TOWARDS THE DEVELOPMENT OF OBSTACLE DETECTION IN RAILWAY TRACKS USING THERMAL IMAGING

V. Vivek[*], *J. Hemalatha*[*], *T.P. Latchoumi*[†], *S. Mohan*[‡]

Abstract: To prevent collisions between trains and objects on the railway line, rugged trains require an intelligent rail protection system. To improve railway safety and reduce the number of accidents, studies are underway. Machine learning (ML) had progressed rapidly, creating new perspectives on the subject. A technique called speed up robust features (SURF) is proposed by researchers to collect regionally and globally relevant information. In addition, taking advantage of the Ohio State University (OSU) heat walker benchmarking dataset, the effectiveness of this technique was examined under various lighting scenarios. This technology could help in reducing train accident rates and financial costs. The findings of the proposed methodology are very specific, in addition to the ability to quickly identify items (obstacles) on the railway line, both of which contribute to rail security. The proposed faster region based convolutional neural network (FR-CNN) with 2D singular spectrum analysis (SSA) improves the performance analysis of an accuracy of 90.2%, recall 95.6% and precision 94.6% when compared with an existing system YOLOv2 and YOLOv5 with 2D SSA.

Key words: *intelligent railway, safety, thermal image, faster region based convolutional neural network, performance*

Received: March 10, 2023 **DOI:** 10.14311/NNW.2023.33.019
Revised and accepted: June 1, 2023

1. Introduction

A train often moves at a speed of one hundred kilometers per hour, which makes it difficult for the train operator to spot irregularities on the very remote railway, particularly at night, due to terrain barriers and harsh temperatures [1]. This has often made it difficult to stop in time to prevent an accident. To detect the presence or absence of a defect on the rail, the railroads were identified. Moreover, the

[*]Veeman Vivek – Corresponding author; Jeyaprakash Hemalatha; Department of Computer Science Engineering, AAA College of Engineering and Technology, Amathur, Virudhunagar, Tamilnadu, India, E-mail: vivekmsucse@gmail.com

[†]Thamarai Pugazhendhi Latchoumi; Department of Computer Science and Engineering, SRM Institute of Science and Technology, Chennai, Ramapuram, Tamilnadu, India

[‡]Sekar Mohan; Department of Mechanical Engineering, AAA College of Engineering and Technology, Amathur, Virudhunagar, Tamilnadu, India

premise that the railway had a consistent curve, which only applies to a limited degree, serves as one of the explanations for why the proposed rail detection technique doesn't operate satisfactorily for longer distances [2]. Locating the railroad remotely would be critical. One explanation would be that if the train stops too late, the huge braking distance makes it impossible to avoid or mitigate the effects of a collision. An approximate representation of a railway track using a model based on graph theory has been done as part of previous research into the problem of rail detection using visual cameras. 60% of tracks are fully or partially identified in a 362-frame dataset with only curved tracks [3, 4]. Instead, 90% of the tracks are fully discovered, and 7% are partially detected in a different sample containing 2044 frames [5]. Therefore, semantic segmentation is not the only obvious way to detect rail on photos, as promising rail models have been developed. Since people could accurately guess the rail location even when that was not in the device's line of vision, it is proposed that improvements be made generally to take the surroundings all around railways into account [6]. In addition, it should be noted that optical cameras have a major problem with lighting variations. Heavy slopes are deleterious, especially when associated with occlusions.

Deep learning (DL) systems have won many competitions in recent times and have also shown enormous achievements in a variety of patterns of identification and identification of objects. Furthermore, it has also been proposed that the main candidate for virtually any image processing challenge should be ML using convolutional neural networks (CNN) [7]. Additionally, it has been successfully applied to lane detection. Regardless of the major distinctions, traffic forecasting and rail recognition issues have some commonalities [8]. As shown, CNNs can be used effectively to identify the aforementioned patterns in the images even when they have been obscured [9, 10].

Restructuring frameworks and programmes appear to be the two main methodologies used in research to tackle the challenge of anomalous identification. The use of deep convolution systems, deep automatic encoders, or conventional fragmented encoding methods to generate descriptive restoration of the dataset [11–13]. The upcoming frame estimation technique uses the same strategy to identify anomalies by evaluating the discrepancies between the projected future frame and the currently obtained image [14].

2. Related works

The latter [15] is based on an approximation of the probability density of the characteristics of the movement and the regular appearances. Image enhancement and tracking using non-parametric and parametric estimation methods were generally used in this case [16]. The adoption of representation that uses DL has risen significantly. Although approaches for anomalous identification with fixed cameras reach the highest accuracies, extremely dynamic situations, such as railway inspections with photos from a drone, present difficulties in reconstructing whatever is regarded as normal [8, 17]. In addition, humans see that the majority of this type of work focuses on the circumstances of camera surveillance.

The use of infrared or ultrasound sensors installed remotely on trains has often been exploited in literary works. For obstacle detection, a programme based on

a distance sensor has been proposed [19]. When an element has been identified within the range of an infrared transmitter at the surface of the train, the lights change [20]. The system proposed by the researchers to detect and prevent obstacles and monitor railways depended on thermal cameras, GSM, and GPS technologies. An infrared sensor was used, similar to the previous search, to identify the barriers in the path of locomotives [21]. Both detection outputs have been combined with a vision system that detects forward obstructions. Recently, a free database called RailSem19 has been published for the interpretation of train and tramway semantic scenarios. The remote sensing techniques aren't mounted on the railway, but rather on the sides of the tracks, and they detect obstructions because there is no link between the transmitters and receptor.

To proposed work knowledge, this was the first database collected for the night rail scenario anomaly detection process. The collection of data takes place at night because inspections were scheduled when trains were not normally operating [22]. In light of the above, three key factors that are immediately related to the automotive environment must be considered when choosing acquisition methods. The use of thermal imaging cameras and external light sources has been employed to address this problem. Since the rail drone is important to highlight the restrictions on the energy use of visible light [23]. The Vesuvio collection contains more than 10000 frames. The presence and location of the barriers were carefully recorded in each frame. There seem to be 11 classes of anomalies or things put on the railway.

3. Problem description

Here's how the methodology works as follows:

1. To fully understand the issue at hand, conduct a literature review of object recognition particularly DL and CNN.
2. Establish an appropriate representation to analyze the performance.
3. Add track labeling capabilities to the existing labeling software according to the chosen representation.
4. Select an appropriate volume of training data and use the labeling software listed above to annotate it.
5. Research the available ML programming framework and choose an appropriate option.
6. Implement the chosen structure of CNN framework.
7. Creates and evaluates an appropriate nonlinear function.
8. Test the proposed model based on performance measures.

4. Proposed framework

Fig. 1 illustrates a deep network as an instance of DL tries to solve problems that seem to be instinctively such as identifying a car in an image. It can be hard for

a human to come up with a mathematically sound solution to these problems. A system can learn to cope with these difficulties by breaking down the aforementioned identification into a hierarchy of information connected in several layers at various levels of abstraction. The graphics get deep when they have been drawn to show the ratio between layers.

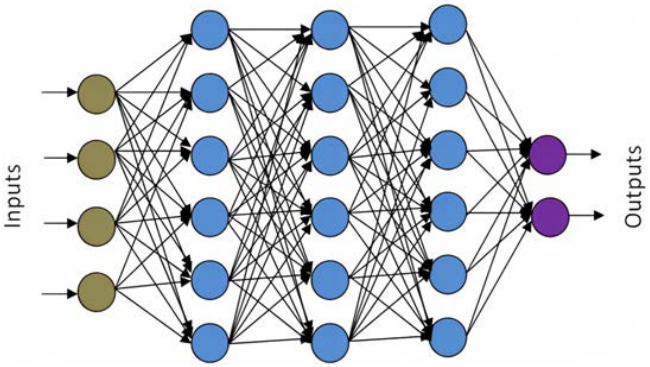


Fig. 1 *DL network framework.*

As illustrated in Fig. 2, the researchers suggest architecture composed of three possible structures, each of which focuses on various tasks: aberration detection, abnormal location, and abnormal categorization. The information for such a structure was frames obtained by the imaging system, which would have climbed into

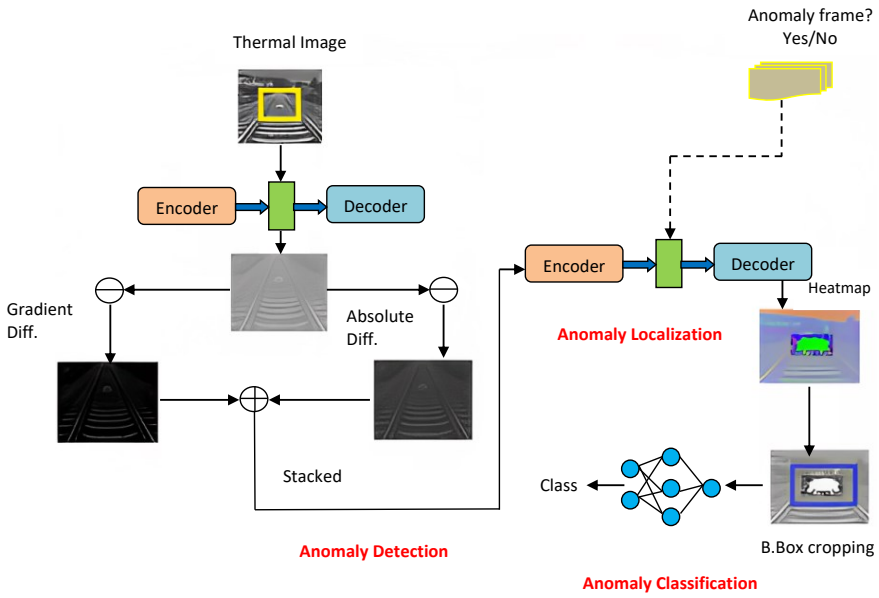


Fig. 2 *The proposed architecture.*

the surface of railway drones. The model consists of two distinct systems, each of which has a size restriction to balance the analysis as an essential part, speed of reasoning, and power consumption. The very first system, here known simply as the "autoencoder," seems to be a deep encoder-decoder design that only accepts normal frames as input and provides reconstructed versions. In that case, such a reconstruction should show a clear image free of any anomalies. Then, using an absolute and gradient differential or a difference estimated only on the gradient, of the image pairs, this reconstruction was matched to the input. The two resulting images would then be combined into a two-channel image and inserted into the module convolution layer, which predicts whether or not anomalies will appear in the frames. This second system utilizes both data about various textures and patching and knowledge about different contours and lines as a result of using different photos as inputs.

This approach consists of three main procedures: the detection of SURF points of interest, the distribution of the main direction and the gathering of information dissemination, and the construction of descriptors [24]. SURF has been used to extract the local data. Edges and cracks in the specimen exhibit high thermal differential than other places in active thermal imaging, shape information from edges was employed to depict relevant data is shown in Fig. 3. It should be noted that proposed method improves performance by simultaneously extracting the main

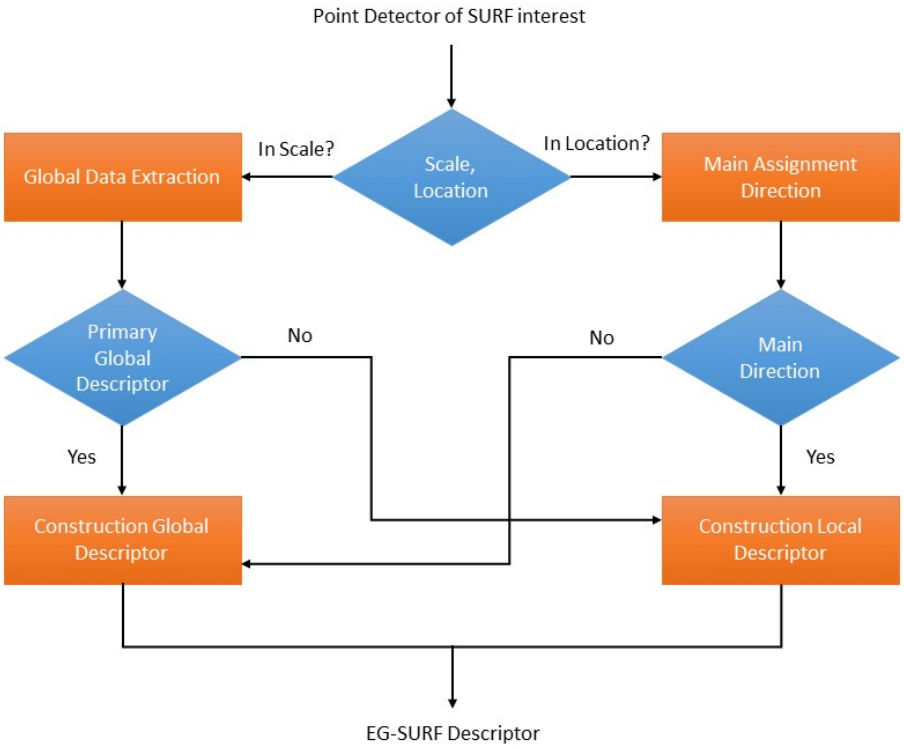


Fig. 3 *A SURF diagram.*

global descriptors and assigning the main management. In the meantime, shape contextual descriptions were produced using the scale and placement of SURF points of interest, which increases the quality of the description.

We use the same local detectors based in Hessen, known as SURF. The Hessian matrix $\mathbf{H}(x)$ in (i, j) at scale for a point (i, j) in an image with size $i \times j$ is:

$$\mathbf{H}(x) = \begin{bmatrix} L_{ii}(x) & L_{ij}(x) \\ L_{ji}(x) & L_{jj}(x) \end{bmatrix}. \tag{1}$$

By quantizing the Hessian matrix's determinant, interest points were discovered. SURF also uses Laplacian of Gaussian (LoG) can filters and approximations to get the scaling constant. Therefore, the determinants yield,

$$\det(\mathbf{H}) \approx D_{ii}D_{jj} - (\alpha D_{ij})^2, \tag{2}$$

$$\alpha = \frac{\|L_{ij}(1.2)\|_{\mathrm{F}}\|D_{ii}(1.2)\|_{\mathrm{F}}}{\|L_{ii}(1.2)\|_{\mathrm{F}}\|D_{ij}(1.2)\|_{\mathrm{F}}} \approx 0.9, \tag{3}$$

where $\|\cdot\|_{\mathrm{F}}$ – Frobenius norm; L – current pixel of image with scaling constant 1.2; D – vicinity of image; α – Gaussian centred; s – stack; the normalized filter response with size $L \times L$ represented as

$$L = ((s + 1) + 1). \tag{4}$$

To obtain highly precise objects of interest using the 2nd order Taylor series approximations, 3D quadratic extrapolation has been discovered.

$$\det(\mathbf{H})|_i \approx \det(\mathbf{H})|_{i_0} + \frac{\partial\left(\det(\mathbf{H})\,|_i\right)^2}{\partial i} + \frac{1}{2}i^{\mathrm{T}}\frac{\partial^2(\det(\mathbf{H})|_i)}{\partial i^2}i. \tag{5}$$

By decomposing Eq. (5) set equal to 0, the interpolated location and scale were determined.

$$i_{\max} = \left[\frac{\partial^2 \det(\mathbf{H})|_i}{\partial i^2}\right]^{-1}\left[\frac{\partial \det(\mathbf{H})|_i}{\partial i}\right]. \tag{6}$$

By a wide margin, the position and dimensions of each object of interest were known; humans represent them as $P = [i\, j]^{\mathrm{T}}$, where P is the location of object of interest.

Every point of interest has been assigned a major orientation, to ensure invariance. First, wavelet packet filters, frequency hoping (FH) are used to convoke pixels in the vicinity of a point within a radius of 6 in both the i and j directions. Every pixel's filter reaction was graded using a Gaussian function with a frequency of 2 based on how far away it stands from the location of the object of interest, P. Gaussian filters are used in image processing because they have a property that their support in the time domain, is equal to their support in the frequency domain. This comes about from the Gaussian being its own Fourier transform. The mathematical notation for this procedure seems to be:

$$R_i = N_{6\sigma}^{(i)} \cdot F_{\mathbf{H}}G\left(I/6\sigma, 2\sigma\right), \tag{7}$$

$$R_j = N_{6\sigma}^{(j)} \cdot F_{\mathbf{H}}G\left(I/6\sigma, 2\sigma\right), \tag{8}$$

where R – velocities of pixels; $F_{\mathbf{H}}G$ – histogram entry; N – state variables of the filter; I – image. After that, $\pi/3$ sector has been used to aggregate the weighted Haar responses defined as

$$mw_k = \sum_{i=k}^{n} R_i + \sum_{j=k}^{n} R_j. \tag{9}$$

These red spots represent Haar response; the longer the arrowhead, the higher the reaction, and the more spots encompassed in the area shown in Fig. 4. The primary vector of the current object of interest seems to be the angle of rotation with the largest magnitude.

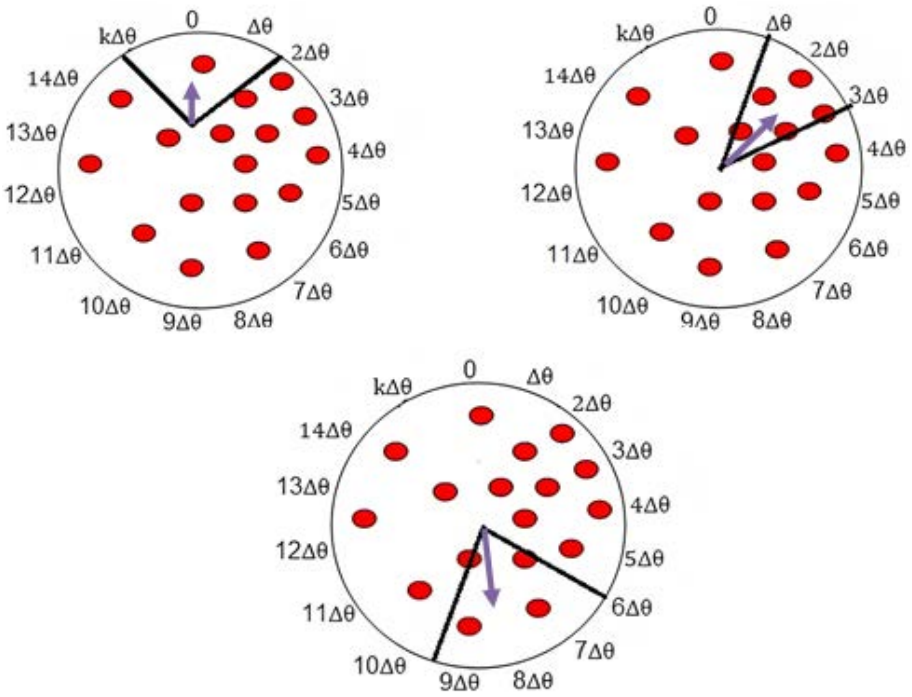


Fig. 4 *Major distribution of directions graphically.*

4.1 Track detection

The identification of railway tracks is the initial and most important component of the proposed work. Using the singular 2D spectrogram, the frame was split into several image and discriminatory elements. Horizontal, vertical, and diagonal information about the scene is contained racist and discriminatory elements [25]. The element of the runway arrangement with diagonal features has the most similar. From the image and discriminatory elements of the specific frame, xth element was chosen, which has a significant trend for the railroad track:

$$S_{\text{track}} = S^{(X)}|_{X=l}, \tag{10}$$

where $S^{(X)}|_{X=l}$ represents railway track information based on discriminative pattern component.

Then, utilizing background subtraction segments, the locations of the railroad lines are recovered from the xth element. The following seems to be the track's characteristics:

$$S_{\text{seg}}(x,y) = \begin{cases} 1 & \text{if } T_{\text{low}} \le S_{\text{track}}(x,y) \le T_{\text{high}}, \\ 0 & \text{else.} \end{cases} \tag{11}$$

4.2 Deep network for object recognition

The FR-CNN is made up of two components: the deep fully CNN to suggest regions, and FR-CNN to identify objects [26]. It seems to be anchored with three scales and three ratios at the position of the image in the standard arrangement. Finally, the classification is done using the FR-CNN network, which seems to have two completely connected layers. Proposals have been classified into a different

Algorithm 1 Object recognition and railway track detection.

Input: Sequence of railway track thermal video, g: the number of component to extract m and n: window size (the embedding dimension).
Output: Railway tracks and obstacles.
Step 1. Extract the frames $X(x,y)$ from the video sequence.
Step 2. Decompose the frame $X(x,y)$ into elementary components using automatic SSA.
$T \leftarrow \text{BuildTrajectoryMatrix}(X,m,n)$
$R \leftarrow \text{Rank}(T)$
Apply StandardDeviation(T)
for $k \leftarrow 1$ to R **do**
$\{T_k \leftarrow \lambda_k E_k E_k^{\text{T}}$
$t^{(k)} \leftarrow \text{DiagAveraging}(T_k)$
$l_x \leftarrow \text{StandardDeviation}(\{t^{(1)}, t^{(2)}, \ldots, t^{(R)}\}, g)\}$
end for
for $x \leftarrow 1$ to g **do**
$\{|\text{S}^{(x)} \leftarrow \sum_{x \in l_x} t^{(k)}$
$\text{Return}\{S^{(1)}, S^{(2)}, \ldots, S^{(g)}\}\}$
end for
Case 1: Railway Track Detection
Step 3. Select the ith component having railway track features from the g output components.
Step 4. Extract the railway track features from the ith component.
Case 2: Object Recognition
Step 5. Implement the DL to identify the object from the kth obstacle containing component.
Step 6. Coordinate the railway track object information to generate the final output.

Class N1 by a single layer. For better optimization of the boundary box for the "N" classes using regression prediction, another completely connected layer was used. The specified regression model must be constrained.

$$F_{\text{loss}}\left(m_a, n_a\right) = \frac{1}{C}\sum_{a}^{n} L_c\left(m_a, m_a^*\right) + \gamma\frac{1}{S}\sum_{a}^{n} m_a^* L_s\left(n_a, n_a^*\right), \tag{12}$$

where m_a – object predicted probability; n_a – predicted bouncing box parameterized coordinates; L_c – classification loss and L_s – regression loss respectively.

5. Experimental setup

The images used in this room were taken from a series of thermo-video railroad rails. To assess the performance of the proposed work, a total of 749 thermal video images have been used.

5.1 Decomposition of railway track

Before compensating for the interference, the temperature production order of a railway image was first recovered. Fig. 5(a) shows the movement artifact removal results. Then, using 2D SSA, every image was divided into several parts. The window size, $m \times n$, and also the number of elements g, to be deconstructed from the input frame were two important variables that influence the SSA performance in Fig. 5(b). When the window size was $m \times n^{1/4}$ 8×8 and also the number of discriminating elements was $g = 15$, the most informative component is deconstructed. The effectiveness of the proposed approach was examined for various values of m, n and g. Accordingly, each frame has been divided into 15 specific items that are seen in Fig. 5(c). The various parts of the input image that were deconstructed reveal the different details of the scene that the image captured. The required track data and obstructions are used in the situation to improve the in-depth assessment.

5.2 The output of the proposed framework

The researchers use two different dataset configurations to study the proposed system. In the former, all abnormal frames—that is, frames with an object—are grouped. Lastly, humans randomly select an equivalent amount of regular images from the test and training selections from the data set. Researchers call this typical information arrangement "80-20". In the second set, which may be more difficult, we choose 3 classes that are different from one another in terms of appearance for the testing set and also the remaining 8 classes for something like the given dataset shown in Fig. 6. Then, like in the last instance, humans select an equivalent number of normal frames for both the testing and training portions of the database.

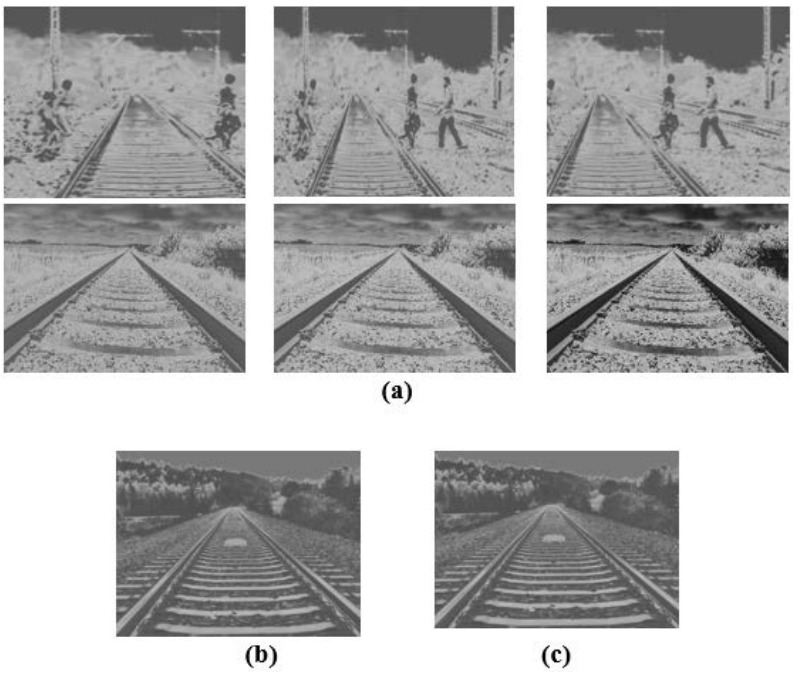


Fig. 5 *Track data and necessary barriers are used in the situation to improve the thorough assessment. (a) Input images, (b) after moving feature extraction, the image, (c) single components of the input frame have been deconstructed.*

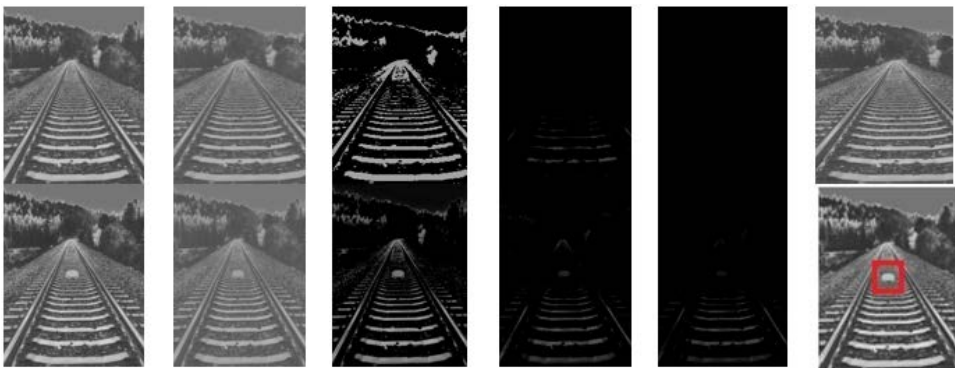


Fig. 6 *The proposed framework results.*

Tab. I lists the results for the abnormal localization, categorization, and reconnaissance tasks. The experimental results generally support the notion that the proposed framework was able to address things that were not encountered during the training phase. The categorization of the different images indicates a reasonable method for locating anomalies on the railways, and temperature information was almost certainly an excellent choice. Since the classifier has been trained and eval-

Modules based on anomaly	80-20 rule				Cross-class rule			
	Accuracy	Precision	Recall	F1-score	Accuracy	Precision	Recall	F1-score
Detection	0.968	0.988	0.955	0.974	0.849	0.904	0.778	0.836
Localization	0.902	0.742	0.990	0.848	0.786	0.522	0.998	0.685
Classification	0.972	0.796	0.795	0.786	–	–	–	–

Tab. I *Outcomes of experiments using the proposed framework.*

uated in all categories, humans just keep the "80-20" parameter for the anomalous classification problem. The findings for each category were presented in Tab. II, while the overall results were outlined in Tab. II.

Class type	Accuracy	Precision	Recall	F1-score
Electrical insulator	0.95	0.55	0.46	0.49
Fuel tank	0.97	0.66	0.77	0.72
Rail signal	0.98	1.00	0.78	0.88
Pickaxe	0.96	0.57	0.93	0.71

Tab. II *The outcomes of the anomalous categorization module.*

The magnitude of the anomaly hurts the recall and accuracy rate as demonstrated by the fact that electromagnetic insulation would be the lowest anomaly in the data set. Low ratings were also given to elements that resemble components often found visually on railways, such as pickaxes and markers. In addition, the researchers point out that the use of shallow arrays for the classification problem does not materially affect the results.

There should be a higher significance on the edge inside the size of the point of interest. All edge pixels are first considered to be 1, while unmarked pixels were considered to be 0. Recalling the Gaussian function compared to different number of frames λ is shown in Fig. 7. With small, the function's output decrements quickly. It is acceptable for proposed work weighing needs because of this characteristic. The histogram would then be computed using the graded edges, as shown in Fig. 8. The histograms have been divided into the same pitch size, while the circles are used. On frames intensity E, the red blotches represent edge pixels. As seen on the right-hand side of Fig. 8, all regions REs were segmented using a vector decomposition method to their adjacent cells. In comparison with non-binding approaches, this vector deconstruction approach provides more reliable and accurate contextual information.

To further verify the effectiveness of the proposed approach, Fig. 9 shows a horizontal assessment of several registration techniques. Precision is reserved only when the corresponding numbers are above 4. For each procedure, the same number of positions of interest (POI) was generated. The proposed approach, 350, displays the highest accuracy curves since it considers the SURF and form context as a

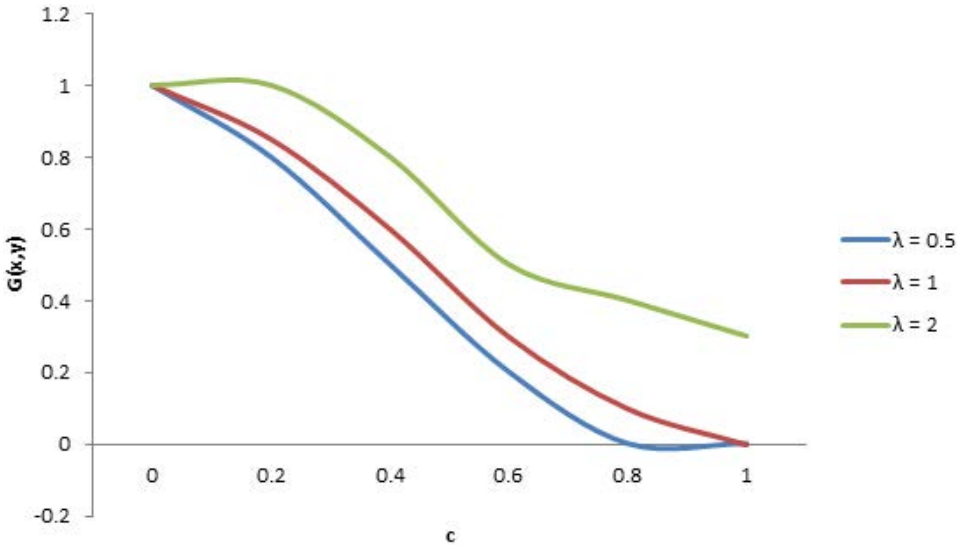


Fig. 7 *Gaussian function under different* λ.

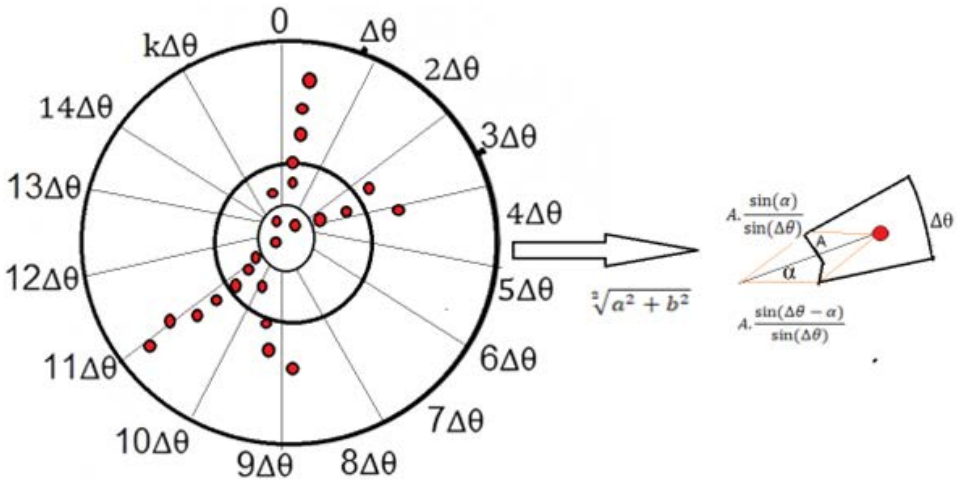


Fig. 8 *Graphic depiction of form context.*

special instance. The SIFT form context, and accuracy of extended orientation scale invariant feature transform (EOHSIFT) are all below 0.5 for all distance thresholds, making them virtually unusable for the recording of cross-spectrum track images.

5.3 Railway track detection

It must be discovered that the element stack = $s(4)$ of every frame includes significant data about the railroad tracks after every element deconstructed from each

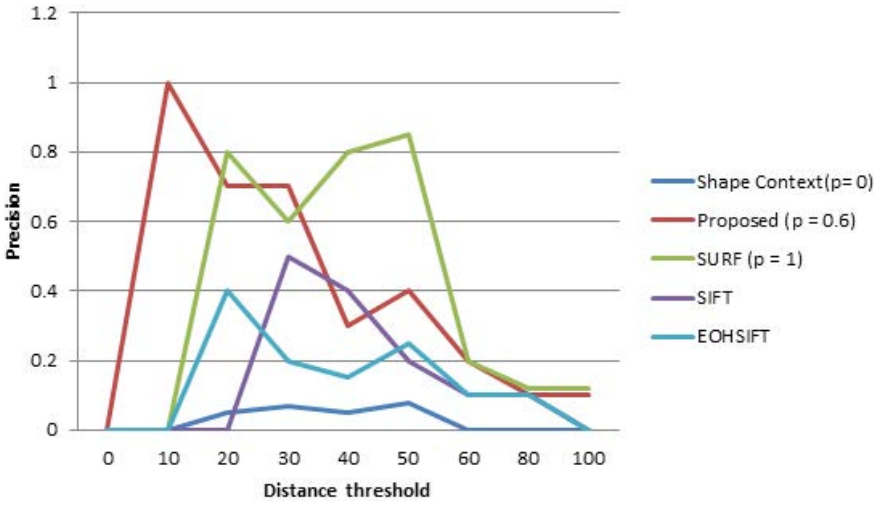


Fig. 9 *Compares the distance threshold and precision of various techniques.*

frame was studied to detect the railway track. As a result, in this work, the identification of the railway track was done using the fourth component of each frame. The properties of the fourth piece of the track are then separated and used to identify the track. To illustrate the tracks, these features were finally brought to the screen is shown in Fig. 10(a). Fig. 10(b), (c), and (d) illustrate the fourth

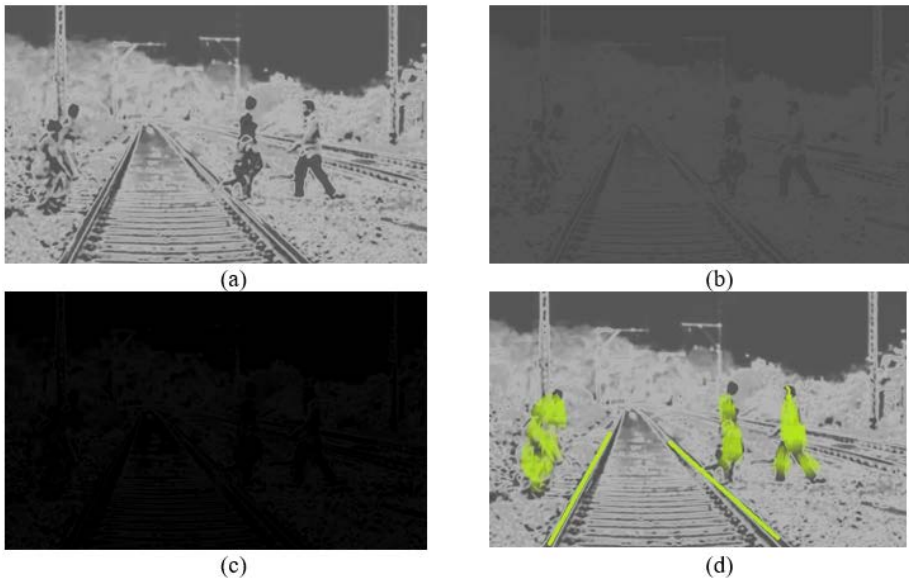
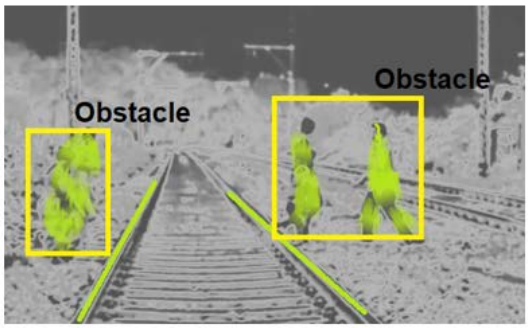


Fig. 10 *The input frame, the fourth element of the input frame, which contains the majority of the data about the railway line, the fourth element's track attributes, and the recognized railway track.*

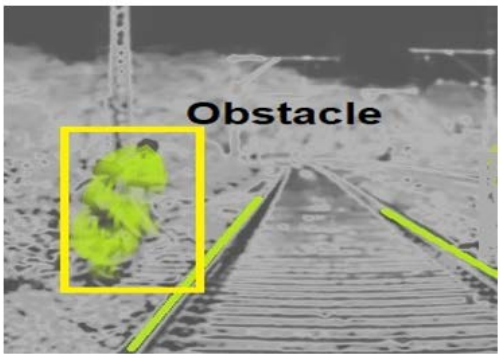
component of the frames, the characteristics retrieved from the elements, and also the recognized models finalized accordingly. Similarly, how the railway track was recognized in all frames, this enables the identification of the railway track.

5.4 Performance evaluation

To control the speed of the train and prevent collisions, obstacles on the track must be seen at an adequate distance. By increasing the intersection over union (IoU) setting, the FR-CNN gives better results for timely recognition of the barrier. IoU values ranging from 0.7 to 0.9 are generally considered to be a sufficient range for accurate determination using FR-CNN network. The preliminary results of the proposed method are obtained from a recognized array using a threshold of 0.7 as shown in Fig. 11(a), but the barrier was detected closer to the railway. A compromise between the recognition rate and the range of the train at the time of identification was reached by analyzing system performance for different IoU combinations. To address this problem, it has been found that the FR-CNN



(a)



(b)

Fig. 11 *Results of track classification and obstruction identification using FR-CNN at IoU = 0.7 and 0.5 respectively.*

achieves high accuracy and timely obstacle recognition at the IoU threshold level of 0.5. Fig. 11(b) shows the result of the track recognition and barrier assessment at IoU = 0.5.

Tab. III lists the variables that were assessed for various IoU values. At IoU "0.5," measurements of accuracy, precision, and recall are higher. Fig. 12 illustrates the performance of the proposed strategy. The data lead to the conclusion that the planned approach works best at IoU = 0.5 with a precision of about 85%.

IoU	Accuracy	Precision	Recall
0.5	90.2	94.5	95.6
0.7	87.0	92.0	94.0

Tab. III *Proposed technique performs for various IoUs.*

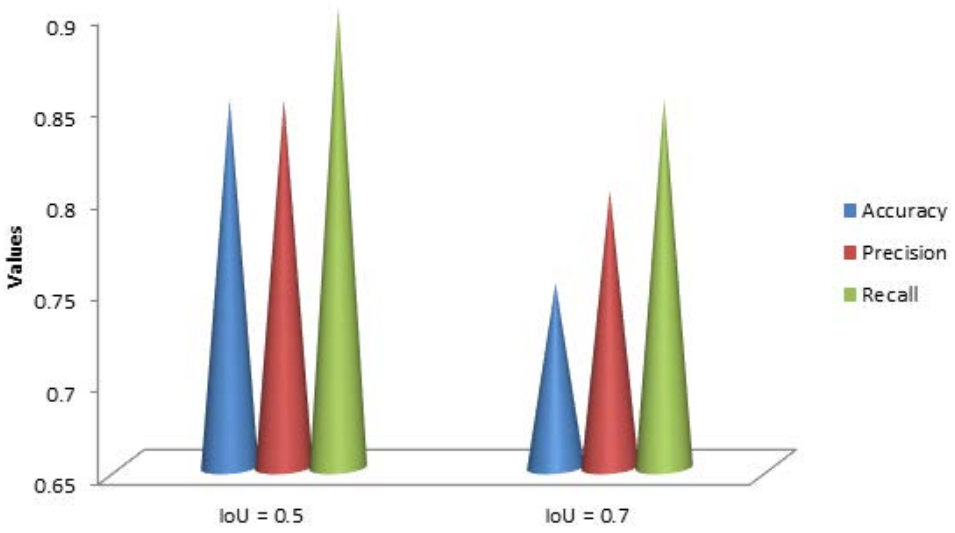


Fig. 12 *Performance measures of IoU levels.*

5.5 Performance comparison

For thermally railroad track information with these settings, Tab. IV compares the effectiveness ofproposed workproposed technique with the other object detection methods.

Method	Accuracy	Precision	Recall
YOLOv2	0.723	0.717	0.744
YOLOv5 + 2D SSA	0.848	0.844	0.882
FR-CNN + 2D SSA	0.900	0.946	0.956

Tab. IV *Comparison of performance measures.*

The FR-CNN detector design was utilized for testing and training purposes. Fig. 13(a) depicts a thermal image taken in the afternoon during light rain, with the following conditions: temperatures of 68 °F, a minimum dew point of 58 °F, a humidity of 70%, and vision of 9.0 miles. Given that the second element significantly incorporates pedestrian data, the image was dissected using a 2D SSA, as illustrated in Fig. 13(b). The FR-CNN would then be used together with the second thermal element to identify pedestrians. The result of the FR-CNN with the detected pedestrians may be seen in Fig. 13(c). Fig. 13(d) shows the pedestrians identified in the light rain source images. The proposed method with thermal imaging functions effectively in the rain situations

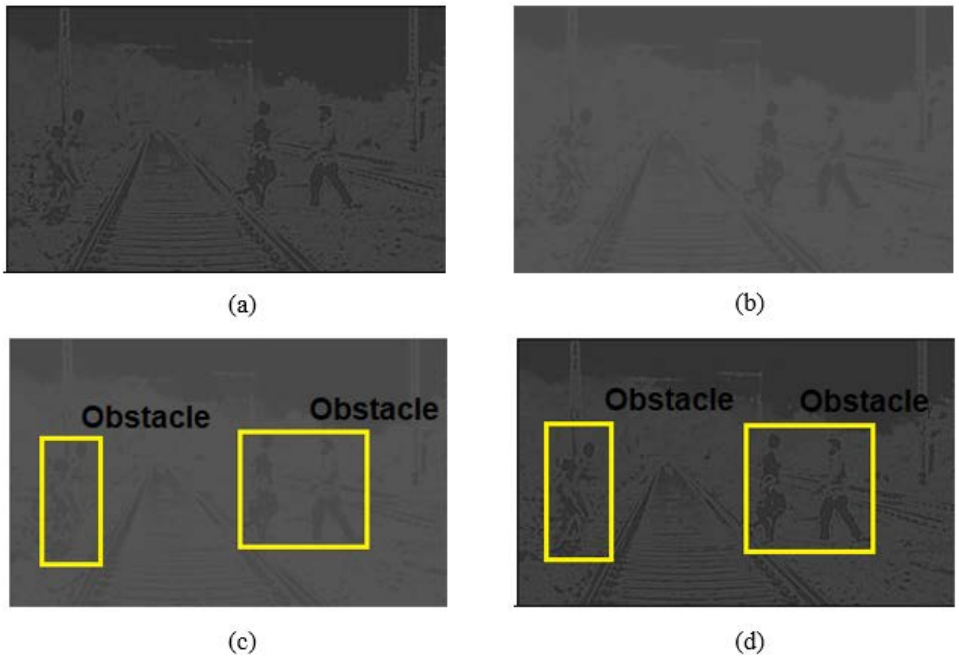


Fig. 13 *(a) Input thermal image in a light rain situation, (b) the input image's second component with a predominance of pedestrian data, (c) identified walkers in the input thermal imaging of the second process, (d) conclusion.*

6. Conclusions

This study would present a new and effective method for identifying railway tracks and barriers. In the proposed method, the thermal image was divided into multiple information-carrying elements using the 2D SSA, and each element would then be employed by the FR-CNN algorithm to recognize the obstruction on the railway lines. Another relevance of this work is the identification of tracks using the SSA. The FR-CNN with 2D SSA were combined to provide a more effective and reliable method of developing an advanced warning system to prevent rail accidents and increase rail safety. The experimental results provide evidence of the utility of the proposed approach. By balancing a low computational load and excellent reliability, surface architectures provide a performance benefit. Future research should concentrate on the use of brightness images and stereo information from the dataset in combination with heat information to increase the performance measures.

Conflict of interest Authors do not have conflict of interest

Research involving human participants and/or animals This article does not contain any studies with human participants or animals performed by any of the authors.

Funding statement No funding

Data availability statement No data associated with this manuscript

Author contribution Vivek Veeman – Conceptualization and validation,
Hemalatha Jeyaprakash – Validation,
Thamarai Pugazhendhi Latchoumi – Methodology,
Sekar Mohan – Training and testing.

References

[1] STYPUŁKOWSKI K., GOŁDA P., LEWCZUK K., TOMASZEWSKA J. Monitoring system for railway infrastructure elements based on thermal imaging analysis. *Sensors*, 2021, 21(11), 3819.

[2] JING G., QIN X., WANG H., DENG C. Developments, challenges, and perspectives of railway inspection robots. *Automation in Construction*, 2022, *138*, 104242.

[3] ŁUKASIAK J., ROSIŃSKI A., WIŚNIOS M. The Impact of Temperature of the Tripping Thresholds of Intrusion Detection System Detection Circuits. *Energies*, 2021, *14*(20), 6851.

[4] GARIKAPATI P., BALAMURUGAN K., LATCHOUMI T.P., MALKAPURAM R. A Cluster-Profile Comparative Study on Machining AlSi 7/63% of SiC Hybrid Composite Using Agglomerative Hierarchical Clustering and K-Means. *Silicon*, 2021, *13*, pp. 961–972.

[5] WACHNIK B., KŁODAWSKI M., KARDAS-CINAL E. Reduction of the information gap problem in Industry 4. 0 projects as a way to reduce energy consumption by the industrial sector. *Energies*, 2022, *15*(3), 1108.

[6] JAKUBOWSKI K., PAŚ J., ROSIŃSKI A. The Issue of Operating Security Systems in Terms of the Impact of Electromagnetic Interference Generated Unintentionally. *Energies*, 2021, *14*(24), 8591.

[7] PAS J., KLIMCZAK T. Rosinski A. *Stawowy, M. The Analysis of the Operational Process of a Complex Fire Alarm System Used in Transport Facilities. Build. Simul*, 2022, *15*, pp. 615–629.

[8] NGUYEN T.X.B., CHAHL J. Sparse Optical Flow Implementation Using a Neural Network for Low-Resolution Thermal Aerial Imaging. *Journal of Imaging*, 2022, *8*(10), 279.

[9] PAŚ J., ROSIŃSKI A., WIŚNIOS M., STAWOWY M. Assessing the Operation System of Fire Alarm Systems for Detection Line and Circuit Devices with Various Damage Intensities. *Energies*, 2022, *15*(9), 3066.

[10] ŁUKASIAK J., ROSIŃSKI A., WIŚNIOS M. The Issue of Evaluating the Effectiveness of Miniature Safety Fuses as Anti-Damage Systems. *Energies*, 2022, *15*(11), 4013.

[11] RISTIĆ-DURRANT D., HASEEB M.A., BANIĆ M., STAMENKOVIĆ D., SIMONOVIĆ M., NIKOLIĆ D. Smart on-board multi-sensor obstacle detection system for improvement of rail transport safety. *Proceedings of the Institution of Mechanical Engineers, Part F: Journal of Rail and Rapid Transit*, 2022, *236*(6), pp. 623–636.

[12] ANAND M., BALAJI N., BHARATHIRAJA N., ANTONIDOSS A. A controlled framework for reliable multicast routing protocol in mobile ad hoc network, Materials Today: Proceedings, 2021, ISSN 2214-7853.

[13] RISTIĆ-DURRANT D., FRANKE M., MICHELS K. A review of vision-based on-board obstacle detection and distance estimation in railways. *Sensors*, 2021, *21*(10), 3452.

[14] BADRLOO S., VARSHOSAZ M., PIRASTEH S., LI J. Image-Based Obstacle Detection Methods for the Safe Navigation of Unmanned Vehicles: A Review. *Remote Sensing*, 2022, *14*(15), 3824.

[15] RAHIMI M., LIU H., CARDENAS I.D., STARR A., HALL A., ANDERSON R. A Review of Technologies for Localisation and Navigation in Autonomous Railway Maintenance Systems. *Sensors*, 2022, *22*(11), 4185.

[16] PELESKA J., HAXTHAUSEN A.E., LECOMTE T. Standardization considerations for autonomous train control. In: *International Symposium on Leveraging Applications of Formal Methods*. Springer, Cham, 2022, pp. 286–307.

[17] GUAN L., JIA L., XIE Z., YIN C. A Lightweight Framework for Obstacle Detection in the Railway Image based on Fast Region Proposal and Improved YOLO-tiny Network. *IEEE Transactions on Instrumentation and Measurement*, 2022, *71*, pp. 1–16.

[18] KYATSANDRA A.K., SAKET R.K., KUMAR S., SARITA K., VARDHAN A.S.S., VARDHAN A.S.S. Development of TRINETRA: A Sensor-Based Vision Enhancement System for Obstacle Detection on Railway Tracks. *IEEE Sensors Journal*, 2022, *22*(4), pp. 3147–3156.

[19] FRANKE M., GOPINATH V., RISTIĆ-DURRANT D., MICHELS K. Object-Level Data Augmentation for DL-Based Obstacle Detection in Railways. *Applied Sciences*, 2022, *12*(20), 10625.

[20] VIVEK V., MURUGAN D., SINGARAVELAN S. Development of Obstacle Detection and Collision Avoidance System using Thermal Imaging Technique,*Asian Journal of Research in Social Sciences and Humanities, indianjournals,* 2017, 7(11), pp. 128–148.

[21] LATCHOUMI T.P., SWATHI R., VIDYASRI P., BALAMURUGAN K. Develop New Algorithm To Improve the Safety Of WMSN In Health Disease Monitoring. In: *2022 International Mobile and Embedded Technology Conference (MECON)* IEEE, 2022, pp. 357–362.

[22] AHMED W., HANIF A., KALLU K.D., KOUZANI A.Z., ALI M.U., ZAFAR A. Photovoltaic panel classification using isolated and transfer-learned deep neural models using infrared thermographic images. *Sensors*, 2021, *21*(16), 5668.

[23] DI TOMMASO A., BETTI A., FONTANELLI G., MICHELOZZI B. A Multi-Stage model based on YOLOv3 for defect detection in PV panels based on IR and Visible Imaging by Unmanned Aerial Vehicle. *Renewable Energy*, 2022.

[24] KUMAR R.E., TIIHONEN A., SUN S., FENNING D.P., LIU Z., BUONASSISI T. Opportunities for ML to accelerate halide-perovskite commercialization and scale-up. *Matter*, 2022, *5*(5), pp. 1353–1366.

[25] BOMMES L., HOFFMANN M., BUERHOP-LUTZ C., PICKEL T., HAUCH J., BRABEC C., MARIUS PETERS I. Anomaly detection in IR images of PV modules using supervised contrastive learning. *Progress in Photovoltaics: Research and Applications*, 2022, *30*(6), pp. 597–614.

[26] BAI T., YANG J., XU G., YAO D. An optimized railway fastener detection method based on modified FR-CNN. *Measurement*, 2021, *182*, doi: 10.1016/j.measurement.2021.109742.