

AI-driven Innovations for Sustainable Hospital Waste Management in Developing Countries

Teddy Ako Eyonganyoh

African Institute for Mathematical Sciences (AIMS), P.O. Box 608 Limbe, Cameroon and Department of Computer Engineering, Faculty of Engineering and Technology, the University of Buea, PO Box 63, Cameroon

E-mail: teddy.eyonganyoh@aims-cameroon.org

ORCID iD: <https://orcid.org/0009-0002-3874-3224>

Samuel Kotva Goudougou*

African Institute for Mathematical Sciences (AIMS), P.O. Box 608 Limbe, Cameroon and Department of Mathematics and Computer Science, Faculty of Science, the University Ngaoundere, Cameroon, PO Box 454 Ngaoundere, Cameroon

E-mail: samuel.kotva@aims-cameroon.org, samuel.kotva@univ-ndere.cm

ORCID iD: <https://orcid.org/0009-0005-0674-6588>

*Corresponding author

Justin Moskolai Ngossaha

Department of Mathematics and Computer Science, Faculty of Science, the University of Douala, Cameroon, PO Box 24 157 Douala, Cameroon

E-mail: jmoskolai@fs-univ-douala.cm

ORCID iD: <https://orcid.org/0000-0002-9228-4543>

Paul Dayang

Department of Mathematics and Computer Science, Faculty of Science, the University Ngaoundere, Cameroon, PO Box 454 Ngaoundere, Cameroon

E-mail: pdayang@univ-ndere.cm

ORCID iD: <https://orcid.org/0000-0002-6566-5447>

Received: January 14, 2026; Revised: March 2, 2026; Accepted: April 3, 2026; Published: August 8, 2026

Abstract: Population growth and pandemics like COVID-19 have led to the depletion of natural resources and an increase in hospital waste generation. This issue is particularly pressing in developing countries, where innovative solutions are needed to address the environmental and health risks associated with improper waste disposal. Traditional waste sorting methods, which rely on human intervention, are time-consuming and pose a significant risk of infection. Moreover, different categories of hospital waste require specific treatment methods. This study proposes an artificial intelligence-based approach for classifying and sorting hospital waste using Convolutional Neural Networks (CNNs). The proposed CNN model effectively identifies and categorizes various types of hospital waste, providing a sustainable solution that enhances regulatory compliance. The model leverages K-fold cross-validation and data augmentation on a publicly available, modest-sized dataset to handle variability in waste categories. The model achieved a peak classification accuracy of 97.08%, along with high precision, recall, and F1-scores, despite class imbalances and the presence of visually similar waste categories. These results highlight the potential of the model to improve hospital waste management practices, thereby reducing environmental impact and health risks.

Index Terms: Artificial Intelligence, Hospital Waste Management, Developing Countries, Sustainability, Convolutional Neural Network (CNN).

1. Introduction

In recent decades, rapid economic development, population growth, and urbanization have placed increasing

pressure on healthcare systems worldwide, leading to a significant rise in the generation of healthcare waste. This surge in medical waste, particularly in densely populated regions, poses a substantial environmental and public health challenge. Africa, the second most populous continent, experiences the fastest rate of population growth, further intensifying these issues. Developing countries face distinct difficulties in managing healthcare waste due to limited access to resources such as electricity, insufficient infrastructure, and ineffective waste management systems [1]. These shortcomings have led to a reliance on traditional waste disposal methods, with a significant portion of waste being deposited in public landfills, contributing to environmental degradation and posing risks of disease transmission [2].

The COVID-19 pandemic exacerbated the already critical situation of healthcare waste management, particularly in developing regions. The pandemic generated unprecedented amounts of highly infectious medical waste, creating new complexities in safe disposal [3]. Prior to the pandemic, over 50% of the global population was already exposed to environmental pollution and public health risks due to improper waste disposal practices [4]. The rapid increase in waste production during the pandemic exposed the limitations of traditional waste management systems and underscored the urgent need for more efficient and sustainable waste management practices.

According to recent projections, municipal solid waste is expected to rise from 2.01 billion tonnes in 2018 to 3.4 billion tonnes by 2050 [5]. Without proper management, this growing volume of waste will have severe environmental consequences, including soil and water contamination, as well as the increased transmission of diseases. The global medical waste management market is expected to grow from USD 12.5 billion in 2022 to USD 23.7 billion by 2032, driven by the increasing volume of healthcare waste and rising awareness of its environmental impact [6]. However, despite this growth, traditional waste management methods—such as manual sorting, incineration, and landfill disposal—are proving inadequate in addressing the volume and complexity of modern healthcare waste, especially in developing nations.

Waste management systems in developing countries are often characterized by inefficiency and a lack of infrastructure. More than 90% of solid waste in these regions is either incinerated or deposited in landfills, both of which contribute to environmental pollution and pose serious public health risks [7]. The infectious and hazardous nature of healthcare waste requires specialized handling and disposal, which traditional methods are ill-equipped to provide. Moreover, manual waste sorting methods expose workers to a high risk of infection, further complicating the management process. These issues underscore the need for innovative solutions that can enhance the efficiency of healthcare waste management and promote sustainability.

In response to these challenges, artificial intelligence (AI) has emerged as a promising technology that can revolutionize waste management practices. AI has the potential to improve efficiency, reduce operational costs, and optimize waste sorting, collection, and disposal processes. By leveraging machine learning algorithms and data-driven decision-making, AI can enhance the identification and segregation of different types of healthcare waste, minimizing human error and promoting safer handling procedures [8]. Furthermore, AI can predict waste generation patterns and optimize the logistics of waste collection and transportation, making waste management systems more adaptive and resilient.

This study aims to explore the potential of AI-driven innovations in addressing the challenges of healthcare waste management, particularly in developing countries. By analyzing current applications of AI in waste management and investigating emerging trends, this research seeks to propose strategic solutions that integrate AI technologies into healthcare waste management systems. The study also addresses the technical and infrastructural barriers to implementing AI in developing regions and offers recommendations for overcoming these obstacles. In doing so, this work contributes to the global effort to promote sustainable healthcare waste management and reduce the environmental and public health risks associated with improper waste disposal.

The rest of the paper is organized as follows: Section 2 provides a review and synthesis of existing research on artificial intelligence concepts, highlighting their application on waste management. Section 3 describes the methodological framework proposed, including the techniques and tools used. Section 4 presents the resulting model based on the case study, including the main steps in the application of the methodology. The final section summarizes the main findings of the study, highlighting the contributions of the proposed framework. It also provides directions for future developments, including potential extensions, as well as potential applications in other domains.

2. Related Works

Hospital waste management refers to the systematic handling, treatment, and disposal of several categories of waste produced in healthcare facilities. As discussed by Odonkor and Mahami [9], these facilities, including hospitals, clinics, and medical laboratories inevitably generate a diverse range of healthcare waste, which is an integral part of their health services or operations. Some of these categories of hospital waste according to the standards of the World Health Organization (WHO) [10] are infectious, pharmaceutical, sharps, and general waste. Each category presents unique challenges and requires specific handling, treatment, and disposal methods to mitigate associated risks effectively. Table 1 presents an overview of the various categories of waste with their corresponding modes of treatment and disposal.

Among these waste types, pharmaceutical waste and contaminated needles and syringes (sharps) stand out due to their nature, necessitating careful disposal to prevent environmental contamination and public health hazards. With

general waste constituting about 85% of the total healthcare waste generated, pharmaceutical, sharps, infectious, toxic, and radioactive waste constitute the remaining 15% [9,10]. When indiscriminately disposed of, pharmaceutical waste (e.g., expired drugs) is harmful to the public if it re-enters the community, and openly burning this type of waste poses environmental risks due to harmful residuals and emissions [1]. Additionally, improper disposal of sharps can lead to the transmission of infectious diseases. According to statistics from the WHO [9], an individual who is wounded by a needle that was previously used on an infected patient has a 30% chance of contracting hepatitis B, a 1.8% chance of contracting hepatitis C, and a 0.3% chance of contracting HIV. Unsafe injections have been identified as the cause of approximately 1.7 million cases of hepatitis B, 315,000 cases of hepatitis C, and 33,800 new HIV infections. Hospital waste management is a huge challenge worldwide, especially in developing countries. According to several studies, some of these challenges include inadequate waste segregation practices, a lack of proper disposal facilities, and limited awareness of health and safety standards among healthcare workers. Improper waste disposal has huge implications on human health and the environment, such as groundwater contamination, child mortality, and congenital disabilities[11]. Approximately 40% of healthcare professionals sustain injuries when handling waste, which may result in skin infections, musculoskeletal disorders, and eye injuries[12].

Multiple studies have shown that in developing countries' settings, where proper methods for disposing of medical waste are lacking or poorly established, hospital waste is frequently combined with regular household waste and treated as municipal solid waste either by incineration or by disposing of it in open landfills. These treatment methods are used for more than 90% of hospital waste. Incineration, which remains a prevalent method of waste treatment, offers benefits such as a 90% reduction in waste volume and effective energy recovery. However, concerns persist regarding the emission of harmful air pollutants if not properly managed. The other most common method of disposing is through illegal dumping where waste is disposed in areas not designated for waste disposal [13]. Table 1 presents the different disposal methods, their risk and proposed solutions.

Table 1. Categories of hospital wastes, treatment, and disposal methods according to WHO [4]

Waste Category	Constituents	Treatment and Disposal Options
Risk waste	Infectious waste, infected sharps, and genotoxic waste	Incineration.
Non-risk waste	Paper, plastic, cardboard and packaging	Anaerobic digestion - Pyrolysis.
Sharps	Needles, syringes, scalpels, infusion sets, saws and knives, blades and broken glass	Autoclaving/microwave treatment and recycling. Plasma gasification.
Pharmaceutical waste	Expired or unused pharmaceutical products, surplus drugs, vaccines or sera and discarded items used in handling pharmaceutical waste, such as bottles, boxes, gloves, masks, tubes or vials	Returned to pharmacies or manufacturers. Recycling. Periodically draining in limited quantities.
Chemical waste	Chemicals from diagnostic and experimental work, cleaning processes, housekeeping and disinfecting procedures, and discarded batteries	Chemicals dilution - batteries recycling
Radioactive waste	Liquid, solid and gaseous waste contaminated with radionuclides generated from in-vitro analysis of body tissue and fluid, in vivo body organ imaging and tumour localisation, and investigation and therapeutic procedures	Treating and burying by appropriate government agency

2.1. Overview of Artificial Intelligence

Artificial intelligence is a rapidly advancing technology that is gaining popularity in various industries [7]. In recent years, the healthcare industry has been increasingly exploring the potential of AI in various domains, and waste management is no exception [4]. As defined by Herath and Mittal [14], AI refers to the training of computers to imitate thinking processes and replicate human behavior. Furthermore, it is a specialized field within computer science that focuses on simulating human intelligence processes. It also involves the development of data-driven systems that allow computers or software to perform tasks and make informed decisions.

AI algorithms, the backbone of AI technologies, are sets of rules or instructions that enable computers to learn autonomously. These algorithms are capable of processing complex data inputs, learning from data, and making decisions or predictions based on that data. These AI algorithms are diverse, ranging from simpler classical algorithms like decision trees (DT) and logistic regression to complex machine learning techniques including neural networks, deep learning, and reinforcement learning. They are used across various applications, natural language processing, robotics, autonomous systems, computer vision, and more.

In the field of AI, Computer vision represents a significant area focused on enabling machines to interpret and understand visual information from the world. It involves developing techniques for machines to extract meaningful information from digital images or videos. Computer vision tasks include image classification, object detection, image segmentation, and more, which are critical for tasks ranging from autonomous driving to medical image analysis. Convolutional Neural Networks (CNNs), a type of deep learning model specially designed for processing grid-like data such as images, are essential to computer vision because they are highly efficient at processing and analyzing visual data.

The following section discusses the fundamentals of deep learning, particularly CNNs, which enable the automation of waste detection and sorting.

2.2. Fundamentals of Deep Learning Algorithm

Deep learning, a subset of machine learning, uses algorithms inspired by the structure and function of the brain called artificial neural networks. At the core of deep learning is the concept of layered representations of data, which are learned through models that consist of multiple layers of computations. This section delves into the core components and architectural principles of deep learning (particularly CNNs), which power many modern AI applications. ANNs are composed of layers of interconnected nodes or neurons, which mimic the neural connections in the human brain. ANNs are fundamental to various fields in AI making them effective in tasks which include pattern recognition, regression and classification.

ANNs are integral to the functionality of CNNs, where they are used to learn hierarchical representations of data. The fundamental unit of an ANN is the artificial neuron, which mimics the function of a biological neuron. An artificial neuron receives input signals, processes them via a weighted sum, and applies a non-linear activation function to produce an output. It uses weights to adjust its output. Figure 1 gives a representation of the artificial neuron, and its mathematical model is:

$$z = \sum_{i=1}^n w_i x_i + b \quad (1)$$

$$y = f(z) \quad (2)$$

where x_i are the input features, w_i the weights, b the bias term, z the linear function (weighted sum plus bias), f the activation function and y the output.

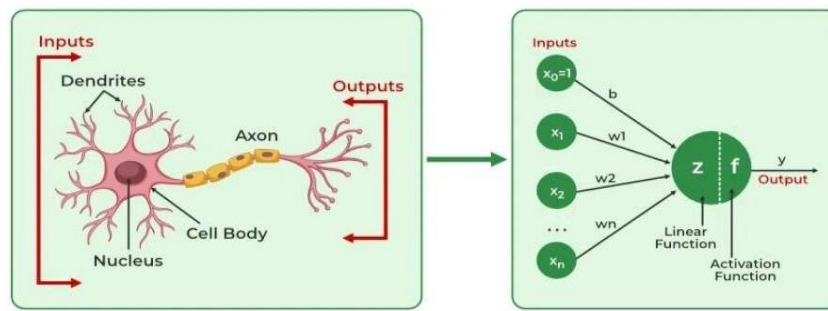


Fig.1. Biological neuron to artificial neuron (Image Source: [15])

When multiple neurons are linked together, they form layers. In an ANN, there are three essential layers: the input, hidden layers, and the output layer. The input layer receives the raw data input, the hidden layer performs computations and feature transformations, and the output layer produces the final decision or prediction. This structure allows ANNs to handle complex pattern recognition tasks by learning from data in a way that mimics human cognitive processes. Deep learning architectures are characterized by their depth, which is the number of layers they contain.

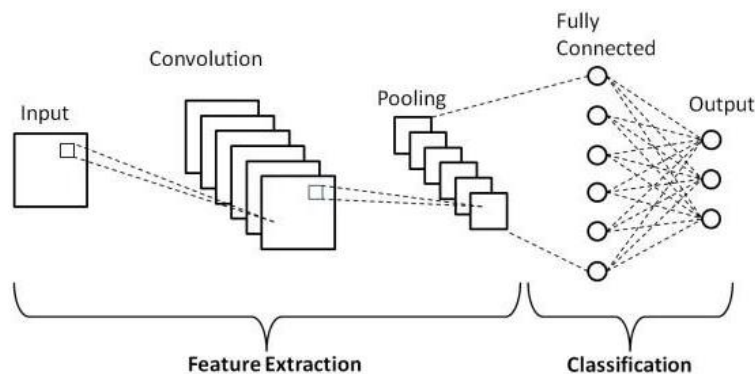


Fig.2. Outline of CNN (Image Source: [16])

Convolutional Neural Networks (CNNs): This specialized class of deep neural networks is particularly powerful for tasks involving grid-like data, such as images. CNNs have shown remarkable success in areas of image recognition, classification, and analysis. CNNs are designed to automatically and adaptively learn spatial hierarchies of features through backpropagation by using multiple building blocks, such as convolutional layers, pooling layers, and fully

connected layers. Each layer of a CNN transforms one volume of activations to another through a differentiable function. This design allows CNNs to effectively handle the variability and complexity of image data, making them particularly suitable for the task of image-based classification where the input is an image and the output is a label classifying the image content.

Figure 2 shows the typical architecture of a CNN for image classification which includes several layers that each play a crucial role.

While ANNs are general-purpose networks used for structured data tasks like regression and classification, CNNs are specialized for processing grid-like data, such as images, using convolutional and pooling layers to capture spatial hierarchies effectively.

Convolutional Layers: These layers perform a convolution operation that filters the input image to create feature maps that summarize the presence of specific features in the input. The first convolutional layer captures lower-level features such as edges and textures, while deeper layers might identify more complex features like shapes or objects. The convolution operation involves sliding a filter or kernel over the input image and computing the dot product of the filter with the image pixels within its field of view. The convolution operation can be mathematically defined as follows:

$$F(i, j) = (I * K)(i, j) = \sum_m \sum_n K(m, n) \cdot I(i - m, j - n) \quad (3)$$

where I is the input matrix, K is the filter, and F is the output feature map at the position (i, j) and m and n iterate over the filter dimensions.

Pooling Layers: This layer, also known as subsampling or downsampling, typically comes after the convolutional layer. It reduces the spatial dimensions or dimensionality (height and width) of each feature map of the input volume but retains the most important information. This reduction is crucial for decreasing computational complexity and controlling overfitting. Max pooling is one of the most common pooling methods, which involves selecting the maximum element from the region of the feature map covered by the filter. In max pooling, the output for each region of the input is given by:

$$P(i, j) = \max_{(m,n) \in W} X(i \cdot \text{stride} + m, j \cdot \text{stride} + n) \quad (4)$$

where W is the area of the window being pooled over, and stride is the step by which the window is slid across the input.

Figure 3 shows an example of max pooling with 2×2 filters and a stride 2.

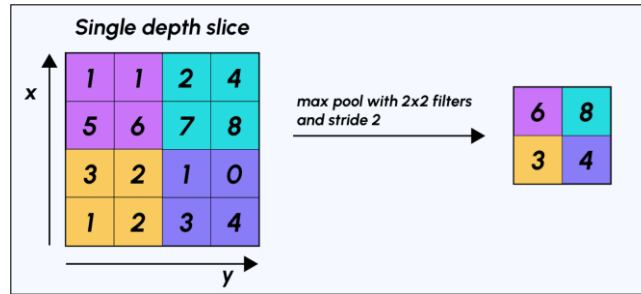


Fig.3. Max-pooling process (Image Source: [17])

Fully Connected Layers: After several convolutional and pooling layers, the high-level reasoning in the neural network is done via fully connected layers. Each neuron in these layers is connected to all activations in the previous layer, helping to classify the image based on the features extracted by the convolutional and pooling layers. Their activation can thus be computed with a matrix multiplication followed by a bias offset.

Activation Functions: After each convolution operation, an activation function is applied to introduce non-linear properties to the system, which allows the network to learn more complex patterns. Common activation functions include the Rectified Linear Unit (ReLU), Sigmoid, and Tanh. ReLU, which is applied after each convolution operation, encourages faster convergence during training by simply thresholding values at zero. Sigmoid and Tanh are functions less used in deep CNNs due to vanishing gradient issues but are still important in the output layers for binary and multi-class classification problems. The ReLU, sigmoid and tanh functions are respectively defined as follows:

$$f(x) = \max(0, x) \quad (5)$$

$$\sigma(x) = \frac{1}{1 + e^{-x}} \quad (6)$$

$$\tanh(x) = \frac{e^x - e^{-x}}{e^x + e^{-x}} \quad (7)$$

Softmax Activation Function: This activation function is commonly used in the output layer of neural networks handling multi-class classification tasks. It converts the raw output scores, known as logits, from the network into probabilities by taking the exponential of each output and then normalizing these values by dividing by the sum of all the exponentials.

Given a vector of raw scores $z = [z_1, z_2, \dots, z_n]$, the softmax function $\sigma(z)$ is defined as:

$$\sigma(z_i) = \frac{e^{z_i}}{\sum_{j=1}^n e^{z_j}} \quad (8)$$

where $z_i \in z$ is the score for class i , e is the base of the natural logarithm, and the denominator is the sum of all probabilities equaling 1.

This ensures that the output values are in the range $[0, 1]$ and sum to 1, making them interpretable as probabilities. The function is particularly useful when the model needs to classify inputs into multiple categories.

$$\sum_{i=1}^n \sigma(z_i) = 1 \quad (9)$$

Loss Functions: The cross-entropy loss function is particularly common in classification problems. This function measures the performance of a classification model whose output is a probability value between 0 and 1. Cross-entropy loss increases as the predicted probability diverges from the actual label. This is mathematically defined as:

$$L(y_c, \hat{y}_c) = -\sum_{c=1}^M y_c \log(\hat{y}_c) \quad (10)$$

where y_c denotes the true probability or target value for class c , $\hat{y}_c$ denotes the predicted probability for class c produced by the model, and M the total number of classes.

Backpropagation: This is used for training the neural network. Backpropagation computes the gradient of the loss function with respect to each weight by the chain rule, updating the weights to minimize the loss.

Gradient Descent: A method to update the weights and minimize the loss function. It adjusts the weights incrementally, using the gradient of the loss function.

$$w_{\text{new}} = w_{\text{old}} - \alpha \frac{\delta L}{\delta w} \quad (11)$$

where α is the learning rate and w is the weight.

Regularization: L2 Regularization, a regularization technique adds a penalty equal to the sum of the squared values of the weights to the loss function, which helps to keep the weights small and improve the generalization of the model.

Dropout: A simple yet effective regularization technique that prevents overfitting by randomly dropping units (and their connections) from the neural network during training.

2.3. AI Applications in Hospital Waste Management

In this section, we provide a comprehensive review of existing applications of AI and CNNs in hospital waste management. First, we examine AI-driven solutions tailored to address challenges in hospital waste management, particularly in resource-constrained developing countries. Next, we review CNN-based approaches for hospital waste management, with a focus on waste sorting. We aim to demonstrate how AI and CNNs are revolutionizing sustainable hospital waste management practices. The study by Vu et al. integrates artificial neural network (ANN) waste prediction models with geographic information system (GIS) waste collection route optimization [12]. Using data from Austin, Texas, USA, they predict waste generation rates for recycling and garbage streams in four city sub-areas for 2023. The ANN model yields mean absolute percentage errors ranging from 10.92% to 16.51%. They then create scenarios reflecting potential future changes in waste composition and use GIS tools to determine optimal waste collection routes. Results show significant changes in travel distance, up to 19.9%, compared to non-modified compositions. Also, dual compartment trucks are compared to single compartment trucks, revealing potential travel distance savings between 10.3% and 16.0%, albeit with a slight increase in collection time.

3. Methodological Approach

The methodological framework consists of four primary phases: data preparation, model development, model testing and validation, and model evaluation. Data preparation focuses on gathering, cleaning, augmenting, normalizing, and splitting the data. This step is crucial as it ensures data quality, thereby yielding reliable model outputs. During the model development phase, the architecture of the CNN is built, compiled, and trained using the prepared data. The neural network layers and parameters are set up to learn from the data. Once the model is developed, it undergoes testing and validation. This testing and validation are performed using a separate subset of data unseen by the model during the training phase. This is done to assess the generalizability and robustness of the model. The final phase involves applying evaluation metrics to determine the model's performance, so as to provide insights into the

effectiveness of the model in real-world scenarios. Figure 4 visually represents the approach, showing the steps taken for the model's creation to its evaluation.

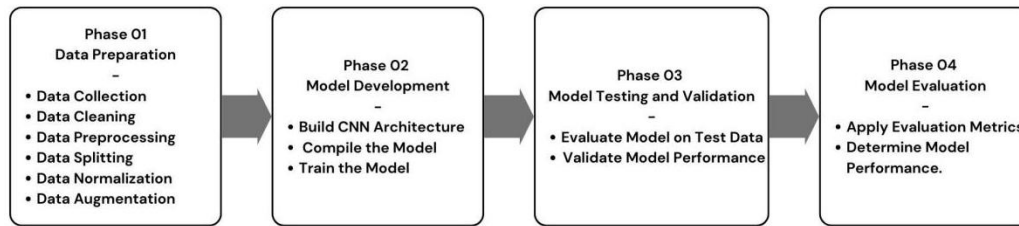


Fig.4. Methodological framework for image classification

3.1. Data Preparation

To build our hospital waste sorting model, we begin with data preparation. This encompasses several other steps which are crucial for training the CNN model to accurately classify hospital waste. The various steps taken in data preparation, from data collection to data splitting and normalization are as follows.

Data Collection: The dataset consists of RGB images collected from numerous healthcare facilities, including hospitals and clinics, ensuring a diverse representation of waste types. These images capture a wide range of hospital waste categories. These color images are represented as three-dimensional matrices known as tensors, with each layer corresponding to a primary color channel (red, green, blue). These matrices encode the intensity values of each color at every pixel location within the image. By processing these matrices, neural networks can extract features and patterns to perform tasks like image classification. This representation facilitates the integration of visual information into the deep learning model's training process, enabling it to interpret and analyze images effectively.

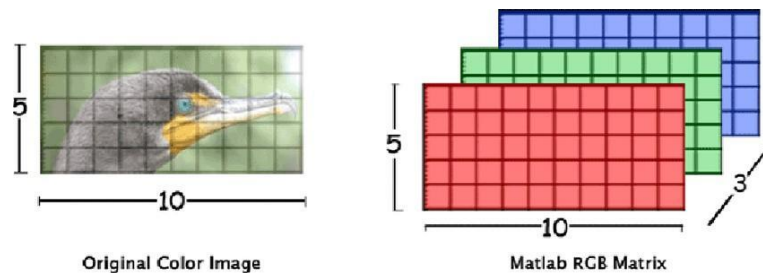


Fig.5. Color image representation and RGB matrix (Image Source [18])

Data Cleaning: The collected data is cleaned to remove samples that could negatively impact the model's performance, accuracy, and generalizability. This step aligns with the principle of "Garbage in, garbage out", where poor-quality data leads to poor model performance. Mislabelled samples reduce accuracy, while duplicated samples cause biased training and overfitting. Eliminating these samples improves the learning process by helping the model focus on meaningful patterns, leading to more accurate predictions.

Data Preprocessing: Data preprocessing involves several steps, including resizing images, encoding class labels, and structuring the data. All images with suitable file extensions (e.g., .jpg, .png) are resized to a uniform dimension of 224 x 224 pixels to ensure consistency and compatibility across the dataset. Resizing images to a uniform dimension reduces the computational burden during training by standardizing the input, thereby improving processing efficiency and training speed. Class labels are encoded using one-hot encoding, which converts categorical data into numerical data for easy interpretation by the model. This ensures that the model can process categorical data without misinterpreting the labels as ordinal. The dataset is also organized into folders, where each folder name corresponds to a specific class label and contains all the relevant images for that class. These preprocessing steps help stabilize the model's training process, improve learning efficiency, and support better generalization.

Data Splitting: To ensure effective model training and evaluation, the dataset undergoes a crucial step: data splitting. The dataset is split into three subsets: the training, validation, and test set. This separation ensures that the model can be trained on a large set of data (training), parameters can be tuned (validation), and its performance can be evaluated on unseen data (test). Augmentation and shuffling are applied to the training set to increase image variations and randomize the order of samples within each epoch. Two split ratios are used. The first is 70-15-15, 70% for training, 15% for validation, and 15% for testing and the second is 80-10-10, 80% for training, 20% for validation, and 10% for testing.

Data Normalization: In original RGB images, each pixel in each color channel is represented by a value corresponding to its intensity. These intensity values typically range from 0 to 255, where 0 represents no intensity (black) and 255 represents maximum intensity (full brightness). The pixel values of the images are normalized to a range of [0,1]. This normalization aids in speeding up the convergence of the neural network during training.

The normalization was performed using the formula:

$$\text{Normalized Pixel Value} = \frac{\text{Original Pixel} - \mu}{\sigma} \quad (12)$$

where μ is the mean of pixel values, and σ is the standard deviation of pixel values.

This standardization method centers the data around zero and scales it to have a standard deviation of 1, which can help improve training stability and convergence in machine learning models.

Data Augmentation: To enhance the robustness of the CNN model, data augmentation introduces variations to images using techniques such as random rotation, flipping, shifting, and zooming. These variations simulate how waste items might appear in different orientations, sizes, and positions in real-world scenarios. These variations introduced by data augmentation increase the effective size of the training dataset. This reduces the risk of overfitting, improving the model's generalization to new, unseen data, especially in cases of limited or imbalanced data. While accuracy offers an initial view of model performance, it may not reflect how well the model handles all classes. A high accuracy may indicate correct classification of majority classes, while the model underperforms on minority classes. To complement accuracy and provide a more comprehensive view of the model's performance, the F1-score is also considered. The F1-score, being the harmonic mean of precision and recall, offers a balanced measure of performance in situations where class imbalance exists.

3.2. Model Development

In the development of a Convolutional Neural Network (CNN) model, several key processes are involved to construct, optimize, and train the model effectively. These processes include building the CNN architecture, compiling the model with appropriate parameters and optimization algorithms, and training the model on the prepared training dataset. This section outlines the sequential steps involved in developing a CNN model.

CNNs were selected for this task due to their ability to automatically extract hierarchical spatial features from images, eliminating the need for manual feature engineering. This makes them effective for large and complex datasets compared to models such as SVMs or KNNs. When evaluated against these models, the CNN was found to be the best choice due to its capacity to learn directly from pixel data, their robustness to variations in image features, and their consistently higher classification performance.

Architecture: The proposed CNN model architecture for classifying hospital waste images is designed with a series of convolutional layers, activation functions, pooling layers, and fully connected layers. This section details the architectures of the proposed CNN models.

Model A:

- **Convolutional Layers:** This model is made up of four convolutional layers. Each of these convolutional layers uses the ReLU function to introduce non-linearity, enabling the model to learn complex patterns. The layers progressively increase the number of 3x3 filters from 32 to 128, improving the model's ability to capture finer details within the images.
- **Pooling Layers:** Following each convolutional layer, a max pooling layer with a pool size of (2x2) is used to reduce the spatial dimensions of the output. This step helps in reducing the computational complexity and also helps in extracting dominant features which are invariant to small translations. Consequently, the model includes a total of four pooling layers.
- **Flatten Layer:** After the final pooling layer, a Flatten layer is used to convert the 2D feature maps into a 1D feature vector. This transformation prepares the data for input into the fully connected layers.
- **Dropout Layer:** A dropout layer is included after the flatten layer. This layer randomly sets a fraction of input units to zero at each update during training time. This helps to prevent overfitting.
- **Fully Connected Layers:** The network includes two fully connected layers. The first has 512 units and uses the ReLU activation, serving as a fully connected layer that processes features extracted from the convolutional layers. The final layer is a dense layer with a softmax activation function. This outputs the probability distribution across the waste categories.

Model A is built using a 70-15-15 split ratio. Table 2 summarizes the architecture of the proposed model.

Model B:

Model B has the same architecture, optimization algorithm, and parameters as Model A. The only difference between the two models is that Model B is trained with an 80-10-10 split ratio, whereas Model A uses a 70-15-15 split.

Model C:

Model C is similar to Model A and B, having the same architecture, optimization algorithm, and parameters. The difference between Model C and the other models is that it employs k-fold cross-validation.

Table 2. Architecture of the proposed CNN model

Layer	Type	Activation Function	Number of Filters	Filter Size
1	Conv2D	ReLU	32	(3x3)
2	MaxPooling2D	-	-	(2x2)
3	Conv2D	ReLU	64	(3x3)
4	MaxPooling2D	-	-	(2x2)
5	Conv2D	ReLU	128	(3x3)
6	MaxPooling2D	-	-	(2x2)
7	Conv2D	ReLU	128	(3x3)
8	MaxPooling2D	-	-	(2x2)
9	Flatten	-	-	-
10	Dropout	-	-	-
11	Dense	ReLU	512	-
12	Dense	Softmax	num_of_classes	-

K-Fold Cross Validation is a robust method used to evaluate the performance of a machine learning model. This involves dividing the initial training set into K equally sized subsets or folds. In this study, K is set to 5, dividing the dataset into five equal subsets. The model is then trained and validated K times and at each time using a different fold as the validation set. The remaining K-1 folds serve as the training set. The results from each fold are averaged to produce a single estimate of model performance. In this particular model with five folds, each training set contains 80% of the initial training set, and each validation set contains 20%.



Fig.6. K-fold cross-validation

Given the metric for each fold is M_i , our overall metric M is:

$$M = \frac{1}{k} \sum_{i=1}^{k=5} M_i \quad (13)$$

Optimizer Algorithm: Several optimizers are commonly used for CNN-based image classification, including Stochastic Gradient Descent (SGD), Root Mean Square Propagation (RMSProp), and Adam. However, this work employs the Adam optimizer due to its adaptive learning rate and faster convergence in high-dimensional spaces.

Adam is a popular choice for training deep learning models due to its efficient computation of learning rates for each parameter. Adam combines the advantages of two other extensions of stochastic gradient descent: Adaptive Gradient Algorithm (Adagrad) that maintains a per-parameter learning rate that improves performance on problems with sparse gradients, and Root Mean Square Propagation (RMSProp) that handles non-stationary objectives. The Adam optimizer adjusts the learning rate by estimating the first (mean) and second (uncentered variance) moments of the gradients.

$$\theta_{t+1} = \theta_t - \frac{\eta \cdot m_t}{\sqrt{v_t + \epsilon}} \quad (14)$$

where θ represents the parameters (weights), η is the learning rate, m_t is the bias-corrected first moment estimate, v_t is the bias-corrected second moment estimate, and ϵ is a small constant for numerical stability.

3.3. Training

The models were trained locally using Jupyter Notebook on a computer equipped with an Intel(R) Core(TM) i7-

6600U CPU @ 2.60GHz and 16 GiB of RAM. The Adam optimizer was used with a learning rate of 0.0001, which was selected based on preliminary tests showing stable convergence at this rate. The models are trained with multiple batch sizes (16, 32, and 64) to evaluate training stability and convergence speed. Smaller batches like 16 yielded more stable training, while a batch size of 64 sometimes led to premature convergence. Each model was trained with two different epoch settings (30 and 50) to balance training time and model performance, while employing techniques to prevent overfitting.

To monitor training and improve performance, two techniques were employed: ModelCheckpoint and EarlyStopping.

- *ModelCheckpoint*: This technique is used to save the model's weight during training, particularly the best performing weights based on a specified metric (usually validation loss or accuracy). The ModelCheckpoint callback is set up to save the model's weights to a specified file path after each epoch only if the new model has achieved a better performance compared to the previously saved best model. This allows for the subsequent loading of the best-performing model for inference or further training.
- *EarlyStopping*: Overfitting occurs when a model learns noise and patterns that are specific to the training data but do not generalize well to unseen data. Early stopping is a technique used to prevent overfitting by monitoring the model's performance on a validation dataset. If the monitored metric (in our case, validation loss) does not improve for a specified number of epochs (patience), training is halted, and the model is restored to the best performing weights observed during training. This helps prevent unnecessary training epochs and ensures that the model does not overfit the training data.

3.4. Model Evaluation

At the end of the training phase, the trained model is used to make predictions on the test dataset, which contains data unseen by the trained model. This is done to assess the model's performance and generalization capabilities. Two important metrics in model testing are test loss and test accuracy.

- *Test Accuracy*: This metric measures the proportion of correctly classified samples in the testing dataset. It provides insight into the overall performance of the model in terms of its ability to correctly classify unseen examples. Higher values of test accuracy indicate better performance.
- *Test Loss*: This metric represents the average loss (or error) computed across all samples in the testing dataset. It indicates how well the model's predictions match the true labels in the testing dataset. Lower values of test loss indicate better performance, as they imply that the model's predictions are closer to the true labels.

4. Case Study and Validation

4.1. Context of the Case Study

This study uses the Medical Waste 4.0 Dataset, which consists of hospital waste images collected by Bruno et al[5], accessible via a Github repository. The dataset is a valuable resource for testing computer vision methods for hospital waste sorting. The images were acquired using an OAK-D camera, and each sample consists of three images: an RGB image and a stereo pair. The resolution of the RGB images is 1920 x 1080, while that of the grayscale images is 640 x 400. The full dataset is divided into two sub-datasets, named Dataset A and Dataset B. Since Dataset B was derived from Dataset A, Dataset A was used as the primary dataset for training, validation, and testing the hospital waste classification models.

The Medical Waste 4.0 Dataset was selected for its availability, and diverse representation of hospital waste categories aligned with WHO classification standards. It includes diverse visual representations of hospital waste, captured through RGB and stereo imaging, which simulate real-world conditions in resource-limited hospitals. While some class imbalances exist and the dataset does not fully represent every WHO-defined waste category, it remains one of the most relevant publicly available resources for this task. Medical waste types found in developing countries are generally consistent with those in more developed healthcare systems, further supporting the dataset's applicability. To address potential biases and improve model generalization, data augmentation techniques were applied during preprocessing. Future validation on additional datasets or in real hospital environments is recommended to further assess model generalizability.

The dataset contains a total of 5,520 images across 13 categories, namely: gauze, glove pair latex, glove pair nitrile, glove pair surgery, glove single latex, glove single nitrile, glove single surgery, medical cap, medical glasses, shoe cover single, shoe cover pair, test tube, urine bag.

After thorough data cleaning, duplicate or very similar images in separate categories were discarded. The following pairs were merged: "glove single nitrile" and "glove pair nitrile" into "glove nitrile", "glove single latex" and "glove pair latex" into "glove latex", "glove single surgery" and "glove pair surgery" into "glove surgery", and "shoe cover single" and "shoe cover pair" into "shoe cover". The final dataset used for model development comprised 3,903 images across 9 categories.

The image distribution, i.e., the number of files per category, is shown in Figure 8. This gives an idea of how

balanced or imbalanced the dataset is across the different categories.



Fig.7. Sample images of the dataset

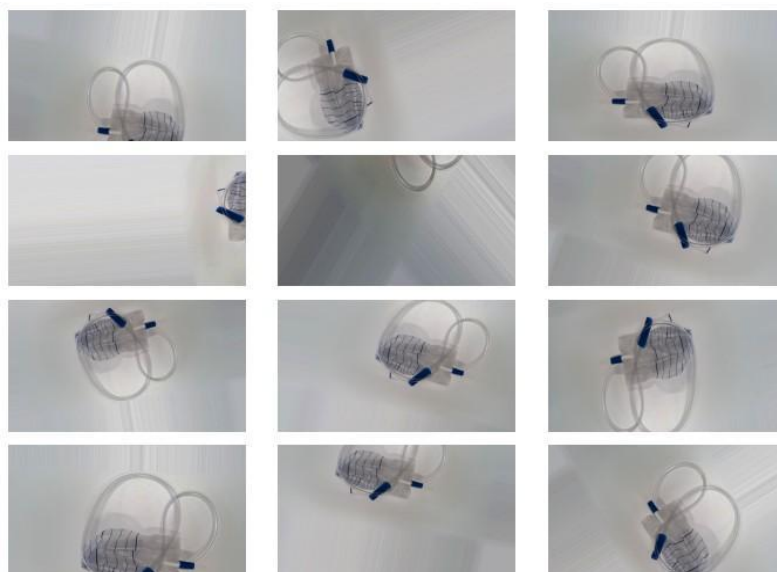


Fig.8. Data augmentation on a training set image

After going through data cleaning, all images were resized to 224 x 224 for size consistency throughout the dataset. The data was then split into two different ratios: 70-15-15 and 80-10-10. Figure 9 shows the counts of images in each subset for the two different ratio splits. The data is then normalized after the split, and the following data augmentation techniques were applied on the training set: brightness range of 0.1 to 0.7, rotation range of 90°, width and height shift range of 0.2, shear range of 0.2, zoom range of 0.2, horizontal and vertical flip set as True.

4.2. Results and Performance Evaluation

In this Results and Performance section, the built models are compared across different metrics and visualized through graphs to determine which model performs better. Multiple CNN models, built using various hyperparameters, and optimization algorithms have been evaluated. Table 3 summarizes the development parameters of the various

models to facilitate a comparison of their accuracies.

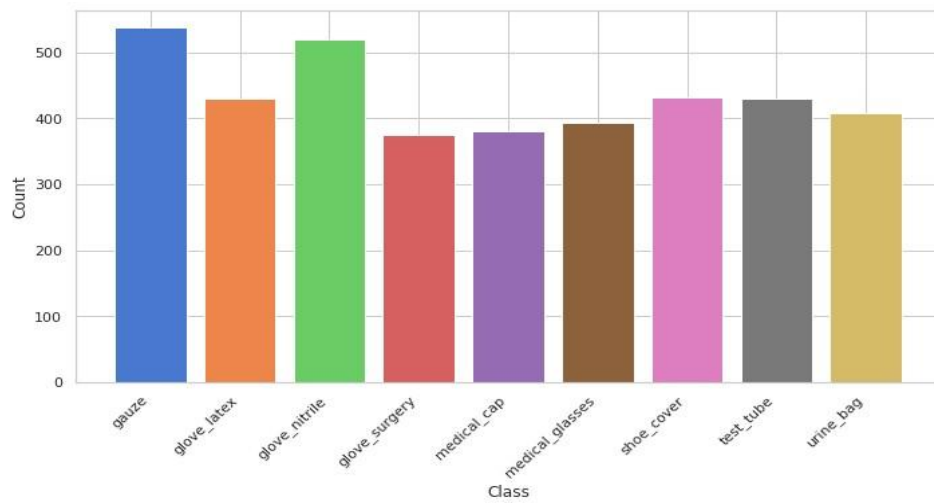


Fig.9. Dataset image distribution of categories

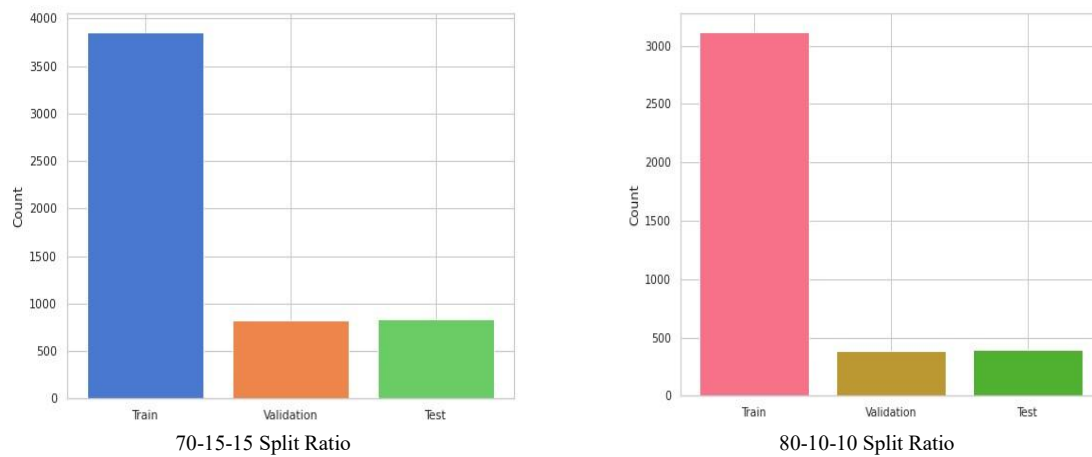


Fig.10. Data distribution per split ratio

Accuracy: The performances of CNN Models A and B are summarized in Table 3. In contrast, Model C, which employs k-fold cross-validation, is summarized in Table 4. Model A and Model B show a clear differentiation based on training data splits, with Model A being trained with splits of 70-15-15 and Model B on 80-10-10.

Table 3. Comparison of CNN models

CNN Model	Optimizer	Batch Size	Epochs	Accuracy
Model A (70-15-15 Split)	Adam	16	30	84.18%
	Adam		50	91.58%
	Adam	32	30	81.65%
	Adam		50	90.74%
	Adam	64	30	74.24%
	Adam		50	78.96%
Model B (80-10-10 Split)	Adam	16	30	89.47%
	Adam		50	84.96%
	Adam	32	30	84.71%
	Adam		50	82.46%
	Adam	64	30	83.46%
	Adam		50	85.46%

Model A, with a more balanced split showed varying levels of accuracy depending on the batch size and number of epochs. The best performance was observed with a batch size of 16 over 50 epochs, achieving an accuracy of 91.58%.

However, increasing the batch size consistently led to a drop in accuracy, which shows the sensitivity of the model performance to batch size. Model B, with a larger training set and smaller validation and test set, generally exhibited lower accuracies across most configurations compared to model A. The highest accuracy of Model B was 89.47%, achieved with a batch size of 16 over 30 epochs. Similar to model A, performances tend to decrease with larger batch sizes.

While Model A configurations with 50 epochs generally outperformed those with 30 epochs, Model B configurations with 30 epochs generally performed better. Techniques such as EarlyStopping and ModelCheckpoint were used, which allowed training to halt at the optimal point once overfitting began. The comparison of these two models' results shows the critical impact of data distribution and the training dynamics on the effectiveness of CNN models in classifying complex datasets like medical waste.

Our third model, Model C, was developed using the same architecture and was evaluated using K-fold cross-validation. In each iteration, the training set was 80% of the dataset and the validation set was 20%. The results of each iteration are summarized in Table 4.

Table 4. K-folds accuracies in model C

CNN Model	K-Folds	Accuracy
Model C	Iteration 1	96.16%
	Iteration 2	96.93%
	Iteration 3	98.85%
	Iteration 4	95.26%
	Iteration 5	98.21%

Model C was trained with a batch size of 16, an Adam optimizer and each iteration consisted of 10 epochs. The accuracy of Model C at each iteration was higher than that of Models A and B in their best configurations. One possible reason is that Model C utilized more training and validation data. The minimum accuracy was seen at Iteration 4 with 95.26% and the maximum accuracy was seen at Iteration 3 with 98.85%. The accuracy of Model C is the average of all five accuracies which gives us 97.08%. Figure 11 visualizes the accuracies of the various models, showing that Model C outperforms Models A and B.

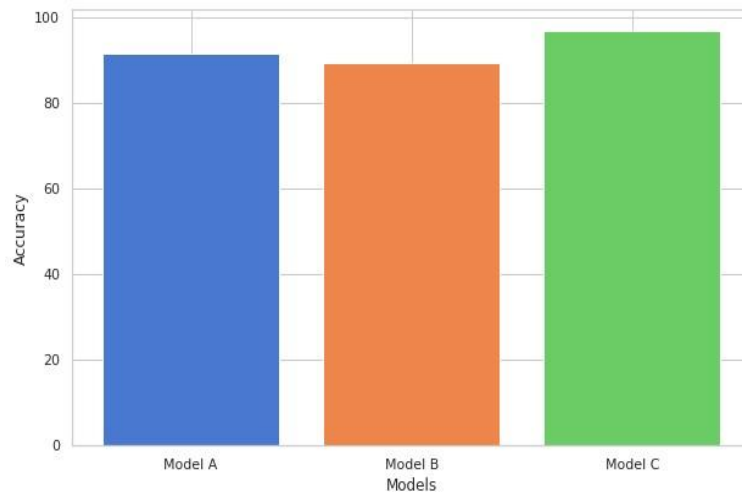


Fig.11. Comparison of model A, B, and C accuracies

Precision, Recall, F1-Score Results: The detailed performance metrics for the best configurations of Models A and B reveal insights into their ability to classify various hospital waste items from our dataset. Based on Model A results found in Table 5, we observe exceptionally high precision, recall, and F1-scores in most categories. This suggests a well-tuned model capable of generalizing effectively across different types of hospital waste. Nevertheless, the categories "gauze" and "glove_latex" showed lower performance in precision, recall, and consequently F1-score, with the lowest F1-score being 0.76. This may indicate challenges in distinguishing these items from each other due to visual similarities. Another reason could be the way these items appear within the dataset. These edge cases highlight the model's limitation in distinguishing between visually similar waste items. All other categories in Model A exhibit near-perfect scores, reflecting that the model can reliably identify and classify these items with high accuracy.

Model B also shows a good performance across many categories, although generally lower when compared to Model A. Categories such as "gauze" and "glove_latex" showed lower performance than other categories, with F1-scores of 0.75 and 0.52, respectively. These results reinforce the classification difficulty posed by these visually similar

items.

Model C, which was trained using a larger validation split (20%) and K-fold cross-validation, demonstrates the most robust and consistent performance. The model benefits from improved generalization and higher support across categories. Most precision, recall, and F1-score values exceed 0.90, including significant improvement in distinguishing “gauze” and “glove_latex”, achieving F1-scores of 0.97 and 0.96 respectively. This indicates a strong ability to separate even visually similar classes, suggesting more effective feature extraction and learning under cross-validation.

The use of F1-score as an evaluation metric was particularly important due to the presence of class imbalance in the dataset. F1-score provides a balanced view of precision and recall, offering a more informative measure of model performance than accuracy alone. Overall, Model C stands out as the best-performing model, combining high per-class performance with improved handling of difficult edge cases.

We next examine the confusion matrices for each of these models.

Table 5. Precision, recall, F1-score, and support on best model A

Class	Precision	Accuracy	F1-Score	Support
medical_cap	1.00	1.00	1.00	58.00
glove_nitrile	0.99	1.00	0.99	79.00
shoe_cover	1.00	0.97	0.98	66.00
glove_surgery	0.97	1.00	0.98	57.00
test_tube	0.97	0.98	0.98	65.00
urine_bag	0.86	0.98	0.92	62.00
medical_glasses	0.98	0.83	0.90	60.00
glove_latex	0.70	0.85	0.76	65.00
gauze	0.85	0.68	0.76	82.00

Table 6. Precision, recall, F1-score, and support on best model B

Class	Precision	Accuracy	F1-Score	Support
medical_cap	1.00	1.00	1.00	39.00
glove_nitrile	0.97	1.00	0.99	38.00
shoe_cover	1.00	0.93	0.96	44.00
glove_surgery	0.93	1.00	0.96	40.00
test_tube	0.98	0.95	0.96	42.00
urine_bag	0.95	0.95	0.95	44.00
medical_glasses	0.96	0.94	0.95	53.00
glove_latex	0.65	0.89	0.75	55.00
gauze	0.72	0.41	0.52	44.00

Table 7. Precision, recall, F1-score, and support on best model C

Class	Precision	Accuracy	F1-Score	Support
medical_cap	1.00	1.00	1.00	83.00
glove_nitrile	1.00	1.00	1.00	83.00
shoe_cover	1.00	1.00	1.00	73.00
glove_surgery	0.99	1.00	1.00	113.00
test_tube	1.00	0.96	0.98	75.00
urine_bag	0.97	0.99	0.98	88.00
medical_glasses	0.97	0.97	0.97	92.00
glove_latex	0.92	1.00	0.96	83.00
gauze	1.00	0.92	0.96	90.00

Confusion Matrix: The confusion matrix of each model offers detailed insights into the classification performance of each model across different categories of hospital waste. The diagonals in the confusion matrix represent correctly classified instances for each class, i.e., True Positives (TP).

Model A shows very good classification with a very high number of True Positives across the different categories. Categories like "glove_nitrile", "glove_surgery", and "medical_cap" have the highest true positive rates indicating effective recognition and classification. Other categories had at most one or two misclassifications. However, the categories "medical_glasses", "gauze", and "glove_latex" exhibited more misclassifications. The categories "gauze" and "glove_latex" showed the lowest TP counts, indicating the highest misclassification rates. Occasionally,

"medical_glasses" were misclassified as "urine_bag", while "urine_bag" was misclassified only once as "glove_nitrile".

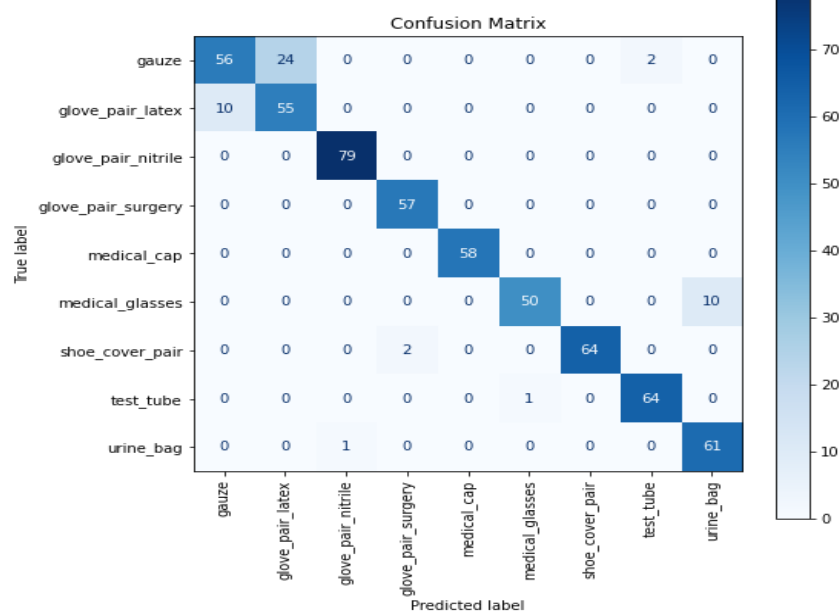


Fig.12. Confusion matrix - best A

Model B, generally shows a good performance as well, particularly in the categories "glove_nitrile", "glove_surgery", "medical_glasses" and "medical_cap" which still have a TP rate of 100%. There were few misclassifications in other categories, except for "gauze" and "glove_latex", which had a higher number of misclassifications despite having fewer test samples compared to Model A.

While Model A tended to maintain a more balanced performance across most categories, Model C showed near-perfect classification for most hospital waste categories, even though Models A and B struggled with "gauze" and "glove_latex". Despite the larger number of test samples, Model C still demonstrated better performance than Models A and B. This suggests that a larger training and validation set is crucial for achieving a good performance.

The categories that posed significant classification challenges for our model were "gauze" and "glove_latex". Examining sample images from these two categories provides insight into the classification difficulties.

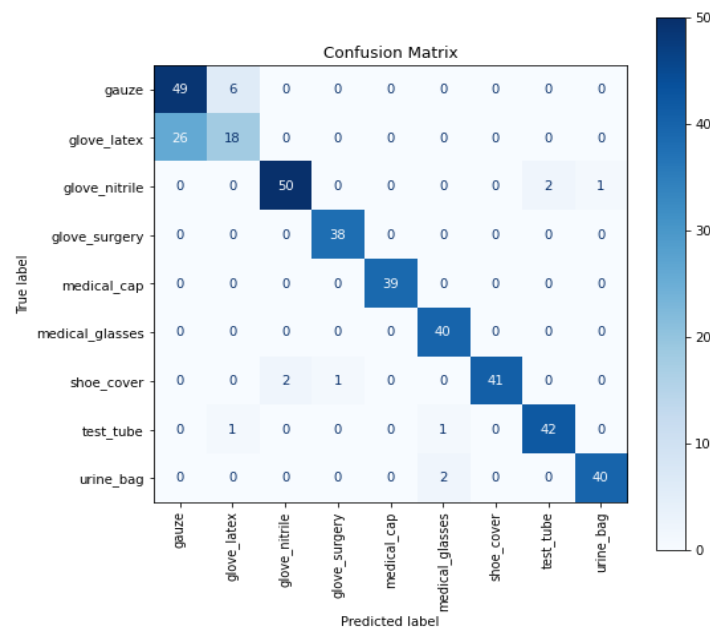


Fig.13. Confusion matrix - best B

Analyzing the images in the "gauze" and "glove_latex" categories helps explain the classification difficulties, as these items appear visually similar.

Model C demonstrated robust feature extraction capabilities, successfully classifying "gauze" and "glove_latex"

images despite their visual similarities.

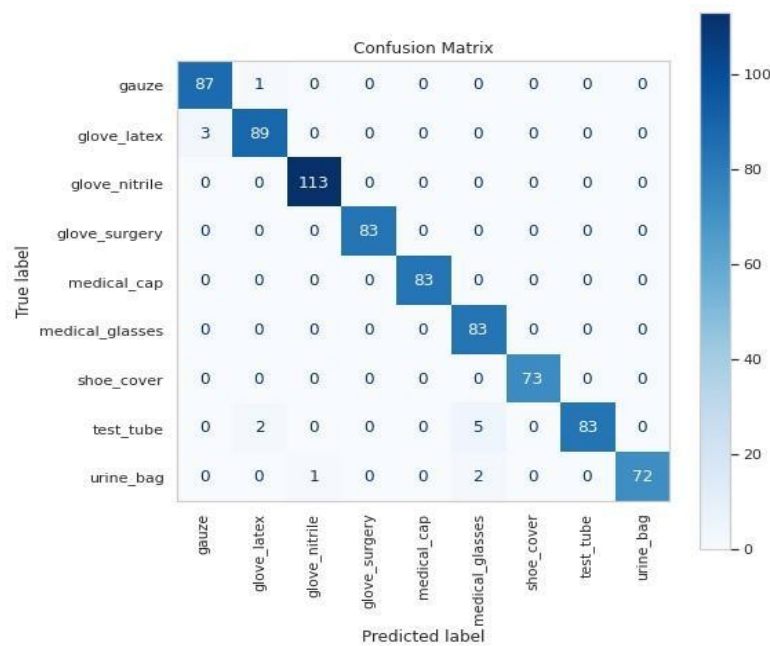


Fig.14. Confusion matrix - best C

Learning Curve: Model A shows a steady increase in training accuracy over 30 epochs, reaching a peak and maintaining stability, which indicates effective learning without significant overfitting. However, the validation accuracy showed slight fluctuations but remained close to the training curve, indicating good generalization. The loss plot shows a sharp initial decrease, stabilizing as epochs progress. There were also periodic spikes in validation loss, indicating moments of learning adjustments.

Model B shows a more rapid convergence in accuracy, achieving higher validation accuracy peaks compared to Model A. This suggests that the increased proportion of training data improved the model's ability to generalize, allowing for convergence in fewer epochs. However, the loss plot for Model B shows higher variability, especially in validation loss, which could indicate potential overfitting despite higher accuracy scores.

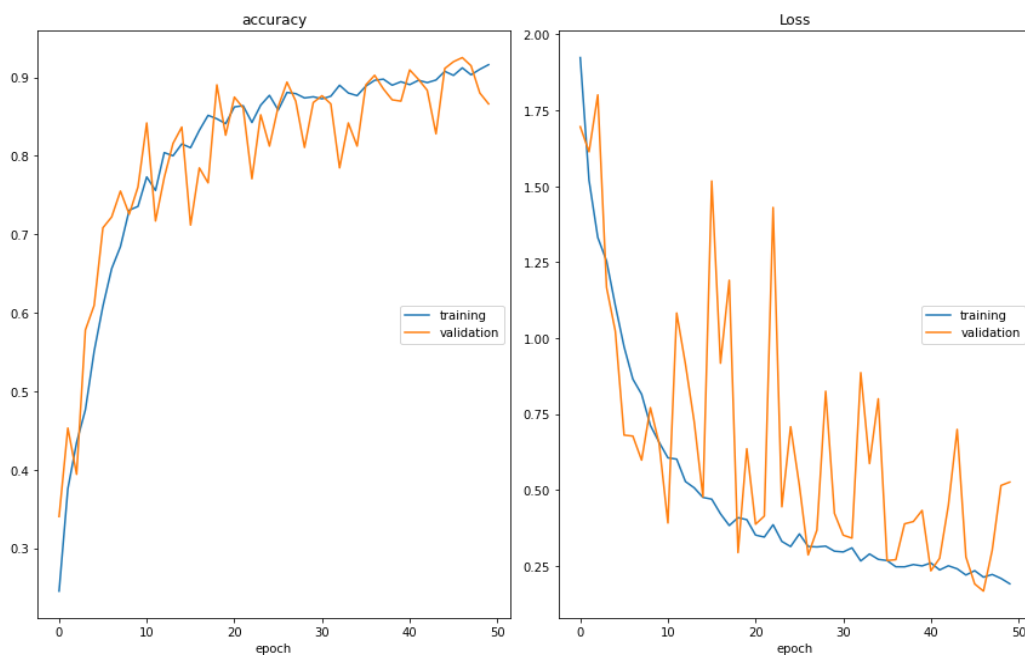


Fig.15. Learning curve - best A

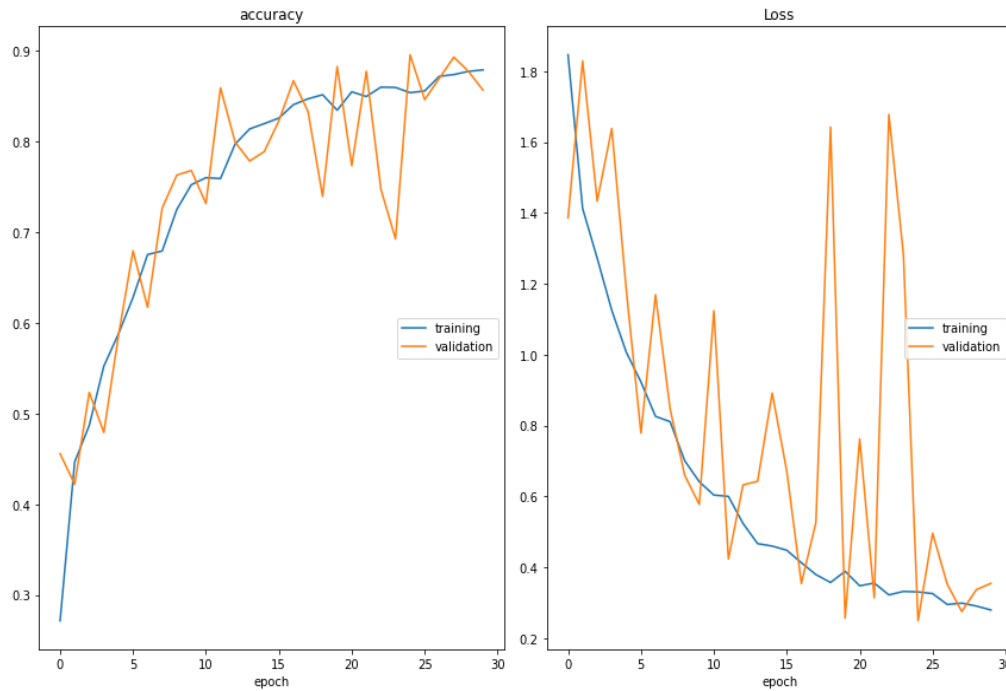


Fig.16. Learning curve - best B

Model C shows a steady increase over the epochs, indicating that the model is learning effectively from the training data. The validation accuracy also improved, although it showed more variability. Both curves exceed 0.95, suggesting the model is not overfitting and has good generalization performance.

The training loss decreases consistently, indicating the model is reducing its error on the training data. The validation loss follows a similar pattern. The final loss values are low, below 0.2. This indicates the model has learned effectively and is performing well on both the training and validation sets.

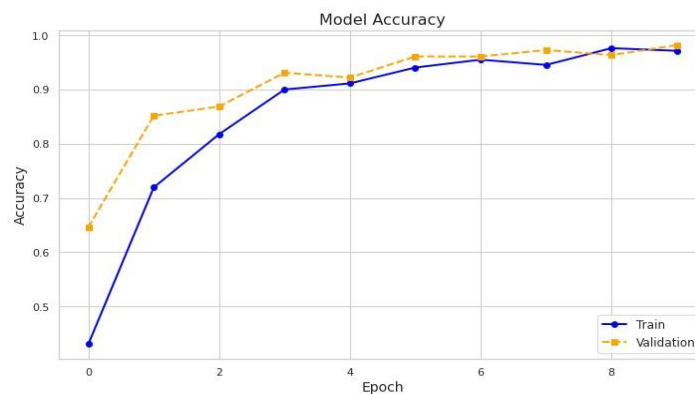


Fig.17. Accuracy learning curve of model C

The evaluation of our models in this study shows the significant potential of CNNs in improving hospital waste management. While Model C demonstrated superior performance in most configurations, with an accuracy of 97.08%, the other models highlighted the critical role of hyperparameter tuning in achieving high classification accuracy. This research paves the way for future developments in AI applications for sustainable waste management, emphasizing the need for further improvements in model robustness. The insights from this study are particularly important for advancing hospital waste management solutions in developing countries and improving public health and environmental protection.

4.3. Discussion and Perspective

The findings of this study highlight the practical viability and strong performance of AI-based approaches particularly CNNs in classifying and sorting hospital waste. With an accuracy of 97.08% and high F1-scores, the proposed model demonstrates strong potential for improving waste management efficiency, reducing dependency on manual sorting, and minimizing human exposure to hazardous materials. These results validate the applicability of deep learning techniques in addressing real-world environmental and health challenges, especially in settings where

traditional waste management methods are inadequate or unsafe.

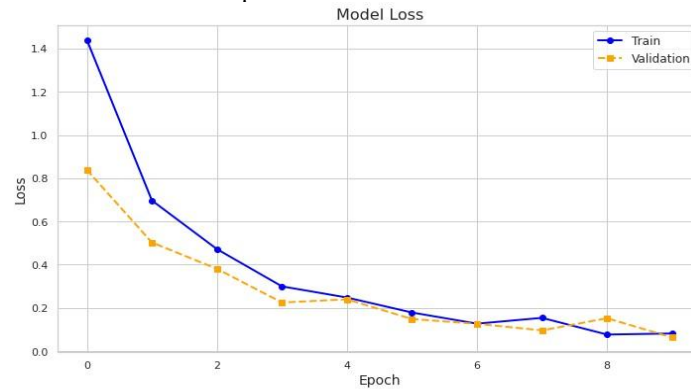


Fig.18. Loss learning curve of model C

Despite the promising results, the implementation of such AI-based systems in developing countries remains constrained by several challenges. These include limited access to digital infrastructure, inconsistent data quality, lack of skilled personnel, and the high cost of deploying and maintaining intelligent systems. In addition, AI applications in healthcare and environmental sectors raise ethical concerns related to data privacy, fairness, and accountability. Without addressing these limitations, there is a risk that such technologies may fail to deliver meaningful impact or could even exacerbate existing inequalities in healthcare service delivery and waste management practices.

Advancing toward sustainable adoption of AI in healthcare waste management requires a comprehensive, multi-pronged approach. Key components include the development of well-defined regulatory frameworks to guide AI integration, the promotion of public-private partnerships to harness financial and technical resources, and the creation of context-specific solutions tailored to the operational realities of local hospital settings. Strengthening local capacity through education and skill development is vital to ensure the effective operation, maintenance, and scalability of these technologies. Additionally, the phased rollout of pilot initiatives, with active engagement from both communities and institutions, can provide valuable insights for refining AI models and evaluating their long-term viability. With strategic planning and inclusive governance, AI technologies can significantly enhance the safety, efficiency, and sustainability of healthcare waste management systems throughout the Global South.

5. Conclusions

This work focused on harnessing the capabilities of Convolutional Neural Networks (CNNs) to improve the classification and segregation of hospital waste. This is a crucial aspect of healthcare management that impacts both public and environmental health, especially in resource-constrained and developing countries. By leveraging deep learning techniques, we aimed to develop a model that could accurately identify and categorize various types of hospital waste from images (visual data). Our objective was to design and test CNN models that could enhance the efficiency and accuracy of hospital waste management systems. We sought to evaluate these models across different configurations and training splits to understand their performance dynamics and identify optimal conditions for effective waste classification. Although this study was based on a specific dataset, future validation on external or real-world datasets is necessary to ensure robustness across diverse hospital environments. The results demonstrated that our CNN models, particularly Model C which used k-fold cross validation, achieved an accuracy of 97.08% in classifying medical waste types. Due to class imbalance present in certain categories, the F1-score was analyzed to provide a more balanced evaluation of the model's performance. Model C recorded excellent F1 scores ranging from 0.96 to 1 across various classes, reinforcing its status as the best-performing model, reflecting the model's true general performance. The performance disparities between the categories "gauze" and "glove_latex" suggested the need for more robust feature extraction techniques or enhancing datasets that have better representations of these various categories. The performance evaluation revealed that specific model parameters and training conditions significantly influenced the models' effectiveness. This shows the importance of robust dataset preparation in developing practical AI applications for hospital waste management. Scalability to larger datasets and adaptability to more complex hospital waste scenarios remain critical next steps. Real-world implementation challenges such as infrastructure limitations, equipment availability, and staff training must also be considered when deploying such AI systems in hospital environments.

All the Declarations and Statements

Authors Contributions Statement

Teddy Ako Eyonganyoh - Methodology, Data Curation, Software, Validation, Formal Analysis, Visualization, Writing - Original

Draft: Designed the methodological framework, performed data cleaning and preprocessing, implemented the research model, trained and validated models, conducted comparative performance evaluation using standard evaluation metrics, analyzed experimental results, prepared performance visualizations, drafted the initial manuscript, and conducted literature review.

Samuel Kotva Goudougou - Investigation, Validation, Formal Analysis, Visualization, Writing - Review & Editing: Provided technical guidance on model implementation, assisted in performance evaluation, contributed to result interpretation and visualization, and reviewed and edited the manuscript for clarity and coherence.

Justin Moskolai Ngossaha - Conceptualization, Data Curation, Supervision: Proposed the research topic and ideas, provided the dataset, constructed the overall framework, and supervised project execution.

Paul Dayang - Project Administration, Writing - Review & Editing: Coordinated project milestones and timelines, reviewed and refined the manuscript for clarity and coherence.

All authors have read and agreed to the published version of the manuscript.

Conflict of Interest Statement

The authors declare no conflicts of interest.

Funding Declaration

None

Data Availability Statement

This study analyzed publicly available datasets. The datasets can be found here: <https://zenodo.org/records/7709001>, accessed on Mar. 25, 2026.

Ethical Declarations

Not applicable. This study does not involve human or animal subjects.

Acknowledgements

The authors express their deepest gratitude to the African Institute for Mathematical Sciences (AIMS) Cameroon for providing the essential academic environment, facilities, and resources that made this research possible. We sincerely thank the journal reviewers and experts for their professional evaluation and valuable recommendations, which have contributed to improving the quality of the manuscript and the clarity of the results.

Declaration of Generative AI in Scholarly Writing

The authors used AI-assisted tools to strengthen the language and readability of the manuscript. All outputs were reviewed and edited with human oversight, and the authors take full responsibility for the final content.

Abbreviations

The following abbreviations are used in this manuscript:

AI - Artificial Intelligence
ANN - Artificial Neural Network
CNN - Convolutional Neural Network
DT - Decision Tree
GIS - Geographic Information System
KNN - K-Nearest Neighbors
ReLU - Rectified Linear Unit
RGB - Red, Green, Blue
RMSProp - Root Mean Square Propagation
SGD - Stochastic Gradient Descent
SVM - Support Vector Machine
WHO - World Health Organization

References

- [1] Chisholm, J. M., Zamani, R., Negm, A. M., Said, N., Abdel Daiem, M. M., Dibaj, M., and M. Akrami, "Sustainable waste management of medical waste in African developing countries: A narrative review," *Waste Management & Research*, vol. 39, no. 9, pp. 1149-1163, 2021, doi: 10.1177/0734242X211029.
- [2] Ranjbari, M., Esfandabadi, Z. S., Shevchenko, T., Chassagnon-Haned, N., Peng, W., Tabatabaei, M., and M. Aghbashlo, "Mapping healthcare waste management research: Past evolution, current challenges, and future perspectives towards a circular economy transition," *Journal of Hazardous Materials*, vol. 422, p. 126724, 2022, doi: 10.1016/j.jhazmat.2021.126724.
- [3] Andeobu, L., Wibowo, S., and S. Grandhi, "Artificial intelligence applications for sustainable solid waste management practices in Australia: A systematic review," *Science of The Total Environment*, vol. 834, p. 155389, 2022, doi: 10.1016/j.scitotenv.2022.155389.
- [4] Singh, N., Ogunseitan, O. A., and Y. Tang, "Medical waste: Current challenges and future opportunities for sustainable

- management," *Critical Reviews in Environmental Science and Technology*, vol. 52, no. 11, pp. 2000-2022, 2022, doi: 10.1080/10643389.2021.1885325.
- [5] Fang, B., Yu, J., Chen, Z., Osman, A. I., Farghali, M., Ihara, I., et al., "Artificial intelligence for waste management in smart cities: A review," *Environmental Chemistry Letters*, vol. 21, no. 4, pp. 1959-1989, 2023, doi: 10.1007/s10311-023-01604-3.
 - [6] Justin, M. N., Ali, Y., Apollinaire, B. B., Rodrigue, D., Samuel, B. T., and A. E. Adolphe, "Towards a Flexible Urbanization Based Approach for Integration and Interoperability in Heterogeneous Health Information Systems: Case of Cameroon," in *International Conference on Safe, Secure, Ethical, Responsible Technologies and Emerging Applications*, pp. 258-275, 2023, doi: 10.1007/978-3-031-56396-6_16.
 - [7] Bruno, A., Caudai, C., Leone, G. R., Martinelli, M., Moroni, D., and F. Crotti, "Medical waste sorting: A computer vision approach for assisted primary sorting," in *2023 IEEE International Conference on Acoustics, Speech, and Signal Processing Workshops (ICASSPW)*, pp. 1-5, 2023, doi: 10.1109/ICASSPW59220.2023.10193520.
 - [8] Abdallah, M., Talib, M. A., Feroz, S., Nasir, Q., Abdalla, H., and B. Mahfood, "Artificial intelligence applications in solid waste management: A systematic research review," *Waste Management*, vol. 109, pp. 231-246, 2020, doi: 10.1016/j.wasman.2020.04.057.
 - [9] Odonkor, S. T., and T. Mahami, "Healthcare waste management in Ghanaian hospitals: Associated public health and environmental challenges," *Waste Management & Research*, vol. 38, no. 8, pp. 831-839, 2020, doi: 10.1177/0734242X20914748.
 - [10] Kaushal, R., Rohit, and A. K. Dhaka, "A comprehensive review of the application of plasma gasification technology in circumventing the medical waste in a post-COVID-19 scenario," *Biomass Conversion and Biorefinery*, vol. 14, no. 2, pp. 1427-1442, 2024, doi: 10.1007/s13399-022-02434-z.
 - [11] Janik-Karpinska, E., Brancaloni, R., Niemcewicz, M., Wojtas, W., Foco, M., Podogrocki, M., and M. Bijak, "Healthcare waste - a serious problem for global health," *Healthcare (Basel)*, vol. 11, no. 2, p. 242, 2023, doi: 10.3390/healthcare11020242.
 - [12] Vu, H. L., Bolingbroke, D., Ng, K. T. W., and B. Fallah, "Assessment of waste characteristics and their impact on GIS vehicle collection route optimization using ANN waste forecasts," *Waste Management*, vol. 88, pp. 118-130, 2019, doi: 10.1016/j.wasman.2019.03.037.
 - [13] Xia, W., Jiang, Y., Chen, X., and R. Zhao, "Application of machine learning algorithms in municipal solid waste management: A mini review," *Waste Management & Research*, vol. 40, no. 6, pp. 609-624, 2022, doi: 10.1177/0734242X211033716.
 - [14] Herath, H. M. K. K. M. B., and M. Mittal, "Adoption of artificial intelligence in smart cities: A comprehensive review," *International Journal of Information Management Data Insights*, vol. 2, no. 1, p. 100076, 2022, doi: 10.1016/j.jjimei.2022.100076.
 - [15] De Hitta, M. A. S., Castro, E. C., Felix, M. J. B., and A. C. Blanco, "Estimation of PM concentrations in the Philippines using Sentinel-5P and artificial neural network," in *Remote Sensing of the Atmosphere, Clouds, and Precipitation VIII*, vol. 13262, pp. 129-142, 2025, doi: 10.1117/12.3042769.
 - [16] Kwie, M. Lie., "Authenticity vs. Idealization: Unravelling the Impact of TikTok Influencers on Dutch Young Women's Well-being through Serial Mediation of upward Social Comparison and Relative Deprivation." (2023).
 - [17] Gomes, D., "Classification of food objects using deep convolutional neural network using transfer learning," *International Journal of Education and Management Engineering*, vol. 14, no. 2, pp. 53-60, 2024.
 - [18] Jamaludin, J., Rahim, R. A., Rahim, H. A., Rahiman, H. F., Muji, S. Z. M., Fadzil, N. S. M., et al., "Introducing an application of a charged coupled device (CCD) in an optical tomography system," *Jurnal Teknologi (Sciences & Engineering)*, vol. 73, no. 3, 2015, doi: 10.11113/jt.v73.4253.

Authors' Profiles



Teddy Ako Eyonganyoh received a Master of Science degree in Mathematical Sciences with a major in Data Science from the African Institute for Mathematical Sciences (AIMS), Cameroon, in 2025. His research interests include machine learning, deep learning, artificial intelligence, multi-agent systems and software engineering. He is currently pursuing a second master's degree in software engineering at the University of Buea, Cameroon.



Samuel Kotva Goudougou received a Master's degree in Software and Systems in Distributed Environments from the University of Ngaoundéré, Cameroon, in 2021. His research interests include the use of machine learning techniques to solve certain problems, Sustainable Systems and Intelligent mobility systems. He is currently a PhD student in computer science at The University of Ngaoundere. His research activities deal with Interoperability, Distributed Systems, Opportunities Networks, Decision Support Systems, Knowledge Engineering, Sustainable Systems, Intelligent mobility systems and Cybersecurity.



Justin Moskolai Ngossaha is Senior Lecturer in the Mathematics and Computer Science Department of the Faculty of Science of the University of Douala (Cameroon). He received the PhD degree in Computer Sciences from the National Polytechnic Institute of Toulouse in France in 2018. His research activities deal with Interoperability, Distributed Systems, Decision Support Systems, Knowledge Engineering, Sustainable Systems and Intelligent mobility systems.



Paul Dayang received the Ph.D. degree in computer science in 2014 from the University of Bremen, Germany. He is an Assistant Professor at the University of Ngaoundéré, Cameroon. His research interests include the area of information systems, intelligent systems, and information retrieval techniques for natural language processing with focus on resource-scarce African languages.

How to cite this paper

Teddy Ako Eyonganyoh, Samuel Kotva Goudoungou, Justin Moskolai Ngossaha, Paul Dayang, "AI-driven Innovations for Sustainable Hospital Waste Management in Developing Countries", International Journal of Intelligent Systems and Applications(IJISA), Vol.18, No.4, pp.205-225, 2026. DOI:10.5815/ijisa.2026.04.11