

BAB III

METODE PENELITIAN

Pada bab ini, akan dijelaskan mengenai metodologi penelitian yang digunakan dalam Tugas Akhir ini. Berdasarkan latar belakang yang telah dijelaskan pada bab 1, diketahui bahwa penelitian ini menggunakan subjek berupa dataset jenis batu permata yang terdiri dari citra yang diperoleh melalui citra yang diunduh dari website *Kaggle*. Objek dari penelitian ini adalah klasifikasi jenis batu permata.

3.1 ALAT YANG DIGUNAKAN

Pada penelitian ini membutuhkan beberapa perangkat diantaranya yaitu perangkat keras (*hardware*), perangkat lunak (*software*), dan dataset. Berikut adalah perangkat keras (*hardware*) dan perangkat lunak (*software*) yang akan digunakan disebutkan dalam Tabel 3.1.

Tabel 3. 1 Perangkat keras dan perangkat lunak.

No.	Alat	Spesifikasi	Jumlah
1	Laptop	•RAM 8 GB •Intel(R) Core(TM) i3-3217U CPU @ 1.9GHz (8 CPUs), ~1.8GHz •NVIDIA GeForce 820M	1
2	Operating System	Windows 10 Home 64-bit	1
3	<i>Visual Studio Code</i>	1.79.2	1
4	<i>Python Preprocessing Library</i>	• <i>Numpy</i> • <i>Matplotlib</i>	1

3.1.1 Perangkat Keras (*Hardware*)

Perangkat keras yang digunakan dalam proses penelitian ini adalah sebuah Personal Computer (PC) dengan spesifikasi sebagai berikut :

- 1) Intel core i3
- 2) Windows 10 Pro 64 – Bit
- 3) RAM 8 GB

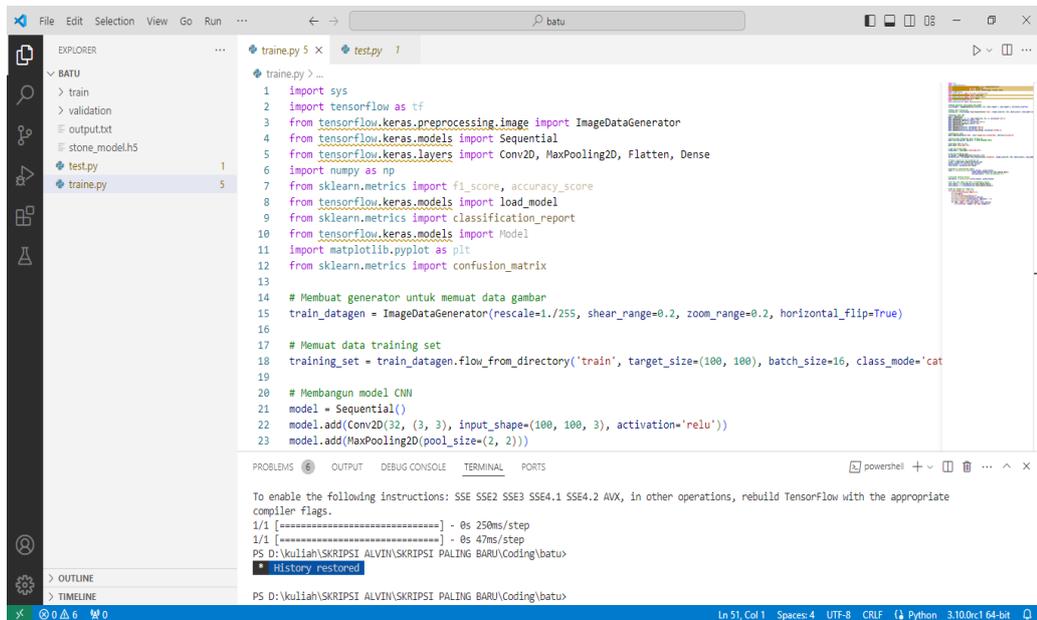
4) Memory Harddisk

3.1.2 Perangkat Lunak (*Software*)

Perangkat lunak (*software*) yang digunakan dalam sistem ini adalah :

1) *Visual Studio Code* versi 1.79.2

Visual Studio Code adalah editor kode sumber yang ringan namun kuat yang berjalan di desktop dan tersedia untuk *Windows*, *macOS* dan *Linux*. Muncul dengan dukungan *built-in* untuk *JavaScript*, *TypeScript* dan *Node.js* dan memiliki ekosistem ekstensi yang kaya untuk bahasa dan runtime lain seperti (*C++*, *C #*, *Java*, *Python*, *PHP*, *Go*, *.NET*). Pada penelitian ini menggunakan *Visual Studio Code* versi 1.79.2. Tampilan aplikasi dapat dilihat pada Gambar 3.1



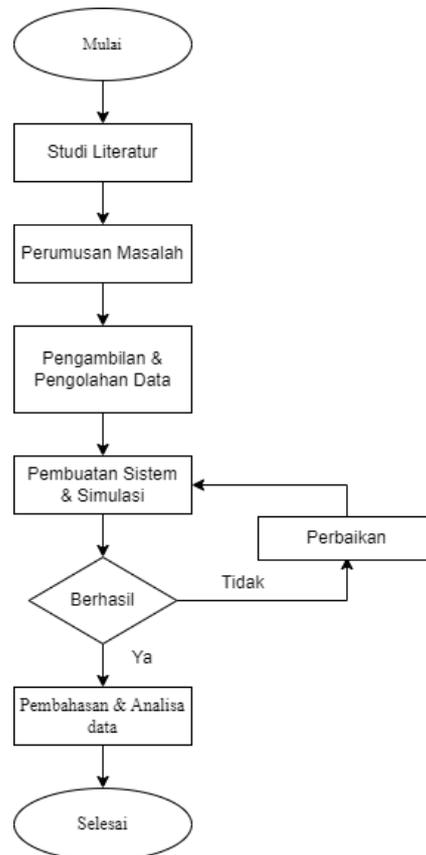
Gambar 3.1 Tampilan *Visual Studio Code*

2) Bahasa Pemrograman *Python*

Python adalah bahasa pemrograman populer yang dapat digunakan untuk berbagai hal, mulai dari membangun website, menganalisis data, *scripting*, hingga pembuatan game. *Python* adalah bahasa pemrograman yang bersifat *open source*, sehingga dapat menggunakannya secara gratis. *Python* menjadi salah satu opsi tepat untuk membuat sistem *machine learning* karena dinilai sederhana, mudah dimengerti, fleksibel, efisien, dan multi-implementasi. Pada penelitian ini menggunakan *Python* versi 2024.4.1.

3.2 ALUR PENELITIAN

Penelitian dilakukan dalam beberapa tahap yaitu tahap perancangan sistem, tahap pembuatan simulasi, tahap pengujian simulasi, dan yang terakhir adalah tahap analisis dari hasil pengujian simulasi. Diagram alur penelitian yang akan dilakukan penulis dapat dilihat pada Gambar 3.2.



Gambar 3. 2 Diagram Alur Penelitian

Berdasarkan pada Gambar 3.2 yang merupakan diagram alur penelitian maka peneliti akan menjelaskan proses setiap tahap pada poin-poin berikut.

3.2.1 Studi Literatur

Penelitian ini dilakukan dengan meninjau kajian pustaka yang sudah dikumpulkan dan dipelajari oleh penulis untuk menentukan konsep percobaan dan perancangan ini berdasarkan penelitian sebelumnya. Tujuan dari kajian pustaka adalah untuk membedakan penelitian dengan penelitian yang pernah dilakukan oleh orang lain dan menjadi referensi dalam pengembangan ataupun penulisan selanjutnya.

3.2.2 Perumusan Masalah

Perumusan masalah dilakukan setelah mendapatkan konsep dari kajian pustka yang telah ditentukan. Tujuan dari perumusan masalah adalah menentukan langkah-langkah penyelesaian masalah yang dikaji dan menjadi acuan untuk analisa, perancangan sistem, dan pembahasan, serta kesimpulan yang didapat dari hasil penelitian yang telah dilakukan.

3.2.3 Pengambilan dan Pengolahan data

Penulis selanjutnya melakukan pengambilan data lalu melakukan pengolahan data.

3.2.3.1 Dataset

Dalam hal ini penulis mengambil data dataset menggunakan dataset jenis batu permata yang bersumber dari kaggle.com yang memiliki 3.200 dataset, data citra tersebut dikelompokkan menjadi 87 kelas yang sudah dibagi menjadi data latih dan uji. Disini penulis menggunakan 1.002 citra yang di kelompokkan menjadi 30 kelas sebagai data latih dan data uji pada penelitian ini.

Setiap citra memiliki label yang menandakan jenis pada batu permata. Dataset ini merupakan sumber data utama yang digunakan dalam penelitian ini untuk melatih dan menguji model klasifikasi jenis batu permata menggunakan *Convolutional Neural Network* (CNN). Dengan menggunakan dataset ini, diharapkan dapat meningkatkan pemahaman dan kemampuan dalam mengenali serta mengklasifikasikan jenis batu permata berdasarkan citra batu yang diperoleh. Dataset jenis batu permata yang digunakan dalam penelitian ini dapat di lihat pada Tabel 3.2.

Tabel 3. 2 Jenis dan jumlah dataset batu permata

No.	Jenis Batu Permata	Data Latih	Data Uji	Jumlah
1.	Alexandrite	27	7	34
2.	Almandine	25	6	31
3.	Amazonite	25	7	32
4.	Amber	25	6	31
5.	Amethyst	28	7	35
6.	Andalusite	25	7	32

No.	Jenis Batu Permata	Data Latih	Data Uji	Jumlah
7.	Andradite	25	6	31
8.	Aquamarine	29	7	36
9.	Aventurine Green	36	9	45
10.	Aventurine Yellow	27	7	34
11.	Benitoite	25	6	31
12.	Beryl Golden	29	7	36
13.	Bixbite	29	7	36
14.	Bloodstone	28	7	35
15.	Blue Lace Agate	29	8	37
16.	Carnelian	27	6	33
17.	Cats Eye	25	6	31
18.	Chalcedony	24	6	30
19.	Chalcedony Blue	24	6	30
20.	Chrome Diopside	25	7	32
21.	Chrysoberyl	25	6	31
22.	Chrysocolla	29	7	36
23.	Chrysoprase	24	6	30
24.	Citrine	32	8	40
25.	Coral	25	6	31
26.	Danburite	25	7	32
27.	Diamond	25	6	31
28.	Diaspore	25	6	31
29.	Dumortierite	25	7	32
30.	Emerald	29	7	36
Total		801	201	1002

Berdasarkan Tabel 3.2 yang merupakan dataset batu permata dengan total citra 1002, pada setiap jenis batu masing-masing memiliki 30-45 data citra dengan 30 jenis batu permata. Lalu setiap jenis citra batu di kelompokkan menjadi 2 bagian yaitu data latih dan data uji dengan rasio 80% sebagai data latih dan 20% sebagai data uji.

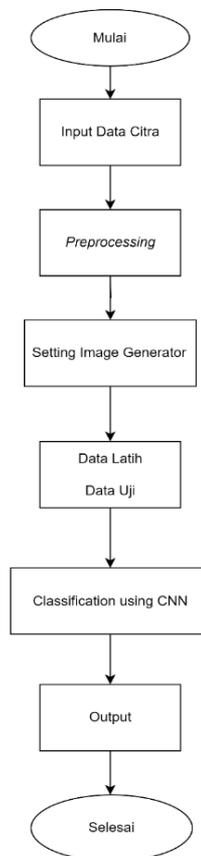
3.2.4 Pembuatan Sistem

Setelah data terkumpul, data tersebut akan dipreproses dan dipersiapkan untuk analisis lebih lanjut dan dilakukan simulasi agar mengetahui sistem ini dapat bekerja dengan baik atau tidak. Hasil dari setiap algoritma akan dievaluasi menggunakan metrik evaluasi yang relevan, seperti akurasi, *presisi*, *recall*, dan *F1-score*. Tahap selanjutnya adalah pembahasan dari hasil simulasi apabila simulasi berhasil dilakukan, jika tidak maka simulasi dilakukan kembali hingga statusnya berhasil.

3.2.5 Analisa

Kemudian di tahap ini akan dilakukan analisa terhadap akurasi *presisi*, *recall*, dan *F1-score* pada model klasifikasi jenis batu permata dengan CNN dalam memprediksi data citra dari dataset. Penelitian akan melakukan metode yang sama seperti pada tahap evaluasi model yaitu membandingkan hasil prediksi kelas jenis batu permata oleh model CNN dengan kelas yang sebenarnya.

3.3 PEMODELAN SISTEM



Gambar 3. 3 Blok Diagram Perancangan Sistem

Gambar 3.3 merupakan gambar blok diagram pada penelitian ini. perancangan sistem yang terdiri dari tiga bagian yaitu perancangan *input*, proses, dan *output* dan pada sistem kerja alat mempunyai fungsi yang berbeda-beda. Diagram *input* diatas menentukan data yang akan dimasukkan yaitu: nama jenis batu permata. Pada bagian proses, data batu permata dikumpulkan lalu diklasifikasi menggunakan metode dan algoritma yang sudah ditentukan, dan hasil outputnya merupakan hasil untuk klasifikasi jenis batu permata.

3.3.1 Input Data Citra

Tahap input Citra merupakan proses memasukkan data citra atau gambar ke dalam *source code*.

3.3.2 Preprocessing

Data yang digunakan untuk pelatihan model deep learning biasanya perlu dipreproses agar lebih mudah dikelola dan digunakan. Ini bisa termasuk konversi format data, normalisasi nilai, augmentasi data, dan lain-lain. Tujuan dari preprocessing data adalah untuk meningkatkan kualitas data dan akurasi model.

3.3.3 Setting Image Generator

Pada tahap ini *setting Image Generator* untuk memuat data gambar atau citra, yang di bagi menjadi data *labelling* dan data *augmentation*.

3.3.3.1 Data Labelling

Proses pelabelan data dilakukan pada citra (gambar) yang dihasilkan oleh image generator. Setiap gambar diberikan label yang sesuai dengan klasifikasi yang telah disusun sebelumnya. Hal ini memungkinkan pengidentifikasi kelas atau kategori pada setiap data gambar. Setelah pelabelan, dilakukan verifikasi dan koreksi terhadap hasil pelabelan. Hal ini penting untuk memastikan keakuratan dan konsistensi dalam proses pelabelan. Jika terdapat kesalahan pelabelan, perlu dilakukan koreksi agar dataset yang digunakan memiliki label yang benar dan konsisten. Terakhir, dataset dapat dipisahkan menjadi set pelatihan, set validasi, dan set pengujian sesuai dengan kebutuhan.

3.3.3.2 Augmentation Dataset

Dataset *training* akan diaugmentasi agar menghindari overfitting pada pelatihan klasifikasi. Hal tersebut bertujuan agar model dapat mengenali variansi dasar pada citra obyek daun Stroberi. Transformasi yang digunakan dalam augmentasi dataset adalah rotasi, refleksi secara horizontal dan vertikal, *zoom*, *scale*, *width* dan *height shift*, serta *shear*. Hasil data *training* yang sudah diaugmentasi dan data yang telah tervalidasi disimpan pada *storage* data.

3.3.3.3 Data Latih

Pada tahap ini, model mempelajari pola-pola dalam data dan menghasilkan model yang dapat digunakan untuk melakukan prediksi. Pelatihan model dilakukan dengan menggunakan metode backpropagation dan optimisasi pada *loss function*.

3.3.3.4 Data Uji

Setelah model berhasil dilatih, tahap terakhir adalah pengujian model pada data uji. Pada tahap ini, model diuji untuk melihat apakah dapat mengklasifikasikan data uji dengan akurat atau tidak.

3.3.3.5 Klasifikasi Menggunakan CNN

Pengklasifikasian jenis batu permata menggunakan *Convolutional Neural Network* (CNN) yang dilakukan dengan beberapa tahap yakni:

1. *Preprocessing Data*: Data yang digunakan untuk pelatihan model *deep learning* biasanya perlu dipreproses agar lebih mudah dikelola dan digunakan. Ini bisa termasuk konversi format data, normalisasi nilai, augmentasi data, dan lain-lain. Tujuan dari *preprocessing* data adalah untuk meningkatkan kualitas data dan akurasi model.
2. Implementasi Arsitektur CNN: Selanjutnya, Arsitektur CNN digunakan dengan menggunakan *library framework tensorflow* dan Keras [25]. Tahapan ini juga melibatkan penentuan jumlah *layer*, jumlah filter pada tiap *layer*, fungsi aktivasi, dan lain sebagainya.
3. Pelatihan Model: Setelah pembuatan model, model dilatih dengan menggunakan data pelatihan. Pada tahap ini, model mempelajari

pola-pola dalam data dan menghasilkan model yang dapat digunakan untuk melakukan prediksi. Pelatihan model dilakukan dengan menggunakan metode backpropagation dan optimisasi pada *loss function*.

4. *Testing Model*: Setelah model berhasil dilatih, tahap terakhir adalah pengujian model pada data uji. Pada tahap ini, model diuji untuk melihat apakah dapat mengklasifikasikan data uji dengan akurat atau tidak.
5. *Confusion Matrix*: Tahap *confusion matrix* membantu mengevaluasi kinerja model dan memahami seberapa baik model dapat mengklasifikasikan data dengan benar. Ini penting untuk mengukur keefektifan model dan meningkatkan kinerjanya jika perlu.
6. *Best CNN Models*: Jika model telah dilatih dan diuji dengan baik, model dapat diekspor ke dalam format yang dapat digunakan oleh aplikasi. Dalam kesimpulannya, untuk melatih sebuah model CNN, perlu melakukan beberapa tahap, yaitu *preprocessing* data, pembuatan model, pelatihan model, validasi model, pengujian model, dan ekspor model. Tahaptahap ini penting untuk mendapatkan model yang akurat dan dapat digunakan pada aplikasi yang diinginkan.