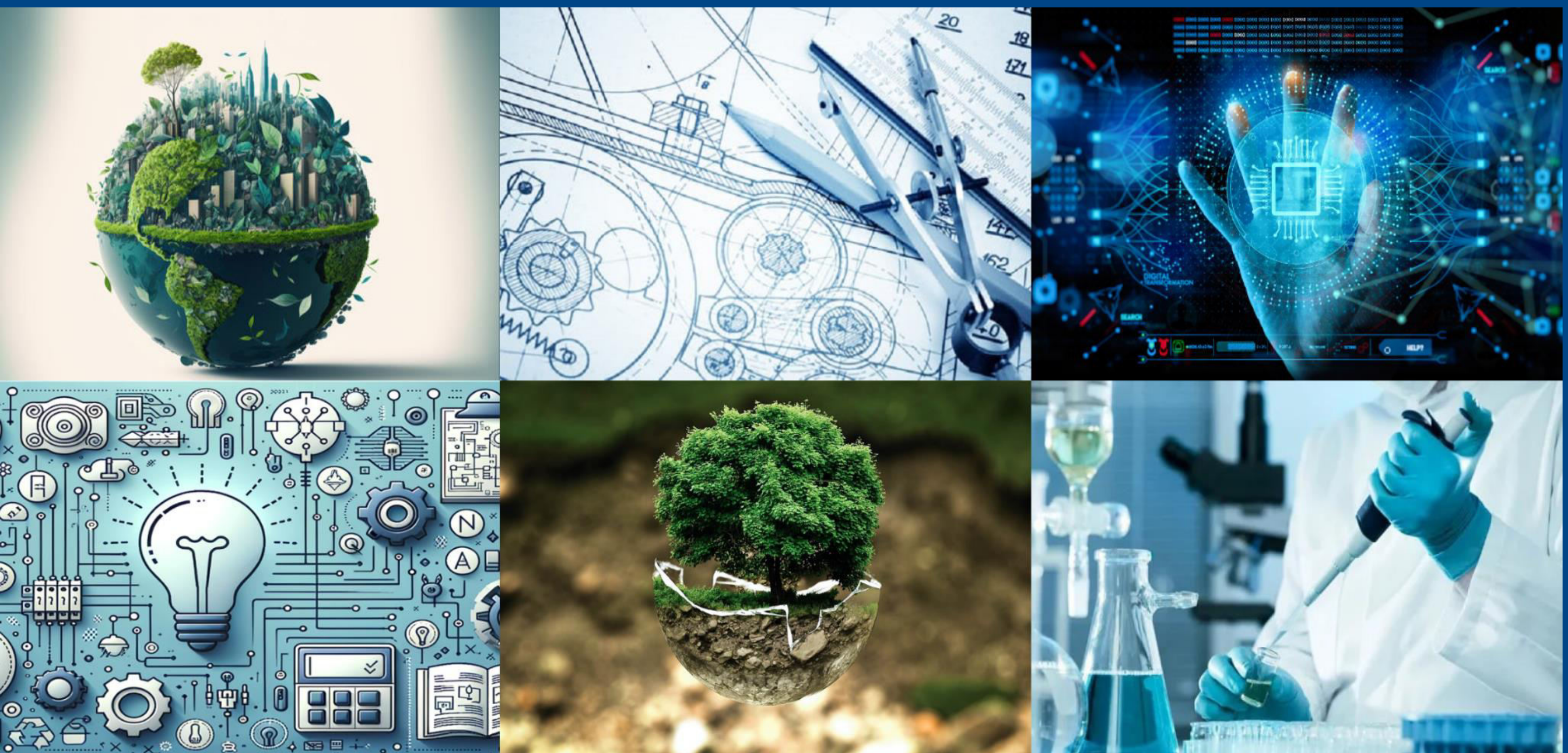
ISSN: 2582-7219

International Journal of Multidisciplinary Research in Science, Engineering and Technology

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)



Impact Factor: 8.206

Volume 8, Issue 4, April 2025

www.ijmrset.com   ijmrset@gmail.com   +91 6381 907 438   +91 63819 07438

© 2025 IJMRSET | Volume 8, Issue 4, April 2025| DOI: 10.15680/IJMRSET.2025.0804096

ISSN: 2582-7219 | www.ijmrset.com | Impact Factor: 8.206| ESTD Year: 2018|

International Journal of Multidisciplinary Research in Science, Engineering and Technology (IJMRSET)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

Pedestrian Detection Using CNN

Suriyaa. B V, Ms. A. Roselin

III-B. Sc. Department of Computer Science with Data Analytics, Dr. N.G.P. Arts and Science College, Coimbatore, India

Assistant Professor, Department of Computer Science with Data Analytics, Dr. N.G.P. Arts and Science College, Coimbatore, India

ABSTRACT: Pedestrian detection is a critical task in computer vision with wide-ranging applications in autonomous vehicles, surveillance systems, and smart cities. This project focuses on developing an efficient pedestrian detection system using Convolutional Neural Networks (CNNs). The system is designed to accurately identify and localize pedestrians in images and video frames under varying environmental conditions and occlusions. CNNs are well-suited for visual recognition tasks due to their ability to automatically learn spatial hierarchies of features. In this project, a deep learning model is trained on benchmark datasets such as INRIA or Caltech Pedestrian Dataset, leveraging convolutional layers to extract features and fully connected layers for classification. The model is capable of distinguishing pedestrians from background objects, and a bounding box regression mechanism is employed to localize pedestrians within the frame.

KEYWORDS: Pedestrian Detection, CNN, Image classification, Stream-lit deployment.

I. INTRODUCTION

Pedestrian detection stands as a vital component in the field of computer vision, serving as a standard problem within object detection tasks. Its applications are widespread, including autonomous driving, pedestrian safety systems, robotics, and intelligent video surveillance.

However, accurately detecting pedestrians presents several challenges. These include variability in human shapes, dynamic background environments, diverse walking poses, and frequent occlusions. To tackle these complexities, recent studies have shifted focus towards deep learning-based methods, particularly Deep Convolutional Neural Networks (DCNNs), due to their ability to capture high-level features through hierarchical learning.

DCNNs have shown remarkable improvements in pedestrian detection tasks. Although traditional handcrafted features sometimes outperform deep features in certain benchmarks, deep learning methods, when combined with suitable feature refinement strategies, tend to consistently deliver superior results.

For instance, older approaches using handcrafted features and shallow models like SquaresChnFtrs (34.81% miss rate), InformedHaar (34.60% MR), and Katamari (22.49% MR) have performed better than earlier deep learning models such as MultiSDP (45.39% MR) and SDN (37.87% MR). Nonetheless, more recent deep learning-based algorithms like CompACT-Deep (11.75% MR), DeepParts (11.89% MR), and TA-CNN (20.86% MR) have outperformed the former, establishing DCNNs as a strong solution in modern pedestrian detection systems.

There are two dominant approaches to using DCNNs in object detection tasks: the **Sliding Window-based CNN (SWCNN)** and the **Region-based CNN (R-CNN)**. The SWCNN method involves applying multi-scale sliding windows across the image to extract features via DCNN. On the other hand, the R-CNN method relies on object proposal algorithms (OPAs), such as selective search, edge boxes, and BING, to generate region proposals which are then analyzed by the CNN for classification and localization.

While OPAs are effective in identifying potential object locations, they often lack the precision needed for tight bounding box localization. Enhancements to the R-CNN framework have improved both training and testing efficiency by optimizing computation sharing.

ISSN: 2582-7219 | www.ijmrset.com | Impact Factor: 8.206| ESTD Year: 2018|

International Journal of Multidisciplinary Research in Science, Engineering and Technology (IJMRSET)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

Contemporary DCNN-based pedestrian detection techniques continue to advance and differ in their strategies for feature extraction and network design. Both SWCNN and R-CNN approaches are actively used to fine-tune network parameters for better detection accuracy.

Several benchmark datasets are widely used for evaluating pedestrian detection algorithms. The most prominent among them include **Caltech-USA**, **INRIA**, **KITTI**, and **ETH**. Caltech-USA and KITTI are large-scale, complex datasets that pose significant challenges. INRIA, while older and smaller, includes diverse scenes such as streets, beaches, and mountainous regions. The ETH dataset, of medium size, offers the added benefit of stereo imagery.

II. TRAINING APPROACHES FOR PEDESTRIAN DETECTION BASED ON DCNN

Training strategies for pedestrian detection using Deep Convolutional Neural Networks (DCNNs) typically fall into two primary categories: **Sliding Window-based CNN (SWCNN)** and **Region-based CNN (R-CNN)** approaches. Both methods leverage the powerful feature learning capabilities of DCNNs, but they differ in how they process images and localize potential pedestrians.The system presented in this project consists of two key stages: 1) Pathogen Detection and 2) Classification. In the first stage, morphological operations are applied to detect infected regions by extracting edges and filtering out healthy areas. The second stage uses a CNN-based classification model to categorize the pathogen type. Finally, the results are displayed through a Streamlit web interface, providing real-time pathogen detection and classification.

Pedestrian detection using SWCNN

The Sliding Window-based Convolutional Neural Network (SWCNN) approach is one of the earliest and most straightforward methods for object detection using deep learning, particularly in pedestrian detection tasks. This technique involves scanning the input image with a fixed-size window that moves across the image at different positions and scales. Each windowed region is then passed through a Convolutional Neural Network (CNN) to determine whether it contains a pedestrian.

In this approach, the CNN acts as a binary classifier, distinguishing pedestrian regions from the background. Since pedestrians can appear at various locations and scales within an image, multiple window sizes and aspect ratios are used to improve detection accuracy. The method ensures that potential pedestrian regions are not missed by exhaustively examining the image.

Despite its simplicity and effectiveness, the SWCNN approach can be computationally expensive, as it requires feature extraction and classification for a large number of overlapping regions. This often results in slower processing times, making it less ideal for real-time applications without significant optimization.

To address the performance limitations, various enhancements have been proposed, such as using region proposals to reduce the number of sliding windows or incorporating more efficient CNN architectures. Additionally, techniques like hard negative mining and non-maximum suppression (NMS) are often employed to improve detection accuracy and reduce false positives.

Pedestrian Detection using various RCNN approaches

Region-based Convolutional Neural Networks (R-CNN) have played a significant role in advancing pedestrian detection through deep learning. These approaches aim to improve both the accuracy and efficiency of object detection by focusing on specific regions in an image that are likely to contain objects, rather than analyzing the entire image exhaustively.

1. R-CNN (Region-based CNN):

The original R-CNN method begins by generating a set of object proposals using algorithms like selective search. Each proposed region is resized to a fixed size and passed through a CNN to extract deep features. These features are then classified using a separate classifier such as a Support Vector Machine (SVM), and bounding box regressors are used to refine localization. While this approach significantly improves accuracy over traditional methods, it is computationally expensive due to redundant operations for overlapping regions.

ISSN: 2582-7219 | www.ijmrset.com | Impact Factor: 8.206| ESTD Year: 2018|

International Journal of Multidisciplinary Research in Science, Engineering and Technology (IJMRSET)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

2. Fast R-CNN:
Fast R-CNN improves upon the original R-CNN by processing the entire image through a CNN just once to generate a feature map. Instead of individually feeding each region proposal into the network, regions of interest (RoIs) are extracted from the shared feature map and classified. This reduces redundancy, improves speed, and allows end-to-end training using a single network.

3. Faster R-CNN:
Faster R-CNN further optimizes the process by replacing external region proposal methods with a built-in Region Proposal Network (RPN). The RPN shares convolutional features with the detection network, allowing the entire pipeline—from proposal generation to classification and localization—to be trained jointly. This significantly boosts both accuracy and speed, making it suitable for real-time pedestrian detection applications.

4. Mask R-CNN (for extended use cases):
Though primarily used for instance segmentation, Mask R-CNN builds on Faster R-CNN and adds a branch for predicting object masks. It can be adapted for detailed pedestrian detection in scenarios requiring pixel-level accuracy, such as in crowded scenes or complex environments.

These R-CNN variants have been widely used with pedestrian detection datasets like Caltech, KITTI, and INRIA, and they continue to evolve with improvements in backbone networks, feature fusion, and context modeling. Their ability to balance precision and speed makes them a strong choice for pedestrian detection tasks across various real-world applications.

III. ANALYZING DEEP PEDESTRIAN DETECTION ALGORITHMS

In recent years, deep learning has revolutionized pedestrian detection by enabling models to learn complex visual patterns and representations directly from data. Among these, deep pedestrian detection algorithms based on Convolutional Neural Networks (CNNs) have demonstrated significant improvements in both accuracy and robustness compared to traditional methods.

These algorithms vary in terms of network architecture, feature extraction methods, training strategies, and performance metrics. Early deep learning models faced challenges such as high miss rates and limited localization accuracy, but newer architectures have addressed these limitations through better design and training optimizations.

Modern approaches incorporate advanced techniques like multi-scale feature fusion, context-aware learning, and part-based detection, allowing them to handle occlusion, pose variation, and cluttered backgrounds more effectively. For example, models like **DeepParts** and **TA-CNN** specifically focus on improving detection in complex scenarios by learning features from different body parts or by integrating semantic attributes.

Another critical aspect of analyzing these algorithms is their computational efficiency. While some models achieve high accuracy, they may not be suitable for real-time applications due to processing delays. Lightweight models and optimization techniques such as pruning, quantization, and knowledge distillation are often explored to make these algorithms faster and more resource-efficient.

Performance evaluation is typically conducted using benchmark datasets like **Caltech**, **INRIA**, and **KITTI**, with metrics such as miss rate (MR), precision, recall, and inference time being commonly used. The results show a consistent trend: deep learning models, especially those trained on large, diverse datasets and fine-tuned using pedestrian-specific data, significantly outperform earlier handcrafted feature-based methods.

In conclusion, deep pedestrian detection algorithms have evolved rapidly, offering impressive detection accuracy and adaptability across various conditions. Continued research focuses on improving speed, reducing computational costs, and enhancing detection in real-world environments with heavy occlusion or low visibility.

© 2025 IJMRSET | Volume 8, Issue 4, April 2025| DOI: 10.15680/IJMRSET.2025.0804096

ISSN: 2582-7219 | www.ijmrset.com | Impact Factor: 8.206| ESTD Year: 2018|

International Journal of Multidisciplinary Research in Science, Engineering and Technology (IJMRSET)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

IV. CATEGORIZATION OF DCNN-BASED PEDESTRIAN DETECTION

Pedestrian detection methods utilizing Deep Convolutional Neural Networks (DCNNs) can be broadly categorized based on their architectural design, detection strategy, and use of contextual or part-based information. These categories help in understanding how different approaches tackle the challenges of pedestrian detection, such as occlusion, scale variation, and complex backgrounds.

1. Full-Body Detection Models:
bounding box. They typically work well in clear, unobstructed scenes and are trained to recognize full pedestrian silhouettes. While efficient in simple environments, they may struggle with partial occlusions or varied body poses.

2. Part-Based Detection Models:
To overcome the limitations of full-body models, part-based detection methods break down the human body into components such as head, torso, and legs. These parts are detected separately and then combined to infer the presence of a pedestrian. Models like DeepParts fall into this category and are particularly useful in scenarios where pedestrians are partially visible.

3. Multi-Scale and Multi-Stage Models:
These methods enhance detection performance by analyzing features at multiple scales or through a staged process. They are designed to recognize pedestrians of different sizes, often caused by varying distances from the camera. Multi-stage models may use an initial network to generate proposals and another for refinement, improving both localization and classification accuracy.

4. Context-Aware Models:
Contextual information, such as surrounding objects or scene layout, is leveraged in this category to improve detection accuracy. By understanding the environment in which a pedestrian is likely to appear, these models reduce false positives and improve detection in cluttered scenes.

5. Region Proposal-Based Models:
These methods use region proposal algorithms or networks to identify areas of interest before applying DCNNs for classification and localization. Faster R-CNN and its derivatives are key examples, balancing detection precision with efficiency.

Each category has its strengths and is suitable for different application scenarios. Combining elements from multiple categories, such as using part-based features with contextual learning, often leads to better performance and robustness.

V. CONCLUSION

Deep Convolutional Neural Networks (DCNNs) have significantly transformed the landscape of pedestrian detection, offering powerful tools to address long-standing challenges such as occlusion, scale variation, and complex backgrounds. Through various approaches—ranging from sliding window techniques to region-based models, and from full-body detection to part-based and context-aware methods—DCNNs have demonstrated superior accuracy and adaptability compared to traditional handcrafted methods.

The continuous evolution of network architectures, training techniques, and dataset utilization has resulted in models that not only detect pedestrians more precisely but also perform efficiently across real-world scenarios. Despite ongoing challenges like real-time performance and dense crowd detection, the integration of deep learning into pedestrian detection systems holds great promise for applications in autonomous vehicles, surveillance, and public safety systems.

As research progresses, combining multiple strategies and optimizing deep models for both speed and accuracy will be key to building reliable, intelligent pedestrian detection systems for next-generation smart technologies.

ISSN: 2582-7219 | www.ijmrset.com | Impact Factor: 8.206| ESTD Year: 2018|

International Journal of Multidisciplinary Research in Science, Engineering and Technology (IJMRSET)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

REFERENCES

1. P. Dollár, R. Appel, S. Belongie, and P. Perona, "Fast feature pyramids for object detection," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 36, no. 8, pp. 1532–1545, 2014.
2. W. Ouyang and X. Wang, "Joint deep learning for pedestrian detection," Proceedings of the IEEE International Conference on Computer Vision (ICCV), 2013, pp. 2056–2063.
3. J. Hosang, M. Omran, R. Benenson, and B. Schiele, "Taking a deeper look at pedestrians," Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2015, pp. 4073–4082.
4. S. Zhang, R. Benenson, and B. Schiele, "Filtered channel features for pedestrian detection," Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2015, pp. 1751–1760.
5. S. Zhang, R. Benenson, M. Omran, J. Hosang, and B. Schiele, "How far are we from solving pedestrian detection?" Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2016, pp. 1259–1267.
6. R. Girshick, "Fast R-CNN," Proceedings of the IEEE International Conference on Computer Vision (ICCV), 2015, pp. 1440–1448.
7. S. Ren, K. He, R. Girshick, and J. Sun, "Faster R-CNN: Towards real-time object detection with region proposal networks," Advances in Neural Information Processing Systems (NeurIPS), 2015, pp. 91–99.
8. L. Zhang, L. Lin, X. Liang, and K. He, "Is Faster R-CNN doing well for pedestrian detection?" European Conference on Computer Vision (ECCV), 2016, pp. 443–457.
9. M. Mathias, R. Benenson, R. Timofte, and L. Van Gool, "Handling occlusions with franken-classifiers," Proceedings of the IEEE International Conference on Computer Vision (ICCV), 2013, pp. 1505–1512.
10. P. Dollar, C. Wojek, B. Schiele, and P. Perona, "Pedestrian detection: An evaluation of the state of the art," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 34, no. 4, pp. 743–761, 2012.

INNO SPACE
SJIF Scientific Journal Impact Factor

INTERNATIONAL STANDARD SERIAL NUMBER INDIA

निस्केयर NISCAIR

INTERNATIONAL JOURNAL OF MULTIDISCIPLINARY RESEARCH IN SCIENCE, ENGINEERING AND TECHNOLOGY

| Mobile No: +91-6381907438 | Whatsapp: +91-6381907438 | ijmrset@gmail.com |

www.ijmrset.com