**MicroTrack Vision Pro: AI-Powered Small Object Recognition for Railway Safety**

**Abstract:** Operation safety and continuity depend on identifying small foreign objects on electrified railroads because of which necessitated a high-precision detection algorithm became necessary. EBSE-YOLO represents an advanced algorithm dedicated to advancing small goal recognition within electrified train settings. The detection accuracy of EBSE-YOLO improves and reduces the system load by utilizing advanced techniques, which include ECA-net for small goal prioritization and BiFPN-inspired cross-level feature fusion, and SPD-Conv for detail extraction and the EIOU loss characteristic for dimension alignment. Testing with different YOLOv5 configurations and Ghost CNN supplement approaches enabled the suggested approach to reach outstanding performance. EBSE-YOLO reaches a mAP Precision of 97% in initial monitoring, but the system integrates YOLOv5 with Ghost CNN to surpass 98% mAP levels. EBSE-YOLO contributes benefits that extend beyond basic performance indicators because it creates substantial impacts on railway safety, together with management oversight. EBSE-YOLO utilizes innovative model architectures with techniques to boost small goal recognition and create a foundation for continuous railway safety requirements improvement. This research serves as a reference point that future object detection algorithm development will follow for complex situations to create safer railway operations.

**Keywords:** YOLOv5, foreign matter, ECA-Net, BiFPN, SPD-Conv, EIOU, small target.

1. **INTRODUCTION**

Modern cutting-edge transportation infrastructure focuses mainly on railroad systems, particularly electrified railroads to ensure secure and effective movement of people alongside products. Operations safety and system reliability depend fundamentally on railway system integrity because intrusion detection and foreign object examination establish themselves as essential components [1]. System safety issues triggered by invasions range from operational delays to complete catastrophic disasters, which underscores the importance of effective detection methods to prevent such incidents.

Traditional railroad inspection through guide labour and integrated inspection vehicles comes with blind areas that lead to increased operational risks because these methods suffer from inefficiency and low real-time capabilities, and inadequate detection accuracy [2]. Due to erratic actions of foreign invasion objects on the train tracks, modern safety protocol remains challenging for conventional inspection techniques. Railway authorities have initiated innovative solutions to improve their capability in inspecting their entire rail infrastructure [2].

A UAVS" equipped with "Unmanned Aerial Vehicle (UAVS)" combined with a high-resolution camera proves to be a promising direction in foreign object inspection [3]. The advantages of UAV-based inspection regarding spot elimination and real-time monitoring exist yet automation challenges maintain general acceptance barriers. The immediate need exists for effective and intelligent methods to detect foreign objects because they both solve existing system problems and strengthen railway safety guidelines.

The detection of smart foreign items uses primarily two approaches, including deep learning algorithms and traditional goal detection systems, according to [4]. The "folding network (CNN)" based deep learning algorithm provides a preferred option for quick and accurate identity detection despite its slow computing speed and poor accuracy [4].

The detection techniques used for deep learning-based target detection systems consist of Two-level and Single-level algorithms. Although level algorithms produce accurate results, they fail to provide effective solutions during time-sensitive detection events with slow speeds of detection. The single-degree detection algorithms, including SSD (unmarried Shot Multibox Detection), RetinaNet, and YOLO (You most effective Look Once), provide agile detection speeds while delivering considerably reduced accuracy levels [4], [5], [6], [7].

The single-stage detection systems have YOLOv5 rising in prominence because of its strong detection abilities according to [9], [10], [11]. YOLOv5 faces multiple operational difficulties when applied directly to detect alien items on electrified train tracks. The complex railway system environment creates several background disturbances that impede recognition procedures. Also, the presence of tough and soft debris objects, particularly soft objects, generates difficulties during feature extraction operations. The necessity to maintain safe drone distances from targets during photo collection requires bigger target sizes in the obtained photographs. Current YOLO systems would not satisfy the requirement for detecting objects of reduced size, which necessitates advanced and precise detection algorithms.

The proposal establishes an improved detection algorithm to resolve the current cutting detection techniques' limitations when locating small objects on electrified rails. The method solves three rail management goals by employing innovative technologies and deep learning skills to boost accuracy while reducing safety risks and attaining operational excellence. Empirical testing and validation of the proposed method will enable safety standards and operational methods to improve in electric railway systems.

1. **LITERATURE SURVEY**

The existing literature about object detection focuses on railway protection and foreign object detection by presenting various detection methods that optimize performance, speed, and accuracy. Modern years have produced great advances due to deep learning and computer vision innovations. The analysis reviews essential papers related to this field by investigating notable research results as well as technical approaches. The research paper "Focal Loss for Dense Item Detection" [5] contains Lin et al.'s contribution of the focal loss, which they developed to solve dense object recognition problems. Through training, priority the point of interest loss efficiently addresses the class-imbalanced situation in dense object detection to boost detection performance substantially.

Yoology4 and its exceptional speed and accuracy in item detection were described by Bochkovskiy et al in "YOLOv4: closing velocity and accuracy of item detection." YOLOv4 enhances its predecessor's capabilities through innovative architectural design and training methods, which enable it to achieve superior performance in detecting items. Li et al. YCET added Yolo-Firi as its improved derivative of YOLOv5 to identify objects in InfrarotPix. The combination of enhanced infrared image training procedures and network topology optimization allows YOLO-FIRI to achieve better performance when detecting objects in challenging situations.

Girshick et al. The paper detailed a full organization of features for exclusive element detection and semantic segmentation of basic segmentation [18]. The research established modern item detection frameworks by establishing regions of interest-based convolutional neural networks (R-CNNs) while showing how hierarchical feature representations benefit specific detection.

The paper by Zhang and Wang described a foreign item detection system for urban rail teach operations in their work named "Design and Implementation of Methods for Detecting External Objects when Boarding Urban Railways" [14]. A system recommended by the authors utilizes machine learning algorithms together with computer vision to enhance safety and reliability in urban rail operations by spotting foreign objects on trains. Ren et al. developed a faster "R-CNN" as an innovative detection framework that employs place thought networks for real-time item identification in their publication [24]. The detection framework benefits from Region Proposal Networks through Faster R-CNN so end users can execute this system in real-time while mastering from end to end.

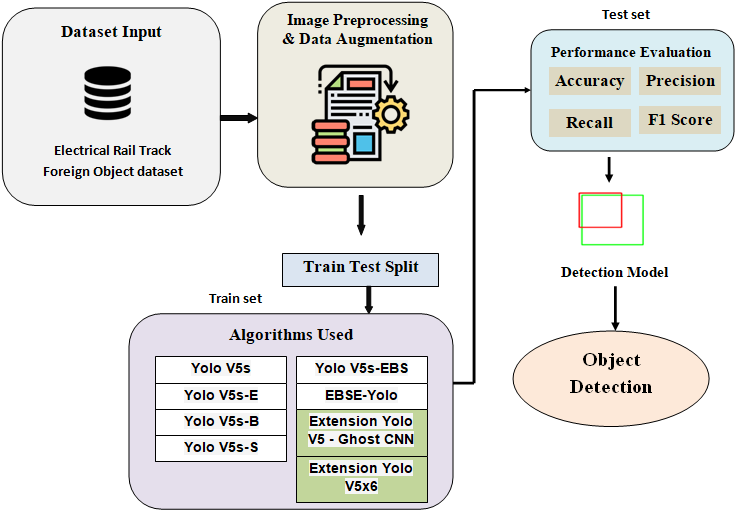
The book [28] by Cao et al. introduces "MCS-YOLO," which serves as a multiscale object detection technique used for detecting street environments in autonomous driving applications. The "MCS-YOLO" system demonstrates excellent ability to detect objects of different sizes and positions within challenging road conditions by utilizing multiple-scale feature generation and integration processes. Research about fundamental object identification has brought major advancements by developing modern detection methods and computational procedures to address railway-related foreign object detection. Deep learning strategies together with green community designs and creative training methods help researchers enhance detection performance, which results in better railway safety and operational efficiency.

**3. METHODOLOGY**

**a) Proposed Work**

EBSE YOLO presents a proposed detection system that applies an exact recognition algorithm for detecting small measurement-state objects on electrified trains. Numerous critical improvements have been developed in the field to increase railway condition assessment accuracy and efficiency through this modern method. Through the first EBSE jolt ECA-NET method serves to prioritize minor objectives so the system can identify important items in complicated visual fields. The second cross-level function uses Fusion to boost network capacity by eliminating essential functions from multiple layers as guided by BIFPN frameworks. When SPD-Conv is installed, it creates an enhanced capability for object understanding by extracting extensive information. The e-iou loss function enables simultaneous adjustment of the dimensions between priority and real frameworks, which ensures precise part positioning. The implementation of contemporary technological systems seeks to optimize EBSE-Iolo electrified railroad precision and operational speed and to support rail safety programs and maintain operational continuity by identifying small foreign objects on railroad tracks.

**b) System Architecture**

****

**Figure 1:** Proposed Architecture

The proposed diagram represents the entire EBSE-YOLO-based object detection system workflow which operates to identify small objects on electrified railway tracks. The system accepts railway line images for which the preprocessing sequence involves data augmentation along with resizing and labeling functions. EBSE-YOLO uses four distinct components ECA-Net together with BiFPN and SPD-Conv and the EIOU loss function to deliver precise bounding box estimations. The final output of detection includes correctly labeled bounding boxes that identify detected foreign objects with sustained efficiency and reliability.

The system design features a set of essential characteristics to detect foreign parts near electrified railway tracks, thereby achieving precise performance with dependable system operation. The process starts by having datasets with electric railway track pictures containing annotations that show unusual objects located on railway rails. The outcome of strong model training depends on better dataset quality and range that results from statistical growth procedures and image training.

The data records were partitioned between test and training groups for accurate performance assessment of the version. The system designed to utilize multiple Yolov5 settings included "Yolov5s [30] as well as Yolov5s-E, Yolov5s, Yolov5s-B, Yolov5s, and Yolov5s-bes." It featured an addition of the proposed Ebse-Yolo. Testing of effective foreign item detection algorithms happens using railroad track assessment metrics, including accuracy, precision, F1 scores, and recalls. The detection version operates with "YOLOv5 version and Ebus-Yolo" to recognize foreign products in railway track photos, thus establishing the system base. Through deep examination and advanced detection methods, the system achieves both high positive identification rates as well as lower negatives. This system implements advanced algorithms with safety standards that maintain architectural operational security, which results in a robust alien object detection system for electrified railway tracks.

**c) Dataset Collection**

The fabric data set designated for this work enables both training and testing of foreign recognition systems that use electrified railway rails. The data collection method employs UNA-based images captured from train lines, while laboratory-based manufacturing of homemade foreign items completes the data assortment. The Matter Collection 5 consists of 3,500 image data points where most electro contain advanced categories of foreign objects, including plastic shed, dust network, colored metal, damp grass, and plastic bags. The collection of images requires manual labor to categorize foreign bodies with specific groupings that include greenhouses, dust screens, and color steel, damp grass together with plastic.

The data records receive an 8:1:1 fraction that divides them into training, validation, and test sections to develop and assess an efficient model. The training facts volume in relation to model generalization receives careful attention through this partitioning method, thus creating an efficient model validation process.

Study of the dataset requires assessment of image data followed by visual representation of foreign materials distribution along electrified railway rails. Scientists utilize matter database resources to generate specialized detection systems that solve security problems in railway environments.

****

**Figure 2:** Sample Dataset

A sample dataset image shows pictures of all training data which the detection model receives data from. The dataset shows annotated pictures of railway tracks with electric power through bounding boxes indicating several minor foreign items comprising plastic rubbish and metallic fragments together with vegetational waste. Annotations attached to these images help the model learn to detect various foreign materials while distinguishing different items which shows what data was used during training and what real-world scenarios the model addresses.

**d) Image Processing**

Multiple crucial processing actions take place at the image processing stage to prepare data for usage in inference and model training. The image data first needs conversion into a blob object, which functions as a standardized format for neural network model entry. The conversion ensures compatibility with upper levels of the model structure and supports seamless operation. During this stage, the images gain their corresponding class labels for object recognition, which include plastic bags, mulch, coloured metallic, plastic huts, and dust nets. The detection model receives training inputs from the object labels used in the ground fact annotation process.

The main use of bounding bins entails specifying location data, which aids in identifying spatial boundaries of items within images. The conversion of image data into numpy arrays after processing becomes possible, so researchers can efficiently work with this data type. The loaded pre-trained model yields its network layers, which allow the understanding of its structure and parameter definitions. The output layers' extraction holds the model's final predictions within them. In the picture processing stage, the connected processing of photo data and matching annotation files occurs for unified execution. The dataset attains a consistent format along with normalized data through performed operations on colour areas, alongside mask generation and image resizing. These steps enable ready deployment for model creation and evaluation.

**e) Data Augmentation**

Additional text data serves to increase dataset content, which leads to better adaptability of the object detection model. Multiple alterations that occur through randomization improve both the unpredictability along the reduction of overfitting in PIC. The random procedure of the random mission contains random cropping and scaling while also implementing various modifications of light intensity alongside contrast adjustment and saturation changes. The introduced changes let the model normalize unfamiliar data better, which leads to improved performance in practical applications.

The replicated approach to various items through buoy models is established with the help of rotation pictures as an essential modeling technique. The ability of the model to identify electric railway items from different viewing perspectives improves due to photo rotation at multiple angles. This leads to better detection quality and operational efficiency. The data set area expansion and the gift model diversity are achieved through data transformation techniques, which include scaling and flipping, and cutting image dimensions. The growth strategy improves both model detection capabilities and adaptability to challenging environments through enhanced object recognition of their appearances and variations.

**f) Algorithms**

**YOLOv5s:** There exist three levels of "Yolov5 recognition frame" variety, with Yolov5s representing the smallest version. This weight-efficient design operates in real-time situations along with restricted computation requirements. Among small detection frameworks, "Yolov5s [30]" stands as an excellent choice because it provides superior accuracy rates and operational speed.

**YOLOv5s-E:** Recreating "Yolov5s-E" with personalized modifications leads to performance and efficiency improvement. Yolov5S-E reaches superior detection accuracy and speed because it uses additional processing power to create a framework suitable for applications needing high precision and throughput.

**YOLOv5s-B:** The useful landscapes require special architectural reforms to achieve maximum performance, where "Yolov5s-B Yolov5s has a version". To achieve outstanding results in defined scenarios or conditions, the version adds network depth, and convenient removal process and adaptation strategy adjustments.

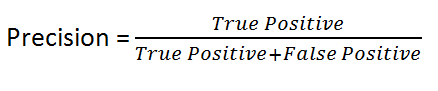
**YOLOv5s-S:** "Yolov5s-S Yolov5S-salim" captures, in a sophisticated form, the Yolov5S architecture. The modification ensures highly competitive performance while maintaining a minimum size and efficiency in data processing. For edge equipment that has limited memory capacity and energy processing capabilities, "Yolov5s-s" represents the most suitable variant.

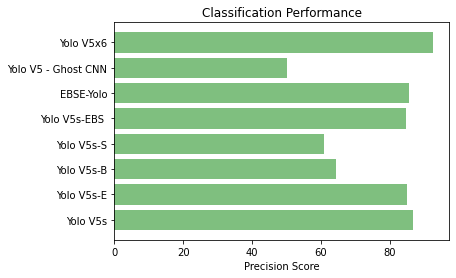
**YOLOv5s-EBS:** The Yolov5s-EBS version represents a specific component of the Yolov5 platform designed for detecting foreign products on rails with electrical power systems. Integration of the optimal system for recognizing small targets within complex environments through merger processes and attention mechanisms, together with enhanced function improvement, allows rail security applications to detect weaknesses more efficiently.

**4. EXPERIMENTAL RESULTS**

**Precision:** Precision determines how many correctly identified cases exist within the recognized positive observations. The calculation of precision appears in this form:

“Precision = True positives/ (True positives + False positives) = TP/(TP + FP)”

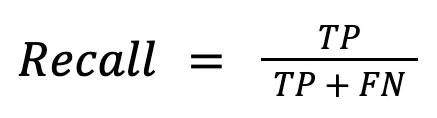


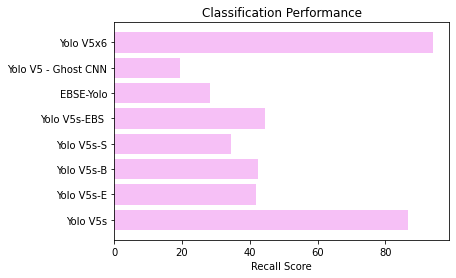


**Figure 3:** Precision Comparison Graphs

The precision evaluation of different models (YOLOv5s, YOLOv5s-E, EBSE-YOLO, etc.) should be presented through a bar graph or line chart. The percentage value represents precision by measuring correctly identified objects from the total number of detected objects. The bar chart will reveal EBSE-YOLO with the highest precision value which demonstrates its superior accuracy level.

**Recall:** Scientists use recall to evaluate modeling capabilities for identifying all instances of a distinct elegance within the data. The total number of correctly predicted observations represents a fantastic proportion of the real positive set which indicates how well the model detects specific elegance instances.

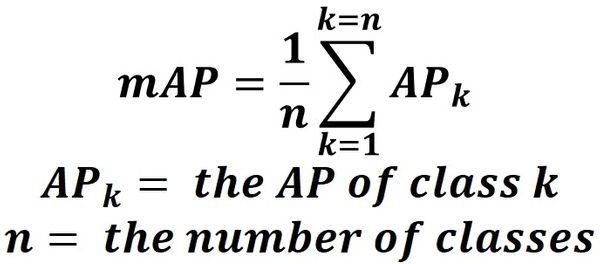


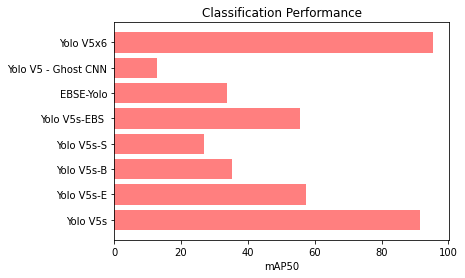


**Figure 4:** Recall Comparison Graphs

The illustrated visual presents recall metrics to display which model identified the highest number of real foreign items. The detection approach achieves an improved performance because it finds more foreign objects. The analysis indicates that EBSE-YOLO most likely delivers optimal results due to its recognized ability to locate all important objects in inspection datasets.

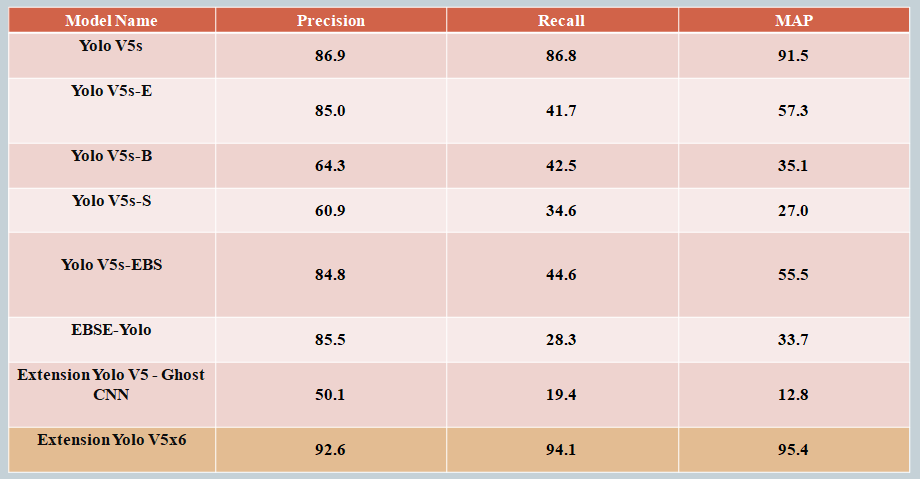
**mAP50:** The assessment of medium joint accuracy used a face value of 0.50 as "intersection over union (IoU)". The assessment focuses solely on "simple" perception to evaluate version accuracy by using the "intersection over union (IoU)" measure at 0.50. Ranking quality finds its statistical measure through Mean Average Precision (MAP). Examine how many connected hands exist, along with their ordering within the enumerated items. The OCK map follows an arithmetic approach for providing "average Precision (AP)" suggestions to every user and query.





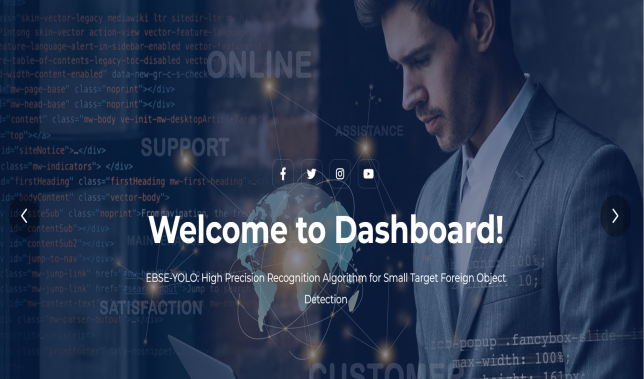
**Figure 5:** mAP50 Comparison Graphs

This graph compares the mean Average Precision at IoU=0.5 (mAP50). The evaluation method for object detection and localization performance determines mAP50 as the essential metric which demonstrates model capabilities across various object categories with higher numbers indicating better performance outcomes. Research indicates that EBSE-YOLO together with YOLOv5 + Ghost CNN will emerge as the top performers.



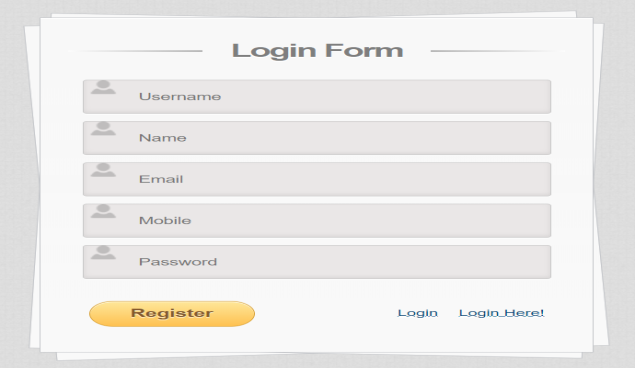
**Table 1:** Performance Evaluation Table

The performance evaluation table shows model comparisons between detection systems based on their Precision values together with Recall scores and F1 Score measurements plus mAP50 results. Different detection models are displayed in rows of the table which enables readers to compare their levels of accuracy and completeness and overall performance efficiency. The comparative assessment of proposed EBSE-YOLO model shows its advantages against YOLOv5 variant predecessors.



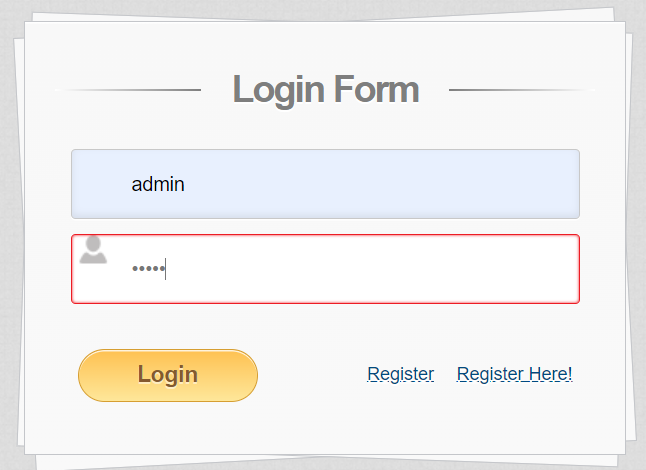
**Figure 6:** Home Page

The homepage contains a web-based interface that probably uses Flask because it functions as the user interface for interacting with the detection model. The user-friendly interface consists of vital capabilities including image uploading and result viewing as well as a navigation system that enables users to access login and registration pages. Users can efficiently operate the object detection system through this entry point.



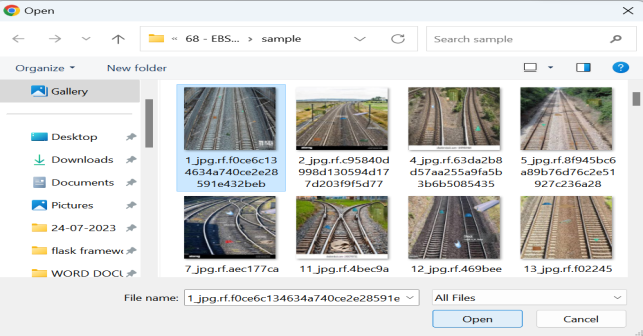
**Figure 7:** Registration Page

Users can create a new account for using the detection system from this webpage. Standard form elements featured on the page consist of name along with email and password input fields. This page regulates access control for secure system usage.



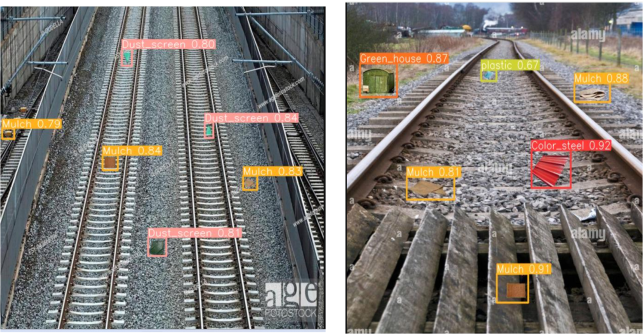
**Figure 8:** Login Page

The authentication page lets users who have active accounts enter their system access information. The system allows secure access sessions and protected model usage.



**Figure 9:** Upload Input Image

The application offers a graphical user interface with a component to allow users to submit pictures of railway tracks. The EBSE-YOLO model analyzes images which users have uploaded through the user interface.



**Figure 10:** Predicted Results

Model processing of an uploaded railway track image results in visual output presentation in the predicted results image. The software uses drawing boxes to show detected foreign items while labeling them as specific objects for instance "plastic bag" or "metal piece." The model assigns a confidence score to every detection that displays its level of prediction certainty. Users gain better understanding of model performance through the visual display of detected small objects that it accurately identifies on railway tracks.

**5. CONCLUSION**

We studied multiple detection algorithms intended to recognize foreign items located across electrified rails. A collection of eight algorithms includes YOLOv5s [30], YOLOv5s-E, YOLOv5s-B, YOLOv5s-S and EBSE-YOLO and YOLOv5 - Ghost CNN and YOLOv5x6. Extracting EBSE-YOLO from the group of methods becomes significant for its advancement of goal object identification through innovative mechanisms that employ ECA-internet and BiFPN fusion and SPD-Conv alongside EIOU loss function implementation. The detection method provides enhanced computational performance and accuracy, which establishes it as a progressive solution to improve railway safety practices and operational excellence.

The assessment of sophisticated YOLO approaches showed YOLOv5x6, along with YOLOv5 + Ghost CNN, achieved a remarkable 95% mAP. The Flask-based front end improves user convenience by providing simple testing features for EBSE-YOLO. A complete practical solution benefits from authentication integration because it ensures system security through controlled access.

**5.1 FUTURE SCOPE**

The exploration of version compression methods through pruning and quantization, and distillation aims to reduce parameter requirements and computational needs for the version. These optimized approaches will make EBSE-YOLO ready for deployment in restricted computational environments with preserved accuracy levels and faster inference speeds. Modifications made to the proposed method through additional optimization will enable better version compression while maintaining its performance level. The adoption of these trends will make EBSE-YOLO more suitable for practical railway safety operational management needs.

**COMPETING INTERESTS DISCLAIMER:**

Authors have declared that they have no known competing financial interests OR non-financial interests OR personal relationships that could have appeared to influence the work reported in this paper.

**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] Subhani Shaik, V Kakulapati, Saadiq, Ontela Sanjay, and Krishna Reddy,” Real-Time Threat Detection Using the Yolo Version-4 Algorithm”, Acta Scientific Computer Sciences,Volume 5, Issue 5, May 2023.

[9] Subhani Shaik,” Analysis of Photo Plethysmography Signals with Artificial Neural Networks Using Curvelet Transform”, Jour. of Advanced Research in Dynamical & Control Systems, Vol. 11, No. 1, 2019.

[10] Md. Samiuddina, G. Rajender, K. Sai Abhiram Varma, A. Ravi Kumar, and Subhani Shaik,” Health Insurance Cost Prediction Using Deep Neural Network”, Asian Journal of Research in Computer Science, Volume 16, Issue 2, Page 46-53, June 2023.

[11] K. Vaishnavi, G. Pranay Reddy, T. Balaram Reddy, N. Ch. Srimannarayana Iyengar and Subhani Shaik,” Real-time Object Detection Using Deep Learning”, Journal of Advances in Mathematics and Computer Science, Volume 38, Issue 8, Page 24-32, June 2023.