

**PREDIKSI KATA BAHASA ISYARAT INDONESIA
(BISINDO) BERBASIS VIDEO MENGGUNAKAN
MULTI-ISYARAT I3D DAN ST-GCN**



SKRIPSI

**CHREALVIN
00000045606**

**PROGRAM STUDI INFORMATIKA
FAKULTAS TEKNIK DAN INFORMATIKA
UNIVERSITAS MULTIMEDIA NUSANTARA
TANGERANG
2025**

**PREDIKSI KATA BAHASA ISYARAT INDONESIA
(BISINDO) BERBASIS VIDEO MENGGUNAKAN
MULTI-ISYARAT I3D DAN ST-GCN**



SKRIPSI

Diajukan sebagai salah satu syarat untuk memperoleh
Gelar Sarjana Komputer (S.Kom.)

**CHREALVIN
00000045606**

**PROGRAM STUDI INFORMATIKA
FAKULTAS TEKNIK DAN INFORMATIKA
UNIVERSITAS MULTIMEDIA NUSANTARA
TANGERANG
2025**

HALAMAN PERNYATAAN TIDAK PLAGIAT

HALAMAN PERNYATAAN TIDAK PLAGIAT

Dengan ini saya,

Nama : Chrealvin
Nomor Induk Mahasiswa : 00000045606
Program Studi : Informatika

Skripsi dengan judul:

Prediksi Kata Bahasa Isyarat Indonesia (BISINDO) Berdasarkan Video Menggunakan Multi-Isyarat I3D dan ST-GCN

merupakan hasil karya saya sendiri bukan plagiat dari laporan karya tulis ilmiah yang ditulis oleh orang lain, dan semua sumber, baik yang dikutip maupun dirujuk, telah saya nyatakan dengan benar serta dicantumkan di Daftar Pustaka.

Jika di kemudian hari terbukti ditemukan kecurangan/penyimpangan, baik dalam pelaksanaan maupun dalam penulisan laporan karya tulis ilmiah, saya bersedia menerima konsekuensi dinyatakan TIDAK LULUS untuk mata kuliah yang telah saya tempuh.

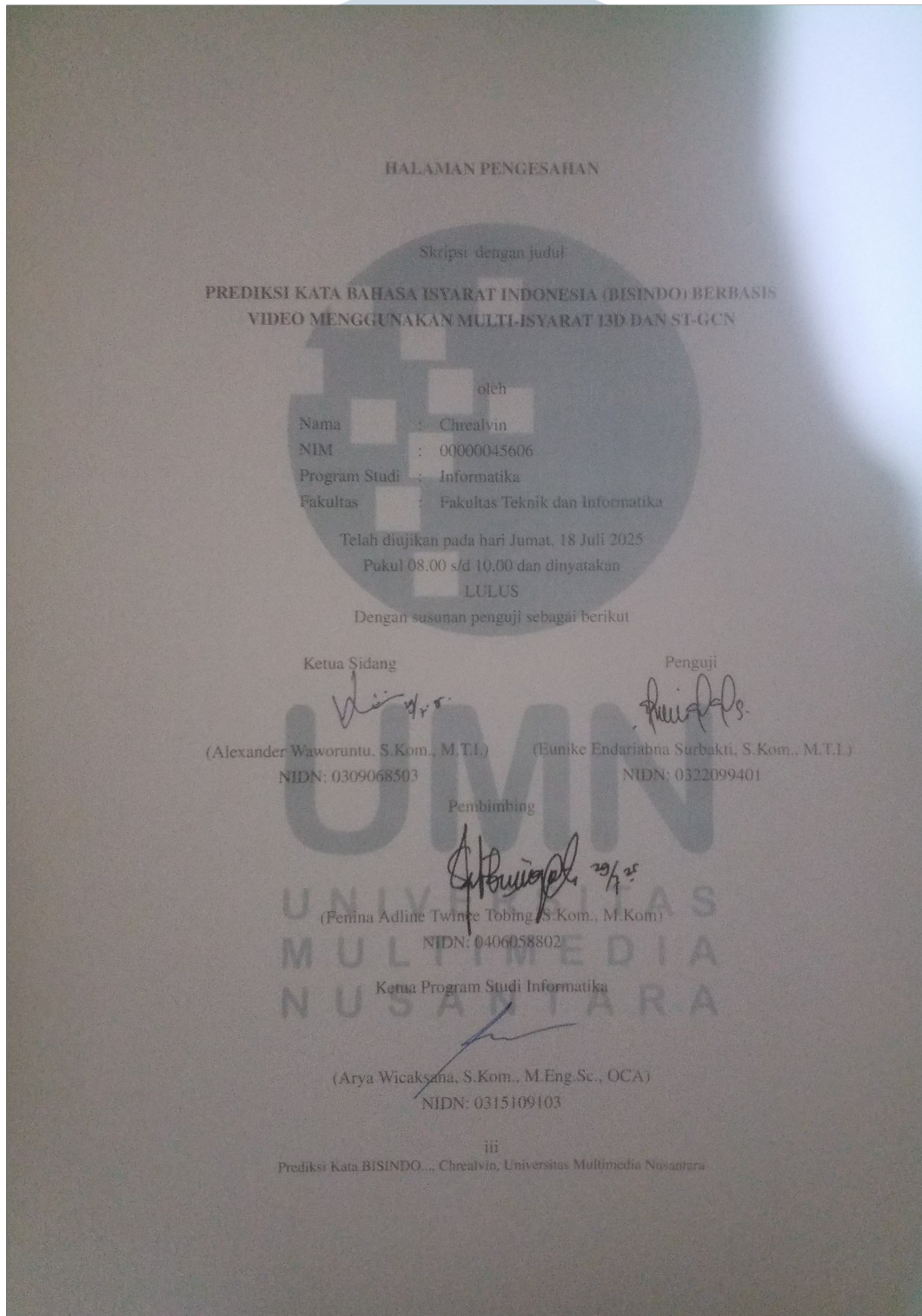
Tangerang, 4 Juli 2025



(Chrealvin)

UMN
UNIVERSITAS
MULTIMEDIA
NUSANTARA

HALAMAN PENGESAHAN



HALAMAN PERSETUJUAN PUBLIKASI

HALAMAN PERSETUJUAN PUBLIKASI KARYA ILMIAH UNTUK KEPENTINGAN AKADEMIS

Yang bertanda tangan di bawah ini:

Nama : Chrealvin
NIM : 00000045606
Program Studi : Informatika
Jenjang : S1
Judul Karya Ilmiah : Prediksi Kata Bahasa Isyarat
Indonesia (BISINDO) Berdasarkan
Video Menggunakan Multi-Isyarat
13D dan ST-GCN

Menyatakan dengan sesungguhnya bahwa saya bersedia (**pilih salah satu**):

- ☒ Saya bersedia memberikan izin sepenuhnya kepada Universitas Multimedia Nusantara untuk mempublikasikan hasil karya ilmiah saya ke dalam repositori Knowledge Center sehingga dapat diakses oleh Sivitas Akademika UMN/Publik. Saya menyatakan bahwa karya ilmiah yang saya buat tidak mengandung data yang bersifat konfidensial.
- ☐ Saya tidak bersedia mempublikasikan hasil karya ilmiah ini ke dalam repositori Knowledge Center, dikarenakan: dalam proses pengajuan publikasi ke jurnal/konferensi nasional/internasional (dibuktikan dengan *letter of acceptance*)^{***}.
- ☐ Lainnya, pilih salah satu:
 - Hanya dapat diakses secara internal Universitas Multimedia Nusantara
 - Embargo publikasi karya ilmiah dalam kurun waktu tiga tahun.

Tangerang, 4 Juli 2025

Yang menyatakan

Chrealvin

HALAMAN PERSEMBAHAN / MOTTO

"And one day, we'll look back to where we started, and be amazed by how far we've come"

Technoblade

UMN
UNIVERSITAS
MULTIMEDIA
NUSANTARA

KATA PENGANTAR

KATA PENGANTAR

Puji dan syukur penulis panjatkan ke hadirat Tuhan Yang Maha Esa karena atas rahmat dan karunia-Nya, penulis dapat menyelesaikan penyusunan skripsi yang berjudul Prediksi Kata Bahasa Isyarat Indonesia (BISINDO) Berdasarkan Video Menggunakan Multi-Isyarat 13D dan ST-GCN dengan baik. Skripsi ini disusun sebagai salah satu syarat untuk memperoleh gelar Sarjana di Universitas Multimedia Nusantara.

Penulis menyadari bahwa penyusunan skripsi ini tidak lepas dari dukungan, bimbingan, dan bantuan dari berbagai pihak. Oleh karena itu, penulis mengucapkan terima kasih kepada semua pihak yang telah membantu. Mengucapkan terima kasih kepada:

1. Bapak Dr. Ir. Andrey Andoko, M.Sc., selaku Rektor Universitas Multimedia Nusantara.
2. Bapak Dr. Eng. Niki Prastomo, S.T., M.Sc., selaku Dekan Fakultas Teknik dan Informatika Universitas Multimedia Nusantara.
3. Bapak Arya Wicaksana, S.Kom., M.Eng.Sc., OCA, selaku Ketua Program Studi Informatika Universitas Multimedia Nusantara.
4. Ibu Fenina Adline Twince Tobing, S.Kom., M.Kom, sebagai Pembimbing pertama yang telah memberikan bimbingan, arahan, dan motivasi atas terselesainya tugas akhir ini.
5. Keluarga saya yang telah memberikan bantuan dukungan material dan moral, sehingga penulis dapat menyelesaikan tugas akhir ini.

Semoga karya ilmiah ini dapat membantu penelitian bahasa isyarat terhadap komunitas tuna rungu dan tuna wicara agar dapat berkomunikasi dengan masyarakat luas

Tangerang, 4 Juli 2025



Chrealvin

PREDIKSI KATA BAHASA ISYARAT INDONESIA (BISINDO) BERBASIS VIDEO MENGGUNAKAN MULTI-ISYARAT I3D DAN ST-GCN

Chrealvin

ABSTRAK

Penelitian ini bertujuan untuk mengembangkan sistem prediksi bahasa isyarat indonesia (BISINDO) berbasis video pada level kata dengan mengintegrasikan fitur *multi-cue* yang terinspirasi pada penelitian sebelumnya. Metode yang dipakai menggabungkan model *Inflated 3D* (I3D) dan *Spatio-Temporal Graph Convolutional Network* (ST-GCN) untuk memprediksi berdasarkan fitur global pada sebuah video yang berupa video instruktur yang didapat menggunakan YOLOv11, pergerakan optik instruktur yang didapat melalui kalkulasi algoritma farneback dan TV-L1, dan pose tubuh yang diprediksi menggunakan MediaPipe. Tahap praproses meliputi deteksi instruktur menggunakan YOLOv11, estimasi pergerakan optik menggunakan metode farneback dan TV-L1, dan prediksi pose menggunakan MediaPipe. Prediksi dari setiap kombinasi model kemudian dirata-ratakan untuk menghasilkan hasil prediksi akhir. Sistem ini dievaluasi menggunakan dataset BISINDO yang terdiri dari 50 kata bahasa Indonesia. Sistem ini mencapai akurasi *top-1* dengan kombinasi rgb+farneback dan rgb+farneback+skeleton, *top-3* dengan kombinasi farneback dan rgb+farneback, dan *top-5* dengan kombinasi farneback masing-masing sebesar 40,82%, 63,27%, dan 77,55%.

Kata kunci: bahasa isyarat, bisindo, i3d, st-gcn



***VIDEO-BASED WORD-LEVEL INDONESIAN SIGN LANGUAGE
(BISINDO) RECOGNITION USING MULTI-CUE I3D AND ST-GCN***

Chrealvin

ABSTRACT

This study aims to develop a video-based word-level recognition for Indonesian Sign Language (BISINDO) by intergrating the multi-cue features inspired by prior research. The proposed method combines Inflated 3D (I3D) model and Spatio-Temporal Graph Convolutional Network (ST-GCN) to do a prediction base on global features: signer video, optical flow, and body posse. Preprocessing method involves signer detection using YOLOv11, optical flow estimation using farneback and TV-L1 algorithms, and pose prediction using MediaPipe. The prediction model will be averaged to create a final prediction. The model is evaluated using BISINDO dataset which consist of 50 Indonesian words. This system achieve top-1 with combinations of rgb+farneback and rgb+farneback+skeleton, top-3 with combinations of farneback and rgb+farneback, and top-5 with farneback combination accuracies of 40,82%, 63.27%, and 77.55%, respectively.

Keywords: bisindo, sign language, i3d, st-gcn



DAFTAR ISI

HALAMAN JUDUL	i
PERNYATAAN TIDAK MELAKUKAN PLAGIAT	ii
HALAMAN PENGESAHAN	iii
HALAMAN PERSETUJUAN PUBLIKASI KARYA ILMIAH	iv
HALAMAN PERSEMBAHAN/MOTO	v
KATA PENGANTAR	vi
ABSTRAK	vii
ABSTRACT	viii
DAFTAR ISI	ix
DAFTAR TABEL	xi
DAFTAR GAMBAR	xii
DAFTAR KODE	xiii
DAFTAR RUMUS	xiv
DAFTAR LAMPIRAN	xv
BAB 1 PENDAHULUAN	1
1.1 Latar Belakang Masalah	1
1.2 Rumusan Masalah	3
1.3 Batasan Permasalahan	3
1.4 Tujuan Penelitian	4
1.5 Manfaat Penelitian	4
1.6 Sistematika Penulisan	4
BAB 2 LANDASAN TEORI	6
2.1 Bahasa Isyarat Indonesia (BISINDO)	6
2.2 Inflated 3D (I3D) Model	6
2.2.1 Jaringan Saraf Konvolusi 3D (Conv3D)	7
2.2.2 3D Max Pooling	9
2.2.3 Inception V1	10
2.2.4 Arsitektur I3D	11
2.3 Algoritma aliran optik	12
2.3.1 <i>Total Variation regularization with L1 data fidelity</i> (TV-L1)	13
2.3.2 Algoritma aliran optik Farneback	14
2.3.3 Estimasi Gerakan	16
2.3.4 Langkah-Langkah Algoritma	16
2.4 Spatial Temporal Graph Convolutional Network (ST-GCN)	16
2.4.1 Konstruksi Graf Pose Tubuh	17
2.4.2 Konvolusi Graf Spatial Temporal	19
2.5 You Only Look Once (YOLO)	21
2.5.1 Arsitektur YOLOv11	22
2.5.2 Backbone	23
2.5.3 Neck	23
2.5.4 Head	24
2.5.5 Blok CBS	24
2.5.6 Lapisan Konvolusi Akhir dan Lapisan Deteksi	24
2.6 MediaPipe	25
2.6.1 Estimasi Pose Tubuh	25
2.6.2 Estimasi Pose Tangan	25
BAB 3 METODOLOGI PENELITIAN	27
3.1 Gambaran Umum Objek Penelitian	27

3.2	Metode Penelitian	27
3.3	Studi Literatur	28
3.3.1	Penelitian Terkait BISINDO pada level kata	28
3.3.2	Penelitian Terkait Bahasa Isyarat Bahasa Asing	29
3.3.3	Analisis Perbandingan dan Celah	30
3.4	Pengumpulan Data	30
3.4.1	Deskripsi Dataset	30
3.4.2	Pengambilan Sampel	30
3.5	Praproses Data	31
3.5.1	praproses instruktur	31
3.5.2	Praproses aliran optik	32
3.5.3	Praproses Pose	33
3.5.4	Data Splitting	33
3.5.5	Labelling	33
3.6	Perancangan Model	33
3.7	Pelatihan Model	34
3.7.1	Augmentasi Data	34
3.7.2	Parameter Tuning	35
3.8	Evaluasi Model	36
3.8.1	Kalkulasi Top-N	36
3.8.2	Penentuan Kombinasi Model	36
BAB 4	HASIL DAN DISKUSI	38
4.1	Spesifikasi Sistem Pengujian	38
4.2	Hasil Perancangan Sistem	38
4.2.1	Praproses	38
4.2.2	Perancangan Model	46
4.2.3	Pelatihan Model	48
4.2.4	Evaluasi Model	55
4.3	Hasil Praproses dan Augmentasi Data	57
4.4	Hasil Pelatihan Model	57
4.5	Prediksi Akurasi Model pada Dataset Test	59
BAB 5	SIMPULAN DAN SARAN	61
5.1	Simpulan	61
5.2	Saran	61
DAFTAR PUSTAKA		62

U N I V E R S I T A S
M U L T I M E D I A
N U S A N T A R A

DAFTAR TABEL

Tabel 2.1	Perbandingan Bahasa Isyarat SIBI dan BISINDO	6
Tabel 2.2	Versi-versi YOLO terdahulu	21
Tabel 3.1	Studi terdahulu terhadap mesin prediksi bahasa isyarat berlevel kata	28
Tabel 3.2	Perancangan kombinasi model	37
Tabel 4.1	Jumlah epoch dimana <i>early stopping</i> terjadi dan akurasi validasinya terhadap setiap model	58
Tabel 4.2	Hasil akurasi kombinasi model	59



DAFTAR GAMBAR

Gambar 1.1	Perbandingan akurasi antara penelitian yang dilakukan maruyama et al dengan penelitian sebelumnya	3
Gambar 2.1	Model framework Inflated 3D (atas) dan Inception module (bawah) pada model Inflated Inception-V1	7
Gambar 2.2	Cara kerja Conv3D dibandingkan dengan Conv2D	7
Gambar 2.3	Mekanisme kerja teknik max pooling 3D	9
Gambar 2.4	Visualisasi algoritma farneback	15
Gambar 2.5	Mekanisme kerja teknik konvolusi pada operasi ST-GCN .	17
Gambar 2.6	Koneksi antar sendi manusia yang berurutan terhadap waktu di ST-GCN	18
Gambar 2.7	Perbandingan akurasi antara YOLOv11 dengan versi-versi YOLO sebelumnya	22
Gambar 2.8	Arsitektur YOLOv11	22
Gambar 2.9	Titik-titik pose tubuh yang diprediksi oleh MediaPipe . . .	25
Gambar 2.10	Titik-titik pose tangan yang diprediksi oleh MediaPipe . .	26
Gambar 3.1	Alur jalannya penelitian	27
Gambar 3.2	Jumlah sampel pada masing-masing kata	31
Gambar 3.3	Metode praproses instruktur pada tahapan pertama	32
Gambar 3.4	praproses algoritma aliran optik farneback dengan visualisasi menggunakan HSV	32
Gambar 3.5	Metode perancangan model pada penelitian ini	34
Gambar 3.6	Metode augmentasi spatial video sebelum diberikan ke model	35
Gambar 4.1	Video frame rgb sebelum dan setelah melakukan augmentasi	57
Gambar 4.2	Grafik perbandingan hasil pelatihan tiap-tiap model terhadap dataset train dan test	58



DAFTAR KODE

Kode 4.1	Teknik pemotongan video	39
Kode 4.2	Kalkulasi aliran optik menggunakan algoritma farneback	40
Kode 4.3	Kalkulasi aliran optik menggunakan algoritma TV-L1	41
Kode 4.4	Teknik load model MediaPipe	42
Kode 4.5	Prediksi pose menggunakan MediaPipe	42
Kode 4.6	Fungsi kalkulasi perhitungan top-N	44
Kode 4.7	Metode split	44
Kode 4.8	Teknik labelling	45
Kode 4.9	Arsitektur kustomisasi ST-GCN	46
Kode 4.10	Arsitektur model I3D	46
Kode 4.11	fungsi pemotongan spatial dan mirroring	48
Kode 4.12	fungsi padding data temporal pada tahap augmentasi ketika ukuran frame kurang dari 64	49
Kode 4.13	fungsi pemotongan temporal untuk ukuran frame lebih dari 64	50
Kode 4.14	fungsi __data_generation untuk augmentasi data rgb dan aliran optik	51
Kode 4.15	fungsi pemotongan spatial dan mirroring	52
Kode 4.16	Fungsi kalkulasi perhitungan top-N	53
Kode 4.17	Fungsi create_callbacks	54
Kode 4.18	Fungsi kalkulasi perhitungan top-N	55
Kode 4.19	Kombinasi Model	56



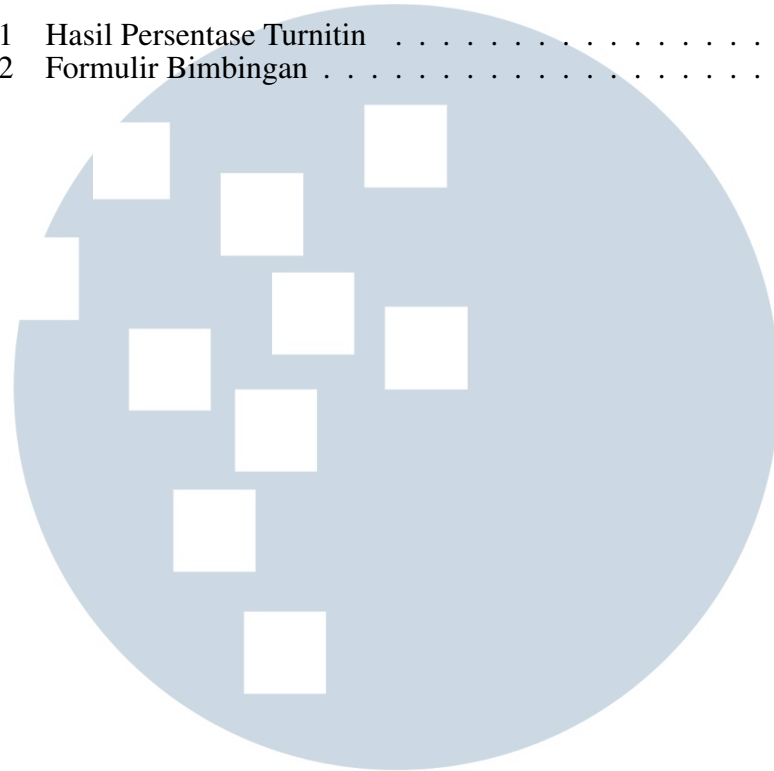
DAFTAR RUMUS

Rumus 2.1	Rumus perhitungan <i>feature map</i> pada konvolusi 2D	8
Rumus 2.2	Rumus perhitungan <i>feature map</i> pada konvolusi 3D	8
Rumus 2.3	Rumus Perhitungan <i>max pooling 3D</i>	10
Rumus 2.7	<i>Brightness Constancy Constraint</i>	12
Rumus 2.8	<i>Brightness Constancy Constraint</i> dengan deret taylor	12
Rumus 2.9	Rumus Akhir <i>Brightness Constancy Constraint</i>	12
Rumus 2.10	Konsep <i>Brightness Constancy Constraint</i> pada Algoritma TV-L1	13
Rumus 2.11	Persamaan Horn-Shnuck	13
Rumus 2.12	Ekspansi Taylor dari Persamaan Horn-Shnuck	13
Rumus 2.13	Minimalisasi Energi dengan Ketentuan L1	14
Rumus 2.14	Minimalisasi Energi dengan Relaksasi Konveks	14
Rumus 2.15	Pengerjaan nilai v ketika u diketahui	14
Rumus 2.16	Pengerjaan nilai u ketika v diketahui	14
Rumus 2.17	Persamaan Polinomial farneback	15
Rumus 2.18	Substitusi Temporal pada persamaan polinomial Farneback	15
Rumus 2.19	Persamaan <i>feature map</i> pada konvolusi CNN biasa	19
Rumus 2.20	Hubungan antar tetangga pada GCN	19
Rumus 2.21	Indeks berat pada <i>feature map</i> di GCN	20
Rumus 2.22	Perhitungan <i>feature map</i> pada GCN	20
Rumus 2.23	Simplifikasi Persamaan GCN	20
Rumus 2.24	Konsep Temporal pada ST-GCN	20
Rumus 2.25	Modifikasi Label Map pada ST-GCN	20



DAFTAR LAMPIRAN

Lampiran 1	Hasil Persentase Turnitin	67
Lampiran 2	Formulir Bimbingan	69



UMN
UNIVERSITAS
MULTIMEDIA
NUSANTARA