ISSN: 2182-2069 / E-ISSN: 2182-2077

Development of CNN and YOLO Models to Detect Smoking Violations of Cadets in the Environment Politeknik Transportasi SDP Palembang

Sumiati[1*], Raden Muhamad Firzatullah[2], Vidila Rosalina[3], and Dentik Karyaningsih[4]

[1*]Department of Informatics Engineering, Universitas Serang Raya, Serang, Indonesia. sumiatiunsera82@gmail.com, https://orcid.org/0000-0002-7347-5379

[2]Land Transportation Management (MTPD, Politeknik Transportasi Sungai Danau dan Penyeberangan, Palembang, Indonesia. firza.bogor@gmail.com https://orcid.org/0009-0002-9146-5657

[3]Informatics Engineering Study Program, Faculty of Information Technology, Universitas Serang Raya, Serang, Indonesia. vidila.suhendarsah@gmail.com, https://orcid.org/0000-0002-4880-0520

[4]Informatics Engineering Study Program, Faculty of Information Technology, Universitas Serang Raya, Serang, Indonesia. karya.tiek@gmail.com, https://orcid.org/0000-0003-2019-9929

Received: November 17, 2024; Revised: December 22, 2024; Accepted: January 15, 2025; Published: February 28, 2025

Abstract

Disciplinary violations occurred in the SDP Palembang Transportation Polytechnic environment, the violations were in the form of many cadets who smoked. This study aims to provide alternative solutions by utilizing video image data obtained from CCTV to develop CNN and YOLO models to detect smoking violations by cadets in the SDP Palembang transportation polytechnic environment. The benefits of research in the scientific field by testing the level of accuracy of the model developed using the YOLO and CNN algorithms on CCTV image data. The results showed that the precision value increased along with the number of epochs that took place, where the final precision value obtained was at a value of 0.90152, while the recall value showed 0.82878. The CNN algorithm Recall value showed 0.9625, while the CNN algorithm Precision value showed 0.7349. The HOG algorithm Precision value showed 0.86562, while the HOG algorithm Recall value showed 0.6715. The combined recall value of the CNN and YOLO algorithms showed 0.94339 while the combined Precision value of the CNN and YOLO algorithms showed 0.89285. The model has a Mean Average Precision (mAP) value of 0.89109, which shows that the model that has been developed has a good level of accuracy.

Keywords: Accuracy, CNN Algorithm, Mean Average Precision (mAP), Precision Value, YOLO Algorithm.

1 Introduction

The Ministry of Transportation State Senior High School is a senior high school under the auspices of the Ministry of Transportation. Students who attend civil service schools are called cadets. Education at the Ministry of Transportation Civil Service School, cadets undergo academic education and character

Journal of Internet Services and Information Security (JISIS), volume: 15, number: 1 (February), pp. 107-129.
DOI: *10.58346/JISIS.2025.I1.008*

*Corresponding author: Department of Informatics Engineering, Universitas Serang Raya, Serang, Indonesia.

development education (Muhayyang, 2020). Character Education is the process of instilling character in students through parenting patterns that include attention and readiness to carry out these traits, both towards God, oneself and others (Insani et al., 2021). Any act that violates the applicable laws and regulations regarding the care of cadets will be recorded and recorded in the cadet pocket book as a consideration in imposing sanctions on the cadets concerned. Disciplinary violations that occurred in the SDP Palembang Transportation Polytechnic environment were violations in the form of many cadets smoking in the SDP Palembang Transportation Polytechnic environment (Thanaraj et al., 2024). Monitoring students via CCTV plays a very important role in reducing the number of acts of violence in the educational environment. This study utilizes video image data obtained from CCTV to monitor students in the Ministry of Transportation State Senior High School environment (Aprilia & Jacky, 2019). This study will utilize video image data obtained from CCTV to supervise students in the Ministry of Transportation Civil Service High School environment (Bashier & Jabeur, 2021). This study aims to apply the YOLO and CNN algorithms to detect smoking activities of cadets on CCTV image data, and to test the accuracy level of the CNN and YOLO models in detecting CCTV image data in detecting smoking activities of cadets in the SDP Palembang Transportation Polytechnic environment (Jyothi & Mary Gladence, 2024).

Apart from using computational methods to detect smoking violations, namely the Convolutional Neural Network (CNN) approach, which is a computational method that works like a biological neural network system by utilizing convolution., where CNN is able to classify images with a high level of curacy by reducing a number of free parameters and reducing deformation of the input image (Andika et al., 2019). YOLO algorithm is an algorithm that is capable of detecting objects with processing speed and a high level of accuracy (Rahma et al., 2021). The CNN algorithm is an algorithm that has implemented two types of frameworks including Darknet and Darkflow with the support of a Graphics Processing Unit (GPU) in detecting grids of objects in image data. The YOLO algorithm detects smoking activity in outdoor environments in real-time with an accuracy rate of up to 95 percent. Several studies using CNN show good results in the object detection process based on image data (Saponara et al., 2021; Abdali et al., 2019). CNN algorithms are generally used in facial recognition, image/video data classification, image analysis and body gesture detection (Charli et al., 2020), One image can have hundreds of layers that represent other colors in the RGB color architecture, while kernel is an operator that is applied to all parts of the image to get an array of values from the image data (Srinivasu & Veeramani, 2021). Khan & Niu, (2021) While invisible images are images in the form of digital files which are represented in the form of bit values, analog images are produced from analog devices such as analog video recorders, analog photo cameras, webcams, CT scans, Roentgen sensors (Purwanto, 2020; Simangunsong, 2020) The digital image formats that are widely used are Binary (Monochrome) images, Grayscale images and True Color images (Simangunsong, 2020). YOLO is an algorithm that is capable of detecting an object in real-time. The image that has the highest score value will be considered as the detected object (Alfarizi et al., 2023), detecting objects in image data, stride is a parameter that states the amount of shift in the carnel filter which is usually used to reduce the output size (Dwiyanto et al., 2022). Images can be processed by computers, the process of converting analog images to digital must be done first. Analog images are produced from analog devices such as analog video recorders, analog photo cameras, webcams, CT Scans, Roentgen Sensors for the Thorax, Short Wave Sensors in radar systems, Ultrasound Sensors in USG systems and so on, visual display, digital data values represent the color of the image to be processed.

This study utilizes video image data obtained from CCTV to monitor students in the Ministry of Transportation State Senior High School environment. This study aims to apply the YOLO and CNN algorithms to detect smoking activities of cadets on CCTV image data, and test the accuracy level of the YOLO and CNN models to detect smoking activities of cadets in the SDP Palembang Transportation Polytechnic environment (Ponduri & Mohan, 2021). The benefits of research in the field of science by testing the accuracy level of the model developed using the YOLO and CNN algorithms on CCTV image data. So, the main research questions of this article are as follows:

- How to apply the YOLO algorithm in detecting smoking activities of cadets on CCTV image data?
- How to apply the CNN algorithm in recognizing the faces of cadets at the SDP Palembang Transportation Polytechnic?
- How is the accuracy level of CNN and YOLO models in detecting CCTV image data in detecting smoking activities of cadets in the SDP Palembang Transportation Polytechnic environment?

Convolutional Neural Network (CNN) is one of the deep learning algorithms that can recognize and classify images based on the image data it has. Several studies using CNN have shown good results in the process of detecting objects based on image data (Alfarizi et al., 2023; Kin et al., 2021; Surendar, 2024). The CNN algorithm is generally used in facial recognition, image/video data classification, image analysis and body gesture detection (Mohammed et al., 2021; Dziurakh et al., 2024). While invisible images are images in the form of digital files represented in the form of bit values, analog images are produced from analog tools such as analog video recorders, analog photo cameras, webcams, CT Scans, Roentgen Sensors (Purwanto, 2020; Madhan & Shanmugapriya, 2024), one image can have hundreds of layers that represent other colors in the RGB color architecture, while kernels are operators that are applied to all parts of the image to obtain array values from the image data (Dwiyanto et al., 2022). The most widely used digital image formats are Binary (Monochrome) images, Grayscale images and True Color images. YOLO is an algorithm that is able to detect an object in real-time. The image that has the highest score value will be considered as the detected object (Dwiyanto et al., 2022; Ali et al., 2022), YOLO applies Neural Network (NN) in detecting objects in image data, stride is a parameter that states the number of filter kernel shifts that are usually used to reduce the output size (Ali et al., 2022; Wei et al., 2020). Deep learning and computer vision have become invaluable tools in agriculture, providing innovative solutions to a variety of challenges. These technologies help in a variety of applications, such as plant disease detection, crop monitoring, and land management (Dhanya et al., 2022; Fracarolli et al., 2020; Fracarolli et al., 2020; Ghazal et al., 2024; Kamilaris & Prenafeta-Boldú, 2018; Kamilaris & Prenafeta-Boldú, 2018; Bharman et al., 2022). Conventional neural networks (CNN) to analyze plant disease images, where the R-CNN model has a detection rate of 99.39% to detect chili plants, so the potential of this learning model is very deep for a revolution in disease management in agriculture (Sharma et al., 2023). CNN model in the use of an innovation to carry out a learning augmentation strategy that produces new data from noisy images in improving the training model, where the approach utilizes the Bayes algorithm to create images using a convolutional autoencoder (Momeny et al., 2022). Integration of YOLOv9 and Swin Transformer models to detect strawberry ripeness with high accuracy. The hybrid model achieved an average precision (mAP) at IoU 0.5 of 87.3%, outperforming the traditional model using only YOLOv9, which recorded an mAP of 86.1% (Mi & Yan, 2024).

1) Method

- **Image Data**

Image Data is a representation/picture, similarity and imitation of an object captured by a camera. Where image data is an output from a data recording system that can be optical, analog or digital. Image data can be grouped into visible images and invisible images. One form of visible image includes photos, paintings, scenery, holograms and images that appear on a monitor or television screen. While invisible images are images in the form of digital files that are represented in the form of bit values. Analog images are continuous images such as vehicle registration numbers, images on television monitors, X-rays, photos printed on paper, paintings, natural scenery, CT scan results, images recorded on cassette tapes and so on. Analog images cannot be represented in a computer so they cannot be processed directly by the computer. Therefore, in order for the image to be processed by a computer, the process of converting analog images to digital must be carried out first. Analog images are produced from analog tools such as analog video recorders, analog photo cameras, webcams, CT scans, Roentgen sensors for Thorax, Shortwave sensors in radar systems, Ultrasound sensors in USG systems and so on (Purwanto, 2020). The value at an intersection between a row and a column at position (x,y) in Figure 1 is called a picture element, image element, pel, or pixel. The last term (pixel) is most often used in digital images. Pixel is the smallest element of an image. Pixel has 2 parameters, namely coordinates and intensity or color. The value contained in the coordinates (x,y) is f (x,y), which is the intensity of the pixel at that point. Based on the matrix form above, digital images can be systematically written as an intensity function f (x,y), where the values x (row) and y (column) are position coordinates and f (x,y) is the function value at each point (x, y) which states the intensity of the image or the gray level or color of the pixel at that point. Based on this explanation, this study will use image data as a data source sourced from CCTV with certain image processing so that it can be used in detecting and supervising cadet discipline.

- **The Convolutional Neural Network (CNN) Approach**

CNN is a type of artificial neural network that is able to process visual data such as images and videos with a good level of accuracy (Suryavanshi, 2024). CNN uses several layers of neural networks in classifying image data including convolution layers, pooling layers and fully connected layers. Where CNN has good effectiveness in processing image data because the convolution layer is used to filter important features in the image, such as edges, colors, and textures. In addition, the pooling layer is useful in reducing data dimensions and helping the network to be invariant to slight changes in object position. As well as the fully connected layer, connecting all neurons in the previous layer and producing the final output, such as class probability for each possible object. Where the image recognition process using the CNN algorithm is shown in Figure 1.

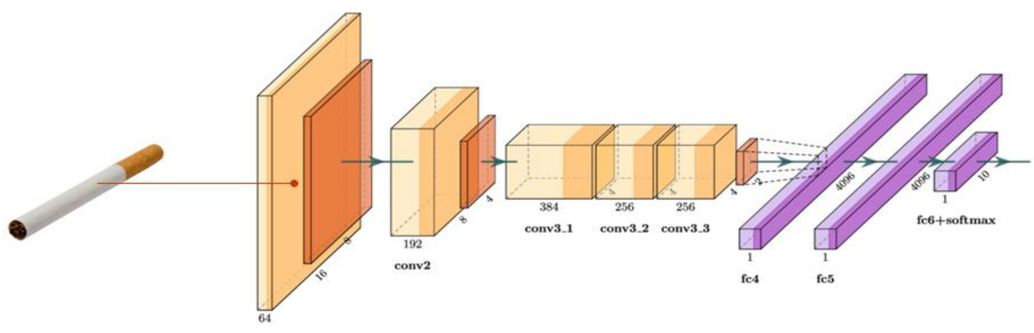


Figure 1: Image Detection Architecture Using CNN Algorithm

(Source: Jiang et al., 2019)

The image classification process using the CNN algorithm begins with processing the input image which is then classified into certain categories such as face shape, sensory devices, or certain colors. The next stage is that the image will be processed into an array form that has a value at each pixel with a certain resolution (channel) consisting of length, width and dimensions. Image channels generally have 3 pieces of image information with each layer representing the colors red, green and blue. In some cases, channels have more than 3 layers, even one image can have hundreds of layers representing other colors in the RGB color architecture (Jiang et al., 2019). Several studies using CNN, including (Saponara et al., 2021) and (Abdali & Al-Tuma, 2019) have shown good results in the process of detecting objects based on image data. Based on this explanation, this study uses the CNN algorithm in creating a model that can detect disciplinary violations of cadets at the Ministry of Transportation College.

The research focuses on creating a CNN model to be able to recognize cadets' disciplinary violation activities. CNN is a deep learning approach that is commonly used in classifying images and grouping them based on similarities and abilities in recognizing activities. CNN has a convolution layer formed from several combinations of convolution layers, pooling layers and fully connected layers. The CNN architecture is shown in Figure 2, which shows that CNN has several types of layers that can be used, namely subsampling layers, convolutional layers, loss layers and fully connected layers.

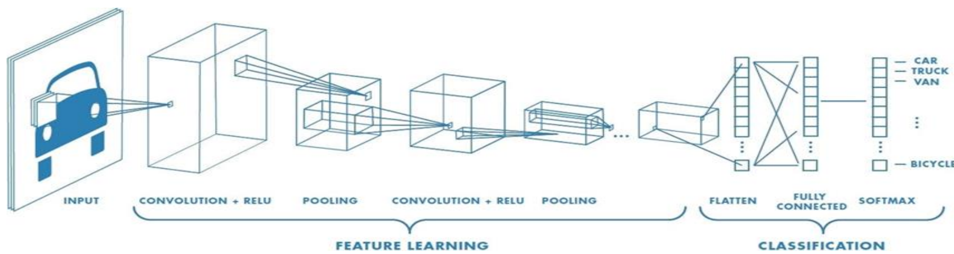


Figure 2: CNN Architecture

(Source: Mulyanto et al., 2021)

The CNN that will be implemented does not use many layers. A simple CNN architecture to complete image classification with a relatively small dataset, considering that there are only a few categories for classification. This technique is implemented in the CNN approach which will affect the performance of the model in training so that overfitting can be reduced. The image classification process using CNN takes place by applying the extraction feature by utilizing the texture extraction feature in recognizing the shape of the cadet's face, the Gabor Filter Characteristics are suitable for use in the application of cadet facial texture recognition where spatially the Gabor Filter is a sinusoidal modulated using the Gauss function. The mathematical equation of the Gabor Filter is shown in Formula 4. Where Θ is the orientation control of the Gabor function, σ is the standard deviation of the Gaussian Envelope and x and y are the coordinates of the Gabor Filter shows in figure 3.

$$g(x,y) = \frac{1}{2\pi\sigma x\sigma \gamma} exp\left[-\frac{1}{2}\left(\frac{\sim x^2}{\sigma x^2} + \frac{\sim y^2}{\sigma y^2}\right)\right] \quad (1)$$

$$\sim x = x\cos\Theta + y\sin\Theta \quad (2)$$

$$\sim y = -x\sin\Theta + y\cos\Theta \quad (3)$$

Figure 3: Filter Gabor

(Source: Boughida et al., 2022)

- **You Only Look Once (YOLO)**

YOLO is an algorithm that is able to detect an object in real-time. Where the detection system is carried out by applying a repurpose classifier or localizer in detecting. The model is applied to image data at several different scales, where the area with the image that has the highest score value will be considered the object being detected (Alfarizi et al., 2023). YOLO applies Neural Network (NN) (Jiang et al., 2019) in detecting objects in image data. YOLOL can divide the image into several regions and predict each restriction block and probability for each prediction. YOLO has advantages over other algorithms in detecting classifier-oriented objects, where YOLO has a high level of prediction accuracy and fast processing speed. YOLO has an architecture similar to CNN but the convolution layer and pooling layer are adjusted according to the number of classes and the number of prediction grids used. Meanwhile, Figure 4 explains how to calculate the bounding box in YOLO in determining the intended prediction object (Du, 2018; Bharathi et al., 2023). YOLO applies Neural Network (NN) (Kin et al., 2021) to detect objects in image data. YOLOL can divide the image into several regions and predict each block of restrictions and probabilities for each prediction. YOLO has advantages compared to other algorithms in detecting object-oriented classifiers, where YOLO has a high level of prediction accuracy and fast processing speed (Fareed et al., 2020). YOLO has an architecture similar to CNN but the convolution layers and pooling layers are adjusted to the number of classes and number of prediction grids used.

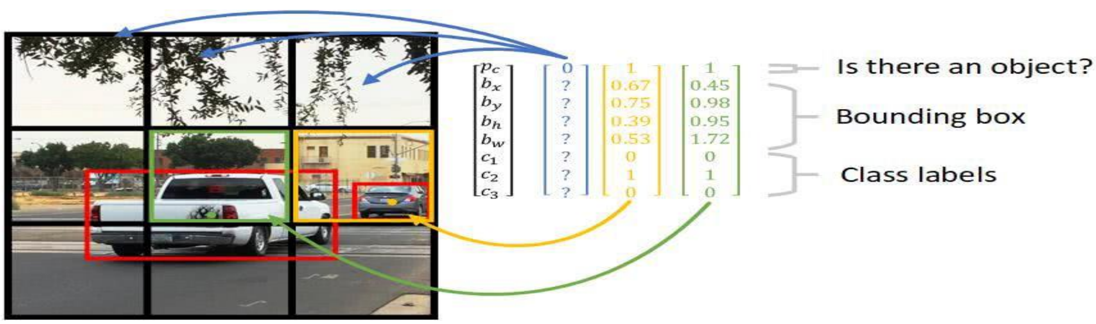


Figure 4: Bounding Box Calculation in YOLO

(Source: Du, 2018)

Where nout is the number of output attributes, while nin is the number of input attributes, k is the size of the convolutional kernel, p is the size of the convolutional padding, while is the size of the convolutional stride. So based on the confidence value of the areas in each bounding box, the class probability of the cigarette object and cigarette pack will be determined, and Figure 10 shows the YOLO research process. Meanwhile, in calculating the output size of the convolution layer in YOLO, Formula 1 is used.

$$n_{out} \frac{n_{in}+2p-k}{s} + 1 \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots (4)$$

Where nout is the number of output attributes, while nin is the number of input attributes, k is the size of the convolutional kernel, p is the size of the convolutional padding, while s is the size of the convolutional stride.

2 Research Method

The research stages can be seen in Figure 5, where the initial stage is the data collection stage which is then preprocessed to obtain uniform data and in accordance with the model creation scenario. After data pre-processing, the next step is to create a model using the CNN method and after the model is created, an evaluation test is carried out to measure the level of accuracy of the model created. In testing the methodology in the study, a dataset was used with labels including cigarettes, cigarette packs and without descriptions with a total of 2,117 images with a resolution of 512 x 512 pixels, where the dataset used Reference in (Gunarsih, 2023) with an explanation of the dataset listed in Table 1. Meanwhile, the dataset used in detecting the faces of cadets at the SDP Palembang Transportation Polytechnic is facial image data of 200 cadets per person with a resolution of 256 x 256 pixels with a grayscale image format as shown in Figure 6.

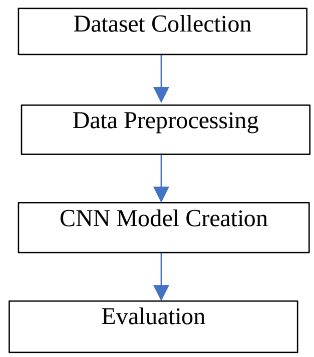


Figure 5: Research Framework

1) Dataset

In testing the research methodology, a dataset was used with labels including cigarettes, cigarette packs and without information with a total of 2,117 images with a resolution of 512 x 512 pixels, where the dataset used refers with an explanation of the dataset listed in Table 1.

Table 1: Dataset Description

Label	Number of Images	Big Data	Type
Cigarette	1.212	± 412 MB	Picture
Wrap	982	± 334 MB	Picture
Without explanation	201	± 68 MB	Picture

(Source: Gunarsih, 2023)

Meanwhile, the dataset used to detect the faces of SDP Palembang Transportation Polytechnic cadets is facial image data of 100 individual cadets with a resolution of 256 x 256 pixels in grayscale image format as in Figure 6.

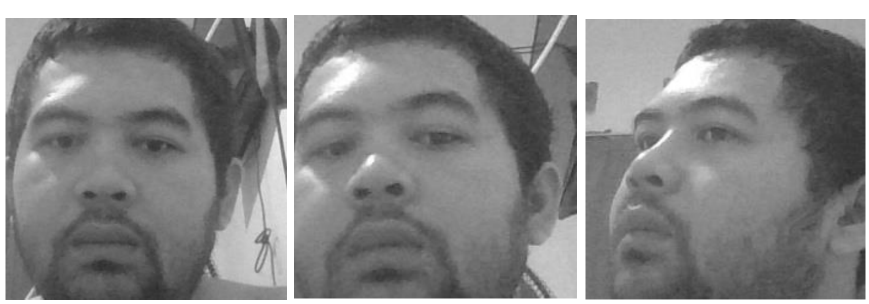


Figure 6: Dataset of Cadet Faces

(Source: Fadnavis, 2014)

2) Data Preprocessing

Data preprocessing is divided into two types of data including face dataset and cigarette dataset. The initial stage of preprocessing the cigarette dataset is by providing markings/annotations to the objects to be detected, where annotation is done by marking each edge point as in Figure 4. After the annotation process is carried out, Next, each object marked in the annotation is given a label for the type of object to be detected. In the example of Figure 7, the purple annotation shows the cigarette smoke object and the yellow annotation shows the cigarette object.

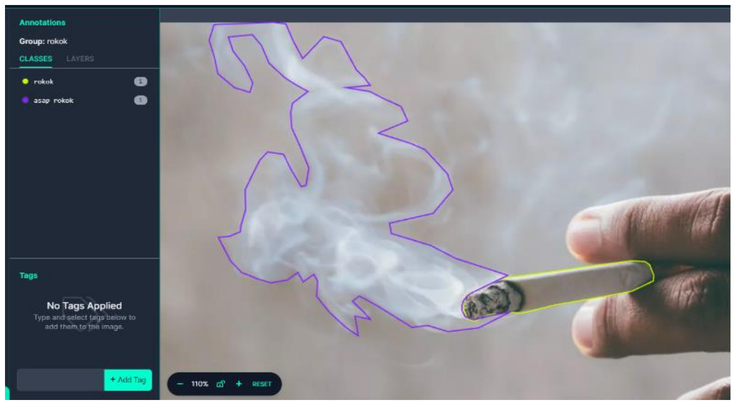
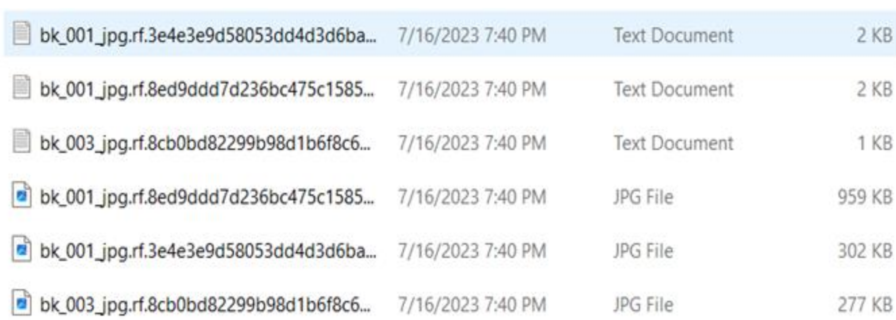


Figure 7: Annotation Process

Next, the object points are combined with metadata, so that the labeling process is carried out as shown in Figure 8.



bk_001_jpg.rf.3e4e3e9d58053dd4d3d6ba...	7/16/2023 7:40 PM	Text Document	2 KB
bk_001_jpg.rf.8ed9ddd7d236bc475c1585...	7/16/2023 7:40 PM	Text Document	2 KB
bk_003_jpg.rf.8cb0bd82299b98d1b6f8c6...	7/16/2023 7:40 PM	Text Document	1 KB
bk_001_jpg.rf.8ed9ddd7d236bc475c1585...	7/16/2023 7:40 PM	JPG File	959 KB
bk_001_jpg.rf.3e4e3e9d58053dd4d3d6ba...	7/16/2023 7:40 PM	JPG File	302 KB
bk_003_jpg.rf.8cb0bd82299b98d1b6f8c6...	7/16/2023 7:40 PM	JPG File	277 KB

Figure 8: Images and Labels

Meanwhile, the label file contains information on the annotation coordinate points on the image along with the object labels. An example of detailed data information in the label file is shown in Figure 9.

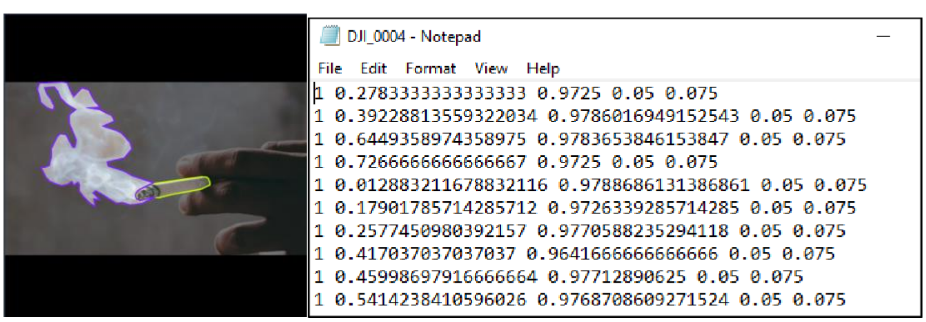


Figure 9: Overview of Annotations and Labels

Next, the data is divided into four parts, including training data, testing data and validation data. Where training data is data used in creating detection models, generally training data has a higher ratio than testing and validation data. Meanwhile, testing data is used to test the capabilities of the resulting model or as a model simulation in the real world. Meanwhile, validation data is used for the model validation process and to prevent overfitting.

Meanwhile, preprocessing of image data in cadet face recognition is done by standardizing the image size to 256 x 256 pixels using the Nearest Neighbor Interpolation algorithm, which multiplies each pixel closest to the center pixel. For example, in Figure 6, there is a 2 x 2 pixel image that will be enlarged to 4 x 4 pixels, where the pixel value at each point will be multiplied vertically and horizontally.

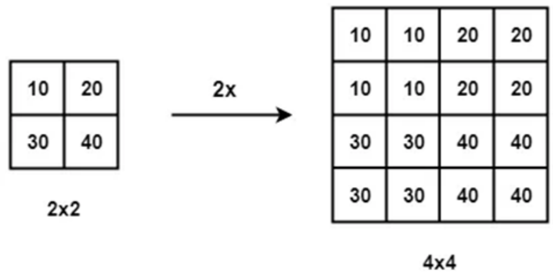


Figure 10: Resize Image

Figure 11: Convert RGB to Grayscale

After the image is resized, the image format is then changed to grayscale using the Luminosity Algorithm, where the image color is represented based on a combination of the red (R), green (G) and blue (B) image color values in each pixel, changed to black and white with a certain value using Formula 2. So that the change in RGB image format to grayscale show figure 11.

$$Grayscale = (0.21 \times R) + (0.72 \times G) + (0.07 \times B) \quad (5)$$

YOLO architecture has a simple architectural structure, where the system will receive image input with the form (448, 448, 3) which is an image that has a size of 448 x 448 with 3 color channels, which will then go through at least one covolution process to produce an output shape (7, 7, 30), with a 7 x 7 image, namely the grid cell size and a value of 30 for the number of bounding boxes multiplied by the sum of the classes and the number of components in one box. Each covolution operation will have parameters for the size of the filter channel, the number of filters and other parameters that affect the shape of the output of the covolution operation, namely padding and stride. YOLO applies the zero-padding method, namely adding 0 to each edge of the input image. In this study, YOLO is used to detect cigarette objects and cigarette packs, the detection process is done by breaking the image down to 1 x 1 pixel, then each bounding box is given a confidence value as in Figure 12. So based on the confidence value of the areas in each bounding box, the class probability of the cigarette object and cigarette pack will be determined, and Figure 12 shows the YOLO research process.

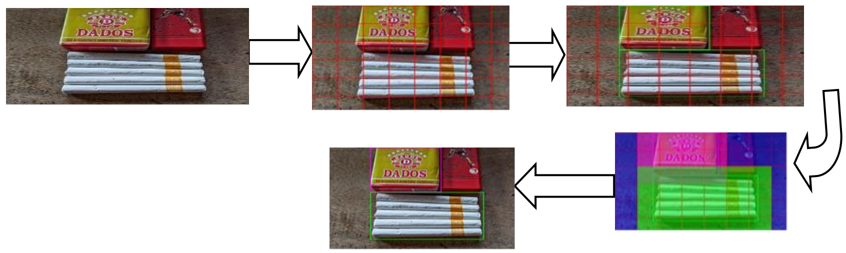


Figure 12: Classification of Cigarettes and Cigarette Packs Using YOLO

(Source: Fadnavis, 2014)

3) Model Building Scenario

Table 2: Testing Scenario

No	Experiment	Model	Batch Size	Epoch	Image Size
1	Experiment 1	YOLOv5s	8	50	160
2	Experiment 2	YOLOv5s	16 (default)	50	160
3	Experiment 3	YOLOv5s	32	50	160
4	Experiment 4	YOLOv5s	64	50	160
5	Experiment 5	YOLOv5s	128	50	160
6	Experiment 6	YOLOv5s	16 (default)	25	160
7	Experiment 7	YOLOv5s	16 (default)	50	160
8	Experiment 8	YOLOv5s	16 (default)	75	160
9	Experiment 9	YOLOv5s	16 (default)	150	160
10	Experiment10	YOLOv5s	16 (default)	300	160
11	Experiment 11	YOLOv5s	16 (default)	50	160
12	Experiment 12	YOLOv5s	16 (default)	50	320
13	Experiment 13	YOLOv5s	16 (default)	50	640
14	Experiment 14	YOLOv5s	16 (default)	25	160
15	Experiment 15	YOLOv5s	16 (default)	50	160
16	Experiment 16	YOLOv5s	16 (default)	75	160
17	Experiment 17	YOLOv5s	16 (default)	150	160
18	Experiment 18	YOLOv5s	16 (default)	300	160

Computing Unit | Batch Size | Number of Epochs | Image-Size

This section explains the model creation scenario, where the creation of the smoking activity detection model is carried out in 18 different algorithm configuration scenarios shown in Table 2. While the configuration indicators used include the number of epochs, batch size, image-size and computational units. While for other configuration indicators listed in Table 3 using general values set in the YOLO version 5 algorithm.

Table 3: Standardized Configuration Indicators

Convolutional Layers	83
Pooling Layers	4
Batch Normalization Layers	201
Activation Layers	83
Linear Layers	3
Jenis Model	YOLOv5s

4) Confusion Matrix

Classification performance measurement generally uses a confusion matrix, where a 2 x 2 matrix explains the results of the binary classification of the dataset used. Where Table 2 shows a simple explanation of the confusion matrix.

Table 4: Confusion Matrix

		Prediction Category	
		Positive	**Negative**
Current Category	**Positive**	TP	FN
	Negative	FP	TN

(Source: Andika et al., 2019)

In Table 4 there is True Positive (TP) which shows the results of correct data prediction with the correct actual value. True Negative (TN) is the result of incorrect data prediction with the wrong actual data value. While False Positive (FP) is the result of incorrect data prediction but the correct predicted result. While False Negative (FN) shows precision and recall data. Meanwhile, accuracy is an evaluation based on the number of correct prediction propositions.

3 Results and Analysis

This section will discuss the testing and drawing of research results consisting of three parts, including the development of the YOLO model in detecting smoking violations, the development of the CNN model in detecting cadet identity and testing the reliability of each model.

1) Effect of Batch-Size

The initial stage of testing focused on measuring the effect of batch-size indicators on the level of precision, recall, GPU memory usage and computing time. Batch size is the number of data samples that usually pass through a neural network at one time, where the batch size determines the number of samples that must be worked on before updating the internal model parameters. In measuring the effect of batch-size on the resulting model, five experiments were carried out with different batch-size conditions, including 8, 16, 32, 64 and 128. Meanwhile, other configuration indicators were standardized by using the YOLOv5s model type, 50 epochs and 160px image size.

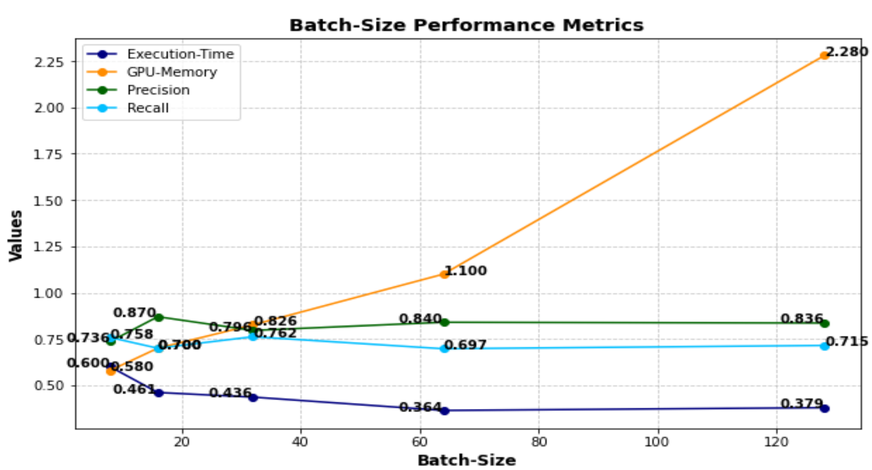


Figure 13: Effect of Batch-size

Based on the test results, where along with the increase in batch size in one training epoch will increase the use of GPU memory as shown in Figure 13, this is because the increasing number of batches, the feature data in the training process in each epoch that is stored will increase. So the use of a high number of batch sizes in a limited computing environment must be considered. Meanwhile, the increase in the number of batch sizes is inversely proportional to the computing time, the higher the batch size (sample) processed in the population in each iteration, the faster the execution time of the data population processing and shorten the execution time. While batch size does not have much effect on the level of accuracy of the resulting model. It can be seen in Figure 12, where the level of precision and recall tends to stagnate and only increases in the number of batch sizes 8 to 16 with an increase value that is not too high. Based on the test of the effect of batch size, it can be concluded that batch size does not have much effect on the level of accuracy of the resulting model. This is in line with the results of research (Radiuk) which shows that the effect of batch size in training models in some conditions will not be significant. Therefore, this study will use a batch size of 16 to measure the influence of other configuration indicators.

2) Effect of Epoch Number

The second stage focuses on measuring the effect of increasing the number of epochs on the resulting model. Where other configuration indicators are standardized by using the YOLOv5s model type, image size 160px, and batch size 16. Epoch is a hyperparameter that determines how many times the neural network algorithm works through the entire dataset both forward and backward.

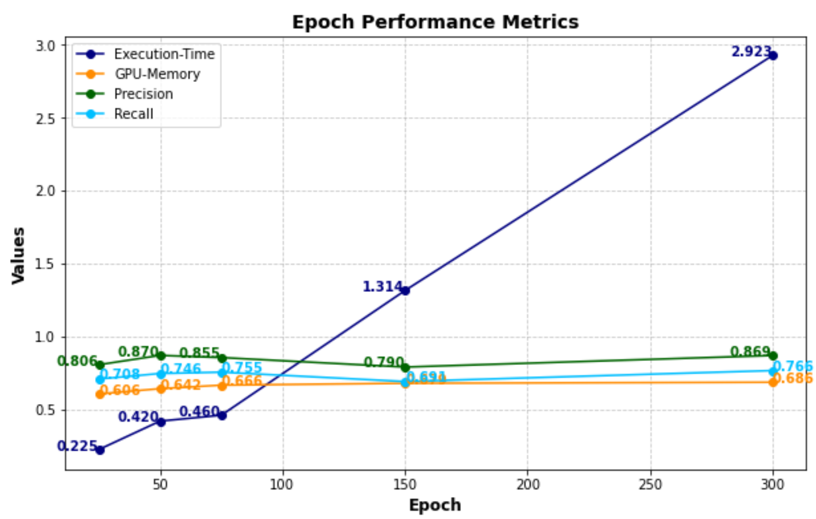


Figure 14: Epoch Effect

Based on the test results as in Figure 14, epoch greatly affects the execution time of model development. Where the more epochs applied, the higher the execution time required. This is due to the increasing number of algorithm layers that must be passed by the dataset population. Meanwhile, the number of epochs does not greatly affect GPU memory usage because in each epoch, the number of dataset populations in each batch will always be the same. Meanwhile, the level of accuracy indicated by precision and recall will have an impact on the increase in the number of epochs, although not too significant, where the increase in accuracy occurs when the number of epochs increases from 25 to 50. This is in line with research which shows that increasing epoch variations will affect the level of model accuracy. Based on this, 50 epochs are used in developing the detection model.

3) Effect of Image Size

The third stage focuses on measuring the effect of image size on the resulting model. Where other configuration indicators are standardized by using the YOLOv5s model type, the number of epochs is 50, and the number of batch-sizes is 16.

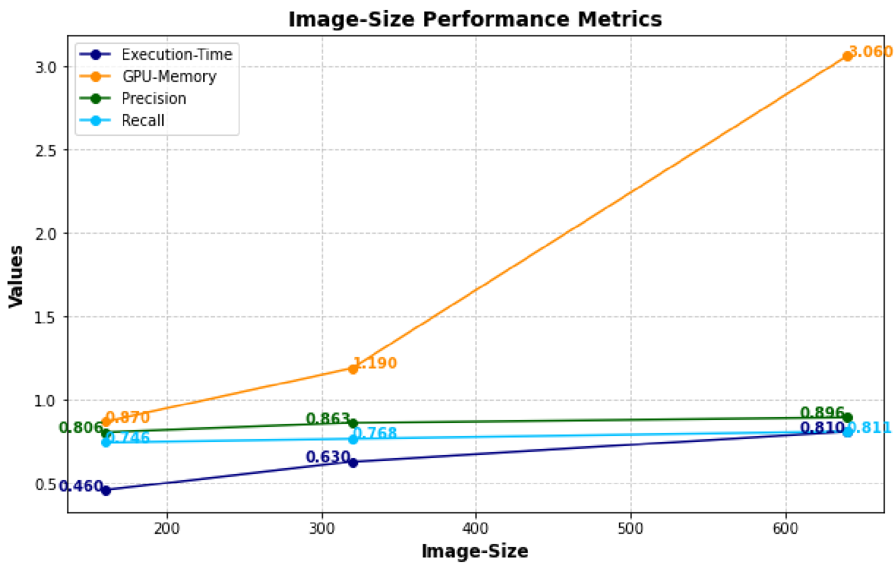


Figure 15: Image Size

From the measurement results, it was found that image size greatly affects the GPU memory used, as in Figure 15. Along with the increase in image size, the GPU memory used also increases. This is because increasing the image size in image data means increasing the number of pixels in each data sample. Meanwhile, increasing image size does not have much effect on execution time, because the amount of data in each epoch does not change and the increase in execution time is only influenced by the increase in the number of feature data due to the increase in image size. Meanwhile, increasing image size increases the level of accuracy indicated by the precision and recall values, this occurs because the validation and testing data used are 640px in size, according to explaining that the closer the image size used in the model training process to the image size used for testing and validation, the more it will affect the increase in accuracy value.

4) Influence of Computing Unit

The last test measures the influence of computing units on the resulting model, there are two types of computing units including model processing using the Central Processing Unit (CPU) while the other type utilizes multi-processor processing owned by the Graphic Processing Unit (GPU) / Multi-GPU processing. While for the configuration indicator using the YOLOv5s model type, the number of epochs is 300, the image-size is 640px, and the number of batch-sizes is 16.

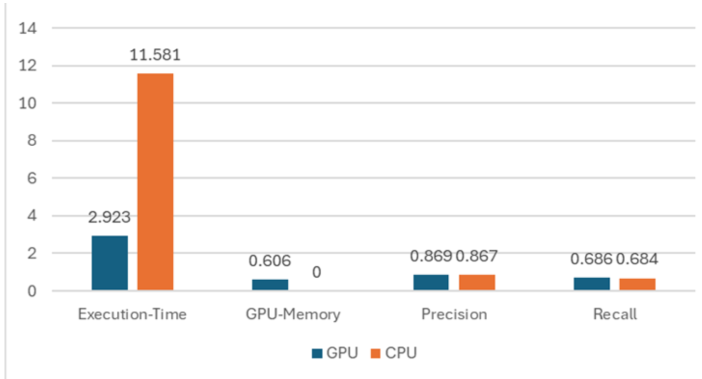


Figure 16: Computing Unit

Figure 16 shows that the performance of the computing unit on execution time is very influential. This is because the number of processing units owned by the CPU in the computing environment setup in this study is only 12 units with each unit having a speed of 2.6 GHz, in contrast to the multi-GPU unit which utilizes 2048 CUDA cores with a speed of 1.057 GHz. This is in accordance with research by (Reddy) showing that multi-GPU performance is much superior in terms of computing speed compared to processing using the CPU. While the type of CPU processing unit does not use GPU memory, this is because the processing results will be stored in Random Access Memory (RAM), while temporary data storage will be stored in the memory cache located on the CPU. While the type of processing unit does not affect the level of accuracy, this is because both on the CPU and GPU, as long as the configuration indicators such as batch-size, number of epochs, type of model and image-size used are the same, they will produce the same results. Based on this, this study will utilize GPU processing so that the computing process takes place faster.

5) Smoking Activity Detection Model

Based on the tests that have been explained in the previous sub-chapter, the model creation is carried out with configuration indicators with specifications including utilizing 16 batch sizes, with 50 processing epochs, utilizing an image size of 640px and a model with the YOLOv5s type.

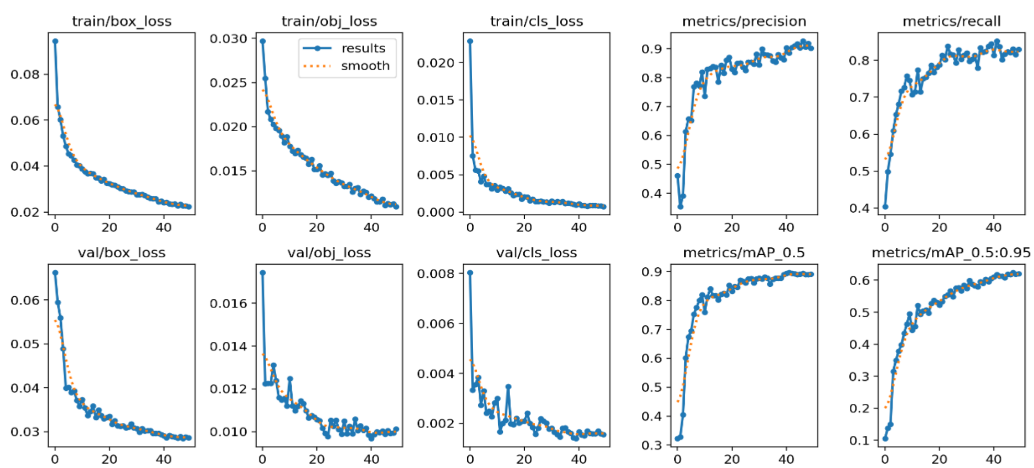


Figure 17: Analysis of the Model Making Process

(Source: Ma et al., 2022)

The analysis of the model creation process is shown in Figure 17, where the graph shows how well the resulting model predicts objects in the bounding box size. Where a lower box loss value indicates a better accuracy value (Ma et al., 2022). Based on the graph, the box loss value decreases with the number of epochs running, where the box loss in the last epoch against the training data has a value of 0.022307 as well as against the validation data with a value of 0.028557, this indicates a good level of accuracy. While object loss is the confidence value of the existence of a missing object. The object loss value decreases with the number of epochs running against the train data showing a value of 0.010121 as well as against the validation data with a value of 0.010944 which indicates a good level of accuracy. While cls loss measures the accuracy of predicting the correct object category in the bounding box. Where cls loss against the train data shows a value of 0.0007255 as well as against the validation data with a value of 0.0015502 which indicates a good level of accuracy. While the precision value is the proportion of the relevant number that can be predicted from all relevant ones, the precision value increases along with the number of epochs that take place, where the final precision value obtained is at a value of 0.90152. The graph of the confidence value against the precision value for each class is shown in Figure 18.

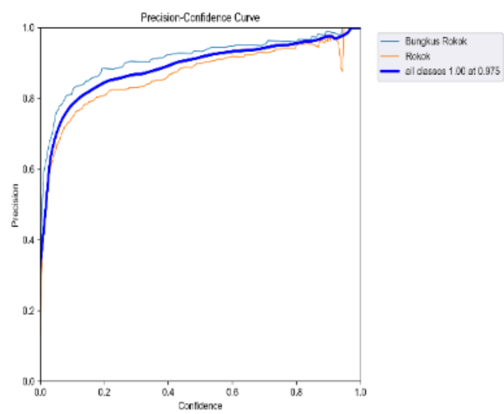


Figure 18: Precision Value Curve for Each Class

Meanwhile, the recall value, which shows the ratio between true positive predictions and true positive global data, shows an increase along with the number of epochs that occur, where the final recall value is at 0.82878. The graph of confidence values against recall values for each class is shown in Figure 19.

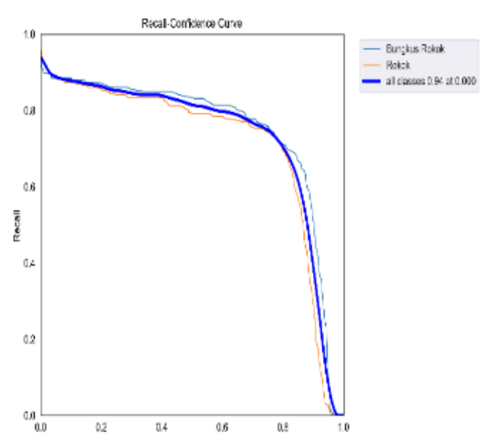


Figure 19: Recall Value Curve for Each Class

Meanwhile, the model has a Mean Average Precision (mAP) value reaching 0.89109, this shows that the model that has been developed has a good level of accuracy in figure 20.

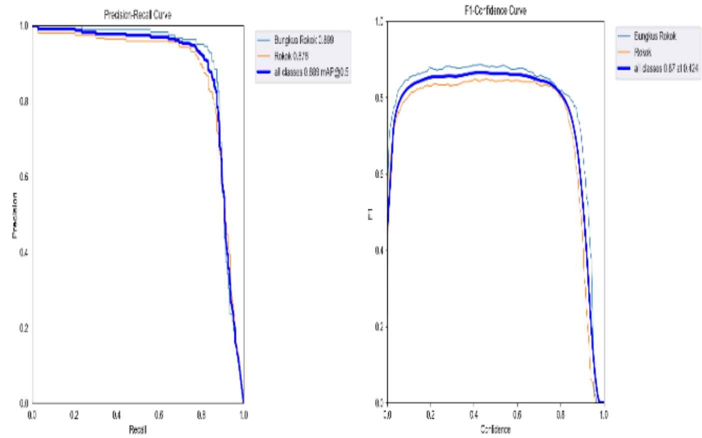


Figure 20: Precision-recall Curve Value Curve for Each Class

(source: results from python program)

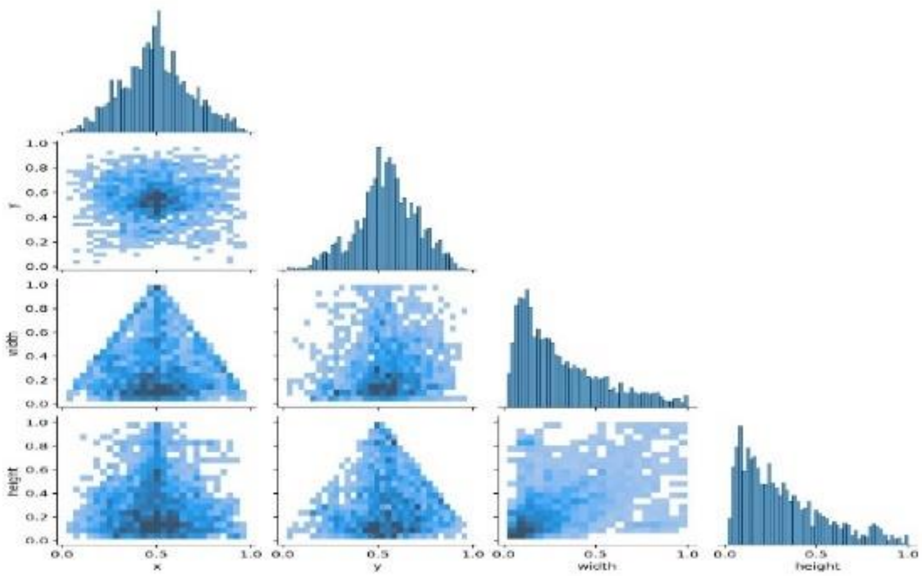


Figure 21: Correlogram Diagram

(Source: Results from python program)

Based on the correlogram table, the Cigarette class label and the Cigarette Pack class label both have a balanced data distribution, although the cigarette class tends to appear more often than the Cigarette Pack class. Based on this graph, the data distribution in the two classes is good shows in figure 21 & 22.

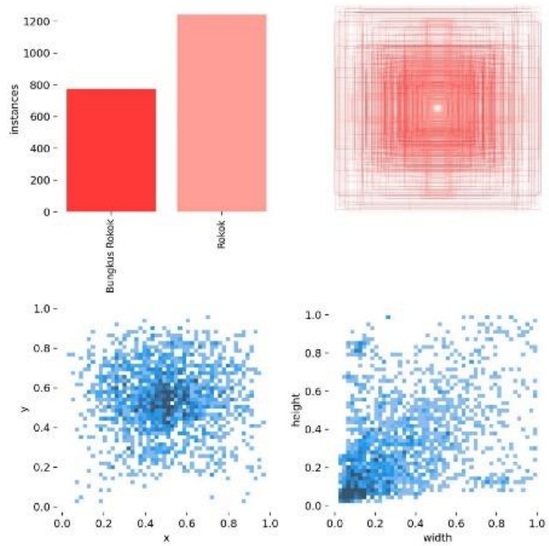


Figure 22: Distribution of Data Between Cigarette Labels and Cigarette Packs

(Source: Results from python program)

6) Confusion Matrix

There is True Positive (TP) which shows the results of correct data prediction with the correct actual value. True Negative (TN) is the result of incorrect data prediction with the wrong actual data value. While False Positive (FP) is the result of incorrect data prediction but the correct predicted result. While False Negative (FN) shows precision and recall data. Next, precision is the proportion of relevant quantities that can be controlled from all those selected. While recall is the proportion of the relevant amount that can be controlled from all that is relevant. The data used is the result of predictions with test data, how to determine the correct and incorrect classes using prediction looping so that the predicted image will be entered into the prediction folder shows in figure 23 & 24.

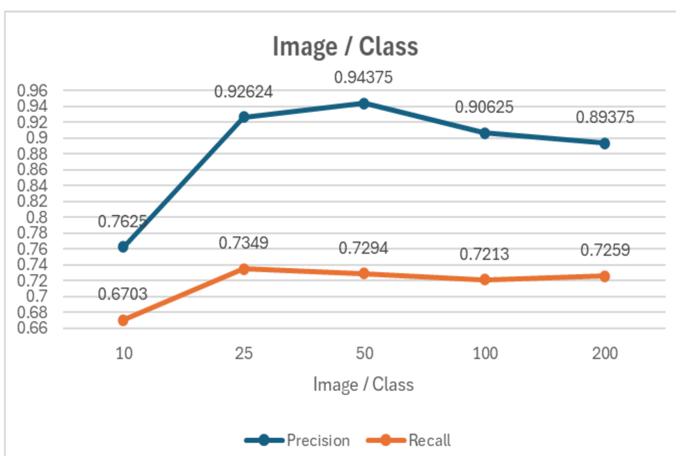


Figure 23: The Effect of Number of Images Classes on CNN Accuracy

(Source: Results from python program)

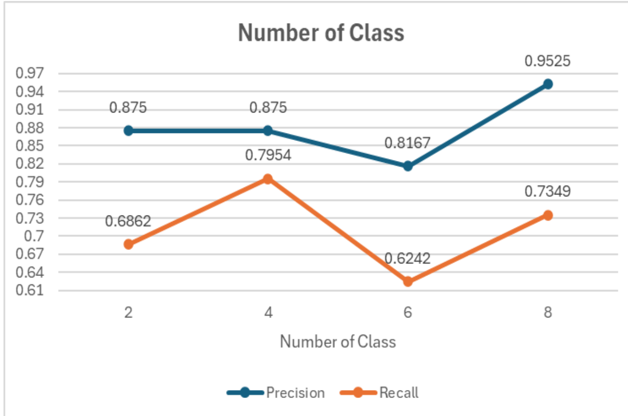


Figure 24: The Effect of Number of Images / Classes on CNN Accuracy

(Source: Results from python program)

7) Comparison of CNN, YOLO, and HOG Algorithms in the Analysis of their Respective Advantages and Disadvantages

CNN has several advantages compared to YOLO and HOG, where CNN has a very good level of accuracy in classifying images with a high level of accuracy, especially for complex tasks such as object recognition and image segmentation compared to YOLO and HOG. In addition, CNN is more flexible in its application to various types of computer vision tasks, from simple image classification to more complex tasks such as object detection and segmentation. And has a very strong feature representation of image data, which makes it very effective in pattern recognition tasks. While CNN has disadvantages compared to HOG and YOLO where it takes a long time to train and requires large computing resources. In addition, CNN also has high complexity which is difficult to understand and implement, and is less efficient for real-time processing in figure 25 & 26 (Kalake et al., 2022).

While YOLO has high speed in detecting images in Real-Time, where YOLO is very fast in processing images and producing detection results. In terms of architecture, YOLO is relatively simpler compared to CNN, making it easier and more flexible to use. However, YOLO has disadvantages where the accuracy obtained is smaller when using images with low sizes, and is sensitive to changes in image scale.

Meanwhile, HOG has invariance to lighting changes, where HOG is quite robust to lighting changes and simple geometric transformations compared to the two previous algorithms. In addition, HOG has more efficient computation, HOG features can be calculated quickly, making it suitable for real-time applications. And HOG can be combined with various classifiers, such as SVM or decision trees. However, HOG is less robust for complex objects and HOG can be affected by noise in the image, which can reduce detection performance. A comparison between CNN, YOLO and HOG algorithms is shown in Table 5 & Table 6.

Table 5: Comparison of CNN, YOLO and HOG Algorithms

Feature	CNN	YOLO	HOG
Accuracy	High	Medium	Medium
Speed	Slow	Fast	Fast
Complexity	High	Medium	Low
Flexibility	High	High	Medium
Application	Classification, detection, segmentation	Real-time detection	Simple object detection

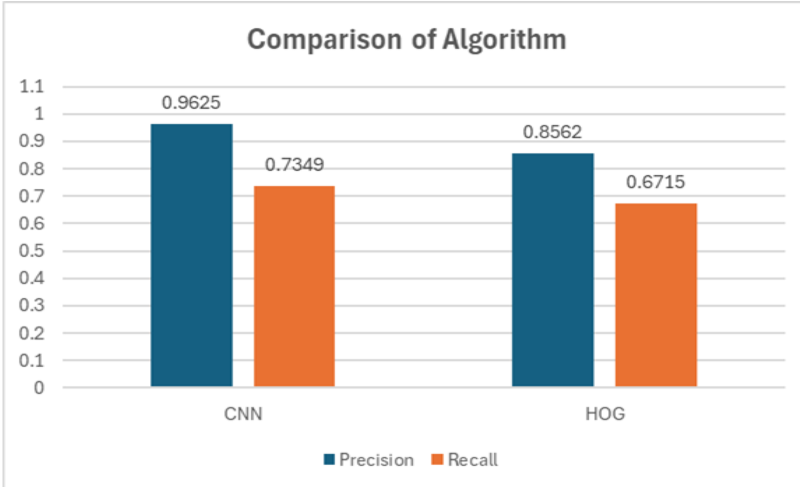


Figure 25: FPS Graphics

(Source: Results from python program)

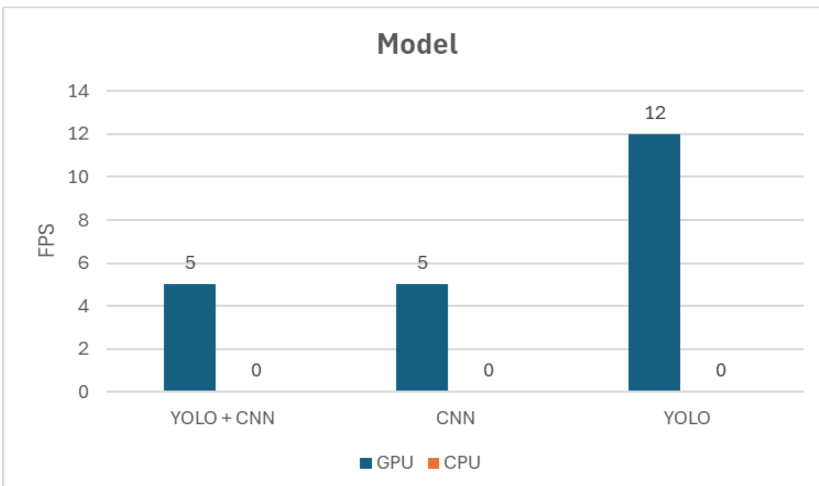


Figure 26: YOLO + CNN Combined Results

(Source: Results from python program)

Table 6: YOLO + CNN Combined Results

PRECISION	0.89285
RECALL	0.94339
Test Scenario	Number of CNN Classes 6 Faces of CNN Agortima + YOLO Number of Classes YOLO 2= cigarettes and cigarette packs

(Source: Results from python program)

8) Potential Limitations or Challenges in Using CCTV Data to Detect Smoking Violations

Challenge in this research, where objects that resemble cigarettes such as pens, pencils, or straws that have similar image attributes such as color, size and texture will be considered as cigarette objects. Where the challenge in further research, is to formulate scenarios and algorithms that are more precise and adaptive in processing CCTV data.

4 Conclusion

This study utilizes video image data obtained from CCTV to monitor students in the Ministry of Transportation State Senior High School environment. This study aims to apply the YOLO and CNN algorithms to detect smoking activities of cadets on CCTV image data, and to test the accuracy level of the YOLO and CNN models to detect smoking activities of cadets in the SDP Palembang Transportation Polytechnic environment. The benefits of research in the scientific field by testing the accuracy level of the model developed using the YOLO and CNN algorithms on CCTV image data. The results of the study showed that the precision value increased along with the number of epochs that took place, where the final precision value obtained was at a value of 0.90152, while the recall value showed 0.82878. The CNN algorithm Recall value showed 0.9625, while the CNN algorithm Precision value showed 0.7349. The HOG algorithm Precision value showed 0.86562, while the HOG algorithm Recall value showed 0.6715. The combined recall value of the CNN and YOLO algorithms showed 0.94339 while the combined Precision value of the CNN and YOLO algorithms showed 0.89285. The model has a Mean Average Precision (mAP) value of 0.89109, which indicates that the developed model has a good level of accuracy.

Limitations and Future Directions

The next research plan is to develop an internet of things device in detecting violations of cadets in the SDP Palembang Transportation Polytechnic Environment.

Acknowledgement

We would like to thank KEMDIKBUDRISTEK as the funder of the 2024 Grant for the funds provided so that this research can run smoothly, and we would like to thank Serang Raya University for the support provided, especially LPPM UNSERA.

References

[1] Abdali, A. M. R., & Al-Tuma, R. F. (2019, March). Robust real-time violence detection in video using cnn and lstm. *In 2019 2nd scientific conference of computer sciences (SCCS)* (pp. 104-108). IEEE. https://doi.org/10.1109/SCCS.2019.8852616

[2] Alfarizi, D. N., Pangestu, R. A., Aditya, D., Setiawan, M. A., & Rosyani, P. (2023). Penggunaan Metode YOLO Pada Deteksi Objek: Sebuah Tinjauan Literatur Sistematis. *AI dan SPK: Jurnal Artificial Intelligent dan Sistem Penunjang Keputusan, 1*(1), 54-63.

[3] Ali, L., Alnajjar, F., Khan, W., Serhani, M. A., & Al Jassmi, H. (2022). Bibliometric analysis and review of deep learning-based crack detection literature published between 2010 and 2022. *Buildings, 12*(4), 432. https://doi.org/10.3390/buildings12040432

[4] Andika, L. A., Pratiwi, H., & Handajani, S. S. (2019). Klasifikasi penyakit pneumonia menggunakan metode convolutional neural network dengan optimasi adaptive momentum. *Indonesian Journal of Statistics and Its Applications, 3*(3), 331-340. https://doi.org/10.29244/ijsa.v3i3.560

[5] Aprilia, V., & Jacky, M. (2019). Relasi kuasa dan CCTV di sekolah menengah atas negeri di Bojonegoro. *Paradigma*, *7*(4).

[6] Bashier, E., & Jabeur, T. B. (2021). An Efficient Secure Image Encryption Algorithm Based on Total Shuffling, Integer Chaotic Maps and Median Filter.

[7] Bharathi, S., Senthilarasi, M., & Hari, K. (2023). Key frame extraction based on Real-Time

person availability using YOLO. *Journal of Wireless Mobile Networks, Ubiquitous Computing, and Dependable Applications, 14*(2), 31-40. https://doi.org/10.58346/JOWUA.2023.I2.003

[8] Bharman, P., Saad, S. A., Khan, S., Jahan, I., Ray, M., & Biswas, M. (2022). Deep learning in agriculture: a review. *Asian Journal of Research in Computer Science, 13*(2), 28-47. https://doi.org/10.9734/ajrcos/2022/v13i230311

[9] Boughida, A., Kouahla, M. N., & Lafifi, Y. (2022). A novel approach for facial expression recognition based on Gabor filters and genetic algorithm. *Evolving Systems, 13*(2), 331-345. https://doi.org/10.1007/s12530-021-09393-2

[10] Charli, F., Syaputra, H., Akbar, M., Sauda, S., & Panjaitan, F. (2020). Implementasi Metode Faster Region Convolutional Neural Network (Faster R-CNN) Untuk Pengenalan Jenis Burung Lovebird. *Journal of Information Technology Ampera, 1*(3), 185-197. https://doi.org/10.51519/journalita.volume1.isssue3.year2020.page185-197

[11] Dhanya, V. G., Subeesh, A., Kushwaha, N. L., Vishwakarma, D. K., Kumar, T. N., Ritika, G., & Singh, A. N. (2022). Deep learning based computer vision approaches for smart agricultural applications. *Artificial Intelligence in Agriculture, 6,* 211-229. https://doi.org/10.1016/j.aiia.2022.09.007

[12] Du, J. (2018, April). Understanding of object detection based on CNN family and YOLO. In *Journal of Physics: Conference Series* (Vol. 1004, p. 012029). IOP Publishing. https://doi.org/10.1088/1742-6596/1004/1/012029

[13] Dwiyanto, R., Widodo, D. W., & Kasih, P. (2022, November). Implementasi Metode You Only Look Once (YOLOv5) Untuk Klasifikasi Kendaraan Pada CCTV Kabupaten Tulungagung. In *Prosiding SEMNAS INOTEK (Seminar Nasional Inovasi Teknologi)* (Vol. 6, No. 3, pp. 102-104). https://doi.org/10.29407/inotek.v6i3.2669

[14] Dziurakh, Y., Kulyniak, I., Sarkisian, H., Zhygalo, I., Chepil, B., & Vaskovych, K. (2024). Intrusion Detection Systems for Smart Tourism Platforms: Safeguarding Food Safety and User Privacy. *Journal of Internet Services and Information Security, 14*(4), 484-498. https://doi.org/10.58346/JISIS.2024.I4.030

[15] Fadnavis, S. (2014). Image interpolation techniques in digital image processing: an overview. *International Journal of Engineering Research and Applications, 4*(10), 70-73.

[16] Fareed, A., Ghazi, A., Al-Dawoodi, A., Aljunid, S. A., Idrus, S. Z. S., Rashidi, C. B. M., ... & Fadhel, I. E. I. (2020, January). Comparison of Laguerre-Gaussian, Hermite–Gaussian and linearly polarized modes in SDM over FMF with electrical nonlinear equalizer. In *AIP Conference Proceedings* (Vol. 2203, No. 1). AIP Publishing.

[17] Fracarolli, J. A., Pavarin, F. F. A., Castro, W., & Blasco, J. (2020). Computer vision applied to food and agricultural products. *Revista Ciência Agronômica, 51*(spe), e20207749. https://doi.org/10.5935/1806-6690.20200087

[18] Ghazal, S., Munir, A., & Qureshi, W. S. (2024). Computer vision in smart agriculture and precision farming: Techniques and applications. *Artificial Intelligence in Agriculture.* https://doi.org/10.1016/j.aiia.2024.06.004

[19] Gunarsih, N. A. (2023). Aplikasi Deteksi Larangan Merokok di Tempat Larangan Merokok (Doctoral dissertation, Politeknik Harapan Bersama).

[20] Insani, G. N., Dewi, D., & Furnamasari, Y. F. (2021). Integration of character education in civic education learning to develop the character of elementary school students. *Jurnal Pendidikan Tambusai, 5*(3), 8153-8160. https://doi.org/10.31004/jptam.v5i3.2313

[21] Jiang, J., Liu, F., Xu, Y., & Huang, H. (2019). Multi-spectral RGB-NIR image classification using double-channel CNN. *IEEE Access, 7,* 20607-20613. https://doi.org/10.1109/ACCESS.2019.2896128

[22] Jyothi, B., & Mary Gladence, L. (2024). Enhanced Accuracy for Lung Adenocarcinoma (LUAD) Prediction based UMAP Feature Using Artificial Neural Network. *Journal of Wireless Mobile Networks, Ubiquitous Computing, and Dependable Applications, 15*(4), 380-394.

https://doi.org/10.58346/JOWUA.2024.I4.026

[23] Kalake, L., Dong, Y., Wan, W., & Hou, L. (2022). Enhancing detection quality rate with a combined hog and cnn for real-time multiple object tracking across non-overlapping multiple cameras. *Sensors, 22*(6), 2123. https://doi.org/10.3390/s22062123

[24] Kamilaris, A., & Prenafeta-Boldú, F. X. (2018). Deep learning in agriculture: A survey. *Computers and electronics in agriculture, 147,* 70-90. https://doi.org/10.1016/j.compag.2018.02.016

[25] Khan, Z. Y., & Niu, Z. (2021). CNN with depthwise separable convolutions and combined kernels for rating prediction. *Expert Systems with Applications, 170,* 114528. https://doi.org/10.1016/j.eswa.2020.114528

[26] Kin, N. W., Asaari, M. S. M., Rosdi, B. A., & Akbar, M. F. (2021). Fpga Implementation of CNN for Defect Classification on CMP Ring. *Jurnal Teknologi, 83*(5), 101-108. https://orcid.org/0000-0002-0225-4819

[27] Ma, Y., Yang, J., Li, Z., & Ma, Z. (2022, March). YOLO-cigarette: An effective YOLO network for outdoor smoking real-time object detection. *In 2021 Ninth International Conference on Advanced Cloud and Big Data (CBD)* (pp. 121-126). IEEE. https://doi.org/10.1109/CBD54617.2021.00029

[28] Madhan, K., & Shanmugapriya, N. (2024). Efficient Object Detection and Classification Approach Using an Enhanced Moving Object Detection Algorithm in Motion Videos. *Indian Journal of Information Sources and Services, 14*(1), 9–16. https://doi.org/10.51983/ijiss-2024.14.1.3895

[29] Mi, Z., & Yan, W. Q. (2024). Strawberry Ripeness Detection Using Deep Learning Models. *Big Data and Cognitive Computing, 8*(8), 92. https://doi.org/10.3390/bdcc8080092

[30] Mohammed, M. A., Han, Z., & Li, Y. (2021). Exploring the detection accuracy of concrete cracks using various CNN models. *Advances in Materials Science and Engineering, 2021*(1), 9923704. https://doi.org/10.1155/2021/9923704

[31] Momeny, M., Jahanbakhshi, A., Neshat, A. A., Hadipour-Rokni, R., Zhang, Y. D., & Ampatzidis, Y. (2022). Detection of citrus black spot disease and ripeness level in orange fruit using learning-to-augment incorporated deep networks. *Ecological Informatics, 71*, 101829. https://doi.org/10.1016/j.ecoinf.2022.101829

[32] Muhayyang, M. (2020). Peran Pendidikan Karakter Dalam Membentuk Kedisiplinan Taruna Ditinjau Dari Aspek Kebudayaan Indonesia. *Jurnal Venus, 8*(1), 78-89. https://doi.org/10.48192/vns.v8i1.275

[33] Mulyanto, A., Susanti, E., Rossi, F., Wajiran, W., & Borman, R. I. (2021). Penerapan Convolutional Neural Network (CNN) pada Pengenalan Aksara Lampung Berbasis Optical Character Recognition (OCR). *JEPIN (Jurnal Edukasi Dan Penelitian Informatika), 7*(1), 52-57. https://doi.org/10.26418/jp.v7i1.44133

[34] Ponduri, V., & Mohan, L. (2021). Reliable multiple object detection on noisy images by using YOLOV3. International *Journal of communication and computer Technologies, 9*(1), 6-9.

[35] Purwanto, A. (2020). Penerapan Algoritma Marr-Hilderth Untuk Pendeteksian Tepi Pada Citra CT-Scan. *Bulletin of Information Technology (BIT), 1*(1), 28-35.

[36] Rahma, L., Syaputra, H., Mirza, A. H., & Purnamasari, S. D. (2021). Objek Deteksi Makanan Khas Palembang Menggunakan Algoritma YOLO (You Only Look Once). *Jurnal Nasional Ilmu Komputer, 2*(3), 213-232. https://doi.org/10.47747/jurnalnik.v2i3.534

[37] Saponara, S., Elhanashi, A., & Gagliardi, A. (2021). Real-time video fire/smoke detection based on CNN in antifire surveillance systems. *Journal of Real-Time Image Processing, 18,* 889-900. https://doi.org/10.1007/s11554-020-01044-0

[38] Sharma, R., Kukreja, V., & Bordoloi, D. (2023, May). Deep learning meets agriculture: A faster RCNN based approach to pepper leaf blight disease detection and multi-classification. *In 2023 4th International Conference for Emerging Technology (INCET)* (pp. 1-5). IEEE.

https://doi.org/10.1109/INCET57972.2023.10170692

[39] Simangunsong, M. (2020). Perbandingan Algoritma Elias Delta Code Dan Unary Coding Dalam Kompresi Citra Forensik. *Resolusi: Rekayasa Teknik Informatika dan Informasi, 1*(1), 18-26. https://doi.org/10.30865/resolusi.v1i1.9

[40] Srinivasu, L. N., & Veeramani, V. (2021). Curve-Let Transform Based'Text-In-Image Steganography'By Using Huffman Coding. *International Journal of communication and computer Technologies, 9*(1), 10-14.

[41] Surendar, A. (2024). Emerging Trends in Renewable Energy Technologies: An In-Depth Analysis. *Innovative Reviews in Engineering and Science, 1*(1), 6-10. https://doi.org/10.31838/INES/01.01.02

[42] Suryavanshi, A., Mehta, S., Jain, A., Choudhary, A., & Jain, V. (2024, May). Empowering a Violence Prevention with CNN-SVM based Classification. *In 2024 5th International Conference for Emerging Technology (INCET)* (pp. 1-6). IEEE. https://doi.org/10.1109/INCET61516.2024.10593166

[43] Thanaraj, M. S., Poomalai, R., Vibin, R., Kamalakar, V., Gupta, S., Perumal, M. P. R., & Vadivel, M. (2024). Sustainable energy integration for seawater desalination. *Journal of Environmental Protection and Ecology, 25*(6), 2086–2096.

[44] Wei, X., Wei, D., Suo, D., Jia, L., & Li, Y. (2020). Multi-target defect identification for railway track line based on image processing and improved YOLOv3 model. *IEEE Access, 8*, 61973-61988. https://doi.org/10.1109/ACCESS.2020.2984264

Authors Biography

Sumiati, ST., MM., Ph.D was born in Pekalongan, Indonesia in 1982. Research areas in Intelligent Computing, machine learning, deep learning, Artificial Intelligence.

Raden Muhamad Firzatullah, M.Kom, receive his B.Ed. in the field of Informatics and Computer Engineering Education from State University of Jakarta. He receive M.Kom in Computer Science from Bogor Agricultural University. He is a lecturer and researcher from the Inland Water and Ferries Transport Polytechnic of Palembang. His research focus in Image Recognition, Skyline Query, Software Development and Neural Network.

Vidila Rosalina, M.Kom. is a lecturer and researcher in Informatic Department, Universitas Serang Raya. Educational backround in IS, CRM, DSS, and DM.

Dentik Karyaningsih M.Kom., Dentik Karyaningsih received a Master's degree in Computer Science from Universitas Budi Luhur Jakarta and is currently a lecturer at Universitas Serang Raya. He is particularly interested in the application base mobile android, information system, data mining and E-Commerce".