TUGAS PROYEK PENGENALAN POLA LAPORAN AKHIR PROYEK

Dosen Pengampu:

Imam Much Ibnu Subroto, ST., M.Sc., Ph.D



Disusun oleh:

KELOMPOK 9

Muhammad Fikri Maulana	(32602100076)
Muhammad Hibatullah Arkaan	(32602100078)
Muhammad Iko Wardhana	(32602100079)
Muhammad Maulana Fathul Muin	(32602100082)

PRODI TEKNIK INFORMATIKA FAKULTAS TEKNOLOGI INDUSTRI UNIVERSITAS ISLAM SULTAN AGUNG SEMARANG 2024

A. Collection Data

Dataset Aksara Jawa merupakan kumpulan gambar yang mewakili aksara-aksara Jawa. Dataset ini penting untuk melatih model machine learning yang dapat mengklasifikasikan aksara Jawa. Preprocessing data merupakan langkah penting sebelum melatih model, untuk memastikan data berkualitas baik dan siap untuk digunakan.

Berikut sumber Dataset dan Hasil Pengerjaannya:

GitHub:

- https://github.com/vzrenggamani/aksarajawa-hanacaraka

Kaggle:

- https://www.kaggle.com/code/mpwolke/aksara-jawa-javanese-script-images/input

Dataset Gdrive:

- https://drive.google.com/drive/folders/1J1fxbcAKZszxZWu66o7BTRa36G9
Hz4r ?usp=sharing

Google Colab:

- https://colab.research.google.com/drive/1vbcoNVOj3ig6id1VXM0udyKjQ
https://colab.research.google.com/drive/1vb

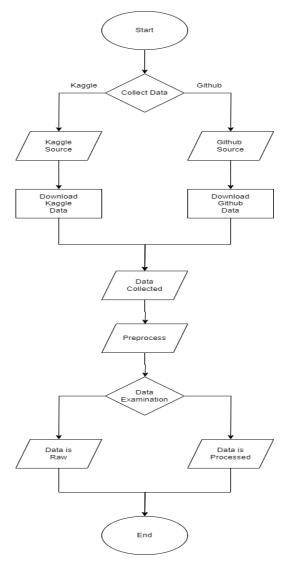
Link Gdrive:

- <u>https://drive.google.com/drive/folders/1J1fxbcAKZszxZWu66o7BTRa36G9</u>

<u>Hz4r_?usp=sharing</u>

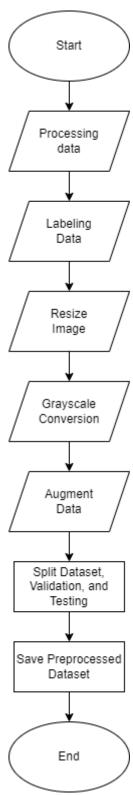
B. Flowchart Model

a. Flowchart Collection Data



Gambar 1. 1 Gambar Flowchart Collection Data

b. Flowchart Preprocessing Data



Gambar 1. 2 Gambar Preprocessing Data

C. Preprocessing Data

a) Pelabelan Data

Setiap gambar yang dikumpulkan dari sumber-sumber tersebut sudah dilabeli dengan benar. Label-label ini menunjukkan karakter Aksara Jawa yang ada pada setiap gambar. Langkah ini penting untuk memastikan data yang digunakan dalam pelatihan model adalah data yang terstruktur dengan baik.

b) Classification

Data yang telah dilabeli akan digunakan untuk melatih model klasifikasi karakter Aksara Jawa. Model ini akan dapat mengenali dan mengklasifikasikan karakter-karakter Aksara Jawa dari gambar input.

c) Resizing Gambar

Semua gambar akan diubah ukurannya menjadi dimensi yang seragam, misalnya 64x64 piksel. Ini dilakukan untuk memastikan semua input ke dalam model memiliki ukuran yang sama, yang akan membantu dalam proses pelatihan model.

d) Grayscale Conversion

Konversi gambar ke format grayscale akan dilakukan untuk mengurangi kompleksitas data. Langkah ini juga membantu dalam meningkatkan efisiensi komputasi.

e) Augmentasi Data

Data yang sudah di-preprocessing akan diperkaya melalui augmentasi seperti rotasi, pergeseran, dan pembesaran gambar. Ini akan membantu dalam meningkatkan variasi dan kualitas dataset.

f) Pembagian Dataset

Dataset akan dibagi menjadi tiga bagian utama: set pelatihan, set validasi, dan set pengujian. Pembagian ini penting untuk mengevaluasi kinerja model dengan benar.

g) Penyimpanan Dataset

Dataset yang sudah dipreproses akan disimpan dalam format yang sesuai untuk digunakan dalam proses pelatihan dan pengujian model.

D. Langkah-langkah Preprocessing Data

a) Menyiapkan Kode untuk import Dataset dari Drive

```
**#kode untuk menyambungkan gdrive tempat dataset yang di simpan.
from google.colab import drive
drive.mount('/content/drive')

**Drive already mounted at /content/drive; to attempt to forcibly remount, call drive.mount("/content/drive", force_remount=True).
```

Gambar 1. 3 Gambar Import Dataset

Kode di atas berfungsi untuk mengimpor dataset dari Google Drive ke Google Collab

b) Import Libary

```
import os
import random
import shutil
import tensorflow as tf
from tensorflow.keras.preprocessing.image import ImageDataGenerator
from shutil import copyfile
import matplotlib.pyplot as plt
import numpy as np
from tensorflow.keras.utils import load_img, img_to_array
```

Gambar 1. 4 Gambar Import Library

Kode di atas adalah mengimpor library di Python, import library sendiri adalah proses memasukkan modul atau paket eksternal ke dalam program Anda agar Anda bisa menggunakan fungsionalitas yang ada dalam library tersebut.

c) Mengecek dataset yang ada

```
# Define source paths
source_path = '/content/drive/MyDrive/dataset'
source_path_ba = os.path.join(source_path, 'ba')
source_path_ba = os.path.join(source_path, 'ba')
source_path_ca = os.path.join(source_path, 'ba')
source_path_da = os.path.join(source_path, 'da')
source_path_dha = os.path.join(source_path, 'dha
source_path_ga = os.path.join(source_path, 'ga')
source_path_ha = os.path.join(source_path, 'ha')
source_path_ka = os.path.join(source_path, 'ja')
source_path_la = os.path.join(source_path, 'ka')
 source_path_la = os.path.join(source_path,
 source_path_ma = os.path.join(source_path,
source_path_na = os.path.join(source_path,
 source_path_nga = os.path.join(source_path,
 source_path_nya = os.path.join(source_path,
 source_path_pa = os.path.join(source_path,
source_path_ra = os.path.join(source_path,
 source_path_sa = os.path.join(source_path,
 source_path_ta = os.path.join(source_path,
 source_path_tha = os.path.join(source_path,
 source_path_wa = os.path.join(source_path,
 source_path_ya = os.path.join(source_path,
 print(f"There are {len(os.listdir(source_path_ba))} images of ba.")
print(f"There are {len(os.listdir(source_path_ca))} images of ba.")
print(f"There are {len(os.listdir(source_path_ca))} images of ca.")
print(f"There are {len(os.listdir(source_path_da))} images of da.")
print(f"There are {len(os.listdir(source_path_da))} images of dha.")
print(f"There are {len(os.listdir(source_path_da))} images of ga.")
print(f"There are {len(os.listdir(source_path_ha))} images of ha.")
print(f"There are {len(os.listdir(source_path_ha))} images of ja.")
print(f"There are {len(os.listdir(source_path_ka))} images of ka.")
                                                                                                                                         images of la.")
images of ma.")
 print(f"There are {len(os.listdir(source_path_la))}
print(f"There are {len(os.listdir(source_path_la))} images of ta.")
print(f"There are {len(os.listdir(source_path_ma))} images of ma.")
print(f"There are {len(os.listdir(source_path_ma))} images of na.")
print(f"There are {len(os.listdir(source_path_ma))} images of nga.")
print(f"There are {len(os.listdir(source_path_ma))} images of pa.")
print(f"There are {len(os.listdir(source_path_ma))} images of ra.")
print(f"There are {len(os.listdir(source_path_ma))} images of ra.")
print(f"There are {len(os.listdir(source_path_ma))} images of ta.")
print(f"There are {len(os.listdir(source_path_ma))} images of ta.")
 print(f"There are {len(os.listdir(source_path_ta))} images of ta.")
print(f"There are {len(os.listdir(source_path_tha))} images of tha.";
 print(f"There are {len(os.listdir(source_path_wa))} images of wa.
 print(f"There are {len(os.listdir(source_path_ya))}
```

Gambar 1. 5 Gambar mengecek dataset

Kode diatas adalah untuk mengecek data yang ada dalam dataset, diatas mengimport data dari google drive ke google collab dengan alamat '/content/drive/MyDrive/dataset'

d) Hasil pengecekan Dataset Mengecek dataset yang ada

```
→ There are 114 images of ba.
    There are 108 images of ca.
    There are 108 images of da.
    There are 108 images of dha.
    There are 108 images of ga.
    There are 102 images of ha.
    There are 108 images of
                            ka.
    There are 108 images of
    There are 108 images of la.
    There are 108 images of ma.
    There are 108 images of na.
    There are 102 images of nga.
    There are 108 images of nya.
    There are 108 images of pa.
    There are 108 images of
    There are 108 images of sa.
    There are 108 images of ta.
    There are 108 images of tha.
    There are 108 images of wa.
    There are 108 images of ya.
```

Gambar 1. 6 Gambar Hasil cek Dataset

Kode di atas adalah hasil dari pengecekan data pada gambar sebelumnya, data diatas menampilkan jumlah gambar pada data Ha, Na, Ca, Ra, Ka.

e) Penghitungan Jumlah Gambar dalam Dataset Pelatihan dan Validasi

```
print(f"There are {len(os.listdir(TRAINING_ta_DIR))} images of ta for training")
print(f"There are {len(os.listdir(TRAINING_tha_DIR))} images of tha for training")
print(f"There are {len(os.listdir(TRAINING_wa_DIR))} images of wa for training")
print(f"There are {len(os.listdir(TRAINING_ya_DIR))} images of ya for training")

print(f"There are {len(os.listdir(VALIDATION_ba_DIR))} images of ba for validation")
print(f"There are {len(os.listdir(VALIDATION_ca_DIR))} images of da for validation")
print(f"There are {len(os.listdir(VALIDATION_da_DIR))} images of dha for validation")
print(f"There are {len(os.listdir(VALIDATION_dha_DIR))} images of dha for validation")
```

Gambar 1. 7 Gambar Menghitung jumlah Gambar

Kode ini menampilkan jumlah gambar untuk berbagai kategori aksara (seperti ta, tha, wa, ya) dalam direktori pelatihan dan validasi. Fungsi os.listdir() digunakan untuk menghitung file dalam setiap direktori, memberikan gambaran distribusi data antar kategori dan set.

f) Struktur Direktori untuk Dataset

```
# Membuat Struktur Direktori Training dan Validation untuk Klasifikasi Gambar
root_dir = '/content/drive/MyDrive/Untitled Folder'

def create_train_val_dirs(root_path):
    os.makedirs(os.path.join(f'{root_path}/training', 'ba'))
    os.makedirs(os.path.join(f'{root_path}/training', 'ca'))
    os.makedirs(os.path.join(f'{root_path}/training', 'da'))
    os.makedirs(os.path.join(f'{root_path}/training', 'da'))
    os.makedirs(os.path.join(f'{root_path}/training', 'da'))
    os.makedirs(os.path.join(f'{root_path}/training', 'la'))
    os.makedirs(os.path.join(f'{root_path}/training', 'la'))
    os.makedirs(os.path.join(f'{root_path}/training', 'ka'))
    os.makedirs(os.path.join(f'{root_path}/training', 'la'))
    os.makedirs(os.path.join(f'{root_path}/training', 'na'))
    os.makedirs(os.path.join(f'{root_path}/training', 'na'))
    os.makedirs(os.path.join(f'{root_path}/training', 'nya'))
    os.makedirs(os.path.join(f'{root_path}/training', 'ra'))
    os.makedirs(os.path.join(f'{root_path}/training', 'ra'))
    os.makedirs(os.path.join(f'{root_path}/training', 'ra'))
    os.makedirs(os.path.join(f'{root_path}/training', 'ta'))
    os.makedirs(os.path.join(f'{root_path}/training', 'ta'))
    os.makedirs(os.path.join(f'{root_path}/training', 'ta'))
    os.makedirs(os.path.join(f'{root_path}/training', 'ta'))
    os.makedirs(os.path.join(f'{root_path}/training', 'ya'))

    os.makedirs(os.path.join(f
```

Gambar 1. 8 Gambar Struktur Direktori Dataset

Kode di atas membuat struktur direktori untuk dataset training dan validation dalam klasifikasi gambar menggunakan fungsi makedirs dari modul os.

g) Fungsi untuk Memisahkan Data Gambar menjadi Set Pelatihan dan Validasi

Gambar 1. 9 Gambar Fungsi Pisah Data

Kode di atas mengacak urutan data dalam direktori sumber dan memisahkannya menjadi set pelatihan dan validasi menggunakan fungsi split_data. Fungsi ini menggunakan random.sample untuk mengacak data, lalu membagi data sesuai dengan proporsi yang ditentukan oleh SPLIT SIZE.

h) Mengatur dan Memisahkan Data Gambar ke Direktori

```
# Skrip untuk Mengatur dan Memisahkan Data Gambar ke Direktori Pelatihan dan Validasi

# Jalur untuk mengakses dataset
ba_SOURCE_DIR = source_path_ba
ca_SOURCE_DIR = source_path_ca
da_SOURCE_DIR = source_path_dha
ga_SOURCE_DIR = source_path_dha
ga_SOURCE_DIR = source_path_ha
ja_SOURCE_DIR = source_path_ha
ja_SOURCE_DIR = source_path_la
ma_SOURCE_DIR = source_path_la
ma_SOURCE_DIR = source_path_la
ma_SOURCE_DIR = source_path_na
na_SOURCE_DIR = source_path_na
na_SOURCE_DIR = source_path_na
na_SOURCE_DIR = source_path_na
na_SOURCE_DIR = source_path_ra
sa_SOURCE_DIR = source_path_ta
ta_SOURCE_DIR = source_path_ta
tha_SOURCE_DIR = source_path_ta
tha_SOURCE_DIR = source_path_ta
tha_SOURCE_DIR = source_path_ya

TRAINING_DIR = os.path.join(root_dir, 'training')
VALIDATION_DIR = os.path.join(TRAINING_DIR, 'ba')
VALIDATION_ba_DIR = os.path.join(TRAINING_DIR, 'ca')
VALIDATION_ca_DIR = os.path.join(TRAINING_DIR, 'ca')
VALIDATION_ca_DIR = os.path.join(TRAINING_DIR, 'ca')
```

Gambar 1. 10 Gambar Atur dan Pisah Data

Kode di atas mengatur jalur untuk mengakses dataset dan memisahkan data gambar ke dalam direktori pelatihan dan validasi. Setiap kelas memiliki direktori sumber yang sesuai, seperti ba_SOURCE_DIR untuk kelas 'ba'. Kode ini juga menentukan direktori pelatihan dan validasi untuk setiap kelas dengan menggunakan fungsi os.path.join untuk memastikan struktur direktori yang konsisten.

i) Pemisahan Data untuk Pelatihan dan Validasi Model

```
TRAINING_ya_DIR = os.path.join(TRAINING_DIR, 'ya')
VALIDATION_ya_DIR = os.path.join(VALIDATION_DIR, 'ya')

# Tentukan proporsi gambar yang digunakan untuk pelatihan
split_size = 0.7

# Menjalankan fungsi di atas
split_data(ba_SOURCE_DIR, TRAINING_ba_DIR, VALIDATION_ba_DIR, split_size)
split_data(ca_SOURCE_DIR, TRAINING_ca_DIR, VALIDATION_ca_DIR, split_size)
split_data(da_SOURCE_DIR, TRAINING_da_DIR, VALIDATION_da_DIR, split_size)
split_data(dha_SOURCE_DIR, TRAINING_dha_DIR, VALIDATION_dha_DIR, split_size)
split_data(ga_SOURCE_DIR, TRAINING_ga_DIR, VALIDATION_ga_DIR, split_size)
```

Gambar 1. 11 Gambar Pisah Data Pelatihan dan Validasi Model

Kode tersebut membagi dataset menjadi set pelatihan dan validasi untuk berbagai kategori data dengan proporsi 70:30. Fungsi split_data() digunakan untuk memisahkan data dari direktori sumber ke direktori pelatihan dan validasi yang telah ditentukan.

j) Statistik Gambar Pelatihan per Kelas

```
Original sa directory has 108 images
Original ta directory has 108 images
Original tha directory has 108 images
Original wa directory has 108 images
Original ya directory has 108 images
There are 79 images of ba for training
There are 75 images of ca for training
There are 75 images of da for training
There are 75 images of ga for training
There are 75 images of ha for training
There are 71 images of ha for training
There are 75 images of ja for training
There are 75 images of ka for training
There are 75 images of ka for training
There are 75 images of la for training
```

Gambar 1. 12 Gambar Statistik Pelatihan

Gambar di atas menampilkan jumlah gambar dalam setiap direktori untuk data pelatihan berbagai kelas. Setiap direktori berisi gambar yang akan digunakan dalam proses pelatihan model. Informasi ini menunjukkan bahwa semua direktori awal memiliki 108 gambar, sementara jumlah gambar yang akan digunakan untuk pelatihan bervariasi per kelas, dengan jumlah gambar antara 71 hingga 79.

k) Membuat Generator Data Pelatihan dan Validasi

Gambar 1. 13 Gambar Membuat Generator Data

Kode di atas membuat generator data pelatihan dan validasi menggunakan ImageDataGenerator dari TensorFlow. Generator ini digunakan untuk mengelola gambar pelatihan dan validasi dengan melakukan augmentasi data seperti rescaling, rotasi, zoom, dan shear. Kode ini menghasilkan dua generator: train_generator untuk data pelatihan dan validation_generator untuk data validasi, yang siap digunakan untuk melatih model.

1) Definisi dan Kompilasi Model CNN

Gambar 1. 14 Gambar Definisi dan Kompilasi Model CNN

Kode di atas mendefinisikan dan mengkompilasi model Convolutional Neural Network (CNN) menggunakan TensorFlow. Model tersebut memiliki tiga lapisan Convolutional dengan ukuran kernel 3x3 dan fungsi aktivasi ReLU, masingmasing diikuti oleh lapisan MaxPooling dengan ukuran 2x2. Model ini diakhiri dengan dua lapisan Dense, satu dengan 512 unit dan fungsi aktivasi ReLU, serta yang terakhir dengan 20 unit dan fungsi aktivasi Softmax. Model dikompilasi menggunakan optimizer Adam dengan laju pembelajaran 0.001 dan menggunakan loss function categorical_crossentropy serta mengukur akurasi.

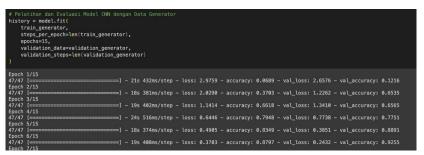
m) Menampilkan Arsitektur dan Detail Parameter Model CNN

Layer (type)	Output	Shape	Param #
conv2d_15 (Conv2D)	(None,	62, 62, 16)	
max_pooling2d_15 (MaxPooling2D)	(None,	31, 31, 16)	
conv2d_16 (Conv2D)	(None,	29, 29, 32)	4640
<pre>max_pooling2d_16 (MaxPooli ng2D)</pre>	(None,	14, 14, 32)	
conv2d_17 (Conv2D)	(None,	12, 12, 64)	18496
max_pooling2d_17 (MaxPooling2D)	(None,	6, 6, 64)	
flatten_5 (Flatten)	(None,	2304)	
dense_10 (Dense)	(None,	512)	1180160
dense_11 (Dense)	(None,	20)	10260

Gambar 1. 15 Gambar menampilkan Arsitektur dan Parameter CNN

Gambar di atas menampilkan arsitektur dan detail parameter dari model Convolutional Neural Network (CNN) yang digunakan. Model ini dibuat menggunakan fungsi create_model(), dan terdiri dari beberapa lapisan convolutional, max-pooling, flatten, dan dense. Tabel menunjukkan jenis lapisan, bentuk keluaran (output shape), dan jumlah parameter yang dapat dilatih (trainable params). Total parameter model adalah 1.214.004, yang semuanya dapat dilatih.

n) Pelatihan dan Evaluasi Model CNN dengan Data Generator

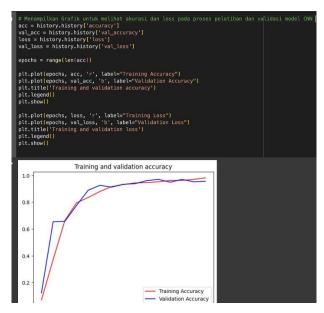


Gambar 1. 16 Gambar Pelatihan dan Evaluasi CNN

Gambar di atas menunjukkan proses pelatihan dan evaluasi model CNN menggunakan data generator. Pada bagian atas gambar, terdapat potongan kode yang digunakan untuk melatih model selama 15 epoch, dengan menggunakan train_generator untuk data pelatihan dan validation_generator untuk data Svalidasi. Pada bagian bawah, ditampilkan hasil pelatihan untuk setiap epoch,

termasuk waktu per step, nilai loss, dan akurasi untuk data pelatihan serta data validasi.

o) Visualisasi Akurasi dan Loss Selama Pelatihan dan Validasi Model CNN



Gambar 1. 17 Gambar Visualisasi model CNN

Gambar di atas menunjukkan grafik akurasi dan loss selama proses pelatihan dan validasi model CNN. Grafik pertama menampilkan akurasi pelatihan dan validasi, di mana garis merah menunjukkan akurasi pelatihan dan garis biru menunjukkan akurasi validasi. Grafik kedua menunjukkan loss pelatihan dan validasi, dengan garis merah menunjukkan loss pelatihan dan garis biru menunjukkan loss validasi. Hasil ini membantu dalam memahami bagaimana model berperforma selama pelatihan dan validasi, serta apakah terjadi overfitting atau underfitting.