ORIGINAL PAPER

Thermal image processing-based intelligent technique for object detection

Sayantan Nath[1] · C. Mala[1]

Abstract

The Thermal Image Processing Technique (TIPT) is the most prominent tool for observing an object at night. It is vastly used in many domains like security, healthcare, process control, and surveillance especially in defense vehicles where visualization at night would also mandatory for checking. In this paper, a thermal imaging camera is proposed to only be felicitated in regular as well as automated vehicles for better identification of objects especially at night when visibility is very less. Due to the huge variation in grayscales and pseudo-coloring values in the thermal image, a fuzzy-based CNN [FCNN] model is proposed to be applied to identify the boundaries of the objects. In this technique, the correlation between the thermal images of the moving object and its types is proposed to be trained with the novel FCNN model. The framed methodology is not only implementable in benched marked video datasets but also applicable in real-life conditions on the ground experimental scenario on a live feed video streaming. The results significantly indicate the enhancement of visual capacity in the TIPT compared to normal visual technique.

1 Introduction

In today's fast-developing world, communication plays the most crucial role in our day-to-day lives and transportation [1] is one of the vital among them. In this round-the-clock working lifespan, transportation is required to being run throughout the night [2] also. However, the exponentially increasing road accidents, especially in India, are the main concern of road safety [3] today, and at night, the rate of rush driving is spiked. In this condition, automated vehicles are extremely important to control traffic and accident rates [4]. Though the visual camera in the object detection could be able to acquire the objects in front of the car in a day, it is completely useless at night [5]. The headlights of the car also have limited perceptibility [6] up to a few meters only. So, it is instantly essential to implement thermal imaging cameras in object detection [7] for better identification of objects at night.

A thermographic video camera [8] is proposed to be implemented as the data acquisition tool to recognize the objects in front of the car. This technology would be used to train [9] for four types of objects: pedestrian, vehicles, two-wheeler, and cattle. The basic data generation process is proposed to be modeled by the training [10] of a deep-learning-based identification technique. The modeled data and the trained process will be executed in the real-life scenario [11] on-road and capture live thermographic video of an object moving in front of the car. From the trained model [12], the objects would be classified into four groups as indicated above. The basic detection system would be trained to recognize the objects moving in front of the car. Due to the huge variation [13] in grayscale and pseudo-coloring values in thermal images, a fuzzy-based edge detection process is planned to be applied for distinguishing the boundaries [14] of the objects.

In this paper, the correlation between the thermal signature [15] of the moving object and its type is planned to be trained with the time-aspect ratio. The data collected [16] from the different points of implementation are also ideating to compare with the driving experts. By this linking technique, a real-time [TIPT] [17]-based object identifica-

✉ C. Mala
mala@nitt.edu

Sayantan Nath
sayantan@nitt.edu

[1] Post-Doctoral Fellow, National Institute of Technology,
    Tiruchirappalli, India

tion model would be designed. The produced results would also be compared with the road safety transportation [18] board for further analysis. In this modern working lifespan, transportation is required to being run throughout the night also. However, the exponential [19] increase in road accidents especially in India is the main concern of road safety today [20], and during the night, the rate of rash driving gets spiked. In this condition, automated vehicles are extremely important to control traffic and accident rates.

Thermography, thermal imaging, and thermal video work on the basic principle of infrared radiation. Thermal cameras usually detect radiation in the long-infrared range [21] of the electromagnetic spectrum (roughly from 9 to 14 μm) and produce images of that radiation signature, called thermogram [22]. Since infrared radiation is emitted by all objects with any temperature absolute zero (− 273 °C) according to the black body radiation law [23], thermography makes it possible to observe any object beyond visible illumination. The object detection is only taken as the applicable area in this paper. To measure the temperature patterns of an object using an infrared imager [24], it is necessary to estimate or determine the object's emissivity. To get a more accurate temperature measurement, a thermographer may apply high emissivity to the surface of the object. It shows a visual picture so temperatures [25] over a large area can be compared. It is capable of catching moving targets in real-time scenarios. It is used to measure or observe in areas inaccessible or visually blind for other normal methods. The basic working methodology for the proposed thermographic image processing-based object detection [26] is framed into four basic stages.

The learning process of the proposed model to correlate between different levels of temperature patterns and structure [27] of the specific objects would be developed further in the steps. The foundation of the proposed model proposed is analyzed to an abstractive numerical technique for the basic structural integrity [28]. The connectivity among the Fuzzy Logic and Convolution Neural Network [29] is designed to produce the core model. The classification of the obtained items in the thermal video would be integrated [30] within the object-frame by the "Jaccard Index" method. The fuzzy-based convolution neural network (F-CNN) method [31] was described to predict the traffic flow which was applicable for the dataset only. This method is completely incompatible in image processing where numerical datasets are not the primary investigative constraints. The factor [31] of temperature was not very effective for the block size of 32×32 indicated in this paper since it may vary more frequently in real-life scenarios of thermal images on road.

A deep learning approach targeting an object to track and classify into its respective category without reconstruction of any frame was investigated. The approach had two parts basically: tracking and classification [32, 34, 35]. The tracking had been conducted by using YOLO technique, and the classification was done by using the Residual Network proposed as [ResNet]. The experiments using mid-wave and long-wave infrared videos had demonstrated the efficacy of a high-performing approach to track and classify directly the object in their respective domains. By skipping the time-reconstruction stage to allow performing real-time tracking and classification based on combination of: YOLO and ResNet, have been innovated certainly.

An efficient way to enhance detection of small target in long-range and low-quality infrared videos by unsupervised, modular and flexible methods was investigated. Though the indicated approach was suitable where training data were limited, the inter-connection between temperature and the tracking object are not discussed. The experimental video using low and medium infrared clearly demonstrated the efficiency but the inter-correlation of the technique especially with the thermal image was not investigated. Though their comparative approach between ResNet and YOLO might find the better results, the inter-dependency between functionality of the object frame and temperature value of the object was not found yet. Tracking and classification of object especially for compressive measurement using pixel in the video frames were investigated by deep learning approach via integrated with YOLO and ResNet. The potential development of this model to integrate the demonstrated approach with real-time tracking and classification directly had missing the integration of temperature with the methodology.

An unsupervised, modular and flexible method to detect small objects in long and low-quality infrared videos using motion information extracted from optical flow methods [33] had been investigated. The optical flow methods combined with contrast enhancement and component analysis were found effective for target detection. Though the experiments conducted on long- and mid-wave infrared video dataset obtained from DSIAC clearly demonstrated the efficacy, the temperature relation with the detection technique was not evaluated. It clearly demonstrated that the proposed approach under different conditions especially in deep learning-based approach was more accurate but the correlation between temperature and detection model was missing.

The technique of vehicle detection and classification at the presence of human targets were investigated by pixel-wise code [36] exposure (PCE) camera. The combination between two deep learning algorithms was used for detection in mid-wave infrared (MWIR) videos obtained SENSIAC. Though the experimental result showed that the framework was capable for target detection up to 1500 m, the temperature factor in this scenario was completely ignored.

A combined deep learning approach: YOLO and ResNet were used to achieve for realistic optical and MWIR videos. Though the approach was modular and capable to detect multiple targets simultaneously, however the target up to 500 m

for small human was not effective. Because of this limitation, the temperature factor was completely overlooked.

2 Proposed working methodology

The basic working methodology for the [TIPT]-based object detection is illustrated in this section. At the initial stage, the abstractive structure of the proposed fuzzy-based model is designed and mapped with the convolution neural network (CNN). The mapped technique is compared through the "Intersection over Union" (IoU) or the Jaccard Index [37] to frame out the object in the image. The error value is also measured and used for self-learning for the network. Then, the algorithmic design of the entire proposed system is coined sequentially and designed through the Unified Modeling Language (UML). Four basic property diagrams (component, sequence, use case, activity) are demonstrated to elaborate the operational structure of the proposed model.

The simplified procedures for the proposed model are tentatively structured into multiple stages described as sequentially:

- Development of a fuzzy-based learning model: The correlation between different levels of temperature and structure of the objects is defined in this stage. The CNN-based learning model is merged with the fuzzy set to incorporate the temperature-based object detection technique. All the mathematical development of the coined methodology is fundamentally designed. The sub-groups of the complete procedure is shown in following steps:

  - *Abstract Foundation-Based Modeling*: The mathematical foundation of the proposed model is analyzed according to abstractive technique where the relation between the object detection function (g) and classification technique function (f) is merged together through convolution.
  - *The relationship between Fuzzy and Convolution Neural Network*: The inter-connectivity among the Fuzzy Logic and the Convolution functions is mapped through the ANN.
  - *Functional Integration:* After completion of classification and detection in thermal image, the recognized object is outlined within a frame by the "Jaccard Index".
  - *Complete structure for the FCNN model*: The complete architectural design of the proposed model is completed through the following steps.

- Error calculation technique: The error value of the object detection method by the camera is measured for further back-propagation to the neural network.

- Algorithmic designing: The complete algorithmic procedures of the proposed system are molded into sequential structure.
- System modeling: The connective architecture of the proposed model is designed by UML in multiple diagrams indicated below.

  - *Component diagram:* The connectivity among different functioning components of the proposed model is shown in this diagram.
  - *Use case diagram:* The direct relationship between the objects and detection system of the thermal images is shown in this diagram.
  - *Sequence diagram:* The stage-by-stage processes from beginning of acquisition of thermal image to framing of object in image are shown in this diagram sequentially.
  - *Activity diagram:* The series of activity of the proposed model concerning different inputs during the overall operation is shown in this diagram.

The comparison among the planned methodology and purpose of analysis for the proposed technique is shown in Supplementary Table 1.

The proposed technique is designed with the help of an abstract algebraic method to inter-connect the working functions into a single formation. The basic connections are separated into four stages: Target Object → (Thermal Camera → Proposed Image Processing Model) → Final Result. The main research focus of this paper is to produce a more robust processed thermal image of objects for object detection only. The respective diagram of the fundamental design of the proposed [TIPT]-based intelligent object detection model is shown in Supplementary Fig. 1.

2.1 Complete structure of the FCNN model.

The overall procedure of the proposed *TIPT* for object detection is designed on back-propagation-based supervised learning technique. The equations drawn from the mathematics of the back-propagated neural network are modified according to the proposed fuzzy-based convoluted neural network model. The step-by-step algorithmic structure of the proposed system defined according to the multi-layered back-propagation neural network (BPN) model [38]. The steps of the BPN enhanced with the fuzzy convolution model are being begun with the obtaining of initial weight and complete at the ending. The respective phases are denoted as the main executing module of the proposed technique:

Main ().

[

– The initial weight to the BPN network and its basic learning rate would be defined according to the temperature value of the thermal image;

– The loop of While (epoch $==h$ ∥ output $==$ targeted result, t_r) [∀ h = number of epochs to reach t_r] would be executed until the desired results would be obtained.
[

– Now, in the inner layer of the Receiving input $=x_i$ in the BPN network and the weight is sent from the initial layer to the hidden layer unit [∀ 1 to n ∣ n = total no. of input units] at the next level.

So, the total input measured at the single jth unit at hidden layer appeared from the previous layer with the bias (b_{0j}) value is calculated according to the feed forward process of the BPN network shown in Eq. (1):
where the parameters in Eq. (1) are defined as:

• b_{0j} = propagating bias value to jth no. of unit of the hidden layer.
• v_{ij} = weight at the 'j' number of units of the inputs appeared from 'i' number of units at the input layer.

• $j=1$ to p where p = total number of units present at hidden layers in BPN.
• $\tilde{A}$ = total no. of synaptic of individual unit (i) in the initial layers according to the fuzzy summation.

– So, the net output from the jth no. of unit of the hidden layers, $Q_j = R(Q_{in_j})$ where Q_{in_j} = total input at the jth no. of unit of the hidden layers and $R(a)$ = activation function for arbitrary variable 'a'.
– Now, total value transmitted in the hidden layers is being sent to the output layers of the BPN.
– So, the total input (Y_{in_k}) measured at the single kth unit at outer layer (Y) appeared from the previous hidden layer with the bias (b_{1_k}) value is calculated according to the BPN network model shown in Eq. (2):

where the parameters in Eq. (2) are defined as:

• b_{1k} = propagating bias value to kth no. of unit of the outer layer
• w_{jk} = weight at the 'kth' number of units of the inputs appeared from 'jth' number of units at the hidden layer
• $k=1$ to m where m = total number of units present at outer layers in BPN network

– So, the net output from the kth no. of unit of the outer layers, $Y_k = S(Y_{in_k})$ where Y_k = total input at the kth no. of unit of the outer layers and $S(b)$ = activation function for arbitrary variable 'b'.

– Now, the calculation of the error values in the BPN network according to the proposed FCNN structure is indicated:
– The measurement of **Error** (φ_k) for the input values obtained from hidden layers to the output layers is calculated according to error propagation process of the BPN network, as shown in Eq. (3) where the parameters in Eq. (3) are defined as:

• $\Delta\varphi_k$ = the amount of external error measured at the 'k' no. of unit of the output layer backtracked towards the hidden layer.
• t_r = error-correcting-term for the BPN network.
• Differential factor of activation function, $S(b)$ with respect to temperature, T in thermal image.

– Now, the change in the weights (w_{jk}) of the neurons transmitting from the hidden layers to the output layer is calculated according to the updating weight process of the BPN network with the learning rate, β. The respective mathematical orientation is shown in Eq. (4):

where the parameters in Eq. (4) are defined as:

– Changes in weights (w_{jk}) from the 'j' no. of units in hidden layers to the kth unit of the output layer.
– β = learning rate between those layers.
– Y_{in_k} = total input measured at the single kth unit at outer layer of the BPN network.
– So, the changes in the values of the bias (b_{1_k}) of the input weights to the kth unit of the output layers from hidden layers with the learning rate, β, are calculated according to the updation process of the BPN network, as shown in Eq. (5):

where the parameters in Eq. (5) are defined as:

• Δb_{1_k} = changes in the values of the bias (b_{1_k}) of the input weights to the kth unit of the output layers.
• Y_{in_k} = total input at the kth no. of unit of the outer layers.

– Therefore, the changed weights (w_{jk}) (new) of the neurons transmitting from the hidden layers to the output layer are modified according to the weight updation process of the BPN network as shown in Eq. (6):
– So, the changed bias (b_{1_k}) (new) of the weights to the kth unit of the output layers from hidden layers with the learning rate, β, is calculated according to the updation process of the BPN network, as shown in Eq. (7):
– Now, the changes in learning rate, γ from β to the kth unit of the output layers from hidden layers is modified according to the updation process of the BPN network shown in Eq. (8):

– Then, the changes in the **Error** (δ_j) for the input values obtained from input layers to the hidden layers are calculated according to error propagation process of the BPN network, as shown in Eq. (9):

where the parameters in Eq. (9) are defined as:

- $\Delta\delta_j$ = the amount of internal error measured at the 'j' no. of unit of the hidden layer backtracked towards the input layer.
- $S'(t_r - Y_k)$ = first order of differentiation of the activation function $S(Y_{in_k})$ with respect to temperature, T from the thermal image.

– So, the updated new **Error** (δ_j) values for the FCNN back-propagated model, shown in Eq. (10), to the jth unit of hidden layers:
– Therefore, the new changes in the weights (v_{jk}) of the neurons transmitting from the input layers to the hidden layer are calculated according to the updating weight process of the BPN network with the learning rate, γ shown in Eq. (11): where the parameters in Eq. (11) are defined as:
– Δv_{i_j} = changes in weights (v_{i_j}) from the 'i' no. of units in input layers to the jth unit of the hidden layer.
– γ = updated learning rate from (β) between those layers.
– Then, the newly changed weights (v_{i_j}) of the neurons transmitting from the input layers to the hidden layer are modified according to the weight updation process of the BPN network as shown in Eq. (12):
– So, the changed bias (b_{0_j}) of the input weights to the jth unit of the hidden layers from input layers (0) with the changed learning rate, γ is calculated according to updation of the BPN network, as shown in Eq. (13):
– Now, the changed new bias (b_{0j}) of the input values, shown in Eq. (14), for the jth unit of the hidden layers:
– Then, the changed new learning rate (λ) of the input values, shown in Eq. (15), for the jth unit of the hidden layers:
– Therefore, the newly calculated value of the "Jaccard Index, $J_T(F_F, \Psi)$, shown in Eq. (16), on the measurable parameter temperature (T) is indicated as:

]]

2.2 Error calculation

After processing the FCNN-based proposed TIPT, the calculation method for the error in the back-propagation model is measured by basic square's sum. The basic error calculation, E, of the BPN network from the error-correcting-term t_r is defined in Eq. (17):

where the differential parameters in Eq. (16) defined below are designed according to the chain rule of the partial differentiation to make inter-connection between proposed FCNN model and temperature-based processing technique as shown in Supplementary Fig. 2.

- δ = learning rate for the input values to the unit before output layers obtained from hidden layers
- Ω_j = cumulative form of input of the derivative factor of FCNN concerning Temperature (T)
- Other symbols are already identified in the 'FCNN model structure' earlier

2.3 System modelling

The systematic modeling of the coined mechanism is computed in the UML technique. The UML diagrams of the proposed [TIPT]-based object detection in the object detection are illustrated in different following images. The components diagram of the proposed method indicates the connectivity among different modules functioning in the proposed thermographic system. The component diagram of the proposed system is indicated in Supplementary Fig. 3.

In the component diagram shown in Supplementary Fig. 3, the main parts of the model are the thermal camera unit (TCU), temperature distinguishing unit (TDU), [TIPT] unit (TIPU), etc. The TCU and TIPU are directly connected with the temperature – color relation knot and TDU and temperature level segmentation (TLS) are also connected with the knot proportionally. The observer unit is only linked with the TLS for requiring information. The sub-components of the TIPU, e.g., object tracking, detection, framing, classification, and recognition, etc., are the basic processing functions of the model. The activity diagram of the proposed thermographic system indicates the series of activity and their order of execution concerning different inputs during the whole operation. The activity diagram of the proposed system is indicated in Supplementary Fig. 4.

In the activity diagram shown in Supplementary Fig. 4, the procedure of the proposed model from the acquiring of thermal infrared radiation to the generation of the object detection is indicated step-by-step. In the beginning, the thermal vision is acquired from the infrared emission of the object. Then, different temperature levels on the thermal image are signified by various colors distinguishing technique. Then, the temperature-based segmentation process is applied to detect an object. Then, the object was classified with help of the proposed model and recognized. The communication diagram of the proposed technique indicates the connectivity between different stages of the process beginning from the acquiring of thermal video of the road to the recognition of various objects in the captured visual. The communication diagram of the proposed system is indicated in Supplementary Fig. 5.

In the communication diagram shown in Supplementary Fig. 5, the communication of the proposed model from the acquiring of thermal infrared radiation to the generation

of the object detection is indicated step-by-step. Then, the detected objects are framed and tracked. The use case diagram of the proposed method indicates the direct relationship between the object on-road and the automated vehicle system with a Thermographic imaging system. The use case diagram of the proposed system is indicated in Supplementary Fig. 6.

In the use case diagram shown in Supplementary Fig. 6, the basic use cases are the thermography, color-scale, and video capturing of the object. These use cases are at the preliminary level. Then, the use cases at the processing levels are object tracking, detection, classification, and recognition. The sequence diagram of the proposed method indicates the stage-by-stage processes beginning from the capturing thermal video of objects on road to the transmission of processed images of recognized objects in the video to the automated vehicle controlling system. The sequence diagram of the proposed system is indicated in Supplementary Fig. 7.

In the sequence diagram shown in Fig. 7, the sequence of the proposed model from the acquiring of thermal infrared radiation to the generation of the object detection is indicated step-by-step. In the beginning, the thermal vision is acquired from the infrared emission of the object. Then, different temperature levels on the thermal image are signified by various colors distinguishing technique. Then, the temperature-based segmentation process is applied to detect an object. Then, the object was classified with help of the proposed model and recognized. Then, the detected objects are framed and tracked.

3 Results and performance analysis

The proposed technique has been implemented on a few thermal images and videos of person, pedestrians, vehicles, and two-wheeler obtained from the internet: YouTube and RubTube. Due to the scarcity of standard colored thermal images and videos, there would be left no other option to obtained thermal images and videos except from online video library. Multiple packages from Python are utilized in this experiment: OpenCV, ImageAI, Keras, Numpy, etc. The proposed method is experimented in multiple thermal videos of human figure and roads to evaluate the significance.

3.1 Experiment thermal images

The analysis of thermal image—1 indicating a person in a closed room and its semantic segmented edges are shown in the pseudo-colored thermal Supplementary Fig. 8.

In the left part of Supplementary Fig. 8, a person in a pseudo-scaled thermographic image in a closed room is identified and in the right part, the same person in an edge segmented thermal image has been framed also. The analysis of thermal image—2 indicating several persons in an open space and its segmented edges are shown in the pseudo-colored thermal Supplementary Fig. 9.

In the left part of Supplementary Fig. 9, multiple persons in pseudo-scaled thermographic images in an open space are identified and in the right part, the same persons in an edge segmented thermal image have been framed also. The analysis of thermal image—3 indicating a vehicle and its segmented edges are shown in a pseudo-colored thermal Supplementary Fig. 10. In the left part of Supplementary Fig. 10, multiple persons in a grayscaled thermographic image in an open space scenario have been identified and in the right part, the same per-sons in an edge segmented thermal image are framed also. The analysis of thermal image—4 indicating several vehicles and their segmented edges are also shown in a pseudo-color in Supplementary Fig. 11.

In the left part of Supplementary Fig. 11, several vehicles on an open road scenario in a pseudo-thermographic image are identified and in the right part, the same vehicles in an edge segmented thermal image have been framed also. The analysis of thermal image—5 indicating riding of two-wheeler and its segmented boundaries are also shown in a pseudo-colored thermal Supplementary Fig. 12.

In the left part of Supplementary Fig. 12, a two-wheeler on an open road in a pseudo-thermal image is identified and in the right part, the same two-wheelers in an edge segmented thermal image have been framed also. The comparative study—1 of the thermal image and its respective normal visual images are shown in a gray-colored thermal figure for person detection purpose at night condition in Supplementary Fig. 13.

In the right part of Supplementary Fig. 13, several persons on an open road in the gray-colored thermal image are identified and frames. However, in the left part of the regular visual image, recognition of any persons has been completely failed. Comparative study of thermal image and its respective normal visual images are also shown in an edge segmented thermal image for the same person detection purpose at night condition in Supplementary Fig. 14. Here, the results are as same as the previous one.

The comparative study—2 of thermal image and its respective normal visual images are shown in a gray-colored thermal figure for detection of vehicles in a foggy condition as Supplementary Fig. 15.

In the right part of Supplementary Fig. 15, several vehicles on an open road in the gray-color thermal image are identified and frames clearly. However, in the left part of the regular visual image, recognition of any vehicles has been completely failed due to the foggy condition. The comparative study of thermal image and its respective normal visual images are also shown in an edge segmented thermal image for the same vehicles detection purpose in the foggy condition in Supplementary Fig. 16. Here, the results are as same as the previous one.

3.2 Analysis of the experimented images

The coefficients of the confusion matrix for the detection of an object by TIPT are indicated in SupplementaryTable 2. In process of calculation for the coefficient values of the confusion matrix, any mathematical non-dividable factors are assigned as = 0.

The confusion matrixes for thermal image—2 and its receptive graphical where the *y*-axis indicates the range of numeric value from 0 to 9 are indicated in Supplementary Table 3.

The respective graph of a Confusion matrix is shown in Supplementary Fig. 17 where the *y*-axis denotes the number of objects (person) observed in the thermal image–2.

InSupplementary Fig. 17, the blue-line indicates the variation of the number of objects (person) in the thermal image and the red-line indicates the variation of the number of objects (person) in the edge-segmented image. Though both blue and red lines are at their highest position at the TP point, they came and merged between TN and FP at zero value. The coefficient values related to the parameters of the confusion matrix for the thermal image—2 are calculated in Supplementary Table 4.

The graphical representation of the coefficients related to the parameters of the confusion matrix for the thermal image—2 is shown in Supplementary Fig. 18 in which the y-axis denotes the numeric values of the coefficients calculated in Supplementary Table 4 from the range 0 to 1.

As shown in Supplementary Fig. 18, the blue-line indicates the variation of the coefficient values calculated in Supplementary Table 4 for the thermal image, and the red-line indicates the variation of coefficient values calculated for the edge-segmented image. Both blue and red lines are always varying in-between the range of '0 to 1' where they overlap at some points. From the observation of the graph in Supplementary Fig. 18, it could be concluded that the blue line representing the thermal image has a higher average value than the segmented image shown in the red line. The confusion matrix for thermal image—4 and its receptive graphical where the y-axis indicates the range of numeric value from 0 to 17 are indicated in Supplementary Table 5.

The respective graph of a Confusion matrix is shown in Supplementary Fig. 19 where the y-axis denotes the number of objects (person) observed in the thermal image–4.

In Supplementary Fig. 19, the blue-line indicates the variation of the number of objects (person) in the thermal image and the red-line indicates the variation of the number of objects (person) in the edge-segmented image. Though both blue and red lines are at their highest position at the TP point, they came and merged between TN and FP at zero value. The coefficient values related to the parameters of the confusion matrix for the thermal image—4 are calculated in Supplementary Table 6.

The graphical representation of the coefficient values related to the parameters of the confusion matrix for thermal image—4 is shown in Supplementary Fig. 20 in which the y-axis denotes the numeric values of the coefficients calculated in Table 6 from the range 0 to 1.

As shown in Supplementary Fig. 20, the blue-line indicates the variation of the coefficient values calculated in Supplementary Table 6 for the thermal image, and the red-line indicates the variation of coefficient values calculated for the edge-segmented image. Both blue and red-lines are always varying in-between the range of '0 to 1' where they overlap at some points. From the observation of the graph in Supplementary Fig. 20, it could be concluded that the blue line representing the coefficients of the thermal image has a higher average value compared to the coefficients of the segmented image shown in the red line. The confusion matrix for the comparative study—1 and its receptive graph in which the y-axis denotes the number of objects (persons) observed in the image are indicated in Supplementary Table 7.

The respective graph of the Confusion matrix is shown in Supplementary Fig. 21 where the y-axis denotes the number of objects (persons) observed in the comparative study–1.

In Supplementary Fig. 21, the blue-line indicates the variation of the number of objects (vehicles) in the original thermal image and the red-line indicates the variation of the number of objects in the original visual image. And the green-line indicates the variation of several objects in the edge-segmented thermal image and the violate-line indicates the variation of several objects in the edge-segmented visual image. Though all the blue, red-yellow, and gray lines are at their highest position at the TP point, they came and merged at the point TN only at zero value. The coefficient values related to the parameters of the confusion matrix for the comparative study—1 are calculated in Supplementary Table 8.

The graphical representation of the coefficients related to the parameters of the confusion matrix for the comparative study—1 is shown in Supplementary Fig. 22, in which the y-axis denotes the numeric values of the coefficients calculated in Supplementary Table 8 from the range 0 to 1.

As shown in Supplementary Fig. 22, the blue-line indicates the variation of the coefficient values calculated in Supplementary Table 8 for the original thermal image, and the red-line indicates the variation of coefficient values calculated for the original visual image. And the green-line indicates the variation of the coefficient values calculated for the edge-segmented thermal image and the violate-line indicates the variation of coefficient values calculated for the edge segmented visual image. All the blue, yellow, gray, and red lines are always varying in-between the range of '0 to 1' where they overlap at some points. From the observation of the graph in Supplementary Fig. 22, it could be concluded that the yellow line representing the coefficients of the edge-segmentation of the visual image has spiked twice between

FNR and FOR. The confusion matrix for the comparative study—2 and its receptive graph in which the y-axis denotes the number of objects (vehicles) observed in the image are indicated in Supplementary Table 9.

The respective graph of the Confusion matrix is shown in Supplementary Fig. 23 where the y-axis denotes the number of objects (vehicles) observed in the comparative study—2.

In Supplementary Fig. 23, the blue-line indicates the variation of the number of objects (vehicles) in the original thermal image and the red-line indicates the variation of the number of objects in the original visual image. And the green-line indicates the variation of several objects in the edge-segmented thermal image and the violate-line indicates the variation of several objects in the edge-segmented visual image. Though all the blue, red, violate and green lines are at their highest position at the TP point, they came and merged at the point TN only at zero value. The coefficient values related to the parameters of the confusion matrix for the comparative study—2 are calculated in Supplementary Table 10.

The graphical representation of the coefficients related to the parameters of the confusion matrix for the comparative study—2 is shown in Fig. 24, in which the *y*-axis denotes the numeric values of the coefficients calculated in Supplementary Table 10 from the range 0 to 1.

As shown in Supplementary Fig. 24, the blue-line indicates the variation of the coefficient values calculated in Table 10 for the original thermal image, and the red-line indicates the variation of coefficient values calculated for the original visual image. And the green-line indicates the variation of the coefficient values calculated for the edge-segmented thermal image and the violate-line indicates the variation of coefficient values calculated for the edge segmented visual image. All the blue, yellow, gray, and red lines are always varying in-between the range of '0 to 1' where they overlap at some points. From the observation of the graph in Fig. 24, it could be concluded that the yellow line representing the coefficients of the edge-segmentation of the visual image has spiked twice between FNR and FOR.

3.3 Comparison of the performance of the result

The comparative performance between the confusion matrix of the thermal image—4 and 8 are shown in a tabular format indicated in Table 11.

The respective graph of the Confusion matrix is shown in Supplementary Fig. 25 where the y-axis denotes the number of objects observed in a comparative study between thermal image—2 and 4.

In Supplementary Fig. 25, the blue-line indicates the variation of the number of objects in the original thermal image and the red-line indicates the variation of the number of objects in the original visual image. And the green-line indicates the variation of several objects in the edge-segmented

thermal image and the violate-line indicates the variation of several objects in the edge-segmented visual image. Though all the blue, red-yellow, and gray lines are at their highest position at the TP point, they came and merged at the point TN only at zero value. The coefficient values related to the parameters of the confusion matrix for the comparative study between thermal image—2 and 4 are calculated in Supplementary Table 12.

The graphical representation of the coefficients related to the parameters of the coefficient for the comparative study—2 is shown in Supplementary Fig. 26, in which the *y*-axis denotes the numeric values of the coefficients calculated in Supplementary Table 12 from the range 0 to 1.

As shown in Supplementary Fig. 26, the blue-line indicates the variation of the coefficient values calculated in Supplementary Table 12 for the original thermal image, and the orange-line indicates the variation of coefficient values calculated for the original visual image. And the gray-line indicates the variation of the coefficient values calculated for the edge-segmented thermal image and the yellow-line indicates the variation of coefficient values calculated for the edge segmented visual image. All the blue, yellow, gray, and orange lines are always varying in-between the range of '0 to 1' where they overlap at some points. From the observation of the graph in Supplementary Fig. 24, it could be concluded that the yellow line representing the coefficients of the edge-segmentation of the visual image has spiked twice between FNR and FOR. Now, 3D mesh representation of the coefficient values related to the comparison between the confusion matrix of thermal image—2 and 4 is shown in Supplementary Fig. 27 where the *x*-axis indicates the rows, the *y*-axis indicates the columns and the z-axis indicates the range of the numeric value from 0 to 1:

4 Conclusion

From the development of the paper, it could be concluded that the proposed model of the thermal image progressing is better efficient to incorporate the fuzzy-based convolution neural network architecture with color thermal images than the normal visual images. The produced methodology could be then fed to the object detection for further assistant. From this thermal image processing, the visualization process would also be able to recognize objects in the night and face-up sunlight also which gives more reliability to the system about the detection process in any visual scenario.

References

1. Winter, J., Stein, M.A.: Computer image processing techniques for automated breast thermogram interpretation. Comput. Biomed. Res. **6**(6), 522–529 (1973)
2. Kano, Y., and Wolf E. (1962) Temporal coherence of black body radiation. Proc. Phys. Soci. (1958–1967) **80**(6): 1273
3. Nam, Y., Nam, Y.-C.: Vehicle classification based on images from visible light and thermal cameras. EURASIP J. Image Video Process. **2018**(1), 1–9 (2018)
4. Vojco, K., Aleksandar A., Dragan S., Mile S., and Tatjana K. Intelligent control of object detection: image processing and pathfinding problem. In Proceedings World Automation Congress, 2004., **17**: 131–136. IEEE, 2004.
5. Browne, M., and Ghidary S. S. Convolutional neural networks for image processing: an application in robot vision. In Australasian Joint Conference on Artificial Intelligence, pp. 641–652. Springer, Berlin, Heidelberg, 2003.
6. Shi, R., Ngan K. N., Li S. Jaccard index compensation for object segmentation evaluation. In 2014 IEEE International Conference on Image Processing (ICIP), pp. 4457–4461. IEEE, 2014.
7. Thakur, R. Infrared Sensors for object detection. In Recent Development in Optoelectronic Devices. IntechOpen, 2017.
8. Alam, M.S., Bognar, J.G., Hardie, R.C., Yasuda, B.J.: Infrared image registration and high-resolution reconstruction using multiple translationally shifted aliased video frames. IEEE Trans. Instrum. Meas. **49**(5), 915–923 (2000)
9. Giacomin, J. Thermal: seeing the world through 21st century eyes. Papadakis, 2010.
10. Rossignoli, I., Benito, P.J., Herrero, A.J.: Reliability of infrared thermography in skin temperature evaluation of wheelchair users. Spinal Cord **53**(3), 243–248 (2015)
11. Burigana, L., Magnini, L.: Image processing and analysis of radar and lidar data: new discoveries in Verona southern lowland (Italy). STAR: Sci. Technol. Archaeol. Res. **3**(2), 490–509 (2017)
12. Rakrueangdet, K., Nunak N., Suesut T., Sritham E. Emissivity measurements of reflective materials using infrared thermography. In International MultiConference of Engineers and Computer Scientists (IMECS). 2016.
13. Caillas, C. Thermal imaging for object detection in outdoor scenes. In EEE International Workshop on Intelligent Robots and Systems, Towards a New Frontier of Applications, pp. 651–658. IEEE, 1990.
14. Tsukamoto, T., Tanaka, S.: Patternable temperature sensitive paint using Eu (TTA) 3 for the micro thermal imaging. J. Phys.: Conf. Ser. **476**(1), 012073 (2013)
15. de Bittencourt MMGT, and Gonzaga A. "Digital image processing techniques for uncooled LWIR thermal camera." In Electro-Optical and Infrared Systems: Technology and Applications IX, vol. 8541, p. 85410Z. International Society for Optics and Photonics, 2012.
16. Liu, G., Liu, Z., Liu, S., Ma, J., Wang, F.: Registration of infrared and visible light image based on visual saliency and scale invariant feature transform. EURASIP J. Image Video Process. **2018**(1), 1–12 (2018)
17. LeBeau, T. Thermal imaging for safer object detection. In Infrared Technology and Applications XLV, vol. 11002, p. 110021H. International Society for Optics and Photonics, 2019.
18. Tan, S.-T., Chen, K., Ong, S., Chew, W.: Utilization of spectral vector properties in multivariate chemometrics analysis of hyperspectral infrared imaging data for cellular studies. Analyst **133**(10), 1395–1408 (2008)
19. Nath, S., Agarwal, S., Pandey, G.N.: Mathematical Foundation Based Inter-Connectivity modelling of Thermal Image processing technique for Fire Protection. EAI Endorsed Trans. Creat. Technol. **2**(5), 150097 (2015)
20. Peterson, B.J.: Infrared imaging video bolometer. Rev. Sci. Instrum. **71**(10), 3696–3701 (2000)
21. Lawrence, M., Ashley S. F., Lupton M., McEwen R. K., Wilson M. (2007) Signal processing core for high performance thermal imaging. In Infrared Technology and Applications XXXIII, vol. 6542, p. 654215. International Society for Optics and Photonics.
22. Zhang, J., Jia, X., Li, J.: Integration of scanning and image processing algorithms for lane detection based on fuzzy method. J. Intell. Fuzzy Syst. **29**(6), 2779–2786 (2015)
23. Pal, S. K. "Fuzzy sets in image processing and recognition." In [1992 Proceedings] IEEE International Conference on Fuzzy Systems, pp. 119–126. IEEE, 1992.
24. Iqbal, Q. M., Choudhury J. P., De M. "Image recognition and processing using Artificial Neural Network." In 2012 1st International Conference on Recent Advances in Information Technology (RAIT), 2012.
25. Li T., Wang Y., Chen Z., and Wang R. "Linear feature extraction for infrared image." In Image Extraction, Segmentation, and Recognition, vol. 4550, pp. 281–286. International Society for Optics and Photonics, 2001.
26. Gundogdu, E., Koc A., and Alatan A. A. "Object classification in infrared images using deep representations." In 2016 IEEE International Conference on Image Processing (ICIP), pp. 1066–1070. IEEE, 2016.
27. Smith, M.S., Silver, E.A., Stein, M.K.: Improving instruction in rational numbers and proportionality, vol. 1. Teachers College Press, New York (2005)
28. Zheng, D., Lin X., and Wang X. "Image Segmentation Method Based on Spiking Neural Network with Adaptive Synaptic Weights." In 2019 IEEE 4th International Conference on Signal and Image Processing (ICSIP), pp. 1043–1049. IEEE, 2019.
29. Sharma, S., Sharma, S., Athaiya, A.: Activation functions in neural networks. Towards Data Sci. **6**(12), 310–316 (2017)
30. Galagan, R., and Momot A. "The use of backpropagation artificial neural networks in thermal tomography." In 2018 IEEE First International Conference on System Analysis and Intelligent Computing (SAIC), pp. 1–6. IEEE, 2018.
31. Jiyao, L.F., Hu, M., Chen, W., Zhan, J.: A novel fuzzy-based convolutional neural network method to traffic flow prediction with uncertain traffic accident information. IEEE Access **7**, 20708–20722 (2019)
32. Kwan, C., Chou, B., Yang, J., Rangamani, A., Tran, T., Zhang, J., Etienne-Cummings, R.: Target tracking and classification using compressive measurements of MWIR and LWIR coded aperture cameras. J. Signal Inf. Process. **10**(3), 73–95 (2019)
33. Kwan, C., Chou, B., Yang, J., Tran, T.: Deep learning based target tracking and classification for infrared videos using compressive measurements. J. Signal Inf. Process. **10**(04), 167 (2019)
34. Kwan, C., Budavari, B.: Enhancing small moving target detection performance in low-quality and long-range infrared videos using optical flow techniques. Remote Sens. **12**(24), 4024 (2020)
35. Kwan, C., Chou B., Yang J., Tran T. "Compressive object tracking and classification using deep learning for infrared videos." In Pattern Recognition and Tracking XXX, vol. 10995, p. 1099506. International Society for Optics and Photonics.
36. Kwan, C., Gribben D., and Tran T. "Multiple human objects tracking and classification directly in compressive measurement domain for long range infrared videos." In 2019 IEEE 10th Annual Ubiquitous Computing, Electronics & Mobile Communication Conference (UEMCON), pp. 0469–0475. IEEE, 2019.
37. Sibi, P., Jones, S.A., Siddarth, P.: Analysis of different activation functions using back propagation neural networks. J. Theor. Appl. Inf. Technol. **47**(3), 1264–1268 (2013)
38. Agostinelli F., Hoffman M., Sadowski P., Baldi P. Learning activation functions to improve deep neural networks. arXiv:1412.6830.