

BAB III

METODE PENELITIAN

3.1 Waktu dan Lokasi Penelitian

Penelitian ini dilakukan di *head office* PT Berau Coal pada divisi *mining technology*, terkhusus pada *project mining eyes analytics*. Pengambilan data dilakukan pada prakerin 1 pada tanggal 03 Oktober 2022 - 31 Desember 2022.



Gambar 3. 1 Lokasi PT Berau Coal

Nama Perusahaan : *Head Office* PT Berau Coal
 Alamat : JL. Pemuda, No. 40 Tanjung Redeb, Kabupaten Berau, Kalimantan Timur.
 No. Telepon : (62-554) 23400 – (62-21) 7264 778
 No. Fax : (62-554) 23465 – (62-21) 7268 289
 Email : corporate.communication@beraucoal.co.id

3.2 Subjek dan Objek Penelitian

1. Subjek Penelitian

Dalam penelitian ini yang menjadi subjek penelitian adalah data hasil anotasi yang sudah dilakukan pada Praktik Kerja Industri pertama, data hasil *training* melalui *google colaboratory*, dan

2. Objek Penelitian

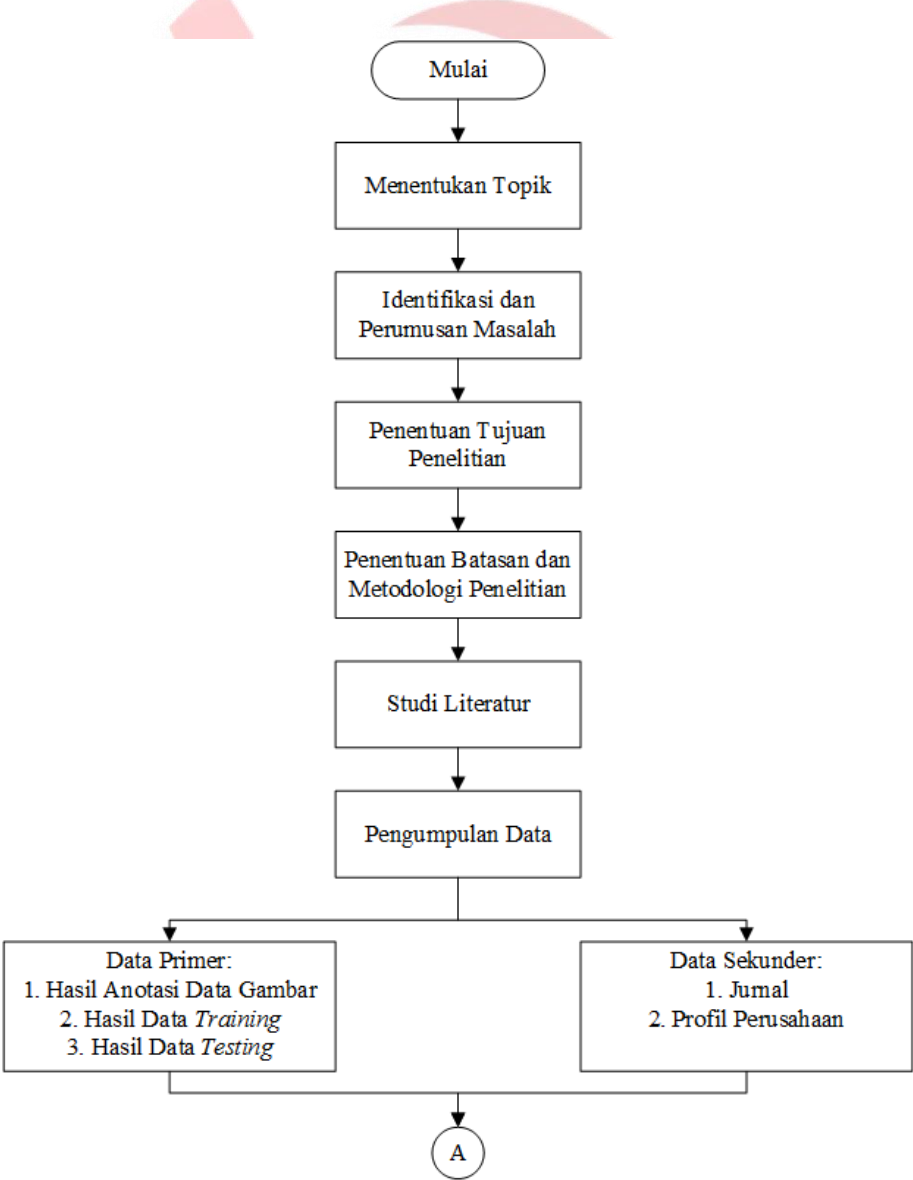
Dalam penelitian ini yang menjadi objek penelitian adalah tingkat akurasi dari beberapa kategori (*classes*) yang sudah menjadi ketentuan pada data hasil anotasi yang dilakukan sebelumnya. Tujuannya untuk mengetahui perbandingan tingkat akurasi dari 3 kategori dan 6 kategori.

3.3 Jenis Penelitian

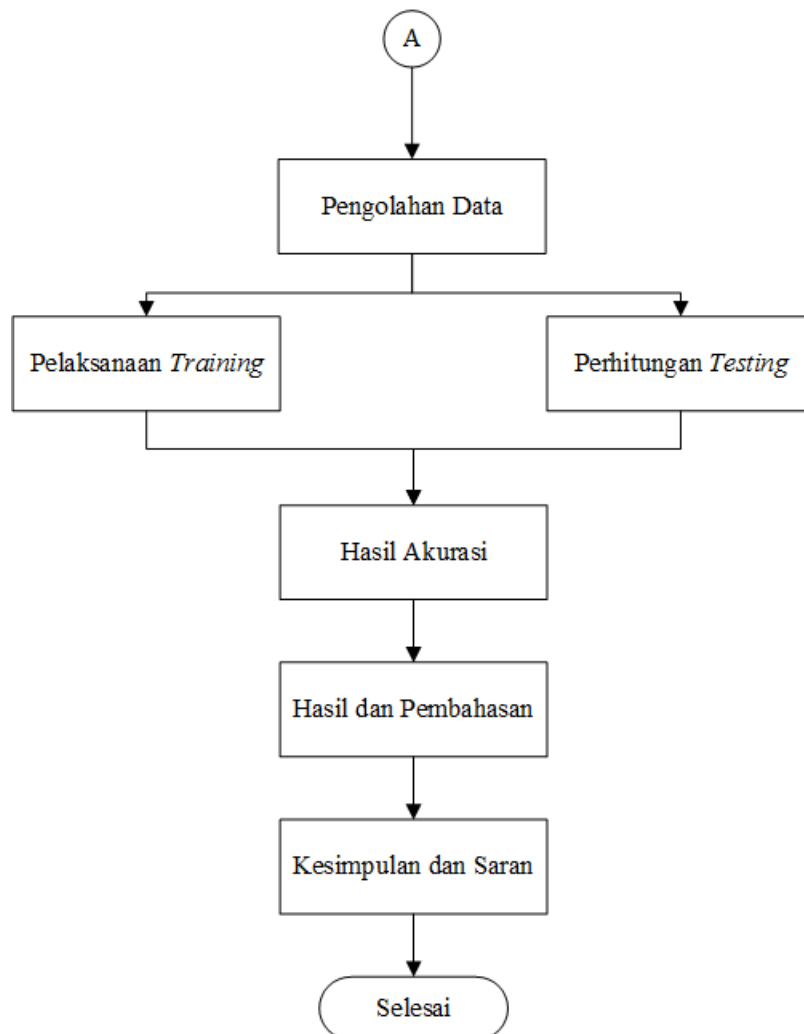
Dalam penelitian ini menggunakan metode kuantitatif yang bertujuan untuk mengetahui tingkat akurasi yang didapatkan dengan melakukan perbandingan dua kategori yaitu 3 kategori dan 6 kategori. Dari hasil akurasi yang didapat maka selanjutnya akan dilakukan perbandingan nilai akurasi yang tertinggi dari dua kategori tersebut. Tujuannya untuk mengetahui seberapa besar pengaruh kategori terhadap akurasi *machine learning*.

3.4 Diagram Alir Penelitian

Tahapan proses yang akan dilakukan dalam penelitian ini digambarkan dalam diagram alir pada gambar 3.1 berikut ini:



Gambar 3. 2 Diagram Alir Penelitian



Gambar 3.3 Lanjutan Diagram Alir Penelitian

Penelitian ini memiliki alur yang menunjang jalannya penelitian pada tahap awal peneliti menentukan topik, melakukan identifikasi serta merumuskan masalah, menentukan tujuan dari penelitian, dan menentukan batasan dan metodologi penelitian. Tahap kedua yaitu melakukan studi literatur yaitu mencari literatur terkait penelitian yang akan diangkat. Tahap ketiga yaitu pengumpulan dan pengolahan data. Dalam pengumpulan data, data primer yang digunakan merupakan hasil anotasi data gambar dan hasil data *training* yang akan menjadi objek penelitian dan data sekunder yaitu profil perusahaan serta jurnal penunjang penelitian. Tahap keempat pengolahan data yaitu menentukan sampel, melakukan *training* serta menghitung akurasi per kategori (*classes*) atau biasa disebut *testing*. Tahap kelima pembahasan mengenai hasil yang didapat dari pengolahan data sebelumnya. Kemudian dari hasil dan pembahasan akan didapatkan kesimpulan dan saran dari penelitian tersebut.

3.5 Populasi dan Sampel

1. Populasi

Populasi dalam penelitian disebut suatu kumpulan yang memenuhi syarat dengan masa penelitian. Populasi adalah sekumpulan orang, hewan, tumbuhan, atau benda yang memiliki karakteristik tertentu yang akan diteliti (Mulyatiningsih dalam Ariyani, 2022). Populasi pada penelitian ini adalah berupa data gambar hasil anotasi berjumlah 3.800 gambar. Dari sekian banyak gambar yang telah dipilih maka hanya beberapa gambar yang akan *ditraining*kan sesuai sampel dengan dua kategori (*classes*) yaitu 3 kategori dan 6 kategori untuk dapat diteliti lebih lanjut.

2. Sampel

Sampel merupakan bagian terkecil dari populasi yang diambil sebagai objek penelitian karena dianggap mampu mewakili dari populasi yang ada. Sampel adalah bagian dari jumlah dan karakteristik yang dimiliki oleh suatu populasi. Pengukuran sampel dilakukan melalui statistik (Sugiyono dalam Ariyani, 2022). Sampel pada penelitian ini adalah data gambar dengan beberapa kategori (*classes*) yaitu 3 kategori dan 6 kategori yang kemudian diambil untuk dilakukan proses *Training*. Dan untuk proses *Training* akan dilakukan di *google colab* dengan bantuan dari Tim *developer* sub divisi *mining technology* PT Berau Coal.

Penelitian ini menggunakan teknik sampling data slovin untuk mencari sampel yang akan digunakan dengan proses *training* untuk mengetahui tingkat akurasi dari data yang telah didapatkan. Rumus slovin adalah sebuah rumus yang dipakai untuk mencari jumlah minimum sampel yang didapatkan dari hasil survei dengan jumlah populasi yang terbatas atau yang dikenal dengan istilah *finite population survey*. Ada beberapa keuntungan yang didapatkan dengan menggunakan rumus Slovin. Pertama, rumus ini sangat mudah digunakan dan diaplikasikan dalam penelitian. Kedua, dengan menggunakan rumus ini, dapat menghemat waktu dan biaya dalam pengambilan sampel. Ketiga, hasil yang diperoleh dari sampel yang diambil dengan menggunakan rumus Slovin cenderung lebih akurat dan representatif (Administrator, 2023). Rumus Slovin dapat memberikan gambaran kasar untuk menentukan jumlah sampel. Namun, rumus non-parametrik ini tidak memiliki ketelitian matematis. Misalnya, tidak ada cara untuk menghitung power statistik (yang memberikan informasi seberapa besar kemungkinan penelitian membedakan efek aktual) (Ryan dalam Akhtar, 2020)

Menurut Sugiyono (2017) adalah sebuah rumus yang digunakan untuk mendapatkan besaran sampel yang dianggap mampu menggambarkan keseluruhan populasi yang ada. Rumus slovin digunakan untuk memperoleh

jumlah sampel dari suatu survei, yang tujuannya untuk mengestimasi proporsi. Proporsi inilah yang kemudian menjadi dasar dari penghitungan varian. Berikut merupakan ketentuan perhitungan rumus slovin:

$$n = \frac{N}{1+N(e)^2} \dots\dots\dots (1)$$

Keterangan:

n = Jumlah sampel

N = Jumlah populasi

e^2 = Batas kesalahan (*margin of error*)

Dalam penelitian ini jumlah populasi yang digunakan sebesar 3800, *margin of error* yang digunakan sebesar 5%, maka jumlah sampel yang didapatkan dapat dilihat pada hasil perhitungan dibawah ini:

$$n = \frac{3800}{1 + 3800(0,05)^2}$$

$$n = \frac{3800}{1 + 3800(0,0025)}$$

$$n = \frac{3800}{1 + 9,5}$$

$$n = 361,90 \approx 362$$

Jadi, sampel yang digunakan dalam penelitian ini sebanyak 362 data gambar yang sudah dianotasi sebelumnya.

3.6 Teknik Analisis Data

Analisa data yang digunakan dalam penelitian ini yaitu teknik analisis inferensial. Hasil yang diperoleh dari perhitungan tersebut digunakan sebagai dasar untuk membuat kesimpulan yang berlaku secara umum. Perhitungan yang digunakan yaitu menggunakan perhitungan akurasi. Salah satu perhitungan akurasi yaitu *confusion matrix*. *Confusion matrix* memberikan informasi perbandingan hasil klasifikasi sebenarnya. *Confusion matrix* berbentuk tabel matriks yang menggambarkan kinerja model klasifikasi pada serangkaian data uji yang nilai sebenarnya diketahui.

Terdapat 4 istilah dari tabel matriks sebagai representasi hasil proses klasifikasi pada *confusion matrix*. Keempat istilah tersebut adalah *True Positive* (TP), *True Negative* (TN), *False Positive* (FP) dan *False Negative* (FN). Berikut merupakan penjelasan dari keempat istilah tersebut:

- 1. *True Positive* (TP) : Merupakan data positif yang diprediksi benar.
- 2. *True Negative* (TN) : Merupakan data negatif yang diprediksi benar.
- 3. *False Positive* (FP) : Merupakan data negatif namun diprediksi positif.
- 4. *False Negative* (FN) : Merupakan data positif namun diprediksi negatif.

Adapun *confusion matrix* dengan 4 kombinasi nilai prediksi dan nilai aktual yang berbeda dapat dilihat pada gambar 3.2

		Actual Values	
		1 (Positive)	0 (Negative)
Predicted Values	1 (Positive)	TP (True Positive)	FP (False Positive) Type I Error
	0 (Negative)	FN (False Positive) Type III Error	TN (True Negative)

Gambar 3. 4 Tabel *Confusion Matrix*

Confusion matrix digunakan untuk menghitung berbagai *performance metrics* untuk mengukur kinerja model yang telah dibuat. *Performance metrics* populer yang umum dan sering digunakan: *accuracy*, *preciission*, dan *recall*. Dalam penelitian ini akan menggunakan *performance metrics recall*. *Recall* menggambarkan keberhasilan model dalam menemukan kembali sebuah informasi. Maka, *recall* merupakan rasio prediksi benar positif dibandingkan dengan keseluruhan data yang benar positif. Nilai *recall* dapat diperoleh dengan persamaan dibawah ini:

$$Recall = \frac{TP}{TP+FN} \dots\dots\dots (2)$$

Keterangan:
TP = *True Positive*
FN = *False Negative*

3.7 Teknik Pengumpulan Data

Terdapat teknik pengumpulan data pada penelitian ini yaitu sebagai berikut:

1. Studi Literatur

Dalam penelitian ini terdapat studi literatur yaitu berkaitan dengan metode pengumpulan data pustaka, membaca dan mencatat, serta mengelola bahan penelitian. Pengumpulan data pustaka dapat diambil dari jurnal, buku, artikel, maupun yang lainnya.

2. Pengumpulan Data Penelitian

Pengumpulan data penelitian ini dilakukan dengan mengumpulkan data gambar dan data gambar yang telah diberi *bounding box*. Kemudian mengumpulkan data dari hasil proses *training* dan proses *testing* yang telah dilakukan dari *google colab*.

3. Pengujian

Pengujian pada penelitian ini dilakukan dengan melihat akurasi dari beberapa kategori. Hasil tingkat akurasi dari 3 kategori dan 6 kategori diperoleh dengan melakukan proses *training* di *google colab*, proses *testing*, dan melakukan perhitungan manual menggunakan rumus *recall*. Hasil akurasi tersebut akan dibuat perbandingan, dari dua kategori tersebut nilai akurasi dari *machine learning* mana yang lebih baik.

4. Analisis Hasil Pengujian

Penelitian ini dilakukan dengan analisis perbandingan hasil dari proses *training* dan hasil proses *testing* yang sudah dilakukan pada *google colab*. Dengan membandingkan hasil kinerja model dalam menangkap informasi yang sudah dipelajari dan membandingkan hasil kinerja model dalam mendeteksi objek.

5. Dokumentasi

Melakukan pengambilan gambar proses *training*, beberapa sample gambar yang teranotasi, serta data-data penunjang lainnya yang dapat menggambarkan proses kerja dari penelitian ini.