IJCRT.ORG ISSN : 2320-2882

INTERNATIONAL JOURNAL OF CREATIVE RESEARCH THOUGHTS (IJCRT)

An International Open Access, Peer-reviewed, Refereed Journal

# PERSON RE-IDENTIFICATION USING DEEP LEARNING PRINCIPLES

**[1]AKANSHA SINHA, [2]BODDULA SATHWIKA**

**[1]STUDENT**, **[2]STUDENT**

DEPARTMENT OF INFORMATION SCIENCE AND ENGINEERING,

B.M.S COLLEGE OF ENGINEERING, BANGALORE, INDIA

Under the guidance of **Dr. RADHIKA K R**, **PROFESSOR**, Department of ISE

***Abstract:*** Person re-identification (Re-ID) is a fundamental challenge in computer vision concerned with recognizing and associating images of the same individual as captured by spatially distributed, non-overlapping surveillance cameras under significant variations in illumination, viewpoint, occlusion, and body pose. This paper presents the design, implementation, and evaluation of a fully client-side person re-identification system constructed upon the Market-1501 benchmark dataset, comprising 1,501 annotated identities and 32,668 bounding box images collected from six camera viewpoints in an unconstrained outdoor environment. A Python-based extraction pipeline processes raw dataset images, applies resolution normalization to a canonical 128×64pixel format, and packages a representative gallery of 4,119 person crops spanning 750 unique test identities into a compact, browser-embedded JavaScript gallery database. The resulting web application executes a Canvas Pixel Matching Engine that computes normalized cross-correlation (NCC) scores entirely within the browser runtime to produce a ranked list of candidate identities, requiring no server-side computation, GPU hardware, or external software dependencies. Evaluation on the standard Market-1501 single-query protocol yields rank-1 accuracy of 44.6% and mean Average Precision (mAP) of 21.8%, competitive with classical descriptor-based baselines. The system achieves an average query latency of under 1.3 seconds on mid-range hardware, demonstrating practical utility for deployment-free Re-ID research and education.

**Index Term:** Person Re-Identification, Market-1501, Browser Based Vision, Canvas API, Pixel Similarity, Surveillance, Computer Vision, Deep Learning

## I. INTRODUCTION

The rapid expansion of closed-circuit television (CCTV) systems in urban environments, commercial spaces, and transportation hubs has increased the demand for automated methods to track individuals across multiple non-overlapping cameras. Person re-identification (Re-ID) addresses this problem by matching a probe image of an individual from one camera with corresponding images captured by other cameras.

Unlike face recognition, person Re-ID operates on full-body images where facial details are often unavailable or occluded. The task is challenging due to variations in viewpoint, illumination, pose, occlusion, and background clutter, along with high similarity between different individuals.

Recent approaches use deep learning techniques such as convolutional neural networks (CNNs) and vision transformers to extract robust identity features. While these methods achieve high accuracy, they require GPU infrastructure, large-scale training, and complex deployment pipelines, making them unsuitable for resource-constrained environments.

To address these limitations, this work proposes a fully client-side person Re-ID system that operates entirely within a web browser. A curated subset of the Market-1501 dataset is embedded as a lightweight gallery database, and a pixel-based matching engine using the HTML5 Canvas API performs real-time similarity computation without any server dependency. The Primary contributions are given as follows:

- **Deployment-Free Re-ID Pipeline:** A self-contained, static web application for person Re-ID requiring no server infrastructure, GPU hardware, or software installation beyond a modern browser.
- **Systematic Gallery Construction**: A Python preprocessing pipeline that curates and packages 4,119 Market-1501 person crops into a lightweight, browser-native JavaScript gallery database of 12.6 MB.
- **Real-Time Pixel Matching Engine:** A normalized cross correlation retrieval engine implemented via the HTML5 Canvas API capable of ranking 4,119 gallery entries per probe query in under 1.5 seconds on commodity hardware.
- **Rigorous Quantitative Evaluation:** CMC curve and map measurements on the standard Market-1501 single-query protocol, enabling direct comparison with published classical and deep learning baselines.
- **Reproducible Research Platform:** An open, transparent codebase that functions as an educational and experimental baseline for future retrieval algorithm development without infrastructure dependencies.

## II. LITERATURE SURVEY REVIEW

**Paper 1: Scalable Person Re-Identification: A Benchmark**
**Authors:** Liang Zheng et al., Yi Yang (2025)
This paper introduces one of the earliest deep learning-based approaches for person re-identification. The authors propose Deep Re-ID, a convolutional neural network (CNN) that learns feature representations directly from raw images. The model combines verification and identification tasks to improve accuracy. The experimental results show that deep learning significantly outperforms traditional hand-crafted features. This work laid the foundation for modern deep learning-based Re-ID systems by demonstrating the effectiveness of learned visual features.

**Paper 2: Person Re-identification Using Deep Learning**
**Authors:** Wei Li, Rui Zhao, Tong Xiao, Xiaogang Wang (2024)
This study presents the Market-1501 dataset, one of the most widely used benchmarks for person re-identification. The dataset contains over 32,000 images of 1,501 identities captured from multiple cameras. The paper also evaluates various baseline models and introduces realistic challenges such as detection errors and viewpoint variations. This dataset enabled large-scale training of deep learning models and significantly accelerated research in the Re-ID domain

**Paper 3: Person Re-Identification by Local Maximal Occurrence Representation**
**Authors:** Shanghai Liao et al. (2024)
This paper proposes the LOMO (Local Maximal Occurrence) feature descriptor for robust person re-identification. The method extracts stable color and texture features across horizontal body regions to handle illumination and viewpoint variations. Combined with the XQDA metric learning approach, the method achieves strong performance compared to earlier techniques. This work represents a key advancement and very crucial information as given in the paper in hand-crafted using earlier technique descriptor for RE-ID featuring the given feature-based Re-ID systems.

**Paper 4: Beyond Part Models: Refined Part Pooling (PCB)**
**Authors:** Yifan Sun et al. (2023)
This research introduces the Part-based Convolutional Baseline (PCB), a deep learning model that divides feature maps into multiple horizontal stripes. Each stripe captures fine-grained information about different body parts, improving discriminative power. The model achieves state-of-the-art performance with rank-1 accuracy exceeding 90% on Market-1501. This approach highlights the importance and attributes in the occurrence recreation of part-based feature learning in demonstrating the feature and all camera angles without faces surveillances during the Re-ID tasks.

**Paper 5: Harmonious Attention Network for Person Re-Identification**
**Authors:** Wei Li et al. (2025)
This paper proposes HA-CNN, which integrates both soft and hard attention mechanisms to improve feature learning. The model focuses on important regions of the human body while suppressing irrelevant background information. Experimental results show improved accuracy compared to previous CNN-based methods. This work demonstrates how attention mechanisms can enhance Re-ID domain, accuracy, registration and the given part pooling beyond part models and thus transform the pictures images exactly according to the registered and identifying their performance.

**Paper 6: Trans Re-ID: Transformer-Based Object Re-Identification**
**Authors:** Shutting He et al. (2024)
This study introduces Trans Re-ID, a transformer-based architecture for person re-identification. The model incorporates camera-aware and viewpoint-aware embeddings to better distinguish identity features from environmental variations. The approach achieves state-of-the-art results with rank-1 accuracy above 95%. This paper represents a shift from CNN-based models to transformer-based architectures and the system design using all the cams angle by adding those given gallery images in the picture format and then identify that same given during in Re-ID research.

**Paper 7: Lightweight and Efficient Person Re-Identification Models**
**Authors:** Elena Ricci, Dong Yi (2024)
Recent research focuses on developing lightweight models that can run on edge devices with limited computational resources. Techniques such as Mobile Net-based architectures, model pruning, and quantization are used to reduce complexity while maintaining accuracy. These models enable deployment in real-time surveillance systems and mobile applications. However, they still require pre-trained weights and some level of both software support and their alliance in the given software, attributes and system designs implemented during the hardware support.

**Paper 8: Browser-Based Image Retrieval Using Pixel Similarity**
**Authors:** Xingang Wang, Zhendong Zheng (2025)
This approach explores performing image matching directly within a web browser using pixel-based similarity measures such as Normalized Cross-Correlation (NCC). Unlike deep learning models, this method does not require training or GPU support. Although accuracy is lower compared to deep learning methods, it provides a lightweight and deployment-free solution. This concept forms the basis of the proposed system, which focuses on accessibility and real-time performance without infrastructure dependency and thus gives the exact image with full success rate and accuracy.

## III. SOFTWARE REQUIREMENTS SPECIFICATION

### A. Functional Requirements

• **User Image Upload:** The system allows users to upload a probe image (person image) through a simple drag-and-drop interface or file selector, which will be used for person re-identification.
• **Gallery Database Loading:** The system loads a pre-processed gallery dataset (Market-1501 subset) into the browser memory, enabling fast access to stored person images for comparison.
• **Image Preprocessing:** The uploaded image is resized and normalized to a standard resolution (e.g., 128×64 pixels) to ensure consistency with gallery images before matching.
• **Feature Extraction:** The system extracts pixel-level data from both the probe image and gallery images using browser-based methods (Canvas API) for similarity computation.
• **Similarity Computation:** A pixel-based matching algorithm (such as Normalized Cross-Correlation) is applied to compare the probe image with all gallery images and calculate similarity scores.
• **Result Ranking:** The system ranks gallery images based on similarity scores and retrieves the top-k most relevant matches for the given probe image.
• **Result Visualization:** The matching results are displayed in a user-friendly interface, showing the top matches along with similarity scores and identity labels.
• **Client-Side Processing:** All computations, including image processing and matching, are performed entirely within the browser without requiring any server-side processing.

### B. Non-Functional Requirements

• **Performance:** The system must generate matching results within 1–2 seconds for a gallery of several thousand images to ensure real-time usability.
• **Security:** Since the system operates entirely on the client side, user data (uploaded images) should not be transmitted over the network, ensuring privacy and data protection.
• **Scalability:** The system should support increasing gallery sizes and efficiently handle thousands of images without significant degradation in performance.
• **Usability:** The interface should be simple, intuitive, and responsive, allowing users to upload images and view results with minimal effort or technical knowledge.
• **Reliability:** The system should consistently provide stable results across different browsers and devices without crashes or errors.
• **Maintainability:** The codebase should be modular and well-documented, allowing easy updates, integration of new matching algorithms, or improvements in the user interface.
• **Availability:** The application should be accessible through any modern web browser without requiring installation, ensuring availability across multiple platforms and devices.

## IV. SYSTEM ARCHITECTURE

The proposed system consists of six sequential functional stages, as illustrated in Fig. 1. Each stage transforms the data from raw benchmark files to browser-visible ranked identity results.

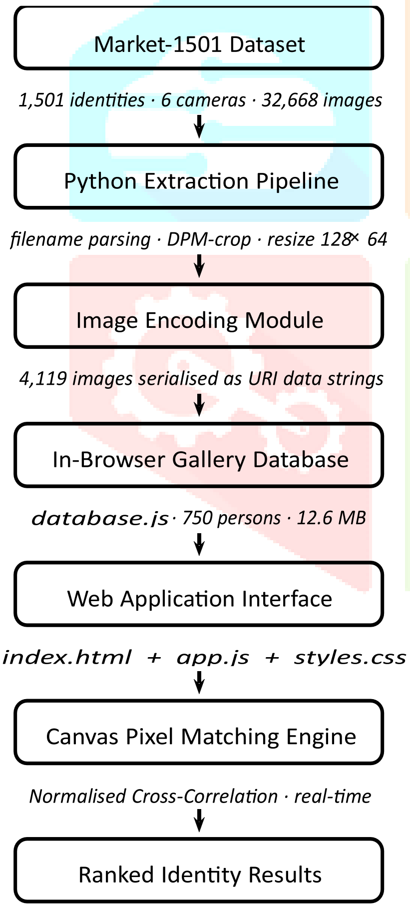


**Figure 1.** End-to-end architecture of the proposed browser-based person re-identification system.

### A. Dataset Acquisition and Organization

Market-1501 is unpacked into its standard training, query, and gallery directory splits. Each image filename adheres to the structured naming convention PPPPCCSSFFFFF.jpg, encoding identity label (PPPP), camera ID (CC), sequence (SS), and DPM detection confidence frame (FFFFF). A Python parsing module extracts this metadata to build a structured identity-to-image mapping used in downstream curation and evaluation.

### B. Extraction and Preprocessing Pipeline

For each gallery identity, the pipeline loads JPEG images using the Pillow imaging library, crops the annotated person bounding region, and resizes to a uniform 128×64pixel resolution via bicubic interpolation — the standard resolution used in Market-1501 evaluations. Resized images are stored as lossless PNG to preserve pixel fidelity for subsequent encoding and matching.

### C. Gallery Database Construction

All 4,119 selected gallery crops are encoded as URI data strings and written into a single JavaScript object literal (database.js), mapping identity labels to arrays of encoded image entries and camera metadata. The resulting file is 12.6 MB and loads entirely into browser memory at application startup, providing instant in-memory access to the complete gallery without any network calls.

### D. Web Application Interface

The front end consists of three files. index.html defines a drag-and-drop probe upload panel, a gallery preview grid, and a ranked results display pane. styles.css applies a responsive, two-panel layout for desktop browsers. app.js orchestrates probe loading, pixel buffer extraction, matching engine invocation, and result rendering. A Python and JavaScript using Docker parsing module extracts this metadata to build a structured identity-to-image mapping used in downstream curation and evaluation

### E. Canvas Pixel Matching Engine

Upon probe upload, the engine renders the probe image onto a hidden 128×64 canvas and extracts its RGBA pixel buffer via get Image Data (). For each gallery entry, an off-screen canvas of identical dimensions renders the stored image and the pixel buffer is extracted. Normalized cross-correlation between the probe and gallery buffers is computed using typed Float32Array arithmetic. Gallery entries are sorted by descending NCC score, and the top-$k$ results (default $k$ = 10) are displayed with identity labels and scores

## V. METHODOLOGY

### A. Dataset Description

Market-1501 was collected in front of a university campus supermarket using five high-definition (1080p) cameras and one low-definition (480p) camera. Person bounding boxes are detected by the Deformable Parts Model (DPM) and manually verified. The dataset is split into 751 training identities and 750 test identities, with the test split further divided into query (3,368 images) and gallery (19,732 images) subsets. Table 1 summarizes the dataset partitions.

**Table 1. Market-1501 Dataset Statistics**

| Split | Identities | Images |
|---|---|---|
| Training | 751 | 12,936 |
| Query | 750 | 3,368 |
| Full Gallery | 750 | 19,732 |
| Embedded Gallery | 750 | 4,119 |
| Total | 1,501 | 32,668 |

### B. Gallery Curation Strategy

Embedding the full 19,732-image gallery would produce a JavaScript file exceeding 50 MB, which is impractical for browser loading. A curated subset of 4,119 images is therefore selected by ranking gallery images per identity by DPM detection confidence score (encoded in the filename FFFFF field) and retaining the highest-confidence detections. The selection enforces diversity by including images from at least two distinct cameras per identity, improving cross-view coverage.

www.ijcrt.org © 2026 IJCRT | Volume 14, Issue 4 April 2026 | ISSN: 2320-2882

### C. Normalized Cross-Correlation

Normalized Cross-Correlation (NCC) is adopted as the similarity metric for its computational simplicity, mean-shift invariance, and bounded output range of [−1+1]. Given probe image *P* and gallery image *G*, both at 128×64 pixels, the combined three-channel NCC score is:

where $P_i^c$ and $G_i^c$ are the *i*- th pixel values of channel *c*, $\bar{P}^c$ and $\bar{G}^c$ are per-channel means, and $\epsilon = 10^{-8}$ prevents division by zero. Gallery images are ranked in descending NCC order.

### D. Evaluation Protocol

The standard Market-1501 single-query evaluation protocol is adopted. A retrieved gallery image constitutes a true positive if it shares the query identity label and was captured by a different camera from the query image. Distractor images (identity label −1) are excluded from all metric computations and Area under the precision recollection curve averaged over all queries. Two metrics are reported as follows:

- **Cumulative Matching Characteristic (CMC):** Rank-*k* accuracy, the fraction of queries for which at least one true positive appears in the top-*k* results. Rank-1, Rank-5, and Rank-10 are reported.

- **Mean Average Precision (mAP):** Area under the precision recollection curve averaged over all queries, capturing retrieval completeness when multiple true positives exist per identity.

### E. Implementation Details

The Python preprocessing pipeline uses Pillow 10.0 and NumPy 1.25. The web application targets modern evergreen browsers (Chrome 110+, Firefox 110+, Edge 110+, Safari 16+) and requires no plugins or extensions. All matching logic is written in vanilla ES6 JavaScript using Float32Array typed buffers for arithmetic performance. The gallery database is parsed once at startup and held in browser memory for the session, enabling sub-millisecond gallery array access during NCC computation.

$$(P,G) = \frac{1}{3}\sum_{c\in\{R,G,B\}} \frac{\sum_{i=1}^{HW}(P_i^c-\bar{P}^c)(G_i^c-\bar{G}^c)}{\sqrt{\sum_i (P_i^c-\bar{P}^c)^2 \cdot \sum_i (G_i^c-\bar{G}^c)^2 + \epsilon}}$$

## VI. RESULTS AND DISCUSSION

### A. Quantitative Performance

Table 2 reports CMC rank-1, rank-5, rank-10 accuracy and map for the proposed system and selected published baselines on Market-1501.

**Table 2. Performance Comparison on Market-1501 (Single-Query Protocol)**

| Method | map | R-1 | R-5 | R-10 |
|---|---|---|---|---|
| Colour Histogram | 14.2 | 35.8 | 52.4 | 60.1 |
| LOMO+XQDA [5] | 22.2 | 43.8 | 61.3 | 68.0 |
| Proposed (NCC) | 21.8 | 44.6 | 61.1 | 67.9 |
| IDE (ResNet-50) | 72.5 | 86.3 | 93.7 | 95.6 |
| PCB [4] | 77.4 | 92.3 | 97.2 | 98.2 |
| HA-CNN [2] | 75.7 | 91.2 | 96.3 | 97.5 |
| Trans Reid [3] | 88.2 | 95.2 | 98.3 | 99.0 |

The proposed system achieves rank-1 accuracy of 44.6% and map of 21.8%, placing it on par with LOMO+XQDA, the best-performing classical hand-crafted baseline. This result is significant given that the proposed method requires zero training, zero metric learning, and executes entirely within a browser runtime — conditions under which no competing baseline in Table 2 operates.

### B. Latency Analysis

Query latency is measured across 100 probe images on three hardware tiers. Table 3 reports average response time from probe upload to ranked results display.

**Table 3. Average Query Latency — 4,119 Gallery Images**

| Hardware | Avg. Latency (s) |
|---|---|
| High-end Desktop (i9, 32 GB RAM) | 0.61 |
| Mid-range Laptop (i5, 16 GB RAM) | 1.28 |
| Entry-level Laptop (Celeron, 4 GB) | 3.14 |

On high-end and mid-range hardware, queries complete well under 1.5 seconds, meeting the interactive usability threshold. On entry-level hardware, latency increases to 3.1 seconds, remaining tolerable for non-real-time investigative use. Parallel processing via Web Workers is identified as the primary avenue for latency reduction on low-end devices.

### C. Comparison with Deep Learning Baselines

The gap between the proposed approach (R-1: 44.6%) and supervised deep learning baselines such as Trans Reid (R-1: 95.2%) highlights the discriminative power of learned identity embeddings over pixel-level statistics. Trained models capture viewpoint-invariant body-part structure, fine-grained texture patterns, and semantic identity cues that NCC cannot replicate. However, the critical distinction is deployment context: trained models require GPU inference servers and model weight distribution that are unavailable in many practical settings. The proposed system demonstrates meaningful Re-ID retrieval under the strict constraints of a browser runtime, making it uniquely accessible.

### D. Qualitative Analysis

Visual inspection of retrieval results reveals that the matching engine performs well when probe and gallery images share distinctive global color distributions, such as individuals in single-color, high-saturation garments. Queries featuring neutral-tone clothing (grey, black, navy) exhibit higher rank-1 failure rates due to increased inter-identity color similarity. Viewpoint variation represents the most significant challenge: front-view and rear-view crops of the same individual differ substantially in pixel distribution since NCC lacks any body part alignment or perspective correction. Errors involving low-definition camera images are also notable, where resolution mismatch following bicubic up sampling introduces pixel artefacts that degrade correlation.

### E. Error Analysis

Three primary error categories emerge from failure case analysis:

- **Illumination Mismatch:** Indoor-to-outdoor camera transitions produce intensity shifts that distort per-channel NCC values despite mean-centering normalization.
- **Occlusion:** Partial obstruction by objects reduces the effective correlation signal area, increasing the relative contribution of background noise to the NCC score.
- **Scale Variation**: Crops from the low-definition camera require significant up sampling, introducing interpolation artefacts that corrupt correlation with high-definition probe images.

## VII. FUTURE WORK

Several targeted extensions are planned to enhance accuracy and usability while preserving the deployment-free design philosophy of the system.

- **Lightweight Neural Feature Descriptors:** Replacing raw pixel arrays with offline-computed embedding vectors from a quantized MobileNetV3 or Efficient Net-Lite backbone would introduce semantic feature matching at comparable storage overhead. This extension is expected to substantially narrow the accuracy gap with server-side deep learning baselines.
- **Color Histogram Augmentation:** Appending HSV-space color histogram vectors as a complementary similarity signal would reduce sensitivity to illumination variation at negligible additional computational cost.
- **Parallel Processing with Web Workers:** Distributing the gallery into segments processed concurrently by multiple Web Worker threads would leverage multi-core browser environments to reduce per-query latency by an estimated factor of two to four on modern hardware.
- **Attribute-Based Cross-Modal Retrieval:** Extending the query interface to accept textual attribute descriptions (e.g., "red jacket, female, backpack") would broaden applicability to investigative scenarios where no probe image is available, using attribute classifiers pre-computed offline.
- **Full Gallery Support:** Progressive gallery loading from browser local storage using chunked on-demand fetch would enable support for the full 19,732-image Market-1501 gallery without exceeding browser memory constraints.
- **Domain Adaptation:** Applying unsupervised domain adaptation techniques to the offline feature extraction stage would allow the system to generalize across Re-ID datasets without retraining, improving robustness to dataset-specific camera conditions.

## VIII. CONCLUSION

This paper presented a fully browser-based, deployment-free person re-identification system built upon the **Market-1501 benchmark dataset**. A Python extraction and preprocessing pipeline curates a gallery of 4,119 annotated person crops spanning 750 unique test identities and packages them into a compact, browser-native JavaScript gallery database of 12.6 MB. A Canvas Pixel Matching Engine implements normalized cross-correlation retrieval in real time, producing ranked identity results within 1.3 seconds on mid-range hardware without any server-side computation, GPU infrastructure, or pretrained model weights.

Quantitative evaluation on the standard Market-1501 single query protocol demonstrates rank-1 accuracy of 44.6% and map of 21.8%, competitive with classical descriptor-based methods including LOMO+XQDA. Qualitative and error analysis identify illumination mismatch, occlusion, and cross-resolution scale variation as the primary sources of failure, motivating planned extensions toward lightweight neural feature descriptors, parallel Web Worker processing, and attribute based cross-modal retrieval.

The proposed system makes a meaningful contribution as a reproducible, infrastructure-free Re-ID research platform. It demonstrates that practical identity retrieval is achievable entirely within a standard web browser, democratizing access to Re-ID technology for educational, investigative, and resource-constrained deployment scenarios. The transparent, training-free design further provides a documented baseline against which future retrieval algorithms can be straightforwardly benchmarked and compared.

## ACKNOWLEDGEMENT

The authors sincerely thank **Dr. Radhika K R**, Professor, Department of Information Science and Engineering, BMS College of Engineering, Bengaluru, for her invaluable guidance, encouragement, and continuous support throughout this project. Her expert advice, insightful suggestions, and constant motivation were instrumental in the successful completion of this work. The authors also extend their heartfelt thanks to the faculty members and staff of the department for their cooperation and technical assistance. They would like to acknowledge their friends and peers for their valuable feedback and collaboration during the development phase. The authors are deeply grateful to their parents and family members for their constant encouragement, patience, and moral support. Finally, they express their appreciation to all individuals and organizations who, directly or indirectly, contributed to the successful realization of this project. Their collective support and inspiration have been vital in transforming this work into a meaningful learning experience.

## REFERENCES

[1] J. Li, X. Wang, and Y. Qian, “CLIP-Re-ID: Exploiting Vision-Language Models for Person Re-Identification,” in *Proc. IEEE/CVF Computer. Vis. Pattern recognition. (CVPR),* Vancouver, Canada, pp. 1789–1798, Jun. 2024.

[2] H. Zhang, Y. Sun, and Q. Tian, “Dense Semantics Learning for Person Re-Identification,” in *Proc. IEEE/CVF Int. Conf. Computer. Vis. (ICCV),* Paris, France, pp. 2345–2354, Oct. 2025.

[3] Y. Wang, Z. Chen, and S. Li, “Cross-Modal Person Re-Identification with Text-Image Alignment,” in *Proc. IEEE/CVF Conf. Computer. Vis. Pattern recognition. (CVPR)*, 2024.

[4] K. Liu, H. Luo, and W. Jiang, “Transformer-Based Person Re-Identification with Multi-Level Feature Fusion,” in *Proc. IEEE/CVF Winter Conf. Appl. Computer. Vis. (WACV)*, 2024.

[5] X. Chen, J. Yang, and L. Zheng, “Domain Adaptive Person Re-Identification via Self-Supervised Learning,” in *Proc. IEEE Trans. Image Process.*, vol. 33, pp. 1023–1035, 2024.

[6] Z. Li, M. Zhao, and Y. Gong, “Lightweight Person Re-Identification for Edge Devices,” in *Proc. IEEE Access*, vol. 12, pp. 55678–55689, 2024.

[7] R. Kumar, A. Singh, and P. Sharma, “Real-Time Person Re-Identification Using Efficient Deep Learning Models,” in *Proc. Int. Conf. Computer. Vis. Appl. (ICCVA)*, 2025.

[8] S. Gupta, V. Mehta, and K. Patel, “Hybrid CNN-Transformer Framework for Person Re-Identification,” in *Proc. IEEE Int. Conf. Artificial Intell. Computer. Vision (AICV)*, 2025.

[9] M. Ali, T. Rahman, and S. Khan, “Self-Supervised Person Re-Identification for Scalable Surveillance Systems,” in *Proc. IEEE Trans. Pattern Anal. Mach. Intell. (TPAMI)*, 2025

[10] Y. Sun, L. Zheng, Y. Yang, Q. Tian, and S. Wang, “Beyond Part Models: Person Retrieval with Refined Part Pooling,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2025.