IJIREEICE

ISSN (O) 2321-2004, ISSN (P) 2321-5526

International Journal of Innovative Research in Electrical, Electronics, Instrumentation and Control Engineering

Impact Factor 8.021 ⪤ Peer-reviewed & Refereed journal ⪤ Vol. 14, Issue 8, August 2026

DOI: 10.17148/IJIREEICE.2026.14810

Illumination-Adaptive Multi-Branch Feature Fusion for Robust Person Re-Identification

R Sivani[1], Rakshitha S N[2], Shreya Sathapathi[3], Charulatha R.T[4]

B.Tech Student, Department of CSE, SRM Institute of Science and Technology, Chennai, India[1-3]

Guide, Department of CSE, SRM Institute of Science and Technology, Chennai, India[4]

Abstract: Person re-identification (Re-ID) across camera networks remains sensitive to illumination variation, since a single global appearance encoding is rarely robust to both bright and poorly lit capture conditions. This paper presents an adaptive illumination-robust Re-ID pipeline that estimates a continuous per-crop illumination score, generates four parallel enhanced views of each detected person (raw, CLAHE, gamma-corrected, and multi-scale Retinex), extracts an OSNet embedding for each view, and fuses the four embeddings using illumination-conditioned weights before performing cosine-similarity gallery retrieval via FAISS. The system is evaluated on Market-1501 under a stratified protocol comparing a Normal split against a synthetically generated Lowlight split, and comparing the proposed adaptive fusion against a static/average-fusion ablation baseline. Results show Rank-1 accuracy saturated at 98-99.6% across all four conditions, while mean Average Precision (mAP) is more discriminative, falling from approximately 71% on the Normal split to 44-46% on the Lowlight split. Contrary to the design hypothesis, the heuristic adaptive fusion does not outperform static average fusion in this evaluation, and static fusion is marginally better under Lowlight conditions. These findings are presented transparently as evidence that the current rule-based fusion weighting requires learning-based replacement, motivating a concrete direction for future work rather than an unqualified robustness claim.

Keywords: person re-identification, illumination robustness, adaptive feature fusion, OSNet, YOLOv8, Retinex, CLAHE, FAISS, Market-1501

I. INTRODUCTION

Person re-identification (Re-ID) is the task of matching a detected person across non-overlapping camera views by comparing appearance-based feature embeddings, and it underpins applications such as surveillance, retail analytics, smart-city pedestrian tracking, and missing-person search. A persistent challenge in real-world Re-ID deployments is illumination variation: the same individual can appear markedly different in brightness, contrast, and color balance depending on the camera, time of day, and local lighting, which degrades the reliability of a single fixed feature representation.

Existing approaches to illumination robustness in Re-ID broadly fall into two categories: illumination-invariant representation learning, which attempts to learn features that are inherently insensitive to lighting changes, and image-level enhancement, which pre-processes the input to normalize illumination before feature extraction. This work adopts the second strategy and extends it with an adaptive, illumination-conditioned fusion step, on the hypothesis that different enhancement techniques compensate best for different lighting regimes and that combining them intelligently should outperform any single fixed enhancement or a naive average of all of them.

The system described here is deliberately framed as “illumination-robust” rather than “day/night” Re-ID: the primary evaluation dataset, Market-1501, exhibits only moderate inter-camera illumination variation and is not a dedicated nighttime benchmark. A synthetic low-light evaluation split was constructed specifically to provide meaningful illumination-stratified evidence in the absence of real nighttime captures of the same identities, and this limitation is stated explicitly rather than implied away.

The contributions of this work are:

- A 7-stage modular Re-ID pipeline combining YOLOv8-based detection, continuous LAB-based illumination scoring, four-branch parallel enhancement (Raw, CLAHE, Gamma, multi-scale Retinex), OSNet embedding extraction, illumination-conditioned adaptive fusion, and FAISS-backed gallery retrieval with confidence thresholding.
- An illumination-stratified evaluation methodology (Normal vs. synthetic Lowlight splits, each tested under Adaptive vs. Static/Average fusion) that is reusable for testing robustness claims even without access to real paired day/night data.

IJIREEICE

ISSN (O) 2321-2004, ISSN (P) 2321-5526

International Journal of Innovative Research in Electrical, Electronics, Instrumentation and Control Engineering

Impact Factor 8.021 ⁒ Peer-reviewed & Refereed journal ⁒ Vol. 14, Issue 8, August 2026

DOI: 10.17148/IJIREEICE.2026.14810

A transparent, honestly reported empirical result: the heuristic adaptive fusion mechanism, as currently tuned, does not measurably outperform simple static averaging, which is analyzed and used to motivate concrete future work rather than being obscured.

II. RELATED WORK

Illumination-invariant Re-ID has been approached through several complementary directions in the literature. Retinex-based methods decompose an image into illumination and reflectance components and use the illumination-invariant reflectance component for matching. Illumination-identity disentanglement approaches explicitly separate identity-relevant appearance features from lighting-dependent nuisance factors during representation learning, typically using adversarial or generative objectives. Illumination-unification methods instead normalize all gallery and query images toward a canonical lighting condition prior to feature extraction, which is conceptually closest to the multi-branch enhancement strategy used in this work. Separately, dedicated nighttime Re-ID benchmarks and methods (e.g., illumination distillation frameworks evaluated on nighttime-specific datasets) exist precisely because standard daytime benchmarks such as Market-1501 do not capture extreme low-light degradation, which reinforces the decision in this work to construct a synthetic low-light split rather than overclaim day/night generalization from Market-1501 alone.
Relative to this body of work, the present system does not propose a new learned illumination-invariant representation; instead it combines four established, individually well-studied enhancement techniques (CLAHE, gamma correction, multi-scale Retinex, and the unmodified raw image) at the feature level via an illumination-conditioned fusion step, and evaluates that specific design choice empirically rather than assuming it is beneficial.

III. PROPOSED SYSTEM ARCHITECTURE

The system follows a 7-stage sequential pipeline, executed per input frame or image, as summarized in Table I.

Stage	Function	Description
1. Detection	Person localization	YOLOv8 (nano variant, yolov8n.pt) detects people in the frame; each bounding box is cropped and resized to 128×256, the standard Re-ID aspect ratio.
2. Illumination Analysis	Continuous lighting estimate	The crop is converted to LAB color space; L-channel mean gives brightness and pixel standard deviation gives contrast, combined into a continuous illumination_score ∈ [0,1]. A coarse Day/Evening/Night label is derived only for display.
3. Multi-Branch Enhancement	Illumination compensation	Four parallel versions of the crop are produced: Raw (unmodified), CLAHE (local contrast boost), Gamma-corrected (power-law brightness adjustment, gamma adaptable to illumination_score), and Retinex (multi-scale Retinex with sigma scales [15, 80, 250]).
4. Feature Extraction	Embedding generation	Each of the four enhanced images is passed independently through a pretrained OSNet backbone, yielding four separate L2-normalized embedding vectors.
5. Adaptive Feature Fusion	Illumination-conditioned combination	The illumination_score selects a weight for each of the four branches via a 3-bucket heuristic rule table; the weighted sum of the four embeddings produces one fused embedding.
6. Gallery Matching	Retrieval	The fused embedding is compared against a gallery of known identity embeddings via cosine similarity, using a FAISS-backed index for efficient nearest-neighbor search; the top-k most similar identities are returned.
7. Confidence Estimation	Match acceptance	A similarity threshold (default 0.7) determines whether the best match counts as a confident identification or an unmatched/unknown result, optionally factoring in the top-1/top-2 margin.

DOI: 10.17148/IJIREEICE.2026.14810

A. Detection and Illumination Scoring

Person detection uses YOLOv8n for its favorable accuracy-to-latency trade-off on consumer hardware. Rather than relying on a discrete, hand-labeled illumination bucket, the pipeline computes a continuous illumination_score directly from LAB color statistics, which is the signal actually consumed by the fusion stage; the Day/Evening/Night label exists purely for human-readable display and does not affect the control logic.

B. Multi-Branch Enhancement and Feature Extraction

Combining four enhancement strategies (Raw, CLAHE, Gamma, Retinex) is intended to provide redundancy: no single correction technique is relied upon to work well across every lighting condition. Each enhanced view is independently encoded by OSNet (osnet_x1_0, loaded via Torchreid), fine-tuned on the 751 identities of the Market-1501 training split (checkpoint best_model.pth.tar), producing four L2-normalized embeddings per detection.

C. Adaptive Feature Fusion

The four branch embeddings are combined into a single fused embedding via a weighted sum, where the per-branch weights are selected from a 3-bucket heuristic rule table conditioned on illumination_score (thresholds at approximately >0.6, 0.3-0.6, and <0.3), with each bucket favoring a different subset of branches. This fusion function is not a trained model; it is a hand-designed placeholder, explicitly documented in the codebase as intended for future replacement by a learned gating network. This is stated plainly because it is directly relevant to interpreting the results in Section VI: the fusion mechanism evaluated here is a heuristic, not a learned component, and the results should be read accordingly.

D. Gallery Matching and Confidence Estimation

The fused embedding is matched against a FAISS-backed gallery index using cosine similarity, returning the top-k candidates. A configurable similarity threshold (default 0.7) separates confident identifications from unmatched results, optionally incorporating the margin between the top-1 and top-2 similarity scores to reduce false positives in ambiguous cases.

IV. MODELS AND TECHNOLOGY STACK

Component	Role	Notes
YOLOv8n	Person detection	Lightweight, real-time-capable variant pretrained on COCO (generic person class).
OSNet (osnet_x1_0)	Feature/embedding extraction	Loaded via Torchreid; fine-tuned on 751 Market-1501 identities. Chosen for its multi-scale omni-scale feature design, suited to fine-grained Re-ID matching.
Fusion weighting function	Combines 4 branch embeddings	Hand-designed heuristic (if/elif rule table on illumination_score buckets), not a trained model; documented placeholder for a future learned gating network.
FAISS	Approximate nearest-neighbor search	Facebook AI Similarity Search library; an efficient vector index, not a learned model.

The system was developed and evaluated on Apple Silicon (M4) using PyTorch with the MPS backend for GPU acceleration, without requiring a dedicated CUDA/NVIDIA GPU, which supports the real-time-capable framing of the detection and embedding stages.

V. EXPERIMENTAL SETUP

A. Dataset

Evaluation uses Market-1501, comprising approximately 19,732 gallery images, 3,368 query images, and 751 identities captured across 6 camera views, following the standard protocol with junk/distractor images excluded. Market-1501 has only moderate inter-camera lighting differences and is not a dedicated day/night Re-ID benchmark, which motivates the synthetic evaluation design below rather than an unsupported day/night generalization claim.

International Journal of Innovative Research in Electrical, Electronics, Instrumentation and Control Engineering
Impact Factor 8.021 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 14, Issue 8, August 2026
DOI: 10.17148/IJIREEICE.2026.14810

B. Illumination-Stratified Evaluation Design

To obtain meaningful illumination-stratified evidence in the absence of real paired day/night captures, two test conditions were constructed: a Normal split (the original, unmodified Market-1501 test set) and a Lowlight split (a synthetically darkened copy of the same test set), generated via gamma darkening (gamma tested in the range ≈2.0-3.0), reduced brightness and contrast, and added light Gaussian sensor noise. Both splits were evaluated under two fusion strategies: the proposed Adaptive Fusion, and a Static/Average Fusion ablation baseline using equal weights across all four branches regardless of illumination. This yields a 2×2 stratified design (Normal/Lowlight × Adaptive/Static), reported using Rank-1 accuracy and mAP, the standard Re-ID metrics.

As a sanity check on the illumination score itself, the unmodified Normal split showed illumination scores clustering narrowly around 0.4-0.6 (mostly classified "Evening"), confirming that Market-1501 lacks strong natural illumination diversity and directly supporting the need for the synthetic Lowlight split. The Lowlight split, by contrast, produced scores consistently in the 0.1-0.3 range ("Night"), confirming that the synthetic degradation procedure worked as intended.

VI. RESULTS AND DISCUSSION

Evaluation Split	Fusion Type	Rank-1 Accuracy	mAP
Normal	Adaptive Fusion	99.61%	71.76%
Normal	Static/Average Fusion	99.64%	71.04%
Lowlight	Adaptive Fusion	98.46%	43.54%
Lowlight	Static/Average Fusion	98.57%	45.43%

Table III reports Rank-1 accuracy and mAP across all four evaluated conditions. Three findings are drawn from these results and reported plainly, without overstating the benefit of the proposed adaptive mechanism.

First, Rank-1 accuracy is saturated at 98-99.6% across every condition tested, which makes it a poor discriminator between fusion strategies at this dataset's difficulty level: the retrieval task is simply too easy at Rank-1 to expose meaningful differences. Second, mAP is substantially more informative, dropping from approximately 71% on the Normal split to 44-46% on the Lowlight split for both fusion types, which confirms that the synthetic Lowlight condition genuinely increases task difficulty and validates the stratified evaluation design as a whole. Third, and most importantly, adaptive fusion did not outperform static/average fusion in this evaluation; static fusion was in fact marginally better on both Rank-1 and mAP under Lowlight conditions (mAP: 45.43% static vs. 43.54% adaptive).

The root cause is attributed to the fusion weighting function being a hand-designed heuristic rather than a learned mechanism: this evaluation shows that the heuristic, as currently tuned, provides no measurable advantage over naive averaging. This is presented as a finding rather than a failure, and it motivates a specific, testable direction for future work — replacing the rule-based weight table with a learned gating network — rather than a vague call for "more work."

VII. LIMITATIONS

- Fusion weighting is heuristic, not learned: hand-picked weight values per illumination bucket, which this evaluation shows do not outperform simple static/average fusion, indicating the current weights are not well-calibrated.
- Dataset lacks real day/night variation: Market-1501 has only moderate inter-camera lighting differences; the Lowlight condition is synthetically generated (gamma darkening plus noise), not a real nighttime capture, so results may not fully transfer to real nighttime surveillance footage.
- Rank-1 metric saturation (98-99.6% across all conditions) makes it a poor discriminator between fusion strategies at this dataset's difficulty level; mAP is more informative but still shows no clear adaptive-fusion advantage.
- Coarse illumination labeling: the Day/Evening/Night classification is a simple threshold-based bucketing of a continuous score, not a learned or validated categorical classifier.
- Demo gallery is not the evaluation gallery: an interactive demo would use one representative embedding per identity (enrollment-style, ≈749 identities), while the formal Rank-1/mAP evaluation used the full multi-image gallery (≈19,732 images); the two are not directly comparable and are discussed separately.
- 4× inference cost: running detection, enhancement, and embedding extraction across four parallel branches multiplies computational cost roughly fourfold relative to a single-branch baseline, with no proven accuracy benefit demonstrated so far.

DOI: 10.17148/IJIREEICE.2026.14810

- No cross-dataset validation: the system has only been evaluated on Market-1501; generalization to other Re-ID datasets or real deployment scenarios remains untested.

VIII. FUTURE WORK

- Replace the heuristic fusion with a learned gating network — a small trainable MLP or attention mechanism that learns optimal per-branch weights from data, conditioned on illumination_score and/or image features directly, rather than fixed rule-based buckets.
- Evaluate on real nighttime Re-ID datasets (e.g., Night600, NightReID, or similar benchmarks with genuine low-light captures) to test whether the current findings hold beyond synthetic degradation.
- Extend the interactive demo to a multi-image gallery per identity, matching the formal evaluation protocol so live demo results are more representative of true Re-ID performance.
- Explore efficiency optimization by learning to select or weight only the one or two most relevant enhancement branches per image instead of always computing all four.
- Run an ablation isolating which of CLAHE, Gamma, Retinex, and Raw contributes most under which lighting condition, to potentially simplify the architecture.
- Test cross-dataset generalization on additional Re-ID benchmarks (e.g., DukeMTMC-reID, CUHK03, MSMT17) beyond Market-1501.

IX. APPLICATIONS

- Surveillance and security systems, re-identifying persons of interest across multiple cameras with varying lighting conditions such as entrances, parking lots, and mixed indoor/outdoor corridors.
- Retail analytics, tracking customer movement patterns across a store's camera network without relying on facial recognition.
- Smart-city and public-safety infrastructure, supporting cross-camera pedestrian tracking across city-wide CCTV networks operating under varying times of day.
- Missing-person and forensic investigation support, searching archived camera footage for a specific individual across different lighting conditions and camera angles.
- Access control systems, as a supplementary identity-verification layer combined with other biometric methods.

Sports and crowd analytics, tracking specific individuals across broadcast or venue camera feeds under variable stadium or arena lighting.

X. CONCLUSION

This paper presented an adaptive illumination-robust person re-identification pipeline combining YOLOv8 detection, continuous LAB-based illumination scoring, four-branch parallel enhancement, OSNet feature extraction, illumination-conditioned fusion, and FAISS-backed gallery retrieval. An illumination-stratified evaluation on Market-1501, comparing Normal and synthetic Lowlight splits under Adaptive and Static/Average fusion, showed that Rank-1 accuracy is saturated and uninformative at this dataset's difficulty level, while mAP confirms the Lowlight condition meaningfully increases task difficulty. Critically, the proposed adaptive fusion did not outperform simple static averaging, and this result is reported transparently rather than minimized. The modular, interpretable design of the pipeline and the reusable illumination-stratified evaluation methodology remain useful contributions independent of this specific finding, and the identified gap between heuristic and learned fusion motivates a concrete, testable next step: replacing the rule-based weight table with a learned gating network and validating on real nighttime Re-ID data.

ACKNOWLEDGMENT

The authors extend sincere gratitude to **Ms. Charulatha R.T**, Project Guide, Department of Computing Technologies, SRM Institute of Science and Technology, for her sustained mentorship, technical direction, and constructive feedback throughout every phase of this project's development and documentation.

REFERENCES

[1]. Y. Huang, Z.-J. Zha, X. Fu, and W. Zhang, "Illumination-Invariant Person Re-Identification", Proc. 27th ACM Int. Conf. Multimedia (MM), pp. 365-373, 2019.
[2]. Z. Zeng et al., "Illumination-Adaptive Person Re-identification", IEEE Trans. Multimedia, vol. 22, no. 12, pp. 3064-3074, 2020.

International Journal of Innovative Research in Electrical, Electronics, Instrumentation and Control Engineering

Impact Factor 8.021 ⪤ Peer-reviewed & Refereed journal ⪤ Vol. 14, Issue 8, August 2026

DOI: 10.17148/IJIREEICE.2026.14810

[3]. Z. Zhong, L. Zheng, Z. Zheng, S. Li, and Y. Yang, "Camera Style Adaptation for Person Re-identification", Proc. IEEE Conf. Computer Vision and Pattern Recognition (CVPR), pp. 5157-5166, 2018.

[4]. A. Lu et al., "Illumination Distillation Framework for Nighttime Person Re-Identification and a New Benchmark", IEEE Trans. Multimedia, vol. 26, pp. 406-419, 2023.

[5]. K. Zhou, Y. Yang, A. Cavallaro, and T. Xiang, "Omni-Scale Feature Learning for Person Re-Identification", Proc. IEEE/CVF Int. Conf. Computer Vision (ICCV), pp. 3701-3711, 2019.

[6]. M. Yaseen, "What is YOLOv8: An In-Depth Exploration of the Internal Features of the Next-Generation Object Detector", arXiv preprint arXiv:2408.15857, 2024.

[7]. M. Yaseen, "What is YOLOv8: An In-Depth Exploration of the Internal Features of the Next-Generation Object Detector", arXiv preprint arXiv:2408.15857, 2024.

[8]. J. Johnson, M. Douze, and H. Jégou, "Billion-Scale Similarity Search with GPUs", IEEE Trans. Big Data, vol. 7, no. 3, pp. 535-547, 2017.