FOREX Publication
Open Access | Rapid and quality publishing

International Journal of Electrical and Electronics Research (IJEER)
Research Article | Volume 13, Issue 3 | Pages 393-401 | e-ISSN: 2347-470X

An Efficient Method for Object Detection with Automatic Illumination Correction

Chinthakindi Kiran Kumar[1], Gaurav Sethi[2], and Kirti Rawal[3]

[1]*School of Electronics & Electrical Engineering, Lovely Professional University, India; ckkmtech11@gmail.com*
[1]*Department of Electronics & Communication Engineering, Malla Reddy College of Engineering & Technology, Hyderabad, India;*
[2]*School of Electronics & Electrical Engineering, Lovely Professional University, India; gaurav.11106@lpu.co.in*
[3]*School of Electronics & Electrical Engineering, Lovely Professional University, India; kirti.20248@lpu.co.in*

*Correspondence: *ckkmtech11@gmail.com*

ABSTRACT- Recent advancements in artificial intelligence and computer vision have led to the automation of many conventional surveillance techniques, especially in smart city applications. However, object detection systems often struggle in poorly lit environments and may suffer from slow processing speeds. To address these restrictions, this paper proposes an adaptive illumination correction technique using an enhanced Logarithmic Image Processing model. The method improves object visibility in low-light video frames and enhances detection performance. Additionally, a customized deep convolutional neural network is developed to accurately detect objects after applying the illumination correction. The combined framework is evaluated on standard datasets and demonstrates superior robustness to varying illumination conditions compared to existing state-of-the-art methods. The results show significant improvements in metrics such as accuracy, recall, precision, and F-measure, proving the effectiveness of the proposed approach.

Keywords: Illumination correction, LIP model, Deep Network, Object detection.

ARTICLE INFORMATION

Author(s): Chinthakindi Kiran Kumar, Gaurav Sethi, Kirti Rawal;
Received:09/02/25; **Accepted:** 26/06/25; **Published:** 25/08/25;
E- ISSN: 2347-470X;
Paper Id: IJEER 0902-03;
Citation: 10.37391/ijeer.130302
Webpage-link:
https://ijeer.forexjournal.co.in/archive/volume-13/ijeer-130302.html

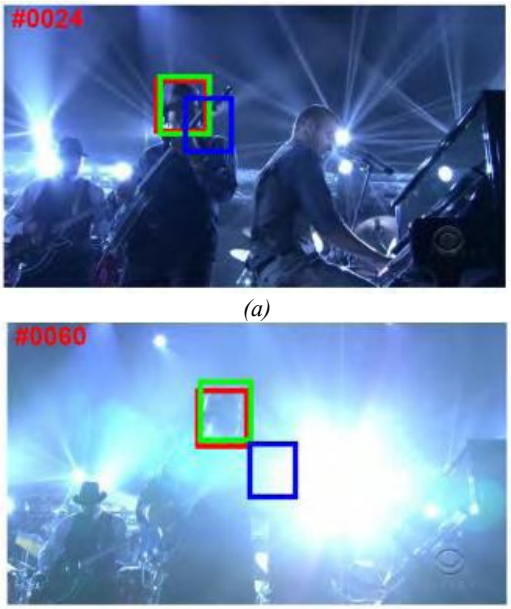
Publisher's Note: FOREX Publication stays neutral with regard to jurisdictional claims in Published maps and institutional affiliations.

1. INTRODUCTION

With the fast growth of smart cities, there is a growing mandate for reliable and intelligent surveillance systems that can detect and track objects effectively to ensure public safety. Intelligent Video Surveillance (IVS) has emerged as a crucial component in this domain, enabling automated monitoring, anomaly detection, and alert generation in real-time scenarios [1]. IVS systems typically consist of several stages: video acquisition [2], target recognition [3], object tracking [4], anomaly detection [5], and data analysis [6].

When deployed on embedded platforms, IVS solutions often utilize fog and edge computing to enable low-latency decision-making. These systems are particularly useful in public areas where prompt identification of unusual behavior is essential. Object detection and tracking, which form the core of IVS, have been extensively studied, with many algorithms developed to improve accuracy, reduce false positives, and adapt to environmental challenges [7-10].

Despite these advancements, a major limitation of current detection and tracking algorithms is their poor performance under varying illumination conditions. Real-world video footage often suffers from lighting inconsistencies due to factors such as shadows, glare, or low-light environments. As shown in *figure 1*, changes in illumination can lead to incorrect detection results even with advanced models.



(a)

(b)

Figure 1. (a) Object detection red box represents ground truth; blue and green boxes represent detection outputs (b) The impact of illumination change is shown as object is wrongly identified

FOREX Publication
Open Access | Rapid and quality publishing
International Journal of Electrical and Electronics Research (IJEER)
Research Article | Volume 13, Issue 3 | Pages 393-401 | e-ISSN: 2347-470X

To address this challenge, recent works have attempted to use Discriminative Correlation Filters (DCF) along with Histogram of Oriented Gradients (HOG) features for detection [11]. However, such methods still exhibit significant performance degradation in low-light scenarios. Therefore, an effective pre-processing technique that corrects illumination before detection is crucial for improving robustness.

This paper introduces a two-part solution: (*i*) an adaptive illumination correction algorithm based on an improved Logarithmic Image Processing (LIP) model, and (*ii*) a deep convolutional neural network tailored for object detection in videos. The proposed framework enhances contrast in underexposed frames while preserving high-intensity regions, and leverages the corrected frames to achieve more accurate detection and segmentation results.

2. ADAPTIVE ILLUMINATION CORRECTION (AIC) MODEL FOR VIDEO SEQUENCES

The Adaptive Illumination Correction (AIC) model is developed to enhance image visibility in video frames affected by poor and uneven lighting. These frames often suffer from limited dynamic range, spatial inconsistencies, and increased noise levels, which collectively degrade the performance of object detection and tracking systems. The AIC process involves a sequence of transformations to boost low and mid-level intensities, while preserving or slightly reducing higher ones to avoid saturation.

Images or video frames captured under low-light conditions often exhibit a limited dynamic range, which may not align well with the capabilities of the sensors or display devices. To address this mismatch, it is advisable to normalize the pixel intensity values at the initial stage. In many cases, a simple linear normalization of the pixel values to the range [0 1] is sufficient. This can be achieved using a transformation based on the following *eq. (1)*.

$$L_0(i) = \frac{L_i(k) - \min(L_i(k))}{[\max(L_i(k)) - \min(L_i(k))]} \quad (1)$$

Where Lo is the output normalized intensity values of the input L_1 intensities. To compress the higher intensity values, a logarithmic transformation can be applied to the normalized pixel values. This operation can be expressed using *eq. (2)*.

$$L_l(i) = \frac{\max(L_0(i))}{\log(\max(L_0(i)) + 1)} \log(L_0(i) + 1) \quad (2)$$

This function is comparable to the alteration achieved by human retina and this type of adjustment increases intensity values of dark pixels while decreasing intensity of high values [12]. In order to additional adjust the above equation a nonlinear exponential mapping function to adjust the native contrast values and reduce the brighter pixel values. An increasing nonlinear function utilized for mapping is defined in *eq. (3)*.

$$T(x) = 1 - \exp(-x) \quad (3)$$

An earlier LIP model was created using calculation of the amount of light travelling though filter. The classical LIP model proposed by Lim has numerous limitations like upper range overspill due to mathematical properties that deals the model. Several researchers [13], [14], [15] have proposed parametric extensions of this model, offered enhanced flexibility and enabled a wide range of applications. Several LIP-based models have been developed to combine features from multiple images. However, in this study, the mathematical framework of the LIP model is specifically utilized. The proposed work focuses to retain the acceptable quality images from various images and video sequences that were acquired under varying illumination. The algorithm aims to enhance the low and middle dark intensity values while not altering the higher intensity values.

Let the original pixel intensities of the input image be denoted by 'x'. These values are first processed using *equations (1)* and *(2)*, and the resulting transformed image is then represented as follows.

$$I_1 = \frac{\max(X)}{\log(\max(X)+1)} * \log(X + 1) \quad (4)$$

Alternatively, the same input values can be adjusted using an exponential function, which helps maintain local contrast while reducing the influence of higher intensity values.

$$I_2 = (1 - \exp(-X)) \quad (5)$$

In this study, the two transformed images, I_1 and I_2, are combined using an appropriate LIP-based model. Specifically, image addition is employed, resulting in the final transformed image as defined in *eq. (6)*.

$$I_3 = \frac{I_1 + I_2}{\lambda + (I_1 * I_2)} \quad (6)$$

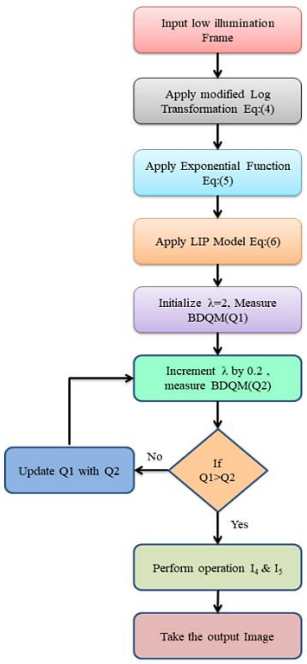


Figure 2. Flow chart of the Adaptive Illumination Correction (AIC) model

FOREX Publication
Open Access | Rapid and quality publishing
International Journal of Electrical and Electronics Research (IJEER)
Research Article | Volume 13, Issue 3 | Pages 393-401 | e-ISSN: 2347-470X

Once the final image I_3 is generated, it undergoes an additional enhancement step. In this work, the parameter λ is adaptively determined using BDQM estimation, as outlined in [16]. Although I_3 exhibits an overall increase in brightness due to additive processing, it still lacks sufficient high-intensity regions. Therefore, further enhancement is necessary to achieve a visually balanced appearance. To address this, an improved cumulative distribution function (CDF) based on the hyperbolic secant distribution [17] is applied. This method, characterized by an S-shaped curve, effectively adjusts the brightness and contrast levels, thereby enhancing the image's overall illumination. As a result, the image, derived using *eq. (7)*, is further transformed as follows.

$$I_4 = \text{erf}(\lambda * atan(exp(I_3)) - 0.5 * I_3) \quad (7)$$

In the given equation, the parameter 'λ' represents the optimized value obtained through an iterative procedure, and I_4 denotes the corresponding enhanced image using this value. This adjustment contributes to improved processing efficiency of the enhancement function, particularly in terms of illumination correction. The inclusion of the error function enhances the curvature of the hyperbolic secant distribution, which proves effective in brightening the darker regions of the image. Furthermore, the constant $\pi/2$ is substituted with the adaptive parameter 'λ' to regulate the intensity of enhancement. Finally, a normalization step is applied to produce the final enhanced image, denoted as I_5.

$$I_5 = \frac{I_4 - \min(I_4)}{\max(I_4) - \min(I_4)} \quad (8)$$

2.1. Pseudo code of the AIC

Input: Low-light image or video frame
Output: Enhanced image
Step 1: Normalize the pixel values of the input image to a standard range.
Step 2: Apply a logarithmic transformation to enhance darker regions.
Step 3: Apply an exponential mapping to improve contrast in mid-tone intensities.
Step 4: Combine the outputs of the logarithmic and exponential enhancements using image addition.
Step 5: Estimate the optimal enhancement parameter (lambda) using a blind image quality metric.
Step 6: Apply global contrast correction using a CDF based on the hyperbolic secant.
Step 7: Further refine brightness using an error function-based transformation.
Step 8: Normalize the final image to ensure output intensity values are in a valid display range.

3. DEEP NETWORKS FOR OBJECT DETECTION

Deep Neural Networks (DNN) for object detection has recently acquired popularity due to their capacity to learn. These networks achieve high accuracy by learning parameters by themselves. DNN's have produced outstanding results in numerous computer vision applications like object identification, segmentation and classification. [18]-[22]. However, due to illumination changes, blur and video defocus other occlusions straightaway associating these images level models to object recognition in videos is problematic. It is very natural to accept video objectors to be more powerful than still pictures detectors since videos comprises more detail evidence about the same object occurrence than image. Regardless of these difficulties it is necessary to develop a model that properly uses temporal information of the videos.

DNNs based on Convolutional Neural Networks have successfully illustrated promising outcomes in Background Separation challenges under difficult environmental circumstances. The authors in [23] have presented CNN to perform segmentation task for video sequences. This approach has the benefit of separating foreground items regardless of their temporal features. Here due to the independence of incoming samples, segmentation using CNN may be accomplished using only spatial attributes. CNN is employed to comprehend the distinctive aspects, which are consequently sent to a classifier to segment stirring objects. Finally, the segmented section is extracted using fully connected framework. Similarly, the authors of [24] described a threesome CNN based solution for multi-stage background feature implanting, utilizing encoder and decoder architecture. In this analysis the encoder part of pre-trained network such as VGG16 is altered within a triplet framework to divide the picture into different scales and extract the features when just a very small number of samples are required.

Several authors have presented various approaches for object recognition including RCNN (Region based CNN) its variation that relay on region and these approaches are precise but time overriding [25]-[26]. The second kind is YOLO (You only look once) and its derivatives which are noted for their speed and effectiveness [27] - [28]. The third type however combines the advantages of first two alternatives and is based on SSD (Single shot multi box detector) [29]-[30]. However, these networks have significant shortcomings when it comes to detecting moving objects. These comprise of processing of high-resolution frames and the removal of immovable object from the video sequence.

Most conventional models (RCNNs, YOLO variants, SSDs) are optimized either for precision (RCNN family) or speed (YOLO), but often fail to handle the temporal continuity and lighting inconsistency in real-world video streams. Our network addresses this by;
• Utilizing only labeled data to avoid overfitting on unlabeled frames.
• Integrating AIC preprocessing, ensuring object visibility is maximized regardless of lighting.
• Employing a Directed Acyclic Graph (DAG) structure for non-linear and parallel feature extraction, allowing richer representations compared to strictly sequential CNNs.

The DAG framework is chosen over traditional sequential CNNs because it supports multi-path feature flow, enabling complex operations like skip connections, multi-scale fusion,

FOREX Publication
Open Access | Rapid and quality publishing
International Journal of Electrical and Electronics Research (IJEER)
Research Article | Volume 13, Issue 3 | Pages 393-401 | e-ISSN: 2347-470X

and fine-grained segmentation without excessive depth or computation.

4. PROPOSED FRAME WORK

The complete procedure flow of the proposed methodology is depicted in *figure 3*. The input video frames may or may not be occluded; nonetheless, the input video is processed for frame separation. and if the frame is poorly illuminated then the frame is processed for automatic illumination correction (AIC), if the contrast values of video frame is ideal then it focused for object recognition using DNN as shown is *figure 4*.

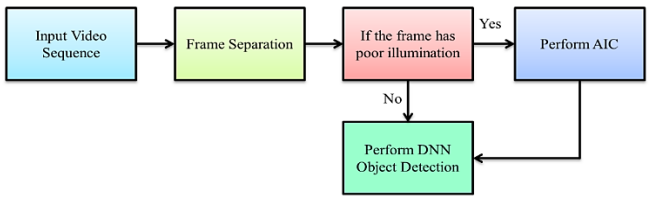


Figure 3. Generalized block diagram of the proposed approach (AICOD)

The collected frames are resized to resolution of 320x240 which corresponds to the full D_1 video standard in true color RGB presentation [31]. The sequence of video frames and its labelled images are supplied parallel to train the DNN. At the beginning, the input images are divided into 16x16 chunks. These blocks in parallel with labelled blocks are forwarded to DAG network which is of 31 layers. For training the deep network huge number of samples are required, which is created during the data augmentation process. During this process the samples are rotated and scaled; so that thousands of samples representing the object region and back ground region can be generated under multiple scenarios. Along with the training samples the respective labeled images are provided to the network making it to differentiate between the background and foreground region.

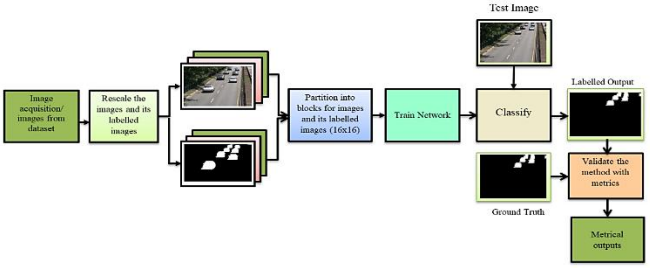


Figure 4. Object detection framework proposed with customizable DNN

The network architecture is represented in *figure 5*; the encoder depth is 2 and consists of many layers where the input frame is fed to the input image layer followed with 2D convolutional layer, Batch normalization, rectified linear unit and max-pooling. The layers input dimension is 16x16x3 and it is transmitted to 2D Convolutional Layer with a filter size of 3x3. The Max-pooling layer is used to do down sampling of components by separating batch elements into rectangular sections that consider utmost of each batch.

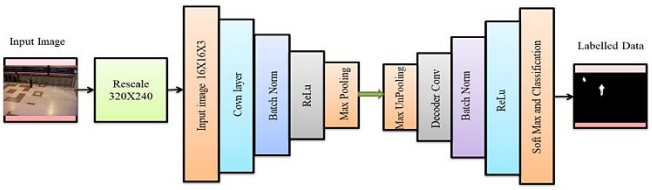


Figure 5. Network design of the constructed DNN

Labelled images are given to the network for training with intensity values of "255" for the foreground region and "0" for the background region. The stochastic gradient descent with momentum kind of solver is used with 50 epochs each of 100 iterations and utilizing a total of 31 layers and 2 encoder depth, the network is trained. The loss function of "cross-entropy" is employed at the soft max layer to categorize the pixels. The *table 1* shows the detailed layer information of the DNN which is designed for object detection.

Table 1. Layer information of the Proposed DNN

Layer No.	Layer Type	Kernel Size / Stride	Input Size	Output Size	Details
1	Input Layer	–	16×16×3	16×16×3	RGB image patch
2	Convolution 2D	3×3 / stride 1	16×16×3	16×16×32	Feature extraction
3	Batch Normalization	–	16×16×32	16×16×32	Normalize feature maps
4	ReLU Activation	–	16×16×32	16×16×32	Non-linearity
5	Max Pooling	2×2 / stride 2	16×16×32	8×8×32	Down sample
6–20	Convolution + BN + ReLU (x5)	3×3 / stride 1	8×8×32 → 8×8×64	8×8×64	Deeper feature maps
21	Up sampling	2×2	8×8×64	16×16×64	Start of decoder
22–28	Convolution + ReLU (x3)	3×3	16×16×64	16×16×64	Refinement
29	Fully Connected (FC)	–	16×16×64	256	Flatten and feed to classifier
30	Softmax Layer	–	256	2 (bg/fg)	Binary classification (pixel-wise)
31	Output Segmentation Map	–	16×16	16×16	Foreground mask

FOREX Publication
Open Access | Rapid and quality publishing
International Journal of Electrical and Electronics Research (IJEER)
Research Article | Volume 13, Issue 3 | Pages 393-401 | e-ISSN: 2347-470X

Unlike YOLO and SSD, which process images individually, our model leverages illumination-corrected frames and frame-wise continuity to maintain segmentation quality even under poor visibility. While not explicitly temporal like RNNs, the architecture's design accommodates temporal coherence by processing illumination-corrected video sequences in batches.

Table 2. Comparison with Standard Models

Model	Speed	Accuracy (F1)	Illumination Robustness	Temporal Awareness
RCNN [25]	Low	High	Poor	No
YOLOv3[27]	High	Moderate	Moderate	No
SSD [29]	Moderate	Moderate	Low	No
Proposed	Moderate	**High**	**High**	**Partial (frame-seq)**

Table 2 shows the comparison between the proposed method and existing models in terms of speed, accuracy, and robustness to illumination.

5. RESULT ANALYSIS

With the change detection benchmark dataset available at [32] the proposed network is assessed, this dataset includes six videos with various effects like object motion blur, shadow and camera jitter. However, in this study static baseline videos with 320X240 true color resolutions are used for analysis. The influence of low illumination and AIC algorithm is shown below in *figure 6*.

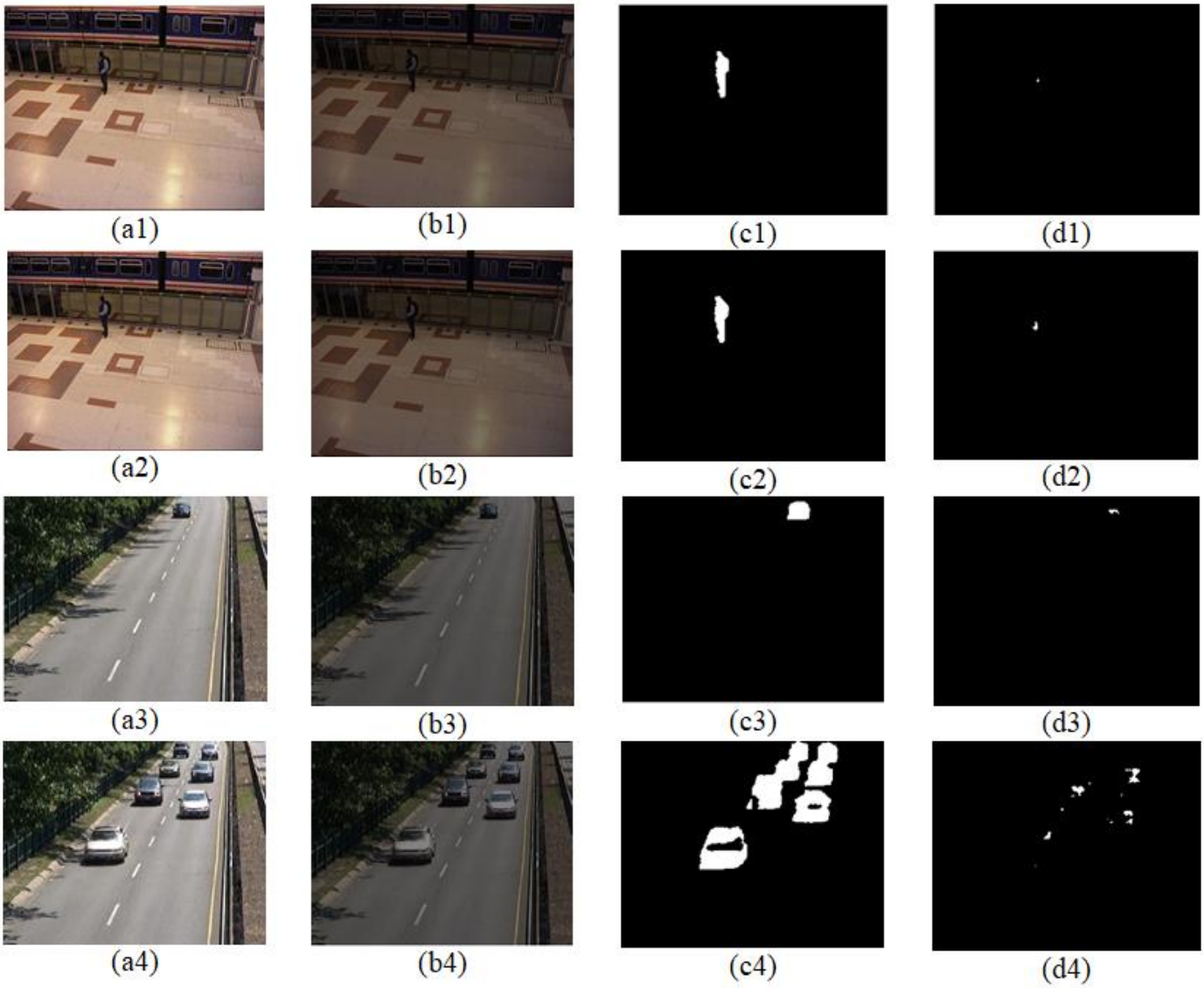


Figure 6. columns (a1) – (a4) represents original frames, (b1)- (b4) poor illuminated frames (c1) -(c4) object detection with AIC and DNN (d1) - (d4) segmented frames without AIC

FOREX Publication
Open Access | Rapid and quality publishing
International Journal of Electrical and Electronics Research (IJEER)
Research Article | Volume 13, Issue 3 | Pages 393-401 | e-ISSN: 2347-470X

The impact of AIC algorithm for object segmentation is depicted in *figure 6* where the original images when subjected for poor illumination may lead to improper segmentation which are shown in (b1-b4) of *figure 6* and the segmented output with AIC algorithm is shown in *figure 6* (c1-c4) and without correction algorithm may attain very poor results as illustrated in *figure 6* (d1-d4). The proposed network hereafter is termed as AICOD network.

5.1. Performance Assessment Metrics

In order to improve the outcomes, a segmentation algorithms performance must be evaluated in order to both assess how precise the segmentation approach is and to modify its parameters. As a result, it is possible to assess the performance measures and the methodology used to produce them. In this work, we evaluated performance metrics like Recall (Re), Precision (Pr), Percentage of wrong classification (PWC), False Negative rate (FNR), False positive rate (FPR), F-measure and Specificity (Sp),[32][33]. The collection contains the true images also includes the ground truth images.

The quality assessment metrics are calculated using procedure outline below. Let call automatic output 'A' and ground truth or manual segment output 'M'. The true positive (TP) is the pixels that are prevalent both the segmented output. The number of pixels that are existing in 'A' absent in 'M' is defined as false positive (FP). False-negative (FN) pixels are those that are present in 'M' absent in 'A' and the number of pixels that are missing from both segmented output is known as true negative (TN).

Table 3. Manual Segmentation and Automatic segmentation Contingency table

		Manual Segmentation/ Ground Truth	
		Negative	**Positive**
Automatic	**Negative**	True Negative	False Negative
	Positive	False Positive	True Positive

Table 4. Metrical calculation based on TP, FP, FN, TN

Metric	Formula
Percentage wrong classification	100*(FN+FP)/(TP+FP+FN+TN)
Recall	TP/(TP+FN)
F-measure	TP/(TP+FP)
Precision	TP/(TP+FP)
Specificity	TN/(TN+FP)
False Negative Rate	FN/(TP+FN)
False Positive Rate	FP/(FP+TN)

The proposed approach effectiveness is compared against VGG16 and ResNet50 networks for multiple frames are shown below.

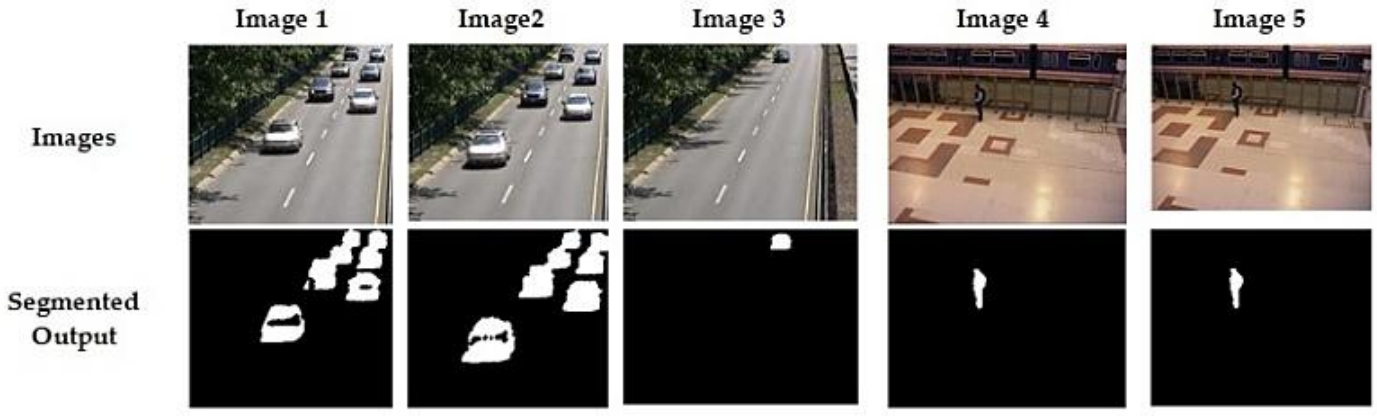


Figure 7. Shows the segmented output using the proposed AICOD Network

Table 5. Outcomes found with proposed AICOD network

Images	Recall	Specificity	FPR	FNR	PWC	Precision	F-measure
Image1	0.988	0.983	0.016	0.011	1.59	0.85	0.918
Image2	0.987	0.986	0.013	0.012	1.31	0.9	0.943
Image3	0.935	0.99	0.0004	0.065	0.08	0.93	0.93
Image4	0.979	0.99	0.00073	0.02	0.08	0.911	0.944
Image5	0.975	0.99	0.000748	0.02	0.09	0.913	0.943
Avg	0.972	0.987	0.006	0.02	0.63	0.901	0.935

FOREX Publication
Open Access | Rapid and quality publishing
International Journal of Electrical and Electronics Research (IJEER)
Research Article | Volume 13, Issue 3 | Pages 393-401 | e-ISSN: 2347-470X

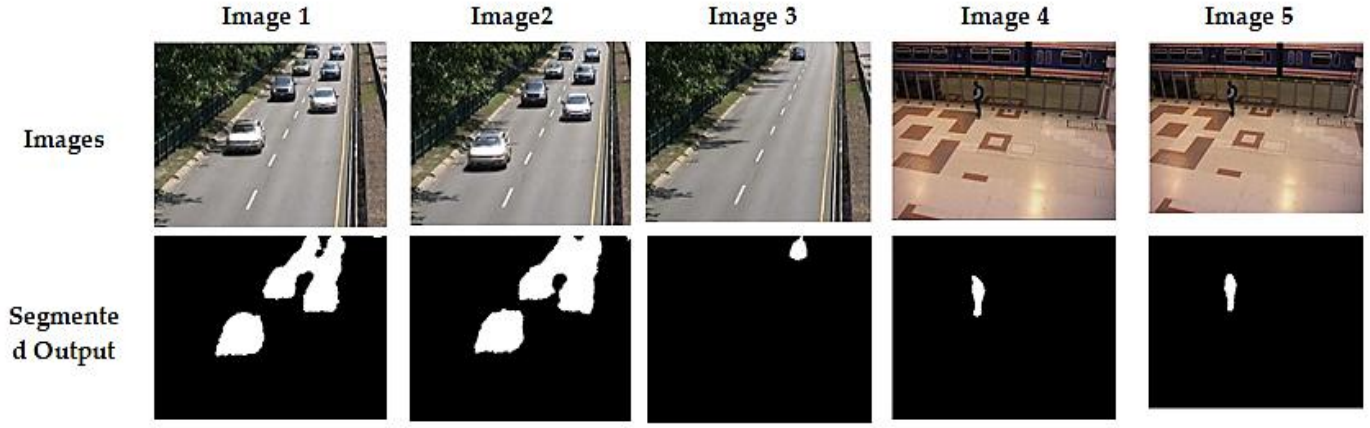


Figure 8. Shows the segmented output using pre-trained VGG16 [34] Network

Table 6. Results obtained with pre-trained VGG16 [34] network

Images	Recall	Specificity	FPR	FNR	PWC	Precision	F-measure
Image1	0.86	0.99	0.003	0.13	1.98	0.97	0.914
Image2	0.78	0.99	0.0014	0.23	3.2	0.98	0.86
Image3	0.87	0.99	0.0011	0.125	0.19	0.84	0.85
Image4	0.914	0.99	0.0011	0.08	0.17	0.88	0.89
Image5	0.916	0.99	0.000	0.08	0.14	0.9	0.91
Avg	0.868	0.99	0.001	0.12	1.13	0.914	0.885

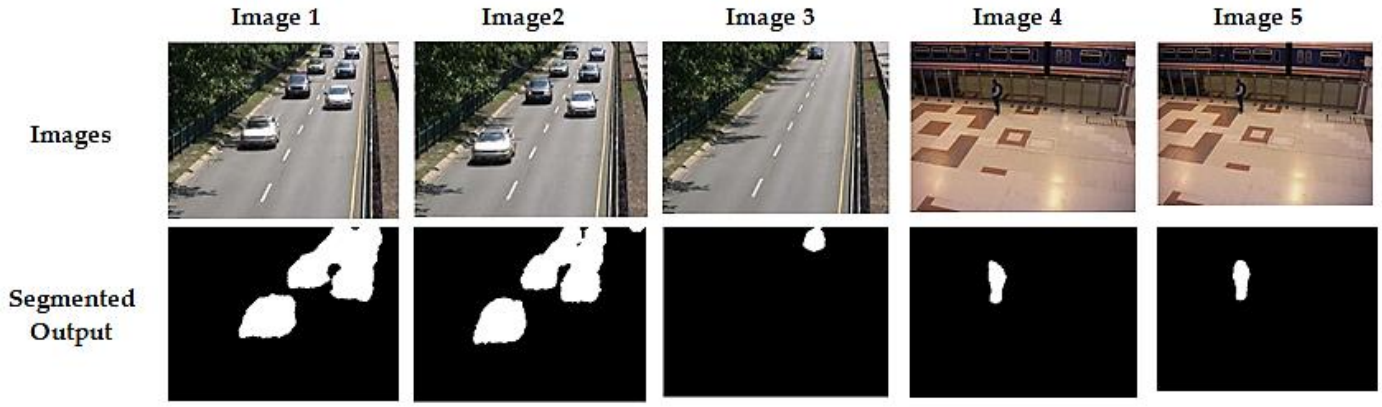


Figure 9. Shows the segmented output using pre-trained ResNet 50 [35] Network

Table 7. Results obtained with pre-trained ResNet 50 [35] network

Images	Recall	Specificity	FPR	FNR	PWC	Precision	F-measure
Image1	0.72	0.99	0.00026	0.3	4.61	0.99	0.81
Image2	0.706	0.99	0.00076	0.29	4.45	0.98	0.82
Image3	0.65	0.99	0.000105	0.34	0.38	0.98	0.78
Image4	0.63	1	0	0.36	0.5	1	0.77
Image5	0.617	1	0	0.38	0.51	1	0.76
Avg	0.66	0.994	0.0025	0.33	2.09	0.99	0.78

FOREX Publication
Open Access | Rapid and quality publishing
International Journal of Electrical and Electronics Research (IJEER)
Research Article | Volume 13, Issue 3 | Pages 393-401 | e-ISSN: 2347-470X

The *figures 7, 8* and *9* shows the pictorial output of the proposed AICOD, VGG16 and ResNet50 Respectively. The performance of the Proposed AICOD, VGG16 and ResNet50 are presented in tables 5,6 and 7 which indicate the average values achieved for state-of-the-art approaches and presented methods in terms of various matrices. In this paper, only few frames' outputs were presented which are compared against each other. *Figure 10* and *11* depicts the proposed methods performance evaluation and its superiority over VGG16 and Resnet50. The proposed technique may be seen to produce higher f-measure, specificity and recall values while upholding very little fault rate in terms of PWC and FNR. The relative study shown in *figures 10*, *11* demonstrates that the method can retain and improve values of 0.1~0.2 for F-measure, 0.2~0.3 for Recall, which is a significant achievement.

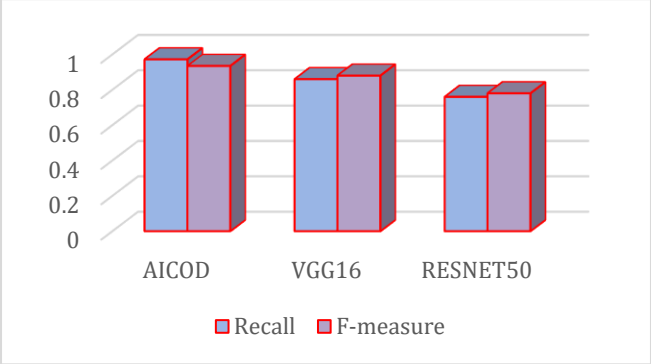


Figure 10. Performance analysis of proposed AICOD and state-of-the-art methods

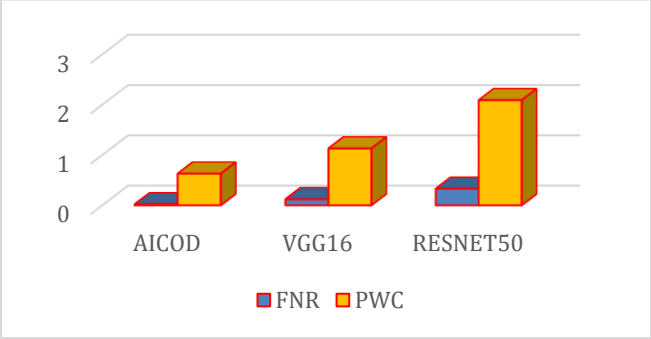


Figure 11. Performance analysis of proposed AICOD and state-of-the-art methods in terms of FNR and PWC

From the experimental result analysis, it can be stated that the proposed AICOD technique could able to precisely identify the existence of objects even in the low illuminated input video frames which is one of the contributions of the work. Secondly, the accuracy of detection is superior to the state of art methods however; this method has a slightly longer processing time which must be leveraged with code optimization, optimal resolution selection.

6. CONCLUSIONS

This paper presents an adaptive illumination correction with object detection mechanism for surveillance videos. The method mainly aims to adaptively correct the input low illuminated frames by verifying the quality of the input frame blindly. The corrective process is an iterative mechanism that aims to adjust the contrast values until the quality of the frame is improved. By doing so, even the objects in the low illuminated frames can be effectively detected. The method is integrated with customized DAG deep neural network that mainly focuses to identify and segment the objects in the frames. The tryouts were carried out with both single object and more than one object in the frames. Based on the observations, it is possible to conclude that even when multi-objects are identified, segmentation spots overlay with near objects. This problem has to be rectified with effective training options and mechanisms. This may not, however represent for severe level in object identification and tracking. Individual objects, on the other hand, are flawlessly detected, demonstrating an unbelievable development in performance examination when compared to state-of-the-art networks.

REFERENCES

[1] G. Han, M. Guizani, J. Lloret, S. Chan, L.Wan, and W. Guibene, "Emerging trends, issues, and challenges in big data and its implementation toward future smart cities," IEEE Commun. Mag., vol. 55, no. 12, pp. 16-17, Dec. 2017.

[2] X. Chang, Y.-L. Yu, Y. Yang, and E. P. Xing, "Semantic pooling for complex event analysis in untrimmed videos," IEEE Trans. Pattern Anal. Mach. Intell., vol. 39, no. 8, pp. 1617-1632, Aug. 2017.

[3] K. Muhammad, J. Ahmad, and S. W. Baik, ``Early fire detection using convolutional neural networks during surveillance for effective disaster management," Neuro computing, vol. 288, pp. 30-42, May 2017.

[4] X. Chang, Z. Ma, M. Lin, Y. Yang, and A. G. Hauptmann, "Feature interaction augmented sparse learning for fast Kinect motion detection," IEEE Trans. Image Process, vol. 26, no. 8, pp. 3911-3920, Aug. 2017.

[5] X. Chang, Z. Ma, Y. Yang, Z. Zeng, and A. G. Hauptmann, "Bi-level semantic representation analysis for multimedia event detection," IEEE Trans. Cybern., vol. 47, no. 5, pp. 1180-1197, May 2017.

[6] X. Chang and Y. Yang, "Semi supervised feature analysis by mining correlations among multiple tasks," IEEE Trans. Neural Netw. Learn. Syst., vol. 28, no. 10, pp. 2294-2305, Oct. 2017.

[7] A. Yilmaz, O. Javed, and M. Shah, ``Object tracking: A survey," ACM Comput. Surv., vol. 38, no. 4, Art. no. 13, 2006.

[8] Z. Pan, S. Liu, and W. Fu, ``A review of visual moving target tracking," Multimedia Tools Appl., vol. 76, no. 16, pp. 1-30, 2017.

[9] Z. Chen, Z. Hong, and D. Tao. (2015). "An experimental survey on correlation filter-based tracking." [Online]. Available: https://arxiv.org/abs/1509.05520.

[10] Z. Kalal, K. Mikolajczyk, and J. Matas, "Tracking-learning-detection," IEEE Trans. Pattern Anal. Mach. Intell., vol. 34, no. 7, pp. 1409-1422, Jul. 2012.

[11] N. Dalal and B. Triggs, ``Histograms of oriented gradients for human detection," in Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit. (CVPR), vol. 1, pp. 886-893. Jun. 2005.

[12] R.Mantiuk, K.Myszkowski, H.P.Seidel, "A perceptual framework for contrast processing of high dynamic range images", ACMTrans.Appl. Percept.3(3), 2006, 286–308.

[13] Vertan, C., Oprea, A., Florea, C. and Florea, L. "A pseudo-logarithmic framework for edge detection", in J.B. Talon, S. Bourennane, W. Philips, D. Popescu and P. Scheunders (Eds.), Advances in Computer Vision, Lecture Notes in Computer Science, Vol. 5259, Springer-Verlag, Juan-les-Pins, pp. 637–644,2008.

International Journal of Electrical and Electronics Research (IJEER)
Research Article | Volume 13, Issue 3 | Pages 393-401 | e-ISSN: 2347-470X

[14] Patras incu, V. and Voicu, I. "An algebraical model for gray level images", Proceedings of the Exhibition on Optimization of Electrical and Electronic Equipment, OPTIM, Brasov, Romania, pp. 809–812,2000.

[15] Panetta, K., Zhou, Y., Agaian, S. and Wharton, E. "Parameterized logarithmic framework for image enhancement", IEEE Transactions on Systems, Man, and Cybernetics, B: Cybernetics 41(2): 460–472,2011.

[16] M. S. Farid, M. Lucenteforte and M. Grangetto, "Blind depth quality assessment using histogram shape analysis," 2015 3DTV-Conference: The True Vision - Capture, Transmission and Display of 3D Video (3DTV-CON), Lisbon, 2015, pp. 1-5.

[17] Z. Li and J. Zheng, "Single Image De-Hazing Using Globally Guided Image Filtering," in IEEE Transactions on Image Processing, vol. 27, no. 1, pp. 442-450, Jan. 2018.

[18] Krizhevsky, A., Sutskever, I., Hinton, G.E.: Imagenet classification with deep convolutional neural networks. In: NIPS. 2012.

[19] Szegedy, C., Liu, W., Jia, Y., Sermanet, P., Reed, S., Anguelov, D., Erhan, D., Vanhoucke, V., Rabinovich, A.: Going deeper with convolutions. In: Computer Vision and Pattern Recognition (CVPR). 2015.

[20] Simonyan, K., Zisserman, A.: Very deep convolutional networks for large-scale image recognition. In: ICLR. 2015.

[21] Bertasius, G., Shi, J., Torresani, L.: Deepedge: A multi-scale bifurcated deep network for top-down contour detection. In: The IEEE Conference on Computer Vision and Pattern Recognition (CVPR) June 2015.

[22] He, K., Zhang, X., Ren, S., Sun, J.: Deep residual learning for image recognition. 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR),2016, 770–778.

[23] Babaee, M., Tung Dinh, D., Rigoll, G.: 'A deep convolutional neural network for video sequence background subtraction', Pattern Recognit., 2018, 76, pp. 635–649.

[24] Lim, L.A., Keles, H.Y.: 'Foreground segmentation using a triplet convolutional neural network for multiscale feature encoding', Pattern Recognit. Lett., 112, pp. 256–262,2018.

[25] S. Ren, K. He, R. Girshick, and J. Sun, ``Faster R-CNN: Towards real-time object detection with region proposal networks," IEEE Trans. Pattern Anal. Mach. Intell., vol. 39, no. 6, pp. 1137-1149, Jun. 2017.

[26] R. Girshick, "Fast R-CNN," in Proc. IEEE Int. Conf. Comput. Vis. (ICCV), Dec. pp. 1440-1448, 2015.

[27] Redmon, S. Divvala, R. Girshick, and A. Farhadi, "You only look once: Unified, real-time object detection", in Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR), pp. 779-788, Jun. 2016.

[28] G. Li, Z. Song, and Q. Fu, ``A new method of image detection for small datasets under the framework of YOLO network," in Proc. IEEE 3rd, Adv. Inf. Technol., Electron. Automat. Control Conf. (IAEAC), pp. 1031-1035, Oct. 2018.

[29] K. Zhao, X. Ren, Z. Kong, and M. Liu, "Object detection on remote sensing images using deep learning: An improved single shot multi-box detector method", J. Electron. Imag., vol. 28, no. 03, p. 1, Jun. 2019.

[30] W. Liu, D. Anguelov, D. Erhan, C. Szegedy, S. Reed, C.-Y. Fu, and A. C. Berg, ``SSD: Single shot multibox detector," in Proc. Eur. Conf. Comput. Vis., 2016, pp. 21-37.

[31] Optiview. (2021, August 4). CCTV Camera Resolution | CCTV Resolution Chart for Cameras. Optiview - Professional Surveillance & Security Solutions Provider. https://optiviewusa.com/cctv-video-resolutions/

[32] N. Goyette, P.-M. Jodoin, F. Porikli, J. Konrad, and P. Ishwar, changedetection.net: A new change detection benchmark dataset,in Proc. IEEE Workshop on Change Detection (CDW-2012) at CVPR-2012, Providence, RI, 16-21 Jun., 2012.

[33] Gala, Nikhil & Desai, Kamalakar., "Abnormal Region Extraction from MR Brain Images using Hybrid Approach", International Journal of Advanced Computer Science and Applications. 9. 10.14569/IJACSA.2018.091230, 2018.

[34] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," 1409.1556, 2014.

[35] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 770–778, June 2016.