International Journal of Engineering

Improving Person Re-Identification Performance Using ESRGAN Image Enhancement

F. Ilbeigi, H. Farsi*, S. Mohamadzadeh, A. Sezavar

*Department of Electrical and Computer Engineering, University of Birjand, Birjand, Iran*

A B S T R A C T

Person re-identification is one of the most important challenges in the field of image processing and computer vision, which has attracted significant attention in recent years. One of the main challenges in this domain is the low quality of images captured by different cameras, which can result in the loss of important details and consequently a reduction in the accuracy of re-identification algorithms. This paper presents a framework that enhances network accuracy and capability in determining individuals' identities by improving the brightness of low-light images using histogram equalization techniques and increasing their resolution using the ESRGAN network. The experimental results show that the proposed method, by improving the brightness and resolution of images which plays a key role in revealing important details, enhances the quality of the re-identification network's input. Also, compared to existing methods, our model increases the accuracy of person re-identification, such that it achieved Rank-1 accuracy of 93.6% for the CUHK01 dataset and 94.5% for the CUHK03 dataset.

doi: 10.5829/ije.2026.39.10a.09

Graphical Abstract
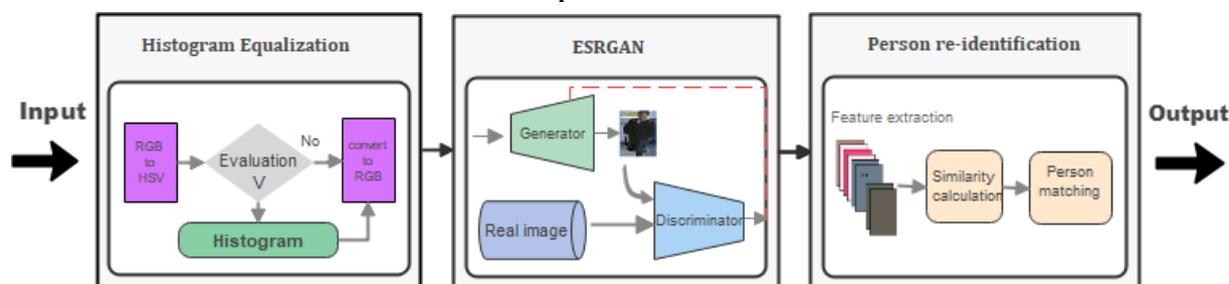
1. INTRODUCTION

Person re-identification is the process of matching and identifying an individual in photos captured by different cameras to retrieve the top-k most similar images to the probe image of a person within a large dataset (gallery) recorded by different cameras (1). In recent years, with the increasing proliferation of surveillance cameras in urban environments, the importance of person re-identification has become more evident than ever. This technique has wide-ranging applications in border control, searching for missing people, video surveillance, traffic management, and urban security. The ability to reliably identify individuals across vast networks of surveillance cameras can play an important role in enhancing security. One of the main challenges of re-identification systems is the low resolution of images, which is particularly exacerbated when cameras capture images from long distances, unfavorable angles, or under adverse weather conditions, turning person re-identification into a complex problem (2). The reduction in image quality makes it difficult to extract accurate and distinctive features that are essential for re-identifying individuals, leading to decreased accuracy and an

*Corresponding Author Email: hfarsi@birjand.ac.ir (H. Farsi)

increased error rate (3). So far, a lot of studies have been done to deal with these problems (4–9). Some methods (10, 11), instead of reconstructing specific appearance features lost due to quality degradation, focus on mapping representations in a defined feature space across different resolutions. This method does not solve the problem of unequal access to information because it is impossible to get high-resolution pixel-level supervision for pictures of people taken from different angles. Some researchers (12, 13) only use handcrafted visual features and did not use neural networks to learn and improve discriminative attributes automatically. Because of this dependence, they have lower efficiency at dealing with complex changes in data, such as changes in angles, lighting, or crowded backgrounds, while artificial neural networks (ANNs) are capable of dealing with complex patterns (14).

This paper proposes a three-step process for improving the accuracy of person re-identification with low-resolution images. The primary focus of this study is on Low-Resolution Person Re-Identification (LR Re-ID). Images with low resolution usually contain scattered information and are sensitive to minor changes (15). These images have low contrast, and their details and edges are blurred and difficult to see, which makes it harder to get basic features like color, pixel intensity, pixel gradients, orientation, and so on (16). These basic features are what allow us to get advanced and semantic attributes, like facial structures or clothing patterns, which are very important for re-identification. The first step in the suggested model is to fix images with brightness levels below a certain threshold (which is based on a certain average pixel intensity) using histogram equalization methods. This first step makes the images bright enough to make hidden details in darker areas easier to see, but the images may still not be very sharp or have a lot of fine details. After that, a deep learning-based super-resolution network processes the images. We used an ESRGAN to improve the network's ability to extract features. It does this by creating a high-resolution image from a low-resolution input, which greatly improves the quality of the images. This enables rapid extraction of distinctive appearance features from the images. Finally, the person re-identification network (17) is used for feature extraction.

2. RELATED WORKS

In this section, the studies conducted in the areas of image super-resolution, person re-identification, and the application of super-resolution for person re-identification are reviewed separately.

2. 1. Image Super Resolution Image Super-Resolution (SR) is a prevalent method to make images clearer by making a high-resolution (HR) version from a low-resolution (LR) image (18). The widespread use of deep learning has sped up the development of super-resolution methods. Zhao et al. (19) were the first to succeed in this area by creating a deep Convolutional Neural Network (CNN) that linked low-resolution (LR) and high-resolution (HR) images from start to finish by lowering reconstruction error. This study laid the groundwork for several other studies in the development of CNN-based super-resolution models. Later, Wang et al. (20) came up with a hybrid model that used clustering of different blur kernels to get a better Peak Signal-to-Noise Ratio (PSNR) than basic models. In addition, image reconstruction was more effective using aggregation-based models like MAP (21) and ESCN (22). However, the majority of these models focus solely on pixel-wise similarity, resulting in output images that lack visual appeal.

Recently Generative Adversarial Networks (GAN) are widely used in image processing. Ledig (23) used a Generative Adversarial Network (24) and came up with a perceptual loss function to make textures that look and feel more real. Later, Zhao et al. (25) and Wang et al. (26) improved the GAN architecture and learning method to provide superior high-resolution outputs. Shamsolmoali et al. (27) used the least squares loss as a discriminator loss function to verify that training was stable. They also came up with a progressive GAN that could recreate more details in images. Subsequently, models such as SRGAN frequently produced outputs with overly smooth textures. The ESRGAN model was introduced to resolve this issue (26). Residual-in-Residual Dense Blocks are used to enhance image restoration and give it a more realistic quality. This study utilizes the ESRGAN model's superior ability to replicate image structures and intricate details. Compared to older models like SRGAN, this model is more successful in enhancing the ability to see and understand images.

2. 2. Person Re-Identification The use of deep networks in person reidentification has recently garnered considerable interest from researchers (28). Early CNN architectures, through the utilization of sequential convolutional layers, were able to significantly enhance the process of feature extraction and learning (29). Sezavar et al. (30) used a CNN-based feature extractor to obtain initial features for person re-identification, followed by a sparse representation for matching individuals across diverse viewpoints that this combination has facilitated the identification of individuals possessing diverse characteristics that integrate well within the group. A combination of an autoencoder and a deep convolutional neural network is used for person re-identification (17). The CNN initiates by identifying visual features. The autoencoder reduces their size and enhances their uniqueness through

compression and reconstruction. This compression reduces noise, enhancing the model's utility across diverse scenarios.

Chen et al. (31) examined a novel deep network known as the Deep Quadruplet Network. This network uses a quadruplet loss function. Instead of the standard triplet loss, this new loss function helps the model handle the distance between positive and negative samples. It also helps create a more structured and discriminative feature space. Zhao et al. (32) introduced a network called Spindle Net that analyzes features based on different human body regions for person reidentification. First, the network divides the person's image into separate body parts. Then, it extracts features from each segment on its own and finally combines these features. Liu et al. (33) introduced a network termed the Comparative Attention Network (CAN), which is trained end-to-end for person reidentification. The network adaptively focuses on critical regions of the images and extracts significant features in a comparative manner to strengthen discrimination between individuals. The use of an attention mechanism enables the model to concentrate on the more relevant parts of the body while ignoring less useful regions. This technique makes re-identification more accurate, especially in complex scenarios with diverse backgrounds.

2. 3. Low-Resolution Person Re-Identification
In recent years, a lot of work has gone into solving the problem of person re-identification in low-resolution images. Early approaches primarily relied on dictionary learning and metric learning, but their performance was limited due to the lack of sufficient fine-grained details in low-resolution images. Many SR-based techniques have been proposed that greatly enhance matching accuracy thanks to progress in neural networks and super-resolution technology. Jiao et al. (34) were the first to put the SRCNN network and a re-identification network together into a single framework and suggest a way to train both networks at the same time. Wang et al. (35) used SR-GAN in a cascading structure by putting a number of SR-GAN modules in a row. This way, each low-resolution picture is sent to a different generator, which makes sure that all of the pictures are the same size. But neither of these methods trains SR modules together, so they don't take into account the extra features that each one finds. This makes it easy to improve them all in one place. They also did not consider generative artifacts, which hurt recognition performance, especially when the scaling factor gets bigger. Cheng et al. (36) improved the SR-ReID framework by using a better training method, which led to better results. Their method makes the two sub-networks work better together by using the natural relationship between super-resolution and re-identification. Mao et al. (10) added two parts: the first is a foreground-centric super-resolution model

(FFSR) that fixes resolution loss in low-resolution input images, and the second is a resolution-invariant person re-identification module that helps with feature extraction. Even though this method has shown promising results, it requires accurate labeling of the foreground region to guide the image recovery learning process. Chen et al. (37) came up with the Resolution-Adaptive and Identity-Preserving Network (RAIN), which uses reconstruction loss to bring the feature distributions of high-resolution and low-resolution images closer together. This helps reduce the disparity between the deep features of images at different resolutions. Li et al. (38) developed a method that integrates resolution-invariant representation learning with image super-resolution to make cross-resolution person re-identification (ReID) work better.

3. PROPOSED METHOD

Figure 1 shows a general image of the proposed network. The model consists of three fundamental components. The first part enhances the image's brightness by equalizing the histogram in the HSV color space. In the second part, the ESRGAN network is used to improve the image resolution. Finally, the reconstructed images are fed into a re-identification network, which pulls out features that are related to a person's identity. The next sections will explain the details of each phase. The proposed method is represented in Algorithm 1.

3. 1. Histogram Equalization   Many surveillance images, particularly those captured at night or in poor weather conditions, suffer from low brightness. This diminishes the contrast, complicating the visibility of essential detail such as edges and facial or bodily features. To alleviate these impacts, the initial stage is the application of a histogram equalization approach within the HSV color space. This color space displays the parts of Hue, Saturation, and Value, making it easy to change the brightness of an image directly and on its own without changing its color. In the proposed method color images are first converted from the RGB color space to the HSV color space. Next, the average brightness of the V channel is found. If the average brightness of the V dark regions. After that, the enhanced images are converted back to the RGB color space, and it is ready Channel is equal to or higher than a set threshold, the original images are sent to the next stage without any changes. Otherwise the images are considered dark, and histogram equalization is applied to enhance details in the for the next step, which is the super-resolution network. Figure 2 shows that the overall brightness of a very dark image can be improved without changing its colors. Histogram equalization alone, on the other hand, can not make images sharper or bring out fine details; it just makes the input better for the next steps in processing.
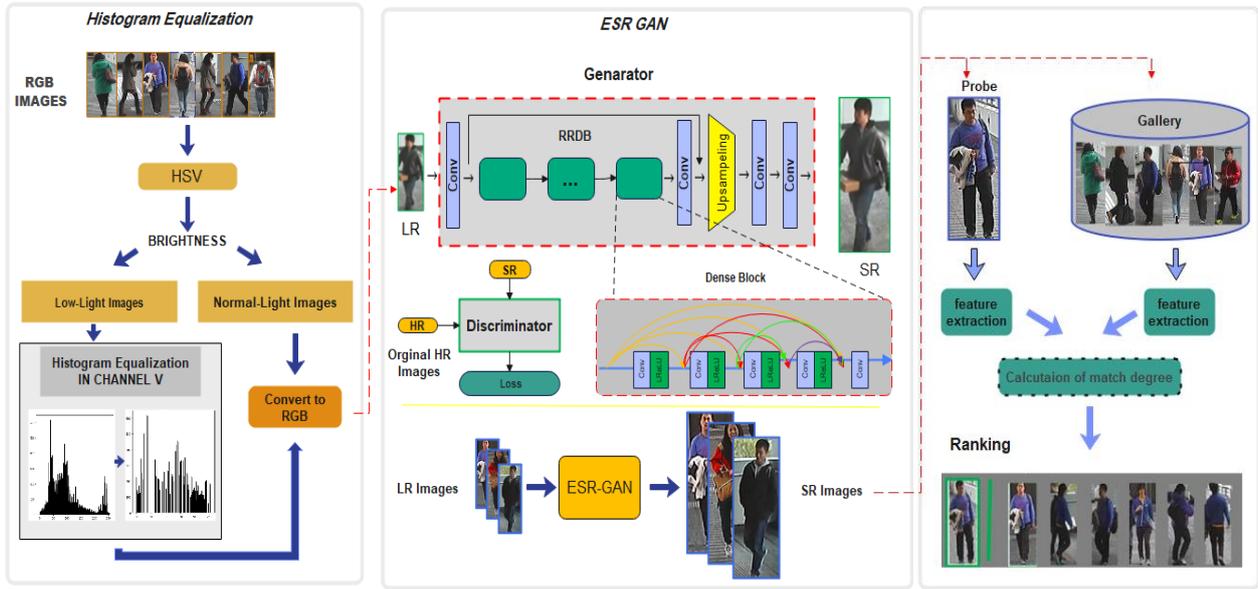
Figure 1. An illustration of the proposed network is presented, which consists of histogram equalization, a super-resolution network (ESRGAN), and a person re-identification network (ReID Network)

Algorithm 1. Algorithm of the proposed method

Step 1 – Histogram Equalization
        Input: RGB image
        Enhanced image
1.   1.Convert input images from RGB → HSV.
2.   Compute the average brightness $\bar{V}$ from the V channel.
3.   If $\bar{V}$ is small
      Apply histogram equalization on the V channel
      Reconstruct HSV (H, S, V′)
      Convert HSV back to RGB
      Else:
      Convert HSV back to RGB (without changes)
Step 2 – Super-Resolution Enhancement (ESRGAN):
        Input: Enhanced image
        Output: Super-resolved image
1.   Training procedure (for each iteration):
2.   Generating an SR image from an LR image
3.   Compute perceptual loss
4.   Compute L1 loss (SR vs HR)
5.   Compute discriminator output on HR and SR (relativistic GAN)
6.   Updating Discriminator
7.   1) Compute discriminator loss L_{Dis}^{Ra}

      2) Update parameters
8.   Updating Generator
      1) Compute generator loss

 $$L_{Gen} = \gamma_1 L_1 + \gamma_2 L_{perceptual} + \gamma_3 L_{Gen}^{Ra}$$

      2) Update parameters
9.   After training, the generator produces a high-resolution image from the enhanced input.
Step 3 –Re-Identification
        Input: Super-resolved image
        Output: Recognition result
1.   Feed the SR image into the Re-ID network to extract a feature vector from the image
2.   Compute similarity score
3.   Rank gallery images based on similarity scores
4.   Re-Identification result
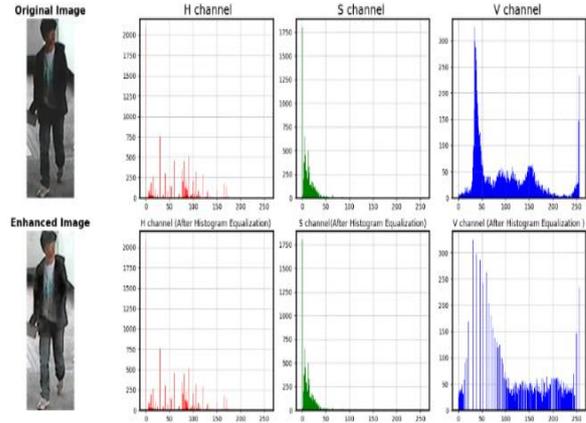
Figure 2. The original RGB image (top image) and the enhanced version after applying histogram equalization on the Value channel in the HSV color space (bottom image)

3. 2. Super Resolution Network   The next step after enhancing the image brightness is to use ESRGAN, a deep super-resolution network based on GAN to make the images clearer. ESRGAN is one of the best ways to super-resolve pictures. The structure of GAN is what this model is based on. The three main parts of ESRGAN are the generator, the discriminator, and the loss function. The generator produces high-resolution images with fine details, and the discriminator tells the difference between the original high-resolution pictures and the ones made by the generator. The loss function is also composed of three parts: pixel-wise loss, perceptual loss, and relativistic GAN loss.

ESRGAN utilizes a relativistic version of the GAN called RaGAN (Relativistic Average GAN) (35). This method uses relative discrimination to tell the difference between real and fake pictures. The discriminator doesn't make a clear decision about whether or not an image is real. Instead, it looks at how real the real image looks compared to the fake one. This function makes the training process more stable and the results better. The perceptual loss is calculated using the feature map of the image that the generator made and that was taken out before the final activation function. This loss makes the final output look more like real photos in terms of how they look and how many small details there are (39). After that, the pixel-wise error is calculated, and the resulting image is sent to the discriminator network. During training, the discriminator is also given the original image, which helps it tell the difference between real and fake images. The generator then changes its adversarial loss and makes a new version of the same image. The process goes on until the discriminator can no longer tell the difference between real and fake images. The total loss of the network is defined as follows:

$$L_{totall} = L_{Gen} + L_{Dis}^{Ra} \qquad (1)$$

where L_{Gen} is the loss of the generator, L_{Dis}^{Ra} is the loss of the relativistic discriminator, and the generator loss is the sum of the next three loss functions:

$$L_{Gen} = \gamma_1 L_1 + \gamma_2 L_{perceptual} + \gamma_3 L_{Gen}^{Ra} \qquad (2)$$

γ_1, γ_2 and γ_3 are weights for different loss types, which control the influence of each loss and are chosen approximately. $L_{perceptual}$ is the perceptual loss, L_{Gen}^{Ra} is the relativistic generator loss, and L_1 is the pixel-wise loss.

3. 2. 1. Pixel-Wise Loss
The network uses a pixel-wise loss function to enhance the pixel-level accuracy of the generated image. This helps ensure that the high-resolution image I_{SR} closely resembles the real image I_{HR} in terms of pixel values. To improve performance and convergence the loss function L_1 is used, as the standard loss function L_2 often leads to overly smooth results. L_1 loss is is computed as follows:

$$L_1 = \sum_x^W \sum_y^H \| I_{SR}(x,y) - I_{HR}(x,y) \| \qquad (3)$$

where W and H are the image's width and height, respectively. This equation finds the L1 distance between the pixels of the high-resolution image that was made and the real image.

3. 2. 2. Perceptual Loss
Perceptual loss functions by contrasting the feature representations of the original and generated images, hence enhancing the preservation of image quality and fine details. Feature maps are

derived from the generator's output prior to the final activation function, retaining high-level semantic information for the calculation of perceptual loss. This method enhances the correlation between the features of the input and the reconstructed images, leading to enhanced reconstruction quality and increased overall model performance (26). The feature maps of the high-resolution original image I_{HR} and the image generated I_{SR} by the generator are compared utilizing the Euclidean distance, and the perceptual loss is calculated based on the following equation:

$$L_{perceptual} = \sum_{x=1}^{W_{i,j}} \sum_{y=1}^{H_{i,j}} (\phi_{i,j}(I_{HR})_{xy} - \phi_{i,j}(I_{SR})_{xy})^2 \qquad (4)$$

The feature map $\phi_{i,j}$ comes from the j-th convolutional layer and the i before the i-th max-pooling layer.

3. 2 .3. Relativistic Gan Loss
The relativistic adversarial generator network estimates the probability of distinguishing between real data and generated data by calculating the following distance:

$$D_{Ra}(I_{HR}, I_{SR}) \qquad (5)$$

The Relativistic Generative Adversarial Network (RGAN) generates images of individuals with enhanced sharpness and superior visual and frequency details relative to the conventional Generative Adversarial Network. This principle is elucidated in Equations 6 and 7:

$$D_{Ra}(Real, Fake) = C(Real) - E(C(Fake)) \to 1 \qquad (6)$$

$$D_{Ra}(Fake, Real) = C(Fake) - E(C(Real)) \to 0 \qquad (7)$$

These two equations show how much more realistic the real image is compared to the fake image and how much less realistic the fake image is compared to the real image. In this case, $E(\cdot)$ is the average of real or fake data in a batch and is the output of the discriminator. This small change makes the model work better than the old discriminator network. Equations 8 and 9 show, respectively, the loss function for the discriminator network and the adversarial loss for the relativistic generative network:

$$L_{Dis}^{Ra} = -E_{I_{HR}}[\log(D_{Ra}(I_{HR}, I_{SR}))] - E_{SR}[\log(1 - D_{Ra}(I_{SR}, I_{HR}))] \qquad (8)$$

$$L_{Gen}^{Ra} = -E_{I_{HR}}[\log(D_{Ra}(I_{HR}, I_{SR}))] - E_{SR}[\log(D_{Ra}(I_{SR}, I_{HR}))] \qquad (9)$$

The relativistic generative adversarial network is a big step up from the standard GAN because it uses a relativistic loss function. This method lets the discriminator compare real and fake images instead of classifying them. In contrast to the conventional GAN, which ceases to learn from real data upon reaching an optimal state, RGAN persists in learning from both real

and synthetic data, as its gradients are influenced by both categories. Furthermore, RGAN produces images characterized by sharper edges, enhanced visual details, and more detailed frequency distribution. Utilizing relative comparison rather than absolute classification enhances the stability of the training process and mitigates problems such as mode collapse.

3. 3. Super Resolution Network   This study uses a pre-trained CNN-based network to re-identify persons. Using CNN as an appropriate deep learning algorithm enables the extraction of meaningful and high-level image features (40). The preprocessed high-resolution images are fed into a CNN utilizing the VGGNet architecture to extract unique feature vectors. These vectors are compared to each other using the Euclidean distance metric. If the distance between two vectors is minimal, the images are presumed to depict the same individual. Otherwise, each image is perceived as depicting a distinct individual. In the ranking stage, the feature vector of the probing image is compared with all gallery images, and individuals are scored according to their similarity. The nearest match is the one with the least distance. The quality of the extracted feature vectors and the improvements made to the preprocessing stages will determine how accurate this stage is. Using a pre-trained network not only saves time on the computer, but it also makes it easier to get to the immense amounts of training data that were used to train the model. Because of this, it works well in the real world (17).

4. EXPERIMENTAL RESULTS

The experiments and evaluations in this study were conducted on two datasets: CUHK01 (41) and CUHK03 (42). The CUHK01 dataset is one of the well-known and widely ued datasets in the field of Person Re-Identification. In this dataset, photos of people are taken from many angles and in different situations. Two different cameras took the pictures, which show 971 different people. Each person is assigned four images, with each image having a resolution of 60×160 pixels.

There are 14,097 photos of 1,467 people in the CUHK03 dataset. These were taken with six different cameras to record individuals from various angles and under different conditions. Figure 3 illustrates some sample frames of these datasets.

To assess the effectiveness of the proposed method, its performance has been compared with several well-known person re-identification methods. The two main ways for evaluating the accuracy of a person re-identification system are the rank-based measure and the Cumulative Matching Characteristic (CMC) curve (43).

The Rank-r measure tells us how likely it is that the correct image that matches the query is one of the top r

results that the algorithm gives you (44). It is very important to know the Rank-1 accuracy because it shows how well the system can find the right person as the first result. When matching is more flexible, higher rank values like Rank-5 and Rank-10 are used to rate performance.

We have comprehensively evaluated the re-identification accuracy for each dataset by showing results at four different rank levels: Rank-1, Rank-5, Rank-10, and Rank-20. Tables 1 and 2 show the results of the experiments. Although the proposed method achieved lower accuracy in the Rank-5 metric compared to some existing methods, it achieves higher accuracy in Rank-1, which holds greater importance in practical applications. For example, the Rank-1 accuracy on the CUHK01 dataset was 93.6%, and on the CUHK03 dataset, it was 95.1%.

The CMC curve indicates how well the system performs at different levels of rank. The horizontal axis of this chart shows the rank, and the vertical axis shows

Figure 3. Sample Images of re-id datasets: CUHK01 and CUHK03

TABLE 1. Numerical results and comparison on CUHK01

Methods	Rank1	Rank5	Rank10	Rank20
DCAE (17)	92.1%	-	97.5%	99.3%
PLCA (45)	46.8%	71.8%	80.5%	88.2%
EDPR (p =100) (46)	88.2%	98.20%	99.35%	-
DSA (47)	90.4%	97.8%	-	-
PersonNet (48)	71.14%	90.07%	95.00%	98.06%
DFSN (49)	83.95%	98.15%	98.97%	-
PN-GAN (50)	67.65%	86.64%	91.82%	-
PSD (51)	83.2%	-	97.1%	98.8%
EFS-Net (52)	70.1%	95.2%	99.1%	99.2%
Improved Quartet (53)	93.4%	98.8%	99.2%	-
The proposed model	93.6%	97.2%	99.4%	99.7%

TABLE 2. Numerical results and comparison on CUHK03

Methods	Rank1	Rank5	Rank10	Rank20
AACN (51)	91.39%	98.89%	99.48%	99.75%
PN-GAN (47)	79.76%	96.24%	98.56%	88.2%
PSD (48)	91.8%	-	99.1%	99.6%
Improved Quartet (53)	89.8%	99.4%	99.8%	-
DCAE(16)	94.4%	-	99.1%	99.65%
DFSN (46)	85.5%	98.74%	99.8%	-
SSM (54)	76.63%	94.59%	97.95%	-
BTL (28)	75.53%	95.15%	99.16%	-
Spindle Net (32)	88.5%	97.8%	98.6%	99.2%
CAN (33)	72.3%	93.8%	98.4%	99.2%
The proposed model	94.5%	96.3%	99.19%	99.75%

how well the system was able to find the right match up to that rank. In other words, this curve represents how likely it is that the correct identity is among the top r results that were calculated as follows:

$$CMC(r) = \frac{1}{N} \sum_{n=1}^{N} E(p_n \leq r) \qquad (10)$$

$$E(P_n \leq r) = \begin{cases} 1 & P_n \leq R \\ 0 & otherwise \end{cases} \qquad (11)$$

In this equation, P_n stands for the position of the nth person on the ranked list, and N stands for the total number of probe IDs in the dataset. The CMC curve is a well-known and often-used statistic in many studies on person re-identification (46, 55, 56). This curve shows how well a model works overall by showing how accurate it is at all rank positions, not just one. Figure 4 shows the CMC curves for the CUHK01 and CUHK03 datasets.

Our experimental results indicate that combining the histogram method with ESRGAN improves person re-

identification accuracy on these two datasets. Histogram equalization enhances brightness and contrast, revealing details hidden in dark areas and preparing the input for ESRGAN, allowing the network to concentrate on the real and clear information instead of on dark or low-light areas. Unlike many traditional super-resolution methods that add noise or edges that don't look natural, ESRGAN, which uses a strong discriminator and an improved perceptual loss function, preserves fine details and edges while generating more realistic and higher-quality images (26). The better quality of reconstruction helps get descriptive features out, which boosts the accuracy of person re-identification systems. Figure 5 shows examples of the initial low-resolution images and their corresponding high-resolution reconstructed versions.
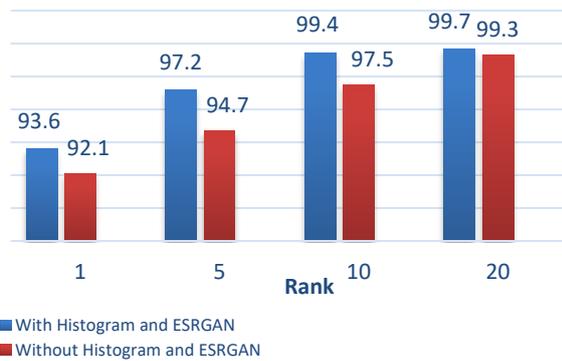

(a)

(b)
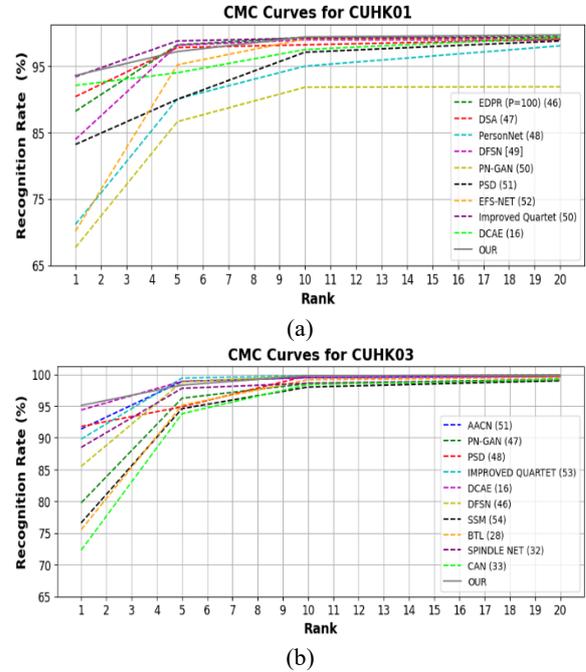
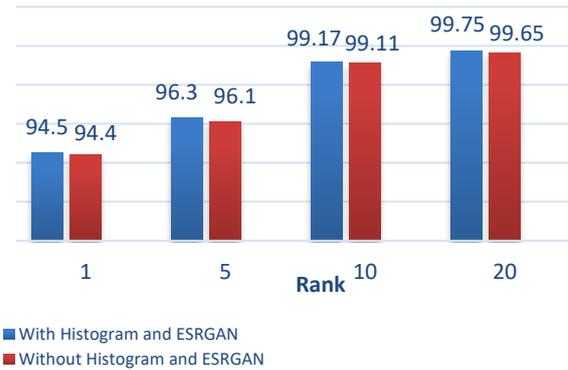**Figure 4.** CMC Curve for CUHK01 and CUHK03

Figure 5. Examples of images with (a) low resolution that have been converted to (b) high-resolution images

Additionally, Figure 6 shows the system's performance in two situations: without and with the use of histogram equalization and ESRGAN. These charts clearly demonstrate that applying the proposed method, particularly at lower ranks such as Rank-1 and Rank-5,

improves person re-identification accuracy. These results highlight the importance of enhancing image resolution, extracting more precise descriptive features, and improving system performance in real-world conditions.

Figure 6. Performance comparison under two conditions, with and without the application of histogram equalization and ESRGAN for (a) CUHK01 and (b) CUHK03 datasets

5. CONCLUSION

In this paper, a framework for person re-identification with the objective of solving the problem of person recognition with low resolution is proposed. Initially, histogram equalization techniques based on the Human Visual System (HVS) have been used to enhance image brightness and contrast. This method, by considering the human eye's sensitivity to intensity variations and details, enhances image perceptibility and improves the clarity of edges and important structures. After this step, the images were input into the ESRGAN network, an advanced GAN-based network that demonstrates remarkable performance in enhancing image resolutions.

And with its advanced architecture, it effectively reconstructs textures and precise structural components, especially edges. As a result, the re-identification network utilized in the final stage is able to successfully extract superior and more distinctive visual features. Comprehensive studies performed using prevalent datasets indicate that the proposed method significantly improves Rank-1 recognition accuracy and demonstrates superior performance compared to other existing approaches.

Ethics Approval and Consent to Participate

This article does not involve any studies with human participants or animals performed by any of the authors. Therefore, ethics approval and consent to participate are not applicable. Competing interests the author declares no financial or organizational conflicts of interest.

Data Availability

The data of this study are available from the corresponding author upon reasonable request.

Declaration of Generative AI and AI-assisted Technologies in the Writing Process

The authors declare that they did not use any AI-assistant technology in writing process.

Author Biosketch

Fatemeh Ilbeigi is currently a Ph.D. student at the University of Birjand, Birjand, Iran. Her research interests include digital image and audio signal processing, deep learning, and artificial intelligence. Her Email is: fatemeh.ilbeigi@birjand.ac.ir

Hassan Farsi is a Professor of Communication Engineering in the Department of Electrical and Computer Engineering at the university of Birjand, Birjand, Iran. He received his Ph.D. from the Centre for Communication Systems Research (CCSR), University of Surrey, Guildford, UK. His research interests include speech, image, and video processing in wireless communications. His Emil is: hfarsi@birjand.ac.ir

Sajad Mohamadzadeh received his Ph.D. in Telecommunication Engineering from the University of Birjand, Birjand, Iran. He is currently an Associate Professor in Faculty of Technical and Engineering of Ferdows, university of Birjand, Birjand, Iran. His research interests include image processing and retrieval, pattern recognition, digital signal processing, and sparse representation. His email is: s.mohamadzadeh@birjand.ac.ir

Amir Sezavar received his Ph.D. degree in Telecommunication Engineering from the University of Birjand, Birjand, Iran. He is currently an Assistant Professor at the Faculty of Technical and Engineering of Ferdows, University of Birjand, Birjand, Iran. His research interests include digital image processing and retrieval, visual signal processing, deep learning, and artificial intelligence. His Email is: a.sezavar@birjand.ac.ir

REFERENCES

1. Zhao K, Cheng D, Kou Q, Li J, Liu R. Sequences consistency feature learning for video-based person re-identification. Electronics Letters. 2022;58(4):142-44. https://doi.org/10.1049/ell2.12382

2. Fu H, Zhang K, Li H, Wang J. MRRNet: Learning multiple region representation for video person re-identification. Engineering Applications of Artificial Intelligence. 2022;114:105108. https://doi.org/10.1016/j.engappai.2022.105108

3. Shahbakhsh MB, Hassanpour H. Empowering Face Recognition Methods using a GAN-based Single Image Super-Resolution Network. International Journal of Engineering, Transactions A: Basics. 2022;35(10):1858-66. https://doi.org/10.5829/ije.2022.35.10a.05

4. Sun Y, Xu Q, Li Y, Zhang C, Li Y, Wang S, Sun J. Perceive where to focus: Learning visibility-aware part-level features for partial person re-identification. In: Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition. 2019:393-402. https://doi.org/10.1109/CVPR.2019.00048

5. Wang Z, Wang Z, Zheng Y, Wu Y, Zeng W, Satoh S. Beyond intra-modality: A survey of heterogeneous person re-identification. IJCAI International Joint Conference on Artificial Intelligence. 2020;2021:4973-80. https://doi.org/10.24963/ijcai.2020/692

6. Yang Q, Yu HX, Wu A, Zheng WS. Patch-based discriminative feature learning for unsupervised person re-identification. In: Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition. 2019:3628-37. https://doi.org/10.1109/CVPR.2019.00375

7. Wang H, Du H, Zhao Y, Yan J. A Comprehensive Overview of Person Re-Identification Approaches. IEEE Access. 2020;8:45556-83. https://doi.org/10.1109/ACCESS.2020.2978344

8. Munjal B, Amin S, Tombari F, Galasso F. Query-guided end-to-end person search. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. 2019:811-20. https://doi.org/10.48550/arXiv.1905.01203

9. Deng W, Zheng L, Ye Q, Kang G, Yang Y, Jiao J. Image-Image Domain Adaptation with Preserved Self-Similarity and Domain-Dissimilarity for Person Re-identification. In: Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition. 2018:994-1003. https://doi.org/10.1109/CVPR.2018.00110

10. Mao S, Zhang S, Yang M. Resolution-invariant person re-identification. arXiv preprint arXiv:1906.09748. 2019. https://doi.org/10.48550/arXiv.1906.09748

11. Chen YC, Huang PH, Yu LY, Huang JB, Yang MH, Lin YY. Deep Semantic Matching with Foreground Detection and Cycle-Consistency. In: Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics). 2019;11363:347-62. https://doi.org/10.1007/978-3-030-20893-6_22

12. Jing XY, Zhu X, Wu F, Hu R, You X, Wang Y, Feng H, Yang JY. Super-Resolution Person Re-Identification with Semi-Coupled Low-Rank Discriminant Dictionary Learning. IEEE Transactions on Image Processing. 2017;26:1363-78. https://doi.org/10.1109/TIP.2017.2651364

13. Li X, Zheng WS, Wang X, Xiang T, Gong S. Multi-scale learning for low-resolution person re-identification. In: Proceedings of the IEEE International Conference on Computer Vision. 2015:3765-73. https://doi.org/10.1109/ICCV.2015.429

14. Imani M. Deep Learning Based Electricity Demand Forecasting in Different Domains. Iranian Journal of Energy and Environment. 2020;11(1):33-39. https://doi.org/10.5829/ijee.2020.11.01.06

15. Khodadadi Z, Owlia MS, Amiri A. Enhancing Fault Detection in Image Analysis: A Combined Wavelet-Fourier Technique for Advancing Manufacturing Quality Control. International Journal of Engineering, Transactions A: Basics. 2024;37(2):387-401. https://doi.org/10.5829/ije.2024.37.02b.14

16. Mortezaie Z, Amiri Z, Mortezaee M. Adaptive Fractional-order Differentiation for Enhanced Image Contrast Utilizing Caputo Masks. International Journal of Engineering, Transactions B: Applications. 2026;39(5):1275-92. https://doi.org/10.5829/ije.2026.39.05b.19

17. Sezavar A, Farsi H, Mohamadzadeh S. A New Model for Person Reidentification Using Deep CNN and Autoencoders. Iranian Journal of Energy and Environment. 2023;14(4):314-20. https://doi.org/10.5829/ijee.2023.14.04.01

18. Al-Kaabi HAH, Aghagolzadeh A, Mikaeeli E. Super-resolution Augmentation of Large Images with Architecture Treatment of Deeply Recursive Convolutional Network. International Journal of Engineering, Transactions B: Applications. 2025;38(5):1042-55. https://doi.org/10.5829/ije.2025.38.05b.08

19. Zhao J, Xiong R, Xu J, Wu F, Huang T. Learning a deep convolutional network for subband image denoising. In: Proceedings - IEEE International Conference on Multimedia and Expo. 2019:1420-25. https://doi.org/10.1109/ICME.2019.00246

20. Wang Y, Wang L, Wang H, Li P, Lu H. Blind single image super-resolution with a mixture of deep networks. Pattern Recognition. 2020;102:107169. https://doi.org/10.1016/j.patcog.2019.107169

21. Jiang J, Yu Y, Wang Z, Tang S, Hu R, Ma J. Ensemble Super-Resolution with a Reference Dataset. IEEE Transactions on

Cybernetics.   2020;50(11):4694-708. https://doi.org/10.1109/TCYB.2018.2890149

22. Wang L, Huang Z, Gong Y, Pan C. Ensemble based deep networks for image super-resolution. Pattern Recognition. 2017;68:191-98. https://doi.org/10.1016/j.patcog.2017.02.027

23. Ledig C, Theis L, Huszár F, Caballero J, Cunningham A, Acosta A, Aitken A, Tejani A, Totz J, Wang Z, Shi W. Photo-realistic single image super-resolution using a generative adversarial network. In: Proceedings - 30th IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2017. 2017:105-14. https://doi.org/10.1109/CVPR.2017.19

24. Goodfellow I, Pouget-Abadie J, Mirza M, Xu B, Warde-Farley D, Ozair S, Courville A, Bengio Y. Advances in neural information processing systems. Curran Associates, Inc. 2014;36:2672-80.   https://doi.org/10.1016/s0925-2312(97)90018-7

25. Zhao L, Bai H, Liang J, Zeng B, Wang A, Zhao Y. Simultaneous color-depth super-resolution with conditional generative adversarial networks. Pattern Recognition. 2019;88:356-69. https://doi.org/10.1016/j.patcog.2018.11.028

26. Wang X, Yu K, Wu S, Gu J, Liu Y, Dong C, Qiao Y, Change Loy C. ESRGAN: Enhanced super-resolution generative adversarial networks. In: Proceedings of the European Conference on Computer Vision (ECCV) Workshops. 2018. https://doi.org/10.48550/arXiv.1809.00219

27. Shamsolmoali P, Zareapoor M, Wang R, Jain DK, Yang J. G-GANISR: Gradual generative adversarial network for image super resolution. Neurocomputing. 2019;366:140-53. https://doi.org/10.1016/j.neucom.2019.07.094

28. Dorrani Z, Farsi H, Mohamadzadeh S. Shadow Removal in Vehicle Detection Using ResUNet-a. Iranian Journal of Energy and Environment. 2023;14(1):87-95. https://doi.org/10.5829/ijee.2023.14.01.11

29. Sahragard E, Farsi H, Mohamadzadeh S. A Deep Learning-based Approach for Accurate Semantic Segmentation with Attention Modules. Iranica Journal of Energy and Environment. 2025;16(4):692-705. https://doi.org/10.5829/ijee.2025.16.04.12

30. Sezavar A, Farsi H, Mohamadzadeh S, Radeva P. A new person re-identification method by defining CNN-based feature extractor and sparse representation. Multimedia Tools and Applications.   2024;83(4):11043-59. https://doi.org/10.1007/s11042-023-15718-x

31. Chen W, Chen X, Zhang J, Huang K. Beyond triplet loss: A deep quadruplet network for person re-identification. In: Proceedings - 30th IEEE Conference on Computer Vision and Pattern Recognition,   CVPR   2017.   2017:1320-29. https://doi.org/10.1109/CVPR.2017.145

32. Zhao H, Tian M, Sun S, Shao J, Yan J, Yi S, Wang X, Tang X. Spindle net: Person re-identification with human body region guided feature decomposition and fusion. In: Proceedings - 30th IEEE Conference on Computer Vision and Pattern Recognition, CVPR   2017.   2017:907-15. https://doi.org/10.1109/CVPR.2017.103

33. Liu H, Feng J, Qi M, Jiang J, Yan S. End-to-end comparative attention networks for person re-identification. IEEE Transactions on Image Processing. 2017;26(7):3492-3506. https://doi.org/10.1109/TIP.2017.2700762

34. Jiao J, Zheng WS, Wu A, Zhu X, Gong S. Deep Low-Resolution Person Re-Identification. Proceedings of the AAAI Conference on   Artificial   Intelligence.   2018;32(1):6967-74. https://doi.org/10.1609/aaai.v32i1.12284

35. Wang Z, Ye M, Yang F, Bai X, Satoh S. Cascaded SR-GAN for scale-adaptive low resolution person re-identification. In: IJCAI International Joint Conference on Artificial Intelligence. 2018:3891-97. https://doi.org/10.24963/ijcai.2018/541

36. Cheng Z, Dong Q, Gong S, Zhu X. Inter-Task Association Critic for Cross-Resolution Person Re-Identification. In: 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition   (CVPR).   2020:2602-12. https://doi.org/10.1109/CVPR42600.2020.00268

37. Chen YC, Li YJ, Du X, Wang YCF. Learning resolution-invariant deep representations for person re-identification. In: 33rd AAAI Conference on Artificial Intelligence, AAAI 2019. 2019;33:8215-22. https://doi.org/10.1609/aaai.v33i01.33018215

38. Li YJ, Chen YC, Lin YY, Du X, Wang YCF. Recover and identify: A generative dual model for cross-resolution person re-identification. In: Proceedings of the IEEE International Conference   on   Computer   Vision.   2019:8089-98. https://doi.org/10.1109/ICCV.2019.00818

39. Adil M, Mamoon S, Zakir A, Manzoor MA, Lian Z. Multi scale-adaptive super-resolution person re-identification using GAN. IEEE   Access.   2020;8:177351-62. https://doi.org/10.1109/ACCESS.2020.3023594

40. Sezavar A, Farsi H, Mohamadzadeh S. A modified grasshopper optimization algorithm combined with CNN for content based image retrieval. International Journal of Engineering, Transactions   A: Basics.   2019;32(7):924-30. https://doi.org/10.5829/ije.2019.32.07a.04

41. Li W, Zhao R, Wang X. Human Reidentification with Transferred Metric Learning. In: Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics). 2013;7724:31-44.   https://doi.org/10.1007/978-3-642-37331-2_3

42. Li W, Zhao R, Xiao T, Wang X. DeepReID: Deep Filter Pairing Neural Network for Person Re-identification. In: 2014 IEEE Conference on Computer Vision and Pattern Recognition. 2014:152-59. https://doi.org/10.1109/CVPR.2014.27

43. Dong H, Lu P, Zhong S, Liu C, Ji Y, Gong S. Person re-identification by enhanced local maximal occurrence representation and generalized similarity metric learning. Neurocomputing.   2018;307:25-37. https://doi.org/10.1016/j.neucom.2018.04.013

44. Hassanpour H, Mortezaie Z, Beghdadi A. Sensing Image Regions for Enhancing Accuracy in People Re-identification. Iranian Journal of Energy and Environment. 2022;13(3):295-304. https://doi.org/10.5829/ijee.2022.13.03.09

45. Matsukawa T, Suzuki E. Person re-identification using CNN features learned from combination of attributes. In: 2016 23rd International Conference on Pattern Recognition (ICPR). 2016:2428-33. https://doi.org/10.1109/ICPR.2016.7900000

46. Guo Y, Cheung NM. Efficient and Deep Person Re-identification Using Multi-level Similarity. In: Proceedings of the IEEE Computer Society Conference on Computer Vision and   Pattern   Recognition.   2018:2335-44. https://doi.org/10.1109/CVPR.2018.00248

47. Zhang Z, Lan C, Zeng W, Chen Z. Densely semantically aligned person re-identification. In: Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition.   2019:667-76. https://doi.org/10.1109/CVPR.2019.00076

48. Wu L, Shen C, van den Hengel A. Personnet: Person re-identification with deep convolutional neural networks. arXiv preprint   arXiv:1601.07255.   2016. https://doi.org/10.48550/arXiv.1601.07255

49. Khatun A, Denman S, Sridharan S, Fookes C. A Deep Four-Stream Siamese Convolutional Neural Network with Joint Verification and Identification Loss for Person Re-Detection. In: Proceedings - 2018 IEEE Winter Conference on Applications of

Computer Vision, WACV 2018. 2018:1292-301. https://doi.org/10.1109/WACV.2018.00146

50. Qian X, Fu Y, Xiang T, Wang W, Qiu J, Wu Y, Jiang YG, Xue X. Pose-normalized image generation for person re-identification. In: Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics). 2019;11213:661-78. https://doi.org/10.1007/978-3-030-01240-3_40

51. Liu Y, Zhang Y, Coleman S, Bhanu B, Liu S. A new patch selection method based on parsing and saliency detection for person re-identification. Neurocomputing. 2020;374:86-99. https://doi.org/10.1016/j.neucom.2019.09.073

52. Sheshkal SA, Fouladi-Ghaleh K, Aghababa H. An Improved Person Re-identification Method by light-weight convolutional neural network. In: 2020 10th International Conference on Computer and Knowledge Engineering, ICCKE 2020. 2020:463-68. https://doi.org/10.1109/ICCKE50421.2020.9303722

53. Khatun A, Denman S, Sridharan S, Fookes C. Joint identification–verification for person re-identification: A four stream deep learning approach with improved quartet loss function. Computer Vision and Image Understanding. 2020;197-8:102989. https://doi.org/10.1016/j.cviu.2020.102989

54. Bai S, Bai X, Tian Q. Scalable person re-identification on supervised smoothed manifold. In: Proceedings - 30th IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2017. 2017:3356-65. https://doi.org/10.1109/CVPR.2017.358

55. Al-Kaabi, H. A. H., Aghagolzadeh, A., Mikaeeli, E. Super-resolution Augmentation of Large Images with Architecture Treatment of Deeply Recursive Convolutional Network. International Journal of Engineering, Transactions B: Applications. 2025;38(5):1042-55. https://doi.org/10.5829/ije.2025.38.05b.08

56. Dey, A., Biswas, S., Abualigah, L. Efficient Violence Recognition in Video Streams using ResDLCNN-GRU Attention Network. ECTI Transactions on Computer and Information Technology. 2024;18(3):329-41. https://doi.org/10.37936/ecti-cit.2024183.255679

Persian Abstract

چکیده

بازشناسی افراد یکی از مسائل مهم در زمینه‌پردازش تصویر و بینایی ماشین است که در سال‌های اخیر توجه زیادی را به خود جلب کرده است. یکی از چالش‌های اصلی در این حوزه، کیفیت پایین تصاویر ثبت‌شده توسط دوربین‌های مختلف است که می‌تواند منجر به از دست‌رفتن جزئیات مهم و در نتیجه کاهش دقت الگوریتم‌های بازشناسی می‌گردد. در این مقاله، چارچوبی برای بهبود دقت بازشناسی افراد پیشنهاد شده است که با بهبود روشنایی تصاویر کم‌نور از طریق تکنیک‌های همسان‌سازی هیستوگرام و افزایش وضوح آن‌ها با استفاده از شبکه‌ی ESRGAN ، دقت و توانایی شبکه را در بازشناسی افراد ارتقا می‌دهد. نتایج تجربی نشان می‌دهد که روش پیشنهادی با بهبود روشنایی و وضوح تصاویر که نقشی کلیدی در آشکارسازی جزئیات مهم دارند، کیفیت ورودی شبکه‌ی بازشناسی را افزایش داده و در مقایسه با روش‌های موجود، دقت بازشناسی افراد را ارتقا می‌دهد، به‌طوری‌که به‌دقت Rank-1 معادل ٪۹۳.٦ بر روی پایگاه‌داده CUHK01 و ٪۹۴.۵ بر روی پایگاه‌داده CUHK03 دست‌یافته است.