

BAB 3

METODOLOGI PENELITIAN

Penelitian ini menggunakan jenis penelitian kuantitatif, dimana penelitian ini didasarkan pada perhitungan analisis statistik yang digunakan untuk menghitung serta mengetahui hubungan antar variabel.

3.1 Kebutuhan Sistem

Pada bagian ini dijelaskan mengenai perangkat lunak serta perangkat keras yang digunakan untuk penelitian.

3.1.1 Perangkat Keras

- Laptop LENOVO_MT_80XG_BU_idea_FM_ideapad 320-14ISK, dengan spesifikasi:
 - Sistem Operasi: Microsoft Windows 10
 - Processor: Intel Core i3-6006U
 - RAM: 4 GB
- Kamera handphone POCO M4 PRO, dengan spesifikasi:
 - Kamera 64MP, 2MP, macro
 - *Auto focus*

3.1.2 Perangkat Lunak

- Sistem Operasi Windows 10
- Google Collab
- Tensorflow
- Keras

3.2 Gambaran Umum Objek Penelitian



Gambar 3.1. Jenis-jenis sampah
sumber : [22]

Gambar 3.1 merupakan gambar dari jenis-jenis sampah yang merupakan objek utama pada penelitian ini. Jenis-jenis sampah yang akan diteliti adalah sampah organik dan sampah anorganik. Sampah organik adalah sampah yang mudah terurai oleh mikroorganisme, sedangkan sampah anorganik adalah sampah yang sulit diuraikan oleh mikroorganisme. Sehingga kedua sampah tersebut harus melalui pemilahan untuk memudahkan proses pembuangan dan pengolahan kembali.

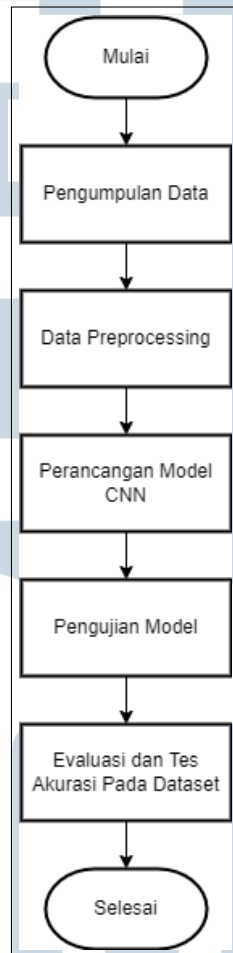
Seperti yang kita ketahui bahwa sampah dihasilkan oleh kegiatan manusia setiap harinya. Jika sampah tersebut tidak melalui proses pemilahan, maka dapat mengakibatkan penumpukan sampah yang akan menjadi sarang bakteri dimana bakteri merupakan penyebab utama dari penyakit. Hal tersebut yang menjadikan sampah sebagai objek utama pada penelitian ini, dengan harapan sistem yang dibangun dapat membantu masyarakat dalam proses pemilahan sampah.

3.3 Jumlah Populasi dan Sampel Pada Penelitian

Pada penelitian ini jumlah populasi dan sampel yaitu jumlah keseluruhan gambar sampah yang telah dikategorikan berdasarkan jenisnya. Terdapat 2 jenis sampah yang akan diteliti yaitu sampah organik dan sampah anorganik. Gambar sampah yang telah dikumpulkan kemudian disortir sebanyak 22.564 gambar un-

tuk data *training* dan 2513 untuk data *test* dari setiap jenisnya. Sehingga jumlah keseluruhan data yang digunakan adalah sebanyak 25.077 gambar sampah.

3.4 Teknik Analisis Data



Gambar 3.2. Alur tahapan analisis data

Pada Gambar 3.2 tersebut merupakan tahapan yang digunakan pada penelitian ini. Pada penelitian ini proses analisis data menggunakan salah satu metode pada *Deep Learning* yang sudah teruji untuk mengklasifikasikan gambar dengan baik dan akurat. Metode yang digunakan adalah Convolutional Neural Network (CNN) dengan menggunakan *software* Google Colab *package* Keras dan Tensor-Flow.

Algoritma CNN akan mengklasifikasikan gambar sampah dengan operasi konvolusi hingga membentuk sebuah pola. Pada CNN terdapat banyak lapisan

karena merupakan algoritma *Deep Learning*, lapisan-lapisan yang terdapat pada arsitektur CNN adalah lapisan *convolution*, *max and pooling*, *flatten*, *dense layer*, dan *output*. Pada seluruh lapisan tersebut dibagi menjadi dua tahapan yaitu *feature learning* dan *classification*. Pembangunan arsitektur CNN yang baik akan berdampak pada hasil klasifikasi gambar yang diperoleh.

- **Pengumpulan Data**

Proses pengumpulan data gambar dilakukan dengan cara mengunduh dari situs web Kaggle. Kaggle merupakan sebuah situs untuk berbagi ide dan aspirasi antar *data scientist*, dimana pada situs tersebut terdapat berbagai macam kumpulan dataset dengan format pada umumnya adalah csv. Dataset yang diunduh yaitu gambar sampah organik dan anorganik yang berjumlah 25.077 yang terdiri dari data *training* sebanyak 22.564 dan data *test* sebanyak 2.513. Selain itu pengumpulan data dilakukan dengan mengambil gambar sampah secara langsung dari lapangan sebanyak 100 data gambar yang terdiri dari 50 gambar sampah organik dan 50 gambar sampah anorganik.

- **Data Preprocessing**

Pada tahap *preprocessing* ini terdiri dari empat tahapan utama yaitu *grayscale*, *thresholding*, segmentasi, dan *resize*. Kemudian data gambar pada dataset akan diolah pada *ImageGenerator* yang terdapat pada *library* Keras.

1. *Grayscale*

Grayscale merupakan gambar yang hanya mempunyai satu *channel* warna, sehingga warna yang ditampilkan hanya derajat keabuan atau nilai intensitasnya saja. Nilai intensitas ini berkisar antara nilai 0 hingga 255, dimana nilai 0 merupakan warna hitam dan nilai 255 adalah putih. Proses *rescale* akan mengecilkan *range* warna gambar menjadi 0 hingga 255 dan memudahkan proses *thresholding* citra menjadi citra biner. Berikut ini merupakan rumus untuk menentukan nilai *grayscale*.

$$y = (a * R) + (b * G) + (c * B) \quad (3.1)$$

*Keterangan

- y = Nilai *pixel grayscale*
- a,b,c = nilai *array* pada setiap *pixel*
- R = Nilai *pixel red*

- G = Nilai *pixel green*
- B = Nilai *pixel blue*

2. *Thresholding*

Thresholding merupakan proses yang akan menghasilkan citra biner yang berasal dari citra *grayscale* atau citra yang memiliki *channel* warna RGB kemudian diatur nilai pikselnya menjadi 0 atau 1 tergantung ambang batasnya apakah ada diatas atau dibawah. Proses *thresholding* dibedakan menjadi dua yaitu *global thresholding* dan *local thresholding*. *Global thresholding* merupakan tahapan mengkonversi citra menjadi hitam putih dimana nilai ambang bergantung pada satu nilai *thresholding*. Sedangkan *local thresholding* bergantung pada dua nilai *thresholding*.

3. *Segmentation*

Kemudian citra hasil dari *thresholding* yang berupa citra biner akan dilakukan pemotongan untuk obyek sampah yang akan diprediksi. Pemotongan ini dilakukan pada setiap baris atau secara vertikal terlebih dahulu, kemudian dilanjutkan pemotongan pada setiap kolom atau secara horizontal.

4. *Resize*

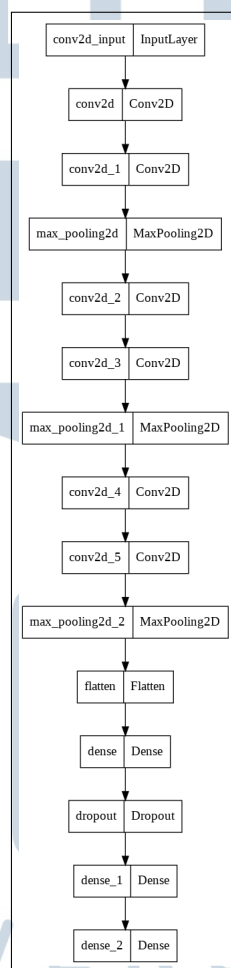
Proses *resize* adalah mengubah ukuran dari gambar menjadi lebih besar atau lebih kecil dari ukuran semula. Proses *resize* ini dilakukan dengan menentukan ukuran yang diinginkan, dimana pada penelitian ini ukuran *input* gambar diatur menjadi 224x224 piksel.

- Perancangan Model CNN

Proses perancangan model CNN merupakan tahap penyusunan sebuah model yang akan digunakan untuk pelatihan data dalam mengenali objek yang dimasukkan. Model yang akan dirancang terdiri dari jumlah layer yang digunakan, penentuan filter yang digunakan, penentuan ukuran kernel, dan penentuan fungsi aktivasi serta ukuran *pooling*.

- Pelatihan Model

Gambar 3.3 berikut merupakan arsitektur yang terbentuk untuk proses pelatihan model yang dilakukan, dimana terdapat 6 tahap konvolusi dan 3 tahap *pooling* dengan menggunakan *max pooling*. Lalu kemudian data multidimensional *array* diubah menjadi vektor tunggal dengan proses *flatten* untuk dimasukkan ke dalam *fully connected layer*. Namun sebelum dimasukkan, akan dilakukan fungsi *dropout* untuk mencegah terjadinya *overfitting* dengan mengurangi neuron secara acak.



Gambar 3.3. Tahapan proses pelatihan model

Setelah tahap perancangan model CNN berhasil dilakukan, tahap selanjutnya adalah proses pelatihan model dengan menggunakan data latih yang telah terkumpul sebelumnya. Proses pelatihan ini dilakukan untuk mendapatkan nilai akurasi yang tinggi dari proses klasifikasi yang dilakukan. Kemudian pada proses pengujian model digunakan sebuah iterasi untuk menentukan be-

rapa kali jaringan akan melakukan proses *training*. Pada proses pelatihan ini terdapat fungsi loss yang digunakan untuk mengukur performa yang dihasilkan oleh model yang telah dirancang dalam memprediksi target. Fungsi loss ini akan bekerja saat terdapat kesalahan yang harus diperhatikan [23].

- Evaluasi dan Pengujian Tingkat Akurasi

Tahap yang terakhir adalah proses pengujian model pada data *test*. Proses ini dilakukan untuk menguji tingkat akurasi yang dihasilkan dari model Convolutional Neural Network (CNN). Tingkat akurasi yang dihasilkan menunjukkan tingkat kebenaran dari proses klasifikasi jenis sampah. Dengan hasil akurasi yang tinggi menunjukkan bahwa model dapat melakukan klasifikasi jenis sampah dengan baik dan akurat. Evaluasi model CNN dilakukan menggunakan *confusion matrix* untuk mengetahui seberapa baik hasil model yang dibangun. *Confusion Matrix* digambarkan dalam bentuk tabel matriks yang menjelaskan kinerja model klasifikasi pada data *test*. Gambar 4.13 di bawah ini merupakan gambaran tabel dari confusion matriks dengan 4 kombinasi prediksi serta nilai aktual yang berbeda.

		Actual Values	
		1 (Positive)	0 (Negative)
Predicted Values	1 (Positive)	<p>TP (True Positive)</p>	<p>FP (False Positive) <small>Type I Error</small></p>
	0 (Negative)	<p>FN (False Negative) <small>Type II Error</small></p>	<p>TN (True Negative)</p>

Gambar 3.4. Tabel confusion matriks

sumber: [24]

Berdasarkan Gambar 4.13 tersebut terdapat empat istilah yang merupakan representasi dari hasil klasifikasi pada *confusion matrix*. Empat istilah tersebut yaitu *True Positive (TP)*, *True Negative (TN)*, *False Positive (FP)* dan *False Negative (FN)*. Diambil dari studi kasus klasifikasi jenis sampah ini, berikut merupakan penjelasan dari keempat istilah tersebut:

1. *True Positive (TP)*

Adalah data positif yang diprediksi benar. Misalnya data sampah organik (class 1) dan model yang telah dibuat memprediksi data sampah tersebut tergolong jenis sampah organik (class 1).

2. *True Negative (TN)*

Adalah data negatif yang diprediksi benar. Misalnya data sampah anorganik (class 2) dan model yang telah dibuat memprediksi data sampah tersebut tergolong jenis sampah anorganik (class 2).

3. *False Positive (FP) — Type I Error*

Adalah data negatif namun diprediksi sebagai data positif. Misalnya data sampah anorganik (class 2) akan tetapi pada model yang telah dibuat memprediksi data tersebut tergolong dalam jenis sampah organik (class 1).

4. *False Negative (FN) — Type II Error*

Adalah data positif namun diprediksi sebagai data negatif. Misalnya data sampah organik (class 1) namun pada model yang dibuat memprediksi sampah tersebut tergolong dalam jenis sampah anorganik (class 2).

Untuk menghitung performa dari matriks agar dapat diketahui kinerja model yang telah dibangun, terdapat beberapa *performance metrics* yang sering digunakan yaitu *accuracy*, *precision*, and *recall*.

1. *Accuracy*

Nilai *Accuracy* merupakan gambaran seberapa akurat model dapat mengklasifikasikan gambar dengan benar. Sehingga *accuracy* ini merupakan rasio prediksi benar secara positif dan negatif dari keseluruhan dataset yang tersedia. Nilai *accuracy* dapat diperoleh dengan Rumus 3.2 berikut.

$$\frac{TP + TN}{TP + TN + FP + FN} \quad (3.2)$$

*Keterangan

- TP = *True Positive*
- TN = *True Negative*

- FP = *False Positive*
- FN = *False Negative*

2. *Precision*

Nilai *precision* merupakan gambaran dari tingkat keakuratan antara data yang diminta dengan hasil prediksi yang dihasilkan oleh model. Sehingga *precision* merupakan rasio prediksi *true positive* dibandingkan dengan keseluruhan hasil yang diprediksi positif. Nilai dari *precision* dapat diperoleh dengan Rumus 3.3 berikut.

$$\frac{TP}{TP + FP} \quad (3.3)$$

*Keterangan

- TP = *True Positive*
- FP = *False Positive*

3. *Recall*

Nilai *recall* merupakan gambaran keberhasilan model dalam mengambil informasi. Sehingga *recall* adalah rasio prediksi *true positive* dibandingkan dengan keseluruhan data yang *true positive*. Nilai *recall* dapat didapatkan dengan menggunakan Rumus 3.4 berikut.

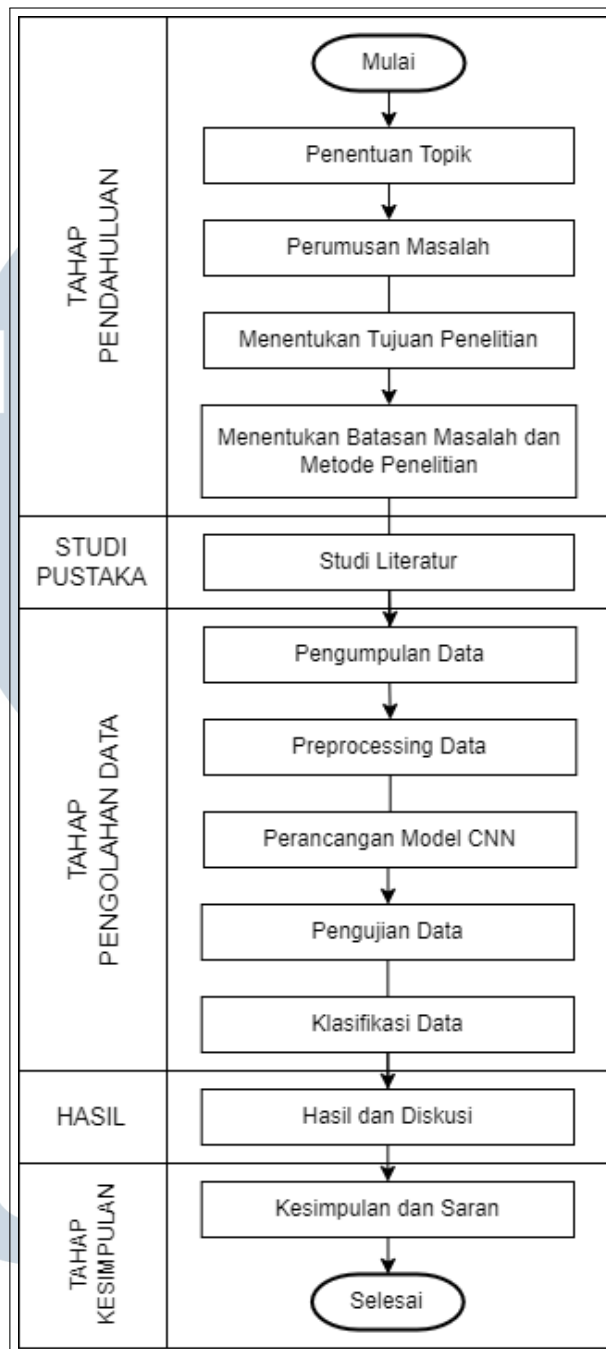
$$\frac{TP}{TP + FN} \quad (3.4)$$

*Keterangan

- TP = *True Positive*
- FN = *False Negative*

3.5 Kerangka Berpikir

Kerangka berpikir pada penelitian ini diadaptasi dari penelitian terdahulu yang kemudian disesuaikan dengan kebutuhan pada penelitian ini. Gambar 3.5 berikut merupakan tahapan penelitian yang dilakukan dari tahap awal hingga akhir.



Gambar 3.5. Alur kerangka berpikir

Berdasarkan *flow diagram* pada Gambar 3.5, diketahui terdapat 5 tahapan penelitian yang dilakukan diantaranya adalah.

1. Pendahuluan merupakan tahapan awal pada penelitian ini, dimana langkah pertama yang dilakukan yaitu menentukan topik penelitian yang diambil berdasarkan permasalahan yang terjadi di lingkungan sekitar. Kemudian

permasalahan yang terjadi tersebut diidentifikasi oleh peneliti untuk menentukan pokok permasalahan yang dapat diselesaikan. Setelah merumuskan sebuah permasalahan, kemudian ditentukan sebuah tujuan penelitian yang disesuaikan dengan pokok permasalahan yang terjadi dengan harapan tujuan penelitian ini mampu menjawab setiap rumusan masalah. Selain itu penting juga ditentukan sebuah batasan masalah dan menentukan metodologi penelitian agar penelitian yang dilakukan dapat terarah pada tujuan penelitian.

2. Tahap selanjutnya adalah studi literatur yang merupakan tahapan dimana peneliti mencari dan memperbanyak sumber referensi pada bidang yang diteliti. Dengan berpacu pada penelitian terdahulu memudahkan peneliti untuk mengembangkan metode yang digunakan.
3. Selanjutnya adalah tahap pengolahan data yang dimulai dari proses pengumpulan data sesuai dengan topik penelitian. Kemudian data yang sudah dikumpulkan dipersiapkan dan diperbaiki sebelum diproses pada algoritma, langkah ini disebut *preprocessing data*. Setelah data sudah siap maka tahap selanjutnya adalah merancang arsitektur CNN, pada umumnya metode CNN ini memiliki 2 tahapan diantaranya adalah tahap *feature learning* dan *classification*. Setelah didapatkan sebuah arsitektur yang terbaik, selanjutnya dilakukan pengujian gambar menggunakan gambar yang telah melalui tahap *preprocessing*. Dari proses pengujian gambar tersebut akan didapatkan hasil akurasi yang terbaik, dimana hasil akurasi tertinggi akan digunakan untuk memproses data uji. Pada akhirnya akan diperoleh hasil akhir untuk pengujian data uji yang digunakan sebagai penentu dari kemampuan CNN untuk mengklasifikasikan jenis-jenis sampah.
4. Tahap keempat yaitu menjabarkan hasil dari tahap pengolahan data menggunakan metode CNN dan melakukan uji coba terhadap sistem yang dibangun. Kemudian hasil yang diperoleh dari model diuraikan dan dianalisis.
5. Tahap yang terakhir yaitu menyimpulkan hasil dari keseluruhan proses penelitian yang telah dilakukan. Kesimpulan dari hasil sesuai dengan tujuan dan rumusan masalah yang telah dipaparkan pada tahapan awal. Selain itu peneliti juga memberikan saran sebagai bahan pertimbangan untuk kekurangan pada penelitian ini agar diperbaiki atau dikembangkan pada penelitian selanjutnya.