International Journal of
INTELLIGENT SYSTEMS AND APPLICATIONS IN ENGINEERING
ISSN:2147-6799 www.ijisae.org Original Research Paper

Person Re-identification Based on Enhanced Position-Regularization Generative Adversarial Network (EPR-GAN) using GLCM, Radon Transform, and DCT

Sharath S. [1*], Rangaraju H. G. [2], Leelavathi G.[3]

Submitted: 24/04/2023 **Revised:** 26/06/2023 **Accepted:** 03/07/2023

Abstract: Person Re-identification (Re-id) is helpful in society for non-invasive biometric person identification, validation, and surveillance, in crowded places. This paper proposes person reidentification based on Enhanced Position-Regularization Generative Adversarial Network (EPR-GAN) using GLCM, radon transform, and DCT. The EPR-GAN model produces output images of the same person with the new postures. A set of eight established poses are defined, and create eight new images for any given picture of a person. The texture analysis and a subspace learning approach are used for features using Gray-Level Co-occurrence Matrix (GLCM), radon transform, and Discrete Cosine Transform (DCT) on EPR-GAN generated images. The GLCM technique is adopted with matrix sizes of 4X4, 8X8, 16X16, 32X32, and 64X64 on the given resized image of 128X64 to discover the local details of the set of images. As the dimension of the GLCM increased, the rank-1 recognition also increased due to the large size; however, the GLCM matric dimension is limited to 64X64, i.e., 4096. The radon transform is the image intensity projection along a radial line oriented at a precise angle applied on the GLCM matrix to enhance the rank one recognition accuracy. The resultant feature values of radon transform applied on 64X64 GLCM matrix is 95X180, i.e., 17100 is very high. Furthermore, the dimensionality reduction algorithm DCT is used on radon transform to decrease the number of final compelling features and enhance the model's performance. As a result, the proposed model's efficiency is better than the existing models.

Keywords*: DCT, GAN, GLCM, Person Re-identification, Radon Transform.*

1. Introduction

Human identification based on appearance, even when their face is hidden, is called person re-identification (Re-id). It retrieves a person of attention across numerous non-overlapping cameras captured in crowded places. It has gained momentum in recent years and has the potential to play essential roles in pedestrian surveillance, attracting greater attention from researchers due to most of its real-world applications in intelligent surveillance, area monitoring, and security.

The cross-scenario Re-ID attempts to identify individuals taken by several cameras at dissimilar times and locations, which demands a more effective identification model across different datasets. Due to the extensive gallery search, different camera settings, low identification rate, changes in viewpoint, pose, illumination, and dense background, Re-ID approaches encounter different barriers. The two critical issues with person Re-ID is metric learning and feature representation.

To address these challenges of Re-ID, a wide range of identification techniques are stated in the significant literature. The primary method is appearance-based, which emphasizes distinctive feature representation [1, 2] and is robust against the variation in pose, illumination, and vision. The other well-known metric learning or optimal subspace learning-based methods [3, 4, 5] are used to learn an optimum metric from a particular set of features. The feature representation uses similarity matching techniques to find the minimum distance between the probe images and the entire gallery set to find an accurate match. In this concern, several similar measuring techniques used are, Euclidian distance [6], Mahalanobis distance [7], and Bhattacharyya distance [8].

Contribution: The EPR-GAN generates images of the same person with the new postures. The texture features are extracted using GLCM from EPR-GAN generated images then radon transform, and DCT is used on the final features. As the dimension of the GLCM increased, the rank-1 recognition also increased due to the large size; however, the GLCM matric dimension is limited to 64X64, i.e., 4096. The dimensionality reduction algorithm DCT is used on radon transform to decrease the number of features.

Organization: The literature survey on existing techniques is given in section 2. The proposed model is discussed in section 3. The result analysis is deliberated in section 4. The final comments are specified in section 5.

[1]*Gec Kr Pet, Vtu Belgaum, India*
[2]*Gec Chamarajanagar, India*
[3]*Gsksjti Bengaluru, India*
[1*]*ss.sharath@gmail.com,*
[2]*rangraju@gmail.com,* [3]*nisargamodini@gmail.com`*

2. Literature Survey

Lei Qi et al., [9] projected an Adversarial Camera Alignment Network (ACAN) for person Re-ID. A Multi-Camera Adversarial Learning (MCAL) camera alignment was developed to map images of dissimilar cameras into a collective subspace. The two schemes are explored, including the current gradient reversal layer (GRL) structure and the projected structure Other Camera Equiprobability (OCE), to deportment the multi-camera confrontational job. The experiments on five large-scale datasets determine the dominance of ACAN over existing procedures that benefit from labeled source domains and generated images by GAN-based models. Yusheng Tao et al., [10] proposed a Diversity Enlarged Mutual Teaching framework (DREAMT). The two strategies are developed in DREAMT: GAN-based source domain augmentation (GSDA) and Cross-Branch Mutual Supervision (CBMS) for dual networks. The GSDA uses two GANs to augment source domain datasets in different ways for pre-training to improve the pre-trained models' results and enlarge the miscellany at the start of target domain adjustment. Each network in the Mutual Teaching framework adopts two branches to extract dissimilar features.

Wenfeng Song et al., [11] proposed a Context-Interactive CNN (CI-CNN) to find the spatial and temporal environments by implanting multi-task Reinforcement Learning (MTRL). The CI-CNN restructures the multitask reinforcement learning by using actor-critic means to get the spatial and temporal information simultaneously. The earlier network was used to determine the prime spatial context region and temporal structure range. De Cheng et al., [12] planned a person re-identification technique based on the superior fusion algorithms using the ranking outcomes that are attained in the test datasets, and the robust combination approach is smeared to attain better re-ranked matching outcomes by considering the recognition capabilities. Guodong Ding et al., [13] presented a similarity-based pseudo-labeling method with two possible label encodings. In the network with cross-entropy loss and with a center regularization term, the discriminative feature is trained simultaneously by predicting pseudo-labels for unlabelled data. Both label encodings are learned and combined to improve the overall performance. Minxian Li et al., [14] presented an unsupervised re-identification deep learning method called Unsupervised Tracklet Association Learning (UTAL). It is used to find the principal re-identification information from the person-tracked data end-to-end and learning within-camera tracklist to maximize the detection of trackless distinctiveness for matching both within and through camera sights.

Tetsu Matsukaway et al., [15] presented a hierarchical Gaussian distribution of pixel features, in which both mean and covariance data are contained within the patch and region-level metaphors. The area is modeled as a set of multiple Gaussian distributions, each representing the presence of a local patch. As the space of Gaussian distribution is not linear, the distribution parameters are embedded into a point of the Symmetric Positive Definite (SPD) matrix. Normalizing the scale of the SPD matrix enriches the graded feature depiction.

Navaneet K L et al., [16] proposed a union function that works on deep feature illustrations of query and gallery matches. An optimization method is customized to increase inquiry representation. Jianjun Lei et al., [17] proposed a steadfast illustration named Semantic Region Representation (SRR) and an effective metric learning called mapping space topology constraint (MSTC). The SRR incorporates semantic depictions to attain an operational correspondence between the consistent regions by analyzing the multiple parts of the body, which emphasizes on the focal point in contradiction of the background intrusion. The MSTC is proposed to consider the topological connections between all the samples in the feature space. Long Wei et al., [18] proposed a Self-Inspirited Feature Learning (SIF) technique to increase the Re-Identification networks' performance from the optimization perspective. A simple confrontational learning system helps the network to learn more about discriminative person depiction. In the proposed technique, an assisting branch is added to the network at the training stage, and the organization of the inventive network remains unaffected during the testing stage.

Mang Ye and Pong C. Yuen [19] proposed a purifyNet to equally enhance the interpreted labels and optimize the neural networks by gradually modifying the predicted labels. The proposed model simultaneously reduces the undesirable noisy labels and concentrates on complex samples with correct labels by increasing a hard-aware instance re-weighting approach. B Hong-Xing Yu et al., [20] an unsupervised asymmetric distance metric based cross-view clustering. The asymmetric distance metric permits detailed feature conversions for each camera view to tackle exact feature misrepresentations. The Deep Clustering-based Asymmetric Metric Learning (DECAMEL) learns a dense cross-view cluster arrangement of Re-identification data and helps to improve the view-specific preference and simplify the potential cross-view discriminative data for unsupervised Re-ID.

3. Proposed Model

The proposed model's outline is illustrated in Figure 1, and it comprises of two main parts: a person re-id feature-learning model and a GAN-based person image creation technique.

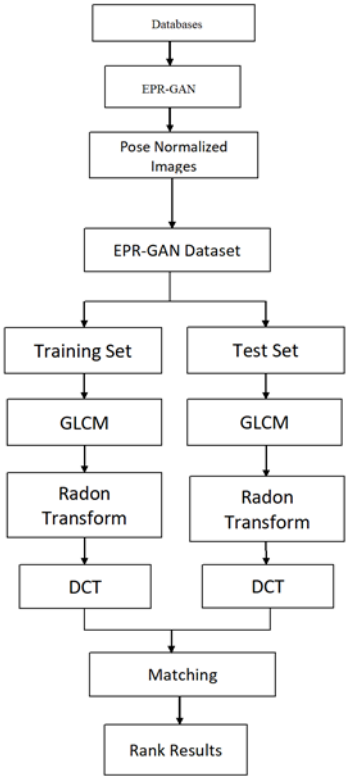

Fig 1 The proposed model

3.1 Image datasets:

Two standard datasets are used to test the projected technique:

3.1.1 Market-1501 [21]

The pictures were taken using six distinct camera angles. It has 1,501 identities among 32,668 bounding boxes. The 750 identities with 19,732 images are utilized for testing, while the 751 with 12,936 images are used for training. The EPR-GAN model is trained using the training set. It is a significant person Re-ID database distributed in 2015. It was collected at Tsinghua University by one low-resolution camera and five high-resolution cameras opposite of a supermarket. For automatic pedestrian bounding box identification, the Market1501 utilizes a pedestrian detector. 32,668 photos of 1,501 pedestrians, each measuring 128 x 64 pixels, make up the entire collection. The sample images are shown in Figure 2. Market-1501 is a more realistic photo and consists of 2793+500k interfering components.

Fig 2 The Market-1501 dataset images [21]

3.1.2 CUHK01 [22]: There are 971 subjects in the collection, 3,884 individually cropped photos, and at least two images from two different camera angles for each individual. The image samples are revealed in Figure 3. Camera *A* in the database consists of multiple perspectives, and poses are used as the probing images, while camera *B* primarily captures the frontal view and the rear end.

Fig 3 Sample images from CUHK01 dataset [22]

3.2 Enhanced Pose-Regularization GAN model architecture

The image generator aims to create images of the same subject in several poses. In order to create a new human image with eight various postures, the input image of a person and the target position image are combined. The image generator has two modules, a Generator, and a Discriminator, much like in any other GAN model. The generator is trained to alter a person's image only available in a specific stance. The discriminator helps improve the created images' quality by distinguishing between the generated and actual data samples.

(i) Position Estimation

The input image and the wanted pose, represented by a skeleton picture, are used in the image-generating process. First, posture estimations [23] are attained by a pre-trained model using the Open Pose toolkit [24]. Then, for a given input image, the pose estimator produces a posed photo, which confines and identifies eighteen anatomical key points with connections. In the model, pose regularization is focused, and only eight established postures are chosen with different angles of poses, as shown in Figure 4.

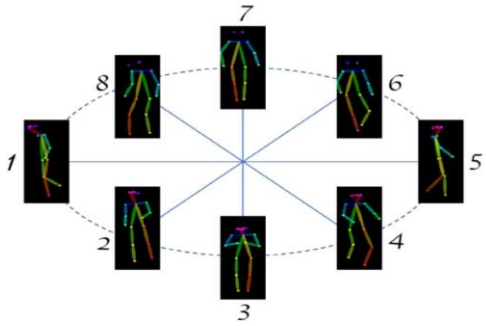

Fig 4 Eight established postures on Market-1501[23]

(ii) Generator

An input image and a target image of an identical person with a different pose are trained to generate a new target pose image, as shown in Figure 5.

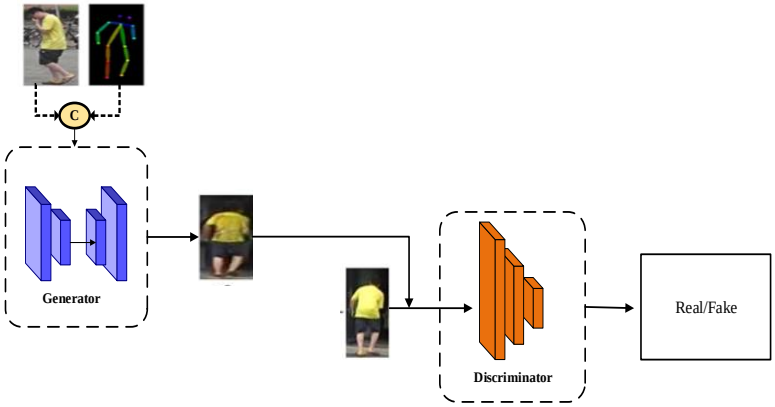

Fig 5 Block diagram representation of EPR-GAN model

The generator consists of encoder-decoder architecture, where the encoder converts the input person image into eigenvectors, and the decoder uses its internal representations and its learning weights to attempt to reassemble the data. The encoder-decoder architecture for the generator is shown in Figure 6.

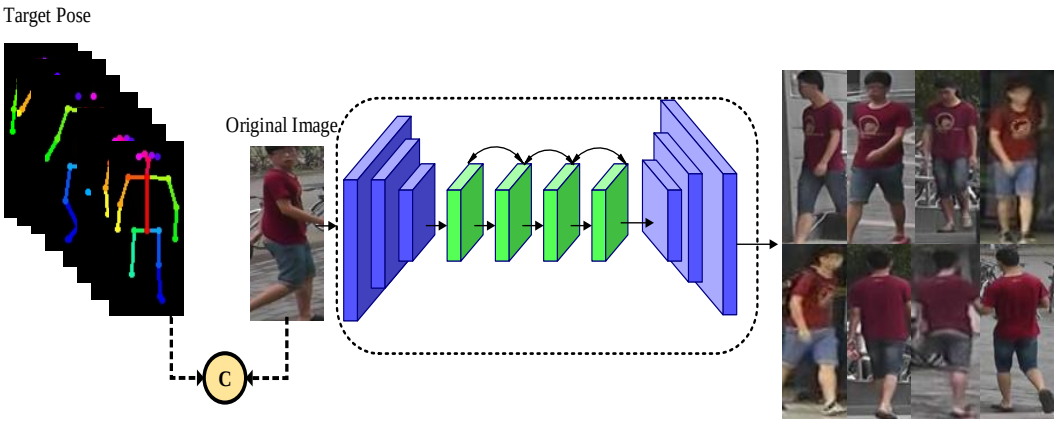

Fig 6 Structure of Generator in EPR-GAN model

The RGB original image and target pose image are of the same size, 128X64 and concatenated, and the resultant image is 256X128. The resultant images are fed as input images to the encoder of the generator. A Deep Convolution Network (DCN) is utilized to obtain the eigenvectors. The encoder and decoder structures are tabulated in Table 1 and Table 2, correspondingly. The Kernel Size is increasing from 3 to 7 for better normalization with batch norm 16.

Table 1 The Encoder Structure

Layers	Input		Kernel Size	Stride	Padding	Output	
	Size	Channel				Size	Channel
1	256X128	3	7	1	0	250X122	64
2	250X122	64	3	2	1	125X61	128
3	125X61	128	3	2	1	63X31	256

Table 2 The Decoder Structure

Layers	Input		Kernel Size	Stride	Padding	Output	
	Size	Channel				Size	Channel
1	63X31	256	3	2	1	125X61	128
2	125X61	128	3	2	1	250X122	64
3	250X122	64	7	1	0	256X128	3

(iii) Discriminator

The discriminator's objective is to learn the generated images to distinguish the input images as real or fake using binary classification. The goal of optimization is to find the optimal generator G_P. The first phase is to iteratively minimize generator L_{G_P} and discriminator L_{D_P} loss functions until convergence.

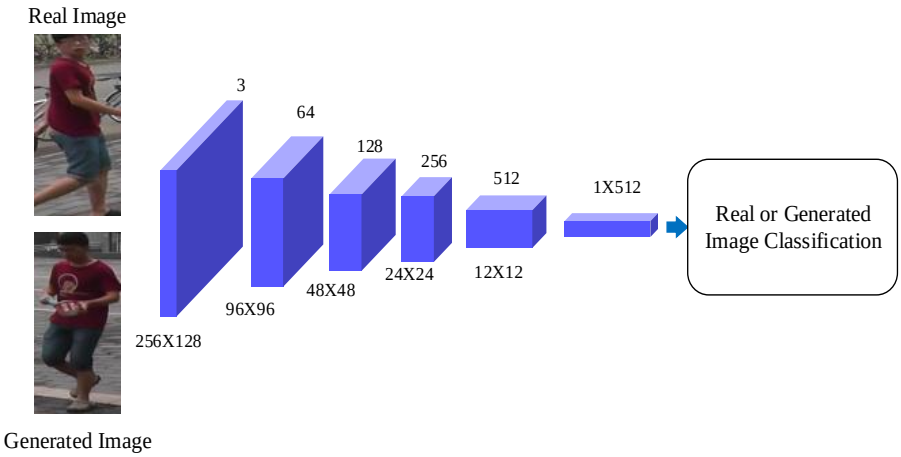

Fig 7 Structure of Discriminator in EPR-GAN model

The 2D Convolution and Leaky ReLu make up the discriminator's encoder. Figure 7 depicts the Discriminator's structure, and Table 3 lists the structural values.

Table 3 The Discriminator Structure

Layers	Input		Kernel Size	Stride	Padding	Output	
	Size	Channel				Size	Channel
1	256X128	3	4	2	1	128X64	64
2	128X64	64	4	2	1	64X32	128
3	64X32	128	4	2	1	32X16	256
4	32X16	256	4	2	1	16X8	512
5	16X8	512	4	2	1	8X4	512
6	8X4	512	4	1	1	7X3	512
7	7X3	512	1	1	0	7X3	512
Adaptive AvgPool2d	7X3	512	--	--	--	1	512

(iv) Generated GAN images

A few examples of generated image pose from EPR-GAN are shown in Figure 8. For a given input image, a realistic image with 8 different poses are generated, with a related visual look as the input person image. The critical contribution of the EPR GAN aims to generate synthesized images to retain the key features and simultaneously remove the occlusion.

Fig 8 Visualization of various poses produced by the EPR-GAN model

(v) Generated GAN images for the Market-1501 dataset

GAN generated 8 different poses for the original image of the Market-1501 dataset as shown in Figure 9.

Fig 9 Different poses produced by the EPR-GAN model for Market-1501 dataset

(vi) Generated GAN images for the CUHK01 dataset

GAN generated 8 poses for the original image of the CUHK01 dataset as revealed in Figure 10.

Fig 10 Different poses produced by the EPR-GAN model for the CUHK01 dataset

3.3 Feature Extraction

3.3.1 Grey Level Co-occurrence Matrix (GLCM)

It is the most traditional second-order statistical technique for analyzing texture. The GLCM [25] is a chart that depicts how frequently certain mixtures of grey levels co-occur in an image or image portion. An image is made up of pixels with an intensity of grey levels. Calculations of texture features employ the information in the GLCM to measure the intensity variation at the target pixel. The GLCM provides information on the likelihood of encountering a pair of pixels or grey levels in a specific orientation over an entire image or area. The GLCM and its related qualities, which are textural characteristics, can describe the spatial dependencies of seismic faces. The number of rows and columns in GLCM equals the grey scale maximum pixel intensity value. The generation of GLCM is shown in figure 11. The horizontal pixel pair (1, 1) occurs only once in the matrix, hence the value one is filled in the 1st row and 1st column. The horizontal pixel pair (6,2) appears three times; hence the value three is supplied in 6^{th} row and 2^{nd} column. Similarly, all horizontal pixel pairs are mapped to generate a GLCM matrix. The process continues till all pixel pairs are exhausted.

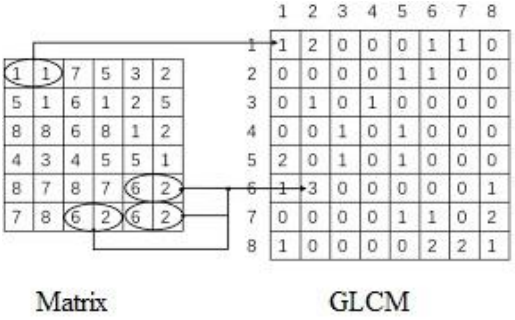

Fig 11 GLCM generation

3.3.2 Radon Transform [26]

The radon function computes an image matrix's projections along predetermined directions. For example, a collection of line integrals makes up a projection of the 2D function f (x, y). By using parallel routes, or beams, in a specific direction, the radon function computes the line integrals from many sources.

3.3.3 Discrete Cosine Transform (DCT)

The DCT is data compression method used to solve numerous glitches in digital image processing. Image compression methods engage the DCT [27] and are widely used in Face Recognition. The DCT is data-independent, separable, tends to concentrate information, and can be implemented effectively using the fast Fourier algorithm. The DCT for a 2D image is specified in Equations (1) and (2).

$$\mathrm{DCT}(i,j) = \frac{1}{\sqrt{2N}} C(i)C(j) \sum_{x=0}^{N-1} \sum_{y=0}^{N-1} \mathrm{pixel}(x,y) \cos\left[\frac{(2x+1)ix}{2N}\right] \cos\left[\frac{(2y+1)jx}{2N}\right] \quad (1)$$

$$C(x) = \frac{1}{\sqrt{2}} \text{ if } x \text{ is } 0, \text{ else } 1 \text{ if } x>0 \quad (2)$$

An illustration of the DCT on the 7X5 matrix

The pixel values of the 7X5 cell considered are shown in Figure 12 for an illustration of DCT.

The DCT is applied to the pixel values of the 7x5 cell to get the DCT coefficients, as shown in Figure12. The DCT coefficients have relatively large values on the upper-left corner of the matrix which represents original image focuses on the upper-left corner of the DCT matrix, and the DCT coefficient values are lesser on the lower-right part of the matrix. The DCT has significant information at the upper-left corner of the matrix, and the non significant information present at lower-right part of the matrix.

$$\begin{bmatrix} 60 & 60 & 62 & 53 & 48 \\ 58 & 73 & 54 & 63 & 62 \\ 68 & 77 & 80 & 67 & 57 \\ 79 & 76 & 56 & 62 & 64 \\ 49 & 69 & 68 & 63 & 79 \\ 43 & 65 & 89 & 79 & 68 \\ 52 & 62 & 67 & 71 & 82 \end{bmatrix}_{7X5} \Rightarrow \begin{bmatrix} 386.2355 & -8.2225 & -15.1651 & -12.6222 & -3.8104 \\ -16.0620 & 31.3364 & 5.9208 & 3.3276 & -5.5680 \\ -13.7872 & -13.6020 & -5.7534 & 3.4427 & 3.8004 \\ -10.8993 & -5.5820 & -3.1918 & -2.6484 & 4.1582 \\ -4.3475 & 13.5171 & 18.9555 & 2.6682 & 0.6123 \\ 6.5589 & 6.9342 & -20.9887 & 10.2078 & 9.3079 \\ 0.5950 & -5.3113 & -6.1824 & -4.3775 & 7.6711 \end{bmatrix}_{7X5}$$

Fig 12. The coefficients of DCT from original matrix

3.4 Matching

In order to learn a metric that highlights inter-personal distance and de-emphasizes intra-personal distance, metric learning is frequently used in person re-identification. The final determination of whether or not a person has been re-identified is made using the learned metric. Similarity measurements are applied in order to determine the distance between two images from various views of image pairs. There is a substantial variance among the feature spaces of two separate perspectives as a result of the diversity of camera angles. The feature spaces between the transformations must be learned to fill in the gaps among the views to enable a more accurate distance computation, which is necessary to compute the similarity of a pair based on the appearances of various characteristics. It is crucial for a person Re-ID task to need an effective distance measure because the changes in samples are naturally very challenging to discover with the obtained high-dimensional features. The extensively used four formulation distance metric functions are used in the proposed model and are discussed.

3.4.1 City block

The distance is also referred to as Manhattan distance which is the distance between two points *a* and *b* with *k* dimensions is computed as shown in Equation (3),

$$\mathrm{CD} = \sum_{j=1}^{k} |a_j - b_j| \quad (3)$$

3.4.2 Euclidean Distance (ED)

It is determined as the square root of the sum of the squared differences between the two vectors and represents the straight-line distance between two pixels as given by Equation (4),

$$\mathrm{ED}(p, q) = \sqrt{\sum_{i=2}^{n} (q_i - p_i)^2} \quad (4)$$

Where p and q are the two points of two matrices or vectors

3.4.3 Cosine

In *n*-dimensional space, it is the cosine of the angle formed by two n-dimensional vectors and is defined as in Equation (5),

$$\text{Cosine Similarity} = \frac{\sum_{i=1}^{n} x_i y_i}{\sqrt{\sum_{i=1}^{n} x_i^2} \sqrt{\sum_{i=1}^{n} y_i^2}} \quad (5)$$

The cosine distance is then defined as, Cosine Distance = 1−Cosine Similarity.

3.4.4 Mahalanobis Cosine

It is the distance among two vectors that display projections of training images and projections of testing images. The matches' scores range from -1.0 to 1.0, with 1.0 representing a perfect score. Mahalanobis Cosine is the cosine of the angle between the images and has been projected in the space of Mahalanobis and standardized by the measurement of the and is given by the Equation (6),

$$d_{MahCosine}(u, v) = -\frac{m.n}{|m||n|} \quad (6)$$

With m and n are two vectors of Mahalanobis space corresponding to u and v. The relation between these vectors is defined as:

$$m_i = \frac{u_i}{\sigma_i} \text{ and } n_i = \frac{v_i}{\sigma_i}$$

With σ_i is the variance along i[th] dimension.

To explore the local details of the set of images, GLCM is adopted with matrix size of 4X4, 8X8, 16X16, 32X32 and 64X64 for the given 128X64 image. As the dimension of the GLCM increased the rank-1 recognition is also increased, due to the high dimension the GLCM matric dimension is limited to 64X64 i.e., 4096. To improve the rank 1 recognition accuracy, Radon Transform is applied on GLCM mq1atrix. The radon transform is the projection of the image intensity along a radial line oriented at a specific angle. The resultant feature values of radon transform applied on 64X64 GLCM matrix is 95X180 i.e., 17100.

The subspace learning method LDA is used to decrease the dimension of the obtained features that have very high dimensions, which could potentially have an impact on the final re-identification outcomes. The LDA functions as a dimensionality reduction algorithm, reducing the original C number of features (where C is the number of classes) to C-1 features. It is important to strike a compromise between dimension reduction and recognition rate, which entails keeping representational information while reducing computation costs. The solution is to boost recognition rate by mapping the retrieved data into a common feature subspace with reduced dimension.

4 Result Analysis

A person Re-ID turns out to be an important mission in surveillance systems to identify a person through multiple camera visions. The cross-scenario Re-ID aims to match an individual automatically captured by multiple cameras at different locations or period, which needs the identification model to be more resourceful over different datasets. In view of this, Re-ID approaches have to face many challenges of extensive gallery search, varied camera settings, low recognition rate, dissimilarities in viewpoint, pose, illumination and cluttered background. The feature representation and metric learning are the major issues for person Re-ID.

4.1 Evaluation Metrics

The person Re-ID evaluation is performed with the quantitative experimental results viz., the standard Cumulated Matching Characteristics (CMC) curves and Rank-k.

(i) CMC curves:

It is the relationship among re-identification correct rate and the rank number. An algorithm rank (score) all the gallery and the query samples according to their distances from small to large, and CMC top-k accuracy is given by the Equation (7),

$$Acc_k = \begin{cases} 1 \; ; \; top\ k\ ranked\ gallery\ samples\ contain\ the\ query\ idenity \\ 0 \; ; \; otherwise \end{cases} \quad (7)$$

The concluding CMC curve is computed by averaging the shifted step function over all the queries and judges the ranking capabilities of the Re-ID model.

(ii) Rank-k:

The likelihood that the top k retrieved results contain the matching accuracy metric determines the right match. An unseen probe image's collected feature vector, x^p, is classed using previously learnt classification procedures to provide its prototype label, c. A gallery/target image's feature vector's distance x^q against it is calculated using the following Equation (8),

$$\text{Dist}(x^p, x^q) = \|(w^c)^T |x^p - x^q|\|_1 \quad (8)$$

A lesser distance outcomes in a higher rank.

4.2 Implementation Details

The experiments are carried out and verified on the four benchmark data sets using various training sizes after features are extracted to assess the robustness and effectiveness of the proposed model. The image samples from each person are divided into two equal-sized groups, one for training and the other for testing, at random. The goal of person reidentification is to use the trained model to find the appropriate matching image of a test image in the gallery set. The first camera's picture sequences are utilized as the probe set and the second camera's image sequences are used as the gallery set, respectively, during the testing phase. The probability of similarity between two objectives increases with decreasing distance between them. A rank list is created by ranking the distance between a given image in the probe set and each image in the gallery set. The perfect scenario is when the correct matching image comes in first on the rank list, indicating that it is the image that comes closest to matching the probe image. At rank 1, it is challenging to discover every target of interest, and the matching precision is typically low. To demonstrate the matching performance, the

average cumulated matching curve at various ranks is employed.

4.3 Results on large-scale dataset

The results are calculated using Market 1501 dataset with variations in pose, random occlusions, and clothing similarities. The generated images from EPR-GAN are used as GAN dataset to perform re-id. The performance of the projected framework with 750 image pairs of the dataset are calculated. The assessment of projected framework with GLCM Matrix of different sizes at rank-1 and rank-5 as offered in Table 4 and Table 5 for four distance matching techniques. The re-id is performed on both GAN dataset and original dataset to check the efficacy of the EPR-GAN model. According to the results, the GAN dataset attained the highest rank 1 matching rate of 99% is attained for 64X64 matrix size with ED matching distance compared with original dataset.

Table 4. Rank rate (%) on Market 1501 dataset using original dataset

Matching Distance	GLCM Matrix	Rank 1	Rank 5
Cityblock	4X4	52.67	88.23
	8X8	67.98	93.09
	16X16	69.39	95.50
	32X32	77.66	96.72
	64X64	85.00	98.41
Cosine	4X4	62.66	85.82
	8X8	72.78	89.39
	16X16	83.40	90.33
	32X32	88.40	97.72
	64X64	92.00	98.56
Euclidean Distance	4X4	72.67	89.22
	8X8	79.90	92.19
	16X16	87.24	95.00
	32X32	93.19	97.72
	64X64	95.00	98.90
Mahalanobis Cosine	4X4	42.88	68.33
	8X8	57.99	77.89
	16X16	66.19	85.99
	32X32	75.66	90.01
	64X64	86.10	93.77

Table 5. Rank rate (%) on Market 1501 dataset using GAN dataset

Matching Distance	GLCM Matrix	Rank 1	Rank 5
Cityblock	4X4	55.07	87.30
	8X8	70.10	89.96
	16X16	76.18	93.24
	32X32	79.76	98.42
	64X64	89.24	98.81
Cosine	4X4	70.29	86.44
	8X8	75.80	91.24
	16X16	88.00	96.60
	32X32	90.81	97.80
	64X64	94.42	98.00
Euclidean Distance	4X4	78.87	90.82
	8X8	81.63	92.89
	16X16	89.47	96.45
	32X32	94.20	98.19

	64X64	96.50	99.00
Mahalanobis Cosine	4X4	49.24	69.30
	8X8	59.20	78.22
	16X16	66.90	80.12
	32X32	79.38	93.00
	64X64	88.50	95.10

4.4 Results on the small-scale dataset

The results are computed using a small dataset CUHK01 with variations in pose, random occlusions, and clothing similarities. The performance of the projected framework with 200 identities is calculated. The assessment of the projected framework with a different matrix sizes at rank-1 and rank-5 as presented in Table 6 and Table 7 for four distance matching techniques. The re-id is performed on both the GAN dataset and the original dataset to check the efficacy of the EPR-GAN model. According to the results, the GAN dataset attained the highest rank 1 matching rate of 69.82% for 64X64 matrix size with ED matching distance compared with the highest rank 1 matching rate of 66.01% for original dataset.

Table 6 Rank rate (%) on CUHK01 dataset using original dataset

Matching Distance	GLCM Matrix	Rank 1	Rank 5
Cityblock	4X4	21.13	55.21
	8X8	38.39	59.24
	16X16	44.80	60.84
	32X32	47.90	72.59
	64X64	49.10	79.00
Cosine	4X4	25.92	49.00
	8X8	36.90	50.05
	16X16	39.28	55.92
	32X32	45.08	68.30
	64X64	51.11	74.21
Euclidean Distance	4X4	29.63	49.34
	8X8	38.01	56.92
	16X16	49.00	79.23
	32X32	58.07	88.92
	64X64	66.01	95.49
Mahalanobis Cosine	4X4	20.73	36.41
	8X8	29.01	44.12
	16X16	37.24	59.13
	32X32	47.03	72.32
	64X64	53.18	84.30

Table 7 Rank rate (%) on CUHK01 dataset using GAN dataset

Matching Distance	GLCM Matrix	Rank 1	Rank 5
Cityblock	4X4	23.73	56.90
	8X8	39.87	62.22
	16X16	47.76	67.44
	32X32	50.20	70.29
	64X64	51.40	82.49
Cosine	4X4	27.82	51.98
	8X8	40.76	56.35
	16X16	45.78	60.95
	32X32	48.24	68.95
	64X64	55.56	75.71
Euclidean Distance	4X4	31.33	52.12

	8X8	39.21	59.12
	16X16	52.18	82.19
	32X32	61.38	89.28
	64X64	69.82	97.92
Mahalanobis Cosine	4X4	22.67	37.21
	8X8	30.31	47.02
	16X16	41.24	61.90
	32X32	50.35	75.55
	64X64	57.88	85.24

4.5 Comparison of Proposed Algorithm with Original dataset and GAN dataset

To authenticate the efficiency of the projected model with the GAN dataset with GLCM matrix size 64X64 using four matching techniques are compared with the original dataset of Market 1501 and CUHK01 as shown in Table 8 for rank 1 matching rate.

Table 8. Top matching rates (%) comparison of the projected framework with original and GAN based datasets.

Datasets	Matching Distance	GLCM Matrix	Market1501	CUHK01
			Rank 1	Rank 1
GAN Dataset	Mahalanobis Cosine	64X64	88.50	57.88
	Cityblock	64X64	89.24	51.40
	Cosine	64X64	94.42	55.56
	Euclidean Distance	64X64	96.50	69.82
Original Dataset (without GAN)	Mahalanobis Cosine	64X64	86.10	53.18
	Cityblock	64X64	85	49.10
	Cosine	64X64	92	51.11
	Euclidean Distance	64X64	95	66.01

The table 8, states that using the GAN dataset results in higher accuracy with less computation as the number of images in the GAN dataset is reduced compared with the original dataset.

4.6 Proposed method comparisons with the current Approaches

To validate the efficiency of the proposed model, the proposed method for 64X64 matrix size with four matching techniques is compared with the existing methods on Market 1501 and CUHK01 dataset as shown in Table 9 for rank 1 matching rate.

Table 9. Top matching rates (%) comparison of the projected framework with the existing methods.

Authors	Rank 1	
	Market 1501	CUHK01
Kunhong Yu et al., [28]	78.33	--
Shuang Liu et al., [29]	93.30	58.20
Proposed model with Mahalanobis Cosine	88.50	57.88
Proposed model with Cityblock	89.24	51.40
Proposed model with Cosine	94.42	55.56
Proposed model with ED	96.50	69.82

The results presented in Table 9, rank 1 matching rate of 96.50% and 69.82% is attained for Market1501 and CUHK01 with proposed method with ED matching distance, which outperforms the previous results stated by Kunhong Yu et al., [28], and Shuang Liu et al., [29]. These findings state the effectiveness of the proposed model over other the existing methods under two different datasets.

5. Conclusion

A person image generation model is proposed by producing an enhanced position-regularization for person re-id to enhance the model's performance. Texture analysis is performed using GLCM, and Radon Transform is applied for data projection in which the images are represented by a very large number of pixel values, hence an DCT is used to decrease the number of features to an added convenient number before the classification process. Extensive testing on two benchmark datasets exhibited that the projected model achieves better performance. The proposed method can be extended for very large database by using artificial intelligence deep learning techniques.

References

[1] Le An, Xiaojing Chen, Shuang Liu, Yinjie Lei and Songfan Yang, "Integrating appearance features and soft biometrics for person re-identification", *Multimedia Tools and Applications*, Vol.76, pp. 12117–12131, 2017.

[2] Muhammad Fayyaz, Mussarat Yasmin, Muhammad Sharif, Jamal Hussain Shah, Mudassar Raza and Tassawar Iqbal, "Multi-level semantic appearance representation for person re-identification system", *Neural Computing and Applications*, Vol.32, pp. 10519–10540, 2020.

[3] Shao-Mei Li, Chao Gao, Jun-Guang Zhu and Chun-Wei Li, "Person Reidentification Using Attribute-Restricted Projection Metric Learning", *IEEE Transactions on Circuits and Systems for Video Technology*, Vol. 28, No. 8, pp. 1765-1776, 2018.

[4] Cairong Zhao, Xuekuan Wang, Wai Keung Wong, Weishi Zheng, Jian Yang and Duoqian Miao, "Multiple metric learning based on bar-shape descriptor for person re-identification", *Pattern Recognition*, Vol. 71, pp. 218-234, 2017.

[5] Qingming Leng, "Co-metric learning for person re-identification", *Advances in Multimedia*, Vol. 2018, pp. 1-9, 2018.

[6] Martin Hirzer, Peter Roth, Martin Kostinger and Horst Bischof, "Relaxed Pairwise Learned Metric for Person Re-identification", *Proceedings of the European Conference on Computer Vision*, pp. 780-793, 2012.

[7] Peter Roth, Martin Hirzer, Martin Kostinger, Csaba Beleznai, Horst Bischof, "Mahalanobis Distance Learning for Person Re-Identification", *Advances in Computer Vision and Pattern Recognition, Springer*, pp. 247-267, 2014.

[8] Cheng-Hao Kuo, Sameh Khamis, Vinay Shet, "Person re-identification using semantic color names and rankboost", *IEEE workshop on applications of computer vision*, pp. 281-287, 2013.

[9] L. Qi, L. Wang, J. Huo, Y. Shi, X. Geng and Y. Gao, "Adversarial Camera Alignment Network for Unsupervised Cross-Camera Person Re-Identification," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 32, no. 5, pp. 2921-2936, May 2022.

[10] Y. Tao, J. Zhang, J. Hong and Y. Zhu, "DREAMT: Diversity Enlarged Mutual Teaching for Unsupervised Domain Adaptive Person Re-Identifcation," *IEEE Transactions on Multimedia*, 2022.

[11] Wenfeng Song, Shuai Li, Tao Chang, Aimin Hao, Qinping Zhao and Hong Qin, "Context-Interactive CNN for Person Re-Identification", *IEEE Transactions on Image Processing*, Vol. 29, pp. 2860-2874, 2020.

[12] De Cheng, Zhihui Li, Yihong Gong and Dingwen Zhang, "Fusion of Multiple Person Re-id Methods with Model and Data-Aware Abilities", *IEEE Transactions on Cybernetics*, Vol. 50, No. 2, pp. 561-571, 2020.

[13] Guodong Ding, Shanshan Zhang, Salman Khan, Zhenmin Tang, Jian Zhang and Fatih Porikli, "Feature Affinity based Pseudo Labeling for Semisupervised Person Re-identification", *IEEE Transactions on Multimedia*, Vol. 21, No. 11, pp. 2891-2902, 2019.

[14] Minxian Li, Xiatian Zhu and Shaogang Gong, "Unsupervised Tracklet Person Re-Identification", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 42, No. 7, pp. 1770-1782, 2020.

[15] Tetsu Matsukaway, Takahiro Okabey, Einoshin Suzuki and Yoichi Sato, "Hierarchical Gaussian Descriptors with Application to Person Re-Identification", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 42, No. 9, pp. 2179-2194, 2020.

[16] Navaneet K L, Ravi Kiran Sarvadevabhatla, Shashank Shekhar, R Venkatesh Babu and Anirban Chakraborty, "Operator-in-the-Loop Deep Sequential Multi-Camera Feature Fusion for Person Re-identification", *IEEE Transactions on Information Forensics and Security*, Vol. 15, pp. 2375-2385, 2020.

[17] Jianjun Lei, Lijie Niu, Huazhu Fu, Bo Peng, Qingming Huang and Chunping Hou, "Person Re-Identification by Semantic Region Representation and Topology Constraint", *IEEE Transactions on Circuits and Systems for Video Technology*, Vol. 29, No. 8, pp. 2453-2466, 2019.

[18] Long Wei, Zhenyong Wei, Zhongming Jin, Zhengxu Yu, Jianqiang Huang, Deng Cai, Xiaofei He, and Xian-Sheng Hua, "SIF: Self-Inspirited Feature Learning for Person Re-Identification", *IEEE Transactions on Image Processing*, Vol. 29, pp. 4942-4951, 2020.

[19] Mang Ye and Pong C. Yuen, "PurifyNet: A Robust Person Re-Identification Model with Noisy Labels", *IEEE Transactions on Information Forensics and Security*, Vol. 15, pp. 2655-2666, 2020.

[20] Hong-Xing Yu, Ancong Wu and Wei-Shi Zheng, "Unsupervised Person Re-identification by Deep Asymmetric Metric Embedding", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 42, No. 4, pp. 956-973, 2020.

[21] Liang Zheng, Liyue Shen, Lu Tian, Shengjin Wang, Jingdong Wang and Qi Tian, "Scalable Person Re-identification: A Benchmark", *IEEE International Conference on Computer Vision*, pp. 1116 -1124, 2015.

[22] Wei Li, Rui Zhao and Xiaogang Wang, "Human Reidentification with Transferred Metric Learning", *in Proceedings of Asian Conference on Computer Vision (ACCV)*, vol 7724, pp. 31–44, 2013.

[23] Xuelin Qian, Yanwei Fu, Tao Xiang, Wenxuan Wang, Jie Qiu, Yang Wu, Yu-Gang Jiang, and Xiangyang Xue, "Pose-normalized image generation for person re-identification", in *Proceedings of the European conference on computer vision (ECCV)*, pp.650-667, 2018.

[24] Cao Z, Simon T, Wei S E, Sheikh Y, "Realtime multi-person 2D pose estimation using part affinity fields", in *proceedings of CVPR*, 2017.

[25] Jiaxing Tan, Yongfeng Gao, Zhengrong Liang, Weiguo Cao, Marc J. Pomeroy, Yumei Huo, Lihong Li, Matthew A. Barish, Almas F. Abbasi, and Perry J. Pickhardt, "3D-GLCM CNN: A 3-Dimensional Gray-Level Co-Occurrence Matrix-Based CNN Model for Polyp Classification via CT Colonography", *IEEE Transactions on Medical Imaging*, vol. 39, no. 6, pp. 2013-2024, 2020.

[26] Minghao Piao, Cheng Hao Jin, Jong Yun Lee, and Jeong-Yong Byun, "Decision Tree Ensemble-Based Wafer Map Failure Pattern Recognition Based on Radon Transform-Based Features", *IEEE Transactions on Semiconductor Manufacturing*, vol. 31, no. 2, pp. 250-257, 2018.

[27] Y. Kong, S. Zhang and S. Liu, "Face recognition based on DCT and multi-scale ∈-LBP," *IEEE International Conference on Computer Science and Automation Engineering (CSAE)*, pp. 275-278, 2012.

[28] Kunhong Yu, Congyan Lang, Songhe Feng and Tao Wang, "Reasonably Assign Label Distributions to GAN Images in Person Re-identification baseline", *IEEE Fourth International Conference on Multimedia Big Data*, pp. 1-5, 2018.

[29] Shuang Liu, Tongzhen Si, Xiaolong Hao and Zhong Zhang, "Semantic Constraint GAN for Person Re-Identification in Camera Sensor Networks", in *IEEE Access*, vol. 7, pp. 176257-176265, 2019.

[30] Verma, D. ., Reddy, A. ., & Thota, D. S. . (2021). Fungal and Bacteria Disease Detection Using Feature Extraction with Classification Based on Deep Learning Architectures. Research Journal of Computer Systems and Engineering, 2(2), 27:32. Retrieved from https://technicaljournals.org/RJCSE/index.php/journal/article/view/29

[31] P, R. H. ., B, S. D. ., M, D. K. ., Sooda, K. ., & B, K. R. . (2023). Transfer Learning based Automated Essay Summarization. International Journal on Recent and Innovation Trends in Computing and Communication, 11(1), 20–25. https://doi.org/10.17762/ijritcc.v11i1.5983