

# NightReID: A Large-Scale Nighttime Person Re-Identification Benchmark

Yuxuan Zhao<sup>1\*</sup>, Weijian Ruan<sup>2,3\*</sup>, He Li<sup>1†</sup>, Mang Ye<sup>1†</sup>

<sup>1</sup>School of Computer Science, Wuhan University

<sup>2</sup>Hangzhou Research Institute, Xidian University

<sup>3</sup>Smart City Research Institute of China Electronics Technology Group Corporation  
zhaoyuxuan@whu.edu.cn, rweij66@163.com, lihe404@whu.edu.cn, yemang@whu.edu.cn

## Abstract

Person re-identification (Re-ID) is crucial for intelligent surveillance systems, facilitating the identification of individuals across multiple camera views. While significant advancements have been made for daytime scenarios, ensuring reliable Re-ID performance during nighttime remains a significant challenge. Given the cost and limited accessibility of infrared cameras, we investigate a critical question: *Can RGB cameras be effectively utilized for accurate Re-ID during nighttime?* To address this, we introduce *NightReID*, a large-scale RGB Re-ID dataset collected from a real-world nighttime surveillance system. NightReID includes 1,500 identities and over 53,000 images, capturing diverse scenes with complex lighting and adverse weather conditions. This rich dataset provides a valuable benchmark for advancing nighttime Re-ID research. Moreover, we propose the *Enhancement, Denoising, and Alignment (EDA)* framework with two novel modules to enhance nighttime Re-ID performance. First, an unsupervised *Image Enhancement and Denoising (IED)* method is designed to improve the quality of nighttime images, preserving critical details while removing noise without requiring paired ground truth. Second, we introduce *Data Distribution Alignment (DDA)* through statistical priors, aligning the distributions between pre-training data and nighttime data to mitigate domain shift. Extensive experiments on multiple nighttime Re-ID datasets demonstrate the significance of NightReID and validate the efficacy, flexibility, and applicability of the EDA framework.

**Code** — <https://github.com/msm8976/NightReID>

## 1 Introduction

Person re-identification (Re-ID) aims to identify a specific person from a gallery of images captured by non-overlapping cameras. This task is crucial for applications such as intelligent surveillance systems and smart cities. While significant progress has been achieved with datasets like Market-1501 (Zheng et al. 2015) and MSMT17 (Wei et al. 2018), alongside methods utilizing CNNs (Sun et al. 2018; Wang et al. 2018; Chen et al. 2019; Ye et al. 2021, 2024b) and Transformers (He et al. 2021; Luo et al. 2021; Li

\*These authors contributed equally.

†Corresponding Author.

Copyright © 2025, Association for the Advancement of Artificial Intelligence (www.aaai.org). All rights reserved.

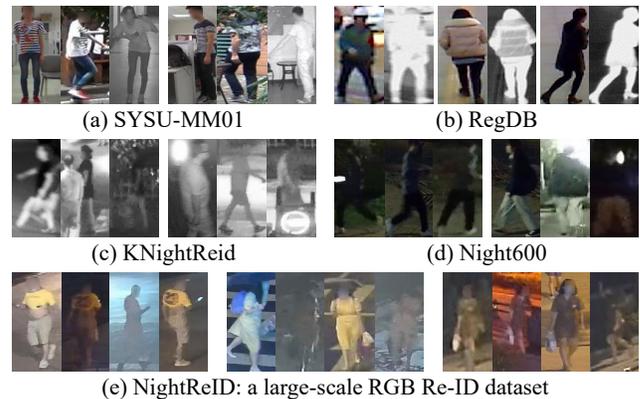


Figure 1: Comparison of existing nighttime Re-ID datasets.

et al. 2022; Ye et al. 2024a; Liu, Ye, and Du 2024), these approaches are predominantly optimized for daytime scenarios. To achieve true robustness, Re-ID systems must perform reliably under all conditions, including nighttime, when the incidence of criminal activity is significantly higher. While infrared cameras can offer improved nighttime visibility, they fall short in capturing fine-grained details such as color and texture, which are crucial for accurate Re-ID. Furthermore, the introduction of infrared fill lights not only increases costs but also has limited effectiveness for distant individuals in outdoor environments, limiting their widespread deployment and applicability across diverse settings. In contrast, RGB (visible) cameras are far more prevalent due to their affordability and accessibility. These considerations lead us to a critical question: *Can RGB cameras be effectively utilized for accurate Re-ID during nighttime?*

Figure 1 illustrates the current landscape of nighttime Re-ID datasets. SYSU-MM01 (Wu et al. 2017) and RegDB (Nguyen et al. 2017) are designed for cross-modal Re-ID, while KNightReid (Zhang, Yuan, and Wang 2019) consists of images from infrared cameras at night. These datasets often feature uniform backgrounds and lack dedicated nighttime RGB images, limiting their utility for RGB-based Re-ID tasks. Although the Night600 dataset (Lu et al. 2023b) is a step forward, being the first to focus on nighttime RGB Re-ID, most images are extremely dark, posing challenges for both human observers and Re-ID methods to distinguish ac-

curately. Additionally, the dark images in Night600 struggle to capture the diversity of real-world scenarios, thereby limiting its practical applicability. Thus, *How to better capture high-quality nighttime data using RGB cameras to reflect realistic and diverse real-world scenarios?*

Beyond data scarcity, nighttime RGB Re-ID presents unique challenges. Researchers have explored integrating low-light image enhancement (LLIE) methods with nighttime Re-ID tasks (Zhang, Yuan, and Wang 2019; Lu et al. 2023b; Du, Du, and Yu 2023). However, Re-ID data typically consist of low-resolution images of individuals affected by various degradation factors such as lighting, weather, blur, and noise, while LLIE tasks usually require high-quality paired data with different exposure durations. Moreover, LLIE methods focus on producing visually appealing images, which may inadvertently introduce noise and compromise the extraction of discriminative features essential for Re-ID tasks. This creates a notable domain gap between LLIE and Re-ID. Al et al. (2022) found that performance in downstream tasks does not always correlate with enhanced visual quality. Thus, *How to design more suitable image enhancement methods specifically for Re-ID tasks?*

Typically, Re-ID methods begin by initializing pre-trained parameters from ImageNet (Deng et al. 2009), followed by image normalization after data augmentation, and feature extraction using a backbone network (Luo et al. 2019). Recent studies suggest pre-training on large-scale human-centered datasets (Fu et al. 2021) to minimize the domain gap between pre-training and target data, thereby enhancing Re-ID performance (Luo et al. 2021; Chen et al. 2023). This methodology assumes that the distribution of target data is similar to the pre-training data, facilitating the effective utilization of learned information. However, nighttime conditions present unique challenges: images captured in low-light environments tend to be darker, with lower and more concentrated pixel values compared to general conditions. This discrepancy in data distribution significantly impacts performance. Thus, *How to better leverage the capabilities of large-scale pre-trained models on nighttime data?*

Recognizing the limitations of the existing datasets, we introduce NightReID, a large-scale real-world nighttime RGB dataset. NightReID includes 1,500 annotated identities and 1,096 unlabeled identities, totaling over 53,000 images captured from six outdoor nighttime RGB cameras. Beyond its scale, NightReID offers diverse gallery sets with varying amounts of testing data and distractors, facilitating comprehensive evaluations of Re-ID methods. NightReID stands out with its complex illumination conditions, including poor light, backlight, and overexposure, as well as diverse scenes, various occlusions, and adverse weather conditions such as rain and fog. Furthermore, we demonstrate that the camera’s exposure gain enhances the capture of higher-quality RGB data at night, making it feasible to use nighttime RGB data for accurate Re-ID. To our best knowledge, NightReID is currently the largest and most realistic nighttime Re-ID dataset, setting a new benchmark for research in this field.

Moreover, we introduce two novel methods specifically designed for nighttime Re-ID, aimed at bridging the domain gap between Re-ID and LLIE tasks, as well as address-

ing the unique distribution characteristics of nighttime data: **1) A No-reference Image Enhancement and Denoising (IED) Method.** Inspired by the no-reference LLIE method Zero-DCE (Guo et al. 2020), our approach dynamically adjusts the enhancement level based on the image’s brightness to avoid over-enhancement. Additionally, it reduces noise while preserving and enhancing details based on noise distribution priors in a no-reference manner. **2) Data Distribution Alignment (DDA) through Statistical Priors.** Nighttime images often exhibit lower and more concentrated pixel values, leading to distribution bias compared to general images. By applying z-score normalization during data loading, we align the data distribution of nighttime datasets with the pre-training data, utilizing prior statistical information to increase overall information entropy and ensure consistency. Finally, we propose a simple multi-branch **Enhancement, Denoising, and Alignment (EDA) framework** to leverage the advantages of both IED and DDA. Without requiring additional training data, our approach significantly improves Re-ID performance on nighttime datasets.

The main contributions are summarized as follows:

- We contribute a large-scale real-world nighttime RGB Re-ID dataset, which features a richer data scale, varied scenes, complex illumination conditions, diverse weather, and higher image quality compared to existing datasets in this domain.
- We propose a no-reference image enhancement and denoising (IED) method specifically designed for Re-ID tasks. Our method better controls the enhancement level and reduces noise while preserving crucial details through a novel no-reference loss function.
- We analyze existing Re-ID methods and propose data distribution alignment (DDA) in nighttime datasets to reduce the distribution gap while preserving information.
- Abundant experiments demonstrate the NightReID benchmark’s significance and validate EDA’s efficacy, flexibility, and applicability in nighttime Re-ID.

## 2 Related Work

### Person Re-Identification in Nighttime

Wu et al. (2017) treat nighttime Re-ID as a cross-modal retrieval task, referred to as Visible-Infrared Re-ID (VI Re-ID), which involves matching and retrieving between daytime RGB and nighttime infrared images. Despite its potential (Zhang and Wang 2023; Ye et al. 2023; Shi et al. 2023; Qiu et al. 2024; Shi et al. 2025, 2024), VI Re-ID encounters challenges in bridging modality gaps and cannot utilize images captured by commonly used RGB cameras at night.

To address illumination variations across different times, Huang et al. (2019) propose an illumination-invariant Re-ID framework combining retinex decomposition with Re-ID. Zeng et al. (2020) design an illumination-identity disentanglement network to preserve identity information while mitigating lighting variations. Zhang et al. (2022) reduce illumination discrepancies by estimating and adjusting the illumination levels of testing images. Large-scale synthetic Re-ID datasets like WePerson (Li, Ye, and Du 2021) and FineGPR

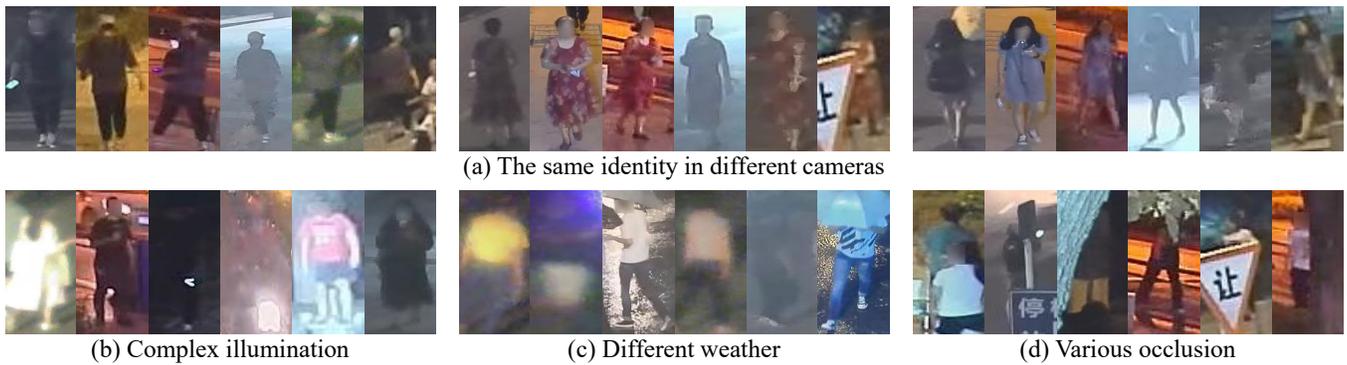


Figure 2: Sample images of the NightReID dataset.

(Xiang et al. 2023) provide various lighting conditions to train more generalized models. However, these methods depend on synthetic data, making it difficult to accurately simulate real-world illumination changes and handle the various degradation factors present in nighttime images effectively.

In real nighttime scenes, Zhang et al. (2019) enhance Re-ID by jointly training a denoising and Re-ID network on an infrared Re-ID dataset. Lu et al. (2023b) introduce a nighttime RGB Re-ID dataset and apply an illumination distillation framework to fuse original and enhanced image features. While these methods contribute to the advancement of nighttime Re-ID research, they may not fully retain and leverage image information during data collection and training, which limits their practical deployment.

### Low Light Image Enhancement

Low-light image enhancement (LLIE) aims to improve the quality of images captured in poorly illuminated environments (Li et al. 2021). Recent deep learning methods have made significant progress in LLIE (Zhang, Zhang, and Guo 2019; Liu et al. 2021b; Cui et al. 2022; Cai et al. 2023; Yan et al. 2024). However, they require paired ground truth for training, which is hard to obtain in practice and may exhibit domain differences with the images needing enhancement.

To overcome this challenge, Jiang et al. (2021) propose EnlightenGAN, an unsupervised method using GANs to map different illumination spaces, constrained by perceptual loss. Guo et al. (2020; 2021) employ image-specific curve estimation for pixel-level adjustment and learning with no-reference loss functions. Zhang et al. (2021b) introduce a maximum entropy-based retinex model for unsupervised re-enhancement and denoising using noise distribution priors.

While LLIE methods excel in visual enhancement, there is a significant gap between their domains and objectives with downstream computer vision tasks. Consequently, they do not always improve the performance of these tasks consistently. Some studies show that LLIE methods enhance face detection effectively (Li et al. 2021; Liu et al. 2021a), but noise and inconsistent enhancement intensities may degrade performance in classification tasks (Al Sobhahi and Tekli 2022). Lu et al. (2023a) address this by using a shared shallow encoder for Re-ID and LLIE, leveraging both real and synthetic datasets for multi-domain learning.

Dataset	#IDs	#Cams	#Ni_Imgs	Distractors	Modality
RegDB	412	1	4,120	No	RGB-IR
SYSU-MM01	491	6	15,792	No	RGB-IR
KnightReid	937	3	315,354	No	IR
Night600	600	8	28,813	No	RGB
<b>NightReID</b>	<b>1,500</b>	<b>6</b>	<b>53,239</b>	<b>Yes</b>	<b>RGB</b>

Table 1: Statistics of nighttime Re-ID datasets. #Ni\_Imgs indicates the number of images captured at night.

## 3 The NightReID Dataset

### Dataset Description

In this paper, we introduce NightReID, a large-scale RGB nighttime Re-ID dataset. NightReID is constructed using six non-overlapping RGB cameras publicly deployed across a campus, with data collection spanning a total of five nights from 7:00 PM to 3:00 AM. The dataset encompasses both high-quality and blurred images, featuring various illumination conditions and diverse weather scenarios. These elements contribute to a rich dataset that accurately represents realistic nighttime surveillance conditions.

To create NightReID, we utilized Faster-RCNN (Ren et al. 2016) and ECO tracker (Danelljan et al. 2017) for initial bounding box extraction. We then manually refined the bounding boxes and excluded images that were indistinguishable to human observers to ensure the dataset’s quality. Additionally, to address privacy concerns, all facial information in the images was blurred. NightReID ultimately comprises 53,239 bounding boxes corresponding to 1,500 annotated identities and 1,096 unlabeled identities. Figure 2 illustrates the diversity of the NightReID dataset, and the statistical comparison of nighttime datasets is shown in Table 1. A more detailed discussion of the NightReID dataset is provided in the technical appendix.

### Dataset Features

NightReID encompasses five unique characteristics:

**Large Size with Distractors:** NightReID surpasses existing nighttime datasets in terms of both the number of identities and images. Additionally, NightReID introduces a distractor set comprising 11,811 images of 1,096 unlabeled identities,

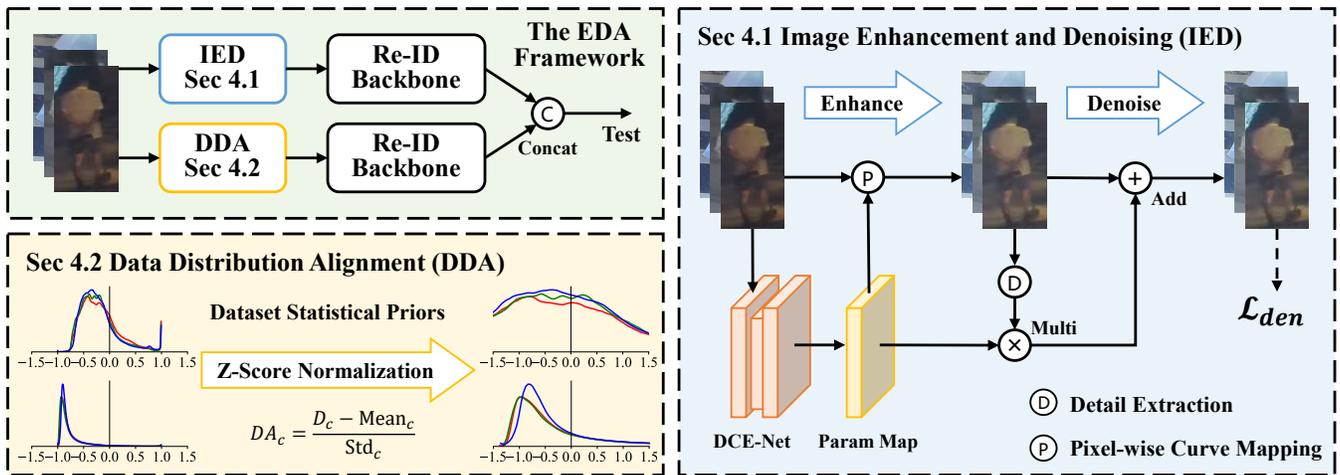


Figure 3: Overview of the Enhancement, Denoising, and Alignment (EDA) framework. The Image Enhancement and Denoising (IED) module enhances nighttime images and removes noise without reference for Re-ID tasks. The Data Distribution Alignment (DDA) module aligns nighttime data more closely with the distribution of pre-training data for richer information.

which consist of bounding boxes without associated identities. This distractor set enables a comprehensive evaluation of the effectiveness and robustness of Re-ID methods.

**Complex Illuminations:** NightReID presents diverse illumination conditions across different cameras and even within subsets from the same camera, including poor light, backlight, and overexposure. These conditions introduce degradation factors such as noise and blur.

**Different Weather:** NightReID captures images under various weather conditions such as rain and fog. Rain can blur the camera lens and cause individuals to use umbrellas, leading to occlusions and degraded image quality. Foggy conditions result in blurry images, complicating the capture of detailed information.

**Various Occlusions:** Suspects may intentionally conceal themselves behind or alongside others or objects, especially at night. NightReID includes numerous images depicting occluded persons, mirroring real-world scenarios encountered in Re-ID tasks.

**Enhanced Image Quality:** During data collection, we utilized ISP-level exposure gain in the cameras to achieve better visibility in nighttime scenes. This technique effectively minimizes information loss due to RGB image quantization, thereby enhancing image quality and improving the performance of downstream tasks without additional costs. Moreover, most RGB cameras in real-world applications can support this capability, offering crucial color details for Re-ID that are often lacking in infrared cameras.

## Evaluation Protocol

For dataset partitioning, we randomly divide the images of annotated 1,500 identities into training and testing sets. Specifically, the training set comprises 500 annotated identities with 15,514 images, while the complete testing set consists of 1,000 annotated identities with 25,914 images. Among the testing set, 528 identities are randomly selected

as a sub-set to simulate different retrieval scenarios. Additionally, 11,811 images from 1,096 unlabeled identities serve as the distractor set. Therefore, the combination of different testing sets and the inclusion or exclusion of distractors results in four evaluation protocols, facilitating a comprehensive analysis of Re-ID methods across varying data scales. The query set consists of 10% randomly selected annotated identity images from the corresponding testing set. Considering the diversity and abundance of irrelevant individuals typically present in real-world surveillance systems, the evaluation protocol that combines the 528 identities with the distractors is adopted as the default.

Aligned with established practices for previous datasets, we utilize the Cumulated Matching Characteristics (CMC) curve and mean Average Precision (mAP) to evaluate Re-ID performance, and gallery images from the same camera and identities as the query will be excluded during testing.

## 4 Methodology

In this section, we introduce two modules specifically designed for nighttime Re-ID tasks and their integration with existing methods, forming the Enhancement, Denoising, and Alignment (EDA) framework. The overall architecture is illustrated in Figure 3.

### Image Enhancement and Denoising

Given the complexity of lighting conditions at night, images often exhibit lower brightness and various degradation factors, making it challenging to extract discriminative features compared to daytime images. While numerous LLIE methods have been developed, most rely on paired ground truth to train, rendering them unsuitable for integration with Re-ID tasks. Inspired by a lightweight unsupervised LLIE method Zero-DCE (Guo et al. 2020), we design an Image Enhancement and Denoising (IED) module to further optimize noise control and enhancement to better align with Re-ID tasks.

Zero-DCE uses a convolutional network DCE-Net to estimate the pixel-wise curve parameters for input images and adjusts each pixel by applying iterative quadratic curves. It also introduces four no-reference losses  $\mathcal{L}_{spa}$ ,  $\mathcal{L}_{exp}$ ,  $\mathcal{L}_{col}$ , and  $\mathcal{L}_{tv_A}$  to collectively optimize the network and assess the quality of enhanced images. However, due to the lack of high-quality ground truth data and the constraints of  $\mathcal{L}_{tv_A}$  on the brightness gradient relationship, it tends to amplify noise during enhancement. This enhanced noise can adversely affect downstream tasks (Al Sobhahi and Tekli 2022). To mitigate this issue, we incorporate detail extraction after enhancement, enabling adaptive pixel-wise blurring or sharpening through a new parameter map from DCE-Net, aimed at removing noise while preserving details:

$$IE(I, x) = I(x) + \mathcal{G}(x)(I(x) - GF(I(x))), \quad (1)$$

where  $I$  is the enhanced image,  $x$  denotes pixel coordinates,  $\mathcal{G}$  is a parameter map, and  $GF$  represents the Gaussian filter. The final output of the IED module is denoted as  $IE$ .

Therefore, determining noise levels without reference becomes a new challenge. Since noise in nighttime images typically follows a Poisson distribution and is mutually independent, the gradient of details in the smoothed image generally exceeds that of noise (Zhang et al. 2021b). Based on this, we use the local normalized gradient of the smoothed image as weights, leveraging the gradient of the enhanced image to evaluate detail intensity at each coordinate within the neighborhood. Regions with low intensity are considered noise and should be removed, while those with higher intensity are regarded as details that need to be preserved or enhanced. By employing a function like  $x * \exp(-\lambda x)$ , which exhibits an initial rise followed by a decline, we can approximately separate noise from details. Consequently, we propose a novel no-reference denoise loss  $\mathcal{L}_{den}$  as follows:

$$\mathcal{L}_{den} = (WN(|\nabla I_{gray}|)) \circ \exp(-\lambda WN(|\nabla I_{gray}|)), \quad (2)$$

where  $I_{gray}$  is the grayscale image of the output,  $\nabla$  is the gradient operations in both horizontal and vertical directions,  $N(x)$  denotes the normalization on  $x$ ,  $\lambda$  is set to 10 (Zhang et al. 2021b),  $W$  is the weight map, defined as:

$$W = N_{local}(|\nabla MF(I_{gray})|), \quad (3)$$

where  $MF$  refers to the mean filter,  $N_{local}(x)$  is the local normalization on  $x$ . Note that the weight  $W$  does not participate in gradient propagation during training.

Furthermore, since the exposure control loss  $\mathcal{L}_{exp}$  uniformly adjusts the brightness to a well-exposed level, some extremely low-light images may be over-enhanced and compromise features. Therefore, we propose an adaptive approach to control the enhancement level of each image based on its brightness and the target exposure level:

$$\mathcal{L}'_{exp} = \frac{1}{M} \sum_{k=1}^M |Y_k - \frac{\bar{I} + 0.5}{2}|, \quad (4)$$

where  $M$  is the number of local regions of size  $16 \times 16$ ,  $Y$  is the average intensity of a local region in the enhanced image, and  $\bar{I}$  is the mean intensity value of the input image.

The overall loss function of the IED module is:

$$\mathcal{L}_{IED} = \lambda_1 \mathcal{L}_{spa} + \lambda_2 \mathcal{L}'_{exp} + \lambda_3 \mathcal{L}_{col} + \lambda_4 \mathcal{L}_{tv_A} + \lambda_5 \mathcal{L}_{den}, \quad (5)$$

where  $\lambda_1, \lambda_2, \lambda_3, \lambda_4, \lambda_5$  denote the weights assigned to each term, which are set to 1, 5, 5, 200, 10 as referenced in (Guo et al. 2020), respectively.

## Data Distribution Alignment

Due to the relatively limited images in Re-ID datasets, existing methods often initialize the backbone network with pre-trained weights from large-scale classification datasets (Luo et al. 2019), which provide the network with essential feature extraction capabilities. While this approach is generally effective, it encounters significant challenges when applied to nighttime datasets, particularly those with extremely low brightness levels. This is because nighttime images are typically captured under unbalanced lighting conditions, resulting in pixel values heavily concentrated in lower range, leading to substantial discrepancies with the data used for pre-training (Neumann et al. 2019). Consequently, the effectiveness of leveraging pre-trained weights for feature extraction is diminished. Therefore, aligning the data distribution of the nighttime dataset with that of the pre-training data can be expected to enhance performance.

CLAHE (Zuiderveld 1994) redistributes pixel values to balance image histogram, while Ye et al. (2019) enhance unsupervised embedding learning via invariant and spreading instance feature. However, rather than treating each image or instance independently, our approach focuses on aligning the overall data distribution of the dataset with pre-training data, while preserving intra-dataset variations and relationships. Given that mean and standard deviation effectively capture dataset characteristics, and large-scale datasets are typically normalized to a standard distribution during pre-training to maximize information content (measured by entropy), we propose a method called Data Distribution Alignment (DDA). DDA leverages statistical priors from the dataset to perform z-score normalization, transforming the dataset into a standard distribution. This process not only enhances the information content of the dataset but also brings its data distribution closer to the pre-training data.

Specifically, for a given set of training images, we first compute the mean and standard deviation of each channel for each image. Then, we calculate the overall mean and standard deviation for each channel by averaging these values across all images:

$$\text{Mean}_c = \frac{1}{N} \sum_{i=1}^N \frac{1}{P_i} \sum_{j=1}^{P_i} I_{ijc}, c \in \{R, G, B\}, \quad (6)$$

$$\text{Std}_c = \frac{1}{N} \sum_{i=1}^N \sqrt{\frac{1}{P_i} \sum_{j=1}^{P_i} (I_{ijc} - \text{Mean}_c)^2}, c \in \{R, G, B\}, \quad (7)$$

where  $N$  is the number of images in the training set,  $P_i$  denotes the number of pixels in each image  $I_i$ ,  $c$  refers to the color channels ( $R, G, B$ ). Mean and Std correspond to the mean and standard deviation of the training set, respectively.

During dataset loading, we first augment the images through resizing, padding, random cropping, and random horizontal flipping. Then, each channel of the augmented

Method	Avenue	528IDs w/ D		528IDs w/o D		1000IDs w/ D		1000IDs w/o D		Night600	
		Rank-1	mAP	Rank-1	mAP	Rank-1	mAP	Rank-1	mAP	Rank-1	mAP
<b>BoT</b>	CVPR19	36.3	25.9	41.0	30.4	27.2	19.7	30.2	22.1	11.6	5.5
<b>AGW</b>	TPAMI21	43.7	32.0	49.0	37.0	34.8	25.1	37.9	27.8	12.2	6.4
<b>MSINet</b>	CVPR23	49.8	35.9	54.4	41.1	39.7	27.8	43.1	34.8	7.8	2.9
<b>HAT</b>	MM21	48.1	35.4	54.0	40.7	39.2	28.7	42.8	31.7	16.7	7.9
<b>TransReID</b>	ICCV21	55.2	41.0	59.9	46.1	46.0	33.2	49.1	36.2	9.7	6.5
<b>TransReID-SSL</b>	arXiv21	<u>65.0</u>	<u>50.8</u>	<u>70.4</u>	<u>56.4</u>	<u>58.3</u>	<u>43.2</u>	<u>61.8</u>	<u>46.7</u>	<u>20.0</u>	<u>9.5</u>
<b>PFD-Net</b>	AAAI22	53.6	39.4	58.7	44.5	44.8	31.9	48.0	34.9	3.9	4.2
<b>DC-Former</b>	AAAI23	38.5	26.6	43.0	31.0	29.4	20.5	31.7	23.0	8.5	5.6
<b>CLIP-ReID</b>	AAAI23	56.1	43.0	61.3	48.7	49.3	35.9	52.5	39.4	17.0	8.1
<b>EDA</b>	AAAI25	<b>67.6</b>	<b>52.3</b>	<b>72.1</b>	<b>57.6</b>	<b>60.1</b>	<b>44.6</b>	<b>63.3</b>	<b>48.0</b>	<b>23.8</b>	<b>11.6</b>

Table 2: Performance of state-of-the-art universal Re-ID methods on nighttime datasets. The left part shows the four variations of NightReID, where IDs denote the number of identities in the testing set, and D refers to distractors. The CNN-based methods are listed in the upper section, while Transformer-based methods are in the lower section.

images is normalized using the mean and standard deviation calculated from the training set, as described in Eq. 8, thereby achieving data distribution alignment. Finally, we introduce standard distribution noise through random erasing (Zhong et al. 2020) before inputting the images into the backbone network for feature extraction.

$$DA_c = \frac{D_c - \text{Mean}_c}{\text{Std}_c}, c \in \{R, G, B\}, \quad (8)$$

where  $DA$  is the dataset  $D$  after data distribution alignment.

### Overall Architecture

Since both of our proposed modules are positioned before feature extraction, they can seamlessly integrate with existing Re-ID methods. To avoid the mismatch when DDA applies original statistical priors to align IED-enhanced images, each module operates independently in separate branches, with features concatenated during testing, forming the EDA framework. In practical applications, competitive results can be achieved even with a single module within a single-branch architecture.

The network is optimized by the commonly employed  $\mathcal{L}_{ID}$  and  $\mathcal{L}_{tri}$  in Re-ID. Specifically, the features extracted by the backbone are first optimized by the triplet loss  $\mathcal{L}_{tri}$ . After batch normalization through BNNeck (Luo et al. 2019), they are further optimized with the cross-entropy loss  $\mathcal{L}_{ID}$  calculated by the classifier. For the multi-branch architecture, features from both branches and their concatenation are all used to compute the losses and then summed together. The overall loss function is defined as:

$$\mathcal{L} = \mathcal{L}_{ID} + \mathcal{L}_{tri} + W_{IED}\mathcal{L}_{IED}, \quad (9)$$

where  $W_{IED}$  is the weight of  $\mathcal{L}_{IED}$ .

## 5 Experiment

### Datasets

In our experiments, we utilized the proposed NightReID dataset along with the Night600 dataset (Lu et al. 2023b).

The Night600 dataset contains 600 identities and 28,813 bounding boxes, captured by eight RGB cameras during

nighttime. The training set comprises 300 identities with 14,462 bounding boxes, while the remaining 300 identities form the gallery set. Additionally, 2,180 images in the gallery set are designated as queries.

### Implementation Details

For our baseline, we employed TransReID-SSL (Luo et al. 2021) as the backbone, which includes the IBN (Pan et al. 2018) style tokenizer and Transformer (Vaswani et al. 2017) encoder, and pre-trained on the LUPerson dataset (Fu et al. 2021). Unless specified otherwise, we adhered to the original implementation’s configurations. All images were resized to  $256 \times 128$  pixels, with a batch size of 64, and the training process spanned 120 epochs on a single NVIDIA RTX3090 GPU. We utilized the SGD optimizer with a momentum of 0.9. The learning rate was set to 0.0004, with a warm-up of 20 epochs followed by cosine decay. Data distribution was aligned based on the training set statistics of each dataset, and the IED module was randomly initialized.

### Performance on NightReID

We evaluated a series of state-of-the-art universal Re-ID methods published in recent years on the nighttime datasets, including BoT (Luo et al. 2019), AGW (Ye et al. 2021), MSINet (Gu et al. 2023), HAT (Zhang et al. 2021a), TransReID (He et al. 2021) TransReID-SSL (Luo et al. 2021), PFD-Net (Wang et al. 2022), DC-Former (Li et al. 2023), and CLIP-ReID (Li, Sun, and Li 2023). To ensure fair comparisons, we employed the official open-source implementations with default configurations.

As summarized in Table 2, the experimental results highlight the significant challenges posed by the NightReID dataset compared to daytime datasets, leading to lower performance across all methods. Despite these challenges, more robust backbones and advanced methods show gradual performance improvements. However, methods like PFD-Net and DC-Former experience performance degradation compared to their baselines due to difficulties in accurate posture estimation and extracting multiple distinguishable features from nighttime images. Additionally, the varying amounts of

Method	NightReID		Night600	
	Rank-1	mAP	Rank-1	mAP
IDF	65.6(+0.6)*	50.7(-0.1)*	17.2(+1.2)	9.2(+0.8)
CENet	-	-	19.2(+3.2) <sup>†</sup>	9.5(+1.1) <sup>†</sup>
EDA	<b>67.6(+2.6)</b>	<b>52.3(+1.5)</b>	<b>23.8(+3.8)</b>	<b>11.6(+2.1)</b>

Table 3: Comparison with nighttime RGB Re-ID methods. \* indicate the results are reproduced by authors, and <sup>†</sup> indicates the performance without additional training data.

Settings	NightReID				Night600	
	BASE	DCE	IED	DDA	R-1	mAP
✓					65.0	50.8
	✓				62.9	47.3
		✓			<u>67.0</u>	51.7
			✓		66.1	51.8
✓			✓		66.7	<u>52.0</u>
	✓		✓		66.5	51.8
		✓	✓		<b>67.6</b>	<b>52.3</b>
					<u>23.6</u>	<u>11.5</u>
					23.4	11.2
					<b>23.8</b>	<b>11.6</b>

Table 4: Ablation study of the proposed method. The upper section presents results from a single branch with one module, while the lower section shows results from the complete multi-branch framework integrating both modules.

identities and distractors in NightReID provide a more comprehensive evaluation of Re-ID methods. For subsequent experiments, we followed the default evaluation protocol involving 528 identities and distractors.

Moreover, the Night600 dataset, with its extremely dark and low-quality images, presents significant challenges for Re-ID methods, often leading to convergence issues and poor performance, thereby limiting its practical applicability. In contrast, NightReID captures high-quality RGB images with exposure gain under similar nighttime, demonstrating the feasibility of effectively utilizing nighttime RGB cameras for accurate Re-ID and providing a valuable benchmark for evaluating Re-ID methods during nighttime.

### Analysis of the EDA Framework

To evaluate the EDA framework, we conducted comparative experiments against two state-of-the-art nighttime Re-ID methods, IDF (Lu et al. 2023b) and CENet (Lu et al. 2023a), as shown in Table 3. Due to the lack of open-source code, we referenced the performance of IDF and CENet on the Night600 dataset from their original papers. Given the variations in backbones and epochs among different methods, we also reported improvements relative to the respective baselines for a fairer comparison. Additionally, we reproduced IDF using the same TransReID-SSL backbone as ours on the NightReID dataset. The results demonstrate that the EDA framework achieved significant improvements on both NightReID and Night600 datasets, surpassing state-of-the-art performance. Moreover, EDA showed greater improvements on Night600 compared to NightReID, which can be attributed to the significantly darker characteristics of the Night600 dataset.

### Ablation Study

To validate the specific contributions of our proposed modules and the advancements of the IED module over vanilla Zero-DCE, we conducted ablation experiments on multiple datasets, as presented in Table 4. The notations BASE, DCE, IED, and DDA represent the baseline, vanilla Zero-DCE, IED module, and DDA module, respectively. The results indicate that the DDA module effectively enhances performance even when applied to a single branch. The IED module also demonstrates improvements over vanilla Zero-DCE, although its enhancement effect may vary within a single branch. In the multi-branch framework, integrating DDA with the baseline leads to further performance gains. Moreover, employing DDA and IED in separate branches achieves the best performance across both datasets.

### Visualization

To further validate the effectiveness of our proposed method, we employed t-SNE (Van der Maaten and Hinton 2008) to visualize the feature distribution of the NightReID dataset, as shown in Figure 4. The results indicate that the integration of our modules leads to more distinct and compact clusters compared to the baseline, which demonstrates the enhanced discriminative power in challenging nighttime Re-ID tasks.

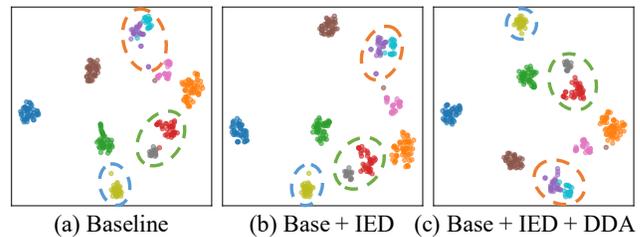


Figure 4: T-SNE visualization of the feature distribution for 10 randomly selected annotated identities in the NightReID dataset. Different colors represent different identities.

## 6 Conclusion

This paper addresses the critical but underexplored task of RGB-based nighttime Re-ID by introducing NightReID, a novel large-scale nighttime Re-ID dataset that outperforms existing datasets in both scale and diversity, providing a robust foundation for further research. Additionally, through comprehensive analysis of nighttime Re-ID data characteristics, we propose the Enhancement, Denoising, and Alignment (EDA) framework with two specialized modules: the Image Enhancement and Denoising (IED) module, which enhances nighttime images without reference, preserving details while removing noise for improved Re-ID suitability; and the Data Distribution Alignment (DDA) module, which aligns nighttime data with pre-training data via statistical priors to enhance data utilization and leverage large-scale pre-trained models. Extensive experiments validate the efficacy, flexibility, and applicability of our approaches in nighttime Re-ID. We believe that NightReID and the EDA framework will significantly advance research in this field.

## Acknowledgements

This work is supported by National Natural Science Foundation of China under Grants (62176188, 62066021, 62361166629, 62402365).

## References

- Al Sobhahi, R.; and Tekli, J. 2022. Comparing deep learning models for low-light natural scene image enhancement and their impact on object detection and classification: Overview, empirical evaluation, and challenges. *Signal Processing: Image Communication*, 109: 116848.
- Cai, Y.; Bian, H.; Lin, J.; Wang, H.; Timofte, R.; and Zhang, Y. 2023. Retinexformer: One-stage retinex-based transformer for low-light image enhancement. In *Proceedings of the ICCV*, 12504–12513.
- Chen, T.; Ding, S.; Xie, J.; Yuan, Y.; Chen, W.; Yang, Y.; Ren, Z.; and Wang, Z. 2019. Abd-net: Attentive but diverse person re-identification. In *Proceedings of the ICCV*.
- Chen, W.; Xu, X.; Jia, J.; Luo, H.; Wang, Y.; Wang, F.; Jin, R.; and Sun, X. 2023. Beyond appearance: a semantic controllable self-supervised learning framework for human-centric visual tasks. In *Proceedings of the CVPR*.
- Cui, Z.; Li, K.; Gu, L.; Su, S.; Gao, P.; Jiang, Z.; Qiao, Y.; and Harada, T. 2022. You Only Need 90K Parameters to Adapt Light: a Light Weight Transformer for Image Enhancement and Exposure Correction. In *33rd British Machine Vision Conference 2022*. BMVA Press.
- Danelljan, M.; Bhat, G.; Shahbaz Khan, F.; and Felsberg, M. 2017. Eco: Efficient convolution operators for tracking. In *Proceedings of the CVPR*, 6638–6646.
- Deng, J.; Dong, W.; Socher, R.; Li, L.-J.; Li, K.; and Fei-Fei, L. 2009. Imagenet: A large-scale hierarchical image database. In *Proceedings of the CVPR*, 248–255. Ieee.
- Du, B.; Du, C.; and Yu, L. 2023. Megf-net: multi-exposure generation and fusion network for vehicle detection under dim light conditions. *Visual Intelligence*, 1(1): 28.
- Fu, D.; Chen, D.; Bao, J.; Yang, H.; Yuan, L.; Zhang, L.; Li, H.; and Chen, D. 2021. Unsupervised pre-training for person re-identification. In *Proceedings of the CVPR*.
- Gu, J.; Wang, K.; Luo, H.; Chen, C.; Jiang, W.; Fang, Y.; Zhang, S.; You, Y.; and Zhao, J. 2023. Msinet: Twins contrastive search of multi-scale interaction for object reid. In *Proceedings of the CVPR*, 19243–19253.
- Guo, C.; Li, C.; Guo, J.; Loy, C. C.; Hou, J.; Kwong, S.; and Cong, R. 2020. Zero-reference deep curve estimation for low-light image enhancement. In *Proceedings of the CVPR*.
- He, S.; Luo, H.; Wang, P.; Wang, F.; Li, H.; and Jiang, W. 2021. Transreid: Transformer-based object re-identification. In *Proceedings of the ICCV*, 15013–15022.
- Huang, Y.; Zha, Z.-J.; Fu, X.; and Zhang, W. 2019. Illumination-invariant person re-identification. In *Proceedings of the 27th ACM international conference on multimedia*, 365–373.
- Jiang, Y.; Gong, X.; Liu, D.; Cheng, Y.; Fang, C.; Shen, X.; Yang, J.; Zhou, P.; and Wang, Z. 2021. Enlightengan: Deep light enhancement without paired supervision. *IEEE transactions on image processing*, 30: 2340–2349.
- Li, C.; Guo, C.; Han, L.; Jiang, J.; Cheng, M.-M.; Gu, J.; and Loy, C. C. 2021. Low-light image and video enhancement using deep learning: A survey. *IEEE transactions on pattern analysis and machine intelligence*, 44(12): 9396–9416.
- Li, C.; Guo, C.; and Loy, C. C. 2021. Learning to enhance low-light image via zero-reference deep curve estimation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 44(8): 4225–4238.
- Li, H.; Ye, M.; and Du, B. 2021. Weperson: Learning a generalized re-identification model from all-weather virtual data. In *Proceedings of the 29th ACM international conference on multimedia*, 3115–3123.
- Li, H.; Ye, M.; Wang, C.; and Du, B. 2022. Pyramidal transformer with conv-patchify for person re-identification. In *Proceedings of the 30th ACM International Conference on Multimedia*, 7317–7326.
- Li, S.; Sun, L.; and Li, Q. 2023. CLIP-ReID: exploiting vision-language model for image re-identification without concrete text labels. In *Proceedings of the AAAI*, volume 37, 1405–1413.
- Li, W.; Zou, C.; Wang, M.; Xu, F.; Zhao, J.; Zheng, R.; Cheng, Y.; and Chu, W. 2023. Dc-former: Diverse and compact transformer for person re-identification. In *Proceedings of the AAAI*, volume 37, 1415–1423.
- Liu, F.; Ye, M.; and Du, B. 2024. Learning a generalizable re-identification model from unlabelled data with domain-agnostic expert. *Visual Intelligence*, 2(1): 28.
- Liu, J.; Xu, D.; Yang, W.; Fan, M.; and Huang, H. 2021a. Benchmarking low-light image enhancement and beyond. *International Journal of Computer Vision*, 129: 1153–1184.
- Liu, R.; Ma, L.; Zhang, J.; Fan, X.; and Luo, Z. 2021b. Retinex-inspired unrolling with cooperative prior architecture search for low-light image enhancement. In *Proceedings of the CVPR*, 10561–10570.
- Lu, A.; Zha, T.; Li, C.; Tang, J.; Wang, X.; and Luo, B. 2023a. Nighttime Person Re-Identification via Collaborative Enhancement Network with Multi-domain Learning. *arXiv preprint arXiv:2312.16246*.
- Lu, A.; Zhang, Z.; Huang, Y.; Zhang, Y.; Li, C.; Tang, J.; and Wang, L. 2023b. Illumination distillation framework for nighttime person re-identification and a new benchmark. *IEEE Transactions on Multimedia*.
- Luo, H.; Gu, Y.; Liao, X.; Lai, S.; and Jiang, W. 2019. Bag of tricks and a strong baseline for deep person re-identification. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition workshops*.
- Luo, H.; Wang, P.; Xu, Y.; Ding, F.; Zhou, Y.; Wang, F.; Li, H.; and Jin, R. 2021. Self-supervised pre-training for transformer-based person re-identification. *arXiv preprint arXiv:2111.12084*.
- Neumann, L.; Karg, M.; Zhang, S.; Scharfenberger, C.; Piegert, E.; Mistr, S.; Prokofyeva, O.; Thiel, R.; Vedaldi, A.; Zisserman, A.; et al. 2019. Nightowls: A pedestrians at night dataset. In *Computer Vision—ACCV 2018, Revised Selected Papers, Part I 14*, 691–705. Springer.

- Nguyen, D. T.; Hong, H. G.; Kim, K. W.; and Park, K. R. 2017. Person recognition system based on a combination of body images from visible light and thermal cameras. *Sensors*, 17(3): 605.
- Pan, X.; Luo, P.; Shi, J.; and Tang, X. 2018. Two at once: Enhancing learning and generalization capacities via ibn-net. In *Proceedings of the ECCV*, 464–479.
- Qiu, L.; Chen, S.; Yan, Y.; Xue, J.-H.; Wang, D.-H.; and Zhu, S. 2024. High-Order Structure Based Middle-Feature Learning for Visible-Infrared Person Re-identification. In *Proceedings of the AAAI*, volume 38, 4596–4604.
- Ren, S.; He, K.; Girshick, R.; and Sun, J. 2016. Faster R-CNN: Towards real-time object detection with region proposal networks. *IEEE transactions on pattern analysis and machine intelligence*, 39(6): 1137–1149.
- Shi, J.; Yin, X.; Chen, Y.; Zhang, Y.; Zhang, Z.; Xie, Y.; and Qu, Y. 2025. Multi-memory matching for unsupervised visible-infrared person re-identification. In *European Conference on Computer Vision*, 456–474. Springer.
- Shi, J.; Yin, X.; Zhang, Y.; Xie, Y.; Qu, Y.; et al. 2024. Learning commonality, divergence and variety for unsupervised visible-infrared person re-identification. In *The Thirty-eighth Annual Conference on Neural Information Processing Systems*.
- Shi, J.; Zhang, Y.; Yin, X.; Xie, Y.; Zhang, Z.; Fan, J.; Shi, Z.; and Qu, Y. 2023. Dual pseudo-labels interactive self-training for semi-supervised visible-infrared person re-identification. In *Proceedings of the ICCV*, 11218–11228.
- Sun, Y.; Zheng, L.; Yang, Y.; Tian, Q.; and Wang, S. 2018. Beyond part models: Person retrieval with refined part pooling (and a strong convolutional baseline). In *Proceedings of the ECCV*, 480–496.
- Van der Maaten, L.; and Hinton, G. 2008. Visualizing data using t-SNE. *Journal of machine learning research*, 9(11).
- Vaswani, A.; Shazeer, N.; Parmar, N.; Uszkoreit, J.; Jones, L.; Gomez, A. N.; Kaiser, Ł.; and Polosukhin, I. 2017. Attention is all you need. *Advances in neural information processing systems*, 30.
- Wang, G.; Yuan, Y.; Chen, X.; Li, J.; and Zhou, X. 2018. Learning discriminative features with multiple granularities for person re-identification. In *Proceedings of the 26th ACM international conference on Multimedia*, 274–282.
- Wang, T.; Liu, H.; Song, P.; Guo, T.; and Shi, W. 2022. Pose-guided feature disentangling for occluded person re-identification based on transformer. In *Proceedings of the AAAI*, volume 36, 2540–2549.
- Wei, L.; Zhang, S.; Gao, W.; and Tian, Q. 2018. Person transfer gan to bridge domain gap for person re-identification. In *Proceedings of the CVPR*, 79–88.
- Wu, A.; Zheng, W.-S.; Yu, H.-X.; Gong, S.; and Lai, J. 2017. RGB-infrared cross-modality person re-identification. In *Proceedings of the ICCV*, 5380–5389.
- Xiang, S.; Qian, D.; Guan, M.; Yan, B.; Liu, T.; Fu, Y.; and You, G. 2023. Less is more: Learning from synthetic data with fine-grained attributes for person re-identification. *ACM Transactions on Multimedia Computing, Communications and Applications*, 19(5s): 1–20.
- Yan, Z.; Zheng, Y.; Fan, D.-P.; Li, X.; Li, J.; and Yang, J. 2024. Learnable differencing center for nighttime depth perception. *Visual Intelligence*, 2(1): 15.
- Ye, M.; Chen, S.; Li, C.; Zheng, W.-S.; Crandall, D.; and Du, B. 2024a. Transformer for object re-identification: A survey. *International Journal of Computer Vision*, 1–31.
- Ye, M.; Shen, J.; Lin, G.; Xiang, T.; Shao, L.; and Hoi, S. C. 2021. Deep learning for person re-identification: A survey and outlook. *IEEE transactions on pattern analysis and machine intelligence*, 44(6): 2872–2893.
- Ye, M.; Shen, W.; Zhang, J.; Yang, Y.; and Du, B. 2024b. Securereid: Privacy-preserving anonymization for person re-identification. *IEEE Transactions on Information Forensics and Security*.
- Ye, M.; Wu, Z.; Chen, C.; and Du, B. 2023. Channel augmentation for visible-infrared re-identification. *IEEE Transactions on Pattern Analysis and Machine Intelligence*.
- Ye, M.; Zhang, X.; Yuen, P. C.; and Chang, S.-F. 2019. Unsupervised embedding learning via invariant and spreading instance feature. In *Proceedings of the CVPR*, 6210–6219.
- Zeng, Z.; Wang, Z.; Wang, Z.; Zheng, Y.; Chuang, Y.-Y.; and Satoh, S. 2020. Illumination-adaptive person re-identification. *IEEE Transactions on Multimedia*, 22(12).
- Zhang, G.; Luo, Z.; Chen, Y.; Zheng, Y.; and Lin, W. 2022. Illumination unification for person re-identification. *IEEE Transactions on Circuits and Systems for Video Technology*, 32(10): 6766–6777.
- Zhang, G.; Zhang, P.; Qi, J.; and Lu, H. 2021a. Hat: Hierarchical aggregation transformers for person re-identification. In *Proceedings of the 29th ACM International Conference on Multimedia*, 516–525.
- Zhang, J.; Yuan, Y.; and Wang, Q. 2019. Night person re-identification and a benchmark. *IEEE Access*, 7.
- Zhang, Y.; Di, X.; Zhang, B.; Li, Q.; Yan, S.; and Wang, C. 2021b. Self-supervised low light image enhancement and denoising. *arXiv preprint arXiv:2103.00832*.
- Zhang, Y.; and Wang, H. 2023. Diverse embedding expansion network and low-light cross-modality benchmark for visible-infrared person re-identification. In *Proceedings of the CVPR*, 2153–2162.
- Zhang, Y.; Zhang, J.; and Guo, X. 2019. Kindling the darkness: A practical low-light image enhancer. In *Proceedings of the 27th ACM international conference on multimedia*.
- Zheng, L.; Shen, L.; Tian, L.; Wang, S.; Wang, J.; and Tian, Q. 2015. Scalable person re-identification: A benchmark. In *Proceedings of the ICCV*, 1116–1124.
- Zhong, Z.; Zheng, L.; Kang, G.; Li, S.; and Yang, Y. 2020. Random erasing data augmentation. In *Proceedings of the AAAI*, volume 34, 13001–13008.
- Zuiderveld, K. 1994. *Contrast limited adaptive histogram equalization*, 474–485. USA: Academic Press Professional, Inc. ISBN 0123361559.