Journal of Theoretical and Applied Information Technology
31st October 2024. Vol.102. No. 20

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

OBJECT DETECTION AT RAILWAY LEVEL CROSSING TO IMPROVE PUBLIC SAFETY

TOMMY WISNU WARDHANA, S.Kom.[1] , RIYANTO JAYADI, S.Kom., Ph.D.[2]

[1]Department Information System Management, Bina Nusantara University, Jakarta 11480, Indonesia

[2]Department Information System Management, Bina Nusantara University, Jakarta 11480, Indonesia

E-mail: [1]tommy.wardhana@binus.ac.id, [2]riyanto.jayadi@binus.edu

ABSTRACT

The most widely used mode of transportation is the train. The growth of Indonesian Railways has greatly benefited the populace in a number of ways, most notably in terms of operations and services. But there hasn't been a corresponding improvement in safety, particularly at railroad level crossings. The paper purpose of this research is to offer ways to improve safety at railway level crossings. The community ultimately benefits from this, as there are fewer accidents at train level crossings. Resources used by making advantage of the CCTV cameras that are stationed at various railroad level crossings. The datasets were created by merging Visdrone datasets, which were supplied by Ultralytics, the company that makes YOLOv8, with additional unique datasets for certain data that are not included in the package. As far as we are aware, this is the first study on object detection at Indonesian level crossings. Although there were some references to comparable research conducted in other nations, none of them made use of YOLOv8, which is now the greatest detection technique The methodology involves making observations and interviews with the division in charge of CCTV surveillance at level crossings. Python programming is done via AnacondaPrompt's Command Line Interface (CLI) tools, while LabelImg is used to provide annotations for custom datasets. The study's findings demonstrate that object identification at railroad crossings is capable of accurately and precisely detecting objects for all object classes, with a high precision level of 98.7%.

Keywords: *Object Detection, Train, Level Crossing, YOLOv8.*

1. INTRODUCTION

In the operation of the train, it is inevitable that the railway tracks intersect with the highway [16]. The intersecting area is called a railway level crossing. A subsidiary of KAI Kereta Api Indonesia), namely KAI Property, provides Crossing Guard workers who are placed at each level crossing equipped with a crossing door. Crossing Guard is tasked with ensuring that trains pass safely, smoothly, and without obstacles. The General of Railways, Ministry of Transportation, Directorate (DJKA) has established three categories of railway level crossings, namely officially guarded, officially unguarded, and illegal (illegal) level crossings. The existence of level crossings, especially illegal ones, is very risky for land vehicle accidents hit by passing trains. As of June 2023, there were 3,039 railway level crossings in Java. About 39 percent of them are officially guarded crossings. However, there are still 16 percent that are illegal or illegal crossings. Improving the safety f railway level crossings is still being pursued. From 2017 to the first semester of 2023, the number of officially unguarded level crossings has more than doubled. The same decline also occurred at illegal crossings from 2016 to June 2023 [8].

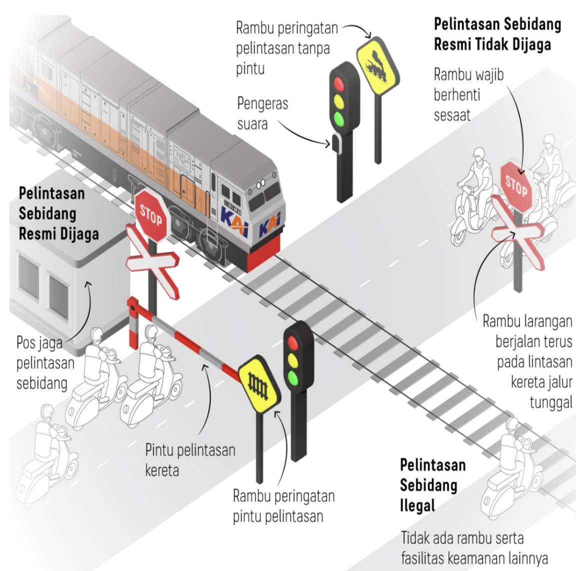

Figure 1: Types of level crossings.

Journal of Theoretical and Applied Information Technology
31st October 2024. Vol.102. No. 20

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

Figure 2: Number of level crossings.

The risk of accidents between trains and other vehicles at level crossings is very high, especially at illegal crossings. In the official press release of PT Kereta Api Indonesia (KAI) in February 2023, the number of accidents at level crossings has increased. Most recently on December 14, 2023, the Whoosh high-speed train feeder train crashed into an online taxi minibus at an unguarded level crossing in Cilame Village, West Bandung Regency, resulting in 5 deaths [8].

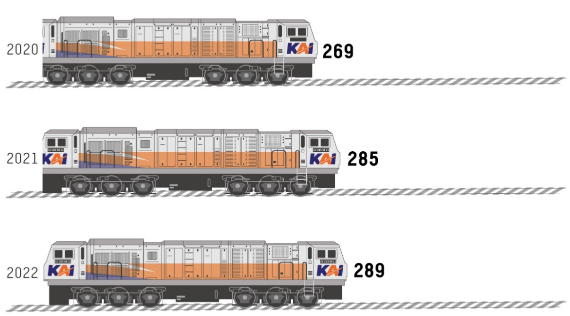

Figure 3: Number of accidents at level crossings.

Drawing from references from a variety of journals, this study examines security measures at railway level crossings that have been implemented in other nations, including China, Japan, Switzerland, and Poland. Specifically, the research examines the use of Artificial Intelligence (AI) via deep learning to monitor objects moving through level crossings in real time. This approach makes use of the resources that are already present at level crossings, such as KAI's CCTV camera equipment, in an effort to lower the likelihood of accidents. Thus, the concept of developing a system that can identify items automatically using actual photos from CCTV camera monitoring emerged. Any object that passes the railway level crossing will be detected by the system. Ultimately, object detection using CCTV cameras allows for the prompt detection of possible hazards.

There are numerous methods for achieving object detection using artificial intelligence and deep learning, such as Region-based Convolutional Neural Networks (R-CNN), Faster R-CNN, Masked R-CNN, Single Shot MultiBox Detector (SSD), Detectron2, EfficientDet, YOLO, and others. Each of them has unique traits, as do the object detection results' speed and accuracy. Thanks to its ease of use, precision, and quick detection time, YOLO is currently the most widely used model. YOLO with its features enables efficient and reliable object recognition in images [4]. YOLO is a convolutional neural network (CNN) based object detection technique. In the context of YOLO, it has been shown to be superior compared to Mask R-CNN and SSD models according to (11) due to these advantages, researchers choose YOLO for object detection. According to [6], the latest literation of YOLO, YOLOv8, outperformed its predecessors in terms of speed and accuracy when it was released in July 2022. The relatively new existence of YOLOv8, as well as the lack of research using CNN for accident detection, are the main reasons for applying YOLOv8 in object detection.

A two-stage object detection model called Faster **R-CNN** (Region-based Convolutional Neural Network) suggests regions of interest first, then classifies them. Although Faster R-CNNs yield reliable results, the two-stage method slows them down. A one-stage object identification model called **YOLO** (You Only Look Once) can forecast bounding boxes and class probabilities in a single pass. Yolo models are renowned for their remarkable speed and real-time inference abilities, but in comparison to other models, they may lose some accuracy. Another one-stage object detection technique that operates at several scales within a single network is called **SSD** (Single Shot MultiBox Detector). It is a well-liked option in many applications since it finds a balance between the accuracy of Faster R-CNNs and the quickness of YOLO.

Table 1: Comparison of object detection models.

Model	Complexity	Speed	Accuracy	Efficiency
Faster R-CNN	High	Slow	High	Less Efficient
YOLO	Medium	Fast	Medium	Efficient
SSD	Low	Fast	High	Efficient

The table illustrates that although Faster R-CNNs are more accurate, their intricate design makes them slower and less effective. YOLO models may lose some accuracy, but they offer speedier inference times. SSD is a flexible option for a variety of applications since it offers an effective balance between speed and accuracy.

Table 2: YOLOv8 performance comparison.

Model	Frame Per Second (FPS)
Faster R-CNN	5-7
YOLO	40-155
SSD	22-46

Performance comparison of Faster R-CNN, YOLOv8, and SSD models. The amount of frames processed per second is represented as FPS. YOLOv8 superior speed over SSD and Faster R-CNN is illustrated in the chart, underscoring its appropriateness for real-time applications.

2. LITERATURE REVIEW

Table 3: Literature review.

No	Research	Title	Problem	Method / Algorithm	State of the art
1	[4]	*Safety Analysis At Level Crossing No. 46 Jl. Kh. Ahmad Dahlan East Jakarta* (Analisis Keselamatan Pada Perlintasan Sebidang No. 46 Jl. Kh. Ahmad Dahlan Jakarta Timur)	Accidents can happen at level crossings because of congestion.	Method HIRARC	It is well recognized that there are deficiencies in the facilities' completion, unsafe traffic volumes, and poor road user discipline.
2	[11]	*A New Form of Train Detection as a Solution to Improve Level Crossing Closing Time*	Lengthy waits for railroad crossings to close.	Method Radar / LIDAR *Geospatial*	Suggested improving the level crossing algorithm without requiring the acquisition of additional train detection equipment.
3	[6]	*Train Object Detection using Faster R-CNN Method with VGG Architecture 16* (Deteksi Objek Kereta Api menggunakan Metode Faster R-CNN dengan Arsitektur VGG 16)	Trains cannot identify objects, especially when there is little light during the day, night, or evening.	Method *Faster* R-CNN Algorithm CNN	Using the light level and object's distance, ascertain whether train objects are present between locomotives and cars.
4	[5]	*Real-Time Obstacle Detection Over Railway Track using Deep Neural Networks*	Because there is insufficient railroad control, accidents at train crossings happen often.	Method CNN Algorithm YOLO v5	Detects all impediments at railroad crossings and notifies all approaching locomotives of their presence.
5	[3]	*A Review of Vision-Based On-Board Obstacle Detection and Distance Estimation in Railways*	The autonomous train tracks are not equipped with any sensors to identify nearby objects or rails.	Method *Computer Vision* Algorithm *Hough Transformation*	A review of vision sensors given their widespread application in the industry.
6	[13]	*Railway Crossing Security System for Highway Traffic Lanes* (Sistem Pengamanan Perlintasan Kereta Api Terhadap Jalur Lalu Lintas Jalan Raya)	At train crossings, there is no safety monitoring system in place.	Method Sensor Arduino	Creating a safety monitoring gadget for railroad crossings that can detect the arrival of a train.
7	[10]	*Analysis of the Railway Accident-Related Damages in South Korea*	Determine the extent of the train accident's damage.	Method ZIG model	Find out how individuals, the high-speed train system, and the death toll are related to each other.
8	[2]	*Study of the Application of Door Technology with Automatic Fences and Yellow Boxes at Level Crossings* (Kajian Penerapan Teknologi Pintu dengan Pagar Otomatis dan Yellow Box di Perlintasan Sebidang)	Unguarded crossings account for the majority of level crossing accidents.	Method Yellow Box & Detection System (YBDS)	The doorstop is automatically closed by means of the Smart Controller, CCTV, and Axle Counter sensors.
9	[16]	*Vehicle Detector for Railway Crossing Safety* (Pendeteksi Kendaraan untuk Keamanan Perlintasan Kereta Api)	Vehicles stopped at railroad crossings have not yet been detected.	Method CNN	Determine which objects or cars are stopped on the train rails.
10	[9]	*Analysis of Railway Accidents' Narratives Using Deep Learning*	The cause of the accident is only narrated and difficult to understand.	Method Word2Vec dan GloVe	Determine the accident's cause with accuracy using the report's narrative.

The numerous investigations that have been carried out make reference to the review of the literature about object detection at railroad level crossings. The most recent version does not appear to make use of the YOLO algorithm. Although the algorithm has reached version 8, the most recent version was only used until version 5. There is a significant difference in the use of YOLOv5 and YOLOv8, the most recent version. In this research, we present an object detection method that uses YOLOv8 and all of its benefits to identify things crossing railroad level crossings. With this strategy, the precision and speed of detection will no longer be a barrier to the use of the prior iteration.

According to [12], the You Only Look Once (YOLO) method is a real-time object detection technique that doesn't take a long time to construct an object detection zone. To identify items in an image, YOLO employs an artificial neural network (ANN) technique. According to [1], this network will segment the image into areas that provide classification as either the intended object or not. At 45 frames per second, this method is incredibly quick; in fact, it has been modified to reach 155 frames per second [7]. The input image is divided into SxS boxes by the You Only Look Once (YOLO) method. In the event that the prediction box contains an object, it will be in charge of identifying it.

[7] state that there are numerous advantages to applying the YOLO method.

1) The YOLO algorithm is incredibly quick. because in order to keep items and other regions apart, this method separates regions into boxes.
2) Unlike other algorithms that evaluate a picture through multiple iterations, the YOLO method examines the complete image during the training and testing phases.
3) The YOLO algorithm discovers representations of objects that are generalizable.

To anticipate each bounding box, the YOLO network incorporates characteristics from the entire image. The YOLO design maintains excellent average precision while enabling quick end-to-end training. [14] states that the YOLO algorithm operates in multiple steps, which include:

1) Dataset
   Used to hold details about the thing that has to be found.
2) Annotation
   This is the process of labeling each image by assigning a bounding box and the class name to each object.
3) Training
   Instruction In the instruction process that needs to be comprehended:
   a) Batch size refers to the quantity of data used in a training session; the higher the power, the better the outcome.
   b) The number of iteration processes is called the epoch.
   c) The speed at which the training process proceeds is known as the learning rate; the higher this rate, the simpler it is to identify things.

YOLOv8, which builds on the innovations of earlier YOLO versions, adds additional features and optimizations that make it the best option for a variety of object identification tasks across a broad range of applications. It is constructed using the subsequent architecture. Among YOLOv8's features are:

- Advanced Backbone and Neck designs: YOLOv8 optimizes feature extraction and object detection performance by utilizing cutting-edge backbone and neck designs.
- Anchor-free Split Ultralytics Head: YOLOv8 uses an anchor-free split Ultralytics head in place of an anchor, which improves detection efficiency and accuracy over anchor-based methods.
- Optimal Accuracy-Speed Tradeoff: YOLOv8 is well-suited for real-time object recognition tasks across a variety of application domains by keeping an eye on preserving an ideal balance between accuracy and speed.
- Variety of Pre-trained Models: YOLOv8 makes it simpler to select the ideal model for your particular use case by providing a selection of pre-trained models that address a range of tasks and performance needs.

According to current sources, YOLOv5 is still in use, however YOLOv8 has not yet been applied to this issue. Because of its stability and the relative novelty of the more recent models, the well-proven YOLOv8 model is still the model of choice in the majority of computer vision use cases. The algorithm utilized for this study is YOLOv8, which is based on the current benefits.

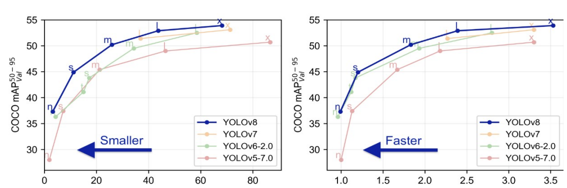

Figure 4: Comparison YOLO version models.

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

3. METHODOLOGY

This research was conducted using a qualitative method and data collection through interviews and observations.

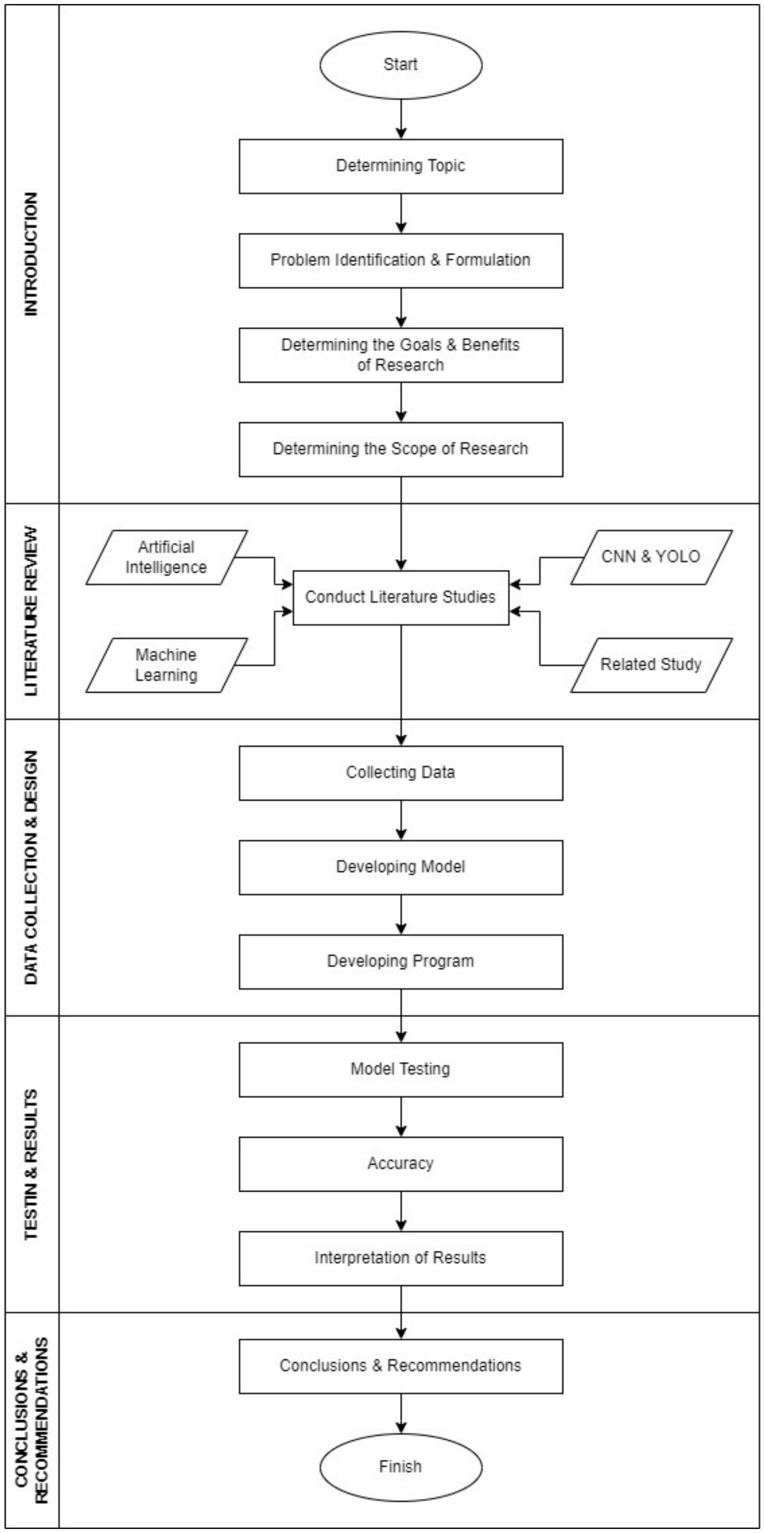

Figure 5: Stages of research.

1.1 Introduction

1.1.1 Determining topic

The frequency of accidents at railroad level crossings is the driving force for this research. Data gathered over the last five years indicates that the number of accidents at railroad level crossings is still rising. This makes it more important for academics to use AI technology to enhance public safety, particularly at railroad level crossings, by identifying preventive measures.

Thus, the subject matter under discussion is associated with the heading "Object Detection at Railway Crossings to Improve Public Safety."

1.1.2 Problem identification and formulation

Researchers use tools such as fishbone diagrams to conduct causal analysis when finding and formulating problems down to the root causes of the problem. The details of fishbone analysis, which uses four primary indicators, are as follows.

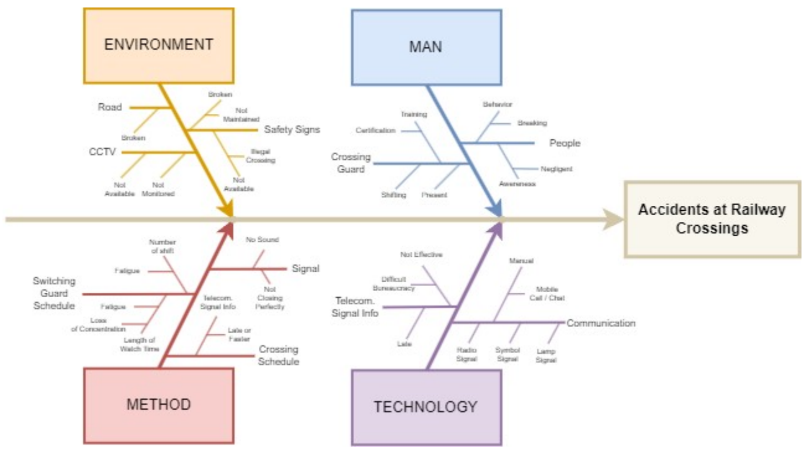

Figure 6: Fishbone analysis.

1.1.3 Determining the goals and benefits of research

The purpose is to offer ways to improve safety at railway level crossings. The benefits that will be obtained from this research include reducing the accident rate, minimizing the losses incurred, and increasing public safety at railway level crossings.

1.1.4 Determining the scope of research

The goal of this project is to use AI technology and in-depth learning to detect things crossing railroad crossings in order to construct a security system that serves as a preventive measure.

1.2 Literature Review

This research refers to a literature review sourced from similar journals from various sources both from within and from abroad. This literature study also draws from knowledge books related to the use of AI and also the latest technology that can be applied to improve public safety at railway level crossings.

1.3 Data Collection and Design

1.3.1 Data collection

Primary data is got from interviews conducted with related parties in the KAI organization including, Directorate of Infrastructure Management, Directorate of Facilities Management, Directorate of Operations and Directorate of Safety and Security. The position level of the interviewees is at the level of Supervisor, Assistant Manager, Manager, and Vice President. The data that has been collected from the interview activities is then observed to see the actual conditions in the field as part of the data validation from the interview activities. This field observation aims to find information in the form of actual conditions at railroad level crossings. Available devices such as CCTV and other supporting tools that can be utilized to improve public safety.

Secondary data used is sourced from a literature review by reviewing previous research, both sourced from within and outside the country. In addition, secondary data is also obtained by reading the current rules both sourced from regulations from the Ministry of Transportation, especially from DJKA and also regulations from KAI. The regulation is a standard reference that must be adhered to and used as a reference in providing solutions in accordance with applicable regulations.

1.3.2 Designing model

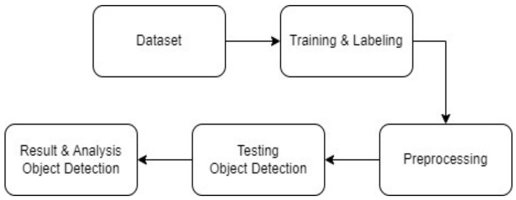

Figure 7: Flow of object detection.

The dataset that will be used is footage sourced from CCTV installed at railway level crossings. This CCTV records all studies that occur in the level crossing area. The image used is a condition image that depicts all passing objects so that later AI can recognize these objects.

Training is the process by which AI, in this case YOLO, learns to recognize items as they are represented in relation to the dataset that is being used. There are two different kinds of datasets: custom datasets made by adding annotations to each object that are not yet included in the Visdrone dataset package, and datasets created using the Visdrone dataset package offered by the YOLOv8 manufacturer, specifically Ultralytics. Labeling is giving AI the name of the thing that appears in the picture so that it has context and meaning and so that the machine learning model can learn from it.

Supervised learning is used, which is a learning strategy in which items are initially recognized by AI for AI to recognize them.

Preprocessing is a stage in the data mining procedure. Before proceeding to the next stage of processing. We'll process the raw data initially. The method that is usually employed is called data preparation, which is another name for data elimination. To help the machine understand the data better, the data will also be modified throughout this process. Data preparation is the process of organizing, reducing, cleansing, and integrating data.

Testing data is information used to evaluate the performance of a machine learning model that has been constructed. YOLOv8 supports the following table outlining the five different model types that can be applied.

Table 4: YOLOv8 models.

Model	size (pixels)	mAP^{val} 50-95	Speed CPU ONNX (ms)	Speed A100 TensorRT (ms)	params (M)	FLOPs (B)
YOLOv8n	640	37.3	80.4	0.99	3.2	8.7
YOLOv8s	640	44.9	128.4	1.20	11.2	28.6
YOLOv8m	640	50.2	234.7	1.83	25.9	78.9
YOLOv8l	640	52.9	375.2	2.39	43.7	165.2
YOLOv8x	640	53.9	479.1	3.53	68.2	257.8

- ***mAP**^val* *values are for single-model single-scale on COCO val2017 dataset.*
  Reproduce by :
  ```
  yolo val detect data=coco.yaml device=0
  ```
- ***Speed** averaged over COCO val images using an Amazon EC2 P4d instance.*
  Reproduce by :
  ```
  yolo val detect data=coco8.yaml batch=1 device=0|cpu
  ```

The YOLOv8m, or Medium Model type, is the model utilized in this study. It is optimized to read speed and accuracy values that are superior to other models and is tailored to the size of the item taken by the CCTV camera. Additionally, this test data is utilized to evaluate how well the produced model handles data that hasn't been seen before or to gauge how well the model can solve problems that are currently being faced.

Results and Analysis are used to find out the results and analyze them to ensure that the model created functions well and can be used with confidence.

1.3.3 Developing program

The YOLOv8 algorithm, which is based on neural networks for object recognition and location determination, is the most generic representation of a single-stage object detection technique. This technique takes the full image as input into the network structure, employs a single CNN model to achieve end-to-end target identification, and computes the bounding box location and object category directly in the output layer [14]. With a strong mix between accuracy and operating speed, the YOLOv8 network represents an ongoing advance over the YOLO series. The four primary modules of

YOLOv8 are input, backbone, head, and prediction. It uses a number of different approaches, including re-parameterized convolution, model scaling for pooling-based models [15], efficient extended layer aggregation network (EELAN), and other methods. The YOLOv8 algorithm's structure can be seen in Figure 3-4 [15].

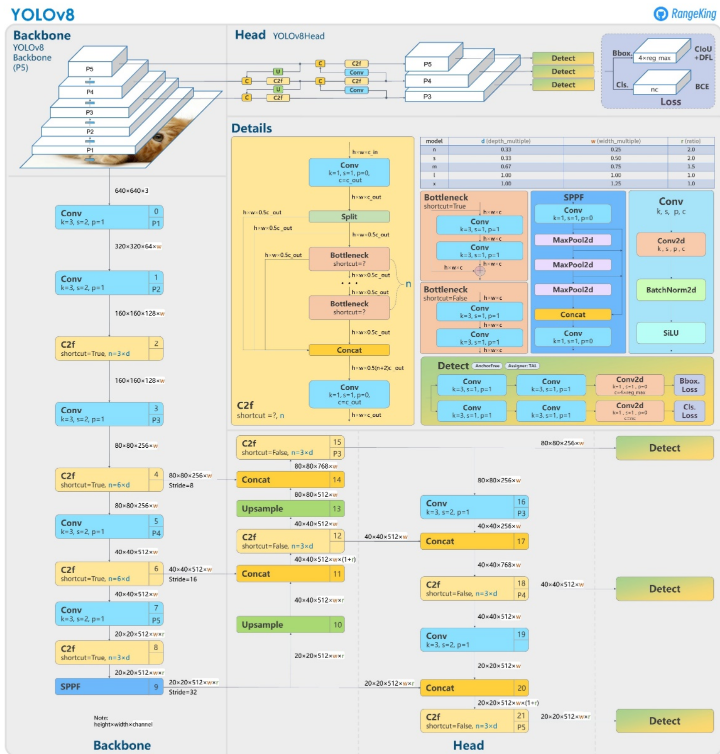

Figure 8: YOLOv8 Architecture.

The YOLOv8 network's central computing unit, or E-ELAN, can instruct clusters of different computing blocks to investigate a wider range of features. Balance is always attained in large-scale local area networks (LAN), irrespective of the gradient direction, path length, and total number of blocks involved [14].

Table 5: Hardware specification.

No	Hardware	Specification
1	Laptop	Asus ROG X Flow 16, AMD Ryzen 9 6900HS, SODIMM DDR5 32GB, NVME SSD 2TB, NVIDIA GeForce RTX 3060 DDR6 6GB.
2	CCTV	HIKVISION DS-2CE16H0T-ITPF, Resolusi 5MP CMOS, Maksimum Resolusi 2560(H)×1944(V), Day & Night mode IR cut filter, Image setting: Brightness, Sharpness, DNR , Mirror, Smart IR.

Table 6: Software specification.

No	Type	Software	Version
1	Operating System	Windows 11	Proffessional
2	Environment	Anaconda Prompt	24.7.1
3	Programming	Phyton Pytorch Pytorch-cuda Torch-mutex Torch 2.4.0 Torchvision Torchaudio LabelImg YOLO	3.10.14 2.4.0 12.4 1.0 2.4.0+cu124 0.19.9+cu124 2.3.1+cu124 1.8.6 8.2.76

Object Detection Algorithm

- **Preparation Environment**

```
# Create Environment
conda create --name ultralytics-env python=3.10 -y
# Activate Environment
conda activate ultralytics-env
# Install Pytorch
conda install pytorch torchvision torchaudio pytorch-cuda=12.4
-c pytorch -c nvidia
# Install Ultralytics YOLOv8
pip install ultralytics
# Install LabelImg
pip install labelImg
```

- **Input: Dataset**
  1. Visdrone Dataset
     https://github.com/VisDrone/VisDrone-Dataset
     Composition: 288 video clips with 261,908 frames and 10,209 static images.
  2. Custom Dataset
     Run LabelImg from AnacondaPrompt use command labelImg. Set ReactBox and annotate it as an object class.
- **Output: Detect Objects**
- **Steps:**

```
# Trainning & Validation Custom Dataset
yolo task=detect mode=train epochs=10 data=data_custom.yaml model=yolov8m.pt imgsz=640 batch=8
# Predict Object
yolo task=detect mode=predict model=yolov8m_custom.pt source=video.mp4 conf=0.5 show=True imgsz=640 line_width=3
```

1.4 Testing and Results

Performance testing was carried out with the aim of determining the level of accuracy in YOLO. This stage is carried out to ensure that the designed function can run well and in accordance with the design objectives. In carrying out performance tests, a test dataset is needed that has been annotated, where the test dataset is a dataset that has never been used as a training dataset. The valid dataset required is around 10% to 20% of the total training dataset.

Experiments require performance indices to evaluate algorithm models. According to the neural network model evaluation index [9]. Utilizing evaluation criteria like recall, accuracy, precision, and F1-score, assess the outcomes. then analyze these results to draw conclusions about the successes and shortcomings of the implemented method. This is important to produce accurate and objective reports.

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}} \tag{1}$$

$$\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}} \tag{2}$$

$$F_1 = \frac{2 \times \text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \tag{3}$$

$$\text{AP} = \int_0^1 \text{Precision} \times \text{Recall dr} \tag{4}$$

Figure 9: YOLO evaluation formula.

In this evaluation, they are True Positive (TP), True Negative (TN), False Positive (FP), and False Negative (FN). True Positive (TP) occurs when the model successfully classifies the data as positive (yes) and the actual answer is also positive (yes). In contrast, True Negative (TN) occurs when the model classifies data as negative (no) and the actual answer is also negative. However, there are also situations where the model can misclassify the data. False Positive (FP) occurs when the model classifies data as positive (yes), but the actual answer is negative (no). Meanwhile, False Negative (FN) occurs when the model classifies data as negative (no), but the actual answer is positive (yes).

The results of performance testing show accuracy that must be in accordance with the model design to get good results. Test results will focus on counting detected objects. The results of this test will be presented in table and confusion matrix form.

Interpretation of results is the final step of this research with the aim of determining relevance based on the research conducted, connecting it with existing knowledge, and forming further research objectives. In this way, the interpretation of the results will guarantee that the findings are valid and trustworthy and contribute to the development of knowledge related to public safety at railway level crossings.

1.5 Conclusions and Recommendations

After describing the research's contents in the preceding chapter, the researcher produces conclusions and recommendations as the study's closing section. The findings of the researcher's investigation will be succinctly explained in the conclusion section. The section on recommendations will list recommendations that the researcher believes are important for parties involved in this research as well as for future research.

4. RESULT AND DISCUSSION

The outcomes of the training program developed with AnacondaPrompt's tools will be discussed in this section. KAI CCTV sources are the data source that was used. Additionally, the data annotation process uses publicly available data with a variety of item categories that are provided by ultralytics, an open source framework that offers YOLOv8 implementation. Visdrone is used in the dataset, and if an object is discovered that is not included in the Visdrone dataset, manual annotation will be done by using LabelImg to customize the dataset.

```
100 epochs completed in 0.154 hours.
Optimizer stripped from C:\Users\19930429\runs\detect\train6\weights\last.pt, 52.0MB
Optimizer stripped from C:\Users\19930429\runs\detect\train6\weights\best.pt, 52.0MB

Validating C:\Users\19930429\runs\detect\train6\weights\best.pt...
Ultralytics YOLOv8.2.76 🚀 Python-3.10.14 torch-2.4.0 CUDA:0 (NVIDIA GeForce RTX 3060 Laptop GPU, 6144MiB)
Model summary (fused): 218 layers, 25,844,971 parameters, 0 gradients, 78.7 GFLOPs
                 Class     Images  Instances      Box(P          R      mAP50  mAP50-95): 100%|██████████| 3/3 [00:00<0
                   all         45         48      0.977      0.947      0.987      0.908
                 Bajaj          5          5      0.925          1      0.995      0.864
               Bicycle          5          5      0.981          1      0.995      0.978
                Busway          5          5      0.972          1      0.995      0.973
                   Car          5          5      0.978          1      0.995      0.938
             Container          5          5      0.976          1      0.995      0.964
               Officer          5          7          1      0.871      0.995      0.891
                 Train          5          6          1      0.654      0.927      0.739
            Motorcycle          5          5      0.978          1      0.995      0.977
                Person          5          5      0.985          1      0.995      0.852
Speed: 0.3ms preprocess, 6.8ms inference, 0.0ms loss, 1.4ms postprocess per image
Results saved to C:\Users\19930429\runs\detect\train6
💡 Learn more at https://docs.ultralytics.com/modes/train
```

Figure 10: Result trainning.

Nine classes of objects 45 for validation and 135 for training have a total of 45 photos that have been specifically annotated. With parameter values of epoch 100, learning rate of 0.01; picture size of 640; batch size of 8; number of photos 45; layers 218; and parameters 25,844,971, YOLOv8 was utilized. Training process took 0.154 hours to complete with 100 epochs.

The FM score is defined as the weighted mean of the recall and precision percentages. As such, this score accounts for both false positives and false negatives. Even though FM is more common than precision, accuracy is not always easy to comprehend. Accuracy performs well when the costs of false positives and false negatives are equivalent.

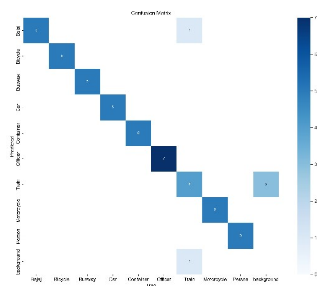

Figure 11: Confussion matrix.

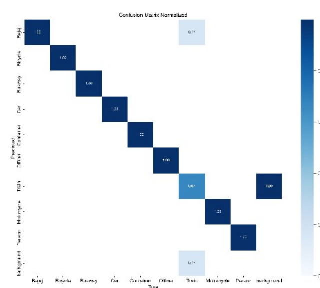

Figure 12: Confussion matrix normalized.

The confusion matrix's results demonstrate that the object detection's accuracy was in line with its class. The tiniest portion displays a detection error because the picture is obscured by a class object that is in front of it, causing the system to either read the background class or overlook another class. Overall though, it indicates that the object discovered is accurate and suitable.

If the costs associated with false positives and false negatives differ, it is desirable to take recall into account in addition to accuracy. Precision in terms of positive findings is the ratio of correctly anticipated observations to all positive finds that were predicted. The percentage of genuine positive forecasts over all real positives is known as the recall.

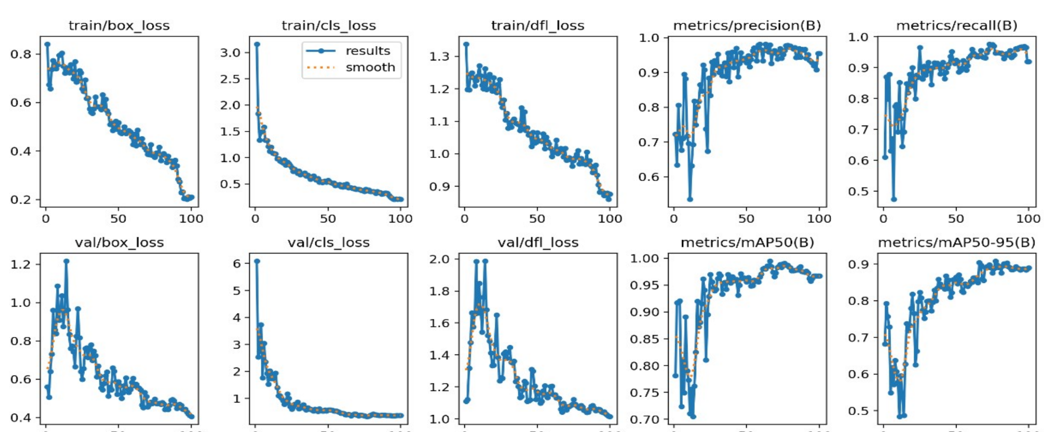

Figure 13: Result of the proposed model.

Journal of Theoretical and Applied Information Technology
31st October 2024. Vol.102. No. 20

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

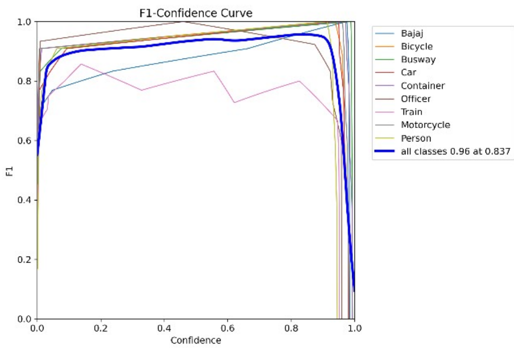

Figure 14: Result of F1 curve.

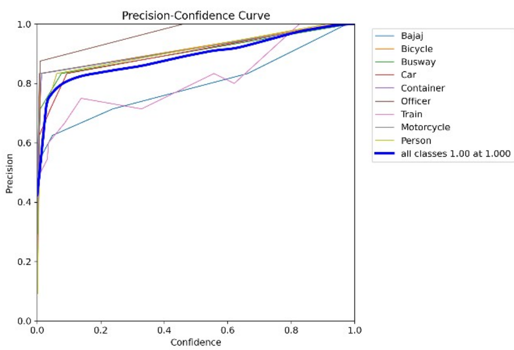

Figure 15: Result of P curve.

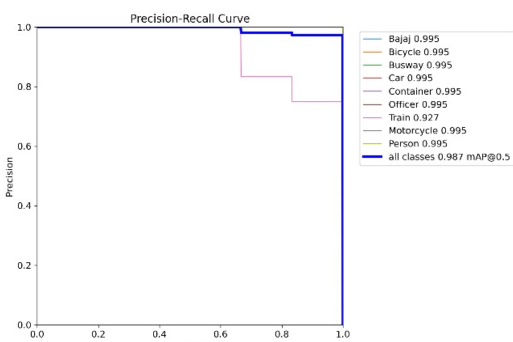

Figure 16: Result of PR curve.

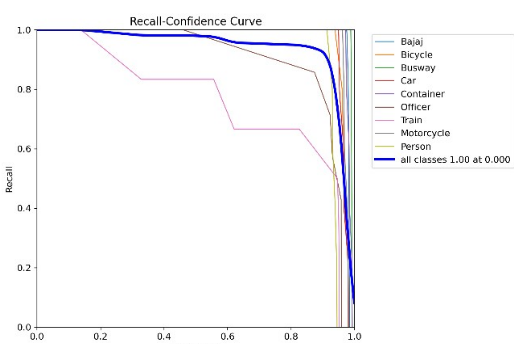

Figure 17: Result of R curve.

The detected object is a CCTV video that records the real situation at the level crossing. The program that has been created will display a box with information in the form of the name of the detected object along with the level of accuracy obtained. In this experiment, we used a size model with medium model (YOLOv8m_custom) so that it can detect objects with the smallest image size on video recordings. Object detection is also carried out during day and night conditions with the aim of determining the level of accuracy produced based on the lighting around the level crossing.

Figure 18: Detection Object by YOLOv8 in Day Condition.

Figure 19: Detection Object by YOLOv8 in Night Condition.

Table 7: Result Evaluation.

Condition	F1	P	R	PR	mAP50
Day	0.96	100%	98.70%	100%	98.70%
Night	0.95	100%	96.50%	100%	96.50%

The results of object detection show that the model has a precision level of 100% for day and night conditions, which means that most of the objects identified by the model are objects that actually exist. Recall of 98.70% during day conditions and 96.50% at night conditions indicates that the model is able to detect most of the objects that actually exist in the dataset. The F1 score of 0.96 during day conditions and 0.95 at night conditions is a good sign that the model has a good balance between precision and recall.

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

According to relevant research sources, the results of this study demonstrate a higher degree of accuracy in object recognition at railroad level crossings when compared to earlier studies. In order to let KAI institutions understand what items are passing at level crossings in real time, YOLOv8 currently offers the best object detection solution. The use of object detection at railroad level crossings has never been done before; this is a new development for organizations who use AI to increase public safety.

5. CONCLUSION AND FUTURE WORK

The implications of the research conducted related to object detection at railroad crossings have been successfully carried out so that every object that passes through the level crossing, whether in the form of a vehicle or any object, can be detected properly as the results of the detection evaluation which show an accuracy value above 90% indicate that object detection has accuracy so that it is in accordance with the real object that passes through.

The results of KAI's object detection will be used in future research to educate crews crossing level crossings so they are aware of the actual conditions before they cross. Furthermore, this data can be utilized to create a sensor that will automatically open and close the level crossing gate. Additionally, the signaling team can utilize it to communicate level crossing information in case of an emergency, enabling them to promptly notify all relevant parties.

Authors Contributions: The authors confirm contribution to the paper as follows: study conception and design: Riyanto Jayadi; data collection: Tommy Wisnu Wardhana; analysis and interpretation of results: Tommy Wisnu Wardhana, Riyanto Jayadi Author; draft manuscript preparation: Riyanto Jayadi, Tommy Wisnu Wardhana. All authors reviewed the results and approved the final version of the manuscript.

Open Data Availability Statement: The authors confirm that the data supporting the findings of this study are available within the article [and/or] its supplementary materials.

REFERENCES:

[1] Aditya Yanuar, "YOLO (You Only Look Once)," 2018.

[2] A. Sianipar, "Study of the Application of Door Technology with Automatic Fences and Yellow Boxes at Level Crossings "(Kajian Penerapan Teknologi Pintu dengan Pagar Otomatis dan Yellow Box di Perlintasan Sebidang), Jurnal Penelitian Transportasi Darat, vol. 22, no. 1, pp. 91–102, Jun. 2020, doi: 10.25104/jptd.v22i1.1603.

[3] D. Ristić-Durrant, M. Franke, and K. Michels, "A Review of Vision-Based On-Board Obstacle Detection and Distance Estimation in Railways," Sensors, vol. 21, no. 10, p. 3452, May 2021, doi: 10.3390/s21103452.

[4] F. I. Kamila, D. Christianto, and H. R. S. Angkat, "Safety Analysis at Level Crossing No. 46 Jl. KH. Ahmad Dahlan East Jakarta" (Analisis Keselamatan Pada Perlintasan Sebidang No. 46 Jl. KH. Ahmad Dahlan Jakarta Timur), JMTS: Jurnal Mitra Teknik Sipil, Aug. 2023, doi: 10.24912/jmts.v6i3.24909.

[5] F. U. Rahman, Md. T. Ahmed, Md. M. Hasan, and N. Jahan, "Real-Time Obstacle Detection Over Railway Track using Deep Neural Networks," Procedia Comput Sci, vol. 215, pp. 289–298, 2022, doi: 10.1016/j.procs.2022.12.031.

[6] J. Pardede and H. Hardiansah, "Train Object Detection using Faster R-CNN Method with VGG Architecture 16" (Deteksi Objek Kereta Api menggunakan Metode Faster R-CNN dengan Arsitektur VGG 16), MIND Journal, vol. 7, no. 1, pp. 21–36, Jun. 2022, doi: 10.26760/mindjournal.v7i1.21-36.

[7] J. Redmon, S. Divvala, R. Girshick, and A. Farhadi, "You Only Look Once: Unified, Real-Time Object Detection," Jun. 2015.

[8] Kompas, "Dangers That Always Lurk at Railway Crossings" (Bahaya yang Selalu Mengintai di Pelintasan Kereta Api), Tim Infografik, Dec. 05, 2023. Accessed: Feb. 12, 2024. [Online]. Available: https://www.kompas.id/baca/kompas_multimedia/bahaya-yang-selalu-mengintai-di-perlintasan-kereta-api

[9] M. Heidarysafa, K. Kowsari, L. Barnes, and D. Brown, "Analysis of Railway Accidents' Narratives Using Deep Learning," in 2018 17th IEEE International Conference on Machine Learning and Applications (ICMLA), IEEE, Dec. 2018, pp. 1446–1453. doi: 10.1109/ICMLA.2018.00235.

[10] M. S. Park, J. K. Eom, J. Choi, and T.-Y. Heo, "Analysis of the Railway Accident-Related Damages in South Korea," Applied Sciences, vol. 10, no. 24, p. 8769, Dec. 2020, doi: 10.3390/app10248769.

[11] M. Zawodny, M. Kruszyna, W. K. Szczepanek, and M. Korzeń, "A New Form of Train

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

Detection as a Solution to Improve Level Crossing Closing Time," Sensors, vol. 23, no. 14, p. 6619, Jul. 2023, doi: 10.3390/s23146619.

[12] Nishad G., "You Only Look Once(YOLO): "Implementing YOLO in less than 30 lines of Python Code", Analytics Vidhya. Accessed: Feb. 12, 2024. [Online]. Available: https://medium.com/analytics-vidhya/you-only-look-once-yolo-implementing-yolo-in-less-than-30-lines-of-python-code-97fb9835bfd2

[13] Pasaribu F. I., Roza I., and Sutrisno O. A., "Railway Crossing Security System for Highway Traffic Lanes" (Sistem Pengamanan Perlintasan Kereta Api Terhadap Jalur Lalu Lintas Jalan Raya), Journal of Electrical and System Control Engineering, vol. 1, pp. 43–52, Aug. 2020, doi: 10.31289/jesce.v4i1.3991.

[14] Rahyagara A., "Object Detection Tutorial Using YOLO" (Tutorial Deteksi Objek Menggunakan YOLO). Accessed: Feb. 12, 2024. [Online]. Available: https://medium.com/@andikirahyagara/tutorial-yolo-you-only-look-once-for-absolutely-noob-c4d5f3751e1f

[15] S. Zhou et al., "An Accurate Detection Model of Takifugu rubripes Using an Improved YOLO-V7 Network," J Mar Sci Eng, vol. 11, no. 5, p. 1051, May 2023, doi: 10.3390/jmse11051051.

[16] Zulkarnain, Zindani A.Y., Amalia A., and Putro F.W., "Vehicle Detector for Railway Crossing Safety" (Pendeteksi Kendaraan untuk Keamanan Perlintasan Kereta Api), Inovasi dan Optimalisasi Teknologi di Bidang Maritim Untuk Mewujudkan Indonesia Emas, vol. 1, no. 1, pp. 35–47, Apr. 2020.