

# Deep Learning Course

## Final Project Report



Disusun oleh:

NURUL ALYA  
REZQIA NURQALBI  
ANDI MUTHIA MULIA PUTRI

H071221009  
H071221025  
H071221041

Universitas Hasanuddin  
Makassar, Indonesia

November 28, 2024

# Contents

|          |                              |           |
|----------|------------------------------|-----------|
| <b>1</b> | <b>INTRODUCTION</b>          | <b>2</b>  |
| <b>2</b> | <b>RELATED WORKS</b>         | <b>3</b>  |
| <b>3</b> | <b>DATASET AND MATERIAL</b>  | <b>4</b>  |
| <b>4</b> | <b>RESULT AND DISCUSSION</b> | <b>10</b> |
| <b>5</b> | <b>CONCLUSION</b>            | <b>13</b> |

# 1 INTRODUCTION

- **Background and context of the project**

Teknologi pengenalan gestur tangan telah berkembang pesat, terutama dengan dukungan machine learning dan deep learning yang semakin canggih. Salah satu aplikasi yang banyak dikembangkan adalah pengenalan gestur hand digit (angka 1 hingga 5), yang memiliki potensi besar untuk meningkatkan interaksi manusia-komputer (Human-Computer Interaction atau HCI). Teknologi ini dapat digunakan untuk berbagai kebutuhan, seperti komunikasi sederhana, kontrol perangkat berbasis IoT, hingga interaksi dengan robot asisten.

Namun, pengembangan teknologi ini menghadapi tantangan besar, terutama dalam hal penyediaan dataset yang cukup bervariasi untuk melatih model deep learning. Variasi dataset diperlukan agar model dapat mengenali gestur tangan dengan akurasi tinggi dalam berbagai kondisi lingkungan, termasuk jarak, pencahayaan, dan latar belakang. Dengan menciptakan dataset yang beragam, teknologi pengenalan gestur tangan dapat diterapkan dalam situasi nyata dan mendukung pengembangan aplikasi berbasis kecerdasan buatan yang lebih andal.

- **Problem statement**

- **Kurangnya Dataset Berkualitas:** Teknologi pengenalan gestur tangan, khususnya untuk hand digit 1 hingga 5, sering terhambat oleh kurangnya dataset yang mencakup berbagai kondisi lingkungan.
- **Pengaruh Faktor Lingkungan:** Variasi jarak, pencahayaan, dan latar belakang yang berbeda dapat memengaruhi akurasi model deep learning dalam mengenali gestur tangan.
- **Keterbatasan Implementasi Luas:** Kendala ini menghambat pengaplikasian teknologi pengenalan gestur tangan secara luas, terutama pada sistem yang membutuhkan keandalan tinggi, seperti robot asisten atau perangkat IoT.

- **Objectives of the research**

- **Menyediakan Dataset yang Bervariasi:** Mengembangkan dataset pengenalan hand digit 1 hingga 5 yang mencakup variasi jarak, pencahayaan, dan latar belakang. Dataset ini dirancang agar model deep learning dapat dilatih untuk mengenali gestur tangan secara andal di berbagai kondisi lingkungan.
- **Meningkatkan Akurasi Model Deep Learning:** Menggunakan dataset baru untuk meningkatkan akurasi model deep learning, seperti Convolutional Neural Networks (CNN), dalam mengenali gestur tangan.
- **Mendukung Pengembangan Aplikasi Berbasis AI** Memfasilitasi pengembangan teknologi berbasis pengenalan gestur tangan untuk aplikasi interaksi manusia-komputer, perangkat IoT, dan robot asisten, khususnya dalam sektor layanan, seperti pemesanan ruangan karaoke.
- **Mendorong Implementasi Teknologi Berbasis Gestur:** Membantu mengintegrasikan teknologi pengenalan gestur tangan dalam aplikasi sehari-hari, sehingga proses menjadi lebih cepat, intuitif, dan bebas sentuhan.

Dengan tercapainya tujuan-tujuan ini, pengembangan teknologi pengenalan gestur tangan diharapkan dapat memberikan manfaat yang signifikan bagi berbagai bidang, dari hiburan hingga interaksi berbasis AI yang lebih luas.

## 2 RELATED WORKS

- **Literature review of relevant papers and articles**

Penelitian seperti yang dilakukan oleh Domahs et al. (2010, dalam Cognition) menunjukkan bahwa kebiasaan menghitung menggunakan jari memengaruhi cara angka diproses, dan mendukung asumsi bahwa efek ini dimodulasi oleh budaya. Namun, tingkat keberagaman budaya dalam kebiasaan menghitung dengan jari secara umum sangat diremehkan di bidang penelitian, yang pada gilirannya membatasi pertanyaan dan desain penelitian.

Dalam hal ini, Bender dan Beller (2012) mengeksplorasi diversitas budaya dalam penggunaan jari untuk menghitung angka. Mereka menemukan bahwa sistem penghitungan jari bervariasi dalam aspek seperti jari yang digunakan untuk memulai, tangan yang digunakan, dan orientasi gerakan (misalnya, dari tangan tertutup ke terbuka atau sebaliknya). Studi ini menegaskan bahwa gestur jari tidak hanya merupakan alat kognitif alami, tetapi juga sangat dipengaruhi oleh kode budaya tertentu, yang mencerminkan kompleksitas interaksi antara faktor biologis dan sosial dalam representasi angka.

Penerapan konsep ini menjadi relevan dalam pengembangan teknologi pengenalan gestur jari, terutama untuk mengenali angka yang ditunjukkan dengan tangan. Teknologi ini merupakan bagian penting dari interaksi manusia-komputer (Human-Computer Interaction atau HCI) dan digunakan dalam berbagai aplikasi, seperti komunikasi non-verbal, kontrol perangkat berbasis IoT, hingga interaksi dengan robot asisten. Sistem ini memanfaatkan teknologi berbasis visi, sensor radar, dan algoritma pembelajaran mesin untuk memahami dan menganalisis gestur tangan.

Namun, salah satu tantangan utama dalam pengembangan teknologi ini adalah menyediakan dataset yang mencakup variasi kondisi lingkungan, seperti pencahayaan, jarak, dan latar belakang yang berbeda. Variasi ini diperlukan untuk melatih model agar mampu mengenali gestur tangan dengan akurasi tinggi dalam berbagai skenario nyata. Dengan mengatasi tantangan ini, pengenalan gestur jari dapat menjadi lebih andal dan dapat diterapkan secara luas dalam berbagai sistem berbasis AI.

- **Comparison of different approaches and their results**

- **Nature and Culture or Finger Counting: Diversity and Representational Effects of An Embodied Cognitive Tool**

Penelitian ini menggunakan pendekatan studi etnografis dan kognitif tentang variasi budaya dalam penghitungan jari yang hasilnya menunjukkan variasi budaya dalam penghitungan jari yang memengaruhi representasi dan pemrosesan angka serta pendekatan pendidikan. Relevansi dan implikasinya menyoroti pentingnya konteks budaya dalam memahami dan mengaplikasikan ges-

tur tangan, relevan untuk desain dataset yang mempertimbangkan variasi budaya.

– **Implementing a Hand Gesture Recognition System Based on Range-Doppler Map**

Penelitian ini menggunakan pendekatan penggunaan sinyal radar berbasis peta Range-Doppler untuk mengenali gestur tangan yang hasilnya mencapai akurasi 98% dalam klasifikasi gestur, dengan fleksibilitas tinggi untuk berbagai skenario penggunaan. Relevansi dan implikasinya menekankan efisiensi pendekatan berbasis radar, namun tantangan pada gestur mikro dan lingkungan beragam tetap relevan untuk dibandingkan dengan metode visual.

– **Hybrid CNN-SVM Classifier for Handwritten Digit Recognition**

Penelitian ini menggunakan kombinasi CNN dan SVM untuk mengenali angka tulisan tangan pada dataset MNIST yang hasilnya akurasi tinggi sebesar 99,28% menunjukkan efektivitas pendekatan hibrida dalam mengenali pola kompleks. Relevansi dan implikasinya menggarisbawahi pentingnya kerangka kerja yang canggih dalam mencapai akurasi tinggi, relevan dengan penelitian pengenalan gestur tangan berbasis visual.

• **Justification for your chosen methodology**

Metodologi yang dipilih untuk proyek ini didasarkan pada kebutuhan untuk mengatasi keterbatasan yang diidentifikasi dalam penelitian sebelumnya terkait keberagaman dataset dan representasi budaya dalam pengenalan gestur. Dengan fokus pada dataset yang menangkap berbagai gestur digit tangan dalam kondisi yang bervariasi, proyek ini bertujuan untuk:

- **Meningkatkan Ketahanan Model:** Dengan menyertakan contoh beragam dari digit tangan dari berbagai sudut pandang, jarak, dan kondisi pencahayaan, dataset akan memungkinkan model untuk lebih baik menggeneralisasi di berbagai skenario dunia nyata.
- **Mendukung Adaptabilitas Budaya:** Mengingat kebiasaan menghitung dengan jari berbeda secara budaya, dataset akan mencakup variasi dalam representasi gestur untuk memenuhi audiens global.
- **Memanfaatkan Teknik Lanjutan:** Menggunakan CNN dan arsitektur deep learning lainnya akan memastikan bahwa sistem mampu mencapai akurasi tinggi dalam mengenali gestur digit tangan, sehingga meningkatkan interaksi manusia-komputer.

### 3 DATASET AND MATERIAL

• **Source of the dataset**

Dataset yang digunakan dalam proyek ini adalah Natural Hand Digit Dataset yang dikembangkan secara mandiri. Dataset ini dirancang untuk mendukung pengenalan gestur tangan berupa angka 1 hingga 5, yang relevan untuk aplikasi berbasis Human-Computer Interaction (HCI) dan IoT.

Metode Pengambilan Data:

- **Kamera:** Pengambilan data dilakukan menggunakan smartphone dengan spesifikasi berikut:
  - \* Resolusi video Ultra HD 4K (3840 x 2160p) pada 30 fps.
  - \* Sistem Operasi: Android 14 dengan chipset Exynos 1480.
- **Variasi Data:** Data dikumpulkan dengan variasi pengambilan berikut:
  - \* Tiga jarak pengambilan: 50 cm, 1 m, dan 2 meter.
  - \* Kondisi pencahayaan rendah untuk meniru situasi dunia nyata.
  - \* Background netral (putih) agar fokus hanya pada gestur tangan.
- **Tinggi pengambilan data:** Kamera dipasang pada ketinggian 125 cm menggunakan tripod untuk mencerminkan tinggi rata-rata bahu manusia, menghasilkan perspektif yang konsisten untuk setiap pengambilan data.
- **Skenario Pengambilan Data:** Pengambilan data dilakukan dengan merekam video angka 1 hingga 5 menggunakan berbagai posisi tangan dari beberapa individu. Hal ini untuk memastikan keberagaman dalam representasi gestur tangan yang dapat digunakan oleh model dalam berbagai situasi.
- **Jumlah Data yang dikumpulkan per pekan :**

## Pekan 11 dan Pekan 12

| Digit             | Jarak | Pekan 11 (Jumlah Data) | Pekan 12 (Jumlah Data) |
|-------------------|-------|------------------------|------------------------|
| 1                 | 50cm  | 90 data                | 30 data                |
| 2                 | 50cm  | 90 data                | 30 data                |
| 3                 | 50cm  | 90 data                | 30 data                |
| 4                 | 50cm  | 90 data                | 30 data                |
| 5                 | 50cm  | 90 data                | 30 data                |
| 1                 | 1m    | 90 data                | 30 data                |
| 2                 | 1m    | 90 data                | 30 data                |
| 3                 | 1m    | 90 data                | 30 data                |
| 4                 | 1m    | 90 data                | 30 data                |
| 5                 | 1m    | 90 data                | 30 data                |
| 1                 | 2m    | 90 data                | 30 data                |
| 2                 | 2m    | 90 data                | 30 data                |
| 3                 | 2m    | 90 data                | 30 data                |
| 4                 | 2m    | 90 data                | 30 data                |
| 5                 | 2m    | 90 data                | 30 data                |
| <b>Total Data</b> |       | <b>1350 data</b>       | <b>900 data</b>        |

## Pekan 13 dan Pekan 14

| Digit      | Jarak | Pekan 13 (Jumlah Data) | Pekan 14 (Jumlah Data) |
|------------|-------|------------------------|------------------------|
| 1          | 50cm  | 90 data                | 90 data                |
| 2          | 50cm  | 90 data                | 90 data                |
| 3          | 50cm  | 90 data                | 90 data                |
| 4          | 50cm  | 90 data                | 90 data                |
| 5          | 50cm  | 90 data                | 90 data                |
| 1          | 1m    | 90 data                | 90 data                |
| 2          | 1m    | 90 data                | 90 data                |
| 3          | 1m    | 90 data                | 90 data                |
| 4          | 1m    | 90 data                | 90 data                |
| 5          | 1m    | 90 data                | 90 data                |
| 1          | 2m    | 90 data                | 90 data                |
| 2          | 2m    | 90 data                | 90 data                |
| 3          | 2m    | 90 data                | 90 data                |
| 4          | 2m    | 90 data                | 90 data                |
| 5          | 2m    | 90 data                | 90 data                |
| Total Data |       | 1350 data              | 1350 data              |

- **Preprocessing steps.**

Proses preprocessing dataset meliputi:

- **Pembagian Data:**

- \* Dataset dibagi menjadi dua subset: training (80%) dan testing (20%) menggunakan fungsi Python untuk memastikan distribusi yang seimbang.
    - \* File dipindahkan secara terstruktur berdasarkan kelas.

- **Resize dan Normalisasi:**

- \* Semua gambar diubah ukurannya menjadi 224x224 piksel.
    - \* Nilai piksel dinormalisasi ke skala  $[0, 1]$  menggunakan `ImageDataGenerator`.

- **Augmentasi Data:**

- \* Augmentasi gambar diterapkan menggunakan `ImageDataGenerator` untuk menambah variasi dataset, seperti flipping, rotasi, dan zoom.

- **Pemeriksaan Data:**

- \* Dataset dianalisis untuk memeriksa distribusi kelas menggunakan plot distribusi data.

\* Contoh gambar dari setiap kelas ditampilkan untuk memverifikasi kualitas dan label.

- **Features and labels included in the data.**

- **Fitur (Features):** Gambar tangan yang menampilkan gestur angka direpresentasikan sebagai matriks nilai piksel berukuran 224x224x3 (RGB). Setiap gambar adalah representasi visual dari angka yang ditunjukkan dengan tangan dalam bentuk gestur.
- **Label (Output):** Kategori angka yang ditunjukkan oleh gestur tangan pada gambar, yaitu angka 1, 2, 3, 4, atau 5. Label ini digunakan sebagai target yang diprediksi oleh model dalam proses pelatihan.

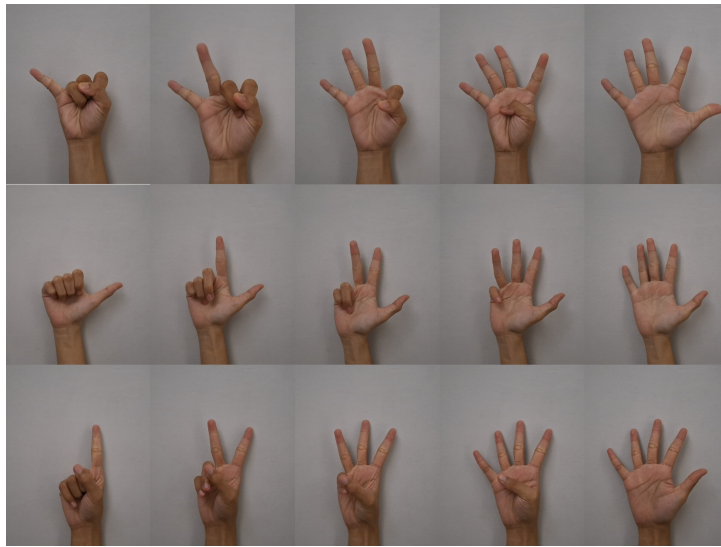


Figure 1: Contoh gambar dari masing-masing kategori angka

- **Tools, Libraries, and Frameworks**

Beberapa alat dan pustaka yang digunakan dalam proyek ini adalah sebagai berikut:

- **Tools:** Google Colab



Figure 2: Logo Google Colab

Google Colab digunakan sebagai platform berbasis cloud untuk menulis, menjalankan, dan mempercepat kode dengan dukungan GPU/TPU gratis.

- **Bahasa Pemrograman:** Python





Figure 3: Logo Bahasa Pemrograman Python

Python digunakan untuk mengembangkan kode yang mencakup pengolahan data, pemrosesan gambar, pembangunan model deep learning, evaluasi performa, dan visualisasi hasil.

– **Library untuk Preprocessing dan Visualisasi:**

```
import random
import shutil
import os
import numpy as np
import matplotlib.pyplot as plt
import matplotlib.image as mpimg

from sklearn.metrics import (
    classification_report,
    confusion_matrix,
    ConfusionMatrixDisplay
)

import tensorflow as tf
from tensorflow.keras.models import Sequential
from tensorflow.keras.layers import Conv2D, MaxPooling2D, Flatten, Dense, Dropout
from tensorflow.keras.preprocessing.image import ImageDataGenerator
from tensorflow.keras.preprocessing.image import load_img, img_to_array
from tensorflow.keras.preprocessing import image
```

Figure 4: Library Yang Digunakan

- \* **Manipulasi File dan Folder:** Pustaka seperti `os`, `shutil`, dan `random` digunakan untuk mengelola file dan folder, seperti memindahkan, menyalin, atau mengacak file dalam direktori. Hal ini berguna saat menyiapkan dataset atau mengatur struktur folder.
- \* **Pengolahan Data dan Visualisasi:** `numpy` digunakan untuk operasi matematika dan manipulasi array yang efisien, sementara `matplotlib` membantu memvisualisasikan data, seperti menampilkan gambar, grafik akurasi, atau matriks kebingungan hasil evaluasi.
- \* **Evaluasi Model:** Pustaka evaluasi seperti `accuracy_score`, `classification_report`, dan `confusion_matrix` dari `sklearn` digunakan untuk mengukur performa model. Anda juga dapat membuat visualisasi seperti matriks kebingungan menggunakan `ConfusionMatrixDisplay`.
- \* **Pemrosesan Gambar:** `tensorflow.keras.preprocessing.image` menyediakan alat untuk membaca, mengubah, dan mengonversi gambar menjadi format array yang kompatibel dengan model deep learning.
- \* **Pembangunan dan Pelatihan Model Deep Learning:** `tensorflow` digunakan untuk membuat arsitektur model seperti CNN (Convolutional Neural Networks), melatih model dengan data, dan melakukan prediksi untuk tugas-tugas seperti klasifikasi gambar.

- \* **Augmentasi Data Gambar:** `ImageDataGenerator` membantu membuat variasi baru dari gambar dataset dengan transformasi seperti rotasi, flipping, atau zoom. Hal ini penting untuk meningkatkan performa model dengan data yang lebih beragam.

- **Framework Machine Learning:**

- TensorFlow/Keras: Untuk membangun dan melatih model CNN.
- Scikit-learn: Untuk evaluasi model menggunakan classification report dan confusion matrix.

## 4 RESULT AND DISCUSSION

- Performance metrics on Precision, Accuracy, Recall, F1-Score and Confusion Matrix).

| Label               | Precision | Recall | F1-Score    | Support    |
|---------------------|-----------|--------|-------------|------------|
| Digit 1 1m          | 1.00      | 0.86   | 0.93        | 66         |
| Digit 1 2m          | 1.00      | 0.64   | 0.78        | 66         |
| Digit 1 50cm        | 0.98      | 0.94   | 0.96        | 66         |
| Digit 2 1m          | 0.86      | 0.86   | 0.86        | 66         |
| Digit 2 2m          | 0.69      | 0.82   | 0.75        | 66         |
| Digit 2 50cm        | 0.93      | 0.98   | 0.96        | 66         |
| Digit 3 1m          | 0.83      | 0.79   | 0.81        | 66         |
| Digit 3 2m          | 0.73      | 0.74   | 0.74        | 66         |
| Digit 3 50cm        | 0.97      | 0.88   | 0.92        | 66         |
| Digit 4 1m          | 0.82      | 0.92   | 0.87        | 66         |
| Digit 4 2m          | 0.71      | 0.45   | 0.56        | 66         |
| Digit 4 50cm        | 0.80      | 0.98   | 0.88        | 66         |
| Digit 5 1m          | 0.97      | 1.00   | 0.99        | 66         |
| Digit 5 2m          | 0.63      | 0.97   | 0.77        | 66         |
| Digit 5 50cm        | 1.00      | 0.88   | 0.94        | 66         |
| <b>Accuracy</b>     |           |        | <b>0.85</b> | <b>990</b> |
| <b>Macro Avg</b>    | 0.86      | 0.85   | 0.85        | 990        |
| <b>Weighted Avg</b> | 0.86      | 0.85   | 0.85        | 990        |

Table 1: Precision, Recall, F1-Score, and Support for Each Label

Tabel 1 menunjukkan hasil evaluasi model klasifikasi berbasis precision, recall, F1-score, dan support untuk setiap label (kategori digit pada jarak tertentu), dengan akurasi total model mencapai 85%, serta rata-rata makro dan berbobot untuk metrik utama (precision, recall, F1-score) juga bernilai 85

| Metric                 | Value  |
|------------------------|--------|
| Tess Loss              | 0.4182 |
| Test Accuracy          | 0.8485 |
| Misclassification Rate | 0.1515 |

Table 2: Hasil Evaluasi Model

Tabel 2 menggambarkan hasil evaluasi model, di mana model memiliki nilai test loss sebesar 0.4182 yang menunjukkan tingkat error prediksi, akurasi pengujian sebesar 84.85% yang mencerminkan proporsi prediksi benar terhadap total data uji, serta tingkat kesalahan klasifikasi (misclassification rate) sebesar 15.15% yang merepresentasikan proporsi data yang salah diklasifikasikan oleh model.

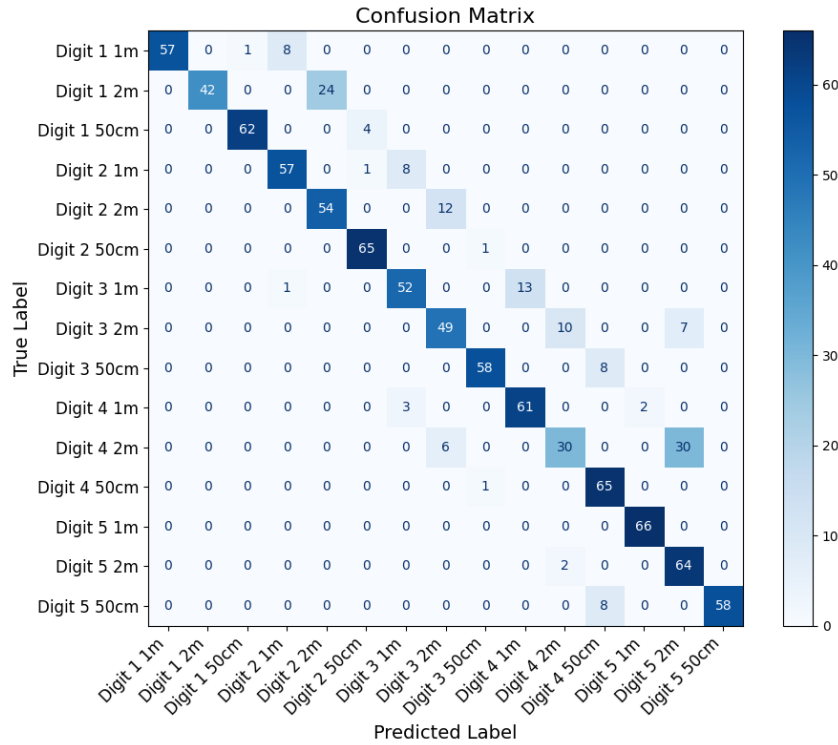


Figure 5: Confusion Matrix

Confusion matrix di atas menunjukkan performa model klasifikasi untuk mendeteksi digit pada berbagai jarak (1 meter, 2 meter, dan 50 cm), di mana model berhasil mengklasifikasikan sebagian besar label dengan benar tetapi masih terdapat beberapa kesalahan prediksi yang terlihat dari nilai non-diagonal, terutama pada True Label "Digit 4 dengan jarak 2 meter", yang memiliki distribusi kesalahan tertinggi dengan salah diklasifikasikan sebagai "Digit 5 dengan jarak 2 meter" sebanyak 30 kali. Dalam confusion matrix:

- \* Warna tergelap menunjukkan nilai prediksi yang paling akurat, yaitu jumlah prediksi yang benar (true positives) untuk label tertentu. Warna ini berada pada diagonal utama matriks (misalnya, "Digit 5 pada 1 meter" dengan 66 prediksi benar).
- \* Warna yang agak terang menunjukkan adanya kesalahan klasifikasi atau prediksi yang salah (false positives atau false negatives), yang berada di luar diagonal utama. Semakin terang warnanya, semakin sedikit jumlah kesalahan untuk kombinasi label tertentu (misalnya, "Digit 2 pada 1 meter" salah diprediksi sebagai "Digit 1 pada 1 meter" sebanyak 8 kali).

- **Visualization of Training and Validation results Accuracy Loss**

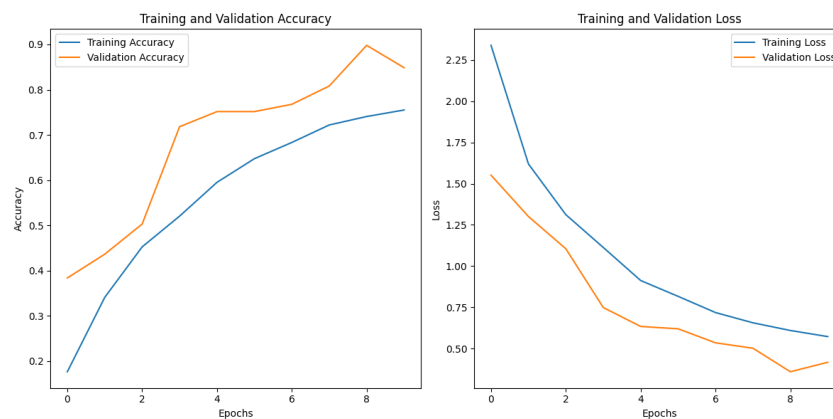


Figure 6: Training and Validation Accuracy & Loss

- \* Training and Validation Accuracy: Akurasi meningkat secara konsisten baik untuk data pelatihan maupun validasi seiring bertambahnya epoch. Namun, pada akhir pelatihan, akurasi validasi sedikit menurun, yang dapat menjadi indikasi overfitting.
- \* Training and Validation Loss: Loss untuk data pelatihan dan validasi terus menurun selama pelatihan, menunjukkan bahwa model belajar dengan baik. Namun, pada akhir pelatihan, loss validasi tampaknya mulai sedikit meningkat, yang juga mendukung adanya kemungkinan overfitting.

- **Discussion of the results, comparing them to related works**

Penelitian sebelumnya yang menggunakan radar berbasis Range-Doppler Map (RDM) dan pada proyek pembuatan dataset natural hand digit yang menggunakan model berbasis CNN untuk pengenalan digit tangan memiliki pendekatan berbeda dengan kelebihan dan kekurangannya masing-masing. Sistem berbasis radar dari penelitian sebelumnya mencapai akurasi tinggi sebesar 98% untuk pengenalan gestur dinamis. Keunggulan utamanya terletak pada privasi, karena radar tidak memerlukan kamera, sehingga cocok untuk lingkungan di mana sensitivitas terhadap pelanggaran privasi menjadi prioritas. Namun, radar memiliki keterbatasan, terutama dalam mendeteksi gerakan halus atau gestur mikro. Selain itu, pengaruh lingkungan terhadap performa radar masih perlu dipahami lebih jauh.

Di sisi lain, proyek ini lebih berfokus pada pengenalan digit tangan statis dari 1 hingga 5 menggunakan kamera dan dataset yang variatif, yang mencakup perubahan jarak, pencahayaan, dan sudut pandang. Dengan akurasi keseluruhan sebesar 85%, pendekatan ini menunjukkan performa yang cukup baik, terutama pada jarak dekat, di mana precision untuk beberapa digit mencapai nilai sempurna. Namun, performa pada jarak yang lebih jauh, seperti pada pengenalan digit 4 di jarak 2 meter, masih memerlukan perbaikan.

Keberagaman dataset yang digunakan dalam proyek ini menjadi salah satu kekuatan utama. Variasi dalam data ini memungkinkan model untuk lebih adaptif terhadap kondisi nyata, sesuatu yang kurang ditonjolkan dalam penelitian berbasis radar. Namun, pendekatan berbasis kamera juga memiliki kelemahan, terutama dalam aspek privasi dan kinerjanya di kondisi pencahayaan yang buruk.

Jika dibandingkan dengan karya-karya terkait lainnya, pendekatan radar seperti dalam penelitian sebelumnya sering dianggap unggul dalam privasi dan kesederhanaan sensor, tetapi terbatas dalam fleksibilitas dan resolusi detail. Sebaliknya, sistem berbasis kamera dan CNN seperti proyek ini memiliki potensi akurasi tinggi dan kemampuan untuk menangkap detail kompleks, namun dengan tantangan dalam menjaga privasi dan performa di lingkungan tertentu.

Dalam hal aplikasi, pendekatan radar lebih sesuai untuk sistem non-visual seperti kontrol kendaraan otonom atau perangkat rumah pintar. Sementara itu, proyek ini membuka peluang besar untuk aplikasi dalam antarmuka manusia-komputer (HCI) dan pengendalian perangkat berbasis IoT, yang membutuhkan pengenalan gestur digit tangan secara spesifik.

## 5 CONCLUSION

- **Recap of the objectives and how they were met**

Proyek ini bertujuan untuk mengembangkan dataset yang mendukung pengenalan gestur tangan, khususnya hand digit dari 1 hingga 5, untuk aplikasi berbasis interaksi manusia-komputer (HCI) dan Internet of Things (IoT). Tujuan tersebut tercapai melalui pengumpulan data yang beragam dalam kondisi pencahayaan dan jarak yang berbeda, serta menggunakan teknik machine learning, khususnya Convolutional Neural Networks (CNN), untuk melatih model pengenalan gestur. Dengan menyediakan dataset yang kaya dan bervariasi, proyek ini berhasil meningkatkan akurasi dalam mengenali gestur tangan di berbagai situasi nyata.

Hasil dari proyek menunjukkan bahwa model yang dilatih dengan dataset ini mencapai akurasi keseluruhan sebesar 85%. Metrik kinerja, termasuk precision dan recall, menunjukkan bahwa model dapat mengenali gestur dengan baik pada jarak yang berbeda. Misalnya, akurasi tertinggi dicapai pada digit 1 dengan jarak 50 cm, mencapai nilai 1.00 untuk precision. Selain itu, variasi dalam dataset memungkinkan model untuk lebih adaptif terhadap kondisi dunia nyata, termasuk pencahayaan rendah dan sudut pandang yang berbeda. Temuan ini menegaskan pentingnya keberagaman dalam dataset untuk meningkatkan performa model pengenalan gestur.

- **Key insights from the results**

- Digit 1 (1m, 50cm) menunjukkan precision dan F1-score yang sangat baik, mencapai 1.00 untuk precision dan 0.93-0.96 untuk F1-score, menunjukkan pengenalan yang sangat baik.
- Digit 2 (2m) memiliki precision yang lebih rendah (0.69) tetapi recall yang lebih baik (0.82), yang menunjukkan bahwa meskipun tidak terlalu presisi, model dapat mengenali banyak gambar dengan benar.
- Digit 4 (2m) menunjukkan nilai precision dan recall yang lebih rendah, terutama pada precision (0.71) dan recall (0.45), yang menunjukkan tantangan dalam mengenali gestur ini pada jarak 2m.

- Digit 5 (50cm) memiliki precision 1.00 dan F1-score yang tinggi, menunjukkan kinerja model yang sangat baik dalam mengenali gestur ini pada jarak dekat (50cm).
- Secara keseluruhan, Digit 1, 2, 3, dan 5 memiliki kinerja yang sangat baik, tetapi Digit 4 pada jarak 2m perlu ditingkatkan.

#### • Future work and recommendations

Ke depan, disarankan agar penelitian ini diperluas dengan menambahkan lebih banyak variasi dalam dataset, terutama menambah data untuk kelas-kelas yang memiliki akurasi rendah, seperti Digit 4 (2m), untuk meningkatkan keakuratan prediksi. Selain itu, melakukan augmentasi lebih lanjut pada kelas-kelas dengan F1-score rendah untuk memperbaiki kemampuan model dalam mengenali variasi gestur pada jarak tertentu. Penelitian lebih lanjut juga dapat mencoba model yang lebih kompleks seperti ResNet atau menggunakan transfer learning untuk meningkatkan kinerja, terutama pada kelas yang sulit dikenali. Dengan terus mengembangkan dan memperbaiki dataset serta metodologi pelatihan, potensi aplikasi teknologi pengenalan gestur tangan dapat dimaksimalkan di berbagai bidang.

## References

- [1] Bender, A., & Beller, S. Nature and culture of finger counting: Diversity and representational effects of an embodied cognitive tool. *Cognition* **124**(2), 156–182. DOI: <https://doi.org/10.1016/j.cognition.2012.05.005>.
- [2] Ahlawat, S., & Choudhary, A. Hybrid CNN-SVM Classifier for Handwritten Digit Recognition. *Procedia Computer Science* **167**, 2554–2560. DOI: <https://doi.org/10.1016/j.procs.2020.03.309>.
- [3] Jhaung, Y. C., Lin, Y. M., Zha, C., Leu, J. S., & Köppen, M. Implementing a Hand Gesture Recognition System Based on Range-Doppler Map. *Sensors* **22**(11). DOI: <https://doi.org/10.3390/s22114260>.