# Accuracy study of image classification for reverse vending machine waste segregation using convolutional neural network

**Tan Hor Yan[1,2], Sazuan Nazrah Mohd Azam[1,2], Zamani Md. Sani[1], Azizul Azizan[3]**

[1]Faculty of Technology and Electrical Engineering, Universiti Teknikal Malaysia Melaka, Melaka, Malaysia
[2]Center for Robotics and Industrial Automation, Universiti Teknikal Malaysia Melaka, Melaka, Malaysia
[3]Advanced Informatics Department, Razak Faculty of Technology and Informatics, Universiti Teknologi Malaysia, Kuala Lumpur, Malaysia

## Article Info

## ABSTRACT

This study aims to create a sorting system with high accuracy that can classify various beverage containers based on types and separate them accordingly. This reverse vending machine (RVM) provides an image classification method and allows for recycling three types of beverage containers: drink carton boxes, polyethylene terephthalate (PET) bottles, and aluminium cans. The image classification method used in this project is transfer learning with convolutional neural networks (CNN). AlexNet, GoogLeNet, DenseNet201, InceptionResNetV2, InceptionV3, MobileNetV2, XceptionNet, ShuffleNet, ResNet 18, ResNet 50, and ResNet 101 are the neural networks that used in this project. This project will compare the F1-score and computational time among the eleven networks. The F1-score and computational time of image classification differs for each neural network. In this project, the AlexNet network gave the best F1-score, 97.50% with the shortest computational time, 2229.235 s among the eleven neural networks.

*Corresponding Author:*

Sazuan Nazrah Mohd Azam
Faculty of Electrical Engineering, Universiti Teknikal Malaysia Melaka
Hang Tuah Jaya, 76100 Durian Tunggal, Melaka, Malaysia
Email: sazuan@utem.edu.my

## 1. INTRODUCTION

In addressing main topic surrounding waste disposal, the suggested countermeasures to encourage greater public participation in the recycling efforts is the recommended introduction of the reverse vending machine (RVM). RVM is simplicity of usage, conveniently located with easy access by the public, and it provides a mechanism for rewarding recyclers for the selected item recycled. The proposed RVM allows the recycling of three beverage containers: polyethylene terephthalate (PET) bottles, aluminium cans, and drink carton boxes. AlexNet, GoogLeNet, DenseNet201, InceptionResNetV2, InceptionV3, MobileNetV2, XceptionNet, ShuffleNet, ResNet 18, ResNet 50, and ResNet 101 are the neural networks that will be used in this project. The performances evaluated among the convolutional neural networks (CNNs) are F1-score and time taken to perform the classification.

Classification technology is the most important technology in the RVM. In previous work, some RVM used hardware classification methods, such as barcode detectors and sensors. The problem with barcode detector is when it is damage and waste is unidentified thus failed to classify the waste. Whereas sensors can only detect certain objects and are limited to the type of sensors used. Sensors and barcode detectors will make the RVM expensive [1]. This paper proposed an image classification method that is a more effective way to perform classification. The image classification method used in this project is transfer

learning with the CNN. CNN-based approached have demonstrated remarkable classification performance in waste categorization tasks [2].

The image classification technique is proficient in handling identification and recognition of particular target objects (such as facial features, handwritten characters, or products), categorizing and labeling images, and evaluating the subjective quality of images [3]. The architecture of a CNN bears resemblance to the interconnected structure of neurons in the human brain, and this association draws inspiration from the functioning of the visual cortex [4]. CNN is analyzed into two groups which are essential layers and secondary layers. The essential layers subsist of convolution, active, pooling, flatten and dense layers [5]. The secondary layers are additional components that can be incorporated into CNNs to increase their resilience against overfitting and improve their ability to generalize. Secondary layers include dropout layers, batch normalization layers, and regularization layers [5]. Figure 1 shows the network architecture of CNN [4].
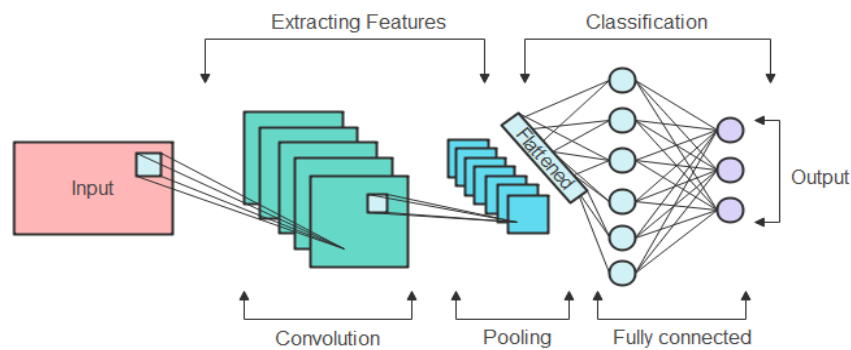


Figure 1. CNN network architecture [4]

Transfer learning is a methodology that focuses on improving the performance of target models on specific domains by leveraging the knowledge acquired from different yet related source domains [6]. Transfer learning offers the advantage of saving resources and time. The key aspect of transfer learning is generalization, which entails transferring only the knowledge that can be effectively applied by new models across diverse scenarios or conditions [7]. Instead of models being strictly tied to a specific training dataset, transfer learning allows for more generalized models. These models can be applied in diverse conditions and with different datasets. Various hardware techniques enable recognition and sorting in RVM machines, rendering fraudulent attempts virtually impossible. In contrast, the traditional waste classification scheme suffers from low efficiency and accuracy [8]. Therefore, a new waste management system is necessary to handle waste materials effectively. Image classification, a computer vision approach, plays a significant role in organizing visual content based on images. Its primary objective is to assign pixels in a digital image to specific classes [9]. Several image classification models have been developed to improve recognition efficiency [10]. One proposed approach focuses on classifying plastic waste based on its resin identification code, facilitating efficient recycling [11]. The classification technique demonstrates high efficiency in discriminating plastic waste, as shown by the WaDaBa database, which contains images of plastic waste.

Furthermore, CNN-based waste-type classifiers have been introduced to classify waste types within municipal solid waste figures, encompassing 9,200 trash figures, utilizing image classification techniques [12]. These classifiers can directly identify various waste types or derive them from waste-item classes, with the highest waste-type classification efficiency reaching 94.86% using the ResNet50 classifier [12]. At present, the combination of transfer learning and CNN has been extensively applied in image recognition. Transfer learning improved accuracy and enhance the robustness of the model [13]. Transfer learning, a technique discusses in reference [6], aims to boost the performance of target models in specific domains by harnessing knowledge from different but related source domains. It has been effectively applied in a range of machine learning tasks, including text sentiment segregation [14], image segregation [15], human activity segregation [16], software defect segregation [17], and multi-language text segregation [18]. Notably, CNN has gained significant popularity in segregation jobs, whether employed as a standalone model or integrated into conventional classifier ensemble CNN architecture [19]. In study [20], a combination of GoogLeNet with SVM displays an encouraging result of 97.86 % based on the TrashNet dataset. A MobileNet variant [21] shows a performance of 96.57 %. Additionally, the optimized DenseNet121 network achieved an impressive 99.6 % accuracy while applying the similar dataset [22].

The outline of this paper is as follows. The related study of image classification, transfer learning, and CNN are presented in section 1. Next, the process details for this project are shown in section 2, followed by the results and discussion of the simulations in section 3. Lastly, this paper is concluded in the final section 4.

## 2. METHOD

This project will be divided into two major steps: preparing the dataset, and image classification. This part discussed the details of the proposed image classification technique using transfer learning with CNN networks. Figure 2 shows the overall flowchart for this project. The overall concept for this project is when the camera captures the beverage containers' image. The captured image will move to the system to recognize the image information. Then, the system will be based on the image information and therefore segregate the object based on the material. The step to perform image processing and classification are the same for all neural networks. The subsequent section will provide an in-depth exploration of these steps, elucidating their intricacies and implications.
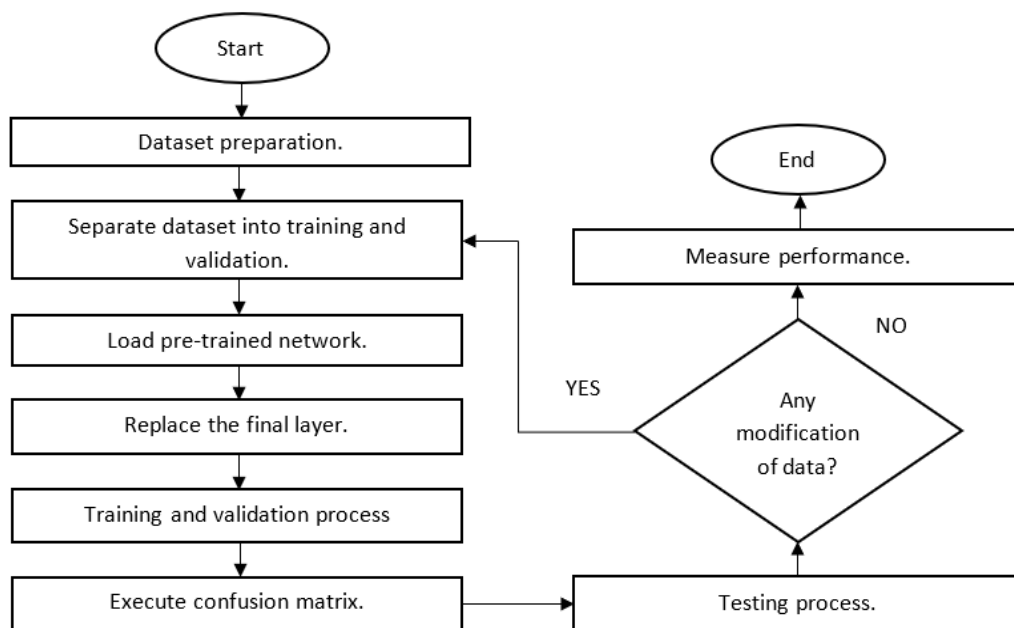
Figure 1. Overall flowchart

### 2.1. Datasets preparation

Before implementing the image classification technique, the critical initial step involves constructing the dataset for the project. The dataset will contain information on three types of beverage containers: drink carton boxes, PET bottles, and aluminium cans. Data will include details on types, sizes, colors, shapes, and images captured from various angles. The collected data will be organized and stored in separate files based on their corresponding materials. All neural networks used in the project will utilize the same dataset to ensure consistency across models. The training dataset is used to fit the model's parameters, while the validation dataset provides an unbiased evaluation and helps tune hyperparameters [23]. Partitioning the datasets prevents overfitting, where the model memorizes patterns in the training data and struggles with unseen data [24].

### 2.2. Image classification algorithms

Transfer learning involves taking pertinent segments of a pre-trained machine learning model and applying them to a fresh, albeit similar, problem. Typically, this entails retaining the model's essential functionality while integrating new elements tailored to address a specific task. For image classification, the process begins by incorporating image datasets and scaling the input data to match the pre-trained model's specified input scale. Then, a CNN model is constructed based on the scaled input data. Parameter transfer is

employed to initialize the model's parameters using the pre-trained CNN image classification model. The final step involves testing the CNN model's performance. Transfer learning in CNN involves retaining the weights and biases of the original model, especially when retraining from the altered version. The primary objective of transfer learning is to leverage extensive datasets for pre-training the model. Subsequently, these trained parameters are fine-tuned using a smaller learning sample. This results in the final trained model benefiting from both swiftness and substantial robustness [9].

Transfer learning is often explained in the context of domains and tasks. A domain, referred to as $G$, composed a feature space denoted as $X$, and a marginal probability distribution, $Z(X)$, where $X = \{x_1, \ldots, x_n\} \in X$. Within a specific domain, $G = \{X, Z(X)\}$, a task consists of two fundamental components: a label space $Y$, and an objective predictive function $f: X \rightarrow Y$. This function $f$ is utilized to make predictions, $f(x)$ of a new instance $x$. The task is represented as $V = \{Y, f(x)\}$, and it is learned from training data that includes pairs of samples $\{x_i, y_i\}$, where $x_i \in X$ and $y_i \in Y$. Given a source domain $G_s$ and learning task $V_s$, and a target domain $G_T$ and learning task $V_T$, transfer learning aims to enhance the learning process of the target predictive function $f(\cdot)$ in a target domain, $G_T$, by leveraging the knowledge acquired from a source domain, $G_s$, and it is corresponding learning task, $V_s$. It should be noted that the source domain, $G_s$, is distinct from the target domain, $G_T$, and the source learning task, $V_s$, is different from the target learning, $V_T$.

Convolutional layers function as feature extractors, responsible for learning feature representations from input images. These layers consist of a collection of filters, often referred to as kernels, applied to the input data before processing [25]. Each kernel has its unique width, height, and weights, enabling it to extract specific features from the input data. The convolutional layers are given as in (1).

$$Y_k = f(A_k \times z) \tag{1}$$

where the input data is denoted as $z$; the convolutional filter associated with the $k^{th}$ feature map is symbolized as $A_k$; the multiplication operation in this context represents the 2D convolutional operator, which computes the inner product of the filter pattern at each position within the input image; and $f(\cdot)$ represents the non-linear activation function.

Pooling layers have the role of lowering the spatial resolution of feature maps, which helps achieve spatial invariance to input distortions and translations. This not only expands the receptive field of convolutional kernels across layers but also lowers computational complexity and memory demands by downsizing the feature maps [26]. This downsampling operation lowers the resolution of the feature maps while preserving essential features needed for the processing of subsequent layers. Pooling layer is given as in (2).

$$Y_{kij} = \max_{(p,q) \in \Re ij} z_{kpq} \tag{2}$$

where the result of the pooling operation for the $k^{th}$ feature map is denoted by $Y_{ki\,j}$, $z_{kpq}$ denotes the element locates at $(p,q)$ within the pooling region $\Re ij$, which excompasses a receptive field centered at the position $(i, j)$. To capture more abstract feature representations, network architectures often incorporate multiple layers of convolution and pooling. After these layers, fully connected layers are employed to interpret these feature representations and enable high-level reasoning.

## 2.3. Model evaluation

The validation dataset is frequently employed alongside a confusion matrix to appraise the effectiveness of a neural network model. It aids in the assessment of metrics like accuracy, precision, recall, and other evaluation criteria. True positive, TP indicates correct positive class predictions, while true negative, TN signifies accurate negative class predictions. False positive, FP refers to incorrect positive class predictions, and false negative, FN represents inaccurate negative class predictions. Figure 3 shows the confusion matrix.

Accuracy and F1-score are commonly used to measure the model's performance, but both will be based on what types of datasets will be used. F1-score is a better evaluation metric when the dataset is imbalanced compared with accuracy [27]. The F1-score takes precision and recall into account and is less affected by class imbalance than accuracy [28]. The accuracy, precision, recall and F1-score are representing in (3)-(6) respectively.

$$Accuracy = \frac{True\ Positive + True\ Negative}{True\ Positive + False\ Positive + True\ Negative + False\ Negative} \tag{3}$$

$$Precision = \frac{True\ Positive}{True\ Positive + False\ Positive} \tag{4}$$

$$Recall = \frac{True\ Positive}{False\ Negative + True\ Positive} \qquad (5)$$

$$F1 - score = \frac{2Precision\ \times Recall}{Precision + Recall} \qquad (6)$$



Figure 3. Confusion matrix

## 3. RESULTS AND DISCUSSION

This section will provide a comprehensive demonstration of the results obtained from the dataset preparation process as well as the image classification algorithms utilized, employing transfer learning with CNNs. The outcome will showcase the effectiveness of the dataset preparation methods in collecting and organizing the necessary data for training and validation. Additionally, the performance of the image classification algorithms, which leverage transfer learning and CNNs, will be evaluated and discussed in detail.

### 3.1. Dataset preparation

Datasets for each category comprise images of various sizes, colors, and shapes, captured from multiple angles, resulting in a total of 500 image datasets. These datasets are then divided into separate sets for validation, training, and testing to prevent overfitting [23]. The datasets will separate into training and validation datasets with the ratio of 70:30. A training partition size ranging from 40% to 80% will provide an optimal balance between accuracy and precision [29]. Figure 4 shows that example of image datasets.



Figure 4. Example of image datasets

### 3.2. Image classification algorithms

The CNN networks are trained and evaluated using MATLAB R2020b. To adapt these CNN networks for new image classification tasks, the final 2-D convolutional layer and the last classification layer of the network are substituted. Table 1 shows the feature vector and feature layer for AlexNet, DenseNet201, GoogLeNet, InceptionResnetV2, InceptionV3, MobileNetV2, XceptionNet, ShuffleNet, ResNet 18, ResNet 50, and ResNet 101.

### 3.3. Model evaluation

In the confusion matrix, the blue color represents the classification effectiveness. The horizontal axis corresponds to the predicted values whereas the vertical axis represents the actual values. Figure 5 shows the confusion matrix for the AlexNet network. Table 2 summarize the performance of the AlexNet is shown in Table 2.

According to Table 2, AlexNet achieved 98.67% accuracy and 1.33% error, with precision, recall, and F1-score all at 0.975%. The processing time for AlexNet was 2220.235 s. The PET bottle datasets in the test set were correctly identified, but two datasets were inaccurately singled out. One dataset from the aluminium can was mistakenly classified under drink carton box, and one dataset from the drink carton box was inaccurately classified under aluminium can classes.

Table 1. Feature vector and feature layer for AlexNet, DenseNet201, GoogLeNet, InceptionResnetV2, InceptionV3, MobileNetV2, XceptionNet, ShuffleNet, ResNet 18, ResNet 50, and ResNet 101

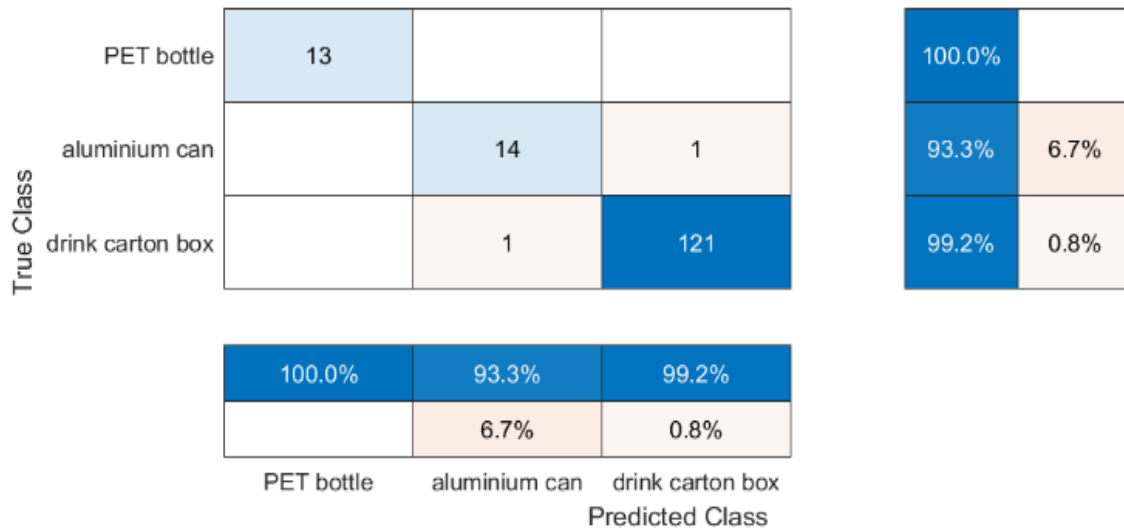| Neural network | Feature vector | Feature layer |
|---|---|---|
| AlexNet | 4,096 | Fc6 |
| DenseNet201 | 1,000 | Fc1000 |
| GoogLeNet | 1,000 | Loss 3-classifier |
| InceptionResnetV2 | 1,000 | Predictions |
| InceptionV3 | 1,000 | Predictions |
| MobileNetV2 | 1,000 | Logits |
| XceptionNet | 1,000 | Predictions |
| ShuffleNet | 1,000 | Node_202 |
| ResNet18 | 1,000 | Fc1000 |
| ResNet50 | 1,000 | Fc1000 |
| ResNet101 | 1,000 | Fc1000 |



Figure 5. Confusion matrix for AlexNet network

Table 2. Performance of the AlexNet network

| Performance | Result |
|---|---|
| Accuracy (%) | 98.67 |
| Error (%) | 1.33 |
| Precision | 0.9750 |
| Recall | 0.9750 |
| F1-Score | 0.9750 |
| Time is taken (s) | 2229.235 |

### 3.3.1. Model evaluation comparison

This project is to develop an image classification method for RVM. The dataset for this project belongs to imbalanced datasets, so the metric will be focused on F1-score rather than accuracy. Thus, all the neural networks will be compared among their F1-score and the computational time to perform the training process. Table 3 shows the model performance among the eleven neural networks.

Based on the Table 3, AlexNet is the preferred CNN network for transfer learning due to its combination of computational efficiency and a reasonably high F1-score of 97.50%. With a processing time of 2229.235 s, it outperforms other networks, including DenseNet201, which has a perfect F1-score but a much longer processing time of 12860.694 s. Additionally, Resnet 18 achieves a higher F1-score than

AlexNet but takes more time to complete the process. Overall, AlexNet is the best choice for the RVM as it achieves a good F1-score and completes the process in the shortest time among the eleven networks considered.

Table 3. Model's performance

| Neural networks | Accuracy (%) | Error (%) | Precision | Recall | F1-score | Computational time (s) |
|---|---|---|---|---|---|---|
| DenseNet201 | 100.00 | 0.00 | 1.00 | 1.00 | 1.00 | 12860.694 |
| ResNet18 | 99.33 | 0.67 | 0.9973 | 0.9792 | 0.9882 | 2982.253 |
| ResNet50 | 98.67 | 1.33 | 0.9608 | 0.9945 | 0.9774 | 6322.046 |
| AlexNet | 98.67 | 1.33 | 0.9750 | 0.9750 | 0.9750 | 2229.235 |
| ShuffleNet | 98.00 | 2.00 | 0.9370 | 0.9918 | 0.9636 | 2831.965 |
| ResNet101 | 98.00 | 2.00 | 0.9508 | 0.9689 | 0.9598 | 10019.215 |
| MobileNetV2 | 97.33 | 2.67 | 0.9216 | 0.9306 | 0.9261 | 4166.495 |
| InceptionResnetV2 | 96.00 | 4.00 | 0.9186 | 0.9022 | 0.9103 | 16930.18 |
| GoogLeNet | 96.00 | 4.00 | 0.8667 | 0.9450 | 0.9042 | 3267.364 |
| InceptionV3 | 95.33 | 4.67 | 0.8882 | 0.9580 | 0.9218 | 9273.491 |
| XceptionNet | 94.00 | 6.00 | 0.8508 | 0.8974 | 0.8735 | 16866.076 |

## 4.    CONCLUSION

A RVM with an image classification technique is proposed to automatically segregate three beverage containers: PET bottles, drink carton boxes, and aluminium cans. Transfer learning with CNN network will be used to perform classification in this research. The convolutional neural networks that used in this project are AlexNet, GoogLeNet, DenseNet201, InceptionResnetV2, InceptionV3, MobileNetV2, XceptionNet, ShuffleNet, ResNet 18, ResNet 50, and ResNet 101. This project needs to evaluate the F1-score of the different neural networks, and the computational time. The neural network with the highest F1-score and the shortest computational time to perform the training process will be used for the RVM. The dataset for this project contains 500 images of beverage containers and is separated into the training and validation sets.

AlexNet is the best CNN network in transfer learning. This is because in terms of computational time, AlexNet takes the shortest time, 2229.235 s, to complete the process. In terms of F1-score, AlexNet achieved 97.50% which is the fourth highest among the eleven networks. Although DenseNet201 achieved 100% F1-score but the computational time to complete the process is 12860.694 s which takes longer than AlexNet. In short, AlexNet is the best CNN network in transfer learning and will be chosen for the RVM.

## REFERENCES

[1]   S. D. Yaddanapudi, B. P. Makala, A. Yarlagadda, C. T. Sapram, S. T. Parsa, and S. Nallamadugu, "Collection of plastic bottles by reverse vending machine using object detection technique," *Materials Today: Proceedings*, vol. 80, pp. 1995–1999, 2023, doi: 10.1016/j.matpr.2021.06.037.
[2]   A. Alsabei, A. Alsayed, M. Alzahrani, and S. Al-Shareef, "Waste classification by fine-tuning pre-trained CNN and GAN," *International Journal of Computer Science and Network Security*, vol. 21, no. 8, pp. 65–70, 2021.
[3]   Z. Yi, "Evaluation and implementation of convolutional neural networks in image recognition," *Journal of Physics: Conference Series*, vol. 1087, Sep. 2018, doi: 10.1088/1742-6596/1087/6/062018.
[4]   C. Aggarwal, "Training a computer to classify distinct images using a convolutional neural network," *MLearning.ai*, 2021. https://medium.com/mlearning-ai/i-taught-my-computer-to-recognise-between-a-dog-and-a-cat-79c0cd529847 (accessed Aug. 01, 2022).
[5]   I. Kandel and M. Castelli, "Transfer learning with convolutional neural networks for diabetic retinopathy image classification. A review," *Applied Sciences*, vol. 10, no. 6, Mar. 2020, doi: 10.3390/app10062021.
[6]   F. Zhuang *et al.*, "A comprehensive survey on transfer learning," *Proceedings of the IEEE*, vol. 109, no. 1, pp. 43–76, Jan. 2021, doi: 10.1109/JPROC.2020.3004555.
[7]   D. Castillo, "Transfer learning for machine learning," *Seldon*, 2021. https://www.seldon.io/transfer-learning (accessed Jul. 07, 2022).
[8]   Q. Zhang, Q. Yang, X. Zhang, Q. Bao, J. Su, and X. Liu, "Waste image classification based on transfer learning and convolutional neural network," *Waste Management*, vol. 135, pp. 150–157, Nov. 2021, doi: 10.1016/j.wasman.2021.08.038.
[9]   K. Sanghvi, "Image classification techniques," Analytics Vidhya, medium.com, 2020. https://medium.com/analytics-vidhya/image-classification-techniques-83fd87011cac (accessed Aug. 01, 2022).
[10]  M. Hussain, J. J. Bird, and D. R. Faria, "A study on CNN transfer learning for image classification," in *Advances in Intelligent Systems and Computing*, Springer International Publishing, 2019, pp. 191–202.

[11] S. Agarwal, R. Gudi, and P. Saxena, "Image classification approaches for segregation of plastic waste based on resin identification code," *Transactions of the Indian National Academy of Engineering*, vol. 7, no. 3, pp. 739–751, Sep. 2022, doi: 10.1007/s41403-022-00324-4.

[12] C. Srinilta and S. Kanharattanachai, "Municipal solid waste segregation with CNN," in *2019 5th International Conference on Engineering, Applied Sciences and Technology (ICEAST)*, Jul. 2019, pp. 1–4, doi: 10.1109/ICEAST.2019.8802522.

[13] M. Duan, G. Wang, and C. Niu, "Method of small sample size image recognition based on convolution neural network," *Computer Engineering and Design*, vol. 39, no. 1, pp. 224–229, 2018.

[14] F. H. Khan, U. Qamar, and S. Bashir, "Enhanced cross-domain sentiment classification utilizing a multi-source transfer learning approach," *Soft Computing*, vol. 23, no. 14, pp. 5431–5442, Jul. 2019, doi: 10.1007/s00500-018-3187-9.

[15] M. Shaha and M. Pawar, "Transfer learning for image classification," in *2018 Second International Conference on Electronics, Communication and Aerospace Technology (ICECA)*, Mar. 2018, pp. 656–660, doi: 10.1109/ICECA.2018.8474802.

[16] R. Ding *et al.*, "Empirical study and improvement on deep transfer learning for human activity recognition," *Sensors*, vol. 19, no. 1, Dec. 2018, doi: 10.3390/s19010057.

[17] E. Rezende, G. Ruppert, T. Carvalho, F. Ramos, and P. de Geus, "Malicious software classification using transfer learning of ResNet-50 deep neural network," in *2017 16th IEEE International Conference on Machine Learning and Applications (ICMLA)*, Dec. 2017, pp. 1011–1014, doi: 10.1109/ICMLA.2017.00-19.

[18] A. Esuli, A. Moreo, and F. Sebastiani, "Funnelling: A new ensemble method for heterogeneous transfer learning and its application to cross-lingual text classification," *ACM Transactions on Information Systems (TOIS)*, vol. 37, no. 3, pp. 1–30, 2019.

[19] R. Tomari, N. S. Razali, N. F. Santosa, A. A. Kadir, and M. F. Hassan, "Reverse vending machine item verification module using classification and detection model of CNN," *International Journal of Advanced Computer Science and Applications*, vol. 12, no. 10, 2021, doi: 10.14569/IJACSA.2021.0121044.

[20] U. Ozkaya and L. Seyfi, "Fine-tuning models comparisons on garbage classification for recyclability," *arXiv preprint arXiv:1908.04393*, Aug. 2019.

[21] D. Ziouzios, D. Tsiktsiris, N. Baras, and M. Dasygenis, "A distributed architecture for smart recycling using machine learning," *Future Internet*, vol. 12, no. 9, Aug. 2020, doi: 10.3390/fi12090141.

[22] W.-L. Mao, W.-C. Chen, C.-T. Wang, and Y.-H. Lin, "Recycling waste classification using optimized convolutional neural network," *Resources, Conservation and Recycling*, vol. 164, Jan. 2021, doi: 10.1016/j.resconrec.2020.105132.

[23] S. Agrawal, "How to split data into three sets (train, validation, and test) and why?" *Towards Data Science*, 2021. https://towardsdatascience.com/how-to-split-data-into-three-sets-train-validation-and-test-and-why-e50d22d3e54c (accessed Aug. 04, 2022).

[24] P. Baheti, "Train test validation split: How to and best practices [2023]," *V7 labs*, 2022. https://www.v7labs.com/blog/train-validation-test-set (accessed Aug. 04, 2022).

[25] M. M. Taye, "Theoretical understanding of convolutional neural network: concepts, architectures, applications, future directions," *Computation*, vol. 11, no. 3, Mar. 2023, doi: 10.3390/computation11030052.

[26] R. Nirthika, S. Manivannan, A. Ramanan, and R. Wang, "Pooling in convolutional neural networks for medical image analysis: a survey and an empirical study," *Neural Computing and Applications*, vol. 34, no. 7, pp. 5321–5347, Apr. 2022, doi: 10.1007/s00521-022-06953-8.

[27] M. Umer, Z. Imtiaz, S. Ullah, A. Mehmood, G. S. Choi, and B.-W. On, "Fake news stance detection using deep learning architecture (CNN-LSTM)," *IEEE Access*, vol. 8, pp. 156695–156706, 2020, doi: 10.1109/ACCESS.2020.3019735.

[28] L. A. Jeni, J. F. Cohn, and F. De La Torre, "Facing imbalanced data-recommendations for the use of performance metrics," in *2013 Humaine Association Conference on Affective Computing and Intelligent Interaction*, Sep. 2013, pp. 245–251, doi: 10.1109/ACII.2013.47.

[29] "Why split data in the ratio 70:30?" *Information Gain Ltd*, 2012. http://information-gain.blogspot.com/2012/07/why-split-data-in-ratio-7030.html (accessed Aug. 29, 2022).

## BIOGRAPHIES OF AUTHORS

**Tan Hor Yan** 🆔 📇 sc ◐ received an engineering degree in electrical engineering from Universiti Teknikal Malaysia Melaka in 2017. She is currently pursuing an M.Sc. degree in electrical engineering at the Technical University of Malaysia, Melaka. She has published one paper in the International Journal of Electrical Engineering and Applied Sciences (IJEEAS) and is recyclable beverages containers for reverse vending machine sorting mechanism using image processing technique. Her research interests include image processing, image classification, reverse vending machines, IoT, Jetson Nano, and convolutional neural networks. She can be contacted at email: m012110024@student.utem.edu.my.

**Sazuan Nazrah Mohd Azam** 🆔 📇 sc ◐ received her B.Sc. Eng. from the Universiti Teknologi MARA (UiTM) in electrical engineering (instrumentation and system) and her M.Sc. from the Universiti Teknologi PETRONAS (UTP) in advanced process control. She completed her Ph.D. thesis and her work is focusing on the model predictive control implementation and analysis specifically for the modified quadruple tank system from DTU Compute, Technical University of Denmark. She is currently attached to Universiti Teknikal Malaysia Melaka (UTeM) as a senior lecturer in Department of Diploma Studies, Faculty of Electrical Engineering focusing on the field of control, instrumentation, and automation. She can be contacted at email: sazuan@utem.edu.my.

**Zamani Md. Sani** (iD) (g) (SC) (C) completed his degree in 2000 at Universiti Sains Malaysia and gained valuable experience during his 6-year tenure at Intel Malaysia Kulim. Subsequently, he pursued his Master's degree at the same university in 2009. Later, he transitioned to the field of education at Universiti Teknikal Malaysia Melaka and successfully earned his Ph.D. from Multimedia Universiti in 2020. His primary research interests lie in image processing and artificial intelligence. He can be contacted at email: zamanisani@utem.edu.my.

**Azizul Azizan** (iD) (g) (SC) (C) obtained his B.Eng. (Hons.) electronics engineering (telecommunications) degree from Multimedia University. He received his Ph.D. qualification in 2009, from University of Surrey in the area of 3.5G physical layer adaptation for satellite systems. He is currently with the Advanced Informatics Department, Razak Faculty of Technology and Informatics, Universiti Teknologi Malaysia. She can be contacted at email: azizulazizan@ieee.org.