



**SISTEM DETEKSI SIMBOL PADA SIBI (SISTEM ISYARAT BAHASA
INDONESIA) SECARA *REAL TIME* MENGGUNAKAN *MEDIAPIPE* DAN
*LSTM***



TUGAS AKHIR

Program Studi

S1 TEKNIK KOMPUTER

UNIVERSITAS
Dinamika

Oleh:

F.X. Lorens Riberu

19410200037

FAKULTAS TEKNOLOGI DAN INFORMATIKA

UNIVERSITAS DINAMIKA

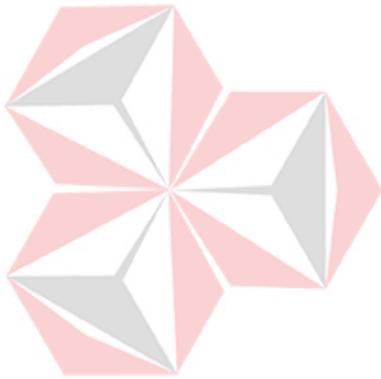
2023

**SISTEM DETEKSI SIMBOL PADA SIBI (SISTEM ISYARAT
BAHASA INDONEISA) SECARA *REAL TIME* MENGGUNAKAN
MEDIAPIPE DAN *LSTM***

TUGAS AKHIR

Diajukan sebagai salah satu syarat untuk menyelesaikan

Program Sarjana Teknik



UNIVERSITAS
Dinamika

Disusun Oleh :

Nama : F.X. Lorens Riberu
NIM : 19410200037
Program Studi : S1 Teknik Komputer

**FAKULTAS TEKNOLOGI DAN INFORMATIKA
UNIVERSITAS DINAMIKA**

2023

TUGAS AKHIR

SISTEM DETEKSI SIMBOL PADA SIBI (SISTEM ISYARAT BAHASA INDONESIA) SECARA *REAL TIME* MENGGUNAKAN *MEDIAPIPE* DAN *LSTM*

Dipersiapkan dan disusun oleh

F.X. Lorens Riberu

NIM: 19410200037

Telah diperiksa, diuji dan disetujui oleh Dewan Penguji

Pada: 6 Januari 2023

Susunan Dewan Pembahas

Pembimbing:

I. Heri Pratikno, M.T., MTCNA., MTCRE.

NIDN: 0716117302

II. Weny Indah Kusumawati, S.Kom., M.MT.

NIDN 0721047201

Pembahas:

Musayyanah S.S.T, M.T.

NIDN: 0730069102


Digitally signed
by Heri Pratikno
Date: 2023.01.06
23:17:24 +07'00'

Universitas
Dinamika
2023.01.06
12:13:06 +07'00'


Digitally signed by Musayyanah
DN: cn=Musayyanah, o=Universitas
Dinamika, ou=ST Teknik Informatika,
ou=Informatika, ou=Universitas Dinamika, ou=ID,
ou=ID
Date: 2023.01.06 12:13:06 +07'00'
KAB: 0730069102

Tugas Akhir ini telah diterima sebagai salah satu persyaratan

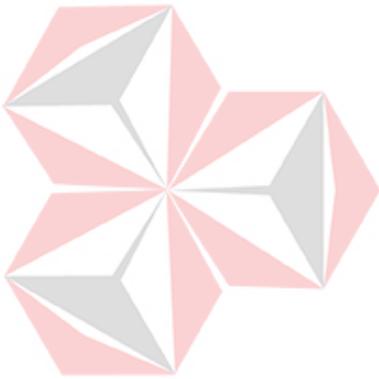
untuk memperoleh gelar sarjana


Digitally signed by
Universitas Dinamika
Date: 2023.01.18
07:21:40 +07'00'

Tri Sagirani, S.Kom., M.MT.

NIDN: 0731017601

Dekan Fakultas Teknologi dan Informatika
UNIVERSITAS DINAMIKA



“What do I wanna say is thank u” – Lorens

UNIVERSITAS
Dinamika



*Dipersembahkan kepada siapapun yang ingin belajar tentang kecerdasan buatan.
Tugas Akhir yang diharapkan dapat membantu untuk pengembangan teknologi ke
arah yang lebih baik bagi manusia.*

UNIVERSITAS
Dinamika

PERNYATAAN
PERSETUJUAN PUBLIKASI DAN KEASLIAN KARYA ILMIAH

Sebagai mahasiswa **Universitas Dinamika**, Saya :

Nama : F.X. Lorens Riberu
NIM : 19410200037
Program Studi : S1 Teknik Komputer
Fakultas : Fakultas Teknologi dan Informatika
Jenis Karya : Laporan Tugas Akhir
Judul Karya : **SISTEM DETEKSI SIMBOL PADA SIBI (SISTEM ISYARAT BAHASA INDONESIA) SECARA REAL TIME MENGGUNAKAN MEDIAPIPE DAN LSTM**

Menyatakan dengan sesungguhnya bahwa :

1. Demi pengembangan Ilmu Pengetahuan, Teknologi dan Seni, Saya menyetujui memberikan kepada **Universitas Dinamika** Hak Bebas Royalti Non-Eksklusif (*Non-Exclusive Royalty Free Right*) atas seluruh isi/sebagian karya ilmiah Saya tersebut diatas untuk disimpan, dialihmediakan, dan dikelola dalam bentuk pangkalan data (*database*) untuk selanjutnya didistribusikan atau dipublikasikan demi kepentingan akademis dengan tetap mencantumkan nama Saya sebagai penulis atau pencipta dan sebagai pemilik Hak Cipta.
2. Karya tersebut diatas adalah hasil karya asli Saya, bukan plagiat baik sebagian maupun keseluruhan. Kutipan, karya, atau pendapat orang lain yang ada dalam karya ilmiah ini semata-mata hanya sebagai rujukan yang dicantumkan dalam Daftar Pustaka Saya.
3. Apabila dikemudian hari ditemukan dan terbukti terdapat tindakan plagiasi pada karya ilmiah ini, maka Saya bersedia untuk menerima pencabutan terhadap gelar kesarjanaan yang telah diberikan kepada Saya.

Demikian surat pernyataan ini Saya buat dengat sebenar-benarnya.

Surabaya, 12 Desember 2022



F.X. Lorens Riberu
NIM : 19410200037

ABSTRAK

Penyandang tunarungu memiliki batasan dalam berkomunikasi secara verbal, dan solusi paling efektif yang dapat digunakan saat ini adalah bahasa isyarat. Di Indonesia sendiri terdapat bahasa isyarat yang digunakan secara resmi yaitu Sistem Isyarat Bahasa Indonesia (SIBI). Namun hanya sedikit orang non penyandang tunarungu yang menguasainya, sehingga akan tetap menyulitkan ketika berkomunikasi. Teknologi *deep learning* dapat menjadi salah satu solusi untuk mengatasi ketimpangan ini. Untuk memudahkan komunikasi antara penyandang tunarungu dengan non penyandang tunarungu yang belum mengetahui bahasa isyarat, maka dibuatlah sistem *deep learning* yang dapat mendeteksi SIBI. Pada penelitian ini digunakan *framework mediapipe* dan model *Long Short Term Memory* untuk membangun sistem SIBI *detection*. Hasilnya didapatkan nilai akurasi setelah training mencapai 100% dan *loss* sebesar $4.5115e-05$. Pada hasil pengujian performa didapatkan nilai *precision* 100%, *recall* 100%, dan *F1-Score* 100%. Sistem kemudian diuji *testing* atau prediksi secara *real time* dengan jarak 1,5, 3, dan 4 meter. Hasilnya nilai rata-rata persentase ketepatan prediksi adalah 83%. Namun prediksi pada jarak 1,5 meter saja memiliki persentase ketepatan prediksi sebesar 100%.

Kata Kunci: Bahasa Isyarat, *Deep Learning*, SIBI, Mediapipe, LSTM



UNIVERSITAS
Dinamika

KATA PENGANTAR

Puji syukur kami panjatkan kehadiran Tuhan Yang Maha Esa, karena atas rahmat dan karunianya-Nya sehingga pelaksanaan pengerjaan Tugas Akhir saya dapat terlaksana dengan baik dan lancar sesuai dengan jadwal yang direncanakan dan dapat diselesaikan tepat pada waktunya. Dalam kesempatan ini saya mengucapkan terima kasih kepada pihak-pihak yang telah banyak membantu dalam penyusunan dan pelaksanaan Tugas Akhir ini. Adapun penyusunan laporan Tugas Akhir ini bertujuan untuk memberikan pengetahuan dan membantu kaum disabilitas di dunia.

Laporan tugas ini, saya susun berdasarkan apa yang telah saya laksanakan di bulan agustus sampai bulan desember. Tugas Akhir ini bertujuan untuk membantu kaum disabilitas dan sebagai ilmu pengetahuan di dunia pendidikan.

Dalam penyusunan laporan ini, saya menyadari masih banyak kekurangan baik dari segi susunan serta cara penulisan laporan ini, karenanya saran dan kritik yang sifatnya membangun demi kesempurnaan laporan ini sangat saya harapkan. Semoga laporan ini bisa bermanfaat bagi para pembaca pada umumnya dan juga bermanfaat bagi penyusun pada khususnya.

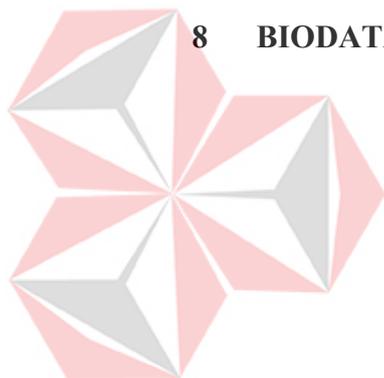
Surabaya, 12 Desember 2023

Penulis

DAFTAR ISI

	Halaman
ABSTRAK.....	vii
KATA PENGANTAR.....	viii
DAFTAR ISI.....	ix
DAFTAR GAMBAR.....	xi
DAFTAR TABEL	xii
DAFTAR LAMPIRAN	xiii
1 BAB I PENDAHULUAN.....	1
1.1 Latar Belakang.....	1
1.2 Rumusan Masalah.....	3
1.3 Batasan Masalah	3
1.4 Tujuan	3
1.5 Manfaat Penelitian	3
2 BAB II TINJAUAN PUSTAKA	5
2.1 Indeks Inklusivitas Indonesia.....	5
2.2 Tunarungu dan Tunawicara	6
2.3 SIBI (Sistem Isyarat Bahasa Indonesia)	7
2.4 <i>Mediapipe</i>	8
2.4.1 <i>Mediapipe Pose</i>	8
2.4.2 <i>Mediapipe Hand Tracking</i>	9
2.4.3 <i>Mediapipe Holistic</i>	10
2.5 Penelitian Sebelumnya.....	12
2.6 <i>Long Short Term Memory (LSTM)</i>	13
2.7 Menghitung Performa	14
2.7.1 Akurasi.....	15
2.7.2 Presisi.....	16
2.7.3 <i>Recall</i>	16
2.7.4 <i>F1 Score</i>	16
3 BAB III METODOLOGI PENELITIAN	17

3.1	Pengumpulan <i>Dataset</i>	17
3.2	Augmentasi Data.....	18
3.3	Proses Training Menggunakan LSTM.....	20
3.4	Proses <i>Testing</i>	21
4	BAB IV HASIL DAN PEMBAHASAN	23
4.1	Hasil Proses <i>Training</i> Model	23
4.2	Pengujian Performa.....	25
4.3	Prediksi	27
5	BAB V PENUTUP.....	30
5.1	Kesimpulan	30
5.2	Saran	30
6	DAFTAR PUSTAKA.....	32
7	LAMPIRAN.....	35
8	BIODATA PENULIS.....	46

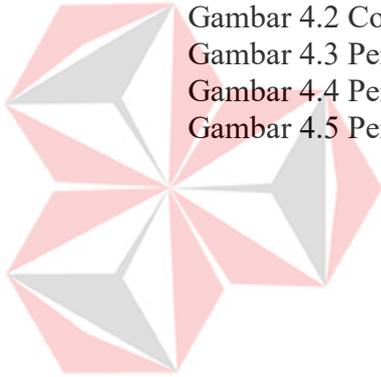


UNIVERSITAS
Dinamika

DAFTAR GAMBAR

Halaman

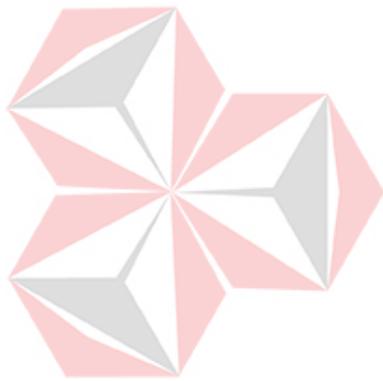
Gambar 2.1 Indeks Inklusivitas Global 2020.....	5
Gambar 2.2 Perbedaan abjad pada BISINDO dan SIBI	8
Gambar 2.3 33 key points Pose landmarks	9
Gambar 2.4 21 key points pada hand landmarks	10
Gambar 2.5 Mediapipe Holistic pipeline	11
Gambar 2.6 Arsitektur Long Short Term Memory (LSTM).....	14
Gambar 2.70 Confusion matrix.....	15
Gambar 3.1 Contoh dataset bahasa isyarat	17
Gambar 3.2 Flowchart proses augmentasi data.....	18
Gambar 3.3 Contoh dataset yang telah di proses random move	19
Gambar 3.4 Contoh dataset yang telah di proses horizontal flip dan random move	19
Gambar 3.5 Flowchart proses training.....	20
Gambar 3.6 Flowchart proses testing.....	22
Gambar 4.1 Hasil categorical accuracy (a) dan loss dari proses training (b).....	24
Gambar 4.2 Confusion matrix.....	26
Gambar 4.3 Pengujian sistem dengan jarak 1,5 meter.....	27
Gambar 4.4 Pengujian sistem dengan jarak 3 meter.....	27
Gambar 4.5 Pengujian sistem dengan jarak 4 meter.....	28



UNIVERSITAS
Dinamika

DAFTAR TABEL

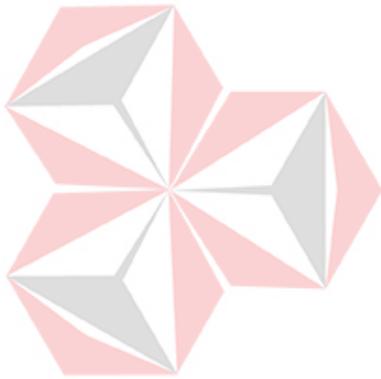
	Halaman
Tabel 3.1 Detail rangkuman dataset.....	18
Tabel 4.1 Hasil training model LSTM.....	23
Tabel 4.2 Hasil performa model LSTM.....	26
Tabel 4.3 Hasil prediksi model LSTM.....	29



UNIVERSITAS
Dinamika

DAFTAR LAMPIRAN

	Halaman
Lampiran 1 Program Augmentasi dan Ekstraksi Landmark	35
Lampiran 2 Training Model LSTM dengan Mediapipe.....	39
Lampiran 3 Program Testing SIBI Detection	42
Lampiran 4 Hasil Turnitin.....	45



UNIVERSITAS
Dinamika

BAB I

PENDAHULUAN

1.1 Latar Belakang

Penyandang disabilitas relatif berdasarkan non penyandang disabilitas di Indonesia masih mengalami kondisi yang memprihatinkan mulai dari permasalahan pendidikan, lapangan pekerjaan, penerimaan masyarakat, dan perlindungan sosial (BPS, 2020). Dua jenis disabilitas adalah tunarungu yang mengalami gangguan pendengaran dan tunawicara dengan gangguan berbicara. Penyandang tunarungu dan tunawicara sama-sama mengalami kesulitan dalam berkomunikasi secara verbal. Saat ini penggunaan bahasa isyarat menjadi solusi paling efektif bagi penyandang tunarungu dan tunawicara untuk berkomunikasi. Meski demikian kedua penyandang jenis disabilitas tersebut masih kesulitan untuk berkomunikasi dengan orang normal, karena tidak semua orang normal dapat memahami bahasa isyarat. Hal ini juga menyulitkan bagi mereka untuk berkomunikasi dan mendapatkan informasi di area publik.

Di Indonesia sendiri terdapat 2 jenis bahasa isyarat yang banyak digunakan yaitu Sistem Isyarat Bahasa Indonesia (SIBI) dan Bahasa Isyarat Indonesia (BISINDO). Namun hanya SIBI lah yang telah ditetapkan dan diakui oleh pemerintah. Disebutkan bahwa SIBI dibuat untuk merepresentasikan tata bahasa lisan Indonesia ke dalam isyarat buatan (Tempo, 2021). SIBI memiliki struktur yang sama dengan tata bahasa lisan Indonesia, seperti adanya awalan dan akhiran. Sedangkan BISINDO meskipun merupakan bahasa isyarat yang banyak digunakan dalam kehidupan sehari-hari namun memiliki variasi yang berbeda-beda pada tiap daerah. Dengan menggabungkan bahasa isyarat, bentuk tangan, orientasi gerak tangan serta ekspresi wajah yang dapat dipahami oleh penderita tunarungu dan orang normal sebagai media tambahan agar kedua belah pihak dapat berkomunikasi dengan baik (Dewi, et al., 2021).

Perkembangan dunia teknologi dewasa ini telah menciptakan berbagai inovasi yang lebih efektif dan efisien di berbagai bidang. Penggunaan *Machine Learning* dan *Deep Learning* pada pengenalan simbol Sistem Isyarat Bahasa Indonesia telah membuka jalan untuk penerapan *Artificial Intelligent* dalam

membantu komunikasi penyandang tunarungu di Indonesia. Pada penelitian yang dilakukan oleh (Putra, 2021) dan (Anam, 2021) menggunakan metode *Convolutional Neural Network* (CNN) dengan masing-masing arsitekturnya adalah 5 layer dan ResNet-50. Meski hasil akurasi *training* yang didapatkan tinggi yaitu masing-masing 90% dan 88%. Namun sistem masih belum bisa digunakan secara *realtime*. Pengembangan berikutnya dilakukan oleh (Hidayatullah, 2022) untuk membuat sistem yang dapat berjalan secara *realtime*. Pada penelitian ini digunakan model arsitektur MobileNet-SSDv2. Hasilnya didapatkan pada proses *training* dengan 20000 *step* memperoleh nilai akurasi mencapai 86,6% dan sistem dapat berjalan secara *realtime* dengan nilai rata-rata *Frame per Second* adalah 17,8 FPS. Hasil akurasi yang didapatkan pada penelitian ini cukup tinggi namun pada hasil deteksi akurasi, didapatkan bahwa akurasi untuk simbol ‘Cinta’ hanya sebesar 56,4%, sehingga dianggap belum cukup.

Sebagai upaya menemukan sistem yang lebih baik untuk pengenalan simbol pada Sistem Isyarat Bahasa Indonesia, penelitian ini akan menggunakan *Mediapipe* untuk melakukan *tracking* tubuh dan *Long Short Term Memory* sebagai model *Deep Learning*. Penggunaan *Mediapipe* dan *LSTM* pernah dilakukan terhadap BISINDO. Pada penelitian tersebut didapatkan akurasi deteksi pada hasil pengujian adalah 92%. Namun dataset yang digunakan merupakan video sehingga proses kalkulasinya menjadi lebih lama (Putri, et al., 2022).

Pengembangan sistem deteksi bahasa isyarat di bidang *machine learning* dan *deep learning* diharapkan dapat meningkatkan hak-hak penyandang disabilitas yang berdampak pada peningkatan kesejahteraan dan indeks pembangunan inklusif di Indonesia. Oleh karena itu, berdasarkan hasil penelitian-penelitian sebelumnya juga, perlu pengembangan lebih lanjut agar teknologi *deep learning* dapat digunakan dan diterapkan oleh penyandang tunarungu untuk berkomunikasi. Dimana pada penelitian ini akan menggunakan *Mediapipe* dan arsitektur *Long Short-Term Memory*, sementara dataset yang akan digunakan dalam bentuk citra gambar. Pada penelitian ini diharapkan dapat meningkatkan hasil akurasi dan nilai FPS ketika digunakan secara *realtime*.

1.2 Rumusan Masalah

Berdasarkan latar belakang diatas, maka dapat dirumuskan masalah pada Tugas Akhir ini sebagai berikut:

1. Bagaimana performa sistem deteksi simbol pada Sistem Isyarat Bahasa Indonesia menggunakan *Mediapipe* dan *Long Short Term Memory*?
2. Bagaimana kinerja proses deteksi simbol Sistem Isyarat Bahasa Indonesia melalui besarnya *Frame per Second* (FPS)?

1.3 Batasan Masalah

Dalam pembuatan Tugas Akhir ini, pembahasan masalah dibatasi pada beberapa hal berikut:

1. Pembatasan hanya 6 simbol Isyarat Bahasa Indonesia untuk *dataset*.
2. Pengujian secara *realtime* menggunakan kamera *webcam*.

1.4 Tujuan

Berdasarkan latar belakang dan rumusan masalah diatas, mendapatkan tujuan pada Tugas Akhir ini sebagai berikut:

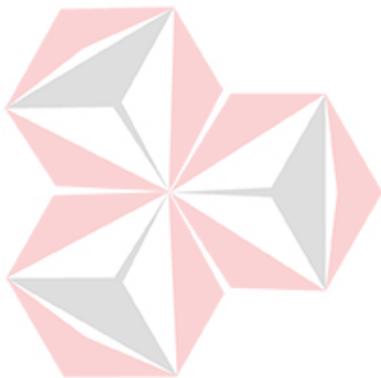
1. Mengetahui performa yang didapatkan pada sistem deteksi simbol Sistem Isyarat Bahasa Indonesia menggunakan *Mediapipe* dan *Long Short Term Memory*.
2. Mengetahui kinerja proses deteksi simbol Sistem Isyarat Bahasa Indonesia melalui besarnya *Frame Per Second* (FPS).

1.5 Manfaat Penelitian

Adapun dari Tugas Akhir ini dapat diperoleh manfaat sebagai berikut:

1. Bagi penulis yaitu untuk menambah pengetahuan dan penerapan mengenai deteksi simbol pada Sistem Isyarat Bahasa Indonesia menggunakan *Mediapipe* dan *Long Short Term Memory*.

2. Bagi mahasiswa yaitu menjadi referensi bagi yang akan melakukan penelitian mengenai deteksi Sistem Isyarat Bahasa Indonesia atau penggunaan model *Long Short Term Memory*.
3. Memudahkan penyandang tuna rungu untuk dapat berkomunikasi dengan orang normal.
4. Bagi pemerintah dapat menjadi solusi untuk pelayanan publik bagi penyandang tunarungu di area publik.



UNIVERSITAS
Dinamika

BAB II TINJAUAN PUSTAKA

2.1 Indeks Inklusivitas Indonesia

Indeks inklusivitas menjadi salah satu faktor acuan perkembangan disabilitas global. Indeks inklusivitas menjadi ukuran holistic dari pembangunan inklusif yang berfokus pada kesetaraan ras/etnis agama, gender, dan disabilitas di ranah perwakilan politik, kekerasan di luar kelompok, ketimpangan pendapatan, tingkat penahanan, dan kebijakan imigrasi dan pengungsi (Bappenas, 2021). Pada tahun 2020 Indonesia secara global menempati peringkat 115 dengan nilai 26,5 dalam pelaksanaan pembangunan inklusif (Menendian, et al., 2020).



Gambar 2.1 Indeks Inklusivitas Global 2020
Sumber: (Menendian, et al., 2020)

Peringkat Indonesia masih jauh tertinggal dari negara-negara maju seperti, Belanda, Selandia Baru, Swedia, UK, dll. Pada Gambar 2.1 diperlihatkan posisi Indonesia dari negara-negara ASEAN, peringkat Indonesia pun masih di bawah Filipina, Vietnam, Singapura, dan Thailand namun masih di atas Malaysia dan Myanmar.

2.2 Tunarungu dan Tunawicara

Tunarungu, seperti dijelaskan oleh Soewito pada buku *Ortho Paedagogik* adalah istilah dimana seseorang yang mengalami ketulian berat sampai total, yang tidak dapat menangkap tutur kata tanpa membaca bibir lawan bicaranya (Nofiaturrehman, 2018). Ketulian adalah kehilangan kemampuan mendengar sehingga proses informasi bahasa melalui pendengaran menjadi terhambat, baik memakai ataupun tidak memakai alat bantu dengar dimana batas pendengaran yang dimilikinya cukup memungkinkan keberhasilan proses informasi bahasa melalui pendengaran.

Komunikasi lisan adalah yang paling banyak digunakan oleh manusia dalam kehidupan sehari-hari. Kemampuan ini didapatkan dari proses imitasi dari mendengar orang lain, terutama lingkungan sekitarnya (Putra, et al., 2016). Bagi individu dengan pendengaran normal mungkin tidak mengalami hambatan dalam melakukan proses imitasi tersebut untuk membangun kemampuan berkomunikasi. Namun lain halnya dengan penyandang tunarungu yang tidak dapat membangun dan mengembangkan kemampuan tersebut. Terlebih bagi penyandang tunarungu sejak lahir, ia tidak bisa mendengar dan tidak bisa melakukan proses imitasi dari lingkungannya. Tunarungu sejak masih lahir inilah yang menyebabkan individu tidak mendapatkan suara yang ditiru sehingga mengakibatkan tunawicara. Tunawicara adalah seseorang yang mengalami kesulitan dalam mengungkapkan pikiran melalui Bahasa verbal, Sehingga dalam berkomunikasi harus menggunakan komunikasi nonverbal (Widiarti & Rahayu, 2020).

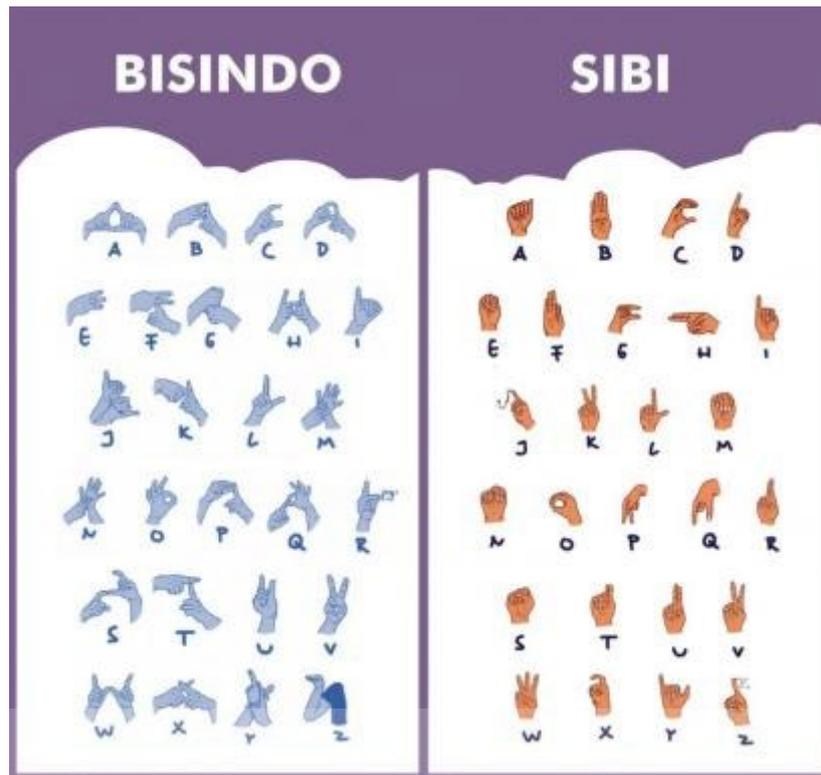
Alternatif media komunikasi yang banyak digunakan oleh penyandang tunarungu dan tunawicara adalah bahasa isyarat. Bahasa isyarat adalah Bahasa yang dilakukan dengan menggunakan gerakan-gerakan badan dan mimik muka sebagai simbol dari makna Bahasa. Di Indonesia, dalam perkembangannya terdapat 2 bahasa isyarat yaitu SIBI (Sistem Isyarat Bahasa Indonesia) dan BISINDO (Bahasa Isyarat Indonesia). BISINDO memiliki keberagaman di setiap daerahnya, sehingga bahasa isyarat ini memiliki perbedaan antara satu daerah dengan daerah lainnya. Dengan pertimbangan Bahasa isyarat yang banyak dan beragam, akhirnya tercipta sistem hasil rekayasa dari orang normal untuk berkomunikasi dengan para tunarungu yang disebut dengan Sistem Isyarat Bahasa Indonesia.

2.3 SIBI (Sistem Isyarat Bahasa Indonesia)

Bahasa Isyarat digunakan sebagai bahasa utama bagi penyandang tunarungu di seluruh dunia. Ini merupakan cara paling efektif untuk antar penyandang tunarungu berkomunikasi (Izzah & Suciati, 2014). Indonesia memiliki 2 bahasa isyarat yang umum digunakan, yaitu Bahasa Isyarat Indonesia (BISINDO) dan Sistem Isyarat Bahasa Indonesia (SIBI). Keduanya memiliki perbedaan gerakan dalam hal bahasa ataupun struktur kalimatnya.

BISINDO merupakan isyarat alamiah yang diciptakan dan digunakan sendiri oleh penyandang tunarungu dengan persepsi mereka terhadap segala sesuatu disekitar mereka. Hal-hal atau benda-benda yang dapat diikonisasi, akan dapat disimbolkan sesuai dengan bentuk, sifat, visual ataupun karakter dari suatu hal tersebut. Akan tetapi, jika hal tersebut terlalu abstrak dan tidak dapat diikonisasi, maka akan direpresentasikan menggunakan abjad jari (Wedayanti, 2019).

Sistem Isyarat Bahasa Indonesia (SIBI) adalah bahasa Isyarat yang ditetapkan oleh pemerintah sebagai bahasa yang resmi digunakan di Indonesia. SIBI merupakan bahasa yang memiliki struktur bahasa dan kalimat yang struktural. Dimana penggunaan subjek objek dan predikat diperhatikan. Dalam perkembangannya Sistem Isyarat Bahasa Indonesia (SIBI) masih menggunakan dari Bahasa isyarat luar negeri yang lalu ditambah dengan isyarat-isyarat lokal Indonesia dan buatan (Nasir, et al., 2021).



Gambar 2.2 Perbedaan abjad pada BISINDO dan SIBI
Sumber: (koran-jakarta.com)

2.4 Mediapipe

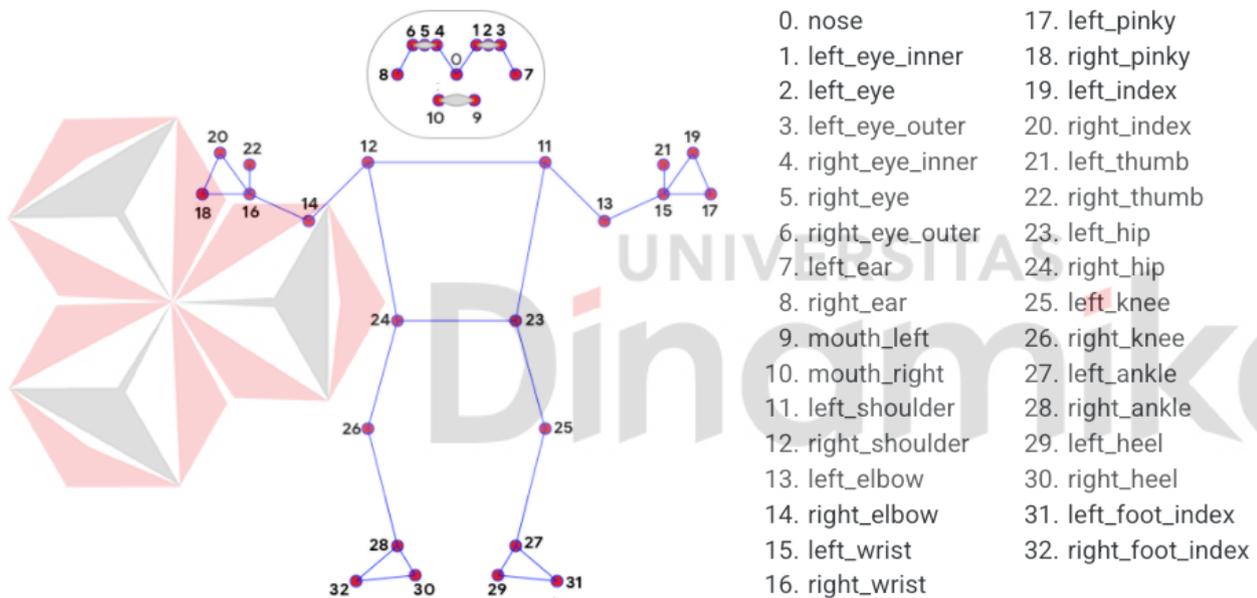
Mediapipe adalah sebuah *framework* yang di kembangkan oleh Google yang dapat membangun *pipelines* untuk memproses data persepsi dari berbagai format audio dan video. Diluncurkan pada 2019, *framework* ini menyediakan sekumpulan *machine learning solutions* seperti *Face Detection*, *Iris*, *Hair Segmentation*, *Holistic*, dan lainnya yang dapat dilihat pada web (MediapipeDev, 2019). *Solutions* yang disediakan kompatibel dengan sistem operasi android dan ios, serta dengan Bahasa C++, Python, JS, dan Coral.

2.4.1 Mediapipe Pose

Human pose estimation dari video telah memainkan peranan penting dalam berbagai teknologi aplikasi seperti mengukur latihan fisik, pengenalan bahasa isyarat, dan control gerakan seluruh tubuh. Misalnya, dapat menjadi dasar untuk aplikasi yoga, menari dan kebugaran. Itu juga dapat mengaktifkan overlay konten dan informasi digital diatas dunia fisik dalam *augmented reality*.

Detektor ini terinspirasi oleh model *BlazeFace* yang digunakan dalam Deteksi Wajah *MediaPipe*, sebagai proksi untuk pendeteksi orang (MediapipeDev, 2019). Ini secara eksplisit memprediksi dua titik kunci virtual tambahan yang dengan tegas menggambarkan pusat, rotasi, dan skala tubuh manusia sebagai lingkaran. Terinspirasi oleh pria *Vitruvian Leonardo*, untuk memperkirakan titik tengah pinggul seseorang, jari-jari lingkaran yang mengelilingi seluruh orang, dan sudut kemiringan garis yang menghubungkan titik tengah bahu dan pinggul.

Model landmark di *MediaPipe Pose* memprediksi lokasi 33 landmark pose. Secara opsional, *MediaPipe Pose* dapat topeng segmentasi seluruh tubuh yang direpresentasikan sebagai segmentasi dua kelas (manusia atau latar belakang).



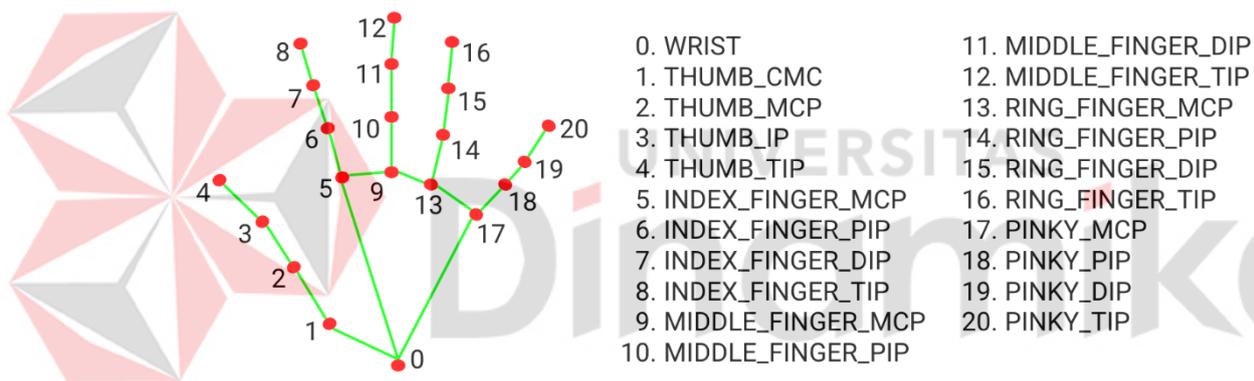
Gambar 2.3 33 *key points Pose landmarks*
(Sumber: MediapipeDev, 2019)

2.4.2 *Mediapipe Hand Tracking*

Kemampuan untuk merasakan bentuk dan gerakan tangan dapat menjadi komponen penting dalam meningkatkan pengalaman pengguna di berbagai domain dan platform teknologi. Misalnya, ini dapat membentuk dasar untuk pemahaman bahasa isyarat dan kontrol gerakan tangan, dan juga dapat memungkinkan *overlay* konten dan informasi digital di atas dunia fisik dalam *augmented reality*. Sementara datang secara alami, persepsi tangan *real-time* yang kuat adalah tugas *computer*

vision yang sangat menantang, karena tangan sering menutup diri sendiri atau satu sama lain (misalnya oklusi jari/telapak tangan dan jabat tangan) dan tidak memiliki pola kontras yang tinggi.

MediaPipe Hands adalah solusi pelacakan tangan dan jari dengan ketelitian tinggi. Ini menggunakan *Machine Learning* (ML) untuk menyimpulkan 21 *landmark* 3D tangan hanya dari satu bingkai. Sementara pendekatan canggih saat ini terutama bergantung pada lingkungan *desktop* yang kuat untuk inferensi, metode kami mencapai kinerja waktu nyata di ponsel, dan bahkan dapat diskalakan ke banyak tangan (*MediapipeDev*, 2019). Kami berharap dengan menyediakan fungsionalitas persepsi tangan ini ke komunitas penelitian dan pengembangan yang lebih luas akan menghasilkan munculnya kasus penggunaan kreatif, merangsang aplikasi baru, dan jalur penelitian baru.



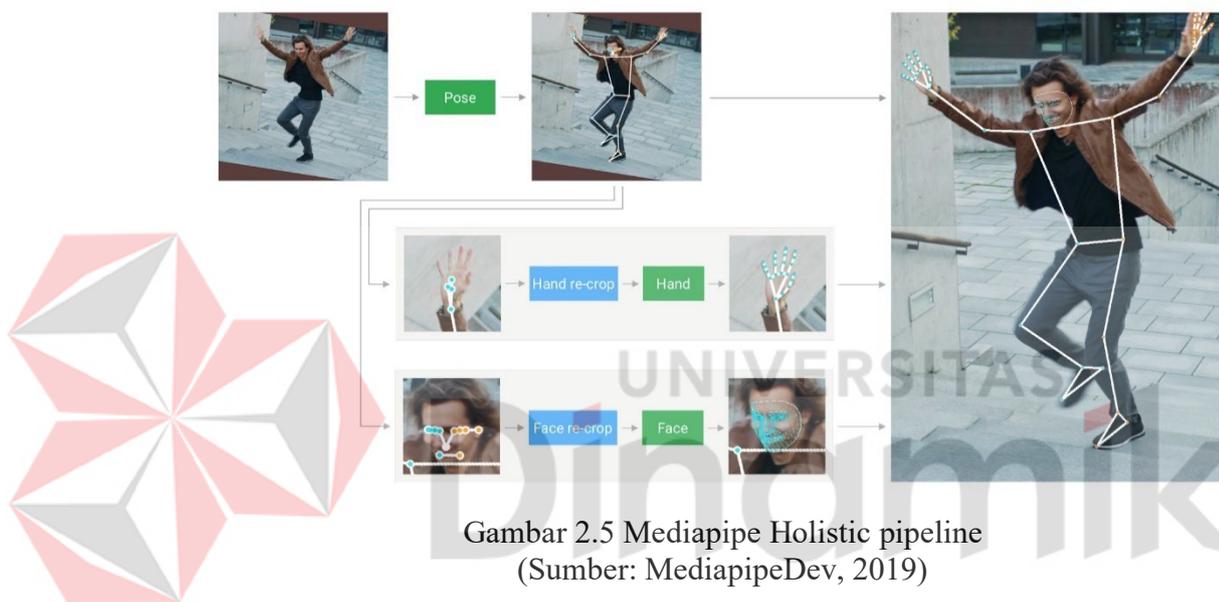
Gambar 2.4 21 *key points* pada *hand landmarks*
(Sumber: *MediapipeDev*, 2019)

2.4.3 *Mediapipe Holistic*

Mediapipe holistic mengintegrasikan beberapa model untuk komponen *pose*, wajah dan tangan, dimana masing-masing dioptimalkan untuk domain khususnya. Karena spesialisasi yang berbeda ini, *input* ke satu komponen menjadi tidak cocok untuk yang lain. Oleh karena itu, *Mediapipe holistic* dirancang dapat menangani berbagai wilayah menggunakan resolusi gambar yang sesuai.

Pertama, *mediapipe holistic* memperkirakan *pose* manusia dengan *detector pose BlazePose* dan model *landmark* berikutnya. Kemudian, dengan menggunakan *landmark pose inferred* diperoleh tiga *crop regions of interest* (ROI) untuk masing-

masing tangan, wajah, dan menggunakan model *re-crop* untuk meningkatkan ROI (MediapipeDev, 2019). Bingkai *input* resolusi penuh kemudian dipotong ke ROI dan menerapkan model wajah dan tangan khusus tugas untuk memperkirakan *landmark* yang sesuai. Semua *landmark* yang telah dideteksi kemudian digabungkan dengan model *pose*. *Mediapipe Holistic* menggunakan model *landmark pose*, wajah, dan tangan di *Mediapipe Pose*, *Mediapipe Face Mesh*, dan *Mediapipe Hands* masing-masing untuk menghasilkan total 543 *landmark* (33 *landmark pose*, 468 *landmark wajah*, dan 21 *landmark tangan* setiap tangan).



Gambar 2.5 Mediapipe Holistic pipeline
(Sumber: MediapipeDev, 2019)

Untuk merampingkan identifikasi ROI pada wajah dan tangan, maka digunakan pendekatan pelacakan yang serupa dengan yang digunakan untuk *pipeline* wajah dan tangan yang berdiri sendiri. Diasumsikan bahwa objek tidak bergerak secara signifikan di antara *frame* dan menggunakan estimasi dari *frame* sebelumnya sebagai panduan ke wilayah objek pada *frame* saat ini. Namun, selama gerakan cepat, pelacak dapat kehilangan target, sehingga detektor harus melokalkannya kembali di gambar. *MediaPipe Holistic* menggunakan prediksi *pose* (di setiap *frame*) sebagai ROI tambahan sebelum mengurangi waktu respon *pipeline* saat bereaksi terhadap gerakan cepat. Hal ini juga memungkinkan model mempertahankan konsistensi semantik di seluruh tubuh dan bagian-bagiannya

dengan mencegah campuran antara tangan kiri dan kanan atau bagian tubuh satu orang dalam bingkai dengan orang lain.

2.5 Penelitian Sebelumnya

Penelitian terhadap sistem deteksi bahasa isyarat telah banyak dilakukan dari berbagai negara dengan berbagai jenis bahasa isyarat. Di Indonesia penelitian dan pengembangan terhadap sistem deteksi SIBI (Sistem Isyarat Bahasa Indonesia) juga dilakukan sebagai upaya untuk membantu penyandang tunarungu mudah berkomunikasi dan berinteraksi. Terdapat 3 penelitian yang sebelumnya menjadikan sistem deteksi simbol pada SIBI sebagai topik penelitian menggunakan *deep learning* (Putra, 2021) (Anam, 2021) (Hidayatullah, 2022), dimana ketiganya mengusulkan metode yang berbeda-beda.

Penelitian yang dilakukan oleh (Putra, 2021) menggunakan metode *Convolutional Neural Network* (CNN) *basic* dengan 5 layer sebagai model *deep learning*. Terdapat 600 citra gambar yang terbagi kedalam 6 simbol yang menjadi dataset untuk *training*. Hasilnya didapatkan bahwa akurasi *training* mencapai angka 90% dan *validation* 91%.

Sementara itu penelitian yang dilakukan oleh (Anam, 2021) menggunakan metode *Convolutional Neural Network* (CNN) arsitektur *ResNet-50*, dengan bantuan dari Mediapipe untuk mendeteksi *landmark* pada tubuh. Model tersebut di-*training* dengan menggunakan 1200 citra gambar sebagai dataset yang terbagi kedalam 6 simbol. Meski didapatkan hasil akurasi training yang lebih rendah sebesar 88% dan *validation* 87%, namun sistem teruji pada dataset yang lebih beragam.

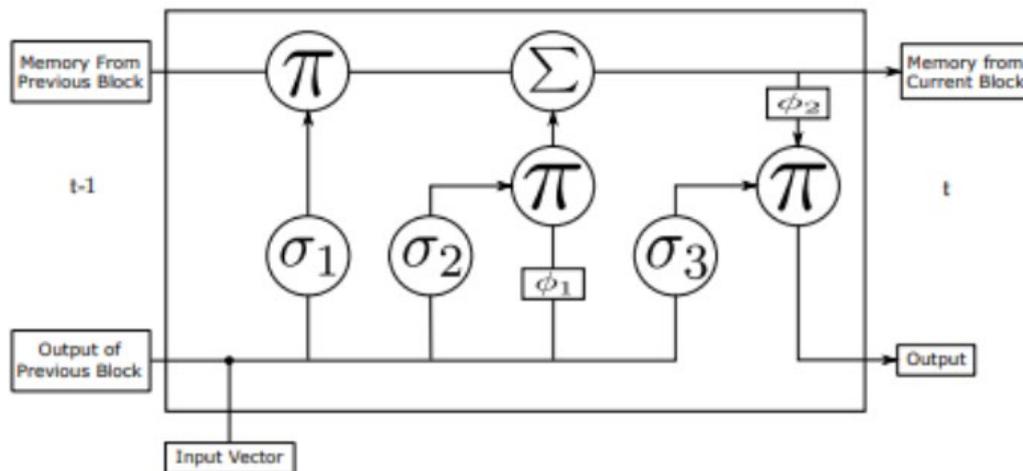
Penelitian selanjutnya oleh (Hidayatullah, 2022) telah memberikan pengembangan yang cukup signifikan, yaitu dengan menambahkan pengujian secara *realtime*. Hal tersebut memperbesar peluang penerapannya untuk penyandang tunarungu. Pada penelitian tersebut digunakan model arsitektur MobileNet-SSDv2, yang terdiri dari 2 jenis Neural Network yaitu *Single Shot Detector* (SSD) dan MobileNet. Hasilnya didapatkan pada proses *training* dengan 20000 *step* memperoleh nilai akurasi mencapai 86,6%.

Penelitian menggunakan *Mediapipe* dan *LSTM* juga pernah dilakukan oleh (Putri, et al., 2022). Pada penelitian tersebut menggunakan bahasa isyarat BISINDO dengan jumlah kosakata sebanyak 30. Dari hasil evaluasi deteksi *real-time* penelitian ini mendapatkan akurasi tertinggi sebanyak 92%. Namun dataset yang digunakan pada penelitian tersebut berupa video dengan setiap video memiliki 30 frame. Setiap frame menjadi *sequence* yang kemudian akan dipelajari oleh *LSTM*.

Pengembangan yang akan dilakukan pada penelitian ini adalah dengan menggunakan *Mediapipe* untuk mendeteksi landmark pada seluruh tubuh dan arsitektur *Long Short-Term Memory* sebagai model *deep learning*. Sementara itu dataset yang akan digunakan dalam bentuk citra gambar agar proses kalkulasi oleh sistem lebih cepat. Sehingga proses prediksi bukan membutuhkan 30 frame untuk setiap prediksinya, melainkan hanya butuh 1 frame saja. Dengan pengembangan ini diharapkan dapat meningkatkan hasil akurasi dan nilai FPS ketika digunakan secara *realtime*.

2.6 Long Short Term Memory (LSTM)

Jaringan *Long Short-Term Memory* merupakan bentuk khusus dari *Recurrent Neural Network* (RNN) yang memiliki *feedback connection*. LSTM baik dalam mengingat informasi untuk waktu yang lama. Karena lebih banyak informasi sebelumnya yang dapat mempengaruhi akurasi model (Kumar, et al., 2018). Jaringan LSTM merupakan jenis jaringan RNN yang mampu mempelajari urutan dalam urutan data. Ini merupakan arsitektur yang diperlukan dalam permasalahan kompleks seperti mesin terjemahan, pengenalan ucapan, dan banyak lagi (Browniee, 2021).



Gambar 2.6 Arsitektur *Long Short Term Memory* (LSTM)
 Sumber: (Kumar, et al., 2018)

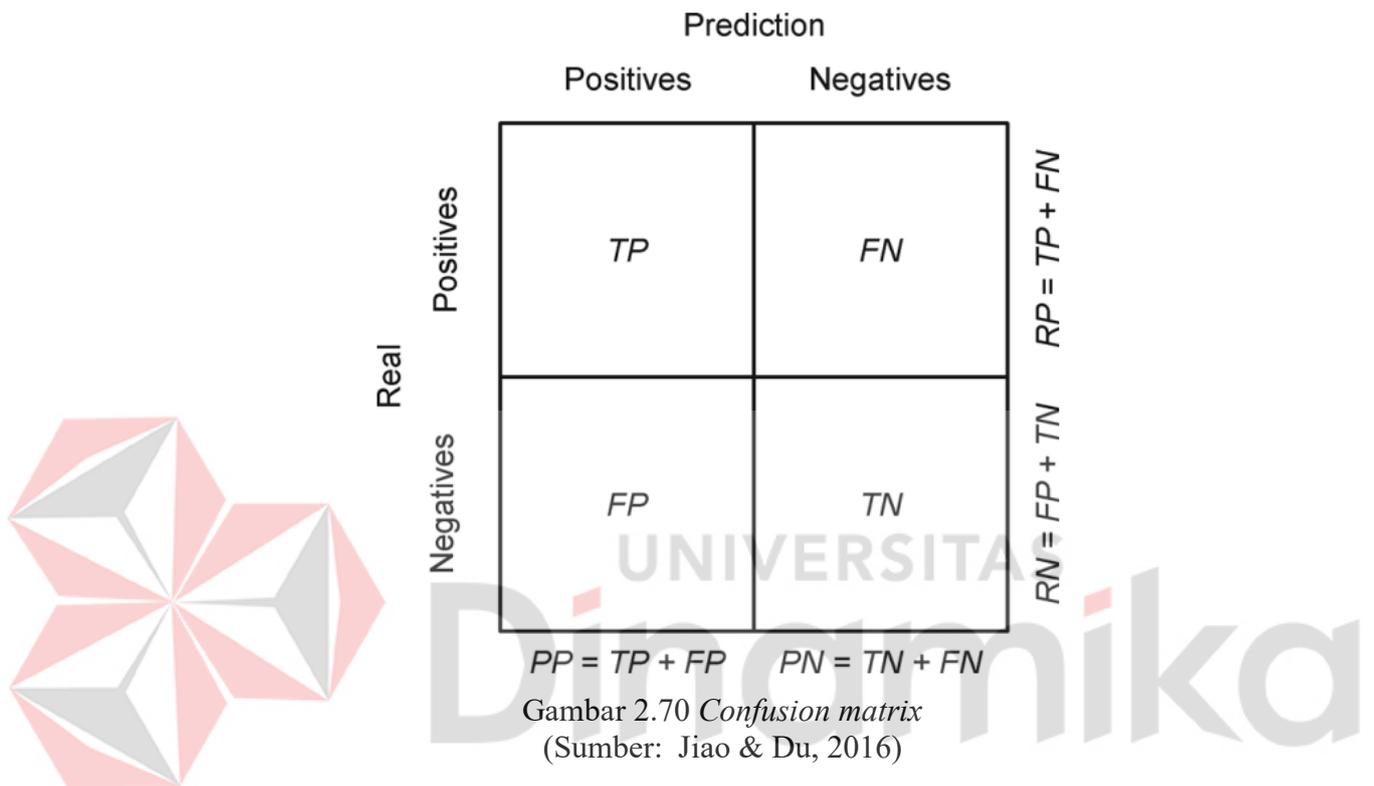
Pada Gambar 2.6 diperlihatkan bagaimana arsitektur *LSTM* menggunakan dua buah memory yaitu *Long Term Memory* (LTM) dan *Short Term Memory* (STM). Proses kalkulasi pada arsitektur LSTM menjadi lebih sederhana dan efektif karena penggunaan konsep *gates*. Terdapat empat jenis *gates* dengan kegunaannya masing-masing yaitu *Forget Gate*, *Learn Gate*, *Remember Gate*, dan *Use Gate*.

2.7 Menghitung Performa

Dalam menentukan atau menghitung performa deep learning, tidak cukup hanya dengan satu angka untuk mengindikasikan probabilitas mendapatkan prediksi yang benar. Pengguna mungkin hanya butuh hasil prediksi dari sampel data yang menurut mereka menarik atau memprediksi data dengan beberapa kondisi saja. Sehingga hasil ini tidak bisa digunakan untuk menentukan atau membandingkan performa. Untuk itu, telah dikembangkan berbagai tipe untuk mengevaluasi performa deep learning salah satunya *confusion matrix*.

Confusion matrix berisi informasi perbandingan antara nilai aktual dan prediksi dari klasifikasi sistem (Santra & Christy, 2012). Terdapat empat data yang umumnya digunakan untuk menghitung performa yaitu *true positives* (TP), *true negatives* (TN), *false positives* (FP), dan *false negatives* (FN) (Jiao & Du, 2016). Keempat data tersebut memiliki arti masing-masing. *True positive* memiliki hasil

prediksi dan nilai aktual positif. *False positives* memiliki nilai aktual positif dan hasil prediksi negatif. *False negative* berarti nilai aktual adalah negatif dan hasil prediksi adalah positif. *True negative* berarti memiliki nilai aktual dan prediksi negatif. Terdapat empat teknik untuk menghitung performa setelah diketahui confusion matrixnya yaitu akurasi, presisi, *recall*, dan *F1 score*.



2.7.1 Akurasi

Akurasi adalah ukuran perhitungan performa yang paling intuitif dan hanya rasio pengamatan yang diprediksi dengan benar terhadap total pengamatan. Akurasi merupakan ukuran yang bagus tetapi hanya jika model menggunakan data simetris dimana jumlah data pada setiap *class* adalah hampir sama. Perhitungan pada nilai akurasi akan menghasilkan nilai antara 0 dan 1, namun ini bisa diartikan juga dengan nilai persentase. Untuk menghitung akurasi menggunakan persamaan (2.1)

$$Akurasi = \frac{TP+TN}{TP+FP+FN+TN} \quad (2.1)$$

2.7.2 Presisi

Presisi adalah rasio prediksi positif yang benar terhadap keseluruhan prediksi positif. Presisi menunjukkan seberapa tepat model yang telah di-*training* dalam memprediksi kejadian positif. Untuk menghitung presisi menggunakan persamaan (2.2).

$$\text{Presisi} = \frac{TP}{TP+FP} \quad (2.2)$$

2.7.3 Recall

Recall adalah rasio prediksi positif yang benar terhadap seluruh data aktual yang positif. Untuk menghitung recall menggunakan persamaan (2.3).

$$\text{Recall} = \frac{TP}{TP+FN} \quad (2.3)$$

2.7.4 F1 Score

F1 Score adalah rata-rata bobot antara presisi dan *recall*. Itulah kenapa nilai ini menghitung *false positive* dan *false negative*. *F1 Score* lebih berguna dibandingkan akurasi terutama jika dataset yang digunakan jumlahnya tidak seimbang untuk setiap *class*-nya. Untuk menghitung *F1 Score* menggunakan persamaan (2.4).

$$F1 \text{ Score} = 2 \times \frac{(\text{recall} \times \text{presisi})}{(\text{recall} + \text{presisi})} \quad (2.4)$$

BAB III METODOLOGI PENELITIAN

3.1 Pengumpulan *Dataset*

Dataset yang digunakan pada penelitian ini merupakan citra gambar manusia yang memeragakan isyarat simbol SIBI (Sistem Isyarat Bahasa Indonesia). Secara keseluruhan terdapat 1500 citra yang terdiri dari 6 simbol (*classes*) dan diperagakan oleh 5 subjek yang berbeda. 3 simbol merupakan kata ganti orang yaitu ‘Aku’, ‘Dia’, dan ‘Kamu’ dan 3 simbol lainnya merupakan kata ganti perasaan ‘Cinta’, ‘Maaf’, dan ‘Sedih’. Menurut jenis kelamin subjek terdiri dari 4 orang laki-laki dan 1 orang perempuan. Contoh *dataset* dapat dilihat pada Gambar 3.1.



Gambar 3.1 Contoh dataset bahasa isyarat

Seluruh *Dataset* diambil menggunakan kamera *Webcam MiBo C1 USB Camera* dengan kualitas resolusi 4k UHD. *Dataset* diambil dengan ukuran resolusi 1280 x 720 dengan posisi *landscape*. Untuk mendapatkan keberagaman dataset maka setiap subjek dipotret dari lokasi yang berbeda-beda dan diminta untuk berpindah posisi relatif terhadap posisi kamera.

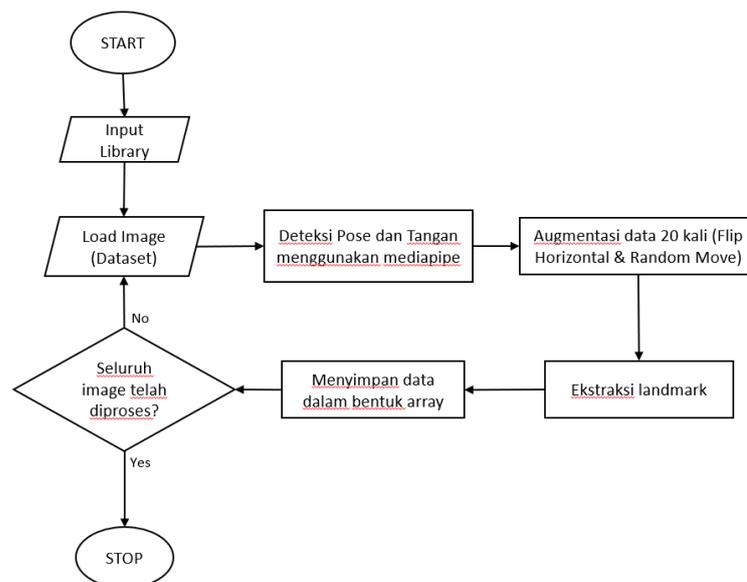
Pada proses training, dataset nantinya akan dibagi menjadi 2 bagian yaitu data *training* dan *validation* dengan perbandingan 80% dan 20%. Dari total 1500 citra gambar didapatkan 1200 sebagai data *training* dan 300 sebagai data *validation*. Tidak ketentuan yang baku dalam pembagian dataset, beberapa rasio pembagian

dataset yang umum digunakan seperti 80:20, 90:10, dan 70:10. Pemilihan rasio 80:20 dipilih karena merujuk dari prinsip Pareto yang menetapkan bahwa 80% konsekuensi berasal dari 20% penyebab, menegaskan hubungan yang tidak setara antara input dan output. Prinsip ini berfungsi sebagai pengingat umum bahwa hubungan antara input dan output tidak seimbang (Sander, 1987). Meskipun prinsip ini banyak digunakan pada bisnis dan ekonomi, namun dapat juga ditemukan penggunaannya pada bidang lain. Secara detail informasi mengenai *dataset* dapat dilihat pada Tabel 3.1.

Tabel 3.1 Detail rangkuman dataset

No	Subjek	Usia	Gender	Tanggal Ambil Data	Jumlah Data (Citra)
1	Subjek 1	23 Tahun	Laki-laki	24 Oktober 2022	Training: 240 Validation: 60
2	Subjek 2	22 Tahun	Perempuan	31 Oktober 2022	Training: 240 Validation: 60
3	Subjek 3	21 Tahun	Laki-laki	21 Oktober 2022	Training: 240 Validation: 60
4	Subjek 4	21 Tahun	Laki-laki	29 Oktober 2022	Training: 240 Validation: 60
5	Subjek 5	26 Tahun	Laki-laki	6 November 2022	Training: 240 Validation: 60

3.2 Augmentasi Data



Gambar 3.2 Flowchart proses augmentasi data

Augmentasi data adalah strategi yang memungkinkan praktisi untuk secara signifikan meningkatkan keragaman data yang tersedia untuk model pelatihan, tanpa benar-benar mengumpulkan data baru (Sanjaya & Ayub, 2020). Pada penelitian ini dilakukan 2 teknik augmentasi pada dataset yang telah disediakan yaitu *horizontal flip* dan *random move*, namun augmentasi ini hanya dilakukan pada setiap lokasi *landmark* yang telah terdeteksi, bukan pada citra. Setiap citra dilakukan proses augmentasi sebanyak 20 kali, sehingga didapatkan total 30000 data *landmark* untuk 1500 citra.



Gambar 3.3 Contoh dataset yang telah di proses *random move*

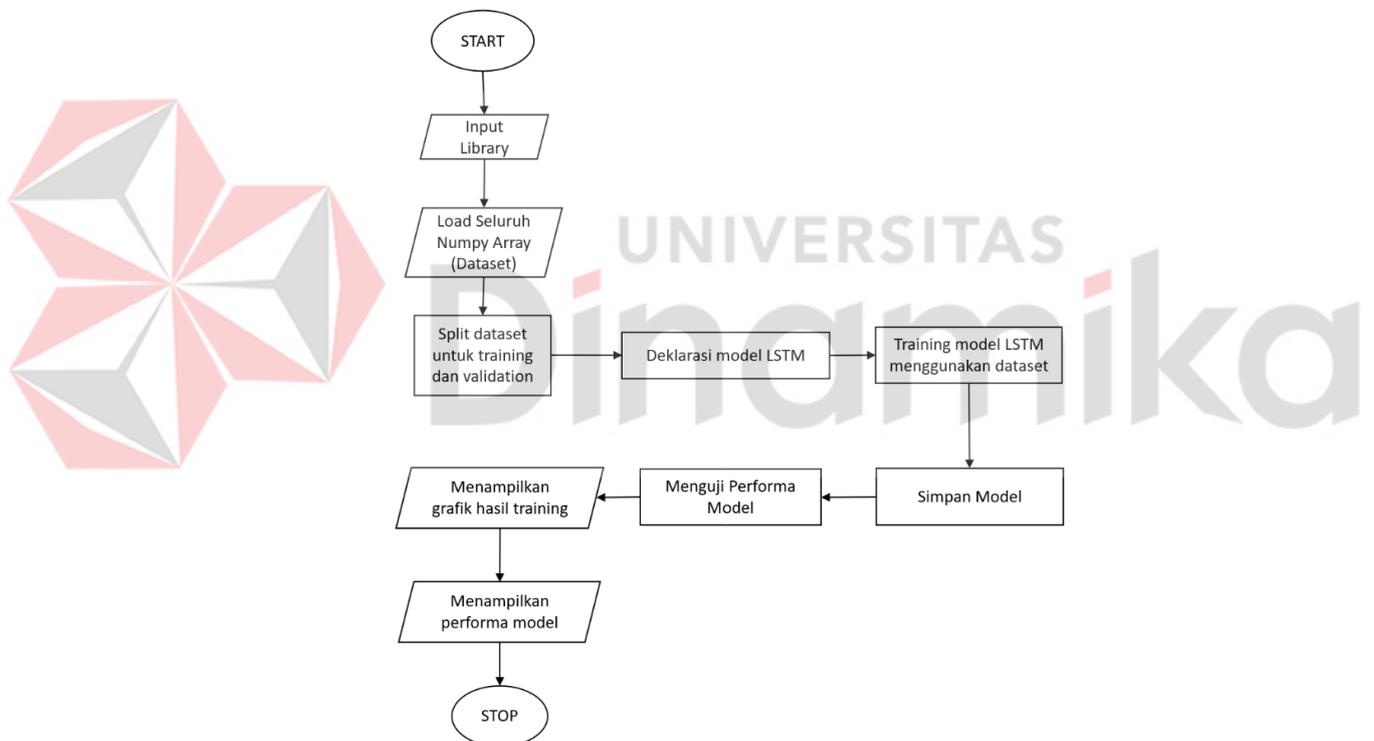


Gambar 3.4 Contoh dataset yang telah di proses *horizontal flip* dan *random move*

Teknik augmentasi ini tidak akan merubah citra asli namun akan merubah nilai lokasi x dan y setiap *landmark* yang telah dihasilkan oleh *mediapipe*. Dapat

dilihat pada Gambar 3.4 dimana posisi *landmark* bergeser terhadap posisi subjek sebenarnya. Perpindahan posisi x dan y pada landmark ini bergeser dengan jarak random oleh program namun tetap berada di dalam frame citra, sehingga tidak mungkin posisi *landmark* berada di luar atau melebihi citra. Sedangkan *flip horizontal* akan merubah nilai posisi x pada *landmark* yang dicerminkan terhadap garis tengah pada citra. Hasil flip horizontal kemudian di proses *random move*. Seluruh data baru hasil augmentasi kemudian disimpan dalam bentuk array numpy dan disimpan pada folder untuk nantinya menjadi data *training*.

3.3 Proses Training Menggunakan LSTM



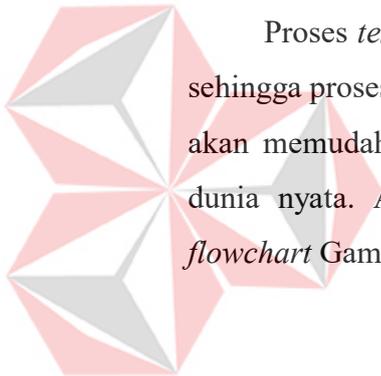
Gambar 3.5 *Flowchart* proses training

Pada Gambar 3.5 ditunjukkan algoritma program *Deep Learning* yang digunakan pada proses *training*. Pertama program menginputkan *library* yang dibutuhkan dan *path* direktori *dataset*. Selanjutnya seluruh dataset hasil ekstraksi landmark yang berupa numpy array di *load*. Seluruh dataset ini di *split* untuk *training* dan *validation* dengan rasio 80% dan 20%. Karena model yang akan

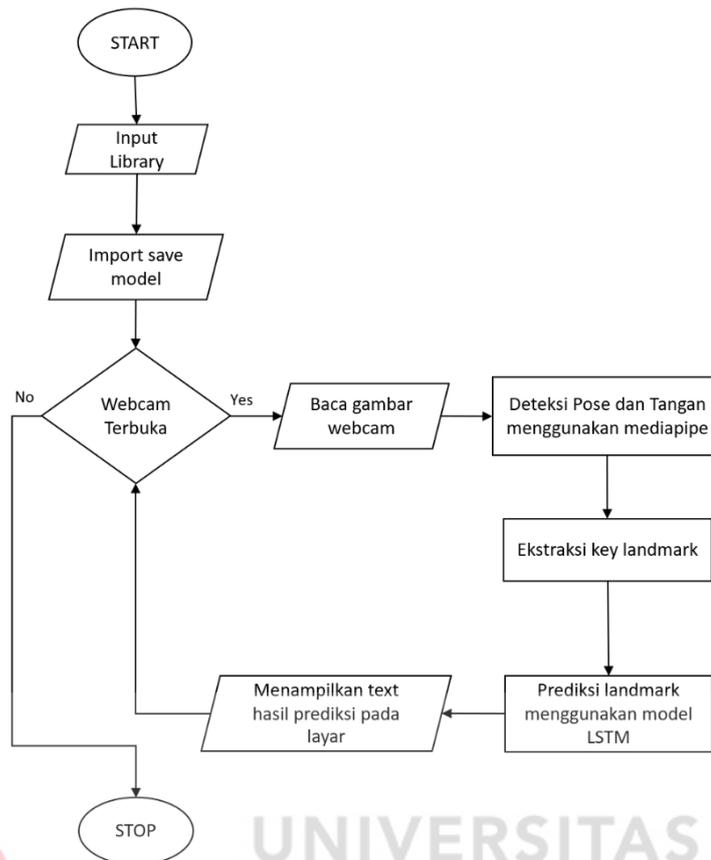
digunakan pada penelitian ini adalah LSTM, maka perlu dibentuk atau di susun model LSTM. Selanjutnya model yang telah dibentuk di training dengan seluruh 80% dataset yang telah di *load*. Model ini kemudian disimpan untuk nantinya dapat digunakan sebagai *transfer learning* pada aplikasi SIBI *detection*. Model kemudian diuji performanya dengan 20% dataset untuk melihat seberapa baik model menurut beberapa variabel uji performa. Grafik hasil training dan tabel hasil uji performa kemudian ditampilkan. Tidak ada persentase pembagian dataset yang optimal untuk *training* dan uji performa. Namun dengan jumlah data *training* yang lebih banyak akan memberikan kinerja model dengan variasi yang lebih banyak.

3.4 Proses *Testing*

Proses *testing* SIBI *detection* pada penelitian ini dilakukan secara *real time*, sehingga proses tersebut dapat dilakukan secara umum oleh siapapun. Hal ini tentu akan memudahkan dan semakin mendekati dalam penerapan SIBI *detection* di dunia nyata. Adapun proses yang dilakukan ketika *testing* ditunjukkan pada *flowchart* Gambar 3.6.



UNIVERSITAS
Dinamika



Gambar 3.6 Flowchart proses testing

Proses *testing* diawali dengan *import* beberapa *library* yang dibutuhkan, seperti *opencv* untuk buka *webcam* dan *mediapipe* untuk deteksi *landmark*. *Import* model yang sebelumnya telah di *training* menggunakan dataset yang telah disediakan. Apabila *webcam* terbuka oleh program kemudian citra dibaca. Setiap *key landmark* dideteksi menggunakan *mediapipe* lalu di ekstrak menjadi suatu data *numpy array*. Data tersebut kemudian diprediksi menggunakan model *LSTM* yang telah di *training*. Hasil prediksi ditampilkan pada layar bersamaan dengan citra yang ditangkap. Proses pembacaan citra, *tracking landmark*, dan prediksi terus dilakukan selama *webcam* terbuka.

BAB IV HASIL DAN PEMBAHASAN

4.1 Hasil Proses *Training* Model

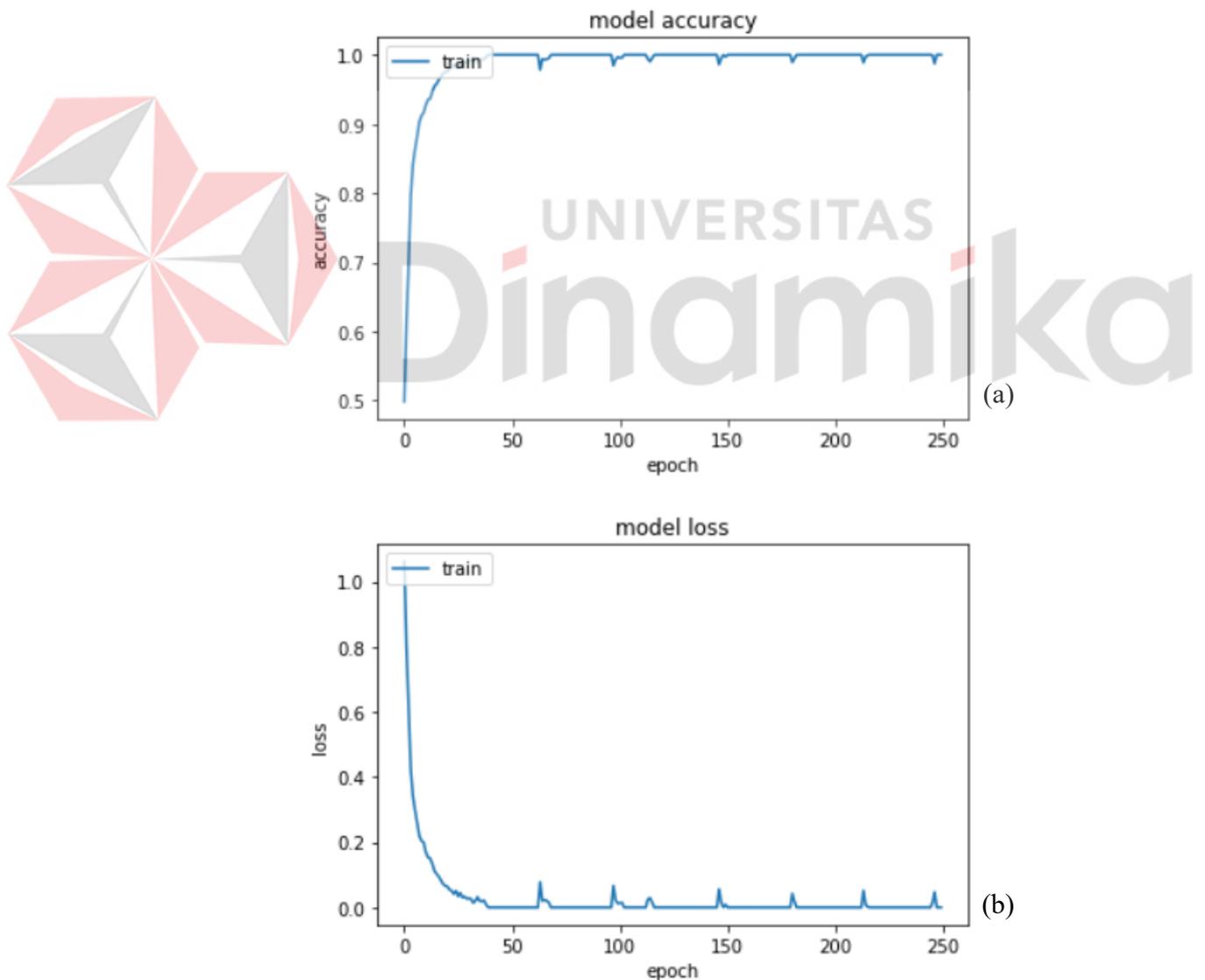
Proses *training* pada model LSTM dilakukan untuk agar model dapat mengenal dan membedakan keenam simbol SIBI pada dataset. Proses *training* dilakukan sebanyak 250 *epochs*, artinya model akan mempelajari seluruh dataset yang sama sebanyak 250 kali agar pengenalan mempelajari terhadap dataset dapat dikenali dengan baik, semakin banyak *epochs* yang digunakan akan semakin baik dalam pengenalan dataset saat *training*. Penentuan *epochs* sendiri tergantung berapa besar dataset yang dimiliki untuk dipelajari. Namun setelah dilakukan percobaan ternyata dengan 50 *epochs* sudah memiliki hasil yang baik. Sedangkan Optimizer yang digunakan pada *training* model adalah 'adam' dan *loss* 'categorical_crossentropy'. Hasilnya didapatkan pada *epochs* ke 50 memiliki akurasi *training* sebesar 100% dan *loss* sebesar 2.4741e-08, hasil ini tidak beda jauh dengan *epoch* ke 250 dengan akurasi *training* sebesar 100% dan *loss* sebesar 4.5115e-05. Sehingga proses *training* model LSTM dengan 50 *epoch* sudah dianggap cukup. Namun hasil akhir dengan *epoch* 250 yang akan disimpan sebagai file hasil *training*. Lebih detail untuk hasil *training* dapat dilihat pada Tabel 4.1.

Tabel 4.1 Hasil *training* model LSTM

Epochs	Categorical Accuracy (%)	Loss
1	49.89	1.0605
25	98.18	0.0413
50	100.00	2.4741e-08
75	100.00	4.8317e-06
100	99.66	0.0124
125	100.00	9.1500e-07
150	99.72	0.0088
175	100.00	5.8611e-10
200	100.00	1.5373e-08
225	100.00	6.5461e-07
250	100.00	4.5115e-05

Pada Tabel 4.1 ditunjukkan 2 hasil parameter yaitu *categorical accuracy* dan *loss*. Perhitungan pada nilai *categorical accuracy* akurasi akan menghasilkan nilai

antara 0 dan 1, namun ini bisa diartikan juga sebagai nilai persentase atau seperseratus, dimana semakin nilai prediksi mendekati 1 maka semakin cocok dengan nilai *actual*. Sementara *loss* merupakan angka untuk menunjukkan seberapa buruk prediksi model pada satu contoh. Pada Tabel 4.1 juga diperlihatkan nilai *categorical accuracy* dan *loss* pada setiap 25 *epoch* nya. Hasilnya didapatkan terjadi peningkatan yang sangat signifikan pada epoch 1 ke 25. Dimana nilai *categorical accuracy* meningkat dari 49.89% menjadi 98.18%. Hal yang sama juga pada nilai *loss* yang mengalami penurunan dari 1.0605 menjadi 0.0413. Hasil tersebut juga ditampilkan dalam bentuk grafik seperti pada Gambar 4.1.

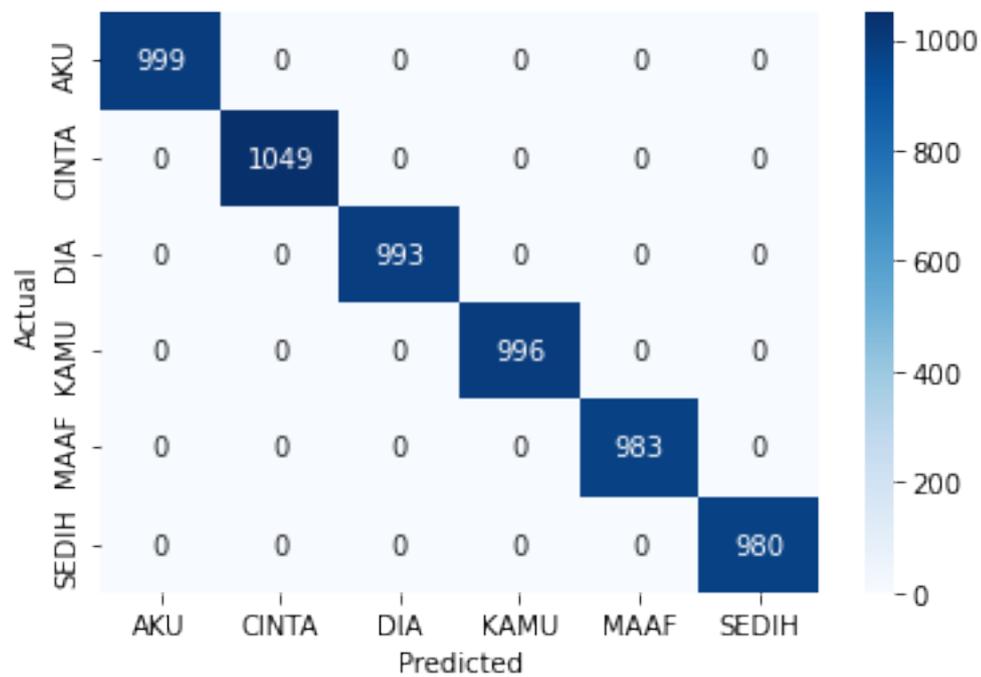


Gambar 4.1 Hasil *categorical accuracy* (a) dan *loss* dari proses *training* (b)

Pada Gambar 4.1 menunjukkan grafik proses *training* pada *categorical accuracy* mengalami kenaikan yang signifikan sejak pada epoch awal. Selanjutnya nilai tersebut berada atau mendekati 100%. Hal yang sama juga pada nilai *loss* dimana pada awal epoch telah mengalami penurunan yang signifikan dan semakin mendekati 0. Hal ini menunjukkan kalau model dapat mengklasifikasikan dataset dengan sempurna dan error yang dihasilkan sangat kecil. Hal ini tentu dipengaruhi oleh beberapa faktor seperti model yang digunakan dan jumlah dataset.

4.2 Pengujian Performa

Model yang telah di *training* kemudian diuji performanya. Pengujian performa merupakan membandingkan hasil prediksi model terhadap nilai *actual* nya. Untuk ini maka digunakan tabel *confusion matrix*, karena dapat menggambarkan dengan lebih detail jumlah data yang diklasifikasikan dengan benar. Pada penelitian ini didapatkan hasil yang sempurna, dimana seluruh dataset *validation* telah terklasifikasikan dengan tepat sesuai nilai *actual* nya seperti ditunjukkan oleh Gambar 4.2. Dimana sebanyak 999 data dengan nilai aktual 'AKU' dapat diprediksi sebagai 'AKU', 1049 data dengan nilai aktual 'CINTA' dapat diprediksi sebagai 'CINTA'. 993 data dengan nilai aktual 'DIA' dapat diprediksi sebagai 'DIA' dan selanjutnya.

Gambar 4.2 *Confusion matrix*

Setelah didapatkan hasil dari tabel *confusion matrix* maka dapat dihitung nilai *performance metrics* yang umum dan sering digunakan yaitu *accuracy*, *precision*, *recall*, dan *F1-score*. Pada Tabel 4.2 ditunjukkan nilai performa dari hasil training model LSTM. Didapatkan nilai setiap *performance metrics* dari nilai rata-rata keenam simbol (*class*) adalah *precision* 100%, *recall* 100%, *F1-score* 100%.

Tabel 4.2 Hasil performa model LSTM

	Precision	Recall	F1-Score	Support
“Aku”	1.00	1.00	1.00	999
“Cinta”	1.00	1.00	1.00	1049
“Dia”	1.00	1.00	1.00	993
“Kamu”	1.00	1.00	1.00	996
“Maaf”	1.00	1.00	1.00	983
“Sedih”	1.00	1.00	1.00	980
Accuracy			1.00	6000
Macro Avg	1.00	1.00	1.00	6000
Weighted Avg	1.00	1.00	1.00	6000

4.3 Prediksi

Model yang telah di-*training* disimpan dengan *extension file* h5, kemudian digunakan untuk melakukan *testing* atau prediksi secara *real time*. Prediksi dilakukan kepada 5 subjek yang berbeda dengan subjek pada dataset. Hal ini bertujuan untuk menguji sistem apakah dapat digunakan oleh banyak orang. Sistem juga diuji dengan 3 jarak yang berbeda yaitu 1,5, 3, dan 4 meter. Sehingga total ada 15 kali pengujian.



Gambar 4.3 Pengujian sistem dengan jarak 1,5 meter dan hasil prediksi benar



Gambar 4.4 Pengujian sistem dengan jarak 3 meter dan hasil prediksi benar



Gambar 4.5 Pengujian sistem dengan jarak 4 meter dan hasil prediksi salah

Dari hasil prediksi didapatkan rata-rata persentase ketepatan prediksi untuk keseluruhan adalah 83%, hal ini dikarenakan dataset yang disediakan memiliki rentang jarak 1-2 meter terhadap kamera. Sehingga pengujian pada jarak 3 dan 4 meter mengalami penurunan akurasi. Sementara rincian nilai persentase untuk masing-masing simbol adalah 'Aku' 86%, 'Kamu' 93%, 'Dia' 100%, 'Cinta' 100%, 'Maaf' 80% dan 'Sedih' 40%. Sebanyak 60% dari pengujian simbol 'Sedih' diprediksi sebagai simbol 'Cinta' dan seluruh kesalahan prediksi ini terjadi pada jarak 3 dan 4 meter. Selain karena dataset yang digunakan pada jarak 1-2 meter, hal ini juga disebabkan kemiripan antara simbol 'Cinta' dan 'Sedih'. Sementara itu prediksi pada jarak 1,5 meter saja memiliki persentase ketepatan prediksi sebesar 100%.

Proses prediksi juga menunjukkan nilai *Frame per Second* (FPS) pada setiap hasil prediksinya. Dari seluruh pediksi didapatkan nilai rata-ratanya adalah 30.5 FPS. Hasil ini sudah dianggap masuk kategori *realtime*.

Tabel 4.3 Hasil prediksi model LSTM

No	Subjek Uji		Jarak Subjek (meter)	AKURASI DETEKSI SIMBOL SIBI						FPS
				Aku	Kamu	Dia	Cinta	Maaf	Sedih	
1	Subjek 1	Video 1	1.5	Aku	Kamu	Dia	Cinta	Maaf	Sedih	32
		Video 2	3	Aku	Kamu	Dia	Cinta	Maaf	Cinta	34
		Video 3	4	Aku	Kamu	Dia	Cinta	Maaf	Cinta	27
2	Subjek 2	Video 1	1.5	Aku	Kamu	Dia	Cinta	Maaf	Sedih	33
		Video 2	3	Aku	Kamu	Dia	Cinta	Maaf	Cinta	33
		Video 3	4	Aku	Kamu	Dia	Cinta	Kamu	Cinta	31
3	Subjek 3	Video 1	1.5	Aku	Kamu	Dia	Cinta	Maaf	Sedih	28
		Video 2	3	Aku	Kamu	Dia	Cinta	Maaf	Cinta	29
		Video 3	4	Dia	Kamu	Dia	Cinta	Kamu	Cinta	27
4	Subjek 4	Video 1	1.5	Aku	Kamu	Dia	Cinta	Maaf	Sedih	32
		Video 2	3	Aku	Kamu	Dia	Cinta	Maaf	Cinta	32
		Video 3	4	Aku	Kamu	Dia	Cinta	Maaf	Cinta	29
5	Subjek 5	Video 1	1.5	Aku	Kamu	Dia	Cinta	Maaf	Sedih	31
		Video 2	3	Aku	Kamu	Dia	Cinta	Maaf	Cinta	32
		Video 3	4	Dia	Dia	Dia	Cinta	Sedih	Cinta	28
Rata-rata				Benar: 13 Salah: 2 Persentase: 86%	Benar: 14 Salah: 1 Persentase: 93%	Benar: 15 Salah: 0 Persentase: 100%	Benar: 15 Salah: 0 Persentase: 100%	Benar: 12 Salah: 3 Persentase: 80%	Benar: 6 Salah: 9 Persentase: 40%	30.5

BAB V PENUTUP

5.1 Kesimpulan

Dari hasil pengujian dan perbandingan data, didapatkan beberapa kesimpulan yaitu:

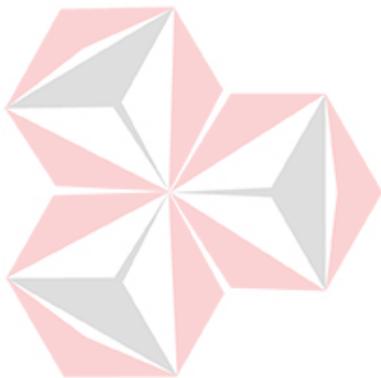
1. Hasil dari proses *training* menggunakan *framework mediapipe* dan model *Long Short Term Memory* dengan *250 epoch* mendapatkan nilai akurasi *training* sebesar 100% dan loss $4.5115e-05$ (0,00000045115).
2. Hasil dari proses pengujian *performance metrics* dengan 20% dataset didapatkan rata-rata nilai performa untuk keenam simbol adalah *Precision* 100%, *Recall* 100%, dan *F1-Score* 100%.
3. Hasil dari prediksi sistem model didapatkan nilai rata-rata persentase ketepatan prediksi adalah 83%, dengan rincian untuk masing-masing simbol sebagai berikut: 'Aku' 86%, 'Kamu' 93%, 'Dia' 100%, 'Cinta' 100%, 'Maaf' 80% dan 'Sedih' 40%.
4. Prediksi sistem model pada masing-masing jarak didapatkan hasil untuk jarak 1,5 meter memiliki akurasi deteksi sebesar 100%, pada jarak 3 meter 83% dan pada jarak 4 meter 66%.
5. Hasil dari pengujian FPS didapatkan nilai rata-ratanya adalah 30.5 FPS. Nilai tersebut sudah dianggap masuk kategori *realtime*.

5.2 Saran

Adapun saran untuk pengembangan pada penelitian ini agar lebih baik, terdapat beberapa saran yaitu:

1. Dataset yang digunakan pada penelitian ini kurang bervariasi. Untuk mendapatkan sistem yang lebih baik pada hasil pengujian maka dataset harus dibuat lebih bervariasi. Beberapa hal yang mungkin bisa menjadi pertimbangan adalah memperbanyak spesifikasi kamera untuk pengambilan dataset, memperbesar range jarak subjek dengan kamera, menggunakan orientasi *portrait* dan *landscape*.

2. Menambahkan jumlah simbol atau *class* pada sistem, agar lebih banyak simbol yang dapat diprediksi oleh sistem. Penambahan simbol tentu akan semakin mendekati pengembangan SIBI *detection* untuk digunakan oleh masyarakat.
3. Membuat aplikasi yang dapat menjalankan sistem SIBI *detection* dengan *user interface*, agar sistem dapat dijalankan dengan lebih mudah oleh masyarakat.



UNIVERSITAS
Dinamika

DAFTAR PUSTAKA

- Anam, N., 2021. Sistem Deteksi Simbol Pada SIBI (Sistem Isyarat Bahasa Indonesia) Menggunakan Mediapipe dan ResNet-50, Surabaya: Repository Dinamika.
- Bappenas, 2021. Tinjauan Peningkatan Akses dan Taraf Hidup Penyandang Disabilitas Indonesia: Aspek Sosioekonomi dan Yuridis, Jakarta: Kementerian PPN/Bappenas.
- BPS, 2020. Indikator Kesejahteraan Rakyat 2020, Jakarta: BPS-Statistics Indonesia.
- Browniee, J., 2021. A Gentle Introduction to Long Short-Term Memory Networks by the Experts. [Online] Available at: <https://machinelearningmastery.com/gentle-introduction-long-short-term-memory-networks-experts/> [Diakses 23 8 2022].
- Dewi, M., Wahyuningrum, T. & Prasetyo, N. A., 2021. Pengenalan Kata Bahasa Isyarat Indonesia (BISINDO) Menggunakan Augmented Reality (AR). *Journal of Informatics, Information System, Software Engineering and Applications*, Volume 3, pp. 53-60.
- Hidayatullah, A. M., 2022. Sistem Deteksi Simbol Pada SIBI (Sistem Isyarat Bahasa Indonesia) Secara Realtime Menggunakan MOBILENET-SSD, Surabaya: Repository Dinamika.
- Izzah, A. & Suciati, N., 2014. Translation of Sign Language Using Generic Fourier Description and Nearest Neighbour. *International Journal on Cybernetics & Informatics*, Volume 3, pp. 31-41.
- Jiao, Y. & Du, P., 2016. Performance Measures In Evaluating Machine Learning Based Bioinformatics Predictors for Classifications. *School of Computer Science and Technology*, 4(4), pp. 320-330.

- Kumar, J., Goomer, R. & Singh, A. K., 2018. Long Short Term Memory Recurrent Neural Network (LSTM-RNN) Based Workload Forecasting Model For Cloud Datacenters. *Procedia Computer Science*, Volume 125, pp. 676-682.
- MediapipeDev, 2019. Live ML anywhere. [Online] Available at: <https://google.github.io/mediapipe/> [Diakses 16 9 2021].
- Menendian, S., Elsheikh, E. & Gambhir, S., 2020. 2020 Inclusiveness Index : Measuring Global Inclusion And Marginality. United States: Othering & Belonging Institute.
- Nasir, M. C., Sudaryanto, E. & Kusumaningrum, H., 2021. Penggunaan Sistem Isyarat Bahasa Indonesia (SIBI) Sebagai Media Komunikasi. Konferensi Nasional.
- Nofiaturrehman, F., 2018. Problematika Anak Tunarungu dan Cara Mengatasinya. *Quality*, Volume 6, pp. 1-15.
- Putra, I. R. W., 2021. Sistem Deteksi Simbol Pada SIBI (Sistem Isyarat Bahasa Indonesia) Menggunakan Convolutional Neural Network, Surabaya: Universitas Dinamika.
- Putra, W. R. M., Azza, A. & Suryaningsih, A., 2016. Hubungan Kemampuan Komunikasi Verbal Dengan Adaptasi Sosial Anak Tuna Rungu Usia 6-12 Tahun di SLB-B Negeri Patrang Kabupaten Jember. *FIKes UNMUH Jember*, pp. 1-11.
- Putri, H. M., F. & Fuadi, W., 2022. Pendeteksian Bahasa Isyarat Indonesia Secara Real -Time Menggunakan Long Short-Term Memory (LSTM). *TTS4.0*, Issue Deteksi Bahasa Isyarat, pp. 13-25.
- Sander, R., 1987. The Pareto Principle: It's Use And Abuse. *Journal of Services Marketing*, 1(2), pp. 37-40.
- Sanjaya, J. & Ayub, M., 2020. Augmentasi Data Pengenalan Citra Mobil Menggunakan Pendekatan Random Crop, Rotate, dan Mixup. *Jurnal Teknik Informatika dan Sistem Informasi*, Volume 6, pp. 311-323.

Santra, A. K. & Christy, C. J., 2012. Genetic Algorithm and Confusion Matrix for Document Clustering. *JCSI International Journal of Computer Science Issues*, 9(1), pp. 322-328.

Tempo, 2021. Sama-sama Bahasa Isyarat, Apa Perbedaan SIBI dan BISINDO?. [Online]

Available at: <https://nasional.tempco.co/read/1535664/sama-sama-bahasa-isyarat-apa-perbedaan-sibi-dan-bisindo#:~:text=Perbedaan%20SIBI%20dan%20BISINDO%20yang,di%20kelompok%20Tuli%20masih%20terpecah.>

[Diakses 19 9 2022].

Wedayanti, N. P. L., 2019. Teman Tuli Diantara SIBI dan BISINDO. *Seminar Riset Linguistik Pengajaran Bahasa*, pp. 137-146.

Widiarti, A. & Rahayu, H., 2020. Pemberdayaan Penyandang Disabilitas Dalam Proses Pembangunan Di Indonesia Ditinjau Dari Perspektif Perundang-undangan. *Jurnal Ilmu Hukum*, 3(2), pp. 274-283.

Wikipedia, 2022. Long short-term memory. [Online]

Available at: https://en.wikipedia.org/wiki/Long_short-term_memory

[Diakses 23 8 2022].