Library Progress International
Vol.44 No. 3, Jul-Dec 2024: P. 16823-16832

Print version ISSN 0970 1052
Online version ISSN 2320 317X

Original Article

Available online at www.bpasjournals.com

A Multi-Scale Conv GCN-FCN Approach with Hybrid Optimization for Enhancing Object Detection in Occlusions

Apurva Kandelkar[1,2], Isha Batra[3],Shabnam Sharma[4], Arun Malik[5]

[1]Lovely Professional University, Phagwara, Punjab, India
[2],Dr. D. Y. Patil Institute of Technology, Pimpri, Pune, Maharashtra, India
[3] Lovely Professional University, Phagwara, Punjab, India
[4] Opentext Corporation
[5]Lovely Professional University, Phagwara, Punjab , India
apurvakandelkar@gmail.com[1, 2], isha.batra2487@gmail.com[3], shabnam09sharma@hotmail.com[4], arumalikhisar@gmail.com[5]

How to cite this article: Apurva Kandelkar, Isha Batra, Shabnam Sharma, Arun Malik (2024) Intersectionality of Gender with Race, Class and Ethnicity and its Impact on Economic Development: A Complex and Multi-Layered Framework. *Library Progress International,* 44(3), 16823-16832

Abstract
Object detection is a fundamental problem in computer vision that involves identifying and localizing instances of objects within digital images. Real-time object detection systems often operate in dynamic environments where the scene are rapidly changing. Noise, poor lighting, background clutter and occlusion are some significant factors that impact the quality of object detection. In real-time object detection, objects can often be overlapped with each other. Occlusion is challenge in which an object is partially obscured by other objects, poses significant challenges for accurately detecting and localizing objects in videos. This common issue hampers the performance of object detection algorithms. Therefore, existing approaches to object detection often struggle with accurately localizing and detecting partially occluded objects. This research offers an in-depth analysis of the most recent state-of-the-art approaches for addressing the occlusion problem in object detection. Unlike previous occlusion detection methods, we introduce a novel approach that utilizes multi-scale context information to handle occlusion in object detection better. We proposed conv GCN-FCN with a hybrid optimization approach and evaluated on two standard datasets: MS COCO Dataset, and PASCAL VOC Dataset. We investigate the efficiency of our novel approach, by using a Graph convolution network (GCN) with a fully convolution network (FCN) and hybridization of the booster algorithm. Our proposed refined selfish herd optimizer enhance the neural network's performance that utilizes multi-scale context information to handle occlusion in object detection better. The experimental results demonstrate our proposed method significantly improves occlusion handling and also improves evaluation parameter such as accuracy, precision, F1 score and sensitivity etc. compared to previous CNN algorithms and methods.
Keywords: Occlusion, Object Detection, Conv GCN-FCN, Multi-Scale Context Information, Neural Network Optimization

Introduction

The occlusion problem in object detection is fundamental problem in computer vision with multiple extensive applications such as autonomous vehicles [1] [2], surveillance, robotics, drones, robot navigation, and mobile robotics [3]. Even with a single picture, occlusion is a frequent occurrence in visual situations. Occlusion is the phenomenon wherein portions of an object in the scene are obscured or obstructed by other components or objects, rendering it partially invisible from a specific perspective. The object detections are rely on objects appearance in the image which becomes challenging when objects are partially visible. In the real world, object detection methods often struggle to efficiently address the complexities of object detection problems. For example, in the case of autonomous driving 2 Dimensional information is not enough to pursue real-world issues, so 3 Dimensional information is required to make accurate decisions The conventional 2 Dimensional object detection techniques often leads to object detection failure in such conditions. In contrast, 3 Dimensional object detection aims to accurately estimate an object's three-dimensional position, orientation, and dimensions, providing a more comprehensive understanding beyond mere detection. Object detection in complex environments with occlusions remains a significant challenge for computer vision systems. To conquer the challenge of detecting objects in occluded scenes, researchers have proposed various approaches, including the integration of convolutional neural networks and graph convolutional networks [5]. In this research paper, we present a novel multi-scale approach that combines convolutional neural networks, graph convolutional networks, and fully convolutional networks. The basic idea is to leverage the strengths of each of these architectures to achieve robust object detection, even in the presence of occlusions. To address this issue, we propose a novel approach called the Multi-Scale Conv GCN-FCN with Hybrid Optimization, which consists of three main components: a multi-scale convolutional backbone, a graph convolutional network. Our proposed method, named the Multi-Scale Conv GCN-FCN with Hybrid Optimization, consists of three main components: a multi-scale convolutional backbone, a graph convolutional network module, and a fully convolutional network for object detection. The multi-scale convolutional backbone [4] is designed to capture features at different scales, which is crucial for handling occlusions of varying sizes. The graph convolutional network module is then used to model the contextual relationships between objects, allowing the system to reason about occluded regions based on the surrounding context. The proposed approach builds upon recent advancements in object detection and deep learning, leveraging a combination of multi-scale convolutional features, graph convolutional networks, and a fully convolutional network for robust occluded object detection. The multi-scale convolutional backbone extracts features at multiple scales to capture both local and global information, while the graph convolutional network module models the spatial relationships between objects to better handle occlusions.

1.1 The Occlusion Problem

The detection of occluded objects remains a significant challenge in computer vision, as occlusions can significantly degrade the performance of object detection algorithms. When one object is close to another, the object gets covered because of another object called an occlusion. The object may be partially or fully covered by other objects, making it difficult to detect the object from the input scene. Because of the noise scanning, occlusion, discrete sampling, and cluttered scenes are challenging. The occlusion problem in object detection is a significant challenge because it leads to the loss of visual information, increased ambiguity, and necessitates sophisticated computational techniques to effectively address it. Existing methods often struggle to accurately identify objects that are partially obscured by other objects or environmental factors [6]. Research on occlusion in object detection is crucial for addressing real-world scenarios where objects may be partially or completely obstructed from view. Many times occlusion occurs due to itself being called Intra-class occlusion and if another object covers the object it is called Inter-class occlusion [8]. Figure 1 shows the occlusion problem in object detection. In object detection, occlusion occurs when one object (Object B) partially or fully blocks another object (Object A), leading to only a partial view of Object A. For example, if a car is partially covered by a tree in an image, the algorithm might struggle to recognize the full car.

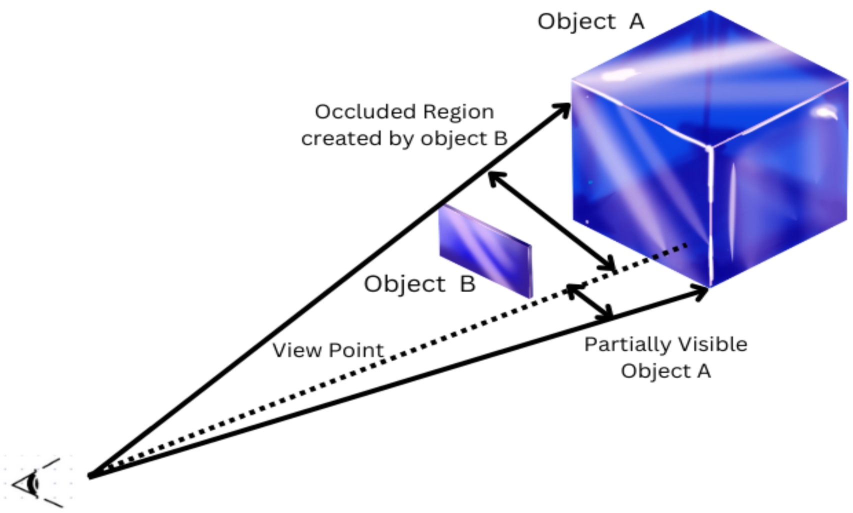

Figure 1 Partial Occlusion Problem (Object A is impeded because of Object B. From this viewpoint object A is partially visible)

The occlusion problem is cause of many segmentation issues, particularly when the objects belong to the same class. This happens because the segmentation method assumes the item in the region of interest (ROI) has a nearly full shape and

predicts each object mask independently. The idea of segmenting images into objects and background has occurred in several contexts in computer vision. Object segmentation helps in highlighting scenes most interesting parts [7].
The key contributions of the study can be summarized in the following points:
1. We propose and develop a fully convolutional one-stage (FCOS) box head and the backbone+ feature pyramid network (FPN) is designed and analyzed by us to identify regions of interest in input images.
2. To optimize the GCN layer we further propose a bidirectional collaboration network (BCNet) for occlusion detection, we advise adjusting the GCN layer's hyper parameters.
3. To improve the performance of neural networks, we suggest a unique structure termed the hybridization of the optimized selfish herd optimizer and booster algorithm which effectively works upon real-world MS COCO and PASCAL datasets.
The paper organization is as follows: Section 1 presented the introduction to object detection and its introduction of the occlusion problem that occurs during object detection in the 2D and 3D environment. In section 2, we present a review of various strategies addressing the occlusion challenge in object detection. Section 3 represents the proposed methodology includes related work regarding occlusion problem in object detection. Section 4 presents experimental analysis of the proposed approaches and this section also covers the dataset with the comparison of the state-of-the-art methods. Section 5 represents conclusion and future scope.

2. Related Work

There are several studies in past literature that use various methods for object detection in occlusion problem. We review relevant literature on object detection occlusion problems and present a unique approach that integrates deep learning techniques and geometric constraints to effectively handle the challenges posed by occlusions in the context of 3D object detection. The related work in addressing the occlusion issue in object recognition highlights the diverse range of approaches and solutions that have been proposed to tackle this challenging issue. Some potential sources to support your literature review on the occlusion problem in object detection are:
The paper [7] contribute to this field with a novel approach through their bilayer network architecture, addressing key issues in occluded object detection and segmentation. The occlusion-aware instance segmentation has gained significant attention due to its challenges in handling overlapping objects in computer vision tasks is discussed in paper. The paper [7] suggests the bilayer network architecture, which is quite innovative and will likely spark further development for some of the more complex approaches to overcoming occlusions in computer vision. Through improvement and building on current models toward newer concepts, such as the bilayer architecture, this paper hence becomes of prime importance toward further development of occlusion-aware instance segmentation techniques.
The paper [9] proposes a new method called OPA-3D, which is Occlusion-Aware Pixel-Wise Aggregation, designed to overcome such challenges. The paper basically discussed about in-depth analysis of occlusion-aware techniques about monocular 3D object detection, especially using OPA-3D to address better occlusion handling. In this paper, they are discussing and evaluating several occlusion-aware strategies over monocular 3D object detection tasks, including attention mechanisms and contextual reasoning. Such valuable insights into how to approach occlusion challenges in object detection is discussed and providing deeper insight into the most advanced approaches in the domain.
The paper [10] suggested technique for monocular 3D object detection, which aids in significantly estimating depth and the detection of occluded objects. It focuses on the aspect of scene complexity and occlusion, thus addressing occlusion through an adaptive instance depth estimation model for scene-level comprehension. MonoSAID is a profound advancement in monocular 3D object detection as an adaptive mechanism to estimate the depth directly addresses the problem of occlusion. Dynamic adaptation, together with leveraging of scene-level context, enhances detection accuracy and depth estimation, especially in complex environments heavy with occlusion.
The author in [11] propose method for enhancing monocular 3D object detection through mechanisms in the presence of occlusions. The authors resolve the common source of error for traditional monocular systems-fouled 3D detection due to occluded objects-by attempting to improve detection performance when partial or full occlusions exist. It contributes to the emerging research on approaches to dealing with occlusion in autonomous driving and robotics applications.
The method proposed in[12] addresses the challenge by proposing a bi-layer network architecture that explicitly models occlusion and foreground-background relationships. The work continues to improve the segmentations of occluded objects as well as partially occluded objects with better-occlusion-handling capability by equipping methods like Mask R-CNN by a novel mechanism in a crowded scene. Compared to the two-stage method of existing instance segmentation approaches, authors modelled the formation of the image as composition of two overlapping layers and proposed Bilayer Convolutional

Network in which the top GCN layer identifies the occluding objects or occluder and the bottom GCN layer infers the partially occluded instance or occludee. The explicit modeling of the occlusion relationship using a bilayer structure naturally decouples the boundaries of both the occluding and the occluded instances, considering interaction between them during mask regression. The authors performs extensive experiments on COCO and KINS show that the occlusion-aware BCNet achieves large and consistent performance gain especially for heavy occlusion cases.

The core contribution of this paper [13] is the introduction of an explicit occlusion model integrated into the 3D object class representation framework. The authors propose a method to explicitly represent and reason about occlusions, improving the detection and recognition of partially occluded objects. The paper extends traditional 2D object detection techniques by incorporating 3D geometric information. The authors develop a framework that represents objects in three dimensions, allowing for more accurate modeling of object shapes and their possible occlusions. Detailed tests on reference datasets including the Pascal 3D+ and KITTI datasets show that explicit occlusion modeling works well. The findings demonstrate notable gains in detection performance over conventional techniques that ignore occlusions.

This [14] paper addresses the problem of detecting partially occluded objects by focusing on visible features and parts of them. In traditional object detection algorithms, it is hard to deal with an occlusion scenario because portions of an object may not be visible. The paper proposed a feature-based model that separates the visible parts from occluded areas with a part-based representation about objects. Focusing on visible parts, the model decreases false detections by missing object components. This paper's contribution lies in enhancing detection accuracy under occlusion, a problem encountered in various real-world applications such as autonomous driving and video surveillance.

The [15] presents an IoU_MDA occluded object detection algorithm to deal with concerns of inaccurate anchor-box matching when under occluded conditions. The key contribution of this paper lies in the Fuzzy Sample Anchor Box IoU Matching Degree Deviation Aware technique, which modifies the traditional dynamical anchor-box matching technique so as to better manage occlusion. While improving the IoU matching process, the proposed approach increases the accuracy in detecting objects partially occluded. Moreover, the application of the fuzziness in the IoU matching accounts for the uncertainty in occlusion due to the proposed algorithm.

In this paper [16] propose an improved real-time detection algorithm for occluded objects, building upon the YOLOv5s framework. The upgrade is focused towards bringing more robustness to the detection model in real-world applications such as self-driving cars or surveillance where objects are partially or fully occluded. For enhancing the model's ability to handle occlusion, authors change the YOLOv5s architecture by adjusting its feature extraction layers as well as its heads. The novel contributions are: using a customized modules to even detect when it is substantially occluded by embedding an optimal anchor-free method.

3. Methodology

We propose approach that aim to enhance the performance of neural networks used for detecting and classifying occluded objects by combining the booster algorithm and the refined selfish herd optimizer. The proposed architecture called Conv GCN-FCN with Hybrid Optimization for detection of occluded objects is shown in Figure 2. In scenarios with heavy occlusion, traditional object detection methods struggle as overlapping objects within a single bounding box can generate confusing instance contours, complicating the distinction between actual object edges and occlusion boundaries. To overcome this limitation we perform the hybrid optimization of Conv GCN-FCN which aims to classify occluded objects with different class labels. We utilize a multi-scale deep feature pyramid based on the core object detection network, such as Mask R-CNN [17]. Figure 2 shows architecture of our Conv GCN-FCN with hybrid optimization.

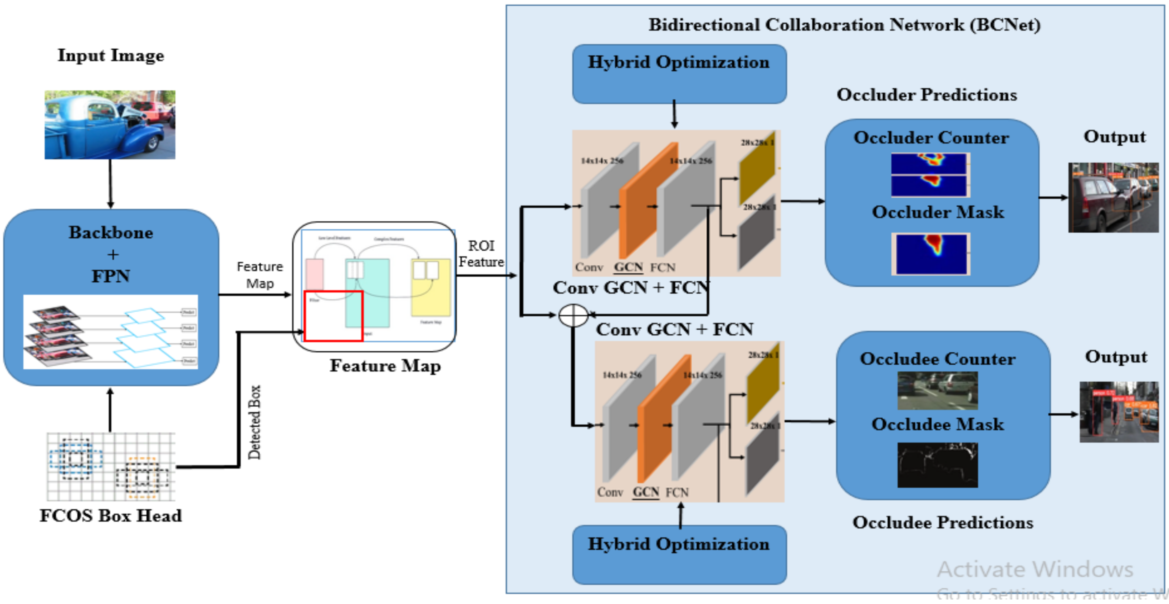

Figure 2 Architecture of our proposed Conv GCN-FCN with Hybrid Optimization

The proposed model is primarily composed of four components: 1) The extraction of features, 2) The mapping of features, 3) Hybridization optimization, and 4) The occluded prediction forecast outcomes. The main methodology of our proposed model is to identify the occlusion patterns that explicitly detect contours and masks using a GCN-FCN layer.

As shown in figure 2, our approach input image is provided to FPN, it is a feature extractor that produces fully convolutional created feature maps at several levels that are properly proportioned, given any single-scale image as input [20]. FPN combines feature maps from different layers through lateral connections in order to improve the detection at multiple scales[18]. Multi-scale feature extraction techniques, such as Feature Pyramid Networks and Dilated Convolutions, are commonly used to improve the feature extraction capability across various scales[19].This approach uses Backbone+Feature Pyramid Network (FPN) which is feature extractor that generates multiple feature map layers (multi-scale feature maps) with better quality information [21]. To extract feature maps from input images, we employed a pre-trained convolutional neural network (CNN) as our main network. In order to locate and identify items inside the pictures, the object detection network uses these feature maps. The core component of the object detection pipeline is the backbone network [22], which works in tandem with the feature pyramid network (FPN) to function as the feature extractor. Graph-FCN is a graph model designed for picture semantic segmentation that is initiated by a fully convolutional network (FCN). FCOS stage in our proposed approach completely avoids the complicated computation related to anchor boxes such as calculating overlapping during training. Next feature map layer generate each feature map corresponds to a specific filter and represents the response of that filter to the input image. This previous stages helps to extract Region of interest from the input image using the backbone+ feature pyramid network (FPN), and the fully convolutional one-stage (FCOS) box head as illustrated in Figure 2.In our proposed model the occlusion pattern is model explicitly by identifying contours and mask with the use of GCN- FCN layer. The FPN-equipped backbone network initially recovers intermediate convolutional features from an input picture in order to prepare it for further processing. The object detection head then crops the ROI feature so that BCNet can create segmentation masks, and predicts bounding boxes with locations and categories for possible occurrences. Figure 3 illustrating object detection utilizing bounding boxes in environment with occlusions.

Figure 3 Object Detection using Bounding Boxes in an Occluded Environment

The goal of the occlusion perception phase is to represent occluding regions by simultaneously detecting contours and masks. It is composed of two convolution layers, namely FCN and the first GCN layer. The refined occlusion feature is element-wise added to the input ROI feature and transmitted to the second GCN, forming a residual connection. Ultimately, the second GCN, which shares a structure with the first GCN, uses this occlusion-aware feature to help it in segmenting the occludee and generates contours and masks for the occluded instance. We aim to enhance the performance of neural networks used for detecting and classifying occluded objects by combining the booster algorithm and the refined selfish herd optimizer. While performing experimental analysis on MS COCO Dataset and Pascal VOC Dataset and evaluated our proposed algorithm with already existing CNN methods.

Following are the four optimization techniques that are applied by our proposed method:
Optimizer 0: GCN+FCN – Our proposed approach applied optimization technique called GCN+FCN which are GCN comprises multiple convolutional and pooling layers to extract features, succeeded by the final fully-connected layers. Fully Convolutional Networks (FCNs) are primarily utilized for semantic segmentation tasks. **Optimizer 1:** GCN + FCN with Booster Algorithms -Boosting is an ensemble learning technique in which multiple weak classifiers are combined to build a strong classifier. **Optimizer 2:** GCN + FCN with Selfish Herd Optimization Model-swarm optimization algorithm for solving global optimization problems. **Optimizer 3:** Proposed GCN + FCN Hybrid Optimization Model - Hybrid optimizations dynamically select, during compilation, which optimization algorithm to apply from a collection of diverse algorithms that achieve the same optimization goal.

4. Experiments

We demonstrate the ability of our approach to infer and identify occluded objects on two datasets. Our framework evaluate the performance aiming on its ability to recover partially occluded objects and finding regions of interest from input images using backbone+ feature pyramid network (FPN), and the fully convolutional one-stage (FCOS) box head. In particular, given an input image containing occluded objects, the proposed approach effectively detects these occluded objects and accurately classifies them according to their correct classes

4.1 Datasets

We evaluated our proposed conv GCN-FCN with hybrid optimization approach on two standard dataset: MS COCO Dataset [23], and PASCAL VOC Dataset [24]. We follow and split both datasets for the training data into a training and validation subset.

For tasks involving segmentation and key point estimate, the MS COCO dataset is utilized [23]. 164K photos divided into sets for testing (41K), validation (41K), and training (83K). This dataset provide various annotations that is task to the process of labelling and tagging data with additional information to provide context or ground truth. This dataset provide this annotation as bounding boxes, key point labelling and instance labelling.

We use the PASCAL Visual Object Classes (VOC) challenge dataset which provide benchmark in detecting a standard for visual object category detection and recognition, offering a standard dataset of images with annotations and standard assessment protocols to the machine learning and vision fields [24]. This PASCAL VOC Dataset provide 20 distinguish classes which helps us to classify occluded images.

4.2. Evaluation Matrices

We compare our approach with other state-of-the-art CNN networks and evaluate our proposed approach on two standard datasets: MS COCO Dataset and Pascal VOC Dataset. We follow evaluation protocols provided by these benchmark for both occlusion detection and classification.

For evaluating the models performance we use confusion matrix [22], which offers an organized method of evaluating the model overall accuracy as well as the different kinds of errors it makes. The matrix consists of four main components: true positives (TP), true negatives (TN), false positives (FP), and false negatives (FN). These elements collectively measure the model's ability to correctly classify instances, while also capturing the misclassifications. The evaluation matrices serve in improving the evaluation of the detection algorithms efficiency of handling occlusions. Accuracy [22] measures the overall correctness of the predictions made by occlusion detection framework.

$$Accuracy = \frac{tp+tn}{tp+tn+fp+fn}$$

The accuracy would reflect the proportion of correctly detected objects, considering both occluded and non-occluded instances [25]. For occlusion detection, precision would indicate how many of the detected objects are correctly identified as non-occluded given the predicted bounding boxes. The ratio of true positive predictions to all positive predictions (true positives plus false positives) is known as precision.

$$Precision = \frac{tp}{tn+fp}$$

F1 score helps detect occluded objects and enhances the accuracy of non-occluded object detection. It represents the harmonic mean of precision and recall, providing a balanced assessment between these two metrics. When comparing the quality of reconstructed images to the original images, PSNR is a statistic that is frequently utilized. The percentage of accurate positive predictions among all real positive events in the dataset is measured by recall. Recall in occlusion detection refers to the algorithm's capacity to identify every occluded item, ensuring that no occluded instance is ignored [26]. Specificity provide details about the degree to which the algorithm detects areas or objects that are not occluded. Similar to recall, sensitivity measures how well an algorithm can identify real positive instances. Sensitivity in occlusion detection refers to how well the algorithm can identify both occluded and non-occluded objects.

4.3 Experimental Results

We apply our proposed method on image dataset from MS COCO and PASCAL to recover objects to be un-occluded. In addition, we compare the performance of our occluded object detection method with existing neural network architectures, demonstrating superior evaluation metrics. The results are illustrated in Table 1 and Table 2 are in terms of accuracy, precision, recall and other evaluation parameters.

Evaluation Results on PASCAL Dataset: In this section we discuss performance of the algorithm from several aspects such as the evaluation parameters and results demonstrate that our method is a significant improvement over the existing state of the art algorithms. Table 1 shows comparative analysis of our approach conv GCN-FCN with hybrid optimization with various CNN architecture on and PASCAL dataset. The table presents experimental result of proposed architecture on PASCAL dataset using evaluation parameters used accuracy, F1 score, and PSNR(Result values for Epochs =200).

Table 2 Experimental Results on PASCAL Dataset (Result values for Epochs=200)

Table 1: Comparison with Results on PASCAL(Result with Epochs=200)

Methods	Accuracy	F1 Score	Precision	PSNR	Recall	Sensitivity	Specificity
RCNN	88.436	88.481	87.889	44.765	88.682	88.682	87.637
Fast RCNN	91.591	90.432	90.675	48.639	90.416	90.416	91.209
Faster RCNN	92.094	92.973	93.149	51.959	93.105	93.105	93.299
Conv GCN-FCN	93.647	93.618	93.371	54.163	93.795	93.795	93.953
Conv GCN-FCN with Booster Algorithm	94.550	94.521	95.103	55.910	95.437	95.437	94.187
Conv GCN-FCN with SHO	95.171	95.125	95.454	56.936	95.523	95.523	95.302
Conv GCN-FCN with Hybrid Optimization(Ours)	96.283	95.931	96.129	60.425	96.018	96.018	96.481

As shown in Table 1 we can say that our approach is showing better performance evaluation parameter such as accuracy values on PASCAL dataset specifically for epoch's values as 200.

Evaluation Results on MS COCO Dataset: Table 2 presents result of proposed architecture on MS COCO Dataset using evaluation parameters used accuracy, F1 score, and PSNR.

Table 1 Experimental Results on MS COCO Dataset (Result values for Epochs =200).

Table 1: Comparison with Results on MS COCO (Result with Epochs=200)

Methods	Accuracy	F1 Score	Precision	PSNR	Recall	Sensitivity	Specificity
RCNN	87.893	88.238	88.994	44.420	88.513	88.513	88.783
Fast RCNN	89.622	91.372	89.451	47.340	90.335	90.335	91.408
Faster RCNN	92.939	92.848	92.155	51.940	92.624	92.624	92.100
Conv GCN-FCN	93.969	94.131	94.554	54.300	93.120	93.120	94.434
Conv GCN-FCN with Booster Algorithm	94.491	95.344	94.902	55.276	94.791	94.791	94.566
Conv GCN-FCN with SHO	95.369	95.369	95.002	56.947	95.065	95.065	95.291
Conv GCN-FCN with Hybrid Optimization(Ours)	96.458	96.437	96.182	60.628	96.367	96.367	95.839

We perform a comparative analysis of various CNN algorithms, including the proposed Conv GCN-FCN with hybrid optimization, on the MS COCO dataset. The MS COCO dataset is a widely used benchmark for evaluating the performance of computer vision models. Figure 3 and Figure 4 showing graph where our approach conv GCN+ FCN with Hybrid Optimization is showing best results for various standardized Epochs=100,200 on MS COCO and PASCAL Dataset for detection of occluded objects. We apply several existing CNN algorithms, including state-of-the-art approaches, and compare their results with our proposed Conv GCN-FCN with hybrid optimization algorithm.

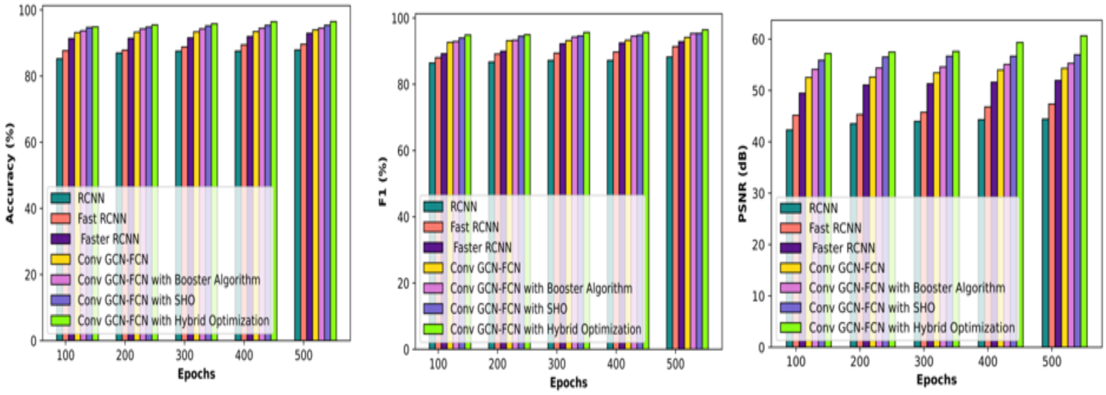

Figure 3 Evaluation Results on MS COCO Dataset. Comparative analysis of various CNN architecture with our conv GCN –FCN with hybrid optimization using evaluation parameters used accuracy, F1 score, and PSNR

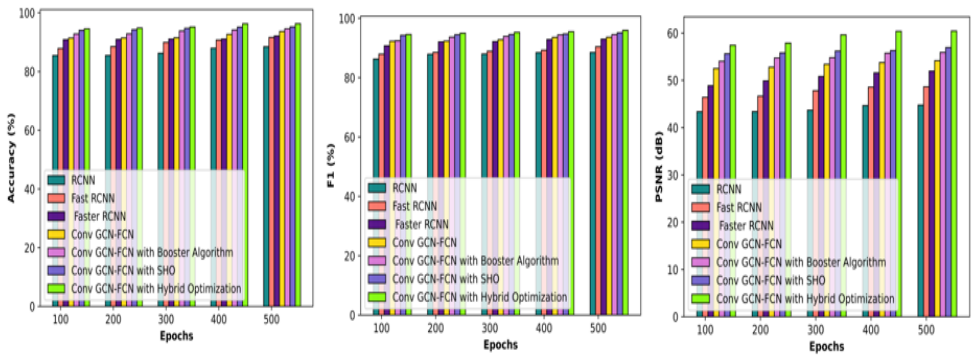

Figure 4 Evaluation Results on PASCAL Dataset. Comparative analysis of various CNN architecture with our conv GCN –FCN with hybrid optimization using evaluation parameters used accuracy, F1 score, and PSNR As shown in Figure 4 graph depicting the performance of our method, conv GCN+ FCN with Hybrid Optimization, which demonstrates superior results across different standardized epochs (100, 200) on the PASCAL Dataset. The outcomes indicate that our approach represents a notable advancement compared to current state-of-the-art algorithms in detecting occluded objects. The experimental analysis is performed on our proposed method and on existing state of art methods. As per comparative study our proposed results of PSNR is improve by 4% as compared to existing state of art methods. The overall performance of our proposed approach is better as compared to existing state of art methods.

5. Conclusion and Future Scope

We propose a novel graph-based architecture for the categorization and detection of occluded objects, combining the strengths of convolutional GCN-FCN with a hybrid optimization approach that leverages the Selfish Herd Optimizer and Booster Algorithm. The experimental findings indicate that our method overtakes traditional CNN methods in performance, particularly in recognizing occluded objects. The efficiency of our system was confirmed using the MS COCO and PASCAL VOC Datasets, where it showed promising results in handling occluded instances. We also seek that our research will explore the adaptation of the proposed conv GCN-FCN architecture into three-dimensional spaces, addressing the challenges of depth perception and complex occlusions in environments represented by benchmarks such as the KITTI 3D dataset. Furthermore, we aim to enhance the robustness of our system against a wider array of occlusion scenarios, pushing the boundaries of object detection performance in increasingly dynamic and unpredictable settings. Through this endeavor, we strive for a comprehensive solution that maintains high accuracy even under severe occlusions, contributing to the advancement of computer vision applications in real-world conditions.

References:

[1] Dai, X., 2019. HybridNet: A fast vehicle detection system for autonomous driving. Signal Processing: Image Communication, 70, pp.79-88.
[2] Qian, R., Lai, X. and Li, X., 2022. 3D object detection for autonomous driving: A survey. Pattern Recognition, 130, p.108796.
[3] Czajewski, W. and Kołomyjec, K., 2017. 3D object detection and recognition for robotic grasping based on RGB-D images and global features. Foundations of Computing and Decision Sciences, 42(3), pp.219-237.
[4] Liu, S., Qi, X., Shi, J., Zhang, H. and Jia, J., 2016. Multi-scale patch aggregation (mpa) for simultaneous detection and segmentation. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (pp. 3141-3149).
[5] Li, G., Xiong, C., Thabet, A. and Ghanem, B., 2020. Deepergcn: All you need to train deeper gcns. arXiv preprint arXiv:2006.07739.
[6] Patel, H., 2023. A Comprehensive Study on Object Detection Techniques in Unconstrained Environments. arXiv preprint arXiv:2304.05295.
[7] Ke, L., Tai, Y.W. and Tang, C.K., 2023. Occlusion-aware instance segmentation via bilayer network architectures. IEEE Transactions on Pattern Analysis and Machine Intelligence.
[8] Arnold, E., Al-Jarrah, O.Y., Dianati, M., Fallah, S., Oxtoby, D. and Mouzakitis, A., 2019. A survey on 3d object detection methods for autonomous driving applications. IEEE Transactions on Intelligent Transportation Systems, 20(10), pp.3782-3795.
[9] Su, Y., Di, Y., Zhai, G., Manhardt, F., Rambach, J., Busam, B., Stricker, D. and Tombari, F., 2023. Opa-3d: Occlusion-aware pixel-wise aggregation for monocular 3d object detection. IEEE Robotics and Automation Letters, 8(3), pp.1327-1334.
[10] Xia, C., Zhao, W., Han, H., Tao, Z., Ge, B., Gao, X., Li, K.C. and Zhang, Y., 2024. MonoSAID: Monocular 3D Object Detection based on Scene-Level Adaptive Instance Depth Estimation. Journal of Intelligent & Robotic Systems, 110(1), p.2.
[11] Wan, Y., Zhao, Q., Hu, Y. and Fang, L., 2023, November. Enhancing Monocular 3D Object Detection with Occlusion Compensation. In 2023 China Automation Congress (CAC) (pp. 7870-7874). IEEE.
[12] Ke, L., Tai, Y.W. and Tang, C.K., 2021. Deep occlusion-aware instance segmentation with overlapping bilayers. In Proceedings of the IEEE/CVF conference on computer vision and pattern recognition (pp. 4019-4028).
[13] Zeeshan Zia, M., Stark, M. and Schindler, K., 2013. Explicit occlusion modeling for 3d object class representations. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (pp. 3326-3333).
[14] Chan, K.C., Ayvaci, A. and Heisele, B., 2015, September. Partially occluded object detection by finding the visible features and parts. In 2015 IEEE International Conference on Image Processing (ICIP) (pp. 2130-2134). IEEE.
[15] Chen, Y., Li, X., He, Z., Chen, H., Chen, J. and Wu, B., 2024. IoU_MDA: An Occluded Object Detection Algorithm Based on Fuzzy Sample Anchor Box IoU Matching Degree Deviation Aware. IEEE Access.
[16] Yuan, H., Li, W. and Cheng, Y., 2024, March. Real-time Detection Algorithm for Occluded Objects Based on

Improved YOLOv5s. In 2024 5th International Seminar on Artificial Intelligence, Networking and Information Technology (AINIT) (pp. 661-666). IEEE.
[17]He, K., Gkioxari, G., Dollár, P. and Girshick, R., 2017. Mask r-cnn. In Proceedings of the IEEE international conference on computer vision (pp. 2961-2969).
[18] Lin, T.Y., Dollár, P., Girshick, R., He, K., Hariharan, B. and Belongie, S., 2017. Feature pyramid networks for object detection. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 2117-2125).
[19] Liu, Y., Sun, P., Wergeles, N. and Shang, Y., 2021. A survey and performance evaluation of deep learning methods for small object detection. Expert Systems with Applications, 172, p.114602.
[20] Xiao, Y., Xu, Y., Zhong, Z., Luo, W., Li, J. and Gao, S., 2021, May. Amodal segmentation based on visible region segmentation and shape prior. In Proceedings of the AAAI Conference on Artificial Intelligence (Vol. 35, No. 4, pp. 2995-3003).
[21] R. Padilla, S. L. Netto and E. A. B. da Silva, "A Survey on Performance Metrics for Object-Detection Algorithms," 2020 International Conference on Systems, Signals and Image Processing (IWSSIP), Niteroi, Brazil, 2020, pp. 237-242, doi: 10.1109/IWSSIP48289.2020.9145130.
[22]Xiao, Y., Tian, Z., Yu, J., Zhang, Y., Liu, S., Du, S. and Lan, X., 2020. A review of object detection based on deep learning. Multimedia Tools and Applications, 79, pp.23729-23791.
[23] Bideaux, M., Phe, A., Chaouch, M., Luvison, B. and Pham, Q.C., 2024. 3D-COCO: extension of MS-COCO dataset for image detection and 3D reconstruction modules. arXiv preprint arXiv:2404.05641.
[24] Everingham, M., Van Gool, L., Williams, C.K., Winn, J. and Zisserman, A., 2010. The pascal visual object classes (voc) challenge. International journal of computer vision, 88, pp.303-338.
[25] R. Padilla, S. L. Netto and E. A. B. da Silva, "A Survey on Performance Metrics for Object-Detection Algorithms," 2020 International Conference on Systems, Signals and Image Processing (IWSSIP), Niteroi, Brazil, 2020, pp. 237-242, doi: 10.1109/IWSSIP48289.2020.9145130.
[26] Nguyen, N.D., Do, T., Ngo, T.D. and Le, D.D., 2020. An evaluation of deep learning methods for small object detection. Journal of electrical and computer engineering, 2020, pp.1-18.