

BAB III METODE PENELITIAN

3.1 Subjek dan Objek Penelitian

3.1.1 Subjek Penelitian

Subjek pada penelitian ini adalah klasifikasi aksara.

3.1.2 Objek Penelitian

Objek pada penelitian ini adalah optimasi hyperparameter pada Convolutional Neural Network (CNN).

3.2 Alat dan Bahan

Dalam penelitian ini digunakan alat dan bahan sebagai penunjang keberhasilan penelitian. Adapun alat dan bahan yang dimaksud yaitu:

3.2.1 Alat

Alat yang digunakan dalam penelitian ini terdiri dari dua jenis yaitu perangkat keras (*hardware*) dan perangkat lunak (*software*), adapun rincian sebagai berikut :

3.3.1 Perangkat Keras (Hardware)

- a. Device : Lenovo Thinkpad T470s
- b. Processor : Processor Intel(R) Core(TM) i7-6600U
- c. RAM : 8 GB

3.3.2 Perangkat Lunak (Software)

- a. Sistem Operasi : *Windows 10 Pro 64-bit*
- b. Bahasa Pemrograman : *Python*
- c. Aplikasi : *Jupyter Notebook* (Membuat Program),
Microsoft Word (Membuat Proposal),
Google Docs (Cek Typo),

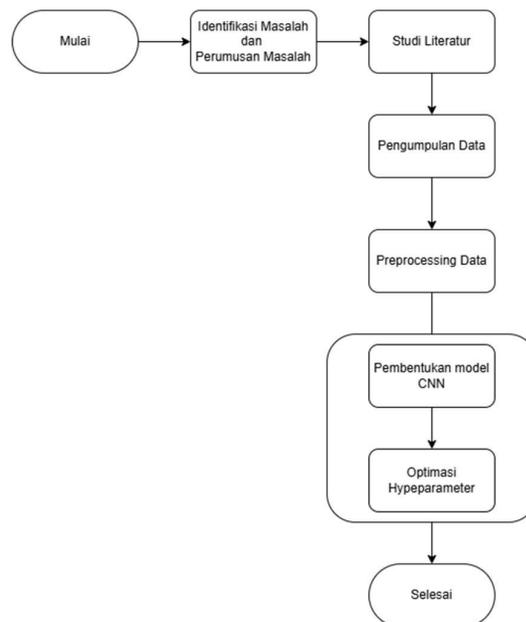
Bing Translator (Penerjemah),
Google Scholar (Mencari Referensi)
Zotero

3.2.2 Bahan

Bahan yang digunakan dalam penelitian ini adalah dataset berupa citra aksara yang mencakup huruf dan kata dari bahasa Jawa, Bali, Sunda dan Lontara dengan format .jpg. Data yang dikumpulkan sebanyak 100 citra pada masing masing kelas sehingga total data yang dikumpulkan sebanyak 400 gambar untuk training dan 400 gambar untuk pengujian. Data tersebut kemudian dijadikan dataset dalam penelitian ini. Selain itu juga dikumpulkan data uji sebanyak 400 gambar dengan masing-masing kelas sebanyak 100.

3.3 Diagram Alir Penelitian

Untuk membuat laporan penelitian ini, terdapat beberapa tahapan dalam melakukan penelitian. Berikut diagram alir penelitian yang dilakukan pada penyusunan laporan ini :



Gambar 3.1 Diagram Alir Penelitian

3.3.1 *Identifikasi dan Perumusan Masalah*

Tahap awal dari penelitian ini adalah mengidentifikasi permasalahan yang akan dibahas oleh penulis, yang telah diuraikan dalam konteks optimasi hyperparameter pada Convolutional Neural Network (CNN) dalam klasifikasi jenis aksara.

3.3.2 *Studi Literatur*

Dalam penelitian ini diperlukan referensi atau sumber data sebagai dasar dan pedoman pengembangan penelitian. Oleh karena itu penulis membaca, meneliti dan memahami konsep dan pembahasan optimasi *hyperparameter CNN* klasifikasi aksara pada jurnal, buku dan penelitian sebelumnya. Hasil yang diperoleh dijadikan sebagai dasar penulisan dan penelitian yang akan dilakukan.

3.3.3 *Pengumpulan Dataset*

Pengumpulan dataset dalam penelitian ini dilakukan melalui metode pengambilan citra aksara daerah yang diperoleh dengan cara *crawling* dari website penerjemah bahasa daerah dan sumber lainnya. Sumber data ini bersifat *open source* dan diperoleh dari beragam sumber yang tersedia secara publik, termasuk repositori online, basis data terbuka, sumber-sumber online, dan sumber data terbuka lainnya. Pendekatan ini memberikan akses kepada peneliti untuk menggali citra-citra aksara yang telah tersebar luas di berbagai platform, memungkinkan analisis serta penelitian yang relevan dalam konteks penelitian ini.

Dataset dibagi menjadi dua jenis, yakni dataset training dan dataset pengujian. Data training digunakan selama proses training model sedangkan data pengujian digunakan untuk menguji model apakah model mengalami overfitting atau tidak. Jumlah dataset yang dikumpulkan untuk penelitian ini sebanyak 400 dataset training dan 400 dataset pengujian, dimana masing-masing akan dibagi menjadi 4 kelas yang memiliki 100 gambar per kelas.

3.3.4 *Preprocessing Data*

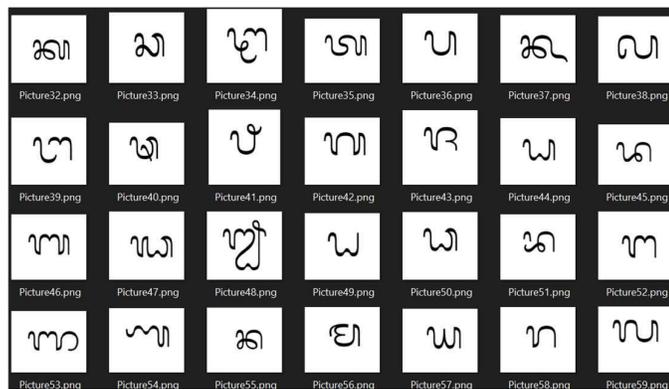
Setelah data diperoleh, tahap selanjutnya adalah preprocessing data. Tujuan dari tahap ini adalah untuk mempermudah dalam ekstraksi ciri yang dibutuhkan

dalam analisis atau pengolahan selanjutnya. Dalam penelitian ini, peneliti melakukan beberapa langkah preprocessing data untuk mempersiapkan gambar-gambar aksara sebelum menggunakannya dalam model klasifikasi.

Langkah awal dari *preprocessing* data pada penelitian ini adalah melakukan pelabelan untuk setiap kelas. Data yang didapatkan dimasukkan ke dalam folder atau kelasnya masing-masing, dimana terdapat empat kelas. Kelas pertama merupakan data citra dari aksara Jawa, kelas kedua merupakan citra dari aksara bali, kelas ketiga merupakan citra dari aksara sunda dan kelas keempat merupakan dari aksara Lontara.



Gambar 3.2 Pembagian Kelas Pada Dataset



Gambar 3.3 Dataset Aksara Bali

dataset yang dikumpulkan memiliki ukuran yang *random* sehingga perlu dilakukan penyesuaian ukuran gambar. Hal ini dapat membantu mengurangi kompleksitas data dan memastikan bahwa model dapat bekerja dengan baik pada berbagai ukuran citra. Pada penelitian ini ukuran citra yang digunakan adalah 256x256 pixel. Selanjutnya, penyesuaian resolusi citra yang dilakukan untuk memastikan bahwa citra memiliki resolusi yang cukup tinggi untuk ekstraksi ciri yang akurat. Peningkatan resolusi citra juga dapat membantu dalam meningkatkan kualitas gambar dan meningkatkan kemampuan model untuk memahami detail-detail penting.

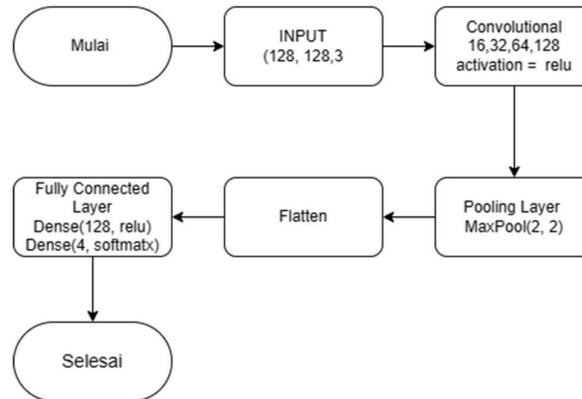
Terakhir, augmentasi pada citra digunakan untuk menciptakan variasi tambahan dalam dataset. Adapun tools yang digunakan dalam augmentasi data adalah *albumentation*. Tahap augmentasi dilakukan dengan transformasi *RGB shift*, *color jitter*, *grid distortion* dan *gaussian noise*. Tujuan melakukan augmentasi ini untuk memperbanyak variasi pada dataset sehingga model tidak hanya mempelajari dataset dengan model yang sama. Dengan dilakukan augmentasi dapat meningkatkan akurasi selama pelatihan. Selama augmentasi akan dilakukan perbanyak dataset sebanyak 1000 gambar per kelas. Yang artinya dataset total sebanyak 4000 gambar. Langkah terakhir adalah pembagian dataset menjadi data *training*, *testing* dan *validation*. Untuk perbandingan antara data *training*, *testing* dan *validation* adalah 60:20:20. Data *training* yang digunakan sebanyak 60%, data *testing* yang digunakan sebanyak 20% dan data *validation* sebanyak 20%.

Dengan langkah-langkah preprocessing ini, data citra aksara menjadi lebih sesuai untuk pelatihan model klasifikasi, dan model dapat memiliki kemampuan yang lebih baik dalam mengidentifikasi dan mengklasifikasikan aksara berdasarkan ciri-ciri yang relevan.

3.3.5 Pembentukan model CNN dan pelatihan Model CNN

Setelah tahap preprocessing data selesai, langkah selanjutnya dalam penelitian ini adalah melakukan pembentukan model CNN (Convolutional Neural Network). Pada tahap ini, peneliti akan merancang arsitektur model CNN yang akan digunakan untuk klasifikasi jenis aksara. Pada tahap ini akan dilakukan pelatihan model untuk mendapatkan hasil pelatihan model CNN tanpa adanya penyesuaian

hyperparameter guna membandingkan hasil model antara model tanpa penyesuaian *hyperparameter* dengan CNN yang disesuaikan dengan *hyperparameter*.



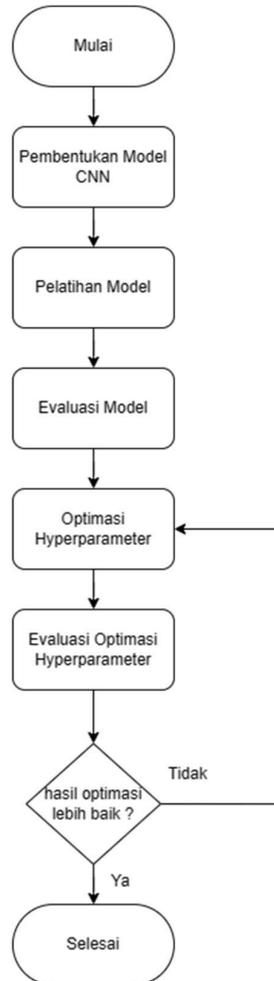
Gambar 3.7 Arsitektur CNN

Untuk arsitektur CNN yang digunakan akan diuraikan sebagai berikut:

Tabel 3.1 Uraian Arsitektur CNN

Parameter	Konfigurasi
Input shape	(128,128,3)
Conv2D	16,(3,3)
Conv2D	32,(3,3)
Conv2D	64,(3,3)
Conv2D	128,(3,3)
Pooling	Max Pooling (2,2)
Dense	128 4
Fungsi aktivasi	Relu Softmax
optimizer	adam
Confusion Matrix	<i>Accuracy, precision, recall dan f1-score</i>

3.3.6 Optimasi Hyperparameter dan pelatihan model CNN



Gambar 3.8 Alur Optimasi *Hyperparameter*

Pada tahap ini, model akan dilakukan optimasi *hyperparameter* dari model CNN yang telah dibentuk sebelumnya. Tujuan dari optimasi *hyperparameter* adalah untuk mengetahui pengaruh akurasi dari model yang telah mengalami optimasi *hyperparameter* dan membandingkannya dengan model yang belum dioptimasi dengan *hyperparameter* dalam klasifikasi aksara. Untuk optimasi *hyperparameter* yang digunakan pada penelitian ini adalah *kernel filter*, *learning*

rate dan *dropout*. Untuk skema konfigurasi *hyperparameter* yang digunakan dapat dilihat pada tabel berikut :

Tabel 3.3 Skema Konfigurasi *Hyperparameter*

Kernel Filter	3 x 3	5 x 5	7x7		
Learning Rate	0,1	0,01	0,001		
DropOut	20%	30%	40%	50%	60%

Untuk skema pelatihan model optimasi *hyperparameter* dapat dilihat sebagai berikut :

Tabel 3.5 Skema Optimasi *Hyperparameter*

skema	Filter Kernel	Learning Rate	Dropout
1	3.3	-	-
2	5.5	-	-
3	7.7	-	-
4	3.3	0.1	-
5	5.5	0.01	-
6	7.7	0.001	-
7	3.3	-	0.2
8	3.3	-	0.3
9	3.3	-	0.4
10	3.3	-	0.5
11	3.3	-	0.6
12	5.5	-	0.2
13	5.5	-	0.3
14	5.5	-	0.4
15	5.5	-	0.5
16	5.5	-	0.6
17	7.7	-	0.2

skema	Filter Kernel	Learning Rate	Dropout
18	7.7	-	0.3
19	7.7	-	0.4
20	7.7	-	0.5
21	7.7	-	0.6
22	3.3	0.1	0.2
23	3.3	0.1	0.3
24	3.3	0.1	0.4
25	3.3	0.1	0.5
26	3.3	0.1	0.6
27	3.3	0.01	0.2
28	3.3	0.01	0.3
29	3.3	0.01	0.4
30	3.3	0.01	0.5
31	3.3	0.01	0.6
32	3.3	0.001	0.2
33	3.3	0.001	0.3
34	3.3	0.001	0.4
35	3.3	0.001	0.5
36	3.3	0.001	0.6
37	5.5	0.1	0.2
38	5.5	0.1	0.3
39	5.5	0.1	0.4
40	5.5	0.1	0.5
41	5.5	0.1	0.6
42	5.5	0.01	0.2
43	5.5	0.01	0.3
44	5.5	0.01	0.4
45	5.5	0.01	0.5
46	5.5	0.01	0.6
47	5.5	0.001	0.2
48	5.5	0.001	0.3
49	5.5	0.001	0.4
50	5.5	0.001	0.5
51	5.5	0.001	0.6
52	7.7	0.1	0.2
53	7.7	0.1	0.3

skema	Filter Kernel	Learning Rate	Dropout
54	7.7	0.1	0.4
55	7.7	0.1	0.5
56	7.7	0.1	0.6
57	7.7	0.01	0.2
58	7.7	0.01	0.3
59	7.7	0.01	0.4
60	7.7	0.01	0.5
61	7.7	0.01	0.6
62	7.7	0.001	0.2
63	7.7	0.001	0.3
64	7.7	0.001	0.4
65	7.7	0.001	0.5
66	7.7	0.001	0.6

Untuk contoh implementasi arsitektur CNN yang dioptimasi sebagai berikut:

Parameter	Konfigurasi
Input Shape	(128,128,3)
Conv2D	16(5.5, activation = relu)
Pooling	Max Pooling(2,2)
Dropout	0.2
Conv2D	32(5.5, activation = relu)
Pooling	Max Pooling(2,2)
Dropout	0.2
Conv2D	64(5.5, activation = relu)
Pooling	Max Pooling(2,2)
Dropout	0.2

Parameter	Konfigurasi
Conv2D	128(5.5, activation = relu)
Pooling	Max Pooling(2,2)
Dropout	0.2
Flatten	Flatten
Dense	(128, activation = relu)
Dropout	0.2
Dense	(4, activation = Softmax)
Optimizer	Adam(Learning Rate = 0.1)

Selama pelatihan menggunakan epoch 100, batch size 25, validation split 20% dan testing 20%. Selama pelatihan, model akan melakukan 100 kali iterasi dengan 25 sample data pelatihan serta menggunakan 20% dari data sebagai data validasi untuk evaluasi model. Setiap model yang dilatih akan dilakukan evaluasi dari hasil pelatihan dengan menggunakan *confusion matrix* yang berisikan nilai *accuracy*, *precision*, *recall* dan *f1-score* menggunakan data uji yang telah disiapkan. Nilai *confusion matrix* dari setiap skema pada optimasi akan dibandingkan dan dipilih model terbaik berdasarkan tingkat akurasi pengujian yang paling tinggi dan mendekati akurasi pelatihan serta nilai *confusion matrix* yang paling baik.

3.3.7 Hipotesis

Pada tahap ini peneliti akan melakukan penarikan hipotesis atau pernyataan sementara dari penelitian yang dilakukan. Untuk itu hipotesis dari penelitian ini adalah sebagai berikut:

1. Optimasi hyperparameter memberikan pengaruh yang cukup besar dalam mengurangi overfitting dan meningkatkan akurasi pada CNN dalam

klasifikasi jenis aksara. Hal ini dapat dilihat dari hasil eksperimen sementara menggunakan dataset dummy.

2. Optimasi hyperparameter tidak memberikan pengaruh yang berarti pada pelatihan CNN dalam klasifikasi jenis aksara.

3.3.8 Kesimpulan

Pada tahap ini, peneliti melakukan perbandingan hasil antara model CNN yang tidak di optimasi menggunakan *hyperparameter* dengan model CNN yang telah di optimasi *hyperparameter*. Dari hasil perbandingan tersebut dianalisis untuk dapat ditarik kesimpulan pengaruh dari optimasi *hyperparameter* pada CNN dalam klasifikasi aksara.