A Method Of Monitoring Cleanliness Of An Area And A System Thereof

Cross-Reference to Related Applications

[0001] The present application claims the benefit of Singapore Patent Application No. 10202200465W filed 17 January 2022, which is incorporated by reference herein.

Technical Field

[0002] The present invention relates to a method of monitoring cleanliness of an area and a system thereof.

Background

[0003] The cleanliness of our streets is an essential facet of the environment and public health. Current methods involve mechanical sweepers moving around according to routine schedules and mobile teams manually confirming if the entire streets, turf, and pavements are cleaned. This manual method takes substantial effort to detect and assess the cleanliness of the street, turf, and pavement, leading to poor efficiency and productivity.

[0004] It would be advantageous to provide a solution to improve the efficiency and productivity of maintaining the cleanliness of our streets.

Summary

[0005] According to various embodiments, a method of monitoring cleanliness of an area is provided. The method includes capturing images of the area to be monitored by a camera system mounted on a mobile unit, retrieving the geo-location of the location where each of the images is captured, processing the images by a processing module on the mobile unit, wherein processing the images comprises classifying the images into a plurality of cleanliness levels, and geo-tagging the images with the geo-location of the images, and transmitting the processed images to a remote server.

[0006] According to various embodiments, the plurality of cleanliness levels may include a clean level and an unclean level, wherein classifying the images comprises classifying the images into the clean level and the unclean level, such that the method may further include selecting processed images from the unclean level, retrieving the geo-locations of the processed images, retrieving cleaning routes by mobile cleaning units and identifying the geo-locations along cleaning routes.

[0007] According to various embodiments, classifying the images may include identifying an asset in each of the images, identifying at least one lapse in each of the images, and categorising the images into the plurality of cleanliness levels.

[0008] According to various embodiments, a system for monitoring cleanliness of an area is provided. The system includes a processing module mounted on a mobile unit, the processing module includes a processor, a memory in communication with the processor for storing instruction executable by the processor, a camera system in communication with the processor for capturing images and configured to transmit the captured images to the processor, a GPS module in communication with processor for obtaining a geo-location and configured to transmit the geo-location to the processor, such that the processor is configured to capture images of the area to be monitored by the camera system mounted on the mobile unit, retrieve the geo-location of the location where each of the images is captured, process the images, such that to process the images, the processor is configured to classify the images into a plurality of cleanliness levels, and geo-tag the images with the geo-location of the images, and transmit the processed images to a remote server.

[0009] According to various embodiments, the plurality of cleanliness may include a clean level and an unclean level, such that classifying the images may include classifying the images into the clean level and the unclean level, such that the processor is further configured to select processed images with the unclean level, retrieve the geo-locations of the processed images, retrieve cleaning routes by mobile cleaning units and identify the geo-locations along cleaning routes.

[0010] According to various embodiments, to classify the images, the processor may be configured to identify an asset in each of the images, identify at least one lapse in each of the images, and categorise the images into the plurality of cleanliness levels.

[0011] According to various embodiments, a cleanliness monitoring system for monitoring the cleanliness of an area is provided. The cleanliness monitoring system includes a mobile unit having a system according to any one of the abovementioned embodiments.

[0012] According to various embodiments, the mobile unit may be vehicular.

Brief Description of Drawings

[0013] Fig. 1 shows an exemplary method of monitoring cleanliness of an area.

[0014] Fig. 2 shows an exemplary embodiment of a system for monitoring cleanliness of area.

[0015] Fig. 3 shows an example of the system architecture of the system.

[0016] Fig. 4 shows an example of the data flow architecture of the system.

[0017] Fig. 5 and Fig. 6 shows examples of the assets and lapses respectively.

[0018] Fig. 7 shows a flow diagram of the 3-stage process to build the AI model for identification of lapse in the images of the area.

[0019] Fig. 8 shows a table of 10 different AI models trained to identify unique and specific assets and lapse.

[0020] Fig. 8A shows examples of the mobile unit.

[0021] Fig. 9 shows a flow diagram of Stage 1 of the process to build the AI model.

[0022] Fig. 10 shows images of the various cleanliness levels.

[0023] Fig. 11-Fig. 13 show conditional filter matrices**.**

[0024] Fig. 14 shows a flow diagram of a conditional filter for identifying a condition of "Turf-Area Moderate" as shown in Fig. 11.

[0025] Fig. 15 and Fig. 16 show screenshots of a dashboard showing various real-time cleanliness levels, i.e. "Clean", "Moderate", "Critical", of a geographical area.

[0026] Fig. 17A and Fig. 17B show labelled images processed by the AI model.

[0027] Fig. 18 shows a graph of prediction accuracy of the assets and lapses.

Detailed Description

[0028] Fig. 1 shows an exemplary method 1000 of monitoring cleanliness of an area. Method 1000 includes capturing images of the area to be monitored by a camera system 120 (as shown in Fig. 2) mounted on a mobile unit 10 (as shown in Fig. 2) in block 1020, retrieving the geo-location of the location where each of the images is captured in block 1040, processing the images by a processing module 110 on the mobile unit in block 1060, such that processing the images includes classifying the images into a plurality of cleanliness levels (or categories), and geo-tagging the images with the geo-location of the images, and transmitting the processed images to a remote server 130 (as shown in Fig. 2) in block 1080. Method 1000 is able to identify and recognize assets, lapses and classify the cleanliness level of an area with a geo-location for every analysed image. The method 1000 enables remote monitoring and real-time data gathering of the state of cleanliness of public assets, e.g. street, turf, pavement, etc. In this way, the method 1000 enables high efficiency and productivity in

monitoring and cleaning of the assets in the area. Method 1000 may be a computer-implemented method.

[0029] Fig. 2 shows an exemplary embodiment of a system 100 for monitoring cleanliness of area. System 100 includes a processing module 110 mountable to the mobile unit 10. Processing module 110 includes a processor 112 for processing the images, a memory 150 in communication with the processor 112 for storing instruction executable by the processor 112, a camera system 120 in communication with the processor 112 for capturing the images and configured to transmit the capture images to the processor 112, a GPS module in communication with the processor 112 for obtaining a geo-location and configured to transmit the geo-location to the processor 112. Processing module 110 may include a power supply 140 configured to supply power to the processing module 110, an I/O interface for providing an interface between the processor 112 and peripheral interface modules. Processing module 110 is configured to communicate with the remote server 130. Processor 112 is configured to capture images of the area to be monitored by the camera system mounted on the mobile unit 10, retrieve the geo-location of the location where each of the images is captured, process the images, such that to process the images, the processor 112 is configured to classify the images into a plurality of cleanliness levels, and geo-tag the images with the geo-location of the images, and transmit the processed images to a remote server 130.

[0030] System 100 enables edge computing where the processing module 110 is wired to the video source, i.e. the cameras, and processing of the image is real-time. Edge computing provides more of localized data intake and analysis performed on the fly for real-time decisions and action. It also saves significant bandwidth time and cost involving connectivity infrastructure, e.g. 4G data, especially in situations where rapid decisions need to be made across multiple devices, i.e. multiple automotive cameras installed on the mobile units 10. Comparatively, cloud computation, where data is transmitted from the mobile unit 10 to a cloud server to be processed, requires bandwidth plus data and computing power, which tends to be internet-centric architecture with 4G as backhaul for remote and mobile applications, such as on-road inspection systems using a cleaning vehicle as a medium. Hence, the system 100 improves inspection and maintenance capabilities and presented new mediums of On-Demand Cleaning. Moreover, the emergence of an IoT powered fleet system

enables real time data analytics and real time monitoring system, which moves the operation planning, one step towards maximized productivity and minimized wastage.

[0031] Each mobile unit 10 may include a mobile cleaning unit 12 or a mobile inspecting unit 14 as shown in Fig. 8A. Mobile cleaning unit may be vehicular, e.g. a sweeper truck. Mobile inspecting unit 14 may be vehicular, e.g. a truck, a car, a motorcycle. An asset is a feature in an area e.g. footpath, litter bin, turf area, road kerb, drop-inlet chamber, illegal dumping, turf area lane closed, water channel, scupper hole, bus stop, etc. A lapse is a failure in cleanliness or items that cause an asset to be unclean, e.g. dry leaves, cigarette butts, dried twigs, plastic wrapper, used masks, tissue paper, silt deposition plastic/glass bottles, metal can, cardboard, branches, palm fronds, etc.

[0032] The plurality of cleanliness levels may be broadly classified into a clean level and an unclean level, such that classifying the images includes classifying the images into the clean level and the unclean level. Method 1000 further includes selecting processed images from the unclean level, retrieving the geo-locations of the processed images, retrieving cleaning routes by mobile cleaning units 12 and identifying the geo-locations along cleaning routes. As the mobile cleaning units 12 are able to clean only the areas which are unclean, the system 100 allows the efficient deployment of resources to only attend the areas of concern during cleaning routes and this helps to save time for the mobile cleaning units 12. During a trial, the mobile cleaning units 12 were able to save an average of 35 manhours per week per team with the aid of the system 100. System 100 enables deployment of ground resources more efficiently to clean only areas of need. Similarly, the mobile inspecting units 14 will be able to use the system 100 for auditing the scheduled cleaned areas for cleaning quality and completion. As the mobile inspecting units 14 are travelling along the cleaning route, the images captured by each unit may be processed by the on-board processing module 110 and the cleanliness of the areas may be assessed real-time. The mobile inspecting units 14 will not need to stop and conduct paper audits or take images to verify work done or otherwise. System 100 may be configured to compare the cleanliness levels of the assets in the same location, i.e. compare images of the same geo-location, before and after cleaning and verify if the asset has been cleaned. Hence, the efficiency and productivity of the inspection team is increased.

[0033] Processing module 110 may be an on-board module of the mobile unit 10, i.e. the processing module 110 may be mounted onto the mobile unit 10. Processing module 110 may be a server and configured to communicate wirelessly with the remote server 130. Remote server 130 may be located at the control centre or at a user premise. Processing module 110 may include an AI module configured to carry out identification of the asset, the lapse, and the cleanliness level. Processing module 110 may be configured to compute and perform Deep Learning Inferencing as the mobile cleaning unit 12 is on the move, i.e. as the mobile unit 10 moves on road. Processing module 110 may be mounted onto a mechanically vibration damping system to minimise the vibration to the processing module 110.

[0034] Fig. 3 shows an example of the system architecture 300 of the system 100. System 100 enables all the different end users, namely the operation team, middle and top management, to be able to access information of the area with cleanliness level status in real-time through an interactive dashboard and mobile app. The camera system 120 and processing module 110 installed on the mobile unit 10 is responsible for performing Deep Learning computation using the pre-trained models to identify and recognize assets, lapses, and classify the cleanliness level with a geo-location for every analysed image. Further, these analysed results are streamed to the backend remote server 130 for performing trend analysis, data visualization, and storage. A wide range of visualization charts may be integrated to generate reports, e.g. Historical report view, Cleaning Feedback Matrix, Street-Route-Sector wise cleanliness level, and a real-time Google™ map view showing the lapse status.

[0035] Fig. 4 shows an example of the data flow architecture 400 of the system 100. The entire data flow architecture 400 is orchestrated as multi-parallel pipeline system for handling and processing a massive concurrent video stream for the plurality of cameras 122 mounted on the mobile units 10. Each frame grabber module received images from their respective camera adaptors and performs a rapid image rescaling to a size of 300 x 300 pixels to fit the trained Deep Learning modules for accurate recognition. The re-sized images are tagged with the geo-location information, its respective camera signature and finally placed into an elastic image queue temporary memory. A state machine designed to handle the input image queue is responsible for reading the images in a First-In-First-Out sequence and passing the images to 10 different parallel Inferencing (AI) threads using 10 different AI models to recognize and

classify different types assets, differ types of lapses with the cleanliness level. Finally, the results are placed into the output result buffer, where a set of additional condition filters are applied to refine the results from the 10 AI threads, by which the prediction accuracy improves. The results are mapped into a custom JSON structure and beamed out to the remote server 130, e.g. cloud, using an encrypted IoT Core connectivity.

[0036] Fig. 5 and Fig. 6 shows examples of the assets and lapses respectively.

[0037] An AI model may be built to recognise the asset, lapse in the asset and cleanliness level of the asset in the image. The AI model is trained using transfer learned Convolution Neural Network (CNN) technique, where the technique is used to train and inference the real-time images of on-road assets, e.g. footpath, litter bin, turf area, road kerb, drop-inlet chamber, illegal dumping, turf area lane closed, water channel, scupper hole, bus stop, etc. The cleaning route is planned along the assets to be cleaned. In other words, the assets are found along the cleaning routes. In addition, the technique is used to train the AI model to identify different lapses, e.g. dry leaves, cigarette butts, dried twigs, plastic wrapper, used masks, tissue paper, silt deposition plastic/glass bottles, metal can, cardboard, branches, palm fronds, etc. found on the respective assets and categorises them into cleanliness levels, e.g. clean, moderate and critical. The architecture for building the AI model is constructed with a 3-stage process, i.e. Stage 1, Stage 2 and Stage 3. In Stage 1, the images are manually labelled and tagged. For example, 500K night and day images are manually tagged based on the type of assets, type of lapse and into the cleanliness levels. Stage 2 comprises training the AI model based on the tagged images. Stage 2 uses a supervised learning technique with SSD MobileNET V2 as the backbone CNN and Stochastic Gradient Descent (SGD) as the synaptic weight optimization algorithm. A COCO dataset was transfer–learned during Stage 1 to generate unique AI models to recognize lapse by categorizing them into assets and lapse as trained. Stage 3 comprises applying a conditional filter matrix to the trained AI model to improve the recognition rate. In Stage 3, a 2nd order custom build conditional filter function is applied to the trained AI models to improve the recognition rate.

[0038] Fig. 7 shows a flow diagram 700 of the 3-stage process to build the AI model for identification of lapse in the images of the area. Stage 1 comprises six process steps, starting with installing the cameras 122 on the mobile units 10 in block 712A and tuning the camera

lens with appropriate sensor parameters to match the pre-defines camera angle and varying atmospheric light conditions in block 712B. Upon setting the camera system 120, real-time images are captured and collected while the mobile units 10 are moving in block 712C. The collected images are manually transferred to the backend cloud annotation server in block 712D. The manual transfer process involves manually removing the external hard drive from the unit server on the mobile unit 10 and connecting it to a stable broadband connection for image upload to the server. The uploaded images are manually annotated with a bounding box in block 712E and, finally, a JSON structure extracted with the annotation information in block 712F. Stage 2 is the computationally intense layer which involves massive parallel computing on the GPU's installed and configured on a RACK server backbone with TensorFlow framework. Stage 2 comprises six process steps, starting with preparing a COCO data set in block 722A. The transfer training was configured with the COCO dataset on an SSD MobileNET V2 in block 722B. The annotated data from Stage 1 was used to compute the optimized gradient with appropriate learning rate and epochs in block 722C. As a final step, the generated frozen graph output from the TensorFlow framework is compiled in block 722D with a model optimized in block 722E, which gives a required 10 different AI models optimized for on-road inspection in block 722F as shown in Fig. 8. The 10 different AI models are trained to identify unique and specific assets and lapses as shown in the Table 800 in Fig. 8. Stage 3 involves building additional feature conditional filter matrix based on the prediction confidence level from the Stage 2 AI models in Block 732A. A conditional matrix is built for each of the cameras in block 732B. From the conditional matrix of each camera, a conditional filter matrix is built in block 732C. The conditional filter matrix is shown in Fig. 11-Fig.13 and a single use case is explained below. The 3-stage process will also be further elaborated below.

[0039] Fig. 9 shows a flow diagram 900 of Stage 1 of the process to build the AI model. Images of the area in the day and night are collected, e.g. 20TB, and manually filtered into an image data set, e.g. 500K. Furthermore, these filtered images are annotated and labelled with the list of assets and lapses. The definition of the cleanliness levels, i.e. clean, moderate and critical, (shown in Fig. 10) were assigned with practical reference by the operation team. A "clean" level includes "clean", and cleaning would not be required. An "unclean" level includes "moderate" and "critical", and cleaning will be required.

[0040] Fig. 11-Fig. 13 show conditional filter matrix 1100,1200,1300. The final architecture stage is a matrix of tuned conditions further integrated with the inference engine while performing real-time detection. The trained AI model from Stage 2 outputs the results in a JSON structure with the assets and the type of lapse. The results are a probability value from 0% to 100% as a measure of the prediction confidence scale or level. Hence, the conditional filter matrix 1100,1200,1300 as shown in Fig. 11-Fig.13 with the confidence range was derived to enhance the recognition accuracy. Confidence range may be defined by the user to fine-tune the recognition accuracy. This conditional filter matrix 1100,1200,1300 is tuned to the optical system with the listed conditions in the conditional filter matrix 1100,1200,1300 and defined for the different light conditions of the day. The uniquely labelled Stage 1 data set, along with the custom tuned conditional filter matrix 1100,1200,1300 of Stage 3, sandwiched with the open standard Stage 2 results in a balanced and tuned combination further improving the performance of the AI model.

[0041] Fig. 14 shows a flow diagram 1400 of the conditional filter matrix 1100,1200,1300 for identifying a condition of “Turf-Area Moderate” as shown in Fig. 11, Rows 10-13. In this example, the asset is the “Turf-Area”, and cleanliness level is “Moderate”. The process starts by filtering out the cameras by the camera signature. As shown, cameras 1 or 4 and 2 or 3 share the same sequence. For images from camera 1 or 4, if the confidence threshold is above 50%, then the condition is true, i.e. the turf area is in moderate condition, else if, the condition is between 20% and 50%, the system 100 checks for Presence and Absence conditions as shown in the flow diagram. A “Presence Condition” is where lapses are identified by the AI model from the image. An “Absence Condition” is where an asset or lapse is not identified by the AI model in the image. As shown in Fig. 14, the Presence conditions include lapse such as DRIED-LEAVES, DRIED-TWING, PALM-FROND, and the Absence conditions excluded assets such as DIVIDER-GATE, ZEEBRA-CROSSING, BARRY-GUARD, PLANTATION and CONSTRUCTION-WORK for the present given example. The Presence and Absence are secondary AI models trained for recognizing assets other than the primary asset list, and these secondary conditions are further used to filter out the final prediction result and refine the accuracy. For example, if the confidence threshold is between 20 and 40 and includes Dried-Leaves but excludes a Divider-Gate, the condition is true. Any condition outside the threshold mentioned above the recognition is considered as false.

[0042] Fig. 15 and Fig. 16 show screenshots 1500,1600 respectively of a dashboard showing various real-time cleanliness level, i.e. “Clean”, “Moderate”, “Critical”, of a geographical area. Using the data collected, it is possible to use the system 100 to plan an effective On-Demand cleaning schedule for the mobile units 10 and attend to areas that require cleaning on given days due to environmental and other localized factors so that it is possible to anticipate and plan the time and resource required for cleaning the route. Instant notification is sent to the respective operation personnel during the identification of critical cleanliness level to follow-up. In addition, every cleaning feedback is linked to a geo-location, e.g. Google™ map location, with the location photo (as shown in the white box in the centre of Fig. 16) for easy identification and follow-up.

[0043] Fig. 17A and Fig. 17B show labelled images 1702,1704 processed by the AI model. As shown in Fig. 17A, the asset identified is “Turf-Area”, the lapse identified is “Dry-Leaves” and the cleanliness level is “Critical”. Referring to Fig. 17B, the asset identified is “Turf-Area”, the lapse identified are “Dry-Leaves” and “Metal Can” and the cleanliness level is “Critical”.

[0044] Fig. 18 shows a graph 1800 of prediction accuracy of the assets and lapses. Based on a randomly selected sample of 250 test images, inferencing is performed using the Stage 1, 2 and 3 trained AI models, for which the obtained asset/lapse prediction accuracy is plotted to the graph 1800. As shown in Fig. 18, the system 100 is able to provide a prediction accuracy of at least 80%.

[0045] Camera system 120 may include plurality of cameras and a camera adapter connected between the processing module 110 and the plurality of cameras. Plurality of cameras include high-speed global shutter cameras for capturing sharp and clear images. Each of the plurality of cameras may be set up, configured, and optimized to capture images of the assets from the different sides of the mobile unit 10. For example, the lens of each camera is sized to match an appropriate focal length from the various sides, e.g. front, side, rear, of the mobile unit 10, the mounting location/height and varying light conditions, to enable a sharp focus on the monitored area. The connectivity between the cameras and processing module 110 is

designed to deliver very high speed, and good quality videos. Each mobile unit 10 may include one or more cameras attached thereto, e.g. 2-4 cameras. Each camera may be high resolution, e.g. 4MP, and a high-speed automotive-grade camera. Images may be captured at 40 frames per second.

[0046] The application running on the processing module 110 is configured to capture the images from the cameras, provide geo-tagging for every captured image, and analyse the captured images to identify various assets, lapses, and the cleanliness level. The obtained results are transmitted to a cloud server 130 wirelessly for storage, machine learning, and data visualization applications.

[0047] System 100 may include a server, a laptop, a computer, etc. Processor 112 typically controls overall operations of the system 100, such as the operations associated with display, data communications. Processor 112 may include one or more processors to execute instructions in the above-described modules to perform all or part of the steps in the above-described methods, e.g. “Xeon Core Intel Processor” based on an industrial grade CPU integrated with 16 core Intel Vector Processing units. Moreover, the processor 112 may include one or more modules which facilitate the interaction between the processor 112 and other modules. System 100 may be in communication with the user devices via the network.

[0048] Memory 150 may be configured to store various types of data to support the operation of the system 100. For example, the data may include instructions for any applications or above methods operated on the system 100, programmes, applications, modules etc. Memory 150 may be implemented using any type of volatile or non-volatile memory devices, or a combination thereof, such as a static random access memory (SRAM), an electrically erasable programmable read-only memory (EEPROM), an erasable programmable read-only memory (EPROM), a programmable read-only memory (PROM), a read-only memory (ROM), a magnetic memory, a flash memory, a magnetic or optical disk.

[0049] Power supply 140 provides power to various modules of the system 100. Power supply 140 may include a power management system, one or more power sources, and any other modules associated with the generation, management, and distribution of power in the system 100. Power supply 140 may include a high-grade vibration-proof railway grade power

supply to power up the entire camera and AI system running on the edge (on the mobile units 10).

[0050] System 100 may include an I/O interface 160 for providing an interface between the processor 112 and peripheral interface modules, such as a keyboard, a click wheel, buttons, and the like. The buttons may include, but are not limited to, a home button, a volume button, a starting button, and a locking button.

[0051] System 100 may include a communication module configured to facilitate communication, wired or wirelessly, between the system 100 and other devices. System 100 may access a wireless network based on a communication standard, such as WiFi, 2G, or 3G, LTE, and 4G cellular technologies or a combination thereof. In one exemplary embodiment, the communication module may receive a broadcast signal or broadcast associated information from an external broadcast management system via a broadcast channel.

[0052] In exemplary embodiments, the system 100 may be implemented with one or more application specific integrated circuits (ASICs), digital signal processors (DSPs), digital signal processing devices (DSPDs), programmable logic devices (PLDs), field programmable gate arrays (FPGAs), controllers, micro-controllers, microprocessors, or other electronic modules, for performing the above described methods.

[0053] The present invention relates to a method of monitoring cleanliness of an area and a system thereof generally as herein described, with reference to and/or illustrated in the accompanying drawings.

Claim

1. A method of monitoring cleanliness of an area comprising:

capturing images of the area to be monitored by a camera system mounted on a mobile unit,

retrieving the geo-location of the location where each of the images is captured,

processing the images by a processing module on the mobile unit, wherein processing the images comprises classifying the images into a plurality of cleanliness levels, and geo-tagging the images with the geo-location of the images, and

transmitting the processed images to a remote server.

2. The method according to claim 1, wherein the plurality of cleanliness levels comprises a clean level and an unclean level, wherein classifying the images comprises classifying the images into the clean level and the unclean level, wherein the method further comprises selecting processed images with the unclean level, retrieving the geo-locations of the processed images, retrieving cleaning routes by mobile cleaning units and identifying the geo-locations along cleaning routes.

3. The method according to claim 1 or 2, wherein classifying the images comprises identifying an asset in each of the images, identifying at least one lapse in each of the images, and categorising the images into the plurality of cleanliness levels.

4. A system for monitoring cleanliness of an area, the system comprising:

a processing module mounted on a mobile unit, the processing module comprises,

a processor

a memory in communication with the processor for storing instruction executable by the processor,

a camera system in communication with the processor for capturing images and configured to transmit the captured images to the processor,

a GPS module in communication with processor for obtaining a geo-location and configured to transmit the geo-location to the processor,

wherein the processor is configured to:

capture images of the area to be monitored by the camera system mounted on the mobile unit,

retrieve the geo-location of the location where each of the images is captured,

process the images,

wherein to process the images, the processor is configured to classify the images into a plurality of cleanliness levels, and geo-tag the images with the geo-location of the images, and

transmit the processed images to a remote server.

5. The system according to claim 4, wherein the plurality of cleanliness levels comprises a clean level and an unclean level, wherein classifying the images comprises classifying the images into the clean level and the unclean level, wherein the processor is further configured to select processed images with the unclean level, retrieve the geo-locations of the processed images, retrieve cleaning routes by mobile cleaning units and identify the geo-locations along cleaning routes.

6. The system according to claim 4 or 5, wherein to classify the images, the processor is configured to identify an asset in each of the images, identify at least one lapse in each of the images, and categorise the images into the plurality of cleanliness levels.

7. A monitoring system for monitoring the cleanliness of an area comprising:

a mobile unit comprising a system according to one of claims 4 to 6.

8. The monitoring system according to claim 7, wherein the mobile unit is vehicular.

9. A non-transitory computer readable storage medium comprising instructions, such that the instructions, when executed by a processor in a system, cause the system to:

capture images of the area to be monitored by a camera system mounted on the mobile unit,

retrieve the geo-location of the location where each of the images is captured,

process the images by a processing module on the mobile unit, wherein processing the images comprises classifying the images into a plurality of cleanliness levels, and geo-tagging the images with the geo-location of the images, and

transmit the processed images to a remote server.