IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
Impact Factor: 7.53
Volume 4, Issue 1, January 2024

The Development of Various Methods for Object Tracking and Classification in Thermal Videos

Shivani Kesharvani[1], Gourav Saxena[2], Kishore Kumar[3]
Research Scholar, Thermal Engineering[1]
Asst. Prof, ,Mechanical Engineering[2,3]
Nagaji Institution of Technology and Management, Gwalior, India

Abstract: *Nowadays, thermal cameras play a major role because of its temperature-based photography in many applications such as video surveillance, monitoring electronics/electrical machines, forest monitoring, monitoring babies/adult patients, and suspicious object detection. Tracking pedestrians in thermal video is a major task for such applications. Thermal cameras usually create images based on temperature emitted by the object only and not on the lighting conditions and outdoor environment conditions. But still thermal images have constraints like no texture or colour information, more number of dead pixels, low resolution, and noticeable visual colour patterns in case of any temperature variations. So the challenge in tracking pedestrians in thermal videos is tracking objects/pedestrians throughout the video without an identity switch by overcoming these constraints which may mislead the tracking process. To overcome these constraints, the proposed system uses tri feature matrix (TFM) as an object descriptor which is used to uniquely identify and represent objects in thermal images. TFM is represented in more compact way as a triple matrix. It is a simple and accurate descriptor suitable for tracking objects in thermal video sequences without an identity switch. The proposed Pedestrian tracking system uses most of the advantages of thermal cameras by overcoming challenges in thermal videos effectively based on a novel descriptor TFM. The proposed system is evaluated with various data sets, and the results are analysed using true positive, true negative, false negative, false positive, accuracy, precision, recall, F-score, global identity mismatch (GMME) and track matching error (TME). The performance metrics such as accuracy, precision, recall, F-score, GMME and TME are computed as 99%, 100%, 99%, 99%, 2.3%, and 2.1%, respectively. From the observation, it is found that the performance of proposed TFM-based system is significantly improved. The experimental result shows that the proposed system achieved more accurate tracking compared to the conventional methods.*

Keywords: Thermal camera, TFM, Convex hull, Convex deficiency, Fourier descriptor

I. INTRODUCTION

Thermal imaging is the process of recording reflected and transmitted thermal radiations in an image format. Thermal images depend on temperature differences of objects that belong to the scenes that are viewed. In an outdoor environment, there may be many sources for thermal radiations and hence thermal imaging becomes a challenging process. As the thermal imaging process is independent of lighting conditions, weather conditions do not affect thermal images. This is the major advantage of the thermal cameras over the conventional visual cameras. Hence, with these advantages, thermal cameras can be efficiently used for object tracking in video surveillance. Object tracking is the process of finding a full track of objects in a video sequence over time. Track of objects need to be defined as a sequence of object positions. The challenges of object tracking in thermal images are: lower resolution, low contrast images and more dead pixels. In thermal images, objects are separable only if there is a variation in temperature between object and its surroundings. Though, thermal imaging can picture a scene in darkness, still there is a challenge in thermal object tracking compared to visual object tracking.

To meet these challenges, the proposed system has the following features:

- Convex hull matrix (CHM) and convex deficiency matrix (CDM). These matrices separate objects which have less temperature variations from its background.

- Discrete Fourier transform—shape matrix (DFT-SM). This matrix takes low-frequency coefficients of an object as a shape descriptor, and hence, no dead pixels are con sidered.
- TFM gives a clear description of an object in low-contrast and low-resolution image. In addition, TFM represents the object as a whole and tracks the object throughout the video sequences without an identity switch.

II. RELATED WORK

Thermal cameras are currently utilized as a part of medicinal services [19] for remote surveillance and safety monitoring by identifying objects that are not visible to human eyes. So thermal cameras act as a strong security tool even in no light condition [3, 10, 12, 14]. A lot of research studies are going on with respect to object tracking in thermal videos. In general, there are two major classification approaches for object tracking in thermal videos. First one is based on artificial neural networks (ANN) such as convolutional neural network (CNN) [15, 20] and residual neural network (RNN). Second one is feature-based object tracking. In [16, 27], a tracker based on fully convolutional neural network (FCNN) and Siamese CNN was developed which used both object tracking model and a scoring model to track objects. Even though the tracker has tracked objects in challenging environments, due to lack of texture information in thermal images the advantage of CNN could not be fully utilized. And also it consumes more time and needs large volume of data to train the network. In [14], a deep neural network CNN is designed which can extract image low-level fea tures for image classification. A synthetic data generation model was developed [26] which required large volume of data to train CNN. The second type of object tracking is based on the features of the object which is used to track the object throughout the video sequences. Features are the small description of an object which describes the spatial or temporal properties of objects [1, 4, 18]. SeckinDemir [22] has used a part-based co-difference matrix as a feature to track the object. In which, the object space has been divided into small parts and it computes a feature for each part of an object. But still this method could not represent an object as a whole, because it uses a single feature to represent a whole object. Li. C et al. [11] have developed a greyscale object
tracking method. This method is based on the Bayesian filtering framework. In which, the feature vector is represented in terms of vector format by concatenating the representation coefficient of two modalities and then the similarity is found by using the sparse representation of the coefficients. The main limitation of this approach is that it needs colour information which is not available in thermal images. MalinBruntha et al. [13] have used FFT and DCT [6] as a feature extractor and investigated about particle swarm and ABC optimization algorithm for feature selection for the application of face recognition.

Kim et al. [8] have developed a method for multi-object tracking which has a single thermal sensor. The main chal lenge in object tracking is the problem of object overlap ping. To handle occlusion problems in a perfect manner, a topographic view of objects is used in their system in which convex and concave points were identified and used to track objects in a topographic view. But the system requires spe cial settings to capture images and is applicable for few types of applications only.

Zhang et al. [25] utilized the edge information to describe the shape of an object without colour information. But in thermal images as the clear edge information is not avail able, one could not describe an object as a whole.

Berg et al. [2] utilized a template-based scheme with the background update to evade tainting of background as far as the object template exists in the thermal video. Kwak et al. [9] have developed an online learning strategy which tracks pedestrian based on the temperature difference in the thermal image.

Talha [23] developed colour-based tracking using particle filtering approach. The key advancement techniques used are consistently relearning local background models for every particle in each imaging modality and its adaptive nature of object tracking. It is a colour-based tracking method. But the drawback is that as the colour information is not available in thermal images, it could not be applied to track objects in thermal images. Hao [5] tracked multiple humans using wireless distributed pyroelectric sensor systems in which an expectation–maximization-Bayesian tracking scheme was developed to track humans in all illuminating situations with an enclosed environment but which could not be used in outdoor conditions.

III. METHODOLOGY

In today‟s modern technology the thermal camera plays a major role in all security applications, surveillance applications etc. and also in pandemic situations the thermal camera plays a major role (Ward et al. 2021), because of its nature of picturizing the scene based on the temperature emitted by the objects in a scene. Hence regardless of any weather conditions and lighting conditions, an object can be visualized. So, in general, there is a huge demand for effective object classification system in thermal videos. The proposed system uses Dynamic STP and TTP based FCNN for object classification in thermal videos. There are two novel methods used in the proposed system. First one is Temperature histogram-based object detection and another one is Dynamic STP and TTP based FCNN for object classification. The nature of FCNN cannot be fully utilized in thermal videos due to the lack of texture information. The dynamic STP and TTP is designed for each object as kernel to extract the STP, TTP feature map of thermal videos. The fully connected layer with soft max classification is used for classification of object. The experimental result shows that the proposed system outperforms the state of art methods.

The aim of the proposed system is to achieve high accuracy in low processing time. In the proposed system, an object is detected using temperature histograms and the captured object is classified using Dynamic STP and TTP based FCNN.

1) Temperature histogram-based object detection that is based on nature of thermal images is the method which is proposed for detecting objects in an image. The detected object region is given as an input for FCNN which reduces processing time.

2) The Dynamic STP and TTP based FCNN classification requires three steps.

i) Dynamic STP and TTP Kernel Generation

The kernels for each object region are generated based on its spatial and temporal correlation of dominant temperature of the object region, so that makes it more accurate.

ii) Novel STP and TTP extraction

The feature STP and TTP extracted from the object region using two layers of convolution operation

iii) Classification

Fully connected layer with soft max classification function used to assign the label for object.

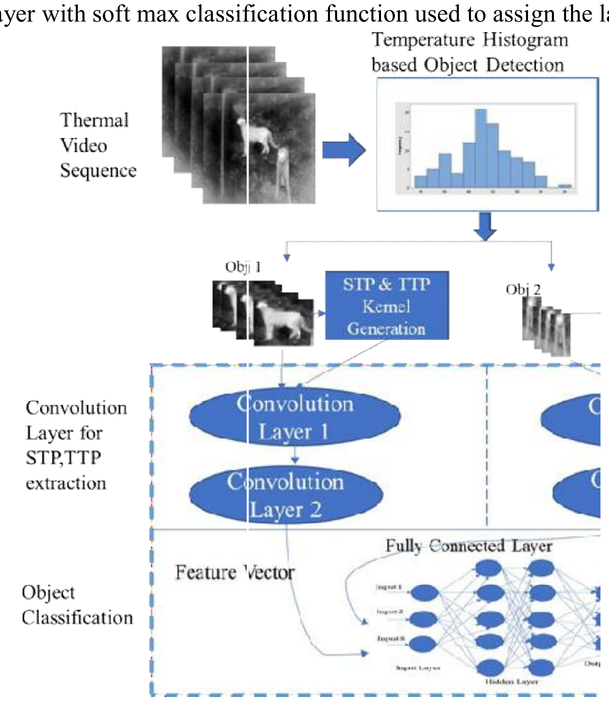

Figure 1 Basic Architecture of STP and TTP based object classification

IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
Impact Factor: 7.53
Volume 4, Issue 1, January 2024

3.1 Proposed System

The proposed system uses STP and TTP based FCNN for classifying objects which are already detected from thermal image. In addition, it uses Temperature-histogram based object detection which is more efficient and it is a perfect method for differentiating foreground from background in thermal images.

3.1.1 Temperature-Histogram based object recognition

Temperature is the key point in a thermal image based on which a thermal camera captures the scene by making use of heat transmitted and observed by an object. In general, any living object will have a higher body temperature than background. Other than any living objects, any moving objects(vehicles) will have a high temperature due to its internal kinetic energy conversion. Below is the proposed algorithm for Temperature histogram based object detection.

3.1.2 Dynamic STP and TTP based FCNN

The proposed system uses Dynamic STP and TTP based FCNN for object classification which calculates the dynamic kernel which is based on spatial, temporal relationship of pixel having mean temperature. The dynamic kernels are generated for each object.

The pixels having mean temperature of an object is taken for generating STP and TTP. Figure 3.2 shows the pixels with mean temperature and its neighbourhoods. All the neighbor pixel temperature values of mean pixel are subtracted from mean temperature, then if the result is greater than 0 it is taken as 1, otherwise it is taken as 0. The center pixel is assigned with mean temperature. The 3x3 dynamic STP kernel is generated by following the above procedure.

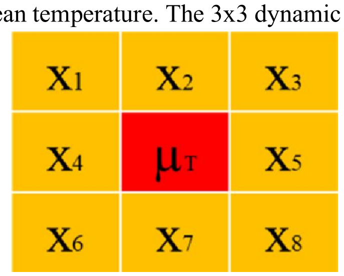

Figure 2 Procedure of STP kernel Generation

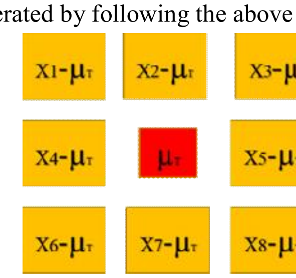

Figure 3 Procedure of STP kernel

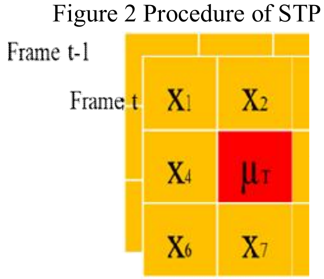

Figure 4 Pixel with mean temperature and corresponding neighbours in previous frame

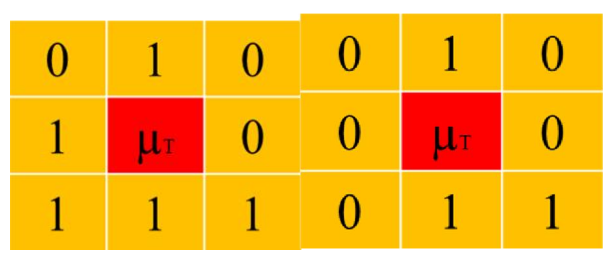

Figure 5 shows the sample STP and TTP kernel.

3.2.Object classification using dynamic STP and TTP based FCNN

In the proposed system, the object pixels are grouped and labeled as object1, object 2. Then the part of image with the appropriate object with corresponding generated STP and TTP kernel is given to input for CNN.

The input object may differ in size, so it is taken as m×n . Then there are two convolution layers with an average pooling layer is used. There are two kernels as STP and TTP generated in the size of 3×3 is applied on m×n image in each convolution layer. Then the 2×2 average pooling layer is applied after convolution operation in each layer. So, after convolution operation and pooling in two layers the image is reduced in terms of $m/4 \times n/4$.

Then it is converted into one dimensional feature vector and given to FCN (Fully Connected Layer) in which classification is performed.

The extracted feature vector defines the STP and TTP of an object.

3.3 Training dynamic STP and TTP based FCNN

Supervised training is used to train the network. In each iteration, the image with a reference object is converted to a temperature matrix and the corresponding matrix is given to convolution layer. The extracted feature vector is given to FCN.

In FCN, there is a single input layer and output layer and two hidden layers. Each bit of single dimensional feature vector is given to single neuron and it is given to each hidden layer neuron through the corresponding weighted path.

In each hidden neuron, the activation is calculated based on the summation of received weighted inputs and activation function

3.4 Testing STP and TTP based FCNN

After so many iterations of training, the network converges to a trained state. After that, the testing process started. In testing the thermal image with referenced objects is given to classification and the performance of proposed system is evaluated.

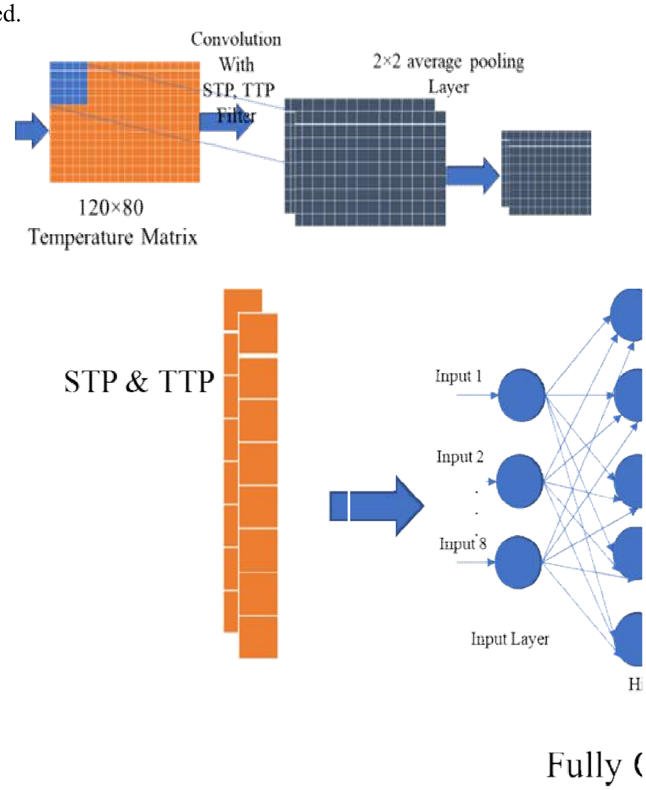

Figure 6 Dynamic STP and TTP based FCNN

The Figure.6 clearly explained about the working procedure of STP and TTP based FCNN.In testing phase also, the image with object is two times convoluted with both STP and TTP filters. The FCN gives the class of the object as output. As all these processes are made on the temperature matrix, the calculation takes less time and complexity.

IV. RESULT AND DISCUSSION

4.1 Performance Evaluation

4.1.1 Experimental results

The proposed system is implemented in Matlab R2020a and evaluated with publicly available data sets of Thermal Cheetah v1 square, Thermal Dogs and People, OTCBVS of that have thermal images with objects like humans, vehicles and animals. For simplicity, the proposed system takes these three classes as output classes.

IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
Impact Factor: 7.53
Volume 4, Issue 1, January 2024

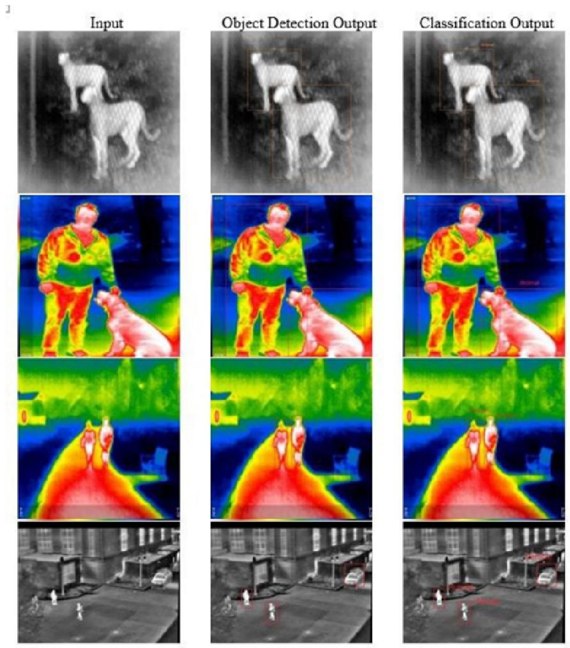

Figure 7 Sample output of STP and TTP based object classification

Figure 7 shows the sample output of STP and TTP based object classification. The data set information is given in Table 1.

Table 1 Dataset Information

Data Set Information			
TrainingSet		Testing Set	
Human	220	Human	180
Animal	280	Animal	225
Vehicle	252	Vehicle	188
Total	752	Total	593

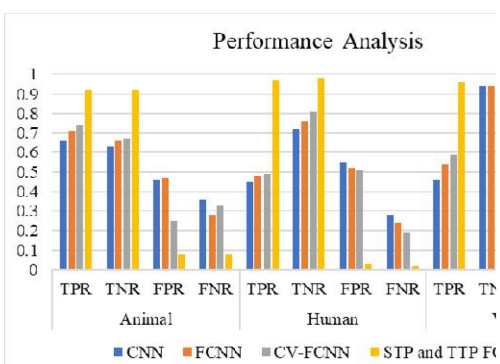

Figure 8: Performance Analysis of CNN (Rodin et al. 2018), FCNN (Pan et al. 2007), CV-FCNN (Yu et al. 2020), proposed STP and TTP-FCNN

The performance analysis of various state of art methods are shown in Figure 3.9. The analysis clearly shows the efficiency of the proposed system. It gives high TPR and TNR and low FPR and FNR for all three classes. In the state of art methods, there is confusion in differentiating human and animal, but the proposed system uses STP, TTP which is invariant to geometric transitions can efficiently differentiate both classes.

IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
Impact Factor: 7.53
Volume 4, Issue 1, January 2024

Conventional architectures like CNN, FCNN are sometimes failed to detect vehicles as foreground objects. But the Temperature-histogram based object detection can efficiently detect all the foreground objects even in bad weather conditions and lighting conditions.

4.1.2 Qualitative evaluation

The accuracy, precision and recall of the system is calculated using the formula

$$Accuracy = \quad TPR+TNR/TPR+TNR+FPR+FNR \tag{1}$$

$$Precision= \ R/TPR+FPR \tag{2}$$

$$Recall= \ R/TPR+FNR \tag{3}$$

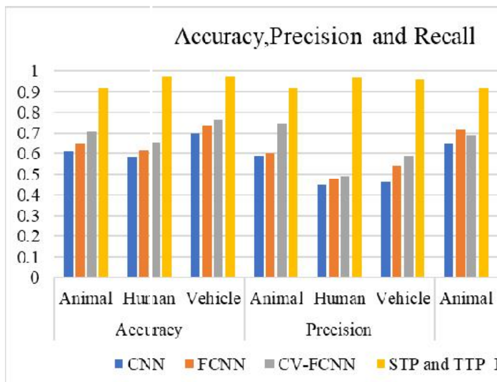

Figure 9 Performance Analysis of CNN (Rodin et al. 2018), FCNN (Pan et al. 2007), CV-FCNN (Yu et al. 2020), proposed STP and TTP-FCNN

The performance analysis shows that the proposed system has high accuracy, precision and recall than the state of art methods.

V. CONCLUSION

The STP and TTP based FCNN are a modified and novel FCNN architecture which can detect multiple objects in a thermal image in less time complexity. Compared to conventional methods like CNN (Rodin et al. 2018), FCNN (Pan et al. 2007), CV-FCNN (Yu et al. 2020), the proposed architecture is different, simple and effectively utilize the nature of thermal image. The proposed system uses temperature-based object detection which is very simple and effective, which does not use any complex equations. This method uses a temperature histogram to differentiate foreground from the background, it does not use any previous frame information for object detection so it will be working fine in the challenging environment of dynamic background.

The pixels in the particular temperature range are grouped based on 8 neighborhood connection component analysis, so it can separate the different objects of same class. If two different class objects are occluded partially means using temperature difference it can be differentiated. So, this method overcomes the partial occlusion problem. The detected objects with boundary box are given to the proposed CNN for classification. So, the problem is reduced and only region of interest is given to the proposed FCNN for processing and the iteration time is reduced and simplified architecture gives an effective result of classification compared to traditional architectures.

The proposed system uses the nature of thermal images for object detection and classification. So, it is possible to detect and classify the object in bad weather conditions and improper lighting conditions by overcoming limitations of thermal images such as low contrast, a lot of dead pixels, not having clear edges for separating foreground and background. The temperature-histogram is used for object detection which is a very simple and efficient method for thermal images. The proposed STP and TTP FCNN classify the detected objects as any one of defined class using the proposed architecture. The proposed architecture is very simple and can-do efficient classification on detected objects compared to the state of art methods.

IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
Impact Factor: 7.53
Volume 4, Issue 1, January 2024

REFERENCES

[1]. Abhishikt K, Andrew J, Martin Sagayam K, Hien DT (2020) Vehi- cle automation and car-following models for accident avoidance. PRZEGLAD ELEKTROTECHNICZNY, pp 118–123

[2]. Berg A, Ahlberg J, Felsberg M (2016) Channel coded distribution field tracking for thermal infrared imagery. In: Proceedings of IEEE conference on computer vision, pattern recognition. Work- shops, pp 1248–1256

[3]. Gade R, Moeslund T (2014) Thermal cameras and applications: a survey. Mach Vis Appl. 25:245–262

[4]. Gundogdu E, Koc A, Solmaz B, Hammoud R, Alatan A (2016) Evaluation of feature channels for correlation-filter-based visual object tracking in infrared spectrum. In: Proceedings of IEEE conference on computer vision, pattern recognition workshops, pp 290–298

[5]. Hao Q, Hu F, Xiao Y (2009) Multiple human tracking and iden- tification with wireless distributed pyroelectric sensor systems. IEEE Syst J 3:428–439

[6]. Hein Dang K, Martin Sagayam P, Malin Bruntha S, Dhanasekar A, Amir Anton J, Rajesh G (2019) Image fusion based on sparse sampling method and hybrid discrete cosine transformation. Int J Sci Technol Res 8(12):1103–1107

[7]. IEEE OTCBVS WS Series Bench; Davis J, Keck M (2005) A two-stage approach to person detection in thermal imagery. In: Proceedings of workshop on applications of computer vision

[8]. Kim W, Beom CY, Lee S (2017) Thermal sensor-based multiple object tracking for intelligent livestock breeding. IEEE Access 5:27453–27463

[9]. Kwak J, Ko BC, Nam J (2017) Pedestrian tracking using online boosted random ferns learning in far-infrared imagery for safe driving at night. IEEE Trans Intell Transp Syst 18:69–81

[10]. Lee JH (2015) Robust pedestrian detection by combining visible and thermal infrared cameras. Sensors 15:10580–10615

[11]. Li C, Sun X, Wang X, Zhang L, Tang J (2017) Grayscale-thermal object tracking via multitask Laplacian sparse representation. Trans Syst Man Cybern Syst 47:673–681

[12]. Ma Y, Wu X, Yu G, Xu Y, Wang Y (2016) Pedestrian detection and tracking from low-resolution unmanned. Sensors 16:446

[13]. Malin Bruntha P, Dhanasekar S, Martin Sagayam K, Immanuel Alex Pandian S (2019) A modified approach for face recognition using PSO and ABC optimization. Int J Innov Technol Explor Eng 8:1571–1577

[14]. Martin Sagayam K, Jude Hemanth D (2019) A probabilistic model for state sequence analysis in hidden Markov model for hand gesture recognition. Comput Intell 35:51–81

[15]. Mhathesh TSR, Andrew J, Martin Sagayam K, Henesey L (2021) A 3D convolutional neural network for bacterial image classification. In: Peter J, Fernandes S, Alavi A (eds) Intelligence in big data technologies—beyond the hype, Advances in intelligent systems and computing, vol 1167

[16]. Mohd AZ, Nikki T (2018) Multiple model fully convolutional neural networks for single object tracking on thermal infrared video. IEEE Access 6:42790–42799

[17]. Naftel A, Khalid S (2006) Motion trajectory learning in the DFT- coefficient feature space. In: Fourth IEEE international conference on computer vision systems (ICVS'06), p 47

[18]. Porikli F, Tuzel O, Meer P (2006) Covariance tracking using model update based on lie algebra. In: Computer vision and pattern recognition, IEEE computer society conference, pp 728–735

[19]. Rajesh G, Raajini XM, Martin Sagayam K, Dang H (2020) A Sta- tistical approach for high order epistasis interaction detection for prediction of diabetic macular edema. Inf Med Unlocked 29:1–9

[20]. Sathish N, Deepak M, Martin Sagayam K, Narain Ponraj D, Ajay Vasanth X, Lawrence H, Chung CH Image fusion in remote sensing based on sparse sampling method and PCNN techniques. DE GRUYTER: Machine Learning for Big Data Analytics, Frontiers in Computational Intelligence

[21]. Seckin Demir H, Cetin AE (2016) Co-difference based object tracking algorithm for infrared videos. In: IEEE international conference on image processing (ICIP), pp 434–438

[22]. Seckin Demir H, Faruk Adil O (2018) Part-based co-differ- ence object tracking algorithm for infrared videos. In: 25 th IEEE international conference on image processing (ICIP), pp 3723–3727

IJARSCT

ISSN (Online) 2581-9429

International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)

International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal

Impact Factor: 7.53

Volume 4, Issue 1, January 2024

[23]. Talha M, Stolkin Rustam (2014) Particle filter tracking of cam- ouflaged targets by adaptive fusion of thermal and visible spectra camera data. IEEE Sens J 14:159–166

[24]. Vinaykumar M, Jatoth RK (2014) Performance evaluation of Alpha-Beta and Kalman filter for object tracking. In: IEEE international conference on advanced communications, control and computing technologies, pp 1369–1373

[25]. Zhang T, Wiliem A, Hemson G, Lovell B (2015) Detecting kan- garoos in the wild: the first step towards automated animal surveillance. In: Proceedings of IEEE international conference on acoustics, speech, and signal processing, pp 1961–1965

[26]. Zhang L, Gonzalez-Garcia A, van de Weijer J, Danelljan M, Khan FS (2019) Synthetic data generation for end-to-end thermal infrared tracking. IEEE Trans Image Processing 28:1837–1850

[27]. Zulkifley MA (2019) Two streams multiple-model object tracker for thermal infrared video. IEEE Access 7:32383–32392.