

DWT based Person Re-Identification using GAN

Arun Kumar D. R.¹, Krishna A. N.¹, Anitha A. C.²

¹ Department of Computer Science and Engineering, SJB Institute of Technology, Bengaluru, India

² Department of Computer Science Engineering, Government S.K.S.J Technological Institute,
K R circle, Bengaluru-560001, Karnataka, India

^{1,2} Affiliated to Visvesvaraya Technological University, Belagavi, Karnataka, India

arunkumardr1987@gmail.com, drkrishnaan@gmail.com, anithakrishna2008@gmail.com

Received: June 29, 2021. Revised: January 3, 2022. Accepted: January 21, 2022.

Published: February 18, 2022.

Abstract—The recent development in person re-identification has challenging task for variations in pose, illumination, expression, and also similar appearance between two different persons. In this paper, we propose Discrete Wavelet Transform (DWT) based person re-identification using Generative Adversarial Network (GAN). The CMU multi-PIE face database with multiple viewpoints and illuminations is considered to test the model. The profile side view face images to be tested are converted into frontal face images using Two-pathway generator adversarial network (TP-GAN). The frontal face images are loaded into the server to create server database. The synthesized TP-GAN images and server database images are pre-processed to convert RGB into grayscale images and also to convert into uniform face image dimensions. The person re-identification is based on feature extraction through DWT, which generates one low frequency LL band and three high frequency bands LH, HL and HH. The LL band coefficients are considered as final features, which are noise-free and compressed number of features. The features of profile side view images and server database images are compared using Normalized Euclidean Distance (NED) and threshold values for person re-identification.

Keywords— Discrete Wavelet Transform, Generative Adversarial Network, Normalized Euclidean Distance, Person Re-identification.

I. INTRODUCTION

The present world filled with Closed-Circuit Television (CCTV) cameras and Internet Protocol (IP) cameras at every corner of the streets to capture the activities for surveillance. The human operators are unable to identify the activities that are generated from surveillance systems as the data is massive. The face detection and face recognitions are vision techniques used to identify persons only when the

persons are nearby and facing towards the camera. However, the CCTV footage may not contain always frontal face images and even cover his/her face from the camera. Hence the new vision technique is explored which is slightly different from face detection and face recognition which leads to person Re-identification (Re-id). The captured videos and images of several persons from one camera are matched with videos and images taken from different cameras to identify a particular person is known as person Re-id. This technique is to authenticate queried person appeared in another camera, placed at different location and time. The person Re-id is based on not only facial features from frontal images but also features from the side angled images and entire body like clothing, height etc. The deep-learning-based approaches for person Re-id increases the performance to a very good extent, there still remain several issues. The training and test sets caused by huge discrepancies in body poses, illuminations, backgrounds and so on, leads to the performance deterioration on the test set. In order to overcome the issues posed by regular deep-learning-based methods, GAN and its modifications have been used in [1], [2], and they are also utilized for data-augmentation of the training samples for person Re-id [3].

As Person Re-id requires comparison between two or more persons based on image features. The DWT transformation plays an important task in concluding whether or not a person appears more than once in the data captured by a network of cameras based on DWT features. A DWT toolbox has been developed for signal processing applications [4]. The toolbox allows the user to choose dissimilar kinds of wavelet functions from three different wavelet families viz., Daubechies, Symlet, and Coiflet, and number of decomposition levels. It plays an energetic role in the computerized video surveillance and has been a vigorous research zone for the past few years. The person Re-id is to find out whether different images captured from multiple cameras are belong to the same person or not. The challenges of person re-identification [5] are:

(i) Person's presence over different cameras gets affected from the visual uncertainty and spatiotemporal i.e., the

dissimilarities in appearance of the same person across different cameras.

(ii) The images are of low resolution and videos are of poor quality containing more unconnected.

(iii) The daylight, shade, reflected light from colored surfaces, indoor lighting can cause the same person to appear in different shades and colors across different cameras.

(iv) The partial or even complete obstruction of persons by others in the crowded environments is a challenge task to extract features

(v) The students with uniform clothing in schools and uniform clothing in workplaces is a challenging for Re-id algorithms

Contribution: In this paper, the person re-identification based on TP-GAN and DWT is proposed. The side view face images to be tested are converted into frontal images using TP-GAN and features are extracted using DWT concept. The frontal face images are stored in the server and the features are extracted using DWT concept. The features extracted of test images are compared with the face images in the server database using NED and threshold for person re-identification. The research paper is planned as follows; literature survey in section 2, the proposed novel model is described in section 3. The simulated result analysis is discussed in section 4. Finally, the paper concludes results and future work in section 5.

II. RELATED WORK

In this section the literature survey on the person re-identification, and GAN are explained. Person re-identification is a subset of study of computer vision and is very useful in person tracking in surveillance videos. Re-identification is a challenging task because appearance of person changes across views, independent of the person's identity due to changes in several co-variate factors. Among these factors, pose plays an important role in causing a change in person's appearance. Qian et al. [6] proposed a pose-normalization GAN model (PNGAN) for alleviating the impact of pose variation. However, there are constraints in human poses estimation and weighting on different parts features. Chi Su et al., [7] presented a pose-driven deep convolution model to learn effective feature representations and adaptive similarity measurements. Considering features of entire human body as a global and body parts as a local, are transformed to normalized state for better feature representation.

Yu Yin et al., [8] presented a Dual-Attention Generative Adversarial Network for photo-realistic face frontalization using contextual needs and local uniformity during GAN training. The local features with their long-range dependencies are integrated for yielding better feature representations for higher pose angles. Changle Rong et al., [9] proposed Feature-Improving GAN for face frontalization to enhance the recognition performance for huge face attitudes. An essential recording among the frontal face and profile face, is presumed and their divergence in deep depiction space is projected. The unit has a dense segment termed Feature-Mapping Block that

assistances to plan the features of profile face images to the frontal face. A discriminator differentiates the features of profile face images and ground true frontal face images that help generator module to deliver good profile faces.

Xiao Luan et al., [10] proposed a Geometry Structure Preserving based GAN, for multi-pose face frontalization and recognition. The generator is a typical auto-encoder, that extracts identity and the decoder synthesize the matching frontal face image. Rui Huang et al., [11] presented a Two-Pathway Generative Adversarial Network for photorealistic frontal view synthesis by observing global structures and local details. Shuang Liu et al., [12] proposed a Semantic Constraint Generative Adversarial Network for person Re-id in camera sensor networks. The method produces several stylishness pedestrians' images with high-level semantic info. However, since face frontalization is not incorporated, there still exists some constraints. Meina Kan et al., [13] studied pose-robust features by demonstrating the non-frontal face images to frontal face images using deep network, named as stacked progressive auto-encoders. Shuren Zhou et al., [14] proposed image-to image translation method with Multi-Camera Transfer GAN on pedestrian dataset by means of one single model. Though it does achieve better accuracy, it still does not make the re-identification process highly effective, and it also cannot be relied on for many other databases.

III. PROPOSED MODEL

Person re-identification has gained a lot of interest especially in the recent times, and many challenges associated with Person re-identification still exist which affect the efficiency of the process. The problem is that the person re-id model trained in one dataset generally does not work fine in alternative dataset. With increasing need to better accuracy in Person Re-identification, no matter how much training is performed using several databases of different identities, there is a huge challenge in recognizing an identity especially if the data captured by cameras are side-profile views. Hence face frontalization plays very important role in order to overcome this problem with the help of GAN which is a Deep Learning technique that provides a lot of benefit. The proposed method is depicted in the Figure 1.

A. Face image Dataset

The aim of the research is to generate the frontal face images from an input side view profile image; hence the proposed model is trained on a number of profile images along with their corresponding frontal images. Therefore, for training the model, the benchmark database with face image database that contains a number of profile images and frontal image of each person are considered. The Carnegie Mellon University - Pose, Illumination, and Expression (PIE) dataset i.e., CMU Multi-PIE face database [15] is considered as shown in Figure 2 to test the proposed model with focal length of 4mm-8mm. This database contains 750,000 images from 337 persons recorded in four sessions completed in the duration of five months. The persons are imaged under 15 view points and 19

illuminations to represent facial expressions with high resolution frontal images.

B. Server Database

The database has frontal images of several persons which is used for the Person Re-id based on certain image features. This database contains 843 frontal images of CMU Multi-PIE face dataset of resized images of 256 x 256 and images are in png format. These images in the dataset are used to compare

with the synthesized images generated by the TP-GAN for input profile images of same persons with different angles.

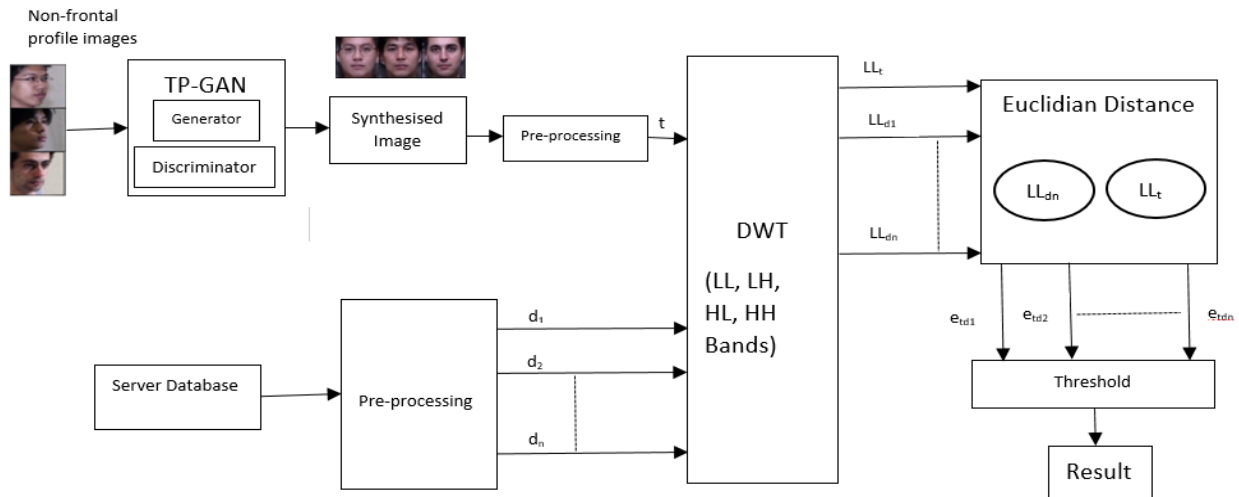


Figure 1. Proposed Method

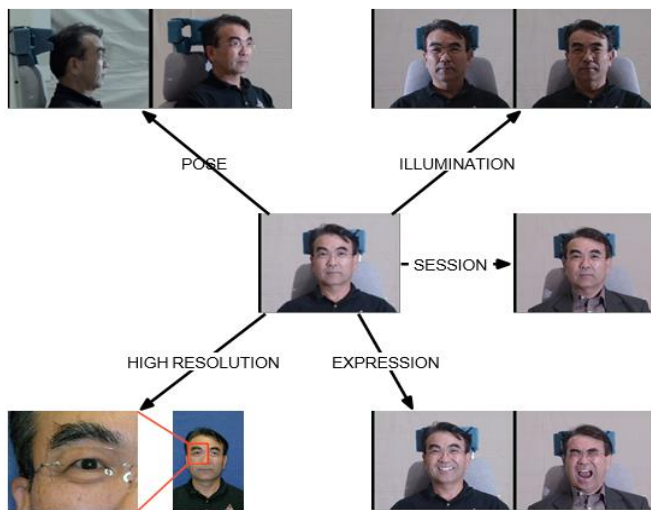


Figure 2. Multi-PIE Face Database [16].

C. Non frontal profile face images captured from CCTV

The profile images with different angles captured are called as test images and let us classify the test images into two sets. The set represents the test face images of a persons that are present in the server database and let us call these test images as Persons-Inside-Server (PIS). These person’s images show positive result for Person re-identification performed between

the test image and the server database. This means that the person in the test image has appeared at-least once previously on another camera/server database. The next set represents the test images of persons that are not present in the server database and let us call these test images as Persons-Outside-Server (POS). These person’s images show negative result for Person-reidentification performed between the test image and the server database. This means the person in the test image has not appeared previously on any camera/server database.

D. Pre-Processing

The colour images of server database and frontal images generated by TP-GAN are converted into grey scale images with uniform size of 256x256. The number of gray scale image features are less and effective compared to colour images which increases speed of computation.

E. Generative Adversarial Network (GAN)

Face frontalization is the process of synthesizing the frontal facing views from a given non frontal profile images with face angles other than frontal view. The frontal face images are synthesized using a GAN [17]. The person Re-ide is built on a lot of changes on GAN have been proposed by many researchers like Cam Style [18], SPGAN [19] and PTGAN [3]. All the aforementioned types are based on CycleGAN [20]. As CycleGAN can only translate images between two domains, but CTGAN (Multi Camera Transfer GAN) [21] based on StarGAN [22], that can convert images from the

source dataset into style of each camera domain in the target dataset.

F. Two Pathway-GAN (TP-GAN)

The input profile images are captured by a number of cameras and are passed into the TP-GAN model to synthesise the frontal image of a person. It is used to frontalize any given profile image or a side-pose image or a non-frontal image [11]. Frontalization is the process of generating a frontal face image of a person given the profile image of the same person. Frontalization has gained a lot of popularity ever since GAN took birth and later on saw tremendous advancement and improvement. Many terrific advancements have been taking place in the field of GAN which is basically used for the generation of a new output, similar yet unique from that of the input dataset. The technology helps in generation of frontal images of a person given the profile image of the same person. TP-GAN synthesises photorealistic frontal view by seeing global structures and local details as shown in Figure 3. The first column is the profile images P1, P2, P3, and P4 captured by CCTV to be tested for re-identification or to be converted into frontal images, and are given as input to TP-GAN. The TP-GAN involves a generator which has two pathways - one with global network dealing the global structure resulting in column 2, and the other, four landmark located patch networks joining to local textures about four facial landmarks resulting in column 3. The column 4 represents the respective synthesised frontal images of the profile images of column 1.

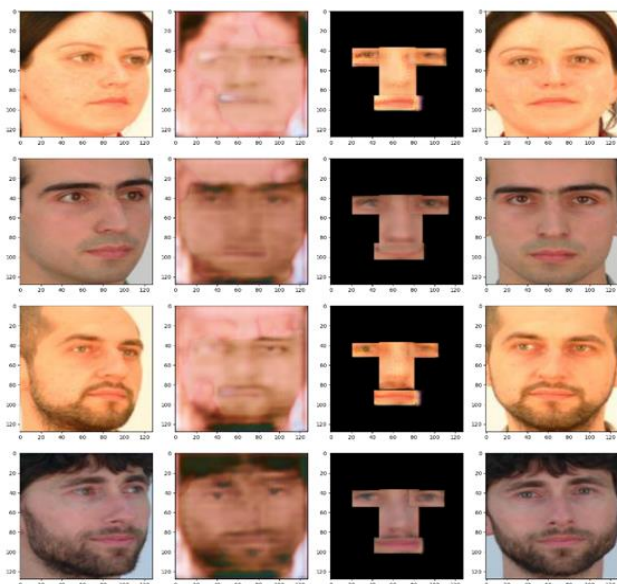


Figure 3. Frontal image generation by TP-GAN

The frontal image samples P1 to P6 stored in the server dataset are as revealed in the figure 4. The profile side images of P1, P2, P3 and P4 are given as input to TP-GAN for conversion.

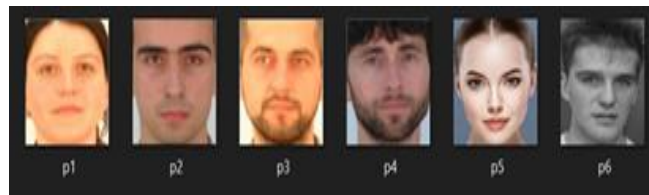


Figure 4. The Frontal Images in the Server Database

G. Person Re-identification using DWT

The process transforms image pixels into wavelets, which are then used for wavelet-based compression and coding [23]. DWT decomposes a given image into a number of sub-bands in the corresponding frequency bands. The process produces 4 sub-bands as output – the low-low band, the low-high band, the high-low band, and the high-high band. In our model, we use the low-low band (the LL band) for the person re-identification purpose, which results in compression of number of features. The synthesized images generated by the TP-GAN are compared with every image in the server database made for person re-identification. The DWT is applied on every resized image, and the LL band of all images is saved and are considered as features.

The DWT with one level decomposition using filters is as shown in Figure 5. The LL band has substantial information of an original image, whereas LH, HL and HH bands represent vertical, horizontal and diagonal insignificant information of an original image. The LL band information is enough to reconstruct original image without other three bands of insignificant information. The final features are considered from LL band coefficients and are converted into 1D-vectors for comparison between images.

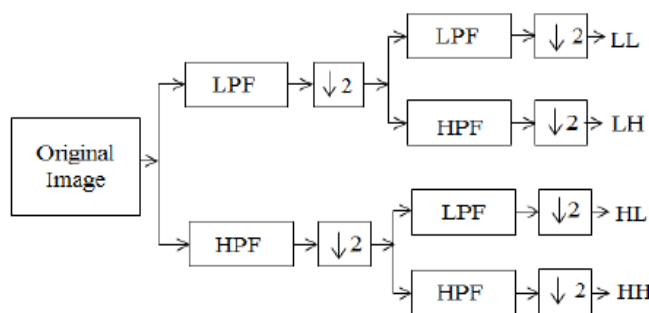


Figure 5. One level 2D-DWT Decomposition.

H. Euclidian Distance (ED)

The Euclidian Distance between the LL band coefficient vectors of the synthesised image and every image of the database, one at a time are computed. The Euclidian Distance for 1D is defined as the distance between two corresponding feature values of two images as given in Equation 1.

$$D(p, q) = \sqrt{\sum_{i=1}^n (q_i - p_i)^2} \tag{1}$$

Where p, q represents the values of two corresponding points in two vectors and n represents total number of final

features.

If two images are exactly same, the ED between the images is zero. For two or more images of the same individual, with the images taken at different view-points, or different lighting, or different facial expressions, or with a slight left-right rotation, the ED between these images is not zero, but minimum. However, the ED between 2 completely different images will be maximum. The purpose of our paper is to achieve person re-identification. This means, we should successfully conclude whether the image of the person in question, appears more than once in different sets of images, with each set captured by different cameras. Hence, the synthesised image, which has the frontal face of a person should be compared with every image in the server database. If there are N number of images in the server database, after performing DWT, there will be N number of LL bands. After finding out the ED between each LL band of the images in the server database, and the LL band of the input frontal image generated by TP-GAN, then there are N number of Euclidian distances.

1. Data Normalization

It is the procedure of altering a set of data points irrespective of their current values, between 0 and 1. This process is also called as Min-Max Normalization. When comparing values of features, then the larger scale values are totally dominate, hence, the normalization of feature values are very important. The goal of normalization is to make every feature value occur in the same scale leads to all feature values are equally important. The most commonly used normalization technique is Min-max to normalize. In normalization the least value is mapped into 0, the highest value is mapped into a 1, and other values mapped between 0 and 1. Every ED value is divided by the largest ED, including the final largest value itself. Therefore, the final largest value, since it gets divided by itself, becomes 1. Hence, the final values of Normalized Euclidian Distances (NED) range between 0 and 1.

Threshold: Now that every normalized Euclidian distance is reduced to values ranging from 0 to 1, depending on the obtained values, it becomes important to conclude that the frontal images of a person in the server database has already appeared in the data captured by a network of cameras, or not. The threshold values, which helps make the decision to identify persons. Smaller the value of ED, higher is the chance that the input image is similar to the corresponding server-database image, and thus, can be concluded that the person to whom the input image belonged to, appeared at least once in a camera. The same way, higher the value of ED, lesser is the chance that the input image is different when compared to the corresponding server-database image, and thus, can be concluded that the person to whom the input image belonged to, did not appear at least once in a camera. However, a threshold value should be fixed such a way, that any normalized ED lesser than this threshold value means that the synthesised image by the TP-GAN has already appeared at least once in a camera. The NED more than the threshold

value means the synthesised image by the TP-GAN has never appeared in any camera.

IV. SIMULATED RESULT ANALYSIS

Python language is used to implement the research with packages of Tensorflow, OpenCV and Matplotlib. Sixteen person’s frontal images are considered in the server database to test the proposed model and the DWT is applied to convert into four low (LL) and higher frequency bands (LH, HL, HH). The LL band images are considered further for Person Re-ide as shown in Figure 6.

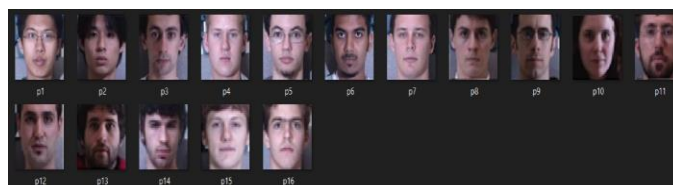


Figure 6. Server Database

A. Person Inside Server Database (PISD) test image analysis

The TP-GAN generated synthesised frontal images are considered and DWT is applied to produce one low and three high frequency bands. The only low frequency LL band is further considered as an image to be tested to perform Person Re-identification. The LL band of synthesised frontal image produced by TP-GAN is considered, which is similar to LL band image in the server database. The LL band image P2, which is generated by TP-GAN corresponding to LL band image of P2 in the server database is known as PISD is as shown in Figure 7.



Figure 7. LL band of PISD Test Image of P2

The Person Re-id performed based on the ED between each server image’s LL band and the PISD test image’s LL band, and normalise every resulting ED. The NED values between server database images and test images are given in Table 1. It is noticed that the values of NED are high in the case of different images compared to similar image. It is clearly indicating that the least NED value is 0.297865554 which corresponds to the server image P2 in the 2nd row in table 1.

Table 1. NED values between every server image and the PISD test image of P2

Sl No.	Server Database	Test Image	NED
1	p1	P2	0.575
2	p2	P2	0.298
3	p3	P2	0.326

4	p4	P2	0.405
5	p5	P2	0.388
6	p6	P2	0.570
7	p7	P2	0.793
8	p8	P2	0.814
9	p9	P2	0.405
10	p10	P2	0.340
11	p11	P2	1.000
12	p12	P2	0.801
13	p13	P2	0.411
14	p14	P2	0.811
15	p15	P2	0.333
16	p16	P2	0.398

However, successful person Re-identification does not only depend on the least NED value. The least value should be less than or equal to the threshold value. The threshold fixed after a lot of trial and error; we have fixed 0.3 as the threshold value. A NED value less than or equal to 0.3 means the PISD test image is similar to the corresponding server image. The similar PISD test image and corresponding server database images are shown in Figure 8 with NED of 0.298.



(a) PISD Test image (b) Matching Server Image –P2
 Figure 8. Similar PISD test and server images

B. Person Outside Server Database (POSD) test image analysis

The LL band of synthesised frontal image generated by TP-GAN is considered, which is not similar to LL band images in the server database. The LL band image P, which is generated by TP-GAN not corresponding to LL band images in the server database is known as POSD test image which means the image does not appear in another camera and hence, not present in the server database as well is as shown in Figure 9.



Figure 9. LL band of POSD Test Image P

The Person Re-identification performed based on the ED between each server image’s LL band and the POSD test image’s LL band, and normalise every resulting ED. The NED values between server database images and test images are given in Table 2. It is noticed that the values of NED are higher in all the cases of images because of POSD. As

discussed earlier, successful person re-identification does not only depend on the least NED value.

Table 2. NED values between server images and the POSD test image

Sl No.	Server Database Images	POSD Test Image	NED
1	p1	P	0.381
2	p2	P	0.961
3	p3	P	0.744
4	p4	P	0.707
5	p5	P	1.000
6	p6	P	0.635
7	p7	P	0.319
8	p8	P	0.345
9	p9	P	0.834
10	p10	P	0.653
11	p11	P	0.321
12	p12	P	0.407
13	p13	P	0.618
14	p14	P	0.353
15	p15	P	0.588
16	p16	P	0.760

The least value should be less than or equal to the threshold value and if the threshold value is fixed at 0.3, then the NED values more than 0.3 means the POSD test image does not exist in the server database.

As far as the matching of the POS test image with a server image is considered, there is obviously one least value out of every set of NED values, and theoretically, the server image corresponding to the least value should be the matched image with the test image without considering threshold values. However, practically, the matching is wrong as the two images (the server image and the test image) are different. This is also justified by the obtained least NED value which exceeds the threshold value. Figure 10 shows that POSD test image P is wrongly matched with the server image P7 by considering only least NED without comparing with threshold value. In general, any POSD test image is expected to not match with the server database.



(a) POSD Test Image (b)Wrongly matched server image - p7
 Figure 10. Dissimilar POSD test and server images

C. Definitions of performance parameters:

(i) *The PISD - True Recognition Rate (PISD-TRR)*: The test images corresponding to the images in the server database are re-identified correctly. It is defined as the ratio of number of PISD images re-identified correctly to the total number of PISD images and is given in Equation 2.

$$\% PISD - TRR = \frac{\text{No.of PISD images that are re-identified correctly}}{\text{Total number of PISD images}} \times 100 \quad (2)$$

(ii) *PISD- False Recognition Rate (PISD-FRR)*: The test images corresponding to the images in the server database are falsely rejected saying that the test image is not present in the database falsely. It is defined as the ratio of number of PISD images falsely rejected to the total number of PISD images and is given in Equation 3.

$$\% PISD - FRR = \frac{\text{No.of PISD images falsely rejected}}{\text{Total number of PISD images}} \times 100 \quad (3)$$

(iii) *POSD- Correct Rejection Rate (POSD-CRR)*: The test images not corresponding to the images in the server database are rejected correctly saying that the test image is not present in the database correctly. It is defined as the ratio of number of POSD images that are rejected correctly to the total number of POSD images and is given in Equation 4.

$$\% POSD - CRR = \frac{\text{No.of POSD images that are rejected correctly}}{\text{Total number of POSD images}} \times 100 \quad (4)$$

Note that for POSD images, the images are expected to not be re-identified, or in simple words, not be matched, with any server images.

(iv) *The POSD- False Accept Rate (POSD-FAR)*:

The test images not corresponding to the images in the server database are falsely accepted saying that the test images are existing in the database falsely. It is defined as the ratio of number of POSD images falsely accepted to the total number of POSD images and is given in Equation 5.

$$\% POSD - FAR = \frac{\text{No.of POSD images that are falsely accepted}}{\text{Total number of POSD images}} \times 100 \quad (5)$$

D. Performance Analysis

The performance of Person-Re-identification using TP-GAN and DWT is evaluated by comparing PISD, POSD and images in the server database. The computed NED values of PISD and POSD images are related with the NED values of images in the server database by varying threshold values. In order to do this, sixteen number of server database images, sixteen number of PISD images, and thirty-five number of POSD images are considered as shown in Figures 11, 12, and 13.

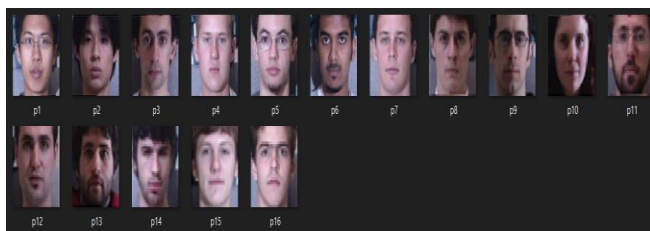


Figure 11. Server Database



Figure 12. PISD Database



Figure 13. POSD Database

The LL bands of all the images in the server database, the PISD database, and the POSD databases are extracted. The NED between every image in the PISD database and all the images in server database is calculated. Since there are sixteen images in the PISD database, every PISD image when compared with the server database images gives sixteen NED values for every PISD image. Thus, we obtain sixteen different sets of NED values, one set for each PISD image. The only the least NED value from each of the sixteen sets obtained are considered. A list of least NED values obtained for every PISD database image when compared each of them with the server database images, is given in Table 3.

Table 3. The Least NEDs when every PISD image is compared with the server images

PISD Images	Least NED
p1	0.219
p2	0.099
p3	0.056
p4	0.003
p5	0.0025
p6	0.005
p7	0.0022
p8	0.171
p9	0.113
p10	0.293
p11	0.126
p12	0.101
p13	0.199

p14	0.173
p15	0.191
p16	0.071

Similarly, once the LL bands of the images in the POSD are extracted, the NED values between every image in the POSD and all the images in server database are calculated. Since there are thirty-five images in the POSD, every POSD image, when compared with the server database gives sixteen NED values, each. Thus, the sixteen different sets of NED values are obtained, one set for each POSD image. The only least NED value from each of the sixteen sets obtained are considered. A list of least NED values obtained for every POSD image when compared each of them with the server database images, is given in Table 4.

The computed least NED values, are compared with the variable threshold values and conclusion is made on whether the test image either from PISD or POSD, is re-identified or not. The effect of threshold values on the performance of proposed method is given in Table 5. It is noticed that as threshold values increases from zero to one the PISD-TRR and POSD-FAR increases from zero to 100%, PISD-FRR and POSD-CRR decreases from 100% to zero.

Table 4. Least NEDs when every POSD image is compared with the server images

POSD Images	Least NED
p1	0.319
p2	0.399
p3	0.479
p4	0.515
p5	0.297
p6	0.319
p7	0.416
p8	0.332
p9	0.296
p10	0.418
p11	0.482
p12	0.406
p13	0.368
p14	0.441
p15	0.256
p16	0.344
p17	0.453
p18	0.538
p19	0.324
p20	0.598
p21	0.320
p22	0.515
p23	0.338
p24	0.379
p25	0.322
p26	0.445

p27	0.382
p28	0.410
p29	0.297
p30	0.442
p31	0.425
p32	0.375
p33	0.298
p34	0.565
p35	0.454

The variations of different recognition rates with threshold values are shown in Figure 14. It is seen that the value of PISD-TRR is zero at zero threshold value since there is a very small difference in features between test images (PISD) and an image in the server database. As the value of threshold increases, PISD-TRR increases to 100%. The value of PISD-FRR is 100% (falsely rejecting) at threshold value zero, since the test image similar to the server database image is rejected falsely as some NED values due to the difference in two images. The POSD-FAR increases as threshold values increases and attains 100% which indicate that the test images not in the server database are also accepted falsely and re-identified correctly which is a wrong process.

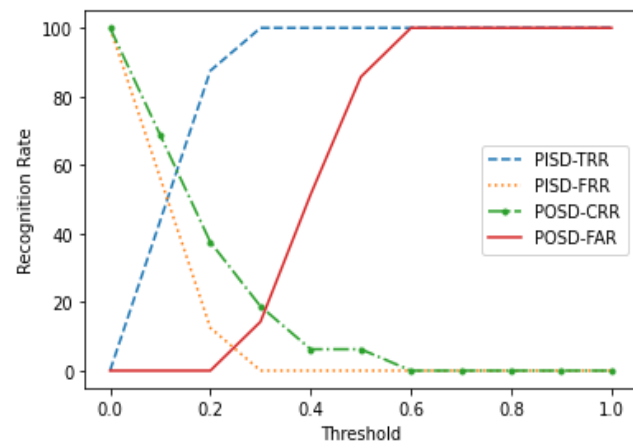


Figure 14. Variations or recognition rates with threshold

E. Comparison with existing methods using ORL face dataset:

A standard Olivetti Research Laboratory (ORL) face dataset images captured between 1992 and 1994 are considered to compare percentage recognition accuracy of proposed technique with the current approaches and the performance outcomes are tabulated in Table 6. The ten different facial expressions like open/closed eyes, smiling/not smiling, with/without glasses, varying lightening conditions of single people are captured. Similarly, face images of forty persons were considered under dissimilar conditions.

Table 5. The performance parameter variations with threshold

Threshold Values	PISD-TRR	PISD-FRR	POSD-CRR	POSD-FAR
0	0	100	100	0
0.1	43.75	56.25	68.75	0
0.2	87.5	12.5	37.5	0
0.3	100	0	18.75	14.28
0.4	100	0	6.25	51.42
0.5	100	0	6.25	85.71
0.6	100	0	0	100
0.7	100	0	0	100
0.8	100	0	0	100
0.9	100	0	0	100
1	100	0	0	100



Figure 15. Ten Face image samples of a single person on the ORL database [24].

An upright frontal and slight tilt of the head positions were present in the dataset. The total number of images in the database are 400 of 40 persons in PGM format with each image size of 92x112. Ten image samples of one person are shown in Figure 15. It is seen that the maximum percentage recognition accuracy of proposed method is better and is 100% compared with the current methods. Hence the method proposed by us is superior as the maximum percentage recognition accuracy is very high for higher threshold values however the limitation of higher threshold value is POSD-FAR is high.

Table 6. Proposed method comparison with existing methods

Sl No.	Authors	Technique	% Maximum Recognition Accuracy
1	Ying Wen [25] (2017)	A discriminative common vector dictionary + sparse representation-based classification	96.54
2	Abuzneid and Mahmood [26] (2018)	PCA+BPNN	96.90
3	Jun Fan et al., [27] (2017)	Locality Preserving Projections + Maximum Margin Criterion for solving general Eigen value	97.10
4	Jun Kong et al., [28] (2018)	circular symmetrical Gabor filter and PCA neural networks	97.50
5	Rangsee et al., [29] (2019)	Bit Endianness + DWT +HOG	98
6	Proposed method	TP-GAN and DWT with NED	99.10

V. CONCLUSION

In this paper, person re-identification based on TP-GAN and DWT is proposed. The side view face images which are used as images to be tested, are given as input to TP-GAN, which converts into synthesized frontal images. The frontal face images are stored in the server database as reference images. The features of frontal face images are extracted by DWT and considered LL band co-efficient as final features. The NED is used to compare features for person re-identification effectively with a smaller number of features. The recognition accuracy is better in our method compare to existing state of art methods. In future, the limitation of our method can be addressed by combining results of converting side angled images to frontal images using GAN and traditional methods of converting side angled images to frontal images.

References

- [1] I. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, and Y. Bengio, "Generative Adversarial Nets," *ACM International Conference on Neural Information Processing Systems*, vol 2, pp 2672–2680, December 2014.
- [2] J Zhu, T Park, P Isola and A AEFros, "Unpaired Image-to-Image Translation using Cycle-Consistent Adversarial Networks," *IEEE International Conference on Computer Vision (ICCV)*, pp. 2242-2251, 2017.
- [3] L Wei, S Zhang, W Gao and Q Tian, "Person Transfer GAN to Bridge Domain Gap for Person Re-identification," *IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 79-88, 2018.

- [4] M. K. Wali, M. Murugappan, R. B. Ahmad and Bong Siao Zheng, "Development of Discrete Wavelet Transform (DWT) Toolbox for Signal Processing Applications," *International Conference on Biomedical Engineering*, pp. 211-216, 2012.
- [5] Gala, Apurva and Shishir K. Shah. "A Survey of Approaches and Trends in Person Re-Identification." *Elsevier Journal of Image and Vision Computing*, vol 32, pp 270–286, 2014.
- [6] Qian, X., Fu, Y., Xiang, T. and Wang, W., "Pose-normalized image generation for person re-identification", In: Proceedings of the European Conference on Computer Vision (ECCV). pp. 650–667, 2018.
- [7] C. Su, J. Li, S. Zhang, J. Xing, W. Gao, and Q. Tian, "Pose-Driven Deep Convolution Model for Person Re-identification," Proceedings of the IEEE International Conference on Computer Vision (ICCV), pp. 3980-3989, 2017.
- [8] Y Yin, S Jiang, J Robinson and Y Fu, "Dual-Attention GAN for Large-Pose Face Frontalization", *IEEE International Conference on Automatic Face and Gesture Recognition*, pp. 249-256, 2020
- [9] C. Rong, X. Zhang and Y. Lin, "Feature-Improving Generative Adversarial Network for Face Frontalization," *IEEE Access*, vol. 8, pp. 68842-68851, 2020.
- [10] X. Luan, H. Geng, L. Liu, W. Li, Y. Zhao and M. Ren, "Geometry Structure Preserving Based GAN for Multi-Pose Face Frontalization and Recognition," *IEEE Access*, vol. 8, pp. 104676-104687, 2020
- [11] R. Huang, S. Zhang, T. Li and R. He, "Beyond Face Rotation: Global and Local Perception GAN for Photorealistic and Identity Preserving Frontal View Synthesis," *IEEE International Conference on Computer Vision (ICCV)*, pp. 2458-2467, 2017.
- [12] S Liu, T Si, X Hao and Z Zhang, "Semantic Constraint GAN for Person Re-Identification in Camera Sensor Networks," *IEEE Access*, vol. 7, pp. 176257-176265, 2019
- [13] M Kan, S Shan, H Chang and X Chen, "Stacked Progressive Auto-Encoders (SPA-E) for Face Recognition Across Poses," *IEEE Conference on Computer Vision and Pattern Recognition*, pp. 1883-1890, 2014.
- [14] S. Zhou, M. Ke, and P. Luo, "Multi-Camera Transfer GAN for Person Re-Identification," *Elsevier Journal of Visual Communication and Image Representation*, vol. 59, pp. 393-400, August 2019.
- [15] T. Sim, S. Baker and M. Bsat, "The CMU Pose, Illumination, and Expression Database," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 25, no. 12, pp. 1615-1618, December 2003.
- [16] Ralph Gross, Iain Matthews, Jeffrey Cohn, Takeo Kanade, and Simon Baker, "Multi-PIE," *Elsevier Journal of Image and Vision Computing*, vol 28, issue 5, pp 807-813, May 2010,
- [17] Zhihong Zhang, Xu Chen, Beizhan Wang, Guosheng Hu, Wangmeng Zuo, Edwin R. Hancock, "Face Frontalization using an Appearance-Flow-based Convolutional Neural Network", *IEEE Transactions on Image Processing*, vol 28, no 5, pp 2187-2199, May 2019.
- [18] Z Zhong, L Zheng, Z Zheng, S Li, and Y Yang, "Camera style adaptation for person Re-identification," *IEEE Conference on Computer Vision and Pattern Recognition*, pp. 5157–5166, 2018.
- [19] W Deng, L Zheng, Q Ye, G Kang, Y Yang, and J Jiao, "Image-Image Domain Adaptation with Preserved Self-Similarity and Domain-Dissimilarity for Person Re-Identification," *IEEE Conference on Computer Vision and Pattern Recognition*, pp. 994–1003, 2018.
- [20] J Y Zhu, T Park, P Isola, and A Afros, "Unpaired Image-to-Image Translation using Cycle-Consistent Adversarial Networks," *IEEE International Conference on Computer Vision*, IEEE Computer Society, pp. 2242–2251, 2017.
- [21] Shuren Zhou, Maolin Ke, and Peng Luo. (2019), "Multi-Camera Transfer GAN for Person Re-Identification", *Elsevier Journal of Visual Communication and Image Representation*, vol 59, pp 393-400, 2019.
- [22] Y. Choi, M. Choi, M. Kim, J. Ha, S. Kim, J. Choo, "Star GAN: Unified Generative Adversarial Networks for Multi-Domain Image-to-Image Translation," *IEEE Conference on Computer Vision and Pattern Recognition*, pp. 8789–8797, 2018.
- [23] H. Demirel and G. Anbar Jafari, "Image Resolution Enhancement by using Discrete and Stationary Wavelet Decomposition," *IEEE Transactions on Image Processing*, vol. 20, no. 5, pp. 1458-1460, May 2011.
- [24] The ORL database, <http://www.cam-orl.co.uk>
- [25] Ying Wen, "A Novel Dictionary based SRC for Face Recognition", *IEEE International Conference on Acoustics, Speech and Signal Processing*, pp. 2582-2586, 2017.
- [26] Mohannad Abuzneid and Ausif Mahmood, "Face Recognition Framework based on Correlated Images and Back-Propagation Neural Network", *IEEE/ACS 15th International Conference on Computer Systems and Applications (AICCSA)*, pp. 1-7, 2018.
- [27] Jun Fan, Qiaolin Ye and Ning Ye, "Enhanced Adaptive Locality Preserving Projections for Face Recognition", *IAPR Asian Conference on Pattern Recognition (ACPR)*, pp. 594-598, 2017.
- [28] Jun Kong, Min Chen, Min Jiang, Jinhua Sun and Jian Hou, "Face Recognition Based on CSGF(2D)2PCANet", *IEEE Access*, Vol.6, pp. 45153-45165, 2018.
- [29] P. Rangsee, K. B. Raja and K. R. Venugopal, "Nibble-Based Face Recognition Using Convolution of Hybrid Features," *IEEE International Conference on Imaging, Signal Processing and Communication (ICISPC)*, pp. 112-116, 2019.

Creative Commons Attribution License 4.0 (Attribution 4.0 International, CC BY 4.0)

This article is published under the terms of the Creative Commons Attribution License 4.0

https://creativecommons.org/licenses/by/4.0/deed.en_US