

BAB III

METODOLOGI PENELITIAN

3.1. Objek Penelitian

Objek pada penelitian ini adalah citra sampah rumah tangga yang terdiri dari 12 kelas yaitu kertas, karton, biologi, logam, plastik, kaca hijau, kaca coklat, kaca putih, pakaian, sepatu, baterai, dan sampah residu. 12 kelas sampah yang diklasifikasikan termasuk kedalam 3 jenis sampah yaitu sampah organik, anorganik, dan berbahaya.

3.2. Diagram Alir Penelitian

Diagram alir penelitian yang berisi tahapan – tahapan penelitian yang dilakukan ditunjukkan pada gambar 3.1.



Gambar 3. 1 Diagram Alir Penelitian

3.3.1. Studi Pustaka

Pada tahapan ini, penulis membaca dan memahami konsep dan permasalahan mengenai *machine learning* yang ada pada jurnal maupun buku. Hasil yang diperoleh akan digunakan sebagai dasar untuk menulis dan melakukan penelitian.

3.3.2. Perumusan Masalah dan Tujuan

Perumusan masalah dilakukan untuk menemukan masalah yang ada dan penyusunan tujuan penelitian dilakukan untuk menentukan tujuan dari penelitian ini.

3.3.3. Mempersiapkan Data

Penelitian ini menggunakan dataset yang diambil dari dataset *Garbage Classification* yang dapat di akses pada <https://www.kaggle.com/mostafaabla/garbage-classification>.

Dataset yang digunakan terdiri dari 15.515 citra sampah rumah tangga yang dibagi menjadi 12 jenis sampah rumah tangga yang diklasifikasikan menjadi 12 kelas yaitu kertas, karton, biologi, logam, plastik, kaca hijau, kaca coklat, kaca putih, pakaian, sepatu, baterai, dan sampah residu. Sampel data sampah yang digunakan ditunjukkan pada gambar 3.2.



Gambar 3. 2 Sampel Data Sampah

3.3.4. Penerapan Metode

Diagram alir penerapan algoritma *Convolutional Neural Network* (CNN) dengan menggunakan model arsitektur *MobileNetV2* ditunjukkan pada gambar 3.3.



Gambar 3. 3 Diagram Alir Penerapan Metode

Berikut ini penjelasan dari diagram alir penerapan metode pada gambar 3.3:

3.3.4.1. *Input Dataset*

Tahap pertama yang dilakukan yaitu dengan menginputkan dataset yang sudah dipersiapkan, yaitu

dataset sampah rumah tangga dengan total 15.515 data citra yang dibagi menjadi 12 kelas. Banyak data citra untuk setiap kelasnya ditunjukkan pada tabel 3.1.

Tabel 3. 1 Rincian data citra sampah

No	Kelas	Total Citra
1	Kertas	1050
2	Karton	891
3	Biologi	985
4	Logam	769
5	Plastik	865
6	Kaca hijau	629
7	Kaca coklat	607
8	Kaca putih	775
9	Pakaian	5325
10	Sepatu	1977
11	Baterai	945
12	Sampah	697
Total		15515

3.3.4.2. Tahap *preprocessing*

Pada penelitian ini ada tiga langkah yang dilakukan pada tahap *preprocessing* yaitu pembagian data, melakukan *resizing*, dan augmentasi data. Langkah pertama dataset akan dibagi menjadi dua yaitu data *training* sebesar 90% dan data *testing* sebesar 10% [9]. Setiap pembagian data, masing masing berisi 12 kelas data sampah rumah tangga.

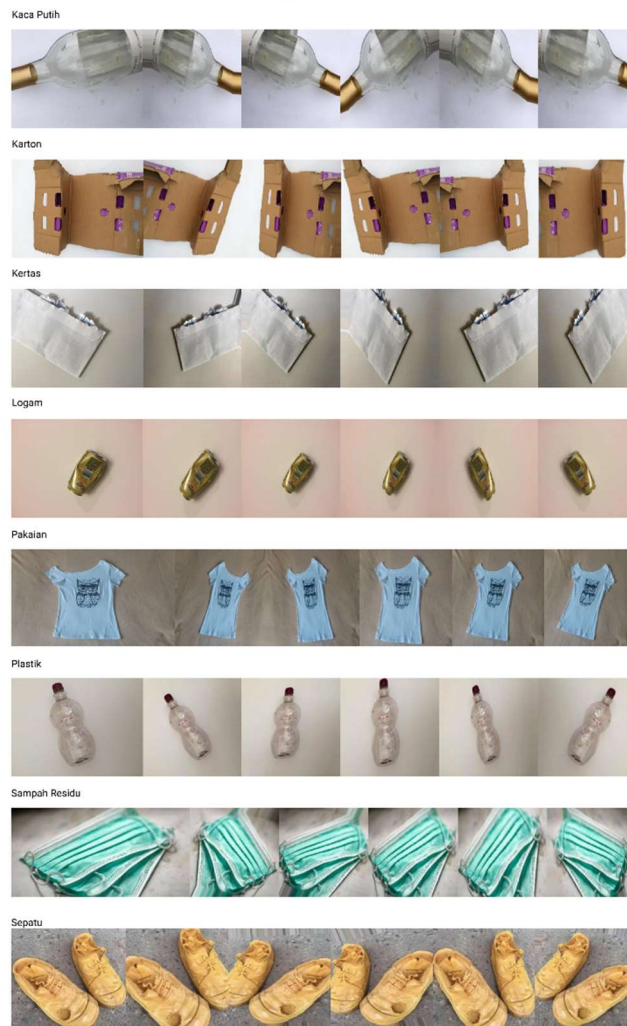
Langkah kedua yaitu melakukan *resizing* dengan mengubah ukuran piksel citra. Resolusi yang disarankan

oleh *mobileNet* adalah 224, 192, 160, atau 128 [45]. Dalam penelitian ini, ukuran pixel citra akan diubah menjadi 224x224 piksel, resolusi tersebut dipilih karena beberapa arsitektur CNN yang terkenal menyarankan ukuran citra (*shape*) tersebut [46].

Langkah ketiga dilakukan augmentasi data untuk memperbanyak data dengan cara *rescaling*, *flipping*, *zooming* dan *rotate*, dengan menggunakan *ImageDataGenerator* yang terdapat pada *library* Keras. Sampel hasil augmentasi ditunjukkan pada gambar 3.4.



Gambar 3. 4 Sampel Data Citra Sampah Hasil Augmentasi

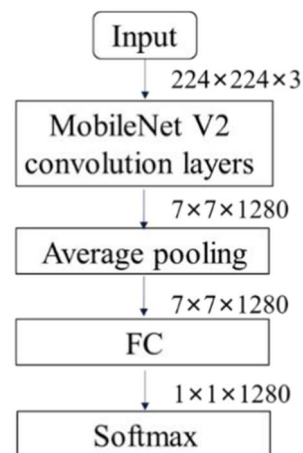


Gambar 3. 4 Sampel Data Citra Sampah Hasil Augmentasi (Lanjutan)

3.3.4.3. Membangun model.

Pada penelitian ini model akan dibangun menggunakan teknik *transfer learning*, dengan menggunakan hasil ekstraksi fitur dan menggunakan *fine tuning* dari model arsitektur *MobileNetV2* (Gambar 3.5) yang sudah dilatih sebelumnya oleh *ImageNet* [47]. Kemudian menambahkan beberapa parameter lainnya,

seperti *Global Average Pooling* yang digunakan untuk menghitung nilai rata-rata dari setiap baris matriks pada hasil konvolusi sebelumnya [48], *flatten* untuk melakukan *reshape feature map* menjadi sebuah *vector* agar bisa digunakan sebagai *input* dari *fully-connected layer* [29], *dropout* untuk mencegah model *overfitting* [1], *dense layer* untuk menggunakan fungsi aktivasi ReLu dan *softmax*. Fungsi aktivasi ReLu digunakan untuk melakukan normalisasi *batch* ketika nilainya negatif maka akan diubah menjadi nol dan ketika tidak negatif nilainya akan tetap [39] , sedangkan fungsi aktivasi *softmax* digunakan sebagai output layer yang akan mengklasifikasi gambar ke dalam banyak kelas [28].



Gambar 3. 5 Arsitektur *MobileNetV2* [14]

Total parameter yang digunakan pada penelitian ini sebanyak 3.582.028. Pada tabel 3.2 ditunjukkan rincian dari arsitektur model yang digunakan. Pada model yang menggunakan hasil ekstraksi fitur MobileNetV2, parameter yang akan digunakan pada saat training sebanyak 1.324.044. Sedangkan pada model yang

menggunakan *Fine Tuning* parameter yang digunakan pada saat training sebanyak 3,547,916.

Tabel 3. 2 Arsitektur model

<i>Layer (type)</i>	<i>Output Shape</i>	<i>Params</i>
<i>MobileNetV2</i>	7,7,1280	2.257.984
3.3.4.4 <i>global_average_pooling2d</i>	1280	0
<i>flatten</i>	1280	0
<i>Dense (Relu)</i>	1024	1311744
<i>dropout</i>	1024	0
<i>Dense (Softmax)</i>	12	12300
<i>Total Params</i>		3.582.028

P

roses *training* yang dilakukan, bertujuan untuk mendapatkan model berdasarkan *input* dan arsitektur yang sudah dibangun. Selama pelatihan, ditambahkan parameter seperti *Adam Optimizer* [40]. *Adam Optimizer* merupakan algoritma optimasi untuk menghitung *learning rate* individu pada parameter yang berbeda, *learning rate* yang direkomendasikan sebesar 0.001 [49]. Parameter *epoch* yang digunakan adalah 10, 20 dan 30 pada setiap model untuk melihat performa akurasi dan loss [39].

3.3.4.5. *Testing*

Testing model dilakukan untuk memastikan model dapat memprediksi dengan baik [20]. Pengujian (*testing*) dilakukan menggunakan data *testing* yang sudah disiapkan.

3.3.4.6. Evaluasi

Evaluasi model yang dilakukan menggunakan *confusion matrix* untuk mengetahui performa algoritma klasifikasi yang digunakan dengan melihat nilai akurasi, presisi, *recall* dan waktu pelatihan [1]. Untuk mencari nilai akurasi, presisi, *recall*, dan *F1-Score* dapat dihitung menggunakan persamaan sebagai berikut:

$$\text{Akurasi} = \frac{TP+TN}{TP+TN+FP+F} * 100\% \quad (3,1)$$

$$\text{Presisi} = \frac{TP}{FP+TP} * 100\% \quad (3,2)$$

$$\text{Recall} = \frac{TP}{TP+FN} * 100\% \quad (3,3)$$

$$\text{F1 - Score} = \frac{2 * \text{Presisi} * \text{Recal}}{\text{Presisi} + \text{Recal}} * 100\% \quad (3,4)$$

Pada nilai *True Negative* (TN) merupakan jumlah data negatif yang terklasifikasi dengan benar. Nilai *False Positive* (FP) merupakan data negatif namun terklasifikasi sebagai data positif. Nilai *True Positive* (TP) merupakan data positif yang terklasifikasi benar. Nilai *False Negative* (FN) merupakan data positif namun terklasifikasi sebagai data negatif.

3.3.5. Kesimpulan dan Saran

Pada tahap ini akan akan diambil kesimpulan untuk mengambil inti dari penelitian yang dilakukan untuk dijadikan pemahaman yang utuh. Kemudian saran yang diberikan digunakan sebagai *referensi* penelitian selanjutnya yang berhubungan dengan masalah yang sama, yang selanjutnya ditulis pada bab kesimpulan dan saran.