**A SYSTEM AND METHOD FACILITATING DETECTION OF NON-FACE MASK USERS**

**TECHNICAL FIELD**

1. The present invention relates to the field of human surveillance system. In particular, it relates to a system and method facilitating detection of non-face mask users, by live image capturing, processing, and classification.

**BACKGROUND**

1. The guidelines of World Health Organization (WHO) regarding the coronavirus disease 2019 (COVID-19) recommends that wearing face masks and maintaining a social distance of at least 2m can avoid the risk of spread of the disease from person-to-person. However, people are hesitating to wear face mask due to the limitations such as poor breathability, interference while talking, fogging of spectacles, and difficulty during food or water consumption. Thus, face mask detection and safe social distance monitoring has become a crucial responsibility for public service providers for maintaining global security.
2. The physical monitoring of face mask detection is time consuming and a tiresome process to the people assigned for doing this purpose. Thus, people are adopting computer aided real-time monitoring techniques for the face mask detection purposes. Techniques such as CCTV surveillance can only detect whether a person is wearing a face mask or not. Such systems are not capable of alerting the person regarding wearing the facemask. Alert generation associated with real time surveillance became possible with the use of deep learning techniques.
3. Recently, intense research is going on deep learning based face mask detection approaches. One of the existing techniques, authored by Y. Shashi, entitled "Deep learning based safe social distancing and face mask detection in public areas for covid-19 safety guidelines adherence", proposes a computer vision based approach focused on the real-time automated monitoring of people to detect both safe social distancing and face masks in public places by implementing the model on raspberry pi4 to monitor activity and detect violations through camera. Another existing technique authored by V. M. Sammy *et. al.*, entitled "Real-time facemask recognition with alarm system using deep learning", discloses a Raspberry Pi-based real-time facemask recognition that alarms and captures the facial image if the person detected is not wearing a facemask. The problem with the existing techniques discussed here are performing detection on persons at a public place and using alarms or any other manual procedures to alert the crowd in public will be unpleasant and disturbing to the rest of the spectators. Also, public places are well lit and it will be easier to detect people without mask in such conditions.
4. In scenarios like movie theatre where there is darkness or limited availability of light, the real time detection of persons without face mask is difficult. One of the existing techniques, authored by L. Yapeng *et. al.*, entitled "A Seat Position Labeling Algorithm Based on Improved Edge Operator and Gray Projection", discloses a seat position labelling algorithm based on improved edge operator and gray projection with the improved Canny operator to detect the seat edge. The problem with the said technique is that seat position labelling was done on empty seats, with no audience sitting in the seats. Also, edge detection does not work in the said method if spectators are seated and are obstructing the seat’s edges.
5. Most of the existing techniques faces problems such as unable to perform detection in low light conditions, real time alerting of the person without mask and low accuracy of detection. Hence, there is a need for a system and a method which can perform detection of people with and without face mask even in low lit conditions with utmost accuracy.

**OBJECTS OF THE PRESENT DISCLOSURE**

1. Some of the objects of the present disclosure, which at least one embodiment herein satisfies are as listed herein below.
2. It is an object of the present disclosure to provide a system and method facilitating detection of non-face mask users.
3. It is another object of the present invention to provide system and method facilitating detection of non-face mask users, which enables detection of face masks in low light situations like movie theatre scenarios, an opera, a gaming area and the like.
4. It is another object of the present invention to provide system and method facilitating detection of non-face mask users, which manage to detect non-face mask anonymous spectators by finding their seat numbers and alerting them without disturbing the rest of the crowd.
5. It is another object of the present invention to provide system and method facilitating detection of non-face mask users, which detects the seat number of non-face mask users and alerts them using mobile Short Message Service (SMS).
6. It is another object of the present invention to provide system and method facilitating detection of non-face mask users, which equips a simple low-cost surveillance camera that captures facial data irrespective of the surrounding lightning.
7. It is another object of the present invention to provide system and method facilitating detection of non-face mask users, which relies on a robust image enhancement algorithm to generate bright images from poorly lighted environments.
8. It is another object of the present invention to provide system and method facilitating detection of non-face mask users, which is capable of detecting every face from the noise present in the image after image enhancement by drawing artificial bounding boxes around the already detected faces.
9. It is another object of the present invention to provide system and method facilitating detection of non-face mask users, provides a low-cost module does the detection, analysis, and alerting for every non-face mask user which thereby reduces the chances of virus spread.

**SUMMARY**

1. The present invention relates to the field of human surveillance system. In particular, it relates to a system and method facilitating detection of non-face mask users, by live image capturing, processing, and classification.
2. An aspect of the present disclosure pertains to a system for facilitating detection of non-face mask users. The system comprising: one or more pre-processor, a memory coupled to the one or more processor. The memory comprises processor-executable instructions to cause the one or more pre-processors to: receive one or more images from at least one image capturing unit associated with one or more users (106). The at least one image data comprise at least one of a CCTV image, a webcam image, a drone cam image and a digital imaging device image. Further, the system can be configured to perform an enhancement process of the one or more images by an enhancement module to obtain an enhanced version of the one or more images. Further, the system can be configured identify one or more faces by a face detection module based on the enhanced version of the one or more images by assigning at least one face coordinate. Further, the system can be configured label one or more position coordinates associated with a positioning arrangement in a region to extract the one or more position coordinates and a contact number associated with the one or more user by a seat labelling module. Furthermore, the system can be configured to classify the one or more users into one or more classes by a face mask detection module based on the at least one face coordinate and the one or more position coordinates, wherein the one or more classes comprises at least one of a masked user and an unmasked user. Finally, the system can be configured to generate and transmit an alert notification to at least one computing device associated with the unmasked user using the at least one contact number by an alert generation module.
3. In an aspect, the one or more images comprise at least one of a CCTV image, a webcam image, a drone cam image, an image from a live video, and an image from digital imaging device, wherein the one or more image capturing unit comprises at least one of a CCTV imaging device, a digital camera, a mobile camera, a drone camera, a camcorder and a web cam, wherein the one or more image capturing unit is located in a region, wherein the region comprising at least one of a theatre, an auditorium, an opera, a gaming area, a party hall, and a public vehicle.
4. In an aspect, the at least one face coordinate pertains to the one or more faces and comprises coordinates x (Top-Left X Coordinate), y (Top-Left Y Coordinate), xw (Bottom-Right X Coordinate), and yh (Bottom-Right Y Coordinate) which forms a rectangle.
5. In an aspect, the face detection module is integrated with a combination of machine learning techniques comprising at least one of a Multi-task Cascaded Convolutional Networks (MTCNN), a Dlib, a Haar Cascades, and an AddBoxes.
6. In an aspect, the one or more faces identified from on the enhanced version of the one or more images are processed by at least one box algorithm, and the box algorithm is configured to detect one or more faces which are undetected by the face detection module.
7. In an aspect, the one or more position coordinates associated with the non-masked user is stored in a dictionary using a deep learning based face mask detection algorithm by the face mask detection module.
8. In an aspect, the one or more position co-ordinates comprises at least one of a row, and a seat number, wherein the positioning arrangement comprises at least one of a chevron style seating, an auditorium style seating, and a banquet seating.
9. In an aspect, the at least one computing device comprises at least one of a mobile phone, a tab, and an android phone.
10. In an aspect, a method facilitating detection of non-face mask users. The method includes steps of receiving, by the system (102), one or more images from at least one image capturing unit associated with one or more users (106). The method includes steps of performing, by the system (102), an enhancement process of the one or more images by an enhancement module to obtain an enhanced version of the one or more images. Further, the method comprises the step of identifying, by the system (102), one or more faces of the one or more users present in the one or more images based on the enhanced version of the one or more images by assigning at least one face coordinate by a face detection module. Further, the method comprises the step of labelling, by the system (102), one or more position coordinates associated with a positioning arrangement in a region to extract the one or more position coordinates and a contact number associated with the one or more user by a seat labelling module. Furthermore, the method comprises the step of classifying, by the system (102), the one or more users into one or more classes by a face mask detection module based on the at least one face coordinate and the one or more position coordinates, wherein the one or more classes comprises at least one of a masked user and an unmasked user. Finally, generating and transmitting, by the system (102), an alert notification to at least one computing device associated with the non-masked user using the at least one contact number by an alert generation module.
11. Various objects, features, aspects, and advantages of the inventive subject matter will become more apparent from the following detailed description of preferred embodiments, along with the accompanying drawing figures in which like numerals represent like components.

**BRIEF DESCRIPTION OF DRAWINGS**

1. The accompanying drawings are included to provide a further understanding of the present disclosure, and are incorporated in, and constitute a part of this specification. The drawings illustrate exemplary embodiments of the present disclosure, and together with the description, serve to explain the principles of the present disclosure.
2. In the figures, similar components, and/or features may have the same reference label. Further, various components of the same type may be distinguished by following the reference label with a second label that distinguishes among the similar components. If only the first reference label is used in the specification, the description is applicable to any one of the similar components having the same first reference label irrespective of the second reference label.
3. FIG. 1 illustrates exemplary network architecture of the proposed system facilitating detection of non-face mask users, in accordance with an embodiment of the present disclosure.
4. FIG. 2 illustrates architecture of the proposed system facilitating detection of non-face mask users, in accordance with an embodiment of the present disclosure.
5. FIG. 3 illustrates an exemplary view of a flow diagram of the proposed method facilitating detection of non-face mask users, in accordance with an embodiment of the present disclosure.
6. FIG.4A illustrates an exemplary view of the working of enhancer algorithm under irregular seating positions of the users, in accordance with an embodiment of the present disclosure.
7. FIG.4B illustrates an exemplary view of the working of enhancer algorithm under sequential positioning of the users, in accordance with an embodiment of the present disclosure.
8. FIG.5A illustrates the exemplary view of assigning seat numbers after face detection and row allotment under irregular seating positions of the users, in accordance with an embodiment of the present disclosure.
9. FIG.5B illustrates the exemplary view of assigning seat numbers after face detection and row allotment under sequential positioning of the users, in accordance with an embodiment of the present disclosure.
10. FIG. 6 illustrates the exemplary view of detection of undetected faces using artificial bounding boxes, in accordance with an embodiment of the present disclosure.
11. FIG.7 illustrates an exemplary computer system in which or with which embodiments of the present invention can be utilized in accordance with embodiments of the present disclosure.

**DETAILED DESCRIPTION**

1. The following is a detailed description of embodiments of the disclosure depicted in the accompanying drawings. The embodiments are in such detail as to clearly communicate the disclosure. However, the amount of detail offered is not intended to limit the anticipated variations of embodiments; on the contrary, the intention is to cover all modifications, equivalents, and alternatives falling within the spirit and scope of the present disclosure as defined by the appended claims.
2. Various aspects of the present disclosure are described with respect to FIG 1-7.
3. FIG. 1 illustrates exemplary network architecture of the proposed system facilitating detection of non-face mask users, in accordance with an embodiment of the present disclosure.
4. In an embodiment, referring to FIG. 1, the system 102 will be connected to a network 104, which is further connected to at least one computing devices 108-1, 108-2, … 108-N (collectively referred as computing device 108, herein) associated with one or more users devices 106-1, 106-2, … 106-N (collectively referred as user 106, herein). The computing device 108 may be personal computers, laptops, tablets, wristwatch or any custom-built computing device integrated within a modern diagnostic machine that can connect to a network as an IoT (Internet of Things) device. Furthermore, the network 104 can be configured with a centralized server 110 that stores compiled data from all the secure transactions.
5. In an embodiment, the system 102 may receive at least one input data from the at least one computing devices 108. A person of ordinary skill in the art will understand that the at least one computing devices 108 may be individually referred to as computing device 108 and collectively referred to as computing devices 108. In an embodiment, the computing device 110 may also be referred to as User Equipment (UE). Accordingly, the terms “computing device” and “User Equipment” may be used interchangeably throughout the disclosure.
6. In an embodiment, the computing device 108 may transmit the at least one received data packet over a point-to-point or point-to-multipoint communication channel or network 104 to the system 102.
7. In an embodiment, the computing device 108 may involve collection, analysis, and sharing of data received from the system 102 via the communication network 104.
8. In an embodiment, the system 102 may execute one or more instruction for facilitating detection of non-face mask users.
9. In an exemplary embodiment, the system 102 may include, but not be limited to, a computer enabled device, a mobile phone, a tablet, a display device, a display projector, a AR/VR/MR, a imaging device, a sensors, a NFC, a network (Wired or Wireless), an apparatus to dispatch gift, prints, ecommerce, instructions, a Remote Detection Service (Detection Device) enabled devices such as iBeacon technologies, NFC, IR/RF services, bluetooth to detect the devices nearby, a connect signs objects, an apparatus, a vending machine, a gift claw machine, a combination of the vending machine, and the gift claw machine, a drone, a robot, an advertisement displays, or some combination thereof.
10. In an exemplary embodiment, the communication network 104 may include, but not be limited to, at least a portion of one or more networks having one or more nodes that transmit, receive, forward, generate, buffer, store, route, switch, process, or a combination thereof, etc. one or more messages, packets, signals, waves, voltage or current levels, some combination thereof, or so forth. In an exemplary embodiment, the communication network 104 may include, but not be limited to, a wireless network, a wired network, an internet, an intranet, a public network, a private network, a packet-switched network, a circuit-switched network, an ad hoc network, an infrastructure network, a Public-Switched Telephone Network (PSTN), a cable network, a cellular network, a satellite network, a fiber optic network, or some combination thereof.
11. In an embodiment, the one or more computing devices 108 may communicate with the system 102 via a set of executable instructions residing on any operating system. In an embodiment, the one or more computing devices 108 may include, but not be limited to, any electrical, electronic, electro-mechanical, or an equipment, or a combination of one or more of the above devices such as mobile phone, smartphone, Virtual Reality (VR) devices, Augmented Reality (AR) devices, laptop, a general-purpose computer, desktop, personal digital assistant, tablet computer, mainframe computer, or any other computing device, wherein the one or more computing devices 108 may include one or more in-built or externally coupled accessories including, but not limited to, a visual aid device such as imaging device, audio aid, a microphone, a keyboard, input devices such as touch pad, touch enabled screen, electronic pen, receiving devices for receiving any audio or visual signal in any range of frequencies, and transmitting devices that can transmit any audio or visual signal in any range of frequencies. It may be appreciated that the one or more computing devices 110 may not be restricted to the mentioned devices and various other devices may be used.
12. In an embodiment, the network 104 is further configured with a centralized server 110 including a database, where the user identity data is used for providing authentication to the users. It can be retrieved based on the requirement.
13. In an embodiment, the system 102 can be configured to receive one or more images from at least one image capturing unit associated with one or more users (106). Thus, a first connection can be established between the system 102 and the user 106 associated with the computing device 108. The one or more images comprise at least one of a CCTV image, a webcam image, a drone cam image, an image from a live video, and an image from digital imaging device. The one or more image capturing unit comprises at least one of a CCTV imaging device, a digital camera, a mobile camera, a drone camera, a camcorder and a web cam, wherein the one or more image capturing unit is located in a region. The region comprising at least one of a theatre, an auditorium, an opera, a gaming area, a party hall, and a public vehicle
14. In an embodiment, the system 102 can be configured to perform an enhancement process of the one or more images by an enhancement module to obtain an enhanced version of the one or more images.
15. In an embodiment, the system 102 can be configured to identify one or more faces by a face detection module based on the enhanced version of the one or more images by assigning at least one face coordinate. The face detection module is integrated with a combination of machine learning techniques comprising at least one of a MTCNN (Multi-task Cascaded Convolutional Networks), a Dlib, a Haar Cascades, and an AddBoxes. The at least one face coordinate pertains to the one or more faces and comprises coordinates x (Top-Left X Coordinate), y (Top-Left Y Coordinate), xw (Bottom-Right X Coordinate), and yh (Bottom-Right Y Coordinate) which forms a rectangle. The one or more position coordinates associated with the non-masked user is stored in a dictionary using a deep learning based face mask detection algorithm by the face mask detection module
16. In an embodiment, the system 102 can be configured to classify the one or more users based on the at least one face coordinate into one or more classes by the face detection module, wherein the one or more classes comprises at least one of a masked user and an unmasked user.
17. In an embodiment, the system 102 can be configured to label one or more position coordinates associated with a positioning arrangement in a region to extract the one or more position coordinates and a contact number associated with the unmasked user by a seat labelling module.
18. In an embodiment, the system 102 can be configured to generate and transmit an alert notification to at least one computing device associated with the unmasked user using the at least one contact number by an alert generation module.
19. Although FIG. 1 shows exemplary components of the network architecture 100, in other embodiments, the network architecture 100 may include fewer components, different components, differently arranged components, or additional functional components than depicted in FIG. 1. Additionally, or alternatively, one or more components of the network architecture 100 may perform functions described as being performed by one or more other components of the network architecture 100.
20. FIG. 2 illustrates architecture of the proposed system facilitating detection of non-face mask users, in accordance with an embodiment of the present disclosure.
21. In an aspect, referring to FIG. 2, the system 102 may comprise one or more processor(s) 202. The one or more processor(s) 202 may be implemented as one or more microprocessors, microcomputers, microcontrollers, edge or fog microcontrollers, digital signal processors, central processing units, logic circuitries, and/or any devices that process data based on operational instructions. Among other capabilities, the one or more processor(s) 202 may be configured to fetch and execute computer-readable instructions stored in a memory 204 of the system 102. The memory 204 may be configured to store one or more computer-readable instructions or routines in a non-transitory computer readable storage medium, which may be fetched and executed to create or share data packets over a network service. The memory 204 may comprise any non-transitory storage device including, for example, volatile memory such as Random Access Memory (RAM), or non-volatile memory such as Erasable Programmable Read-Only Memory (EPROM), flash memory, and the like.
22. Referring to FIG. 2, the system 102 may include an interface(s) 206. The interface(s) 206 may comprise a variety of interfaces, for example, interfaces for data input and output devices, referred to as I/O devices, storage devices, and the like. The interface(s) 206 may facilitate communication to/from the system 102. The interface(s) 206 may also provide a communication pathway for one or more components of the system 102. Examples of such components include, but are not limited to, processing unit/engine(s) 208 and a local database 210.
23. In an embodiment, the processing unit/engine(s) 208 may be implemented as a combination of hardware and programming (for example, programmable instructions) to implement one or more functionalities of the processing engine(s) 208. In examples described herein, such combinations of hardware and programming may be implemented in several different ways. For example, the programming for the processing engine(s) 208 may be processor-executable instructions stored on a non-transitory machine-readable storage medium and the hardware for the processing engine(s) 208 may comprise a processing resource (for example, one or more processors), to execute such instructions. In the present examples, the machine-readable storage medium may store instructions that, when executed by the processing resource, implement the processing engine(s) 208. In such examples, the system 102 may comprise the machine-readable storage medium storing the instructions and the processing resource to execute the instructions, or the machine-readable storage medium may be separate but accessible to the system 102 and the processing resource. In other examples, the processing engine(s) 208 may be implemented by electronic circuitry.
24. In an embodiment, the local database 210 may comprise data that may be either stored or generated as a result of functionalities implemented by any of the components of the processor 202 or the processing engines 208. In an embodiment, the local database 210 may be separate from the system 102.
25. In an exemplary embodiment, the processing engine 208 may include one or more engines selected from any of an image acquisition module 212, image enhancement module 214, face detection module 216, seat labelling module 218, face mask detection module 220, alert generation module 222, and other modules 224 having functions that may include but are not limited to testing, storage, and peripheral functions, such as wireless communication unit for remote operation, audio unit for alerts and the like.
26. In an embodiment, the image acquisition module 212 may include means receiving one or more images from at least one image capturing unit associated with one or more users (106), the one or more images comprise at least one of a CCTV image, a webcam image, a drone cam image, an image from a live video, and an image from digital imaging device, wherein the one or more image capturing unit comprises at least one of a CCTV imaging device, a digital camera, a mobile camera, a drone camera, a camcorder and a web cam.
27. In an embodiment, the image enhancement module 214 may be configured to perform an enhancement process of the one or more images to obtain an enhanced version of the one or more images retrieved from the users 106.
28. In an embodiment, the face detection module 216 may be configured to identify one or more faces based on the enhanced version of the one or more images by assigning at least one face coordinate THe face detection module is integrated with a combination of machine learning techniques comprising at least one of a MTCNN (Multi-task Cascaded Convolutional Networks), a Dlib, a Haar Cascades, and an AddBoxes.
29. In an embodiment, the seat labelling module 218 may be configured to label one or more position coordinates associated with a positioning arrangement in a region to extract the one or more position coordinates and a contact number associated with the unmasked user.
30. In an embodiment, the face mask detection module 220 may be configured to classify the one or more users based on the at least one face coordinate into one or more classes. The one or more classes comprises at least one of a masked user and an unmasked user.
31. In an embodiment, the alert generating module 222 may be configured to generate and transmit an alert notification to at least one computing device associated with the unmasked user using the at least one contact number.
32. FIG. 3 illustrates an exemplary view of a flow diagram of the proposed method facilitating detection of non-face mask users, in accordance with an embodiment of the present disclosure.
33. In an embodiment, referring to FIG. 3, the underlying mechanism of the proposed model is shown in 300. The method utilizes an image acquisition module 212, image enhancement module 214, face detection module 216, seat labelling module 218, face mask detection module 220 and alert generation module 222. The method includes steps of receiving, by the system (102), one or more images from at least one image capturing unit associated with one or more users (106). The method includes steps of performing, by the system (102), an enhancement process of the one or more images by an enhancement module to obtain an enhanced version of the one or more images. Further, the method comprises the step of identifying, by the system (102), one or more faces of the one or more users present in the one or more images based on the enhanced version of the one or more images by assigning at least one face coordinate by a face detection module. Further, the method comprises the step of labelling, by the system (102), one or more position coordinates associated with a positioning arrangement in a region to extract the one or more position coordinates and a contact number associated with the one or more user by a seat labelling module. Furthermore, the method comprises the step of classifying, by the system (102), the one or more users into one or more classes by a face mask detection module based on the at least one face coordinate and the one or more position coordinates, wherein the one or more classes comprises at least one of a masked user and an unmasked user. Finally, generating and transmitting, by the system (102), an alert notification to at least one computing device associated with the non-masked user using the at least one contact number by an alert generation module.
34. FIG.4A illustrates an exemplary view of the working of enhancer algorithm under irregular seating positions of the users, in accordance with an embodiment of the present disclosure.
35. In an embodiment, referring to FIG. 4A, the image enhancement module accepts a snapshot of the one or more users sitting under an irregular seating positioning from the live feed of the movie theatre as input and passes it into our deep learning-based image enhancer algorithm to return an enhanced brighter version of the input as the output. Pperforming enhancement process in the one or more images may results in a slight amount of noise in the output image. Such noise level is significantly high on faces that are far away from the screen; and faces far away from the screen are impossible to detect. Under irregular seating positions of the users detecting faces of all users is not occurring.
36. FIG.4B illustrates an exemplary view of the working of enhancer algorithm under sequential positioning of the users, in accordance with an embodiment of the present disclosure.
37. In an embodiment, referring to FIG. 4B, the image enhancement module accepts a snapshot of the one or more users sitting under a regular seating positioning from the live feed of the movie theatre as input and passes it into our deep learning-based image enhancer algorithm to return an enhanced brighter version of the input as the output. Performing enhancement process in the one or more images may results in a slight amount of noise in the output image. Such noise level is significantly high on faces that are far away from the screen; and faces far away from the screen are impossible to detect. Under regular seating positions of the users detecting faces of all users is easier.
38. FIG.5A illustrates the exemplary view of assigning seat numbers after face detection and row allotment under irregular seating positions of the users, in accordance with an embodiment of the present disclosure.
39. In an embodiment, referring to FIG. 5A, in the case of users sitting in an irregular seat positioning, seat labeling module uses the face coordinate points to achieve seat labelling by assigning row and seat numbers to all the detected faces. The face coordinates x (X Coordinate, Top-Left), y (Y Coordinate, Top-Left), xw (X Coordinate, Bottom-Right), and yh (Y Coordinate, Bottom-Right) forms a rectangle, which holds the face. All faces have different rectangle sizes and positions and for seat assigning the top and bottom boundary coordinates are used. For two users to sit in the same row, one user’s bottom boundary should be greater than the other user’s top boundary, thus seats residing in the same row have *yh(previous seat )*-*y(present seat)>0.* Face detection and row allotment are done concerning the height (Y Coordinate) of each user and for assigning seat numbers under each row the distance (X Coordinate) of user from the row’s starting point is used as a reference. Using this condition seats under each row are sorted, and a sequentially increasing seat number counter is assigned. In case of users sitting under irregular positioning the seats detect are irregular (see FIG. 5A). In the seat labelling step because of irregular seating and camera positions, the Y coordinates of some users got overlapped and the seats assigned to the said users are into rows that are actually on their top.
40. FIG.5B illustrates the exemplary view of assigning seat numbers after face detection and row allotment under sequential positioning of the users, in accordance with an embodiment of the present disclosure.
41. In an embodiment, referring to FIG. 5A, in the case of users sitting in a regular seat positioning, seat labeling module uses the face coordinate points to achieve seat labelling by assigning row and seat numbers to all the detected faces. The face coordinates x (X Coordinate, Top-Left), y (Y Coordinate, Top-Left), xw (X Coordinate, Bottom-Right), and yh (Y Coordinate, Bottom-Right) forms a rectangle, which holds the face. All faces have different rectangle sizes and positions and for seat assigning the top and bottom boundary coordinates are used. For two users to sit in the same row, one user’s bottom boundary should be greater than the other user’s top boundary, thus seats residing in the same row have *yh(previous seat )*-*y(present seat)>0.* Face detection and row allotment are done concerning the height (Y Coordinate) of each user and for assigning seat numbers under each row the distance (X Coordinate) of user from the row’s starting point is used as a reference. Using this condition seats under each row are sorted, and a sequentially increasing seat number counter is assigned.
42. FIG. 6 illustrates the exemplary view of detection of undetected faces using artificial bounding boxes, in accordance with an embodiment of the present disclosure.
43. In an embodiment, referring to FIG. 6, using the detected face data from the face detection module, AddBoxes algorithm generates fake boxes around the already detected faces. Based on the information stored by the system on the number of seats in each row, the system may compute the count of undetected faces by subtracting the faces detected from the total seats. Afeter getting the count of undetected faces in each row, boxes are drawn at at equal distances from the detected faces with the width same as their neighbor’s box. The coordinates of all the fake boxes are appended into a dictionary with row value as the key. The count of undetected faces in each row is stored in the pending seats counter, whenever a new box is drawn the value of the pending seats counter is decremented. When the value of the pending seats counter reaches zero, the sytem may move to the next row, this process gets carried out until all the rows are covered. After iterating over all the rows, the fake boxes dictionary is returned to the main algorithm, where the detected faces and fake boxes under each row are clubbed together into a single dictionary with row value as the key. At the final stage of this process system may have the data of both detected and undetected faces.
44. FIG.7 illustrates an exemplary computer system in which or with which embodiments of the present invention can be utilized in accordance with embodiments of the present disclosure.
45. Referring to FIG. 7, computer system includes an external storage device 710, a bus 720, a main memory 730, a read only memory 740, a mass storage device 750, communication port 760, and a processor 770. A person skilled in the art will appreciate that computer system may include more than one processor and communication ports. Examples of processor 770 include, but are not limited to, an Intel® Itanium® or Itanium 2 processor(s), or AMD® Opteron® or Athlon MP® processor(s), Motorola® lines of processors, FortiSOC™ system on a chip processors or other future processors. Processor 770 may include various modules associated with embodiments of the present invention. Communication port 760 can be any of an RS-232 port for use with a modem based dialup connection, a 10/100 Ethernet port, a Gigabit or 10 Gigabit port using copper or fiber, a serial port, a parallel port, or other existing or future ports. Communication port 760 may be chosen depending on a network, such a Local Area Network (LAN), Wide Area Network (WAN), or any network to which computer system connects.
46. In an embodiment, the memory 730 can be Random Access Memory (RAM), or any other dynamic storage device commonly known in the art. Read only memory 740 can be any static storage device(s) e.g., but not limited to, a Programmable Read Only Memory (PROM) chips for storing static information e.g., start-up or BIOS instructions for processor 770. Mass storage 760 may be any current or future mass storage solution, which can be used to store information and/or instructions. Exemplary mass storage solutions include, but are not limited to, Parallel Advanced Technology Attachment (PATA) or Serial Advanced Technology Attachment (SATA) hard disk drives or solid-state drives (internal or external, e.g., having Universal Serial Bus (USB) and/or Firewire interfaces), e.g. those available from Seagate (e.g., the Seagate Barracuda 7102 family) or Hitachi (e.g., the Hitachi Deskstar 7K1000), one or more optical discs, Redundant Array of Independent Disks (RAID) storage, e.g. an array of disks (e.g., SATA arrays), available from various vendors including Dot Hill Systems Corp., LaCie, Nexsan Technologies, Inc. and Enhance Technology, Inc.
47. In an embodiment, the bus 720 communicatively couples processor(s) 770 with the other memory, storage and communication blocks. Bus 720 can be, e.g. a Peripheral Component Interconnect (PCI) / PCI Extended (PCI-X) bus, Small Computer System Interface (SCSI), USB or the like, for connecting expansion cards, drives and other subsystems as well as other buses, such a front side bus (FSB), which connects processor 770 to software system.
48. In another embodiment, operator and administrative interfaces, e.g. a display, keyboard, and a cursor control device, may also be coupled to bus 720 to support direct operator interaction with computer system. Other operator and administrative interfaces can be provided through network connections connected through communication port 760. External storage device 710 can be any kind of external hard-drives, floppy drives, IOMEGA® Zip Drives, Compact Disc - Read Only Memory (CD-ROM), Compact Disc - Re-Writable (CD-RW), Digital Video Disk - Read Only Memory (DVD-ROM). Components described above are meant only to exemplify various possibilities.
49. In an embodiment, the region wherein the one or more image capturing unit and the one or more user is located may comprise but not limited to at least one of a theatre, an auditorium, an opera, a gaming area, a party hall, and a public vehicle.
50. In an embodiment, the region wherein the one or more image capturing unit and the one or more user is located may comprise but not limited to at least one of a low lit region, a dark region, a dim light region, a night mode region, a night time region, a sufficiently lit region, and a well lit region,
51. In an embodiment, in an exemplary process, a simple surveillance camera is used to capture the images of one or more users from a a region comprising at least one of a dark movie theatre. ^To detect face in a very poor light conditioned environment like movie theatre the live feed Images need brightness enhancement.The image enhancement module configured to accepts a snapshot of the one or more users from the live feed of the movie theatre as input and passes it into one or more deep learning-based image enhancer algorithm which returns an enhanced brighter version of the input as the output. The underlying layers of the enhancer Algorithm use deep learning which is configured to covert the target image into a tensor which is then passed as input to a sequential model. This Sequential model has a combination of six convolution layers, a single Maxpooling layer, and a Relu activation function. A combination of dark images and their corresponding bright images are used to train the Enhancer algorithm. To test the trained model, a collection of Five open-source face mask datasets is used. The faces present in the face mask datasets are captured under ample light conditions, making them futile for a movie theatre purpose. To mimic the movie theatre scenario, a dark dataset by adding a darkness factor of 0.1 into the above combination of five datasets are prepared. On the manually darkened images the enhancer algorithm is tested to produce the corresponding bright images, and the outcome is shown in FIG. 4A and FIG. 4B.
52. In an embodiment, in an exemplary process, The enhanced images are transeffered to the face detection module. This face detection module uses different face detection frameworks like MTCNN(Multi-task Cascaded Convolutional Networks), dlib, Haar Cascades. A combination of all the three frameworks are used to extract the maximum possible number of faces. Face coordinates of the faces detected in the face detection module are stored inside a dictionary data structure for further use. Further, the face coordinates data is passed to the seat labeling module where automatic seat labeling is done.
53. In an embodiment, in an exemplary process, seat labeling module uses the face coordinate points to achieve seat labelling by assigning row and seat numbers to all the detected faces. The face coordinates x (X Coordinate, Top-Left), y (Y Coordinate, Top-Left), xw (X Coordinate, Bottom-Right), and yh (Y Coordinate, Bottom-Right) forms a rectangle, which holds the face. All faces have different rectangle sizes and positions and for seat assigning the top and bottom boundary coordinates are used. For two users to sit in the same row, one user’s bottom boundary should be greater than the other user’s top boundary, thus seats residing in the same row have *yh(previous seat )*-*y(present seat)>0.* If this condition fails, the present seat will be assigned to a new row. Face detection and row allotment are done concerning the height (Y Coordinate) of each user and for assigning seat numbers under each row the distance (X Coordinate) of user from the row’s starting point is used as a reference. Using this condition seats under each row are sorted, and a sequentially increasing seat number counter is assigned. The seats detect are irregular in case of the irregular seating positions of users (see FIG. 5A). In the seat labelling step because of irregular seating and camera positions, the Y coordinates of some users got overlapped and the seats assigned to the said users are into rows that are actually on their top and such users are removed from the seat labeling process.
54. In an embodiment, in an exemplary process, performing enhancement process in the one or more images may results in a slight amount of noise in the output image. Such noise level is significantly high on faces that are far away from the screen; and faces far away from the screen are impossible to detect. Such irregular face detection makes the automatic seat labeling inconsistent and faulty**.**  To remove this problem an AddBoxes function which is configured to detect the undetected faces by adding manual boxes at average seat distances are used.
55. In an embodiment, in an exemplary process, the detected faces data is sent to the AddBoxes algorithm. Using the detected face data, AddBoxes algorithm generates fake boxes around the already detected faces. Theatre management stores the number of seats booked in each row. Based on the information stored by the theatre management on the number of seats booked in each row, the system may compute the count of undetected faces by subtracting the faces detected from the total seats booked. Afeter getting the count of undetected faces in each row, boxes are drawn at at equal distances from the detected faces with the width same as their neighbor’s box. The coordinates of all the fake boxes are appended into a dictionary with row value as the key. The count of undetected faces in each row is stored in the pending seats counter, whenever a new box is drawn the value of the pending seats counter is decremented. When the value of the pending seats counter reaches zero, the sytem may move to the next row, this process gets carried out until all the rows are covered. After iterating over all the rows, the fake boxes dictionary is returned to the main algorithm, where the detected faces and fake boxes under each row are clubbed together into a single dictionary with row value as the key. At the final stage of this process system may have the data of both detected and undetected faces.
56. In an embodiment, in an exemplary process, the faces present in each row will be sent to face mask detection module, which handles the face mask recognition task using the face mask calculator algorithm. In the face mask detection module deep learning-based model is used to search each face for a face mask. The crosschecking is carried out for each and every unit time in the film’s duration. Row and seat numbers of faces with no face mask are appended into a list and this information is stored inside a dictionary *Nomaskfaces* with second as the key. The Dictionary *Nomaskfaces,* which consists of the row and seat numbers as values and second as the key is returned by the main algorithm as the final output. The final outcome is comprised of at least one of a detected face, an undetected faces and the details of face mask violators. Finally, phone numbers corresponding to the stored row and seat numbers of non-face mask users are used to send Alerts by the alert generation module.
57. In an embodiment, in an exemplary process, The face mask detection is performed using seven architectures. Among the seven architectures, four are the proposed Neural Network architectures, two are fine-tuned deep learning architectures (VGG16 and Inception V3), and one is a machine learning algorithm (Support Vector Classifier). The four proposed neural network architectures are sequential models differing in layer count. Each Sequential model consists of a convolution layer followed by normalization, Maxpooling, and a dropout layer. The one-layer neural network architecture has one sequential model, the two-layer neural network has two sequential models, and so on. The sequential models present in each neural network layer differ in filter count. These seven architectures are trained and tested against the enhanced versions of the combination of five face mask images dataset. The combined dataset resulted in a total of 46,644 images, 23,316 Masked and 23,328 unmasked Facial images. 30,464 images were used for training the models and 16,000 images are used for testing them; the results are shown in Table 1. The proposed Two-Layer Neural Network Architecture outperformed the remaining architectures by producing a testing accuracy of 99.4%. The same seven algorithms are trained and tested against the original well-lit face mask images dataset; the results are shown in Table 2. The comparison onTables 1 & 2 gives an insight on how well the enhancer algorithm improved the brightness of darkened images. Once the proposed face mask detection Architecture detects the users with no face mask, the corresponding phone numbers from seat labels are obtaained to to send SMS alerts.

**Table 1: The training and testing result of seven architectures based on the enhanced versions of the combination of five face mask images dataset**

|  |  |
| --- | --- |
| **Architecture** | **Overall Test Accuracy** |
| One Layer Neural Network Architecture | 98.2% |
| Two Layer Neural Network Architecture | 99.4% |
| Three Layer Neural Network Architecture | 99.17% |
| Four Layer Neural Network Architecture | 98.4% |
| Inception V3 Architecture. | 97.7% |
| VGG16 Architecture. | 98.3% |
| Support Vector Classifier. | 92% |

**Table 2: The training and testing result of seven architectures based on the original well-lit face mask images dataset**

|  |  |
| --- | --- |
| **Architecture** | **Overall Test Accuracy** |
| One Layer Neural Network Architecture. | 94.7% |
| Two Layer Neural Network Architecture. | 97.4% |
| Three Layer Neural Network Architecture. | 98.7% |
| Four Layer Neural Network Architecture. | 97.5% |
| Inception V3 Architecture. | 97.68% |
| VGG16 Architecture. | 96.7% |
| Support Vector Classifier. | 91% |

1. If the specification states a component or feature “may”, “can”, “could”, or “might” be included or have a characteristic, that particular component or feature is not required to be included or have the characteristic.
2. As used in the description herein and throughout the claims that follow, the meaning of “a,” “an,” and “the” includes plural reference unless the context clearly dictates otherwise. Also, as used in the description herein, the meaning of “in” includes “in” and “on” unless the context clearly dictates otherwise.
3. It is to be appreciated by a person skilled in the art that while various embodiments of the present disclosure have been elaborated for system and method facilitating detection of non-face mask users. However, the teachings of the present disclosure are also applicable for other types of applications as well, and all such embodiments are well within the scope of the present disclosure. However, the system and method facilitating detection of non-face mask users is also equally implementable in other industries as well, and all such embodiments are well within the scope of the present disclosure without any limitation.
4. Accordingly, the present disclosure provides a system and method facilitating detection of non-face mask users.
5. Moreover, in interpreting the specification, all terms should be interpreted in the broadest possible manner consistent with the context. In particular, the terms “comprises” and “comprising” should be interpreted as referring to elements, components, or steps in a non-exclusive manner, indicating that the referenced elements, components, or steps may be present, or utilized, or combined with other elements, components, or steps that are not expressly referenced. Where the specification claims refer to at least one of something selected from the group consisting of A, B, C….and N, the text should be interpreted as requiring only one element from the group, not A plus N, or B plus N, etc.
6. While the foregoing describes various embodiments of the disclosure, other and further embodiments of the disclosure may be devised without departing from the basic scope thereof. The scope of the disclosure is determined by the claims that follow. The disclosure is not limited to the described embodiments, versions or examples, which are included to enable a person having ordinary skill in the art to make and use the disclosure when combined with information and knowledge available to the person having ordinary skill in the art.

**ADVANTAGES OF THE PRESENT DISCLOSURE**

1. The present disclosure provides a system and method which enables detection of face masks in low light situations like movie theatre scenarios.
2. The present disclosure provides a system and method which manage to detect non-face mask anonymous spectators by finding their seat numbers and alerting them without disturbing the rest of the crowd.
3. The present disclosure provides a system and method which detects the seat number of non-face mask users and alerts them using mobile Short Message Service (SMS).
4. The present disclosure provides a system and method which equips a simple low-cost surveillance camera that captures facial data irrespective of the surrounding lightning.
5. The present disclosure provides a system and method which relies on a robust image enhancement algorithm to generate bright images from poorly lighted environments.
6. The present disclosure provides a system and method which is capable of detecting every face from the noise present in the image after image enhancement by drawing artificial bounding boxes around the already detected faces.
7. The present disclosure provides a system and method provides a low-cost module does the detection, analysis, and alerting for every non-face mask user which thereby reduces the chances of virus spread.

**I/We Claim:**

1. A system (102) facilitating detection of non-face mask users, the system (102) comprising:

one or more processors (202); and

a memory coupled to the one or more processors (202), wherein said memory (204) stores instructions which when executed by the one or more processors (202) cause the system (102) to:

receive one or more images from at least one image capturing unit associated with one or more users (106);

perform an enhancement process of the one or more images by an enhancement module to obtain an enhanced version of the one or more images;

identify one or more faces by a face detection module based on the enhanced version of the one or more images by assigning at least one face coordinate;

label one or more position coordinates associated with a positioning arrangement in a region to extract the one or more position coordinates and a contact number associated with the one or more user by a seat labelling module;

classify the one or more users into one or more classes by a face mask detection module based on the at least one face coordinate and the one or more position coordinates, wherein the one or more classes comprises at least one of a masked user and an unmasked user; and

generate and transmit an alert notification to at least one computing device associated with the unmasked user using the at least one contact number by an alert generation module.

1. The system (102) as claimed in claim 1, wherein the one or more images comprise at least one of a CCTV image, a webcam image, a drone cam image, an image from a live video, and an image from digital imaging device, wherein the one or more image capturing unit comprises at least one of a CCTV imaging device, a digital camera, a mobile camera, a drone camera, a camcorder and a web cam, wherein the one or more image capturing unit is located in a region, wherein the region comprising at least one of a theatre, an auditorium, an opera, a gaming area, a party hall, and a public vehicle.
2. The system (102) as claimed in claim 1, wherein the at least one face coordinate pertains to the one or more faces and comprises coordinates x (Top-Left X Coordinate), y (Top-Left Y Coordinate), xw (Bottom-Right X Coordinate), and yh (Bottom-Right Y Coordinate) which forms a rectangle.
3. The system (102) as claimed in claim 1, wherein the face detection module is integrated with a combination of machine learning techniques comprising at least one of a MTCNN (Multi-task Cascaded Convolutional Networks), a Dlib, a Haar Cascades, and an AddBoxes.
4. The system (102) as claimed in claim 1, wherein the one or more faces identified from on the enhanced version of the one or more images are processed by at least one box algorithm, and the box algorithm is configured to detect one or more faces which are undetected by the face detection module.
5. The system (102) as claimed in claim 1, wherein the one or more position coordinates associated with the non-masked user is stored in a dictionary using a deep learning based face mask detection algorithm by the face mask detection module.
6. The system (102) as claimed in claim 1, wherein the one or more position co-ordinates comprises at least one of a row, and a seat number, wherein the positioning arrangement comprises at least one of a chevron style seating, an auditorium style seating, and a banquet seating.
7. The system (102) as claimed in claim 1, wherein the at least one computing device comprises at least one of a mobile phone, a tab, and an android phone.
8. A method facilitating detection of non-face mask users, the method comprising:

receiving, by the system (102), one or more images from at least one image capturing unit associated with one or more users (106);

performing, by the system (102), an enhancement process of the one or more images by an enhancement module to obtain an enhanced version of the one or more images;

identifying, by the system (102), one or more faces of the one or more users present in the one or more images based on the enhanced version of the one or more images by assigning at least one face coordinate by a face detection module;

labelling, by the system (102), one or more position coordinates associated with a positioning arrangement in a region to extract the one or more position coordinates and a contact number associated with the one or more user by a seat labelling module;

classifying, by the system (102), the one or more users into one or more classes by a face mask detection module based on the at least one face coordinate and the one or more position coordinates, wherein the one or more classes comprises at least one of a masked user and an unmasked user; and

generating and transmitting, by the system (102), an alert notification to at least one computing device associated with the unmasked user using the at least one contact number by an alert generation module.

**ABSTRACT**

**SYSTEM AND METHOD FACILITATING DETECTION OF NON-FACE MASK USERS**

Present invention discloses a system (102) and a method facilitating detection of non-face mask users. System (102) receive one or more images from at least one image capturing unit associated with one or more users (106). System (102) perform an enhancement process of the one or more images by an enhancement module to obtain an enhanced version of the one or more images. System (102) identify one or more faces by a face detection module based on the enhanced version of the one or more images. System (102) classify the one or more users into a masked user and an unmasked user by a face mask detection module. System (102) generate and transmit an alert notification to at least one computing device associated with the unmasked user using the at least one contact number by an alert generation module.

[**FIG. 3 shall be the reference figure**]