AN EFFICIENT FEATURE REPRESENTATION USING CHLOMO AND CNN FOR PERSON RE-IDENTIFICATION

Abstract

The re-identification of a person is defined as determining whether a person has previously been in front of a network of cameras. Efficient features always make perfect learning towards individual re-identification. In this research, a new deep feature learning model called CHLFDLF has been proposed with the mixture of handcrafted features and profound features, extracted using CHannel Integrated LOMO (CHLOMO) and Convolutional Neural Network (CNN) respectively to attain an improved person re-identification system. The extracted features are integrated efficiently and learned using the Cross-view Quadratic Discriminant Analysis (XQDA) similarity metric approach for efficient person re- identification. The performance of this proposed approach is demonstrated by applying it to the different benchmark datasets for re-identification such as VIPeR, CUHK03, and Market1501

Keywords: Handcrafted feature; Convolutional Neural Network; Deep feature; Person Re-identification; Metric Learning.

Authors

Dr. J Stella Janci Rani
Assistant Professor of Computer Science
Sarah Tucker College
Tirunelveli, Tamil Nadu, India
jstellajara17@gmail.com

Dr. M Gethsiyal Augasta
Assistant Professor of Computer Science
Kamaraj College
Thoothukudi, Tamil Nadu, India
augastaglady@gmail.com

Dr. S Princy Suganthi Bai
Associate Professor of Computer Applications,
Hindustan Institute of Technology and Science (Deemed to be University)
Chennai, Tamil Nadu, India
vinodhprincy2003@gmail.com

I. INTRODUCTION

An aspect of image retrieval is person re-identification. This technique is very much significant for community security purposes. While observing big distances, this is important to appropriately equivalent the identical individual in various photographic camera perspectives. An undirected trajectory can be created by corresponding to the destination issue in several camera angles [1]. Deep learning development and the increase in the need for intelligent video surveillance have attracted much more focus in the field of computer vision. Re-identification of a person is a difficult assignment due to the existence of various problems, namely various viewpoints, unconstrained poses, varying low-image resolutions, occlusions, complex camera environments, illumination changes, unreliable bounding box generations, heterogeneous modalities, and background mess [2].

Disjointed and non-overlapping cameras shatter the majority of the individual images. These cameras are fixed in place in an unregulated setting, producing extremely poor-quality images. Extracting a person's representations from shoddy photos is a difficult task. Usually, distance metric learning and feature representation are the two fundamental components of person re-identification methods. Representation of feature and feature extraction techniques are considered very important steps in metric learning. The extracted features in this case are often divided into learned and handcrafted features [3]. The illustration of the feature is necessary since it serves as the basis for metric learning. The efficiency of metric learning is fine recognized with the quality of the feature obtained. Here, the feature extraction process begins with a collection of measured information and then creates a series of derivative values that are intended to be informative and non-redundant.

The majority of feature extraction techniques for person re-identification, it has only employed grayscale textures. The handcrafted feature is improved using the CHannel Integrated LOcal Maximal Occurrence (CHLOMO) [4] approach, which combines the Retinex transform and the Scale Invariant Channel Integrated Statistical Pattern (SICISP) method. The CHLOMO feature is a combined representation of blue, red, and green textures that delivers efficient outcomes for the re-identification of a person. To account for variations in illumination, the CHLOMO approach employs Retinex transforms and the HSV color histogram. The majority of handcrafted features, however, are constrained by their inability to be immediately applied to real-world issues. There have been attempts to automatically learn features for person re-identification in addition to handcrafted features.

In order to train robust representations from an ensemble of localized features, Gray et al. [8] suggested the use of AdaBoost. It has also been reported that deep learning is "state-of-the-art" for a range of tasks, including picture annotation, face recognition, and speech recognition. Given CNN's popularity, CNN-based deep feature learning has recently attracted attention. To improve person re-identification performance, a "Siamese" [21] deep neural network. The suggested model advocated learning the color feature, texture feature, and metric in a single framework. To improve person re-identification performance, a "Siamese" deep neural network is developed, with the proposed model concurrently learning the color feature, texture feature, and metric in a single framework.

Another crucial step in the process of person re-identification is learning a reliable distance or similarity function to handle the challenging matching task. A discriminating metric with the detected features should be learned to match the numerous individual photographs. Features should have been energetically shown through changes in lighting and angle. A variety of similarity measures, including Cross-view Quadratic Discriminant Analysis (XQDA) have been used by researchers for person re-identification.

In this research work, the person re-identification method called CHannel Integrated LOMO features (CHLF) and Deep Learned Features (DLF) based person re-identification model with XQDA (CHLFDLF) is proposed for improving the person re- identification.

The paper is organized as follows: The related studies are presented in Section 2 and the proposed method is discussed in Section 3, and Section 4 offers the proposed algorithm for person re-identification, experimental outcomes and analysis are included in Section 5.

II. LITERATURE REVIEW

There are numerous extant works have focused on increasing strong and refined features in person re-identification. These works also explain features with highly changeable visual appearances produced under considerably various conditions. Wu et al., [1] have applied recent indices for the purpose of person re-identification. Hence, a fair evaluation system is conducted with metric learning methods. The relationships between loss function with deep feature space and metric learning are also considered. There is an analysis of a variety of metric learning features, from handcrafted features to deep features. Moreover, the experimental results show that the elucidation's space is distinct from the space in which the features were initially selected. Ye et al.,[2] have conducted a widespread overview with in detail scrutiny for closed world person re-identification from three various viewpoints, namely deep feature representation learning, deep metric learning, and ranking optimization. Some under-investigated open issues are also analyzed

Liao et al., [5] have designed Local Maximal Occurrence [LOMO] to denote every pedestrian image as a high-dimensional feature. Then, the metric learning technique which is known as Cross-view Quadratic Discriminant Analysis (XQDA) is also used. LOMO feature extracts the horizontal occurrence of local features and maximizes the occurrence to make as Table analysis against viewpoint changes. Retinex transform and a scale-invariant texture operator are also used to overcome illumination problems.

CHLOMO [4] technique is a new Channel Integrated Local Maximal Occurrence (CHLOMO) for attaining sturdiness on variations in lighting and noise in the images by introducing a novel method named Scale-Invariant Channel Integrated Statistical Pattern (SICISP) with the bettered Scale Invariant Local Ternary pattern (SILTP). Normally SILTP is built using a gray channel only. SICISP, a well-known tool for its invariant texture description of lighting, is incorporated into this model from two distinct kinds of patterns for each color conduit. The Channel Integrated Local Maximum Occurrence features are created using the process of min-max fusion by combining HSV color representation with the statistical texture feature representation. This CHLOMO method remains a feature-based person re-identification model with XQDA.

Tao et al., [3] have applied the XQDA metric learning technique. Here, Deep Multi-View Feature Learning (DMVFL) is proposed to provide the collaboration between handcrafted and deep learning features in an easy way. They have used two challenging person re-identification data sets, namely VIPeR and GRID to show that the XQDA is a sturdy algorithm. According to the results obtained, DMVFL shows improvement in current state-of-the-art methods.

The information on both local texture and global color representations has been combined with a raw source image by Jayapriya et al., [6]. By estimating the highest possible chrominance value in the form of HSV, Scale Invariant Local Ternary Pattern (SILTP) for each pixel of the texture, and the source image is used to create the Prioritized Chromatic Texture (PCTimg) technique. In this method, the combined data is used to extract the features using a convolutional neural network (CNN). The Prioritized Chromatic Texture Image (PCTimg) and the initial source image are combined before being delivered to CNN. To re-identify a person, the XQDA similarity metric algorithm is used. The Multiscale Retinex algorithm is also used for pre-processing the images. Shaojun et al.,[7] have combined local features and global features of images through deep learning networks. This method is called the Multi-level Feature Fusion model, which generates more pedestrian descriptors. Specific features are extracted from different network depths using the Part-based Multi-level Net. Additionally, low-to-high-level local characteristics of pedestrian photos are fused using this method. The highest level is extracted using Global- Local Branches, which extract both local and global features.

Moreover, various person re-identification methods DMVFL [3], SLFDLF [13], and PCTimg [6] have been proposed in recent days with the combination of machine learning and deep learning. These methods have utilized both handcrafted and deep-learned features and used XQDA for metric learning. SLFDLF [13] method is the combination of the Splitted LOMO feature and deep learned feature. The local feature has a more discriminative nature to describe the human structural information, hence it is first horizontally divided into size-based grids, and the LOMO features are individually retrieved from each grid. The extracted LOMO features are combined with deep features to represent that image and XQDA metric learning model is used to learn the combination of Splitted LOMO features and deep learned features (SLFDLF).

As LOMO can handle only gray texture images, this research has focused on creating an efficient feature representation model for handling both HSV color and color texture person re-identification images with an effective technique termed CHLFDLF. The proposed CHLFDLF is the combination of the CHannel Integrated LOMO feature (CHLF) and the deeply learned feature (DLF). In other words, CHLFDLF is an integrated channel having deep features and hand-crafted features with red, green, and blue texture features. The XQDA metric learning is applied on the CHLF and the DLF to obtain the distances on various experimental datasets as it is a reliable metric learning algorithm to learn the distance with the ability to achieve outstanding performance on person re-identification challenges.

III. THE RESEARCH METHOD (CHLFDLF)

This section provides an explanation of the CHLFDLF technique, which was developed by combining deep-learned features with enhanced handcrafted features. The suggested CHLFDLF method involves three stages. Using CHLOMO, effective handcrafted features are first retrieved from images of pedestrians. In the subsequent stage, the deep-learned features are extracted from the images using the convolutional neural network (CNN). Finally, utilizing XQDA, the deep-learned features and handcrafted features are effectively integrated to learn the distance:

Stage I: CHLOMO Feature (CHLF)

Local Maximal Occurrence (LOMO) is a well-organized feature representation. The handcrafted features of LOMO are used in person re-identification. Normally, features of LOMO are combined with SILTP features, Retinex transforms, and the HSV color histogram to account for variations in light. Moreover, the LOMO [5] descriptor examines and leverages the flat incidence of constrained characteristics to produce a consistent representation that is robust to changing perspectives. SILTP was proposed for gray channels only. This LOMO feature is robust to changes in illumination and viewpoint. The Splitted Lomo Feature (SLF) [13] method is in the form of a grid-based. Using this technique, handcrafted features are not extracted for the entire image, but from the split images. i.e., all the images are splitted into grid form.

The CHLOMO [4] is an efficient feature representation model with the introduction of SICISP. SICISP is the new operator that is introduced in place of the SILTP. For color texture images, the existing SILTP has been expanded as the Scale-Invariant Channel Integrated Statistical Pattern (SICISP) operator. A SICISP, an HSV color histogram, and Retinex transforms are used in the CHLOMO approach to cope with differences in lighting. In this method, CHLOMO features are extracted for the whole image.

Stage II: Deep Learned Feature (DLF)

The architecture of the proposed technique with CNN for VIPeR, Market1501, and CUHK03 has been depicted. Figure 1 shows the proposed framework of DLF. This model consists of 12 layers. A Dropout layer, Softmax layer, two max- pooling layers, two fully connected layers, two rectified linear unit (ReLU) activation, two convolutional layers, classification layers, and image input layers are included. The pooling layer entails a dropout for alleviating the qualified model. The first convolutional layer utilized a stride parameter size is two and 5×5 kernel size, and the fourth convolutional layer used a stride parameter is one and 5×5 kernel size. For the pooling process, the max-pooling layer is used. The first pooling layer's pool size is 5x5, whereas the second pooling layer's pool size is 3x3. Finally, the second fully connected layer is used to extract the deep features. This method has been used to employ a dropout to steady the learned model and a data augmentation training technique to extend the dataset because most extant re-identification of person databases are small-scale such as the VIPeR dataset. Triplet loss can be used with Convolutional Neural Networks and seen as an entire system. In order to optimize the distances, triplet loss training aims to create an effective feature depiction.

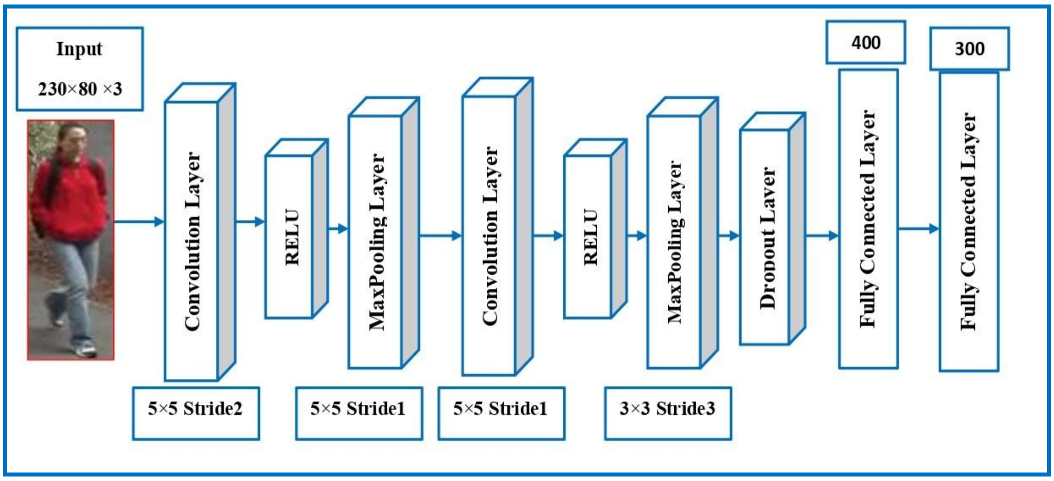

Figure 1: DLF Network Architecture Model

Stage III: Fusion Strategy

It is feasible to combine the encoded complementary information of the two features to overcome their respective shortcomings because the suggested deep feature learning model is directly relevant to real-world issues and the hand-crafted LOMO features are independent of the sample size. XQDA's approach to metric learning is effective and efficient. This approach is utilized to train a low-dimensional subspace with a distance. As a result, only learning the similitude measure will suffice. Figure 2 depicts the technique's fusion strategy. For the training and testing data sets of the suggested method, XQDA is employed to determine the distance C1 amongst handcrafted CHLOMO features and the distance C2 amongst deep-learned features. These two lengths are added to determine the ultimate distance using equation (1). In particular, XQDA is employed for each feature representation to learn the distances denoted by C1 and C2, respectively.

$$C = C1 + \alpha * C2 \tag{1}$$

where α is the trade-off parameter. The matching rank for person re-identification is calculated using the final distance.

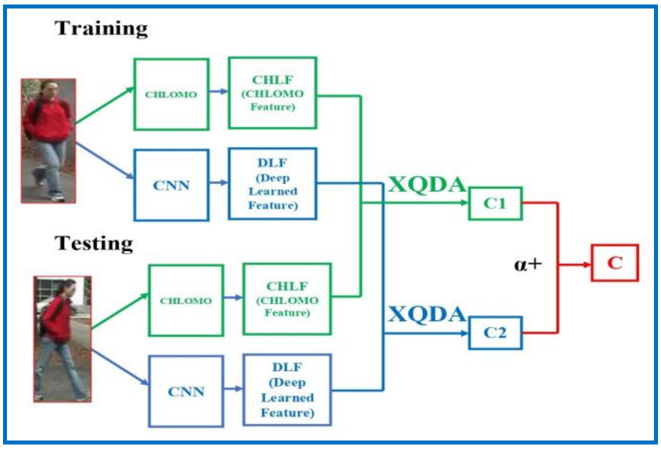

Figure 2: Fusion Process of CHLOMO and Deep Features

The proposed CHLFDLF differs from other person re-identification techniques like dense invariant feature-based support vector ranking (DSVR) in this regard. The primary distinction between the CHLFDLF and the DSVR [22] is that the former uses decision-level fusion as its fusion approach, while the latter simply yet effectively takes use of the teamwork between low level features i.e., handcrafted features and deep learning features. In contrast, the dense invariant features are fused via the feature-level fusion ranking approach identified as DSVR.

IV. THE PROPOSED ALGORITHM (CHLFDLF)

This section describes the step-by-step procedure of proposed CHLFDLF for person re-identification based on CHLOMO features and deep learned features. The detailed proposed algorithm as follows,

Input: HRimg – Person/Human Re-identification Datasets with different angle posed person images.
Notations
HR_{DB} – Human Re-identification $_{Database}$
C1 - XQDA distance between training and testing data's CHLOMO characteristics.
C2 - XQDA distance between training and test data's CNN features.
C - The overall mileage.

Stage I
\# Extract handcrafted features CHLF from HSV color and color textures using CHLOMO ($CHLOMO_{feat}$)

Step 1: for each image img $\in HR_{DB}$
h, w $\longleftarrow$ size(img)
\# h – Height of the image
\# w – Width of the image
$CHLOMO_{feat} \longleftarrow$ extract $CHLOMO_{feat}$ (HR_{DB})
End

Stage II
\# Extract deep features DLF automatically using CNN (CNN_feat)
Step 2: model $\longleftarrow$ train_CNNnetwork (HRDB, epoch)
Step 3: CNN_feat $\longleftarrow$ extract CNN_feat (model, HR_{DB})

Stage III
\# Person /Human re-identification
Step 4: Let $HRDB_{train} \in HR_{DB}$
Step 5: $HRDB_{test} \in HR_{DB}$
Step 6: Learn $CHLOMO_{feat}$ using the XQDA similarity metric learning approach and compute the distance C1.
C1 $\longleftarrow$ XQDA_similarity ($CHLOMO_{feat}$ ($HRDB_{train}$), $CHLOMO_{feat}$ ($HRDB_{test}$))

Step 7: Learn CNN_feat using the XQDA similarity metric learning approach and compute the distance C2

$C2 \longleftarrow$ XQDA_similarity (CNN_feat ($HRDB_{train}$), CNN_feat ($HRDB_{test}$))

Step 8: Compute overall distance C using equation (1)

C (distance) $\longleftarrow$ $C1+\alpha*C2$, $(0 < \alpha<1)$

Step 9: Identified $Human_{index} \longleftarrow$ min_C (distance)

Provide the ***average*** performance in different ranking such Rank1, Rank5, Rank10, Rank15, and so on. Rank the matching accuracy based on the computed distance C.

Output: The re-identified person image

In the proposed algorithm, the step 1 is used to extract the CHLOMO features for each image in the dataset. The CNN is trained in step 2 & 3 in order to extract deep features. The similarity of the XQDA between the training and test images is computed in steps 6 and 7. The Step 8 & 9 determines the matching accuracy depending on the calculated distance C and the output represents the re-identified person.

V. EXPERIMENTAL RESULTS

Every experiment was performed on a personal computer with Windows 10, an Intel®CoreTMi5-7200U Central Processing Unit running at 2.50GHz and 2.71GHz, and 8GB RAM. The suggested technique was created using MATLAB R2017b, which includes a Deep Learning tool as well as a toolkit for image processing. CUHK03 [10], Market 1501 [9], and VIPeR [8] were used as actual datasets to test this technique. Table 1 provides a full overview of the experimental datasets. Only 632 images are included in the VIPeR dataset. The dataset has grown by a factor of 64 as a result of the data augmentation process. Every experimental dataset was scaled to 230x80 pixels for efficiency because each one's size differs. Each class's training and testing images are chosen at random.

Table 1: Datasets used for Evaluating CHLFDLF

Dataset	Classes	Images	Training Images	Testing Images	Image Size
VIPeR	632	1264	316	316	128×48
Market1501	300	5933	4747	1186	128×64
CUHK03	742	7239	5756	1473	vary

1. **Datasets:** Each experimental dataset is thoroughly discussed in this section. On university campuses, a variety of cameras were used to capture the images in all three datasets. Using automated detection on a big scale (for Market-1501 and CUHK03) produces significant image misalignment, Nevertheless, likened to lesser person re-identification databases, this makes these datasets more realistic.

2. **VIPeR Dataset:** The utmost popular dataset for person re-identification is the VIPeR dataset [8], which contains 1264 photos of 632 distinct persons. Each person is made up of two images captured from two distinct random perspectives. It is difficult to re-identify

an individual using the VIPeR dataset, since there are substantial differences in the backdrop, viewpoint, and illumination, making it impossible to match the same person with only two different viewpoints. Here, Images acquired in one sight with Camera B are used as testing, while pictures captured in another sight with Camera A are used as training to evaluate the proposed approach. In order to facilitate training and testing, the 632-person VIPeR dataset is divided in half.

3. **Market1501 Dataset:** The Market 1501 [11] group comprises 32668 pictures of 1501 persons. This dataset contains photos that present a variety of difficulties for person re-identification, including body misalignment, different viewing angles, pose distortion, missing sections, and occlusion. To evaluate the effectiveness of the suggested method, 5933 pictures of 300 persons from the market 1501 dataset were used. 80 percent of each person is chosen arbitrarily for training, while the remaining 20% is chosen arbitrarily for testing.

4. **CUHK03 Dataset:** The 14,097 images in the CUHK03 dataset [10] depict 1,467 different personas. In this dataset, the pedestrians captured in several camera views faced challenges from illumination, positions, points of view, camera background boundaries, picture sizes, and congested backdrops. 7240 photos of 742 persons were taken from the CUHK03 dataset to test the suggested approach. 80 percent of each class/person is chosen at casual for training purposes, while the remaining 20% is chosen at casual for testing purposes.

Evaluation Criterion: The re-identification of person with metric learning algorithms like XQDA can be assessed by metrics such as Mahalanobis distance, Rank and CMC curve.

1. **Mahalanobis Distance:** The distance between the gallery and the probe is calculated using the Mahalanobis distance.

2. **Rank:** In person re-identification, a search user is matched against a group of gallery people in order to create a ranked list based on how closely they match. This process often assumes that the right match is given the top rank, preferably Rank 1. To compares the top n nearest images in the gallery set to one probe image. If a query image's correct match is at the kth position (k = n), this query is considered a success of rank n.

3. **CMC Curve:** Cumulative Matching Characteristics (CMC) curves are the most well-liked evaluation metrics for person re-identification. CMC curve is an accuracy curve that provides identification accuracy for each rank.

VI. RESULT ANALYSIS

In this research, initially, CHLOMO features are obtained as CHLF from every image. Then, each image is erudite using CNN for extracting Deep Features. For VIPeR, Market1501, and CUHK03, the layer of the first fully connected dimension is 400 and the layer of the second fully connected dimension is 632. By reckoning the right match from the topmost m matches, the proposed CHLFDLF's results are evaluated. In relation to the probe set, it assigns a rank to each and every sample in the gallery set. CMC curves are created by repeating this technique ten times. The quantitative evaluation of performance is carried out using the Cumulative Matching Characteristic (CMC) technique [11]. Table 2, Table 3, and Table 4 display the precise top 1, 5, 10, 15, and 20 ranking outcomes of CHLF, DLF, and the proposed CHLFDLF on all experimental datasets. Moreover, the VIPeR, Market1501, and CUHK03 dataset's person re-identification matching ranks using CHLFDLF are plotted in the CMC curve and are shown in Figure 3.

Table 2: Outcomes of the Chlomo Feature (CHLF)

Dataset	R1	R5	R10	R15	R20
CUHK03	92.66	97.83	98.78	99.05	99.32
Market 1501	85.69	95.62	97.81	98.99	99.32
VIPeR	42.41	72.47	83.54	91.46	94.94

Table 3: Outcomes of the Deep Learned Feature (DLF)

Dataset	R1	R5	R10	R15	R20
VIPeR	45.22	57.41	63.26	68.99	90.73
Market 1501	93.56	95.83	96.47	98.69	98.91
CUHK03	72.06	73.78	76.31	78.74	80.72

Table 4: Outcomes of the Proposed Method (CHLFDLF)

Dataset	R1	R5	R10	R15	R20
VIPeR	48.34	75.03	88.48	93.04	97.78
Market 1501	94.29	97.84	98.46	98.98	99.45
CUHK03	96.39	97.61	98.07	98.87	99.1

While comparing the results of CHLF, DLF, and CHLFDLF, the proposed CHLFDLF method provides outstanding results for person re-identification on the experimental datasets.

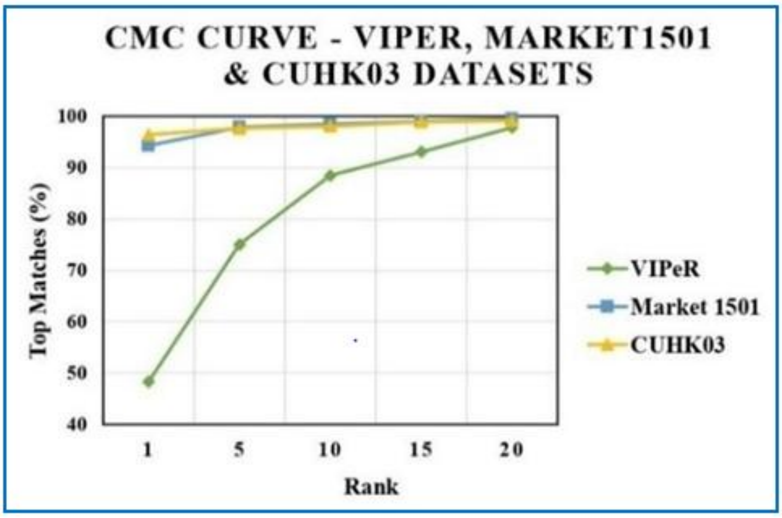

Figure 3: CMC Curve of CHLFDLF on VIPeR, Market1501& CUHK03 Datasets

1. **Evaluation of Performance:** This section compares the effectiveness of the CHLFDLF approach for person re-identification on the Market1501, CUHK03, and VIPeR datasets with that of several feature representation and metric learning approaches.

2. **Comparative Analysis of CHLFDLF on VIPeR dataset:** Comparing the CHLFDLF results on the VIPeR dataset against those of other approaches like CHLOMO [4], LOMO+XQDA [5], PCT-CNN [6], DMVFL [3], ELF [8], and DEep Clustering-based Asymmetric Metric Learning DECAMEL [12]. Table 5 displays the comparative findings from the VIPeR dataset. Table 5 shows that the proposed technique outperforms PCT-CNN and DMVFL by 1.19% and 1.95%, respectively, and reaches the most recent technology with 48.34% at Rank 1 (R1). While considering PCT-CNN by 1.19%, 0.51%, and 0.63% and it outperforms all R1, R10, and R20 the proposed CHLFDLF outperforms the other feature extraction and metric learning algorithms in comparison. The proposed method outperforms all previous methods individually, obtaining 48.34% Rank1 (R1) accuracy. The effectiveness of the proposed method is compared with the existing approaches in Figure 4.

Table 5: The Proposed CHLFDLF method compared with Existing Methods on the VIPeR Dataset

Method	R 1	R 10	R 20
CHLFDLF	**48.34**	**88.48**	**97.78**
PCT-CNN [6]	47.15	87.97	97.15
DMVFL [3]	46.39	86.1	95.32
CHLOMO [4]	42.41	83.54	94.94
LOMO with XQDA [5]	40.0	80	91
ELF [8]	12.0	44.0	61.0
DECAMEL [12]	34.15	-	-

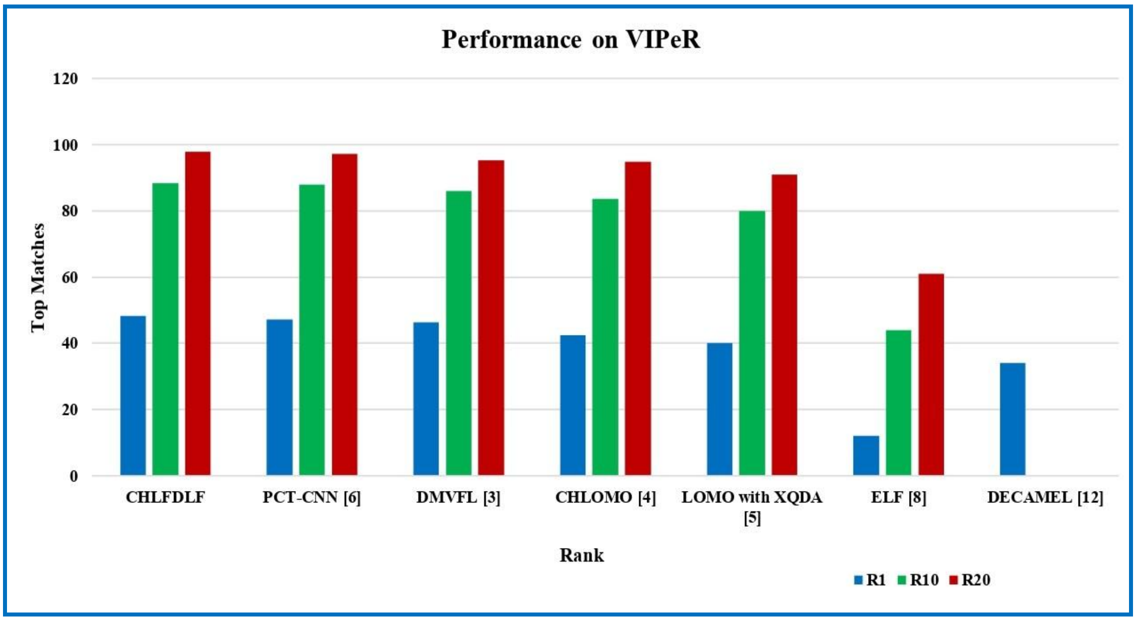

Figure 4: Performance Comparison of CHLFDLF on VIPeR

3. **Comparative Analysis of CHLFDLF on the Market1501 Dataset:** Comparing the CHLFDLF results on the VIPeR dataset against those of other approaches like CHLOMO [4], LF [13], LFDLF [13], DLPR [15], BoW [9], and k-reciprocal Encoding [14]. Table 6 provides a comparison of the proposed approach with previous works. The results of the comparison show that the proposed CHLFDLF approach outperforms other CNN- based methods, including DECAMEL [12], DSML [16], UTAL [17], and PersonNet [18]. With an R1 accuracy of 94.29%, the suggested approach exceeds every other method individually.

Table 6: The Proposed CHLFDLF Method Compared with Existing Methods on the Market1501 Dataset

Method	R1
CHLFDLF	**94.29**
CHLOMO [4]	85.69 CHLOMO [4]
LFDLF [13]	84.15
LF [13]	81.92
DLPR [15]	81
DSML [16]	84.4
k-reciprocal Encoding [14]	77.11
DECAMEL [12]	60.24
UTAL [17]	56.3 UTAL [17]
PersonNet [18]	37.21
BoW [9]	34.4

Table 6 shows that the suggested approach reaches the new state-of-the-art, 94.29% at (Rank1) R1. In comparison to previously studied approaches, the outcome shows that the proposed CHLFDLF features provide better consequences. In Figure. 5, the performance of the suggested approach is contrasted with that of the existing approaches.

Figure 5: Performance comparison of CHLFDLF on Market1501

4. **Comparative Analysis of CHLFDLF on the CUHK03 Dataset:** On the CUHK03 dataset, the results of CHLFDLF are compared to those of CHLOMO [4], Deep Part-Aligned [15], LOMO+XQDA [5], DSML [16], UTAL [17], DECAMEL VCFL [20], k-reciprocal Encoding [14], OSNet [19], FPNN [10], and PersonNet [18]. The performance of re- identification using various methodologies is evaluated using the Rank1(R1), Rank5 (R5), Rank10 (R10), and Rank20 (R1) accuracies. Table 7 shows that the suggested method beats all other approaches in comparison, with a re-identification accuracy of 97.61 % at rank 1. A comparison of the suggested strategy and the earlier approaches is shown in Table 7. Also, approaches based on CNN, like OSNet [19], DSML [16], UTAL [17], and PersonNet [18], are outperforming the proposed CHLFDLF handcrafted features and deep features method. The concert of the suggested method on the dataset for CUHK03 is compared with that of the existing approaches in Figure. 6. The suggested strategy exceeds all other approaches individually with a rank-1 accuracy of 97.61%.

Table 7: The Proposed CHLFDLF Method Compared with Existing Methods on the CUHK03 Dataset

Method	R1	R5	R10	R20
CHLFDLF	**97.61**	**98.07**	**98.63**	**99.16**
CHLOMO [4]	92.66	97.83	98.78	99.32
LOMO+XQDA [5]	91.05	97.05	98.25	99.02
Deep Part- Aligned [15]	85.4	97.6	99.4	99.9
FPNN [10]	20.65	51.32	68.74	83.06
PersonNet [18]	64.8	89.4	94.92	98.2
DSML [16]	88	98	99	
k-reciprocal Encoding [14]	69.9	61.6		
OSNet [19]	72.3	-	-	-
VCFL [20]	70.36	-	-	-
MFF [7]	69.6	-	-	-
UTAL [17]	69.2	-	-	-
DECAMEL [12]	45.82	-	-	-

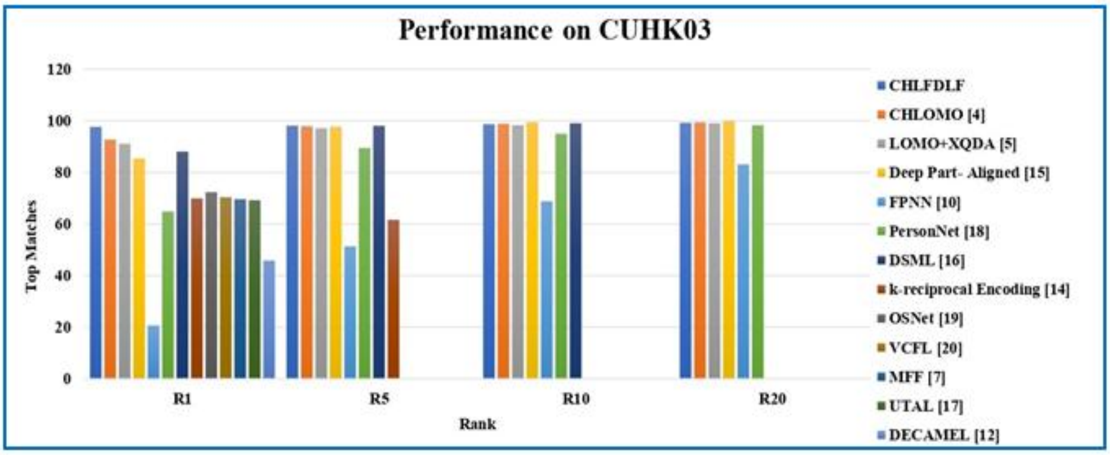

Figure 6: Performance comparison of CHLFDLF on CUHK03

VII. CONCLUSION

The combination of CHLOMO and Deep features (CHLFDLF) has been suggested in this research. It merges with efficient handcrafted features and CNN features to augment the representation. Here, the CHLOMO feature is extracted in the form of a color texture known as CHLF. Besides CHLOMO, deep-learned features are extracted as DLF by applying CNN. In order to effectively learn the distance, the deep features and CHLOMO features are integrated. The outcomes of the suggested CHLFDLF are depicted on VIPeR, CUHK03, and Market 1501 datasets. The experimental analysis represents that data sets used with these enhanced features drastically improve person re-identification and also handle small and large sample size problems in a well-organized way. The Rank1 accuracy of the proposed method CHLFDLF is 48.34 %, 94.29 %, and 96.39%, respectively. In the future, the proposed feature representation method can be implemented with more layers of deep learning networks and evaluated by using small-scale and large-scale datasets.

VIII. ACKNOWLEDGMENT

We sincerely express our gratitude to the editors and reviewers for their insightful comments which helped us to enhance the quality of the manuscript.

IX. NOMENCLATURE

LFDLF–LOMO Feature and Deeply Learned Feature
SLFDLF-Split LOMO Feature and Deeply Learned Feature
LF- LOMO Feature
DLF–Deep Learned Feature
CHLOMO–CHannel Integrated Local Maximal Occurrence
CHLF –CHannel Integrated LOMO Feature
CHLFDLF–Channel integrated LOMO feature and Deep- Learned Feature
SICISP–Scale-Invariant Channel Integrated Statistical Pattern
XQDA–Cross-view Quadratic Discriminant Analysis

REFERENCES

[1] W. Wanyin, Z. Yang, H. Li, J. Cheng, and D. Tao, “Deep features for person re-identification on metric learning,” Elsevier, Pattern recognition, 2020.
[2] M. Ye, S. C. H. Hoi, G. Lin, T. Xiang, L. Shao, and S.Jianbing, “Deep learning for person re-identification: A Survey and Outlook,” Pattern Analysis and Machine Intelligence, 2021.
[3] D. Tao, Y. Guo, Baosheng Yu, Jianxin Pang, and Zhengtao Yu, “Deep Multi-View Feature Learning for Person Re-Identification,” IEEE Transactions onCircuits and Systems for Video Technology, 2017.
[4] J.S.J. Rani and M.G. Augasta, “Human Re- Identification by Local Maximal Occurrence of Color and Scale Invariant Channel Integrated Statistical Pattern,” Springer, Multimedia Tools and applications, 2022.
[5] Zhu, S. Liao, S. Z. Li, and Y. Hu, “Person Re- identification by Local Maximal Occurrence Representation and Metric Learning,” IEEE, pp. 2197– 2206,2015.
[6] K. Jayapriya, J. Jacob, and N. A. B. Mary, “Person re-identification using prioritized chromatic texture (PCT) with deep learning,” Springer, Multimedia Tools and Applications, doi:10.1007/s11042-020-09528-8,2020.
[7] S. Wu and G. Ling, “Multi-Level Joint Feature Learning for person re-identification,” Algorithms,Vol. 13, No. 111, 2020.
[8] D. Gray and T. Hai, “Viewpoint invariant pedestrian recognition with an ensemble of localized features,” In European Conference on Computer Vision, pp.262–275, 2008.

[9] L. Zheng, L. Shen, L. Tian, S. Wang, Q. Tian, and W. Jingdong, "Scalable person re-identification: A benchmark," in Proceedings of the IEEE International Conference on Computer Vision, pp.1116-1124,2015.
[10] W. Li, R. Zhao, T. Xiao, and W. Xiaogang, "DeepReID: Deep Filter Pairing Neural Network for Person Re- Identification," IEEE Conference on Computer Vision and pattern recognition,2014.
[11] D. Gray, S. Brennan, and H. Tao, "Evaluating appearance models for recognition, reacquisition, and tracking," in Proceedings of the International Workshop on Performance Evaluation of Tracking and Surveillance, 2007.
[12] H.X. Yu, A. Wu, and Z. Wei-Shi, "Unsupervised Person Re-identification by Deep Asymmetric Metric Embedding," IEEE Transactions on Pattern Analysis and Machine Intelligence, 2019.
[13] J. S. J. Rani and M. G. Augasta, "A Deep Learning Model for Human Re-Identification with Split LOMO and Deep Features Using XQDA," Test Engineering & Management ISSN:0193-4120pp No.13776–13786, 2020.
[14] Z. Zhong, L. Zheng, D. Cao, and S. Li, "Re-ranking Person Re-identification with k-reciprocal Encoding," The IEEE Conference on Computer Vision and Pattern Recognition, pp: 318-1327, 2017.
[15] L. Zhao, X. Li, Y. Zhuang, and J. Wang, "Deeply- learned part-aligned representations for person re-identification," In IEEE international conference on computer vision, pp3219–3228, 2017.
[16] Z. Leic, X.L. Xub, and C.X. Rena "A Deep and Structured Metric Learning Method for Robust Person Re-Identification," Pattern Recognition, Vol. 96, 2019.
[17] M. Li, X. Zhu, and S. Gong, "Unsupervised Tracklet Person Re-Identification," Pattern Analysis and Machine Intelligence, Vol. 4,2020.
[18] L. Wu, C. Shen, and A. Hengel, "PersonNet: Person Re-identification with Deep Convolutional Neural Networks," arXiv: 1601.07255v2, 2016.
[19] K. Zhou, Y. Yang, T. Xiang, and A. Cavallaro, "Learning Generalisable Omni-Scale Representations for Person Re-Identification," Pattern Analysis and Machine Intelligence, 2021.
[20] F. Liu and L. Zhang, "View Confusion Feature Learning for Person Re-identification," IEEE International Conference on Computer Vision Foundation, 2019.
[21] D. Yi, Z. Lei, S. Liao, and S. Z. Li, "Deep metric learning for practical person re-identification," in Proceedings of the International Conference on Pattern Recognition, 2014, pp. 34-39.
[22] Tan, S., Zheng, F., Liu, L., Han, J., & Shao, L.), "Dense Invariant Feature-Based Support Vector Ranking for Cross-Camera Person Reidentification", 2018.

