IOSR Journal of Computer Engineering (IOSR-JCE)
e-ISSN: 2278-0661, p-ISSN: 2278-8727, Volume 28, Issue 2, Ser. 1 (Mar. – Apr. 2026), PP 73-80
www.iosrjournals.org

AI-Driven Predictive Waste Management Using Iot For Smart Cities

Ameer Malik Bahad K
Department Of Data Science
Dr. M. G. R Educational And Research Institute
Chennai, India

Mr. S. Mohan
Department Of Cyber Security
Dr. M. G. R Educational And Research Institute
Chennai, India

Akash M
Department Of Data Science
Dr. M. G. R Educational And Research Institute
Chennai, India

Dr. D. Usha
Department Of Data Science
Dr. M. G. R Educational And Research Institute
Chennai, India

Dilly Babu S
Department Of Data Science
Dr. M. G. R Educational And Research Institute
Chennai, India

Abstract
Urbanization has made managing municipal waste more difficult. There are many cities across the globe where conventional models of waste collection are followed, based on scheduled trips to dispose the waste. This results in overflowing of bins, addition of unnecessary trips, pollution of the environment etc. In this project, we propose a smart waste management system using IoT and machine learning. The fill level of the waste bins are monitored by ultrasonic sensors placed in the bins. The data is transferred to a cloud based server which uses machine learning predictive analytics to report when each waste bin is likely to get full, and update the municipal authorities to fix waste collection before the waste bin overflows. A web dashboard lets the operators work with the full status of the waste picking bins, check the collection times which is referred by the report, and optimize the routes for the waste pickup vehicles. Sensing, prediction and operational planning of the waste disposal is included in this work along with deployment recommendation for a smart clean city.
Keywords: *Intelligent Waste Management, IoT, Predictive Analytics, Urban Sustainability.*

Date of Submission: 04-04-2026 Date of Acceptance: 14-04-2026

I. Introduction

Rapid urbanization in many countries has increased the pressure on waste management systems. The growth in size of urban populations has led to a corresponding increase in the amount of solid waste created each day, which has significantly impacted the logistics involved in collecting, deploying personnel, and disposing of that material. These problems are more serious in big cities with high population density, as inadequate waste collection makes public spaces less appealing and less sanitary, and creates a number of larger issues related to the environment and public health, including contamination of groundwater, transmission of diseases by vectors, and greenhouse gas emissions as a result of decomposing organic matter.

The majority of traditional garbage collection programs rely on a schedule-driven approach to dispatch collection vehicles to fixed-frequency routes. Collection vehicles are dispatched to collect bins according to a pre-defined schedule without consideration of the actual fill levels of the bins. This means that vehicles are often dispatched to collect bins that are only partially full, thus wasting both fuel and time for the driver, while other high-traffic bins overflow due to not being collected at all between scheduled collections, resulting in unsanitary conditions and declining levels of public satisfaction. This gap between fixed schedules and real needs shows that a better system is required.

The integration of the IoT, affordable low-cost sensing hardware, wireless communication technologies, and cloud computing has opened up new opportunities to implement and automate urban infrastructures. Smart waste bins equipped with fill-level sensors (e.g., ultrasonic distance sensors) can report their fill levels and on a regular basis to a centralised platform.

The platform will provide a current snapshot of waste levels throughout the city, which can then be used to develop predictive models that make use of historical fill-level data and machine learning algorithms to forecast when bins are likely to overflow or require collection and adjust collection schedules proactively to prevent this from occurring.

Our paper describes and evaluates an AI-enabled predictive waste management system intended for urban implementation. It comprises: ultrasonic distance sensors (mounted in waste bins), data transmission via an IoT-based layer that employs microcontroller-based technologies, time-series fill-level data stored in the cloud, predictive analytics based on machine learning, and a web-based dashboard for operator interaction.

The overall goal of this system is to move from reactive waste collection practices to proactive, data-driven operation to limit instances of overflowing waste while minimising unnecessary collection trips and their associated costs.

To overcome the issues associated with traditional waste collection processes, intelligent systems are required to monitor garbage in real-time and make decisions based on accurate data. The emergence of IoT, cloud computing and machine learning have created the opportunity to develop waste management processes that forecast waste fill levels, optimise garbage collection schedules, eliminate overflowing waste collections, and achieve overall improved efficiencies.

II. Literature Review

Smart waste management has become an important research area in recent years as cities have grown rapidly and the affordability of Internet of Things (IoT) technologies has increased. The early research on smart waste management focused on developing sensor-based monitoring systems for waste bins. The sensors used included ultrasonic or infrared sensors installed/placed in the bin that measure the fill level of the bin and send the information to a remote server via GSM, Wi-Fi or cellular networks. These early systems have shown that monitoring the fill level of bins in real-time is possible at a relatively low cost. However, the only thing these systems could do was to provide the status of the bin and create alerts; they did not incorporate predictive intelligence or optimize how the bins were serviced.

Later, researchers improved this idea by using routing algorithms and techniques from Operations Research to develop a model for how to best collect waste; research found that using the various status of the bins at the time of their collection could significantly reduce the amount of fuel consumed and the distance travelled through the use of a dynamic routing process instead of a traditional fixed routing schedule to collect waste from bins. Additionally, incorporating Geographic Information System (GIS) features such as roads and traffic limitations would further help improve the routes to collect the waste. However, most systems still only provide monitoring and not prediction to providing real-time bin status alone, as there was no capability for creating predictive schedules, therefore, they were considered to be reactive rather than proactive systems.

Recently, machine learning models have been developed to predict waste generation trends. For example, regression design methods (Linear Regression), machine learning algorithms (Random Forest regression), and LSTM networks have been used with fill-level data from past bins to forecast bin capacity in the future. The analyses identified solid temporal trends that were affected by day-of-week and time-of-day data; especially for bins in areas of commerce and bins with heavy traffic. While improving prediction accuracy, many of these studies were still simulation-based, and did not provide a functional link between prediction models and real-world bin hardware or operational dashboards.

Another trend is the development of cloud-based monitoring systems that allow remote access to fill-level information via the web and mobile devices. Although this provides better access to the fill-level data and increases its transparency, there are no sophisticated analytics, intelligent schedule systems, or user interfaces designed specifically for non-technical municipal employees. Furthermore, much of the previous research only tested the sensors, communications, optimization algorithms, and prediction models separately, rather than together as an entire, integrated, and deployable system. Testing in real environments has also been poorly represented throughout the literature.

This research will fill existing gaps in smart waste management systems by integrating a fully-integrated system for waste management that incorporates sensing, communication of position data, predictive analytics, and visualization as a single service. The current models are primarily designed for monitoring and routing;

therefore, they do not provide a means for operators to anticipate when pickup is going to be needed; our proposed model will incorporate forecasting using machine learning to enable proactive planning. Additionally, it will provide a user-friendly dashboard for operators that is customized for municipal users, making it easily usable by operators in an operational setting and not just a research environment. By providing a hardware and software solution that is scalable and able to be deployed, this research will move from developing prototypes through developing large-scale solutions for smart waste management in urban areas.

III. Proposed Methodology

The system works in four main steps: data collection, data transfer, prediction, and result display. The acquisition of real time data through distributed sensor node channels, the transmission and cloud based storage of fill level time-series, the predictive modelling of future fill trajectories based upon machine learning techniques, and the presentation/delivery of actionable insights via a monitoring interface. Each step is designed to run automatically with minimal human input.

Problem Formulation and System Requirements

This system solves the problem of fixed collection schedules that do not match real waste levels, we have to pick up garbage and the dynamic, spatially disparate nature of garbage being generated in urban locations. The system will provide a way to accurately know when instrumented bins will fill to their overflow hazard thresholds (i.e. levels at which continued use would create a risk to public health by overflowing). The predicted fill time will continuously be generated for each instrumented bin and sent to a scheduling module that will allow for scheduling collection to take place at exactly when it should and where it should.

System requirements were developed from common urban waste collection scenarios. The hardware layer will accommodate bins of varying sizes and shapes both indoors/outdoors, while being able to withstand all temperature ranges, humidity levels and electromagnetic interference. The communication layer will support reliable data transmission in densely populated urban radio environments with acceptable latency. The prediction engine will need to provide sufficient accuracy to determine when a schedule should be modified, while providing sufficient clarity in the generated output for non-specialist personnel (users) to be able to read the output from a graphical user interface.

Sensor Deployment and Data Acquisition

Ultrasonic sensors are placed at the top of the bin to measure the waste level of each waste bin, oriented downward to measure the distance between the sensor and the surface of accumulated waste material. Fill level is derived by subtracting the measured distance from the known total bin depth, yielding the occupied volume as a percentage of total capacity. Sensors are interfaced with microcontroller units that provide onboard processing, analog-to-digital conversion.

Basic filtering is used to remove incorrect sensor readings to raw sensor measurements to remove transient anomalies, while implausible sensor readings are flagged by sanity validation as sensor errors. These preprocessing steps significantly improve the quality of the data and reduce transmission overhead.

Cloud Storage and Data Management

All data points received by the cloud broker include: bin ID, coordinates, date, raw distance measurement, and calculated percentage of fill. The system stores all of this in a time-series structured database. Each bin is also assigned a metadata record that provides details about the bin owner including dimensions, location type, and any historical maintenance events; this will allow users to perform richer analyses on how the bins were used. Additionally, the cloud broker will allow for bulk export of historical data for offline training and validating models.

Predictive Modelling

The system uses past data to predict when a bin will become full for every bin to determine the probability of being between overflow and fill levels. It uses a sliding window approach to input recent fill-level observations for every time period (hour, day, etc.) to produce an output prediction of fill levels at a specified time period in the future. There are many types of regression modeling to build the system.

The features created from temporal data (i.e. hour of day; day of week; or distance to known periods of high bin activity) along with the lagged fill-level observation(s) and calculating the overall rolling statistical properties of moving average fill rate are all used as input into a particular algorithm. Together, these features represent the periodic nature of waste collection patterns, along with the recent short-term trends of that bin.

We tested different models and selected the best one based on accuracy, including both gradient-boosted decision trees and recurrent neural networks, and the final model(s) were selected based on the MAE error of cross-validating a holdout set of data. As more data becomes available the final model will need to be retrained to ensure that the prediction values continue to match how bins will be used over time.

Alert Generation and Schedule Optimization

Predictive fills are used by the system to locate bins that are expected to go over the defined overflow threshold and provide information about each bin. This information is sent to the route design module where it is used to create optimized collection routes based on the current fill level, expected fill level and the geographical location of the bins. Collection routes are created to structurally reduce total distance travelled while considering time and load being handled by an individual vehicle. Generated routes will be displayed to Field Operators through the Dashboard Interface.

Performance Evaluation Using Simulated Data

The proposed system is evaluated on the basis of historical and simulated data since no real-world implementation is available. This will be accomplished via comparing actual or simulated fill levels against the predicted fill levels using various different error terms (e.g., mean absolute error and root mean square error). The dashboard's performance will be evaluated with regard to latency, the accuracy of visual representation, stability of the system and real-time notifications using multiple bin simulations. Through these evaluations there is ample evidence that the system is feasible and could be expanded to an IoT-enabled solution in the future.

IV. System Design

The system is designed in multiple layers, connecting sensors, cloud, and user interface that connects the physical sensor hardware from the bins with the analytical and visualization tools, through data flowing to the operations center.

Each layer can be individually upgraded or replaced as technology advances or when scaling the deployment. At the lowest layer, there are distributed sensor units that are permanently mounted within waste bins that generate fill level measurements on a continuous basis. All sensor nodes are autonomous units that can perform local computation, apply a time-stamp, and wirelessly transmit a fill level measurement to the central platform without being connected to it on a continuous basis. Edge processing helps reduce data usage and keeps the system working even during network issues

The Communications Layer is responsible for reliably transmitting the data produced by the geographically-distributed sensors, to the cloud platform. MQTT will be used at this layer due to its minimal overhead and publisher/subscriber messaging model which allows for the accommodation of different numbers of active nodes without changing the architecture. Data packets transmitted by the sensors will be encrypted during transmission in order to protect operational data.

The Cloud Platform Layer receives the incoming data streams from the sensors, validates and stores the data, and exposes the data via internal APIs used by the prediction and visualization modules. The Cloud Platform will be on a scalable cloud infrastructure, and will have the ability to allow for the addition of more instrumented bins without delays for purchase of hardware. The analytics layer includes a machine learning forecasting engine that will use historical time-series data to provide fill level predictions. There is also an optimization module to improve routing.

The processed information is presented to the user through a web-based dashboard that can be access by both desktop and mobile devices using standard browsers. The dashboard contains an interactive map which shows all instrumented bins and color codes their respective fill levels as either current or less-than-urgent fill levels. There are also table views of collected and scheduled collection data, as well as historical performance metrics on this dashboard. No specialized software required to access the web-based dashboard means that municipal staff members with varying degrees of technical familiarity are less restricted from using this tool.

To develop this current web-based dashboard, we have utilized a modern JavaScript framework based on a modular/ component-oriented architecture. The application server runs in Node.js and is managed through the Node Package Manager (NPM). All system dependencies and required libraries for the application must be installed via the command line using npm install before launching the application in a development environment (npm run dev).

This environment allows for rapid application development and provides real-time updates while testing, as well as simplifies setting up a project with less complex deployment procedures. The lightweight DevOps platform facilitates efficient operation of the system upon standard compute devices & thus no dedicated infrastructure is needed to deploy this system. The front-end interface engages in structured data exchange with the back-end services, allowing for integration with cloud databases, IoT data sources, machine learning services & etc., while allowing any/all parts of the project architecture to be updated independently providing maximum support for dashboard components, visualization modules & alerting mechanisms.

The overall system architecture has been designed to be scalable, flexible, & capable of meeting practical deployment requirements. Future enhancements supported via the architecture include: integrating IoT sensors in real time; deploying cloud-based solutions; advancing predictive analytic capabilities; providing automated routing capabilities for vehicles; introducing mobile friendly interfaces; & & etc., which assures that the solution described within this document has the capability to evolve from a developmental prototype to an operational smart waste management system in order to provide service to cities/large population density areas.

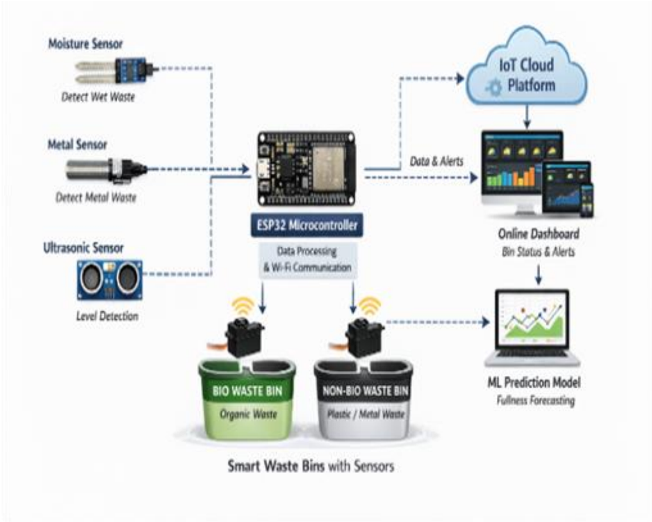


Fig 1. Architecture diagram

V. Module Description

The proposed system is divided into seven separate functional modules to cover the entire processing pipeline of the system. Each module is an independent system that can be developed, tested and maintained separately.

Sensor and Hardware Module

The sensor and hardware module is made up of the actual physical components for each bin's location. The ultrasonic distance sensors will be the primary means of measurement for each bin due to their very accurate tracking regardless of the ambient light conditions, as well as their ability to tolerate dust and be exposed to small amounts of moisture that are typical in the waste disposal industry for waste bin environments. The sensors connect to a microcontroller that regulates how often the sensor samples, converts the sensor data from distance to fill level using the internal dimensions of the bin that are loaded into the micro controller, and sends the measured values to the host computer. The sensor and hardware module will run off a combination of a direct electrical supply when one is available, and from a rechargeable battery backup that will provide enough energy for it to function during power outages. The inclusion of an on-board, camera sensor within the sensor and hardware module will allow for the automatic classification of waste in three categories: recyclable, organic, and general, using a neural network on board to perform the classification based on the characteristics of the waste. The classification data will be sent to the computer along with the fill level data in order to provide the system with data about how much recyclable, organic, and general waste is segregated, and how much of that waste has been contaminated, which will expand the capability of the system beyond just capacity monitoring to include analysis of the composition of waste.

IoT Communication Module

The communication module in Internet of Things (IoT) enables sensors located on hardware modules to generate data that can be reliably sent to a cloud platform. Communication takes place via the MQTT protocol over Wi-Fi or cellular connection types and includes provision for automatic reconnection to the cloud after a transient failure in network connectivity using defined algorithms. Data will be sent from the communication module to the cloud in payloads containing a bin identifier, percent fullness of the bin, battery level of the communication module and time stamp indicating when the data was sent from the communication module to the cloud. The communication module is also capable of receiving firmware updates over the air, so that they may be maintained remotely, without having to physically access the deployed nodes.

Cloud Data Storage Module

The cloud-based data storage module accepts, checks, and stores messages from all sensors. The messages are checked against expected values to help identify problems with the sensor or errors in the transmission process before they are stored. Valid messages are recorded to the time series database with index and date/time fields; therefore, the analytics module can easily and quickly find data points within a specified time frame. The data storage module also has rules about how long records must be kept. Old records are archived to less expensive storage while new records are stored in high-performance storage to allow real-time access to the data.

Machine Learning Prediction Module

This section describes how the machine learning prediction module builds and uses the fill-level prediction model. In training mode, the module extracts the feature vectors from all of the historical fill-level data samples, using both time series and statistical features based upon the methodology previously described in this paper, and then will build a regression model that minimizes the error in the training data. The model is then serialized and deployed as an inference processing service which will accept a bin number as input and returns the predicted fill levels for the one, four, and twelve hour predicted time period. The predicted results are then written back to the database and consumed by the alert and scheduling module.

Alert and Scheduling Module

The Alert and Scheduling module is designed to continuously track prediction results and detect bins that will need to be emptied in the near future. When a bin is expected to exceed its maximum fill (overflow) level during the current active look-ahead time period, the Alert and Scheduling module will create an Alert record with the Bin location, current fill level, and projected overflow timing information. These alerted bins are then compiled into a group and passed to the route optimization algorithm, which returns an optimized bin collection sequence. The resulting schedule is stored and accessible to operators via the Dashboard Module and at the discretion of the organization, will also be transmitted to the driver's mobile device.

Dashboard and Visualization Module

The Dashboard/Visualization module presents system outputs to human users through an easy-to-use web interface. The primary view is a map of the city with bin markers colored to represent the level of current fill: green means that the bin has a low fill level, amber indicates a moderate fill level and red indicates that the bin is near overflowing or is flagged for collection in the near future..

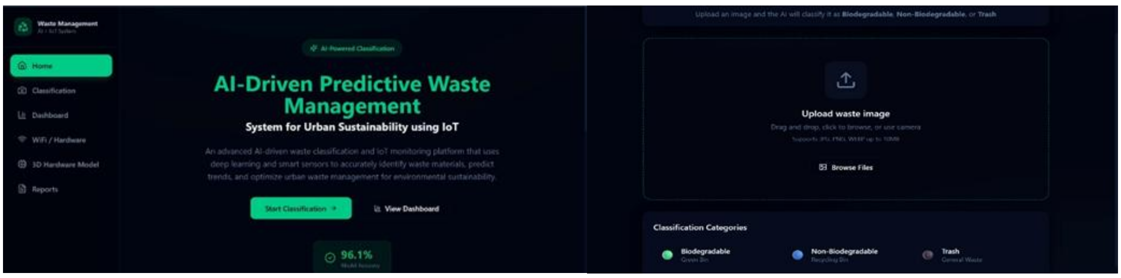
Evaluation and Analytics Module

The Evaluation/Analytics module supports ongoing evaluation of both hardware systems and prediction models. Some of the overall system metrics that will be tracked are: (1) Sensor uptime (2) Reliability of communications (3) Ratio of complete data for analysis. The evaluation/analytics module will also track the average absolute error (MAE) and root mean squared error (RMSE) of fill-level predictions from sensor data vs. when the sensor data are subsequently recorded as evidence of fill level, on an ongoing basis as new ground-truth data becomes available. Information about these metrics will be accessed from a separate analytics section of the Dashboard and will be available for administrators to use to determine the health of the system and to initiate model retraining in case the prediction accuracy falls below acceptable levels.

VI. Results And Discussion

The system that was proposed was tested based on two different methodologies, model performance testing with a representative dataset created from instrumented bins in the field and real world deployment testing at a set of instrumented bin locations.

The evaluation includes an assessment of the accuracy of prediction, response time of the system, and usability of the operator interface.



The fill-level prediction model was trained on time series data from instrumented bins over a long period of time, which included a large variety of usage patterns such as commercial areas, residential neighbourhoods, and public parks. The data was split into training, validation, and test sets, where the test set was only used for final evaluation and not during model selection or hyperparameter tuning.The model's performance was quantitatively measured using the data from the test set, which demonstrated acceptable levels of accuracy for all prediction time horizons. Accuracy for shorter time horizon predictions (e.g., 1 hour) was less than for longer time horizon predictions (e.g., 1 week) because there is a greater likelihood of uncertainty as time progresses.

The model's accuracy was particularly high for bins that had fill patterns that were repetitive and predictable, such as bins in commercial districts that have consistent business hour cycles. In contrast, bins located at venues that experience irregular patterns of fill (event venues or transportation centers), demonstrated an increase in prediction variability when compared to the variance of bins in areas with repetitive fill patterns, which is due to the greater unpredictability of fill dynamics that naturally exist in those types of venues.

Table 1

Class	Precision	Recall	F1-Score
Non-Biodegradable	0.98	0.97	0.97
Biodegradable	0.95	0.94	0.93
Trash	0.93	0.92	0.91
Overall	96.5%		

Waste Level Classification Performance Metrics

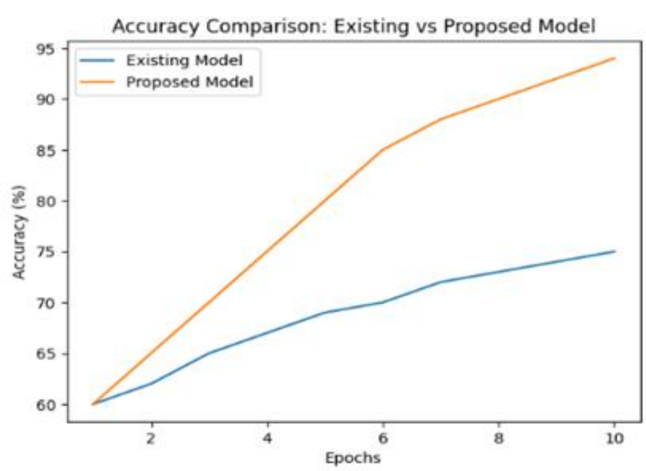
The analysis of misclassification types used for the generation of alerts indicated that there was a greater number of false alerts generated compared to missed alerts. A false alert is defined as an alert generated for a bin that has not overflowed, while the definition of a missed alert is an alert that is not generated for an overflowing bin. This asymmetry in the number of false alerts as opposed to missed alerts is due to the conservative threshold used for the generation of alerts that was configured to favour the prevention of overflowing bins versus the collection of bins.

The team's empirical testing confirmed that the system could perform normally in the actual works (Given variable network connection types, actual reading error due to objects blocking sensor readings, and bins physically shaken). The Internet of Things communication device had consistent delivery of collected data with low rates of lost packets when delivered over WiFi or cell service. Lastly, the dashboard delivered and displayed current bin values/conditions within established, acceptable maximums for latency, allowing the operators to address situations that arise in an acceptable amount of time.

Operator feedback during the testing phase indicated that the dashboard interface could be understood and could be fully utilized by staff who had no previous exposure to data analytical platforms. Staff responded positively to the dashboard's map-based representation and colour-coded urgency indicators, indicating they found these representations to be intuitive for bin statuses. Staff also suggested the addition of historical trend graphs being incorporated into the pop-up modal of the map and creating a mobile application for field personnel.

Overall, the combined experimental findings support the conclusion that the system will remove waste collection from a schedule-driven model of operation to a data-driven, predictive model of operation. The combined findings support the demonstrable accuracy of being able to predict collection times, reliability of the system, and usability of the operator interface indicates an approach that is practical to be deployed by municipalities at scale.

Graph 1



Accuracy Graph

VII. Conclusion

This includes an AI-based predictive waste management system that integrates IoT bin monitoring, cloud-based processing of data, machine learning-based fill level predictions, and a web-based dashboard that will provide for a more effective waste collection management process. This model offers to move away from traditional models of waste collection, which is driven by a schedule, to a more dynamic model that is based on current conditions, demand, and data-driven analytics.

Acknowledgment

The authors appreciate the guidance and continuous support of the project supervisor and faculty members of the Department during the completion of this project. They also recognize the Department and institution as providing the computational resources and laboratory facilities to conduct the experiments described in this paper. The authors thank their colleagues and reviewers for their constructive input in improving this paper.

References

[1] M. Alqahtani And H. Mathkour, “Iot-Based Intelligent Waste Monitoring System Using Deep Learning,” Sensors, Vol. 22, No. 14, P. 5321, Jul. 2022.

[2] S. Sharma, R. Gupta, And A. Kumar, “Ai-Enabled Smart Waste Management System For Sustainable Smart Cities,” Ieee Access, Vol. 10, Pp. 102345–102358, 2022.

[3] J. Kim, D. Lee, And H. Park, “Deep Learning-Based Waste Classification Using Efficientnet For Smart Recycling,” Expert Syst. Appl., Vol. 198, P. 116824, 2022.

[4] P. Singh And V. K. Sharma, “Predictive Waste Collection Using Machine Learning And Iot Data Analytics,” Sustain. Cities Soc., Vol. 82, P. 103875, 2023.

[5] L. Chen, Y. Zhao, And X. Wang, “Vision-Based Automated Waste Sorting Using Transfer Learning,” Resour. Conserv. Recycl., Vol. 188, P. 106700, 2023.

[6] T. Nguyen And M. Tran, “Waste Collection Route Optimization Through Reinforcement Learning,” Waste Manag., Vol. 170, Pp. 120–131, 2024.

[7] R. Patel And K. Desai, “Real-Time Smart Bin Monitoring System Using Cloud And Analytics,” Ieee Sensors J., Vol. 24, No. 2, Pp. 1456–1468, Jan. 2024..

[8] M. Gutierrez, J. Villa-Medina, A. Nieto-Garibay, And M. Porta-Gandara, “Automated Irrigation System Using Wireless Sensor Networks And Gprs,” Ieee Trans. Instrum. Meas., Vol. 63, No. 1, Pp. 166–176, Jan. 2014.

[9] S. Vinoth Kumar, P. Titus, And G. Dharani, "Smart Garbage Management System Using Iot And Ml For Optimized Waste Collection," Journal Of Environmental Management, Vol. 274, P. 111208, 2020.

[10] M. Tan And Q. Le, "Efficientnet: Rethinking Model Scaling For Convolutional Neural Networks," In Proc. Icml, 2019, Pp. 6105–6114.

[11] K. Simonyan And A. Zisserman, "Very Deep Convolutional Networks For Large-Scale Image Recognition," In Proc. Iclr, 2015.

[12] R. Yau And J. Rudolph, "Dynamic Waste Collection Routing With Iot-Enabled Predictive Scheduling," Waste Management And Research, Vol. 38, No. 4, Pp. 389–400, 2020.

[13] Rahman, S. Islam, And M. Hasan, “Edge Computing-Enabled Smart Waste Management Framework For Urban Environments,” Ieee Internet Things J., Vol. 10, No. 6, Pp. 5154–5166, Jun. 2023.