Heliyon 10 (2024) e29966

Contents lists available at ScienceDirect

Heliyon

journal homepage: www.cell.com/heliyon

Research article

Application research of image classification algorithm based on deep learning in household garbage sorting

Jianfei Wang

Suzhou Chien-Shiung Institute of Technology, Taicang, 215411, China

ARTICLE INFO

Keywords:
Household garbage classification
Deep learning
Machine vision
Error-correcting output codes
Capuchin search algorithm
Convolutional neural network

ABSTRACT

The classification of garbage types is an important issue in today's world, and its proper implementation can contribute to environmental conservation and improved efficiency of recycling processes. Unfortunately, the classification of garbage types is currently predominantly performed through human supervision, which leads to high errors and environmental risks. It is crucial to automate this procedure utilizing machine vision methods as a result. This research proposes a revolutionary deep learning-based strategy for classifying domestic waste. The suggested method uses deep learning methods to extract information from images. The Capuchin Search Algorithm (CapSA) is used to improve the hyperparameters of the convolutional neural network (CNN) used as the feature extraction model. Furthermore, for categorizing the retrieved features from the CNN model, a hybrid learning model based on Error-Correcting Output Codes (ECOC) and Artificial Neural Networks (ANN) is used. The classification accuracy may be successfully increased by using this hybrid model, and the benefit becomes more pronounced as the number of target categories rises. The TrashNet and HGCD databases were used to assess the suggested method's effectiveness, and its results in waste type detection were contrasted with those of earlier techniques. Based on the study findings, the suggested approach can identify trash types with an accuracy of 98.81 % and 99.01 % on the TrashNet and HGCD databases, respectively. This is at least a 1.46 % improvement over earlier approaches. The study's conclusions validate that the suggested strategy may be used in real-world scenarios and show how successful the approaches used in it are.

1. Introduction

In today's world, the high volume of daily garbage has become a major challenge and threat to the environment. In every community, households are one of the primary garbage producers. Some countries, such as China and the United States, generate the highest amount of household garbage. In the United States, for instance, the average daily garbage production per person is about 5 pounds, and the average garbage produced by each household is around 18 pounds [1]. However, with the current recycling systems, only 1.5 pounds of the generated garbage per person can be effectively recycled [2]. While some developed countries like Sweden have been able to achieve high recycling rates by reforming societal norms [3], this process still relies on manual garbage separation in the initial stages. Therefore, achieving such a system in densely populated communities would be costly and prone to significant errors [4]. Recent research indicates that by utilizing machine vision techniques and image processing, garbage classification can be performed accurately and efficiently [5,6]. By developing an automated and efficient system for household garbage classification, it is possible to

E-mail address: 2031546@tongji.edu.cn.

https://doi.org/10.1016/j.heliyon.2024.e29966
Received 29 July 2023; Received in revised form 17 April 2024; Accepted 18 April 2024
Available online 19 April 2024
2405-8440/

not only reduce costs and environmental risks but also significantly increase the recycling rate of household garbage [7].

The key requirement for achieving an automated and efficient system for household garbage classification from a machine vision perspective is to fulfill two fundamental prerequisites. Firstly, the input images need to be properly processed, and the descriptive features of the garbage type should be optimally extracted. The execution of this process necessitates the utilization of feature extraction models, which have not received much attention in previous research, and often pre-configured static models (such as CNN models with predefined configurations) have been used for image feature extraction. However, with an optimized configuration of the feature extraction model, its performance can be improved. Secondly, the classification of the extracted features from the images needs to be performed accurately. Considering the large number of target categories (types of garbage) in the research problem, most classification models may not efficiently accomplish this task. It seems possible to solve this sub-problem by combining multiple learning models. In the proposed method, for addressing the first sub-problem, the optimization strategy of configuring CNN based on the Capuchin algorithm (CapSA) will be employed. Additionally, the second sub-problem will be addressed through the combination of Artificial Neural Networks (ANNs) in the form of an Error Correcting Output Codes (ECOC) model. Both of these strategies have formed a novel model for garbage sorting. With these explanations, the contribution of the current paper is as follows:

- To characterize the characteristics pertinent to the different forms of household rubbish, a feature extraction model based on the CNN structure is described in this paper. This method uses the CapSA to optimize the CNN model's hyperparameters. The Capuchin algorithm optimizes the type of pooling functions, number of filters, and also, dimensions for each convolutional layer based on training error as its objective.
- In addition, a hybrid learning model for the categorization of the characteristics collected from the photos is suggested, which is based on ECOC and ANNs. Using this combination approach may help increase the accuracy of categorization, particularly when dealing with situations that include a lot of categories.

This article's continuation is structured as follows: A summary of earlier studies is given in section two. The suggested approach for categorizing domestic trash using deep learning methods is explained in section three. The outcomes of the application and assessment of the suggested approach are covered in section four. The collected findings are examined in Section Five. Lastly, suggestions are made for more study in this field.

2. Literature review

In recent years, the classification of different types of garbage has become a subject of research due to its environmental significance. In this section, we review some of these studies. In Ref. [8], an automatic system based on deep learning for garbage classification is proposed. This research focuses on smartening garbage bins and utilizes the ResNet-34 model for garbage classification. The proposed smart garbage bin in this study uses a solar panel to supply energy to the chip. When garbage is placed in the upper compartment of the bin, its image is processed by a ResNet-34 model, and the recyclability of the garbage is determined. Based on the classification result, the garbage is then thrown into one of the lower compartments labeled as "recyclable" or "non-recyclable". Another model presented in Ref. [9] employs a similar strategy for smartening garbage bins. The proposed model in this study is implemented based on the Raspberry Pi architecture and can categorize garbage into one of the 4 target classes. This approach uses an improved version of MobileNetV3 called GNet for garbage type detection, which combines the structure of the mentioned CNN model with an attention mechanism.

The research conducted in Ref. [10] is one of the first studies that addresses the problem of garbage classification in the real world and utilizes a real-world database. Despite their high performance, deep learning models face challenges in classifying types of garbage under real-world conditions. These challenges include susceptibility to noise, the requirement for a big training set, and high costs. In this method, a novel incremental learning framework called GarbageNet is proposed to address these challenges. The method uses a weakly supervised transfer learning strategy to ensure the quality of feature extraction. Additionally, for classifying new garbage samples, it leverages the nearest neighbor criterion concerning the test samples. Furthermore, it utilizes a combination of training data to mitigate the negative impact of noise and incorrectly labeled data.

In [11], a garbage classification model in public spaces is presented using convolutional neural networks (CNNs). This CNN model, named PublicGarbageNet, can classify garbage images into 4 main classes and 10 sub-classes. The PublicGarbageNet model is based on the Backbone architecture, through which image features are extracted, and then the main/sub-classes are determined based on feature maps. Additionally, in this research, a new database for garbage detection, in line with real-world scenarios, has been collected.

In [12], the performance of various deep learning models for garbage classification is compared. The transfer learning mechanism is applied to RseNet, VGG-16, and Xception models, and the accuracy of these models in classifying household garbage types is compared. According to the results reported in this article, a combination of transfer learning and the Xception model can lead to higher detection accuracy. The method proposed in Ref. [13] can be considered an extension of the research presented in Ref. [12]. In this model, it is first demonstrated that using multiple layers in deep neural networks and employing shortcut connections between layers (which have been widely used in the research literature) cannot effectively improve detection accuracy. Then, a model based on the Xception architecture is proposed for classifying garbage types. This approach extends the network with branch expansions and utilizes additional layers to capture the combination of feature information. According to the reported results, using this method to replace the main structure of the Xception network can improve its performance in household garbage classification.

In [14], deep learning techniques are used for garbage detection and management in smart cities. The proposed deep model in this research is called RefineDet, which consists of two parallel and interconnected convolutional neural networks. Each convolutional component in RefineDet includes 4 convolutional blocks. The first convolutional component is used for binary image classification, allowing the identification of garbage presence in the image. The second convolutional component is used for multi-class image classification, enabling applications such as garbage type recognition. The method presented in Ref. [15] also introduces a hardware-intensive model based on the Raspberry Pi architecture for garbage classification, utilizing deep learning techniques as its detection model. In this research, the performance of four deep models, namely ResNet, AlexNet, VGG-16, and InceptionNet, is compared, and among them, the InceptionNet model demonstrates better performance.

In [16], a combination of transfer learning and convolutional model is used for classifying different types of garbage. The proposed model, called GCNet, first extracts image features and then utilizes EfficientNet, Vision Transformer, and DenseNet models to form the GCNet structure. In this approach, data augmentation is also employed to address the limited number of training samples. The method presented in Ref. [17] uses the DenseNet model for real-time classification of various types of garbage using mobile phones. The proposed DenseNet model consists of 4 dense blocks, and the number of neurons in each block is adjusted using a genetic algorithm. This model can classify garbage images into six predefined categories. Although this model may not have high accuracy, its high processing speed makes it suitable for real-time scenarios.

In [18], a method for detecting the fullness of garbage bins using deep learning techniques is presented. This problem can be important in areas such as process automation in smart cities. The study compares the performance of four models, namely VGG19, MobileNetV3, EfficientNet, and DenseNet, for solving the mentioned problem. The reported results indicate that the DenseNet model can perform the detection with higher accuracy.

In [19], a model is proposed to address the challenges of garbage classification using deep learning techniques. This method gets around the overfitting problem by using Dropout. Additionally, the ReLU layers are used to resolve gradient sparsity during network training, enabling targeted processing of image input, and extracting edge features. The Adagrad adaptive learning rate approach is also used for adjusting the parameters of the deep neural network. Moreover, picture categorization employs a modified probability density function. Researchers in Ref. [20], devised a novel object-recognition method to distinguish and classify discarded digital cameras by examining their labels. This approach attained a 92 % accuracy rate in identifying camera models and a 48 % accuracy rate in recognizing manufacturers.

Researchers [21] introduced a novel optimization model called Stochastic Energy Consumption Disassembly Line Balancing Problem (SEDLBP) to address the challenges of uncertainty in the disassembly process and energy consumption. They also designed an enhanced social engineering optimization algorithm to tackle the SEDLBP efficiently. The developed algorithm proved its effectiveness by outperforming other well-established intelligent algorithms. The examined works are summarized in Table 1.

3. Research methodology

In this section, the database used in this research is described, followed by a description of the steps for classifying household garbage in the proposed approach.

3.1. Database

Two separate databases were used in this research for the classification of household garbage. The first dataset utilized in this study is TrashNet [22]. TrashNet is one of the most commonly used datasets in the field of garbage classification. This database consists of 2527 image samples with the RGB color system, categorizing different types of garbage into six classes: 1- Glass (501 samples), 2- Paper (594 samples), 3- Cardboard (403 samples), 4- Plastic (482 samples), 5- Metal (410 samples), and 6- Trash (137 samples). All samples in the database have dimensions of 512 × 384 pixels. All images were captured under consistent lighting conditions, and the background of all samples was white. Fig. 1 shows some samples from the TrashNet database.

The second dataset used in the current research is HGCD [23]. This database was created to address the limitations of previous similar databases, such as limited sample size and a small number of target categories. The HGCD database contains 15,150 picture samples that divide domestic rubbish into 12 categories: paper, cardboard, biological, metal, plastic, green glass, brown glass, white

Table 1
Overview of the literature works.

Ref.	Year	Model used	limitations
[8]	2020	ResNet-34	Limited to recyclable and non-recyclable garbage
[9]	2021	GNet	Limited to 4 target classes
[10]	2021	GarbageNet	Susceptible to noise, requires large training data, high costs
[11]	2020	PublicGarbageNet	Limited to public spaces, it requires a real-world database
[12]	2020	Xception transfer learning	Sensitive to noise, it requires large training data, high costs
[13]	2020	Multi-branch channel expansion network	Requires large training data
[14]	2022	RefineDet	Can only classify garbage presence in the image
[15]	2022	Hardware-intensive model	Limited application and dependent on Raspberry Pi architecture
[16]	2022	GCNet	Sensitive to noise, it requires large training data, high costs
[17]	2022	DenseNet	Limited to six predefined categories, may not have high accuracy
[18]	2023	DenseNet	Limited to detecting the fullness of garbage bins
[19]	2023	Dropout, Adagrad adaptive learning rate, ReLU layers, modified PDF	Limited evaluation scenarios

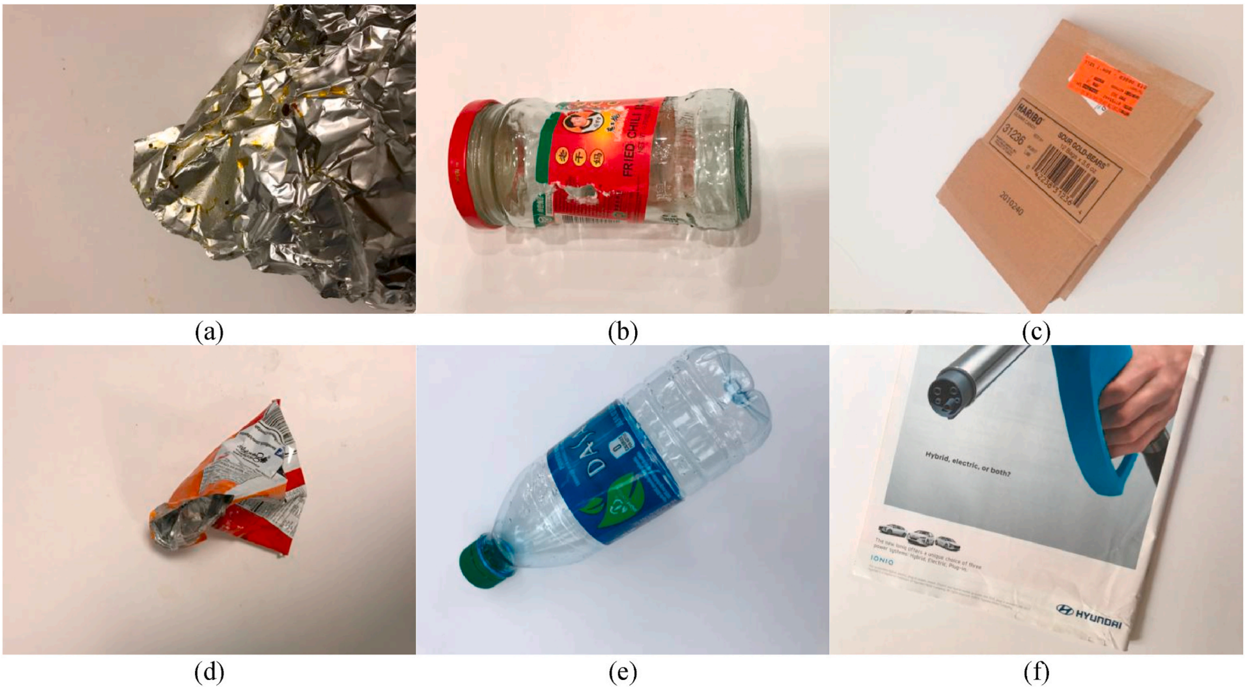


Fig. 1. Some samples from the TrashNet database belonging to categories (a) metal, (b) glass, (c) cardboard, (d) trash, (e) plastic, and (f) paper.

glass, cloth, shoes, batteries, and trash. All photographs in this collection were recorded under varied lighting situations and used the RGB color scheme. Additionally, the dimensions of the samples in this database vary. The diversity in the photography conditions and the larger number of target categories in HGCD make the classification of samples more challenging compared to TrashNet. In this research, the classification of samples from these two databases using the proposed method was studied separately.

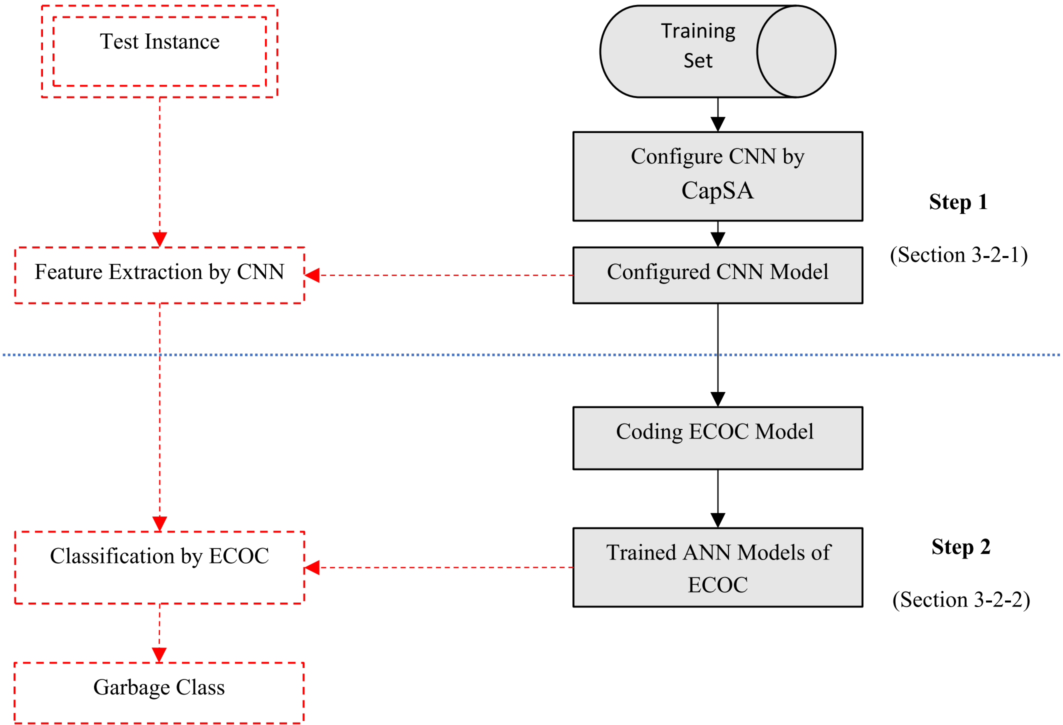


Fig. 2. Structure of household garbage classification in the proposed method.

3.2. Classification of household garbage in the proposed method

The suggested strategy in this paper combines deep learning techniques, ensemble learning, and optimization to create an effective system for identifying and categorizing domestic rubbish. The suggested technique takes a collection of household rubbish photos as input, and the system outputs the assigned category for each kind of domestic waste. The proposed method performs this task in two steps:

1. Feature extraction based on CNN and the CapSA.
2. Classification of the extracted features using ECOC and ANN.

Fig. 2 illustrates the steps of the proposed method. In this figure, the steps related to the training and testing phases are separated. Accordingly, the training phase steps are depicted in the right column, and the testing phase steps are depicted in the left column. According to Fig. 2, the proposed algorithm starts with optimizing the hyperparameters of the CNN through processing the training samples. The optimal configuration of the CNN model refers to determining appropriate values for the convolutional layers and pooling functions. To achieve this, the CapSA is utilized. This optimization algorithm aims to adjust the values of the mentioned parameters in the CNN model to obtain a feature extraction model with minimal training error. The extracted features, through the configured CNN model, serve as input to the classification model. The proposed classification model consists of sets of ANNs that collaborate as an ECOC model. In this classification model, each ANN model is trained as a binary classifier based on the ECOC encoding matrix. Then, by combining the outputs of these models for each training sample and comparing the generated code string with the code that is defined for each target category, the classification result is determined. The multi-class classification issue is broken down into sets of binary (two-class) problems with reduced complexity by the suggested ECOC model. This approach may improve detection accuracy, particularly when used to issues with a high number of target classes.

3.2.1. Feature extraction based on CNN and CapSA

The presented approach begins with the extraction of input image features through a CNN model optimized by CapSA. In this regard, the proposed feature extraction model consists of a basic CNN model, where its hyperparameters are configured using CapSA. The goal of employing CapSA in the presented approach is to obtain a feature extraction model that is tailored to the characteristics of the training samples, allowing for the subsequent detection process to be performed with the minimum possible error. The structure of the basic CNN model used for image feature extraction in the proposed method is illustrated in Fig. 3. The proposed feature extraction model is fed with RGB images from the dataset. For the TrashNet dataset, all samples have the same dimensions, and thus, the input dimensions for each sample are determined based on their actual size ($512 \times 384 \times 3$). On the other hand, for the HGCD dataset, all samples are resized to dimensions of ($200 \times 200 \times 3$). Three convolutional components, each consisting of a convolutional layer, a ReLU layer, and a pooling layer, make up the basic CNN model that is being suggested. The suggested technique uses CapSA to alter the pooling function type and the hyperparameters of each convolutional layer, such as number of filters, width, and length. In addition, three completely linked layers are used to extract picture features after the third pooling layer extracts the feature maps. CapSA is also used in the proposed technique to calculate the number of neurons in these three completely linked layers. With these explanations, the configurable parameters in the proposed basic CNN model are as follows:

- The size of each fully connected layer.
- The function type in pooling layers.
- The number of filters, width, and length in each convolutional layer.

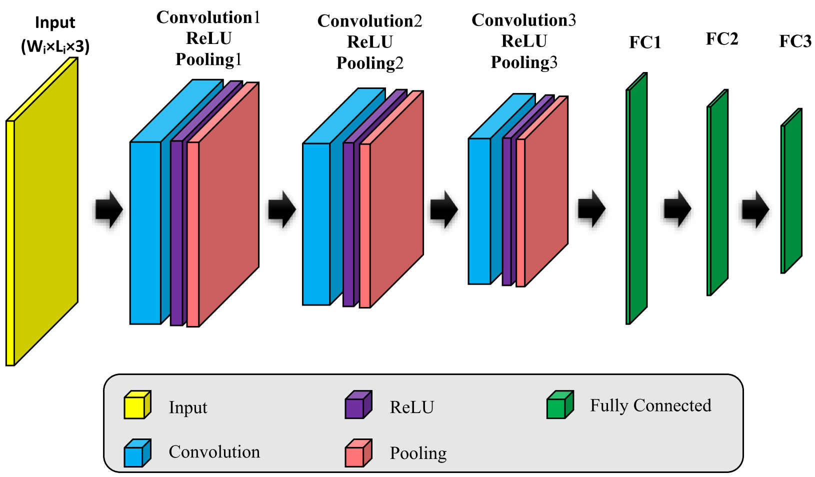


Fig. 3. The basic CNN model's structure, which is utilized in the suggested technique for extracting features from images.

In the proposed method, the optimal configuration of these parameters by CapSA is described. For this purpose, the encoding scheme of the solution vector in the optimization algorithm and the evaluation process for fitting each solution vector are explained. Afterward, the optimization stages of the base CNN model are presented.

Appropriate settings for the convolutional, pooling, and fully connected layers must be found to maximize the configuration of the suggested feature extraction model utilizing CapSA. Every one of the above-listed modifiable parameters is regarded as an optimization variable. It is computationally impractical to investigate all conceivable configurations for the CNN model with current machines. As a result, the suggested technique has a restricted search area for each optimization variable. Accordingly, the search space for each of the mentioned variables in CapSA is defined as follows:

- *The parameters of the convolutional layers*: Each convolutional layer (Convolution1 to Convolution3) is configured through three parameters: length, width, and the number of filters. These three parameters are represented as natural numbers in the solution vector. In this set, the length and width parameters for each filter have the same dimensions and can take values in the range of [3, 9]. Additionally, the number of filters can be a multiple of 8 in the range of [8, 128].
- *The parameters of the pooling function*: Each of the pooling layers (Pooling1 to Pooling3) can choose either the Max pooling or Average Pooling function. In the proposed optimization algorithm, configuring each pooling layer with the Max function is represented by the number 0, and configuring each pooling layer with the Average function is represented by the number 1.
- *The parameters of the fully connected layers*: The fully connected layers (FC1 to FC3) are configured based on the number of neurons. Accordingly, the optimization variables corresponding to these parameters are described as a natural number in the range of [30, 1000].

According to the presented description, each solution vector in the CapSA optimization method for optimizing the configuration of the proposed CNN model will be 12 in length. The first three optimization variables in each solution vector, which may have values ranging from 3 to 9, reflect the length and breadth of the filters in the Convolution1 to Convolution3 layers. The next 3 optimization variables describe the number of filters in these three layers and can take values in the range of [8, 128]. Furthermore, the next 3 optimization variables describe the pooling function type in Pooling1 to Pooling3 layers as binary variables. Lastly, the last 3 optimization variables represent the size of layers FC1 to FC3 as natural numbers in the range of [30, 1000].

Using the supplied structure for each solution vector in the CapSA optimization process, the training error criteria may be used to assess the fitness of various configurations. To do this, the fundamental CNN model is configured using the solution vector x (Fig. 3). Next, a small training set consisting of 25 % of the database samples is used to train this model. Subsequently, a CNN model is set up with a validation set of 25 % of the database samples that do not overlap with the training samples. A SoftMax layer is used to classify the features that the CNN model extracts. Lastly, using the training error as a basis, the fitness of the solution vector x is computed using equation (1):

$$fitness(x) = \frac{F_{CNN(x)}}{N} \tag{1}$$

In the above equation, N represents the total number of validation samples, and $F_{CNN(x)}$ represents the number of classification errors of the configured CNN model through the solution vector x. Based on the provided explanations regarding the structure of the solution vector and its fitness evaluation, the optimization steps for configuring the CNN model using CapSA are as follows [24]:

Step 1: The initial population is randomly determined based on the specified bounds for each optimization variable.
Step 2: The fitness of each solution vector (Capuchin) is calculated using equation (1).
Step 3: The initial velocity of each Capuchin is set.
Step 4: Half of the Capuchin population is randomly selected as leaders, and the rest are determined as followers.
Step 5: If the minimum fitness reaches zero or the number of algorithm iterations reaches K, go to step 13; otherwise, repeat the following steps:
Step 6: The lifetime parameter of CapSA is calculated using the following equation [24]:

$$\tau = \beta_0 \widehat{e}\left(-\frac{\beta_1 k}{K}\right)^{\beta_2} \tag{2}$$

In equation (2), k defines the current iteration count, and parameters β_0, β_1, and β_2 have values of 2, 21, and 2, respectively.

Step 7: Repeat the following steps for each Capuchin agent (leader and followers) with index i:
Step 8: If i is a leader Capuchin, update its velocity according to the following equation [24]:

$$v_j^i = \rho v_j^i + \tau a_1\left(x_{best_j}^i - x_j^i\right) r_1 + \tau a_2\left(F - x_j^i\right) r_2 \tag{3}$$

In equation (3), the index j defines the problem dimensions, and v_j^i represents the velocity of Capuchin agent i in dimension j. x_j^i denotes the position of Capuchin agent i for variable j, and $x_{best_j}^i$ represents the best position of Capuchin agent i for variable j obtained so far. Additionally, r_1 and r_2 are two random numbers in the range [0, 1]. Lastly, c is the parameter controlling the influence of the previous velocity, and its value is set to 0.7.

Step 9: Update the new positions of leader Capuchin agents based on their velocities and movement patterns.
Step 10: Update the new positions of follower Capuchin agents based on their velocities and the position of the leader.
Step 11: Calculate the fitness of the population members based on equation (1).
Step 12: If the positions of the entire population are updated, go to Step 5; else, go to Step 7.
Step 13: Select the solution with the lowest fitness as the CNN's best configuration.

After performing the above steps, the optimal configuration determined by the solution will be applied to the base CNN model, and the resulting model will be used for feature extraction on new images.

3.2.2. Classification based on ECOC and ANN Features

The ECOC model classifies the characteristics extracted using CNN in the second stage of the introduced approach to determine the type of garbage. By breaking down multi-class classification issues into groups of binary-class problems, ECOC is a classification technique that combines binary classifiers to decrease complexity and improve accuracy. This approach breaks down the classes using a binary encoding matrix. The target categories are represented by the rows of the ECOC encoding matrix, and the amount of its columns is the same as the amount of binary classification algorithms that were employed. The binary classification algorithms are trained using the specified coding in each column of the encoding matrix. Every target category is uniquely identified through a binary encoding string [25]. A number of techniques have been put up to build the coding matrix. The coding matrix in the suggested method is created using the ordinal approach. In this technique, the amount of classification models equals $L = C - 1$, where C is the categories count. In this technique, first bit of the code is set to negative (0) and the successive classes are set to positive (1) for the classifier one. The second layer of classifier are encoded negative for the first two classes whereas the other classes are encoded positive by the same layer. This coding process is repeated for the remaining classifiers [26].

ANNs are the learning models that the suggested ECOC model makes use of. In this architecture, there are 10 neurons in each hidden layer of a neural network. The buried layer's activation function is sigmoid tangent in nature. The number of features that the CNN model was able to extract in the previous stage equals the size of the input layer. Two more neurons in the output layer define the ANN model's output code bit in the ECOC coding matrix. Fig. 4 depicts this network's structure. Levenberg-Marquardt back-propagation is the method used to train the neural network [27]. This approach uses the Jacobian matrix to facilitate learning and minimizes the output error to modify the weights of the network.

As a result, the MC × L matrix is built by means of the ordinal encoding technique, where MC stands for the number of classes and L for the number of attributes. The training code matrix is made up of the rows, where each of them corresponds to a certain target class, and for each ANN model, the training is based on the codes present in one of the columns of the training code matrix. The training phase of the employed ANN models is followed by the testing step, in which each trained model processes the features of a new sample, and based on its output, binary information is provided, and this is then used to identify household garbage. The classifiers' code fragments are then combined to generate a code string. The rows of the code matrix are compared with this code text. In the end, the sample will be a member of the class whose binary output from the ANN models is closest to the matching code string. The suggested ECOC model computes the difference between two code strings using the Hamming distance metric.

4. Implementation and results

This section discusses the outcomes that were achieved and provides information on how the suggested strategy was implemented. The introduced model was examined with the aid of MATLAB 2019a, and instances from the TrashNet and HGCD databases were used to gauge how successful it was. A personal computer equipped with an Intel Core i7 3.2 GHz CPU and 16 GB RAM was used for all of the trials. The NVIDIA RTX 2080 Ti graphics processors were used to handle the processing requirements associated with the suggested CNN model. A 10-fold cross-validation approach was used to conduct the trials, in which the database samples were split into 10 non-

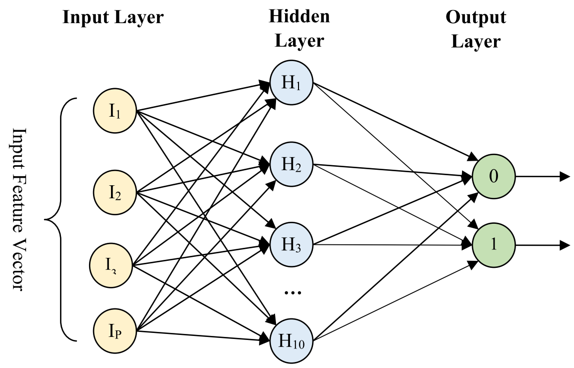


Fig. 4. Structure of binary ANN classifiers in the proposed ECOC model.

overlapping sets and the training and testing stages were performed ten times. Nine sets were utilized to train the CNN and ECOC models in each iteration, with the remaining set being used for model testing and household rubbish identification. The introduced model's performance was evaluated throughout the studies using metrics for F-measure, recall, precision, and accuracy. The ratio of correctly identified test instances to the total number of test instances is represented by the accuracy measure. The accuracy measure evaluates how well the classification algorithm performs in accurately classifying each kind of waste. Conversely, the recall measure shows the proportion of each target class's successfully recognized samples. In addition, the harmonic mean of recall and accuracy is described by the F-measure metric. Precision and recall are measured using equations (4) and (5), respectively. Also, F-measure is calculated using equation (6):

$$Precision = \frac{TP}{TP + FP} \tag{4}$$

$$Recall = \frac{TP}{TP + FN} \tag{5}$$

$$F - Measure = 2 \times \frac{Precision \times Recall}{Precision + Recall} \tag{6}$$

The number of true positive samples that are accurately identified is denoted by TP in the equations above. The number of negative category instances that have been mistaken as positive is shown by the FP. The number of positive category instances that are incorrectly labeled as negative is represented by FN. These measures are determined independently for every class because the research subject involves multiple classes. In this case, the present category has been marked as positive, whereas the other categories have been marked as negative.

The optimization process of the proposed CNN model was performed using CapSA. For this purpose, the optimization process was conducted using samples from the TrashNet and HGCD databases. The population size and the number of iterations in the CapSA algorithm were set to 100 and 400, respectively, throughout the implementation phase. Table 2 displays the designated settings for every layer of the CNN model utilizing the HGCD and TrashNet datasets. Therefore, feature extraction in the proposed method was performed based on CNN models with the provided configurations in Table 2.

The results presented in Table 2 indicate that the CNN model configuration based on the TrashNet dataset leads to the formation of networks with larger dimensions in the convolutional layers. This can be attributed to the larger dimensions of the images in the TrashNet database compared to HGCD. On the other hand, the configuration based on the HGCD database resulted in a higher number of filters in the convolutional layers (more than twice that of TrashNet). This is because the higher number of HGCD samples and the higher complexity of the patterns in it require a larger number of convolutional filters to describe these patterns accurately. This has also resulted in the extracted features for the HGCD database (the size of FC3 layer) being approximately twice as many as those of the TrashNet database.

To assess the effectiveness of each employed technique in the feature extraction and classification steps, the performance of the proposed method has been compared under the following two conditions:

- Proposed (without CapSA): In this condition, the optimization step of the CNN model through the CapSA algorithm is ignored, and a static CNN model is used for feature extraction from the images.
- Proposed (without ECOC): In this condition, the proposed ECOC classification model is ignored, and the extracted features from the optimized CNN model are directly classified using a SoftMax layer.

Also, the performance of the introduced model is evaluated against two models: GCNet in Ref. [16] and MB-Xception in Ref. [13]. Fig. 5 shows a comparison between the mean accuracy of the suggested technique and alternative methods for the TrashNet and HGCD datasets in terms of categorizing various categories of domestic rubbish. Generally, the classification of household garbage in the HGCD dataset achieves higher accuracy compared to the TrashNet dataset. This is because in the HGCD dataset, despite having more classes, the number of training samples is more than five times that of the TrashNet dataset. This characteristic fulfills one of the essential requirements of deep learning models, which is the need for a large number of training samples. Based on this graph, the

Table 2
Configurations defined for each layer of the CNN model using the TrashNet and HGCD databases.

Layer	TrashNet Parameter Setting	HGCD Parameter Setting
Input	513 × 384 × 3	200 × 200 × 3
Convolution1 (W × H@N,Stride)	8 × 8@19, S = 1	7 × 7@43, S = 1
Pooling1	AveragePooling	MaxPooling (2 × 2,S = 1)
Convolution2 (W × H@N,Stride)	5 × 5@17, S = 1	5 × 5@24, S = 1
Pooling2	AveragePooling	MaxPooling (2 × 2,S = 1)
Convolution3 (W × H@N,Stride)	4 × 4@10, S = 1	3 × 3@8, S = 1
Pooling3	AveragePooling	MaxPooling (2 × 2,S = 1)
FC1	934	612
FC2	408	298
FC3	116	195

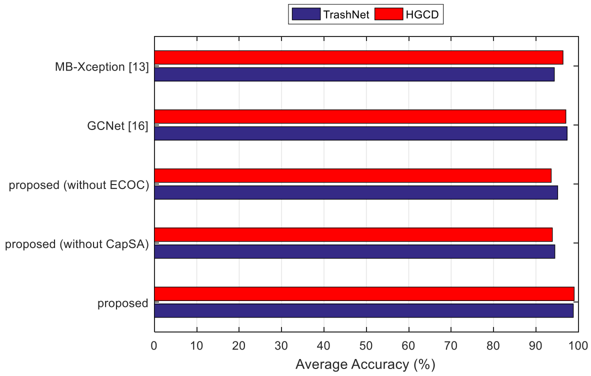


Fig. 5. The suggested method's average accuracy in comparison to alternative approaches in classifying different types of household garbage for the TrashNet and HGCD datasets.

proposed method can classify samples from the TrashNet dataset with an average accuracy of 98.81 %. Furthermore, the suggested method's accuracy for the HGCD dataset is 99.02 %. The suggested solution performs more accurately than the alternatives in both scenarios. A 4.35 % improvement in accuracy may be achieved by refining the feature extraction model in the TrashNet dataset using the CapSA method, according to an analysis of the accuracy measure. For the HGCD dataset, this accuracy gain is 5.14 %. Thus, one of the factors contributing to the increased accuracy of the introduced model can be linked to the application of the CapSA optimization technique for the CNN model. Also, if the ECOC classifier is ignored and the classification of features is done using the SoftMax layer, the classification accuracy for the TrashNet and HGCD datasets will be 95.13 % and 93.60 %, respectively. This means that the ECOC-based classification strategy in the proposed method results in a minimum increase of 3.68 % in accuracy. The decrease in accuracy when removing the ECOC model is higher for the HGCD dataset compared to the TrashNet dataset. This characteristic indicates that the ECOC classifier performs better in problems with a higher number of classes.

The GCNet [16] is the strategy that achieves the closest results to the presented approach, as shown in Fig. 5. This model, combines the EfficientNet, Vision Transformer, and DenseNet models. The confusion matrix is able to provide more thorough insights into how various approaches function when it comes to domestic waste classification. The proposed method's and the GCNet method's confusion matrices for categorizing samples from the TrashNet dataset are shown in Fig. 6.

The test instance true labels are represented by each column in the confusion matrices shown in Fig. 6, while the label predictions made by each classification technique are represented by the rows. For instance, the suggested method accurately classifies 498 samples from the 501 samples that belong to the Glass class (sum of entries in the matrix's first column) in Fig. 6a. Three instances, though, are incorrectly categorized as belonging to the cardboard, metal, and paper categories. Conversely, the suggested approach places 501 instances in the Glass category (which is the sum of entries in the matrix's first row), out of which three of them are in the categories of paper, cardboard, and plastic. The findings for the remaining categories in this matrix are interpreted in a similar way.

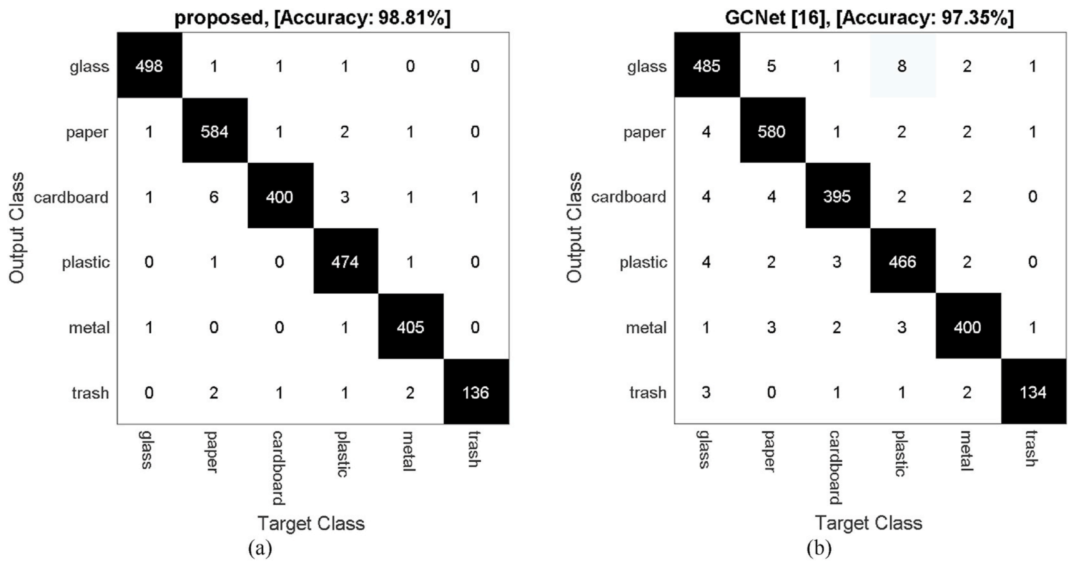


Fig. 6. Confusion matrices (a) for the suggested approach and (b) for the GCNet method [16] while categorizing TrashNet database samples.

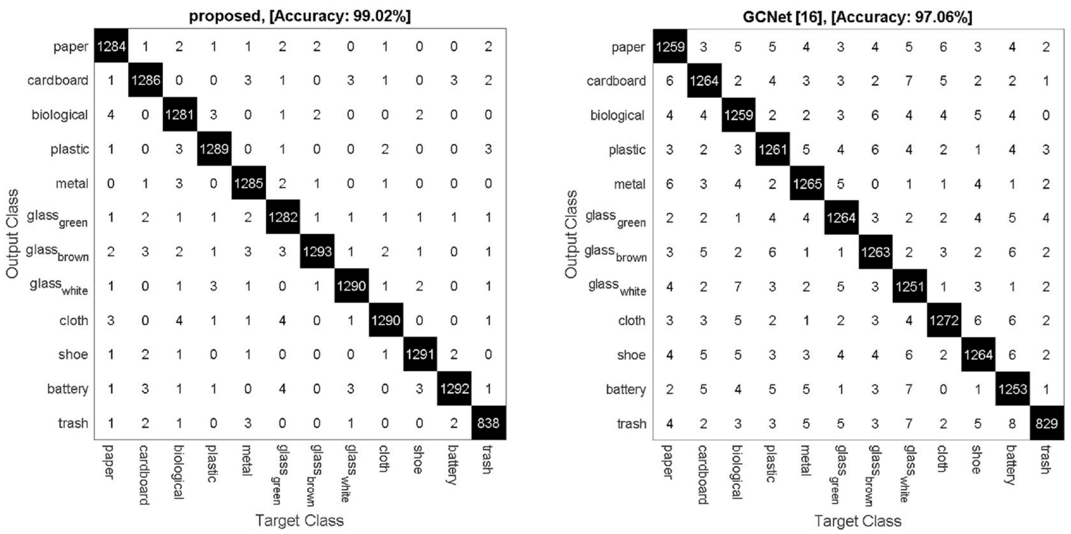
The suggested approach outperforms the GCNet method in identifying all categories, as demonstrated by the comparison of Fig. 6a and b, which results in a 1.46 % gain in accuracy. The matrices of confusion for the suggested approach and the GCNet approach for categorizing samples from the HGCD dataset are also displayed in Fig. 7a and b, respectively.

Analyzing and contrasting the confusion matrices shown in Fig. 7 will lead to comparable findings. According to this figure, the suggested approach may continue to operate as intended even as the number of target categories in the HGCD database increases. When the number of target categories in the HGCD database increases, most methods—including GCNet—seem to lose accuracy, as can be shown by comparing Figs. 5–7. Nonetheless, compared to the TrashNet database, the suggested method's classification accuracy for HGCD samples is greater. There are two reasons for this performance in the suggested method: First, the CNN model may be configured by the CapSA in the suggested approach based on the input data structure. As a result, neither the problem's complexity nor the number of target classes will have an impact on the quality of the features that are retrieved. Second, the number of binary classifiers in the suggested technique can be determined by using the ECOC model in a manner that gets around the multi-class problem's complexity. When compared to other procedures, these variables help explain why the suggested method is more accurate. In Fig. 8, the F-measure, recall, and precision, metrics for the TrashNet and HGCD databases are used to examine how well different approaches perform in categorizing different forms of household rubbish.

In the plotted charts in Fig. 8, the results for the TrashNet database are shown in the left column, and the results for the HGCD database are shown in the right column. Fig. 8a and b shows the precision of different methods obtained from categorizing samples of TrashNet and HGCD, respectively. Fig. 8c and d shows the recall values for these datasets. Also, the F-measure values for these datasets have been presented in Fig. 8e and f. Additionally, each of the plotted charts in the first to third rows represents the values of precision, recall, and F-measure for different classes, respectively. In each chart, the first dimension represents the categories related to different types of household garbage, and the second dimension corresponds to the compared methods. It is evident from looking at the plotted charts in Fig. 8 that the suggested strategy performs better than alternative approaches at classifying various classes. For the TrashNet and HGCD databases, the suggested approach performs better than the approaches that were examined in terms of accuracy, recall, and F-measure. The average accuracy, recall, and F-measure values for TrashNet and HGCD datasets, are computed in Fig. 9a and b, respectively. These graphs show how various approaches fare overall in terms of qualitative categorization measures. Furthermore, Table 3 displays the numerical outcomes of the experiments carried out in this section.

The suggested technique can classify different categories of household rubbish with higher efficiency than the examined methods, as demonstrated in Table 3 and Fig. 9. The introduced model's greater accuracy values when compared to other approaches demonstrate how much more precisely each sort of rubbish can be classified using this technique. Furthermore, the suggested method's superiority in terms of the recall measure shows that it can accurately identify a larger percentage of household waste categories. Fig. 10a and b shows the precision-recall curves resulting from the classification of household garbage for the TrashNet and HGCD databases, respectively. These curves represent the aggregation of precision-recall curves for different classes. In these curves, accuracy values are provided for varying recall thresholds. According to this figure, the presented model can obtain higher accuracy and recall values compared to other methods. The greater area under the curve in the introduced method indicates a higher quality of the classification process. Based on this, it can be concluded that the introduced method is an effective strategy for classifying household garbage under different conditions.

Fig. 11a and b, compare the Receiver Operating Characteristic (ROC) curves of different methods for TrashNet and HGCD datasets, respectively. To provide a clearer comparison, the ROC plots have been enlarged in the upper left corner. These outcomes demonstrate how well the proposed approach performs to lower the FP rate and raise the TP rate. This indicates that for each kind of household



(a) (b)

Fig. 7. Confusion matrices (a) for the suggested approach and (b) for the GCNet method [16] while categorizing HGCD database samples.

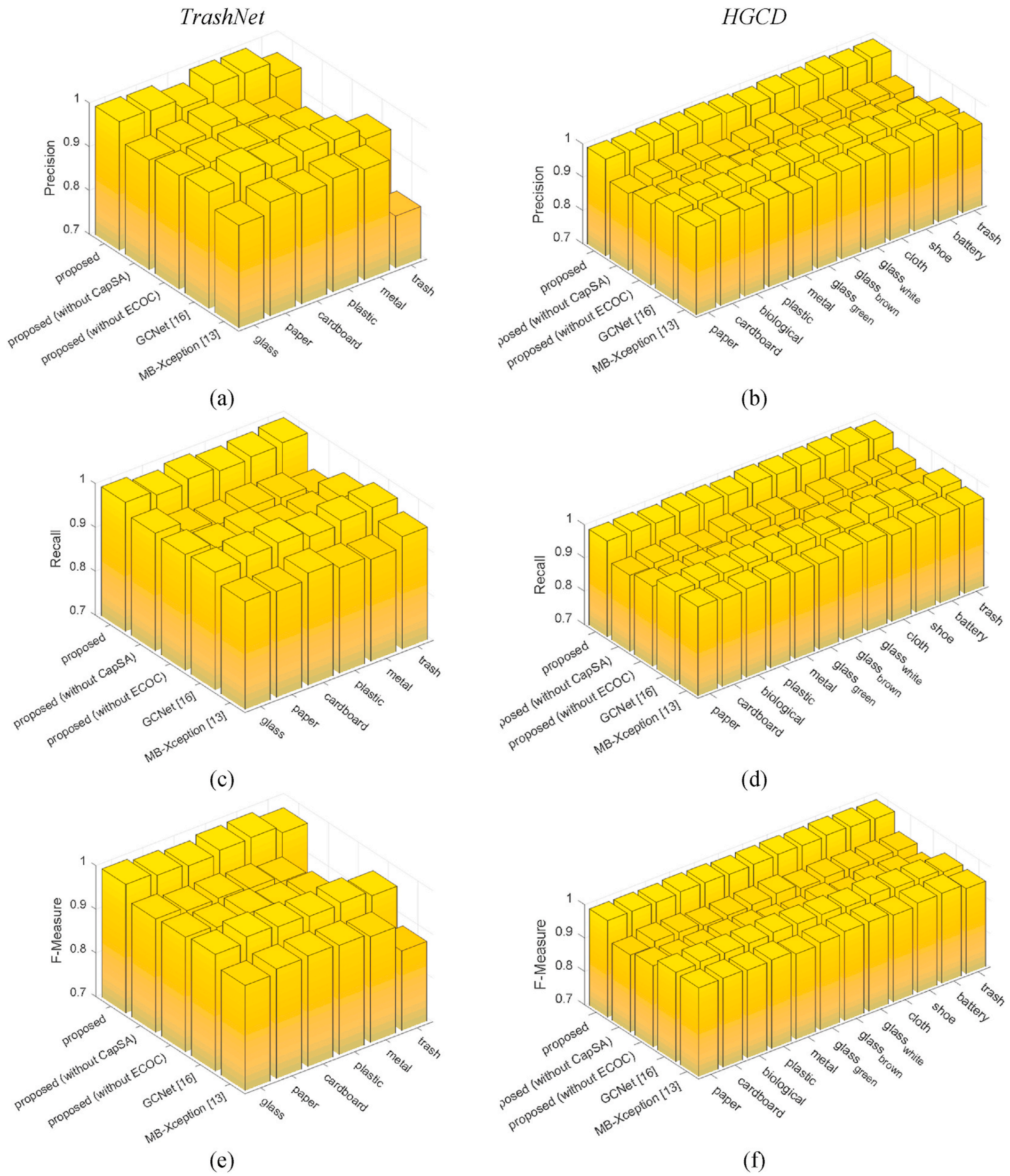


Fig. 8. Comparing the performance of various methods in classifying different types of household garbage based on (a) precision in TrashNet, (b) precision in HGCD, (c) recall in TrashNet, (d) recall in HGCD, (e) F-measure in HGCD, and (f) F-measure in TrashNet.

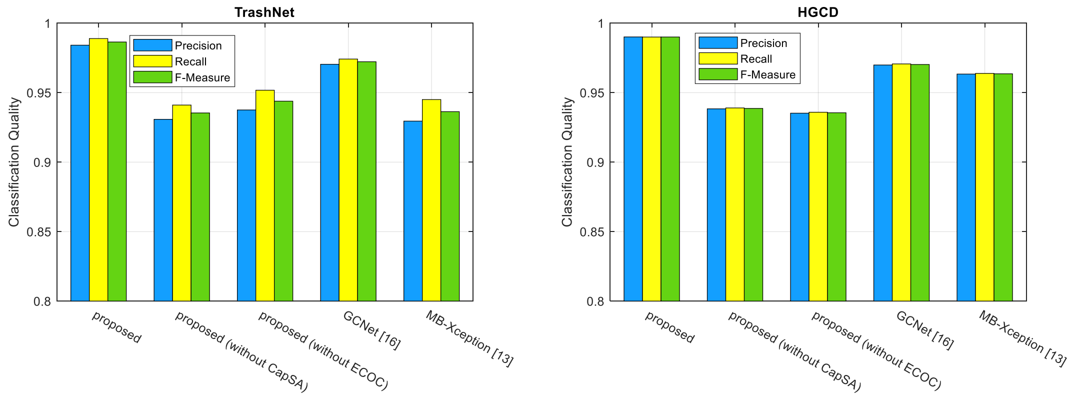
waste, the model that is being provided may offer a more precise forecast.

Table 4 compares the garbage sorting algorithms based on the Area under the precision-recall curve (AUCPR) and the Area Under the ROC Curve (AUC).

The results demonstrate that the introduced model may raise the AUC for the HGCD and TrashNet datasets by at least 0.9 % and 0.8, respectively. Additionally, for the TrashNet and HGCD datasets, the suggested technique might enhance AUCPR by at least 2.19 and 2.66 %, respectively. These statistical findings may demonstrate the suggested method's superiority over other techniques under certain circumstances.

5. Conclusion

One important use of computer vision that may lower environmental risks and greatly increase trash recycling rates is the automatic sorting of domestic waste. This research offered a unique approach to automated home waste sorting. The suggested approach

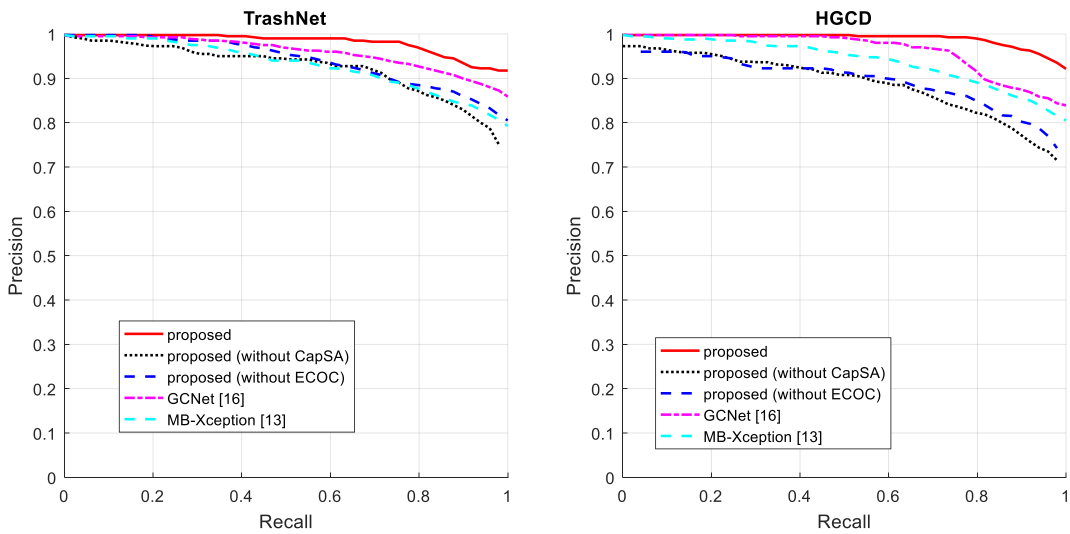


(a) (b)

Fig. 9. The mean values of F-measure, recall, and precision for (a) TrashNet database and (b) HGCD database.

Table 3
Evaluation of the suggested method's performance in comparison to other approaches.

Database	Method	Accuracy	F-measure	Recall	precision
TrashNet	**Proposed**	98.8128 %	0.9865	0.9889	0.9842
	Proposed (Without CapSA)	94.4598 %	0.9354	0.9411	0.9308
	Proposed (Without ECOC)	95.1326 %	0.9439	0.9518	0.9376
	GCNet [16]	97.3486 %	0.9723	0.9742	0.9704
	MB-Xception [13]	94.3411 %	0.9363	0.9450	0.9295
HGCD	**Proposed**	99.0165 %	0.9901	0.9900	0.9901
	Proposed (Without CapSA)	93.8746 %	0.9387	0.9390	0.9384
	Proposed (Without ECOC)	93.6040 %	0.9355	0.9359	0.9352
	GCNet [16]	97.0561 %	0.9703	0.9707	0.9699
	MB-Xception [13]	96.3828 %	0.9636	0.9639	0.9634



(a) (b)

Fig. 10. Precision-Recall curves obtained by classifying household garbage for (a) TrashNet and (b) HGCD databases.

blends optimization, ensemble learning, and deep learning approaches. Two sub-problems were identified in the research issue breakdown: "feature extraction" and "classification." To address the first sub-problem, CNN and CapSA were merged, and to solve the second sub-problem, ECOC and ANN were combined. As a result, this research developed a CNN-based feature extraction model that optimizes the CNN model's configuration using the Capuchin search method. A performance comparison of the suggested feature extraction model with static CNN models revealed that the proposed strategy improves classification accuracy by at least 4.34 %.

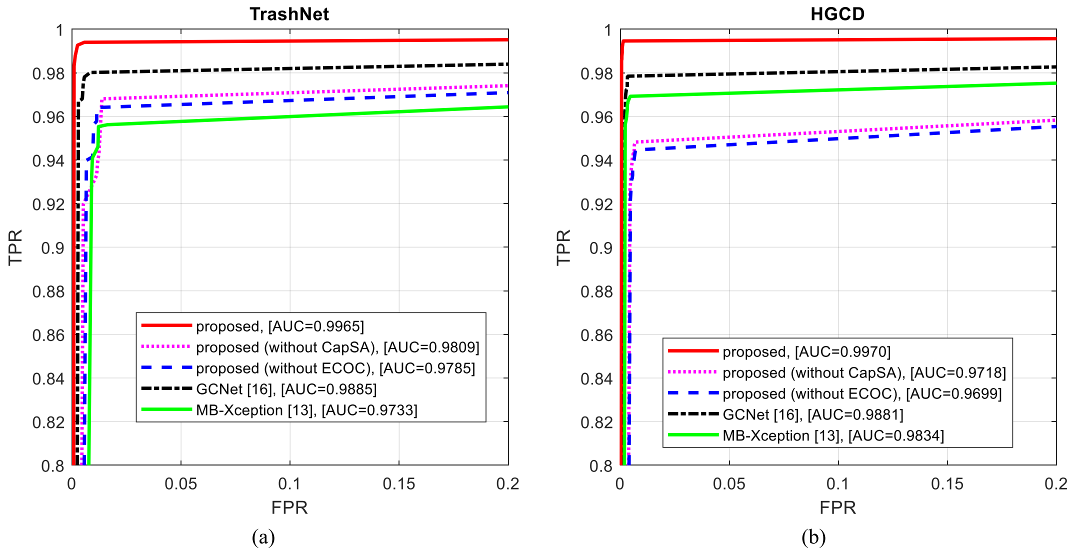


Fig. 11. ROC curves obtained by classifying household garbage for (a) TrashNet and (b) HGCD databases.

Table 4
Comparison of the garbage sorting methods in terms of AUC and AUCPR.

Database	Method	AUC	AUCPR
TrashNet	**Proposed**	0.9965	0.9810
	Proposed (Without CapSA)	0.9809	0.9211
	Proposed (Without ECOC)	0.9785	0.9406
	GCNet [16]	0.9885	0.9591
	MB-Xception [13]	0.9733	0.9303
HGCD	**Proposed**	0.9970	0.9896
	Proposed (Without CapSA)	0.9718	0.9288
	Proposed (Without ECOC)	0.9699	0.9317
	GCNet [16]	0.9881	0.9630
	MB-Xception [13]	0.9837	0.9392

Furthermore, the suggested technique used an ECOC model based on binary ANN classifiers to identify household waste kinds, which is ideal for multi-class issues, particularly when the number of target classes is large. An analysis of the suggested classification model's efficacy in comparison to other models, such as CNN, revealed that the ECOC model can enhance classification accuracy by at least 3.68 %. These findings attest to the fact that every strategy used in the suggested approach improves the precision of the categorization of domestic waste. The TrashNet and HGCD datasets were used to assess how well the suggested strategy performed. The suggested approach demonstrated an improvement of at least 1.46 % over earlier approaches, as seen by the obtained findings, which showed classification accuracies of 98.81 % and 99.01 % for the samples in these two databases, respectively. This is a major increase in the trash categorization setting, where even slight improvements in accuracy may have a big impact on waste management and resource recovery. The lengthy calculation time needed to optimize the CNN model's setup is one of the suggested method's drawbacks. Future studies may investigate the use of parallel processing models to shorten this processing time, even though this step is done offline and has no effect on how quickly test samples are processed. Another aspect worth investigating in future research is the classification model used. By weighting the ANN classifiers in the ECOC model based on their training errors, the effectiveness of each learning model can be determined, and efforts can be made to increase the accuracy of its detection accordingly. Moreover, the potential for further enhanced recognition accuracy lies in the development of more advanced deep learning models and innovative data augmentation techniques, which warrant further research exploration.

Data availability

The datasets used for this study are publicly available. The TrashNet dataset can be accessed at https://github.com/garythung/trashnet. The HGCD dataset can be accessed at https://www.kaggle.com/datasets/mostafaabla/garbage-classification.

CRediT authorship contribution statement

Jianfei Wang: Investigation.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

[1] V. Jayaraman, A.R. Lakshminarayanan, S. Parthasarathy, A. Suganthy, Forecasting the municipal solid waste using GSO-XGBoost model, Intelligent Automation & Soft Computing 37 (1) (2023).
[2] Y. Tong, J. Liu, S. Liu, China is implementing "Garbage Classification" action, Environ. Pollut. 259 (2020) 113707.
[3] T. Erhardt, Garbage in and garbage out? On waste havens in Switzerland, Environ. Resour. Econ. 73 (1) (2019) 251–282.
[4] R. Bai, B. Lin, Are residents willing to pay for garbage recycling: evidence from a survey in Chinese first-tier cities, Environ. Impact Assess. Rev. 95 (2022) 106789.
[5] M.G. Flores, J. Tan, Literature review of automated waste segregation system using machine learning: a comprehensive analysis, Int. J. Simulat. Syst. Sci. Technol. (2019).
[6] S. Shahab, M. Anjum, M.S. Umar, Deep learning applications in solid waste management: a deep literature review, Int. J. Adv. Comput. Sci. Appl. 13 (3) (2022).
[7] A. Namoun, A. Tufail, M.Y. Khan, A. Alrehaili, T.A. Syed, O. BenRhouma, Solid waste generation and disposal using machine learning approaches: a survey of solutions and challenges, Sustainability 14 (20) (2022) 13578.
[8] Z. Kang, J. Yang, G. Li, Z. Zhang, An automatic garbage classification system based on deep learning, IEEE Access 8 (2020) 140019–140029.
[9] B. Fu, S. Li, J. Wei, Q. Li, Q. Wang, J. Tu, A novel intelligent garbage classification system based on deep learning and an embedded linux system, IEEE Access 9 (2021) 131134–131146.
[10] J. Yang, Z. Zeng, K. Wang, H. Zou, L. Xie, GarbageNet: a unified learning framework for robust garbage classification, IEEE Transactions on Artificial Intelligence 2 (4) (2021) 372–380.
[11] M. Zeng, X. Lu, W. Xu, T. Zhou, Y. Liu, PublicGarbageNet: a deep learning framework for public garbage classification, in: 2020 39th Chinese Control Conference (CCC), IEEE, 2020, July, pp. 7200–7205.
[12] S.N. Endah, I.N. Shiddiq, Xception architecture transfer learning for garbage classification, in: 2020 4th International Conference on Informatics and Computational Sciences (ICICoS), IEEE, 2020, November, pp. 1–4.
[13] C. Shi, R. Xia, L. Wang, A novel multi-branch channel expansion network for garbage image classification, IEEE Access 8 (2020) 154436–154452.
[14] F.S. Alsubaei, F.N. Al-Wesabi, A.M. Hilal, Deep learning-based small object detection and classification model for garbage waste management in smart cities and iot environment, Appl. Sci. 12 (5) (2022) 2281.
[15] T. Gupta, R. Joshi, D. Mukhopadhyay, K. Sachdeva, N. Jain, D. Virmani, L. Garcia-Hernandez, A deep learning approach based hardware solution to categorise garbage in environment, Complex & Intelligent Systems (2022) 1–24.
[16] W. Liu, H. Ouyang, Q. Liu, S. Cai, C. Wang, J. Xie, W. Hu, Image recognition for garbage classification based on transfer learning and model fusion, Math. Probl Eng. 2022 (2022).
[17] H.I.K. Fathurrahman, A. Ma'arif, L.Y. Chin, The development of real-time mobile garbage detection using deep learning, Jurnal Ilmiah Teknik Elektro Komputer dan Informatika (JITEKI) 7 (3) (2022) 472.
[18] A. Oğuz, Ö.F. Ertuğrul, Determining the fullness of garbage containers by deep learning, Expert Syst. Appl. 119544 (2023).
[19] Y. Li, W. Liu, Deep learning-based garbage image recognition algorithm, Appl. Nanosci. 13 (2) (2023) 1415–1424.
[20] N. Hayashi, S. Koyanaka, T. Oki, Constructing an automatic object-recognition algorithm using labeling information for efficient recycling of WEEE, Waste Manag. 88 (2019) 337–346.
[21] X. Zhang, H. Zhou, C. Fu, M. Mi, C. Zhan, D.T. Pham, A.M. Fathollahi-Fard, Application and planning of an energy-oriented stochastic disassembly line balancing problem, Environ. Sci. Pollut. Control Ser. (2023) 1–15.
[22] M. Yang, G. Thung, Classification of trash for recyclability status, CS229 project report 2016 (1) (2016) 3. Available online at: https://github.com/garythung/trashnet. Accessed on: April, 2023.
[23] HGCD, Household Garbage Classification Dataset, 2022. https://www.kaggle.com/datasets/mostafaabla/garbage-classification. Accessed on: April, 2023.
[24] M. Braik, A. Sheta, H. Al-Hiary, A novel meta-heuristic search algorithm for solving optimization problems: capuchin search algorithm, Neural Comput. Appl. 33 (2021) 2515–2547.
[25] T.G. Dietterich, G. Bakiri, Error-correcting output codes: a general method for improving multiclass inductive learning programs, in: The Mathematics of Generalization, CRC Press, 2018, pp. 395–407.
[26] H. Joutsijoki, M. Haponen, J. Rasku, K. Aalto-Setälä, M. Juhola, Error-correcting output codes in classification of human induced pluripotent stem cell colony images, BioMed Res. Int. 2016 (2016).
[27] H. Yu, B.M. Wilamowski, Levenberg–marquardt training, in: Intelligent Systems, CRC Press, 2018, p. 12, 1.