

Klasifikasi Sampah Organik dan Non-Organik Menggunakan *Convolutional Neural Network*

<http://dx.doi.org/10.28932/jutisi.v8i1.4314>

Riwayat Artikel

Received: 28 Desember 2021 | Final Revision: 6 Maret 2022 | Accepted: 12 Maret 2022

Abdurrahman Ibnul Rasidi^{#1}, Yolanda Al Hidayah Pasaribu^{#2}, Afzal Ziqri^{#3},
Faisal Dharma Adhinata^{✉#4}

[#] Program Studi Rekayasa Perangkat Lunak, Institut Teknologi Telkom Purwokerto
Jl. DI Panjaitan No.128, Karangreja, Purwokerto Kidul, Kec. Purwokerto Sel., Kabupaten Banyumas, Jawa Tengah,
Indonesia 53147

¹19104032@ittelkom-pwt.ac.id

²19104057@ittelkom-pwt.ac.id

³19104002@ittelkom-pwt.ac.id

⁴faisal@ittelkom-pwt.ac.id

Abstract — Garbage is a unique problem in Indonesia. From ordinary waste to limited emergency plastic waste, Indonesia is the second-largest source of plastic waste in the world. Separate collection and disposal of waste is one way to reduce the waste generated by society and industry in Indonesia. Sorting out the types of waste is the first step before the recycling process. In the field of Computer Vision research, it is difficult to see the type and form of waste with a camera, therefore this study aims to overcome this problem by using Deep Learning technology which is expected to be implemented in the whole of Indonesia starting from some of the largest waste-producing cities. Deep Learning is a computer technique for learning like a human - with experiments being a Part of Machine Learning that can be used to classify images. The method used in this study uses the Convolutional Neural Network (CNN) method which can be used to detect and recognize objects in an image, which can be used to create an automatic waste classification system. Broadly speaking, CNN utilizes the convolution process by moving a convolution kernel (filter) of a certain size to an image, the computer gets new representative information from the results of multiplying that part of the image with the filter used. The test results show that the CNN method can classify inorganic waste with accuracy. 96% and organic waste with an accuracy of 62%.

Keywords— Convolutional Neural Network; Deep Learning; Garbage; Recycling.

I. PENDAHULUAN

Sampah merupakan masalah besar yang dihadapi semua negara, khususnya Indonesia. Jumlah dan jenis sampah di Indonesia terus meningkat setiap tahunnya seiring dengan pertumbuhan jumlah penduduk. Menurut data Sistem Informasi Penanggulangan Sampah Nasional (SIPSN) yang dikutip dari situs website resmi SIPSN Per 2021, jumlah timbulan sampah Indonesia adalah sebanyak 24,67 juta ton/tahun terdapat pengurangan sampah sebanyak 13,38% atau 3,3 juta ton dari tahun sebelumnya akan tetapi penanganan sampah di Indonesia hanya mampu dikelola sebanyak 50,43% atau 12,44 juta ton/tahun [1]. Dengan jumlah tersebut, Indonesia dapat memproduksi sekitar 67.590 ton atau 0,25 kg/orang per hari. Dari data tersebut dapat dikatakan bahwa Indonesia saat ini dalam keadaan darurat sampah. Ini adalah masalah besar dan tantangan bersama dalam menangani sampah. Undang-Undang Republik Indonesia Nomor 18 Tahun 2008 tentang Pengelolaan Sampah. Pasal 1 (1) menyatakan bahwa “sampah adalah sisa kegiatan sehari-hari manusia atau proses alam yang berbentuk padat” [2]. Ada dua jenis sampah, sampah anorganik dan sampah organik, tergantung dari sifatnya. Sampah anorganik adalah sampah umum yang tidak dapat diuraikan, seperti logam/besi, pecahan kaca, dan plastik. Sampah organik biasanya merupakan sampah yang sudah membusuk, seperti sampah dapur, daun-daunan dan buah-buahan [3].

Keadaan sampah yang ada di lingkungan sekitar kita, sekarang selalu dalam keadaan tercampur dan tidak terpilah [4]. Hal itu disebabkan karena kurangnya pengetahuan dan kesadaran pada masyarakat sekitar ketika melakukan kegiatan pembuangan sampah sesuai dengan jenisnya.

Pengelolaan yang tepat, termasuk pengelolaan sampah, khususnya dengan mengelompokkan sampah ke dalam kategori organik atau anorganik, tujuannya agar tidak menimbulkan bau, dan menyebarkan penyakit [5]. Dengan adanya sistem

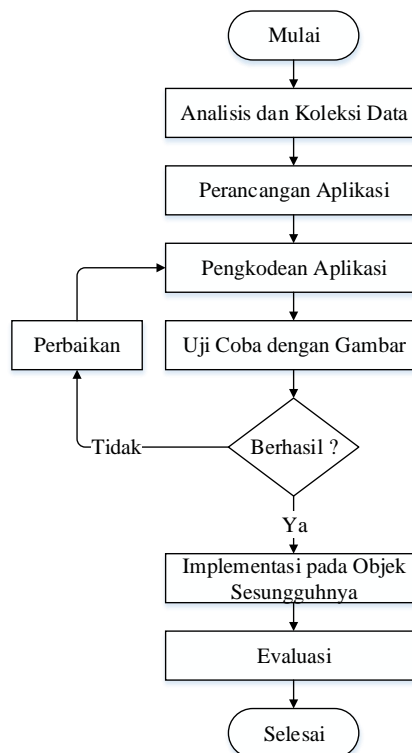
pendeteksi jenis sampah diharapkan mengurangi permasalahan yang ada dan tentunya akan mempermudah pengelolaan sampah untuk diolah kembali atau dimanfaatkan kembali.

Berdasarkan masalah yang ada, mereka perlu dipecahkan dengan cara baru untuk mencapai akurasi yang tinggi, implementasinya adalah terdapat pada TPA (Tempat Pembuangan Akhir) yang nantinya akan berbentuk seperti mesin X-ray, Pemeriksaan Sinar-X atau yang lebih umum dikenal dengan *Rontgen* adalah salah satu teknik pencitraan medis yang menggunakan radiasi elektromagnetik untuk mengambil pencitraan atau foto bagian dalam tubuh, yang dirancang serba otomatis yang mana dapat memilah sampah menggunakan teknologi AI dengan menggunakan pendekatan *deep learning*. *Deep learning* adalah metode pembelajaran oleh mesin dengan meniru cara kerja sistem saraf otak manusia. Menggunakan metode *Convolutional neural networks* (CNN), yang juga dikenal sebagai ConvNets. CNN terdiri dari beberapa lapisan dan sering digunakan untuk pemrosesan gambar dan deteksi objek. Dengan kata lain, melakukan pembelajaran yang mendalam di berbagai tingkatan lapisan. Pendekatan *deep learning* mengkategorikan data menjadi dua sesi, yaitu sesi *training* dan sesi *testing*. Dalam sesi *training*, akan mempelajari cara mengekstrak karakteristik dari setiap data sehingga satu label dapat dibedakan dari yang lain. Selama sesi *testing*, data yang diuji dapat dianalisis dari hasil sesi *training* [6]. Oleh karena itu, diperlukan metode *deep learning* untuk menganalisis dataset berupa gambar jenis sampah organik dan anorganik.

II. METODE PENELITIAN

Pada penelitian kali ini menggunakan model *Convolutional Neural Network* (CNN) untuk mengklarifikasi sampah organik maupun anorganik. Penelitian ini menggunakan dataset *Waste Classification Data*, dataset tersebut kami peroleh dari situs website Kaggle yang memang menyediakan dataset untuk penelitian (6). Data yang dianalisis dan digunakan dalam penelitian ini adalah gambar sampah organik dan anorganik. Alur dari penelitian ini terdapat pada Gambar 1.

Langkah pertama dari penelitian ini adalah mengumpulkan dan menganalisis data yang diperlukan untuk pengembangan program ini, kemudian setelah berhasil mengumpulkan dan menganalisa data kita mulai merancang aplikasi yang mana kita memutuskan program apa yang cocok untuk mengatasi masalah tersebut setelah perancangan kemudian kita masuk ke tahap pengkodean program yang mana kita mulai membangun dan membuat program tersebut setelah selesai maka kita masuk ke tahap pengujian yang mana jika program tersebut gagal atau tidak sesuai maka akan diperbaiki lalu di uji kembali sampai program tersebut benar benar berhasil setelah berhasil uji coba kemudian program tersebut akan kami implementasikan ke objek sesungguhnya yang mana tahap ini bisa disebut pengujian tahap 2 setelah semua berhasil maka perlu adanya evaluasi yang berguna untuk memperbaiki kekurangan dan kelemahan program.



Gambar 1. Tahapan penelitian

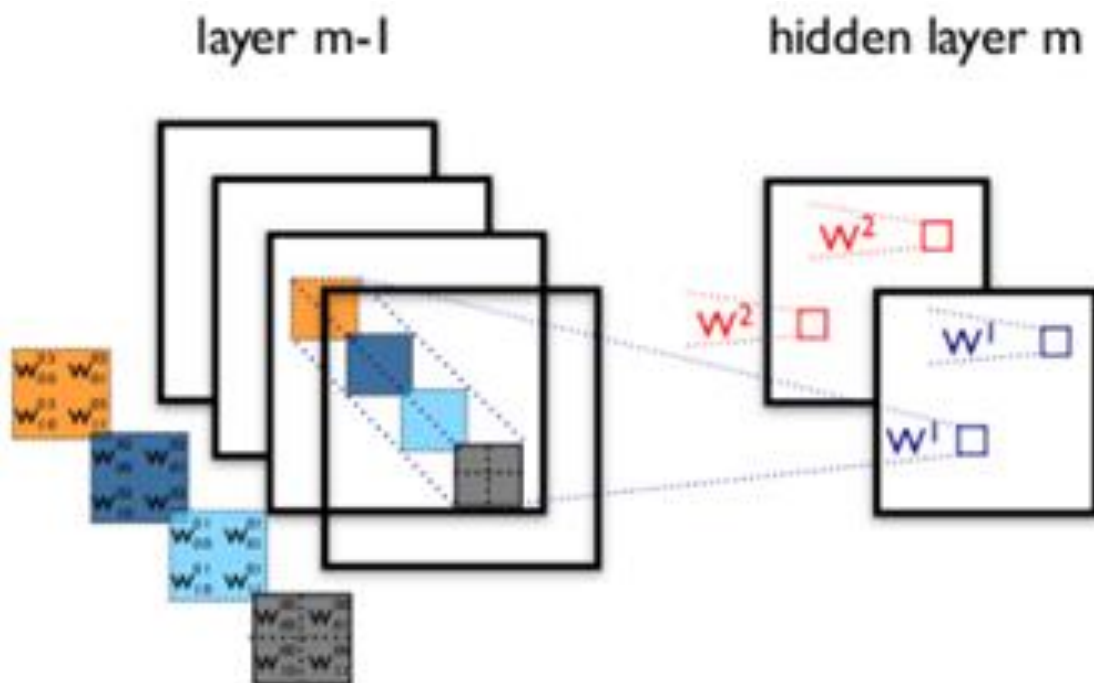
A. Deep Learning

Deep learning adalah metode pembelajaran ekspresi yang memungkinkan model komputer dengan beberapa tingkat pemrosesan untuk mempelajari representasi data dengan beberapa tingkat abstraksi [7]. Model deep learning pada dasarnya dibangun di atas jaringan saraf tiruan (*Neural Network*). Penelitian ini telah berlangsung sejak tahun 1980-an, tetapi baru-baru ini dihidupkan kembali dengan munculnya komputer yang lebih cepat. Deep learning juga merupakan ilmu baru dalam *machine learning* dan akhir-akhir ini banyak digunakan dengan berkembangnya teknologi akselerasi GPU [8].

Deep learning cocok untuk visi komputer, seperti halnya dengan klasifikasi objek dalam gambar. Menerapkan salah satu metode *machine learning* yang dapat digunakan untuk mengklasifikasikan citra objek, yaitu metode *Convolutional Neural Network* (CNN).

B. Convolutional Neural Network

Convolutional Neural Network (CNN) adalah salah satu jenis dari Artificial Neural Network (ANN) yang berfokus pada pemrosesan gambar, video dan suara [9]. CNN memiliki proses kerja yang sama dengan ANN yang menggunakan sistem kerja sel otak manusia, namun CNN setiap neuronnya dipresentasikan dalam bentuk 2 dimensi, CNN dipilih karena prosesnya yang lebih sederhana. Dalam CNN, perilaku linier parameter bobot berbeda dengan perilaku linier CNN karena data yang dikirim melalui jaringan bersifat dua arah. Di CNN, operasi linier menggunakan operasi konvolusi, bobotnya adalah empat dimensi, bukan satu dimensi, yang merupakan kumpulan kernel konvolusi. Gambar 2 menunjukkan set kernel konvolusi.



Gambar 2. Set kernel konvolusi [10]

Urutan algoritma CNN dimulai dengan preprocessing, yang memodifikasi data gambar, seperti ukuran dan pengurangan noise. Selain itu, deteksi region of interest (ROI), penekanan latar belakang, pengambilan fitur gambar, dan pengenalan objek memungkinkan Anda untuk mencocokkan fitur yang diambil dengan objek dalam gambar. Langkah terakhir dalam proses pengambilan keputusan adalah konvolusi gambar. CNN mencakup berbagai jenis lapisan, yaitu lapisan komposit, dan lapisan yang terhubung penuh. Lapisan transformasi dan lapisan agregasi adalah lapisan yang memisahkan ANN dan CNN dalam JST tanpa kedua lapisan tersebut. Kedua lapisan bertanggung jawab untuk melakukan filter gambar untuk menemukan fitur. Namun, karena fitur dan perilaku lapisan CNN yang terhubung penuh adalah perhitungan ANN, ada tugas untuk menentukan hasil fitur menggunakan label yang tersedia.

Banyak penelitian dan pengembangan telah dilakukan pada algoritma CNN. Penelitian sedang dilakukan untuk meningkatkan akurasi, menghemat sumber daya, dan mengurangi kesalahan dalam algoritma CNN. Ini dapat dilakukan dengan mengubah lapisan CNN. Dalam sebuah penelitian, model CNN tertentu dihasilkan dan dihitung untuk mencapai

hasil yang optimal. Model-model ini adalah AlexNet, VGG, Resnet, Inception, MobileNet, dll [11]. Setiap model yang ada memiliki jumlah lapisan kompleksitas dan lapisan bentuk yang berbeda. Dalam penelitian ini, kami melakukan analisis model algoritma CNN yang tepat untuk memastikan klasifikasi sampah yang optimal. Oleh karena itu, algoritma ini dapat langsung digunakan untuk melakukan klasifikasi sampah otomatis oleh sistem [12]. Penelitian ini bertujuan untuk mempersingkat waktu tahapan daur ulang sampah dan mempermudah proses pengumpulan dan pemilahan sampah.

III. HASIL DAN PEMBAHASAN

A. Pengumpulan Data

Koleksi data yang diambil berupa gambar dengan format .jpg yang diperoleh dari sebuah website Kaggle [13]. Total gambar yang dikumpulkan sebanyak 25.077 gambar yang dibagi menjadi 2 data, yaitu untuk data train kelas organik sebanyak 12.565 gambar kemudian untuk kelas recycle pada data train ini terdapat 9.999 gambar. Maka dari itu, pada data train ditemukan 22.564 gambar (90%) dengan 2 class, seperti pada Gambar 3. Sedangkan pada data test, kelas organik terdapat 1.401 gambar, untuk kelas recycle terdapat 1.112 gambar. Maka total pada data test ditemukan 2513 gambar (10%) dengan 2 class.

```
Preprocessing the Training set  
  
In [3]: train_datagen = ImageDataGenerator(rescale=1./255,  
                                           shear_range=0.2,  
                                           zoom_range=0.2,  
                                           horizontal_flip=True)  
training_set = train_datagen.flow_from_directory('DATASET/TRAIN',  
                                                target_size=(64, 64),  
                                                batch_size=32,  
                                                class_mode='binary')  
  
Found 22564 images belonging to 2 classes.  
  
Preprocessing the Test set  
  
In [4]: test_datagen = ImageDataGenerator(rescale=1./255)  
test_set = test_datagen.flow_from_directory('DATASET/TEST',  
                                           target_size=(64, 64),  
                                           batch_size=32,  
                                           class_mode='binary')  
  
Found 2513 images belonging to 2 classes.
```

Gambar 3. Preprocessing data train dan data test

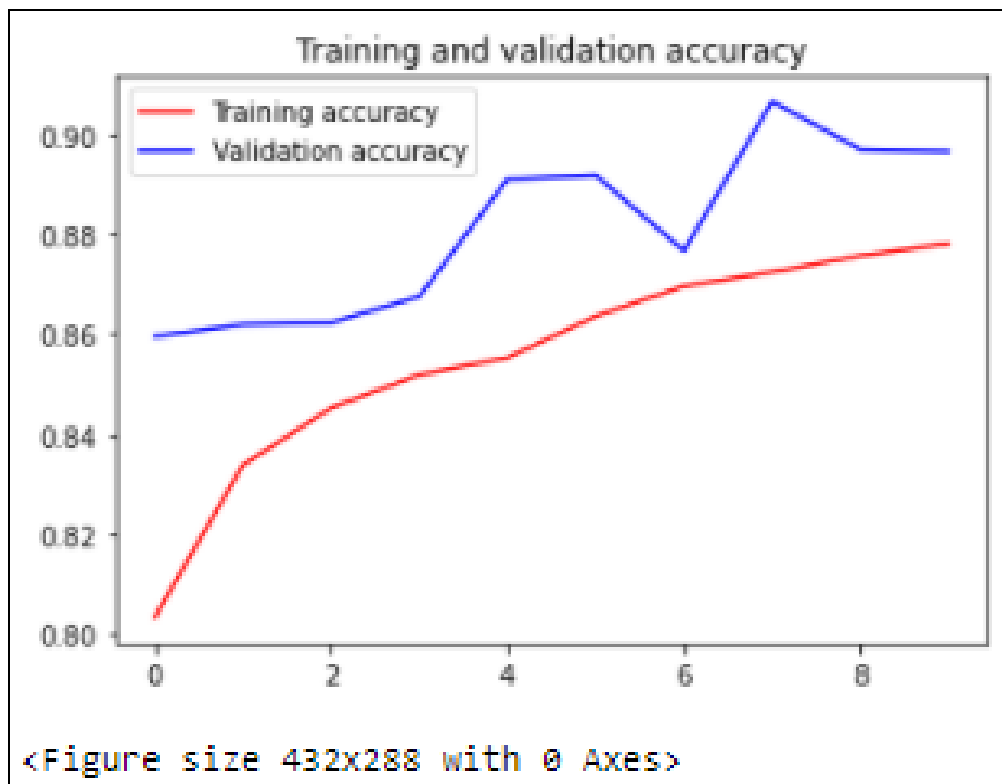
B. Pelatihan Data

Data *training* yang digunakan adalah 90% total data gambar, maka digunakan 22.564 gambar yang terbagi menjadi dua kelas. Proses ini menggunakan fungsi *compile* dengan konfigurasi *Optimizer='adam'*, *loss='categorical_crossentropy'*, *metrics = ['accuracy']*, dan proses *training* yang dapat dikonfigurasi *x = training_set*, *validasi_data = test_set*, *epochs = 10*. Gambar 4 adalah hasil training.

```
Epoch 1/10  
706/706 [=====] - 254s 358ms/step - loss: 0.4355 - accuracy: 0.8034 - val_loss: 0.3752 - val_accuracy:  
0.8595  
Epoch 2/10  
706/706 [=====] - 255s 362ms/step - loss: 0.3810 - accuracy: 0.8338 - val_loss: 0.3340 - val_accuracy:  
0.8619  
Epoch 3/10  
706/706 [=====] - 251s 356ms/step - loss: 0.3650 - accuracy: 0.8452 - val_loss: 0.3331 - val_accuracy:  
0.8623  
Epoch 4/10  
706/706 [=====] - 250s 354ms/step - loss: 0.3474 - accuracy: 0.8519 - val_loss: 0.3464 - val_accuracy:  
0.8675  
Epoch 5/10  
706/706 [=====] - 251s 355ms/step - loss: 0.3390 - accuracy: 0.8553 - val_loss: 0.2951 - val_accuracy:  
0.8910  
Epoch 6/10  
706/706 [=====] - 249s 353ms/step - loss: 0.3243 - accuracy: 0.8636 - val_loss: 0.2811 - val_accuracy:  
0.8918  
Epoch 7/10  
706/706 [=====] - 248s 351ms/step - loss: 0.3146 - accuracy: 0.8696 - val_loss: 0.3227 - val_accuracy:  
0.8766  
Epoch 8/10  
706/706 [=====] - 249s 352ms/step - loss: 0.3041 - accuracy: 0.8724 - val_loss: 0.2662 - val_accuracy:  
0.9065  
Epoch 9/10  
706/706 [=====] - 248s 352ms/step - loss: 0.2976 - accuracy: 0.8757 - val_loss: 0.2775 - val_accuracy:  
0.8969  
Epoch 10/10  
706/706 [=====] - 248s 351ms/step - loss: 0.2920 - accuracy: 0.8781 - val_loss: 0.2802 - val_accuracy:  
0.8965
```

Gambar 4. Hasil training

Grafik berikut menggambarkan *Training* dan *Validation Accuracy* dan *Training* dan *Validation Loss* pada Gambar 5 dan 6. Dapat disimpulkan bahwa ini sangat baik dalam hal akurasi *training*, tetapi tidak terlalu baik dalam akurasi *Validation* karena kurangnya sampel data gambar, hanya 10% dari 25.077 gambar. Solusi yang tersedia adalah dengan menambahkan sampel data gambar pada *validation*, menambahkan beberapa *epochs* dan memodifikasi model CNN.



Gambar 5. Grafik *training* dan *validation accuracy*

Pada Gambar 5 menampilkan hasil grafik akurasi dari model *Convolutional Neural Network* (CNN) yang dibuat dengan tingkat akurasi pada data *training* memperoleh 0.90 dan data *testing* memperoleh akurasi 0.88.




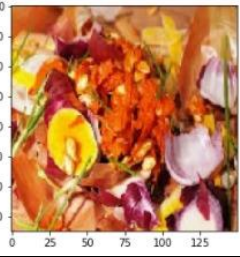
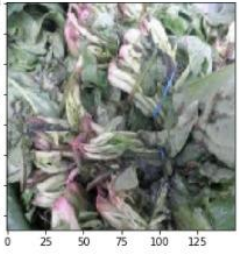
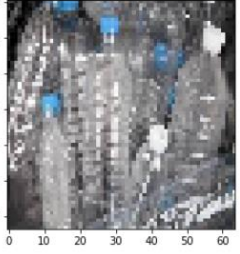
Gambar 6. Grafik *training* dan *validation loss*

Pada gambar 6 menampilkan hasil grafik *loss* dari model *Convolutional Neural Network* (CNN) yang dibuat dengan tingkat *loss* pada data *training* memperoleh nilai *loss* 0.298 dan data *testing* memperoleh nilai 0.282.

C. Implementasi dan Uji Coba

Pada proses uji coba kami menggunakan sampel dari dataset yang berisikan foto atau gambar sampah, Kami mengambil sampel secara acak dari gambar atau foto sampah baik organik maupun anorganik, dan juga dari berbagai kondisi sampah baik yang bentuknya masih terlihat jelas maupun yang bentuknya sudah tidak terlihat jelas atau sudah membusuk. Penerapannya adalah terdapat pada TPA (Tempat Pembuangan Akhir) yang nantinya akan berbentuk seperti mesin X-ray, Pemeriksaan Sinar-X atau yang lebih umum dikenal dengan Rontgen adalah salah satu teknik pencitraan medis yang menggunakan radiasi elektromagnetik untuk mengambil pencitraan atau foto bagian dalam tubuh, yang dirancang serba otomatis yang mana dapat memilah sampah menggunakan teknologi AI dengan menggunakan pendekatan deep learning. Tabel 1 merupakan contoh hasil identifikasi menggunakan metode CNN.

TABEL 1
HASIL IDENTIFIKASI MENGGUNAKAN METODE CNN

Keterangan	Hasil Prediksi
<p>Pada gambar di samping merupakan hasil prediksi bahwa gambar tersebut dikategorikan sebagai sampah anorganik.</p>	<p>DATASET\TEST\R\R_10577.jpg An-Organik</p> 
<p>Pada gambar tersebut merupakan hasil prediksi bahwa gambar tersebut dikategorikan sebagai sampah organik.</p>	<p>DATASET\TEST\O\O_13605.jpg Organik</p> 
<p>Pada gambar tersebut merupakan hasil prediksi bahwa gambar tersebut dikategorikan sebagai sampah organik.</p>	<p>DATASET\TRAIN\O\O_10527.jpg Organik</p> 
<p>Pada gambar di samping merupakan hasil prediksi bahwa gambar tersebut dikategorikan sebagai sampah anorganik.</p>	<p>DATASET\TRAIN\R\R_150.jpg An-organik</p> 

D. Pengujian

Pada fase pengujian, kami memasukkan sebanyak 50 data dari berbagai jenis dan kondisi dari sampah. Hal ini diperuntukkan agar implementasi pada objek yang sesungguhnya dapat berjalan lancar. Dikarenakan kita tidak tahu kondisi sampah yang sebenarnya pada objek sesungguhnya. Tabel 2 menunjukkan hasil pengujian yang menunjukkan akurasi prediksi dan tingkat kesalahan untuk setiap kondisi sampah.

TABEL 2
HASIL PENGUJIAN

Kondisi Sampah	Data	Benar	Salah	Akurasi	Error
Anorganik	50	48	2	96%	4%
Organik	50	31	19	62%	38%

Menurut pengamatan kami hasil pengujian dari pemilahan sampah ini sudah cukup bagus dengan tingkat akurasi 96% pada sampah anorganik dan 62% sampah organik. Namun kekurangannya adalah ketika sistem kami menganalisa data atau gambar yang merupakan sampah organik lebih sering terjadi error mungkin dikarenakan sampah organik yang sangat beragam mulai dari bentuk, warna, dan lain sebagainya. Bila dibandingkan dengan sampah anorganik yang bentuknya tidak bermacam-macam dan mungkin bentuknya hampir sama dengan yang lain. Sebagai contoh sampah botol plastik dan sampah botol kaca keduanya terkadang memiliki bentuk yang mirip dan bila dibandingkan dengan sampah organik yang bentuknya bermacam-macam. Sebagai contoh sampah dapur seperti sisa makanan, sayur dan buah yang sudah busuk, dan lain sebagainya. Hal itu saja mungkin ada perbedaan yang cukup signifikan bila dilihat dari bentuk objeknya.

IV. SIMPULAN

Penelitian ini memberikan solusi pengelolaan sampah dengan mengidentifikasi jenis sampah. Pengelola sampah dapat melakukan daur ulang yang sesuai berdasarkan jenis sampahnya melalui metode *deep learning* dan *Convolutional Neural Network* (CNN). Hasil penelitian ini menunjukkan angka ketepatan sebesar 96% untuk sampah anorganik dan 62% untuk sampah organik. Metode *deep learning* dan CNN yang diimplementasikan pada penelitian ini digunakan untuk mengidentifikasi jenis *waste* dengan cepat, mudah, dan akurat, dibuktikan dengan hasil pada Tabel 2 sehingga hasil dari penelitian ini diharapkan dapat membantu mempermudah pekerjaan dalam pengolahan sampah. Penelitian selanjutnya dapat dilakukan dengan mengubah arsitektur CNN, serta meningkatkan jumlah gambar dalam dataset, sehingga dapat memberikan tingkat akurasi tertinggi, dan mengurangi nilai kesalahan. Kemudian, rencana kedepan adalah dapat merealisasikan pembuatan alat ini guna untuk mengurangi masalah sampah yang ada di Indonesia yang pada tiap tahunnya masalah sampah akan terus bertambah.

DAFTAR PUSTAKA

- [1] SIPSIN, "Sistem Informasi Pengelolaan Sampah Nasional," 2021.
- [2] A. Kahfi, "Tinjauan Terhadap Pengelolaan Sampah," *Jurisprud. Jur. Ilmu Huk. Fak. Syariah dan Huk.*, vol. 4, no. 1, p. 12, 2017, doi: 10.24252/jurisprudentie.v4i1.3661.
- [3] H. Hayat, H., & Zayadi, "Model Inovasi Pengelolaan Sampah Rumah Tangga. JU-ke (Jurnal Ketahanan Pangan)," *JU-ke (Jurnal Ketahanan Pangan)*, vol. 2, no. 2, pp. 131–141, 2018.
- [4] A. E. Widodo, "Otomatisasi Pemilahan Sampah Berbasis Arduino Uno," vol. 6, no. 1, pp. 12–18, 2020.
- [5] M. Yunus, "Rancang Bangun Prototipe Tempat Sampah Pintar Pemilahan Sampah Organik Dan Anorganik Menggunakan Arduino," *Proceeding STIMA*, vol. 1, no. 1, pp. 340–343, 2018.
- [6] L. Marifatul Azizah, S. Fadillah Umayah, and F. Fajar, "Deteksi Kecacatan Permukaan Buah Manggis Menggunakan Metode Deep Learning dengan Konvolusi Multilayer," *Semesta Tek.*, vol. 21, no. 2, pp. 230–236, 2018, doi: 10.18196/st.212229.
- [7] M. R. D. Septian, A. A. A. Paliwang, M. Cahyanti, and E. R. Swedia, "Penyakit Tanaman Apel Dari Citra Daun Dengan Convolutional Neural Network," *Sebatik*, vol. 24, no. 2, pp. 207–212, 2020, doi: 10.46984/sebatik.v24i2.1060.
- [8] W. S. Eka Putra, "Klasifikasi Citra Menggunakan Convolutional Neural Network (CNN) pada Caltech 101," *J. Tek. ITS*, vol. 5, no. 1, 2016, doi: 10.12962/j23373539.v5i1.15696.
- [9] A. Peryanto, A. Yudhana, and R. Umar, "Klasifikasi Citra Menggunakan Convolutional Neural Network dan K Fold Cross Validation," *J. Appl. Informatics Comput.*, vol. 4, no. 1, pp. 45–51, 2020, doi: 10.30871/jaic.v4i1.2017.
- [10] S. R. Suartika E. P., I Wayan, Wijaya Arya Yudhi, "Klasifikasi Citra Menggunakan Convolutional Neural Network (Cnn) Pada Caltech 101," *J. Tek. ITS*, vol. 5, no. 1, p. 76, 2016, [Online]. Available: <http://repository.its.ac.id/48842/>.
- [11] J. Feriawan and D. Swanjaya, "Perbandingan Arsitektur Visual Geometry Group dan MobileNet Pada Pengenalan Jenis Kayu," *Semin. Nas. Inov. Teknol.*, pp. 185–190, 2020.
- [12] H. S. Stephen, Raymond, "Aplikasi Convolution Neural Network Untuk Mendeteksi Jenis-Jenis Sampah," *Explor. - J. Sist. Inf. dan Telemat. (Telekomunikasi, Multimedia, dan Informasi)*, vol. 10, no. 2, pp. 122–132, 2019.
- [13] S. Sekar, "Waste Classification data | Kaggle." 2019.