mahalingam 3.pdf

by Turnitin LLC

Submission date: 02-Aug-2024 06:10PM (UTC+1000)
Submission ID: 2426200768
File name: uploads_5837_2024_08_02_mahalingam_3_e9017910048d3c77.pdf (697.94K)
Word count: 4614
Character count: 25063

Real-Time Dynamic Object Detection Using DNN and GSA

12

Abstract—**This paper presents a novel approach for real-time dynamic object detection using deep belief neural networks (DNN) and the Gravitational Search Algorithm (GSA). The methodology, called DMFE, uses foreground detection to accurately track moving objects in complex surroundings. The approach incorporates an iterative weight factor, achieving a weighted fusion of colour/grey intensities and gradient magnitudes, enhancing detection accuracy. The method is implemented in MATLAB and evaluated on PETS and Hall monitor data sets. Ten metrics are employed to assess the proposed approach, demonstrating its superiority over existing algorithms in terms of accuracy and recall. This paper highlights the effectiveness of DNNs optimized with GSA for real-time dynamic object detection, contributing to advancements in computer vision and intelligent systems.**

Keywords—Real-time object detection, deep belief neural networks, Gravitational Search Algorithm, foreground detection, dynamic object tracking, computer vision.

I. Introduction

The complex field of moving object detection (MoD), which is essential to many applications like security systems, human-computer interfaces, aviation and ocean monitoring, and more, has been the focus of my study [1]. Numerous contemporary technologies, like as cloud computing, IoT, edge computing, automation, intelligent transportation, and even autonomous driving, are useful for the dynamic nature of MoD [2–4].

Nonetheless, there are many difficulties in the MoD environment. Every challenge offers a different riddle to work out, from shifting lighting to strange object movements, from unstable video feeds to complicated motion backdrops [3]. Despite methodological advances [5, 6, 7, 8, 9, 10, 11, 12], dynamic foreground object detection in films with complex backdrops is still a challenging task.

Bharath and Dhivya [13] suggested a method for tracking and classifying objects according to their attributes, especially their speed, in order to overcome these issues. Using this basis as a starting point, Mithun et al. [14] developed a spatial and temporal feature-based vehicle recognition method to address issues with congested vehicle videos and the minute differences between cars and their surroundings. Utilizing contours, textures, and additional features of the motion object separation zone, they improved classification accuracy by utilizing the K-Nearest Neighbor (KNN) algorithm.

Tian et al. [15] investigated object classification by the combination of regular and geometric shape attributes, although this approach required more computing power. Conversely, Dai et al. [16] proposed a strong recognition method for walkers and two-wheelers by extracting four form attributes from object picture areas and feeding them into a BP neural network. However, scalability was still an issue because this approach could only classify a small number of distinct object kinds according to their shapes.

Building upon these principles, our research endeavored to enhance the object identification process by exploring certain object attributes such area, proximity, speed, and length-width ratios [17]. Using the center of mass of object zones and different curve points in the object space, we produced multi-granularity perceptual features to represent moving objects using object area recognition [18]. Using a two-level Support Vector Machine (SVM) classification, our method successfully categorized autos and pedestrians in intricate settings.

In the middle of the abundance of methods for identifying mixed feature integration, we created a brand-new dynamic feature fusion method. Our method efficiently fuses color characteristics with gradient magnitudes (GM) features, in contrast to existing approaches, by giving GM values more weight when they are high. Our Dynamic Magnitude Feature Fusion (DMFE) technique separates moving objects into the foreground and moving backgrounds into the background sector, thereby eliminating false positives and concentrating only on dynamic object detection [19].

Conventional techniques, including frame difference-based approaches, have difficulties when it comes to background/foreground recognition, especially in movies with changing backgrounds and lighting conditions [21]. While statistical and cluster-based algorithms frequently struggle with movies showing significant degrees of background fluctuation, geometric and cluster-based recognition methods have attempted to minimize these issues [8, 22, 23, 24, 25, 26].

Although at a higher processing cost, the introduction of artificial intelligence (AI) and deep learning-based detection techniques [27, 28] has greatly improved detection accuracy. Several methods have used sample consensus methodologies to prevent false positives [29, 30, 19, 31, 32, 33, 34, 35].

II. Literature survey:

The Classification of Systems for Detecting Dynamic Objects includes,

Easy Approach:
Colour fluctuations and subsequent or close frames are used to identify dynamic objects and techniques like shade distinction histogram and consecutive blocks of frames are used.

Statistical Method/Clustering:
Gaussian mixer model (GMM) is used for cluster operations. Also, Balloon estimator approach and KNN kernel have improved accuracy.

Consensus-Driven Approach:
Ensures consistent foreground recognition by fragmenting detected pixels as background.

Learning-Centred Approach:
Machine learning and deep learning designs are essential for classifying items as background or foreground.

Someexamples include spatiotemporal feature-routing networks, incremental maximum margin criterion, and learning from many viewpoints structure.

The Incremental Maximum Margin Criterion (IMMC):IMMC dynamically modifies eigenvalues and eigenvectors to respond to changes and outperforms earlier subspace learning methods, but requires the initial backdrop model.

Object Detection Tracking in Computer Vision:
Azab et al. used a steady cam in a particle filtering system for high detection and tracing rates. Shuai et al. proposed spatial-temporal feature-routing networks for video object recognition.Lin et al. presented an innovative approach to foreground object recognition, emphasizing movement cues.

III. Problem formulation:

Object detection tracking is crucial in real-time applications like criminal action detection and video monitoring. Challenges include object selection, object classification, and detection methods. Techniques include, Equating two frameworks, using colour variance, median-based techniques, Gaussian Motion Modelling (GMM), Bayesian Background Detection System (BGS), and ViBe. However, these methods struggle with locating periodic object activity and integrating objects into scenes. IMMC, an incremental optimal margin standard subspace learning technique, outperforms previous subspace knowledge approaches but requires ground truth. The objective is to separate background and forefront, use DMFE for object segmentation, and achieve real-time object tracing using Optimum DNN based on GSA. However, IMMC requires large data to train for accurate separation results.

IV. Proposed DMFE:

We suggest the Dynamic Multi-Feature Extraction (DMFE) method, which is illustrated in Figure 1 and includes both mathematical and technological steps. In order to maintain movement and define an end-to-end processing flow, the DMFE technique integrates knowledge from prior research to detect moving targets pixel by pixel. Here, we use known techniques, hyperparameters, and experimental settings to briefly outline and go over the main procedural phases of the DMFE approach:

Sequences of input:
Process input frames or movies with uneven surfaces in RGB or grayscale iteratively. RGB/gray structures, Ic, c€{R,G,B}, are the inputs used in our model. Sajid et al. [51] have highlighted the resilience and efficiency of the RGB color model for object detection, which is why we choose it. To prevent noise amplification and the misinterpretation of important alteration information, input sequences are not filtered.

Gradient Scale Framework:
For input frames or movies, compute and normalize the gradient scale frames. The addition of these stabilized gradient series to the intensity arrangements is essential.

Background Setup:
Create backdrop gradients and shades based on the initial intensity series. Neighboring pixels in the background samples determine the background pixels of a perceived pixel.

Classification:
Based on how a seen pixel compares to similar backdrop pixels in the background samples, classify it as background or foreground.

Estimating Errors:
Calculate the separation errors in the classification based on noise from salt and pepper, which will help you customize the background factor and boundary problem at the same time.

Updates to the Background and Threshold Features:
The shortest distance between consecutive background frames and the amount of the detection error are taken into consideration while updating the background and threshold characteristics.

Adjustments to Color and Gradient Threshold:
Color and gradient thresholds should be dynamically computed and adjusted to account for both static and non-static backdrop conditions in order to minimize false detections caused by intensity variations over time.

Updated Background Sample:
Divide the background samples using division pixels and adjust backdrop samples, such as thresholds, based on the changing input settings.

Post-production:
Eliminate flashing noise and close gaps in the object forms of the unfiltered division series.

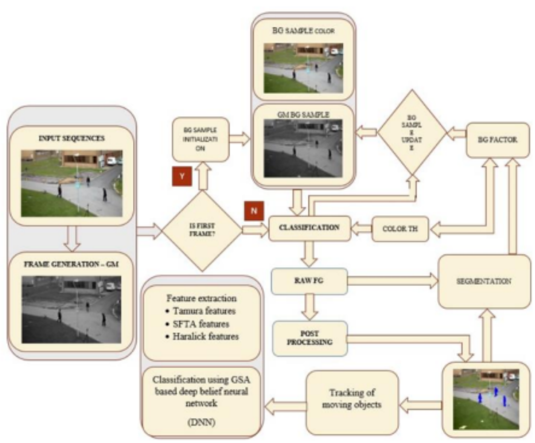

figure 1. The proposed method's flow diagram

Object monitoring is an active optimizing technique that is based on temporal data that are linked to prior structures. Recommending a process with more correctness in intricate situations is a provocation for scientists doing research. Effective target finding and motion tracking videos are suggested in this study. The suggested method includes two key actions: background and foreground dividing apart and tracing motion items.The last operation is to track motion objects; first, the attributes are extracted from the segregated structure using Tamura attributes, SFTA attributes, and Haralick attributes. Following attribute extraction, the objects are traced using a deep belief neural network based on GSA (DNN). Figure 1 depicts the overall method. The recommended method's block layout is shown below.

V. Tracking of moving objects

Following that, the backdrop and foreground are separated, and the items are traced. Object monitoring is the tracking of an item in a video section. Object monitoring is the progression of progressively tracking a motion object (or many things) with a camera. It has various applications, including machine-to-machine communication, safety and security, surveillance, video interaction and compression, amplified realism, mobility controller, clinical imaging, and video series editing. In our recommended technique, two critical operations are employed for monitoring: attribute extraction and object detection using classification.

1. Extraction of attributes:

Following that, a high-quality background was created, and attribute data were extracted from the backdrop image. The proposed technique uses Tamura attributes, SFTA attributes, and Haralick attributes to mine out the attribute value. In the supplementary area, you will get a full depiction of the attribute abstractionprocess.

2. Tamura features:

Coarseness: The goal was to evaluate the differences between rough and finely smoothed surfaces. It includes information about the smoothness components' sizes. By connecting the latter phases, we can determine the image's roughness.

- For every point (a, b) in the input image, averages for the neighborhood have been determined, and their sizes are $2^s \times 2^s$, where $s \in <0,...5>$;

$$A_s(a,b) = \frac{1}{2^{2^s}} \sum_{i=a-2^{s-1}}^{a+2^{s-1}-1} \left(\sum_{j=b-2^{s-1}}^{b+2^{s-1}-1} f(i,j) \right) \qquad (1)$$

- For each point, calculate the average of the uncorrelated neighbors on the opposite side in both the longitudinal and transverse dimensions, as well as the difference between pairs.

For horizontal,

$$E_{s,h} = |A_s(a+2^{s-1},b) - A_s(a-2^{s-1},b)| \qquad (2)$$

For vertical,

$$E_{s,y} = |A_s(a,b+2^{s-1}) - A_s(a,b-2^{s-1})| \qquad (3)$$

- After each directive is taken into account in combination with the others at each factor, determine the dimensionality that produces the highest output value. The middle value of these values throughout the entire image will be the roughness amount for the provided image.

Contrast: The active grayscale scale in the image, the distribution of black and white, the intensity of the sides, and the duration of recurrent patterns all control how distinct the image is from one another. In the limited context, it may also reflect the target feature.

$$Contrast = \frac{\sigma}{\sqrt[4]{\gamma_4}} \; where\; \gamma_4 = \frac{\mu_4}{\sigma^2} \qquad (4)$$

Where, μ_4 is the fourth mean value and σ^2 is the variance value.

Directionality: Rather than focusing on coordination per se, it looks at how coordination occurs in the image. If the only difference between two images is synchronization, then this composition's degree is equal to theirs.

$$\theta = \tan^{-1}\frac{\Delta v}{\Delta h} + \frac{\Pi}{2} \qquad (5)$$

Where $\Delta v \; and \; \Delta h$ is the perpendicular and parallelimplications.

3. SFTA Features:

In SFTA attributes, the input photos have primarily deteriorated and been replaced by a series of binary images that depend on the threshold levels in the threshold set (T). The entire range of threshold values in the threshold set was preferred by the customer. Three texture properties are mined: mean grey degree, dimensionality, and self-similar assessment, since all extracted binary pictures are extracted.

4. Haralick Features:

Contrast: To show structural perfection, the Contrasting characteristic is a metric of imaging intensity comparability or natural fluctuations that are common in a specified area. This criterion is shown by the following equation:

$$Contrast = \sum_{x,y=0}^{N-1} P_{xy}(x-y)^2 \qquad (6)$$

Energy: [0, 1] is the potentiality level. One is said to be the potentiality value for a target that has been healed. The potentiality assessment formula is as follows.

$$Energy = \sum_{x,y} p(x,y)^2 \quad (7)$$

Where, $p(x,y)$ is the pixel value x,y at the textured image's size point($M \times N$)

Entropy: The degeneration helps to determine the propagation alteration in a target region and to stand in for the texture target. The objective problem is effectively approximated by the consistent criteria. The deterioration is assessed using the following Formula.

$$Ent = \sum_{x,y} p(x,y)\log(p(x,y)) \quad (8)$$

Homogeneity: Homogeneity assesses the consistency of accessibility that is not zero. The homogeneousness is defined by the range [0, 1]. The homogeneousness is one if the aim is never altered, and it is higher if the target is only slightly altered.

$$Hom\,ogeneity = \sum_{x,y} \frac{P(x,y)}{[1+(x-y)^2]} \quad (9)$$

As a result, the attribute values are extracted using the procedure described above. The drawn-out properties are then input into the target detection tracing mechanism.

VI. Detection and Tracking of Objects

Our suggested method uses GSA-based deep belief neural networks to identify and track targets (DNN). The foregoing provides a thorough explanation of how to use a deep belief neural network (DNN) based on GSA for target recognition and tracking.

Deep Belief Network Pre-Training Stage (DBN):

We use a feed-forward neural network with many hidden units and an unknown design called the Deep Belief Network (DBN) in the pre-training stage. The DBN's input is given the highest quality possible, and the weights that are found are used as the DBN's starting weights. These weights are then used to pre-train a Deep Neural Network (DNN). Through a process called network approval, the DBN design allows the network to build visible features based on its hidden unit states. This architecture consists of an input unit that hides the input devices, a result unit that has a single device for each assessment class, and Lis series of concealed units.

The biases (Bi) of layer I and the weights Wi between the unit systems make up a DBN's criterion. Creating the standards for a deep neural network is difficult since random initialization frequently results in optimization processes that restrict the search for error task minima and reduce generalization. We use a Restricted Boltzmann Machine (RBM) to tackle this problem.

Restricted Boltzmann Machine (RBM): An RBM Markov structure with one level of unpredictable observable units and one level of unpredictable concealing units that is arbitrarily determined. RBMs are two-dimensional diagrams in which every concealed unit is associated with every observable unit, and there are no observable or hidden linkages. We study the energy processes of the exposed level and the hidden level during the preparatory phase.

A different RBM can be "loaded" on top of the RBM wherever it is used to create a multiple layer system. The input observable level for a new RBM layer is a primer vector, and the current weightiness and biases determine the amounts for the units in earlier RBM layers. The last layer of the previously constructed layers is complex and serves as input to the first RBM. By following the above technique, a state-of-the-art RBM is produced, and the entire procedure can be repeated until the essentially final condition is satisfied. Our opinions on how significant each RBM resolution of declining multiple-objective characteristics is differ. Refinement stage training complicates deep network weightiness.

Procedure for training the DNN:

To begin training the DNN, we convert the observable systems to activation vectors.Following that, we utilise a formula to modify the concealed units while the apparent units are given (11).

- Similarly, we modify the apparent units in concurrently while the concealed units are defined in a formula (12).
- Once the hidden units are upgraded in parallel, the reconstructed observable system with the exact equivalent equation is used as step (13).

A different RBM can be "loaded" on top of the RBM wherever it is used to create a multiple layer system. The input observable level for a new RBM layer is a primer vector, and the current weightiness and biases determine the amounts for the units in earlier RBM layers. The last layer of the previously constructed layers is complex and serves as input to the first RBM. By following the above technique, a state-of-the-art RBM is produced, and the entire procedure can be repeated until the essentially final condition is satisfied. Our opinions on how significant each RBM resolution of declining multiple-objective characteristics is differ. Refinement stage training complicates deep network weightiness.

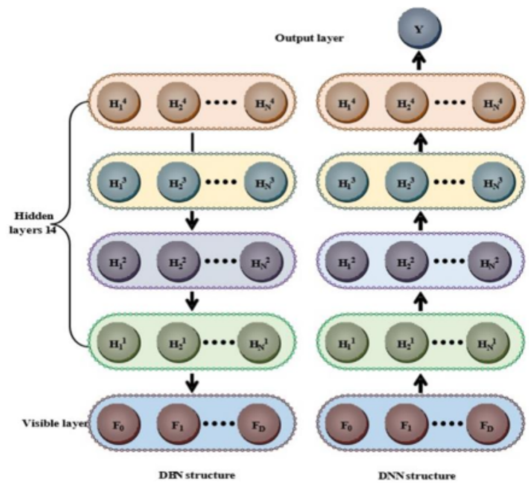

Figure 2. recommended DBN and DNN framework with four hidden layers

The connections between the top two level concealed layers in a DBN architecture are now symmetrical and directionless, while other connections are broadcast top-down. In the DNN configuration, there are only feed-forward associations. The weightiness of a qualified DBN can be used to determine a DNN's weightiness for discriminating fine tuning.

Algorithm for GSA

Gravity and the law of motion both cause the GSA to activate. The methodology is assembled using a population-based approach that connects multiple masses. Based on gravitational density, the masses are distributing information to direct the search to the best location inside the search region. Equations (15) and (16) demonstrate how GSA improved both Newton's rule of motion and his law of gravity.

$$F_{ij}^{d}(t) = G(t)\frac{AM(t)\times PM(t)}{D_{ij}^{2}} \tag{10}$$

$$a_i(t) = \frac{F_{ij}^{d}(t)}{M_i(t)} \tag{11}$$

Where;

AM_i → The agent is connected with dynamic gravitational mass i

PM_j →The passively gravitational mass is related j

$G(t)$ →One gravity invariant at a time t

D_{ij} → The Centroids distance between two agents i and j

$a_i(t)$ →Accelerate

$M_i(t)$ →The inertial mass

The members are now viewed as objects by this technique, and their regimen is assessed based on their masses. Each target is drawn in opposite directions by gravity, which initiates a global arrangement of everything moving toward larger masses. Because of this, some masses make a contribution by offering a direct path of interaction through gravitational force. Greater results are guaranteed at the strategy development stage with slower movement and heavier masses. Four requirements must be met by each mass (agent) in GSA: passively gravitational mass, active gravitational mass, kinematic energy, and localization. The problem is solved by the mass's location, and a fitness factor determines the masses that are centrifugal and gravitational. Each mass thus serves a purpose, and the treatment is focused by appropriately seasoning the centrifugal and gravitational masses. The greatest mass eventually absorbs the other masses, leading to the creation of the best possible option in the search space. The population is denoted by the weights that can be obtained; the position of the N agents is initialized arbitrarily; and the agent's location is represented by the associated connection:

$$w_i = \left(w_i^1, w_i^2, \ldots, w_i^d, \ldots, w_i^n\right) \; for = \left(i = 1,2,\ldots,N\right) \tag{12}$$

Where,

w_i^d → Weight position of i^{th} agent in the d^{th} dimension .

Step 2: Evolution of fitness and optimum fitness computation

The fitness evolution is done as a reduction problem by evaluating the best and vilest physical fitness for each agent at the sum of iterations.

$$Fit(t) = \min\ error \tag{13}$$

$$best(t) = \min_{j\in(1,\ldots,N)} Fit_j(t) \tag{14}$$

$$worst(t) = \max_{j\in(1,\ldots,N)} Fit_j(t) \tag{15}$$

Step 3: Calculate the gravitational constant (G).

To ensure search exactness, the gravity persistent G is engaged at the start and steadily reduced over time. As a consequence, G represents a function of the starting value G_0 and the time (t);

$$G(t) = G(G_0, t) \quad (16)$$

Using health and fitness analysis, we accurately analyze the gravitational and inertial masses. A more substantial weight denotes an especially trustworthy agent. To put it plainly, outstanding spokespeople move more slowly and have extraordinary attraction. As a map of fitness and health, mass values are estimated.

Step 4: Calculate the agent's mass.

Assuming equal rights for gravitational and inertial mass t.

$$AM_i = PM_j = M_i, \text{ where } i = 1, 2, \ldots, N$$

$$m_i(t) = \frac{Fit_i(t) - worst(t)}{best(t) - worst(t)} \quad (17)$$

$$M_i(t) = \frac{m_i(t)}{\sum_{j=1}^{N} m_j(t)} \quad (18)$$

Where,

$Fit_i(t)$ →Signifies the fitness value of the agent i at the time t

Step 5: Accelerations of agent calculation

Then, at time t, the motion of agent i may be stated as:

$$a_i(t) = \frac{F_{ij}^d(t)}{M_i(t)} \quad (19)$$

Step 6: Representative rates and places.

The consequent rate of an agent is considered as a fraction of its current motion and is included in its velocities. Its location and speed are now determined utilized the formulas shown below.

$$V_i^d(t+1) = rand_i \times V_i^d(t) + A_i(t) \quad (20)$$

$$R_i^d(t+1) = R_i^d(t) + v_i^d(t+1) \quad (21)$$

Where,

$rand_i$ →The range contains a uniformly distributed random variable [0, 1]

Step 7: Repetition of stages 2–6.

Steps 2 through 6 are carried out till the maximum limit of the versions is reached. The global health and fitness through the position of the constant representation at delimited metrics is calculated as the overall resolution to that particular problem, which refines the best health and fitness value in the most current edition. Ultimately, the objects in a deep neural network are identified by recycling the idealweightiness.Thus, the aforementioned method is used to monitor the things in the input videostream. The results are highlighted here, and the expected method's presentation is estimated.

VII. RESULT AND DISCUSSION

Our recommended optimum object identification and tracking technique based on DNN and the gravitational search procedure is executed in the MATLAB.Exploration is performed out with diverse datasets of motion, and the effectiveness of our activity is assessed using a variety of metrics.

1. **Description of the dataset:**

   The proposed analysis utilizes two moving object datasets: PETS 2009 and Hall monitor dataset. PETS 2009 is used for crowd image analysis, including population counting and densities. Hall monitor dataset consists of 300 CIF frames with minimal movement.

2. **Experimentation:**

   PETS and Hall monitor footage is used to test the suggested object tracking system; the dataset is divided into several frames for testing. DMFE techniques are used to separate the background from foreground, and then GSA-based DNN is used for object recognition and tracking.

Fig.3: Input video frames (a) PETS dataset (b) Hall monitor

DMFE methods are used to separate the background and foreground at first. Figure 4 depicts the final product.

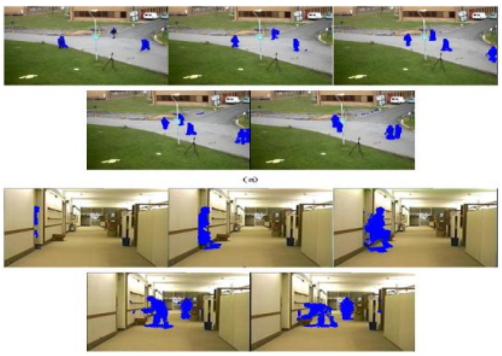

Fig.4: Result of object detection and tracking (a) PETS (b) Hall monitor

The segmented output is identified after the segmentation process is finished. In this instance, items are recognized and tracked using deep belief neural networks based on GSA (DNN). The findings from the motion detection stage are shown in Figure 4.

3. Metrics for Evaluation:

The suggested tracking approach for moving objects is evaluated using four metric values and ten measures, focusing on pixel differences to determine the required evaluation metric data.The ten measures are used to evaluate the effectiveness of this approach.

4. Precision:

The equation estimates how many pixels in the separated image that are separated to be Positive are truly Positive.

$$\Pr ecision = \frac{TP}{FP+TP} \quad (22)$$

5. Rate of Recall or Detection (DR):

Positives are accurately separated and expressed using recall. It is also known as a Detection Rate and is equivalent to Sensitivity.

$$\operatorname{Re} call = \frac{TP}{FN+TP} \quad (23)$$

6. F-Measure:

F-metric Measure's value is presented as a percentage of a harmonic mean for a mix of accurateness and recall metrics.

$$f-measure = \frac{2\left(\Pr ecision \times \operatorname{Re} call\right)}{\Pr ecision + \operatorname{Re} call} \qquad (24)$$

7. Rate of False Positives (FPR):

A proportion of states where the results suggest that the video frame was correctly recognized, although it was not.

$$FPR = \frac{FP}{FP+TN} \qquad (25)$$

8. False Negative Rate (FNR):

The proportion of occurrences in which the result indicates that the video frame was incorrectly categorised but it was really successful.

$$FPR = \frac{FP}{FP+TN} \qquad (26)$$

9. Percentage of Incorrect Classifications (PWC):

PWC is a percentage that represents an improperly segmented video frame

$$PWC = \frac{FP+FN}{TP+FP+TN+FN} \times 100 \qquad (27)$$

10. False Alarm Rate (FAR):

The FAR is a measurement of the sum of pixels that were wrongly identified to the actual quantity of affirmative pixels.

$$FAR = \frac{FP}{FP+TP} \qquad (28)$$

11. Similarity:

This produces a measure of similarity, which is an assessment of pixel similarities between both the separated and ground truth frames.

$$Similarity = \frac{TP}{TP+FP+FN} \qquad (29)$$

12. Specificity:

The fraction of frames that are precisely segmented is used to calculate specificity. i.e., the measurement of how precisely separation is performed for destructive outcomes.

$$Specificity = \frac{TN}{TN + FP} \times 100 \quad (30)$$

13. Accuracy:
Accuracy is defined as the weighted proportion of accurately segmented frames.

$$Accuracy = \frac{TP + TN}{TP + FP + TN + FN} \times 100 \quad (31)$$

14. *Dataset Comparability:*
Our proposed method outperforms existing techniques on CDnet-2014, CMD, PETS, and Hall monitor datasets, achieving high F1-scores, demonstrating its effectiveness in challenging scenarios.

15. *Visual Contrasts:*
Our approach visually compares with DMFE, SBS, and SuperBE proposals, using object silhouette separation to isolate moving objects from input sequences without noise, resulting in superior segmentation object silhouettes.

16. *Evaluation:*
The proposed detection and tracking algorithms for moving objects are evaluated using various criteria, with results assessing the effectiveness of the tracing system on PETS and Hall monitor datasets.

Datasets	Frames	Sp	FPR	FNR	PWC	Re	Pr	F1-score
PETS	1	0.9971	0.0082	0.0416	0.4129	0.9584	0.8981	0.9266
	2	0.9921	0.0073	0.0521	0.4086	0.9654	0.8896	0.9246
	3	0.9968	0.0081	0.0456	0.4099	0.9612	0.8923	0.9386
Hall monitor	6	0.9991	0.0009	0.0512	0.2234	0.9626	0.8964	0.9384
	8	0.9928	0.0146	0.0436	0.2462	0.9712	0.8862	0.9326
	7	0.9865	0.0135	0.0598	0.2564	0.9758	0.8892	0.9692
CMD [53]		0.9948	0.0052	0.0156	0.5774	0.9844	0.9128	0.9473
CDnet-2014[52]		0.9852	0.0148	0.1490	1.9364	0.8510	0.7246	0.7344

Table 1: The proposed method's performance was compared to various datasets, including CDnet-2014, CMD, PETS, and Hall monitor

Our technique on PETS and Hall monitor datasets produces high average ratings, with the suggested DMFE achieving over 92.77 percent F1-scores across all videos. The improved performance is due to adaptive fusion, which efficiently identifies edge information, reducing false positives and false negatives. The adaptive threshold is calculated based on the backdrop type, reducing erroneous detection, and background pixels are replaced to reduce false positives and increase false negatives.

XIII. Conclusion:

The proposed dynamic multiple feature fusion approach (DMFE) improves detection accuracy by combining colour/gray intensity and gradient magnitude. This enhances the recognition of object shapes and outlines, leading to more accurate detection. The approach also improves segmentation performance through adaptive threshold, background sample update, and quasi colour-gradient input. The technique achieves the highest accuracy on PETS Hall monitor videos, using DNN and GSA algorithms for object detection and tracking. Experimental results show the proposed approach outperforms the state of the art, achieving the highest accuracy and recall values for both PETS and Hall monitor videos.

mahalingam 3.pdf

ORIGINALITY REPORT

13%	8%	12%	1%
SIMILARITY INDEX	INTERNET SOURCES	PUBLICATIONS	STUDENT PAPERS

PRIMARY SOURCES

1	T. Mahalingam, M. Subramoniam. "Optimal object detection and tracking in occluded video using DNN and gravitational search algorithm", Soft Computing, 2020 Publication	5%
2	link.springer.com Internet Source	2%
3	T. Mahalingam, M. Subramoniam. "ACO–MKFCM: An Optimized Object Detection and Tracking Using DNN and Gravitational Search Algorithm", Wireless Personal Communications, 2019 Publication	2%
4	S Renuga, S S K Jagatheeshwari. "Efficient Privacy-Preserving Data Sanitization over Cloud Using Optimal GSA Algorithm", The Computer Journal, 2018 Publication	1%
5	Shahin Azali, Mansour Sheikhan. "Intelligent control of photovoltaic system using BPSO-GSA-optimized neural network and fuzzy-	<1%

	based PID for maximum power point tracking", Applied Intelligence, 2015 Publication	
6	eprints.utm.my Internet Source	<1%
7	vdoc.pub Internet Source	<1%
8	ijece.iaescore.com Internet Source	<1%
9	www.researchgate.net Internet Source	<1%
10	usir.salford.ac.uk Internet Source	<1%
11	tjee.tabrizu.ac.ir Internet Source	<1%
12	Mario I. Chacon-Murguia, Huber E. Orozco-Rodriguez, Juan A. Ramirez-Quintana. "Self-Adapting Fuzzy Model for Dynamic Object Detection Using RGB-D Information", IEEE Sensors Journal, 2017 Publication	<1%
13	mdpi-res.com Internet Source	<1%
14	Andrés Jerez, Mireille Lavagna, Damien Bensimon. " Strong-coupling fixed point	<1%

instability in a single-channel Kondo model ", Physical Review B, 2003
Publication

15	Submitted to Institute of Technology, Nirma University Student Paper	<1 %
16	www2.mdpi.com Internet Source	<1 %

Exclude quotes	Off	Exclude matches	Off
Exclude bibliography	Off		