Scientific Journal of Informatics
Vol. 11, No. 2, May 2024

p-ISSN 2407-7658 https://journal.unnes.ac.id/journals/sji/index e-ISSN 2460-0040

Comparative Analysis of YOLOv5 and YOLOv8 Cigarette Detection in Social Media Content

Cinantya Paramita[1*], Catur Supriyanto[2], Amalia[3], Khalivio Rahmyanto Putra[4]
[1,2,4]Study Program in Informatics Engineering, Faculty of Computer Science, Universitas Dian Nuswantoro, Indonesia
[3]Study Program in Industrial Engineering, Faculty of Engineering, Universitas Dian Nuswantoro, Indonesia

Abstract.
Purpose: Addresses the pressing public health concern of tobacco product portrayal on social media, which significantly influences the younger demographic by glamorizing smoking culture. The purpose is to compare the capabilities of YOLOv5 and YOLOv8 models in detecting and censoring cigarette-related imagery on social media platforms, aiming to reduce exposure among children and teenagers.
Methods: Employing a dataset of 2,188 images collected from Twitter, this research undertook a comprehensive methodology involving data preprocessing, YOLOv5 and YOLOv8 model training, and rigorous evaluation. The study utilized mean Average Precision (mAP) and F1-Score metrics to evaluate the performance of YOLOv5 and YOLOv8 models, focusing on their precision, recall, and efficacy in detecting cigarette and cigarette pack objects.
Result: The analysis highlighted YOLOv8's superiority, with a marginally higher mAP value of 0.933 compared to YOLOv5's 0.919, alongside enhanced precision and recall rates. This result underscores YOLOv8's advanced object detection capabilities, owing to its architectural innovations and anchor-free detection system. Additionally, the study confirmed the absence of significant overfitting or underfitting issues, indicating robust learning processes of the models.
Novelty: Innovates in digital public health by using YOLOv5 and YOLOv8 models to automatically censor tobacco-related content on social media, effectively reducing youth exposure to such imagery. YOLOv8, in particular, exhibits marginally superior detection capabilities. The evaluation results surpass those of previous research on cigarette and cigarette burning detection, underscoring the study's significant contribution to future research and public health initiatives.

Keywords: Yolov5, Yolov8, Deep learning, Detection, Cigarette
Received March 2024 / **Revised** May 2024 / **Accepted** May 2024

INTRODUCTION

In this era of technological advancement, social media has emerged as a highly popular and influential platform in daily life. Regarding to the information cited comes [1] considering the tobacco industry's history of attempting to manipulate public health and interfere with scientific research, it is alarming that the industry has significant influence over tobacco-related content on social media platforms. Millions of people worldwide use social media platforms to interact, share information, and access multimedia content. Individuals of all ages, from young to old, can access various types of content on social media. However, not all content on social media has positive value. Content creators and influencers often unknowingly display negative behaviors through the content they upload on social media, one of which is the culture of smoking frequently portrayed by content creators in their content. Furthermore, data released by Statista indicates that 12.5% of active social media users are teenagers aged 13-17 years old [2]. At this age, teenagers tend to be more easily influenced by the content they consume.

Based on the RISKESDAS (Riset Kesehatan Dasar) report published by the Ministry of Health, there has been an increase in the prevalence of smokers in Indonesia. In 2018, the prevalence rate reached 33.8%, with 9.1% of them being adolescents aged 10-18 years old [3]. People smoke due to various factors, including the influence of cigarette advertisements and content on social media that depict people smoking and cigarettes. In a study involving 145 respondents, where 96 respondents were under 18 years old and the remainder were aged 18 years and above, it indicating interest in cigarettes due to cigarette-related

*Corresponding author.
Email addresses: cinantya.paramita@dsn.dinus.ac.id (Paramita)
DOI: 10.15294/sji.v11i2.2808

content, 87.4% from the internet and 85.4% from social media [4]. Tobacco use, poses a significant public health threat, with over 5 million U.S. adolescents using cigarettes in 2019 including 10.5% of middle school students and 27.5% of high school students. Social media plays a pivotal role, with adolescents exposed to tobacco-related content, influencing attitudes and behaviors[5]. However, this exposure poses a serious challenge as it can contribute to the normalization and glamorization of smoking among youth. To address this issue, a solution is needed to reduce children and teenagers' exposure to content displaying tobacco products. One solution that can be implemented is censoring methods for tobacco products in social media content. However, the identification and censorship process [6], usually done manually, which tend to be inefficient due to the significant time and effort required. Therefore, an automated system is needed to censor images [6] on social media that display cigarette products. By implementing an automatic censorship mechanism, it is hoped that this can be an effort to mitigate the influence of tobacco-related content on the adolescent population.

Computer vision [7] is capable of performing image classification, object detection, and object tracking[8]. Object detection is a method that proves beneficial for identifying objects in images using computers, akin to the processes executed by humans. In the advancement of object detection technology, several methods are utilized to detect objects in images. One of the latest methods is the You Only Look Once (YOLO) method. A study on the utilization of the YOLO method for object detection indicates that YOLO yields rapid and efficient results for object detection[9][10]. YOLO is an object detection method that combines classification and localization concepts. Classification predicts the class of a specific object in the image, while localization determines the object's location in the image. YOLO employs neural networks for real-time object detection [11].

In a separate study, the object detection process was compared across the YOLO, Faster R-CNN, and SSD methods. The findings of this investigation revealed that the YOLO method demonstrated superior performance in object detection. The accuracy level of the YOLO method surpassed that of both Faster R-CNN and SSD [12]. Furthermore, in another study, models utilizing the YOLO method exhibited smaller file sizes and higher Frames Per Second (FPS) compared to the other methodologies [13][14]. Consequently, in the research concerning automatic safety risk detection, a comparison between YOLOv5 and YOLOv8 was conducted to attain superior outcomes. The results of the study indicated that YOLOv8 outperformed YOLOv5 [15][16], thus warranting the application of this method in the current research endeavor.

Based on previous research findings, it is well-established that the YOLO method exhibits favorable performance for detecting cigarette product objects in images. Taking into consideration the exposition provided in the preceding paragraphs, this study aims to conduct object detection and censorship of cigarette products present in images sourced from social media platforms by comparing YOLOv5 and YOLOv8, employing the YOLO method and integrating it into the system.

METHODS

The employment of research instruments is crucial for the smooth execution and optimal results of a study, facilitating effective data collection and processing using appropriate techniques. This research utilizes a combination of hardware and software instruments. The hardware comprises an ASUS laptop equipped with an Intel Core-i5 5200U processor, 8GB RAM, 256GB SSD, and NVIDIA GeForce 840M VGA. The software suite includes the Windows 10 64-bit operating system, PyCharm, OpenCV Library, Google Colab, and RoboFlow, all utilizing Python. This integration of tools is anticipated to efficiently support the research activities. The data collection method involved manual searches on Twitter using keywords relevant to cigarette products, covering the period from January 1, 2019, to March 1, 2023. Images of cigarette objects were downloaded to form the training dataset for the YOLOv5 and YOLOv8 models.

Cigarette object data required preprocessing prior to being used for model training, encompassing steps such as data cleaning, data annotation, and data division. Data cleaning involved collecting and filtering cigarette images to ensure their quality. Data annotation was conducted by labeling [17] each image to identify cigarette objects. Data division was executed by splitting the data into training, validation, and testing sets, with the aim of avoiding overfitting [18]. This issue was addressed by zeroing out (setting to zero) the weights of a specific percentage of hidden units to prevent overfitting [19]. The division ratio of the data was set at 80:10:10 for the training, validation, and testing sets, respectively. Through this technique, the study was able to prepare a dataset for training and testing both YOLO models [20] within a system designed for censoring cigarette products.

To conclude this research, it was necessary to propose a suitable method as the foundation for the development of an automatic censorship system for cigarette products. The intended workflow for this purpose is as follows:

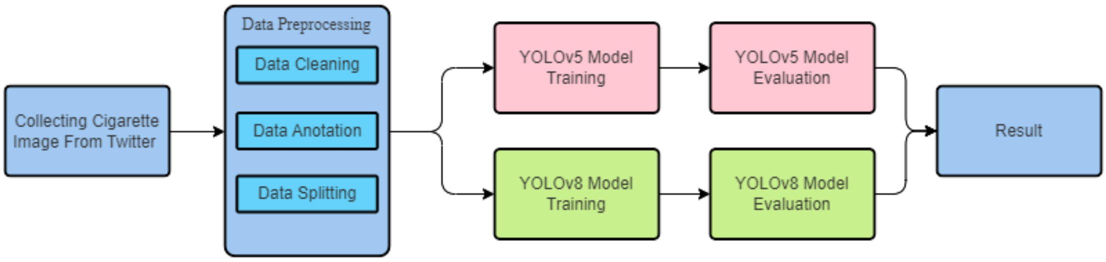


Figure 1. Method flow YoloV5 and YoloV8

Collecting Cigarate Image

Images of cigarettes and cigarette packages were collected through the Twitter platform using its manual search feature. Twitter was chosen because its search functionality allows for the addition of attributes to keywords, enabling the retrieval of results relevant to the research needs. The dataset must include a sufficient variety of cigarette products to allow the model to accurately recognize different types of cigarette products. To obtain image data for cigarette products from Twitter, collection was carried out using the previously mentioned using keywords related to cigarette products such as "udud", "sebat", and "cigarettes". A total of 2188 images were successfully collected. With a diverse dataset, the model can be trained to recognize and distinguish between various types of cigarettes and cigarette packages with greater accuracy.

Preprocessing

Preprocessing is a crucial stage in ensuring the cleanliness, structure, and readiness of data for analysis or further processing. This stage encompasses data cleaning, data annotation, and data splitting [21]. Data cleaning involves removing duplicates and irrelevant or low-quality images that could interfere with model training. It is a key phase in preparing the dataset for modeling, beginning with the elimination of duplicate images and the cleaning of images not relevant to the object of interest, such as cigarettes and their packaging. Objects obscured by other objects by 50% or more are removed. From the initial total of 2188 images, 2117 images remained after the preprocessing process. In the data annotation phase, objects in images are labeled or classified to provides information about the location and boundaries of objects to guide the YOLO model in object recognition, using RoboFlow, where 1620 cigarette objects and 1330 cigarette package objects were identified. The data is divided into a training set (1691 images), a validation set (210 images), and a testing set (216 images) with an 80:10:10 ratio to effectively allocate data during the training, validation, and testing processes. Thus, in data splitting, the dataset is processed into subsets such as training data and testing data for model evaluation and testing. Preprocessing ensures the quality, consistency of data, and prepares it for the next steps in data processing.

Training

The training process of the YOLOv5 and YOLOv8 models is a crucial step in the development of a system for identifying objects in images. There are two YOLO detection object models used in the training process, including the YOLOv5-m model and the YOLOv8-m model. Google Colab is used as the training platform with the annotated dataset. The YOLOv5 and YOLOv8 models train on features related to objects, including their shape, size, position, and class. Through various iterations, the models gradually improve their ability to detect objects with higher accuracy.

Evaluation

After the training process is completed, the model's performance is evaluated using the mean Average Precision (mAP) and F1-Score metrics. mAP is a common metric in object detection for evaluating a model's accuracy in identifying and placing bounding boxes around objects in images. To calculate mAP, the Average Precision (AP) value is calculated for each class, and then the average of these values is taken. The calculation of AP is based on the Precision-Recall (PR) [22] curve formed from the model's predictions. Detection areas with an Intersection over Union above a certain threshold are considered True Positives, while others are False Positives. Precision and Recall are calculated based on the concepts of the confusion matrix [23]. The values of Precision and Recall are calculated as follows [23]:

$$\text{Precision} = \frac{\text{True Positive}}{True\ Positive + False\ Positive} \quad (1)$$

$$\text{Recall} = \frac{\text{True Positive}}{\text{No. of Ground Truth Boxes}} \quad (2)$$

Where, True Positive is an object that are correctly detected, and False Positive is an object that are incorrectly detected by the model [23]. Precision can be described as the proportion of true positives (TP) out of all predicted positives, with false positives (FP) acting as the divisor in the calculation, as presented in Equation (1). Conversely, recall is the proportion of true positives (TP) relative to the total actual positive cases, where false negatives (FN) serve as the divisor in the denominator, as depicted in Equation (2).

Additionally, the Intersection over Union (IoU) formula measures the accuracy of object detection models by calculating the ratio of the overlap area between the predicted and ground truth bounding boxes to their union area. The Intersection over Union (IoU) between the predicted bounding box (B_p) and the ground truth bounding box (B_{gt}) is calculated using the following equation (3).

$$\text{IoU} = \frac{\text{area of overlap}}{\text{area of union}} \quad (3)$$

Where, IoU is an Intersection over Union [24], Area of Overlap is the area of overlap between the two bounding boxes, and Area of Union is the total area encompassed by the overlapping bounding boxes[25].

The AP (Average Precision) value is obtained by calculating the average precision across the PR (Precision-Recall) curve where recall is within the range [0, 0.1, 1]. The calculation of the AP value used in PASCAL VOC [26] can be expressed using equation (4).

$$\text{AP} = \frac{1}{11}\sum_{r\in R} Pinterp(r) \quad (4)$$

Where, AP is the average value of precision, and Pinterp(r) the interpolated precision at a specific recall level (r). Furthermore, to calculate precision at each recall level in the precision-recall curve with the aim of evaluating the performance of object detection models, formula (5) is required.

$$Pinterp(r) = {}^{\max p(r')}_{r'r'\geq r} \quad (5)$$

Where, Pinterp(r) is interpolated precision at a specific recall (r), Max p(r') is maximum value of precision across all recalls, and r' is recall values at different points within the range of 0 to 1.

Where $p(r')$ represents the measured precision at recall r. Based on the formula above, the value of Mean Average Precision (mAP) can be calculated as follows:

$$mAP\frac{1}{n}\sum_{i=1}^{n} APi \quad (6)$$

Where, mAP is Mean Average Precision, n is total number of predictions within a class object, and i is threshold value. Additionally, precision and recall values can be used to calculate the F1-Score using the formula in equation (7)

$$F1 - Score = 2 \times \frac{R \times P}{R + P} \qquad (7)$$

Where, R is Recall, P is Precision [27]. F1-score is used as an indicator that the object detection model exhibits high levels of precision and recall.

Arsitektur YOLOv5 dan YOLOv8

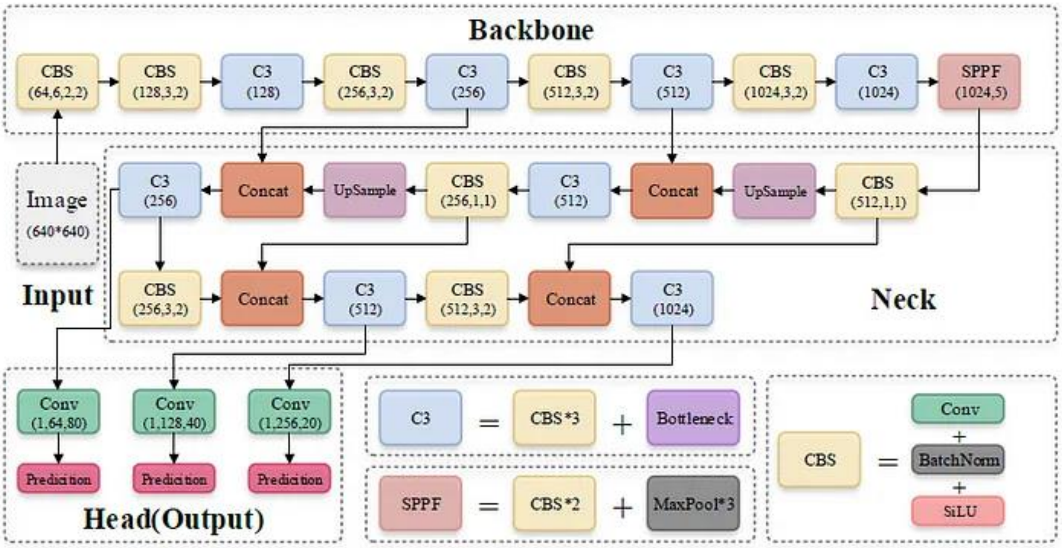


Figure 3. YOLOv5 architecture

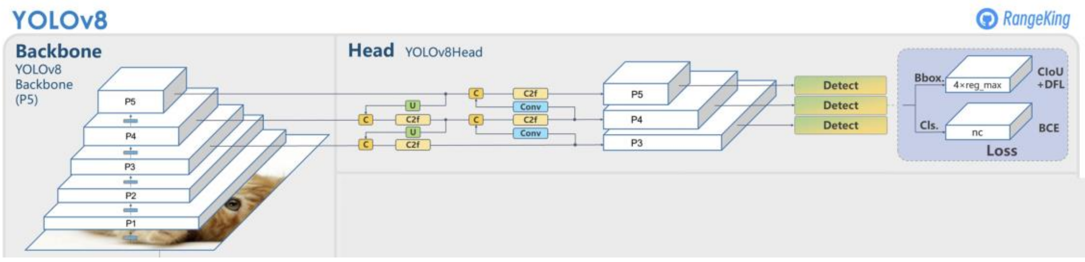


Figure 2. YOLOv8 architecture

As shown in figure 2 and 3, YOLOv5 and YOLOv8 share foundational backbone architectures for object detection, integrating modules like CSPDarknet53 [28] and SPPF, but YOLOv8 diverges with notable modifications. It updates the initial convolutional module from YOLOv5's 6x6 kernel to a 3x3 kernel and replaces the C3 module with a C2f module, which aggregates outputs from all bottleneck blocks, unlike C3's reliance on the last bottleneck block output. Both models conclude their backbone architecture with an SPPF module, maintaining a core structural similarity. In terms of neck architecture, they employ FPN and PAN for feature fusion and scale detection, enhancing the ability to detect objects of varying sizes by fusing and upscaling feature maps.

Distinctly, YOLOv8's detection process introduces significant innovations, utilizing two convolutional modules and a Conv2d layer for improved object detection. It shifts from the conventional anchor-based to an anchor-free detection system, predicting object centers directly. This shift simplifies the detection process by reducing the number of predicted boxes, thereby optimizing the Non-Maximum Suppression (NMS) process[24][29]. These modifications in YOLOv8 not only signify advancements over YOLOv5 but also highlight the continuous evolution of YOLO models to increase detection efficiency and accuracy [30], addressing complex detection tasks with refined methodologies.

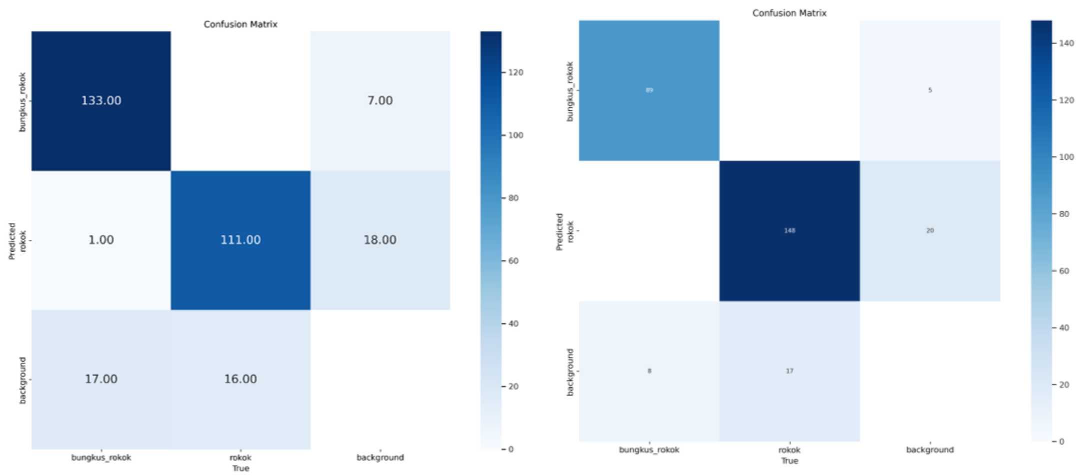
RESULTS AND DISCUSSIONS

In this research, the methodology for detecting cigarette-related content and packaging incorporates the utilization of the Python programming language, facilitating the deployment of the system. This approach is further augmented by the integration of several computational tools, including Google Colab, RoboFlow, and the PyCharm Community Integrated Development Environment (IDE). The employment of these tools, in conjunction with the dataset, necessitates the configuration of a data.yml file, serving the critical function of configuring the dataset for the training process.

The process of cloning the repository for the application of the YOLOv5 and YOLOv8 methodologies results in the generation of all requisite files for the training phase. Through the repositories of YOLOv5 and YOLOv8, the training procedure is executed by meticulously setting several parameters. These include the input image resolution (416 pixels), batch size (16), number of epochs (150), dataset configuration, and configurations pertinent to the preservation of the trained model.

Subsequent to the training phase and the acquisition of a model from said phase, the derived model is subjected to evaluation using the F1-Score and mean Average Precision (mAP) evaluation metrics. The evaluation encompasses 216 images, incorporating a total of 278 instances of cigarettes and cigarette packaging. Among these instances, 151 are attributed to cigarette packaging, while 127 pertain to cigarettes. The evaluation procedure entails the detection of objects across the 216 images, wherein the trained models predict the location and class of objects within each image. The outcomes of these predictions are juxtaposed with pre-defined bounding boxes, facilitating a comparative analysis.

This comparative analysis between the detection outcomes and the bounding boxes is instrumental in calculating the evaluation metrics, such as the F1-Score and mAP. The evaluation, predicated on this comparative framework, aims to ascertain the quality and efficacy of the YOLOv5 and YOLOv8 models as implemented within the scope of this research. The evaluation process is executed through the invocation of the val.py script, located within the YOLOv5 and YOLOv8 repositories. Execution of the val.py script mandates the adjustment of several parameters, including weights, data, img, and task, to align with the specific requirements of the evaluation phase.



(a) YOLOv5 (b) YOLOv8

Figure 4. Confusion matrix model YOLOv5 andYOLOv8

During the model evaluation process, prediction data was obtained and presented in a confusion matrix as depicted in the above figure. For the YOLOv5 model in Figure 4 (a), it was found that in the "Cigarette Pack" object category, there were 133 objects correctly detected (True Positive/TP), 18 objects missed (False Negative), and 7 other objects falsely detected as cigarette packs (False Positive). Similarly, in the "Cigarette" object category, there were 111 objects correctly detected (True Positive/TP) and 16 objects missed (False Negative). Additionally, 19 objects were falsely detected as cigarette objects in the background image (False Positive). On the other hand, for the YOLOv8 model in Figure 4 (b), in the "Cigarette Pack" object category, there were 89 objects correctly detected (True Positive/TP), 8 objects missed (False Negative), and 5 other objects falsely detected as cigarette packs (False Positive). Likewise, in the "Cigarette" object category, there were 148 objects correctly detected (True Positive/TP) and 17 objects missed (False Negative). Furthermore, 20 objects were falsely detected as cigarette objects in the background image (False Positive).

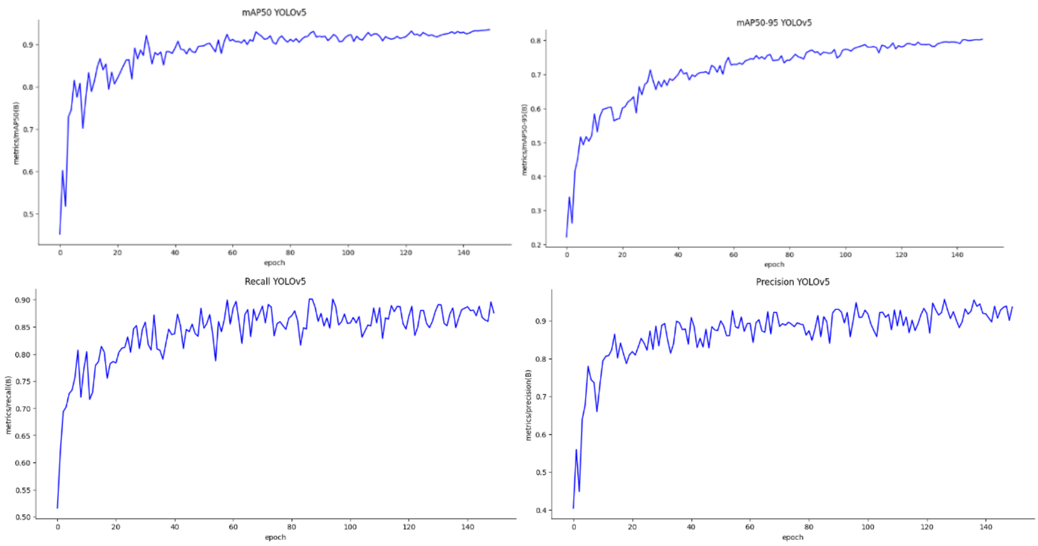


Figure 5 YOLOv5 performance result

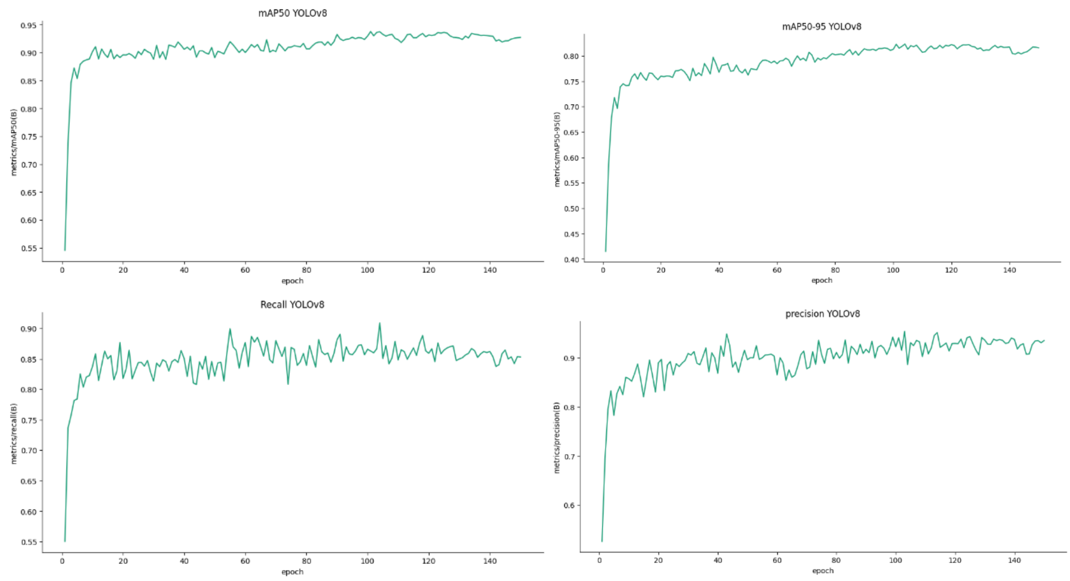


Figure 6. YOLOv8 performance result

The graphical depictions presented in Figures 6 and 7 delineate the performance metrics of precision, recall, mAP50, and mAP50-95 for both the YOLOv5 and YOLOv8 models. These figures manifest a consistent enhancement in performance values observed across the training epochs. Upon analysis of the features portrayed in Figures 6 and 7, no evidence suggesting overfitting or underfitting during the training process is discernible. This observation underscores the seamless and successful learning process experienced by the deep learning models, devoid of any significant issues.

Mean average precision YOLOv5 and YOLOv8

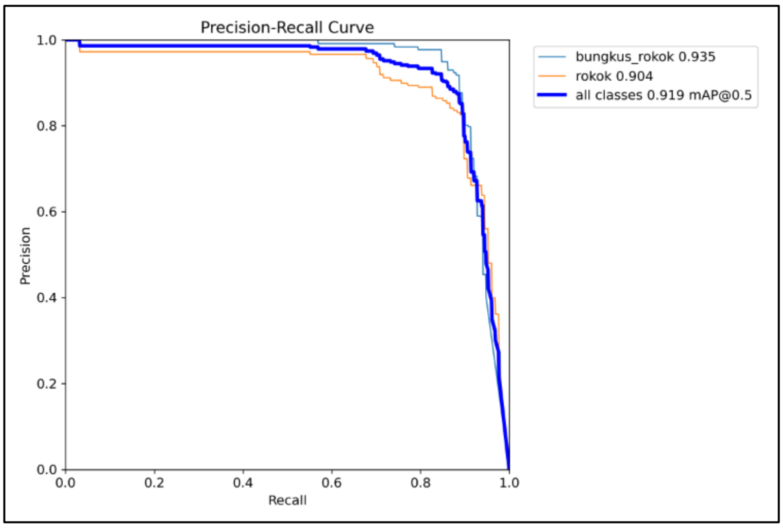


Figure 7. Precision-recall curve graph YOLOv5

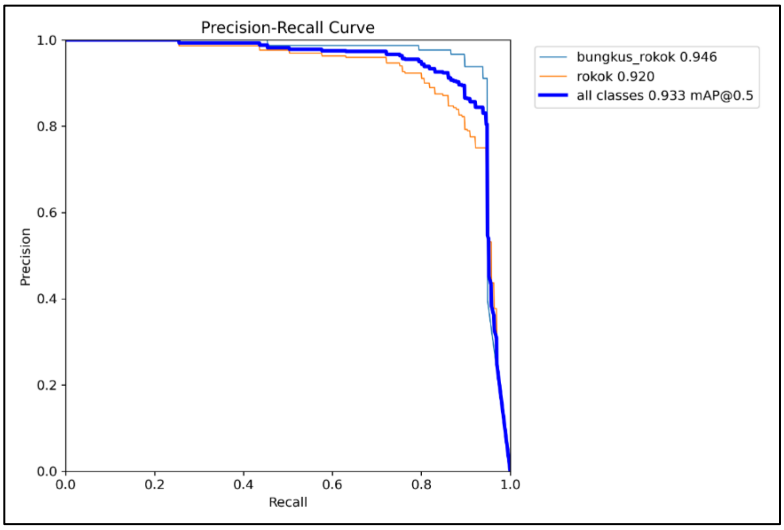
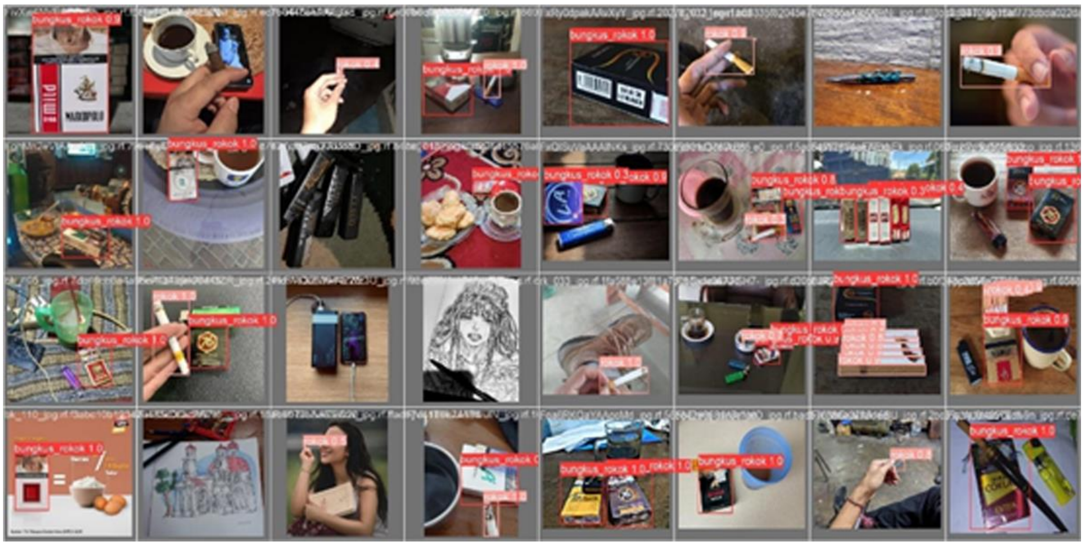


Figure 8. Precision-recall curve graph YOLOv8

Figure 8 illustrates the precision-recall curve of the test results, where the Average Precision (AP) values for the "cigarette_pack" class are 0.935 and for the "cigarette" class are 0.904. Thus, the mAP of the cigarette detection model can be calculated from equation (6). From the aforementioned mAP calculation, the evaluation mAP value of the YOLOv5 model in detecting cigarette objects is 0.919. Figure 9 demonstrates the PR curve of the test results, where the AP value for the "bungkus_rokok" class is 0.946 and for the "rokok" class is 0.920. Consequently, the mAP of the cigarette detection can be computed from equation (6). From the above mAP calculation, the evaluation mAP value of the YOLOv8 model in detecting cigarette objects is 0.933.



(a) YOLOv5 model (b) YOLOv8 model

Figure 9. Model detection result YOLOv5 and YOLOv8

Therefore, based on the values of True Positive, True Negative, False Positive, and False Negative from the previous confusion matrix, calculations can be performed to obtain precision and recall values. Precision and recall serve as measures to determine the value of the F1-Score matrix, with the computed results as follows:

Table 1. Comparison of model performance evaluation

Model	Precision	Recall	F1-Score	mAP
YOLOv5	0.901	0.877	0.888	0.919
YOLOv8	0.913	0.907	0.909	0.933

Table 2. Comparison of model performance evaluations for each object

Object	YOLOv5				YOLOv8			
	Precision	Recall	F1-Score	mAP	Precision	Recall	F1-Score	mAP
Cigarette Packs	0.95	0.880	0.913	0.935	0.946	0.917	0.931	0.946
Cigarette	0.853	0.874	0.864	0.904	0.880	0.896	0.887	0.920

Observing Table 1 and Table 2, where the calculations of precision, recall, and F1-Score for the YOLOv5 model concerning cigarette packs and cigarettes are presented, intriguing results are obtained. The precision for cigarette packs is 0.95, while for cigarettes, it is 0.853, with an average precision of 0.901. Similarly, the recall for cigarette packs is 0.88, and for cigarettes, it is 0.874, resulting in an average recall of 0.877. As for the F1-Score, the values for cigarette packs and cigarettes are 0.913 and 0.864, respectively, yielding an average F1-Score of 0.888. Meanwhile, in the precision calculation for the YOLOv8 model, it is found that the precision values for cigarette packs and cigarettes are 0.946 and 0.88, respectively. Thus, the average precision for both objects is 0.913. The recall calculation shows that the recall values for cigarette packs and cigarettes are 0.917 and 0.896, respectively, resulting in an average recall of 0.907 for both objects. The F1-Score values for cigarette packs and cigarettes are 0.931 and 0.887, respectively, resulting in an average F1-Score of 0.909 for both objects.

The test results above show that the model performance is good. this can be decided after comparing with previous research where Z Zhang et al. detect smoking activity by detecting cigarette objects in an image. In this research, a custom deep learning model was used that used the CSP + CBAM backbone model framework and produced a model with mAP value of 86.32% [30]. The effectiveness of the YOLO model in detecting content related to smoking as seen from the test results data above, this model can be used to help carry out automatic detection and censorship of content spread on social media or broadcasts on television.

CONCLUSION

This study conducted a detailed comparison between the YOLOv5 and YOLOv8 models in detecting cigarette and cigarette pack objects within a dataset of 216 images. The YOLOv5 model demonstrated a high degree of accuracy, with True Positive rates of 133 for cigarette packs and 111 for cigarettes, while the YOLOv8 model showed True Positive rates of 89 for cigarette packs and 148 for cigarettes. The evaluation metrics, including precision, recall, F1-Score, mAP50, and mAP50-95, indicated consistent performance improvements across training epochs for both models, with no evidence of overfitting or underfitting. The precision-recall curves revealed that the YOLOv5 model achieved a mAP value of 0.919, whereas the YOLOv8 model attained a slightly higher mAP value of 0.933, suggesting superior performance in object detection accuracy. These results highlight the effectiveness of both models in detecting cigarette-related content, with YOLOv8 demonstrating a slight edge in overall performance metrics [31].

In addition to the technical findings, it is important to address the ethical considerations inherent in implementing an automated system for censoring cigarette content on social media platforms. The use of these models raises concerns about the potential for censorship of non-broadcast content, thereby inadvertently limiting freedom of expression. Due to the Datasets used for model training and evaluation may not fully capture the variability and complexity of tobacco-related imagery on social media platforms. Future research is expected to address these limitations by using larger and more diverse datasets and conducting rigorous sensitivity analyzes as model performance may be influenced by factors such as image quality, lighting conditions, and object occlusion.

ACKNOWLEDGEMENT

The author extends deep appreciation to the Research and Service Community Institute (LPPM) at Universitas Dian Nuswantoro for their support. Financial backing for this study came from the Basic Research Grant Scheme, under contract number 109/A.38-04/UDN-09/XI/2023. Additionally, sincere thanks are offered to the entire team for their contributions

REFERENCES

[1] Kathy Crosby, “Industry influencer: how tobacco content is infiltrating social media,” Truth Initiative. [Online]. Available: https://truthinitiative.org/research-resources/tobacco-pop-culture/industry-influencer-how-tobacco-content-infiltrating-social

[2] Hanadian Nurhayati-Wolf, “Breakdown of social media users by age and gender in Indonesia as of January 2021 Breakdown of social media users by age and gender Indonesia 2021.” [Online]. Available: https://www.statista.com/statistics/997297/indonesia-breakdown-social-media-users-age-gender/

[3] Y. Achmad, “Rencana Aksi Program,” KEMENKES RI 2024. [Online]. Available: https://e-renggar.kemkes.go.id/file_performance/1-465827-01-3tahunan-178.pdf

[4] K. Laili, S. P. Ratih, R. W. Gayatri, and S. Adi, “The impact of exposure to cigarette advertising and promotion on youth smoking behavior in Malang Regency (Indonesia) during the COVID-19 Pandemic,” J. Public Health Afr., vol. 13, no. s2, 2022, doi: 10.4081/jphia.2022.2409.

[5] P. Cavazos-rehg et al., “Exploring How Social Media Exposure and Interactions Are Associated With ENDS and Tobacco Use in Adolescents From the PATH Study,” pp. 487–494, 2021, doi: 10.1093/ntr/ntaa113.

[6] T. Hassan et al., “Meta-transfer learning driven tensor-shot detector for the autonomous localization and recognition of concealed baggage threats,” Sens. Switz., vol. 20, no. 22, pp. 1–25, 2020, doi: 10.3390/s20226450.

[7] K. Han, Y. Wang, Q. Tian, J. Guo, C. Xu, and C. Xu, “GhostNet: More features from cheap operations,” Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit., pp. 1577–1586, 2020, doi: 10.1109/CVPR42600.2020.00165.

[8] D. Wang, J. G. Wang, and K. Xu, “Deep learning for object detection, classification and tracking

in industry applications," Sensors, vol. 21, no. 21, pp. 3–5, 2021, doi: 10.3390/s21217349.
[9] T. Diwan, G. Anirudh, and J. V. Tembhurne, "Object detection using YOLO: challenges, architectural successors, datasets and applications," Multimed. Tools Appl., vol. 82, no. 6, pp. 9243–9275, 2023, doi: 10.1007/s11042-022-13644-y.
[10] H. Deshpande, A. Singh, and H. Herunde, "Comparative analysis on YOLO object detection with OpenCV," Int. J. Res. Ind. Eng., vol. 9, no. 1, pp. 46–64, 2020, doi: 10.22105/riej.2020.226863.1130.
[11] P. P. Khan, Mohammad Ayoub, Rijwan Khan, Artificial Intelligence Applications in Agriculture and Food Quality Improvement. Engineering Science Reference, 2022. doi: 10.4018/978-1-6684-5141-0.
[12] J. A. Kim, J. Y. Sung, and S. H. Park, "Comparison of Faster-RCNN, YOLO, and SSD for Real-Time Vehicle Type Recognition," 2020 IEEE Int. Conf. Consum. Electron. - Asia ICCE-Asia 2020, pp. 8–11, 2020, doi: 10.1109/ICCE-Asia49877.2020.9277040.
[13] L. Tan, T. Huangfu, and L. Wu, "Comparison of YOLO v3, faster R-CNN, and SSD for real-time pill identification," arXiv, 2021.
[14] M. Ş. Gündüz and G. Işık, "A new YOLO-based method for social distancing from real-time videos," Neural Comput. Appl., vol. 35, no. 21, pp. 15261–15271, 2023, doi: 10.1007/s00521-023-08556-3.
[15] K. Kim, K. Kim, and S. Jeong, "Application of YOLO v5 and v8 for Recognition of Safety Risk Factors at Construction Sites," Sustainability, vol. 15, no. 20, p. 15179, 2023, doi: 10.3390/su152015179.
[16] I. P. Sary, S. Andromeda, and E. U. Armin, "Performance Comparison of YOLOv5 and YOLOv8 Architectures in Human Detection using Aerial Images," Ultima Comput. J. Sist. Komput., vol. 15, no. 1, pp. 8–13, 2023, doi: 10.31937/sk.v15i1.3204.
[17] L. He, X. Ren, Q. Gao, X. Zhao, B. Yao, and Y. Chao, "The connected-component labeling problem: A review of state-of-the-art algorithms," Pattern Recognit., vol. 70, pp. 25–43, 2017, doi: 10.1016/j.patcog.2017.04.018.
[18] S. Alqethami, S. Alghamdi, T. Alsubait, and H. Alhakami, "RoadNet: Efficient Model to Detect and Classify Road Damages," Appl. Sci. Switz., vol. 12, no. 22, 2022, doi: 10.3390/app122211529.
[19] O. A. Montesinos López, A. Montesinos López, and J. Crossa, Multivariate Statistical Machine Learning Methods for Genomic Prediction. 2022. doi: 10.1007/978-3-030-89010-0.
[20] R. Zhao, S. H. Tang, E. E. Bin Supeni, S. B. A. Rahim, and L. Fan, A Review of Object Detection in Traffic Scenes Based on Deep Learning, vol. 9, no. 1. Multimedia Tools and Applications, 2024. doi: 10.2478/amns-2024-0322.
[21] D. E. Birba, "A Comparative study of data splitting algorithms for machine learning model selection," Degree Proj. Comput. Sci. Eng., vol. 2020, no. 1, pp. 1–23, 2020.
[22] H. Zhu, H. Wei, B. Li, X. Yuan, and N. Kehtarnavaz, "A review of video object detection: Datasets, metrics and methods," Appl. Sci. Switz., vol. 10, no. 21, pp. 1–24, 2020, doi: 10.3390/app10217834.
[23] Y. Zhang, T. Zuo, L. Fang, J. Li, and Z. Xing, "An Improved MAHAKIL Oversampling Method for Imbalanced Dataset Classification," IEEE Access, vol. 9, pp. 16030–16040, 2021, doi: 10.1109/ACCESS.2020.3047741.
[24] H. Liang, T. Wu, Q. Zhang, and H. Zhou, "Non-Maximum Suppression Performs Later in Multi-Object Tracking," Appl. Sci. Switz., vol. 12, no. 7, 2022, doi: 10.3390/app12073334.
[25] M. S. Rana, A. Nibali, and Z. He, "Selection of object detections using overlap map predictions," Neural Comput. Appl., vol. 34, no. 21, pp. 18611–18627, 2022, doi: 10.1007/s00521-022-07469-x.
[26] R. Padilla, W. L. Passos, T. L. B. Dias, S. L. Netto, and E. A. B. Da Silva, "A comparative analysis of object detection metrics with a companion open-source toolkit," Electron. Switz., vol. 10, no. 3, pp. 1–28, 2021, doi: 10.3390/electronics10030279.
[27] Z. Ning, X. Wu, J. Yang, and Y. Yang, "MT-YOLOv5: Mobile terminal table detection model based on YOLOv5," J. Phys. Conf. Ser., vol. 1978, no. 1, 2021, doi: 10.1088/1742-6596/1978/1/012010.
[28] D. Caballero-Ramirez, Y. Baez-Lopez, J. Limon-Romero, G. Tortorella, and D. Tlapa, "An Assessment of Human Inspection and Deep Learning for Defect Identification in Floral Wreaths," Horticulturae, vol. 9, no. 11, pp. 1–21, 2023, doi: 10.3390/horticulturae9111213.
[29] D. Kumawat, D. Abhayankar, and S. Tanwani, "INTELLIGENT SYSTEMS AND APPLICATIONS IN ENGINEERING Exploring Object Detection Algorithms and implementation

of YOLOv7 and YOLOv8 based model for weapon detection," vol. 12, no. 3, pp. 877–886, 2024.
[30] Z. Zhang, H. Chen, R. Xiao, and Q. Li, "Research on Smoking Detection Based on Deep Learning," J. Phys. Conf. Ser., vol. 2024, no. 1, p. 012042, Sep. 2021, doi: 10.1088/1742-6596/2024/1/012042.
[31] E. Arulprakash and M. Aruldoss, "A study on generic object detection with emphasis on future research directions," J. King Saud Univ. - Comput. Inf. Sci., vol. 34, no. 9, pp. 7347–7365, 2022, doi: 10.1016/j.jksuci.2021.08.001.