Garbage Bin Classification Using Convolutional Neural Networks with Clean and Noisy Bin Images

Raj Kumar Sharma [1[0000-0002-5612-1616]], Mukta Sharma[2[0000-0003-4400-8568]]

[1]KK MODI UNIVERSITY, Mahmara, Jalbandha Road, Durg, Chhattisgarh-491001

[2]Trinity Institute of Professional Studies, Dwarka Sector 9, Institutional Area, New Delhi-110075

[1]profrksharma07@gmail.com

[2]drsharmamukta@gmail.com

Abstract

This research focuses on Garbage Bin Classification. It investigates the efficacy of Convolutional Neural Networks (CNN) when employed on 2D garbage bin images belonging to both clean and noisy datasets. The study encompasses the collection of a primary dataset comprising 2D images, subsequently introducing four distinct types of noise, such as the Gaussian, Salt & Pepper, Poisson, and Uniform. The evaluation entails the application of various prominent CNN architectures, including AlexNet, VGG16, a four-layer CNN (CNN4), and a five-layer CNN (CNN5), to both datasets. Concurrently, during model training, loss and accuracy graphs are generated to assess model performance. The CNN5 model performed exceptionally well when evaluated using the standard dataset, showcasing an impeccable classification performance. Across various key metrics, the model's ability to accurately classify instances within the Empty and Full classes has been resoundingly demonstrated. An accuracy score of 1.0 attests to the model's error-free predictions, while a perfect R2 Score of 1.0 signifies an exact match between the model's predictions and observed data. Additionally, scores of 1.0 for Precision, Recall, and F1 stress the model's thorough and accurate optimistic predictions. The CNN4 model with noisy data showcases a commendable R2-score of 0.85, often indicative of a robust fit for various applications. This implies that the model can provide substantive insights into the target variable, augmenting prediction accuracy and comprehension of variable relationships. Ultimately, the findings advocate the suitability of the CNN5 model for clean datasets while endorsing the CNN4 model as an optimal choice for handling noisy datasets, hence proposing these models to identify the garbage bin class in building smart cities' garbage management systems.

Keywords: AlexNet; VGG16; CNN; Garbage Bin; Image Classification; Smart City

Introduction

The global garbage crisis requires urgent attention and collaboration from individuals, communities, industries, and governments. A holistic approach involving city authorities, waste management agencies, residents, and businesses can tackle overflowing bins, promote sustainable waste management, and create healthier urban environments. The East Asia and Pacific region accounts for 23% of global garbage production, according to estimates by the World Bank. In contrast, only 6% of all garbage is produced in the Middle East and North Africa. At least 33% of the 2.01 billion tonnes of municipal solid garbage produced annually worldwide must be managed environmentally sound. On average, each person worldwide generates 0.74 kilogrammes of garbage daily, although this figure varies between 0.11 and 4.54 kilogrammes. Surprisingly, high-income nations account for 34% of global garbage production despite having just 16% of the world's population. By 2050, it is anticipated that there will be 3.40 billion tonnes of waste generated worldwide, which is a significant increase over the estimated growth in the human population during that time. It is worth noting that waste production tends to increase along with rising income levels.

It's important to note that WHO's recommendations and statements on the garbage problem might have evolved or been updated since my last knowledge update in September 2021 (for more information, visit the official World Health Organization website). Illustrates the significant shift expected in trash production by 2050, with low-income nations projected to experience a threefold increase compared to high-income countries. The Middle East and North Africa (ME&N), Sub-Saharan Africa (SSA), Latin America and the Caribbean (LA&C), North America (NA), South Asia (SA), Europe and Central Asia (E&CA), and East Asia and the Pacific (EA&P) are only a few of the places where waste generation is very common (see Figure 1). Unfortunately, over half of the waste generated in these areas is unlawfully disposed of. The projected surge in waste poses severe threats to the environment, public health, and economic growth, necessitating immediate and drastic actions to combat this issue.

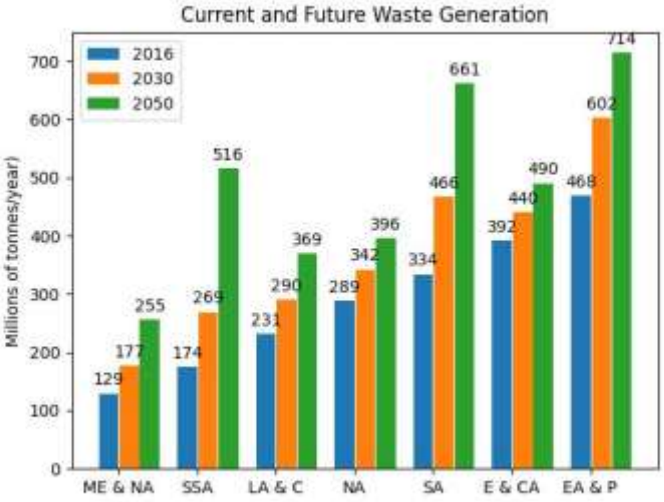


***Figure 1**. Forecasted regional garbage production (in metric tonnes per year)*

Figure 2 reveals that the amount of waste varies according to a country's income level due to distinct consumption patterns. More dry recyclable garbage (such as plastic, paper, cardboard, metal, and glass) is produced in high-income nations, whereas less food and green waste (32%) are made in low-income countries. As economic development levels decline, organic waste increases, reaching 53% in middle-income countries and 57% in low-income countries. It is important to note that in low-income nations, recyclable materials only account for 20% of the waste stream. In addition to income differences, there are slight variations in waste streams across different regions. Everywhere else save Europe, Central Asia, and North America produces at least half as much organic garbage as those three regions combined.

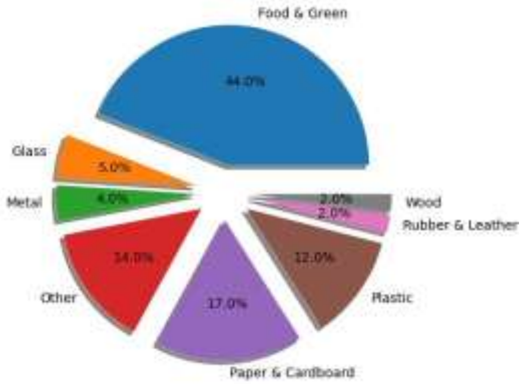


***Figure 2**. Global waste composition*

There are a variety of waste management systems on the market today. Several different types of sensors, including ultrasonic, infrared, and weight sensors, can be placed on trash cans to keep tabs on how full they are. Waste management agencies can better plan pickup routes and times using the sensors' real-time information. Radio-frequency identification (RFID) tags on trash cans monitor efficiency and location, enabling trucks to check bin locations and collection statuses. A Global Positioning System (GPS) enabled trash can and truck tracking devices to improve management. This allows for optimised container route planning and real-time monitoring. Trash cans can be remotely monitored using IoT (Internet of Things) solutions by utilising a network of sensors, data analytics, and connectivity. These apparatuses report relevant data to organisations responsible for trash management, including fill levels, temperatures, and humidity. Only some cities or regions will have access to the same monitoring tools due to differences in waste management infrastructure and technological advancements.

Literature Review

In the research [1], a garbage sorting system is built using convolutional neural networks trained with data from a deep learning dataset. The system uses image and object detection algorithms to categorise and recognise distinct waste types properly. Algorithms like ResNet, MobileNetV2, and several YOLOv5 models for trash object detection are then trained and tested on the garbage classification data. Combining the findings of the studies conducted on these methods, the consensus voting algorithm boosts the recognition rate for picture classification by 2%. The practical deployment of the system has proved to be a substantial improvement in the identification rate for trash picture categorisation, reaching over 98%. Additionally, the system has been successfully transferred to a Raspberry Pi microprocessor, with the expected results.

The research [2] aims to develop reliable methods for estimating and forecasting DW volumes using machine learning (ML) techniques. This paper analyses time series using various machine learning algorithms. The investigation includes two case studies. In this first case study, we will gather information on household garbage from the Saudi Arabia and Bahrain governments from 2010 through 2021. The second case study involves a 30-day investigation into the garbage habits of a large family of eleven. The family produced 1.7–7.9 kg of biodegradable garbage and 0.0–2.0 kg of non-biodegradable waste. The study shows encouraging results using time series analysis and carefully choosing input predictors. Validation and evaluation of trained models using residuals, mean square error, root mean square error, and R2-Score, ensuring reliable waste quantities. ML algorithms show promise in developing smart waste management systems, reducing environmental, economic, and societal impacts.

The study presents a system in which all trash cans are equipped with cameras and motion detectors. At regular intervals, [3] these devices take 2D pictures of the trash. A multi-layer Neural Network processes the collected 2D photos of the rubbish level within the bins. An ideal solution can be found by identifying the best iteration level and classifying the bins as empty or filled. The major goal of this study is to facilitate prompt interventions for trash can emptying, leading to a cleaner, safer community. Using computer vision, the Internet of Things (IoT), and machine learning (ML) techniques like regression, classification, clustering, and correlation rules based on the interpretation of solid waste photos, we present a tried-and-true approach to managing municipal solid waste (MSW) in the literature[4]. ResNet V2 is a set of image classification models developed from the ground up and trained via transfer learning. ResNet V2 allows deep neural networks to be trained on large datasets, which increases accuracy and decreases the error rate while mapping identities. We use batch normalisation and mixed hybrid pooling to strengthen CNN further and get top-tier results. The proposed model can distinguish between biodegradable and non-biodegradable garbage, allowing for more precise collection. The success of the proposed model can also be ensured by keeping an eye on performance measures like accuracy and loss. Compared to state-of-the-art models, the proposed ResNet-based CNN achieves 19.08% better performance with 34.99% lower loss during garbage classification.

The Internet of Things (IoT) [5] ushers in a new era of urban development by making garbage collection a model of efficiency and sustainability. Waste management is vital for protecting the environment and human health as industrialisation and urbanisation speed up. Existing solutions prioritise efficient monitoring at the expense of real-time functionality. In this research, we suggest a system that uses Internet of Things-based automation to facilitate smart trash management. The proposed approach employs machine learning classification techniques to track the volume of garbage, the presence of hazardous materials, and the quality of air throughout a metropolis. The haphazard forest method outperforms other classification approaches, with 92.15% accuracy at a runtime of 0.2 ms.

The research suggests a deep learning architecture-based hardware solution for waste segregation [6]. The suggested deep learning-based hardware solution SmartBin uses image classification using a Convolutional Neural Network System Architecture running on a Real-time embedded system to sort rubbish into biodegradable and non-biodegradable categories. Image classification for garbage detection seeks to speedily classify items in the trash can. In contrast to object detection of a particular entity, when photographs of things of that entity share some common qualities and traits, rubbish might be of any dimension, shape, colour, or texture, making this a challenging operation. The goal of this study is to analyse the effectiveness of hardware components (PiCam, raspberry pi, infrared sensors, etc.) used for garbage detection in the bin and to compare their performance to that of various pre-trained Convolution Neural Networks, including AlexNet, ResNet, VGG-16, and InceptionNet. The proposed model's most significant performance indicator was the InceptionNet Neural Network, which achieved an accuracy of 98.15 per cent and a loss of 0.10 on the training set and 96.2 per cent and 0.13 on the validation set, respectively.

The research [7] provides an Internet of Things (IoT)-based smart city solid waste management solution to address deficiencies of current approaches. Public Bin Level Monitoring Units (PBLMUs) and Home Bin Level Monitoring Units (HBLMUs) are two end sensor nodes incorporated into the suggested design. These nodes track the accessibility of both public and private trash cans. Fill levels and tank locations are transmitted to a command centre via the PBLMUs and HBLMUs. An easy-to-use graphical user interface (GUI) allows waste management specialists to see how complete each garbage can is. Multiple large-scale tests were carried out to demonstrate the proposed system architecture's effectiveness. PBLMUs were installed in the first eight bins and linked to a LoRa WAN network; HBLMUs were deployed in the second eight bins and connected to a Wi-Fi network. Users piled trash into the bins to varying heights, and the intuitive UI kept tabs on their progress. Second, we measured the sleep current and active current contributions of a PBLMU in a controlled experiment to derive its average current consumption. Finally, it was determined that a PBLMU would have a lifespan of roughly 70 days under ideal conditions. Smart towns could benefit from an IoT-based garbage management system with real-time monitoring and an easy-to-use graphical user interface.

There are several obstacles to effectively managing electronic trash (E-waste) [8], such as needing more technical expertise, inadequate infrastructure, inadequate financial assistance, and a lack of community involvement. Improper E-waste recycling practices harm human health and the environment, and this article explores initiatives to address these issues in depth. To correctly manage E-waste, it is essential to create a registry of obsolete electronic items, which can be aided by introducing recycling policies that minimise negative environmental impacts. One strategy has been offered to help both rich and developing nations better manage their electronic trash. It is possible to reduce adverse effects while promoting a sustainable and resilient ecosystem if developing countries prioritise implementing systematic management practices for E-waste and embrace best practices.

The research [9] proposes a trash monitoring system with an integrated waste sorting mechanism, providing a straightforward and inexpensive means of categorising and cataloguing garbage generated at home. The plan aims to separate metal, plastic, and glass from regular trash. After sorting, the trash generated by each household is reported to a central database. Adopting this approach can contribute to a more pristine environmental ecosystem and more effective waste management procedures.

To facilitate more efficient waste management, recyclables are separated into the following six categories: plastic, glass, paper/cardboard, metal, fabric, and other recyclables [10]. Deep-learning convolutional neural networks (CNNs) were employed for efficient garbage classification. This research looks at seven cutting-edge CNN models and various data preprocessing techniques for trash sorting. For all nine classes in the validation set, these models achieve an accuracy of between 91.9% and 94.6%. MobileNetV3 stands out due to its low storage requirements (49.5 MB), quick processing time (261.7 ms), and high classification accuracy (94.26 per cent).

The goal is to design a waste management infrastructure that is both efficient and sustainable, with full compatibility with other municipal databases. The study [11] demonstrates the usefulness of the data collected by sensor nodes and the system's effectiveness, giving municipal administrators new perspectives. One major takeaway from this research is that a low-overhead, innovative city app can aid municipal administrations in their daily tasks. In addition, the feasibility of establishing a low-power wide-area network in Istanbul, Turkey, will be studied in detail. Finding a reliable alternative to cellular networks for usage in smart cities is the primary motivation behind this study.

The article [12], aims to improve our daily lives by aiding garbage disposal. Opportunities for remote device management are plentiful in today's age of pervasive mobile connectivity. Waste disposal in rural areas, urban centres, and everywhere in between has become a significant problem that affects people's health and spreads disease. As a result, it's more important than ever to use technology to improve garbage collection. Manual monitoring and disposal operations are commonplace in the current waste management system, which can be inefficient and time-consuming. The Internet of Things enables easier monitoring of garbage cans, supporting conventional methods.

The paper [13] focuses on designing, implementing, and validating a dynamic weighing system created to measure municipal waste as bins are dumped. The paper introduces and explores two different methods: the amplitude technique, which measures the most significant change in amplitude of the signal, and the period approach, which measures the time it takes for the lifting device to oscillate until it is damped. Because of its small size, the system can be mounted directly on the lifting apparatus and fitted with FBG sensors. Within the 10-100 kilogramme range for conventional containers used for municipal waste, the presented system showed high accuracy during four months of pilot testing. The maximum variance was 4.04 kg. The technique did not rely on garbage trucks. Therefore, it could be easily incorporated into existing practices. The paper develops garbage collection systems using Arduino UNO, ultrasonic, and moisture sensors [14]. Using image processing, we can accurately calculate a landfill's waste index. In addition, a hardware prototype has been created to show how the suggested framework might be implemented. This strategy will result in clean, pollution-free cities due to improved waste management.

The investigation [15] explores the idea of a smart trash can whose capacity may be updated in real time via a smartphone app. Both the garbage collector and the user will be able to monitor the contents of the can thanks to the two-way communication between the can and the user's phone. When the garbage collector's bin gets too full, they can easily handle the surplus waste by shifting it to a larger landfill container. This trash can has been put through rigorous testing, showing that it consistently performs as advertised. It quickly senses the presence or absence of items and reacts by opening or shutting its cover accordingly. In addition, it communicates the garbage can's capacity to the user's mobile device in a flash, taking only 0.45 to 0.47 seconds.

Using the LoRa protocol for wireless data transmission and a deep learning model built using TensorFlow, this study aspires to create an intelligent garbage collection and disposal system [16]. The sensor data is transmitted over LoRa, and Tensorflow recognises and categorises objects in real time. The servo motors precisely control the bin's various metal, plastic, paper, and general trash compartments. Using a pre-trained object detection model, the TensorFlow framework allows us to quickly and accurately recognise and label objects as trash. With a Raspberry Pi 3 Model B+ as its brain, this object identification model trains on junk photographs to provide a static inference graph for camera-based object recognition. Each trash can has an ultrasonic sensor built in so that the fullness may be detected at a glance. The built-in GPS module may track the bin's location and status in real time. Information regarding the bin's background, current situation, and capacity is transmitted in real time over the LoRa protocol. The in-built RFID module enables easy identification of trash management workers.

The study [17] investigates the feasibility of simultaneously keeping tabs on many trash cans in different communities. Each trash can has sensors that track the amount of waste, metal, and hazardous gases. Trash collection and disposal predictions were evaluated using support vector machines, neural networks, random forests, decision trees, and K-nearest neighbours. According to the findings, the RF algorithm performed best, with an overall alert message prediction accuracy of 85.29 percent. Overall, the study will affect the reduction of pollution in "smart cities" and encourage "green technologies." This study aids in the development of ecologically friendly practices by optimising waste management.

It has become clear after carefully examining the available research publications that. These studies use deep learning models for trash identification, such as ResNet, MobileNetV2, YOLOv5, and InceptionNet. Several research has demonstrated IoT-based garbage monitoring systems. However, these have only been subjected to limited testing periods, and evidence also suggests that sorting garbage into biodegradable and non-biodegradable components is helpful. Studies have focused on the technology but a fuller understanding of the benefits and drawbacks of innovative waste management systems. Understanding how to make a smooth transition while using the benefits of both traditional and modern waste management approaches can be gained by integrating new intelligent waste management technology with existing garbage collection and disposal systems. Several research employs various data-transfer protocols (like LoRa). Selecting optimal communication methods could be aided by a thorough comparison of alternative protocols' performance, range, energy economy, and interference potential in realistic urban environments.

The research gaps in the existing literature focus on promising avenues for more investigation to advance smart waste management and deliver more efficient and long-lasting solutions. Models' robustness and generalizability across various real-world scenarios, lighting conditions, and waste types could benefit from more investigation. More high-quality investigation into the most precise and time-saving approaches for differentiating different kinds of garbage bins using machine learning models would be a great boon to efforts to lessen trash. This study investigates a novel class of CNN models for use in a Full or Empty garbage bin detection system that uses images as input for the models. It provides a workable solution for dealing with such situations.

Methodology

Observing Figure 3, this study captures images of garbage bins and sorts them into two categories: ones with garbage in them and ones without. Images taken from the gathered dataset are processed to provide enhanced images with noise. Two data sets have been transformed into a matrix that performs computations with machine learning models. The study analyses the results and behaviour of the datasets (clean and noisy) with the AlexNet, VGG16, CNN4 (CNN with four layers, each layer has a set of two convolutional layers), and CNN5 (CNN with five convolutional layers) models for the study of garbage bin status (Empty or Full). The models are trained with the preprocesses dataset, and the accuracy & loss of training models are visualised with line graphs. The testing dataset creates a confusion matrix, and results (accuracy, precision, recall, F1-Score and ROC area) are computed and analysed. Based on the result analysis, a suitable model/s will be proposed.

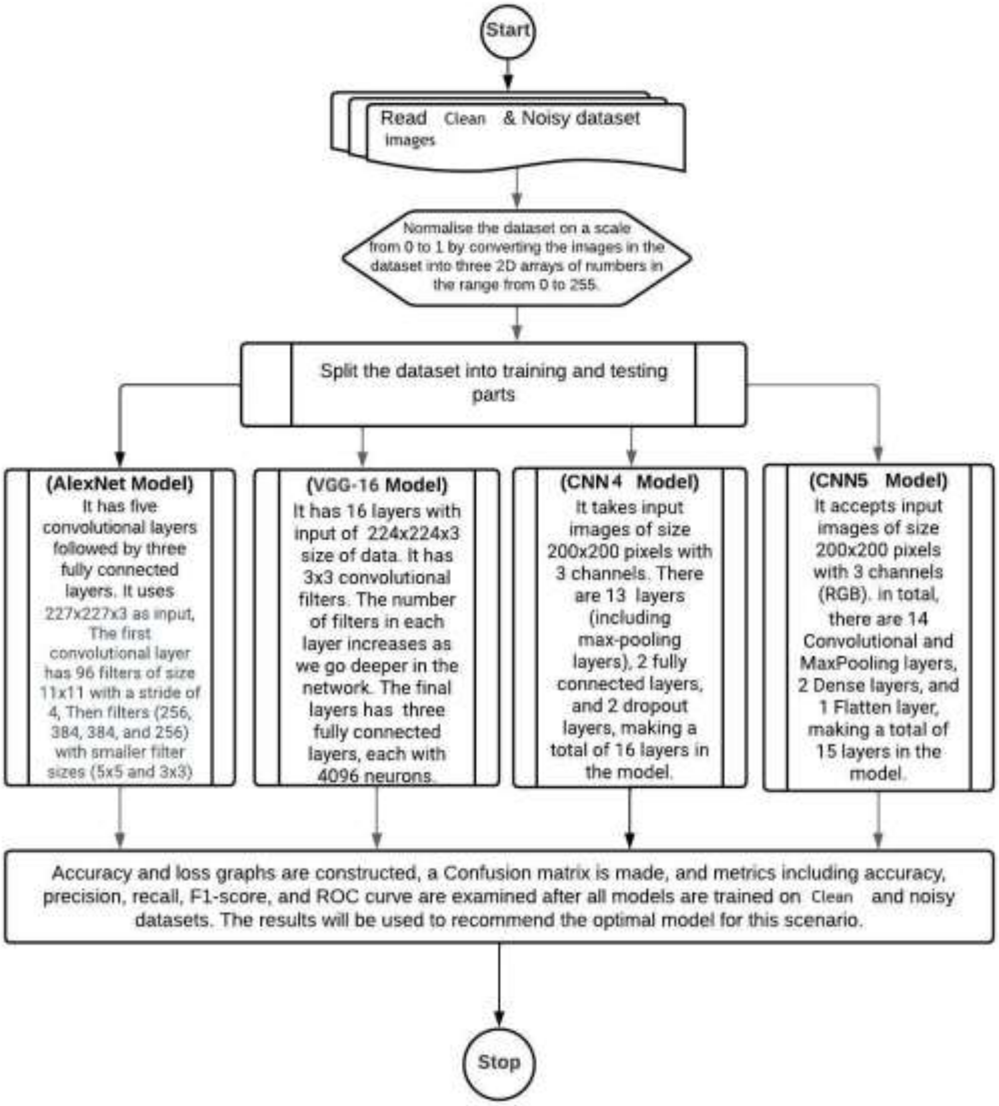


***Figure 3**. Proposed Methodology*

Data Preprocessing

The core dataset used in this study comprised images of garbage bins taken in the experiment lab. Each record in the collection is either "Empty Bin" or "Full Bin." Python libraries create noisy dataset images with Gaussian, Salt & Pepper, Poisson, and Uniform noise. The RGB dataset images are converted into a NumPy array of integers. The array of images standardises the input data to a range between 0 and 1. See Eq. (1).

$$x_{normalise} = \frac{x - \mu}{\sigma} \tag{1}$$

where μ is mean and σ is standard deviation.

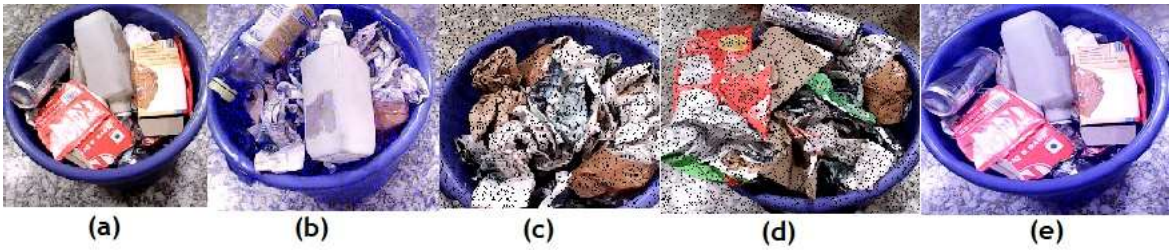


Figure 4. Bin images (a) Clean (b) Gaussian noise (c) Salt & pepper noise (d) Poisson noise (e) Uniform noise

In this analysis, images are formed from the standard dataset (see Fig. 4). To add Gaussian noise; an image is generated by randomly selecting numbers from a normal distribution with a fixed mean and standard deviation. The Gaussian distribution [18] can be described by its mean and standard deviation, see Eq. (2). These are combined with the actual pixel values in the image. The calculation is represented mathematically by Eq. (3).

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}} \quad (2)$$

$$I_{noise} = I + N(0, \sigma^2) \quad (3)$$

Where "x" is the random variable (the intensity of a pixel), is the distribution's mean (often 0 for adding noise), and is the standard deviation that determines how far or how intense the noise is dispersed. Where "I" is the clean original image, "I_noisy" is the noisy version, and $N(0, \sigma^2)$ is a Gaussian random sample with zero mean and two standard deviations.

Impulse noise [19], commonly known as salt and pepper noise, corrupts pixels in an image randomly, changing them to either the highest (salt) or lowest (pepper) value within the permissible pixel intensity range. This noise is frequently employed to mimic image sensor flaws such as pixel dropout. Two parameters, P_0 and P_1 govern the overall salt and pepper noise level. These values characterise how likely a salt and pepper noise will influence a given pixel. Randomly selecting pixels and assigning them the maximum intensity (salt) or the least intensity (pepper) depending on the stated probabilities is how salt and pepper noise is introduced into an image. Here, P_0 and P_1 determine how likely it is for salt and pepper noise to appear in each pixel.

Poisson noise [20] is a form of statistical noise that manifests itself over a wide range of phenomena, and it is most commonly connected with the occurrence of uniformly distributed random events throughout time and/or space. This paper's noisy dataset is the result of adding images of garbage bins to it using Poisson noise. As seen in Eq. (4), the Poisson distribution is defined.

$$P(X = k) = \frac{(\lambda^{k} * e^{-\lambda})}{k!} \quad (4)$$

The likelihood of seeing k occurrences during a certain time frame is denoted by $P(X + k)$. The λ denotes the average number of occurrences per given time or distance. The number e , named after Euler, is roughly equivalent to 2.71828. The nonnegative integer k denotes the number of events. In a Poisson distribution, the variance is the same as the mean. $Var(X) = \lambda, and\ so\ does\ E(X).$

Uniform noise [21] is used in garbage bin images to create a noisy dataset. Generating a random and unbiased distribution of pixel values is especially useful. Mathematically, the probability density function of a uniform distribution is defined in Eqs. (5, 6).

$$f(x) = \frac{1}{b-a}\ where\ a \le x \le b \quad (5)$$

$$f(x) = 0\ for\ x < a\ or\ x > b \quad (6)$$

Where "a & b" represent the distribution's minimum and maximum values. The probability density function at location "x" is denoted by the function $f(x)$. The mean $\mu = \frac{(a+b)}{2}$ of a normal distribution between two given limits (a, b). The variance of uniform distribution is $\sigma^2 = \frac{(b-a)^2}{12}$.

Proposed Models

Whether garbage bins are empty or full is the topic of this study. To this end, it is suggested that models for determining whether or not the bin is full be developed using Convolutional Neural Networks (CNN) [22]. The convolution operation lies at the heart of any convolutional neural network. It includes performing element-wise multiplications followed by a sum at each step as a small filter (called a kernel) is slid over the input garbage bin image. see Eq. (7). This method helps recognise patterns by extracting them from the input garbage can image. To examine the models utilising the input image and learn complex patterns while avoiding the vanishing gradient problem, this research makes use of the Rectified Linear Unit (ReLU), defined as f(x) = max(0, x). Max pooling is performed on pooling layers to reduce the number of sampled spatial points in the feature maps without losing information. This helps avoid overfitting and reduces computing complexity.

Fully connected layers follow convolutional and pooling layers; see Eq.(8) before the output is flattened into a one-dimensional vector. These layers are analogous to those found in a classic feedforward neural network and are responsible for acquiring knowledge via complex feature combinations. The Cross-Entropy Loss function quantifies the distinction between the model's predictions and the actual target labels, see Eq. (9). Backpropagation is used to train the CNN, and its adjustments to the network weights depend on the loss function's gradients that have been determined for the model parameters. During training, the weights are updated using an optimisation process called Stochastic Gradient Descent (SGD) or Adam, see Eq. (10). In this investigation, we examine the AlexNet, VGG16, CNN4 and CNN5.

$$Output_{size} = \left(\frac{(Input_{size} - Filter_{size} + 2 * Padding)}{Stride}\right) + 1 \tag{7}$$

$$Pooling\ Output_{size} = \frac{Input_{size} - Pooling_{size}}{Stide} + 1 \tag{8}$$

If the input feature map is square, the $Input_{size}$ is its width or height. The convolutional filter's (kernel's) width or height, referred to as $Filter_{size}$ is applied to the input. The amount of zero-padding pixels used on the edges of the input feature map is known as $Padding$. It helps maintain spatial dimensions and prevent information loss during convolution. The $Stride$ is the amount of movement the filter makes over the input. It shows how far the filter advances after each convolution operation, horizontally or vertically. where $Pooling_{size}$ is the width or height of the pooling window. The number of filters utilised is proportional to the output depth.

$$Entropy_{loss} = -\frac{1}{N}\sum_{i=1}^{N} \sum_{j=1}^{C} y_{true_{ij}} \log\left(y_{pred_{ij}}\right) \tag{9}$$

The batch size is denoted by N. The total number of categories is C. $y_{true_{ij}}$ is the true label for the i^{th} sample and j^{th} ' class. It is 1 if the i^{th} sample belongs to j^{th} class, otherwise 0. $y_{pred_{ij}}$ is the predicted probability for the i^{th} sample and j^{th}class.

$$\theta = \theta - \eta \nabla J(\theta) \tag{10}$$

where θ is model parameters, updated using the gradient calculated, η is learning rate, $\nabla J(\theta)$ is the loss function's gradient relative to the modelled variables.

AlexNet Model

An innovative convolutional neural network (CNN) model called AlexNet was put forth to investigate the classification of garbage bin images (whether they are Empty or Full). The model's performance will be analysed using two datasets (clean and noisy). the AlexNet [23] model has five convolutional layers comprise its structure, followed by three fully linked layers. It was deeper and more intricate than earlier CNN models, allowing it to learn hierarchical characteristics from garbage bin images successfully. It uses the ReLU activation function and includes overlapping max-pooling. This pooling technique uses a stride size smaller than the pooling window size to boost generality and spatial resolution. The training data was effectively made more diverse using data augmentation techniques like random cropping and horizontal flipping to prevent overfitting. It employs dropout regularisation in the fully linked layers to lessen overfitting. For thorough architecture, see Figure 5.

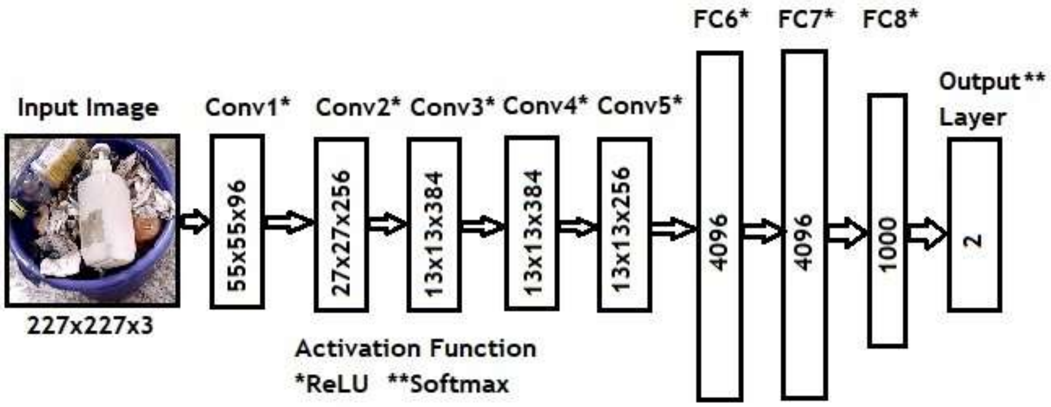
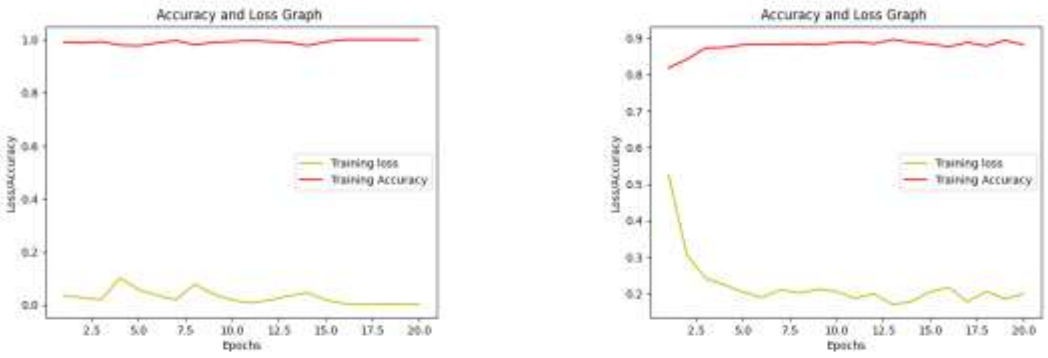


Figure 5. AlexNet Model

To implement the model, the Keras library with TensorFlow is used in the backend and is constructed using the "Sequential" API, which allows us to stack layers sequentially to build the neural network. With an input shape of (227, 227, 3) representing a trash bin picture, the first convolutional layer employs 96 filters with a kernel size of (11, 11), strides of (4, 4), and an activation function of ReLU. The second convolutional layer utilises a MaxPool of (3, 3) and strides of (2, 2). The MaxPool layer comes after a second layer that employs 256 filters of size (5, 3). The third layer employs 384 filters with a 3x3x1 ReLU activation function kernel and a 1x1 stride. The 256-filter fourth layer employs a ReLU activation function with a 3-by-3-by-1 kernel and a 1-by-1 stride. After a MaxPool Layer with a size of (3, 3) and a stride of (1, 1), the fifth layer employs 256 filters with the same activation function (ReLU).

After the first five convolutional layers, a Flatten Layer is used to flatten the resulting 3D feature maps into a 1D vector. The ReLU activation function is used in the 4096-neuron first and second dense layers with a dropout rate of 0.4%. The third dense layer employs 1000 neurons with ReLU activation function and the same dropout rate. The output layer calculates the probabilities of the output classes with a Softmax activation function. There are 20 training epochs for this model. Training loss and accuracy recorded and corresponding graphs are created with clean and noisy dataset as shown in Figure 6.



(a) (b)

Figure 6. Graph showing AlexNet model's (a) accuracy and (b) loss on clean and noisy datasets, respectively.

VGG-16 Model

The VGG16 model is presented to examine the classification of garbage bin status (Empty or Full) using clean and noisy 2D images from the dataset [24]. Figure 7, depicts the input image's number of colour channels (often 3) and the size of the waste bin pictures (224, 244). A maximum pooling layer and two convolutional layers comprise each of the five network blocks. The input picture is processed by two 64-filter convolutional layers (conv1 and conv2) in Block 1. The ReLU activation function is used with 3x3 filters in the convolutional layers. Downsampling the spatial dimensions by a factor of 2 is accomplished by applying a max-pooling layer (pool1) after each convolutional layer.

The second block employs 128-filter convolutional layers (conv3 and conv4) and a second max-pooling layer (pool2). Following a max-pooling layer (pool3), Block 3 features three more 256-filter convolutional layers (conv5, conv6, and conv7). Like the preceding blocks, Block 4 consists of three 512-filter convolutional layers (conv8, conv9, and conv10) and a max-pooling layer (pool4). Block 5 concludes with the final max-pooling layer (pool5) and three 512-filter convolutional layers (conv11, conv12, and conv13). Layer sizes in the maximum pool are permanently adhered to (3, 3).

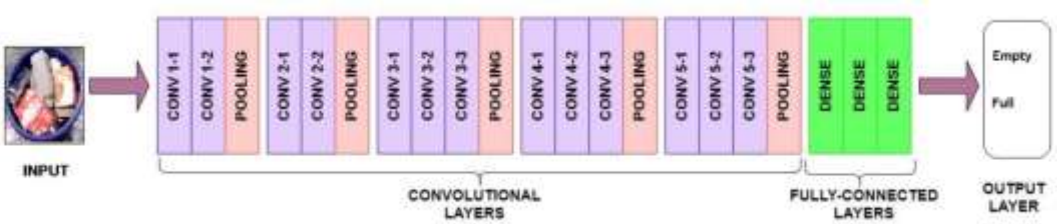
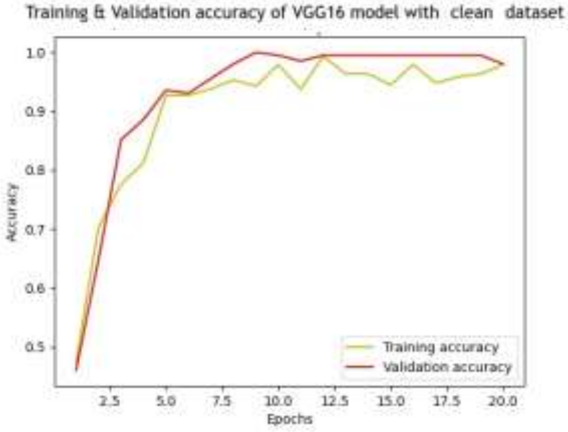
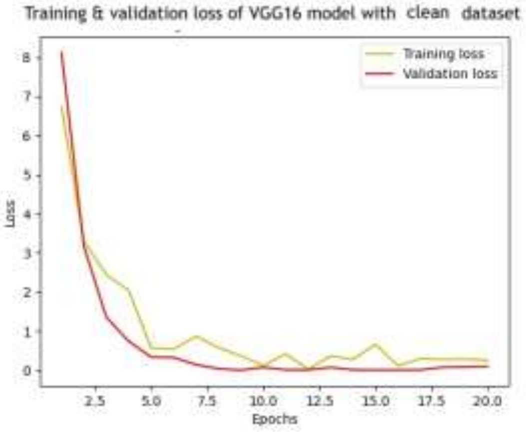


Figure 7. VGG-16 model

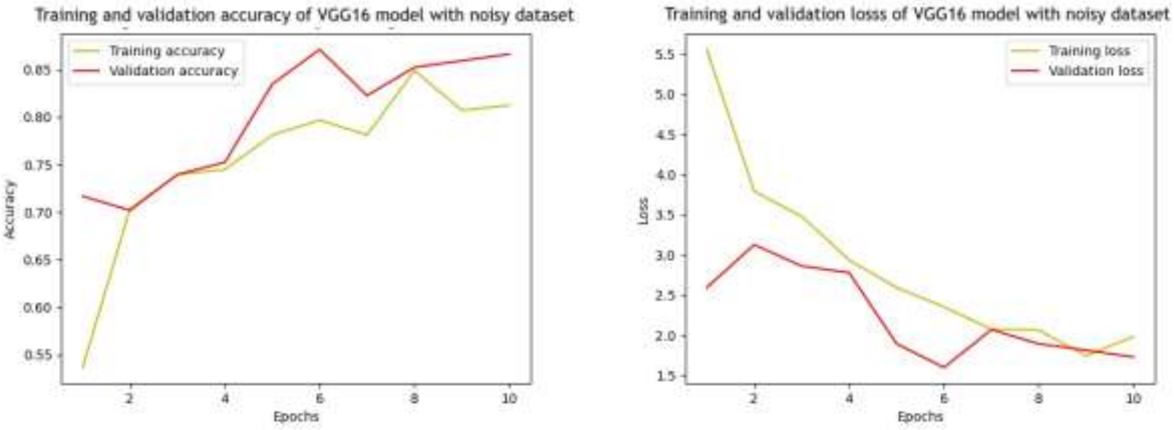
After the fifth and final pooling layer, the output is flattened into a one-dimensional vector. After that, there are two dense layers (dense1 and dense2), each containing 4096 units and activated by the ReLU function. These thick layers process the characteristics extracted by the convolutional layers and serve as a classifier. The two output classes from the classification task are represented by units in the final fully connected layer (output). The softmax activation function is well-suited to this category of classification problems since it transforms raw scores into class probabilities. There are 20 training epochs for this model. In Fig. 8, we have a graph depicting training loss and accuracy for clean and noisy datasets.



(a)



(b)



(c) (d)

Figure 8. VGG-16 model's graph shows (a) accuracy and (b) loss on the clean dataset, (c) accuracy and (d) loss on the noisy dataset.

CNN4 model

See Figure 9 to illustrate how the CNN4 model was explicitly developed for Full/Empty garbage bin image classification. The input layer architecture requires an RGB image of 200x200 pixels in dimensions and three colour channels as input. The first layer of the model is a Conv2D layer with 32 filters of size (3, 3). The ReLU activation function guarantees that the output has the exact spatial dimensions as the input. The spatial dimensions of the feature maps are reduced by half using a MaxPool2D layer with a pool size of (2, 2) and then another Conv2D layer with 32 filters. Two sets of Conv2D layers (each with 64 filters), ReLU activation, and MaxPool2D are used. In the succeeding Conv2D layers, filters are increased to 128 before ReLU activation and MaxPool2D is followed. A pair of Conv2D layers with 256 filters each, ReLU activation, and MaxPool2D are used to complete the feature extraction and dimensionality reduction.

$$\sigma(x) = \frac{1}{1+e^{-1}} \quad (11)$$

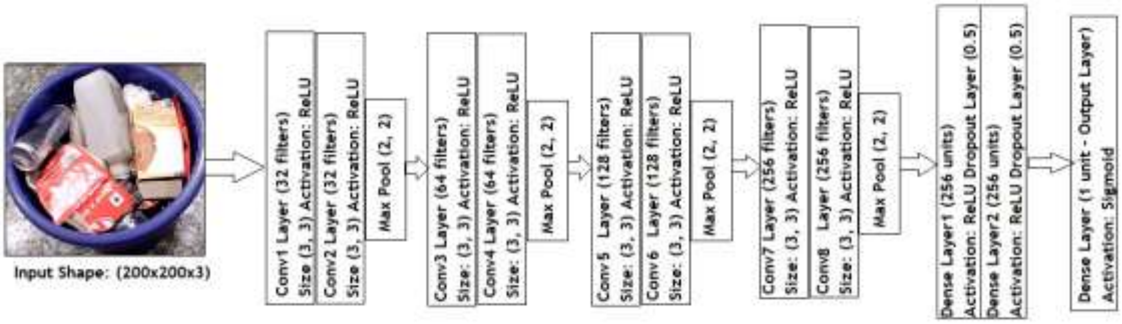
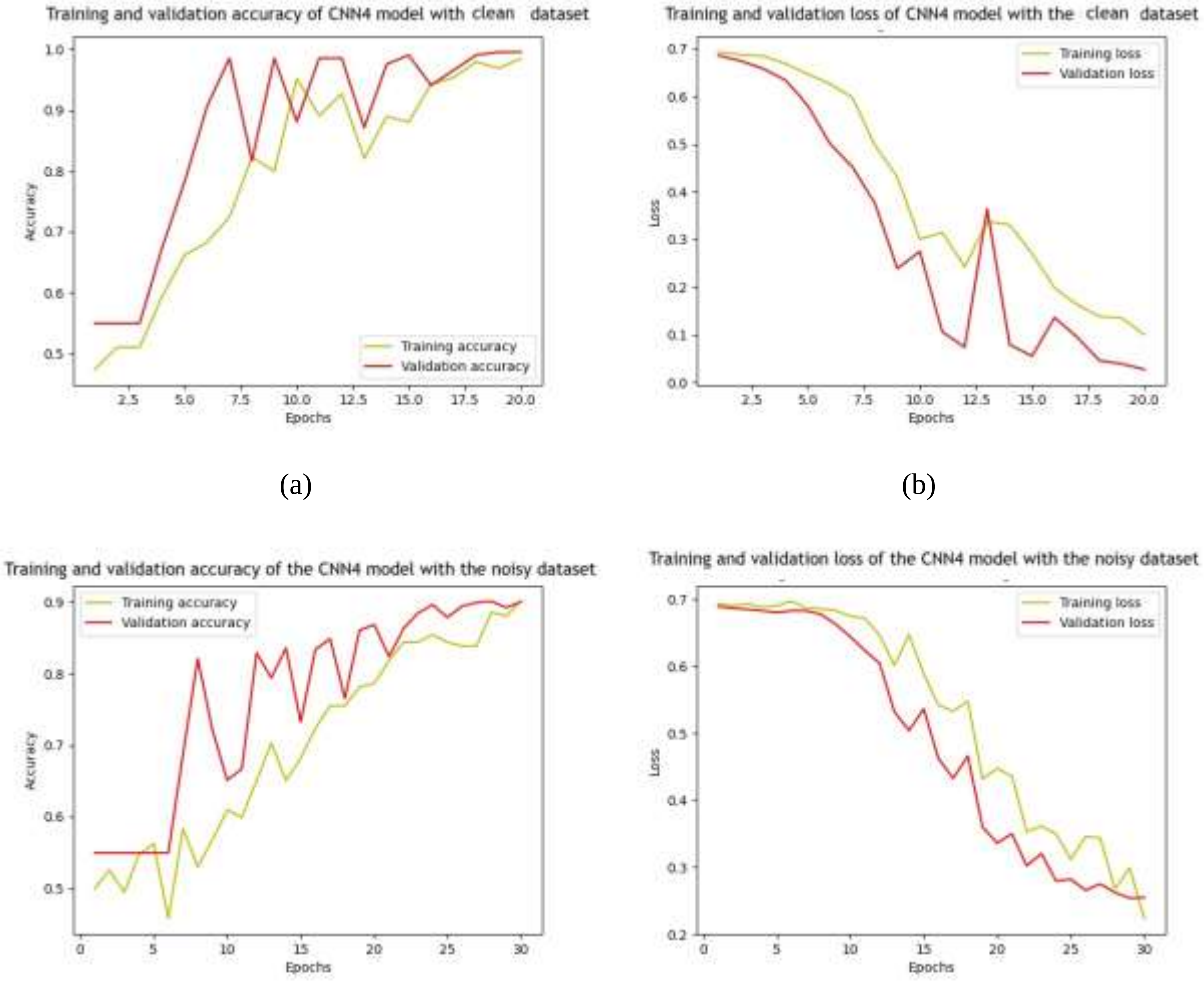


Figure 9. CNN4 model

The feature maps are flattened to a 1D vector to feed into fully connected layers. We implement a Dense layer that has 256 units and uses the ReLU activation function. This layer captures features at a higher level in the flattened vector. After the initial Dense layer, a Dropout layer is deposited at a rate of 0.5. Overfits can be avoided with dropout, which randomly sets a percentage of input units to 0 during training. After the initial Dropout layer, a second Dense layer with 256 units and ReLU activation is added. One sigmoid activation function unit (1 for Full class, 0 for Empty category) represents the binary classification task in the final Dense layer, as shown in Eq. (11). The model can then be trained with tagged garbage images to adjust its weights and biases for the classification problem at hand. As an optimiser, it employs Adam. The Adaptive Moment Estimation (Adam) optimisation algorithm is widely used in the

0.0001-per-second training of neural networks. Binary classification problems often employ the standard loss function. During training, it compares the actual labels to the anticipated probabilities and seeks to minimise the gap between them. Figure 10 visually represents the training and validation process, focusing on accuracy and loss metrics, across two distinct datasets: one composed of clean, standard data, and the other containing data with added noise or irregularities.



(a) (b)

(c) (d)

Figure 10. The graphs for the CNN4 model illustrate the following aspects: (a) Accuracy using the clean dataset, (b) Loss using the clean dataset, (c) Accuracy using the dataset with added noise, and (d) Loss using the dataset with added noise.

CNN5 model

The suggested CNN5 architecture uses five convolutional and pooling layers to gradually extract characteristics from garbage bin input images. Filters of ever-higher complexity are used in these layers to gather information across many scales. After the output has been flattened, it is fed into fully linked layers so that the model may make predictions. Since the model is built for binary classification tasks, the sigmoid activation function is used in the final layer. This setup demonstrates how effectively the CNN5 model can automatically acquire such representational hierarchy from an image.

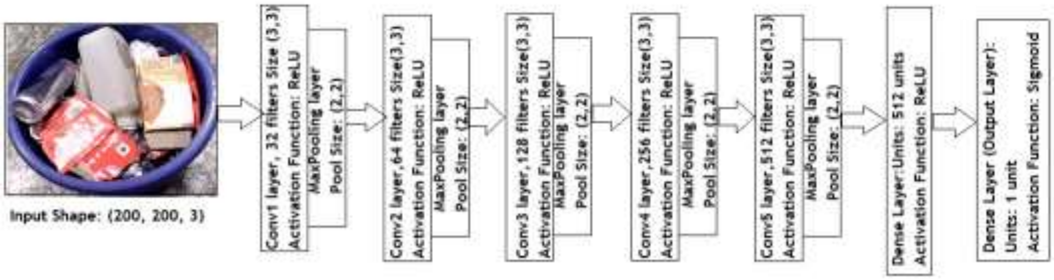
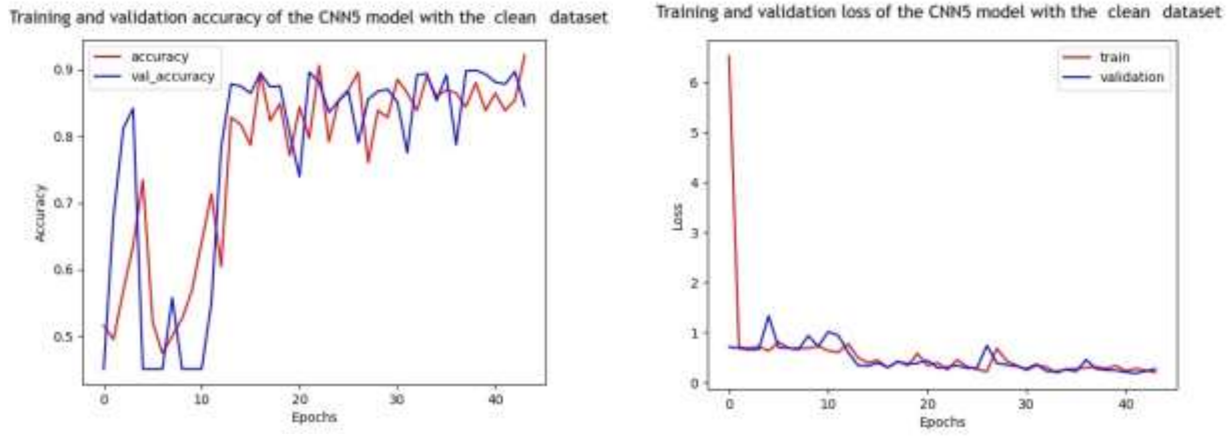


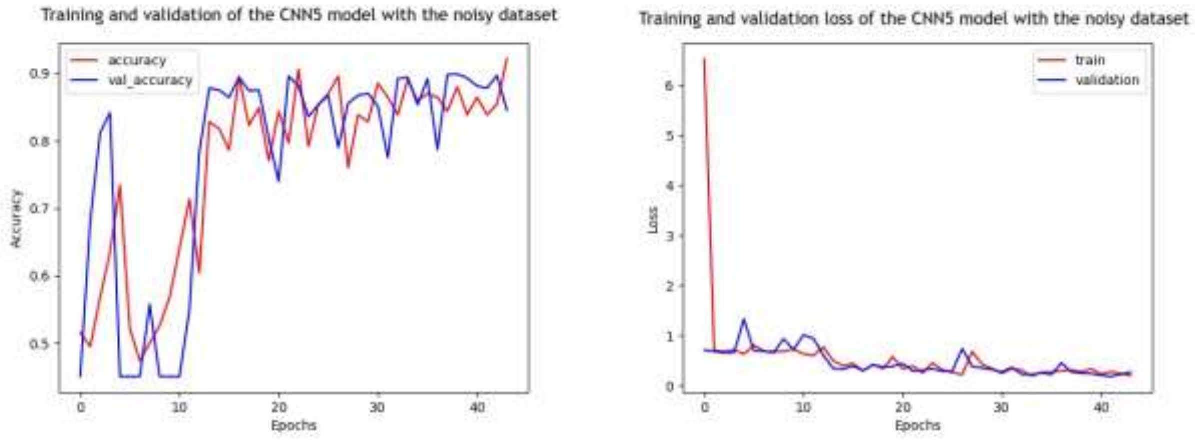
Figure 11. CNN5 model

As shown in Figure 11, layer 1 of the model receives an input data image with the dimensions (200, 200, 3). The model uses a 'Conv2D' layer with 32 filters of size (3, 3) and a ReLU activation function, and finally, a 'MaxPool2D' layer with a pooling size of (2, 2). At layer 2, we have a 'Conv2D' layer with 64 filters, followed by a 'MaxPool2D' layer with a pooling size of (2, 2). The 'MaxPool2D' layer comes after the third layer's 128 filters, the fourth layer's 256 filters and the fifth layer's 512 filters. To make the data suitable for the fully connected layers, the 'Flatten' layer flattens the multi-dimensional vector produced by the convolutional layers. There are new 512-unit dense layers that use the ReLU activation function. The model's output, which represents the binary classification result, is generated by a final 'Dense' layer consisting of a single unit and a sigmoid activation function.

During training, the model will employ the loss function, optimisation, and metric you choose to determine whether to change its starting weights. To maximise the accuracy of the training data while minimising the loss function, the model will iteratively tweak its parameters. The loss parameter defines the objective function against which the model is trained. The 'binary_crossentropy' function is used here. In this case, we employ the RMSprop optimiser, which modifies the training speed according to the various parameters. It speeds up convergence and enhances the effectiveness of training. The optimiser's step size in the weight space for each update is 0.001 because of the 'learning_rate' option, which is set to that value. It determines the rate at which the model learns from experience. See In Figure 12, we visually depict the training and validation procedure, emphasising accuracy and loss metrics, for two separate datasets: one consisting of clean data and the other with additional noise or irregularities.



(a) (b)



(c) (d)

Figure 12. Accuracy and Loss graphs of CNN5 (a) Accuracy with the clean dataset (b) Loss with the clean dataset (c) Accuracy with the noisy dataset (d) Loss with the noisy dataset.

Performance Evaluation

An essential part of evaluating the performance and quality of a prediction or classification model is the computation of the confusion matrix. The fraction of correct predictions (including true positives and true negatives) relative to the total number of instances is the overall accuracy, which may be calculated using Eq. (12). Precision (Positive Predictive Value) measures the model's ability to forecast positive cases accurately, See Eq. (13). The percentage of true positive cases a model accurately identifies is the recall (also known as sensitivity or true positive rate). This is a must when a false negative might have dire repercussions, as it reduces the likelihood of missing true positives, see Eq. (14). The F1-Score is the mathematical middle ground between precision and recall. It gives a statistic that considers false positives and negatives, striking a compromise between precision and recall. When the distribution of classes is irregular or noisy, as in the case of Eq. (15), this method excels.

$$Overall\ Accuracy = \frac{TP+TN}{TP+FP+TN+FN} \tag{12}$$

$$Precision = \frac{TP}{TP+FP} \tag{13}$$

$$Recall = \frac{TP}{TP+FN} \tag{14}$$

$$F1_{Score} = \frac{2}{\frac{1}{Recall}+\frac{1}{Precision}} \tag{15}$$

True positive (TP), true negative (TN), false positive (FP) type I error, and false negative (FN) type II error.

These performance indicators collectively provide a more comprehensive understanding of a garbage bin model's strengths and weaknesses. A high overall accuracy may not guarantee a good model, as it can be misleading in imbalanced classes. Therefore, this paper considers metrics like recall, precision, and F1-Score to gain a deeper insight into the model's performance, especially when misclassification costs vary between classes.

Results and Discussion

As depicted in Table 1, the performance of AlexNet, VGG16, CNN4, and CNN5 models stands out remarkably across clean and noisy datasets. Regarding accuracy, it becomes evident that CNN5 attains the highest accuracy on clean datasets, whereas CNN4 exhibits superior accuracy when dealing with noisy datasets.

The provided results of the CCN5 model with a clean dataset (accuracy 1.0, R2-Score 1.0, Precision 1.0, Recall 1.0, F1-Score 1.0 and area under ROC 1.0) reflect the performance of a classification model. The outstanding results from the CNN5 model's performance on the clean dataset underscore its exceptional classification capabilities. These metrics offer insights into how well the model performs regarding the Empty and Full classes and how well the model's predictions match the actual data variability. In the noisy dataset, a negative R2 score, such as -0.103, suggests that the model's predictions do not explain the data variability, which might indicate that the model is not well-suited for this dataset.

The provided results of the CCN4 model with a noisy dataset (accuracy 0.96, R2-Score 0.85, Precision 0.97, Recall 0.96, F1-Score 0.96 and area under ROC 0.96) reflect the performance of a classification model. This shows its commendable classification performance. The model demonstrates a balanced interplay between accuracy, precision, recall, F1-score, and AUC-ROC, signifying its capability to handle structured data and make accurate predictions. The R2-score of 0.85 is considered relatively high and is often considered a good fit for many applications. It implies that the model's predictions provide valuable insights into the target variable and are likely helpful for making accurate predictions or understanding the relationships between variables. Table 1 provides a more in-depth analysis.

Table 1. Comparisons of AlexNet, VGG16, CNN4, and CNN5 model's performance

Model Type	*Data Type*	*True Empty*	*False Empty*	*False Full*	*True Full*	*Accuracy*	*R2 Score*	*Precision Score*	*Recall Score*	*F1 Score*	*ROC Curve*
AlexNet	Clean	80	1	0	64	0.99	0.97	0.99	0.99	0.99	0.99
	Noisy	377	0	73	271	0.89	0.59	0.92	0.89	0.90	0.89
VGG16	Clean	75	0	1	74	0.99	0.97	0.99	0.99	0.99	0.99
	Noisy	260	30	6	441	0.95	0.93	0.98	0.94	0.96	0.95
CCN4	Clean	71	4	0	75	0.97	0.89	0.97	0.97	0.97	0.97
	Noisy	**265**	**25**	**0**	**447**	**0.96**	**0.85**	**0.97**	**0.96**	**0.96**	**0.96**
CCN5	**Clean**	**75**	**0**	**0**	**75**	**1.0**	**1.0**	**1.0**	**1.0**	**1.0**	**1.0**
	Noisy	288	192	2	255	0.74	-0.10	0.99	0.57	0.72	0.80

Conclusions

When applied to the clean dataset, CNN5 produces results that are nothing short of remarkable. Accuracy, R2-score, precision, recall, F1-score, and area under the receiver operating characteristic (AUC-ROC) all being flawless exemplify CNN5's unparalleled proficiency in dealing with clean, well-structured data. These findings have far-reaching ramifications in fields like image analysis, NLP, and science, where precise and trustworthy predictions are of the utmost significance. The CNN4 model performed exceptionally well when applied to the noisy dataset. Accuracy, R2 score, precision, recall, F1 score, and AUC-ROC are only some of the metrics that highlight how well the model handles noisy data. Considering the pervasive presence of noise in today's data landscape, the consequences of these discoveries are far-reaching, touching on fields as diverse as garbage identification, medical diagnosis and economic forecasting. Hence, this paper recommends the CNN5 model best suited for clean datasets and CNN4 for noisy datasets. While impressive results are undoubtedly promising, the broader context and potential limitations must be considered to ensure reliable and meaningful outcomes in practical scenarios. Hence may be used to monitor garbage bins and used as an application in smart cities. Despite progress, problems, such as uneven data distribution, environmental fluctuations, and more clarity in the models, still need fixing. Additional work might be done to integrate the system into real-world garbage management scenarios, investigate transfer learning techniques, and incorporate real-time data sources.

Acknowledgement

Our sincere gratitude goes out to the lab assistants, whose steadfast commitment and technical know-how were crucial in the development and precise processing of the main dataset for this experiment. Their dedication and unwavering support provided a solid foundation for our study. We also like to express our sincere gratitude to the hardworking library assistants who enhanced our research process. The breadth and depth of our comprehension were considerably enhanced by their relentless efforts to locate and provide pertinent literature, research papers, and resources whenever required. These people have exemplified the true meaning of teamwork and shared purpose through their cooperation and support. We appreciate their contributions, which were crucial in determining the outcome of this project.

References

[1] Z. Yang, Y. Bao, Y. Liu, Q. Zhao, H. Zheng, and Y. L. Bao, “Research on deep learning garbage classification system based on fusion of image classification and object detection classification,” *Mathematical Biosciences and Engineering*, vol. 20, no. 3, pp. 4741–4759, 2023, doi: 10.3934/mbe.2023219.

[2] A. Lakhouit *et al.*, “Machine-learning approaches in geo-environmental engineering: Exploring smart solid waste management,” *J Environ Manage*, vol. 330, no. November 2022, p. 117174, 2023, doi: 10.1016/j.jenvman.2022.117174.

[3] R. K. Sharma and M. Jailia, “IoT based smart bin system using multilayer neural network classifier for smart cities,” in *AIP Conference Proceedings*, AIP Publishing, 2023.

[4] S. S. Mookkaiah, G. Thangavelu, R. Hebbar, N. Haldar, and H. Singh, “Design and development of smart Internet of Things–based solid waste management system using computer vision,” *Environmental Science and Pollution Research*, vol. 29, no. 43, pp. 64871–64885, 2022, doi: 10.1007/s11356-022-20428-2.

[5] G. Uganya, D. Rajalakshmi, Y. Teekaraman, R. Kuppusamy, and A. Radhakrishnan, “A Novel Strategy for Waste Prediction Using Machine Learning Algorithm with IoT Based Intelligent Waste Management System,” *Wirel Commun Mob Comput*, vol. 2022, 2022, doi: 10.1155/2022/2063372.

[6] T. Gupta *et al.*, “A deep learning approach based hardware solution to categorise garbage in environment,” *Complex and Intelligent Systems*, vol. 8, no. 2, pp. 1129–1152, 2022, doi: 10.1007/s40747-021-00529-0.

[7] S. Vishnu *et al.*, “IoT-enabled solid waste management in smart cities,” *Smart Cities*, vol. 4, no. 3, pp. 1004–1017, 2021, doi: 10.3390/smartcities4030053.

[8] R. Rautela, S. Arya, S. Vishwakarma, J. Lee, K. H. Kim, and S. Kumar, “E-waste management and its effects on the environment and human health,” *Science of the Total Environment*, vol. 773. Elsevier B.V., Jun. 15, 2021. doi: 10.1016/j.scitotenv.2021.145623.

[9] A. Suvarnamma and J. A. Pradeepkiran, “SmartBin system with waste tracking and sorting mechanism using IoT,” *Clean Eng Technol*, vol. 5, Dec. 2021, doi: 10.1016/j.clet.2021.100348.

[10] C. Wang, J. Qin, C. Qu, X. Ran, C. Liu, and B. Chen, “A smart municipal waste management system based on deep-learning and Internet of Things,” *Waste Management*, vol. 135, pp. 20–29, Nov. 2021, doi: 10.1016/J.WASMAN.2021.08.028.

[11] E. Aktay and N. Yalçın, “A smart city application: A waste collection system with long range wide area network for providing green environment and cost effective and low power consumption solutions,” *IET Smart Cities*, vol. 3, no. 3, pp. 142–157, Sep. 2021, doi: 10.1049/smc2.12014.

[12] T. M. N. Vamsi, G. Kalyan Chakravarthi, P. Lanka, and B. Divakar, “An IoT Based Smart Garbage Monitoring and Disposal Support System,” in *Proceedings - 5th International Conference on Computing Methodologies and Communication, ICCMC 2021*, Institute of Electrical and Electronics Engineers Inc., Apr. 2021, pp. 438–442. doi: 10.1109/ICCMC51019.2021.9418289.

[13] M. Fajkus *et al.*, “Fiber-optic bragg system for the dynamic weighing of municipal waste: A pilot study,” *IEEE Access*, vol. 9, pp. 99050–99059, 2021, doi: 10.1109/ACCESS.2021.3095219.

[14] “Khan R, Kumar S, Srivastava AK, Dhingra N, Gupta M, Bhati N, Kumari P (2021) Machine learning and IoT-based waste management model. Comput Intelli Neurosci. https://doi.org/10.1155/2021/5942574.”

[15] N. H. Latifah, S. Sitangsu, S. Rasyad, and A. S. Handayani, “Real Time Garbage Bin Capacity Monitoring,” *Computer Engineering and Applications Journal*, vol. 9, no. 2, pp. 127–134, 2020, doi: 10.18495/comengapp.v9i2.340.

[16] T. J. Sheng *et al.*, “An Internet of Things Based Smart Waste Management System Using LoRa and Tensorflow Deep Learning Model,” *IEEE Access*, vol. 8, pp. 148793–148811, 2020, doi: 10.1109/ACCESS.2020.3016255.

[17] S. Dubey, M. K. Singh, P. Singh, and S. Aggarwal, “Waste Management of Residential Society using Machine Learning and IoT Approach,” *2020 International Conference on Emerging Smart Computing and Informatics, ESCI 2020*, pp. 293–297, 2020, doi: 10.1109/ESCI48226.2020.9167526.

[18] A. Khmag, “Additive Gaussian noise removal based on generative adversarial network model and semi-soft thresholding approach,” *Multimed Tools Appl*, vol. 82, no. 5, pp. 7757–7777, 2023, doi: 10.1007/s11042-022-13569-6.

[19] F. Dong, Y. Chen, D. X. Kong, and B. Yang, “Salt and pepper noise removal based on an approximation of l0 norm,” *Computers and Mathematics with Applications*, vol. 70, no. 5, pp. 789–804, 2015, doi: 10.1016/j.camwa.2015.05.026.

[20] M. H. Syed, K. Upreti, M. S. Nasir, M. S. Alam, and A. Kumar Sharma, “Addressing image and Poisson noise deconvolution problem using deep learning approaches,” *Comput Intell*, no. February, 2022, doi: 10.1111/coin.12510.

[21] C. Wu, F. Wu, T. Qi, and Y. Huang, “NoisyTune: A Little Noise Can Help You Finetune Pretrained Language Models Better,” *Proceedings of the Annual Meeting of the Association for Computational Linguistics*, vol. 2, pp. 680–685, 2022, doi: 10.18653/v1/2022.acl-short.76.

[22] C. Srinilta and S. Kanharattanachai, “Municipal solid waste segregation with CNN,” *Proceeding - 5th International Conference on Engineering, Applied Sciences and Technology, ICEAST 2019*, pp. 1–4, 2019, doi: 10.1109/ICEAST.2019.8802522.

[23] T. F. Gonzalez, “Handbook of approximation algorithms and metaheuristics,” *Handbook of Approximation Algorithms and Metaheuristics*, pp. 1–1432, 2007, doi: 10.1201/9781420010749.

[24] Karen Simonyan∗ & Andrew Zisserman+, “VERY DEEP CONVOLUTIONAL NETWORKS FOR LARGE-SCALE IMAGE RECOGNITION Karen,” *American Journal of Health-System Pharmacy*, vol. 75, no. 6, pp. 398–406, 2018, [Online]. Available: https://click.endnote.com/viewer?doi=arxiv%3A1409.1556&token=WzQxNTM2OSwiYXJ4aXY6MTQwOS4xNTU2Il0.1ERLUjt9JINMPEYW74pwJFdB7PA