International Journal of Advances in Applied Sciences (IJAAS)
Vol. 13, No. 1, March 2024, pp. 188~196
ISSN: 2252-8814, DOI: 10.11591/ijaas.v13.i1.pp188-196

Detecting and identifying occluded and camouflaged objects in low-illumination environments

Gaytri Bakshi, Alok Aggarwal
School of Computer Science, University of Petroleum and Energy Studies, Dehradun, India

Article Info

Article history:

Received Sep 4, 2023
Revised Nov 4, 2023
Accepted Dec 27, 2023

Keywords:

Camouflage
Computer vision
Low-light
Neural network
Object detection

ABSTRACT

One of the prevailing areas of contemporary research involves the differentiation and identification of diverse objects within a given scene through automated systems. The field of study under consideration presents a multitude of obstacles, including but not limited to issues such as diminished lighting conditions, occlusion, and camouflage. The captured image exhibits variations in illumination, resulting in uneven brightness, reduced contrast, and the presence of noise. The fundamental basis of computer vision algorithms lies in the process of extracting features from datasets and subsequently discerning these features through neural networks. The task of extracting distinct feature key points from images captured under low lighting conditions is exceedingly challenging. To address this issue, the present study seeks to employ deep learning models to implement image enhancement techniques specifically designed for low-light conditions. The primary emphasis lies in obtaining key feature points that are differentiable, thereby enabling the utilization of this annotated data for specific tasks such as object detection. The task of identifying occluded and camouflaged objects has been successfully accomplished, yielding an impressive accuracy rate of 93% in total. The mean average precision has been achieved as 85% which is reasonably high compared to many earlier works.

Corresponding Author:

Gaytri Bakshi
School of Computer Science, University of Petroleum and Energy Studies
Bidholi, Dehradun, Uttarakhand 248007, India
Email: gaytri.b88@gmail.com

1. INTRODUCTION

Researchers have been engaged in the investigation of low-illuminated images for a period exceeding ten years. Attempting to derive an optimal methodology for addressing the challenges of occlusion, concealment, camouflage, and image reconstruction, as well as incorporating additional machine learning algorithms such as object detection, poses a laborious undertaking. The identification and recognition of human presence is a crucial task performed across various domains, including the security sector. This sector encompasses the safeguarding of residential areas, encompassing both internal pathways and houses, as well as community gardens and local roadways. This field encompasses the detection of facial features [1], [2] as well as the detection of humans and their activities. Blind spots and inadequate low-light illumination pose significant challenges in these particular regions. Criminal elements exploit these challenges and employ tactics to conceal their identities, thereby impeding the accurate detection capabilities of existing systems. Another area of focus is disaster management. While this field encompasses both natural and anthropogenic disasters, the optimization of these processes necessitates the implementation of intelligent and automated rapid technologies in order to safeguard human lives. This study primarily focuses on the detection of humans at long distances and in different climatic conditions [3]–[5].

Traffic analysis and management systems constitute a significant area of study, wherein researchers endeavor to effectively regulate road traffic and devise strategies to mitigate congestion and prevent vehicular accidents. Various attempts have been undertaken to establish systems aimed at standardizing traffic regulations and apprehending individuals who violate them [6], [7]. Autonomous vehicles represent a prominent forthcoming solution. The defense sector encompasses the domain that consistently necessitates the utilization of automated systems to ensure the security of border areas, battlefields, and rescue missions. This primarily pertains to the identification of individuals in complex situations [8]. In various domains, there are instances where images are captured under conditions of low illumination or predominantly in darkness. Conversely, numerous challenges arise during the detection process of the image due to the limited dynamic range exhibited by these images. The factors that contribute to a decrease in the detection confidence rate in low-light images encompass various concerns, including contrast, low brightness, excessive darkness, occlusion, and camouflage.

To counter these concerns many efforts have been made to improve the degraded image, but they either lose the quality or take too much time to go through different stages. The same goes for object detection, where speed is compromised when accuracy is in question. The method suggested in this study strikes a good balance between speed and precision. The two goals of this effort are to improve object visibility in images and to execute object detection using cutting-edge models for increasing detection accuracy. The task of identifying occluded and camouflaged objects has been successfully accomplished, yielding an impressive accuracy rate of 93% in total. The mean average precision has been achieved as 85% which is reasonably high compared to many earlier works. The remainder of this paper is structured as follows. The study conducted by preceding scholars in the subject area is succinctly summarized in section 2. Section 3 provides a description of the suggested methodology. The implementation of the suggested strategy is described in section 4. Results along with its validation are shown in section 5. The framework's performance analysis is shown in section 6. In section 7, the quantitative analysis is described. Section 8 finally concludes the work.

2. LITERATURE REVIEW

There have been several crude ways for object detection, such as hand-engineered filters or a cascade classifier that uses binary feedback and follows a single sequence like a cascade. ConvNets are an alternative to it and in case given sufficient training, they can pick up these filters and properties [9]. ConvNet’s architecture has characteristics comparable to the linked network of neurons in the human brain and has been molded by how the visual cortex is structured. One of the most innovative technologies in machine learning and artificial intelligence, particularly for image processing, is deep neural networks (DNN) [10]. To make smart systems, models were designed for embedded systems that were light in nature such as MobileNet. It is a compact DNN that uses depth-wise separable convolutions as part of its streamlined architecture [11].

Another convolutional neural network (CNN)-based technique called single shot detector (SSD) was proposed for object detection. In one pass, the single convolution network used in the SSD design learns to predict bounding box locations and classify those places [12]. To make up for the accuracy losses, SSD introduces a few enhancements like default boxes and multi-scale features. It has a high IoU rate, particularly when numerous objects are present in a group [13]. SSD and you look only once (YOLO) use a single shot to detect many items within the image as opposed to other algorithms based on quicker region-based convolutional neural networks (R-CNN) and traditional techniques like the harr cascade classifier. While YOLOv3 is quick accuracy needs to be improved. Similar to YOLO, SSD has a multi-box architecture and can distinguish between several classes of items in a group and extract more characteristics [14]. Even though models were generated that were based on CNN the issue with low illumination existed for autonomous systems. There have been a lot of studies done on human identification by autonomous systems, which have been suggested to enhance photographs, especially in low-light situations. While pixel-wise inversion, haze reduction, and histogram equalization are all effective methods, they are all filter-based approaches that make use of basic primitives [15].

Following that, there are a number of neural network-based applications that use CNN and generative adversarial networks (GAN). These approaches, however multi-scale, were unable to maintain the quality of the original image since the discriminator could break down and stop working [16], [17]. MIRNet features an interactive architecture despite being completely convolutional. Hence, they can identify a few objects correctly and cannot identify others. This work is an effort in the same direction to come up with solutions so that the proposed framework could be able to detect objects in low light and even counter the problems of occlusion, camouflage, and complex background with a good amount of precision as well as speed.

3. METHODOLOGY

To detect objects in a dark environment the proposed method firstly improves the image brightness while recovering the color and features, and secondly, the class of the object existing in it is predicted using an end-to-end learning method. This enables effective object identification and detection from the images taken in low-illumination situations. The methodology comprises two subnetworks in series. Both are based on convolutional architecture. The first is a multi-scale MIRNet interaction architecture and the other is a multi-scale, multi-box SSD architecture.

3.1. MIRNet

It is a feature extraction model that maintains the original high-resolution features to preserve fine spatial details while computing a complementary collection of features at various spatial scales. It is a frequently occurring information exchange process where the characteristics from several multi-resolution branches are gradually combined for better representation learning. A novel method for fusing features from different scales utilizing a selective kernel network that correctly maintains the original feature information at each spatial level while dynamically combining varying receptive fields. A recursive residual design enables the building of very deep networks by gradually decomposing the input signal to streamline the overall learning process.

3.2. Mobile Net single shot detector

As an effective CNN architecture created for mobile and embedded vision applications, MobileNet is an object detector that was introduced as a design that constructs lightweight DNN using tested depth-wise separable convolutions. The underlying MobileNetV2 network with an SSD layer that categorizes the detected image makes up the first section. In essence, the SSD layer uses the Mobile Net base network as a feature extractor to classify the object of interest. SSD [12] is a one-shot detector and is a neural network architecture created for detection purposes, which entails both classification localization (bounding boxes) simultaneously.

MobileNet [11], introduced by Google, is an efficient architecture that switches the standard convolution filters with the depth-wise convolution filters. The feature map generated by the standard convolution layer is shown by (1) [15]. Later to eliminate the interaction between the number of output channels and the size of the kernel, the feature map produced by depth-wise separable convolutions reduces the computational cost. The input channel with depth-wise convolution can be written as (2) [15].

$$G_{k,l,n} = \sum_{i,j,m} K_{i,j,m,n} \cdot F_{K+i-1,l+j-1,m}) \tag{1}$$

$$\hat{G}_{k,l,n} = \sum_{i,j,m} \hat{K}_{i,j,m,n} \cdot F_{K+i-1,l+j-1,m} \tag{2}$$

This depth-wise separable convolution layer is divided into both depth-wise and point-wise convolution layers. Each input channel receives a single filter using depth-wise convolutions. Pointwise convolution, one of the fundamental 11 convolution layers, is then used to integrate the output of the depth-wise layer linearly. MobileNets use batch norm and ReLU nonlinearities for both layers. This makes a lightweight hybrid model combining advanced methods and gives a good amount of speed and accuracy.

4. IMPLEMENTATION

4.1. Pre-processing module

The LOw-Light (LOL) dataset is acquired to perform the image enhancement under this module. The LOL dataset has 500 low-light images. The dataset offers 485 training photos and 15 test images. A low-light input image and its associated well-exposed reference image make up each pair of images in the dataset. To create a tensor flow dataset the input dataset images are pre-processed. The dataset images are resized with a resolution of 128×128 to be sent to the enhancement module.

4.2. Image enhancement

The pre-processed image dataset is enhanced using the MIRNet model under this module. The MIRNet model is trained for low-light image enhancement on a pre-processed tensor flow dataset created from the LOL dataset. The selective feature fusion kernel (SKFF) function adjusts receptive fields dynamically using the FUSE and SELECT actions. By maintaining high resolution while accepting the rich contextual data from low resolution, the multiscale residual block (MRB) is employed to produce an output that is spatially consistent. In order to retain the residual model, MRB essentially executes down-sampling and up-sampling procedures. Adam optimizer with a learning rate of 1e-4 and char bonnier loss as the loss function are used to train MIRNet. Peak signal noise ratio (PSNR) is another term for the ratio of a signal's maximum possible value to the strength of distorted noise that degrades the quality of an image. The saved model is used as a pre-trained model to obtain prediction and enhancement results for a low-light image.

4.3. Object detection module

A hybrid MIRNet+MobileNet SSD framework is designed under this module, where the enhanced image from MIRNet is passed as an input image for object detection. The MobileNet SSD is pre-trained for classification on the common objects in context (COCO) dataset. The classification of objects in an image is done frame by frame. For classification, the SSD algorithm is used in MobileNet architecture, and the pre-trained weights file is saved as a TensorFlow Inference graph passing through the network architecture. The feature extraction algorithm passes all images and classifies them into various classes. Suppressing the detection with an accuracy score of less than 50% [net.detect (img,confThreshold=0.5)]. The accuracy is also increased by implementing non-maximum suppression. Finally, bounding boxes are created and displayed with class names and detection confidence percentages. The architecture of the framework is represented in Figure 1 and the workflow of the framework is presented in Figure 2.

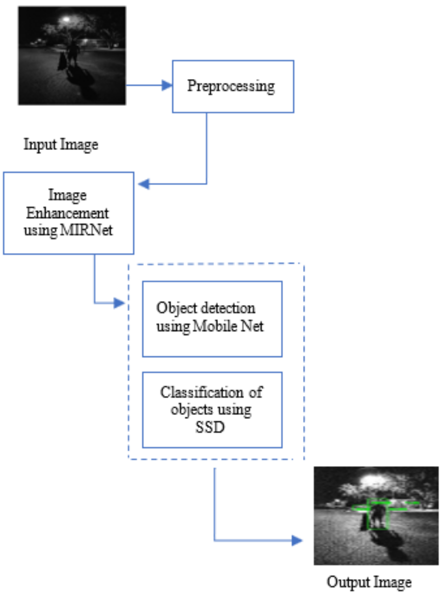


Figure 1. The architecture of the hybrid MIRNet+MobileNet SSD framework

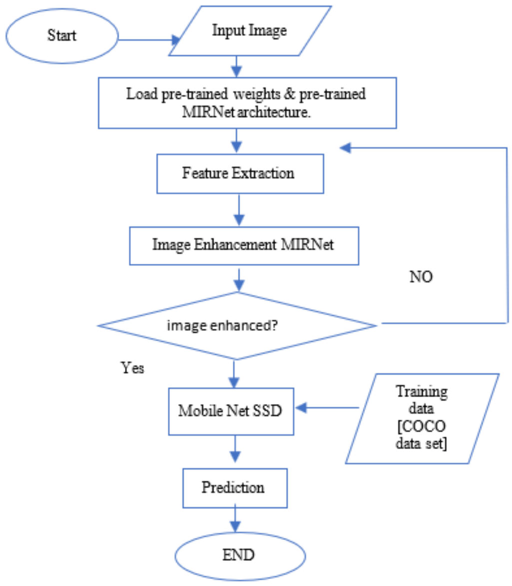


Figure 2. Workflow of hybrid MIRNet+MobileNet SSD model framework

5. VALIDATION AND RESULTS

The image enhancement module is trained on 200 images from the pre-processed dataset, and it is then validated on 10 test images of low-light outdoor scenarios. The training progress is saved after every epoch in TensorFlow checkpoints. After minimizing the loss function sufficiently, the training was stopped, and the model was saved. The model has been trained over 50 epochs minimizing loss function and maximizing PSNR to 0.1090 and 67.5139 respectively and the same for the validation loss to 0.1124 and 67.1488 respectively. The following training and validation losses are shown in Tables 1 and 2. The exclusively dark (ExDark) dataset is used to train and test the entire hybrid MIRNet+MobileNet SSD framework. The ExDark dataset consists of 7,363 low-light photos, with 10 different situations ranging from very low light to twilight [18] Out of these images, a total of 1,000 images were taken in an outdoor environment which was divided in an 80:20 ratio for both training and testing. So, the entire network was trained on 800 images and tested on 200 low-light images. Figures 6 to 8 show the evaluated results of normal dark images converted to enhanced images as well as object detection.

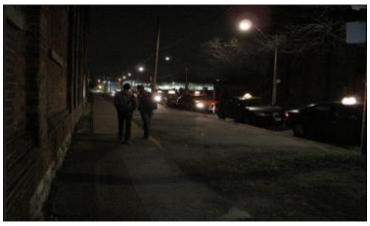
Figure 3 displays the enhanced image with detection and classification of objects with probabilities of 65%, 54% as a person, and 51%, 50% as a car. Figure 4 showcases the image enhancement with the detection and classification of objects into classes. The hybrid model detected a person in a scene with probabilities of 82% and 70%, a car with 66%, and misclassification of a car to a bus with 59%. Figure 5 displays a case of misclassification where the model identifies the shadow of a person as a skateboard with probabilities to detect humans is 77% and the misclassification after image enhancement is 50%.

Table 1. Training data of enhancement module

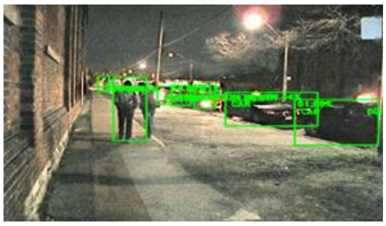
Epochs	Train_Loss	Train_PSNR
5	0.1651	63.6555
10	0.1539	64.3999
15	0.1340	65.5611
20	0.1273	66.0817
25	0.1288	66.0734
30	0.1275	66.3542
35	0.1191	66.7690
40	0.1125	67.1694
45	0.1076	67.6359
50	0.1090	67.5139

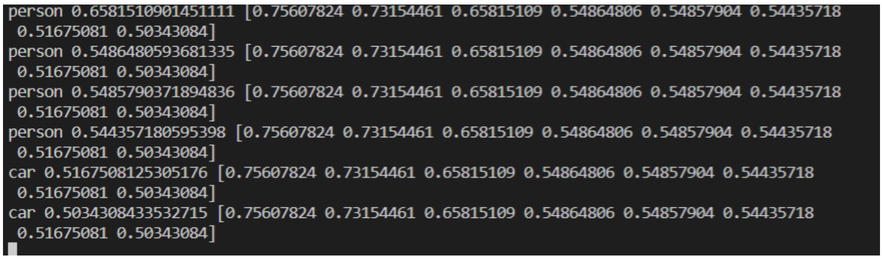
Table 2. Validation data of enhancement module

Epochs	Val_Loss	Val_PSNR
5	0.1333	65.6338
10	0.1220	66.7203
15	0.1111	67.2009
20	0.1185	67.0208
25	0.1027	67.9508
30	0.1034	67.4624
35	0.1043	67.4840
40	0.1034	67.6437
45	0.1103	67.2720
50	0.1124	67.1488


(a)


(b)


(c)



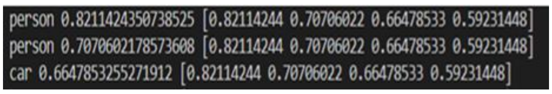
(d)

Figure 3. Resulted from image 1 passed by the hybrid model (a) original low illuminated input image, (b) enhanced image, (c) object detection after enhancement, and (d) detection score of objects within a scene


(a)

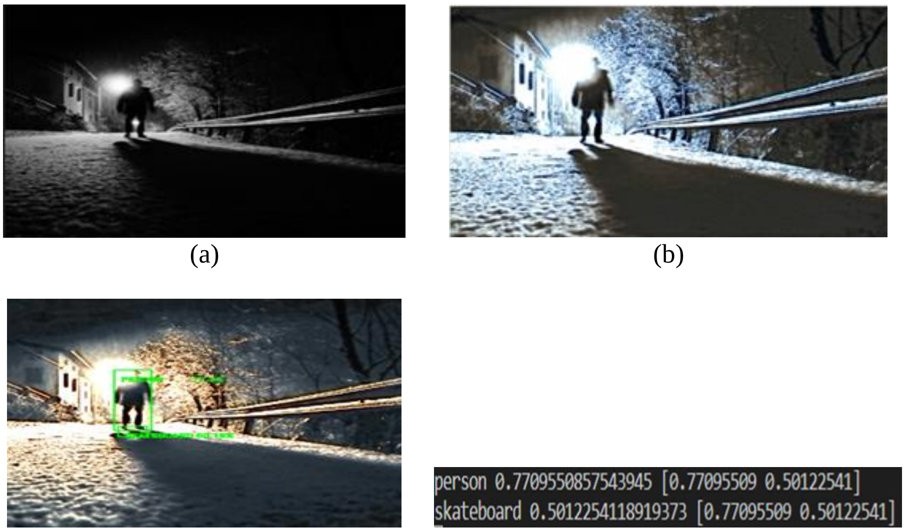

(b)


(c)



(d)

Figure 4. Resulted from image 2 passed by the hybrid model (a) original low-illuminated input image, (b) enhanced image, (c) object detection after enhancement, and (d) detection score of objects within a scene



(a) (b) (c) (d)

Figure 5. Resulted from image 2 of image enhancement with object detection (a) original input image, (b) enhanced image, (c) object detection after enhancement, and (d) detection score of objects within a scene

6. PERFORMANCE ANALYSIS

6.1. Confusion matrix

Firstly, it is evaluated on the confusion matrix which is shown in Figure 6. This gives an estimation of how efficiently this multiclass model works for the detection of each class. The diagonal blue colored boxes represent the true positive evaluations of each class with the number of times the model detects the particular class of object. The confusion matrix even describes the cases of misclassification for some classes such as car is misclassified as truck and bus.

6.2. Precision

Another metric is based on precision calculation. The mean average precision achieved by this hybrid model is 85% as shown in Table 3. with precision values of each class. The precision obtained for each class by the model is represented by the graph shown in Figure 7.

6.3. Accuracy and F1 score

Thirdly the overall accuracy is evaluated by an alternative machine learning evaluation statistic called F1 score which evaluates a model's prediction ability by focusing on performance concerning each class. This is well described in Table 4 and Figure 8 concerning all the classes. Another metric is accuracy. The accuracy statistic counts the number of times a model accurately predicted the whole dataset. Table 4 well depicts that the overall accuracy achieved is 93%.

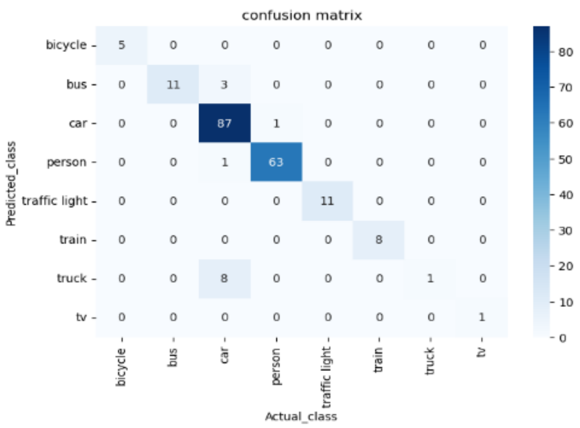


Figure 6. Confusion matrix

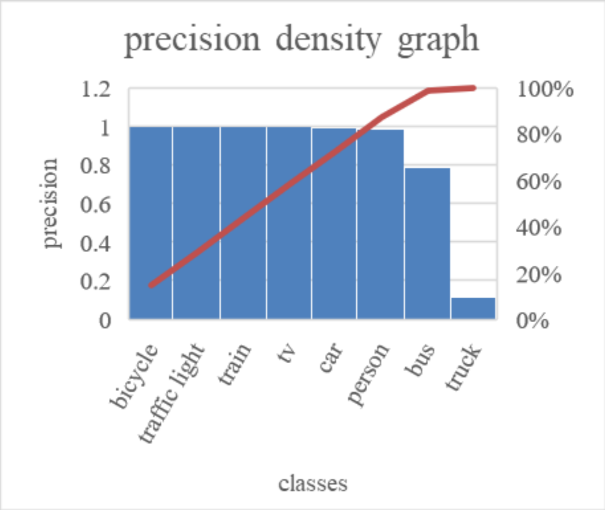


Figure 7. Precision density graph

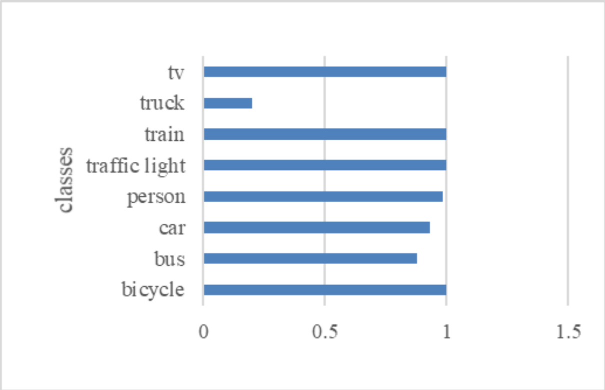


Figure 8. Class-wise F1 score evaluation

Table 3. Mean average precision

Classes	Precision	mAP
Bicycle	1	
Bus	0.785714	
Car	0.988636	
Person	0.984375	0.85873
Traffic light	1	
Train	1	
Truck	0.111111	
Tv	1	

Table 4. Metrics of hybrid model

Classes	Precision	Recall	F1-score	Support
Bicycle	1	1	1	5
Bus	0.785714	1	0.88	11
Car	0.988636	0.878788	0.930481	99
Person	0.984375	0.984375	0.984375	64
Traffic light	1	1	1	11
Train	1	1	1	8
Truck	0.111111	1	0.2	1
Tv	1	1	1	1
Accuracy	0.935	0.935	0.935	0.935
Macro avg	0.85873	0.982895	0.874357	200
Weighted avg	0.973145	0.935	0.949988	200

7. QUANTITATIVE ANALYSIS

Illumination is one of the prominent traits in images. With the variation in illumination in any landscape, detection by autonomous systems gets affected and can even become difficult. This work proposes a hybrid model that would tackle these issues as it has a fusion of an enhancement model with a state-of-the-art object detection model. Table 5 provides a comparative analysis of the state-of-the-art models with the proposed hybrid model on the metrics of mean average precision.

Table 5. Comparative analysis of the hybrid model with state-to-the-art model

S. No.	Methodology	mAP (%)
1	NVD+FPN [19]	35.3
2	Faster R-CNN [20]	77.8
3	Retina Net [21], [22]	75.2
4	YOLOv3- enhanced [23]	78.0
5	RetinaMFANet [24]	78.9
6	RetinaNet [25]	58.1
7	Faster-RCNN [26]	54.3
8	LFA+FFN+CSDH [27]	64.6
9	Our hybrid model	85

8. CONCLUSION

The ability of automated systems to discern and identify various objects within a given scene is considered to be a highly significant area of research. Additionally, the system encounters various challenges such as inadequate illumination, occlusion, and the potential for objects to blend into their surroundings. The image acquired exhibits the presence of noise, diminished contrast, and inconsistent brightness due to the fluctuating lighting conditions. Images captured under poor lighting conditions pose significant challenges for the system to accurately extract the salient features. The accurate identification and prediction of specific feature key points in photographs captured under poor lighting conditions pose a significant challenge for automated systems. The present study employs deep learning models to achieve image enhancement in low-light conditions and endeavors to propose a hybrid model for enhancing low-light images and subsequently detecting objects within a scene. The primary objective is to obtain key feature points that are differentiable, as this enables the utilization of labeled data in more specialized tasks such as object detection. This approach presents a novel methodology for surmounting challenges and attaining enhanced outcomes in terms of precision. An overall accuracy rate of 93% has been achieved in the detection of obscured and disguised objects. The mean average precision has been achieved as 85% which is reasonably high compared to many earlier works.

ACKNOWLEDGMENTS

Without the assistance of the University of Petroleum and Energy Studies, this work would not have been feasible. To ensure the effective implementation and validation of the research contributions outlined, we would like to express my gratitude to the Machine Intelligence Research Centre (MIRC) of the University for their help in the form of knowledge exchange, discussions, and computing GPU infrastructure.

REFERENCES

[1] A. Kumar, A. Kaur, and M. Kumar, "Face detection techniques: a review," *Artificial Intelligence Review*, vol. 52, no. 2, pp. 927–948, Aug. 2019, doi: 10.1007/s10462-018-9650-2.

[2] G. Bakshi, A. Aggarwal, D. Sahu, R. R. Baranwal, G. Dhall, and M. Kapoor, "Age, gender, and gesture classification using open-source computer vision," *Lecture Notes in Networks and Systems*, vol. 490, pp. 63–73, 2023, doi: 10.1007/978-981-19-4052-1_8.

[3] A. S. Mohammed, A. Amamou, F. K. Ayevide, S. Kelouwani, K. Agbossou, and N. Zioui, "The perception system of intelligent ground vehicles in all weather conditions: a systematic literature review," *Sensors (Switzerland)*, vol. 20, no. 22, pp. 1–34, Nov. 2020, doi: 10.3390/s20226532.

[4] G. Bakshi and A. Aggarwal, "Computational approaches to detect human in multifaceted environmental conditions using computer vision and machine intelligence-a review," in *Proceedings of the 2022 3rd International Conference on Intelligent Computing, Instrumentation and Control Technologies: Computational Intelligence for Smart Systems, ICICICT 2022*, IEEE, Aug. 2022, pp. 1547–1554. doi: 10.1109/ICICICT54557.2022.9917742.

[5] S. Sambolek and M. Ivasic-Kos, "Automatic person detection in search and rescue operations using deep CNN detectors," *IEEE Access*, vol. 9, pp. 37905–37922, 2021, doi: 10.1109/ACCESS.2021.3063681.

[6] Gaytri, R. Kumar, and U. Rajnikanth, "A smart approach to detect helmet in surveillance by amalgamation of IoT and machine learning principles to seize a traffic offender," 2021, pp. 701–715. doi: 10.1007/978-981-15-7533-4_55.

[7] X. Wang *et al.*, "Real-time and efficient multi-scale traffic sign detection method for driverless cars," *Sensors*, vol. 22, no. 18, pp. 1–12, Sep. 2022, doi: 10.3390/s22186930.

[8] D. K. Singh and D. S. Kushwaha, "Automatic intruder combat system : away to smart border surveillance," *Defence Science*

Journal, vol. 67, no. 1, pp. 50–58, Dec. 2017, doi: 10.14429/dsj.67.10286.
[9] S. Guennouni, A. Ahaitouf, and A. Mansouri, "A comparative study of multiple object detection using haar-like feature selection and local binary patterns in several platforms," *Modelling and Simulation in Engineering*, pp. 1–8, 2015, doi: 10.1155/2015/948960.
[10] K. G. Kim, "Book review: deep learning," *Healthcare Informatics Research*, vol. 22, no. 4, pp. 351–354, 2016, doi: 10.4258/hir.2016.22.4.351.
[11] A. G. Howard *et al.*, "MobileNets: efficient convolutional neural networks for mobile vision applications," *Computer Vision and Pattern Recognition*, vol. 14, no. 2, pp. 53–57, 2009.
[12] W. Liu *et al.*, "SSD: single shot multibox detector," in *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, vol. 9905 LNCS, 2016, pp. 21–37. doi: 10.1007/978-3-319-46448-0_2.
[13] J. Huang *et al.*, "Speed/accuracy trade-offs for modern convolutional object detectors," *Proceedings - 30th IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2017*, pp. 3296–3305, 2017, doi: 10.1109/CVPR.2017.351.
[14] J. Du, "Understanding of object detection based on CNN family and YOLO," *Journal of Physics: Conference Series*, vol. 1004, no. 1, pp. 1–9, Apr. 2018, doi: 10.1088/1742-6596/1004/1/012029.
[15] Z. Gu, C. Chen, and D. Zhang, "A low-light image enhancement method based on image degradation model and pure pixel ratio prior," *Mathematical Problems in Engineering*, pp. 1–19, Jul. 2018, doi: 10.1155/2018/8178109.
[16] L. W. Wang, W. C. Siu, Z. S. Liu, C. T. Li, and D. P. K. Lun, "Deep relighting networks for image light source manipulation," in *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, vol. 12537 LNCS, 2020, pp. 550–567. doi: 10.1007/978-3-030-67070-2_33.
[17] L. W. Wang, Z. S. Liu, W. C. Siu, and D. P. K. Lun, "Lightening network for low-light image enhancement," *IEEE Transactions on Image Processing*, vol. 29, pp. 7984–7996, 2020, doi: 10.1109/TIP.2020.3008396.
[18] Y. P. Loh and C. S. Chan, "Getting to know low-light images with the exclusively dark dataset," *Computer Vision and Image Understanding*, vol. 178, pp. 30–42, Jan. 2019, doi: 10.1016/j.cviu.2018.10.010.
[19] Y. Xiao, A. Jiang, J. Ye, and M. W. Wang, "Making of night vision: object detection under low-illumination," *IEEE Access*, vol. 8, pp. 123075–123086, 2020, doi: 10.1109/ACCESS.2020.3007610.
[20] 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*, pp. 91–99, 2015.
[21] B. Yue, L. Chen, H. Shi, and Q. Sheng, "Ship detection in SAR images based on improved RetinaNet," *Journal of Signal Processing*, vol. 38, no. 1, pp. 128–136, 2022.
[22] L. Ge, Z. Yelong, and Z. Meirong, "Vehicle information detection based on improved RetinaNet," *Journal of Computer Applications*, vol. 40, no. 3, p. 854, 2020.
[23] J. Redmon and A. Farhadi, "YOLOv3: an incremental improvement," 2018, doi: 10.48550/arXiv.1804.02767.
[24] R. Ma, K. Bao, and Y. Yin, "Improved ship object detection in low-illumination environments using RetinaMFANet," *Journal of Marine Science and Engineering*, vol. 10, no. 12, pp. 1–16, Dec. 2022, doi: 10.3390/jmse10121996.
[25] T. Y. Lin, P. Goyal, R. Girshick, K. He, and P. Dollar, "Focal loss for dense object detection," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 42, no. 2, pp. 318–327, 2020, doi: 10.1109/TPAMI.2018.2858826.
[26] K. Wang and M. Z. Liu, "Object recognition at night scene based on DCGAN and faster R-CNN," *IEEE Access*, vol. 8, pp. 193168–193182, 2020, doi: 10.1109/ACCESS.2020.3032981.
[27] S. Wu, Z. Liu, H. Lu, and Y. Huang, "Shadow hunter: low-illumination object-detection algorithm," *Applied Sciences (Switzerland)*, vol. 13, no. 16, pp. 1–17, Aug. 2023, doi: 10.3390/app13169261.

BIOGRAPHIES OF AUTHORS

Gaytri Bakshi is an Assistant Professor at the University of Petroleum and Energy Studies, Dehradun (UK) India. Currently pursuing a Ph.D. Earned master's degree (M.Tech.) in Computer Science and Engineering in 2015 and a Bachelor's degree in Computer Science in 2012 from Kurukshetra University. Skilled in research areas of IoT, machine learning, and deep learning. She can be contacted at email: gaytri.b88@gmail.com.

Alok Aggarwal is currently spearheading efforts as a Professor (CSE), with the University of Petroleum and Energy Studies, Dehradun (UK) India. Earned Ph.D. degree in Mobile Computing area from IIT Roorkee, Roorkee India in 2010, Master's degree (M.Tech.) in Computer Science and Engineering in 2001, and Bachelor's degree in Computer Science in 1995. Contributing over 20+ years in Teaching (CSE and IT) and as a software engineer for 4 years. He has contributed 250+ research articles in various journals and conference proceedings and authored 5 books with an H-index of 26. His areas of specialization are mobile computing, mobile Ad-hoc networks, lightweight cryptography, and the Internet of Things. He can be contacted at email: alok.aggarwal@ddn.upes.ac.in.