An AIoT Device for Raising Awareness about Trash Classification at Source

Vo Ngoc Sang[1,2], Nguyen Ngoc Thanh Xuan[1,2], Le Gia Phat[1,2], Nguyen Lam Tam Nhu[1,2], Ho Tri Khang[1,2], Tran Tien Phat[1,2], and Hoang-Anh Pham[1,2,∗][0000−0002−5806−5910]

[1] Ho Chi Minh City University of Technology (HCMUT), 268 Ly Thuong Kiet Street, Ward 14, District 10, Ho Chi Minh City, Vietnam
[2] Vietnam National University Ho Chi Minh City (VNU-HCM), Linh Trung Ward, Ho Chi Minh City, Vietnam
[∗] Corresponding author: anhpham@hcmut.edu.vn

Abstract. Waste segregation is a critical issue for environmental protection and sustainable growth. In Vietnam, public awareness and action on waste separation at source remain limited, highlighting the importance of engaging individuals, particularly students, in transforming waste disposal practices. Modern technologies, including the Internet of Things (IoT) and Artificial Intelligence (AI), have revolutionized various aspects of our lives and offer promising solutions to raise public awareness on this issue. This paper proposes an IoT device named BEG (BACHKHOA Eco-friendly Guide) integrating AI-based Computer Vision technology to classify waste via a camera. Unlike existing smart trash cans, which classify and dispose of the trash automatically, our device provides information about the waste type to guide users on proper disposal, thus reinforcing awareness of garbage classification at source. We also introduce the BEGNet, a Convolutional Neural Network (CNN) employing RegNetY120 as its backbone, which demonstrates superior performance in accuracy compared to other approaches on both the Trashnet dataset and our custom dataset - BKTrashImage. The proposed BEG device will improve knowledge about waste segregation, reduce improperly disposed waste, and foster a thriving circular economy.

Keywords: Artificial Intelligence· Computer Vision· Internet of Things· Human Awareness· Waste Segregation

1 Introduction

Promoting sustainable practices and mitigating the environmental impact of improper waste management requires individuals worldwide to be aware of waste sorting at source. Vietnam, however, still lacks awareness and action in this area. In [21], the survey results conducted on 1513 students at the Vietnam Maritime University showed that 95.85% of the students have an understanding of the harmful effects of single-use plastic, but only 57% of the students could identify groups of waste sorting. The student's concern and sense of the meaning of plastic codes and labels are still limited. Based on this survey results, most students are aware of the harms of plastic waste, but this awareness is still insufficient. Besides, according to a new study from the Ministry of Natural Resources and Environment [22], Ho Chi Minh City first drew out plans in 1999 to categorize waste. The largest city in Vietnam declared in 2018 that it intended to require all families to separate garbage properly by 2020 and to penalize those who did not. After 2020, all households, companies, and organizations in the city were to separate their garbage into three categories: recyclables, organic waste, and others. Still, not much has changed in this regard. Moreover, the majority of students (about 90%) have a basic awareness of the many forms of garbage and the significance of waste sorting at source, according to survey results from 500 Thai Nguyen University residents who live in dorms A, B, and K at Thai Nguyen University [17]. Nonetheless, as many as 82% of pupils still don't understand how to separate rubbish. This is because most students (71.4%) do not know how to separate garbage and the school has not yet controlled and structured waste sorting at source, the majority of students still gather all waste categories together and place them in the trash can in each dorm. Eventually, they have not yet understood their duty to preserve the shared environment, just 22.4% of students frequently take part in cleaning their living area. Therefore, it is essential to implement propaganda, dissemination, and education to raise students' awareness about waste sorting, reduction, reuse, recycling, and treatment.

In the contemporary age, the application of scientific knowledge and technological advancements, particularly artificial intelligence (AI), has become an essential component in numerous facets of human existence. This is a testament to how far we have come in our understanding and manipulation of the world around us. Artificial intelligence, a broad and complex field, encompasses several sub-disciplines, one of which is deep learning. Deep learning, in recent years, has experienced significant progress and has established itself as an influential instrument in a variety of applications. This is largely due to its unique ability to autonomously learn and decipher intricate patterns from extensive datasets, a feature that sets it apart from other AI techniques. As we continue to explore its potential, it is clear that deep learning will play a pivotal role in shaping the future of artificial intelligence and, by extension, our world.

In that scenario, the integration of cutting-edge technologies holds significant promise in promoting sustainable practices. This study introduces a revolutionary prototype called BEG (Bachkhoa Eco-friendly Guide) that utilizes computer vision and IoT to educate people about environmental issues and improve waste management. This study makes the following significant contributions.

- Develop the BEGNet, an efficient trash image classification model.
- Propose an IoT-based architecture for effectively implementing and deploying the trash bin system.
- Build a BKTrashImage dataset containing 6205 images of five categories: paper cups, aluminum cans, milk boxes, PET bottles, and foam boxes.

– Conduct experiments to demonstrate the proposed model's effectiveness compared to state-of-the-art trash classification approaches on an existing Trashnet dataset and our BKTrashImage dataset.

2 Related Works

2.1 Trash Image Classification on the Trashnet Dataset

Several studies have focused on developing robust Convolutional Neural Network (CNN) models for trash image classification using the Trashnet dataset [25], which includes six categories: glass, paper, cardboard, plastic, metal, and trash. The dataset was captured using various mobile devices, such as the Apple iPhone 7 Plus, Apple iPhone 5S, and Apple iPhone SE.

Aral et al. [6] explored transfer learning models based on popular CNN architectures, including Densenet121, DenseNet169, InceptionResnetV2, MobileNet, and Xception, for trash classification on the Trashnet dataset. They split the dataset into 70% for training, 13% for validation, and 17% for testing, using a batch size of 8 and an input size of 224×224. The Densenet121 model achieved the highest accuracy of 95%.

Ruiz et al. [19] investigated the effectiveness of various CNN models, such as VGG, Inception, and ResNet, for automatic trash categorization. They used 80% of the Trashnet dataset for training, 10% for validation, and the remaining 10% for testing. The Inception-ResNet combination achieved an average accuracy of 88.66% on the Trashnet dataset.

Vo et al. [23] proposed a robust Trash Classification model (DNN-TC) using Deep Neural Networks. They modified the original ResNext-101 model by adding two fully connected layers after the global average pooling layer. The authors used 60% of the Trashnet dataset for training, 20% for validation, and the remaining 20% for testing. DNN-TC achieved an accuracy of 94%.

Lam et al. [16] used computer vision technology to transform a conventional trash bin into an intelligent one. Their method, MobileNetV2, achieved 85.59% and 71.00% accuracy on Dataset 1 and Dataset 2, respectively.

In addition to the above-mentioned methods for trash classification, several well-known CNN models, such as RegNet [18] [8], ResNext [24], ImageNet [15], VGG [20], ResNet [9], and DenseNet [11], which were originally developed for image classification, can also be used as baseline models for trash classification.

2.2 Previous Studies on Smart Trash Cans

Existing research on smart waste management has emphasized the advantages of leveraging state-of-the-art technologies, including the Internet of Things (IoT) and Artificial Intelligence (AI), to enhance and modernize current waste management approaches.

Lam et al. [16] proposed an architectural design for a trash bin system that employs cameras to take pictures of refuse and a central processing unit that analyzes these images to identify the correct bin for disposal.

Ali et al.[5] presented a municipal solid waste management system that incorporates IoT-based smart waste bin monitoring. This system is designed to tackle the escalating waste production challenges in developing nations, which stem from rapid population growth and urbanization. Their system employs ultrasonic sensors (HC-SR04) and temperature-humidity sensors (HW-505) to detect fires in bins, a feature that could save lives and mitigate economic losses.

Kanade et al. [14] developed a smart garbage monitoring system that employs an ultrasonic sensor to measure waste levels. The collected data are then uploaded to the Cloud using the Firebase platform, enabling remote monitoring of the bins' fill status.

Jia et al. [13] proposed a smart trash can (STC) system that uses sensors, edge computing, and NB-IoT communication to enable intelligent monitoring and management of trash cans in cities. The system has smart trash cans with sensors to detect fill level, harmful substances, location, etc. Edge nodes preprocess data before sending it via NB-IoT to a central server. The server platform analyzes the data to optimize trash collection.

Huh et al. [12] proposed an Internet of Things (IoT) based smart trash bin model for efficient waste separation and management in smart cities. The system uses sensors, image processing, and spectroscopy techniques to automatically classify and separate trash into recyclable categories. It checks trash levels, compresses trash to increase capacity, and sends alerts when full.

2.3 Author's comments of related works

Significant advancements have been achieved in the field of trash image classification, where numerous studies have focused on creating robust Convolutional Neural Network (CNN) models specifically designed for the Trashnet dataset [25]. The high accuracy rates attained by these studies suggest a promising future for their integration into intelligent waste management systems. However, despite significant advancements in smart waste management research [5][14][16][13][12], current systems fall short in providing features that would improve user engagement with trash classification at the source, a critical component of effective waste management.

3 The proposed BEG and Its Implementation

3.1 The Proposed Trash Image Classification Model (BEGNet)

RegNetX and RegNetY, introduced in [18], offer a network design space aimed at decreasing computation and the number of epochs needed for training. This design space has the added benefit of being more easily interpretable. RegNetY specifically refers to a convolutional network design space that generates simple and regular models with parameters including depth (d), initial width ($w_0 > 0$), and slope ($w_a > 0$), and produces a unique block width for each block. The RegNet model is subject to a key constraint, which is a linear parameterization of block widths, as specified in Equation 1.

$$u_j = w_0 + w_a.j \tag{1}$$

In addition to RegNetX, a modification for RegNetY that involves incorporating Squeeze-and-Excitation blocks was also introduced in [10]. These blocks are a type of architectural unit designed to enhance the representational capabilities of a network by allowing for dynamic channel-wise feature recalibration. The structure of the Squeeze-and-Excitation block is illustrated in Fig. 1, and its operation can be described as follows:

- A convolutional block is taken as input.
- The input channels are "squeezed" into a single scalar value using average pooling.
- A dense layer, followed by a ReLU activation function, is applied to add non-linearity, and the output channel complexity is reduced by a certain ratio.
- Another dense layer, followed by a sigmoid activation function, produces a smooth gating function for each channel.
- Finally, the feature maps of the convolutional block are weighted based on the output of the side network, which is referred to as the "excitation".

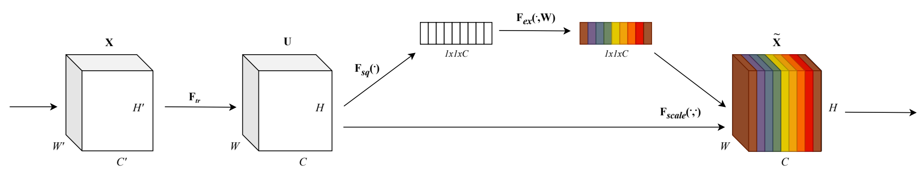


Fig. 1: A Squeeze-and-Excitation block.

To classify images, a robust framework for trash classification, namely BACHKHOA Eco-friendly Guide Network (BEGNet), which utilizes Deep Neural Networks, is introduced. The BEGNet employs the RegNetY120 pre-trained on the ImageNet dataset [7] as its backbone, followed by a Global Average Pooling layer, a Dropout layer, a Dense layer indicating the number of classes, and finally a customized activation function that is inspired by the Sigmoid activation function and defined by Eq. 2.

$$Customized\ Activation\ Function = \frac{1}{1 + e^{-\frac{1}{5}x}} \tag{2}$$

The selection of the Dropout Rate and Customized Activation Function configurations is based on empirical evaluation, as discussed in Section 4. Fig. 2 provides further clarification on our proposed method, where N_class represents the number of trash labels.

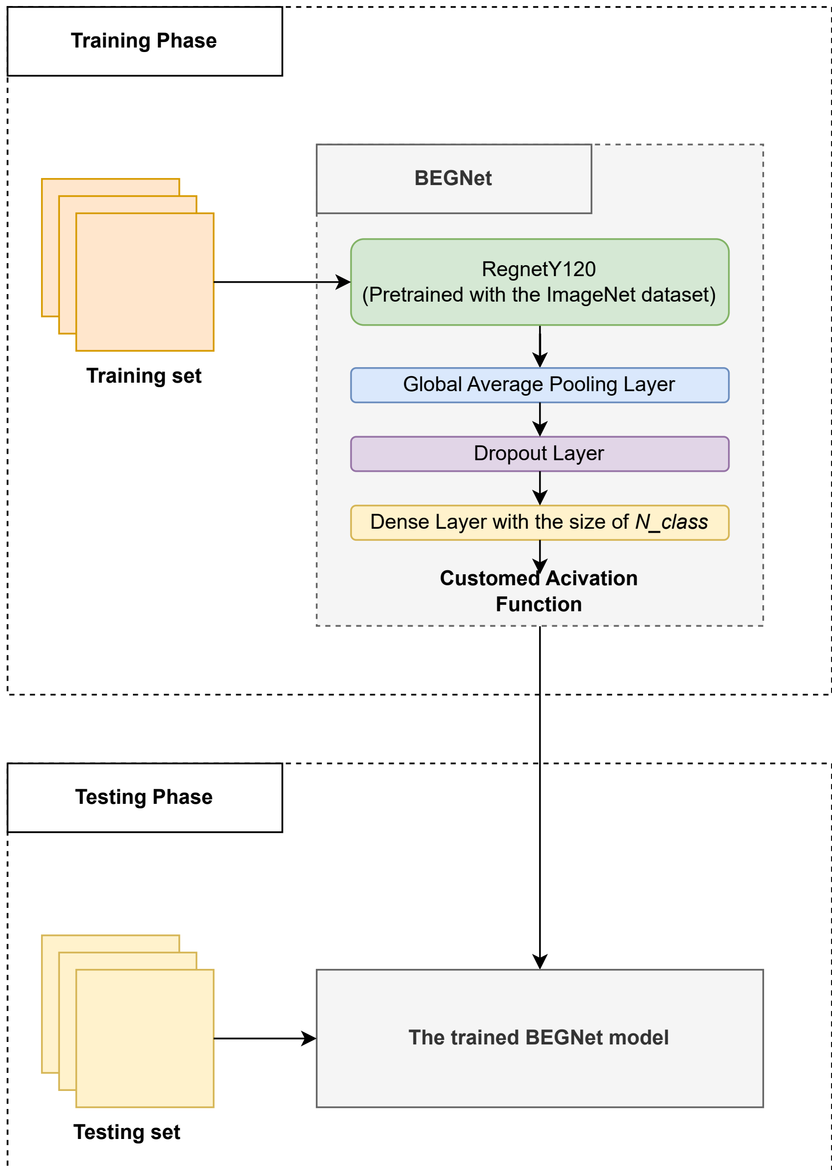


Fig. 2: BEG's approach involves using the RegNetY120 backbone, which was pre-trained with the ImageNet dataset. This is followed by a Global Average Pooling layer, a Dropout layer, and a Sigmoid activation function.

3.2 The Proposed IoT-based Architecture for Prototyping BEG Device

Our proposed BEG system is architecturally categorized into three principal components, each designed to fulfil specific roles within the waste management ecosystem:

1. **Server-side component:** This includes both a server and a database service functioning as the communication backbone between two user groups: individuals disposing of waste (referred to as users) and personnel responsible for the operational management of the waste bins system (referred to as administrators). Leveraging the robust infrastructure provided by Amazon Web Services (AWS) [4], we incorporate several AWS solutions. The server is powered by Amazon Elastic Compute Cloud (Amazon EC2) [2], specifically the `p3.2xlarge` instance type, chosen for its high-performance computing capabilities essential for machine learning tasks. Additionally, we utilize FastAPI [3], a contemporary web framework designed for crafting RESTful APIs in Python, to process incoming requests efficiently. For data persistence, Amazon DynamoDB [1] is employed, a serverless NoSQL database offering support for both key-value and document data structures.
2. **Admin-side component:** The administrative interface is a web application crafted using Bootstrap to ensure a responsive and intuitive user experience across various devices. Administrators gain access through a secure login mechanism, enabling them to oversee the system's status effectively. The dashboard provides critical insights, including real-time metrics of waste volume per bin, imagery of disposed items, and temporal and locational data of disposal events.
3. **User-side component:** Situated at disposal locations, the user-side setup encompasses intelligent trash bins segregated into three categories—Recyclable, Non-recyclable, and Organic waste. Each bin is equipped with a suite of hardware: a Jetson Nano embedded computing board, an IMX219 8MP camera module paired with the Jetson Nano, a 10.1-inch capacitive touch screen LCD with a resolution of 1024×600, and a quartet of HC-SR04 ultrasonic sensors. The sensors are distributed with one for each bin to measure the waste level through distance sensing inside the bin, and an additional sensor is coupled with the camera module to detect trash presence, thereby optimizing the image capture process for the AI models.

The product workflow consists of several key steps. First, the IR sensor detects the presence of trash in front of the camera, triggering the camera module to capture an image and temporarily store it in the Jetson Nano's storage. The Jetson Nano then sends a request to the Amazon EC2 server, where the BEGNet model is hosted, to analyze the image and obtain the classification result. Upon determining the trash type, the screen will display guidance for proper disposal. Additionally, an IR sensor is installed in each trash bin to measure the volume of waste present. Once a bin reaches its capacity, a signal is transmitted to the Amazon EC2 server and this server will then store this information in Amazon Dynamondb, while at the same time updating the web application to notify the administrators immediately. Fig. 3 provides a visual representation of the entire architecture.

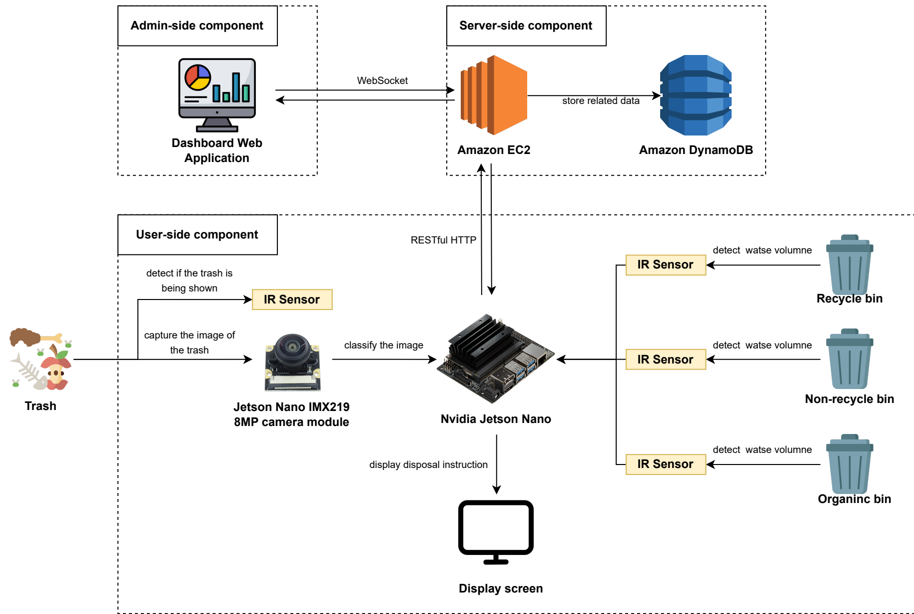


Fig. 3: The deployment of BEG system.

4 Evaluation

4.1 Datasets

The Trashnet dataset is a collection of 2,527 images, each taken with a mobile device and sorted into one of six categories: glass, paper, cardboard, plastic, metal, and trash. The images feature the respective items set against a uniform white background, with lighting provided by natural sunlight, artificial room light, or a combination of both. The distribution of images across the different categories is summarized in Table 1. Representative images from each category are presented in Figure 4.

Table 1: The statistic of the Trashnet dataset.

No	Classes	Number of images
1	Glass	501
2	Paper	594
3	Cardboard	403
4	Plastic	482
5	Metal	410
6	Trash	137

To address the challenge of enhancing students' understanding and awareness of source-separated waste classification, this study aims to create an extensive dataset that encompasses categories of refuse typically discarded by students. This endeavor involved augmenting the existing Trashnet dataset [1] with supplementary images obtained from the internet, as well as photographs manually taken of various trash items. The newly formed dataset, designated as the BK-TrashImage dataset, is composed of six distinct categories: Paper Cups, Aluminum Cans, Milk Boxes, PET Bottles, and Foam Boxes, amounting to a total of 6,205 images. Efforts were made to ensure a relatively even distribution of images across these categories. The class-wise breakdown of the BKTrashImage dataset is detailed in Table 2.

Fig. 4: Samples of Trashnet dataset.

To have more insight into how BKTrashImage looks, Fig. 5 shows some samples of each class of the dataset.

4.2 Experimental Setting

In this section, we conduct a quantitative evaluation of the proposed image classification model, BEGNet, alongside other existing methodologies. Our assessment utilizes two datasets previously described: Trashnet and BKTrashImage.

Table 2: The statistic of the BKTrashImage dataset.

No	Classes	Number of images
1	Paper Cup	1556
2	Aluminum Can	1531
3	Milk box	1624
4	PET bottle	1499
5	Foam box	1395

Fig. 5: Samples of BKTrashImage dataset.

For both datasets, we allocate 80% of the data for training the models and reserve the remaining 20% for testing their performance.

The experimental methodologies were deployed in Python 3.10.11 and executed on the TensorFlow framework, an open-source library tailored for deep learning applications in Python. These experiments were carried out on a system running Ubuntu 20.04.5 LTS, equipped with an NVIDIA A100-SXM4-40GB GPU and 83.48 GB of RAM to facilitate the computational processes.

In this study, we apply advanced classification techniques for sorting trash, as detailed in Section 2. These techniques include Densenet121_Aral[6], Inception-ResNet_Ruiz[19], MobileNetV2_Lam[16], and DNN-TC_Vo[23], which serve as benchmarks to evaluate the efficacy of our proposed trash image classification method. Before training, we standardize the images by resizing them to 224x224 pixels and normalizing the pixel values, scaling them by dividing each pixel value by 255 for computational efficiency.

For optimization, we employ Stochastic Gradient Descent (SGD) across all methods, with a learning rate α of 0.0001. Our training process is conducted with a batch size of 128, spanning 20 epochs. Model performance is assessed after every epoch using the test set. In the testing phase, we measure and compare the classification accuracy of the various methods on the test set to determine their relative performance.

4.3 Experimental Results

BEGNet Model Hyperparameter Selection To begin with, we present experimental results regarding the selection of the Dropout Rate and activation function for the BEGNet model. The performance of BEGNet, in terms of accuracy on the Trashnet dataset, was evaluated based on the joint settings of these two variables. We tested four discrete values of Dropout Rate: 0.0, 0.2, 0.5, and 0.8. Additionally, we examined four activation functions: Softmax, Sigmoid, Hard Sigmoid, and a self-customized activation function. The initial experiment indicated that BEGNet achieved optimal performance with a 0.2 Dropout Rate and the self-customized activation function. Table 3 presents the accuracy of BEGNet for each joint setting.

Table 3: The accuracy of BEGNet on each joint setting of Dropout Rate values and activation functions.

	0.0	**0.2**	**0.5**	**0.8**
Sigmoid	94.26	94.46	93.86	93.86
Softmax	93.47	94.06	94.06	93.47
Hard Sigmoid	89.31	90.50	88.51	91.41
Customized Activation Function	94.65	**95.45**	94.85	93.86

BEGNet Model in comparison with other methods The accuracy of the proposed BEGNet model, as well as that of other experimental methods, when evaluated on the Trashnet dataset, is displayed in Table 4. Our proposed approach outperforms the competing methods, with an accuracy of 95.45%. In comparison, Densenet121_Aral, Inception-ResNet_Ruiz, DNN-TC_Vo, and MobileNetV2_Lam obtained accuracies of 93.66%, 93.27%, 93.07%, and 88.12%, respectively.

To further demonstrate the effectiveness of our proposed approach, we evaluate the experimental methods on the BKTrashImage dataset, which includes types of trash commonly discarded by students. As shown in Table 5, our method achieved the highest accuracy of 98.09% compared to all other experimental methods. The next best performing models were Inception-ResNet_Ruiz (96.45%) and Densenet121_Aral (96.32%). Although MobileNetV2_Lam and DNN-TC_Vo achieved high accuracies of 95.73% and 95.79%, respectively, they still fell short by over 1% compared to our proposed model.

Table 4: The accuracy of the experimental methods on the Trashnet dataset.

No	Methods	Accuracy on the testing set (%)
1	Densenet121_Aral [6]	93.66
2	Inception-ResNet_Ruiz [19]	93.27
3	MobileNetV2_Lam [16]	88.12
4	DNN-TC_Vo [23]	93.07
5	BEGNet	**95.45**

Table 5: The accuracy of the experimental methods on the BKTrashImage dataset.

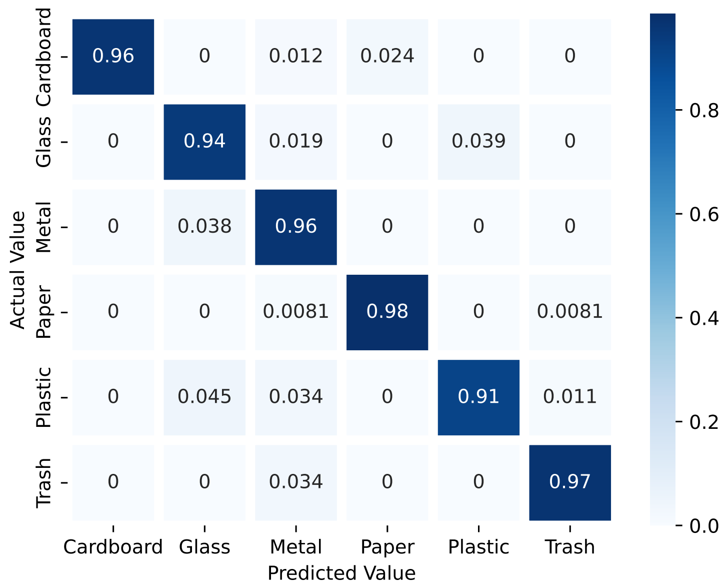
No	Methods	Accuracy on the testing set (%)
1	Densenet121_Aral [6]	96.32
2	Inception-ResNet_Ruiz [19]	96.45
3	MobileNetV2_Lam [16]	95.73
4	DNN-TC_Vo [23]	95.79
5	BEGNet	**98.09**

Further evaluation of BEGNet Model on the two datasets To provide a comprehensive evaluation of our BEGNet Model's classification accuracy, we present confusion matrices for two datasets in Figure 6. These matrices have been normalized by the total number of instances per class to facilitate a direct comparison of class-specific Recall, indicated by the diagonal elements. Additionally, for a more detailed overview, we compile Precision, Recall, and F1 Score for each class in tabular form. The performance metrics for the Trashnet dataset can be found in Table 6, while those for the BKTrashImage dataset are displayed in Table 7. This format ensures that the model's predictive performance is both transparent and easily interpreted.

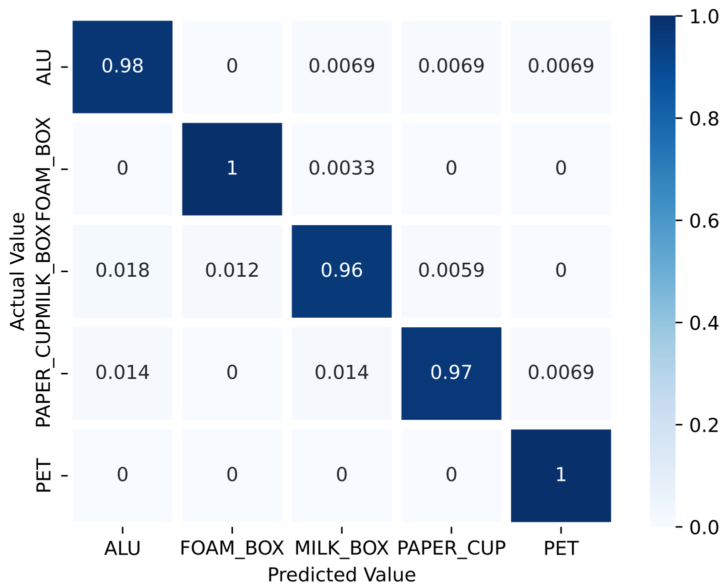
	Precision	Recall	F1 Score	Support
Cardboard	1.000	0.9639	0.9816	83
Glass	0.9327	0.9417	0.9372	103
Metal	0.9036	0.9615	0.9317	78
Paper	0.9839	0.9839	0.9839	124
Plastic	0.9524	0.9091	0.9302	88
Trash	0.9333	0.9655	0.9492	29

Table 6: The evaluation of BEGNet in terms of Precision, Recall, F1 Score and Support of each class on the Trashnet dataset.

On the Trashnet dataset, BEGNet exhibits robust classification performance across the board. Notably, the Plastic class shows commendable results with Precision, Recall, and F1 Score values of approximately 0.95, 0.91, and 0.93,



(a) Confusion matrix of Begnet on the Trashnet dataset.



(b) Confusion matrix of Begnet on the BKTrashImage dataset.

Fig. 6: Confusion matrix of Begnet on the two datasets.

respectively. Despite these strong metrics, the performance for Plastic is slightly lower compared to other classes. This can be attributed, in part, to the physical resemblance between Plastic and Glass samples, such as their shared transparent properties. The confusion matrix reveals a higher rate of misclassification between these two classes when contrasted with others.

	Pricision	**Recall**	**F1 Score**	**Support**
Aluminum Can	0.966	0.9793	0.9726	290
Foam Box	0.9868	0.9967	0.9917	299
Milk Box	0.9789	0.9644	0.9716	337
Paper Cup	0.9859	0.9655	0.9756	290
PET Bottle	0.9871	1.0000	0.9935	305

Table 7: The evaluation of BEGNet in terms of Precision, Recall, F1 Score and Support of each class on the BKTrashImage dataset.

Turning to the BKTrashImage dataset, BEGNet's results are even more impressive, with most classes surpassing an F1 Score of 0.97. Outstandingly, Foam Box and PET Bottle classes both exceed an F1 Score of 99%. Other classes, including Paper Cup, Aluminum Can, and Milk Box, also showcase high F1 Scores of 0.9756, 0.9726, and 0.9716, respectively. Furthermore, the BKTrashImage dataset is characterized by a more balanced sample distribution compared to the Trashnet dataset, enhancing the reliability of these results.

4.4 Discussion

In our study, the proposed BEGNet model was benchmarked against leading methods using the Trashnet and BKTrashImage datasets. The former is a more compact dataset predominantly comprising images of individual objects, whereas the latter includes a variety of trash items commonly discarded by Vietnamese students and is similarly structured. Our empirical analysis of BEGNet yielded impressive outcomes, with an accuracy of 95.45% on the Trashnet dataset and 98.09% on the BKTrashImage dataset, surpassing other state-of-the-art techniques.

The performance of BEGNet was rigorously assessed using confusion matrices along with Precision, Recall, and F1 Score metrics. Although the model's ability to differentiate the Plastic category from other categories such as Glass on Trashnet was modestly challenged, it showcased exceptional precision in identifying classes like Foam Box and PET Bottle on the BKTrashImage dataset. In addition, the balanced nature of the BKTrashImage dataset further reinforces the dependability and generalizability of BEGNet's classification efficacy.

5 Conclusion

In this study, we introduce an Artificial Intelligence of Things (AIoT) device that incorporates AI-powered Computer Vision for the classification of waste materials. The cornerstone of this research is the creation of the BEGNet model, which is an enhanced version of the RegNetY120 architecture, specifically optimized for waste classification tasks. BEGNet demonstrates a marked improvement in accuracy over existing state-of-the-art models on both the Trashnet and our newly developed BKTrashImage datasets. Additionally, we have developed an IoT framework to facilitate the real-world application and deployment of our smart trash bin system.

Our proposed system diverges from typical smart waste management solutions by not only classifying waste but also by educating users about proper disposal techniques, thereby promoting source-separated recycling. The BACHKHOA Eco-friendly Guide (BEG) aims to yield extensive benefits across environmental, societal, and economic fronts. By enhancing source classification knowledge among students, the system aims to instigate a shift in waste disposal behaviors, empowering students to become pivotal agents of change in the community's waste management practices.

In summary, the BEG initiative is taking proactive steps to overcome the challenges in waste segregation by cultivating environmental responsibility among the student body. Nevertheless, it necessitates ongoing enhancements to realize its full potential. Future work includes broadening the dataset to capture a more diverse array of waste types, refining the Deep Neural Network to allow simultaneous detection of multiple waste objects, and improving the web-based application to foster better engagement with the entire waste management process. These improvements are vital for the progression of this environmentally impactful project.

Acknowledgement

This research is funded by the Office for International Study Programs (OISP) under grant number SVOISP-2023-KH&KTMT-46. We acknowledge Ho Chi Minh City University of Technology (HCMUT), VNU-HCM for supporting this study.

References

1. Amazon DynamoDB — aws.amazon.com. https://aws.amazon.com/pm/dynamodb/, [Accessed 03-01-2024]
2. Amazon EC2 — aws.amazon.com. https://aws.amazon.com/pm/ec2/, [Accessed 03-01-2024]
3. FastAPI — fastapi.tiangolo.com. https://fastapi.tiangolo.com/, [Accessed 03-01-2024]

4. Overview of Amazon Web Services - Overview of Amazon Web Services — docs.aws.amazon.com. https://docs.aws.amazon.com/whitepapers/latest/aws-overview/introduction.html, [Accessed 03-01-2024]
5. Ali, T., Irfan, M., Alwadie, A.S., Glowacz, A.: Iot-based smart waste bin monitoring and municipal solid waste management system for smart cities. Arabian Journal for Science and Engineering **45**, 10185–10198 (2020)
6. Aral, R.A., Keskin, Ş.R., Kaya, M., Hacıömeroğlu, M.: Classification of trashnet dataset based on deep learning models. In: 2018 IEEE International Conference on Big Data (Big Data). pp. 2058–2062. IEEE (2018)
7. Deng, J., Dong, W., Socher, R., Li, L.J., Li, K., Fei-Fei, L.: Imagenet: A large-scale hierarchical image database. In: 2009 IEEE conference on computer vision and pattern recognition. pp. 248–255. Ieee (2009)
8. Dollár, P., Singh, M., Girshick, R.: Fast and accurate model scaling. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. pp. 924–932 (2021)
9. He, K., Zhang, X., Ren, S., Sun, J.: Deep residual learning for image recognition. In: Proceedings of the IEEE conference on computer vision and pattern recognition. pp. 770–778 (2016)
10. Hu, J., Shen, L., Sun, G.: Squeeze-and-excitation networks. In: Proceedings of the IEEE conference on computer vision and pattern recognition. pp. 7132–7141 (2018)
11. Huang, G., Liu, Z., Van Der Maaten, L., Weinberger, K.Q.: Densely connected convolutional networks. In: Proceedings of the IEEE conference on computer vision and pattern recognition. pp. 4700–4708 (2017)
12. Huh, J.H., Choi, J.H., Seo, K.: Smart trash bin model design and future for smart city. Applied sciences **11**(11), 4810 (2021)
13. Jia, G., Zhu, Y., Han, G., Chan, S., Shu, L.: Stc: an intelligent trash can system based on both nb-iot and edge computing for smart cities. Enterprise Information Systems **14**(9-10), 1422–1438 (2020)
14. Kanade, P., Alva, P., Prasad, J.P., Kanade, S.: Smart garbage monitoring system using internet of things (iot). In: 2021 5th International Conference on Computing Methodologies and Communication (ICCMC). pp. 330–335. IEEE (2021)
15. Krizhevsky, A., Sutskever, I., Hinton, G.E.: Imagenet classification with deep convolutional neural networks. Communications of the ACM **60**(6), 84–90 (2017)
16. Lam, K.N., Huynh, N.H., Ngoc, N.B., Nhu, T.T.H., Thao, N.T., Hao, P.H., Kiet, V.V., Huynh, B.X., Kalita, J.: Using artificial intelligence and iot for constructing a smart trash bin. In: Future Data and Security Engineering. Big Data, Security and Privacy, Smart City and Industry 4.0 Applications: 8th International Conference, FDSE 2021, Virtual Event, November 24–26, 2021, Proceedings 8. pp. 427–435. Springer (2021)
17. Phan Thi Thu Hang, Hoang Thi Thanh Hien, N.T.T.: Current status of awareness and behaviour of Thai Nguyen University students about waste and waste classification. http://tailieudientu.lrc.tnu.edu.vn/Upload/Collection/brief/brief_41504_45275_9520149225537.pdf, [Accessed 17-01-2024]
18. Radosavovic, I., Kosaraju, R.P., Girshick, R., He, K., Dollár, P.: Designing network design spaces. In: Proceedings of the IEEE/CVF conference on computer vision and pattern recognition. pp. 10428–10436 (2020)
19. Ruiz, V., Sánchez, Á., Vélez, J.F., Raducanu, B.: Automatic image-based waste classification. In: From Bioinspired Systems and Biomedical Applications to Machine Learning: 8th International Work-Conference on the Interplay Between Nat-

ural and Artificial Computation, IWINAC 2019, Almería, Spain, June 3–7, 2019, Proceedings, Part II 8. pp. 422–431. Springer (2019)
20. Simonyan, K., Zisserman, A.: Very deep convolutional networks for large-scale image recognition. arXiv preprint arXiv:1409.1556 (2014)
21. Thi, D.P., Thuy, H.D.T.: Survey and assessment of the students' awareness and behaviors about waste sorting and single use plastic consumption habits, https://jmst.vimaru.edu.vn/index.php/tckhcnhh/article/view/21
22. VnExpress: Vietnam's waste sorting plan goes unsupported - VnExpress International — e.vnexpress.net. e.vnexpress.net/news/news/environment/vietnam-s-waste-sorting-plan-goes-unsupported-4675376.html, [Accessed 17-01-2024]
23. Vo, A.H., Vo, M.T., Le, T., et al.: A novel framework for trash classification using deep transfer learning. IEEE Access **7**, 178631–178639 (2019)
24. Xie, S., Girshick, R., Dollár, P., Tu, Z., He, K.: Aggregated residual transformations for deep neural networks. In: Proceedings of the IEEE conference on computer vision and pattern recognition. pp. 1492–1500 (2017)
25. Yang, M., Thung, G.: Classification of trash for recyclability status. CS229 project report **2016**(1), 3 (2016)