JURNAL FASILKOM
Volume 14 No. 3 | Desember 2024: 761-765

P-ISSN : 2089-3353
E-ISSN : 2808-9162

Implementation Of Machine Learning To Identify Types Of Waste Using CNN Algorithm

Matsnan Haqqi[1], Lailatur Rochmah[2], Arisanti Dwi Safitri[3], Rizki Adhi Pratama[4], Tarwoto[5]
[1,2,3,4,5]Information System Departmen, Faculty of Computer Science, Amikom Purwokerto University
[1]21sa2125@mhs.amikompurwokerto.ac.id, [2]21sa2118@mhs.amikompurwokerto.ac.id,
[3]21sa2115@mhs.amikompurwokerto.ac.id, [4]21sa2157@mhs.amikompurwokerto.ac.id, [5]tarwoto@purwokerto.ac.id

Abstract

Waste management remains a significant challenge globally, particularly in Indonesia, where the annual waste generation reached 24.67 million tonnes in 2021, with only 50.43% properly managed. To address the issue of mixed organic and inorganic waste and the lack of public awareness regarding waste separation, this study applied machine learning, specifically the Convolutional Neural Network (CNN) algorithm, to classify waste types. The research aimed to develop an effective automated waste classification model to improve waste management processes. The research involved collecting a dataset of 2,848 images representing six waste categories: glass, cardboard, paper, metal, organic, and plastic. Preprocessing techniques such as cropping, noise reduction with Gaussian filters, and data augmentation were applied to enhance data quality. The dataset was divided into training, validation, and testing subsets in a 70:20:10 ratio. The CNN model employed feature extraction through convolution, activation, and pooling layers, followed by classification using a fully connected layer and a softmax function. Model performance was evaluated using accuracy, precision, recall, and F1-score metrics. The model achieved an overall accuracy of 95%, with an average precision, recall, and F1-score of 0.95 across all classes. These results demonstrate the CNN model's ability to reliably classify waste types. Compared to previous studies, this research achieved higher accuracy through the use of enhanced preprocessing and CNN optimization. This study highlights the potential of CNN-based models for automated waste classification, contributing to sustainable waste management practices and fostering environmental awareness in the future research.

Keywords: convolutional neural network; garbage types; classification

1. Introduction

Waste is a major problem faced by all countries, especially in Indonesia. The amount and type of waste in Indonesia continue to increase yearly, along with the population increase. According to data from the National Waste Management Information System (SIPSN) quoted from the official SIPSN website as of 2021, the amount of Indonesia's waste generation is 24.67 million tonnes/year, there is a reduction in waste of 13.38% or 3.3 million tonnes from the previous year, but waste handling in Indonesia can only be managed as much as 50.43% or 12.44 million tonnes/year [1].

Around our environment, the waste situation is still quite bad because organic and non-organic waste are mixed and not separated. This is due to the lack of public awareness of the need to dispose of waste according to its respective types. Proper waste management activities such as classifying waste according to organic and non-organic categories have the aim of not causing the spread of disease and causing odors that disturb the community [1]. The problem of understanding this type of waste can be solved by using a machine learning model.

This research is to apply machine learning to identify the type of waste using the CNN algorithm. CNN is one type of neural network that is commonly used on image data. CNN is a method that can self-learn features in complex images, which can receive input in the form of images, and determine what aspects or objects in an image can be used by machines to "learn" to recognize images and distinguish between one image and another [2], [3].

In previous research conducted by Asep Marzuki, with the title "Analysis of Bottle Waste Classification Models Based on Image Processing and Machine Learning in the Design of Automatic Bottle Waste Exchange Applications". This research was conducted to determine the accuracy value of the performance of the machine learning model in waste classification which resulted in the machine learning model successfully identifying the type of bottle waste with an average accuracy of 57.5%. From the results of the research conducted, it is still necessary to increase the accuracy value in further research both by increasing the number of datasets and optimizing the convolutional neural networks (CNN) algorithm to increase the accuracy value in classifying the type of bottle waste [4].

P-ISSN : 2089-3353
E-ISSN : 2808-9162

Other research was conducted by Devi and Hadi with classification using the CNN model. Research by dividing training data by 80% and test data by 20%, using CNN with dense as many as 128 layers and epochs as many as 50 produces full accuracy in training data and test data. So that by using the CNN model, the classification of inorganic and organic waste images can be done [11].

Previous research conducted by Cindy et al. applied the Convolutional Neural Network (CNN) algorithm with LeNet-5 architecture to detect the condition of fruit, whether fresh or rotten, through an Android-based application. The resulting model was able to achieve an accuracy of 93% and support the creation of a healthier lifestyle by helping the process of determining the quality of fruit automatically [12].

Harun Mukhtar's research with classification using the CNN method can classify mangrove fruit maturity with an accuracy of 96% with the number of images used influential during the training process. The greater the number of images used in the training process, the better the accuracy of the system created [5].

2. Research Methods

Figure 1. Research Stages

In the first stage, waste image data collection is carried out which is then divided into test data and training data. After the image data is collected, preprocessing is carried out such as cropping and applying filters to improve quality and remove noise in the image. The next stage is the application of the CNN algorithm which includes training and testing processes. The result obtained is the labeling of the input image.

2.1. Data Collection

The data collected is in the form of garbage images from open source, Kaggle entitled Garbage Image Dataset. The reason for using this dataset is because this dataset is representative and sample data. The quality of the data greatly affects the model that will be generated by the system. There are 2848 image files divided into 6 subclasses, namely glass (453), cardboard (404), paper (610), metal (411), organic (420), and plastic (550).

2.2. Preprocessing

To increase the amount of data without actually collecting new data, data augmentation is performed on the images, such as cropping, padding, and horizontal flipping. This augmentation aims to increase the variety of data by randomly taking images from different angles, thus strengthening the model to be built [6].

The waste data collected is in the form of images. The images are not all clear and noise-free. Therefore, preprocessing is done to improve image quality and reduce noise. Preprocessing is done through the process of cropping the image and applying a Gaussian filter to reduce noise in the image [7].

2.3. Splitting data

The process of dividing data into subsets is an important step in the development of an effective machine learning model. In this study, the data is divided into three main parts use split folder: training data, validation data, and test data. The division was carried out with a ratio of 70% for training data, 20% for validation data, and 10% for test data.

2.4. Application of CNN Algorithm

Convolutional Neural Network (CNN) is one of the most popular algorithms used in deep learning, where the learning model is specifically designed to perform live classification on two-dimensional media such as images, videos, text, or sound.

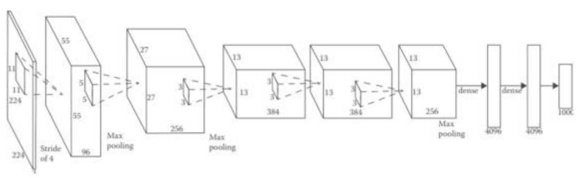

Figure 2. CNN Alexnet architecture [8].

Figure 2 shows the AlexNet architecture. This architecture was used by Alex et al. in the ImageNet Large Scale Visual Recognition Challenge 2012 competition. In the competition, the CNN method proposed by Alex et al. successfully outperformed other deep learning methods such as Support Vector Machine (SVM) in the case of image classification.

P-ISSN : 2089-3353
E-ISSN : 2808-9162

In general, CNN architecture is divided into two stages, namely feature learning and classification. Feature learning consists of three layers of processes often called convolution, activation, and pooling. In the classification stage, there is a fully connected layer consisting of layers of neurons and ends with a softmax function to determine the class of the input image.

2.5 Evaluation

$$Accuracy = \frac{TP+TN}{TP + FP + TN + FN}$$

Precision is a metric that shows the proportion of positive predictions that are correctly positive.

$$Prescision = \frac{TP}{TP + FP}$$

Recall is a metric that measures the number of positive cases that have been correctly identified by a prediction.

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

F1-Score is an evaluation measure that integrates precision and recall values to provide an overall picture of a model's performance.

$$F1\text{- } score = 2 \times \frac{precision \times recall}{precision + recall}$$

Macro-averaged-precision and recall are the precision and user-acquired values, respectively, it is calculated individually for each user as specified in, respectively [9].

3. Results and Discussion

3.1 Spliting data

At this stage, the collected dataset needs to be split into three main subsets for training, validation, and testing purposes. To perform data division automatically and efficiently, the Python split folders library is used. The following code is used to split the dataset based on the predefined ratio:

```
splitfolders.ratio(dataset_path,
                   output="data_sampah",
                   seed=42,
                   ratio=(.7, .2, .1),
                   group_prefix=None)
```

Using the split folders library, the process of splitting the dataset into subsets for training, validation, and testing can be done automatically and consistently.

3.2. Implementation of CNN

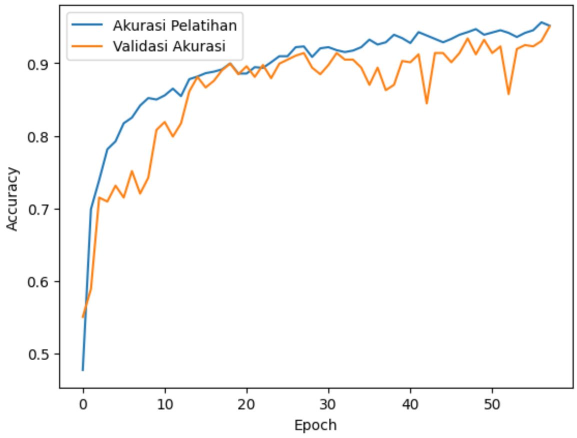

Figure 3. Accuracy training and validation

In the curve image above, the results of the training accuracy and validation accuracy process get pretty good results because the accuracy is increasing and the validation accuracy results are not too far from the training accuracy process as shown in Figure 3.

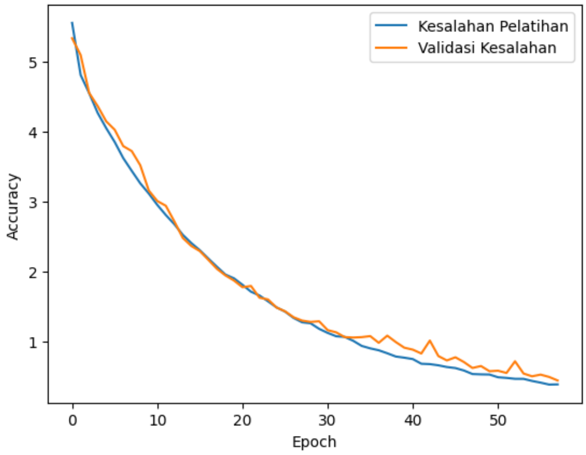

Figure 4. Loss training and validation

In the curve image below, the results of the training loss and validation loss process get pretty good graphic results because the training loss processing is decreasing and for validation loss, the processing results are also decreasing as well as in Figure 4.

P-ISSN : 2089-3353
E-ISSN : 2808-9162

3.2. Evaluation

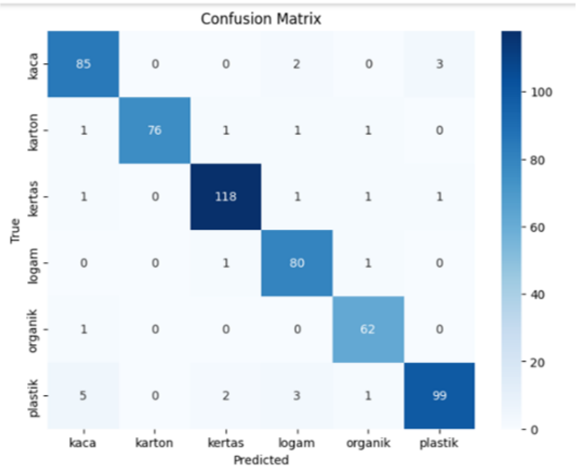

Figure 5. Confusion Matrix

The model generated from the training process has been tested using test data that is different from the training data. Based on the results of testing with the confusion matrix, the model showed strong performance on all evaluation metrics. The overall accuracy reached 95%, with precision, recall, and F1-score values averaging 95% for each class. Each class achieved high levels of accuracy and precision, with the highest precision in the cardboard class (100%) and the highest recall in the metal class (98%). F1-score values also showed consistency, ranging from 0.93 to 0.97 across all classes.

Table 1. Classification Report

Model	Accuracy	Precision	Recall	F1-Score
CNN	0.95	0.95	0.95	0.95

The litter classification model using the Convolutional Neural Network (CNN) algorithm achieved excellent performance based on evaluation metrics. With an overall accuracy of 95%, the model demonstrated a reliable ability to automatically identify waste types. The average precision, recall, and F1-score values of 0.95 each indicate that the model has a high ability to classify each type of waste with minimal error. This research is better than the previous research classification using the CNN model. Research by dividing training data by 80% and test data by 20%, using CNN with 128 dense layers and 50 epochs produces full accuracy on training data and test data [11].

4. Conclusion

The model training process was conducted for 100 epochs and stopped at the 58th epoch. Each epoch produced a model with kernel parameter values, weights, and biases, and recorded loss and accuracy values, as shown in Table 1. From the table, it can be seen that the model accuracy tends to increase at each epoch, with the highest training accuracy of 95.64% and the smallest loss value of 0.3782 at the 57th epoch. The resulting our model was then tested using different test data from the training data. Based on the results of testing with the confusion matrix, the model showed strong performance on all evaluation metrics, with overall accuracy reaching 95% and average precision, recall, and F1-score values of 95% in each class. F1-score values showed consistency, ranging from 0.93 to 0.97 across all classes. These results indicate that the model can classify waste types well and reliably, making it potentially optimal for use in automated waste classification applications.

Acknowledgment

The authors express their sincere gratitude for the research funding, which is entirely supported by the Lembaga Penelitian dan Pengabdian Masyarakat (LPPM) at Universitas Amikom Purwokerto.

Daftar Rujukan

[1] A. Ibnul Rasidi, Y. A. H. Pasaribu, A. Ziqri, and F. D. Adhinata, "Klasifikasi Sampah Organik dan Non-Organik Menggunakan Convolutional Neural Network," *Jurnal Teknik Informatika dan Sistem Informasi*, vol. 8, no. 1, Apr. 2022, doi: 10.28932/jutisi.v8i1.4314.

[2] R. Kurniawan, P. B. Wintoro, Y. Mulyani, and M. Komarudin, "IMPLEMENTASI ARSITEKTUR XCEPTION PADA MODEL MACHINE LEARNING KLASIFIKASI SAMPAH ANORGANIK," *Jurnal Informatika dan Teknik Elektro Terapan*, vol. 11, no. 2, Apr. 2023, doi: 10.23960/jitet.v11i2.3034.

[3] J. Homepage, S. Sahibu, I. Taufik, and P. Studi Sistem Komputer, "MALCOM: Indonesian Journal of Machine Learning and Computer Science Implementation of the Convolutional Neural Network Algorithm for Classifying Types of Organic and Non-Organic Waste Implementasi Algoritma Convolutional Neural Network untuk Klasifikasi Jenis Sampah Organik dan Non Organik," vol. 4, no. 3, pp. 840–852, 2024, doi: 10.57152/malcom.v7i2.1346.

[4] A. Marzuki, A. Zaky, A. Cahayani Adha, T. Mohammad Yoshandi, U. Awal Bros, and K. Pekanbaru, "JURNAL MEDIA INFORMATIKA [JUMIN] Analisis Model Klasifikasi Sampah Botol Berbasis Image Processing Dan Machine Learning Dalam Rancang Bangun Aplikasi Penukaran Sampah Botol Otomatis," 2024.

[5] H. Mukhtar *et al.*, "Deep Learning Untuk Klasifikasi Kematangan Buah Mangrove Berdasarkan Warna".

[6] J. Sanjaya and M. Ayub, "Augmentasi Data Pengenalan Citra Mobil Menggunakan Pendekatan Random Crop, Rotate, dan Mixup," *Jurnal Teknik Informatika dan Sistem Informasi*, vol. 6, no. 2, Aug. 2020, doi: 10.28932/jutisi.v6i2.2688.

[7] "8 ANALISA PERBANDINGAN METODE FILTER GAUSSIAN, MEAN DANMEDIAN TERHADAP REDUKSI NOISE".

[8] R. Av Venkatesan and B. Li, "Convolutional Neural Networks in Visual Computin."

P-ISSN : 2089-3353
E-ISSN : 2808-9162

[9] A. Darvishy, H. Ibrahim, F. Sidi, and A. Mustapha, "HYPNER: A Hybrid Approach for Personalized News Recommendation," *IEEE Access*, vol. 8, pp. 46877–46894, 2020, doi: 10.1109/ACCESS.2020.2978505.

[10] I. Muhammad Erwin, E. Prakasa, and B. Sugiarto, "KAYU7NET: IDENTIFIKASI DAN EVALUASI F-MEASURE CITRA KAYU BERBASIS DEEP CONVOLUTION NEURAL NETWORK (DCNN)", doi: 10.25126/jtiik.

[11] O. D. S. Sunanto and P. H. Utomo, "Implementasi Deep Learning Dengan Convolutional Neural Network Untuk Klasifikasi Gambar Sampah Organik Dan Anorganik," *Pattimura Proceeding Conf. Sci. Technol.*, vol. 1, no. 2, pp. 335–340, 2022

[12] D. C. Agustin, M. A. Rosid, and N. Ariyanti, "Implementasi Convolutional Neural Network Untuk Deteksi Kesegaran Pada Apel," *J. Fasilkom*, vol. 13, no. 02, pp. 145–150, 2023.