block.

Image Recreating Layers (IRL):

Image recreating layers are injected between the Backbone network layers. In this process, the input image passes through several sections in parallel with kernel sizes from different layers. The kernel sizes in the first layer of each section are 5, 7, and 9 respectively, and in the last section, two layers with kernel sizes of 7 and 5 are used. In subsequent layers of each section, a kernel size of 3 is used, and the output of each section is used among the blocks of the Backbone network.

Table 1 reports details of the data used in training the network. This section contains the total count of images in the dataset and also accounts for the query and gallery images. Training images are typically divided into two groups: the training set and the evaluation set, which helps in validating the performance of the model. The table lists the number of classes utilized in training the network.

Table 1: Information related to the data used in training the network

Dataset	Number of image	Train image	Valid image	Quary image	Gallery image	Num train ID
Market-1501	36036	12185	751	3368	19732	751
Cuhk03	14097	6260	1105	1400	5332	767

Results

After applying the mentioned changes, the results obtained from the implemented experiments are summarized in Table 2 For the purpose of comparison, architectures like DenseNet and ResNet50 are pitted against MobileNetV2 along with fully connected layers.

In Fig. 4, the visual output is presented. A single image from the Query set is chosen, and other images from the same class or those closely related are selected. Finally, it's determined which image belongs to the class of the original image. This distinction is indicated by the colors green and red.

Table 2: Comparison of methods for market1501 and cuhk03 datasets

Methods	Dataset	Rank@1	Rank@5	mAP
DenseNet	market	0.877	0.955	0.698
	Cuhk03	0.498	0.517	0.362
ResNet50	market	0.869	0.943	0.702
	Cuhk03	0.502	0.518	0.373
ResNet50 +IRL	market	0.862	0.938	0.701
	Cuhk03	0.508	0.526	0.401
MobileNetV2	market	0.685	0.854	0.444
	Cuhk03	0.407	0.467	0.213
MobileNetV2 +RPP+IRL	market	0.686	0.857	0.451
	Cuhk03	0.437	0.473	0.261

Fig. 4: The visual output of the utilized method.

Additionally, the Table 2 includes results for these architectures with the added image recreating layers for CUHK03 and Market-1501 datasets. All these results were obtained under similar conditions for the Market-1501 and CUHK03 datasets.

A noteworthy point regarding the use of more complex networks like ResNet50 and DenseNet in this context is that these architectures have been pre-trained on large datasets. Moreover, they involve more layers, which significantly increases the training time as a result.

The issue of processing time is highly critical to ensure the feasibility of using these systems, especially in resource-constrained environments. Therefore, a trade-off between accuracy and processing time is considered. Although the fundamental issue in this architecture lies in the number of its layers, it has been designed to increase speed for real-world applications and for use on hardware-constrained devices. Consequently, the accuracy of this architecture is lower compared to other architectures. The incorporation of the image recreation layer alongside it can bring about a partial improvement in the network's performance. If we use the image reformation layer in larger architectures, the results may not change significantly, and in some cases, it might deviate from the primary goal. According to the results in Table 2, it can be stated that the image reformation layer is more effective on architectures with fewer layers.

According to Table 3, the execution time for each epoch during the training of various methods is presented in seconds. All these stages for all methods have been obtained under similar conditions with comparable hardware. The use of the MobileNetV2 architecture is faster by more than 50%, and this speed increase also holds true concerning the addition of the image recreation layer.

In Fig. 5 and Fig. 6, the loss and accuracy curves related to the training are presented. These curves pertain to both the training and evaluation datasets.

During training, all the mentioned cases have used the CrossEntropy loss function. The number of epochs for training is set to 60, and a batch size of 32 is considered.

Table 3: Implementation time of different methods

Dataset	Dens-Net	Res-Net50	ResNet +RPP	Mobil-NetV2	Mobile-NetV2 +RPP+IRL
Market-1501	63.7 (s)	50.3 (s)	85.5 (s)	23.2 (s)	29.3 (s)
Cuhk03	90.4 (s)	79.9 (s)	110 (s)	50.9 (s)	50.2 (s)

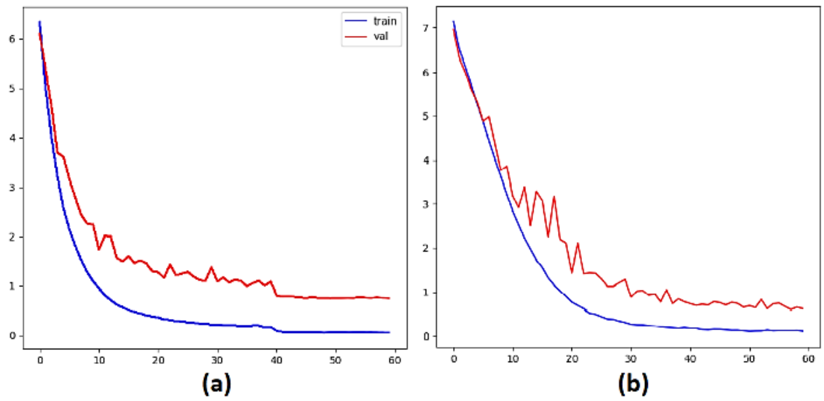

Fig. 5: The loss charts for training and evaluation data. Section (a) is for market-1501 and (b) is for cuhk03.

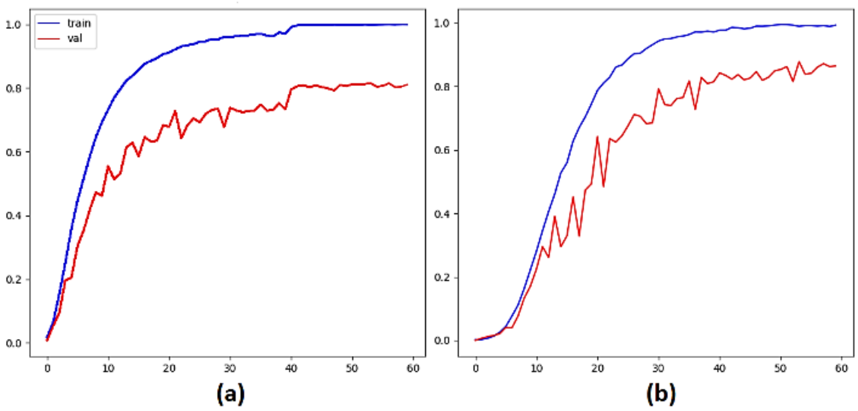

Fig. 6: The accuracy charts for training and evaluation data. Section (a) is for market-1501 and (b) is for cuhk03.

Conclusion

Person ReID is a problem with numerous practical applications in our daily lives. It is crucial to employ methods that can consider its applicability by reducing computational costs. The trade-off between speed and accuracy is also a challenge that should be addressed, and the field should be flexible enough to use various methods depending on specific needs.

Various challenges and issues exist in improving person ReID methods. One of these concerns is the limited quantity of data compared to other machine vision domains. Given the significance of input data in deep learning, challenges related to classification, identification, etc., involve vast datasets, while the person ReID problem is relatively constrained in terms of data. Furthermore, important architectures in this field have already been trained on these datasets. In the case of the MobileNetV2 architecture, it has been fully implemented and not pretrained. Another significant challenge is the quality of the data. Since these data are collected by cameras that often lack decent quality, using methods capable of establishing meaningful connections between image components is of special importance.

The use of appropriate methods such as data augmentation and the application of generative adversarial networks (GANs), among others, can be employed to increase input data and enhance the performance of networks. Additionally, the adoption of techniques like Vision Transformers, which can establish meaningful relationships across the entire image, holds promise for future endeavors.

In general, this work aims to establish an interactive relationship between hardware constraints and adequate accuracy. It is inevitable to consider that many existing methods may achieve higher accuracy, but they often come at the cost of increased computational requirements. This study strives to explore certain modifications and their implementations to effectively enhance results while taking into account this trade-off between accuracy and computational cost.

Author Contributions

Dr. Zahiri has drawn the general road map. R .Iranpoor has searched for important articles in this field. Then, by checking the results and collecting the necessary data, the implementation of the proposed method has been done. Dr. Zahiri reviewed the results and made changes in the way of implementation and final editing of the work.

Acknowledgment

The authors would like to thank the editor and anonymous reviewers.

Conflict of Interest

The authors declare no potential conflict of interest regarding the publication of this work. In addition, the ethical issues including plagiarism, informed consent, misconduct, data fabrication and, or falsification, double publication and, or submission, and redundancy have been completely witnessed by the authors.

Abbreviations

ReID	Person re-Identification
IRL	Image Recreating Layers
PPLR	Pseudo Label Refinement
HAB	Holistic Attention Branch
PAB	Partial Attention Branch
PCB	Part-based Convolutional Baseline
RPP	Random Partitioning Pooling
CMC	Cumulative Matching Characteristics
mAP	mean Average Precision

References

[1] W. Wei, W. Yang, E. Zuo, Y. Qian, L. Wang, "Person re-identification based on deep learning—An overview," J. Visual Commun. Image Represent., 82: 103418, 2022.

[2] M. Farenzena, L. Bazzani, A. Perina, V. Murino, M. Cristani, "Person re-identification by symmetry-driven accumulation of local features," in Proc. 2010 IEEE Computer Society Conference on Computer Vision and Pattern Recognition: 2360-2367, 2020.

[3] W. S. Zheng, S. Gong, T. Xiang, "Person re-identification by probabilistic relative distance comparison," in Proc. CVPR 2011: 649-656, 2011.

[4] D. Wu et al., "Deep learning-based methods for person re-identification: A comprehensive review," Neurocomput., 337: 354-371, 2019.

[5] Y. Sun, L. Zheng, Y. Yang, Q. Tian, S. Wang, "Beyond part models: Person retrieval with refined part pooling (and a strong convolutional baseline)," in Proc. the European Conference on Computer Vision (ECCV): 480-496, 2018.

[6] Z. Zhong, L. Zheng, Z. Zheng, S. Li, Y. Yang, "Camera style adaptation for person re-identification," in Proc. the IEEE Conference on Computer Vision and Pattern Recognition: 5157-5166, 2018.

[7] H. J. Mohammed et al., "ReID-DeePNet: A hybrid deep learning system for person re-identification," Math., 10(19): 3530, 2022.

[8] Y. Zhu et al., "Multiscale global-aware channel attention for person re-identification," J. Visual Commun. Image Represent., 90: 103714, 2023.

[9] L. Zhao, X. Li, Y. Zhuang, J. Wang, "Deeply-learned part-aligned representations for person re-identification," in Proc. the IEEE International Conference on Computer Vision: 3219-3228, 2017.

[10] K. Zhu et al., "Aaformer: Auto-aligned transformer for person re-identification," IEEE Trans. Neural Networks Learn. Syst., 2023.

[11] Y. Cho, W. J. Kim, S. Hong, S. E. Yoon, "Part-based pseudo label refinement for unsupervised person re-identification," in Proc. the IEEE/CVF Conference on Computer Vision and Pattern Recognition: 7308-7318, 2022.

[12] Y. Chen, H. Wang, X. Sun, B. Fan, C. Tang, H. Zeng, "Deep attention aware feature learning for person re-identification," Pattern Recognit., 126: 108567, 2022.

[13] D. Gray, H. Tao, "Viewpoint invariant pedestrian recognition with an ensemble of localized features," in Proc. ECCV 2008: 262-275, 2008.

[14] C. C. Loy, T. Xiang, S. Gong, "Multi-camera activity correlation analysis," in Proc. 2009 IEEE Conference on Computer Vision and Pattern Recognition: 1988-1995, 2009.

[15] W. Li, R. Zhao, X. Wang, "Human reidentification with transferred metric learning," in Proc. 11th Asian Conference on Computer Vision, Part I 11: 31-44, 2013.

[16] W. Li, R. Zhao, T. Xiao, X. Wang, "Deepreid: Deep filter pairing neural network for person re-identification," in Proc. the IEEE Conference on Computer Vision and Pattern Recognition: 152-159, 2014.

[17] L. Zheng, L. Shen, L. Tian, S. Wang, J. Wang, Q. Tian, "Scalable person re-identification: A benchmark," in Proc. the IEEE International Conference on Computer Vision: 1116-1124, 2015.

[18] E. Ristani, F. Solera, R. Zou, R. Cucchiara, C. Tomasi, "Performance measures and a data set for multi-target, multi-camera tracking," in Proc. European Conference on Computer Vision: 17-35, 2016.

[19] X. Wang, G. Doretto, T. Sebastian, J. Rittscher, P. Tu, “Shape and appearance context modeling,” in Proc. ICCV 2007: 1–8, 2007.

[20] M. Ye, J. Shen, G. Lin, T. Xiang, L. Shao, S. C. Hoi, "Deep learning for person re-identification: A survey and outlook," IEEE Trans. Pattern Anal. Mach. Intell., 44(6): 2872-2893, 2021.

[21] J. Redmon, S. Divvala, R. Girshick, A. Farhadi, "You only look once: Unified, real-time object detection," in Proc. the IEEE Conference on Computer Vision and Pattern Recognition: 779-788, 2006.

[22] Z. Li, C. Peng, G. Yu, X. Zhang, Y. Deng, J. Sun, "Detnet: A backbone network for object detection," arXiv preprint arXiv:1804.06215, 2018.

[23] S. Targ, D. Almeida, K. Lyman, "Resnet in resnet: Generalizing residual architectures," arXiv preprint arXiv:1603.08029, 2016.

[24] S. Xie, R. Girshick, P. Dollár, Z. Tu, K. He, "Aggregated residual transformations for deep neural networks," in Proc. the IEEE Conference on Computer Vision and Pattern Recognition: 1492-1500, 2017.

[25] A. G. Howard et al., "Mobilenets: Efficient convolutional neural networks for mobile vision applications," arXiv preprint arXiv:1704.04861, 2017.

[26] J. Zang, L. Wang, Z. Liu, Q. Zhang, G. Hua, N. Zheng, "Attention-based temporal weighted convolutional neural network for action recognition," in Proc. 14th IFIP WG 12.5 International Conference Artificial Intelligence Applications and Innovations (AIAI 2018): 97-108, 2018.

[27] F. N. Iandola, S. Han, M. W. Moskewicz, K. Ashraf, W. J. Dally, K. Keutzer, "SqueezeNet: AlexNet-level accuracy with 50x fewer parameters and< 0.5 MB model size," arXiv preprint arXiv:1602.07360, 2016.

[28] C. Fran, "Deep learning with depth wise separable convolutions," in Proc. IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2017.

[29] M. Sandler, A. Howard, M. Zhu, A. Zhmoginov, L. C. Chen, "Mobilenetv2: Inverted residuals and linear bottlenecks," in Proc. the IEEE Conference on Computer Vision and Pattern Recognition: 4510-4520, 2018.

[30] Y. Sun, L. Zheng, Y. Li, Y. Yang, Q. Tian, S. Wang, "Learning part-based convolutional features for person re-identification," IEEE Trans. Pattern Anal. Mach. Intell., 43(3): 902-917, 2019.

[31] Y. Zhu, S. Newsam, "Densenet for dense flow," in Proc. 2017 IEEE International Conference on Image Processing (ICIP): 790-794, 2017.

Biographies

Seyed Hamid Zahiri received the B.Sc., M.Sc. and Ph.D. degrees in Electronics Engineering from Sharif University of Technology, Tehran, Tarbiat Modarres University, Tehran, and Mashhad Ferdowsi University, Mashhad, Iran, in 1993, 1995, and 2005, respectively. Currently, he is a Professor with the Department of Electronics Engineering, University of Birjand, Birjand, Iran. His research interests include pattern recognition, evolutionary algorithms, swarm intelligence algorithms, and soft computing.

- Email: hzahiri@birjand.ac.ir
- ORCID: 0000-0002-1280-8133
- Web of Science Researcher ID: NA
- Scopus Author ID: NA
- Homepage: NA

Rasool Iranpoor was born on July 19, 1991. He received M.Sc. degree in Electronic Engineering from Birjand University, Birjand, Iran, in 2018. He is currently a Ph.D. student at Birjand University to receive a Ph.D. degree in Electronics Engineering. His research interests include Machine Learning, Image Processing, Computer Vision, and Deep Learning Algorithms.

- Email: rasool.iranpoor@birjand.ac.ir
- ORCID: 0000-0002-7769-259X
- Web of Science Researcher ID: NA
- Scopus Author ID: NA
- Homepage: NA

How to cite this paper:
S. H. Zahiri, R. Iranpoor, “Image recreating in improving the performance of architectures for person re-identification,” J. Electr. Comput. Eng. Innovations, 12(2): 401- 408, 2024.

DOI: 10.22061/jecei.2024.10446.706

URL: https://jecei.sru.ac.ir/article_2094.html