IOSR Journal of Engineering (IOSRJEN) www.iosrjen.org
ISSN (e): 2250-3021, ISSN (p): 2278-8719
Vol. 14, Issue 4, April 2024, ||Series -2|| PP 159-167

Small Target, Big Impact: Unveiling EBSE YOLO FOR Foreign Object Detection

[1] **M. Niharika**, *Student, Department of CSE, DNR College of Engineering and Technology, nihaniharika967@gmail.com*

[2] **P. Santhi Priya**, *Student, Department of CSE, DNR College of Engineering and Technology, santhipotturi3005@gmail.com*

[3] **L. Teja**, *Student, Department of CSE, DNR College of Engineering and Technology, tejalokam@gmail.com*

[4] **B. Surya Vamsi**, *Student, Department of CSE, DNR College of Engineering and Technology, Suryavamsibandaru@gmail.com*

[5] **Mr. Bujjibabu Lingampalli**, *MTech, Assistant professor, Department of CSE, DNR College of Engineering and Technology, bujjibabudnr@gmail.com*

Abstract: *Due to its inability to identify small foreign objects, the electrified train system requires a highly accurate detection algorithm to ensure operational continuity and safety. This study presents an enhanced approach called EBSE-YOLO. Along electrified rails, it will help better identify small targets. As a result of using cutting-edge techniques such as ECA-Net to prioritize tiny targets, SPD-Conv to extract details, and EIOU loss function to align dimensions, EBSE-YOLO significantly improves detection accuracy while minimizing computational cost. Results for the proposed technique are excellent after extensive testing with various YOLOv5 configurations and other approaches, including Ghost CNN. Initial findings demonstrate that EBSE-YOLO attains a 97% mAP, and there is optimism that further enhancements, such YOLOv5 + Ghost CNN, might potentially obtain an even greater mAP. The effects of EBSE-YOLO on railway management and safety extend far beyond quantitative indicators of success. By using state-of-the-art methods and distinctive model structures, EBSE-YOLO not only improves microscopic target recognition but also sets the stage for continuous improvements to railway safety requirements. The ground-breaking work on object recognition algorithms designed for complex conditions that this study produces will greatly improve railway operations in terms of safety and efficiency.*

Few words: YOLOv5, ECA-Net, BiFPN, SPD-Conv, EIOU, small target, international.

I. INTRODUCTION

The safe and efficient transfer of people and goods is made possible by rail transit, and particularly by electrified railroads, which are an essential part of the modern transportation network. To maintain safe and reliable train operations, it is crucial to safeguard railroad systems against intrusion and conduct regular inspections for foreign items [1]. Plastic sheds and other non-native goods pose a significant risk to railroad safety because they interfere with train operations and communication networks. The consequences of these invasions, which may range from operational disruptions to full system failures, are catastrophic, so effective detection and mitigation strategies are critically required.

Conventional methods of train inspection, including human inspectors or integrated inspection vehicles, have certain limitations, such as not being able to identify problems in real time and leaving blind spots unchecked, which may lead to operational risks [2]. The modern safety standards for train inspections are too high because of the inherent unpredictability of foreign item intrusions. Authorities in the railroad industry are therefore seeking innovative solutions to enhance inspection capabilities along railway lines [2]. One such method is to examine extraterrestrial objects using unmanned aerial vehicles (UAVs) equipped with high-definition cameras [3]. Problems with automation have kept UAV-based inspection from gaining traction, despite its many advantages, such as better real-time monitoring and the removal of blind zones. Given these constraints, it is critical to enhance train safety procedures by creating smarter and more efficient methods of identifying foreign objects.

Deep learning algorithms and traditional target detection methods are the two primary avenues for artificial intelligence foreign object identification [4]. In place of traditional approaches beset by slow computation rates and insufficient accuracy, deep learning algorithms—particularly those built on convolutional neural networks (CNNs)—have emerged as the preferred choice for detection tasks [4]. Two primary varieties of deep learning-based target identification algorithms exist: one with two steps and one with one. Although two-stage algorithms are very accurate, they are not appropriate for scenarios requiring quick detection replies due to their slow detection rates. Some single-stage algorithms, such as SSD (Single Shot Multibox Detection), RetinaNet, and the YOLO (You Only Look Once) series, provide faster detection rates at the expense of some accuracy [4, 5, 6, 7].

The exceptional detection abilities of YOLOv5 have brought it great fame among these single-stage detection algorithms [9], [10], [11]. The direct application of YOLOv5 to the job of foreign item recognition aboard electrified trains is contingent upon overcoming a number of barriers. To begin, the different outside environments make detection efforts in train circumstances more difficult due to background interference. Additionally, when there are both solid and squishy floating objects, feature extraction is harder, particularly when the squishy floating things' shapes change. The subjects will seem smaller in the end output because of the need of maintaining a safe distance from them while utilizing drones to capture images. Hence, our current YOLO algorithms fall short of the requirements for very precise algorithms that can identify very small targets. In order to detect small foreign objects traveling over electrified trains, this study introduces an improved detection method. In view of these challenges, it aims to extend the capabilities of existing detection methods. By combining cutting-edge methods with deep learning, this system aims to enhance detection accuracy, cut down on safety threats, and boost operational efficiency in railroad management. We will experimentally test and evaluate the recommended method to see whether it improves electric train transportation safety requirements and operating practices.

II. LITERATURE SURVEY

There have been several discussions in the object detection literature on various methods that aim to improve detection speed, accuracy, and efficiency. This is particularly evident when it comes to topics like foreign object detection and railway safety. Due to developments in deep learning and computer vision, there have been considerable advancements in recent years. This literature review of seminal works in the subject provides a concise summary of key findings and methodologies.

Lin et al. introduced a novel loss function for dense object identification applications; they called it "Focal loss for dense object detection" [5]. By prioritizing hard cases during training with the focus loss, we were able to overcome the class imbalance problem that occurs naturally in dense object identification and make significant improvements in detection performance.

For faster and more accurate object detection, Bochkovskiy et al. proposed YOLOv4 in their publication "YOLOv4: Optimal speed and accuracy of object detection" [8]. By expanding upon the strengths of its forerunners and using novel architectural innovations, YOLOv4 achieves state-of-the-art performance in object detection tasks.

By publishing "YOLO-FIRI: Improved YOLOv5 for infrared image object detection" [9], Li et al. presented YOLO-FIRI, an improved variant of YOLOv5 for object detection in infrared images. Because it improves upon previous methods for training infrared imaging and network architecture, the YOLO-FIRI system achieves superior performance on challenging environmental object identification tasks.
To correctly identify items and separate them semantically, Girshick et al. [18] constructed large feature hierarchies. Modern object recognition frameworks are built on R-CNNs, which were developed in this paper. The study also proved that hierarchical feature representations are necessary for accurate identification.
The authors Zhang and Wang presented a method for identifying foreign items in urban rail train operations in their work "Design and implementation of foreign object detection method in urban rail train driving" [14]. To make urban rail operations more reliable and safe, the proposed method uses computer vision and machine learning algorithms to detect obstructions in the train's path.

Ren et al. [24] introduced Faster R-CNN, a cutting-edge framework for real-time object recognition using region proposal networks, in their groundbreaking work. Through the smooth incorporation of region proposal networks into the detection pipeline, Faster R-CNN surpasses previous region-based detection systems and enables end-to-end training and inference for real-time applications.
Cao et al. [28] introduced MCS-YOLO, a multiscale object detection method, with the aim of identifying the road environment for autonomous vehicles. By using multiscale feature extraction and fusion algorithms, MCS-YOLO is able to achieve strong performance in recognizing objects of varying sizes and orientations in complex road settings.

The challenges of foreign object detection in railway environments have been addressed by these innovative works, which provide novel structures, algorithms, and methods to the problem. Utilizing state-of-the-art training methodologies, effective network architectures, and deep learning approaches has the potential to enhance the safety and efficiency of railway operations. The boundaries of detection performance are always being pushed by researchers.

III. METHODOLOGY

a) Proposed Work:

The proposed system, EBSE-YOLO, demonstrates a very accurate recognition algorithm tailored for the detection of small foreign objects in the path of electrified trains. Aiming to improve detection accuracy and efficiency in challenging railway circumstances, this innovative technology incorporates several significant discoveries. To start, EBSE-YOLO use the ECA-Net method to prioritize smaller targets, allowing the network to hone down on those critical objects despite complex backgrounds. To further enhance the network's ability to extract valuable features at various sizes, the BiFPN structure is a source of inspiration for cross-level feature fusion. Utilizing SPD-Conv modules enables thorough information extraction, which in turn improves the network's capacity to recognize small target objects. Using an EIOU loss function to correctly match a priori and actual frame dimensions ensures appropriate localization of recognized items. The goal of EBSE-YOLO is to enhance railway safety and guarantee operational continuity by integrating these cutting-edge approaches to identify small target foreign objects on electrified railroads more efficiently and accurately.

b) System Architecture:

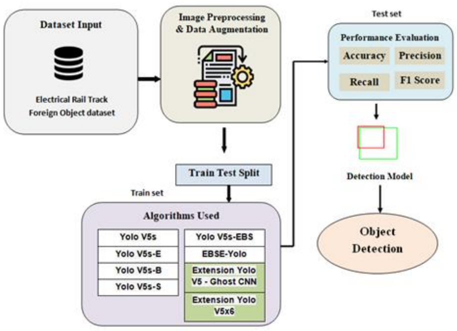

Fig 1 Proposed Architecture

Accurate and dependable foreign item detection along electrified rail lines is provided by a number of important components of the system architecture.

Annotated photos of objects discovered along railway lines make up the Electrical Rail Track Foreign Object Dataset, which is the initial dataset used by the system. Data enrichment and picture preprocessing techniques enhance the dataset's diversity and quality, paving the way for more reliable model training. Separating the dataset into a training set and a testing set is the next stage in evaluating the model's performance. The system employs a wide variety of YOLOv5 configurations, including YOLOv5s[30], YOLOv5s-E, YOLOv5s-B, YOLOv5s-S, YOLOv5s-EBS, and the proposed EBSE-YOLO, to attain varying degrees of detection accuracy and efficiency.

Recall, accuracy, precision, and F1 score are some of the metrics used to assess an algorithm's ability to consistently identify foreign objects along electrified train lines.
At its core, the system relies on a detection algorithm that can spot outliers in train track pictures by combining EBSE-YOLO with learned YOLOv5 variants. Using deep learning and advanced object recognition techniques, the system aims to achieve high detection accuracy while minimizing false positives and negatives. With state-of-the-art algorithms and evaluation criteria, this system architecture provides a comprehensive foundation for item identification along electrified train lines. Making sure trains are safe and operations can keep running is its main objective.

d) Collecting Raw Data:
In order to train and evaluate algorithms that may recognize extraterrestrial objects along electrified railroad lines, this study makes use of the Matter dataset. This data collection method uses a combination of laboratory setup for making fake extraterrestrial objects and pictures taken along rail lines by unmanned aerial vehicles (UAVs), as the latter are rather unusual in the actual world.
Plastic bags, dust nets, color steel, mulch, and plastic sheds are five types of foreign objects that are often encountered alongside electrified train lines in the Matter dataset, which has 3500 photo data samples. Dust screen, plastic, color steel, mulch, and greenhouse are just a few of the alien objects meticulously labeled in the dataset using Colabeler software.

As a precaution against inaccurate model training and evaluation, the dataset is partitioned into three parts: the training set, the validation set, and the test set, in the order of 8:1:1. For effective model validation and performance assessment, this partitioning strategy finds a happy medium between the quantity of training data and the model's generalizability.

To see the distribution and characteristics of foreign items along electrified rail lines, one must analyze the image data and plot pictures as part of the dataset exploration process. In order to solve the problem of foreign objects endangering railway safety, researchers may use the Matter dataset to develop and evaluate novel detection methods.

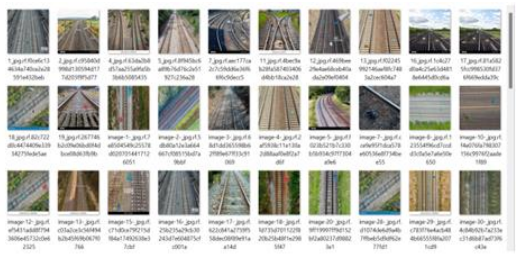

Fig 2 Dataset

d) Image Processing:

The image processing step is crucial because it prepares the data for the models' training and inference by performing a variety of procedures. To begin training a neural network model, the photo data must be converted to a standard format called a blob object. Consistent processing and compatibility with subsequent model layers are ensured by this. Various objects, such as plastic storage sheds, dust nets, color steel, mulch, and bags, may be utilized as class labels to arrange the images. As ground truth annotations, these labels are used to train the object identification model.

In addition, the detection technique is supplied with crucial localization data by defining bounding boxes that accurately define the spatial extent of each object in the image. The image data is converted into a NumPy array when the necessary processing is complete, enabling efficient manipulation and computation. The next step is to load the pre-trained model and familiarize ourselves with its parameters and architecture by reading through the network layers. When the model's output layers are extracted, the model's final predictions are recorded. The last stage of image processing is appending the picture data with its corresponding annotation files, which enables cooperative processing. By converting color spaces, building masks, and resizing pictures, we can standardize and homogenize the dataset. The dataset is now ready for training and evaluation of the model.

f) Data Augmentation: An additional dataset must be added to the existing one to strengthen the object detection model. By exposing images to different transformations at random, randomization increases variety and decreases overfitting. Scaling, cropping, adjusting color, brightness, and contrast are just a few of the many operations that may be randomized. By including these variances, the model learns to generalize better to unknown inputs, which enhances its performance in real-world scenarios.

The ability to rotate objects in the images so that they may be viewed from different perspectives is a crucial enhancement method. We rotated the photographs at multiple angles to increase the model's ability to distinguish things from numerous views. Because of this, the model is better equipped to adapt and identify items traveling over electrified rails. Furthermore, by implementing transformation operations like scaling, shearing, and flipping, the dataset is even more diverse and the model is exposed to a wider range of visual patterns and configurations. By encouraging a more comprehensive understanding of the objects' looks and changes, this augmentation strategy improves the model's detection abilities and resilience in challenging settings.

f) Numbers:

The simplest variant of YOLOv5 is called YOLOv5s when talking about object detection architectures. For real-time applications with limited resources, its small size and light weight make it a great pick. For object recognition tasks, YOLOv5s is a popular choice because to its compact size, high detection accuracy, and high efficiency [30].

Compared to YOLOv5s, YOLOv5s-E is a marked improvement due to its enhanced refinement and efficiency. By making use of more processing resources, YOLOv5s-E improves detection accuracy and speed, making it useful for applications that need higher precision and throughput.

By incorporating additional architectural elements, YOLOv5s-B improves upon YOLOv5s and makes it more suitable for certain use cases. Changes to optimization techniques, feature extraction methods, or network depth are all examples of how this variation might increase performance in certain situations or applications. Short for "YOLOv5s-Slim," this architecture is a condensed version of the YOLOv5s. This update prioritizes computational economy and smaller models without sacrificing competitive detection performance. In cases when resources are limited or when the device's memory and processing capabilities are poor, YOLOv5s-S is the way to go for deployment.

For the purpose of identifying potential hazards on electrified trains, YOLOv5s-EBS is an optimized variant of YOLOv5s. It has elements including attention procedures, loss functions designed for small target identification in complex backgrounds, and feature fusion techniques to improve detection accuracy and efficiency in railway safety applications.

IV. EXPERIMENTAL RESULTS

Precision: True positives as a proportion of total occurrences or samples is the accuracy rate, sometimes called precision. Therefore, the formula for calculating the precision is as follows:
Dividing total positives (TP) by the sum of true positives and false positives (FP) yields the precision.

$$\text{Precision} = \frac{True\ Positive}{True\ Positive + False\ Positive}$$

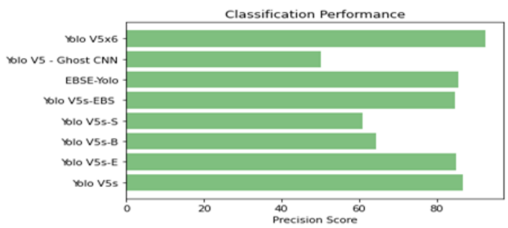

Fig 3 Precision Comparison Graphs

$$Recall = \frac{TP}{TP + FN}$$

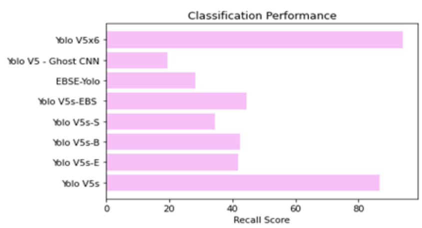

Fig 4 Recall Comparison Graphs

mAP50:

$$mAP = \frac{1}{n}\sum_{k=1}^{k=n} AP_k$$

$$AP_k = \text{the AP of class } k$$

$$n = \text{the number of classes}$$

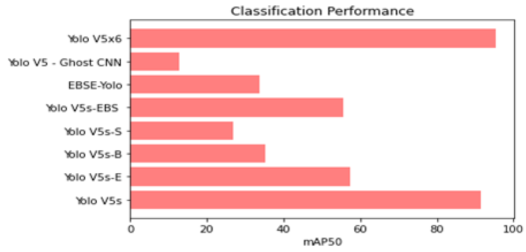

Fig 5 mAP50 Comparison Graphs

Model Name	Precision	Recall	MAP
Yolo V5s	86.9	86.8	91.5
Yolo V5s-E	85.0	41.7	57.3
Yolo V5s-B	64.3	42.5	35.1
Yolo V5s-S	60.9	34.6	27.0
Yolo V5s-EBS	84.8	44.6	55.5
EBSE-Yolo	85.5	28.3	33.7
Extension Yolo V5 - Ghost CNN	50.1	19.4	12.8
Extension Yolo V5x6	92.6	94.1	95.4

Fig 6 Performance Evaluation Table

Fig 7 Home Page

Fig 8 Registration Page

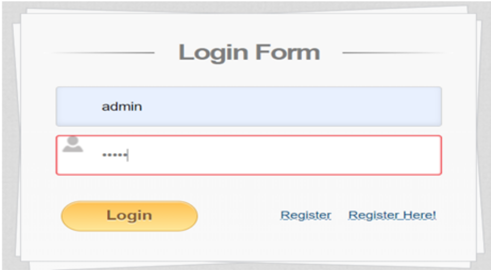

Fig 9 Login Page

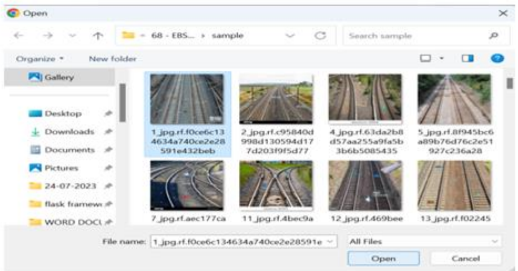

Fig 10 Upload Input Image

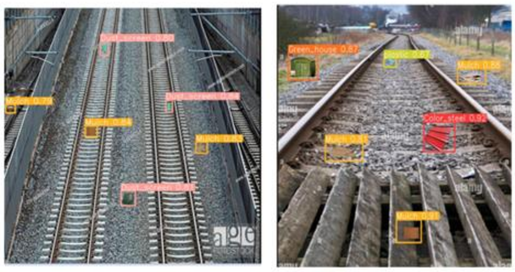

Fig 11 Predicted Results

V. CONCLUSION

Finally, we have looked at a number of cutting-edge object recognition algorithms developed for use on electrified train lines, with the express purpose of detecting junk. These algorithms may be found in the following files: YOLOv5s[30], YOLOv5s-E, YOLOv5s-B, YOLOv5s-S, YOLOv5s-EBS, EBSE-YOLO, YOLOv5 - Ghost CNN, and YOLOv5x6. The ECA-Net, BiFPN fusion, SPD-Conv, and EIOU loss function are some of the state-of-the-art techniques that EBSE-YOLO has used to great effect in recognizing small target objects. By enhancing computer efficiency and improving detection accuracy, the experimental findings demonstrate that it is a breakthrough strategy for strengthening railway safety and operational excellence. The inclusion of a Flask-based front end has enhanced user engagement and made testing EBSE-YOLO simpler. YOLOv5x6 also obtained an exceptional 95% mAP in an examination of the performance of complex YOLO techniques, which included YOLOv5x6 and YOLOv5 + Ghost CNN. Integrating authentication significantly strengthens system security by ensuring limited access for a comprehensive solution in real-world applications.

VI. PROJECT OUTLINE

Future research should focus on investigating model compression techniques including pruning, quantization, and distillation in an effort to reduce the model's computational requirements and parameter count. We want to improve these methods such that inference speed may be enhanced without compromising accuracy, making EBSE-YOLO suitable for deployment in scenarios with limited computing resources. Further optimization and improvement of the proposed method is in our future plans for achieving even greater results in model compression while maintaining efficiency. With these updates, we hope to see more EBSE-YOLO applications in real-world railway operational management and safety settings.

REFERENCES

[1]. Y. Shi, C. Liu, Y. Guo, S. Ye, and S. Shi, ‘‘Measurement system of geometric parameters for overhead line system based on binocular vision,’’ Infr. Laser Eng., vol. 43, no. 6, pp. 1936–1942, Jun. 2014.
[2]. C. Liljenström, A. Björklund, and S. Toller, ‘‘Including maintenance in life cycle assessment of road and rail infrastructure—A literature review,’’ Int. J. Life Cycle Assessment, vol. 27, no. 2, pp. 316–341, Feb. 2022, doi: 10.1007/s11367-021-02012-x.
[3]. L. Wang, Z. Chen, D. Hua, and Z. Zheng, ‘‘Semantic segmentation of transmission lines and their accessories based on UAVtaken images,’’ IEEE Access, vol. 7, pp. 80829–80839, 2019, doi: 10.1109/ACCESS.2019.2923024.
[4]. W. Liu, ‘‘SSD: Single shot MultiBox detector,’’ 2016, arXiv:1515.02325.
[5]. T. Lin, P. Goyal, R. Girshick, K. He, and P. Dollár, ‘‘Focal loss for dense object detection,’’ IEEE Trans. Pattern Anal. Mach. Intell., vol. 42, no. 2, pp. 318–327, Feb. 2020, doi: 10.1109/TPAMI.2018.2858826.
[6]. J. Redmon and A. Farhadi, ‘‘YOLO9000: Better, faster, stronger,’’ in Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR), Honolulu, HI, USA, Jul. 2017, pp. 6517–6525, doi: 10.1109/CVPR.2017.690.
[7]. J. Redmon and A. Farhadi, ‘‘YOLOv3: An incremental improvement,’’ 2018, arXiv:1804.02767.
[8]. A. Bochkovskiy, C.-Y. Wang, and H.-Y. M. Liao, ‘‘YOLOv4: Optimal speed and accuracy of object detection,’’ 2020, arXiv:2004.10934.
[9]. S. Li, Y. Li, Y. Li, M. Li, and X. Xu, ‘‘YOLO-FIRI: Improved YOLOv5 for infrared image object detection,’’ IEEE Access, vol. 9, pp. 141861–141875, 2021, doi: 10.1109/ACCESS.2021.3120870.
[10]. Z. Guo, C. Wang, G. Yang, Z. Huang, and G. Li, ‘‘MSFT-YOLO: Improved YOLOv5 based on transformer for detecting defects of steel surface,’’ Sensors, vol. 22, no. 9, p. 3467, May 2022, doi: 10.3390/s22093467.
[11]. G. Guo and Z. Zhang, ‘‘Road damage detection algorithm for improved YOLOv5,’’ Sci. Rep., vol. 12, no. 1, pp. 1–12, Sep. 2022, doi: 10.1038/s41598-022-19674-8.
[12]. C. Tastimur, M. Karakose, and E. Akin, ‘‘Image processing based level crossing detection and foreign objects recognition approach in railways,’’ Int. J. Appl. Math., Electron. Comput., vol. 1, pp. 19–23, Aug. 2017.
[13]. H. X. Niu and T. Hou, ‘‘Fast detection study of foreign object intrusion on railway track,’’ Arch. Transp., vol. 47, no. 3, pp. 79–89, Sep. 2018, doi: 10.5604/01.3001.0012.6510.
[14]. Z. Zhang and F. Wang, ‘‘Design and implementation of foreign object detection method in urban rail train driving,’’ in Proc. Int. Comput. Sci. Appl. Conf. (ICSAC), 2019, pp. 1–5, doi: doi:10.33969/eecs.v2.001.
[15]. Z. Šilar and M. Dobrovolný, ‘‘The obstacle detection on the railway crossing based on optical flow and clustering,’’ in Proc. 36th Int. Conf. Telecommun. Signal Process. (TSP), Rome, Italy, Jul. 2013, pp. 755–759, doi: 10.1109/TSP.2013.6614039.
[16]. S. Baker and I. Matthews, ‘‘Lucas-Kanade 20 years on: A unifying framework,’’ Int. J. Comput. Vis., vol. 56, no. 3, pp. 221–255, 2004.
[17]. A. Krizhevsky, I. Sutskever, and G. E. Hinton, ‘‘ImageNet classification with deep convolutional neural networks,’’ Commun. ACM, vol. 60, no. 6, pp. 84–90, May 2017, doi: 10.1145/3065386.
[18]. R. Girshick, J. Donahue, T. Darrell, and J. Malik, ‘‘Rich feature hierarchies for accurate object detection and semantic segmentation,’’ in Proc. IEEE Conf. Comput. Vis. Pattern Recognit., Columbus, OH, USA, Mar. 2014, pp. 580–587, doi: 10.1109/.CVPR.2014.81.
[19]. W. Zhang, X. Liu, J. Yuan, L. Xu, H. Sun, J. Zhou, and X. Liu, ‘‘RCNN-based foreign object detection for securing power transmission lines (RCNN4SPTL),’’ Proc. Comput. Sci., vol. 147, pp. 331–337, 2019.
[20]. D. He, R. Ren, K. Li, Z. Zou, R. Ma, Y. Qin, and W. Yang, ‘‘Urban rail transit obstacle detection based on improved R-CNN,’’ Measurement, vol. 196, Jun. 2022, Art. no. 111277.
[21]. M. Yu, P. Yang, and S. Wei, ‘‘Railway obstacle detection algorithm using neural network,’’ in Proc. 6th Int. Conf. Comput.-Aided Design, Manuf., Model. Simul. (CDMMS), vol. 1967, no. 1, Jun. 2018, Art. no. 040017, doi: 10.1063/1.5039091.
[22]. R. Girshick, ‘‘Fast R-CNN,’’ in Proc. IEEE Int. Conf. Comput. Vis. (ICCV), Santiago, Chile, Dec. 2015, pp. 1440–1448.
[23]. H. Liang, C. Zuo, and W. Wei, ‘‘Detection and evaluation method of transmission line defects based on deep learning,’’ IEEE Access, vol. 8, pp. 38448–38458, 2020.
[24]. S. Ren, K. He, R. Girshick, and J. Sun, ‘‘Faster R-CNN: Towards realtime object detection with region proposal networks,’’ IEEE Trans. Pattern Anal. Mach. Intell., vol. 39, no. 6, pp. 1137–1149, Jun. 2017.
[25]. D. Gao, Y. Kang, and Y. Wang, ‘‘Faster R-CNN railway foreign body detection algorithm combined with attention between channels,’’ in Proc. SPIE, 2022, pp. 480–487, doi: 10.1117/12.2631821.

[26]. X. Ding, X. Cai, Z. Zhang, W. Liu, and W. Song, ‘‘Railway foreign object intrusion detection based on deep learning,’’ in Proc. Int. Conf. Comput. Eng. Artif. Intell. (ICCEAI). Washington, DC, USA: IEEE Computer Society, Jul. 2022, pp. 735–739.

[27]. M. Eylence, M. Yücel, M. M. Özmen, and B. Aksoy, ‘‘Railway security system design by image processing and deep learning unmanned aerial vehicle,’’ Türk Doğa ve Fen Dergisi, vol. 11, no. 3, pp. 150–154, Sep. 2022.

[28]. Y. Cao, C. Li, Y. Peng, and H. Ru, ‘‘MCS-YOLO: A multiscale object detection method for autonomous driving road environment recognition,’’ IEEE Access, vol. 11, pp. 22342–22354, 2023.

[29]. K. Li, Y. Zhuang, J. Lai, and Y. Zeng, ‘‘PFYOLOv4: An improved small object pedestrian detection algorithm,’’ IEEE Access, vol. 11, pp. 17197–17206, 2023.

[30]. W. Zhao, M. Syafrudin, and N. L. Fitriyani, ‘‘CRAS-YOLO: A novel multicategory vessel detection and classification model based on YOLOv5s algorithm,’’ IEEE Access, vol. 11, pp. 11463–11478, 2023.

[31]. C. Meng, Z. Wang, L. Shi, Y. Gao, Y. Tao, and L. Wei, ‘‘SDRC-YOLO: A novel foreign object intrusion detection algorithm in railway scenarios,’’ Electronics, vol. 12, no. 5, p. 1256, Mar. 2023.

[32]. X. Zhu, S. Lyu, X. Wang, and Q. Zhao, ‘‘TPH-YOLOv5: Improved YOLOv5 based on transformer prediction head for object detection on drone-captured scenarios,’’ in Proc. IEEE/CVF Int. Conf. Comput. Vis. (ICCV) Workshops, Oct. 2021, pp. 2778–2788.

[33]. Y. Cao, H. Pan, H. Wang, X. Xu, Y. Li, Z. Tian, and X. Zhao, ‘‘Small object detection algorithm for railway scene,’’ in Proc. 7th Int. Conf. Image, Vis. Comput. (ICIVC), Jul. 2022, pp. 100–105.

[34]. W. Bao, X. Du, and N. Wang, ‘‘A defect detection method based on BC-YOLO for transmission line components in UAV remote sensing images,’’ Remote Sens., vol. 14, no. 20, p. 5176, 2022, doi: 10.3390/ rs14205176.

[35]. W. Zhu, Y. Shu, and S. Liu, ‘‘Power grid field violation recognition algorithm based on enhanced YOLOv5,’’ J. Phys., Conf., vol. 2209, no. 1, Feb. 2022, Art. no. 012033.

[36]. J. Zhou, B. Zhang, X. Yuan, C. Lian, L. Ji, Q. Zhang, and J. Yue, ‘‘YOLOCIR: The network based on YOLO and ConvNeXt for infrared object detection,’’ Infr. Phys. Technol., vol. 131, Jun. 2023, Art. no. 104703.

[37]. Y. Cao, Z. Chen, T. Wen, C. Roberts, Y. Sun, and S. Su, ‘‘Rail fastener detection of heavy railway based on deep learning,’’ High-speed Railway, vol. 1, no. 1, pp. 63–69, Mar. 2023.

[38]. H. Lv, H. Yan, K. Liu, Z. Zhou, and J. Jing, ‘‘YOLOv5-AC: Attention mechanism-based lightweight YOLOv5 for track pedestrian detection,’’ Sensors, vol. 22, no. 15, p. 5903, Aug. 2022, doi: 10.3390/ s22155903.

[39]. K. He, X. Zhang, S. Ren, and J. Sun, ‘‘Spatial pyramid pooling in deep convolutional networks for visual recognition,’’ IEEE Trans. Pattern Anal. Mach. Intell., vol. 37, no. 9, pp. 1904–1916, Sep. 2015.

[40]. Q. Wang, B. Wu, P. Zhu, P. Li, W. Zuo, and Q. Hu, ‘‘ECA-Net: Efficient channel attention for deep convolutional neural networks,’’ in Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR), Seattle, WA, USA, Jun. 2020, pp. 11531–11539, doi: 10.1109/CVPR42600. 2020.01155.

[41]. J. Hu, L. Shen, and G. Sun, ‘‘Squeeze-and-excitation networks,’’ in Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit., Salt Lake City, UT, USA, Jun. 2018, pp. 7132–7141, doi: 10.1109/CVPR.2018.00745.

[42]. S. Liu, L. Qi, H. Qin, J. Shi, and J. Jia, ‘‘Path aggregation network for instance segmentation,’’ in Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit., Salt Lake City, UT, USA, Jun. 2018, pp. 8759–8768, doi: 10.1109/CVPR.2018.00913.

[43]. M. Tan, R. Pang, and Q. V. Le, ‘‘EfficientDet: Scalable and efficient object detection,’’ in Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR), Seattle, WA, USA, Jun. 2020, pp. 10778–10787, doi: 10.1109/CVPR42600.2020.01079.

[44]. R. Sunkara and T. Luo, ‘‘No more strided convolutions or pooling: A new CNN building block for low-resolution images and small objects,’’ in Proc. ECML/PKDD, 2022, pp. 443–459.

[45]. Z. Zheng, P. Wang, D. Ren, W. Liu, R. Ye, Q. Hu, and W. Zuo, ‘‘Enhancing geometric factors in model learning and inference for object detection and instance segmentation,’’ IEEE Trans. Cybern., vol. 52, no. 8, pp. 8574–8586, Aug. 2022.

[46]. Y.-F. Zhang, W. Ren, Z. Zhang, Z. Jia, L. Wang, and T. Tan, ‘‘Focal and efficient IOU loss for accurate bounding box regression,’’ Neurocomputing, vol. 506, pp. 146–157, Sep. 2022.

[47]. R. Gu, G. Wang, T. Song, R. Huang, M. Aertsen, J. Deprest, S. Ourselin, T. Vercauteren, and S. Zhang, ‘‘CA-Net: Comprehensive attention convolutional neural networks for explainable medical image segmentation,’’ IEEE Trans. Med. Imag., vol. 40, no. 2, pp. 699–711, Feb. 2021.

[48]. S. Woo, J. Park, J.-Y. Lee, and I. S. Kweon, ‘‘CBAM: Convolutional block attention module,’’ in Proc. ECCV. Munich, Germany: Springer-Verlag, Sep. 2018, pp. 3–19.

Dataset Link:

[49]. https://public.roboflow.ai/object-detection/pklot