

Special Issue: Measurement, control, and analysis of motion using ICT and AI

Article

Study on Automatic Detection of Dust Mask Wearing Status in Factories

Wenyuan Jiang ^{1*}, Yuhei Yamamoto ², Hajime Tachibana ³, Keisuke Nakamoto ³, Kunihiro Katai ³, Naoya Nakagishi ⁴ and Hikaru Muranaka ⁴

¹Faculty of Architectural and Environmental Design, Osaka Sangyo University, 3-1-1 Nakagaito, Daito-shi, Osaka 574-8530, Japan

²Faculty of Environmental and Urban Engineering, Kansai University, 3-3-35 Yamate-cho, Suita-shi, Osaka 564-8680, Japan

³Komaihaltec Inc., 1-19-10 Ueno, Taito-ku, Tokyo 110-8547, Japan

⁴Formerly with Faculty of Engineering, Osaka Sangyo University, 3-1-1 Nakagaito, Daito-shi, Osaka 574-8530, Japan

Abstract

In construction and industrial work environments, workers are mandated to wear dust masks to ensure their health and safety. However, in actual field conditions, many workers neglect this requirement due to breathing discomfort and the heat and humidity within the factory. To address this issue, it is necessary to detect workers who are not wearing masks in real time and prompt them to put them on. Therefore, this study proposes a method to determine the wearing status of dust masks using deep learning based on video footage captured within the factory.

Keywords: Automatic safety management; Dust mask; Deep learning; Wearing status classification.

1. Introduction

In construction and industrial work environments, the Industrial Safety and Health Act was enacted in 1972 (Ministry of Health, Labour and Welfare, 2024) to prevent health hazards among workers. Under this law, the use of protective equipment is mandatory to prevent exposure to dust and hazardous substances. In particular, tasks such as concrete drilling and demolition generate large amounts of dust, and it has been pointed out that prolonged exposure can lead to serious occupational diseases such as pneumoconiosis and lung cancer. As a countermeasure, wearing dust masks is considered one of the most fundamental and important safety practices. However, in actual worksites, some workers neglect to wear masks due to discomfort when breathing and the intense heat and humidity of factory interiors and outdoor environments during the summer season. It is also common for workers to return to work after a break without putting their masks back on. These situations accumulate over time, resulting in workers being continuously exposed to dust and eventually developing respiratory illnesses.

In recent years, such health issues have increasingly led to official recognition as work-related injuries or even legal disputes. In fact, there have been court cases (Supreme Court of Japan, 2020; Takamatsu District Court, c. 2024) in which the development of illness due to the lack of dust mask usage became the central issue, with the employer's failure to ensure safety being called into question. Despite the critical importance of dust masks in preventing health hazards, managing their proper usage in the field remains a difficult challenge. Against this background, increasing attention has been given in recent years to the enhancement of safety management at worksites through the application of ICT (Information and Communication Technology) (JNIOH, 2021). Studies involving the use of cameras and sensors to track worker positions, recognize actions, and monitor the use of protective equipment have become more active. With the help of AI-based automatic detection systems, it is becoming possible to monitor and record the conditions of workers in real time.

Received: date, 3 July 2025, Revised: 28 July 2025, Accepted: 21 August 2025, Published: 17 October 2025

* Correspondence: kyo@aed.osaka-sandai.ac.jp

Publisher's Note: JOURNAL OF DIGITAL LIFE. stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © SANKEI DIGITAL INC. Submitted for possible open access publication under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

Previous studies have proposed mask detection methods based on deep learning models such as MobileNet (Elnady et al., 2021) and YOLO (Loey et al., 2021), which are capable of detecting the presence or absence of mask usage with a certain level of accuracy. However, at construction sites, instances of improper mask wearing, such as the nose or mouth being exposed, have been observed. Although models such as MobileNet and YOLOv3 are relatively lightweight, they have limitations in representational capacity, making it difficult to extract complex features, such as the degree of mask displacement or the boundary between the mask and the face. In particular, their architectures are less effective for recognizing small objects, which is disadvantageous for classifying masks that cover only part of the face. Therefore, it is considered challenging to apply these techniques directly.

Therefore, this study investigates a method for automatically determining the dust mask wearing status of workers using deep learning-based detection techniques.

2. Related Works

In recent years, the application of deep learning technologies for detecting mask wearing conditions has been actively studied in various research fields. In particular, the spread of COVID-19 and the growing emphasis on safety in work environments have accelerated research on mask detection using surveillance footage and images.

Elnady et al. (2021) proposed a lightweight mask detection model based on MobileNetV2 (Sandler et al., 2018), which achieved high classification accuracy on frontal facial images.

Loey et al. (2021) developed a real-time mask detection method combining YOLOv3 (Redmon et al., 2018) and transfer learning, demonstrating its effectiveness in surveillance video analysis.

Ullah et al. (2022) introduced an integrated model using DeepMaskNet, capable of both detecting the presence or absence of masks and performing face recognition even when a mask is worn, thus maintaining high recognition accuracy under masked conditions.

Pan et al. (2023) proposed a deep learning-based binary classification model to determine mask wearing status from facial images. Based on YOLOv5 (Ultralytics, 2020), the model was optimized for lightweight implementation and accuracy improvement, achieving practical applicability in actual worksites settings.

However, these studies are limited to binary classification—detecting whether a mask is worn or not—and do not address the issue of “improper mask wearing” (e.g., nose uncovered or mask worn under the chin), which is frequently observed in actual work environments. Improper wearing is often excluded from training datasets, resulting in a high likelihood of being misclassified as “properly worn.”

To address this issue, Campos et al. (2022) proposed a ternary classification model that includes “improper wearing” as a third category and demonstrated its effectiveness in reducing misclassification. However, their model was designed for images in which the subject’s head appears relatively large (typically 500 to over 1000 pixels), making it unsuitable for wide-angle surveillance footage in which the head size may be as small as 100 pixels. Moreover, there are various patterns of improper mask wearing, and under the complex conditions typical of actual worksites environments, the classification accuracy remains insufficient for practical implementation.

As described above, while existing mask detection studies have shown a certain level of effectiveness under controlled conditions, they fall short in addressing actual challenges such as improper mask usage at construction and industrial sites. Thus, their applicability in practical scenarios remains limited.

Such improper wearing has been observed in approximately 20–30% of cases (Kagawa Occupational Health Promotion Center, 2002; Ministry of Health, Labour and Welfare, 2020). In addition, in locations where there is a risk of dust exposure, mask wearing from the entrance is required, making entrance checks necessary.

Therefore, in this study, we assume the scenario of checking mask wearing status at the entrance as a dust control measure, and aims to develop an AI-based system to automatically detect mask wearing conditions in the context of dust protection in work environments. The primary focus is on detecting difficult cases such as “improper wearing” (e.g., nose exposed, chin wearing), which are commonly observed on-site and are challenging to detect using existing technologies.

3. Methods

3.1. Overview of Proposed Method

In this study, we propose a method for determining the mask wearing status of workers in videos captured at construction sites and similar environments. In particular, the method needs to account for multiple patterns of "improper wearing." Considering the potential future expansion in the number of such patterns, it would be difficult to handle all cases using a single detection model alone.

Therefore, it is considered difficult to realize the judgment of multiple wearing states solely with detection techniques such as YOLO. In this study, we propose a two-stage approach in which head detection and mask-wearing state classification are separated, incorporating a flexible classification method capable of handling complex wearing states (correctly worn, below nose, below mouth and no mask).

In addition to ensuring classification accuracy, this study aims to reduce implementation costs by utilizing surveillance cameras, while achieving quasi-real-time alerts for improper mask wearing. Furthermore, on-site managers have requested processing at least once per second, making it essential to adopt a method that offers a well-balanced trade-off between accuracy and processing speed.

3.2. Method Selection

Object detection methods such as the YOLO (Redmon et al., 2016) series, SSD (Liu et al., 2016) and RetinaNet (Lin et al., 2017) are widely used. In addition, two-stage detection methods such as Faster R-CNN (Ren et al., 2015) are often employed to improve accuracy; however, these approaches have the drawback of high computational cost and limited real-time performance. As near-real-time processing is also required in practical settings, it is necessary to adopt a method that offers a good balance between accuracy and speed. Among these, the YOLOv5–v9 series achieves both high speed and high accuracy, making it advantageous for deployment in surveillance systems and edge devices that require real-time processing. In particular, YOLOv8 (Ultralytics, 2023) supports multi-scale learning, which improves detection performance for small objects such as faces and masks, and has recently seen increased application in personal protective equipment detection.

On the other hand, classification methods mainly utilize deep convolutional neural networks such as MobileNet (Howard et al., 2017), ResNet (He et al., 2016), DenseNet (Huang et al., 2017), VGGNet (Simonyan and Zisserman, 2015), and Vision Transformer (ViT) (Dosovitskiy et al., 2021).

MobileNet is lightweight and easy to implement; however, it tends to have limited representational power for complex features. VGGNet, while more computationally intensive, has a simple architecture and consistently demonstrates high classification performance, making it effective for tasks involving visually distinctive features. In particular, VGG19 (Simonyan and Zisserman, 2015), with its deeper architecture, enables higher-level feature extraction and is considered well-suited for distinguishing subtle differences in improper mask wearing conditions (e.g., nose exposed or chin wearing). Moreover, since actual worksite applications in this study do not require frame-by-frame real-time processing, the high-accuracy VGG19 is considered the most suitable option.

Based on the above considerations, in this study, to achieve quasi-real-time performance with processing every second, we adopt YOLOv8 in the detection stage for its excellent balance between detection accuracy and processing speed for small objects, and VGG19 in the classification stage for its high accuracy in identifying complex mask wearing conditions, including improper wearing.

3.3. Proposed System

3.3.1 Overview of Proposed System

The proposed system is shown in Fig. 1. The proposed system is divided into a Training Component and a Classification Component. The Training Component consists of two functions: Head Training Data Generation Function and Mask Classification Training Data Generation Function. The Classification Component consists of three functions: Head Detection Function, Head Region Extraction Function and Mask wearing Classification Function.

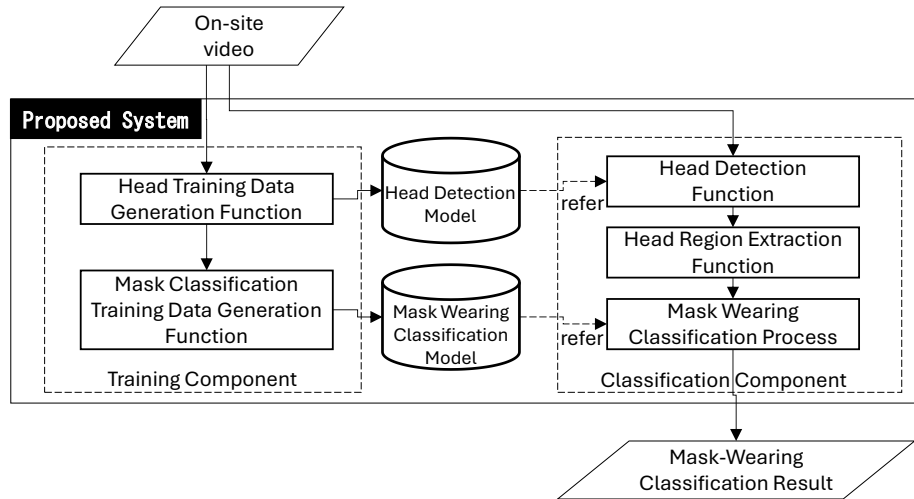


Fig.1.Proposed System

3.3.2 Head Training Data Generation Function

In this function, training data for head detection is generated by manually annotating the heads of workers in video images, as shown in Fig. 2. Then, a Head Detection Model is trained using YOLOv8 based on the annotated data.



Fig.2. Image for Annotation

3.3.3 Mask Classification Training Data Generation Function

In this function, head images are first manually cropped from each frame of the video using the head training data generation function. Next, based on the opinions of on-site factory managers, each head image is labeled with one of four commonly observed categories (Fig.3): Worn Correctly, Worn below Nose, Worn below Mouth, and No Mask. Then, a Mask wearing Classification Model is trained using VGG19 based on these labeled head images.



Fig.3. Four Mask Wearing Status Categories

3.3.4 Head Detection Function

In this function, the heads of workers are detected from site videos in order to determine their mask wearing status, and their position coordinates and sizes are obtained. This is accomplished using a YOLOv8-based Head Detection Model.

3.3.5 Head Region Extraction Function

In this function, head images are extracted to enable the classification of mask wearing status. Specifically, the detected position and size information from the head detection function are used to extract the head regions from the full images.

3.3.6 Wearing Classification Function

In this function, the head images are analyzed using the Mask wearing Classification Model to determine the mask wearing status. Specifically, the model classifies each head image into one of four categories: worn correctly, worn below nose, worn below mouth, or No mask.

4. Results

4.1 Overview of Experiment

To verify the effectiveness of the proposed method, this study conducts three experiments. The first is the Factory Entrance Verification Experiment. The second is the Factory Interior Verification Experiment. The third is the Generalization Verification Experiment.

In this study, when surveillance camera footage was used as training data, many scenes contained only a single person or none at all, making it difficult to collect data covering all patterns. Furthermore, since workers are engaged in their tasks, they face the camera only for short periods, making it challenging to obtain sufficient footage from various angles. Therefore, to ensure an adequate quantity and variety of training data, newly recorded videos were used.

4.2 Factory Entrance Verification Experiment

This experiment simulates a scenario in which mask wearing status is inspected at the entrance of a factory. An example of experimental image is shown in Fig. 4. The video was recorded in Full HD at 30 fps.



Fig.4. Image of Experimental Video at Factory Entrance

The head detection training data consisted of annotated head images extracted from a 58 seconds video in which the five individuals shown in Fig. 5 moved freely within the factory. One frame was sampled every second, resulting in 58 images.

For the mask classification training data, as shown in Fig. 5, each of the five subjects was recorded while rotating 360 degrees under four different mask-wearing conditions. From these videos, 20 head images per class were manually extracted for each individual, yielding a total of 400 images, which were used as the training data.



Fig.5. Sample Image of Training Data Video at the Factory Entrance

The verification video, as shown in Fig. 4, was recorded by capturing five individuals passing through an entrance in various mask wearing conditions. Among them, two workers were wearing their masks below mouth, while the other three represented the remaining categories, one for each. For the verification, 22 images were extracted from the entrance-passing scenes, ensuring representation of all four mask wearing categories. A maximum of three images per individual was used. The purpose was to evaluate whether the system could correctly classify the mask wearing status. The results of the verification are shown in Table 1.

Table 1. Results of Factory Entrance Verification

Prediction \ Ground truth	Worn correctly	Worn below nose	Worn below mouth	No mask
Worn correctly	3	3	0	0
Worn below nose	0	3	0	0
Worn below mouth	0	0	3	0
No mask	0	1	2	7

In the YOLO detection results, the heads of all 22 workers appearing in the 22 images were successfully detected. Therefore, it was confirmed that YOLO can be used to detect workers' heads at the factory entrance.

In the VGG verification results, it showed that the four-class classification achieved an accuracy of 0.73. In addition, since no cases of other patterns were misclassified as "worn correctly," it can be considered that improper wearing and no mask were detected without omission. The primary causes of misclassification were that the mask color was similar to the background, as shown in Fig. 6, and that the color of the worn mask differed from the training data, as shown in Fig. 7. To address these issues, it is possible to improve the system by incorporating temporal data for sequential analysis or by adding training data with masks of various colors.



Fig.6. Examples Where Background and Mask are of Similar Color



Fig.7. Example Where Mask Color Differs from Training Data

4.3 Factory Interior Verification Experiment

From the results of the second experiment, it was found that when the back of the head—where the mask is not visible—is captured, accurate classification is difficult. In addition, when the side of the head is captured, a decrease in classification accuracy may also be observed. Therefore, in this experiment, we verify the range of head orientations (angles) from which the mask-wearing status can be accurately determined, in order to confirm how much the mask's visibility affects classification accuracy.

This experiment simulates a scene in which the mask wearing status of workers is inspected during work inside a factory. An example of the test video is shown in Fig. 8. The video is approximately 20 seconds long and was recorded in Full-HD at 30 fps. For evaluation, a total of 122 images were extracted from the video at intervals of every five frames. The purpose of the verification is to evaluate the detection accuracy of YOLO and the classification accuracy of VGG. In this scene, there are consistently 1–3 workers, with a total of 282 detection instances obtained.

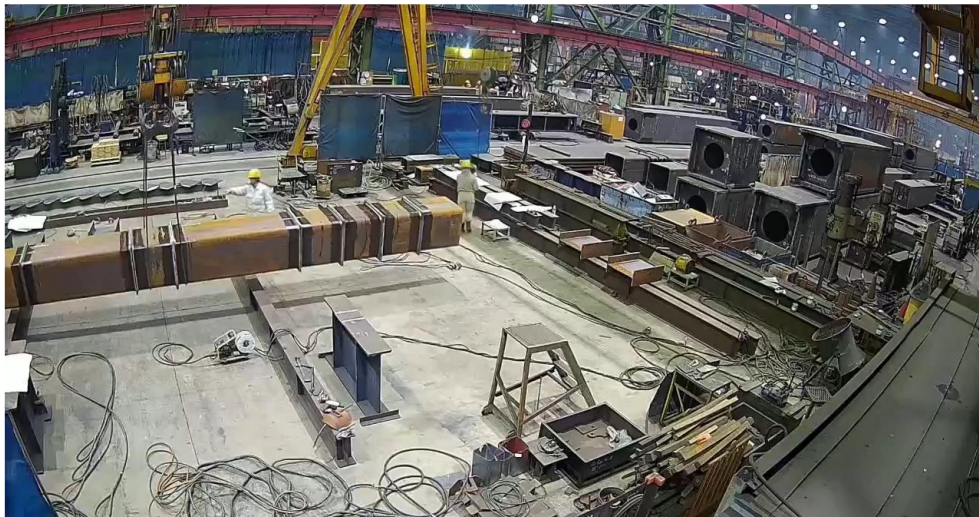


Fig.8. Image of Experimental Video in Factory

The YOLO detection results showed that out of 122 frames, only 49 frames were detected as heads. Accordingly, 31 heads were correctly detected out of a total of 282 individuals. There were also 23 false detections. For the three workers involved, their heads were detected 5, 26, and 0 times, respectively. The detection performance yielded a recall of 0.11, a precision of 0.57, and an F1-score of 0.18. These results indicate that the detection accuracy of YOLO is low in this context. A possible cause is that, in the images captured by the actual surveillance cameras, the workers' heads (average size: approximately 28×28 pixels) are smaller than those in the training data (average size: approximately 42×42 pixels) of the same color, and that the factory background is complex, with the rust color of the pipes (average RGB values: 209, 180, 116) being similar to the color of the helmets (average RGB values: 189, 181, 144).

The classification accuracy of VGG is shown in Table 2.

Table 2. VGG Classification Accuracy in Factory

Prediction Ground truth	Worn correctly	Worn below nose	Worn below mouth	No mask
Worker A (No mask)	0	0	0	5
Worker B (Worn correctly)	6	0	0	20

In terms of classification accuracy using VGG, the correct prediction rate for the 31 detected head images was 0.35. Upon closer examination, these 31 head images corresponded to two workers. For one worker, five frames were detected, all of which were correctly classified. For the other worker, only 6 out of 26 detected frames were classified correctly.

An analysis of the misclassified images revealed that all of them showed the back or side of the workers' heads (Fig. 9.). This suggests that accurate classification is possible when the majority of the mask is visible in the image. On the other hand, considering that misclassification may lead to false alarms, it is deemed necessary to introduce a new category such as "unclassifiable" for cases where the mask cannot be seen clearly, such as when only the back of the head is visible.



Fig.9. Example of a Captured Back of Head

4.4 Generalization Verification Experiment

When the side of the head is captured, the mask may or may not be recognizable. Therefore, in this experiment, we examine the range of head orientations (angles) that allow accurate determination of mask wearing status.

The experiment is conducted in a laboratory setting, as shown in Fig. 10. As illustrated in Fig. 11, the verification is carried out at four distances from the camera: approximately 4 m, 6 m, 8 m, and 10 m. The camera is installed at a height of about 2.4 m. Although the footage is recorded in 4K at 30 fps, the resolution is downscaled to Full HD to match the conditions of actual worksites.



Fig.10. Location of Generalization Verification Experiment

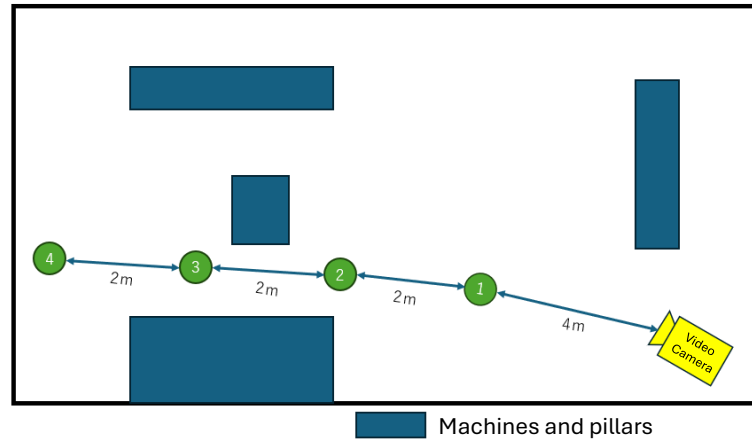


Fig.11. Points of Generalization Validation

At each location, the test subject rotates 360 degrees in 15-degree increments, starting from a position facing the camera (0 degrees). Then, the head region of each subject was visually inspected, manually cropped, and used as the validation data. For the training data, as shown in Fig. 12, images were captured of five participants wearing four different mask wearing conditions while rotating 360 degrees. Additionally, images were extracted at one-second intervals from the recorded videos and used for training. Furthermore, 20 head images per class for each individual were manually extracted from the videos, yielding a total of 400 images, which were employed as the training dataset.

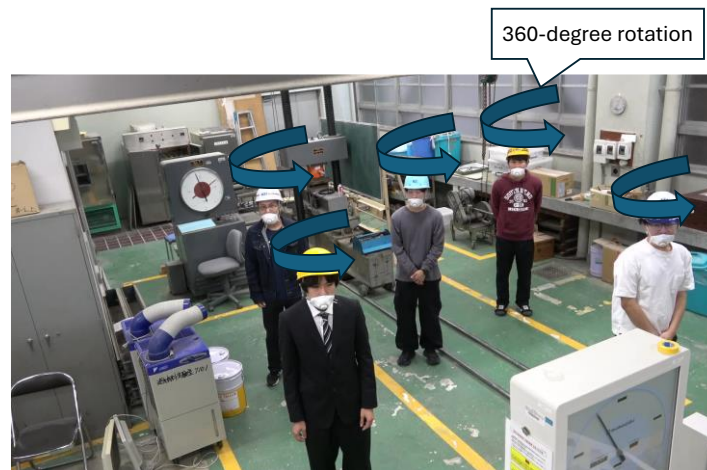


Fig.12. Method of Capturing Training Data

The experimental results are shown in Table 3. From the table, it can be observed that mask classification is difficult within the angular range of approximately 100 to 260 degrees. On the other hand, this range of angles frequently appears in factory environments, which significantly contributes to the decrease in accuracy. Therefore, in future work, a new category labeled "back of the head" will be added to address this unrecognizable range, and further validation will be conducted.

Table 3. Results of the Generalization Verification Experiment

	Accurately classify	Cannot be accurately classified	Accurately classify
1. 4m	0~105 degrees	120~195 degrees	210~345 degrees
2. 6m	0~135 degrees	150~210 degrees	225~345 degrees
3. 8m	0~150 degrees	165~210 degrees	225~345 degrees
4. 10m	0~105 degrees	120~285 degrees	300~345 degrees

5. Conclusions

In this study, we developed a technology to classify the mask wearing status of workers into four categories (properly worn, worn below nose, worn below mouth, and no mask) for automatic safety management at construction sites and factories. The experimental results showed that while the mask wearing status can be accurately classified at the entrance, the classification accuracy within the factory was found to be low. The primary cause is likely the small size of the workers' heads in the footage captured inside the factory. Therefore, we plan to collect real factory data and build a new model. Additionally, to address misclassifications of the back of the head, we plan to add patterns where the mask is not visible, such as the back and side of the head. However, such patterns exhibit substantial differences in visual features depending on the viewing angle, suggesting that consolidating them into a single category may degrade accuracy. Accordingly, future work will explore the optimal number of categories and seek to improve classification performance. Furthermore, we also plan to develop a technology to distinguish between dust masks and general masks. With these improvements, we aim to secure a high-precision mask wearing status classification technology for use in actual worksites environments.

Author Contributions

Conceptualization, W.J., Y.Y., H.T., K.N. and K.K.; methodology, W.J., Y.Y., H.T., K.N. and K.K.; software, W.J., Y.Y., N.N. and H.M.; validation, W.J., N.N. and H.M.; formal analysis, W.J., N.N. and H.M.; investigation, W.J., Y.Y., N.N. and H.M.; data curation, W.J., Y.Y., H.T., K.N., K.K., N.N. and H.M.; writing—original draft preparation, W.J.; writing—review and editing, W.J. and Y.Y.; supervision, W.J. and Y.Y.; project administration, W.J., Y.Y., H.T., K.N. and K.K.; funding acquisition, W.J., Y.Y., H.T., K.N. and K.K.

Funding

This research received no external funding.

Institutional Review Board Statement

Not applicable.

Informed Consent Statement

Not applicable.

Data Availability Statement

Not applicable

Acknowledgments

Not applicable.

Conflicts of Interest

The authors declare no conflict of interest.

Appendix

Not applicable.

References:

- Campos, A., Melin, P. and Sánchez, D. (2023). Multiclass Mask Classification with a New Convolutional Neural Model and Its Real-Time Implementation. *Life*, 13(2), 368. (<https://doi.org/10.3390/life13020368>)
- Dosovitskiy, A., Beyer, L., Kolesnikov, A., Weissenborn, D., Zhai, X., Unterthiner, T., Dehghani, M., Minderer, M., Heigold, G., Gelly, S., Uszkoreit, J. and Houlsby, N. (2021). An image is worth 16x16 words: Transformers for image recognition at scale. *International Conference on Learning Representations*. (<https://arxiv.org/abs/2010.11929>)
- Elnady, A. and Almghraby, M. (2021). Face Mask Detection in Real-Time Using MobileNetv2. *International Journal of Engineering and Advanced Technology*, 10(6), 104–108. (DOI:10.35940/ijeat.F3050.0810621)
- He, K., Zhang, X., Ren, S. and Sun, J. (2016). Deep Residual Learning for Image Recognition. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. (<https://doi.org/10.48550/arXiv.1512.03385>)

- Howard, A. G., Zhu, M., Chen, B., Kalenichenko, D., Wang, W., Weyand, T., Andreetto, M. and Adam, H. (2017). MobileNets: Efficient Convolutional Neural Networks for Mobile Vision Applications, *arXiv*. (DOI:10.48550/arXiv.1704.04861)
- Huang, G., Liu, Z., Van Der Maaten, L. and Weinberger, K. (2017). Densely Connected Convolutional Networks. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2261-2269. (DOI: 10.1109/CVPR.2017.243)
- Kagawa Occupational Health Promotion Center. (2002): Wearing Status of Dust Masks among Welders in 2002, A Survey on Proper Wearing Conditions. Retrieved from <https://www.kagawas.johas.go.jp/researchStudy/entry-54.html> (accessed July 2, 2025).
- Lin, T. Y., Goyal, P., Girshick, R., He, K. and Dollár, P. (2020). Focal Loss for Dense Object Detection. *Transactions on Pattern Analysis and Machine Intelligence*, 44(2), pp. 318-327. (DOI: 10.1109/TPAMI.2018.2858826)
- Liu, W., Anguelov, D., Erhan, D., Szegedy, C., Reed, S., Fu, C. and Berg, A. (2016). SSD: Single Shot Multibox Detector. *Proceedings of the European Conference on Computer Vision*, 21-37. (https://doi.org/10.1007/978-3-319-46448-0_2)
- Ministry of Health, Labour and Welfare. (2020). Study on the Effects of Powered Air-purifying Respirator (PAPR) with an Electric Fan on Respiratory Burden: Report on the Clinical Research Grant for Industrial Injury and Disease Prevention. Retrieved from <https://www.mhlw.go.jp/content/000680034.pdf> (accessed July 2, 2025).
- Pan, X., Liang, X., Ma, Z. and Deng, P. (2023). A Lightweight YOLOv5 Real-time Mask-wearing Detection Algorithm for the Post-pandemic Era, *Future Sensor Engineering*, 3(8), 21-30. (<https://doi.org/10.54691/fse.v3i8.5523>)
- Redmon, J. and Farhadi, A. (2018). YOLOv3: An Incremental Improvement. *arXiv*. (<https://doi.org/10.48550/arXiv.1804.02767>)
- Redmon, J., Divvala, S., Girshick, R. and Farhadi, A. (2016). You Only Look Once: Unified, Real-time Object Detection. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 779-788. (DOI: 10.1109/CVPR.2016.91)
- Ren, S., He, K., Girshick, R. and Sun, J. (2017). Faster R-CNN: Towards Real-time Object Detection with Region Proposal Networks. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 39(6), pp. 1137-1149. (DOI: 10.1109/TPAMI.2016.2577031)
- Sandler, M., Howard, A., Zhu, M., Zhmoginov, A., and Chen, L. C. (2018). MobileNetV2: Inverted Residuals and Linear Bottlenecks. *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 4510-4520. (<https://doi.org/10.48550/arXiv.1801.04381>)
- Simonyan, K. and Zisserman, A. (2015). Very Deep Convolutional Networks for Large-scale Image Recognition. *International Conference on Learning Representations*. (<https://arxiv.org/abs/1409.1556>)
- Supreme Court of Japan (August 31, 2020). Ruling on Asbestos-related Health Damage and State/Manufacturer Duty to Warn and Protect Workers. Retrieved from https://www.rodod.co.jp/precedent/precedent_tag/労働安全衛生法 (accessed July 2, 2025).
- Takamatsu District Court (c. 2024). Decision on Employer Liability for Pneumoconiosis Caused by Asbestos Dust Exposure. Retrieved from https://ascope.net/asbestos_law/damage/石綿粉じんが飛散する工場に勤務していた被災者 (accessed July 2, 2025).
- Ullah, N., Javed, A., Ghazanfar, M., Alsufyani, A. and Bourouis, S. (2022). A Novel DeepMaskNet Model for Face Mask Detection and Masked Facial Recognition, *Journal of King Saud University Computer and Information Sciences*, 34(10), 9905-9914. (<https://doi.org/10.1016/j.jksuci.2021.12.017>)
- Ultralytics. (2020). YOLOv5. Retrieved from GitHub repository. <https://github.com/ultralytics/yolov5> (accessed July 2, 2025).
- Ultralytics. (2023). YOLOv8: Next-generation Object Detection and Segmentation. GitHub repository. Retrieved from <https://github.com/ultralytics/ultralytics> (accessed July 2, 2025).