Design of Integrated Artificial Intelligence Techniques for Video Surveillance on IoT Enabled Wireless Multimedia Sensor Networks

Romany F. Mansour[1]*, Carlos Soto[2], Roosvel Soto-Díaz[3], José Escorcia Gutierrez[4,5], Deepak Gupta[6], Ashish Khanna[6]

[1] Department of Mathematics, Faculty of Science, New Valley University, El-Kharga, (Egypt)
[2] Mechanical Engineering Program, Universidad Autónoma del Caribe, 08001 Barranquilla (Colombia)
[3] Mechatronic Engineering Program, Universidad Simón Bolívar, 08001 Barranquilla (Colombia)
[4] Electronics and Telecommunications Engineering Program, Universidad Autónoma del Caribe, 08001 Barranquilla (Colombia)
[5] Biomedical Engineering Program, Corporación Universitaria Reformada, 08001 Barranquilla (Colombia)
[6] Department of Computer Science & Engineering, Maharaja Agrasen Institute of Technology, Delhi (India)

Received 1 August 2021 | Accepted 20 June 2022 | Published 8 August 2022

Abstract

The recent advancements in the Internet of Things (IoT) and Wireless Multimedia Sensor Networks (WMSN) made high-speed multimedia streaming, data processing, and essential analytics processes with minimal delay. Multimedia sensors used in WMSN-based surveillance applications are beneficial helpful in attaining accurate and elaborate details. However, it has become essential to design an effective and lightweight solution for data traffic management in WMSN owing to the massive quantities of data, generated by multimedia sensors. The development of Artificial Intelligence (AI) and Machine Learning (ML) techniques can be leveraged to investigate, collect, store, and process multimedia streaming data for decision-making in real-time scenarios. In this aspect, the current study develops an Integrated AI technique for Video Surveillance in IoT-enabled WMSN, called IAIVS-WMSN. The proposed IAIVS-WMSN technique aims to design a practical scheme for object detection and data transmission in WMSN. The proposed IAIVS-WMSN approach encompasses three stages: object detection, image compression, and clustering. The Mask Regional Convolutional Neural Network (Mask RCNN) technique is primarily utilized for object detection in the target region. Besides, Neighbourhood Correlation Sequence-based Image Compression (NCSIC) technique is applied to reduce data transmission. Finally, Artificial Flora Algorithm (AFA)-based clustering technique is designed for the election of Cluster Heads (CHs) and construction clusters. The design of object detection with compression and clustering techniques for WMSN shows the novelty of the work. These three processes' designs enable one to accomplish effective data transmission in IoT-enabled WMSN. The researchers conducted multiple simulations to highlight the supreme performance of the IAIVS-WMSN approach. The simulation outcomes inferred the enhanced performance of the IAIVS-WMSN algorithm to the existing approaches.

Keywords

Artificial Intelligence, Intelligent Surveillance, Internet Of Things, Object Detection, WMSN.

DOI: 10.9781/ijimai.2022.08.005

I. Introduction

THE advancements made in miniaturization technology allow the incorporation of heterogeneous sensing devices on single sensing platforms. A single prolific advantage of the miniaturization process is the accessibility of a Complementary Metal Oxide Semiconductor (CMOS) camera. When the latter is incorporated into conventional Wireless Sensor Networks (WSN), it transforms the WSN into Wireless Multimedia Sensor Networks (WMSN). This network evolution allow the execution of multi-dimension signal processing methods on sensing platforms [1]. Further, it also provides advanced services compared to conventional WSN. WMSN includes various relay nodes and camera motes while the latter needs to be placed in the target region to monitor the existence of intruders and transfer the data to the monitoring location. Currently, WMSN is utilized in several surveillance applications such as monitoring elders, identifying anomalies in secure regions, traffic monitoring, among others. In the event of a surveillance application, the camera mote captures the video and transfers it to the monitoring place through relay nodes [2]. If no change is observed in the scene, it is not necessary to transfer the whole video in a resource-limited environment such as WMSN. Fig. 1 shows the common framework of cluster-based WMSN.

* Corresponding author.

E-mail address: romanyf@sci.nvu.edu.eg

WMSN enables both individuals and organizations to stream video/audio data and still transmit the images together with scalar sensor data. Therefore, it has become possible to transmit the group of data from distinct modalities under a complex environment which was not the case in WMSN in previous years [3]. The significant differences between WMSN and conventional WSNs are the size and nature of the data collected, processed, and transmitted in the former [4]. In general, scalar data is the major kind of data in WSN that can be transmitted and processed in a simple manner. Nevertheless, multimedia data like audio, video and image have a complex structure that makes data processing, a highly complex task. Further, multimedia data is also prominent in size, which incurs high wireless transmission costs and is decided based on network bandwidth and consumed energy [5].

Furthermore, they can be utilized in resource-constrained, challenging, and unattended regions. However, some specific constraints in terms of data processing must be overcome when developing novel methods in WMSN. These constraints include limited processing, limited memory, limited battery power, and narrow bandwidth [6]. Hence, the cognition of amendment in conventional multi-dimension signal processing is an early-stage development prior to real-world execution in WMSN. In order to achieve this coherency, the current study emphasizes adjusting the conventional data processing model from WMSN. Mainly, the aim is to extend the lifespan of nodes by mitigating costs incurred upon in-node processing. As a result, tracking processes and visual object identification are performed at the sink node, where the realization of object geometry allows the remote node to decide the additional operation in its sensing region.

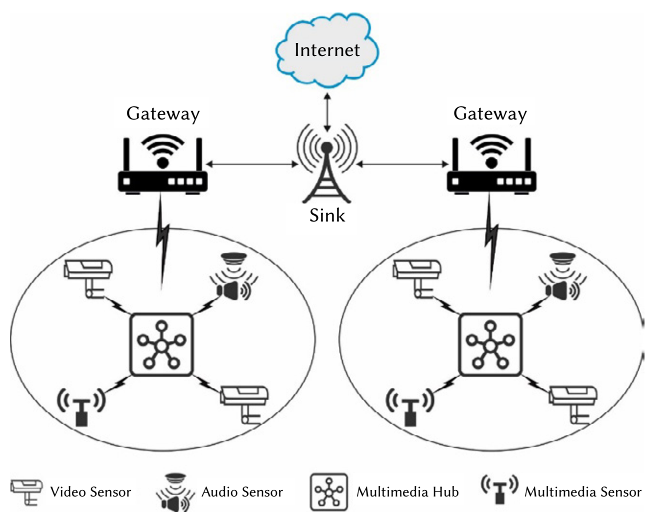

Fig. 1. The general structure of cluster-based WMSN.

In other words, each kind of WSN node has constrained resources like battery power, CPU, RAM, secondary storage, and communication bandwidth. Such limitation paves the way for developing practical algorithms and techniques to be incorporated into WMSN. To minimize the data size of the transferred multimedia, various researchers have proposed image compression algorithms or modified the previous algorithms for WMSN [7]. Most of the studies were aimed at reducing the quantity of data to be transferred, when utilizing energy-effective and simple image compression algorithms. The results show that the objective could be obtained to a specific level by compression algorithms [8]. However, it is still impossible to claim that resource limitation problems can be overcome in WMSN. Other solutions were also presented to solve the aforementioned problems based on extraction and detection of the objects present in the image [9]. This method aims to process multimedia data and extract more useful data, i.e., small informative data than massive raw multimedia data, at the node level. Since the data that needs to be transferred is smaller in size, there is a parallel reduction observed in energy consumption for data transmission. Finally, processing effort at the sensors level might get altered based on various aspects.

To the best of the researcher's knowledge, no works conducted earlier have integrated all three major processes of WSN such as object detection, compression, and clustering. The current study develops an Integrated AI technique for Video Surveillance in IoT-enabled WMSN, called IAIVS-WMSN. The proposed IAIVS-WMSN approach contains three significant processes: object detection, image compression, and clustering. A Mask Regional Convolutional Neural Network (Mask RCNN) technique is utilized for object detection in target region. Moreover, the Neighborhood Correlation Sequence-based Image Compression (NCSIC) technique is applied to minimize the quantity of data transmission. At last, the Artificial Flora Algorithm (AFA)-based clustering technique is used in Cluster Heads (CHs) and constructing clusters. A comprehensive experimental analysis was executed to showcase the superiority of the proposed IAIVS-WMSN algorithm.

II. Related Works

In Rehman et al. [10], new object detection and image transmission method were presented for WMSN. An image segment is broadcasted instead of whole images. In this sense, minimal image content transmission and in-node energy conservation to the sink node are assured. The efficiency of the presented system was determined based on the reconstructed image and in-node energy consumption level. In Guo [11], an effective compressive sensing method-based customized memory gradient pursuit approach using earlier terminations in WSN was proposed. This method was able to strike compelling tradeoffs amongst energy dissipation for wireless communication, specific kinds of minimum storage, and bandwidth. Later, the presented method adapted an unscented particle filter to predict the target position.

Koyuncu et al. [12] explored the effect of combining audio-visual multimedia and scalar data gathered by sensors from WMSN in order to achieve energy-effective and precise object classification and detection. To perform this method, the researchers presented a wireless multimedia sensor with video and audio processing and capturing capabilities along with ordinary or traditional scalar sensor nodes. Multimedia sensor nodes are maintained in sleep mode to save energy until they get activated by the scalar sensor nodes which are active all the time. Sukumaran et al. [13] constructed a CS-BS architecture using a new threshold approach for anomaly detection through less measurement in a protected indoor environment. In CS-BS architecture, CS is implemented on variance frame i.e., sparse, which reduces the usage of bandwidth, energy, and memory. In this architecture, a foreground thresholding is presented based on a measurement matrix to extract the motion objects from a scene.

Wang et al. [14] proposed EDACR for WMSN by taking limitations of energy consumption and QoS into account. In this study, the researchers primarily designed an RL-based method to ensure energy-balanced routing and QoS based on the knowledge of reliability and delay. The experimental results infer a decline in energy consumption while in parallel, QoS is guaranteed in terms of conventional and distributed adoptive cooperative protocols. Akter et al. [15] introduced a comprehensive tracking and localization method in WSN. Considering the limitations of static cluster, an energy-effective incremental clustering approach was proposed in this study, after which the Gaussian adoptive resonance concept was presented at the border area. The presented work was permitted to learn, create, update and retain the clusters gradually through online learning to adopt continuous motion patterns. At last, the trilateration method was employed for the accurate location of a dynamic object through a sensor network.

Shao [16] proposed a dynamic clustering target tracking method for moving trends. In this method, a dynamic cluster is formed in networks whereas the CH dynamically schedules the nodes to track the target collaboratively. The tracking approach primarily consists of two phases. In the first phase, the CH establishes a 'noighbor node set' within its transmission range, and the noighbor node is selected in the 'noighbor node set' based on the distances between the node and the targets to create an 'intracluster member set' to perform on the target. In Sathyaprakash and Prakasam [17], an RBMLCA approach was presented using a randomized method. It involves optimising QoS parameters, cluster head formation, and clustering of nodes. The adapted clustering method helped in data transmission on multimedia sensor networks, whereas the RBMLCA method yielded an optimal quality of evolutional parameters.

Heng et al. [18] proposed a holistic WMSN framework for image transmission that performs well on different images. This method was proposed based on standard deviation and leveraging its saliency features. Then, FLS was utilized to determine the suitable features while the samplings were assigned, and all the respective blocks were resized with CS. The integrated FLS & BCS approaches were executed by SPL recreation to determine the convergence speed. Alqaralleh et al. [19] presented a novel Reliable MultiObject Tracking Model with DL and Energy-Effective WMSN. At first, the FL method was applied to determine the CHs and achieve energy efficacy. Then, a new tracking approach RNN-T was proposed using RNN with tumbling effects. The presented RNN-T model was implemented in all the sensors and CH executed the tracking approach to track all the animals. Lastly, the tracking outcomes were transferred to cloud server for research purposes.

III. The Proposed Model

The proposed IAIVS-WMSN approach comprises three stages Mask RCNN-based object detection, AFA-based clustering, and NCSIC-based image compression. In the first stage, the objects in the target regions are detected with the help of the Mask RCNN technique. Next, the AFA technique is implemented to determine the CHs and optimally construct the images. Furthermore, the NCSIC technique is applied to compress the images at CMs and CHs prior to data transmission to the BS. A detailed discussion of these processes is offered in the upcoming subsections.

A. System Model

This section deals with the network's topology that is utilized to transmit the image. The analysis of in-network energy utilization inspires it during image frame broadcasting. In image broadcasting, the energy utilization occurs continuously on the superior side like in-node processing. Then, the vital goal of presenting the topology has to define the entire energy utilization when broadcasting the group of image frames. Assume that energy utilization occurs arbitrarily in the WMSN node. All the WMSN nodes are equipped with the restricted resource. Consequently, there is no channel impairment in the network. For a reliable transmission of the image within networks and to avoid collision, due to simultaneous broadcast of image data by more than two nodes, all the WMSN nodes are separated into two classes such as Cluster head (CH) and Monitoring Node (MN). The network topology is demonstrated in Fig. 2.

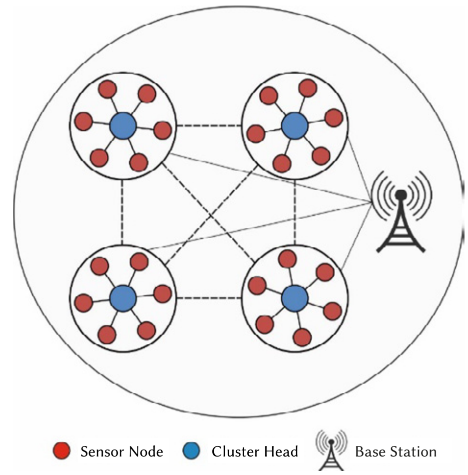

Fig. 2. Network Model.

B. Design of Object Detection Technique

In this primary step, the Mask-RCNN technique is applied to detect the presence of objects in the target region. In general, Mask-RCNN is theoretically a flexible and simple architecture, for instance, segmentation, object recognition, and detection. It can effectively identify the objects in an image, when producing a high-quality segmentation mask during all instances. The FPN for object identification, the primary block framework of Mask-RCNN, is accountable for removing the features. The RPN and Mask RCNN share full image convolution features with the detection network. Thus, it allows near cost-free region suggestions. Afterward, fast RCNN is prolonged to make Mask RCNN by including a branch for prediction of object mask like the present branch for bound box detection. RPN is utilized for masking the RCNN before 'elective searching' so that the RPN can share the convolution feature of the whole map with the detection network. It can predict the object score as well as boundary position at all the locations while it is also an FCN.

Fast RCNN exploits RPN as a regional generation network to generate the candidate region. FPS depends on the Fast RCNN approach as much as five times while its MAP, verified on VOC 2012, improved by 70.4%. Further, to enhance the detection accuracy of the target, Mask RCNN exploits a bilinear interpolation approach in which ROI align is used rather than ROI pool, based on Fast RCNN [20]. ROI align approach is utilized to determine the precise values of the input features, according to bilinear interpolations at four regular sampled places from every ROI bin and the result is aggregated. This technique enhances the precision of Mask RCNN to 10%. In order to ensure Mask RCNN executes the mask function and attains highly accurate instance segmentation in pixel-to-pixel alignment, the mask branch is included in RCNN. Mask RCNN has the ability to execute three processes such as target detection, segmentation, and recognition. Its detection speed can still attain 5 FPS. The diagram of Mask RCNN is displayed in Fig. 3.

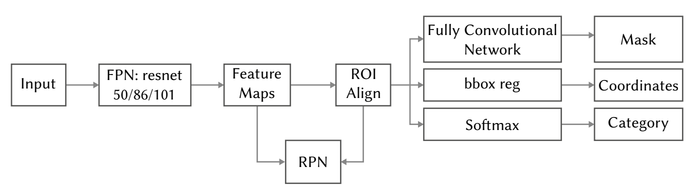

Fig. 3. Structure of Mask RCNN.

Next, the candidate regions are integrated with the feature map so that the scheme can attain the target's classification, detection, and mask. Depending on Mask RCNN, a technique named SF-FPN was proposed earlier to improve the speed and detection accuracy using Resnet86. This work enhanced RPN parameter settings, dataset, and FPN structure. The enhanced technique, projected in this work, cloud perform segmentation, recognition, and detection of the targets simultaneously.

C. Design of Clustering Technique

AFA is designed based on the migration and reproduction of the flora. The procedure of identifying an optimum survival place for flora to flourish, is employed to define an optimum solution for the problem. Offspring Plant (OSP), Original Plant (OP), Propagation Distance (PD), and plant location are four essential components of MOAF. Spreading, selection, and evolution behavior are three major behavioral forms. Primarily, this process arbitrarily makes the OP. Later, it spreads seeds to the position within an arbitrary spreading scope. The spreading scope is determined using PD. Then, the fitness of a seed in a specific location is evaluated based on the objective function whereas 'fitness' denotes the quality of the solution. Eventually, roulette is employed to decide the survival seeds. These survival seeds develop a novel OP. The iterations are repeated till the termination condition is arrived at. This process includes external documents to save the optimal solution.

In this study, each decision variable of the test function contains an upper limit $\vec{X}^{max} = [X_1^{\max}, X_2^{\max}, \dots, X_D^{\max}]^T$ and a lower limit $\vec{X}^{min} = [X^{\min}, X^{\min}, \dots, X_D^{\min}]^T$. Primarily, this process makes N OP to be dependent upon the decision variable's upper and lower limits, arbitrarily. It uses i row and j column matrix P_{ij} to represent the location of OP, in which $i = 1, 2, \dots, D$ indicates the dimension and $j = 1, 2, \dots, N$ signifies the amount of OP [21]: The population is give in (1):

$$P_{ij} = rand(0,1) \cdot \left(X_i^{\max} - X_i^{min}\right) + X_i^{min} \tag{1}$$

Whereas $rand(0,1)$ refers to uniform distribution amount in the range of 0 and1.

OP spreads their offspring within a specific scope using radius i.e., PD while novel PD stimulates the PD of the grandparent and parent plants as given in (2).

$$d_j = d_{1j} \cdot rand(0,1) \cdot c_1 + d_{2j} \cdot rand(0,1) \cdot c_2 \tag{2}$$

Here, c_1 & c_2 denote the learning coefficients, d_{lj} and d_{2j} represent the PD of parent and grandparent plants correspondingly, and rand (0, 1) stands for uniform distribution amount in the range of 0 and1. Parent PD becomes the novel grandparent PD, as provided in (3).

$$d'_{1j} = d_{2j} \tag{3}$$

The regular AF optimization approach uses the standard deviation between the locations of OP and OSP as novel parent PD, as defined in (4).

$$d'_{2j} = \sqrt{\sum_{i=1}^{N} (P_{ij} - P'_{ij})^2 / N} \tag{4}$$

To maintain the data of optimal solution arrived at till now, MOAF optimized approach uses the plant in an external document. Equation (5) shows the novel parent PD, i.e. the difference between the location of the plant in external document P_{id}^* and OSP P'_{id}:

$$d'_{2j} = P_{ij}^* - P'_{ij} \tag{5}$$

The spreading behavior creates the OSP based on OP position and novel PD as given in (6).

$$P'_{i,j\cdot b} = G_{i,j\cdot b} + P_{i,j} \tag{6}$$

Here, $b = 1, 2, \dots, B$, B refers to the number of OSP where the OP could transmit, $P'_{i,j\cdot b}$ represents the position of OSP, $P_{i,j}$ represents the location of OP and $G_{i,j\cdot b}$ denotes the arbitrary amount utilizing Gaussian distribution with mean 0 and variance j. A novel OP is generated based on Eq. (2), when there are no OSP endures.

In the standard AF approach, the survival likelihood is defined based on the survival of OSP. The survival likelihood is calculated through (7):

$$p = \left|\sqrt{F\left(P'_{i,j\cdot b}\right)/F_{max}}\right| \cdot Q_x^{(j\cdot b-1)}. \tag{7}$$

Here, Q_X denotes the electing likelihood range between 0 and1. $F_{\max}$ indicates the fitness of OSP using maximum fitness. $F(P'_{i,j\cdot b})$ represents the fitness of the $(j \cdot b)th$ solution. The computation equation of fitness is the function of the main problem. In the MOAF approach, Pareto dominance relationships are employed. The survival likelihood is given in (8).

$$p = 0.9 \cdot \frac{domi(j \cdot b)}{B} + 0.1 \tag{8}$$

Here, domi $(j \cdot b)$signifies the number of solutions dominated using the solution, $(j \cdot b)$. B indicates the number of OSP that OP can transmit. The objective function allows the effectual election of CHs, including energy, distance, delay, link lifetime, inter-cluster and intra-cluster distances as shown in (9).

$$\begin{aligned} F = {} & \omega_1 \times E + \omega_2 \times (1 - D) + \omega_3 \times \left(1 - \partial^{\text{inter}}\right) \\ & + \omega_4 \times \partial^{\text{intra}} + \omega_5 \times (1 - d) + \omega_6 \times L \end{aligned} \tag{9}$$

where ω_1, ω_2, ω_3, ω_4, ω_5, and ω_6 denote the weights. The energy of the nodes is denoted as E, D denotes the delay in transmission, ∂^{inter} represents the intercluster distance, ∂^{intra} specifies the intracluster distance, d indicates the distance between two nodes, and l corresponds to link time.

D. Design of Image Compression Technique

The presented technique is a reliable image compression method mainly established to save energy-constrained sensor nodes in WMSN. It functions on two levels: bit decrease utilizing the NCSIC technique and encoding by LZMA. NCSIC technique creates the optimum codeword to all the individual pixel values based on the 'bit traversal method utilizing 0's and 1's' [22]. With the help of 0's based and 1's based traversal, the NCSIC technique creates two codewords for all individual pixel values. The codeword with minimal bits is elected as the optimum codeword. Assuming that G is an input image with pixel $\phi_{m,n}$ and is demonstrated as a 2D array as provided in (10).

$$G = \begin{bmatrix} \phi_{0,0} & \phi_{0,1} & \cdots & \phi_{0,n-1} \\ \phi_{1,0} & \phi_{1,1} & \cdots & \phi_{1,n-1} \\ \vdots & \vdots & \ddots & \vdots \\ \phi_{m-1,0} & \phi_{m-1,1} & \cdots & \phi_{m-1,n-1} \end{bmatrix} \tag{10}$$

Here, m and n refer to the height and width of the input image, G. A value of $m \times n = N$ gives the full resolution of the image G, and $\phi_{m,n}$ implies the place of pixels in the m^{th} row and n^{th} column of grayscale image, G. The entire count of bits required to store the optimum codeword Opt_{size} of the input image is estimated in (11).

$$\text{Opt}_{\text{size}} = \sum_{i=1}^{N} \text{NCSIC}_{\text{opt}}(i) + \text{control bits} \tag{11}$$

Here, NCSIC_{opt} denotes the number of current bits in codewords. Further, the NCSIC technique requires eight control bits for an optimum amount of bits in the compressed data. Especially, the quantuty of bits required on average to store a separate pixel in an image is defined in (12).

$$\text{NCSIC}_{\text{ch\_av}} = \frac{\text{Opt}_{\text{size}}}{N}, \quad 0 \le \text{NCSIC}_{\text{ch\_av}} \le 5 \tag{12}$$

As the RGB values of grayscale images are established to be similar, it is sufficient to encode some RGB values to all the pixels of the individual. Therefore, in principle, the NCSIC technique needs less than 0 bits and a maximum of five bits for representing some possible pixel values. In order to further improve the compression model, an optimum codeword is encoded by LZMA, and the compressed file is created.

Primarily, the NCSIC technique reads the pixel value from the original image and converts the pixel values to an equivalent binary method. If the initial bit is 0, 0's based-traversal occurs and saves the control bit like 00 or 01. After determining the value of 0's, the equivalent places (p) are saved as the portion of optimum codeword ($00 - p_1$). These processes continue still each 0's in the order are recognized and the places are upgraded ($00 - p_1, p_2, \ldots$) respectively. However, 1's-based traversal appears to occur in 1's, while 0's based traversal looks towards the incidence of 0's in binary form. To all the pixel values, two codewords are created based on 0's and 1's based traversals. Next, the codeword with fewer bits is elected as a better codeword. Then, LZMA is implemented to encode an optimum codeword created by the NCSIC technique. The presented technique follows a symmetrical manner in which the compression and decompression are inverse functions of one another.

IV. Performance Validation

The current section provides the performance analysis results of the proposed technique under different aspects. The proposed model was simulated using the MATLAB tool. The results were inspected in three distinct ways. At first, the object detection performance of the Mask RCNN model was determined by identifying two objects namely 'Human' and 'Vehicle'. Table 1 shows the object detection outcomes of the Mask-RCNN approach and other existing techniques.

Fig. 4 depicts the object detection outcomes accomplished by Mask.

RCNN technique and other existing approaches on the applied object, 'Human'. The figure depicts that the Support Vector Machine with Speeded Up Robust Feature (SVM-SURF) method attained poor object detection outcomes with a $prec_n$ of 49%, $reca_l$ of 46%, and an $F_{measure}$ of 47%. Likewise, the k-nearest neighbors (k-NN) technique gained a slightly increased performance with a $prec_n$ of 51%, $reca_l$ of 35%, and an $F_{measure}$ of 42%. In the meantime, the Support Vector Machine with Structural Features (SVM-SF) technique achieved a moderate outcome with a $prec_n$ of 55%, $reca_l$ of 32%, and a $F_{measure}$ of 40%. Along with that, k-NN+SVM and SVM-SF+SVM-SURF techniques demonstrated near-optimal outcomes. However, the Mask RCNN approach reached the maximum object detection performance with a high $prec_n$ of 61%, $reca_l$ of 65%, and an $F_{measure}$ of 64%.

Fig. 5 showcases the object detection results achieved by the Mask RCNN approach and other techniques on the applied object, 'Vehicle'. The figure shows that the k-NN+SVM algorithm attained the least object detection result with a $prec_n$ of 84%, $reca_l$ of 96%, and a $F_{measure}$ of 90%. Likewise, the k-NN technique reached somewhat improved performance with a $prec_n$ of 85%, $reca_l$ of 95%, and an $F_{measure}$ of 90%.

Meanwhile, SVM-SF and SVM-SF+SVM-SURF methods produced moderate and similar outcomes with a $prec_n$ of 86%, $reca_l$ of 93%, and an $F_{measure}$ of 89%. In addition, the SVM-SURF technique portrayed near-optimal outcomes with a precision of 87%, $reca_l$ of 68%, and a $F_{measure}$ of 76%. Finally, the Mask RCNN approach achieved the highest object detection performance with its $prec_n$ value of 89%, $reca_l$ value of 98%, and $F_{measure}$ value of 94%.

Next, the image compression performance of the NCSIC technique was examined in terms of Compression Ratio (CR), Compression Factor (CF), and Space Saving (SS), and the results are shown in Table II. The results portray that the NBWT system reached a minor compression outcome with CR, CF, and SS values such as 0.2598, 3.8494, and 74.022% respectively. Also, the LZMA technique attained a somewhat increased compression performance with CR, CF, and SS values such as 0.2469, 4.0502, and 75.310%, respectively. Moreover, the JPEG technique produced a moderately closer compression outcome with CR, CF, and SS values such as 0.2332, 4.2877, and 76.677% respectively. However, the NCSIS technique resulted in an effective outcome with a CR of 0.2159, CF of 4.6308, and an SS of 78.406%.

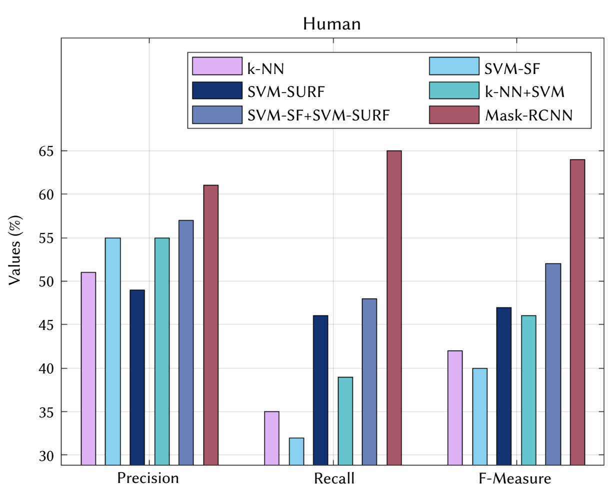

Fig. 4. Object Detection Analysis Results of Mask RCNN model on Human.

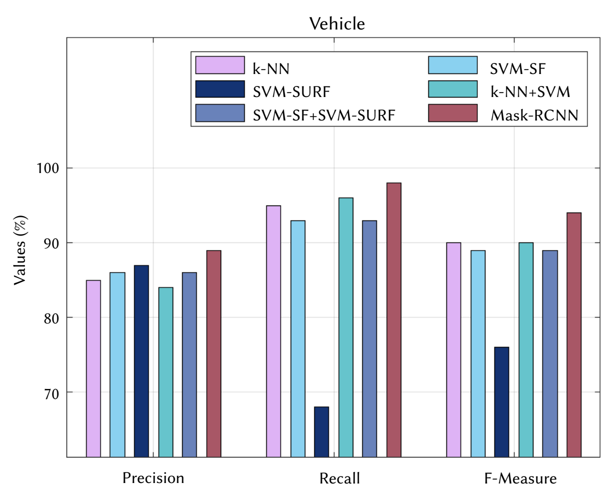

Fig. 5. Object detection analysis Results of Mask RCNN model on Vehicle.

TABLE I. Results of the Analysis of Existing Techniques and the Proposed Mask-RCNN Method Under Distinct. Measures

Models	Human			Vehicle		
	Precision	Recall	F-Measure	Precision	Recall	F-Measure
k-NN	51.00	35.00	42.00	85.00	95.00	90.00
SVM-SF	55.00	32.00	40.00	86.00	93.00	89.00
SVM-SURF	49.00	46.00	47.00	87.00	68.00	76.00
k-NN+SVM	55.00	39.00	46.00	84.00	96.00	90.00
SVM-SF+SVM-SURF	57.00	48.00	52.00	86.00	93.00	89.00
Mask-RCNN	61.00	65.00	64.00	89.00	98.00	94.00

TABLE II. Analysis Results of the Proposed and Existing Methods Under Different Measures

Methods	CR	CF	Space Savings (%)
NCSIC	0.2159	4.6308	78.406
JPEG	0.2332	4.2877	76.677
LZMA	0.2469	4.0502	75.310
BWT	0.2598	3.8494	74.022

Afterward, the performance analysis of the proposed IAIVS-WMSN approach was conducted consernîng to First Node Die (FND), Half Node Dies (HND), and the Total Remaining Energy (TRE) and the results are shown in Table III and Fig. 6. The outcomes show that Clamped Homogenous Electric Field (CHEF) method obtained the least compression outcome with FND, HND, and TRE values such as 10, 139, and 160J, respectively.

TABLE III. Analysis Results of the Proposed IAIVS-WMSN Method and Existing Methods Under Distinct Measures

Methods	FND (# of rounds)	HND (# of rounds)	TRE (J)
CHEF	10	139	160
EEUC	12	151	220
MOFCA-Original	16	155	233
MOFCA-Optimized	16	167	243
TTDFP	18	179	259
IAIVS-WMSN	25	184	278

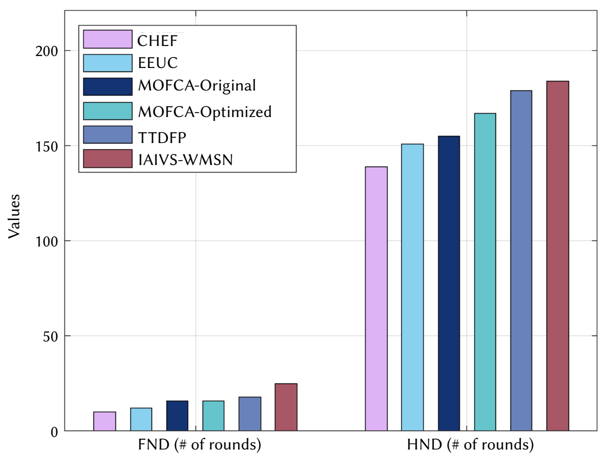

Fig. 6. Results of the Analysis of IAIVS-WMSN Model under Distinct Measures.

Simultaneously, Energy Efficient Unequal Clustering (EEUC) approach gained somewhat improved compression performance results with FND, HND, and TRE values such as 12, 151, and 220J respectively. The Multiobjective Fuzzy Clustering Algorithm (MOFCA)-an original technique demonstrated even more enhanced results with FND, HND, and TRE values such as 16, 155, and 223J respectively. Besides, the MOFCA-Optimized method reached somewhat higher performance with FND, HND, and TRE values such as 16, 167, and 243J correspondingly. JPEG algorithm produced a moderately closer compression result with FND, HND, and TRE values such as 18, 179, and 259J correspondingly. Eventually, the IAIVS-WMSN methodology produced effectual outcomes with an FND of 25, HND of 184, and a TRE of 278J.

Table IV shows the comparative analysis results achieved by the IAIVS-WMSN model and other existing techniques in terms of Packet Delivery Ratio (PDR), End to End Delay (ETED), and throughput. Fig. 7 shows the PDR analysis results of the IAIVS-WMSN algorithm under several nodes. The figure implies that the proposed IAIVS-WMSN system showcased better outcomes with a maximum PDR over other algorithms. For sample, with 100 nodes, the IAIVS-WMSN system reached an enhanced PDR of 91%, whereas TTDFP, MOFCA-opt., MOFCA-Org., EEUC, and CHEF offered fewer PDR values such as 90%, 90%, 90%, 90%, and 89% respectively. In the meantime, with 300 nodes, the proposed IAIVS-WMSN manner gained a superior PDR of 92%, whereas TTDFP, MOFCA-opt., MOFCA-Org., EEUC, and CHEF obtained the least PDR values such as 90%, 89%, 89%, 88%, and 86% correspondingly. Eventually, with 500 nodes, the proposed IAIVS-WMSN algorithm achieved the highest PDR of 90%, whereas TTDFP, MOFCA-opt., MOFCA-Org., EEUC, and CHEF accomplished the least PDR values such as 89%, 88%, 87%, 86%, and 84% correspondingly.

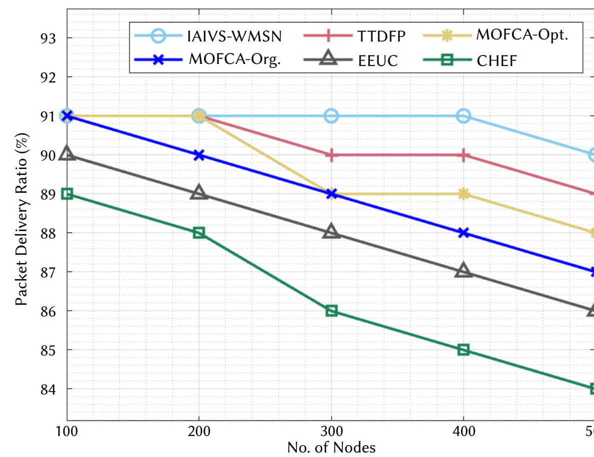

Fig. 7. PDR Analysis Results of IAIVS-WMSN Model and other Existing Techniques.

Fig. 8 shows the ETED analysis results of the proposed IAIVS-WMSN technique under several counts of nodes. The figure portrays that the proposed IAIVS-WMSN technique provided an effective outcome with the least ETED over other systems. For a sample, with 100 nodes, the IAIVS-WMSN algorithm achieved a minimum ETED of 3.70s, whereas TTDFP, MOFCA-opt., MOFCA-Org., EEUC, and CHEF achieved the maximum ETED values such as 3.80s, 3.85s, 3.90s, 4.50s, and 5.80s respectively. Moreover, with 300 nodes, the IAIVS-WMSN method demanded a minimal ETED of 4.60s, whereas TTDFP, MOFCA-opt., MOFCA-Org., EEUC, and CHEF demanded high ETED values such as 5.00s, 5.30s, 6.50s, 7.80s, and 8.10s respectively. Furthermore, with 500 nodes, the proposed IAIVS-WMSN approach demanded the least ETED of 6.80s, whereas TTDFP, MOFCA-opt., MOFCA-Org., EEUC, and CHEF required high ETED values such as 7.50s, 7.70s, 7.90s, 10.60s, and 11.00s correspondingly.

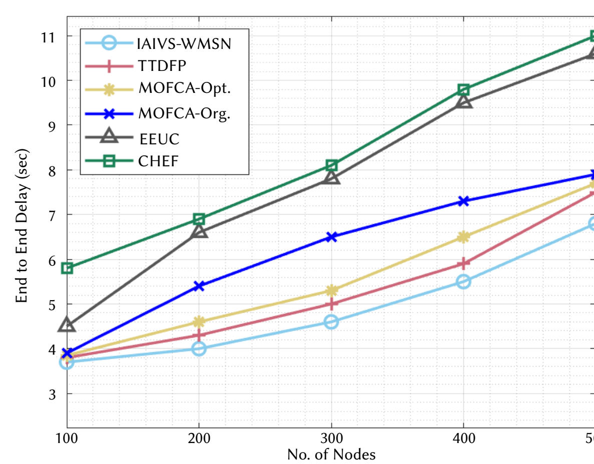

Fig. 8. ETED Analysis Results of IAIVS-WMSN Model and other Existing Techniques.

TABLE IV. Analysis Results of the Proposed and Existing Methods Under Different Measures

No. of Nodes	Packet Delivery Ratio (%)					
	IAIVS-WMSN	TTDFP	MOFCA-opt.	MOFCA-Org.	EEUC	CHEF
100	91.00	90.00	90.00	90.00	90.00	89.00
200	92.00	91.00	91.00	90.00	89.00	88.00
300	92.00	90.00	89.00	89.00	88.00	86.00
400	93.00	90.00	89.00	88.00	87.00	85.00
500	90.00	89.00	88.00	87.00	86.00	84.00
No. of Nodes	End to End Delay (sec)					
	IAIVS-WMSN	TTDFP	MOFCA-opt.	MOFCA-Org.	EEUC	CHEF
100	3.70	3.80	3.85	3.90	4.50	5.80
200	4.00	4.30	4.60	5.40	6.60	6.90
300	4.60	5.00	5.30	6.50	7.80	8.10
400	5.50	5.90	6.50	7.30	9.50	9.80
500	6.80	7.50	7.70	7.90	10.60	11.00
No. of Nodes	Throughput (Mbps)					
	IAIVS-WMSN	TTDFP	MOFCA-opt.	MOFCA-Org.	EEUC	CHEF
100	94.00	93.00	92.00	91.00	87.00	76.00
200	91.00	87.00	81.00	78.00	74.00	65.00
300	85.00	80.00	72.00	68.00	65.00	56.00
400	77.00	72.00	66.00	61.00	54.00	49.00
500	74.00	68.00	61.00	57.00	49.00	42.00

Fig. 9 shows the throughput analysis results achieved by the IAIVS-WMSN approach and other techniques under various node counts. The figure reveals that the proposed IAIVS-WMSN technique outperformed all other techniques with optimum results and maximal throughput. For a sample, with 100 nodes, the IAIVS-WMSN algorithm gained a superior throughput of 94Mbps, whereas TTDFP, MOFCA-opt., MOFCA-Org., EEUC, and CHEF achieved the least throughput values such as 93Mbps, 92Mbps, 91Mbps, 87Mbps, and 76Mbps correspondingly. Besides, with 300 nodes, the proposed IAIVS-WMSN system gained an improved throughput of 85Mbps, whereas TTDFP, MOFCA-opt., MOFCA-Org., EEUC, and CHEF achieved the least throughput values such as 80Mbps, 72Mbps, 68Mbps, 65Mbps, and 56Mbps correspondingly. Finally, with 500 nodes, the proposed IAIVS-WMSN technique achieved the maximum throughput of 74Mbps, whereas TTDFP, MOFCA-opt., MOFCA-Org., EEUC, and CHEF attained the least throughput values such as 68Mbps, 61Mbps, 57Mbps, 49Mbps, and 42Mbps correspondingly. From the results mentioned above and the discussion, it is clear that the proposed IAIVS-WMSN algorithm is an excellent tool for effective communication in WMSN.

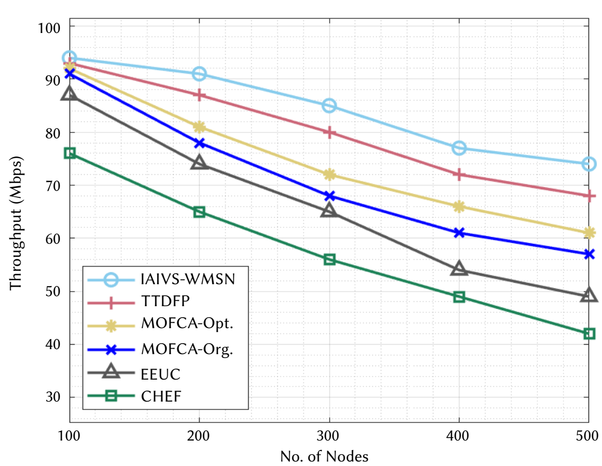

Fig. 9. Throughput Analysis results of IAIVS-WMSN Model and other Existing Techniques.

V. Conclusion

In this study, a novel IAIVS-WMSN algorithm has been designed and developed to accomplish effectual data transmission in WMSN. The presented IAIVS-WMSN system comprises three stages: Mask RCNN-based object detection, NCSIC-based image compression, and the AFA-based clustering. In addition, AFA system derived a Fitness Function containing various input parameters for effective selection of CHs and proper construction of clusters in WMSN. A comprehensive experimental analysis was conducted to validate the superiority of the proposed the IAIVS-WMSN technique. The simulation outcomes confirmed the enhanced performance of IAIVS-WMSN approach to the existing algorithms. In future, effective background subtraction and multipath route planning techniques can be designed to improve network efficiency.

References

[1] I. F. Akyildiz, T. Melodia, and K. R. Chowdhury, “A survey on wireless multimedia sensor networks,” Computer networks, vol. 51, no. 4, pp. 921–960, 2007.

[2] W. Fang, D. H. Wang, and Y. Wang, “Energy-efficient distributed target tracking in wireless video sensor networks,” International Journal of Wireless Information Networks, pp. 1– 11, 2015.

[3] M. Koyuncu, A. Yazici, M. Civelek, A. Cosar and M. Sert, “Visual and auditory data fusion for energy-efficient and improved object recognition in wireless multimedia sensor networks,” IEEE Sensors Journal, 19(5), pp.1839-1849, 2018.

[4] H. ZainEldin, M. A. Elhosseini and H. A. Ali, “Image compression algorithms in wireless multimedia sensor networks: A survey,” Ain Shams Engineering Journal, vol. 6, pp. 481–490, 2015.

[5] Q. Lu, W. Luo, J. Wanga and B. Chen, “Low-complexity and energy efficient image compression scheme for wireless sensor networks,” Computer Networks, vol. 52, pp. 2594–2603, 2008.

[6] Donoho DL. Compressed sensing. IEEE Trans Inform Theory2006;52:1289–306.

[7] T. Sheltami, M. Musaddiq and E. Shakshuki, “Data compression techniques in Wireless Sensor Networks,” Future Generation Computer Systems, vol. 64, pp. 151-162, 2016.

[8] D. M. Pham and S. M. Aziz, “Object extraction scheme and protocol for energy efficient image communication over wireless sensor networks,”

Computer Networks, vol. 57, pp. 2949–2960, 2013.
[9] M. Civelek and A. Yazici, "Automated moving object classification in wireless multimedia sensor networks," IEEE Sensors Journal, vol. 17, pp. 1116-1131, 2017.
[10] Y.A.U. Rehman, M. Tariq and T. Sato, "A novel energy efficient object detection and image transmission approach for wireless multimedia sensor networks," IEEE Sensors Journal, 16(15), pp.5942-5949, 2016.
[11] Q. Guo, "Enhanced compressed sensing for visual target tracking in wireless visual sensor networks," Journal of Electronic Imaging, 26(6), p.063028.
[12] M. Koyuncu, A. Yazici, M. Civelek, A. Cosar and M. Sert, "Visual and auditory data fusion for energy-efficient and improved object recognition in wireless multimedia sensor networks," IEEE Sensors Journal, 19(5), pp.1839-1849, 2018.
[13] A.N. Sukumaran, R. Sankararajan, and M. Swaminathan, "Compressed sensing based foreground detection vector for object detection in wireless visual sensor networks," AEU-International Journal of Electronics and Communications, 72, pp.216-224, 2017.
[14] D. Wang, J. Liu, D. Yao and I.E.E.E. Member, "An energy-efficient distributed adaptive cooperative routing based on reinforcement learning in wireless multimedia sensor networks," Computer Networks, 178, p.107313.
[15] M. Akter, M.O. Rahman, M.N. Islam, M.M. Hassan, A. Alsanad, and A.K. Sangaiah, "Energy-efficient tracking and localization of objects in wireless sensor networks," IEEE Access, 6, pp.17165-17177.
[16] Y. Shao, "Wireless multimedia sensor network video object detection method using dynamic clustering algorithm," Multimedia Tools and Applications, 79(23), pp.16927-16940.
[17] P. Sathyaprakash, and P. Prakasam, "Boltzmann randomized clustering algorithm for providing quality of evolution in wireless multimedia sensor networks," Wireless Personal Communications, 112(4), pp.2335-2349.
[18] S.Heng, Aimtongkham, P., Vo, V.N., Nguyen, T.G. and So-In, C., 2020. Fuzzy Adaptive-Sampling Block Compressed Sensing for Wireless Multimedia Sensor Networks. Sensors, 20(21), p.6217.
[19] B. A. Y. Alqaralleh, S. N. Mohanty, D. Gupta, A. Khanna, K. Shankar and T. Vaiyapuri, "Reliable Multi-Object Tracking Model Using Deep Learning and Energy Efficient Wireless Multimedia Sensor Networks," in IEEE Access, vol. 8, pp. 213426-213436, 2020, doi: 10.1109/ACCESS.2020.3039695.
[20] C. Xu, G. Wang, S. Yan, J. Yu, B. Zhang, S. Dai, Y. Li and L. Xu, "Fast Vehicle and Pedestrian Detection Using Improved Mask R-CNN," Mathematical Problems in Engineering, 2020.
[21] X. Wu, H. Shao, S. Wang and W. Wang, "A Multi-objective Artificial Flora Optimization Algorithm," In International Conference on Simulation Tools and Techniques (pp. 267-277). Springer, Cham.
[22] J. Uthayakumar, M. Elhoseny and K. Shankar, "Highly reliable and low-complexity image compression scheme using neighborhood correlation sequence algorithm in WSN," IEEE Transactions on Reliability, 69(4), pp.1398-1423, Dec. 2020.

Romany Fouad Mansour

Romany Fouad Mansour received a Ph.D. degree from the University of Assiut in 2009, He is currently working as an Associate Professor at Department of Mathematics, Faculty of science, New Valley University, Egypt. His research interests include Artificial Intelligence, Computer Vision, Computer Networks, Soft computing, Image Processing and Bioinformatics.

Carlos Soto

Carlos Soto received a B.S. degree in Mechanical Engineering from the Universidad del Norte (Colombia) in 2005. He was holder an Erasmus Mundus Scholarship in a Master of Advanced Materials Science and Engineering in 2007, by Universidad Politécnica de Cataluña (Spain) and Luleå Technology University (Sweden). Currently, he is a full-time professor in the Engineering Faculty of Universidad Autónoma del Caribe since 2016.

Roosvel Soto-Díaz

Roosvel Soto-Díaz received the B.S. and M.Sc. degrees in Materials Engineering and Mechanical Engineering from Universidad Autónoma del Caribe, respectively, in 2016 and 2018. He obtained meritorious and scientific distinction from his undergraduate and master's degree work. At present, he is a professor at the Universidad Simón Bolívar since 2021.

José Escorcia-Gutierrez

José Escorcia-Gutierrez received a Ph.D. degree in Computer Engineering and Mathematics of Security from the Universitat Rovira i Virgili in 2021 (Spain). In 2009 and 2011, he received the B.S. and M.Sc. degrees in Electronic Engineering from the Universidad del Norte (Colombia). His research interests include image processing, pattern recognition, computer vision, machine learning, and medical image analysis. Currently, He is a full-time professor in the Electronics and Telecommunication Engineering program of the Universidad Autónoma del Caribe, and he works under contract at the Corporación Universitaria Reformada.

Deepak Gupta

Dr. Deepak Gupta received a B.Tech. degree in 2006 from the Guru Gobind Singh Indraprastha University, Delhi, India. He received M.E. degree in 2010 from Delhi Technological University, India and Ph. D. degree in 2017 from Dr. APJ Abdul Kalam Technical University (AKTU), Lucknow, India. He has completed his Post-Doc from National Institute of Telecommunications (Inatel), Brazil in 2018. He has co-authored more than 145 journal articles including 109 SCI papers and 45 conference articles. He has authored/edited 53 books, published by IEEE-Wiley, Elsevier, Springer, Wiley, CRC Press, DeGruyter and Katsons. He has filled four Indian patents. He is convener of ICICC, ICDAM & DoSCI Springer conferences series. Currently he is Associate Editor of Expert Systems (Wiley), and Intelligent Decision Technologies (IOS Press). He is the recipient of 2021 IEEE System Council Best Paper Award. He has been featured in the list of top 2% scientist/researcher database in the world [Table-S7-singleyr-2019]. He is also working towards promoting Startups and also serving as a Startup Consultant. He is also a series editor of "Elsevier Biomedical Engineering" at Academic Press, Elsevier, "Intelligent Biomedical Data Analysis" at De Gruyter, Germany, "Explainable AI (XAI) for Engineering Applications" at CRC Press. He is appointed as Consulting Editor at Elsevier.

Ashish Khanna

Dr. Ashish Khanna [IEEE Member'2019, Senor Member'2020] has expertise in Teaching, Entrepreneurship, and Research & Development with specialization in Computer Science Engineering Subjects. He received his Ph.D. degree from National Institute of Technology, Kurukshetra in March 2017. He has completed his M. Tech. in 2009 and B. Tech. from GGSIPU, Delhi in 2004. He also has completed his Post Doc from INATEL, Brazil and Universiity of Valladolid, Spain. He has around 170 accepted and published research papers and book chapters in reputed SCI, Scopus journals, conferences and reputed book series including 90 papers accepted and published in SCI indexed Journals. He also has 8 published Patents and design patents to his credit. Additionally, He has co-authored, edited and editing around 45 books. He has also served as a keynote speaker, resource person and given several invited. He was resource person speaker and delivered a lecture at the online ATAL program at JNU in 2020. He has worked as Guest Editor in several reputed SCI and SCOPUS journals. He is also serving as Series Editor in publishing houses like De Gruyter (Germany) of "Intelligent Biomedical Data Analysis" series, Elsevier of "Intelligent Biomedical Data Analysis" and CRC Press of "Intelligent Techniques in Distributed Systems: Principles and Applications". He is also acting as a consulting editor for Elsevier. His research interest includes, Distributed Systems and its variants (MANET, FANET, VANET, IoT), Machine learning. He is the recipient of 2021 IEEE System Council Best Paper Award. He is serving as convener in some springer international conferences series like ICICC, ICCCN, ICDAM, and DoSCI. He is senior IEEE member

(SMIEEE) and ACM member too. He has played key role in promoting and initiating several startups and is also a startup consultant. He has also played the key role in the establishment of Bhavya Books (A Delhi Based Publishing House) and also promoted Smart India Hackathon at MAIT. He is key member and originator of Universal Inovator an International Research Lab. He initiated the first ever of its kind event in India "WHERE STARTUP MEETS INVESTOR" in collaboration with Universal Inovator and SIIF SSCBS, DU under the banner of ICICC-2020 international conference. First time a Research based conference was conducted along with a Startup funding event. He also played the key role in the India's first conference on Patents (ICIIP-2021) in association with UI and IGDTUW Delhi.