ARCNet: A Novel Deep Learning Model for Robust Solid Waste Image Classification in Waste Management Systems

Velmurugan Lingamuthu
School of Computing Science and Engineering, VIT Bhopal University India
vel.erode@gmail.com

Devaraju Sellappan
School of Computing Science Engineering and Artificial Intelligence VIT Bhopal University, India
devamcet@gmail.com

Sakthivel Balu
Department of Electronics and Communication Engineering, PSYEC under Anna University, India
786sakthivel@gmail.com

Sheik Mohammed Abdul Raghabu
Department of Electronics and Communication Engineering, PSYEC under Anna University, India
786sheik.mohd@gmail.com

Aanjankumar Sureshkumar
School of Computing Science and Engineering, VIT Bhopal University India
itsec1990@gmail.com

Prabu Kanna Govindarajan
School of Computing Science Engineering and Artificial Intelligence, VIT Bhopal University, India
gpkanna@gmail.com

Abstract: *Accurate classification of solid waste types is essential for efficient waste management. It also supports resource recovery and environmental sustainability. Image processing-based waste classification techniques have received greater attention due to their accuracy and reliability. Existing techniques failed to capture the complex features of waste images. This work proposes Attention-Enhanced Residual Capsule Network (ARCNet), an innovative deep-learning architecture for solid waste image classification in waste management systems. ARCNet integrates Residual Attention Blocks (RABs) and Capsule Network (CapsNet) layers to improve feature extraction and capture the spatial hierarchies within waste images. In addition, a Global Context Attention Module (GCAM) is incorporated for the refined analysis of waste images. The model's parameters are optimized using the Black Eagle Optimizer (BEO). BEO is inspired by the hunting and flight behaviors of the black eagles. Experimental results show that ARCNet outperforms traditional models in terms of accuracy, precision and recall rates. This approach supports automation in waste management and offers a reliable solution for improving sorting efficiency in recycling and disposal processes.*

Keywords: *Solid waste management, classification, ARCNet, global context attention module.*

Received April 11, 2025; accepted July 18, 2025
https://doi.org/10.34028/iajit/22/6/6

1. Introduction

Solid waste management is a crucial aspect of urbanization, environmental sustainability, and resource recovery. It involves the processes of collecting, treating, and disposing of solid waste. Waste is collected from various sources and managed through disposal processes [3]. Solid waste refers to materials that are neither liquid nor soluble. It consists of complex and hazardous materials from industries, electronic waste, household waste and agricultural waste [17].

The types of waste management processes are landfill, incineration, composting, recycling and vermicomposting. The waste management process mainly depends on the accuracy of waste segregation. Conventional methods for waste segregation involve manual sorting. However, it is more time-consuming and leads to inefficiency in the recycling and disposal process. Recently, the integration of automated systems using Machine Learning (ML) and Deep Learning (DL) techniques has emerged as a promising solution to optimize waste management systems [7]. By using image classification models, these systems can autonomously classify waste materials. It improves sorting efficiency, accuracy and operational throughput.

Image-based waste classification techniques involve three main steps: preprocessing, feature extraction and classification [9] as shown in Figure 1. The preprocessing techniques use different filters to improve the appearance and resolution of waste images. Feature extraction processes identify visual features of texture, shape, and colours for model training. The classification model classifies them based on extracted features. The well-known ML and DL models are Support Vector Machine (SVM), AdaBoost, Convolutional Neural Network (CNN), ResNet and DenseNet models.

However, solid waste materials exhibit complex and heterogeneous characteristics that challenge traditional image classification models. Variations in texture, size, and appearance among different waste types make it difficult to distinguish them effectively. To address these limitations, in this work, a new DL model called Attention-Enhanced Residual Capsule Network (ARCNet) is proposed. The DL model integrates

Residual Attention Blocks (RABs) and Capsule Network (CapsNet) layers to improve the local and global feature discrimination ability of waste images.

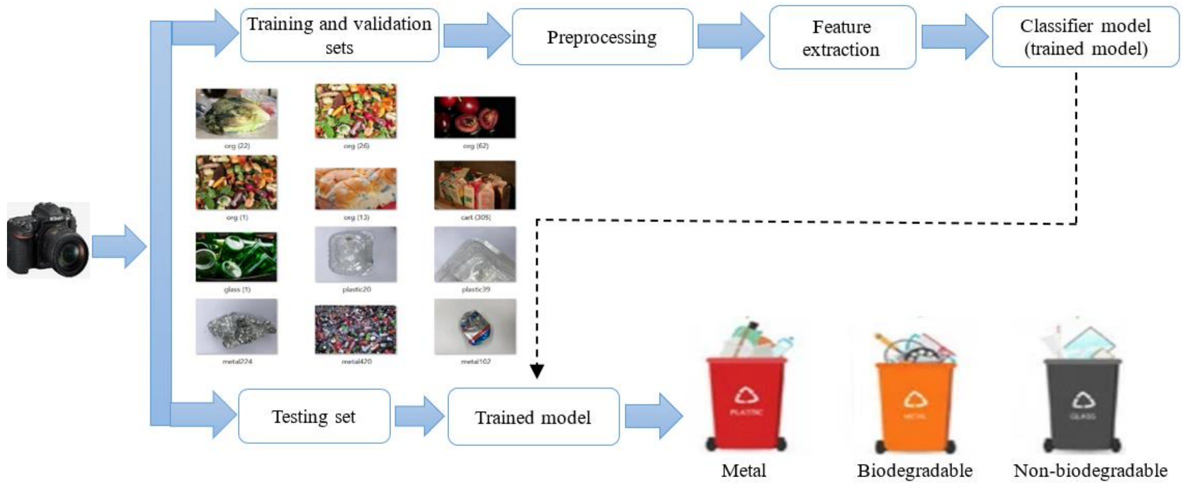

Figure 1. Image processing-based waste management system.

The RABs are designed to refine feature extraction by selectively focusing on key regions of interest in the image. It is used for the model to learn more robust and discriminative features. The CapsNet, on the other hand, is used to preserve the spatial relationships between features. It allows the model to better understand the hierarchical structure of waste materials. Additionally, the inclusion of a Global Context Attention Module (GCAM) further refines the model's ability to prioritise important regions of the image.

The paper is organised as follows: in section 2, which carries a literature review on existing systems, section 3 discusses the proposed model that includes its workflow, section 4 explores the result and discussion of the proposed model with an existing work comparison and section 5 summarises the work with a conclusion.

2. Related Work

Prakash *et al*. [18] explore the use of Q-learning for efficient path planning between waste-generating industries and waste-processing facilities. This approach supports real-time industrial waste management with higher reliability. Similarly, Rutqvist *et al*. [20] developed an ML combined embedded system for waste management by detecting recycling container emptying using sensor data. By using the Random Forest classifier, the accuracy improved from 86.8% to 99.1% and recall from 47.9% to 98.2% when compared to existing models.

Moni *et al*. [13] propose an Internet of Things (IoT)-based smart waste management system. It consists of an embedded system with sensors to monitor bin levels and notify authorities when thresholds are reached. In addition, it includes an Android app for coordinating activities and optimizing waste collection routes using the travelling salesman algorithm. Likewise, Nirde *et al*. [16] proposed a management system with a cloud interface. The status of a bin is updated using GSM technology.

Nijitha *et al*. [15] investigate the use of Hyper-Spectral Imagery (HSI) and spectral unmixing for urban solid waste. It applies sub-pixel detection and characterisation for waste classification. Compared to other systems, HSI HSI-based system shows less accuracy.

The hybrid CNN model-based solid waste management is proposed by Murugan *et al*. [14]. The hybrid model consists of recurrent layers in classification to improve the accuracy. Chowdhury *et al*. [4] address the challenges of object detection-based waste management in Bangladesh. They focused on real-time solid waste detection using optimized DL models. Experimental results on open-source and self-collected annotated images show that the optimized DL model achieves a classification accuracy of 73% and an F1-score of 0.729.

Li *et al*. [11] propose a Multi-model Cascaded Convolutional Neural Network (MCCNN) for detecting and classifying domestic waste. It combines YOLOv4 and recurrent layers to reduce false positives. Experimental results on the Large-Scale Waste Image Dataset (LSWID) with 52 waste categories show that the MCCNN model achieved a 10% improvement in detection precision.

Yang *et al*. [25] developed the Texture and Pixel Gradient Convolutional Network (TPCNet) to improve construction waste classification. This model uses a fusion strategy to integrate texture features with RGB images. Additionally, a Two-stage Feature Pyramid Network (TFPN) improves spatial multi-scale contextual information. Results on a large-scale real-world dataset show that TPCNet improves classification accuracy for similarly shaped waste by using texture features.

To detect hazardous waste, the DL model-based on the YOLOv8n network is proposed by Xiao *et al*. [24].

It includes a bi-level routing attention module to extract long-range dependencies and reduce false positives and negatives. Yang *et al*. [26] presented a vision-based garbage recognition system called GarbageNet for the classification of domestic waste. GarbageNet uses a weakly-supervised transfer learning model to mitigate mislabeled data effects.

Williams *et al*. [23] propose a new technique for discriminating nonferrous metals in commercial waste streams. It uses Magnetic Induction Spectroscopy (MIS) to measure how metal fragments scatter an excitation magnetic field at various frequencies. The results show that MIS alone achieves purity and recovery rates greater than 80% for most metal groups and exceeds 93% for stainless steel.

Ahmad *et al*. [1] proposed a double fusion model for waste classification. This model combines multiple DL models using both feature-level and score-level fusion methods. The double fusion scheme optimally integrates the capabilities of deep models through an early and late fusion strategy. Results on real-time images show that the fusion model achieves a significant improvement of 3.58% over existing techniques.

To avoid an unexpected mixing of different waste types during transportation and collection, He *et al*. [8] propose an improved supervision approach based on modified YOLOv3. In addition, the standard convolution is replaced with depth-wise separable convolution to improve the model's location accuracy and performance in multi-target scenarios. Results show that the modified YOLOv3 outperforms other detection algorithms and achieves on mean Average Precision (mAP) of 98.5%, which is 0.7% and 1.1% higher than YOLOv5l and EfficientDet-B0, respectively. Tian *et al*. [21] introduce the Convolutional Block Attention Module (CBAM) to improve spatial feature perception of the DL models in waste classification. They propose a lightweight model using CBAM called MobileNetV3 to improve the accuracy. Experimental results show a 3.6% improvement in recognition accuracy (96.55%) with reduced memory consumption and a 26.4ms recognition time per image.

Rupok *et al*. [19] introduce a DL model called ElectroSortNet for waste classification. ElectroSortNet is a modified CNN architecture which uses channel-based attention through squeeze and excitation networks. It also uses residual connections to improve accuracy and address vanishing gradients. The model achieves 97.44% test accuracy on a new dataset and outperforms previous architectures. Masand *et al*. [12] propose a DL model-based on the EfficientNet architecture for trash classification. They also created a new dataset of 8135 images by combining various source images. Implementation results show that the EfficientNet model achieved 92.87% accuracy for the classification. Gangopadhyay and Zhai [6] propose a DL model called Compost Green-Brown Network (CGBNet) to differentiate green and brown compost. It uses transfer learning to achieve a 95% accuracy in classification.

To overcome minimum data set images, the data augmentation combined DL model is proposed by Yang and Li [27] for waste segregation. Initially, the data set image counts are increased by using data augmentation techniques like flipping, rotation and inversion. Then, the DL model of Waste Segregation Network (WasNet) is used for classification. Compared to another model, the tunable parameter counts are very low in the WasNet model. Tran *et al*. [22] proposed a municipal solid waste classification model using transformer-based models. In transformer models, the attention mechanism is used to capture long-range dependencies without using sequential processing.

3. Attention-Enhanced Residual Capsule Network (ARCNet)

This work proposed a new DL model, ARCNet, for solid waste image classification to improve waste management efficiency. ARCNet combines RABs with CapsNet layers to strengthen feature extraction and effectively capture spatial hierarchies. This is used to differentiate different materials and textures in solid waste images. To further enhance the model's focus on critical image features, a GCAM is used for refined feature representation. The model's parameters are optimized using the Black Eagle Optimizer (BEO) as shown in Figure 2.

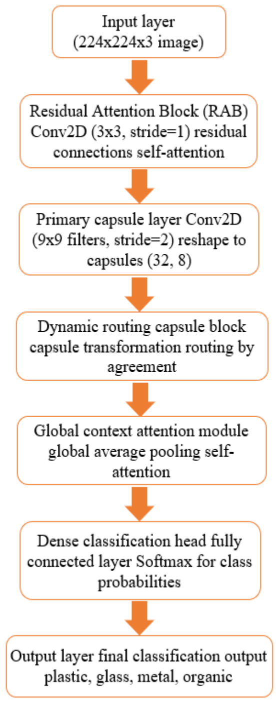

Figure 2. ARCNet architecture.

3.1. Residual Attention Block (RAB)

The RAB combines a residual structure with an attention mechanism. The residual connection allows the model to maintain information across layers and attention highlights relevant parts of the image. The structure of RAB is shown in Figure 3.

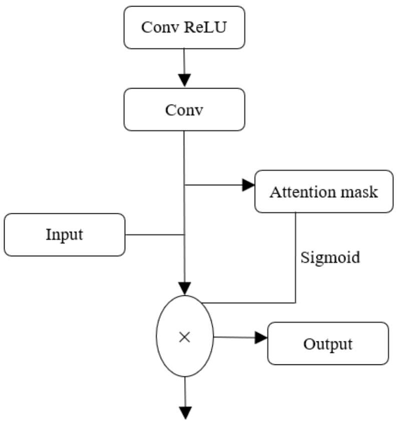

Figure 3: RAB structure.

The output of the residual block is calculated as follows:

$$x_{conv1} = ReLU(Conv(x, W_1)) \quad (1)$$

$$x_{conv2} = Conv(x_{conv1}, W_2)) \quad (2)$$

where σ is the sigmoid activation function. The attention layer assigns a weight to each feature map location. The attention weights are computed as follows:

$$attention = \sigma(Conv(x_{conv2}, W_{attn})) \quad (3)$$

3.2. Black Eagle Optimizer (BEO)

Optimization algorithms are used in fields of engineering to achieve the objective function [5, 10]. The BEO is a nature-inspired optimization algorithm based on the hunting and flight behaviors of the black eagle. BEO mimic natural processes to solve complex optimization problems [2, 28]. Compared to other optimization algorithms, BEO shows better performance in terms of convergence accuracy, stability, and global optimization ability.

The black eagle is known for its hunting prowess. It exhibits several key behaviors that have been adopted in the BEO algorithm: Initial search and exploration, prey encirclement, attack and capture and adaptability. The black eagle starts its hunt by soaring over a large area to search for potential prey. During this phase, it covers a wide region to maximize the chances of finding food. In prey encirclement, it begins to circle around it. This process reduces the distance between itself and the prey. Finally, the eagle dives toward its prey with precision to capture it. This behavior is modelled as the attack phase in the BEO algorithm. The solution moves towards the optimal region with high accuracy in this phase. The black eagle adjusts its flight and hunting strategy based on the environment and the behavior of the prey. Similarly, the BEO algorithm adapts its search behavior based on the landscape of the objective function. It allows the balance between exploration and exploitation. This process is mathematically expressed as follows:

- **Position Update**

The position of each candidate solution is updated iteratively. It can be modelled as follows:

$$X_i^{t+1} = X_i^t + r_1\left(X_{gb} - X_i^t\right) + r_2 . randn(0,1) \quad (4)$$

Where, X_i^t is the current position of the i-th particle at time t. X_{gb} is the global best solution found so far. r_1 *and* r_2 are random coefficients controlling the exploration behavior. $randn$ represents a random Gaussian noise term that adds variability to the position update.

- **Prey Encirclement**

When the search focuses on narrowing down the search region, the radius of exploration shrinks as the algorithm gets closer to the optimal solution. The radius R_t decreases with each iteration:

$$R_t = R_0\left(1 - \frac{t}{T}\right) \quad (5)$$

where R_0 is the initial radius, t is the current iteration, and T is the total number of iterations.

- **Attack Phase**

When the global best solution is located, the algorithm attacks by moving the candidate solutions directly toward it. It is used to speed up the convergence toward the optimal solution. It can be modelled as follows:

$$X_i^{t+1} = X_{gb} + \gamma . \text{rand}\,(-\delta, \delta) \quad (6)$$

Where, γ and δ control the magnitude and direction of the attack. It allows a more aggressive convergence toward the best solution.

- **Termination Phase**

The algorithm terminates once the maximum number of iterations T is reached or if a satisfactory fitness value is obtained. The algorithm checks if the fitness has stabilized or if the maximum number of iterations is met:

$$stop\ if\ \left|f\left(X_{gb}^{t+1}\right) - d\left(X_{gb}^t\right)\right| < \epsilon \quad (7)$$

where ϵ is a small threshold value that indicates when further optimization is unlikely to produce significant improvements.

- **Primary Capsule Layer**

The primary capsule layer extracts capsules from the feature map produced by the RAB. Each capsule represents various properties like position, orientation, and texture of objects in the image. The feature map from the RAB is transformed into capsules using a convolution operation as follows:

$$x_{caps} = Conv(x_{RAB}, W_{caps}) \quad (8)$$

where W_{caps} are the weights of the convolutional layer in the primary capsule. The resulting feature map is reshaped into capsules as follows:

$$capsules = x_{caps}.reshape(N, C, H, W) \dashrightarrow Primary\ capsules \quad (9)$$

Where N, C, H, and W are the batch size, number of capsules, height, and width, respectively. Each capsule vector retains spatial relationships and orientation information, essential for waste object detection.

- **Dynamic Routing Capsule Block**

This block refines the representation from the primary capsules by routing the capsules to higher-level features. Each primary capsule u_i is transformed into higher-level capsules through a transformation matrix W_{ij} as follows:

$$\widehat{u_{j|i}} = W_{ij} u_i \quad (10)$$

where W_{ij} is the transformation matrix, u_i is the output of the primary capsule i, and $\widehat{u_{j|i}}$ is the predicted output capsule j. The routing mechanism adjusts the weights between capsules c_{ij} to ensure that capsules agree on the final output as follows:

$$c_{ij} = \frac{\exp(b_{ij})}{\sum_k \exp(b_{ik})} \quad (11)$$

$$s_j = \sum_i c_{ij} \widehat{u_{j|i}} \quad (12)$$

The squash function ensures the output vector length is between 0 and 1, representing the likelihood of a specific feature being present as follows:

$$v_j = squash(s_j) = \frac{||s_j||^2}{1 + ||s_j||^2} \cdot \frac{s_j}{||s_j||} \quad (13)$$

where c_{ij} are the coupling coefficients, s_j is the weighted sum of input capsules, and v_j is the output of capsule j after applying the squash nonlinearity. The coupling coefficients are updated iteratively to improve agreement as follows:

$$b_{ij} = b_{ij} + \widehat{u_{j|i}} \cdot v_j \quad (14)$$

3.3. Global Context Attention Module (GCAM)

The GCAM improves the feature map by assigning weights across the entire feature map which improves context awareness. The structure of GCAM is shown in Figure 4. The feature map is condensed into a vector using Global Average Pooling (GAP) which computes the mean of each feature map as follows:

$$x_{gap} = \frac{1}{H \times W} \sum_{i=1}^{H} \sum_{j=1}^{W} x_{i,j} \quad (15)$$

where x_{gap} represents the average-pooled feature vector. The attention weights are computed from the GAP vector as follows:

$$attn = \sigma(Conv(x_{gap}, W_{gcam}) \quad (16)$$

The feature map is multiplied with the attention weights to focus on significant areas as follows:

$$x_{GCAM} = x.attn \quad (17)$$

The final dense layer generates the classification output. The output from GCAM is flattened into a 1D vector as follows:

$$z = Flatten(x_{GCAM}) \quad (18)$$

The flattened vector passes through a dense layer followed by a softmax activation to generate probabilities for each class as follows:

$$y_{logits} = W_{fc}.z + b \quad (19)$$

$$y = Softmax(y_{logits}) \quad (20)$$

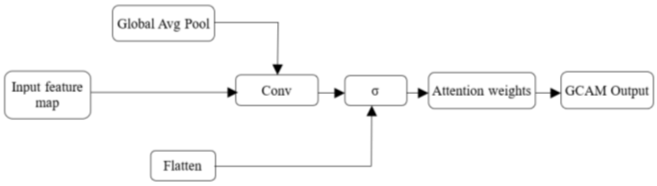

Figure 4. GCAM structure.

3.4. Parameter Tuning

In ARCNet, the accuracy of the model depends on various hyperparameters. It needs to be tuned for optimal performance. BEO is used to fine-tune these hyperparameters. The tuned parameters include Learning Rate (LR), number of RAB, capsule dimensions, dynamic routing iterations and GCAM Parameters. The LRs control how much the weights are updated during training. The capsule dimensions define the number of dimensions in each capsule vector. The higher dimensions can capture more complex features, but increase computational cost. The GCAM parameters include kernel sizes and attention coefficients that determine how the attention is applied across feature maps.

- **Pseudocode for Parameter Tuning in ARCNet with BEO**

BEGIN: ARCNet Parameter Tuning using BEO

- *Step* 1: Initialize parameters.

  Define the total population size (P).
  Set the maximum number of iterations (Tmax).
  Define the boundaries for each hyperparameter:
  Randomly initialize a population of candidate solutions (hyperparameters).
  Set the exploration radius (R0) and decay rate for narrowing the search space.

- *Step* 2: Evaluate initial fitness.

  FOR each candidate solution in the population:
  Configure the ARCNet model using the candidate's hyperparameters.
  Train ARCNet on the training dataset.
  Compute fitness for each candidate (validation accuracy).
  END FOR
  Identify and store the global best solution (X_{gb}) with the highest fitness value.

- *Step* 3: Optimization loop

  FOR t=1 to Tmax:
  DISPLAY “Iteration:”, t

  a) Exploration Phase

     FOR each candidate solution (Xi):
     -Update the position of the candidate based on global best solution (X_{gb}).
     -Incorporate Gaussian noise for diversity in exploration.
     END FOR

  b) Encirclement phase

     -Narrow the search radius (Rt) as iterations progress to refine solutions.
     -Adjust candidate solutions to move closer to the global best (Xgb).

  c) Attack phase

     -Intensify the search around X_{gb} to accelerate convergence.
     -Move candidates aggressively towards X_{gb} while maintaining diversity.

  d) Fitness evaluation

     FOR each updated candidate solution:
     -Configure the ARCNet model with updated hyperparameters.
     -Train ARCNet on the training dataset.
     -Compute fitness for the updated solution.
     END FOR
     -Update X_{gb} if a better solution is found.

  e) Log progress:

     -Record the current best fitness value and corresponding hyperparameters.

  END FOR

- *Step* 4: Termination

  IF maximum iterations reached OR fitness improvement is negligible:
  DISPLAY "Optimization Complete.
  -Return the best hyperparameter set (X_{gb}).
  END IF

- *Step* 5: Final model training

  -Train ARCNet on the full dataset using the optimized hyperparameters (X_{gb}).
  -Validate and evaluate ARCNet’s performance on the test dataset.

END

A population of candidate solutions is created with each candidate representing a set of hyperparameters. Each candidate is randomly initialized within predefined boundaries. The total population size, maximum iterations, and the initial exploration radius are also set at this stage. The fitness of each candidate is then evaluated by training ARCNet with the corresponding hyperparameters on the training dataset.

The optimization process proceeds in iterative phases. During the exploration phase, candidate solutions are updated based on their positions relative to the global best. In the encirclement phase, the search radius shrinks with each iteration. The attack phase follows, where candidate solutions are aggressively moved toward X_gb to accelerate convergence. After each phase, the updated candidates are re-evaluated by training ARCNet with their respective hyperparameters and recalculating their fitness values. If a better solution is found, X_gb is updated accordingly. The process continues until either the maximum number of iterations is reached or the improvement in fitness becomes converged. The final optimized hyperparameter set (X_gb) is then used to train ARCNet on the complete dataset. Table 1 below summarizes the hyperparameters with their boundary ranges.

Table 1. Hyperparameters with their boundary ranges.

Parameter	Description	Range
LR	Step size for model training	[0.0001, 0.01]
Number of RABs	Count of RABs	[1, 5]
Capsule dimensions	Dimensions per capsule vector	[8, 16]
Routing iterations	Iterations for dynamic routing	[1, 5]
GCAM kernel size	Convolution kernel size for GCAM	[3, 7]

4. Results and Discussions

The dataset is collected from Mendeley Data (https://data.mendeley.com/datasets/n3gtgm9jxj/2). It consists of 2,827 coloureds .jpg images categorised as 329 biodegradables, 1,210 metal, and 1,288 non-biodegradable images. The images vary in orientation and resolution, ranging from 191 pixels to 264 pixels. The dataset is divided into training and test subsets with a 70:30 ratios. The test set consisted of 82 biodegradables, 299 metal, and 403 non-biodegradable images. All images were resized to a uniform dimension of 224×224 pixels using bilinear interpolation. Although no augmentation is applied during training, standard preprocessing steps, including normalization are used to scale pixel values between 0 and 1. The visualisation of the dataset images is shown in Figure 5.

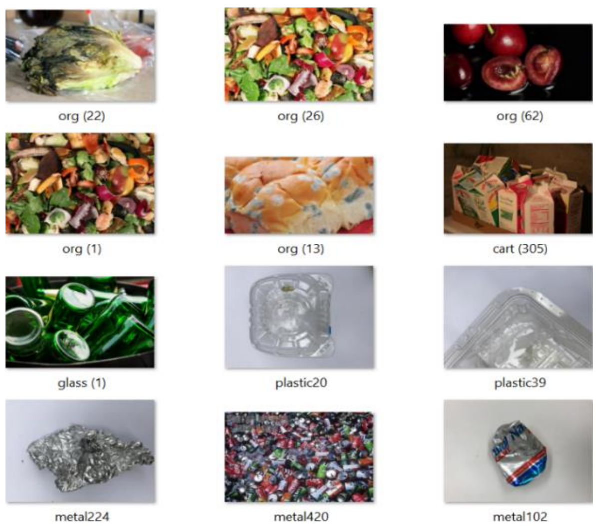

Figure 5. Dataset visualization.

The proposed model is coded in Python and evaluated in Python IDLE 3.8. The TensorFlow 3.10 package is used on an NVIDIA GeForce RTX GPU for accelerated computations. For parameter optimization, the total population size was set to 20 candidate solutions, and the optimization process was run for 30 iterations**.** For model training, BEO searched over a defined range of candidate values for each hyperparameter: capsule dimensions [8, 16] and routing iterations [1, 5]. The optimal configuration of 8 capsules with 16 dimensions and 3 routing iterations is identified based on its superior validation accuracy and stability across folds. The accuracy and loss evaluation of the proposed model for both training and testing are given in Figures 6 and 7. It is observed that the model has learned most of the patterns from the training data. The model shows good generalization performance. Both the training and testing accuracies show an upward trend, indicating that the model is learning and improving its predictions on both the training and unseen testing data.

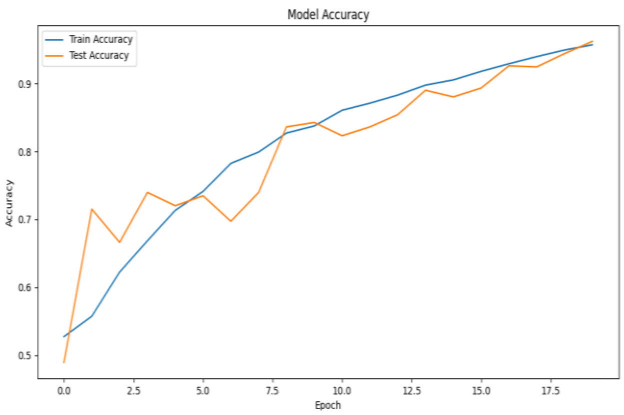

Figure 6. Model accuracy evaluation.

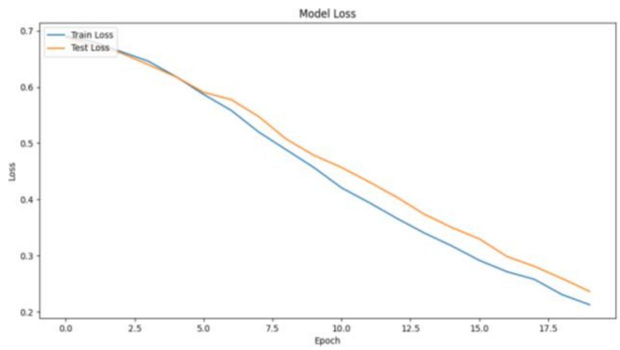

Figure 7. Model loss evaluation.

To assess the performance of models, the metrics like sensitivity, accuracy, F1-score, and precision. These metrics are derived from the confusion matrix that includes True Negatives (TN), True Positives (TP), False Negatives (FN) and False Positives (FP). Then, the performance metrics can be computed as follows:

$$Accuracy = \frac{TP + TN}{(TP + TN + FP + FN')} \tag{21}$$

$$Recall = \frac{TP}{(TP + FN')} \tag{22}$$

$$Precision = \frac{TP}{(TP + FP')} \tag{23}$$

$$F1\ score = 2.\frac{Precision\,.\,Recall}{(Precision + Recall\,)} \tag{24}$$

The performance of the models is given in Table 2. All models were trained and evaluated on the same hardware environment with an 85:15 ratio. ARCNet with BEO achieved the highest accuracy, precision, recall and F-score rate of 98.34%, 98.28, 98.9, and 98.64 respectively. It proves the model's ability to classify solid waste images correctly with balanced performance. In contrast, ARCNet without BEO achieved the accuracy, precision, recall and F-Score rate of 96.04%, 96.03, 96.04 and 96.03, respectively. It is evident that the usage of BEO in a model for higher accuracy. Other models, such as MCCNN and GarbageNet show moderate performance with accuracies of 94.40% and 92.6%, respectively. These models failed to match the robustness and adaptability of ARCNet. Overall, the integration of BEO into ARCNet not only enhances its classification accuracy but also increases its precision, recall, and F1-score rates. The performance results are graphically shown in Figure 8.

Table 2. Comparison with other models.

Method	Accuracy	Precision	Recall	F1-score
ARCNet with BEO	98.34	98.28	98.9	98.64
ARCNet without BEO	96.04	96.03	96.04	96.03
MCCNN	94.40	94.50	94.40	94.44
GarbageNet	92.6	92.73	92.66	92.43
ElectroSortNet	87.5	87.8	87.2	86.8
CBAM	85.3	85.6	85.3	85.4

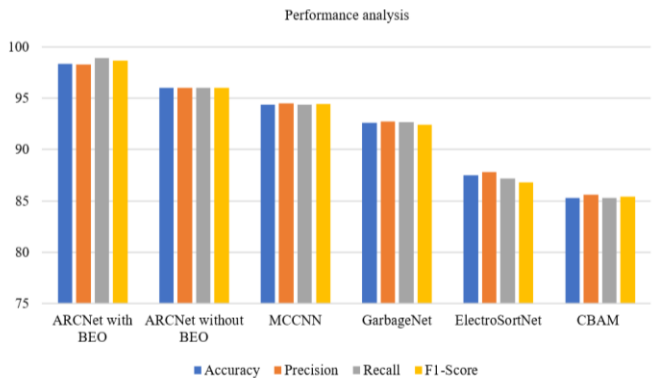

Figure 8. ROC analysis.

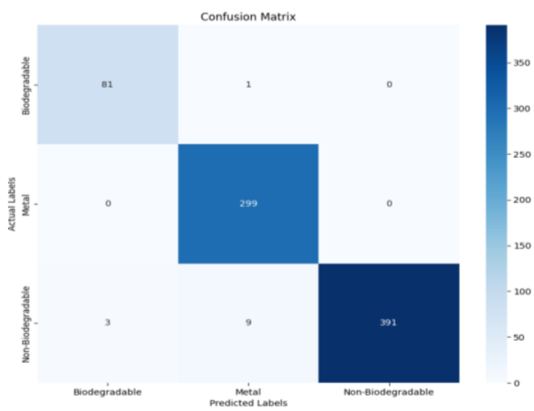

Figure 9. Confusion matrix of the proposed model.

The confusion matrix of the proposed model is given in Figure 9. The model correctly classifies 81 out of 82 biodegradable objects with only 1 misclassification as non-biodegradable. The model correctly classifies all 299 Metal objects, with no misclassifications. The model correctly classifies 391 out of 394 non-biodegradable objects with 3 misclassifications as Biodegradable. The confusion matrix indicates that the model has high accuracy and precision with minimal misclassifications. The validation results of the wastes are given in Figure 10.

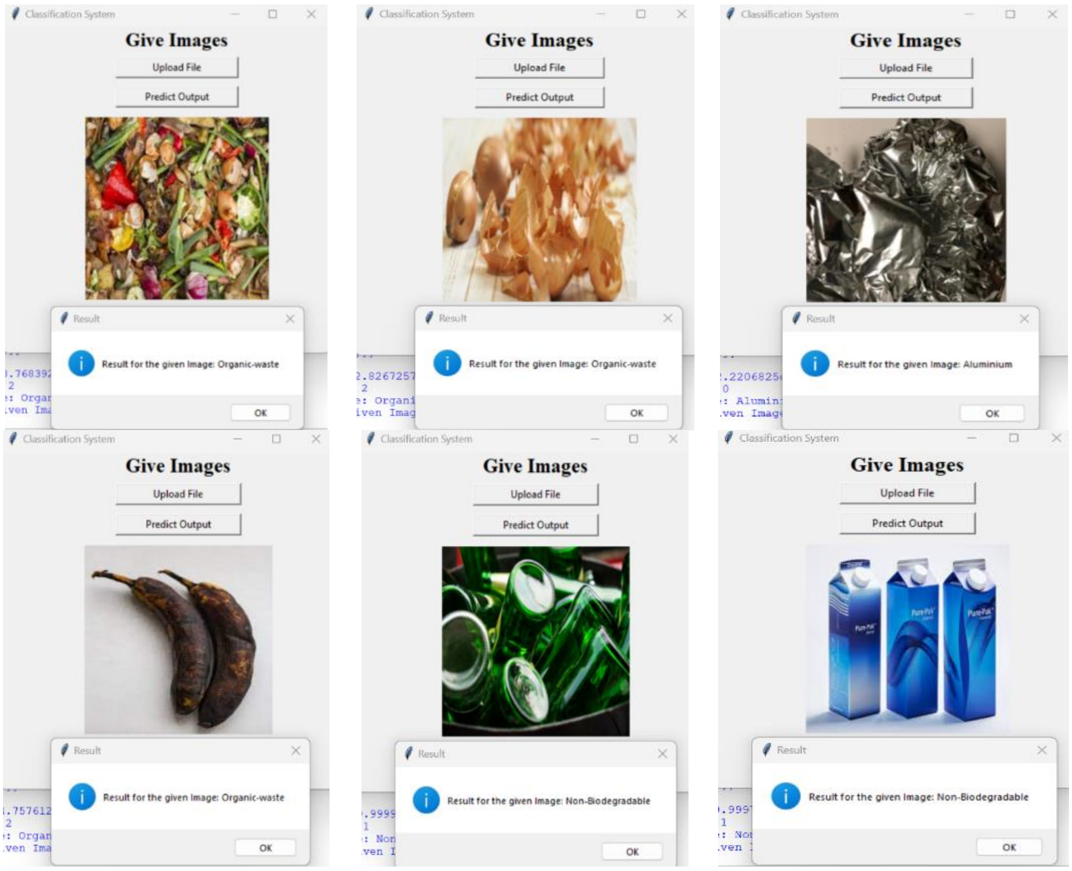

Figure 10. Validation results

To assess the stability and robustness of ARCNet, five independent training runs are performed using different random seeds and dataset shuffles. The mean accuracy achieved was 98.21% with a standard deviation of ±0.12. It is observed that ARCNet shows consistent performance across different initializations and data splits. Similarly, the F1-score exhibited low variance with an average of 98.52% ± 0.15. It also proves the model's robustness and generalizability.

Table 3. Ablation study of ARCNet components.

Model variant	Accuracy (%)	Precision (%)	Recall (%)	F1-score (%)
ARCNet (Full model with BEO)	98.34	98.28	98.90	98.64
Without GCAM	96.87	96.75	97.01	96.88
RAB replaced with standard residual block	95.92	95.78	96.14	95.96
Without CapsNet (using CNN instead)	94.65	94.48	94.72	94.60
Without BEO (static hyperparameters)	96.04	96.03	96.04	96.03

The ablation study of the proposed model is given in Table 3. The full model achieved the highest accuracy of 98.34%. The accuracy values drop to 96.87% when GCAM is removed. It proves the role of GCAM in increasing global contextual awareness. Likewise, the replacement of RAB with a standard residual block reduced accuracy to 95.92%, which confirms the importance of attention mechanisms in improving feature refinement. In addition, the static hyperparameters instead of BEO optimisation reduced accuracy to 96.04%. It proves the importance of an optimizer in fine-tuning model performance.

5. Conclusions

This work proposes a novel DL-based architecture for solid waste image classification in waste management systems. By integrating attention blocks and CapsNet layers, the proposed model achieves superior feature extraction and hierarchical representation of waste images. The use of the BEO further enhances the model's performance by effectively optimizing its parameters. Experimental results show that ARCNet overcomes existing models in accuracy, precision, and recall rates. The proposed approach contributes to efficient waste management and environmental sustainability.

Ethical Approval

Ethical approval was not required for this study as it does not involve human participants, animal studies, or sensitive data. The study used publicly available air quality datasets for training and testing the proposed model.

Consent to Participate

Informed consent was not required for this study, as the data used for training the air quality prediction model was publicly available and did not involve human subjects. No personal data was collected, and the study adheres to ethical guidelines for the use of publicly available datasets.

Data Availability Statement

The datasets used in this study are collected manually. For evaluation, the data set is used from Kaggle website.

References

[1] Ahmad K., Khan K., and Al-Fuqaha A., "Intelligent Fusion of Deep Features for Improved Waste Classification," *IEEE Access*, vol. 8, pp. 96495-96504, 2020. https://ieeexplore.ieee.org/document/9096360

[2] Allam M. and Malaiyappan N., "Hybrid Feature Selection based on BTLBO and RNCA to Diagnose the Breast Cancer," *The International Arab Journal of Information Technology*, vol. 20, no. 5, pp. 727-737, 2023. https://doi.org/10.34028/iajit/20/5/5

[3] Arnold E., Fernandez P., Larcena J., Mission M., and et al., "Household Awareness and Participation on Waste Disposal: An Effective Solid Waste Management Amidst COVID-19 Pandemic," *in Proceedings of the IEEE 13th International Conference on Humanoid, Nanotechnology, Information Technology, Communication and Control, Environment, and Management*, Manila, pp. 1-6, 2021. https://ieeexplore.ieee.org/document/9731888

[4] Chowdhury T., Sinthiya N., Shanta S., Hasan T., and et al., "Object Detection Based Management System of Solid Waste Using Artificial Intelligence Techniques," *in Proceedings of the IEEE 13th Annual Ubiquitous Computing, Electronics and Mobile Communication Conference*, New York, pp. 19-23, 2022. https://ieeexplore.ieee.org/document/9965643

[5] Ganesamoorthy N., Sakthivel B., Subbramania D., and Balasubadra K., "Hen Maternal Care Inspired Optimization Framework for Attack Detection in Wireless Smart Grid Network," *International Journal of Informatics and Communication Technology*, vol. 13, no. 1, pp. 123-130, 2024. https://ijict.iaescore.com/index.php/IJICT/article/view/20588

[6] Gangopadhyay S. and Zhai A., "CGBNet: A Deep Learning Framework for Compost Classification," *IEEE Access*, vol. 10, pp. 90068-90078, 2022. https://ieeexplore.ieee.org/document/9864578

[7] George A., Mankar S., Suryanarayana N., Gupta M., and et al., "Deep Learning Based Identification of Solid Waste Management in Smart Cities through Garbage Separation and Monitoring," *in Proceedings of the 5th International Conference on Inventive Research in Computing Applications*, Coimbatore, pp. 337-342, 2023. https://ieeexplore.ieee.org/document/10220690

[8] He Y., Li J., Chen S., Xu Y., and Liang J., "Waste Collection and Transportation Supervision Based on Improved YOLOv3 Model," *IEEE Access*, vol. 10, pp. 81836-81845, 2022. https://ieeexplore.ieee.org/document/9848790

[9] Jin L., "A Type-II Fuzzy Optimization Model for Municipal Solid Waste Management," *in Proceedings of the 8th International Conference on Fuzzy Systems and Knowledge Discovery*, Shanghai, pp. 55-59, 2011. https://ieeexplore.ieee.org/document/6019534

[10] Kavitha V., Sakthivel B., Deivasigamani S., Jayaram K., and et al., "An Efficient Modified Black Widow Optimized Node Localization in Wireless Sensor Network," *IETE Journal of Research*, vol. 70, no. 12, pp. 8404-8413, 2024. https://doi.org/10.1080/03772063.2024.2384486

[11] Li J., Chen J., Sheng B., Li P., and et al., "Automatic Detection and Classification System of Domestic Waste via Multimodel Cascaded Convolutional Neural Network," *IEEE Transactions on Industrial Informatics*, vol. 18, no. 1, pp. 163-173, 2022. https://ieeexplore.ieee.org/document/9444810

[12] Masand A., Chauhan S., Jangid M., Kumar R., and Roy S., "ScrapNet: An Efficient Approach to Trash Classification," *IEEE Access*, vol. 9, pp. 130947-130958, 2021. https://ieeexplore.ieee.org/document/9531608

[13] Moni J., Jacob P., Pawar S., Varghese R., and et al., "An Intelligent Smart Bin System for Solid Waste Management in Smart Cities," *in Proceedings of the International Conference on Decision Aid Sciences and Applications*, Chiangrai, pp. 1083-1087, 2022. https://ieeexplore.ieee.org/document/9765020

[14] Murugan P., Muneeswaran V., Vardhan N., Venkatesh B., and et al, "Deep Learning-based Hybrid Image Classification Model for Solid Waste Management," *in Proceedings of the International Conference on Cognitive Robotics and Intelligent Systems*, Coimbatore, pp. 168-172, 2024.

https://ieeexplore.ieee.org/document/10534019

[15] Nijitha P., Kumar M., Pankaj D., and Nidamanuri R., "Solid Waste Detection and Waste-Material Characterization in Urban Environment at Subpixel Level in Airborne Hyperspectral Imagery," *in Proceedings of the 13th Workshop on Hyperspectral Imaging and Signal Processing: Evolution in Remote Sensing*, Athens, pp. 1-3, 2023. DOI: 10.1109/WHISPERS61460.2023.10430865

[16] Nirde K., Mulay P., and Chaskar U., "IoT Based Solid Waste Management System for Smart City," *in Proceedings of the International Conference on Intelligent Computing and Control Systems*, Madurai, pp. 666-669, 2017. https://ieeexplore.ieee.org/document/8250546

[17] Oralhan Z., Oralhan B., and Yigit Y., "Smart City Application: Internet of Things (IoT) Technologies Based Smart Waste Collection Using Data Mining Approach and Ant Colony Optimization," *The International Arab Journal of Information Technology*, vol. 14, no. 4, pp. 11-15, 2017. https://www.iajit.org/upload/files/Smart-City-Application-Internet-of-Things-IoT-Technologies-Based-Smart-Waste-Collection-Using-Data-Mining-Approach-and-Ant-Colony-Optimization.pdf

[18] Prakash A., Agarwalla N., Priyadarshini R., and Misra R., "A Novel Approach Towards Industrial Waste Management Using Q-Learning," *in Proceedings of the International Conference on Intelligent Computing and Remote Sensing*, Bhubaneswar, pp. 1-6, 2019. https://ieeexplore.ieee.org/document/9555893

[19] Rupok H., Sourov N., Anannaya S., Afroz A., and et al., "ElectroSortNet: A Novel CNN Approach for E-Waste Classification and IoT-Driven Separation System," *in Proceedings of the 3rd International Conference on Advancement in Electrical and Electronic Engineering*, Gazipur, pp. 1-6, 2024. https://ieeexplore.ieee.org/document/10561784

[20] Rutqvist D., Kleyko D., and Blomstedt F., "An Automated Machine Learning Approach for Smart Waste Management Systems," *IEEE Transactions on Industrial Informatics*, vol. 16, no. 1, pp. 384-392, 2020. https://ieeexplore.ieee.org/document/8709695

[21] Tian X., Shi L., Luo Y., and Zhang X., "Garbage Classification Algorithm Based on Improved MobileNetV3," *IEEE Access*, vol. 12, pp. 44799-44807, 2024. https://ieeexplore.ieee.org/document/10478519

[22] Tran H., Le T., Pham H., Vu M., and Dao S., "An Integrated Learning Approach for Municipal Solid Waste Classification," *IEEE Access*, vol. 12, pp. 176569-176585, 2024. https://ieeexplore.ieee.org/document/10750192

[23] Williams K., O'Toole M., and Peyton A., "Scrap Metal Classification Using Magnetic Induction Spectroscopy and Machine Vision," *IEEE Transactions on Instrumentation and Measurement*, vol. 72, pp. 1-11, 2023. https://ieeexplore.ieee.org/document/10147806

[24] Xiao Z., Yang G., Wang X., Yu G., and et al., "DBS-YOLO: A High-Precision Object Detection Algorithm for Hazardous Waste Images," *IEEE Transactions on Instrumentation and Measurement*, vol. 73, pp. 1-15, 2024. https://ieeexplore.ieee.org/document/10716561

[25] Yang J., Wang G., Sun Y., Bai L., and Yang B., "Deep Pixel-Wise Textures for Construction Waste Classification," *IEEE Transactions on Automation Science and Engineering*, vol. 21, no. 3, pp. 4732-4742, 2024. https://ieeexplore.ieee.org/document/10322658

[26] Yang J., Zeng Z., Wang K., Zou H., and Xie L., "GarbageNet: A Unified Learning Framework for Robust Garbage Classification," *IEEE Transactions on Artificial Intelligence*, vol. 2, no. 4, pp. 372-380, 2021. https://ieeexplore.ieee.org/document/9435085

[27] Yang Z. and Li D., "WasNet: A Neural Network-Based Garbage Collection Management System," *IEEE Access*, vol. 8, pp. 103984-103993, 2020. https://ieeexplore.ieee.org/document/9107232

[28] Zhang H., San H., Chen J., Sun H., and et al., "Black Eagle Optimizer: A Metaheuristic Optimization Method for Solving Engineering Optimization Problems," *Cluster Computing*, vol. 27, pp. 12361-12393, 2024. https://link.springer.com/article/10.1007/s10586-024-04586-1

Velmurugan Lingamuthu is an Associate Professor, School of Computing Science and Engineering, VIT Bhopal University, India. He attained his undergraduate degree in Electronics (1994) from Bharathiar University in Coimbatore, India, under the guidance of the Professor Pasuthathi and Professor Venkatachalem. Velmurugan had received his Master of Computer Application from the Madurai Kamaraj University at Madurai (India) in 1997 under the supervision of the Professor Petchimuthu. Velmurugan had received his Ph.D. in Computer Science from the Chennai University at Chennai (India) in 2013 under the supervision of the Professor P.Thangaraj. He has worked in various institutions in the position of lecturer, assistant professor and associate professor. His research fields include Mobile Computing, Distributed Computing and Artificial Intelligence.

Devaraju Sellappan is a Senior Assistant Professor in the School of Computing Science and Engineering (SCSE) at VIT Bhopal University. He received his Doctorate from Anna University, an M.Phil in Computer Science from Periyar University and an MCA in Computer Applications from Periyar University. He has an academic experience of 19 years and 2 years of Industry at various levels. He has 20 publications in peer reviewed International/National journals and has 20 publications in various International/National Conferences. He has organized/Attended a number of Conferences, Workshops and FDP. He has published 2 patents, 5 Book Chapters. He is a lifetime member of IACSIT and ISTE, India. His area of research includes Network Security, Intrusion Detection, Soft Computing, and Wireless Communication.

Sakthivel Balu received his BE degree in Electronics and Communication Engineering from SACS MAVMMM Engineering College, Madurai, India and ME degree and PhD in VLSI design from Anna University, Chennai, India. He is currently working as an Associate Professor in PSY Engineering College at Sivagangai, India. His areas of interest are VLSI Architecture Design, particularly as applied to data path circuits like adders and multipliers. In addition to his expertise in VLSI, he has made significant contributions to the field of Artificial Intelligence (AI), Deep Learning (DL), and Data Science. He has worked extensively on AI and DL models applied to Image Processing, Time-Series Prediction, and Medical Diagnostics, with a focus on applications such as Medical Image Segmentation, Air Quality Prediction, and Environmental Monitoring.

Sheik Raghabu received his BE degree in Electronics and Communication Engineering from M.I.E.T, Madurai, in 2015. He later completed his ME degree in VLSI Design from M.I.E.T in 2019. Currently, he is pursuing his PhD part-time at Anna University, specializing in VLSI design. He started his academic career as an Assistant Professor at M.I.E.T Engineering College, Madurai, India. He has been working as an Assistant Professor at PSY Engineering College, Sivagangai, India. He is particularly focused on Leveraging AI models to Optimize Agricultural Processes, such as Precision Farming, Crop Prediction, and Pest Detection, using IoT and Sensor Data. His research interests include VLSI Architecture Design, Particularly Focusing on Optimizing Circuits like adders and multipliers.

Aanjankumar Sureshkumar is presently working as an Assistant Professor at VIT Bhopal University, India, and holds a Doctorate from Anna University in the topic of Botnet Security and an M.E. in Software Engineering. He has twelve years of academic experience by serving as Associate Professor, UG-HoD, and other positions. He has 11 articles in international conferences conducted in and outside India and 14 papers in peer-reviewed national and international journals with high impact factors. He has authored a book titled Graph Theory and Applications under the Anna University syllabus. He is also a reviewer for SCI journals, and he has 4 SCI journal publications with an impact factor above 3 and one IEEE transaction which impact factor is 10.9, 26 Scopus-indexed publications, 6 book chapters, 2 Indian patents, and 1 UK design patent that he has authored and published. His main areas of research include Artificial Intelligence, Software Engineering, Cyber Forensics, Network Security, and Human-Computer Interaction.

Prabu Govindarajan is a Senior Assistant Professor Grade-II in the school of Computing Science and Engineering, VIT Bhopal University, Madhya Pradesh. He has received his Ph.D. in 2019 and has done his Master degree in Information Technology in 2011 from Kalasalingam Academy of Research and Education. He has served different academic and administrative roles in various institutions since 2011. Dr. Prabu Kanna G has several publications, in peer reviewed International and National journals with high impact factor. He has contributed a number of Conferences, Workshops and Faculty Development Program. His area of research includes Cloud Computing, Big Data, Artificial Intelligence, and Machine Learning.