Int. J. Advanced Networking and Applications
Volume: 17, Issue: 02, Pages : 6856-6863 (2025) ISSN: 0975-0290
https://doi.org/10.35444/IJANA.202517208

SNK-PU: Leveraging Multi-View CCTV Footage for Improved Person Re-Identification

Punit Sohanvi
University Institute of Engineering and Technology, Panjab University, Chandigarh, India
Email : punit7nov@gmail.com
Naveen Aggarwal
University Institute of Engineering and Technology, Panjab University, Chandigarh, India
Email : navagg@gmail.com
Nirmal Kaur
University Institute of Engineering and Technology, Panjab University, Chandigarh, India
Email : nirmaljul19@gmail.com

ABSTRACT

Person re identification (ReID) is a crucial computer vision task aimed at matching individuals across different camera views despite variations in viewpoint, lighting, and occlusions. Even though CCTV cameras are everywhere, person reidentification remains challenging due to their inefficiency in accurately capturing faces when they are covered or blurred, as most cameras are optimized for capturing faces from a single direction. The lack of datasets that focus on multi-view capturing of individuals from all four sides makes this issue worse, with few models addressing the need to reidentify individuals based on multiple images from various angles. To tackle this problem, we have employed a novel technique that leverages the keypoints and gait of individuals from a multi-dimensional perspective, rather than relying solely on facial or biometric images. Our proposed novel SNK-PU dataset addresses this gap by including images of individuals from the four specific orientations: front, back, left, and right. Additionally, it incorporates supplementary parameters such as height and age for ReID. We trained a Siamese model with a triplet loss function on the proposed SNK-PU dataset, utilizing our innovative image generation technique. When the entire dataset is used for both training and testing, the model achieves an accuracy of 95.11%. However, with an 80:20 train-test split, the accuracy reaches 86.45%, reflecting a more realistic evaluation scenario. Unlike traditional methods that rely on uniform image recognition, our approach leverages easily accessible keypoints, enabling more consistent and robust person re-identification across diverse conditions.

Keywords **-About five key words in alphabetical order, separated by commas.**

Date of Submission: July 16, 2025 Date of Acceptance: August 21, 2025

1. INTRODUCTION

Person reidentification (ReID) [1] involves identifying and tracking the same individual across multiple camera feeds, even when conditions like viewpoint, lighting, and occlusions vary. This task is essential for enhancing security and forensic capabilities. Existing methods [2] for person ReID are divided into two main categories: traditional methods and modern deep learning approaches. Traditional methods rely on hand-crafted feature extraction and metric learning to distinguish individuals based on color, texture, and shape descriptors. However, these methods frequently encounter challenges due to significant appearance variability from environmental changes and the limited availability of multi-view images of individuals.

In recent years, deep learning has revolutionized person ReID by leveraging convolutional neural networks (CNNs) [3] and other advanced techniques to automatically learn discriminative features from large datasets. Some of the most recent methods [4, 5] incorporate attention mechanisms and multi- scale feature extraction to further enhance performance. For example, Multi-Scale Context-Aware Network (MSCAN) [6] achieved a remarkable 95.4%Rank-1accuracy on the Market-1501 dataset, significantly outperforming earlier approaches by focusing on both local and global features and their interdependencies. Additionally, the use of generative adversarial networks (GANs) [7] to augment training data has also shown promise in improving ReID performance by simulating various challenging conditions. Despite these advancements, person ReID faces several challenges [8]. Variations in lighting conditions, occlusions, changes in pose, and background clutter can all negatively impact the accuracy of ReID systems. Additionally, lack of datasets focusing on multi-orientation views of individuals limits the ability to address these challenges effectively. These challenges make it difficult for models to consistently identify individuals across different scenarios.

In this paper, we present a novel approach to address the limitations posed by the lack of multi-view datasets in person re-identification. We introduce a purpose-built, multi-view dataset that not only fills this gap but also improves the ability to re-identify individuals who are partially obscured, including those with covered faces. By focusing on gait representation and the dynamic changes in keypoints across the different viewpoints, we develop a method that leverages these less obtrusive features for accurate-identification. This dual strategy of generating a diverse dataset and utilizing gait and keypoint information aims to improve robustness and reliability in scenarios where traditional identification methods fall short. The primary contributions of study include:

1) A novel approach for image representation by capturing multi-view perspectives of a person from CCTV footage to improve person re-identification.
2) A novel multi-view dataset "Siamese Network based on Keypoints for Panjab Univesrity Students" (SNK-PU) is designed to provide a comprehensive, multidimensional view of individuals from various perspectives, including left, right, front, and back. In addition to these diverse viewpoints, SNK-PU also incorporates additional parameters such as height, age, and gender, offering a detailed representation of each individual for more accurate person Re-ID.
3) Siamese network model is developed based on keypoints extraction and height of an individual. Proposed approach achievesan accuracy of 95.11% for SNK-PU dataset for complete and 86.45% partial variant.

This paper is organized to systematically address the topic of person reidentification and present advancements in the field. Section 2 begins with an introduction of publicly available datasets and ends by examining their limitations. In Section 3, we outline our proposed methodology, which utilizes multi-view CCTV footage to enhance person identification. Section 4 introduces a novel dataset SNK-PU that incorporates multi-view orientation and additional relevant parameters. Section 5 presents the improved results of our approach in comparison to existing methods on both our dataset and established datasets. Finally, Section 6 concludes the paper by summarizing the findings and discussing future research directions.

2. EXISTING DATASETS

Public re-identification (ReID) datasets are integral to the progress of research in the field of computer vision, particularly in the development of sophisticated algorithms de- signed to identify individuals across different camera views. These datasets serve as standardized benchmarks, enabling researchers to evaluate and compare the performance of various ReID methods in a consistent manner. By providing a diverse and extensive collection of images and videos of individuals captured under varying conditions, these datasets are essential in training models that can handle the complexities of real-world scenarios, such as changes in lighting, pose, and occlusion. Some of the key re-identification datasets that have significantly contributed to the field are shown in Table 1 and discussed below:

- **VIPER (2007)** [9]: The VIPER dataset features 1,264 images of 632 identities captured from different viewpoints. It challenges ReID algorithms to handle significant viewpoint variations, essential for testing robustness in real-world surveillance.

- **PRID-2011 (2011)** [10]: PRID-2011includes 1,134 image sequences of 200 individuals from overlapping outdoor camera views. It tests ReID performance under variations in lighting, viewpoint, and background clutter.

- **CUHK01 (2012)** [11]: CUHK01 contains 3,884 images of 971 identities from two camera views. It focuses on pose variation and moderate lighting changes, evaluating ReID algorithms under these conditions.

- **CUHK03 (2014)** [12]: CUHK03 provides 28,325 images of 1,467 identities from six cameras, with both manually labeled and automatically detected images. It tests ReID algorithms against detection errors, pose variations, and lighting changes.

- **iLIDS-VID (2014)** [13]: The iLIDS-VID dataset comprises 600 video sequences of 300 individuals captured in an airport setting, focusing on challenges such as motion, varying lighting, and occlusions.

- **Market1501 (2015)** [14]: Market1501 includes 32,668 images of 1,501 identities captured by six cameras. It addresses occlusions, background clutter, and lighting variations, serving as a benchmark for large-scale ReID evaluations.

- **MARS (2016)** [15]: MARS features over 1.1 million frames of 1,261 identities from six cameras on a university campus. It provides complex motion patterns, lighting variations, and occlusions for testing continuous tracking and ReID.

- **DukeMTMC-ReID (2017)** [16]: DukeMTMC-ReIDconsists of 36,411 training and 15,913 testing images of 1,812 identities from eight campus cameras. It challenges algorithms with diverse poses, lighting, and backgrounds.

- **DukeMTMC-VideoReID (2017)** [16]: This extension of DukeMTMC-ReID adds video sequences, incorporating temporal dynamics and motion patterns to enhance ReID performance.

- **MSMT-17 (2018)** [17]: MSMT-17 is one of the largest datasets with 126,441 images of 4,101 identities from 15 cameras in varied indoor and outdoor scenes, testing ReID algorithms under diverse environmental conditions.

TABLE 1: Person re-identification datasets (publicly available)

ataset(Year)	Environ ent	IDs	Cam s
IPER (2007) [9]	ampus	632	2
RID-2011(2011) [10]	utdoor	200	2
UHK01(2012) [11]	ampus	971	2
UHK03(2014) [12]	ampus	1360	10
IDS-Vid(2014) [13]	irport	300	2
larket1501(2015) [14]	ampus	1501	6
IARS(2016) [15]	ampus	1261	6
ukeMTMC-ReID(2017) [16]	ampus	1404	8
ukeMTMC-ideoReID(2017) [16]	ampus	1404	8
ISMT-17(2018) [17]	ampus	4101	15

Most of the current person-re-identification datasets are usually incomplete due to their failure to include some essential keypoints which are listed below which decrease their usability.

1) Most of the datasets do not divide individuals' orientation into front, back, left, and right views. Consequently, it is impossible for any model that requires understanding or prediction of human orientation to accurately re-identify various people from different angles.

2) Many of these sets lack height attributes of people in the images, which can help differentiate between individuals who look alike but have different heights. This makes such current datasets limited in terms of useful applications where a person's identity must be determined completely and with great accuracy.

3) Existing datasets do not include essential demographic parameters for each person, such as age and gender, which are critical for improving identification accuracy and model robustness.

Based on above points we have decided to create our own dataset to ensure comprehensive coverage of human orientations and inclusion of height, age and gender parameters thereby, facilitating more accurate and reliable person re- identification model development.

3. PROPOSED METHODOLOGY

The proposed methodology is Siamese based deep learning model for person Re-ID based on keypoints and height of the individuals, as shown in Figure 1.The captured person images within the dataset are categorized into four distinct folders: front, back, right, and left. Each folder corresponds to data from 74 individuals, with each individual represented by four images of each orientation. Subsequently, the following sequential steps are executed:

1) **Initial Key Point Extraction:** Images within these designated folders undergo keypoint extraction utilizing YOLOv8, followed by the computation of differences among these key points.
2) **Key Point Generation:** Key points are generated for various orientations, including front, back, left, right, front-back, front-right, front-left, back-right, back-left, and right-left, along with a comprehensive model encompassing all orientations.
3) **Key Point Representation:** The extracted key points are then transformed into images or adjusted to a specific aspect ratio by means of upscaling, leveraging the PIL library in Python.

4) **Siamese Network Integration:** These transformed images are subsequently inputted into a Siamese network with triplet loss function, structured around two distinctive categories: partial and complete. Partial category models are tailored to discern new individuals, while complete category models are designed for the purpose of person reidentification.

The image preparation for our person re-identification model involved several critical steps to ensure that data is ready for analysis. Initially, the images captured and stored in individual folders are passed through the YOLOv8 model to generate keypoints for each image. These keypoints, which represent crucial features of the individuals, are then extracted for the further processing. The differences in keypoints across the images are analyzed to enhance the accuracy and reliability of the person re- identification process.

This systematic approach facilitates robust analysis and recognition within the dataset, thereby contributing to enhanced accuracy and efficacy in subsequent tasks such as individual identification and reidentification.

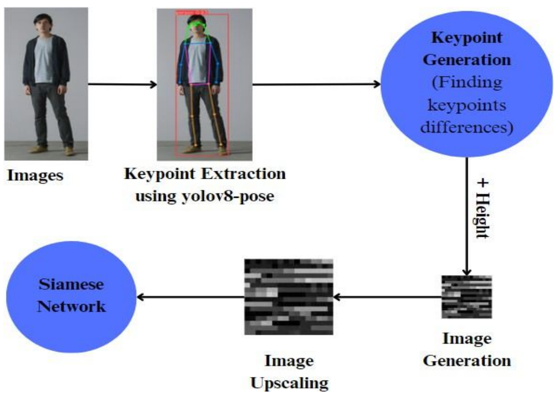


Fig. 1: Proposed approach for person ReID

4. SNK-PU DATASET

Person re-identification datasets are crucial for developing and evaluating algorithms aimed at recognizing individuals across different camera views and time instances in surveillance systems. However, the availability of such datasets is often limited, which poses challenges for researchers and practitioners in this field. In response to this limitation, developed SNK-PU dataset comprises 74 individuals, each with a comprehensive set of data captured from a single viewpoint as shown in Figure 2.

This novel dataset addresses the scarcity of person re-identification data by offering a diverse collection of individuals captured in various poses and under different environmental conditions, as illustrated in Table 2. The dataset consists of a video sequence and four images per person, each captured from distinct viewpoints: front, back, left, and right. These viewpoints provide comprehensive coverage of the individuals' appearances, facilitating robust person re-identification across different camera angles and orientations. It also includes the real height of individuals. The data collection process involves recording individuals using Hikvision camera positioned at a single point approximately 500 cm away from the subjects. This setup ensures consistency in the capture environment, minimizing variations in lighting, camera settings, and other factors that could affect the dataset's quality and reliability. By providing researchers and developers with access to this dataset, we aim to accelerate advancements in person reidentification technology. The availability of rich and diverse data is crucial for training and evaluating machine learning models, benchmarking algorithms, and fostering innovation in the field of video surveillance and security systems. In summary, our newly created dataset fills a gap in the availability of person re-identification data by offering a comprehensive collection of individuals captured from multiple viewpoints. We believe that this dataset will serve as a valuable resource for researchers and practitioners working on person re-identification algorithms, enabling them to develop more accurate and robust solutions for real-world applications.

(a) Front Side

(b) Left Side

(c) Back Side (d) Right Side

Fig.2: Person images captured from multiple views

TABLE 2: Dataset Description

Parameter	Description
Number of individuals	74
Number of Key points	17 per person
umber of Images and videos	14800 images and 74 videos
Video format and duration	Mp4, 30s sec each
mage format and resolution	BMP(1080*1920)
ther recorded parameters	eight, Age and Gender

4.1 Environment setup

The environment setup for our person re-identification dataset is meticulously arranged using a Hikvision camera, strategically placed 500 cm away from each individual. The camera, mounted at a height of 200 cm, is connected to a laptop equipped with Hikvision software to ensure seamless integration and data capture. Efforts were made to minimize obstructions between the camera and the subjects, resulting in a clear line of sight. The dataset, comprising images and videos of 74 individuals, is recorded under optimal conditions with proper lighting and minimal movement, ensuring high-quality data capture.

4.2 Data acquisition

UsingaHikvisioncamerapositioned500cmawayandat a height of 200 cm, we captured a around 200 images for each of the 74 individuals, ensuring comprehensive coverage with four images each from the front, back, left, and right sides. Additionally, a 30-second video of each individual was recorded, showcasing all four sides. The images, in BMP format, are captured at a resolution of 1080x1920 pixels and a frame rate of 20 fps, while the videos were recorded in MP4 format. The recording environment was optimized with proper lighting, minimal obstruction, and very little movement from the individuals to ensure high-quality data acquisition.

4.3 Data preprocessing

The dataset preprocessing for our person re-identification study is designed to enhance data quality and storage efficiency. Initially, the images and videos captured using the Hikvision camera is reviewed to ensure minimal noise and optimal clarity. The original BMP images, which were recorded at a resolution of 1080x1920 pixels and 20fps, were subsequently converted to JPG format to facilitate easier storage, as BMP files tend to occupy significant space. This conversion maintained image quality while reducing file size. Additionally, the videos, originally in MP4 format, are verified for consistency and clarity, ensuring they accurately represented each individual from all four sides. This preprocessing step was crucial in preparing a streamlined, high- quality dataset for further analysis.

4.4 Storage

The storage of our person re-identification dataset was meticulously organized to ensure easy access and efficient data management. Each individual’s data, consisting of both images and a video, was stored in a dedicated folder named after the individual. Within these folders, the photos were systematically labeled to indicate the side and the height of the individual, resulting in filenames such as ”left_height”, ”right_ height”, ”back_height”, and ”front_height” for the four images captured from each angle. The videos, which provided a comprehensive view of all four sides of the individual, were also stored in the respective folders. This structured organization facilitated straightforward retrieval and analysis of the dataset, ensuring that each individual’s data was clearly labeled and easily identifiable.

4.5 Dataset Division

The proposed approach is divided into two categories based on the amount of dataset used for training and testing:

Partial Dataset: The dataset is partitioned into training and testing sets utilizing an 80:20 split ratio. Consequently, the training set comprised images of 66 individuals, while the test set contained images of 8 individuals.

Complete Dataset: The entire dataset is utilized for both training and evaluation, encompassing images of all 74 individuals.

The reason for dividing the proposed approach into two categories (partial dataset and complete dataset) lies in their distinct applications. Models trained on partial datasets are optimized for recognizing new individuals, making them particularly useful in security contexts such as theft prevention. Conversely, models trained on complete datasets are designed to excel at re-identifying individuals already recorded in the system. This capability is beneficial for applications such as attendance tracking and worker counting in factories. By addressing these two specific needs, the approach ensures that the model is adept at both recognizing new

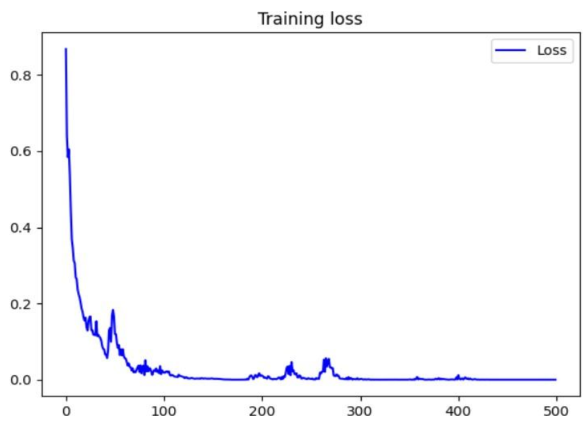
Int. J. Advanced Networking and Applications
Volume: 17, Issue: 02, Pages : 6856-6863 (2025) ISSN: 0975-0290
https://doi.org/10.35444/IJANA.202517208

people and accurately identifying those previously recorded.

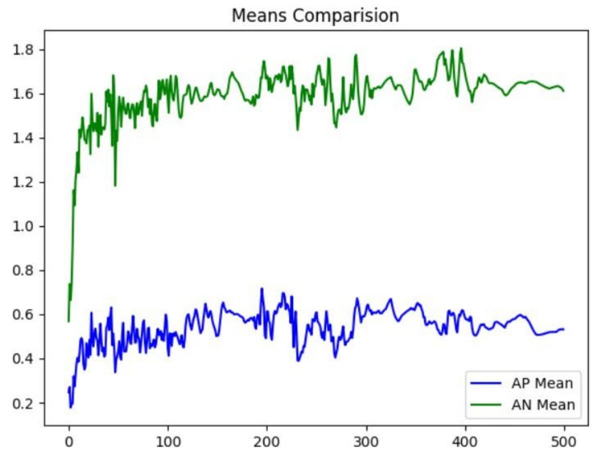
5. EXPERIMENTAL RESULTS ON SNK-PU DATASET

The publicly available Siamese network, utilizing triplet loss for person re identification (ReID), demonstrates a practical application in distinguishing individuals. Trained extensively on this dataset, the model achieves an impressive accuracy of 98.83%, illustrating its ability to reliably differentiate individuals based on the dataset's characteristics. The precision rate of 100% and a robust recall of 97.66% underscore the model's effectiveness in providing precise and comprehensive person matching. With an exceptional F1 score of 98.81%, the model maintains an excellent balance between precision and recall, which is crucial for real-world applications. However, despite these impressive results, the model may face limitations in practical scenarios where individuals' faces are occluded. For effective-identification, the model needs to accommodate multiple orientations, such as front, left, back, and right views.

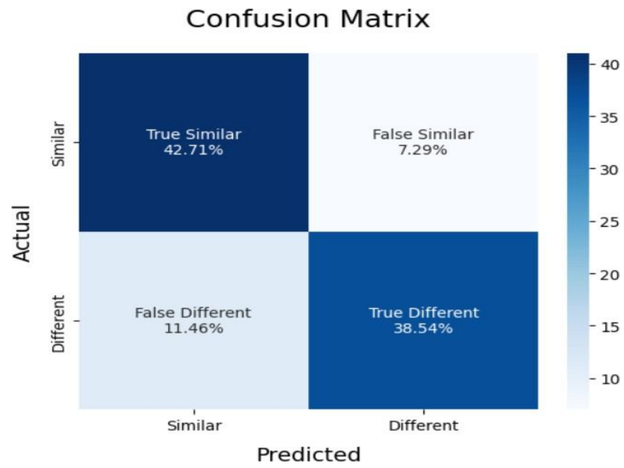
To address these limitations, the proposed approach leverages person keypoints extracted from CCTV footage. This method re-identify individuals based on multiple keypoint views, including front, back, left, and right orientations. Using this keypoint-based approach in conjunction with the Siamese network, we observed promising results. We trained the Siamese model on the SNK-PU dataset using our innovative image generation method. With the complete dataset accuracy began at 89.64% after 64 epochs, improved to 92.18% after 100 epochs, peaked at 95.11% after 200 epochs, and slightly adjusted to 92.77% after 500 epochs. When the dataset was split into an 80:20 ratio for training and testing, accuracy started at 76.04% after 64 epochs, rose to 78.13% after 100 epochs, reached 86.45% after 200 epochs, and slightly adjusted to 81.25% after 500 epochs.



(a) Training loss



(b) Mean comparison



(c) Confusion matrix

Fig 3: (a-c) Performance evaluation of proposed approach on partial SNK-PU dataset

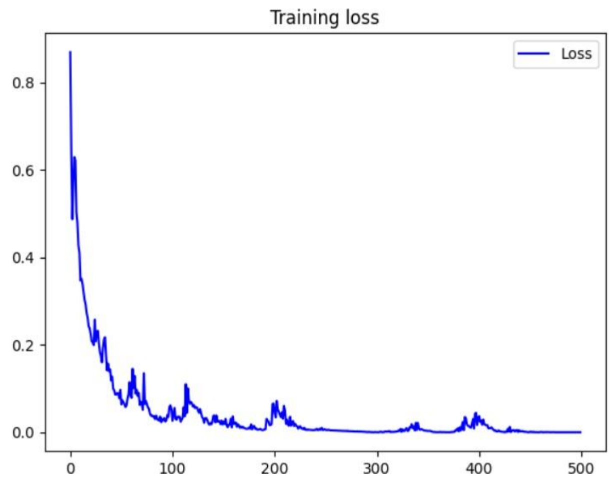
Figure 3 and Figure 4 shows the results achieved on the SNK-PU dataset using our proposed approach. The loss function converged to zero during training, indicating effective loss minimization and robust training. The mean comparison analysis demonstrates the network's ability to differentiate between similar and dissimilar images, ideally assigning low distances to anchor and positive images and higher distances to negative ones. This grouping ensures effective reidentification by clustering similar individuals and separating dissimilar ones. The confusion matrix further highlights the model's high true positive rate, showing strong performance in distinguishing positive from negative images, although it occasionally struggles with edge cases, suggesting areas for improvement. These results demonstrate the efficacy of the Siamese network in handling keypoint-based data, with consistently high accuracy across varying training durations, reflecting its robustness and suitability for more comprehensive person re-identification tasks.

Volume: 17, Issue: 02, Pages : 6856-6863 (2025) ISSN: 0975-0290
https://doi.org/10.35444/IJANA.202517208

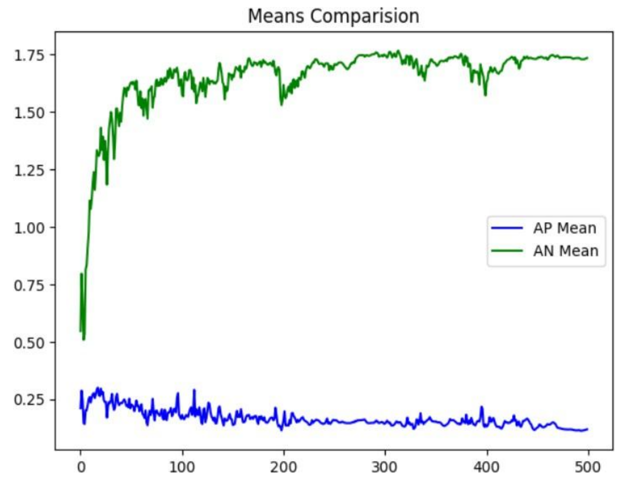
6. CONCLUSIONS

Despite notable progress in person ReID, challenges [8] such as illumination changes, occlusions, pose variations, and background clutter continue to hinder performance, further compounded by the lack of multi-view datasets. To bridge this gap, we introduced a purpose-built multi-view dataset that captures diverse perspectives from CCTV footage. Also, we have discussed a novel deep learning-based approach is implemented based on Siamese [18] network with triplet loss function. This helps to improve accuracy and efficiency in recognizing individuals/persons across different camera views or time instances. For this purpose, images are grouped into four separate folders according to their orientation: front, back, right and left. Each folder contains data from 74 persons, with four images for each orientation per person. Initially, keypoints are extracted from the images. The differences between these keypoints are then computed to facilitate image generation. Keypoints are generated for various orientations, including front, back, left, and right. These keypoints are subsequently transformed into images and fed into a Siamese network using a triplet loss function, with the dataset categorized into two types: partial and complete. Models trained on partial datasets are optimized for recognizing new individuals, useful for security applications like theft prevention. Models using complete datasets focus on re-identifying known individuals, beneficial for attendance tracking and worker counting. This methodical system enables the dataset to analyse and recognize more effectively, thereby enhancing accuracy of individual person identification or person re-identification.

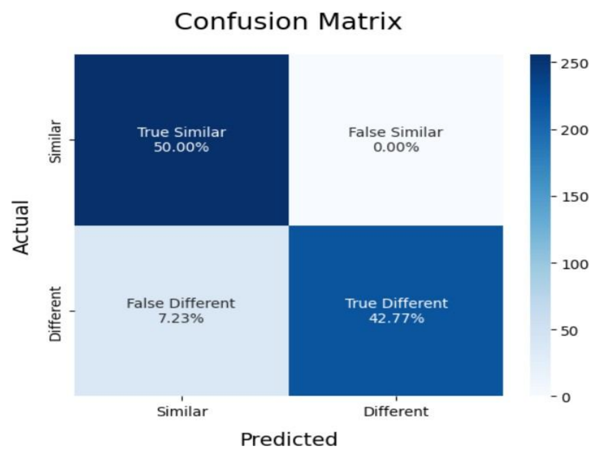
A novel dataset named "SNK-PU" is also designed to improve model accuracy in recognizing individuals from different angles. Dataset is collected for 74 persons, each captured from four distinct viewpoints: front, back, left, and right. These multiple viewpoints help in identifying individuals across different camera perspectives. Apart from distinct camera viewpoints, height of the person is also included for person re-identification. The proposed approach is capable of determining the similarity between people in images by using keypoints, specific viewpoint and height of person, Thorough experiments conducted on SNK-PU dataset have achieved accuracy of 86.45% under partial categories and 95.11% under complete categories. These achievements clearly indicate that Siamese network-based technique is robust especially when dealing with difficult situations such as person re-identification and individual recognition.



(a) **Training loss**



(b) **Mean comparison**



(c) **Confusion matrix**

Fig. 4: (a-c) Performance evaluation of our proposed approach on complete SNK-PU Dataset

Author Contributions

Punit Sohanvi contributed to data curation, resources, formal analysis, design, and draft writing. Naveen Aggarwal was responsible for conceptualization, supervision, dataset creation, methodology, and final version of paper. Nirmal Kaur contributed to conceptualization, supervision, resources, data curation, and visualization. All authors reviewed and approved the final version of the manuscript and agree to be accountable for all aspects of the work.

REFERENCES:

[1]. Ye, Mang, Jianbing Shen, Gaojie Lin, Tao Xiang, Ling Shao, and Steven CH Hoi. ”Deep learning for person re-identification: A survey and outlook.” IEEE transactions on pattern analysis and machine intelligence 44, no. 6 (2021): 2872-2893.

[2]. Zheng, Liang, Yi Yang, and Alexander G. Hauptmann. ”Person re identification: Past, present and future.” arXiv preprint arXiv:1610.02984 (2016).

[3]. Wu, Lin, Chunhua Shen, and Anton van den Hengel. ”Personnet: Person re-identification with deep convolutional neural networks.” arXiv preprint arXiv:1601.07255 (2016).

[4]. Sarker, Prodip Kumar, Qingjie Zhao, and Md Kamal Uddin. ”Transformer-based person re-identification: a comprehensive review.” IEEE Transactions on Intelligent Vehicles (2024).

[5]. Zhu, Kuan, Haiyun Guo, Shiliang Zhang, Yaowei Wang, Jing Liu, Jinqiao Wang, and Ming Tang. ”Aaformer: Auto-aligned transformer for person re-identification.” IEEE Transactions on Neural Networks and Learning Systems (2023).

[6]. Huang, Zhi-yong, Wen-cheng Qin, Fen Luo, Tian-hui Guan, Fang Xie, Shu Han, and Da-ming Sun. ”Combination of validity aggregation and multi-scale feature for person re-identification.” Journal of Ambient Intelligence and Humanized Computing (2023): 1-16.

[7]. Liu, Deyin, Lin Wu, Richang Hong, Zongyuan Ge, Jialie Shen, Farid Boussaid, and Mohammed Bennamoun. ”Generative metric learning for adversarially robust open-world person re-identification.” ACM Transactions on Multimedia Computing, Communications and Applications 19, no. 1 (2023): 1-19.

[8]. Zahra, Asmat, Nazia Perwaiz, Muhammad Shahzad, and Muhammad Moazam Fraz. ”Person re-identification: A retrospective on domain specific open challenges and future trends.” Pattern Recognition 142 (2023): 109669.

[9]. Gray, Douglas, and Hai Tao. ”Viewpoint invariant pedestrian recognition with an ensemble of localized features.” In Computer Vision–ECCV 2008: 10th European Conference on Computer Vision, Marseille, France, October 12-18, 2008, Proceedings, Part I 10, pp. 262-275. Springer Berlin Heidelberg, 2008.

[10]. Hirzer, Martin, Csaba Beleznai, Peter M. Roth, and Horst Bischof. ”Person re-identification by descriptive and discriminative classification.” In Image Analysis: 17th Scandinavian Conference, SCIA 2011, Ystad, Sweden, May 2011. Proceedings 17, pp. 91-102. Springer Berlin Heidel berg, 2011.

[11]. Li, Wei, Rui Zhao, and Xiaogang Wang. ”Human reidentification with transferred metric learning.” In Computer Vision–ACCV 2012: 11th Asian Conference on Computer Vision, Daejeon, Korea, November 5-9, 2012, Revised Selected Papers, Part I 11, pp. 31-44. Springer Berlin Heidelberg, 2013.

[12]. Li, Wei, Rui Zhao, Tong Xiao, and Xiaogang Wang. ”Deepreid: Deep f ilter pairing neural network for person re-identification.” In Proceedings of the IEEE conference on computer vision and pattern recognition, pp. 152-159. 2014.

[13]. Wang, Taiqing, Shaogang Gong, Xiatian Zhu, and Shengjin Wang. ”Person re-identification by discriminative selection in video ranking.” IEEE transactions on pattern analysis and machine intelligence 38, no. 12 (2016): 2501-2514.

[14]. Zheng, Liang, Liyue Shen, Lu Tian, Shengjin Wang, Jingdong Wang, and Qi Tian. ”Scalable person re-identification: A benchmark.” In Pro ceedings of the IEEE international conference on computer vision, pp. 1116-1124. 2015.

[15]. Zheng, Liang, Zhi Bie, Yifan Sun, Jingdong Wang, Chi Su, Shengjin Wang, and Qi Tian. ”Mars: A video benchmark for large-scale person re-identification.” In Computer Vision–ECCV 2016: 14th European Con ference, Amsterdam, The Netherlands, October 11-14, 2016, Proceedings, Part VI 14, pp. 868-884. Springer International Publishing, 2016.

[16]. Zheng, Zhedong, Liang Zheng, and Yi Yang. ”Unlabeled samples generated by gan improve the person re-identification baseline in vitro.” In Proceedings of the IEEE international conference on computer vision, pp. 3754-3762. 2017.

[17]. Wei, Longhui, Shiliang Zhang, Wen Gao, and Qi Tian. ”Person transfer gan to bridge domain gap for person re-identification.” In Proceedings of the IEEE conference on computer vision and pattern recognition, pp. 79-88. 2018.

[18]. Guo, Qing, Wei Feng, Ce Zhou, Rui Huang, Liang Wan, and Song Wang. ”Learning dynamic siamese network for visual object tracking.” In Proceedings of the IEEE international conference on computer vision, pp. 1763-1771.