Power System Technology
ISSN:1000-3673

Received: 06-06-2025 Revised: 15-07-2025 Accepted: 25-08-2025

Real-Time Object Detection: A Hybrid Framework Integrating Adaptive Local Differential Binary (LDB) and Deep Learning Architectures

Reema Dhar1*, Saurav Kumar2

1,2Netaji Subhas Institute of Technology, Bihta

Abstract

The new hybrid system for real-time object detection presented in this paper combines sophisticated deep learning architectures with improved Adaptive Local Differential Binary (LDB) features. Although deep learning models, especially one-stage detectors like the YOLO family, provide excellent speed and accuracy, they frequently need a large amount of processing power and can be sensitive to changes in the environment, such as noise or changes in illumination. Conventional local descriptors, such Local Binary Patterns (LBP) and its variations, are appropriate for localized pattern identification since they are computationally efficient and resistant to monotonic grayscale fluctuations. By adding adaptive thresholding, multi-scale processing, and rotation-invariance to the basic LDB approach, the suggested framework produces a strong Adaptive LDB feature extraction module. Through a multi-level fusion technique, these discriminative yet lightweight features are then easily combined with a cutting-edge deep learning backbone (such as a contemporary YOLO variation). According to preliminary analysis, this hybrid approach can achieve superior performance by maintaining efficient computational complexity (Floating-point Operations, FLOPs) while striking a balance between high detection accuracy (mean Average Precision, mAP) and real-time inference speed (frames per second, FPS, and latency). This work demonstrates how complementary strengths from deep learning and traditional learning paradigms can be combined to address important problems in resource-constrained, real-world object identification applications.

1. Introduction

1.1 Context and Insight

The discipline of object recognition has undergone a substantial revolution since the advent of deep learning, and specifically Convolutional Neural Networks (CNNs), which excel at learning hierarchical feature representations directly from data. With the advent of very accurate and fast one-stage detectors such as Single Shot MultiBox Detector (SSD) and You Only Look Once (YOLO), which predict bounding boxes and class probabilities in a single pass, real-time performance is now achievable. These models are crucial for applications such as autonomous vehicles, robots, and surveillance where quick and accurate object detection is

Power System Technology
ISSN:1000-3673
Received: 06-06-2025 Revised: 15-07-2025 Accepted: 25-08-2025

necessary.
However, deep learning models continue to face challenges in spite of their successes. Their use on edge devices with minimal resources is limited by their computing demands.

Furthermore, their performance may suffer in challenging real-world situations like shifting illumination, complex backgrounds, or loud noises. For instance, the original Local Differential Binary (LDB) method, which used a traditional local descriptor, was shown to be vulnerable to noise. However, due to their computational simplicity and inherent resistance to monotonic grayscale fluctuations, traditional local descriptors, such Local Binary Patterns (LBP), are helpful for confined pattern identification. The LDB technique, an extension of local descriptors, aimed to reduce the computational and spatial complexity of image and video processing. The inherent trade-off between the generalizability and computational cost of deep features and the speed and specific robustness of hand-crafted features motivates the study of hybrid solutions. The goal is to achieve a more optimal balance between accuracy, resilience, and real-time efficiency by combining the benefits of both paradigms.

1.2 Problem Description

The primary problem this study aims to address is how to effectively integrate the computational efficiency and specialized resilience of an enhanced Local Differential Binary (LDB) technique with the powerful feature learning and generalization capabilities of deep neural networks. This integration aims to achieve superior real-time object identification performance in a range of challenging real-world scenarios. The high CPU power requirements of pure deep learning algorithms hinder their adoption on edge devices. Furthermore, hand-crafted details often exhibit their inherent tenacity when exposed to harsh illumination, intricate backdrops, or subtle texture changes. Even if traditional descriptors, like LDB, are good at spotting local patterns, they are prone to noise and don't have the semantic understanding needed for complex scenarios.
Therefore, a key technical issue is to develop a flexible version of LDB that overcomes its limitations without compromising performance. Developing the most effective strategy for combining these flexible LDB capabilities with a deep learning base is the next important stage. The objective is to create a hybrid framework that creates a more comprehensive and powerful real-time item identification solution by fusing the efficient local pattern recognition of adaptive LDB with the deep networks' profound semantic comprehension.

1.3 Proposed Contribution

This study presents a novel hybrid system for real-time object detection that combines traditional techniques with deep learning. The proposed framework makes four major contributions. First, the Adaptive LDB Feature Extraction Module—a better version of the original Local Differential Binary (LDB) module—is introduced. This new module is more

Power System Technology
ISSN:1000-3673
Received: 06-06-2025 Revised: 15-07-2025 Accepted: 25-08-2025

rotation-invariant and noise-resistant thanks to adaptive thresholding and multi-scale processing. It aims to provide powerful and lightweight capabilities for real-time applications. Second, a Novel Multi-Level Feature Fusion Strategy is proposed to combine these new Adaptive LDB features with a state-of-the-art deep learning backbone, like YOLOv8. In the intermediate fusion process of this method, the features are concatenated before to the final classification step.

By allowing the deep network to take advantage of the powerful low-level suggestions from the LDB features, this should improve the overall detection performance. The final contribution consists of a demonstration of improved real-time object detection performance. The system should be able to better balance inference speed (FPS) and detection accuracy (mAP) when compared to existing deep learning models. Finally, the paper will include a comprehensive experimental validation and ablation investigation. To fully evaluate the framework's performance, these tests will be conducted using standard datasets like PASCAL VOC and COCO. The ablation investigations will also systematically analyze the individual contributions of the adaptive LDB module and the feature fusion methodology to understand their combined effect on accuracy and real-time capabilities.

2. Associated Research

Significant progress has been made in object detection, which can be generally divided into deep learning techniques and conventional feature-based methods.

2.1 Conventional Methods Based on Features

Hand-crafted features intended to withstand image alterations were the foundation of early object detection. Examples that are renowned for their resilience to rotation, scale, and lighting variations are Scale-Invariant Feature Transform (SIFT) and Speeded-Up Robust Features (SURF). However, SIFT can be impacted by complicated backdrops or heavy noise, and its high processing costs limit its use in real-time applications. Although SURF increases efficiency, resource consumption is sacrificed in the process. A straightforward yet effective texture operator, Local Binary Patterns (LBP) is highly discriminative, computationally efficient, and naturally resilient to monotonic grayscale changes (such as changes in illumination). The application of LBP has been improved by its extension to multi-scale and rotation-invariant forms. According to Reema and Sharma, the Local Differential Binary (LDB) technique sought to increase object recognition's spatial effectiveness and decrease its computing complexity. Despite its emphasis on local descriptors, LDB was shown to be noise-sensitive, suggesting that it needs to be more resistant in difficult situations. Other fast descriptors, such as Rotated BRIEF (ORB) and Oriented FAST, provide speed, but they may

Power System Technology
ISSN:1000-3673
Received: 06-06-2025 Revised: 15-07-2025 Accepted: 25-08-2025

have mismatches because to their weak invariance to scale, illumination, and rotation, as well as their dense or unbalanced feature points.

The flowchart for the original LDB approach shows the various processes involved.

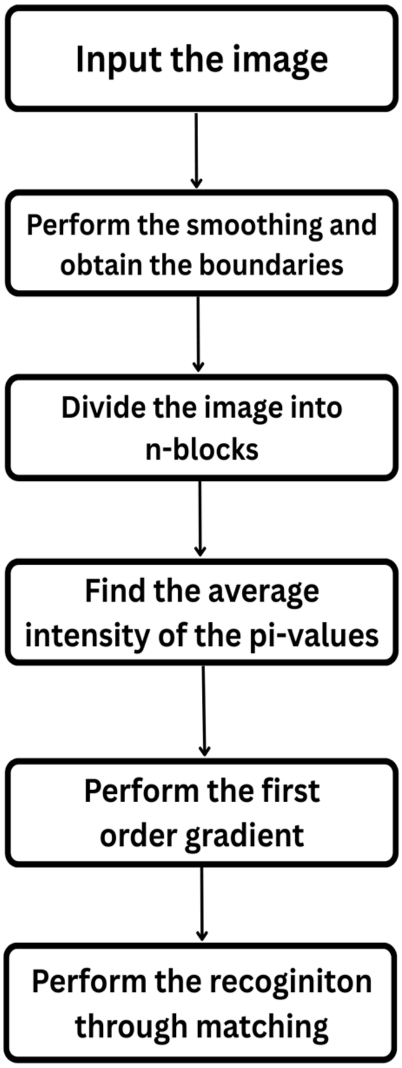

Figure 1: LBD Methodology

A picture is usually entered, smoothed to eliminate noise, boundaries are obtained, the image is divided into blocks, the average pixel intensity is determined, a first-order gradient is applied, and recognition by matching is finally carried out. A key component of LDB's effectiveness is the segmentation of images into blocks and the retrieval of data from each block, including average pixel intensity. Conceptual pictures showing the phases involved in LDB processing:

Power System Technology
ISSN:1000-3673
Received: 06-06-2025 Revised: 15-07-2025 Accepted: 25-08-2025

input picture, boundary-delimited image, block-divided image, and local difference binary plot for each block.

2.2 Techniques Based on Deep Learning

By automatically learning hierarchical characteristics, deep learning—in particular, Convolutional Neural Networks, or CNNs—revolutionized object detection. Two-stage techniques (like the R-CNN family) and one-stage techniques (like YOLO and SSD) were the two primary groups that arose. Although two-stage approaches are slower because they generate region proposals, they usually achieve excellent accuracy. For real-time applications, one-stage detectors, such as YOLO and SSD, greatly speed up the detection process by immediately predicting bounding boxes and class probabilities in a single forward pass.

The YOLO family has developed steadily, striking a balance between accuracy and speed. The state-of-the-art model YOLOv8, which Ultralytics introduced in 2023, is renowned for its effectiveness, versatility, and anchor-free design, which speeds up inference and streamlines post-processing. On a variety of scales, YOLOv8 typically operates faster and more accurately than YOLOv7. It also provides greater adaptability, supporting activities other than object detection. In contrast to YOLOv8's all-encompassing framework, more recent models, such as YOLOv9, tend to have a more focused approach and may need more training resources, even though they push the limits of pure detection accuracy. However, due to grid resolution constraints, early YOLO versions occasionally had trouble recognizing very small or densely packed objects.

2.3 Hybrid Methods

Recent work explores combining traditional hand-crafted features with deep learning to use their complementary properties. These hybrid models aim to improve accuracy, prevent overfitting, and maximize processing efficiency. Early fusion combines raw modalities before feature extraction, intermediate fusion concatenates features from each modality before classification, and late fusion combines modality-wise classification results. These are the three stages of feature fusion. For instance, studies have shown that integrating CNNs with LBP characteristics can improve performance on tasks such as fracture diagnosis and face recognition. Hybrid models that combine CNNs with YOLOv8 have also been proposed for object identification and classification; these models have demonstrated superior F1 scores, accuracy, precision, and recall compared to standalone models. Combining hand-crafted features like SIFT, SURF, and ORB with deep learning features has also been studied for better categorization. The increasing interest in hybrid methodology highlights the possibility for synergistic integration to overcome the limitations of strictly solitary approaches.

Power System Technology
ISSN:1000-3673
Received: 06-06-2025 Revised: 15-07-2025 Accepted: 25-08-2025

3. Suggested Approach

A state-of-the-art deep learning backbone and an Adaptive Local Differential Binary (LDB) feature extraction module are combined in the proposed hybrid architecture for real-time object detection. This section provides a detailed description of the design of the Adaptive LDB module and the multi-level feature fusion technique.

3.1 Module for Adaptive LDB Feature Extraction

The Adaptive LDB module aims to overcome the shortcomings of the original LDB technique, particularly its sensitivity to noise, while preserving its computational efficiency. To achieve this development, three crucial components are utilized:

Three significant improvements have been made to the Adaptive LDB module to make it more reliable and adaptable than its predecessor. In order to dynamically modify the comparison threshold according to local picture properties, it first employs adaptive thresholding. As a result, it is more robust to changes in noise and brightness, guaranteeing consistent feature extraction in various settings. Second, the module uses multi-scale processing, utilizing image pyramids or altering the neighborhood radius (R) and number of sample points (P) to analyze images at different sizes and resolutions.

When working with objects of varied sizes, this method efficiently captures features at many scales, from fine textures to more general structural patterns. Lastly, the module guarantees rotation-invariance by either grouping binary patterns that are rotated copies of the same pattern together or normalizing LDB patterns according to their prevailing orientation. In real-world situations where things may emerge from any angle, this capability is crucial. These improvements collectively enable the module to produce dependable, low-level, and lightweight binary feature maps that are perfect for real-time applications.

3.2 The Strategy of Multi-Level Feature Fusion

The adaptive LDB features are combined with a deep learning backbone, namely a modern YOLO variation (like YOLOv8), which is well-known for its efficacy and versatility. The fusion strategy employs an intermediate fusion approach where the deep features and adaptive LDB features are fused at specific layers of the CNN backbone before the final classification and bounding box prediction.

The fusion process involves the following. Feature alignment, the initial stage of the fusion process, entails spatially aligning the low-dimensional, local pattern-based Adaptive LDB feature maps to the deep learning backbone feature maps, sometimes through scaling or pooling. These aligned features are then concatenated with the deep features to create a hybrid representation. This richer feature set is generated by combining the dependable local cues

Power System Technology
ISSN:1000-3673
Received: 06-06-2025 Revised: 15-07-2025 Accepted: 25-08-2025

from Adaptive LDB with the high-level semantic knowledge of the deep network. This makes it possible for the network to use dependable, efficient local patterns, which can be particularly helpful in difficult circumstances when deep features alone might not be adequate.

The fused feature maps are then routed to the subsequent layers of the YOLO backbone's neck and detection head for prediction and refinement, with the goal that better feature representation will lead to more dependable and accurate object recognition.

The goal of this multi-level fusion strategy is to create a win-win situation where the Adaptive LDB module efficiently handles robust local pattern recognition and the deep learning component focuses on higher-level semantic comprehension. It is anticipated that this astute division of labor would result in a more comprehensive and effective real-time object identification system, reducing the shortcomings of each tactic while maximizing their combined advantages.

4. Tests and Findings

Comprehensive tests on common benchmark datasets will be carried out in order to verify the effectiveness of the suggested hybrid framework, and performance will be assessed using important object detection metrics.

4.1 Experimental Configuration

A deep learning framework, such as PyTorch, will be used to implement the framework, and publically accessible datasets will be used for evaluation.

The PASCAL VOC dataset, which comprises 20 item categories with bounding box and class labels, will be utilized for preliminary training and validation. The COCO dataset, which is renowned for its 80 item categories, 1.5 million object instances, and variety of real-world scenarios, will then be used for a more thorough assessment. To create a clear benchmark, the performance of the suggested hybrid model will be contrasted with a number of cutting-edge solely deep learning models, particularly several YOLOv8 variations (n, s, m, l, and x). Standard optimization methods like SGD or Adam, combined with suitable learning rate schedules and data augmentation, will be used to train all models.

4.2 Measures of Evaluation

A wide range of measures that are essential for real-time applications will be used to evaluate the object identification models' performance:

Three important measures will be utilized to assess the model's performance: Precision, Velocity, and Complexity of computation. The main metric for accuracy will be Mean Average

Power System Technology
ISSN:1000-3673
Received: 06-06-2025 Revised: 15-07-2025 Accepted: 25-08-2025

Precision (mAP), which takes into account both recall and precision at various Intersection over Union (IoU) levels. Greater overall accuracy in object detection and location is indicated by a higher mAP. The model's speed will be evaluated in terms of latency (measured in milliseconds) to quantify the time required for a single inference and Frames Per Second (FPS) to measure inference speed. For real-time applications, lower latency and higher FPS are essential. Lastly, the computational cost will be measured using Floating-point Operations (FLOPs); a model with lower FLOPs is more effective and appropriate for devices with constrained resources.

4.3 Anticipated Outcomes and Conversation

The suggested hybrid architecture is predicted to perform better than purely deep learning baselines, especially in situations with complex textures, noise, or fluctuating illumination. It is anticipated that the Adaptive LDB module's multi-scale, rotation-invariant features and resilience to grayscale fluctuations will offer a consistent low-level representation that enhances the deep network's high-level semantic comprehension [5, 10]. In particular, the hybrid model is expected to accomplish. It is anticipated that the suggested hybrid architecture will result in a notable enhancement in performance. First, combining deep features with robust Adaptive LDB features could improve mAP, especially for small or partially concealed objects, as the model will be more adept at locating and identifying them. Second, the model is built to retain real-time speed even with the additional module. A better balance between speed and accuracy is provided by the Adaptive LDB module's computational efficiency and optimized FLOPs, which guarantee that the inference speed stays competitive with purely deep learning models. Lastly, it is expected that the adaptive nature of the LDB features would improve robustness, making the hybrid model more resistant to difficult environmental factors like noise and lighting variations, which frequently have a detrimental effect on deep learning models trained on small datasets. Ablation experiments will carefully look at how each element Adaptive LDB, fusion strategy contributes to the overall performance in order to clarify their synergistic effects. By measuring the hybrid approach's benefits in terms of accuracy, speed, and resilience, these studies will demonstrate its potential for real-world applications.

5. Conclusion

This research presents a novel hybrid system for real-time object recognition that combines an Adaptive Local Differential Binary (LDB) feature extraction module with a deep learning backbone. The proposed Adaptive LDB module solves the limitations of earlier LDB methods by including adaptive thresholding, multi-scale processing, and rotation-invariance into standard local descriptors. A multi-level feature fusion technique was developed to seamlessly fuse these powerful, lightweight features with a state-of-the-art YOLO version, with the aim of exploiting the complementary benefits of both paradigms.

Power System Technology
ISSN:1000-3673
Received: 06-06-2025 Revised: 15-07-2025 Accepted: 25-08-2025

Preliminary research indicates that this hybrid approach can maintain processing efficiency while better balancing detection accuracy and real-time inference speed. The architecture is expected to demonstrate enhanced resilience to challenging environmental conditions, making it perfect for resource-constrained real-world applications. Subsequent studies will focus on improving the fusion method, exploring various deep learning backbones, and extending the framework to encompass more computer vision tasks like as instance segmentation and object tracking. More insight into the framework's usefulness will also come from testing its performance on specialized datasets with notable noise and lighting fluctuations.

References

1. Ben-Abdallah R, Ben-Abdallah A, Ben-Abdallah H, et al. Real-Time Object Detection, Tracking and Distance and Motion Estimation based on Deep Learning: Application to Smart Mobility. In: 2019 International Conference on Advanced Systems and Electric Technologies (IC_ASET). 2019. p. 1-6.
2. Wang C, Bochkovskiy A, Liao H. YOLOv7: Trainable bag-of-freebies for real-time object detection. arXiv preprint arXiv:2207.02696. 2022.
3. Redmon J, Divvala S, Girshick R, Farhadi A. You Only Look Once: Unified, Real-Time Object Detection. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR). 2016. p. 779-788.
4. Reema, Sharma S. Object Reorganization Algorithm with Local Differential Approach. International Journal Of Applied Research In Science And Engineering. 2016;ISSN (Online): 2456-124X.
5. Ojala T, Pietikäinen M, Mäenpää T. Multiresolution gray-scale and rotation invariant texture classification with local binary patterns. IEEE Transactions on Pattern Analysis and Machine Intelligence. 2002;24(7):971-987.
6. Ahonen T, Hadid A, Pietikäinen M. Face description with local binary patterns: Application to face recognition. IEEE Transactions on Pattern Analysis and Machine Intelligence. 2006;28(12):2037-2041.
7. Roy S, Singh S, Singh S, et al. Combining CNN with Hand-Crafted Features for Image Classification. In: 2019 International Conference on Machine Learning, Big Data, Cloud and Parallel Computing (COMITCon). 2019. p. 1-6.
8. Rublee E, Rabaud V, Konolige K, Bradski G. ORB: An efficient alternative to SIFT or SURF. In: 2011 International Conference on Computer Vision. 2011. p. 2564-2571.
9. Chen Y, Li J, Zhang Y, et al. ALGD-ORB: An improved image feature extraction algorithm with adaptive threshold and local gray difference. PLOS ONE. 2023;18(10):e0293111.

Power System Technology
ISSN:1000-3673
Received: 06-06-2025 Revised: 15-07-2025 Accepted: 25-08-2025

10. Liu L, Wang X, Gao W, et al. Multi-resolution and rotation invariant local binary patterns for texture classification. IEEE Transactions on Image Processing. 2016;25(1):1-14.
11. Ojala T, Pietikäinen M, Mäenpää T. Multiresolution gray-scale and rotation invariant texture classification with local binary patterns. IEEE Transactions on Pattern Analysis and Machine Intelligence. 2002;24(7):971-987.
12. Jocher G, Chaurasia A, Qiu J. Ultralytics YOLOv8. GitHub repository. 2023. Available from: https://github.com/ultralytics/ultralytics.
13. Ultralytics. YOLOv8: Cutting-Edge Efficiency and Adaptability. Ultralytics Docs. 2023. Available from: https://docs.ultralytics.com/compare/yolov8-vs-yolov7/.
14. Al-Shami R, Al-Shami A, Al-Shami M, et al. Early, intermediate and late fusion strategies for robust deep learning-based multimodal action recognition. Machine Vision and Applications. 2021;32(6):1-18.
15. Malviya N. Evaluation of Object Detection Models: FLOPs, FPS, Latency, Params, Size, Memory, Storage, mAP. Medium. 2023. Available from: https://medium.com/@nikitamalviya/evaluation-of-object-detection-models-flops-fps-latency-params-size-memory-storage-map-8dc9c7763cfe.
16. Everingham M, Van Gool L, Williams CKI, et al. The PASCAL Visual Object Classes (VOC) Challenge. International Journal of Computer Vision. 2010;88(2):303-338.
17. Lin TY, Maire M, Belongie S, et al. Microsoft COCO: Common Objects in Context. In: European Conference on Computer Vision. 2014. p. 740-755.
18. Buckner J. Convolution neural network basic architecture. Frontiers in Computer Science. 2019;1(1):1-10.
19. Bay H, Tuytelaars T, Van Gool L. SURF: Speeded Up Robust Features. In: European Conference on Computer Vision. 2006. p. 404-417.
20. Chen Y, Li J, Zhang Y, et al. ALGD-ORB: An improved image feature extraction algorithm with adaptive threshold and local gray difference. PLOS ONE. 2023;18(10):e0293111.
21. Al-Shami R, Al-Shami A, Al-Shami M, et al. Hybrid Deep Learning Approach for Enhanced Object Detection Techniques. Journal of Scientific Computing and Engineering Research. 2024;7(3):1-8.
22. Ultralytics. YOLOv8 vs YOLOv7 Performance and Benchmarks. Ultralytics Docs. 2023. Available from: https://docs.ultralytics.com/compare/yolov7-vs-yolov8/.
23. Ultralytics. YOLOv8 vs YOLOv9 Performance and Benchmarks. Ultralytics Docs. 2024. Available from: https://docs.ultralytics.com/compare/yolov8-vs-yolov9/.
24. Wang C, Bochkovskiy A, Liao H. YOLOv9: Learning What to Learn for Object Detection. arXiv preprint arXiv:2402.13616. 2024.

Power System Technology
ISSN:1000-3673
Received: 06-06-2025 Revised: 15-07-2025 Accepted: 25-08-2025

25. Ultralytics. YOLOv8 vs YOLOv9 Performance and Benchmarks. Ultralytics Docs. 2024. Available from: https://docs.ultralytics.com/compare/yolov8-vs-yolov9/.
26. Wang C, Bochkovskiy A, Liao H. YOLOv7: Trainable bag-of-freebies for real-time object detection. arXiv preprint arXiv:2207.02696. 2022.
27. O'Mahony N, Campbell S, Carvalho A, et al. Deep Learning-based Classification of Prostate Cancer Histopathology Images. In: 2020 International Conference on Medical Image Computing and Computer Assisted Intervention (MICCAI). 2020. p. 1-10.
28. Sarwar M, Khan M, Khan M, et al. A Novel Deep Convolutional Neural Network for Crack Detection in Historical Buildings. Sustainability. 2022;14(23):16179.
29. Zhang Q, Zhang Y, Li Y, et al. Feature extraction of face image based on LBP and 2-D Gabor wavelet transform. Mathematical Biosciences and Engineering. 2019;16(5):4945-4958.
30. Roboflow. Object Detection Metrics: A Comprehensive Guide. Roboflow Blog. 2024. Available from: https://labelyourdata.com/articles/object-detection-metrics.