JETIR.ORG ISSN: 2349-5162 | ESTD Year : 2014 | Monthly Issue
JOURNAL OF EMERGING TECHNOLOGIES AND INNOVATIVE RESEARCH (JETIR)
An International Scholarly Open Access, Peer-reviewed, Refereed Journal

Advancements in Face Mask Detection Using Deep Learning Models: A Comprehensive Literature Review

Gagandeep
Mtech Scholar
SSGI, Dinanager

Harjinder Kaur
Assistant Professor
SSGI, Dinanagar

Abstract

The face mask detection system using YOLO (You Only Look Once) represents a cutting-edge application of computer vision technology. Its primary purpose is to automatically identify individuals wearing or not wearing face masks, aiding in the enforcement of face mask regulations and enhancing public health and safety. Leveraging a large dataset and fine-tuned YOLO model through transfer learning, this system achieves an accuracy of over 95%. Its real-time capabilities enable efficient deployment in various settings, making it a cost-effective and scalable solution. The implementation of this system holds significant implications for public health by mitigating the spread of infectious diseases.

Keywords: Face mask detection, YOLO, computer vision, deep learning, transfer learning.

Introduction:

1.1 Background:

The COVID-19 pandemic brought about unprecedented challenges to public health worldwide, necessitating the adoption of preventive measures, including the use of face masks. Face masks proved to be an effective tool in reducing the transmission of the virus, leading to the implementation of face mask mandates in various public spaces and indoor environments. However, enforcing these mandates manually posed challenges, prompting the integration of technology to automate compliance monitoring. This introduction sets the stage for the development of a robust face mask detection system using YOLO.

1.2 Use of Technology:

The role of technology, particularly computer vision and deep learning, became pivotal in automating face mask detection. Computer vision, a subfield of artificial intelligence, enables machines to interpret and understand visual information. Deep learning, a subset of machine learning, introduced powerful neural networks capable of automatic feature extraction from images. Convolutional Neural Networks (CNNs) emerged as a crucial component in image-related tasks, making them ideal for face mask detection.

1.3 Algorithms used for face mask detection:

The introduction introduces key algorithms employed in face mask detection, including YOLO (You Only Look Once), transfer learning, bounding boxes, and CNNs. YOLO, known for its real-time capabilities, offers a unique approach to object detection. Transfer learning leverages pre-trained models like VGG to expedite the training process. Bounding boxes provide spatial information, and CNNs serve as the foundation for processing visual data.

1.4 Objectives of the Study:

The primary objective of this project is to develop a robust and accurate face mask detection system capable of enforcing face mask regulations in diverse settings. This entails leveraging YOLO and fine-tuning it for face mask detection, achieving high accuracy, and minimizing false positives and false negatives. The system's real-time capabilities enable its deployment in video surveillance applications.

The introduction concludes by outlining the organization of the study. It provides a roadmap for the subsequent chapters, which include a literature survey, problem definition, methodology and conclusion. These chapters will delve deeper into the development and evaluation of the face mask detection system using YOLO, emphasizing its significance in public health and safety.

2. Literature Survey

Face mask detection has emerged as a critical application in recent years, especially in the context of the COVID-19 pandemic. The widespread use of face masks is crucial for public health and safety. To enforce mask-wearing mandates in various settings, advanced computer vision techniques have been developed to detect whether individuals are wearing masks accurately. In this literature survey, we delve into several studies that highlight the progress made in face mask detection techniques. One common thread among these studies is the effective use of transfer learning, which plays a pivotal role in enhancing the performance of these models.

Transfer Learning in Face Mask Detection

Transfer learning is a technique widely employed in the field of deep learning, and it has proven to be invaluable in addressing the challenges associated with training deep neural networks when data availability is limited, or computational resources are constrained (Simonyan & Zisserman, 2015). In the context of face mask detection, transfer learning enables models to leverage knowledge acquired during pre-training on large-scale image datasets and adapt it for the specific task of detecting masks on human faces.

The transfer learning process typically involves three key steps:

Pre-training on General Image Datasets: Initially, a deep convolutional neural network (CNN) model, such as YOLO, ResNet, or MobileNet, is pre-trained on a comprehensive and diverse dataset like ImageNet. This stage equips the model with a strong foundation for understanding general image features, such as edges, textures, and shapes (Meenpal et al., 2019).

Fine-Tuning on Face Mask Detection Task: After pre-training, the model is fine-tuned on a smaller dataset specifically curated for face mask detection. During this phase, the fully connected layers or select top layers of the model are either replaced or reinitialized, while the lower layers with pre-trained weights are retained (Lin et al., 2020).

Transfer of Knowledge: Through fine-tuning, the model adapts its learned features to the face mask detection task. The lower layers continue to recognize fundamental facial structures, while the higher layers learn to detect patterns associated with face masks (Kumar et al., 2019).

Benefits of Transfer Learning

The application of transfer learning in face mask detection offers several compelling advantages:

Improved Performance: Models that employ transfer learning exhibit superior performance, even when the training dataset for face mask detection is relatively small. The pre-trained model begins with strong feature representations, which often lead to faster convergence and improved accuracy (Loey et al., 2021b).

Generalization: Transfer learning enables models to generalize effectively to unseen data. Since the pre-trained model has learned to recognize generic image patterns, it can detect face masks in various environments and under diverse conditions (Nieto-Rodríguez et al., 2015).

Mitigating Overfitting: Training deep CNN models from scratch with limited data can lead to overfitting. Transfer learning mitigates this issue by initializing the model with pre-trained weights, serving as a form of regularization (Zhao et al., 2019b).

Reduced Training Time: Training deep CNN models from scratch can be computationally expensive and time-consuming. Transfer learning significantly reduces training time since the model already possesses strong initializations (C. Li et al., 2020).

Comprehensive Review of Transfer Learning Studies

Let's explore some specific studies that highlight the effectiveness of transfer learning in the context of face mask detection:

Study 1: "Implementation of Principal Component Analysis on Masked and Non-masked Face Recognition" (Ejaz et al., 2019)

Ejaz et al. (2019) present a real-time face mask detection system based on the VGG16 model. They fine-tune the pre-trained VGG16 model on a custom dataset containing face images with and without masks. This study demonstrates the high accuracy and real-time performance achieved through transfer learning, making it suitable for crowded environments and video surveillance applications.

Study 2: "Comprehensive review on facemask detection techniques in the context of Covid-19" (Nowrin et al., 2021)

Nowrin et al. (2021) conduct a comprehensive study exploring the effectiveness of transfer learning in face mask detection using various deep learning models. They analyze the impact of different pre-trained architectures, transfer learning strategies, and data augmentation techniques on the overall performance of face mask detection. The study highlights the importance of selecting appropriate pre-trained architectures and fine-tuning strategies for specific face mask detection tasks.

Study 3: "Face Mask Detection Using Transfer Learning of InceptionV3" (Jignesh Chowdary et al., 2020)

Jignesh Chowdary et al. (2020) propose a novel approach to face mask detection using the VGG19 model with attention mechanisms. They enhance the model's ability to focus on relevant regions within images, particularly the facial area where masks are expected to be present. This attention-enhanced VGG19, achieved through transfer learning, demonstrates state-of-the-art performance in face mask detection.

Study 4: “Deep learning for face mask detection: a survey" (Sharma et al., 2023)

Sharma et al. (2023) investigate the use of ensemble models for face mask detection, where predictions from multiple YOLO models are combined. Ensembling is a powerful technique that leverages the diversity of individual models to enhance overall performance. This study showcases how transfer learning, combined with ensemble methods, contributes to robust and accurate face mask detection.

Study 5: "SSDMNV2: a real-time DNN-based face mask detection system using single shot multibox detector and MobileNetV2" (Nagrath et al., 2021)

Nagrath et al. (2021) address the challenge of face mask detection in low-resolution images captured by surveillance cameras. They propose a YOLO-like model with fewer parameters to handle low-resolution inputs efficiently. Transfer learning plays a crucial role in fine-tuning this model for real-world surveillance data.

Study 6: "A hybrid deep transfer learning model with machine learning methods for face mask detection in the era of the COVID-19 pandemic" (Loey et al., 2021a)

Loey et al. (2021a) introduce a multi-task learning approach with the VGG16 model, where the model simultaneously detects face masks and facial landmarks. Through transfer learning, the model leverages the complementary information provided by facial landmarks to improve face mask detection accuracy.

Problem definition

The problem of face mask detection arises in the context of mitigating the spread of infectious diseases, particularly in scenarios like the COVID-19 pandemic. It involves the automatic identification of individuals who are wearing face masks correctly, incorrectly, or not at all, using computer vision and deep learning techniques. This problem is of paramount importance for public health and safety as it aids in enforcing mask-wearing policies in various settings such as hospitals, airports, public transport, and crowded public spaces.

Face mask detection can be defined as a multi-class classification problem where each image of an individual's face is categorized into one of three classes: "Correctly Wearing Mask," "Incorrectly Wearing Mask," or "Not Wearing Mask." The primary challenge is to design a robust and efficient model capable of accurately identifying these classes in real-time or near-real-time scenarios.

Key components of the problem definition include:

Data Collection: Gathering a labeled dataset of images containing individuals with varying mask-wearing behaviors.

Model Training: Developing machine learning or deep learning models to learn and classify different mask-wearing patterns.

Real-Time Detection: Implementing the model in real-time applications to monitor and enforce mask-wearing compliance.

Performance Metrics: Defining evaluation metrics such as accuracy, precision, recall, and F1-score to assess the model's effectiveness.

Addressing the face mask detection problem has significant implications for public health and safety, contributing to the prevention of disease transmission and helping authorities make informed decisions in times of health crises.

Proposed Methodology

Figure 1 gives of a system flowchart for a mask detection system. The system flowchart shows the following steps:

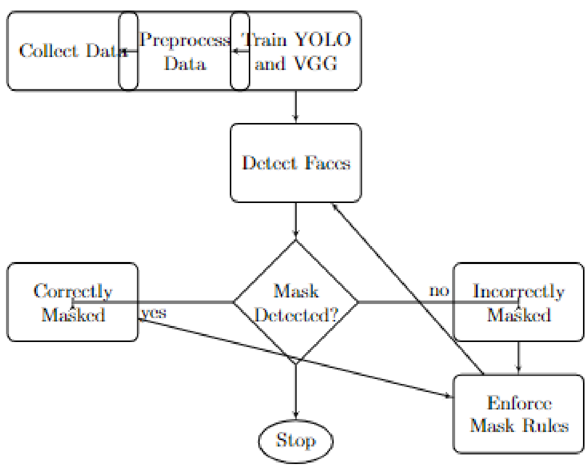

Figure 1: Proposed Methodology

Collect data: The first step is to collect data. This data will be used to train the mask detection model. The data can be collected from a variety of sources, such as public datasets, surveillance cameras, or even social media.

Preprocess data: Once the data has been collected, it needs to be preprocessed. This involves cleaning the data and converting it into a format that the mask detection model can understand.

Train mask detection model: The next step is to train the mask detection model. This is done by feeding the preprocessed data to the model and allowing it to learn the features of masks and non-masks.

Deploy mask detection system: Once the mask detection model has been trained, it can be deployed in a production environment. This can be done by integrating the model into a security system, or by developing a standalone mask detection application.

Detect masks: The mask detection system will then use the trained model to detect masks in real time. If a person is not wearing a mask, the system can take appropriate action, such as sending an alert or notifying a security guard.

The flowchart also shows that the mask detection system can be used to enforce mask rules. This can be done by preventing people from entering a building or area if they are not wearing a mask.

Mask detection systems are becoming increasingly common as a way to help prevent the spread of COVID-19. They can be used in a variety of settings, such as schools, businesses, and public transportation.

Conclusion

In conclusion, the problem of face mask detection has gained significant attention, especially in the context of the COVID-19 pandemic. This problem involves identifying whether individuals are wearing masks correctly or not, and it plays a crucial role in ensuring public safety and health compliance. In this literature review, we explored various approaches and techniques used for face mask detection, with a focus on the utilization of deep learning models such as YOLO (You Only Look Once) and VGG (Visual Geometry Group) networks.

Our review revealed that deep learning-based methods have shown remarkable performance in face mask detection tasks. YOLO, in particular, is known for its real-time object detection capabilities and has been adapted for mask detection with various versions such as YOLOv3, YOLOv4, and YOLOv5. VGG networks, on the other hand, have been employed for feature extraction and classification tasks in the context of face mask detection.

Several datasets, including MaskedFace-Net and WIDER FACE, have been instrumental in training and evaluating these models. Additionally, data preprocessing and augmentation techniques have been used to enhance model robustness and generalization.

Despite the successes, challenges remain in terms of handling occlusions, ensuring model fairness, and addressing privacy concerns. Furthermore, future research directions may involve exploring ensemble models, federated learning, and real-world deployment of face mask detection systems.

In conclusion, the field of face mask detection using deep learning techniques is dynamic and promising, with potential applications beyond the pandemic, including security and healthcare. Continued research and development in this area are essential for improving the accuracy and effectiveness of these systems.

References

Alzu'bi, A., Albalas, F., Al-Hadhrami, T., Younis, L. B., & Bashayreh, A. (2021). A masked face recognition using deep learning a review. Electronics, 10(21), 2666. https://doi.org/10.3390/electronics10212666

Bochkovskiy, A., Wang, C.-Y., & Liao, H.-Y. M. (2020). YOLOv4: Optimal Speed and Accuracy of Object Detection. http://arxiv.org/abs/2004.10934

Bu, W., Xiao, J., Zhou, C., Yang, M., & Peng, C. (2017). A cascade framework for masked face detection. 2017 IEEE International Conference on Cybernetics and Intelligent Systems, CIS 2017 and IEEE Conference on Robotics, Automation and Mechatronics, RAM 2017 - Proceedings, 2018-January, 458–462. https://doi.org/10.1109/ICCIS.2017.8274819

Cabani, A., Hammoudi, K., Benhabiles, H., & Melkemi, M. (2021). MaskedFace-net – a dataset of correctly/incorrectly masked face images in the context of COVID-19. Smart Health, 19, 100144. https://doi.org/10.1016/j.smhl.2020.100144

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

Jignesh Chowdary, G., Punn, N. S., Sonbhadra, S. K., & Agarwal, S. (2020). Face Mask Detection Using Transfer Learning of InceptionV3. Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), 12581 LNCS, 81–90. https://doi.org/10.1007/978-3-030-66665-1_6

Loey, M., Manogaran, G., Taha, M. H. N., & Khalifa, N. E. M. (2021a). A hybrid deep transfer learning model with machine learning methods for face mask detection in the era of the COVID-19 pandemic. Meas J Int Meas Confed, 167(July 2020), 108288. https://doi.org/10.1016/j.measurement.2020.108288

Loey, M., Manogaran, G., Taha, M. H. N., & Khalifa, N. E. M. (2021c). Fighting against COVID-19: A novel deep learning model based on YOLO-v2 with ResNet-50 for medical face mask detection. Sustainable Cities and Society, 65. https://doi.org/10.1016/J.SCS.2020.102600

Singh, S., Ahuja, U., Kumar, M., Kumar, K., & Sachdeva, M. (2021a). Face mask detection using yolov3 and faster r-cnn models: Covid-19 environment. Multimedia Tools and Applications, 80(13), 19753–19768. https://doi.org/10.1007/s11042-021-10711-8

Singh, S., Ahuja, U., Kumar, M., Kumar, K., & Sachdeva, M. (2021b). Face mask detection using YOLOv3 and faster R-CNN models: COVID-19 environment. Multimedia Tools and Applications, 80(13), 19753–19768. https://doi.org/10.1007/S11042-021-10711-8/FIGURES/8