Location-aware hazardous litter management for smart emergency governance in urban eco-cyber-physical systems

Amirhossein Peyvandi¹ · Babak Majidi^{1,2} · Soodeh Peyvandi³ · Jagdish C. Patra⁴ · Behzad Moshiri^{5,6}

Received: 28 October 2020 / Revised: 21 September 2021 / Accepted: 28 September 2021 /

Published online: 3 January 2022

© The Author(s), under exclusive licence to Springer Science+Business Media, LLC, part of Springer Nature 2021

Abstract

Smart city management is facing a new challenge from littered face masks during COVID-19 pandemic. Addressing the issues of detection and collection of this hazardous waste that is littered in public spaces and outside the controlled environments, usually associated with biomedical waste, is urgent for the safety of the communities around the world. Manual management of this waste is beyond the capabilities of governments worldwide as the geo-spatial scale of littering is very high and also because this contaminated litter is a health and safety issue for the waste collectors. In this paper, an autonomous biomedical waste management framework that uses edge surveillance and location intelligence for detection of the littered face masks and predictive modelling for emergency response to this problem is proposed. In this research a novel dataset of littered face masks in various conditions and environments is collected. Then, a new deep neural network architecture for rapid detection of discarded face masks on the video surveillance edge nodes is proposed. Furthermore, a location intelligence model for prediction of the areas with higher probability of hazardous litter in the smart city is presented. Experimental results show that the accuracy of the proposed model for detection of littered face masks in various environments is 96%, while the speed of processing is ten times faster than comparable models. The proposed framework can help authorities to plan for timely emergency response to scattering of hazardous material in residential environments.

Keywords Smart city · COVID-19 pandemic · Edge surveillance · Waste management · Emergency response

✉ Babak Majidi
b.majidi@khatam.ac.ir

Extended author information available on the last page of the article

1 Introduction

Since the emergence of the COVID-19 pandemic and the significant growth in usage of personal protective equipment (PPE), mostly face masks and gloves, PPE waste is becoming a significant issue for governments worldwide. A major issue caused by the substantial increase in the usage of face masks is the littering problem. Millions of discarded face masks are littered in various urban and natural environments all around the world on a daily basis. These littered masks are a major source of microplastic fibres contamination of environment [17, 52]. This new type of waste leads to significant new challenges for urban waste management [59]. The discarded face masks can lead to the spread of the virus by multiple careers [4, 28]. The COVID-19 virus can remain infectious on the face masks for up to a week, the spread of the virus through biological and non-biological careers poses potential risks. Furthermore, since many people, including street cleaners and children, may be unaware of the dangers of the discarded masks, they can also be in danger.

The issue of safe management and treatment of biomedical waste is discussed significantly in the literature [27, 58]. However, until now, the production of hazardous medical waste is always concentrated at clinics and laboratories. Handling hazardous medical waste in these controlled environments allowed the waste collectors to plan for careful and safe treatment of the dangerous and infectious materials. However, due to the COVID-19 pandemic, significant quantities of highly contagious discarded face masks are littered in streets, parks, office buildings, residential areas, and natural environments. This issue creates an urgent challenge for the waste management community to address the problems of detection, collection, and treatment of this significant amount of biomedical waste in residential areas. Another major issue in dealing with this problem is the scale and urgency of the current situation. The discarded face masks are littered in almost every urban area around the world, and they are spreading to the natural environment through wind and running water. Recently, the discarded face masks are considered a significant pollution source for the rivers and marine ecosystems. Due to the scale of the problem, manual management of this issue is challenging, and the technology should be harnessed to handle this problem.

Real-time video monitoring and surveillance are gradually become a part of smart cities. The edge computing and System on Chips (SoC) based smart cameras allow computationally intensive vision systems to run at the edges of the Internet of Things (IoT) networks [35]. A potential application for these technologies is litter monitoring and waste management in smart cities. A solution for timely detection and management of hazardous biomedical litters such as COVID-19 infected face masks is this real-time video monitoring system. In this paper, a framework that uses edge video surveillance and location intelligence for autonomous detection of littered face masks is proposed. A new dataset of littered face masks, called MaskNet, is collected and a new deep neural network architecture, called LitterNet, for fast detection of discarded face masks in various environments is proposed. Using the location of the detected littered masks and a proposed location-based model that uses smart city geo-location datasets, the areas with higher probability of hazardous biomedical litter can be predicted for preventive smart waste management. This research has several novel contributions:

- (1) The novel multi-location littered face mask dataset, called MaskNet, gathered for training machine learning models for autonomous detection and classification of discarded face masks in various urban and natural environments. This dataset is collected in Steyr, Austria, and Tehran, Iran, and to the best of our knowledge, is the first dataset of littered

medical face masks in different environments. Although currently there are several face mask datasets available online, the experimental results in this research prove that the detection of littered face mask using these datasets leads to low accuracy. In order to have an accurate deep neural network model for littered mask detection, face masks with various degrees of degradation and deformation, covered in various levels of dirt and in various natural backgrounds under different light conditions should be presents in the dataset. The proposed MaskNet dataset addresses these requirements.

- (2) The second contribution of the paper is a new deep neural network architecture, called LitterNet, which is able to automatically detect and classify environments with discarded face masks with the accuracy of 96%. The speed of proposed model is 10 times faster than comparable architectures, which is an important factor in processing the required footage on the SoC processing units of edge surveillance nodes in the smart city with limited computational resources.
- (3) The third contribution of this research is a location intelligence model which combines the output of LitterNet with various models interpolated from smart city geo-location datasets in order to provide a series of proposals for hazardous litter management intelligence in smart cities for emergency response management.

The rest of the paper is organized as follows: Sect. 2 provide a review of the related works to this research. Section 3 presents the proposed framework, methodology of this research and the experimental setup. In Sect. 4, experimental results are detailed and discussed. Finally, Sect. 5 concludes the paper.

2 Litter management in smart eco-cyber-physical systems

In recent years, smart cities analyse every behaviour of the citizens and make smart decisions for autonomous management of various complex governance tasks. The visual surveillance and monitoring systems are also gradually integrated into many business environments for security and Business Intelligence (BI) based behaviour analysis. Artificial intelligence-based algorithms can enable smart cities to model human behaviour with high accuracy and provide decision support for governments. In this context, the smart city can be considered as an eco-cyber-physical system (ecoCystem) [30, 34] which is a combination of the living and nonliving components of the ecosystem in conjunction with the cyber-physical sensors and intelligent agents in the environment, interacting as a system. In many countries around the world, street cleaning requires sweepers to visit different locations in the city and manually verify if the street or location needs to be cleaned. This approach is time-consuming and is not cost-effective [8]. The governance intelligence platforms can use the collected data from surveillance camera networks and robotic agents for autonomous modelling of the litter patterns. The city management can use this data to model the littering behaviours and take precautionary measures to protect the environment and the safety of citizens. Furthermore, law enforcement can automatically identify and fine the people impacting the community's health with littering.

Artificial intelligence and robots were able to help the waste management community to increase the speed and scale of waste management significantly [33, 42, 54]. Alvarez et al. [5] proposed a process for electronic equipment recycling, which uses robots. Zhihong et al. [62] proposed Region Proposal Generation (RPN) and VGG-16 model for object recognition and pose estimation to sort littered bottles. Karbasi et al.

[26] investigated the sorting of used button cell batteries by utilizing robotic arms, and they proposed three different detection techniques for sorting this type of waste. Furthermore, in [54], Wang et al. used neural networks to assist a robot in construction sites and utilized Faster R-CNN methods to detect scattered nails and screws in real-time to enable the robot to automatically collect discarded nails and screws.

Visual litter analysis is one of the primary methods for modelling the quantity and characterization of litter. Autonomous sorting of the litter using image processing helps to separate various types of waste from each other for recycling [9, 20, 31, 41, 44]. Also, deep neural networks are able to provide practical visual modelling tools for detection and classification of various types of waste [10, 12, 31, 37, 44]. Sakr et al. [41] proposed that for reducing the waste contamination by other materials, the sorting and separation must be done as early as possible, and the need to automate this process is a significant facilitator for waste companies. To this end, the authors made use of convolution neural networks (CNN) and Support Vector Machines (SVM) for sorting plastic, paper, and metal wastes, using a Raspberry Pi 3 acting as the controller. In [9], the Bircanoglou et al. utilized different deep neural network architectures for intelligent waste sorting. Sreelakshmi et al. [44] investigated the segregation of plastic and non-plastic wastes.

During the current pandemic, Medical Waste Management (MWM) has become a major challenge for governments around the world. Kalantary et al. [25] assessed five hospitals in Iran and discussed that the COVID-19 pandemic led to an average of 102.2% increase of medical waste in both private and public hospitals. They further discussed that the ratio of infectious waste has increased by 9% in medical waste composition and 121% compared to pre-pandemic levels. In [51], the Tsai et al. analysed and compared the status of the medical waste generation before and after the COVID-19 pandemic in Taiwan and found that the difference was significant in the second half of 2020. Furthermore, the author suggested that all generated waste from healthcare facilities must be disposed of in officially certified incineration plants, with proper temperature to effectively kill viruses. This process will ensure safe and complete destruction of the COVID-19 virus and prevent further spread from medical wastes, such as face mask. Hartanto and Mayasari [21] took a different path and utilized the analytic hierarchy process (AHP) in order to determine appropriate material for making environmentally friendly face masks. In [43], Sari et al. investigated medical waste from COVID-19 intensive care hospitals and disposable face mask waste generation in Indonesia, and the results show that several tons of medical and face mask waste are being generated daily, and they must be safely transported to treatment facilities. Tagle et al. [47] addressed the lack of legislation for correct disposal of waste management in the current pandemic in Mexico. They further discuss that it is necessary to create a protocol for the MWM of households which during the pandemic can be classified as medical waste.

Torres and De-la-Torre [50] addressed the issue of face mask waste management in Peru, and their estimation suggests that about 15 million face masks waste are generated daily in their country. A significant portion of these masks are being mismanaged, thus polluting streets, beaches and marine environment of regions like Lima, Ancash, La Libertad, and Piura Regions. Moreover, in [40], it is discussed that medical mask waste can be found along with household waste without special treatment and separation, which can prove to be hazardous, and they require a unique treatment process. Joes et al. [24] presented an environment-friendly design of a trash bin for the current pandemic. Their proposed product is designed to treat and manage the medical mask wastes by changing the mask's shape using a hole puncher and further disinfecting them using two disinfectants provided in the storage box inside the trash bin. Even though the

proposed product disinfects face masks, the issues of littered masks in different environments and streets remains unsolved.

Tirkolaee and Aydin [7] focused on vehicle operations in waste management and minimizing the total costs of transportation, outsourcing, and use of vehicles during the pandemic by using CPLEX solver. Also, in [49], Tirkolaee et al. addressed the location-routing problem and proposed a mixed-integer linear programming (MILP) model along with fuzzy chance-constrained programming to investigate the routing problem for MWM in the current pandemic. Dharmaraj et al. [16] argued that since face masks are made of petroleum-based non-renewable polymers that are non-biodegradable, they are hazardous to the environment and create health issues. Littered face masks are ending up in rivers and open waters, thus endangering the marine ecosystem and their fauna and flora population.

However, until now, all the researches are focused on either the biomedical waste in controlled environments such as hospitals or dealing with collected waste in waste management sites. The issue of a significant quantities of littered face masks in streets, parks, office buildings, residential areas, and natural environments is a very important issue and this research deals with this issue for the first time. The scale and urgency of this problem necessitate the use of technology for a scalable and safe solution. Using mobile edge computing and installing high-resolution cameras on garbage trucks and other municipality vehicles, street cleaning can be performed more efficiently and faster [61]. Therefore, in this research a solution for addressing these issues is presented. Table 1 presents a summary of investigated related researches.

3 Location-aware hazardous litter detection

In smart cities, the cloud is commonly located far away from end-users, and therefore, it cannot offer low latency connections. In order to solve this issue, different edge computing methods are employed. The method suitable for the proposed framework is multi-access edge computing (MEC). In this approach, the data is processed between the cloud and the end-user. In our proposed framework, fixed and mobile surveillance cameras across the smart city, acting as modular rapidly deployable decision support agents (MoRaD DSA) [1, 2], provide the system with street footage as well as location and time. The MEC hosts can either be embedded in data collection agents (drones, robots, trucks), or they can be placed in fixed locations. The MEC host uses the proposed LitterNet model to process the surveillance videos and to determine the location of the discarded face masks in the streets. Meanwhile, the respective geo-location data (latitude and longitude) of the input video will also be collected. The geo-location data will be transmitted to the smart city cloud, which is responsible for performing further analysis. By having access to the geolocation data, using the proposed location intelligence model, the smart city authorities can make plans for waste collection and street cleaning and predict which districts, blocks, streets, and locations are going to be most littered with discarded face masks. Therefore, the smart city authorities will be able to make plans for the efficient routing of garbage trucks and street sweepers. If required by the smart city management, more accurate and computationally intensive models will be used on the cloud for better face mask detection or detection of other types of trash. The overall architecture of the proposed framework is presented in Fig. 1.

The proposed framework consists of two main phases: (1) *Data Collection & Preparation* and (2) *Model Training & Evaluation*. In the upcoming sections, the

Table 1 Summary of related works

Issue	Remarks	Refs
Medical waste quantity	The medical waste generation has significantly increased after the pandemic	[25, 43, 51] [7, 21, 47, 49]
Medical waste treatment and recycling	Medical wastes must be transported to treatment and incineration plants Environment-friendly masks can replace the current models The transportation routes of medical wastes can be optimized Regulations for households are also required	
Littered medical waste	Litter damages the marine ecosystem Significant increase in the number of face masks discarded on beaches and streets Littered face masks can increase the spread of the virus	[16, 24, 40, 50] [26, 31, 37, 44, 45, 54]
Litter management	Robots can be utilized to sort garbage Most works focused on sorting one or two types of trash Littered face mask detection has not been investigated

Fig. 1 Overview of the proposed framework for litter management

aforementioned phases will be discussed. In the proposed framework, first, input images with different resolutions will be captured by surveillance systems and along with the location of the place that the input image was acquired will be sent to the host. Since the input images come in different resolutions, they will be resized to the input shape (step 2.1 – 540×720 size pixels) of the LitterNet or DenseNet201 models. The surveillance footages with QCIF, CIF, 2CIF, 4CIF, and D1 standards with size of 176×120 , 352×240 , 704×240 , 704×480 , and 720×480 pixels, respectively, should be upscaled in order to be used in the proposed framework. Other standard formats such as 720p HD, 960p HD, 1.3 MP, 2MP, and other standards with higher resolutions can be used in their native resolution in the proposed system. After the input image is resized, it will be sent to the LitterNet model (2.2) in order to find out whether there are any littered face masks in the images or not. If a littered mask is detected (2.3) the resized input image, along with coordinates of the location of this littered face mask (2.4) will be sent to the cloud. In this case as the masks are detected on the edge devices the network latency is very low.

In the proposed framework the authorities should provide the required level of accuracy for litter detection. The experimental results show that the accuracy of the proposed computationally light LitterNet architecture, which is suitable for running on edge nodes, is 96%. If from authorities' point of view this accuracy is not acceptable, or if detection and classification of other forms of litter is required, the deployment of more computationally intensive neural models is essential. In this case, the input with its original resolution (before the resizing in 2.1) will be sent to the cloud. Next, in

the cloud, the littered mask image will be fed to the face mask detection model (3.1) in order to perform detection and predict the bounding box with high accuracy. Then, based on the input coordinates, proximity locations within 10 to 15 m of the actual location will be generated (3.2) to perform further geolocation analysis. These locations will be stored in a centralized server (3.3) for usage in later steps. Finally, by integrating all the acquired data, further Exploratory Data Analysis (EDA) and visualizations (3.6) can be done to provide authorities and human operators in smart city with more accurate information and tools to enable them to design different plans and optimize garbage collection.

It is important to mention that in case of transmission of the recorded footage by edge nodes to the cloud the scalability of the proposed framework will require significantly high bandwidth and only in cases of emergency this high accuracy system can be used. In these conditions the network latency is also high. However, advanced SoCs with neural computing capabilities may soon make this issue irrelevant. Also various methods for detection, re-identification, and tracking across multiple cameras without the need for storing the streaming data are proposed with may help in alleviating this problem [35].

3.1 Data Collection and Preparation

Due to the fact that before this research there was not any dataset of littered medical masks, as the first step in the proposed framework a new dataset containing images of littered face masks, called MaskNet, was collected. The MaskNet dataset was collected in Austria and Iran in July 2020. The dataset was collected daily for seven days in Steyr, Austria, and Tehran, Iran, during different times of day from 6 A.M. up to 6 P.M. from different environments such as streets, parks, riverbanks, inside buildings and offices. It consists of 1061 surgical masks images that are littered on the streets and other urban areas. In addition, images of streets without discarded masks were also collected. The original size of the dataset is 6.9 GB, and the photos have an image resolution of 3024×4032 pixels. The images are captured by three smartphones: An Apple iPhone 11, a Samsung Galaxy S8, and a Samsung Galaxy S10.

For sampling, the images were acquired with the default zoom of the aforementioned smartphones with the mask at the centre, top, bottom, corners of the image, and discarded masks were left in different states of deformation, various levels of dirt and near different objects, such as bins and bushes. In order to efficiently train the model used in the proposed system and normalize the dataset, the size of the images is reduced to the image size of 540×720 pixels, which reduces the size of the dataset to 220 MB. In order to train the models, 70% of images are used for training the LitterNet deep neural network, 13% for cross-validation, and 17% for testing. Figure 2 shows a few samples of the MaskNet dataset.

In order to model the realistic surveillance videos, prevent the model from overfitting and increasing the amount of data in the dataset, data augmentation (DA) is used. DA increases the number of data samples from each image by using various techniques such as changing the object's location in the picture, the scale of the image, rotating the image, and other transformations. The importance of using this technique, in addition to increasing the amount of data in the dataset, is preventing overfitting by allowing the model to see various images of the same object. This process will increase the accuracy of the neural network model. The DA was performed using Keras's *ImageDataGenerator* class. To define the best parameters for DA, various combinations of augmentation parameters are examined and the final DA parameters, presented in Table 2,

Fig. 2 Samples from the MaskNet dataset

Table 2 DA parameters

Parameter	Value
Rotation_range	20
Width_shift_range	0.1
Height_shift_range	0.1
Shear_range	0.1
Zoom_range	0.2
Vertical_flip	True
Horizontal_flip	True

were chosen. This data augmentation makes the data relevant for video monitoring applications as the changes in scale and rotation simulates the fixed and moving surveillance camera footage.

An important issue in collection of datasets is randomness of data. For deep learning datasets it is essential to have a high-quality data and the data should cover as many cases as possible and be a good representation of reality. A model trained on high quality dataset can capture the underlying characteristics of the data and makes the predictions more generalize. In this research, the DA is used for oversampling minority classes and generating new data to ensure high data quality dataset that doesn't have under-represented classes. One of the future directions of this research is making better datasets using Generative Adversarial Networks (GANs).

Another dataset used in this research is the TrashNet dataset [56]. The images in this dataset are photos of garbage with a white background. The original dataset is approximately 3.5 GB, with an image size of 2448×3264 . However, the resized version is around 40 MB in size and 512×384 size images. The dataset consists of 2527 images distributed in six categories of cardboard (403 samples), glass (501 samples), metal (410 samples), paper (594 samples), plastic (482 samples), and trash (137 samples). The trash category does not have sufficient samples, thus, in this work, we decided it is best to remove this category. The final dataset consisted of 2390 images. Just as before, we used 70% of images for training, 17% for testing, and the remaining 13% for cross-validation. Figure 3 shows several samples from this dataset.

The augmentation parameters were utilized for both MaskNet and TrashNet datasets with the same values. In the next section, the training process of the models utilized in this research will be investigated.

Fig. 3 Samples from the TrashNet dataset

3.2 Model training & evaluation

In this research, for rapid littered face mask detection, a light CNN-based architecture called LitterNet is proposed. The LitterNet architecture consists of Conv2D and Max-Pooling2D layers, as presented in Fig. 4. In this network, layers close to the input have fewer convolutional filters, while layers that are located deep in the network and immediate to the output have more filters. Therefore, the first Conv2D layer has only 16 filters, with a 3×3 strides and ReLU activation function. Then there is a MaxPooling2D layer with a 2×2 pool size. After that, the Conv2D's filter expands to 32, while other arguments are held as the previous layer. The third and final Conv2D layer has 64 filters, which are then passed to a Flatten layer after passing through the final MaxPooling2D layer. The goal of the design of the architecture was simplicity and speed while keeping the accuracy as high as possible. To prevent overfitting, a Dropout layer with the probability of 0.5 is used. A 16-unit dense layer is embedded in the architecture with rectified linear unit (ReLU) activation, whose

Fig. 4 The architecture of LitterNet and the process of mask and trash detection

output is then passed to the last dense layer, which utilizes SoftMax as its activation function to identify whether there is a littered face mask in the photos or not. Two optimizers, Adam [60] and RMSProp [64], are used as optimization algorithms. If LitterNet finds any discarded mask in the input, the respective geolocation of the input image will be marked.

For evaluation of the proposed model, several state-of-the-art deep neural network architectures are investigated for the proposed application. DenseNet [22] is a deep neural network architecture in which connections between input and output layers are very short. Since it utilizes densely connected layers, this model requires fewer layers. In DenseNets, the connections between input and output layers are very short. Following the feed-forward nature of the network, each layer in a dense block receives feature maps from all the other layers and passes its output to all subsequent layers. Feature maps are fused through concatenation. These connections form a dense circuit of pathways that allow better gradient flow, and each layer has direct access to the gradients of the loss function and the original input signal. Because of these dense connections, the model requires fewer layers, as there is no need to learn redundant feature maps, allowing the collective knowledge (features learned collectively by the network) to be reused.

Another investigated architecture is Inception-V3 [46], which is a CNN that is trained on more than a million images from the ImageNet dataset. This architecture is consisted of 48 layers and can classify images into 1000 different object categories. The next architecture used in this study for comparison of results is Neural Architecture Search Network (NASNet) [63], in which the cells are not predefined and instead designed by the reinforcement learning. In this architecture a building block is optimized on a small dataset and then it will be transferred to a larger dataset. The best convolutional layer on CIFAR-10 is selected and then these layers will be applied to the ImageNet by stacking several copies of them together. The next architecture is a modified depth-wise separable convolution called Xception [11]. This architecture is a pointwise convolution followed by a depth-wise convolution in which depth-wise convolution is the channel-wise $n \times n$ spatial convolution. For object detection architecture, EfficientDet [48] is used, that is based on a weighted Bi-directional Feature Pyramid Network (BiFPN) and trained on the MS-COCO dataset. Finally, for transfer learning, DenseNet121, DenseNet169, DenseNet201, InceptionV3, Xception, NASNetLarge, and NASNetMobile architectures were utilized.

After choosing the best hyperparameters, the final setup and hyperparameters were the same for all trained models, both on MaskNet and TrashNet datasets. The models were trained for 50 epochs and a batch size of 16 using Adam and RMSProp optimizers and DA. Furthermore, the default image size of each dataset (540×720 for MaskNet, and 512×384 for TrashNet) was used as the default input shape of every trained model. The EfficientNet object detection network was retrained using TensorFlow Object Detection API. At first, 40,000 steps were set for training. However, after 18,000 steps, the loss began to increase slowly, and therefore we had to stop the process. Moreover, the loss was below 0.1 (mostly between 0.08 and 0.095) for the final few hundred steps. The experiments of this study were performed using Keras library with the TensorFlow backend (Version 2.5) on the Google Colab platform. The experiments were conducted on an Nvidia Tesla K80 GPU.

In total, 31 different models were trained on both MaskNet and TrashNet datasets. For the novel MaskNet dataset, 14 models (7 models with Adam optimizer, and 7 using RMSProp) were trained using transfer learning and aforementioned architectures, plus 2 models by utilizing the proposed LitterNet architecture (1 with Adam, and 1 with RMSProp), and one object detection model, summing up to 17 models in total. For the TrashNet dataset, 14 models in total were trained by utilizing the previously discussed

methods and architectures. After the models were trained on both datasets, their performance needed to be compared and evaluated in order to choose a final model for each task (face mask and trash detection). For this task, the performance of the models on the test set was compared. The results of this comparison are presented in experimental results section.

3.3 Location intelligence for predictive biomedical waste modelling

Using the predictive geospatial location of litter, the smart city management can efficiently plan the routing of garbage trucks and rapidly remove the hazardous biomedical litter from the streets. The geo-location data generated by LitteNet can validate and suggest the exact points where the face masks are littered. However, the possibility of locating a significant portion of littered masks using visual surveillance is not practical. Therefore, the proposed framework uses a collection of smart city data along with the locations provided by the LitterNet in order to predict the litter cluster locations.

The first dataset which can help the location intelligence for litter cluster prediction is the locations of crowds. The crowd is usually the main source of litter and the number of people gathered in a place have correlation with the amount of litter in that place. The dataset of geo-locations of people favourite sites in the city, as show in Fig. 5a, can indicate the possibility of litter clusters in the areas with high visitor density.

A dataset of locations of buildings with public-interest is also available in the smart city. This geospatial dataset, presented in Fig. 5b is obtained by applying height of the structure to better visualize the larger buildings, comes in three categories: (a) list of buildings that are of national/international importance, (b) buildings of regional importance, and finally,

Fig. 5 **a** Clusters of people favourite locations in the city, **b** height map and location of buildings with public-interest, **c** the clusters of locations of places where people eat publicly and shop, **d** the geospatial clusters of public litter bins

(c) buildings of local importance. It should be noted that most of the restaurants and public litter bins are also located in the vicinity of these buildings.

Another important dataset is the locations of restaurants, pubs, cafes, takeaways, hotels and other public places, as well as supermarkets, which is presented in Fig. 5c. These locations are another important factor for density of litter in the streets. Another important factor regarding littering is that some people, due to carelessness, discard their trash outside of litter bins. Also, in many cases the litter bins are not emptied in time. Therefore, in many places there are trash clusters in vicinity of litter bins. Therefore, the final geospatial dataset utilized in this work is public litter bin's location dataset as shown in Fig. 5d. Combining these geospatial locations, probable location of litter clusters can be generated. This information combined with validation and reporting of actual litter obtained from the proposed vision system enables the authorities to plan routing of garbage trucks and street sweepers.

4 Experimental results and discussion

4.1 Mask detection and classification using LitterNet

For training the best model on the MaskNet dataset, different hyperparameters were tested. Epochs ranging from 30 to 50, with batch sizes of 8, 16, and 32, were tested. The best models were trained for 50 epochs since epochs below that led to under-fitting, and models trained for 60 epochs were over-fitted (10–12%). The final models were trained using a batch size of 16. The best-performed model for classifying MaskNet dataset was the DenseNet201 + Adam trained model, which achieved a test accuracy of 98.13% using DA and Adam optimizer. Furthermore, this model offered the least overfitting, with a training accuracy of 98.91%. After that, the DenseNet169 + Adam trained model showed a test accuracy of 97.04%. The NASNetLarge + RMSProp and Xception + RMSProp offered similar classifications, with 94.39% and 95.33% test accuracies. Furthermore, the NASNet-Mobile and DenseNet121 trained models performed similarly to each other. The worst performance was observed from the InceptionV3 + Adam trained model, with a test accuracy of 88.79%.

The proposed LitterNet model offered a test accuracy of 94.39% for both Adam and RMSProp models. However, the Cross-entropy loss was lower for the Adam trained model (0.14) compared to RMSProp model (0.32). The LitterNet shows a promising performance, even though it is a simple network compared to other pre-trained and complex models and networks utilized for comparison, and it outperformed the NASNetMobile, InceptionV3, and DenseNet121 architectures for the given task. The results of the classification of the MaskNet dataset by various deep neural network architectures are presented in Table 3.

Table 4 shows the total number of the parameters for LitterNet, DenseNet201 and DenseNet169 models. As presented, the proposed LitterNet model has a lower number of parameters, which boosts the performance, and the time for prediction is much lower than DenseNet201 trained model. Since the goal of the framework is to deployable on the edge nodes and the models have to work with a stream of videos, lightness and speed are the most vital performance measures for the selection of the neural network architecture. Therefore, when the LitterNet model presents a much faster performance, the 2–4% difference in the accuracy between LitterNet and other models such as DenseNet201 may be tolerable.

Table 3 Training and test results for MaskNet dataset

Model	Adam		RMSProp	
	Training (%)	Test (%)	Training (%)	Test (%)
DenseNet121	95.73	92.52	95.66	93.46
DenseNet169	96.90	97.04	95.0	96.26
DenseNet201	98.91	98.13	98.45	96.83
InceptionV3	95.47	88.79	94.88	93.46
Xception	97.04	91.59	94.91	95.33
NASNetLarge	97.67	93.46	98.25	94.39
NASNetMobile	94.79	91.17	94.78	92.84
LitterNet	95.26	94.39	95.95	94.39

The performance comparison of LitterNet + Adam model (a) with the DenseNet201 + Adam trained model (b) is presented in Fig. 6

Fig. 6 Performance comparison of LitterNet (a) and DenseNet201 (b) on MaskNet**Table 4** Total number of parameters of LitterNet, DenseNet201, and DenseNet169 models

Model	LitterNet	DenseNet201	DenseNet169
Total no. parameters	5880,914	18,325,826	12,646,210

In order to further evaluate the performance of the proposed LitterNet, its prediction time on CPU was compared with that of the DenseNet201 + Adam and DenseNet169 + Adam trained models, and the results are presented in Table 5. The testing was conducted on Google Colab, which uses a single core hyper-threaded Xeon Processors @ 2.00Ghz (1 core, 2 threads). For this task, a 7-s video of littered masks and a clean street was captured using a Samsung Galaxy S10, and the framerate (FPS) was reduced to a common FPS of 24 for convenience. From this footage, 168 photos were extracted and fed to the models for prediction.

Table 5 Prediction time comparison of LitterNet, DenseNet201, and DenseNet169

Model	LitterNet	DenseNet201	DenseNet169
CPU time	User: 4.98 s Sys: 150 ms Total: 5.13 s	User: 52.6 s Sys: 1.96 s Total: 54.6 s	User: 43.9 s Sys: 1.59 s Total: 45.5 s
Normal time	3.22 s	30 s	64 s

In this table, the normal time is measured by a stopwatch manually, and starts from the time the one-line of code was submitted to the CPU (to predict the set of 168 photos using LitterNet model) to the time when the process completed. User is the amount of CPU time spent in user-mode code (outside the kernel) within the process, and this is only the actual CPU time used in executing the process. Sys is the amount of CPU time spent in the kernel within the process. This means executing CPU time spent in system calls within the kernel, instead of library code, which is still running in user-space. Finally, the Total indicates how much actual CPU time the process of prediction is used.

As presented, the LitterNet model, with a total CPU time of 5.13 s, performs significantly faster (almost 10 times faster) compared to DenseNet201 and DenseNet169 models, which predict the same set of 168 photos using the same CPU and setup in 54.6 and 45.5 s, respectively. Therefore, the results suggest that utilizing the LitterNet model for edge surveillance can be computationally cost-effective, and therefore, its marginally lower accuracy may be tolerated. Figure 7 shows LitterNet results, specifying in which pictures a littered face mask is found and in which picture no discarded face mask exists.

However, the model did fail to correctly detect and classify a few instances in the dataset. As presented in Fig. 8, the model was unable to accurately predict images in which the mask was further away than other samples or a sample that had captured a littered tissue (left).

Fig. 7 LitterNet classification results

Fig. 8 A few samples of wrong predictions of LitterNet

As for the EfficientNet detection model, the loss was below 0.1 (mostly between 0.08 and 0.095) for the final few hundred steps. Overall, the trained mask detection model was promising since it was able to detect surgical masks in different and complex environments. In Fig. 9, different losses of the mask detection model, such as classification loss (a), localization loss (b), regularization loss (c), and total loss (d), visualized by TensorBoard, are presented.

Fig. 9 Different losses of the mask detection model

In Fig. 10, a collection of correctly detected littered face masks from MaskNet dataset using EfficientNet are presented. The masks are in different backgrounds, including dirt road, paved road, between leaves, on stairs, and tangled in a fence. All the masks are detected with high accuracy.

The next stage of experiments was testing the proposed model on images it has never seen before. A series of littered face mask images are collected from Google Image service, and the automatic detection of the face masks in these images using the proposed model is investigated. As presented in Fig. 11, in most cases, the proposed system managed to detect the mask with high accuracy. The model even managed to detect an N95 mask, which was not available in the dataset and was never seen by the model while it was being trained.

As previously discussed, the littered face masks endanger the marine ecosystem, and detecting and collecting these litters can save the ecosystem. As presented in Fig. 12, the proposed system was able to detect the littered face masks in the water with high accuracy. It is worth noting that there were not any samples of littered masks under the water in the dataset, and the model had not seen such instances while it was being trained yet managed to detect these litters correctly. Underwater object detection is a challenging task and usually image enhancement algorithms are required before further processing stages. Colour correction and contrast enhancement are required to compensate for scattering and absorption of light by water and the underwater suspended particles which have severe effects on image quality. Images of objects underwater usually have degradations such as colour fading, contrast reduction, and detail blurring [32].

However, there are few instances in which the model either was unable to detect the masks correctly or was only able to find a percentage of the masks presented in the images. Figure 13 shows several samples of such cases. As the figure shows, the detection model

Fig. 10 Samples of correctly detected littered face masks from MaskNet dataset using EfficientNet

Fig. 11 Face mask detection on images from Google Image service

Fig. 12 Underwater detection of the face masks

fails to detect all littered masks correctly in densely mask littered areas. There are also mislabelled cases with complex backgrounds, a large number of overlapping face masks, and circumstances in which the face masks are mislabelled as gloves.

Fig. 13 The samples in which the proposed system was unable to detect the masks correctly and completely

There are certain elements influencing the accuracy of the proposed system. The limitation of computer vision algorithms, including reliance on good environmental light and weather, is a factor that may limit the applicability of the proposed system in lousy weather or close to sunset and during the night. In the proposed framework, detection and classification of other forms of litter is also considered. To find the best model for this part of the framework, in the next section, the results of various deep models trained on the TrashNet dataset are analysed.

4.2 Classification of other forms of litter

For classifying the TrashNet dataset, the formerly investigated pre-trained models were investigated. Same as before, different hyperparameters with different values were researched, and the best parameters and values were chosen for the given task. The final models were trained for 50 epochs with a batch size of 16. DenseNet201+RMSPProp trained model, with a test accuracy of 94.04%, presented the best performance among other models. After that, the DenseNet169+Adam model provided 92.72% of accuracy on the test set, and NASNetLarge presented 92.09% and 91.46% for RMSPProp and Adam trained models, respectively. NASNetMobile, DenseNet121, and InceptionV3, with 88%~89% test accuracy, offered the worst performance among other trained models. Most of the models are overfitted by 7 to 8 percent, and the best model is overfitted by 4%, which is acceptable, given the complex task of classifying five different categories of trash. The proposed LitterNet has a comparably low accuracy of 81% for classification of TrashNet dataset. The reason for this low accuracy is similarity of samples in different classes of TrashNet dataset and the simplicity of the LitterNet architecture. The results of the classification of the TrashNet dataset by various deep neural network architectures are presented in Table 6.

Table 6 Training and test results for TrashNet dataset

Model	Adam		RMSProp	
	Training (%)	Test (%)	Training (%)	Test (%)
DenseNet121	97.35	89.87	97.33	89.87
DenseNet169	98.66	92.72	98.47	92.09
DenseNet201	99.12	91.77	98.34	94.04
InceptionV3	96.93	88.92	97.12	89.87
Xception	98.76	90.19	97.99	89.87
NASNetLarge	99.56	91.46	99.06	92.09
NASNetMobile	93.08	88.61	92.09	87.34

The performance of the top two trained models (DenseNet201 + RMSProp & DenseNet169 + Adam) on the TrashNet dataset is presented in Fig. 14. Furthermore, Fig. 15 shows the confusion matrix of the DenseNet201 + RMSProp trained model. As shown, the model confuses some samples of glass trashes with metal and cardboard with paper.

The confusion of the aforementioned categories is quite understandable since there are samples that, as shown in Fig. 16, might be confused even by humans. Examples of these samples are the crumpled foil, which is confused with a paper, or cardboard confused with a piece of paper.

There have been several studies of machine vision-based trash classification in the past few years. In the RecycleNet project [9], various architectures and optimizers are compared, and the best model for trash classification is selected. The test accuracy in the RecycleNet project is 81%, and for the DenseNet121 model in their project is 95% after 200 epochs, whereas the DenseNet201 model in this study achieves 94% accuracy after only 50 epochs. Aral et al. [6] achieved test accuracies of 94% and 95% after 120 + epochs, whereas the same was accomplished in this study after only 50 epochs. Therefore, the accuracy of the proposed model is comparable with similar models in the

Fig. 14 Performance comparison of DenseNet201 (a) and DenseNet169 (b) on TrashNet

Fig. 15 Performance comparison of DenseNet201 (a) and DenseNet169 (b) on TrashNet

Predicted: paper / Actual: cardboard

Predicted: paper / Actual: cardboard

Predicted: metal / Actual: paper

Predicted: paper / Actual: metal

Predicted: metal / Actual: glass

Predicted: plastic / Actual: glass

Fig. 16 Mislabelled predictions in TrashNet dataset

research. In the next section, location intelligence for predictive biomedical waste cluster detection will be investigated.

4.3 Location intelligence for predictive biomedical waste cluster detection

For the location intelligence dataset, different geo-location dataset publicly provided by Dundee City Council, Scotland, UK, are utilized. In a survey in Dundee city, people were asked to pinpoint their favourite location in the city, and the result was the *Dundee Places* dataset [13]. These locations are the most populated in the city, which have the potential to also be the most polluted in the city. A list of historic-interest buildings of Dundee city is

Fig. 17 Clusters of probable locations of litter in the smart city

Fig. 18 Per district litter location intelligence

also publicly provided [14] and used in this study. The locations of restaurants, pubs, takeaways, hotels, as well as supermarkets and other food shops in Dundee city, are provided by the food standards agency (FSA) [18]. The Public Litter Bins dataset, which consists of locations of all public bins within Dundee city [15], is used for litter bin locations. As demonstrated in Fig. 17, combining these geospatial locations, probable location of litter clusters can be generated. This information, combined with validation and reporting of actual litter from LitterNet enables authorities to plan the routing of garbage trucks and street sweepers.

In order to make the location intelligence map more usable for waste management, as presented in Fig. 18, the litter location information can be clustered to provide per district information for the smart city authorities.

This information enables smart city management to offer optimized solutions for collecting garbage across different areas, blocks, and streets of the city. At the first level of the multi-level assessment model (Level 1), the waste management authorities have the litter

information for the entire city. In Level 2, the per district analysis of the litter information is presented. At Level 3, adjacent blocks and, in Level 4, a single block is analysed. Level 5 focuses on a single point of interest (PoI) which can be either a restaurant, supermarket, pub, grocery store, takeaway, hotel, or even a public litter bin. The points provided by LitterNet, are used for validation and adding information at Level 6. The multi-level analysis of littering probability in Dundee city is presented in Fig. 19.

The measures taken for the management of the littering problem can be categorized into four main groups: the preventive measures, the mitigating measures, the removing measures, and the behaviour-changing measures. The proposed system can help to identify the mask litterers and help to provide preventive, mitigating, and educational solutions for these individuals.

Due to the large scale of the biomedical littering problem, labour-based management of this issue is very hard, expensive, and time-consuming. The waste management decision-makers require detailed and real-time knowledge regarding issues such as magnitude and sources of biomedical litter. Using this knowledge, the smart city management may appropriately intervene to develop rapid cleaning programs, enforce the law and establish educational plans in the area. The city management can use the automatic biomedical litter classification system for rapid and large-scale city-wide litter analysis and modelling. The proposed system allows the reduction in the cost and increase in the scale and speed of litter surveying in smart cities. The close to the real-time geospatial distribution of biomedical litter and up-to-date region level models can be visualized for smart city management and to provide an encompassing view of the health issues from face mask littering affecting the region. Automatic face recognition and person tracking [39] can be used for the

Fig. 19 Multi-level biomedical litter predictive modelling

enforcement of littering laws. In this case, the privacy of the citizens becomes an issue for which autonomous artificial intelligence based privacy persevering methods provide accurate solutions [3].

Another application of the LitterNet is as the vision system of street cleaning robots. The current street sweeping robots can be equipped with a LitterNet based vision system to automatically detect and collect the littered face masks. A robotic trash sorting machine can use the LitterNet system for sorting the face masks and separating them from other forms of trash. The LitterNet can also be used for automatic counting and volume estimation of the face mask in bins and dumps. Various issues regarding the implementation of camera network for this task such as implementation using Software Defined Networks [36] can also be the topic of future research. Furthermore, the issues of privacy preserving in the collection of data from the smart city is an important issue and machine learning as service model can be used for training necessary models while keeping the privacy of citizens [38].

The collection of a larger dataset and training the LitterNet by all the situations it may encounter in real-world scenarios is essential for making the proposed system more reliable. This dataset needs to be updated at regular intervals in order to provide an accurate view of the littered face mask patterns. The LitterNet can also help in the automation of sorting the litter and finding the hazardous face masks without posing health risks to the waste collectors.

4.4 Performance comparison with related works

There are many different face mask datasets publicly available on the internet, such as the Face Mask Detection~12 K Images Dataset (MD-12 K, Available at: <https://www.kaggle.com/ashishjangra27/face-mask-12k-images-dataset>) and Face Mask Lite Dataset (FML, Available at: <https://www.kaggle.com/prasoonkottarathil/face-mask-lite-dataset>). The MD-12 K dataset consists of six thousand images of people wearing face mask that are scrapped from google search, and images without the face mask from the CelebFace dataset (Available at: <https://www.kaggle.com/jessicali9530>), summing up to approximately twelve thousand samples in total. The FML dataset consists of a total of twenty thousand samples, ten thousand for each category (with/without mask), and all samples (all human faces) are generated by style-based GAN architecture.

These datasets are mainly focused on identifying whether a person is wearing a face mask or not. In this part of experimental results, it is investigated that whether a deep neural network model trained on these datasets is suitable to littered face mask detection. Therefore, the LitterNet model was trained on MD-12 K and FML datasets, and the performances were compared with the model trained on MaskNet dataset. For training the new models, the same hyperparameters and set ups as training the LitterNet model on the MaskNet dataset were

Table 7 Performance of models trained on face mask detection datasets for littered face mask detection

Train dataset	MD-12 K	FML
Test dataset	MaskNet	MaskNet
Accuracy	53.19%	51.77%

Dataset: MaskNet

Pred.: No Mask / Act.: Mask

Pred.: No Mask / Act.: Mask

Pred.: No Mask / Act.: Mask

Fig. 20 Wrong predictions from models trained on available face mask detection datasets

used. A model comparison was performed for the current simulation and the results are presented in Table 7.

As presented in Table 7, the MD-12 K trained model yielded 53.19% test accuracy for detecting littered face masks. Also, the trained model on FML dataset offered an accuracy of 51.77% for detection of littered face masks. It can be concluded that the mask detection datasets in previous research are not suitable to detecting littered face masks in different environments. A few samples of wrong predictions are presented in Fig. 20.

It can be concluded that in order to have an accurate deep neural network model for littered mask detection, face masks with various degrees of degradation and deformation should be present in the dataset. Furthermore, the masks should be covered in various levels of dirt and also various natural environments and different light conditions should be present in the dataset. The proposed MaskNet dataset addresses these requirements. In order to present the advantages and contributions of this research the proposed framework is compared with some of the proposed methods in 2021 for waste management and face mask detection in Table 8.

5 Conclusion and future works

Due to the COVID-19 pandemic, the littered medical masks are currently endangering different cities and environments. In this paper, a biomedical litter management framework that uses video surveillance and location intelligence for detection and predictive modelling of littered face masks is proposed. The proposed LitterNet model achieved an accuracy of 96% for detecting medical masks with optimized hyperparameter tuning. The simplicity of the model allows implementation on MEC and high speed of processing. The accuracy of the proposed model is more than enough for practical applications. A location intelligence based multi-level predictive litter model for enabling authorities to efficiently offer an optimized plan for cleaning streets of the smart city is also presented in this research. For future direction of research acquisition of more littered medical masks samples from a wide variety of environments, such as beaches, along with other types of face masks is necessary. Furthermore, by using GANs on the MaskNet more samples can be generated. Building a robot or a drone capable of detecting and cleaning littered medical masks in

Table 8 Comparison of the proposed framework with existing works

References	Dataset	Model	Location intelligence
Proposed framework	Littered face masks in different conditions are considered in the dataset	Customized deep neural network architecture for fast face mask detection on edge nodes	Location intelligence for predictive modelling of hazardous waste patterns
Gondal et al. [19]	The dataset for litter detection does not contain face masks	The model is capable of classification of various forms of litter, however have limited applicability on edge	Not considered
Hussain et al. [23]	Face mask dataset does not contain littered face mask	The model cannot detect littered face mask with high accuracy	Not considered
Loey et al. [29]	Face mask dataset does not contain littered face mask	The model cannot detect littered face mask with high accuracy	Not considered
Nazerdeylami et al. [34]	The dataset for litter detection does not contain face masks	The model is capable of classification of various forms of litter, however have limited applicability on edge	Not considered
Yang et al. [55]	The dataset for litter detection does not contain face masks	The model is capable of classification of various forms of litter, however have limited applicability on edge	Not considered
Wang et al. [53]	The dataset for litter detection does not contain face masks	The model is capable of classification of various forms of litter, however have limited applicability on edge	Not considered
Yu et al. [57]	Face mask dataset does not contain littered face mask	The model cannot detect littered face mask with high accuracy	Not considered

different environments is a topic of interest. Finally, in future, the proposed system will be utilized in a large-scale real-life scenario.

Declarations

Conflict of interest The authors declare that they have no conflict of interest.

Ethical approval This article does not contain any studies with human participants or animals performed by any of the authors.

Informed consent Informed consent was obtained from all individual participants included in the study.

References

1. Abbasi MH, Majidi B, Manzuri MT (2018) Deep cross altitude visual interpretation for service robotic agents in smart city. In: 2018 6th Iranian Joint Congress on Fuzzy and Intelligent Systems (CFIS). IEEE
2. Abbasi MH, Majidi B, Manzuri MT (2018) Glimpse-gaze deep vision for modular rapidly deployable decision support agent in smart jungle. In: 2018 6th Iranian Joint Congress on Fuzzy and Intelligent Systems (CFIS). IEEE
3. Abbasi MH et al (2019) Deep visual privacy preserving for internet of robotic things. In: 2019 5th Conference on Knowledge Based Engineering and Innovation (KBIE)
4. Almendros A (2020) Can companion animals become infected with Covid-19? *Vet Record* 186(12):388–389
5. Alvarez-de-los-Mozos E, Renteria A (2017) Collaborative robots in e-waste management. *Procedia Manuf* 11:55–62
6. Aral RA et al (2018) Classification of trashnet dataset based on deep learning models. In: 2018 IEEE International Conference on Big Data (Big Data)
7. Babaei Tirkolaee E, Aydın NS (2021) A sustainable medical waste collection and transportation model for pandemics. *Waste Manage Res*. <https://doi.org/10.1177/0734242X211000437>
8. Balchandani C et al (2017) A deep learning framework for smart street cleaning. In: 2017 IEEE Third International Conference on Big Data Computing Service and Applications (BigDataService). IEEE
9. Bircanoğlu C et al (2018) RecycleNet: intelligent waste sorting using deep neural networks. In: 2018 Innovations in Intelligent Systems and Applications (INISTA)
10. Bircanoğlu C et al (2018) RecycleNet: Intelligent waste sorting using deep neural networks. in 2018 Innovations in Intelligent Systems and Applications (INISTA). IEEE
11. Chollet F (2017) Xception: deep learning with depthwise separable convolutions. In: 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)
12. Chu Y et al (2018) Multilayer hybrid deep-learning method for waste classification and recycling. *Comput Intell Neurosci* 2018:1
13. City D (2019) Dundee Places. <https://data.dundee.gov.uk/dataset/places-database>. Accessed 17 July 2021
14. City D (2019) Listed Buildings. <https://data.dundee.gov.uk/dataset/listed-buildings>. Accessed 17 July 2021
15. City D. Public Litter Bins 2019. <https://data.dundee.gov.uk/dataset/public-bin-locations>. Accessed 17 July 2021
16. Dharmaraj S et al (2021) The COVID-19 pandemic face mask waste: a blooming threat to the marine environment. *Chemosphere* 272:129601
17. Fadare OO, Okoffo ED (2020) Covid-19 face masks: a potential source of microplastic fibers in the environment. *Sci Total Environ* 737:140279
18. FSA DC. Food Hygiene Information Scheme 2019. <https://data.dundee.gov.uk/dataset/food-hygiene-information-scheme>. Accessed 17 July 2021
19. Gondal AU et al (2021) Real time multipurpose smart waste classification model for efficient recycling in smart cities using multilayer convolutional neural network and perceptron. *Sensors* 21(14):4916

20. Gundupalli SP, Hait S, Thakur A (2017) A review on automated sorting of source-separated municipal solid waste for recycling. *Waste Manage* 60:56–74
21. Hartanto BW, Mayasari DS (2021) Environmentally friendly non-medical mask: an attempt to reduce the environmental impact from used masks during COVID 19 pandemic. *Sci Total Environ* 760:144143
22. Huang G et al (2017) Densely connected convolutional networks. In: 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)
23. Hussain S et al (2021) IoT and deep learning based approach for rapid screening and face mask detection for infection spread control of COVID-19. *Appl Sci* 11(8):3495
24. Joes S et al (2021) The design development of an ergonomic public trash bin for COVID-19 medical mask waste
25. Kalantary RR et al (2021) Effect of COVID-19 pandemic on medical waste management: a case study. *J Environ Health Sci Eng* 19(1):831–836
26. Karbasi H et al (2018) Robotic sorting of used button cell batteries: utilizing deep learning. In: 2018 IEEE Conference on Technologies for Sustainability (SusTech). IEEE
27. Kargar S, Paydar MM, Safaei AS (2020) A reverse supply chain for medical waste: a case study in Babol healthcare sector. *Waste Manage* 113:197–209
28. Leroy EM, Gouilh MA, Brugère-Picoux J (2020) The risk of SARS-CoV-2 transmission to pets and other wild and domestic animals strongly mandates a one-health strategy to control the COVID-19 pandemic. *One Health* 10:10013
29. Loey M et al (2021) Fighting against COVID-19: a novel deep learning model based on YOLO-v2 with ResNet-50 for medical face mask detection. *Sustain Cities Soc* 65:102600
30. Majidi B et al (2021) Geo-spatiotemporal intelligence for smart agricultural and environmental eco-cyber-physical systems. In: *Enabling AI Applications in Data Science*. Springer, pp 471–491
31. Melinte DO et al (2019) Deep learning computer vision for sorting and size determination of municipal waste. In: *International Conference of Mechatronics and Cyber-Mixmechatronics*. Springer
32. Musić J et al (2020) Detecting underwater sea litter using deep neural networks: an initial study. In: 2020 5th International Conference on Smart and Sustainable Technologies (SpliTech)
33. Nazerdeylami A, Majidi B, Movaghar A (2019) Smart coastline environment management using deep detection of manmade pollution and hazards. In: 2019 5th Conference on Knowledge Based Engineering and Innovation (KBEI). IEEE
34. Nazerdeylami A, Majidi B, Movaghar A (2021) Autonomous litter surveying and human activity monitoring for governance intelligence in coastal eco-cyber-physical systems. *Ocean Coastal Manage* 200:105478
35. Neff C et al (2019) REVAMP 2 T: real-time edge video analytics for multicamera privacy-aware pedestrian tracking. *IEEE Internet Things J* 7(4):2591–2602
36. Norouzi A, Majidi B, Movaghar A (2018) Reliable and energy-efficient routing for green software defined networking. In: 2018 9th International Symposium on Telecommunications (IST). IEEE
37. Nowakowski P, Pamuła T (2020) Application of deep learning object classifier to improve e-waste collection planning. *Waste Manage* 109:1–9
38. Peyvandi A et al (2021) Computer-aided-diagnosis as a service on decentralized medical cloud for efficient and rapid emergency response intelligence. *New Gener Comput*. <https://doi.org/10.1007/s00354-021-00131-5>
39. Qezavati H, Majidi B, Manzuri MT (2019) Partially covered face detection in presence of headscarf for surveillance applications. In 2019 4th International Conference on Pattern Recognition and Image Analysis (IPRIA)
40. Rozana K et al (2021) Microbiome analysis of medical mask waste as an early step to prevent environmental pollution due to unstandardized waste treatment strategies. In: *IOP Conference Series: Earth and Environmental Science*. IOP Publishing
41. Sakr GE et al (2016) Comparing deep learning and support vector machines for autonomous waste sorting. In: 2016 IEEE International Multidisciplinary Conference on Engineering Technology (IMCET)
42. Sarc R et al (2019) Digitalisation and intelligent robotics in value chain of circular economy oriented waste management—a review. *Waste Manage* 95:476–492
43. Sari GL et al (2021) Infectious waste management as the effects of Covid-19 pandemic in Indonesia. *Asian J Soc Sci Manage Technol* 3:62
44. Sreelakshmi K et al (2019) Capsule neural networks and visualization for segregation of plastic and non-plastic wastes. In: 2019 5th International Conference on Advanced Computing & Communication Systems (ICACCS)

45. Sreelakshmi K et al (2019) Capsule neural networks and visualization for segregation of plastic and non-plastic wastes. In 2019 5th international conference on advanced computing & communication systems (ICACCS). IEEE
46. Szegedy C et al (2016) Rethinking the inception architecture for computer vision. In: 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)
47. Tagle MG, Cilia-López VG (2021) The massive misuse of face mask as a risk to COVID-19 pandemic in Latin American: the case of Mexico
48. Tan M, Pang R, Le QV (2020) Efficientdet: scalable and efficient object detection. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition
49. Tirkolaee EB, Abbasian P, Weber G-W (2021) Sustainable fuzzy multi-trip location-routing problem for medical waste management during the COVID-19 outbreak. *Sci Total Environ* 756:143607
50. Torres FG, De-la-Torre GE (2021) Face mask waste generation and management during the COVID-19 pandemic: an overview and the Peruvian case. *Sci Total Environ* 786:147628
51. Tsai W-T (2021) Analysis of medical waste management and impact analysis of COVID-19 on its generation in Taiwan. *Waste Manage Res* 39(1_suppl):27–33
52. Vanapalli KR et al (2020) Challenges and strategies for effective plastic waste management during and post COVID-19 pandemic. *Sci Total Environ* 750:141514
53. Wang H et al (2020) Smartphone-based bulky waste classification using convolutional neural networks. *Multimed Tools Appl* 79(39):29411–29431
54. Wang Z, Li H, Zhang X (2019) Construction waste recycling robot for nails and screws: computer vision technology and neural network approach. *Autom Constr* 97:220–228
55. Yang J et al (2021) GarbageNet: a unified learning framework for robust garbage classification. *IEEE Trans Artif Intell*. <https://doi.org/10.1109/TAI.2021.3081055>
56. Yang M, Thung G (2016) Classification of trash for recyclability status. CS229 Project Report, vol 2016.
57. Yu J, Zhang W (2021) Face mask wearing detection algorithm based on improved YOLO-v4. *Sensors* 21(9):3263
58. Zamparas M et al (2019) Medical waste management and environmental assessment in the Rio University Hospital, Western Greece. *Sustain Chem Pharm* 13:100163
59. Zand AD, Heir AV (2020) Emerging challenges in urban waste management in Tehran, Iran during the COVID-19 pandemic. *Resour Conserv Recycl* 162:105051
60. Zhang Z (2018) Improved adam optimizer for deep neural networks. In: 2018 IEEE/ACM 26th International Symposium on Quality of Service (IWQoS). IEEE
61. Zhang P et al (2019) Urban street cleanliness assessment using mobile edge computing and deep learning. *IEEE Access* 7:63550–63563
62. Zhihong C et al (2017) A vision-based robotic grasping system using deep learning for garbage sorting. In: 2017 36th Chinese Control Conference (CCC). IEEE
63. Zoph B et al (2018) Learning transferable architectures for scalable image recognition. In 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition
64. Zou F et al (2019) A sufficient condition for convergences of adam and rmsprop. In: Proceedings of the IEEE conference on computer vision and pattern recognition

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Authors and Affiliations

Amirhossein Peyvandi¹ · **Babak Majidi**^{1,2} · **Soodeh Peyvandi**³ ·
Jagdish C. Patra⁴ · **Behzad Moshiri**^{5,6}

¹ Department of Computer Engineering, Faculty of Engineering, Khatam University, Tehran, Iran

² Emergency and Rapid Response Simulation (ADERSIM) Artificial Intelligence Group, Faculty of Liberal Arts & Professional Studies, York University, Toronto, Canada

³ Business Intelligence, University of Applied Sciences Upper Austria, Steyr, Austria

⁴ Faculty of Science, Engineering and Technology, Swinburne University of Technology, Melbourne, Australia

⁵ School of Electrical and Computer Engineering, College of Engineering, University of Tehran, Tehran, Iran

⁶ Department of Electrical and Computer Engineering, University of Waterloo, Waterloo, Canada