

TUGAS AKHIR

**Rancang Bangun Sistem Pengenalan Gerakan Bibir pada Kata Dasar
Bahasa Indonesia Berbasis *LSTM***



FRISKA ADISTI MAHARDINI

NPM 22120081

Dosen Pembimbing

Dr. Ir. Anang Kukuh Adisusilo, ST., MT.

**PROGRAM STUDI INFORMATIKA
FAKULTAS TEKNIK
UNIVERSITAS WIJAYA KUSUMA SURABAYA
SURABAYA
2026**

TUGAS AKHIR

**Rancang Bangun Sistem Pengenalan Gerakan Bibir pada Kata Dasar
Bahasa Indonesia Berbasis *LSTM***



FRISKA ADISTI MAHARDINI

NPM 22120081

Dosen Pembimbing

Dr. Ir. Anang Kukuh Adisusilo, ST., MT.

**PROGRAM STUDI INFORMATIKA
FAKULTAS TEKNIK
UNIVERSITAS WIJAYA KUSUMA SURABAYA
SURABAYA
2026**

LEMBAR PENGESAHAN

Tugas Akhir Disusun Untuk Memenuhi Salah Satu

Syarat Memperoleh Gelar

Sarjana Komputer (S.Kom)

Di

Universitas Wijaya Kusuma Surabaya

Oleh :

FRISKA ADISTI MAHARDINI

22120081

Hari / Tanggal Sidang : Kamis, 16 Juli 2026

Menyetujui,
Dosen Pembimbing



Dr. Ir. Anang Kukuh Adisusilo, ST., MT.

NIP : 197802152015041001

Ketua Program Studi
Informatika



Nonot Wisnu Karyanto, ST., M.Kom.

NIK : 11563 - ET



Dekan
Fakultas Teknik

Dr. Ir. Anang Kukuh Adisusilo, ST., MT.

NIP : 197802152015041001

LEMBAR PENGESAHAN REVISI

Judul : Rancang Bangun Sistem Pengenalan Gerakan Bibir pada Kata Dasar Bahasa Indonesia Berbasis *LSTM*

Nama : Friska Adisti Mahardini

NPM : 22120081

Telah Diuji Pada :

Hari : Kamis

Tanggal : 16 Juli 2026

Tempat : Ruang Asistensi, Universitas Wijaya Kusuma Surabaya

Disetujui Oleh,

Dosen Penguji I



Ir. Maslihah, MT.
NIK 12643 - ET

Dosen Penguji II



Dr. Noven Indra Prasetya, S.Kom. M.Kom.
NIK 09414 - ET

Mengetahui,
Dosen Pembimbing



Dr. Ir. Anang Kukuh Adisusilo, ST., MT.
NIP 197802152015041001

PERNYATAAN KEASLIAN TULISAN

Yang bertanda tangan dibawah ini saya :

Nama : Friska Adisti Mahardini
NPM : 22120081
Program Studi : Informatika
Fakultas : Teknik
Universitas Wijaya Kusuma Surabaya

Menyatakan dengan sebenarnya bahwa Skripsi Tugas Akhir yang saya tulis dengan judul “Rancang Bangun Sistem Pengenalan Gerakan Bibir pada Kata Dasar Bahasa Indonesia Berbasis LSTM”, benar-benar hasil karya saya sendiri dan bukan merupakan pengambilan tulisan atau pikiran orang lain yang saya akui sebagai tulisan atau pikiran saya sendiri. Apabila di kemudian hari dapat dibuktikan bahwa Skripsi Tugas Akhir ini adalah hasil jiplakan, maka saya bersedia menerima sanksi atas perbuatan tersebut.

Surabaya, 21 Juli 2026

Yang membuat pernyataan,



Friska Adisti Mahardini

Rancang Bangun Sistem Pengenalan Gerakan Bibir pada Kata Dasar Bahasa Indonesia Berbasis *LSTM*

Friska Adisti Mahardini
Program Studi Informatika
Fakultas Teknik
Universitas Wijaya Kusuma Surabaya
fam081@mhs.uwks.ac.id

Abstrak

Komunikasi verbal merupakan aspek fundamental dalam interaksi manusia, namun efektivitasnya kerap terhambat oleh gangguan lingkungan seperti kebisingan atau keterbatasan fisiologis seperti gangguan pendengaran. Sebagai solusi alternatif, teknologi *Visual Speech Recognition (VSR)* atau *lip reading* dikembangkan untuk mengenali gerakan bibir tanpa bergantung pada sinyal audio. Meskipun teknologi *VSR* telah berkembang, penelitian *VSR* untuk Bahasa Indonesia masih terbatas karena ketersediaan *dataset* dan pengembangan model yang belum banyak dilakukan. Oleh karena itu, penelitian ini bertujuan untuk mengembangkan model pengenalan gerakan bibir pada kata dasar Bahasa Indonesia serta mengimplementasikannya ke dalam sistem berbasis web sebagai media pengujian model. Penelitian ini menggunakan *dataset* yang terdiri atas 898 video dari 12 kelas kata. Metode yang digunakan adalah pendekatan *deep learning* dengan *Long Short-Term Memory (LSTM)* yang dikombinasikan dengan *Convolutional Neural Network (CNN)*. *CNN* digunakan untuk mengekstraksi fitur spasial dari urutan *frame* area mulut, sedangkan *LSTM* digunakan untuk mempelajari pola temporal perubahan bentuk bibir selama proses pengucapan kata. Model yang telah dilatih kemudian diimplementasikan ke dalam sistem berbasis web dan diuji menggunakan 12 video uji serta rekaman langsung melalui *webcam*. Hasil penelitian menunjukkan bahwa model terbaik diperoleh pada *epoch* ke-51 dengan akurasi sebesar 92,22% pada data validasi, dengan *macro average precision* sebesar 0,93, *recall* sebesar 0,92, dan *F1-score* sebesar 0,92. Pada pengujian sistem, penggunaan video uji menghasilkan prediksi yang sesuai pada seluruh kelas kata, Sedangkan pengujian melalui *webcam* dengan maksimal tiga kali percobaan pada setiap kata menunjukkan bahwa kestabilan prediksi masih dipengaruhi oleh kualitas rekaman, pencahayaan, posisi wajah, kejelasan gerakan bibir, jarak terhadap kamera, waktu mulai dan berhenti perekaman, serta kemiripan *visem* pada bunyi huruf /m/, /b/, dan /p/.

Kata Kunci: *Visual Speech Recognition, Lip reading, LSTM, CNN, Bahasa Indonesia.*

KATA PENGANTAR

Segala puji dan syukur penulis panjatkan ke hadirat Allah SWT. atas limpahan rahmat, taufik, serta hidayah-Nya, sehingga penulis dapat menyelesaikan Tugas Akhir dengan baik. Laporan ini disusun sebagai salah satu syarat untuk memperoleh gelar Sarjana Komputer (S.Kom) pada Program Studi Informatika, Fakultas Teknik, Universitas Wijaya Kusuma Surabaya. Pada kesempatan ini, penulis menyampaikan penghargaan dan ucapan terima kasih yang sebesar-besarnya kepada semua pihak yang telah memberikan dukungan, baik secara moral maupun material, selama proses penyusunan proposal ini, khususnya kepada:

1. Allah SWT. yang senantiasa memberikan kemudahan, kekuatan, dan kelancaran dalam setiap langkah dalam proses ini.
2. Ibu, kakak, dan keluarga tercinta yang selalu memberikan doa, dukungan, serta motivasi yang tiada henti.
3. Bapak Dr. Ir. Anang Kukuh Adisusilo, ST., MT. selaku Dekan Fakultas Teknik Universitas Wijaya Kusuma Surabaya dan Dosen Pembimbing Tugas Akhir, yang telah memberikan bimbingan, arahan, serta motivasi selama penyusunan laporan ini.
4. Bapak Nonot Wisnu Karyanto, ST., M.Kom., selaku Ketua Program Studi Informatika Universitas Wijaya Kusuma Surabaya.
5. Seluruh Dosen Program Studi Informatika Universitas Wijaya Kusuma Surabaya yang telah membekali penulis dengan ilmu dan wawasan selama masa perkuliahan.
6. Sahabat dan teman-teman seperjuangan yang tidak dapat disebutkan satu per satu, terima kasih atas kebersamaan, dukungan, dan semangat yang selalu diberikan.
7. Partisipan yang telah berkontribusi dalam proses pengembangan *dataset* penelitian.
8. Para dosen penguji yang telah memberikan masukan berharga dalam menyempurnakan penelitian ini.

Penulis juga menyampaikan apresiasi kepada seluruh pihak lain yang telah berkontribusi namun tidak dapat disebutkan satu per satu. Penulis menyadari bahwa laporan ini masih memiliki keterbatasan, sehingga kritik dan saran yang konstruktif sangat diharapkan demi penyempurnaan penelitian. Akhir kata, semoga laporan ini dapat memberikan manfaat bagi para pembaca.

Surabaya, Juli 2026

Friska Adisti Mahardini

DAFTAR ISI

Abstrak	ii
KATA PENGANTAR.....	iii
DAFTAR ISI	iv
DAFTAR GAMBAR.....	vii
DAFTAR TABEL	ix
BAB 1 PENDAHULUAN.....	1
1.1 Latar Belakang	1
1.2 Rumusan Masalah.....	3
1.3 Batasan Masalah	3
1.4 Tujuan Penelitian	3
1.5 Manfaat Penelitian	4
1.6 Sistematika Penulisan	4
BAB 2 TINJAUAN PUSTAKA.....	5
2.1 Penelitian Terdahulu	5
2.2 <i>Computer Vision</i>	6
2.3 <i>Visual Speech Recognition</i>	6
2.4 Pengolahan Citra Digital dan Video	7
2.5 Ekstraksi Fitur Wajah dan <i>Region of Interest</i>	7
2.6 <i>Dataset</i>	8
2.7 <i>Machine Learning</i> dan <i>Neural Network</i>	9
2.8 <i>Deep Learning</i>	10
2.9 <i>Convolutional Neural Network</i>	10
2.10 <i>Recurrent Neural Network</i>	11
2.11 <i>Long Short-Term Memory</i>	11
2.12 Evaluasi Kinerja Model	13
2.13 Bagan Alir	15
2.14 <i>Unified Modeling Language</i>	15
2.15 Bahasa Pemrograman <i>Python</i>	16
2.16 Pustaka <i>OpenCV</i>	16
2.17 <i>Framework MediaPipe</i>	16
2.18 <i>Framework Tensorflow/Keras</i>	17
2.19 Pustaka <i>React</i>	17
2.20 <i>Framework FastAPI</i>	18
2.21 Pengujian Perangkat Lunak	18
BAB 3 METODOLOGI PENELITIAN	19

3.1	Pengumpulan Data	20
3.2	Pra-pemrosesan Data	21
3.2.1	Seleksi Data	21
3.2.2	Ekstraksi <i>ROI</i> Mulut	22
3.2.3	<i>Uniform Sampling</i>	22
3.2.4	Konversi <i>Grayscale</i>	22
3.2.5	Normalisasi Piksel	23
3.2.6	Standardisasi Piksel	23
3.2.7	Pembentukan <i>Dataset</i>	23
3.3	Perancangan Arsitektur Model	24
3.4	Pelatihan Model	25
3.5	Evaluasi Kinerja Model	26
3.5.1	<i>Classification Report</i>	26
3.5.2	<i>Confusion Matrix</i>	26
3.5.3	Kurva <i>ROC</i> dan Nilai <i>AUC</i>	27
3.6	Pengembangan Sistem Web	27
3.6.1	Halaman Utama	28
3.6.2	Fitur Pengujian dengan Video Uji	29
3.6.3	Fitur Pengujian dengan <i>Webcam</i>	32
3.7	Pengujian Sistem	34
3.7.1	Pengujian dengan Video Uji	34
3.7.2	Pengujian dengan <i>Webcam</i>	35
3.8	Lingkungan Pengembangan	35
3.9	Penyusunan Laporan	35
BAB 4 HASIL DAN PEMBAHASAN		36
4.1	Hasil Pengumpulan Data	36
4.2	Hasil Pra-pemrosesan Data	37
4.2.1	Hasil Seleksi Data	38
4.2.2	Hasil Ekstraksi <i>ROI</i> Mulut	39
4.2.3	Hasil <i>Uniform Sampling</i>	41
4.2.4	Hasil Konversi <i>Grayscale</i>	42
4.2.5	Hasil Normalisasi Piksel	43
4.2.6	Hasil Standardisasi Piksel	45
4.2.7	Hasil Pembentukan <i>Dataset</i>	46
4.3	Hasil Arsitektur Model	49
4.4	Hasil Pelatihan Model	50
4.5	Hasil Evaluasi Kinerja Model	55

4.5.1	<i>Classification Report</i>	55
4.5.2	<i>Confusion Matrix</i>	56
4.5.3	Kurva <i>ROC</i> dan Nilai <i>AUC</i>	57
4.6	Hasil Pengembangan Sistem Web	58
4.6.1	Halaman Utama	59
4.6.2	Halaman Pengujian dengan Video Uji	62
4.6.3	Halaman Pengujian dengan <i>Webcam</i>	68
4.7	Hasil Pengujian Sistem	75
4.7.1	Pengujian dengan Video Uji.....	75
4.7.2	Pengujian dengan <i>Webcam</i>	82
BAB 5 KESIMPULAN		92
5.1	Kesimpulan	92
5.2	Saran	92
DAFTAR PUSTAKA.....		93
LAMPIRAN		Lampiran 1

DAFTAR GAMBAR

Gambar 2.1 Representasi ekstraksi fitur wajah	8
Gambar 2.2 Arsitektur <i>Artificial Neural Networks</i>	9
Gambar 2.3 Arsitektur <i>Convolutional Neural Network</i>	10
Gambar 2.4 Arsitektur <i>Recurrent Neural Network</i>	11
Gambar 2.5 Arsitektur <i>Long Short-Term Memory</i>	12
Gambar 2.6 Simbol bagan alir.....	15
Gambar 2.7 <i>MediaPipe Face Mesh landmarks</i>	17
Gambar 3.1 Diagram alir pelaksanaan penelitian.....	19
Gambar 3.2 Ilustrasi posisi partisipan saat perekaman.....	21
Gambar 3.3 Alur pra-pemrosesan data	21
Gambar 3.4 Contoh hasil ekstraksi <i>Region of Interest (ROI)</i> pada area mulut	22
Gambar 3.5 Rancangan arsitektur model	24
Gambar 3.6 Arsitektur Sistem	27
Gambar 3.7 <i>Use Case Diagram</i>	28
Gambar 3.8 Rancangan Halaman Utama	29
Gambar 3.9 <i>Activity Diagram</i> Pengujian dengan Video Uji	30
Gambar 3.10 <i>Class Diagram</i> Fitur Pengujian dengan Video Uji.....	31
Gambar 3.11 Rancangan Halaman Pengujian dengan Video Uji.....	31
Gambar 3.12 <i>Activity Diagram</i> Pengujian dengan <i>Webcam</i>	32
Gambar 3.13 <i>Class Diagram</i> Fitur Pengujian dengan <i>Webcam</i>	33
Gambar 3.14 Rancangan Halaman Pengujian dengan <i>Webcam</i>	34
Gambar 4.1 Hasil Ekstraksi <i>ROI Mulut</i>	40
Gambar 4.2 Hasil <i>Uniform Sampling</i>	42
Gambar 4.3 Hasil Konversi <i>Grayscale</i>	43
Gambar 4.4 Hasil Normalisasi Piksel.....	44
Gambar 4.5 Histogram Nilai Piksel Sebelum dan Sesudah Normalisasi	44
Gambar 4.6 Hasil Standardisasi Piksel.....	46
Gambar 4.7 Histogram Nilai Piksel Sebelum dan Sesudah Standardisasi	46
Gambar 4.8 Distribusi <i>Dataset</i> Berdasarkan Kelas	48
Gambar 4.9 Distribusi Kelas pada Data <i>Training</i> dan Data Validasi.....	48
Gambar 4.10 Arsitektur Model.....	49
Gambar 4.11 Grafik <i>Accuracy</i> , <i>Loss</i> , dan <i>MSE</i>	54

Gambar 4.12 <i>Confusion Matrix</i>	56
Gambar 4.13 Kurva <i>ROC</i> dan Nilai <i>AUC</i>	57
Gambar 4.14 Struktur Penyimpanan <i>File</i> Sistem	58
Gambar 4.15 Hasil Implementasi Halaman Utama	60
Gambar 4.16 <i>Section</i> Hero dan Navigasi.....	60
Gambar 4.17 <i>Section</i> Informasi Sistem	61
Gambar 4.18 <i>Section</i> Alur Sistem	61
Gambar 4.19 <i>Section</i> Kata yang Didukung dan <i>CTA</i>	62
Gambar 4.20 Halaman Pengujian dengan Video Uji	63
Gambar 4.21 Daftar Video Uji	64
Gambar 4.22 Hasil Pengujian dengan Video Uji	67
Gambar 4.23 Halaman Pengujian dengan <i>Webcam</i>	68
Gambar 4.24 Hasil Tampilan Kamera Aktif	70
Gambar 4.25 Hasil Pengujian dengan <i>Webcam</i>	74
Gambar 4.26 Tampilan Pengujian <i>Webcam</i> Tidak Berhasil.....	74
Gambar 4.27 Hasil Pengujian Video Uji Kelas Apel	75
Gambar 4.28 Hasil Pengujian Video Uji Kelas Bantu	76
Gambar 4.29 Hasil Pengujian Video Uji Kelas Bola	76
Gambar 4.30 Hasil Pengujian Video Uji Kelas Halo	77
Gambar 4.31 Hasil Pengujian Video Uji Kelas Haus.....	77
Gambar 4.32 Hasil Pengujian Video Uji Kelas Jeruk	78
Gambar 4.33 Hasil Pengujian Video Uji Kelas Lapar.....	78
Gambar 4.34 Hasil Pengujian Video Uji Kelas Makan.....	79
Gambar 4.35 Hasil Pengujian Video Uji Kelas Minum	79
Gambar 4.36 Hasil Pengujian Video Uji Kelas Saya	80
Gambar 4.37 Hasil Pengujian Video Uji Kelas Tolong	80
Gambar 4.38 Hasil Pengujian Video Uji Kelas Tomat	81

DAFTAR TABEL

Tabel 4.1 Distribusi Data Mentah.....	36
Tabel 4.2 Distribusi Data Berdasarkan <i>Speaker</i>	37
Tabel 4.3 Hasil Seleksi Data	38
Tabel 4.4 Hasil Pembagian <i>Dataset</i>	48
Tabel 4.5 Ringkasan Arsitektur Model.....	50
Tabel 4.6 <i>History Model Training</i>	52
Tabel 4.7 Hasil <i>Classification Report</i>	55
Tabel 4.8 Ringkasan Hasil Pengujian dengan Video Uji	81
Tabel 4.9 Hasil Pengujian <i>Webcam</i> Kelas Apel	82
Tabel 4.10 Hasil Pengujian <i>Webcam</i> Kelas Bantu	83
Tabel 4.11 Hasil Pengujian <i>Webcam</i> Kelas Bola	83
Tabel 4.12 Hasil Pengujian <i>Webcam</i> Kelas Halo	84
Tabel 4.13 Hasil Pengujian <i>Webcam</i> Kelas Haus.....	84
Tabel 4.14 Hasil Pengujian <i>Webcam</i> Kelas Jeruk	85
Tabel 4.15 Hasil Pengujian <i>Webcam</i> Kelas Lapar	86
Tabel 4.16 Hasil Pengujian <i>Webcam</i> Kelas Makan.....	87
Tabel 4.17 Hasil Pengujian <i>Webcam</i> Kelas Minum	88
Tabel 4.18 Hasil Pengujian <i>Webcam</i> Kelas Saya.....	89
Tabel 4.19 Hasil Pengujian <i>Webcam</i> Kelas Tolong	89
Tabel 4.20 Hasil Pengujian <i>Webcam</i> Kelas Tomat	90
Tabel 4.21 Ringkasan Hasil Pengujian dengan <i>Webcam</i>	91