Urban Street Cleanliness Assessment Using Mobile Edge Computing and Deep Learning

Pengcheng Zhang, *Member*, *IEEE*, Qi Zhao, Jerry Gao, Wenrui Li and Jiamin Lu

Abstract—During the process of smart city construction, city managers always spend a lot of energy and money cleaning street garbage due to the random appearances of street garbage. Consequently, visual street cleanliness assessment is particularly important. However, existing assessment approaches have some clear disadvantages, such as the collection of street garbage information is not automated, and street cleanliness information is not real-time. To address these disadvantages, this paper proposes a novel urban street cleanliness assessment approach using mobile edge computing and deep learning. First, the high-resolution cameras installed on vehicles collect street images. Mobile edge servers are used to store and extract street image information temporarily. Second, these processed street data is transmitted to the cloud data center for analysis through city networks. At the same time, Faster R-CNN (Faster Region-Convolutional Neural Network) is used to identify street garbage categories and count the number of garbage. Finally, the results are incorporated into the street cleanliness calculation framework to ultimately visualize street cleanliness levels, which provides convenience for city managers to arrange clean-up personnel effectively. The overall approach is illustrated and visualized using the street images of Jiangning District in Nanjing, China. The practical application shows the feasibility and usability of the approach.

Index Terms—Smart cities; Street cleaning; Garbage detection; Deep learning; Mobile edge computing.

1 INTRODUCTION

A smart city [12] is an urban area that uses state-of-the-art technologies such as the Internet of Things (IoT) [30], Cloud computing [2] and other information technologies to manage and assess the resources and environment of a city in an efficient way [8]. The smart city concept integrates information and communication technology, and various physical devices connected to the network to optimize the efficiency of city operations and services [7], [17], [26]. However, due to the rapid development of a smart city, city managers are facing huge challenges in how to develop and maintain urban infrastructure. Street cleanliness [29] represents the spiritual outlook and humanistic atmosphere of a city. Keeping the streets clean is good for the development of modern cities. Currently, many major cities regard urban street cleanliness as one of the primary tasks of urban civilization [27]. If the urban street cleanliness level does not pass the pre-defined standard, it will have a serious effect on citizen's satisfaction and also affect the overall reputation of the city. The European city cleaning network summit also points out that cleaning streets timely is an effective way to improve city cleanliness [6].

- *P. Zhang, Q. Zhao and J. L are with College of Computer and Information, Hohai University, Nanjing, P.R.China E-mail: pchzhang@hhu.edu.cn; zqqizhao@163.com; jiamin.luu@hhu.edu.cn*
- *J. Gao is with Department of Computer Engineering, San Jose State University, San Jose, CA, USA E-mail: jerry.gao@sjsu.edu*
- *W. Li is with school of Information Engineering, Nanjing Xiao Zhuang University, Nanjing, Jiangsu, 211171, China E-mail: wenrui_li@163.com*

Manuscript received XXXX XXXX; revised XXXX, XXXX.

At present, the large number of streets make the amount of garbage on streets uncontrollable. Meanwhile, the process of garbage detection on streets is not automated and always requires human intervention at almost every level [4]. Citizens check the location of garbage manually and submit reports to city administrators, then city administrators arrange nearby city personnel to sweep garbage. Some cities even set up cameras at the crossroads of the streets to see if there is any garbage in the area. However, these manual solutions cannot grasp garbage cleanliness of all the streets of the city in time. For this reason, researchers [21], [23] around the world are studying automated approaches, using a cleaning vehicle with cameras to capture the streets regularly and collect street information, such as street pictures, geographical location, date and time. Besides, existing object detection algorithms are used to detect images in the remote cloud platform. Finally, the detection results are sent to the city managers for decision making.

Towards this research direction, this paper proposes a novel urban street cleanliness assessment model using mobile edge computing [24] and deep learning [20]. The high-resolution cameras installed on the vehicle collect street images. Meanwhile, the edge servers located at the edge of the network are used to store and process the street image information temporarily, and then these processed data is transferred to the remote cloud center through city network. Faster R-CNN (Faster Region-Convolutional Neural Network [28]) is used to identify street garbage categories and count the number of garbage. The results are sent to the street cleanliness level assessment model for evaluation. Finally, the approach visualizes street cleanliness level, which provides convenience for city managers to arrange cleaners in time.

In summary, the main contributions of this paper are described as follows:

- We describe a novel edge computing framework. There is an edge layer between cloud servers and mobile terminals. We configure edge servers (micro-data centers) to handle a part of services from mobile devices at the edge layer. It can also store data resources temporarily and transmit data resources in time.
- Faster R-CNN is used to identify street garbage categories and count the number of garbage. A multi-layer assessment model across different layers is used. The whole city is divided into 5 layers: *city*, *area*, *block*, *street*, *point*. Every layer will carry out street cleanliness calculation.
- We provide a public garbage data set [1]collected by ourselves, which can be used as a benchmark for evaluating street garbage detection and street cleaning. Furthermore, we use the data set to give a visual street cleaning map for Jiangning District, Nanjing, China. The application validates the feasibility and usability of the proposed approach. The results are useful for improving and optimizing city street cleanliness.

The rest of this paper is organized as follows: Existing work and their limitations are discussed in Section 2. Section 3 provides some preliminary knowledge including mobile edge computing, multi-layer assessment model, and deep network. Urban street garbage detection and cleanliness assessment approach is provided in Section 4. In Section 5, we use street images collected from Jiangning District to validate our approach. Finally, Section 6 concludes the paper and looks into future work.

2 Related work

2.1 Smart Cities

Smart city construction has become the focus of the whole society. Smart cities use intelligent methods to sense and handle urban activities through the Internet of Things, cloud computing and other technologies, which can improve the quality of service in all aspects of society and economy [1], [3], [14]. Meanwhile, smart cities can also achieve the purpose of reducing costs and resource consumption. Currently, many scholars in the world have done many researches related to smart cities. Zygiaris et al. [33] proposed a planning framework called "Smart City Reference Model". Urban planners can use the framework to define the smart city concept and apply an urban layout to green, interconnected, open, integrated, smart, and innovative concepts. The framework provides an idea for realizing sustainable development of a smart city. The recent practical application is to analyze smart city planning in big cities such as Barcelona, Amsterdam, and Edinburgh. Hefnawy [13] et al. combined a smart city and life cycle concept to create a suitable information and knowledge sharing platform in a smart city. It aims to solve the problem of unreasonable arrangement, lacking planning and internal coordination of large activities in the city, which can achieve the goal of organizational consistency and efficiency.

In addition, Large companies also attempt to put into the research of the smart city. China Telecommunication carried out the development plan of smart city, focusing on 12 theme applications including smart community, smart transportation, smart energy, smart medical services and etc. IBM [10] launched the Watson "Big Data and Analysis Platform" to help solve smart city problems such as smart transportation and air pollution. Microsoft [10] launched the "Future City" plan to solve challenges such as environmental deterioration and traffic congestion by acquiring, integrating and analyzing multiple heterogeneous big data in the city.

However, to the best of our knowledge, there is no specific research topic on urban cleanliness for the construction of a smart city.

2.2 Street Garbage Detection and Street Garbage Detection

Mittal [25] et al. launched a street garbage project that aims to segment a pile of garbage roughly in the images. They label these images and divide two parts of images that contain garbage or do not contain garbage, then they use CNN (Convolutional Neural Network) to segment the area containing garbage in the image. Besides, they use the Bing Image Search API to create their data set and get an accuracy of 87.69%, a sensitivity of 83.96% and a specificity of 90.06%. Their method focuses on the segmentation of a pile of garbage, but there are many errors in segmentation judgment and they do not provide details of the garbage type. Rad et al. [27] proposed a fully automated computer vision application based on garbage quantification, They collect different types of garbage images from streets and sidewalks through a data acquisition system established on the top of a vehicle. Then they use classification detection algorithm OverFeat-GoogLeNet which is based on deep CNN to train different types of garbage that they label, and finally, they can detect the garbage that appears on the street accurately. However, at present, they are only able to detect street garbage, and they have not carried out an urban street cleanliness assessment.

Also, researchers are thinking about how to use technology to achieve urban cleaning and urban street cleanliness assessment. Borozdukhin [5] et al. proposed a method to solve the optimization of garbage disposal in big cities. The method searches for the time-optimized dynamic route for garbage collection trucks by establishing a mathematical model of dynamic optimal paths, which can make the garbage collection trucks spend shorter time from the garbage collection area to the landfill area. However, the system only considers the route selection of garbage collection trucks and it does not consider the urban cleanliness assessment. Clean Street LA [19] is an initiative by the London city Mayor and the system uses ESRI and GIS tool to map and plot the street cleanliness status block by block. Multiple layers and grids are created to reflect different parts of the city. Cleanliness information on the streets with a cleanliness score is visualized on a map. This information is used to decide on the area that requires

1. https://pan.baidu.com/s/1aVQ8lLA4AmRBF1Sga_etWg

attention. However, the limitation of the system is that the monitoring is limited to garbage bins and cannot be extended to monitor the streets. In [23], a mobile app was developed to evaluate street cleanliness and waste collection. This app is based on a plan of indicators that can be used to evaluate the street cleanliness and waste collection service of Santander municipality. Specific methodologies for calculating and evaluating 59 indicators have been developed to obtain information regarding the status of the different elements of the service. Pearson correlation coefficient results suggest that an inverse relationship between the street cleanliness index values and the frequency street cleanliness services/population density ratio exists.

In short, although researchers are concerned about street cleanliness, they have not yet applied mobile edge computing, cloud computing, and deep learning to assess street cleanliness in time.

3 Preliminaries

3.1 Mobile Edge Computing

With the rapid construction of smart cities, the Internet generates a large amount of data. Traditional cloud computing requires that data must be transmitted to the cloud center for centralized processing. Remote cloud is a smart brain for processing big data [31]. Since the cloud center is usually far away from end users, it is largely unable to provide low latency. In order to solve this problem, mobile edge computing has been proposed to deploy computing resources to devices close to the terminal. The European Telecommunications Standards Institute (ETSI) [9] defines mobile edge computing (MEC) as a distributed mobile cloud computing (MCC) system. The computing resources are close to mobile devices, and functions such as computing, storage, and processing are added to the wireless network side. In fact, mobile edge computing is based on cloud computing. It only calculates a small part of service. It is especially important for big data analysis. For example, when a user uploads a video or makes a comment, he/she can send it to a remote server through an edge virtual server. The edge virtual server can extract the video content and estimate the possibility that other people want to watch the video. If the probability is high, the edge server will cache this video locally so that anyone interested in this video can get the video directly from its cache instead of receiving it from a remote server, which saves transmission resources and reduces latency. In this paper, we use mobile edge computing to process street images in advance and filter out pictures that meet our needs, which has a good effect on recognition efficiency.

3.2 Multi-level Assessment Model

To measure the cleanliness of the urban streets, our street cleanliness assessment approach provides a multi-level assessment model across different layers. This model can be divided into five layers as shown in Fig. 1. Layer 1 is the first layer, it is defined as the city area and sets the scope of assessment. Layer 1 covers all the streets in the city. Layer 2 is the second layer where a city is divided into multiple areas and each region is an administrative area. Layer 3

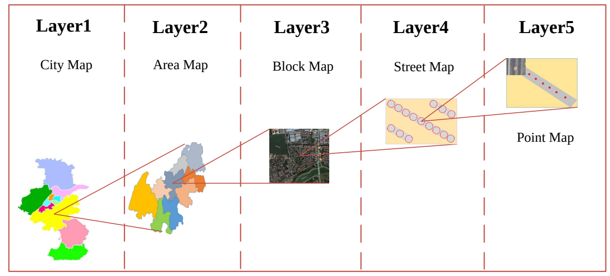


Fig. 1. Multi-level Assessment Model based on Nanjing

is the third layer where each area is divided into multiple blocks according to the sub-administrative area. Each block is uniquely identified by a combination on administrative area and block name. Layer 4 is the fourth Layer where each block has several streets. Layer 5 is the bottom layer where each street has several data collection points.

3.3 Deep Network

Deep learning originates in artificial neural networks. By establishing multiple hidden layers and training large amounts of data, useful features can be learned to achieve the expected classification effect. In recent years, deep learning has become a hot topic in the field of Object Detection. Girshick et al. [28] designed a deep learning object detection algorithm called Faster R-CNN based on region proposal. The algorithm has two main modules: the Region Proposal Network (RPN) proposal box extraction module and the Fast R-CNN detector module [11]. RPN is a fully convolutional neural network [22]. Its function is to find the possible object proposals in the map and extract the proposal box. Fast R-CNN is a proposal detector based on RPN extraction and it identifies the object of the proposal box. RPN shares the same convolutional layers by using a convolutional neural network based on object detection and a convolutional neural network that generates a suggestion window.

- The image is input to the convolutional neural network, and spread to the shared convolutional layer to get the feature map;
- The feature map extracted by the shared convolutional layer generates a suggestion window through RPN network, and gives region suggestions and region scores;
- The feature map of the first step is input to the pooling layer in Fast R-CNN to extract area features. Combined with region suggestions and region scores, classification probabilities and bounding box regression are trained, the classification scores of the region are output, and the results are finally tested.

Faster R-CNN is considered as one of the most precise image detection approaches. It has high detection accuracy and speed. Consequently, the street garbage detection approach in this paper adopts Faster R-CNN (Regional-Convolutional Neural Network) as the underlying model to detect the type and quantity of street garbage.

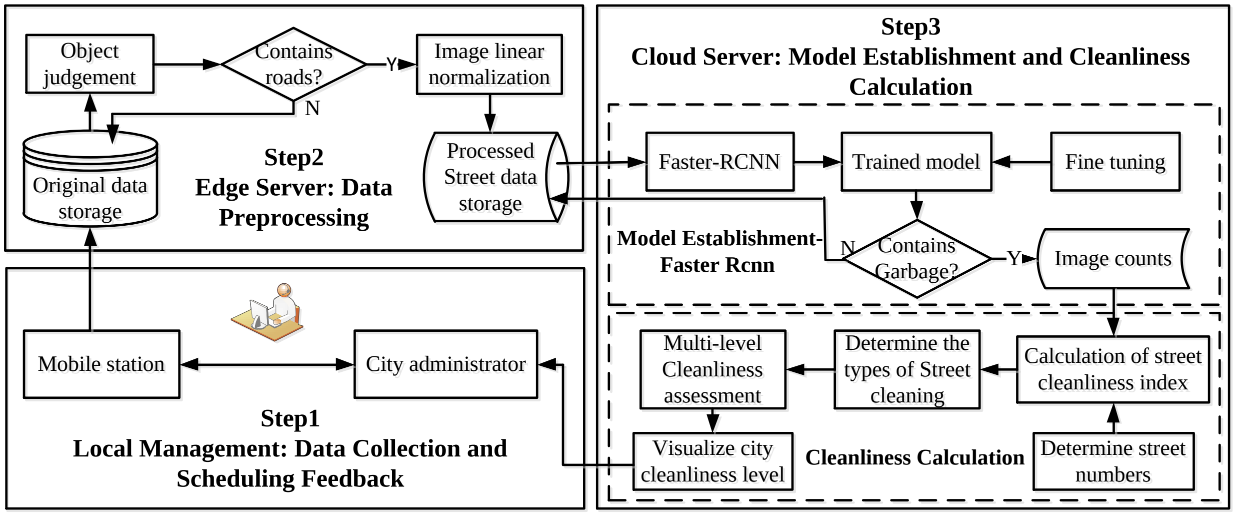


Fig. 2. Urban street garbage detection and cleanliness assessment

4 Urban Street Garbage Detection and Cleanliness Assessment Approach

In this section, we first give a general overview of our approach in Section 4.1. Second, data collection and mobile edge processing are introduced in Section 4.2. Third, the deep learning algorithm description of the approach is presented in Section 4.3. Finally, Section 4.4 offers multi-level cleanliness assessment of our approach.

4.1 Approach Overview

Edge computing can reduce latency and resources. Compared with traditional cloud computing, the main difference is that some services are processed on the edge in advance when a large amount of data is generated. R-CNN is also widely used in image recognition. Based on the above work, we design a novel urban street garbage detection and cleanliness assessment approach. The approach combines mobile edge computing and R-CNN to detect urban street garbage. Based on the above detection results, we use the street cleaning standard to calculate street cleanliness. Fig. 2 shows the main process of our assessment approach. The approach is mainly composed of three parts, as described in the following:

- The first step is data collection and scheduling feedback in the local management. The city administrators control the mobile station to collect the street garbage image data and respond to the level of street cleanliness presented by cloud center in real time. Then municipal cleaning personnel is arranged nearby.
- The second step is called data preprocessing. During this step, we use the edge server to store the image data captured by the mobile station temporarily and carry out road judgment of the images from the mobile station in advance. Then, the edge server filters out the images containing road areas. We use linear normalization to get the same size images and these images are sent to the cloud center for garbage detection.
- The third step is the model establishment and cleanliness calculation. During this step, the cloud server provides an object detection algorithm. Then a model is trained by selecting appropriate parameters and iterations to detect garbage on the street. In the garbage detection stage, we design a counting function to count the quantities of garbage detected. Finally, based on the results of the above detection, street cleanliness level is calculated with respect to different levels.

4.2 Data Collection and Mobile Edge Processing

4.2.1 Data Collection

During the data collection stage, the main task is to collect garbage and street images needed by the assessment approach. When a vehicle equipped with a high-resolution camera is in a city street environment, the information collected includes mainly two parts: *street image information* and *local management information*. For *street image information*, the cleaning vehicle equipped with a high-resolution camera is shot on each street according to the administrator's assignment. The distance between adjacent shooting points is set by the administrator, and the cleaning vehicle takes pictures at each shooting point according to the four directions including *left*, *front*, *right* and *back*. The shooting range is $150 - 300m^2$. For mobile stations, the following rules are set: 1) fixed image resolution; 2) vehicle speed is approximately 25 kilometers per hour; 3) shooting points are a fixed distance; 4) there are 4 pictures in each shooting points. For *local management information*, the mobile station needs to report the location to the city manager regularly. The administrator responds in time and arranges cleaning staff to clean.

4.2.2 Mobile Edge Processing

We use edge servers to complete two tasks. The first task is to improve the performance of the entire system. During this stage, when object detection is performed, image data collected is first input into the CNN network and then the

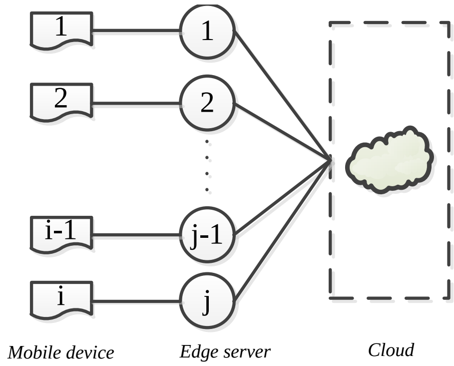


Fig. 3. Street image data preprocessing in the edge environment

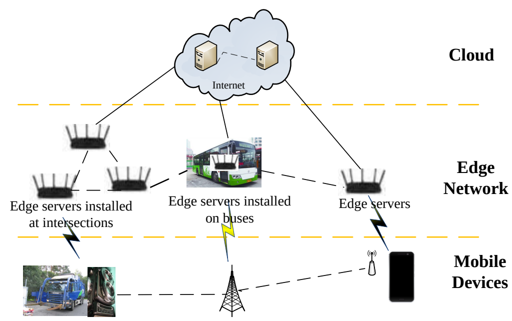


Fig. 4. Street image data acquisition architecture based on mobile edge computing

size of pictures is modified to the suitable size. We believe that if image data is preprocessed in the edge server, it can reduce the overall time of the entire system. We design an algorithm to modify the size of images automatically in the edge server when image data is transmitted to the edge server. Fig. 3 shows street image data preprocessing in mobile edge environment. Edge server j receives street image data from mobile device i. We define t_{i-j} as the time street image data transferred from mobile device i to edge server j. t_{j-c} is the time from edge server j to cloud c. t_{edge} is described as the time picture processed in the edge server. Total time of mobile edge processing is calculated using the following formulae:

$$T = t_{i-j} + t_{edge} + t_{j-c} \tag{1}$$

The second task is to filter out valuable data in advance through edge servers. Sometimes, the data collected by a mobile station may be useless. For example, there are some problematic pictures that include house, car or camera shooting angle causes the street blocked in 4 pictures collected at a certain shooting site. For the whole garbage detection system, pictures without full street images are obviously useless. In order to reduce the consumption of resource and time, we design the edge data processing layer. When the layer receives street images from mobile stations, street image information is temporarily saved for road detection. Image data that contains city roads is filtered and passed to the cloud center for street cleanliness assessment. The idea is actually “computation migration” and it is the key idea of edge computing. Fig. 4 shows main architecture based on mobile edge computing, and it includes the following main parts:

Mobile station. There is a specific garbage collection vehicle in the city. We install cameras with high resolution, high pixel and network transmission capabilities on the top of the garbage collection vehicle. The camera faces the ground and covers the front 50 meters. The garbage collection vehicle takes photos regularly in city streets every day according to a specific line and these data is transmitted into edge servers in time. At the same time, urban citizens can also act as garbage collection vehicles. They can collect street garbage data with their own mobile devices and transmit the collected data to edge servers.

Edge server. Edge server is at the edge of the network. It directly connects to nearby mobile devices through a wireless data link to handle a portion of service requested from mobile devices. It also has the ability to temporarily store data from mobile devices. These edge servers are placed on street intersections and highly mobile buses, which can work better within the mobile device network.

Cloud. This layer is used to create training models and perform street garbage detection tasks. Meanwhile, the cloud server presents urban street cleanliness level in time and feedback relevant information to city managers.

4.3 Image Detection Using Neural Network (R-CNN)

In Section 3.3, we have already introduced that our street garbage detection is based on the Faster R-CNN algorithm. Below, we describe the detection algorithm in detail from three parts: *network design*, *network training*, and *street garbage detection*.

4.3.1 Network Structure

The main task of this part is to select and design the network structure. We first input any size pictures to the CNN network to prepare for getting feature map. The CNN network we choose is the ZF-Net proposed by Matthew D.Zeiler [32]. The input layer is a 224×224 3-channel RGB image, and the first layer contains 96 convolution kernels. In order to avoid the first layer convolution kernels mixing high-frequency, low-frequency information and there is no intermediate frequency information. The filter size is set to 7×7 in the first layer. Then the maximum pooling operation is performed, the stride is set to 2. The normalized operations are compared, and 96 different feature templates produced are 55×55 size. Layer 2, 3, 4 and 5 have similar operations. The layer outputs 256 feature maps of size 6×6. Layer 6 and layer 7 are fully connection layers. Finally, the layer 5’s result sample is input to the classifier and the bounding box regression. The classifier gives the category of the region proposal, and the bounding box regression gives the position information of the region proposal.

4.3.2 Network Training

After we design the garbage detection network, it is obviously necessary to train the network to learn the characteristics of the street garbage. The specific network training process for our application is divided into four steps:

- RPN pre-training is performed. The RPN is initialized by using ImageNet [18] to train network parameters. The Gaussian distribution with a standard

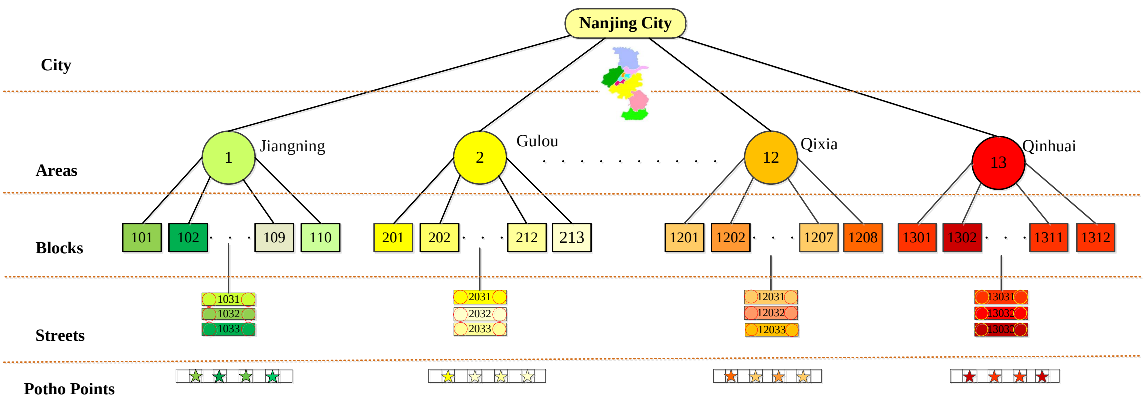


Fig. 5. Multi-level Assessment Model

deviation of 0.01 and a mean of 0 are used to initialize the additional layers. Then the end-to-end fine-tuning task is used for the region proposal.

- Fast R-CNN pre-training is performed and using the proposals obtained by the step 1 to perform end-to-end fine-tune training of Fast R-CNN, and the ImageNet model is used to initialize network parameters.
- Re-initialize the RPN training with the network fine-tuned by Fast R-CNN in step 2 and fix the shared convolutional layers. That is, the learning rate is set to 0.
- The shared convolutional layer is fixed in step 3, and using the region proposal obtained in step 3 to fine-tune the fully connection layer of Fast R-CNN.

4.3.3 Street Garbage Detection

In this stage, we use the trained model to detect garbage on the street. These street images are input to the CNN, and then the CNN reflects the features of images to the feature map by calculating. Each proposal region network can calculate a proposal region corresponding to each other. The input images generate 300 region proposal boxes. Then the classification layer and the regression layer display the region proposal box where the garbage is located. Here, we set a counting function. As shown in formulae 2, every time, a region proposal box is generated, and the bounding box is automatically counted once. That is, the value of the count function is incremented by one and finally, we count the categories and quantities that are detected in the region proposal box.

$$C(f;D) = \sum_{i=1}^{m} \mathbb{I}(f(x_i) = y_i) \tag{2}$$

C is a counting function for generating a proposal box, that is, the number of garbage detected. f is a result function detected by the garbage model, D is a test sample set, x is a test sample, and y is a real garbage label.

4.4 Multi-level Cleanliness Assessment Model

Based on the layered model, the cleanliness values of the city streets are evaluated in five levels below.

Definition 1. *City*={*City*, *Area*, *Block*, *Street*, *Point*} where

- *City* is the geographical area of a city. A city corresponds to one map.
- *Area* is an internal part of a city. Usually, a city map is divided into many areas.
- *Block* is a part of an area. Usually, an area is divided into many blocks.
- *Street* is composed of roads divided into several blocks in the city. Each street belongs to one corresponding block, and each street has a number of grid points.
- *Point* is a collection of shooting points on a street, and it is the most basic scope of assessment.

Fig. 5 shows the hierarchical view of City Nanjing. From the figure, we can see that Nanjing is divided into 13 different areas. Every area has a corresponding number and these numbers represent how many areas divided there are. For example, there are 13 administrative regions in Nanjing. Consequently, this city has 13 numbers. Jiangning District is numbered 1 and Qinhuai District is numbered 13. Similarly, every area is divided into different blocks. The number of sub-administrative regions determines the number of blocks. For example, there are 10 sub-administrative regions in Jiangning District. Consequently, this area has 10 corresponding numbers. Following this way, other layers are also numbered similarly. Color represents the cleanliness level of every layer. Area Value (AV) below is indicated with an average of results from each block within the area.

Definition 2. *Grid*. Formally, a grid of a city is hierarchically defined as a Quad (*PicturePoint*, *GridPoint*, *INode*, *SNode*), where

- *PicturePoint*. Multiple Images are captured in each direction (Front, Back, Left, Right) at some point on the street and sent to Cloud along with location data.
- *GridPoint*. GridPoint represents logical radius to assess the urban street. It can have one or more Picture Points collectively produce the cleanliness level across the street. Normally, a GridPoint has two picture points and each picture point captures images in four direction (Front, Back, Left, Right).

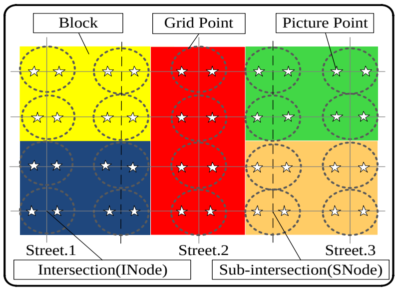


Fig. 6. Grid Architecture

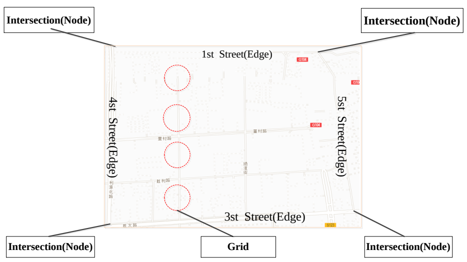


Fig. 7. Graph mapping to block

- *INode*. It is used to represent intersections between different streets. There are several picture points between nodes.
- *SNode*. Often INode is not enough to divide two streets. SNode representing sub intersection is used to divide the area between two streets into more Grid Points.

Fig. 6 shows a grid architecture. In this figure, all blocks are divided into several *Grid Point* and *INodes*. Then *INodes* are also divided into subsections with *SNodes*. Each grid point has at most two picture points and the radius of each grid point is $50m$.

Directed Graph (DG) is used to represent the entire city and each road has a direction. Fig. 7 shows the correlation of graph block with actual city block. The definition of Graph is shown in the following:

Definition 3. A graph is defined as $G = (N, E)$, where N is the set of nodes and E is the set of edges. Each edge would have a starting node and ending node represented as below.

$$e = (n_{start}, n_{end}) \tag{3}$$

Each grid point has a value from the machine learning system represented by PV (shown as Point Shooting in Fig. 8) between n_{start} andn_{end}. Once PV is calculated, the grid point value (GV) (shown as circle in Fig. 8) using:

$$GV(n_{start} - n_{end}) = PV \tag{4}$$

where GV is a grid point value, PV is a photo point value in a grid point.

In general, the urban cleanliness assessment model is divided into multiple area assessments, further divided into block assessments, and finally focuses on the assessment of each street. Below we define the detail assessment approach of each level.

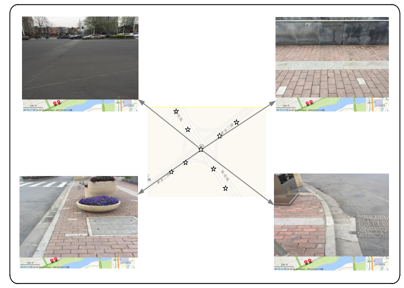


Fig. 8. Point shooting

4.4.1 *Street Layer Assessment*

The street layer assessment is the basis of the level assessment model and it plays a key role in the entire assessment system. At this layer, each street has several photo points, and at least four photos are taken at each point. The distance between the two photo points is determined by City administrators. Fig. 8 shows a street layer assessment. Each photo point is shot in Front, Back, Left, and Right directions. If there is garbage in the photos, street layer cleanliness can be assessed through these pictures. As we can see from Fig. 6, a street passes through multiple grid points. The assessment value of each street is the average value of grid points. Here, the assessment value of each street is obtained by formulae 5.

$$SV = \frac{1}{n}\sum_{i=1}^{n} GV_i \tag{5}$$

where SV (Street Value) is the assessment value of a street, GV is the assessment value of each grid, and n is the total number of grids.

4.4.2 *Block Layer Assessment*

The Block layer assessment is based on street layer assessment. There are many streets in a block. The assessment value of each block is the average value of all street assessments. Here, the assessment value of each block is obtained by formulae 6.

$$BV = \frac{1}{n}\sum_{i=1}^{n} SV_i \tag{6}$$

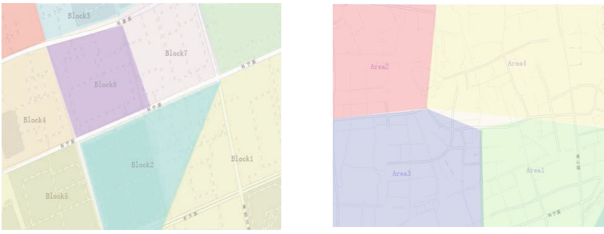
where BV (Block Value) is the assessment value of a block, SV is the assessment value of each street, and n is the total number of streets in the block. Fig. 9 (a) shows a block layer assessment where the different colors in the block represent different assessment values of each block.

4.4.3 *Area Layer Assessment*

After the block layer analysis, an area layer assessment is performed. There are multiple blocks in one area. Here, the assessment value of each area is obtained by formulae 7.

$$AV = \frac{1}{n}\sum_{i=1}^{n} BV_i \tag{7}$$

where AV (Area Value) is the assessment value of an area, BV is the assessment value of each block, and n is the total number of blocks in the area. Fig. 9 (b) shows an area layer assessment where the different colors in the area represent different assessment values of each area.



(a) Block level assessment (b) Area level assessment

Fig. 9. Level assessment

TABLE 1
Garbage classification and weighting values

Litter classification		Weighting values
Inorganic	Small	1
	Medium size	2
	Large	4
Organic	Small	2
	Medium size	4
	Large	6
Trash	with litter	3
	Empty	0
Tree basins	Uncleaned	6
	Clean	0
Uncollected sweeping waste		6
Tree leaves		1
Animal droppings		3
Sticky residue on the pavement		2

4.4.4 City Layer Assessment

A city has multiple areas. Based on the above three levels of assessment, the assessment value of a city is obtained by formulae 8.

$$CV = \frac{1}{n}\sum_{i=1}^{n} AV_i \tag{8}$$

where CV (City Value) is the assessment value of a city, AV is the assessment value of each region in the city, and n is the total number of areas in the city. Finally, through the level assessment, the city cleanliness can be measured comprehensively and accurately.

4.4.5 Cleanliness Assessment Calculation

Jang et al. [15] pointed out that urban street garbage mainly includes *plastic packaging*, *leaves*, *peels*, *cans*, *plastic bottles*, *animal hair*, and etc. They are usually scattered in any corner of the street. The key factor affecting the city street cleanliness is street garbage. Consequently, researchers have been discussing how to measure the cleanliness level of urban streets according to the amount of street garbage. We select a random sampling method and take some streets from all streets of the city as samples, which can eliminate subjectivity and minimize the impact on accuracy results. The goal of the sampling design is to determine the minimum number of streets surveyed, formulae 9 represents the minimum street sample size.

$$n = \frac{k^2 \cdot p \cdot q \cdot N}{e^2 \cdot (N-1) + k^2 \cdot p \cdot q} \tag{9}$$

where n is the minimum sample size and k is the sampling interval. To ensure a 95% confidence interval, we set k as 1.96, and p represents the probability that an event will take place. q is equal to $1-p$, $p = q = 0.5$, N is the total number of city streets, and e is the estimation error, where $e = 0.1$.

TABLE 2
Values of correction factor λ

Type of pavement	Conservation	Wind and rain Low	Wind and rain Medium	Wind and rain High
Asphalt	Good	1	0.8	0.6
	Bad	0.9	0.7	0.5
Sand		0.8	0.6	0.5

TABLE 3
Street cleanliness level classification

Cleanliness Index	Cleanliness level	Level Color
$SV < 70$	Very high	
$70 \leqslant SV < 100$	High	
$100 \leqslant SV < 150$	Medium	
$150 \leqslant SV < 200$	Low	
$SV \geqslant 200$	Very low	

The paper in [23] provides a method for calculating the urban street cleanliness. This method was determined by the Spanish Federation of Municipalities and Provinces to be able to measure different street cleaning services. Formulae 10 is the calculation method of the urban street cleanliness index. Once we obtain the cleanliness index value, each street can be classified according to this value(SV) calculated by Formulae 4, 5. As shown in Table 3, the lower the SV value, the cleaner the streets.

$$PV = \frac{\lambda \times C}{n \times S} \times 100 \tag{10}$$

where S is the observation area, it is the angle of view taken by the camera installed on the top of the garbage collection vehicle and it is set to $S = 150m^2$. and n and λ are correction factors. There are many factors that affect the cleanliness index, such as weather conditions, the type of street pavement. C is the weighted quantity of litter in the street, and the classification in Table 1 is considered. The quantity of litter of each type was multiplied by a weighting coefficient that depends on the litter classification, in which inorganic and organic litters correspond to three subcategories: *small*, *medium* and *large*. Here, the subcategories are measured by quantities. *small* represents that the number of garbage in the image are between 1 and 4, *medium* represents that the number of garbage in the image between 5 and 10, *large* represents that the number of garbage in the image is more than 10. Table 2 shows the values of the correction factor λ in various situations. For example, when the street is an asphalt road, the road surface is flat and the weather is fine, the correction factor λ is 1. n represents the change in the quantity of garbage in special circumstances, and it is often between 1 and 2. For example, the existence of a bus stop increases the quantity of garbage. This factor can have a value between 1 and 2. The shooting scene does not include factors such as at the bus stop, and n is set to 1.

5 EXPERIMENTAL EVALUATION

5.1 Research Questions

In this section, we conduct a set of experiments to validate the proposed approach using street image data collected by ourselves. The experiments are designed to investigate the following three main research questions:

RQ1: Is the time reduced after adding mobile edge layer?

RQ2: Is Faster R-CNN able to detect street garbage?

RQ3: Is the street level cleanliness assessment model useful?

Our approach is used to provide street cleanliness real-time information for city managers. Consequently, for real-time, we design *RQ1* to investigate whether will the processing time of our approach be reduced after the mobile edge layer is added, compared to cloud computing. We design *RQ2* to investigate whether this algorithm can detect urban street garbage. *RQ3* is designed to investigate whether street level cleanliness assessment model can provide visual street cleanliness information for city managers

5.2 Experimental Design

5.2.1 Experimental Environment

The experimental environment is a Lenovo PC with Inter Core i5-7500 CPU 16G RAM. The operating system is Ubuntu16.04. Our experiment selects the popular deep learning framework CAFFE [16] and builds a Faster R-CNN. The ConvNet is trained on a single NVIDIA GeForce GTX 1050ti GPU with 4 GB of memory.

5.2.2 Data Set

The experimental data include *garbage model training data* and *street detection data*. First, to train a garbage model, we take garbage pictures from lots of streets in Nanjing to make training data set. Due to the diversity of urban street garbage types, we classify common street garbage into the following nine categories: *waste paper*, *plastic bag*, *plastic bottle*, *peel*, *cigarette butts*, *waste cloth*, *cigarette case*, *leaves*, and *cans*, and these basically contain common garbage types in the streets. Next, we label and classify every image with garbage according to the format of the VOC2007 data set. A total of 681 image data are collected, and the size of an image is 420×400 pixels. We then divide collected data into 3 parts, 321 images are the training set, 260 images are the test set, and 100 images are used as the verification set.

According to the multi-level assessment model, the entire Nanjing City is the first layer. Then according to the administrative division, we divide the entire Nanjing into 13 areas. We randomly select Jiangning District as the second layer of the study from 13 administrative areas. There are 10 administrative streets in Jiangning District. Based on these, we divide Jiangning District into 10 blocks. The 10 administrative blocks are the third layer of this study. The fourth layer is a number of streets in the block. We set up a shooting point every 50 meters on the street and a grid point is composed of two shooting points. The street is captured in four directions: *front*, *back*, *left* and *right* at every shooting point. The shooting angle is about $150m^2$. Therefore, we randomly stipulate that we set a sample of 1 km for each street, That is, garbage collection vehicles takes 80 Street images on each street. We learn that there are approximate 3,875 streets in the Jiangning District from the Jiangning street Management Office. Based on formulae 9, we conclude that minimum number of street sample is 96. Finally, we choose 100 streets as our experimental sample data.

The street detection data mainly comes from 9 administrative regions in Jiangning District. We collect about 8,000 street images. Since city staff may clean street garbage every day in the morning, this may impact the results of street cleanliness calculation. To avoid this effect, we collect street images from 11 am to 16 pm every day. The experimental data can be obtained through Baidu Cloud. [2]

TABLE 4
The time consumption comparison between the mobile edge and the cloud environment

Counts of pictures	Environment	Processing time
8000	Cloud server	379.206s
	Mobile edge server	3.686s

TABLE 5
Quantities of pictures preprocessed at the edge

Street name	Quantities of pictures	Quantities of pictures containing road
Shangyuan Street	80	72
Danyang Street	80	77
Lize Street	80	68
Yunlong Street	80	80
Focheng East Street	80	80
Focheng West Street	80	71
Baoshuiyuan Street	80	74
Xincheng Street	80	77
Guanghua Street	80	72
......	......	......
Liyuan North Street	80	76
Liyuan South Street	80	74
Qiancao Street	80	80
Beiyan Street	80	73
Nanwanying Street	80	80
Shuanglong Street	80	79
Jiyin Street	80	75
Tongxia Street	80	65
Wuge Street	80	78
Total	8000	7503

5.3 Experimental Results

5.3.1 Mobile Edge Processing

To address **RQ1**, we select NAS network memory as the edge server of the edge data processing layer. NAS network memory has the advantages of faster response speed, higher data bandwidth, high sharing, and support for Internet connectivity. First, in our experiments, we place 100 edge servers in the 100 streets and each street has an edge server. Every edge server process 80 street pictures, it's like distributed processing. Table 4 shows the time consumption comparison between the mobile edge and the cloud environment. When 8000 street pictures are preprocessed in the edge environment, the consumed time is only 3.686s. However, when these street pictures are transmitted into the remote Cloud server, the cloud server will use ZF network to process the pictures. This process takes about

2. https://pan.baidu.com/s/1aVQ8lLA4AmRBF1Sga_etWg

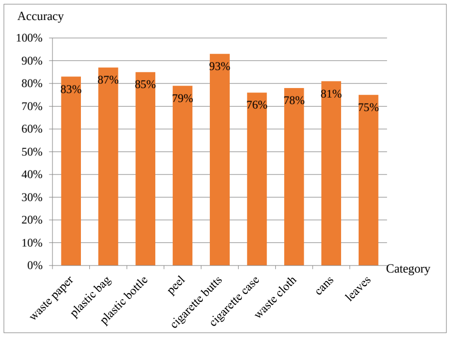


Fig. 10. Results of all kinds of garbage detection accuracy

Fig. 12. Garbage image detection result

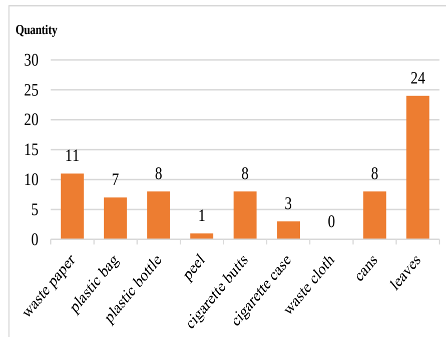


Fig. 13. Garbage classification and testing counts based on Fochengxi road

379.206s. Finally, we find that compared to the traditional cloud server processing 8000 pictures centrally, 100 edge servers processing 8000 pictures simultaneously take less time (3.686s vs 379.206s). Second, the city administrators can check the pictures preprocessed by the NAS network memory through the network, filter out the pictures that the street is obscured. Table 5 shows the results of street pictures preprocessed. We can find that 80 pictures are taken on each street and a total of 7503 pictures containing road are found at 100 streets by manual screening.

5.3.2 Image Detection Results

To address **RQ2**, we optimize the model iterative through three stages of training and select ZF network to extract garbage image features. Since our training data is small, and in order to avoid over-fitting, we use ImageNet pre-trained model weights as the initial input value of the garbage detection model. After 50000 times of iterations, the final model parameters are described as follows. Training impulse is 0.9, weight attenuation parameter is 0.0005 and training rate is 0.001.

Fig. 10 shows the different categories of garbage detection accuracy. The Y axis shows the percent of correctly recognized images from tested images. The X axis shows the categories of garbage. From the figures, we can see that the accuracy of *plastic bag*, *cigarette butts*, *waste paper*, *plastic bottle* and *cans* can reach 81% and 93%. Because these objects have a regular shape and their texture is clear, which can be easily identified. Other categories have lower accuracy because of their limited properties and relatively fewer image sets. The test result shows that the model can reach 82% detection accuracy. Besides, to describe the performance of garbage model trained in detail, Fig. 11 shows the Precision and Recall of different garbage categories. Consequently, our model meets the basic requirements of street garbage detection.

We send road images collected on every street to the Faster R-CNN classifier and use the trained garbage model to detect the garbage on the street. As shown in Fig. 12, the garbage objects identified are marked in rectangular boxes and the value on the rectangle represents a probability of determining whether the object is a garbage category. We find that the recognition rate of *leaves* is 0.998, the recognition rate of *cigarette butt* is 0.995, and the recognition rate of *peel* is 0.988. For a better calculation of the street cleanliness, when a rectangular box is generated, we automatically count the number of rectangular boxes. Finally, the number of rectangular boxes is the amount of garbage detected by the classifier. Fig. 13 shows the results of garbage detection based on the 1km street of Focheng West Road in Jiangning District, Nanjing, China.

5.3.3 Cleanliness Assessment and Exhibition

To address **RQ3**, we classify *waste paper*, *plastic bag*, *plastic bottle*, *cigarette case*, and *cans* into inorganic garbage category, and *peel*, *cigarette butts*, and *waste cloth* into organic garbage category. According to the garbage categories and weights in the Table, we begin to look for weighting factors that correspond to the garbage categories detected by the classifier.

Through the calculation, we can obtain the total weighting of garbage in every street. Table 6 shows the weighted amount of garbage in Focheng West Road and we can see 79 street garbage in this road. At this time, according to the specific environment of Focheng west Road, we set $C = 87, S = 150, \lambda = 1, n = 1$. The final cleanliness value of the Focheng west Road is 58 (shown as in formulae 11). According to Table 3, the cleanliness level of the Focheng west Road is the highest at this time, which means that this street is very clean.

$$SV = \frac{(1 \times 87)}{(1 \times 50)} = 58 \tag{11}$$

In a similar way, we calculate the cleanliness value of 100 streets in Jiangning District one by one based on the image of the city streets collected. Then, we provide a visual street cleanliness road map. This map can shows some information of street cleanliness clearly. The results

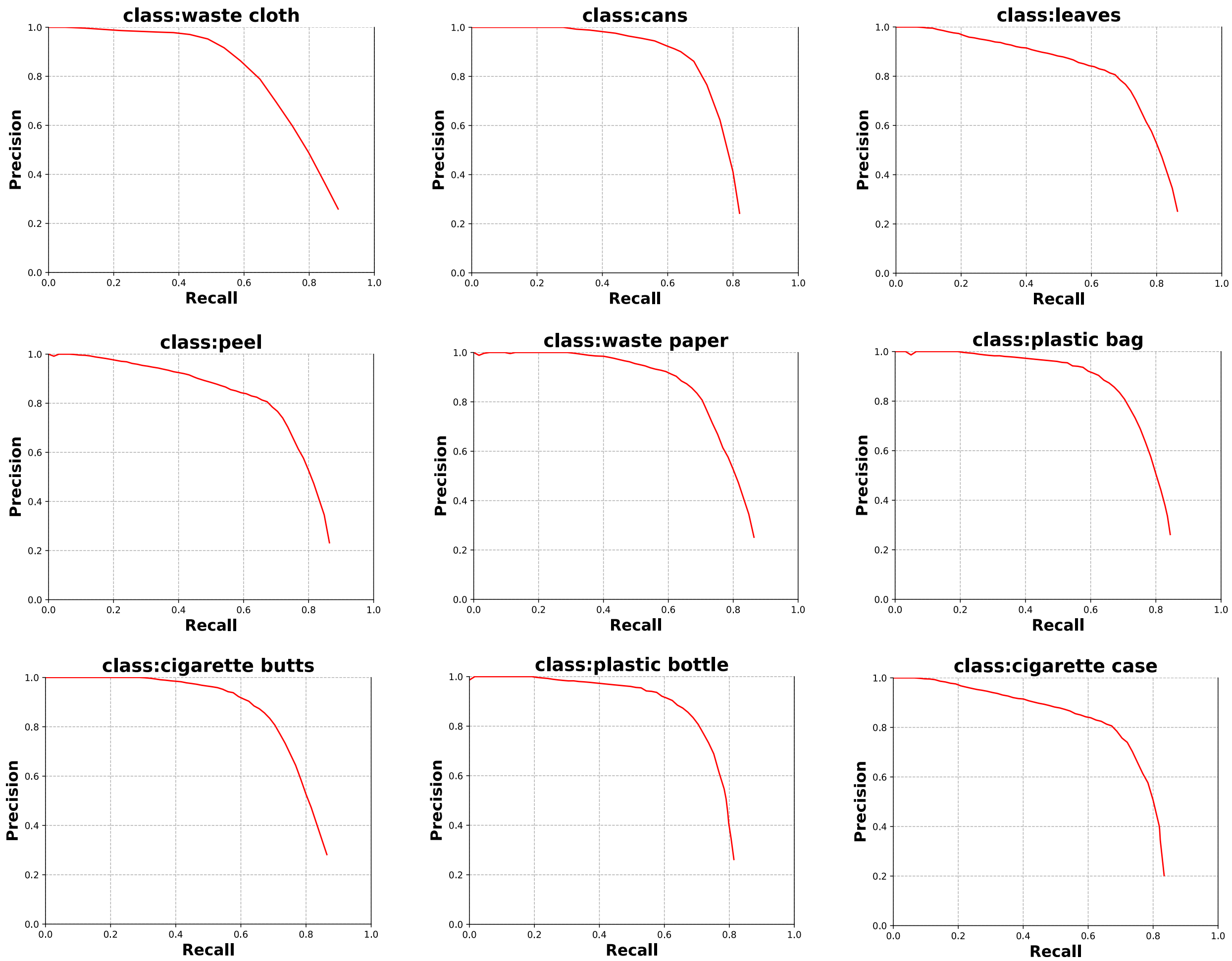


Fig. 11. Precision and Recall of different garbage categories: (a) Class: waste cloth, (b) Class: cans, (c) Class: leaves, (d) Class: peel, (e) Class: waste paper, (f) Class: plastic bag, (g) Class: cigarette butts, (h) Class: plastic bottle, (i) Class: cigarette case.

TABLE 6
Total amount of garbage weighted

Garbage classification		Weighting coefficient	Counting results	Total amount
	Small	1	47	47
Inorganic	Medium	2	0	0
	Large	4	0	0
	Small	2	8	16
Organic	Medium	4	0	0
	Large	6	0	0
Tree leaves		1	24	24
Total				87

are shown in Fig. 14. There are 5 different colored lines. The blue lines represent that the value of SV is less than 70, and street cleanliness level is very high. The green lines represent that the value of SV is between 70 and 100, and the street cleanliness level is high. The yellow lines represent that the value of SV is between 100 and 150, and the street cleanliness level is medium. The orange lines represent that the value of SV is between 150 and 200, and the street cleanliness level is low. The red lines indicate that the value of SV is greater than 200, and the street cleanliness level is very low. Therefore, the red lines and the orange lines show that there is a large amount of garbage in the street, and the city administrators should arrange for cleaning personnel immediately.

After we assess the street cleanliness layer, we start block layer analysis based on the multi-level assessment model. The Block layer assessment is based on street layer assessment. We divide Jiangning district into 9 blocks according to sub-administrative areas. Every block has several streets. Here, the cleanliness of a block is determined by the street cleanliness in this block. By referring to Equation 6, the cleanliness assessment values of 9 blocks can be calculated. Different colors represent different street cleanliness values. As shown in Fig. 15, the green and the blue blocks represent a lower cleanliness value and the pavement in the blocks is clean. The yellow blocks represent that cleanliness is medium and there is garbage on the local roads. For example, the value of *Moling* street cleanliness belonging to the block layer is 119 and the cleanliness level is medium. Consequently, this block needs the attention of city administrators.

Finally, city administrators can perform different levels of cleanliness assessment according to their own needs and obtain city cleanliness values based on various levels in time, which can help them to reasonably arrange cleaning personnel.

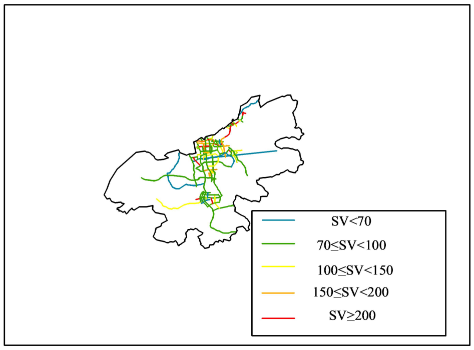


Fig. 14. Exhibition of street level cleanliness assessment based on Jiangning

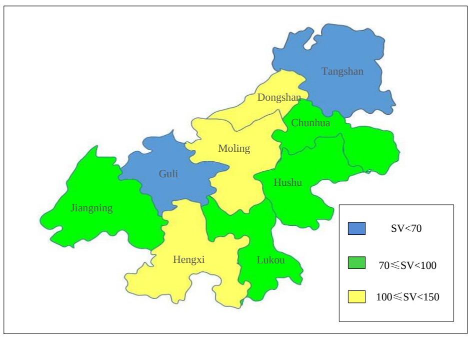


Fig. 15. Exhibition of block level cleanliness assessment based on Jiangning district

6 Conclusions and Suggestions for Future Work

The development of novel technologies has driven a number of cities into the way to smart cities. Street cleanliness is one of the concerns for smart cities. Consequently, this paper proposes a novel urban street cleanliness assessment approach using mobile edge computing and deep learning. A visual street cleanliness road diagram is presented, such an automated system can help city administrators to know the cleaning state of the street easily.

Several directions for future work are possible. These directions are described as follows:

- We plan to develop a solution that can automatically implement image filtering preprocessing at the mobile edge because manual filtering greatly affects the real-time transmission and wastes time.
- Our model contains common street garbage data. However, the model does not play a great role in the uncommon garbage data. Thus, the training data needs to be further expanded to improve the accuracy of the model.
- Our model is always used on sunny days, and the cleanliness on rainy days is also worth studying in the future.

7 Acknowledgements

The work is supported by the National Key R&D Program of China under Grant No. 2018YFC0407901, National Natural Science Foundation of China under Grant Nos. 61572171, and the Fundamental Research Funds for the Central Universities under Grant No. 2019B15414.

References

[1] U. Aguilera, O. Peña, O. Belmonte, and D. López-de Ipiña, "Citizen-centric data services for smarter cities," *Future Generation Computer Systems*, vol. 76, pp. 234–247, 2017.
[2] M. Armbrust, A. Fox, R. Griffith, A. D. Joseph, R. Katz, A. Konwinski, G. Lee, D. Patterson, A. Rabkin, I. Stoica *et al.*, "A view of cloud computing," *Communications of the ACM*, vol. 53, no. 4, pp. 50–58, 2010.
[3] C. Badii, P. Bellini, D. Cenni, A. Difino, P. Nesi, and M. Paolucci, "Analysis and assessment of a knowledge based smart city architecture providing service apis," *Future Generation Computer Systems*, vol. 75, pp. 14–29, 2017.
[4] C. Balchandani, R. K. Hatwar, P. Makkar, Y. Shah, P. Yelure, and M. Eirinaki, "A deep learning framework for smart street cleaning," in *IEEE Third International Conference on Big Data Computing Service and Applications*, 2017, pp. 112–117.
[5] A. Borozdukhin, O. Dolinina, and V. Pechenkin, "Approach to the garbage collection in the smart clean city project," in *Information Science and Technology (CiSt), 2016 4th IEEE International Colloquium on*. IEEE, 2016, pp. 918–922.
[6] L. J. C. Brinez, A. Rengifo, and M. Escobar, "Automatic waste classification using computer vision as an application in colombian high schools," in *Networked and Electronic Media*, 2017, pp. 10 (5 .)–10 (5 .).
[7] N. T. Buck and A. While, "Competitive urbanism and the limits to smart city innovation: The uk future cities initiative," *Urban Studies*, vol. 54, no. 2, pp. 13–43, 2017.
[8] A. Cocchia, "Smart and digital city: A systematic literature review," pp. 13–43, 2014.
[9] ETSI, "Eurropean Telecommuunication Standards Institute (ETSI), Mobile Edge Computing[EB/OL]," http://www.etsi.org/technologies-clusters/technologies/mobileedge-computing, 2016, [Online; accessed 3-December-2016].
[10] FengLei, "IBM, Microsoft, ali and other giants use AI to explore smart cities," http://news.ifeng.com/a/20170622/51302864_0.shtml, 2017, [Online; accessed 22-July-2017].
[11] R. Girshick, "Fast r-cnn," *Computer Science*, pp. 1440–1448, 2015.
[12] I. A. T. Hashem, V. Chang, N. B. Anuar, K. Adewole, I. Yaqoob, A. Gani, E. Ahmed, and H. Chiroma, "The role of big data in smart city," *International Journal of Information Management*, vol. 36, no. 5, pp. 748–758, 2016.
[13] A. Hefnawy, A. Bouras, and C. Cherifi, "Integration of smart city and lifecycle concepts for enhanced large-scale event management," in *IFIP International Conference on Product Lifecycle Management*. Springer, 2015, pp. 687–697.
[14] R. G. Hollands, "Will the real smart city please stand up? intelligent, progressive or entrepreneurial?" *City*, vol. 12, no. 3, pp. 303–320, 2008.
[15] Y. C. Jang, P. Jain, T. Tolaymat, B. Dubey, and T. Townsend, "Characterization of pollutants in florida street sweepings for management and reuse," *Journal of Environmental Management*, vol. 91, no. 2, pp. 320–327, 2010.
[16] Y. Jia, E. Shelhamer, J. Donahue, S. Karayev, J. Long, R. Girshick, S. Guadarrama, and T. Darrell, "Caffe: Convolutional architecture for fast feature embedding," in *Acm International Conference on Multimedia*, 2014, pp. 675–678.
[17] R. Kitchin, "The real-time city? big data and smart urbanism," *Geojournal*, vol. 79, no. 1, pp. 1–14, 2014.
[18] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "Imagenet classification with deep convolutional neural networks," in *International Conference on Neural Information Processing Systems*, 2012, pp. 1097–1105.
[19] LA, "Clean Streets LA Challenge," http://cleanstreetsla.com/clean-streets-challenge/, 2017, [Online; accessed 5-October-2017].
[20] Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," *nature*, vol. 521, no. 7553, p. 436, 2015.
[21] W. Li, B. Bharat, and J. Gao, "A multiple-level assessment system for smart city street cleanliness," in *The Thirtieth International Conference on Software Engineering and Knowledge Engineering (SEKE 2018)At: Hotel Pullman, Redwood City, San Francisco Bay, California, USA*, 2018, pp. 675–681.

[22] J. Long, E. Shelhamer, and T. Darrell, “Fully convolutional networks for semantic segmentation,” in *IEEE Conference on Computer Vision and Pattern Recognition*, 2015, pp. 3431–3440.

[23] I. López, V. Gutiérrez, F. Collantes, D. Gil, R. Revilla, and J. L. Gil, “Developing an indicators plan and software for evaluating street cleanliness and waste collection services,” *Journal of Urban Management*, vol. 6, no. 2, pp. 66–79, 2017.

[24] Y. Mao, C. You, J. Zhang, K. Huang, and K. B. Letaief, “A survey on mobile edge computing: The communication perspective,” *IEEE Communications Surveys & Tutorials*, vol. 19, no. 4, pp. 2322–2358, 2017.

[25] G. Mittal, K. B. Yagnik, M. Garg, and N. C. Krishnan, “Spotgarbage:smartphone app to detect garbage using deep learning,” in *ACM International Joint Conference on Pervasive and Ubiquitous Computing*, 2016, pp. 940–945.

[26] P. Neirotti, A. D. Marco, A. C. Cagliano, G. Mangano, and F. Scorrano, “Current trends in smart city initiatives: Some stylised facts,” *Cities*, vol. 38, no. 5, pp. 25–36, 2014.

[27] M. S. Rad, A. Von Kaenel, A. Droux, F. Tieche, N. Ouerhani, H. K. Ekenel, and J. P. Thiran, “A computer vision system to localize and classify wastes on the streets,” *Liu M., Chen H., Vincze M. (eds) Computer Vision Systems*, vol. 10528, pp. 195–204, 2017.

[28] S. Ren, K. He, R. Girshick, and J. Sun, “Faster r-cnn: towards real-time object detection with region proposal networks,” *IEEE Transactions on Pattern Analysis & Machine Intelligence*, no. 6, pp. 1137–1149, 2017.

[29] G. G. Van Ryzin, S. Immerwahr, and S. Altman, “Measuring street cleanliness: A comparison of new york citys scorecard and results from a citizen survey,” *Public Administration Review*, vol. 68, no. 2, pp. 295–303, 2008.

[30] F. Xia, L. T. Yang, L. Wang, and A. Vinel, “Internet of things,” *International Journal of Communication Systems*, vol. 25, no. 9, pp. 1101–1102, 2012.

[31] I. Yaqoob, E. Ahmed, A. Gani, S. Mokhtar, M. Imran, and S. Guizani, “Mobile ad hoc cloud: A survey,” *Wireless Communications and Mobile Computing*, vol. 16, no. 16, pp. 2572–2589, 2016.

[32] M. D. Zeiler and R. Fergus, “Visualizing and understanding convolutional networks,” in *European Conference on Computer Vision*, vol. 8689, 2014, pp. 818–833.

[33] S. Zygiaris, “Smart city reference model: Assisting planners to conceptualize the building of smart city innovation ecosystems,” *Journal of the Knowledge Economy*, vol. 4, no. 2, pp. 217–231, 2013.