New normal: cooperative paradigm for COVID-19 timely detection and containment using Internet of things and deep learning

FAROOQUE HASSAN KUMBHAR, SYED ALİ HASSAN, SOO YOUNG SHIN∗
Department of IT Convergence, Kumoh National Institute of Technology, Korea

Received: 13.05.2021 • Accepted/Published Online: 09.08.2021 • Final Version: 04.10.2021

Abstract: The spread of the novel coronavirus (COVID-19) has caused trillions of dollars of damages to the governments and health authorities by affecting the global economies. It is essential to identify, track and trace COVID-19 spread at its earliest detection. Timely action can not only reduce further spread but also help in providing an efficient medical response. Existing schemes rely on volunteer participation, and/or mobile traceability, which leads to delays in containing the spread. There is a need for an autonomous, connected, and centralized paradigm that can identify, trace and inform connected personals. We propose a novel connected Internet of Things (IoT) based paradigm using convolution neural networks (CNN), smart wearable, and connected E-Health to help governments resume normal life again. Our autonomous scheme provides three-level detection: inter-object distance for social distancing violations using CNN, area-based monitoring of active smartphone users and their current state of illness using connected E-Health, and smart wearable. Our exhaustive performance evaluation identifies that the proposed scheme with CNN YOLOv3 achieves up to 90% mean average precision in detecting social distancing violations, as compared to existing YOLOv2 achieving 70% and faster R-CNN with 76%. Our evaluation also identifies that wearing protective gear reduces spread by 50%, and timely actions in the first hour can help avoid three times COVID-19 exposure.

Key words: Convolution neural network, contagious diseases, internet of things, smart city, tracking

1. Introduction
The coronavirus (COVID-19) is a contagious virus from the severe acute respiratory syndrome (SARS) family, which affects the respiratory system of the host with high fever, cough, and breathing problems. The virus has spread to 213 countries with devastating effects, starting in December 2019 with more than 174 million reported cases and 3.7 million deaths till 9 June 20211 The contagious virus spreads by direct exposure to the infected hosts (humans, animals) through a cough or sneeze droplets which can travel up to 2 meters (6 feet). The risk of getting infected increases with a close encounter with the confirmed patient without any intermediate preventive coverings such as face masks, glass shields, eye protectors, etc [1]. Various regular business and social activities such as having coffee at a coffee shop, getting haircuts from barbers, walking in a group, taking public transports, etc. lead to human-to-human infection spread. Various governments took extreme measures of countrywide lockdown and time-based curfews to reduce the number of new cases. These steps however cost too much in terms of business, government financial support, and personal psychological impacts. Stopping the spread of COVID-19 is essential, but resuming daily life activities is also important. Standard operating

∗Correspondence: wdragon@kumoh.ac.kr


This work is licensed under a Creative Commons Attribution 4.0 International License.
procedures (SOPs) with smart lockdowns to minimize the possible spread of infection can help achieve the goal of restoring economies in a safe environment.

To open businesses and economies, it is essential to conduct mass screening [2], infection tracing, confirmed patient monitoring, and timely detection to reduce further spread by taking appropriate actions. We believe that the Internet of Things (IoT) with smart wearable, fog computing [3], and the smart connected city has a lot to offer [4]. Contact tracing using mobile applications is already being used by various countries such as Australia, Singapore, and South Korea to track people with the confirmed disease and newcomers in self-isolation [5, 6]. Google+Apple and Facebook have stepped up to provide their service for contact tracing as well [7]. These government and social media applications rely on user participation and willingness. A partial solution in [8] proposes a headset-like wearable device that can track COVID-19 symptoms. However, practicality of continuous monitoring using a headset-like device is questionable. Authors in [9] propose a surveillance system to monitor social distancing between a group of people using convolution neural network-based object detection technique, YOLO (you only look once) v3 [10]. An interesting study in [11] proposed to identify regions with high mobility using cellular handovers by mapping the relationship of cellular mobility to disease spread. The authors argue that it is essential to identify the area in danger to better contain the virus spread. The solution combats the spread partially; however, it can be coupled with surveillance or smart wearable to provide better results. Moreover, there are advancements in extracting images from video streams to perform various detection using YOLOv2 [12], which can provide accurate detection. We believe that the nature of the COVID spread and social impact requires multiple aspects of containment, instead of relying on single information [13]. A study in [14] proposes a fuzzy system for mortality prediction of COVID-19 and outlines risk, clinical and miscellaneous factors. The research discusses post-COVID analysis to identify and reduce the health challenges. Authors in [15] present short-term forecasting for electric energy considering COVID-19 and health centers using artificial neural networks. The study focuses on the availability of electricity in hospitals and health centers to ensure pervasive medical care to COVID-19 patients. Using CT scanned images of COVID-19 patients, [16] present a detection tool and experiment on IBM quantum computer for COVID-19 detection. These studies explore the post-spread impact of COVID-19 that can be avoided altogether with better proactive schemes. An intriguing idea in [17] suggests a connected architecture to proactively trace COVID-19 symptoms using wearable and takes immediate actions with connected E-Health. Each device monitors health vitals and feeds to a centralized system that continuously monitors and tracks people with high symptoms and confirmed patients.

A major part of existing research focuses on the post-COVID medical effects and how to tackle them. Moreover, most of the earlier work trace confirmed patients employing either their cellular devices or volunteer updates. On the other hand, it is as much essential to identify indirect and asymptomatic spread, which is usually under the radar for current tracking mechanisms. A huge chunk of literature is dedicated to the identification of mobility patterns and cellular traces. There have been several incidents where governments loosen the restriction resulting in sudden spikes in new COVID-19 cases. We believe that there exists no solution, which considers a multi-level and heterogeneous approach to track, trace, and contain the COVID-19 spread. There is a dire need of a proactive, smart and connected paradigm, which not only restricts further virus spread but also allows people to resume their life. In this study, we propose a connected IoT-based paradigm that targets two major agendas; 1) timely detection and appropriate actions to stop the spread and 2) new normal with connected and informed resumption of daily life activities. First, a connected surveillance camera detects violations of social distances by detecting objects and reporting for the possible spread of viruses [12]. Thereafter,
Algorithm 1 Proposed IoT and YOLOv3 based algorithm.

1: Camera share continuous video stream with the server
2: for all Video frames do
3: Detects objects and assigns bounding boxes using CNN based object detection (YOLO v2 or YOLO v3)
4: for all Pair of detected objects do
5: if Exists \( D_{ij} < D_{th} \) then
6: Server identifies area by communicating with BSs (associated with camera)
7: for all Active users in BS(s) area do
8: Active user health update using health center information
9: if Exists a user with confirmed disease then
10: Immediate health and safety actions
11: else
12: end if
13: end for
14: end if
15: end for
16: end for

where \( D_{ij} \) is the inter-object distance between center of the boundary boxes of detected objects, \( D_{th} \) is the allowed minimum distance threshold.

the fog node-based server [18] traces connected cellular devices for active confirmed patients in the reported area. Each wearable device based on IoT assists in identifying people with high symptoms in the reported area. If a confirmed patient or a person with high symptoms is in the reported area, the system trigger warnings and/or actions to contain the spread of the virus, using the connected health care system [19]. Not only does the proposed IoT and YOLOv3 based paradigm restricts the spread of the virus, but it also allows a more secure and informed environment to restart new normal. The novelty and innovation of the proposed scheme lie in the following facts: 1) the envision of a connected paradigm for the detection of social distance breaches using YOLO v2 and YOLO v3 based on CNN, 2) The IoT-based detection and tracking of active diseased persons and highly symptomatic person also lead to containing the spread sooner, 3) The proposed solution also triggers timely actions and warnings (quarantine or self-isolation) toward persons at risk, 4) We propose a centrally connected three-level-based detection paradigm, social distancing violation, area-based tracking, and symptom-based tracing, 5) As a whole, the proposed paradigm is one of its kind, which proactively restricts the spread of the virus and can help governments combat this deadly pandemic.

The rest of the paper is organized as follows. Section 2 outlines our proposed architecture and algorithm with three levels. Section 3 evaluates the proposed scheme using video stream feed in CNN-based object detection and python-based tracking system. Section 4 concludes the paper.

2. Proposed IoT and YOLOv3 based paradigm

COVID-19 spreads over time through aerosols or droplets containing the virus, which comes into contact with the eyes, mouth, or nose. The chances of transmission increase in close human encounters, contacts, and gatherings. However, preventive measures, protective gear, and swift actions can significantly reduce the spread. It is always advisable to contain the virus before it spreads in the community by self-isolation, quarantine, and
immediate medical care. However, there exists no solution which tracks and traces the spread and helps people make informed decisions. In this study, we innovate by proposing a multi-level IoT and CNN YOLOv3 based paradigm which triggers proximity-based tracking and tracing of people with high symptoms and confirmed patients. By taking autonomous immediate actions, the proposed solution can keep a constant check on daily life and reduce the chances of further virus spread. We propose an efficient three-level tracking and detection scheme, i.e. social distancing violation, area-based tracking, and disease/symptom-based actions. Figure 1 illustrates that a continuous video stream is fed to a CNN-based trained model, which detects social distancing violations. The second level of tracking identifies the crowded area and traces all connected cellular users. The active users’ data is matched with the connected E-Health and smart wearable to identify confirmed patients or anyone with high symptoms related to COVID-19. The proposed multi-level scheme ensures efficient detection and triggers necessary actions leading to an informed and safe community.

![Figure 1](image.png)

**Figure 1.** Proposed IoT and YOLOv3 (CNN) based COVID-19 containment scheme, spanning over three major levels of timely detection.

### 2.1. Social distancing violation detection

A CNN-based multi-object recognition model (YOLOv2 and YOLOv3) can identify multiple trained objects in a frame. We consider two widely used object detection methods, YOLO v2 and YOLO v3 [20]. YOLO models accurately and quickly detect objects classified by multiple labels by applying a single neural network to the image. Each input image is divided into several regions where each detected object is assigned a bounding box with features, such as center coordinates, height and width, confidence, and object class shape accuracy. The YOLO v3 is an improved incremental approach (over YOLO v2) with 53 convolutional layers and a deep network darknet-53 backbone. Given a test image, the CNN-based model produces
a collection of all the detected objects. Each bounding box is represented using corners pixel values, i.e. $(x_{\text{min}}, y_{\text{max}}), (x_{\text{max}}, y_{\text{max}}), (x_{\text{min}}, y_{\text{min}}), (x_{\text{max}}, y_{\text{min}})$. The center $(x_o, y_o)$ of the bounding box (detected object) can be estimated as given below:

$$(x_o, y_o) = \left(\frac{x_{\text{min}} + x_{\text{max}}}{2}, \frac{y_{\text{min}} + y_{\text{max}}}{2}\right)$$  \hspace{1cm} (1)

The inter-distance ($D^{i,j}$) between the centers of two detected objects ($(x^i, y^i)$ and $(x^j, y^j)$) can be calculated as:

$$D^{i,j} = \sqrt{(x_o^i - x_o^j)^2 + (y_o^i - y_o^j)^2}$$  \hspace{1cm} (2)

We design an autonomous social distancing detection tool using a fog node-based object identification method. We deploy CNN based model in the fog node, which is fed by a continuous surveillance video stream. The detection model is trained using 2000 images, marked with person class. For every input frame, the CNN model returns the bounding box values for each detected object. Using Equation 1 and 2, we identify that the inter-distance is not violating any predefined rules e.g., $D^{i,j} > D_{\text{th}}$. Figure 2 illustrates few examples of our training dataset containing person class. The testing example shows that a violation is identified when two detected bounding boxes have an inter-distance smaller than a pre-defined threshold ($D^{i,j} < D_{\text{th}}$). The $D_{\text{th}}$ is a programmable threshold that can be controlled by the operators and implementing authorities.

![Figure 2](image-url)  

**Figure 2.** Training and testing data snippets of CNN based object detection (YOLOv2 and YOLOv3).
2.2. Area at risk tracking and disease/ symptoms based actions

In addition to social distancing violation detection, the proposed scheme includes two modules for active user tracking in the area at risk and for disease/symptom based actions. Each surveillance camera is also mapped to a specific area, which subsequently correlates the actively connected users and operating base stations (BS). In a case of social distancing violation detected by the fog node-based autonomous CNN model, the target area is investigated for active users. Subsequently, identified devices are cross-referenced with the local E-Health system to identify any active confirmed patients in the area. Specifically, the active mobile users are scrutinized on two aspects:

1. Confirmed disease cases: We assume that the health centers have data of confirmed active cases. If a cellular user in the reported area is a diagnosed active case, then the health officials are notified. Notification to the active users in the area informs about a possible exposure that also requires them to self-isolate or to contact a medical center. Moreover, the active confirmed patient is immediately quarantined.

2. People having high symptoms for COVID-19: The people who have been in contact with active confirmed patients or have been traveling might not show symptoms immediately. We consider that each person is either wearing an IoT-based health monitor or getting health vitals checked frequently using handheld equipment. Each disease, specifically COVID-19, has widely known symptoms that can help identify a possible infection. The IoT-based smart wearable can monitor major observed symptoms for COVID-19, such as fever (98.6% cases), fatigue (70% cases), and dry cough (60% cases). A person having symptoms higher than a specified threshold for a minimum specific time is also a major concern. If a person with high symptoms is recorded regardless of the reported area, then immediate notifications are dispatched. The person and the people in close proximity are promptly asked to self-isolate.

It should be noted that the area that has no violations, confirmed patients, or a person with high symptoms, is not at risk and can resume activities in the social life.

2.3. Proposed new normal paradigm

We propose a new normal paradigm that employs three autonomous and connected modules, i.e. social distancing violation detection, area at risk tracking, and disease/symptom-based identification and actions. Algorithm 1 discusses the overall orchestra of the proposed autonomous scheme, which monitors, detects, and informs against contagious diseases such as COVID-19. Algorithm 1 outlines the proposed IoT and YOLOv3 based algorithm where surveillance cameras continuously monitor to detect any social distancing violations. Each surveillance video frame is an input to the pre-trained CNN-based detection model. The CNN-based model returns all the detected objects (person class) in the given picture with boundary boxes. Using Equation 1 and 2, we identify social distancing violations, which subsequently trigger further inquisition. Fig 2 shows the test images, where three objects are detected and assigned boundary boxes. For every pair of detected objects, inter-object distance is compared with a pre-defined and programmable distance threshold ($D_{th}$). In the case of social distancing violation detection, area-based tracking collects information of all active users in the respective area. The active user data are cross-referenced with the connected local E-Health network, which identifies any registers confirmed patient or users with high persistent symptoms. The health authorities take immediate action if a confirmed patient or a suspected patient is identified in the area. All active users at risk are notified.
and informed about the situation and suggested to take actions such as isolation. Moreover, if a person has high symptoms, he/she is asked to take immediate precautionary self-isolation and to inform health officials.

The aim of the proposed autonomous scheme is to take required action at the early stages of a virus spread and keep communities safe and informed. Our proposed scheme allows people to perform day-to-day business and operations using smart wearables to keep track of documented patients and potential infections. We believe that the proposed system creates a new normal environment, which can help reopen economies and reduce the widespread of the disease.

### Table. Simulation parameters.

<table>
<thead>
<tr>
<th>Parameters</th>
<th>Values</th>
<th>Parameters</th>
<th>Values</th>
<th>Parameters</th>
<th>Values</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>CNN based object detection (YOLO v2 and YOLO v3 training parameters)</strong></td>
<td></td>
<td><strong>Python based virus spread simulation parameters</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Optimizer</td>
<td>SGD (Stochastic gradient descent) [20]</td>
<td>Iterations</td>
<td>10,000 [20]</td>
<td>Time</td>
<td>300 min</td>
</tr>
<tr>
<td>Learning rate</td>
<td>0.001 [20]</td>
<td>Channels</td>
<td>3 [20]</td>
<td>Area</td>
<td>1000 × 1000 meters</td>
</tr>
<tr>
<td>Batch size</td>
<td>64 [20]</td>
<td>Decay</td>
<td>0.0005 [20]</td>
<td>Mobility</td>
<td>Random walk</td>
</tr>
<tr>
<td>Subdivisions</td>
<td>4 [20]</td>
<td>Momentum</td>
<td>0.9 [20]</td>
<td>Mobility speed</td>
<td>[1,10] m/min</td>
</tr>
<tr>
<td>Stride</td>
<td>1 [20]</td>
<td>Hue</td>
<td>.1 [20]</td>
<td>Distance threshold</td>
<td>5 m</td>
</tr>
<tr>
<td>Saturation</td>
<td>1.5 [20]</td>
<td></td>
<td></td>
<td>Symptom threshold</td>
<td>0.9</td>
</tr>
</tbody>
</table>

Loss function: The proposed scheme expands on the accurate distancing violation detection using trained CNN models. Efficient social distancing detection using YOLOv3 can be evaluated using the loss function, which is also known as sum-square error. YOLO utilizes a layered approach where a loss function is the sum of all errors such as coordinates error, intersection over union (IOU) errors, and classification errors as given below:

\[
\text{loss} = \sum_{i=0}^{g^2} \text{Err}_{\text{coord}} + \text{Err}_{\text{iou}} + \text{Err}_{\text{cls}},
\]

where \(g\) is the number of grids in the frame. The overall loss function described above in its simplest form considers each loss during the model training. During the training, the model can face unstable reactions and have classification errors constant with the coordinate error. The absolute loss function for YOLO model
training can be estimated as:

\[
\text{loss} = \lambda_{\text{coord}} \sum_{i=0}^{B} \sum_{j=0}^{C} l_{ij}^{\text{obj}} \left[ (a_i - \hat{a}_i)^2 + (b_i - \hat{b}_i)^2 \right] + \lambda_{\text{coord}} \sum_{i=0}^{B} \sum_{j=0}^{C} l_{ij}^{\text{obj}} \sum_{c=0}^{C} \sum_{c \in \text{class}} \left( R_i(c) - \hat{R}_i(c) \right)^2
\]

where \( B \) and \( C \) represent cell numbers and confidence of the prediction boxes, respectively. The object confidence in the class is represented by \( R \). \((a,b)\) defines the center of the coordinates of each cell, and \( h, w \) shows its width and height. The weights \( \lambda_{\text{coord}}, \lambda_{\text{noobj}} \) are defined for loss function position and classification loss function.

Nevertheless, the loss function is an efficient metric to evaluate the classification models.

3. Experiments and performance evaluation

The proposed solution relies on efficient social distancing violation detection and containing the spread. We comparatively evaluate the proposed IoT and YOLOv3 scheme and existing YOLOv2 in two parts. First, we exhaustively train a model using YOLOv2 and YOLOv3 for social distancing violation detection. We collect and test real-time video feed to identify social distancing violations. Subsequently, we simulate scenarios where a number of smart devices are geographically distributed with planted confirmed patient devices. The simulation identifies how the spread occurs and the timeline to act immediately. Moreover, we consider the impact of wearing protective gear such as face masks and shields.

CNN based object detection: We comparatively evaluate faster region-based convolution neural networks (R-CNN) [21], YOLOv2 [12], and the proposed IoT and YOLOv3 based system. The darknet-19 for YOLO v2 and darknet-53 for YOLO v3 are trained on our dataset [20]. The model is trained with 2000 images each having a resolution of 416 × 416, with a batch size of 64 and a subdivision of 8. The learning rate of 0.001 ensures faster model convergence with early training time [20]. We ran 10,000 iterations with stochastic gradient descent (SGD) optimizer. A list of all training parameters is outlined in Table. The mean average precision (mAP) is calculated with the train/test split set of 80/20 ratio. Figure 3a and Figure 3b illustrate the comparative results generated by the trained CNN-based models of YOLO v2 and YOLO v3, respectively. After 3000 iterations, both YOLOv2 and v3 models experience relatively small loss fluctuations and changes. However, the YOLO v3 outperforms YOLO v2 by achieving the mAP value of up to 90 % as opposed to mAP value of 73 %. It can also be observed that the mAP results of YOLOv3 have fewer fluctuations after 1000 iterations, whereas YOLOv2 performs poorly with several inconsistencies and achieves merely 73 % mAP. On the other hand, faster R-CNN [21] have a mAP value of 76 %, which is higher than the YOLOv2 (73 %) and lower than YOLOv3 (90 %). YOLO variations are trained to classify objects and apply bounding boxes at the same time, which makes it distinctly different than faster R-CNN. The faster R-CNN considers a region-based network, which is further utilized in assigning boundary boxes and classify objects. With more layers of the darknet-53 in YOLOv3, the proposed IoT and YOLOv3 based system efficiently and accurately identifies person class, which is subsequently used to detect social distancing violations.
Python simulation for virus spread: Considering that the CNN-based object detection module generates a violation warning in a particular area of $2000 \times 2000$ meters having 500 to 2500 active users. Each user depicts random mobility between $[0,20]$ m/min. Our python-based simulation assigns a wearable health monitoring device to each active user, which continuously reports health vitals for an observation period of 500 min. We plant 50 confirmed patients in random locations in the reported area. A user is exposed to the disease by being in the proximity (distance threshold 3 m) of a confirmed patient or a person with significant symptoms. However, the probability of getting infected when everyone wears a mask or does not wear a mask is set to an arbitrary programmable value of 0.5 and 0.2, respectively. Moreover, the symptoms-based detection identifies a possible infected person if he/she has symptoms more than a predefined threshold (0.9), for an hour duration.

Figure 4a illustrates the potentially infected count of exposed persons to the people having confirmed disease. With the increase in the number of people from 500 to 2500, the risk of getting exposed to virus also increases. However, wearing protective gear like a face mask or shields substantially reduces the probability of getting infected by 50%. On the other hand, Figure 4b shows that the number of people exposed to a person having high symptoms (potential carrier) increases from 20 to 350+ with the increase in the total active users from 500 to 2500, respectively. The protective gears reduce the risk of getting infected after exposure by more than 50%. Nevertheless, the proposed scheme’s timely detection of exposed or infected helps take appropriate actions. Only the detected people are required to be isolated or tested, whereas others carry on with their social lives in the new normal. Moreover, Figure 5a shows that uninterrupted diseased person exponentially spreads the virus and puts more lives in danger over time. In a sparse environment of 250 people, the spread doesn’t have as much impact as in a dense environment of 1500 people. The simulation shows that within an hour or two, the spreads can include up to 16 people / 1 diseased person. Moreover, the results are effective when making decisions, such as containment of the area or extending the tracking radius. The proposed scheme identifies social distancing violations using IoT and YOLOv3 based paradigm and proposes to take immediate actions. Figure 5b demonstrates that the proposed IoT-based paradigm has almost 90% social distancing violation detection precision, as compared to earlier systems based on YOLOv2, which shows a maximum of 70% precision. Nevertheless, the proposed scheme is capable enough to detect violations and takes necessary actions of quarantine, isolation, and warning.
Our evaluation identifies that the proposed IoT and YOLOv3 based paradigm detects violations in 0.02 ms later than the earlier method of YOLOv2, as shown in Figure 6. The delay is incurred by the additional CNN layers, which substantially improves the detection precision. Both YOLOv2 and YOLOv3 based systems are fed with a video stream where a violation occurs in the fifth frame, which is immediately detected by both solutions. The 0.02 ms delay is minuscule and an acceptable trade-off for better and efficient gains. Nevertheless, the proposed IoT and YOLOv3 based solution provides a connected paradigm that ensures timely action and containment of the deadly COVID-19 virus. We believe that proactive and precautionary solutions are the key to reduce further dangers of such contagious diseases.
4. Conclusion

In this study, we propose a connected paradigm using IoT-based health monitoring and CNN-based object detection methods. The proposed scheme aims to contain the spread at early stages and allow people to continue with their social activities. Our novel solution identifies the social distancing violations using CNN-based YOLOv3 object detection and tracks exposed or infected people using smart wearables. The YOLO v3 darknet-53 model based on CNN achieves 90% accuracy in object detection to identify inter-object distance and violation of social distancing. In addition, the python simulation successfully traces all exposed people with the likelihood of getting infected with or without protective gears. Our comparative analysis shows that the proposed IoT and YOLOv3 based scheme outperforms existing YOLOv2 and faster R-CNN schemes. We believe that the proposed autonomous and connected paradigm is the stepping stone towards a well-informed and safe community.

Acknowledgment

This research was supported by Kumoh National Institute of Technology (2018104114).

References


