INTELLIGENT SURVEILLANCE FOR MASK REGULATION IN HEALTHCARE USING THE YOLOV11 ALGORITHM

Afu Ichsan Pradana[1*], Harsanto[2], Burhanuddin Bin Mohd Aboobaider[3], Malika Harsanto[4]
Duta Bangsa University[1,2], Universiti Teknikal Malaysia Melaka[3], Universitas Muhammadiyah Surakarta[4]
*Correspondence Email : afu_ichsan@udb.ac.id

ABSTRACT

The use of face masks in healthcare settings is a crucial measure in preventing the spread of infectious diseases, particularly since the outbreak of the COVID-19 pandemic. However, public compliance with mask-wearing remains a challenge despite the implementation of various regulations. This study aims to design and develop an automatic mask-wearing detection system by leveraging the YOLOv11 algorithm, which is renowned for its superior speed and accuracy in object detection. The methodology involved collecting a dataset of facial images with and without masks, data labeling, model training using YOLOv11, and evaluating the system's performance in real-world conditions. Test results demonstrate that the system can perform real-time mask detection with a mean Average Precision (mAP) of 0.9, establishing it as an effective solution for supporting health protocol monitoring in medical facilities. Consequently, this system not only enhances monitoring efficiency but also has the potential to minimize the risk of infection spread through an intelligent technological approach.

KEYWORDS mask detection, YOLOv11, computer vision, healthcare facilities, health protocols

This work is licensed under a Creative Commons Attribution-ShareAlike 4.0 International

INTRODUCTION

The COVID-19 pandemic, which has impacted the world since late 2019, has had significant repercussions on various aspects of life, including public health[1]. One of the most effective preventive measures to reduce viral transmission is the use of face masks. According to the World Health Organization (WHO), wearing a mask can reduce virus transmission via droplets produced when talking, coughing, or sneezing[2]. n Indonesia, the government has issued various policies to encourage public compliance with health protocols, including the use of masks in public areas and healthcare facilities[3]. However, despite strict regulations, many individuals still neglect to wear masks, especially in crowded places such as clinics and hospitals.

Given this problem, it is crucial to develop a system that can automatically detect mask usage to improve compliance with health protocols. One applicable technology is a camera-based detection system utilizing machine learning algorithms[4]. Since the onset of the COVID-19 pandemic, detection systems for face mask usage have seen significant advancements in the domains of image processing and computer vision[5], as computer vision has proven to be a revolutionary aspect of modern technology[6]. Currently, several approaches are being applied for face mask detection, including Convolutional Neural Networks (CNN)[5], [7], [8], [9], Viola-Jones or Haar Cascade [10], [11], [12], [13], [14], [15], [16], *You Only Look Once (YOLO)* [17], and Hybrid Deep Transfer Learning techniques[5], [18]. Among the various methods used for face detection and image recognition in the context of mask-wearing, the YOLO method has proven effective in various applications, including face and object detection [19]. With technological advancements, the latest version of this algorithm, YOLOv11, offers improvements in detection speed and accuracy[20], making it an ideal tool for implementation in a mask detection system. By leveraging the capabilities of YOLOv11, this system can assist healthcare workers and medical facility managers in efficiently monitoring mask-wearing compliance.

Based on this problem discussion, this research aims to develop a system that can accurately detect mask usage to improve public compliance in wearing masks within healthcare settings[21], In practical terms, the developed system can be used by clinics and other medical facilities to enhance adherence to health protocols, thereby reducing the risk of virus spread [3] within these environments.

RELATED WORKS

Various technological approaches have been developed to help monitor and improve compliance with health protocols, particularly the use of face masks. One widely used approach is computer vision-based detection systems supported by deep learning algorithms. Research by Xie, Y., et al. [22] developed a mask-wearing detection method using YOLOv5 with data augmentation techniques and an adapted network architecture, achieving high detection accuracy. Similarly, Xu, et al. [23] introduced an enhancement to YOLOv5 by integrating an attention mechanism, resulting in a lightweight yet accurate model for mask detection. In an effort to increase system robustness against data variations, Yang, et al. [24] proposed a Triplet-Consistency Representation Learning method that proved effective even with limited training data.

Meanwhile, Al-Tamimi and Ali [25] utilized the YOLOv5s model for mask detection, yielding promising results when applied to a combined dataset. Kowalczyk and Rumiński [26] introduced a unique approach by using thermal imaging for mask detection based on YOLOv5 nano, demonstrating the method's effectiveness in low-light conditions. Research by Fan and Jiang [27] introduced RetinaFaceMask, a single-stage detection system capable of distinguishing between correct and incorrect mask usage, thereby expanding the scope of mask surveillance systems more accurately.

Furthermore, a study by Singh, et al. [28] compared the effectiveness of the YOLOv3 model with Faster R-CNN in the context of the pandemic and concluded that YOLOv3 was superior for real-time detection. Wang, et al. [29] integrated YOLOv5 with ResNet-50 to build a unified monitoring system capable of detecting mask usage and social distancing simultaneously. Additionally, Atrey, et al. [30] evaluated the real-world performance of

YOLOv6 and demonstrated its capability to enhance compliance with health protocols. Dewi, et al. [31] also implemented the YOLOv7 model, noting a significant improvement in the speed and accuracy of real-time face mask detection.

From these various studies, it is evident that YOLO technology has advanced rapidly and is widely used in face mask detection systems. However, to date, few studies have specifically examined the effectiveness of its latest version, YOLOv11, in detecting mask usage, particularly in healthcare settings which have unique characteristics and stricter monitoring requirements. Therefore, this research is relevant for addressing this gap by exploring and implementing the YOLOv11 algorithm in a mask detection system that is accurate, efficient, and contextualized to the needs of medical facilities in Indonesia.

RESEARCH METHOD

This study employs a deep learning approach utilizing YOLOv11, chosen for its significant advantages in real-time applications. The research was conducted using the Python 3 programming language, leveraging cloud computing on Google Colab with an Nvidia Tesla T4 GPU (15 GB capacity) and 13 GB RAM. The research methodology is illustrated in Figure 1.

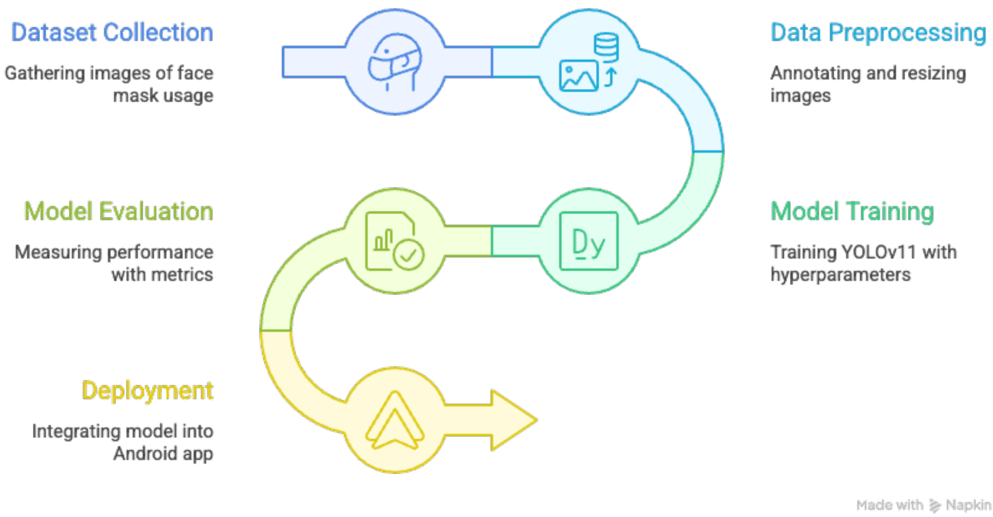

Figure 1. Research methodology

Figure 1 illustrates the research methodology for this paper, which includes dataset collection, data preprocessing, model development using the YOLOv11 algorithm, model evaluation, and deployment to an Android-based mobile application.

a. Dataset Collection

The dataset used in this study is a public dataset containing images of face mask usage. The dataset consists of 4,850 images: 3,779 images of faces with masks and 1,071 images without masks. The dataset was split with a ratio of 80% for training, 10% for validation, and 10% for testing.

b. Preprocessing

Before model training, preprocessing was performed to ensure input quality through the following steps:

1. **Annotation:** In this stage, each image in the dataset was labeled accordingly to help the model distinguish objects from different classes using the Roboflow Annotate labeling tool.
2. **Resizing:** This stage involved resizing the input images to 640 x 640 pixels.

c. Model Training

In this stage, the labeled dataset was fed into the YOLOv11 model for training. The YOLO detection model can be formulated as follows:

$$Pr(object) * IOU_{pred}^{truth}$$

(1)

where IoU (Intersection over Union) is the ratio of the intersection area to the union area between the predicted bounding box and the ground truth box. IoU is a value between 0 and 1; a value closer to 1 indicates the predicted bounding box is more accurate. For the same intersection, along with the bounding box, each grid cell also specifies conditional class probabilities (C). The class-specific probabilities for each grid cell [32] are defined as follows:

$$Pr(class_i|object) * Pr(object) * IOU_{pred}^{truth} = Pr(class_i) * IOU_{pred}^{truth}$$

(2)

The confidence score represents the probability that a class appears in the box and how well the predicted box fits the object. The next step involved selecting the YOLOv11 weights used, specifically YOLOv11n, with the following hyperparameter tuning: batch size = 16, learning rate = 0.01, momentum = 0.937, and epochs = 30.

d. Model Evaluation

Model evaluation was used to measure the performance of the YOLO algorithm using three evaluation metrics: Precision (P), Recall (R), and mean Average Precision (mAP) [33], which are formulated as follows:

1) $$Precision \frac{True\ Positives\ (TP)}{True\ Positives\ (TP) + False\ Positive\ (FP)} \quad (3)$$

2) $$Recall \frac{True\ Positives\ (TP)}{True\ Positives\ (TP) + False\ Negatives\ (FN)} \quad (4)$$

3) *Mean Average Precision(mAP)*

$$AP = \int_0^1 P(R)dR \quad (5)$$

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

e. Deployment

The final stage involved integrating the trained model into an Android-based mobile application.

RESULT AND DISCUSSION

a. Result

The training process using YOLOv11 produced a mask usage detection model. The performance metrics for each class are presented in Table 1

Table 1. Model Training Results

Class	Precision	Recall	mAP
No Mask	0.942	0.714	0.843
With Mask	0.916	0.885	0.954

Based on the data in Table 1, for the "No Mask" class, the precision value reached 0.942. This indicates that 94.2% of all detections predicted as "no mask" were correct, meaning the model has a low error rate in misclassifying objects that do not belong to this class (low false positives). However, the recall for this class is 0.714, meaning that approximately 28.6% of the objects that were actually not wearing masks were not detected by the model (high false negatives). The mAP (mean Average Precision) value for this class is 0.843, reflecting a reasonably good overall detection performance, though there is room for improvement, particularly in recall.

Conversely, for the "With Mask" class, the model demonstrated superior performance with a precision of 0.916 and a recall of 0.885. This indicates that the vast majority of the model's predictions are correct (91.6%) and its detection coverage is extensive (88.5%), resulting in very few undetected cases. The mAP value for this class reached 0.954, signifying a very high level of detection accuracy and consistency.

Overall, the model performs more optimally in detecting individuals who are wearing masks compared to those who are not. In conclusion, this detection system is sufficiently reliable for monitoring mask-wearing compliance, particularly in healthcare service environments. The higher false negative rate for the "No Mask" class suggests a potential area for future work, such as data augmentation focused on uncovered faces or model fine-tuning, to improve the recall without significantly compromising the high precision.

b. Discussion

1) Precision, Recall, and mAP

The accuracy of object detection is highly dependent on the balance between precision and recall, where the system is required to recognize objects quickly without compromising accuracy. The graphs also indicate that model performance tends to stabilize when the confidence level is very high[34].

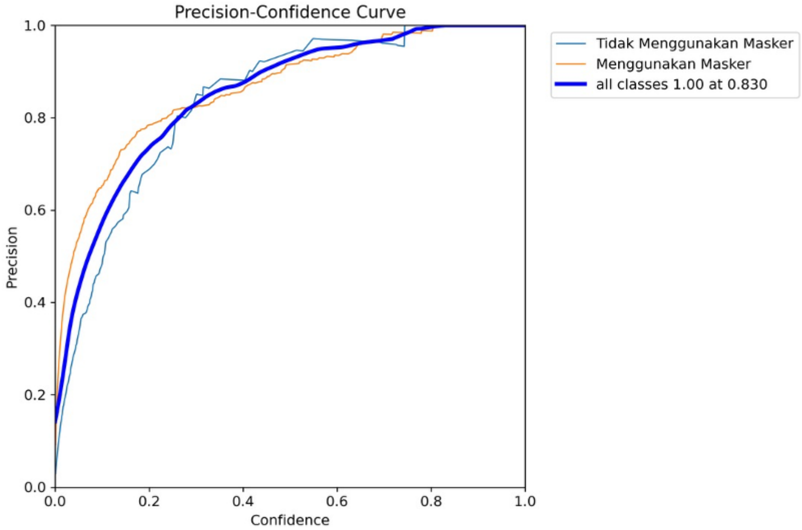

Figure 2. Precision-Confidence Curve

Figure 2 shows that as the confidence value increases, the precision for both classes also increases significantly. At confidence values around 0.8 and above, precision approaches the maximum value (1.0), indicating that the model's predictions at high confidence levels are highly accurate.

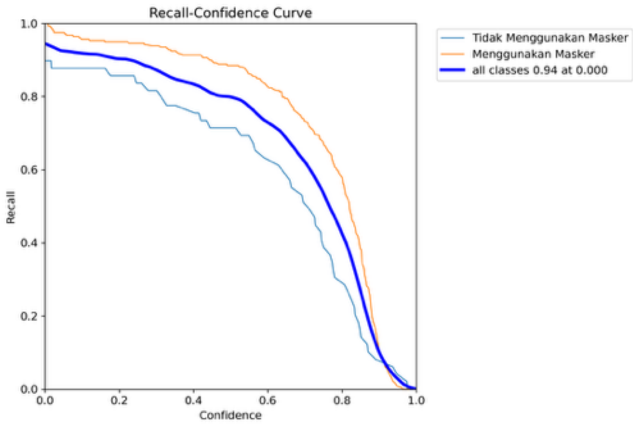

Figure 3. Recall-Confidence Curve

Figure 3 demonstrates that when the confidence value is low, the recall value is high, meaning the model successfully detects almost all target objects, albeit with potentially more false positives. However, as the confidence value increases, recall decreases sharply. The orange curve represents the "With Mask" class, which consistently has higher recall values compared to the light blue curve representing the "Without Mask" class. This indicates that the model performs better at detecting individuals wearing masks compared to those not wearing masks, especially at higher confidence levels. The dark blue curve shows the average recall across all classes, which peaks at 0.94 at a confidence of 0.00, indicating that the model is most comprehensive in detecting objects when no confidence filtering is applied.

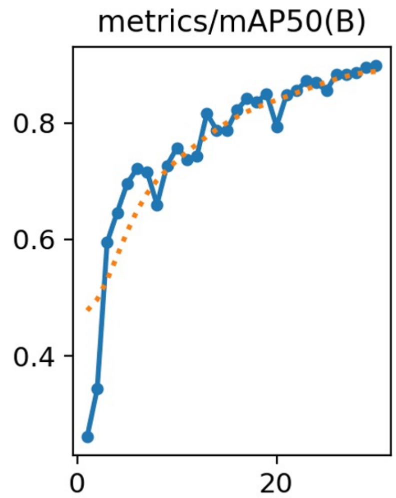

Figure 4. mAP Metrics Graph

From Figure 4, it can be observed that the mAP value increases significantly in the initial epochs. This indicates that the model quickly learns to recognize important patterns from the training data. After around the 10th epoch, the rate of improvement begins to slow, but consistent progress continues until it approaches the maximum value of 0.9 in the final epochs. This steady increase in mAP50 indicates that the model underwent effective training without clear signs of overfitting. The high mAP value confirms that the model possesses accurate detection capabilities at a bounding box similarity threshold of 50% or higher, which aligns with general evaluation standards for object detection tasks.

2) Confusion Matrix

The following Confusion Matrix provides a detailed analysis of the model's performance in classifying each class[35].

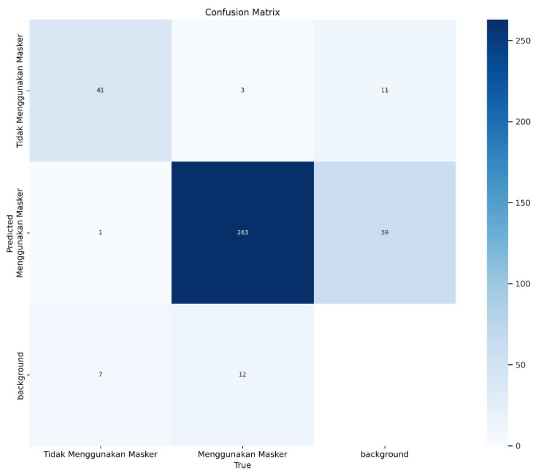

Figure 5. Confusion Matrix

The model correctly classified 41 samples for the "Without Mask" class and 263 samples for the "With Mask" class. This indicates that the model performs exceptionally well in detecting mask usage but requires improvement in distinguishing background areas to reduce irrelevant detection errors. The high number of true positives for the "With Mask" class underscores its reliability, while the lower true positive count for the "Without Mask" class suggests a need for enhanced feature learning or data balancing to improve detection of uncovered faces. This analysis aligns with the precision-recall metrics, confirming the model's strength in identifying compliance and its potential for refinement in identifying non-compliance.

3) Deployment

Following the completion of model training, the subsequent step was to convert the model into a format compatible with Android devices, namely TFLite. The conversion process yielded a TFLite model with the specifications presented in Table 2.

Table 2. Model Conversion Results

Model	Size
YOLOv11	26.6 MB
TFLite	10.2 MB

The next step involved integrating the resulting TFLite file into an Android application. This research utilized Android Studio as the Integrated Development Environment (IDE) with the Kotlin programming language. The interface of the final application is shown in Figure 5.

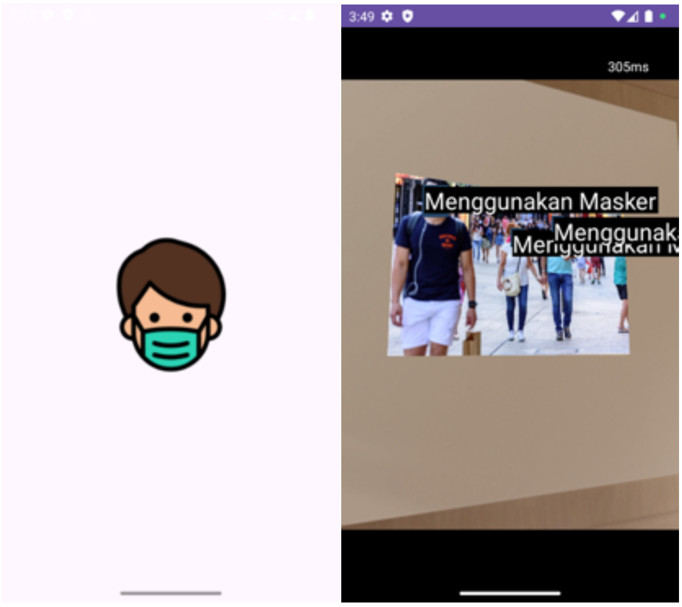

Figure 5. Android Application Interface

The application interface displays a live camera feed with real-time bounding box detection and class labels ('With Mask' / 'Without Mask') over detected faces. A confidence score is displayed for each detection.

CONCLUSION

This research successfully developed a mask-wearing detection system using the YOLOv11 algorithm, capable of high-accuracy real-time detection in healthcare environments. The testing results demonstrate that the system is effective in distinguishing between individuals who are wearing masks and those who are not, thereby supporting the automated enforcement of health protocols. Consequently, this technology holds significant potential to enhance monitoring and compliance with mask-wearing policies, ultimately minimizing the risk of infection spread in medical facilities.

REFERENCES

N. F. Abdul Hassan, A. A. Abed, and T. Y. Abdalla, “Face mask detection using deep learning on NVIDIA Jetson Nano,” *Int. J. Electr. Comput. Eng.*, vol. 12, no. 5, pp. 5427–5434, 2022, doi: 10.11591/ijece.v12i5.pp5427-5434.

Nena Ayu Sabrina, Yuliska Uswatun Hasanah, and Meilina Hafifah, “Efektivitas Penggunaan Masker dalam Mencegah Penyebaran Infeksi Saluran Pernapasan Atas di Lingkungan Sekolah,” *J. Vent.*, vol. 2, no. 2, pp. 36–44, Jun. 2024, doi: 10.59680/ventilator.v2i2.1160.

S. Balovsyak, O. Derevyanchuk, V. Kovalchuk, H. Kravchenko, and M. Kozhoka, “Face Mask Recognition by the Viola-Jones Method Using Fuzzy Logic,” *Int. J. Image Graph. Signal Process.*, vol. 16, no. 3, pp. 39–51, 2024, doi: 10.5815/ijigsp.2024.03.04.

L. B. Adrianto, M. I. Wahyuddin, and W. Winarsih, “Implementasi Deep Learning untuk Sistem Keamanan Data Pribadi Menggunakan Pengenalan Wajah dengan Metode Eigenface Berbasis Android,” *J. JTIK J. Teknol. Inf. Dan Komun.*, vol. 4, no. 2, p. 89, Jan. 2021, doi: 10.35870/jtik.v5i1.201.

P. Nagrath, R. Jain, A. Madan, R. Arora, and P. Kataria, “SSDMNV2: A real time DNN-based face mask detection system using single shot multibox detector and MobileNetV2,” *Sustain. Cities Soc.*, no. January, 2021.

A. Chavda, J. Dsouza, S. Badgujar, and A. Damani, “Multi-Stage CNN Architecture for Face Mask Detection,” *2021 6th Int. Conf. Converg. Technol. I2CT 2021*, 2021, doi: 10.1109/I2CT51068.2021.9418207.

P. R. Togatorop and A. Fauzi, “Klasifikasi Penggunaan Masker Wajah Menggunakan Squeezenet,” *JATISI J. Tek. Inform. Dan Sist. Inf.*, vol. 9, no. 1, pp. 397–406, 2022, doi: 10.35957/jatisi.v9i1.642.

P. Nyoman and Putu Kusuma Negara, “Deteksi Masker Pencegahan Covid19 Menggunakan Convolutional Neural Network Berbasis Android,” *J. RESTI Rekayasa Sist. Dan Teknol. Inf.*, vol. 5, no. 3, pp. 576–583, 2021, doi: 10.29207/resti.v5i3.3103.

R. A. A. Ramdhani Reza Rizqi, Adam Riza Ibnu, “Deep Learning Implementation For Face Mask Detection,” *J. Inf. Technol. Comput. Sci. INTECOMS*, vol. 4, no. 2, 2021.

T. Arifianto, “Penerapan Algoritma Viola-Jones Untuk Deteksi Masker Covid-19 Di Politeknik Perkeretaapian Indonesia Madiun,” *JATISI J. Tek. Inform. Dan Sist. Inf.*, vol. 8, no. 4, pp. 2030–2040, 2021, doi: 10.35957/jatisi.v8i4.1106.

F. Luthfillah Ahmad, A. Nugroho, and dan Alfa Faridh Suni, “Deteksi Pemakai Masker Menggunakan Metode Haar Cascade Sebagai Pencegahaan COVID 19,” *Edu Elektr. J.*, vol. 10, no. 1, pp. 13–18, 2021.

C.-D. Politeknik and P. Indonesia, “Penerapan Algoritma Viola-Jones Untuk Deteksi Masker,” *J. Tek. Inform. Dan Sist. Inf.*, vol. 8, no. 4, pp. 2030–2040, 2021.

R. P. M. D. Labib, S. Hadi, and P. D. Widayaka, “Low Cost System for Face Mask Detection Based Haar Cascade Classifier Method,” *MATRIK J. Manaj. Tek. Inform. Dan Rekayasa Komput.*, vol. 21, no. 1, pp. 21–30, 2021, doi: 10.30812/matrik.v21i1.1187.

G. Aprilian Anarki, K. Auliasari, and M. Orisa, “Penerapan Metode Haar Cascade Pada Aplikasi Deteksi Masker,” *JATI J. Mhs. Tek. Inform.*, vol. 5, no. 1, pp. 179–186, 2021, doi: 10.36040/jati.v5i1.3214.

A. Thariq and R. Y. Bakti, “Sistem Deteksi Masker dengan Metode Haar Cascade pada Era New Normal COVID-19,” *J. Sist. Dan Teknol. Inf.*, vol. 9, no. 2, pp. 241–244, 2021, doi: 10.26418/justin.v9i2.44309.

M. Ulum, M. Imaduddin, H. Sukri, and A. F. Ibadillah, “Deteksi Suhu Tubuh Dan Masker Otomatis Dengan Metode Haar Casecade Sebagai Solusi Pencegahan Penularan Covid-19,” *J. Ris. Rekayasa Elektro*, vol. 3, no. 2, pp. 119–126, 2021.

S. Singh, U. Ahuja, M. Kumar, K. Kumar, and M. Sachdeva, “Face mask detection using YOLOv3 and faster R-CNN models: COVID-19 environment,” *Multimed. Tools Appl.*, vol. 80, no. 13, pp. 19753–19768, 2021, doi: 10.1007/s11042-021-10711-8.

G. Jignesh Chowdary, N. S. Punn, S. K. Sonbhadra, and S. Agarwal, “Face Mask Detection Using Transfer Learning of InceptionV3,” *Lect. Notes Comput. Sci. Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinforma.*, vol. 12581 LNCS, pp. 81–90, 2020, doi: 10.1007/978-3-030-66665-1_6.

Z. Han, H. Huang, Q. Fan, Y. Li, Y. Li, and X. Chen, “SMD-YOLO: An efficient and lightweight detection method for mask wearing status during the COVID-19 pandemic,” *Comput. Methods Programs Biomed.*, vol. 221, 2022, doi: 10.1016/j.cmpb.2022.106888.

R. Khanam and M. Hussain, “YOLOv11: An Overview of the Key Architectural Enhancements,” Oct. 23, 2024, *arXiv*: arXiv:2410.17725. doi: 10.48550/arXiv.2410.17725.

C. Dewi and H. J. Christanto, “Automatic medical face mask recognition for COVID-19 Mitigation: Utilizing YOLO V5 object detection,” *Rev. Intell. Artif.*, vol. 37, no. 3, pp. 627–638, 2023, doi: 10.18280/ria.370312.

Y. Xie, Z. Hu, and J. Yu, “Face Mask Wearing Detection Based on YOLOv5,” *Int. J. Adv. Netw. Monit. Controls*, vol. 7, no. 2, pp. 67–75, Jan. 2022, doi: 10.2478/ijanmc-2022-0017.

S. Xu, Z. Guo, Y. Liu, J. Fan, and X. Liu, “An Improved Lightweight YOLOv5 Model Based on Attention Mechanism for Face Mask Detection,” Sep. 11, 2022, *arXiv*: arXiv:2203.16506. doi: 10.48550/arXiv.2203.16506.

C.-W. Yang, T.-H. Phung, H.-H. Shuai, and W.-H. Cheng, “Mask or Non-Mask? Robust Face Mask Detector via Triplet-Consistency Representation Learning,” Oct. 01, 2021, *arXiv*: arXiv:2110.00523. doi: 10.48550/arXiv.2110.00523.

M. AlTamimi and F. Mohammed Ali, “Face mask detection based on algorithm YOLOv5s,” *Int. J. Nonlinear Anal. Appl.*, no. Online First, Aug. 2022, doi: 10.22075/ijnaa.2022.28178.3824.

N. Kowalczyk and J. Rumiński, “Mask Detection and Classification in Thermal Face Images,” *IEEE Access*, vol. 11, pp. 43349–43359, 2023, doi: 10.1109/ACCESS.2023.3272214.

X. Fan and M. Jiang, “RetinaFaceMask: A Single Stage Face Mask Detector for Assisting Control of the COVID-19 Pandemic,” Dec. 15, 2021, *arXiv*: arXiv:2005.03950. doi: 10.48550/arXiv.2005.03950.

C. Dewi and H. J. Christanto, “Automatic Medical Face Mask Recognition for COVID-19 Mitigation: Utilizing YOLO V5 Object Detection,” *Rev. Intell. Artif.*, vol. 37, no. 3, pp. 627–638, Jun. 2023, doi: 10.18280/ria.370312.

I. S. Walia, D. Kumar, K. Sharma, J. D. Hemanth, and D. E. Popescu, “An Integrated Approach for Monitoring Social Distancing and Face Mask Detection Using Stacked ResNet-50 and YOLOv5,” *Electronics*, vol. 10, no. 23, p. 2996, Dec. 2021, doi: 10.3390/electronics10232996.

J. Atrey, R. Regunathan, and R. Rajasekaran, "Real-world application of face mask detection system using YOLOv6," *Int. J. Crit. Infrastruct.*, vol. 20, no. 3, pp. 216–240, Jan. 2024, doi: 10.1504/IJCIS.2024.138785.

C. Dewi and H. J. Christanto, "Automatic Medical Face Mask Recognition for COVID-19 Mitigation: Utilizing YOLO V5 Object Detection," *Rev. Intell. Artif.*, vol. 37, no. 3, pp. 627–638, Jun. 2023, doi: 10.18280/ria.370312.

G. Oreski, "YOLO*C — Adding context improves YOLO performance," *Neurocomputing*, vol. 555, p. 126655, Oct. 2023, doi: 10.1016/j.neucom.2023.126655.

Y. Cui, D. Guo, H. Yuan, H. Gu, and H. Tang, "Enhanced YOLO Network for Improving the Efficiency of Traffic Sign Detection," *Appl. Sci.*, vol. 14, no. 2, p. 555, Jan. 2024, doi: 10.3390/app14020555.

J. Redmon and A. Farhadi, "Advancements in YOLO Object Detection for Real-World Applications," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 45, no. 6, pp. 1254–1265, 2023.

D. Chicco and G. Jurman, "The Advantages of the Matthews Correlation Coefficient (MCC) over F1 Score and Accuracy in Binary Classification Evaluation," *BMC Genomics*, vol. 21, no. 1, p. 6, 2020.