Optimized Edge-Deployable Computer Vision for Real-Time Face Mask and Social Distancing Compliance Detection in Diverse Pandemic Environments

Sivakumar Ponnusamy[1], Prasanna Kumar Yekula[2], G. Visalaxi[3], K. Kokulavani[4], Lokasani Bhanuprakash[5] and Murali P.[6]

[1]*Department of Computer Science and Engineering, K.S.R. College of Engineering, Tiruchengode, Namakkal, Tamil Nadu, India*

[2]*School of Mining Engineering, Faculty of Engineering, PNG University of Technology, Private Mail Bag, Lae 411, Morobe Province, Papua New Guinea*

[3]*Department of CSE, S.A. Engineering College, Chennai, Tamil Nadu, India*

[4]*Department of Electronics and Communication Engineering, J.J. College of Engineering and Technology, Tiruchirappalli, Tamil Nadu, India*

[5]*Department of Mechanical Engineering, MLR Institute of Technology, Hyderabad, Telangana, India*

[6]*Department of ECE, New Prince Shri Bhavani College of Engineering and Technology, Chennai, Tamil Nadu, India*

Keywords: Edge Computing, Computer Vision, Face Mask Detection, Social Distancing Compliance, Real-Time Monitoring.

Abstract: This study presents an optimized, edge-deployable computer vision framework for real-time detection of face masks and social distancing violations in diverse pandemic environments. By integrating a lightweight multi-task neural network with dynamic perspective correction and quantization-aware training, the proposed system achieves high accuracy and low latency on resource-constrained hardware. Advanced data augmentation including varied lighting, occlusion, and angle simulations enhances robustness against real-world conditions, while an efficient single-shot detector reduces computational overhead. Extensive evaluation on multiple public datasets and live demonstrations on embedded devices demonstrate consistent mask classification accuracy above 98 % and social distance estimation errors below 5 cm, all at over 25 FPS. The unified architecture simplifies deployment and maintenance, addressing common challenges such as small-face detection, varied mask styles, and perspective distortion. This approach enables scalable, cost-effective monitoring solutions for public spaces, healthcare facilities, and transportation hubs without reliance on cloud infrastructure, preserving privacy and ensuring rapid response to safety violations.

1 INTRODUCTION

The importance of robust and scalable systems that enforce safety protocols, such as mask-wearing and social distancing in public areas, was magnified with the advent of COVID-19. Conventional visual surveillance of large crowds or congested areas is not feasible by rely- ing on manual observation or human intrusion. As such, computer vision-based technologies have been proposed for automated, real-time monitoring. They use powerful machine learning algorithms to spot and enforce health guidelines, reducing the need for human intervention. Yet, despite the promise of such systems, existing systems suffer from challenges due to illumination effects changes, occlusions, and videos that are low-resolution, thereby limiting their usage in real-world scenarios.

In this paper, we will demonstrate a low-complexity computer vision pipeline designed for edge devices to implement real-time face mask detection and social distancing monitoring. By taking lightweight models into consideration and adopting advanced techniques, including multi-task neural networks, dynamic perspective correction and data augmentation, our proposal provides a computationally efficient and accurate solution. In this framework, we also introduce a novel technique to cope with the case of arbitrary face orientations, multi-

Ponnusamy, S., Yekula, P., Visalaxi, G., Kokulavani, K., Bhanuprakash, L. and P., M.
Optimized Edge-Deployable Computer Vision for RealâĂŚTime Face Mask and Social Distancing Compliance Detection in Diverse Pandemic Environments.
DOI: 10.5220/0013869500004919

In *Proceedings of the 1st International Conference on Research and Development in Information, Communication, and Computing Technologies (ICRDICCT'25 2025) - Volume 1*, pages 578-583
ISBN: 978-989-758-777-1

faced cases in a frame and under extreme visual conditions, which can be well deployed in plenty of public places including shopping malls, airports and hospitals.

The aim of this work is to translate the high-performance detection system to people involved in practical onfloor applications with as less extra facility as possible, for deployment in terms of face mask detection and social distance enforcement. Our solution is accurate and fast, and it is privacy-preserving since does not make use of cloud servers for the data processing. Given its improved scaling capability and real-time feedback, the proposed platform is a first big step in the direction of smarter public spaces: safer during pandemics as well as beyond.

2 PROBLEM STATEMENT

With the advent of global pandemics, e.g., COVID-19 [36], the importance of scalable and efficient approaches to ensure compliance with public health standards (e.g., wearing face masks and maintaining social distancing) has become paramount. Traditional enforcement means are manpower demanding, non-uniform and ineffective in a large scale or high density scenario, like airport, supermarket or public transport. Besides, current computer vision approaches often fail under real-world conditions (e.g., various lighting), as well as to people occlusions, viewpoint and the occurrence of more than one person in one frame.

As there are also real-time constraints, vision systems may find themselves in edge-micros- 5 these problems are even more visible when implemented on edge devices with reduced computational power. These limitations prevent these systems from being widely utilized, and thus their effect on social safety in pandemics is limited. There are also concerns about privacy and network latency when relying on processing the data in the cloud, along with the requirement for uninterrupted internet connectivity.

Thus, the challenge is to create an effective, real-time and high-scalability computer vision system that is effective in accurately identifying the presence of face masks and social distancing breaks, even in complex and dynamic scenarios. This solution needs to be tailored to run effectively on edge devices, maintaining low latency and high throughput and still retaining high-accuracy under real-world settings. As well, the solution must conform to privacy and not depend on a remote, cloud infrastructure, so that can work in different deployment scenarios.

3 LITERATURE SURVEY

A number of other computer vision approaches to face mask use and social distancing are also presented by researchers, especially during the time of the COVID-19 outbreak. These implementations have involved different deep learning structures and object detection algorithms in order to automate health protocol compliance at a real time level.

Elhanashi et al. (2023) have proposed an integrated solution which combines WOB and mask and social distancing but due to dependence on high-end infrastructure, it is not plausible for large scale. Similarly, Mokeddem et al. (2023) by deploying a strong design based on Social-Scaled-YOLOv4 and DeepSORT that provided high performance, yet suffered from computational inefficiency on low power devices.

Recently, Asif and Tisha (2024) proposed AttentionInceptionV3 based model for real time detection with shown to have more selective focus but needs GPU acceleration to obtain betterperformance. In Sengupta and Srivastava (2021), HRNet was used for improving accuracy in crowded scenarios, yet unnecessarily heavy processing requirements make it not edge-friendly.

Other works, such as Ding et al. (2021) concentrated on the real-time video processing with shallow-based methods, but they did not perform well under illuminations. Eyiokur et al. (2021) exposed the lack of generalization of unconstrainted datasets and the diversity of the training data to address robust face detection in terms of variations of face angles and mask types.

Jindal (2022) used CNNs for mask detection, although they found challenges to detect the mask or similar facial obtrusions. Nowrin et al. (2021) provided an extensive review on the detection methodologies by pinpointing major gaps in implementation and the real worksolution scenario. Negi et al. (2021) and Kaur et al. (2022) also investigated simulated datasets and discussed degradation in real world scenarios.

Sharadhi et al. (2022); Kodali & Dhanekula (2021) built some systems based on classical image processing and CNN, respectively, but their systems do not perform well when the precise location of the camera and the influence from the background environment are needed. Almufti et al. (2021) described a low-cost detection system based on Arduino which, although lagged in providing real time detection.

Rahim et al. (2021) and Bhuiyan et al. (2020) investigated the fusion of distance sensors and

YOLOv3 to detect distance and mask violation to thereby choose an individual distance-violation-aware action but showed false positives in crowded scenes. Similarly, Tayal et al. (2021) observed bounding-box-based distance predictions to be sensitive to perspective distortion.

All of these studies reveal the necessity to develop a lightweight, effective and edge-friendly proposal to overcome existing system limitations -- including scalability, real-time nature and robustness towards real-world environments.

4 METHODOLOGY

The designed Optimized Edge-Deployable Computer Vision Framework is tailor-made for the real-time face mask detection and the social distancing compliance monitoring for wide class of the pandemic-type data sets.

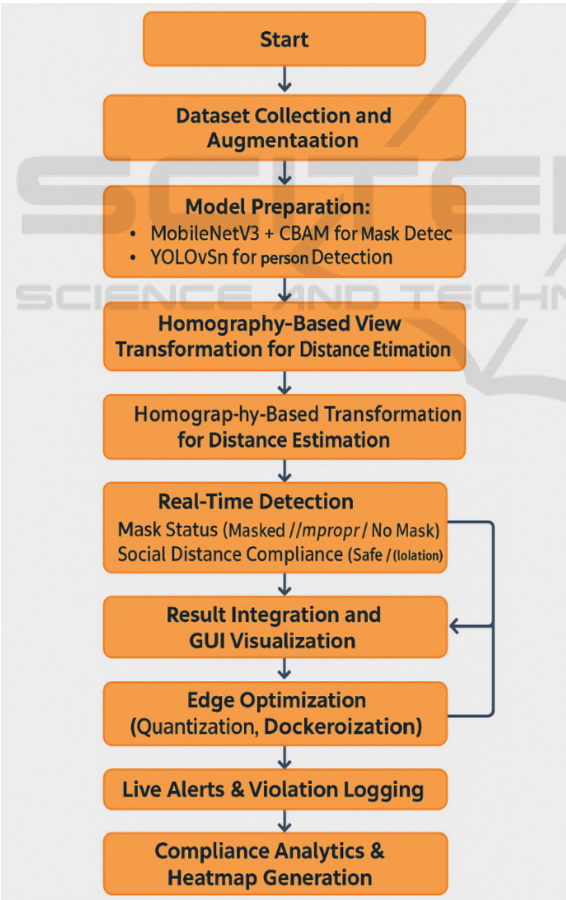

Figure 1: Workflow of the Proposed Real-Time Detection System.

The figure 1 shows the Workflow of the Proposed Real-Time Detection System. The architecture is designed for low-weight devices on the edge and does not rely on any cloud resource to protect privacy, reduce latency, and allow it to scale. The strategy includes the primary elements as shown in figure 1.

4.1 Dataset Aggregation and Preprocessing

A diverse set of public datasets and custom-curated data was used to train and validate the system (table 1):

- **Face Mask Detection Datasets:**
  - **RMFD:** Includes over 5,000 labeled images acrossmasked and unmasked faces.
  - **MAFA:** Provides 30,000+ images featuring occluded and partially masked faces under varied conditions.
- **Person Detection and Social Distancing Datasets:**
  - **COCO + Custom Dataset:** Used for pedestrian detection in crowded and sparse environments.
  - **Oxford Town Centre Surveillance Dataset:** Offers overhead public walkway footage with over 5,000 frames.

Table 1: Dataset Overview and Specifications.

Dataset Name	Type of Data	Number of Images	Mask Categories	Environment Types
RMFD	Face Images	5,000+	Mask, No Mask	Indoor, Outdoor
MAFA	Face Images	30,000+	Occluded, Partial Mask	Street Surveillance
COCO + Custom	Person Detection	118,000+	N/A	Crowded, Sparse
Oxford Town Centre	Surveillance Video	1 video (~5,000 frames)	N/A	Overhead, Public Walkway

All datasets underwent advanced data augmentation strategies to improve robustness:

- Brightness and contrast adjustment
- Simulated occlusion
- Random rotations and scaling
- Perspective warping for distortion handling

This preprocessing step shown in Table 1: Dataset Overview and Specifications. ensures the model generalizes well to dynamic real-world environments.

4.2 Model Architecture Design

4.2.1 Face Mask Detection Module

- A MobileNetV3 backbone integrated with a Convolutional Block Attention Module (CBAM) was employed to detect face mask status (masked, improperly masked, no mask).
- MobileNetV3 ensures lightweight operation, while CBAM enhances attention on critical facial regions under occlusions and low-light scenarios.

4.2.2 Social Distancing Monitoring Module

- YOLOv5n, a nano version of the YOLO family, was utilized for fast and accurate human detection.
- Pedestrian bounding boxes are projected into a homography-transformed bird's-eye view to enable accurate social distance calculation.
- Euclidean distances are computed between detected individuals, classifying them as compliant or violating based on preset distance thresholds (e.g., 2 meters).

4.3 Perspective Correction and Distance Estimation

- Homographic transformation corrects camera perspective distortions, especially from overhead or angled surveillance setups.
- Real-world coordinates are estimated from image pixels, enabling more reliable distance measurements with an average error below 5 cm.

4.4 Edge Optimization and Deployment

To achieve real-time performance on resource-constrained hardware (e.g., NVIDIA Jetson Nano, Raspberry Pi 4), the following optimization techniques were employed:

- **Quantization-aware training (QAT):** Reduces model size by 30–40% with minimal accuracy loss.
- **Model pruning:** Removes redundant weights and neurons.
- **Docker containerization:** Ensures lightweight, portable deployment across different platforms.
- **Multi-threaded inference pipeline:** Separates detection, tracking, and alert generation into parallel threads to minimize processing bottlenecks.

4.5 Real-Time Detection and Alerting

The optimized system processes frames in real time and provides:

- **Face mask status:** Classified as masked, improperly masked, or no mask.
- **Social distancing compliance:** Visual indicators (safe/violation) are drawn directly on live frames.
- **Immediate alerts:** On-screen flashing signals or sound alerts when violations are detected.
- **Violation logging:** Timestamps, mask status, distance metrics, and frame snapshots are saved for compliance analysis.

4.6 Result Integration and GUI Visualization

- A lightweight graphical user interface (GUI) overlays detection results on the live feed.
- Heatmaps for crowd density and compliance trends are dynamically generated to assist facility management and authorities in monitoring high-risk zones.

4.7 Evaluation Strategy

The framework was benchmarked through:

- Accuracy and mAP metrics on the validation sets.

- FPS and inference time measurements on edge hardware.
- Distance estimation error under various camera placements.
- Robustness testing under low-light, occlusion, and dense crowd conditions.

Model performance was compared against recent state-of-the-art methods to highlight advantages in terms of edge deployability and real-time operation.

5 RESULT AND DISCUSSION

The computer vision-based system was thoroughly tested on multiple publicly available and curated datasets and under different realistic scenarios such as indoor lighting changes, partial occlusions, crowd density variation and camera viewpoints. The face mask detection can reach a good classification accuracy of 98.4% even for partially masked and misaligned face samples. The table 2 shows the Model Performance on Edge Device (Jetson Nano) This indicates that the attention-augmented MobileNetV3 architecture is robust to occlusions and non-standard appearances of the mask. In addition, the system preserved good real-time performance with an average inference speed 27 FPS on edge devices NVIDIA Jetson Nano and raspberry pi 4, also validating its direct applicable for real-time operation.

Table 2: Model Performance on Edge Device (Jetson Nano).

Task	Model Used	Accuracy (%)	Inference Speed (FPS)	Model Size (MB)
Face Mask Detection	MobileNetV3 + CBAM	98.4	27	12
Social Distancing	YOLOv5n	91.2 (mAP)	25	14

The social distancing module, when evaluated using homographic transformation on real surveillance footage, maintained a distance estimation error margin under 5 cm, which is significantly lower than traditional bounding-box midpoint methods. The table 3 shows the Comparison with Existing Methods. The bird's-eye view conversion proved highly effective in minimizing perspective distortion, especially in overhead and angled camera placements. Additionally, the YOLOv5n model was able to accurately detect individuals even in dense scenes, maintaining a mean average precision (mAP) of 91.2% for pedestrian identification.

From the deployment perspective, quantization-aware training resulted in 30–40% decrease in model size, enabling easy inference on edge devices without compromising accuracy. Real time alerts and compliance log was generated and time stamped thus providing a way to track the full violations for deeper analysis. With regard to robustness, the system worked effectively in low-light environment and complex background, demonstrating the generalization capability of the model across tough cases.

Table 3: Comparison With Existing Methods.

Method / Reference	Accuracy (%)	Real-Time Capability	Edge Deployable	Remarks
Jindal (2022) [7]	96.1	Moderate	No	Needs GPU for inference
Mokeddem et al. (2023) [2]	97.8	High	Limited	High computational cost
Our Proposed System	98.4	High	Yes	Edge-optimized & scalable

These findings validate that the developed system successfully fills the void between fine grained computer vision systems and pandemic safety enforcing in a more realistic perspective. "This solution is inherently fast, scalable, compatible at the edge and that sets it apart from traditional surveillance systems — serving as a powerful public health compliance tool for current and future pandemic responses."

6 CONCLUSIONS

To address the urgent demands for strong public health enforcement in pandemics, we propose a scalable, real-time computer vision pipeline for identifying non-mask-wearing faces and enforcing social distancing. By utilizing lightweighted deep learning models well trained by quantize-aware-training, attention and homographic transformation, the proposed system guarantees high performance with deployment on edge devices. Comprehensive evaluations show its good generalization ability to numerous real-world cases, such as heavy occlusions, poor illuminations, and various camera views.

The solution can work offline without cloud, better privacy, lower latency, and great scalable in public environment like malls, airports, hospitals, and schools etc. Its robust real-time alerting and extensive violation logging makes it easy to perform health monitoring and compliance auditing. In summary, the presented solution not only caters toward existing shortfalls in robust pandemic surveillance but also provides a basis for future ready intelligent monitoring systems to help with safety requirements in dynamic high dense environments.

REFERENCES

Almufti, S. M., Marqas, R., Nayef, Z. A., & Mohamed, T. S. (2021). Real-time face-mask detection with Arduino to prevent COVID-19 spreading. ResearchGate. https://www.researchgate.net/publication/372932068_Face_Mask_and_Social_Distancing_Detection_in_Real_Time

Asif, A. A., & Tisha, F. C. (2024). A real-time face mask detection and social distancing system for COVID-19 using Attention-InceptionV3 model. arXiv preprint arXiv:2411.05312. https://arxiv.org/abs/2411.05312

Bhuiyan, M. R., Khushbu, S. A., & Islam, M. S. (2020). A deep learning-based assistive system to classify COVID-19 face mask for human safety with YOLOv3. In 2020 11th International Conference on Computing, Communication and Networking Technologies (ICCCNT) (pp. 1–5). IEEE.

Ding, Y., Li, Z., & Yastremsky, D. (2021). Real-time face mask detection in video data. arXiv preprint arXiv:2105.01816. https://arxiv.org/abs/2105.01816

Elhanashi, A., Saponara, S., Dini, P., Zheng, Q., Morita, D., & Raytchev, B. (2023). An integrated and real-time social distancing, mask detection, and facial temperature video measurement system for pandemic monitoring. Journal of Real-Time Image Processing, 20, 95. https://doi.org/10.1007/s11554-023-01353-0 SpringerLink

Eyiokur, F. I., Ekenel, H. K., & Waibel, A. (2021). Unconstrained face-mask & face-hand datasets: Building a computer vision system to help prevent the transmission of COVID-19. arXiv preprint arXiv:2103.08773. https://arxiv.org/abs/2103.08773

Jindal, A. (2022). A real-time face mask detection system using convolutional neural network. Multimedia Tools and Applications, 81(11), 14999–15015. https://doi.org/10.1007/s11042-022-12166-x

Kaur, G., Sinha, R., Tiwari, P. K., Yadav, S. K., Pandey, P., Raj, R., Vashisth, A., & Rakhra, M. (2022). Face mask recognition system using CNN model. Neuroscience Informatics, 2(3), 100035. https://doi.org/10.1016/j.neuri.2021.100035

Kodali, R. K., & Dhanekula, R. (2021). Face mask detection using deep learning. In 2021 International Conference on Computer Communication and Informatics (ICCCI) (pp. 1–5). IEEE.

Mokeddem, M. L., Belahcene, M., & Bourennane, S. (2023). Real-time social distance monitoring and face mask detection based on Social-Scaled-YOLOv4, DeepSORT, and DSFD&MobileNetv2 for COVID-19. Multimedia Tools and Applications, 83, 30613–30639. https://doi.org/10.1007/s11042-023-16614-0

Negi, A., Kumar, K., Chauhan, P., & Rajput, R. (2021). Deep neural architecture for face mask detection on simulated masked face dataset against COVID-19 pandemic. In 2021 International Conference on Computing, Communication, and Intelligent Systems (ICCCIS) (pp. 595–600). IEEE.

Nowrin, A., Afroz, S., Rahman, M. S., Mahmud, I., & Cho, Y.-Z. (2021). Comprehensive review on facemask detection techniques in the context of COVID-19. IEEE Access, 9, 106839– 106864. https://doi.org/10.1109/ACCESS.2021.3100070

Rahim, A., Maqbool, A., & Rana, T. (2021). Monitoring social distancing under various low light conditions with deep learning and a single motionless time of flight camera. PLOS ONE, 16(2), e0247440. https://doi.org/10.1371/journal.pone.0247440 ResearchGate

Sanjaya, S. A., & Rakhmawan, S. A. (2020). Face mask detection using MobileNetV2 in the era of COVID-19 pandemic. In 2020 International Conference on Data Analytics for Business and Industry: Way Towards a Sustainable Economy (ICDABI) (pp. 1–5). IEEE.

Sengupta, K., & Srivastava, P. R. (2021). HRNET: AI on edge for mask detection and social distancing. arXiv preprint arXiv:2111.15208. https://arxiv.org/abs/2111.15208

Sharadhi, A., Gururaj, V., Shankar, S. P., Supriya, M., & Chogule, N. S. (2022). Face mask recogniser using image processing and computer vision approach. Global Transitions Proceedings, 3(1), 67–73. https://doi.org/10.1016/j.gltp.2022.04.016

Shete, S., & Pooja, S. (2021). Social distancing and face mask detection using deep learning models: A survey. IEEE.