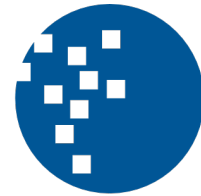


**DETEKSI VIDEO DEEFAKE MENGGUNAKAN
INCEPTIONV3 DAN LSTM BERDASARKAN METODE
PEMBUATAN VIDEO DEEFAKE**



UMN
UNIVERSITAS
MULTIMEDIA
NUSANTARA

SKRIPSI

JASON SEBASTIAN TJOANG
00000055443

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

**DETEKSI VIDEO DEEFAKE MENGGUNAKAN
INCEPTIONV3 DAN LSTM BERDASARKAN METODE
PEMBUATAN VIDEO DEEFAKE**



SKRIPSI

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

**JASON SEBASTIAN TJOANG
00000055443**

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

HALAMAN PERNYATAAN TIDAK PLAGIAT

Dengan ini saya,

Nama : Jason Sebastian Tjoang

Nomor Induk Mahasiswa : 00000055443

Program Studi : Informatika

Skripsi dengan judul:

Deteksi Video Deepfake Menggunakan InceptionV3 dan LSTM Berdasarkan Metode Pembuatan Video Deepfake

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, 04 Juli 2025



(Jason Sebastian Tjoang)

U M M N
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

HALAMAN PENGESAHAN

Skripsi dengan judul

DETEKSI VIDEO DEEPPAKE MENGGUNAKAN INCEPTIONV3 DAN LSTM BERDASARKAN METODE PEMBUATAN VIDEO DEEPPAKE

oleh

Nama : Jason Sebastian Tjoang
NIM : 00000055443
Program Studi : Informatika
Fakultas : Fakultas Teknik dan Informatika

Telah diujikan pada hari Kamis, 17 Juli 2025

Pukul 10.00 s/s 12.00 dan dinyatakan

LULUS

Dengan susunan penguji sebagai berikut

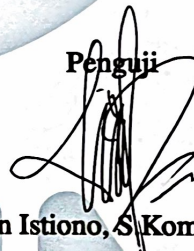
Ketua Sidang



(Eunike Endariahna Surbakti, S.Kom.,
M.T.I.)

NIDN: 0322099401

Penguji



(Wirawan Istiono, S.Kom., M.Kom.)

NIDN: 0313048304

Pembimbing



(Alexander Waworuntu, S.Kom., M.T.I.)

NIDN: 0309068503

Ketua Program Studi Informatika,



(Arya Wicaksana, S.Kom., M.Eng.Sc., OCA)

NIDN: 0315109103

**HALAMAN PERSETUJUAN PUBLIKASI KARYA ILMIAH UNTUK
KEPENTINGAN AKADEMIS**

Yang bertanda tangan di bawah ini:

Nama : Jason Sebastian Tjoang
NIM : 00000055443
Program Studi : Informatika
Jenjang : S1
Judul Karya Ilmiah : Deteksi Video Deepfake
Menggunakan InceptionV3 dan LSTM
Berdasarkan Metode Pembuatan
Video Deepfake

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, 04 Juli 2025

Yang menyatakan

Jason Sebastian Tjoang

**Jika tidak bisa membuktikan LoA jurnal/HKI, saya bersedia mengizinkan penuh karya ilmiah saya untuk dipublikasikan ke KC UMN dan menjadi hak institusi UMN.

HALAMAN PERSEMBAHAN / MOTTO

"Done is better than perfect."

Sheryl Sandberg



KATA PENGANTAR

Puji syukur atas terselesaikannya penulisan laporan skripsi ini yang berjudul: Deteksi Video Deepfake Menggunakan InceptionV3 dan LSTM Berdasarkan Metode Pembuatan Video Deepfake untuk memenuhi salah satu syarat untuk mencapai gelar Sarjana Komputer Jurusan Informatika Pada Fakultas Teknik dan Informatika Universitas Multimedia Nusantara. Saya menyadari bahwa, tanpa bantuan dan bimbingan dari berbagai pihak, dari masa perkuliahan sampai pada penyusunan skripsi ini, sangatlah sulit bagi saya untuk menyelesaikan skripsi ini. Oleh karena itu, saya 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. Bapak Alexander Waworuntu, S.Kom., M.T.I., sebagai Pembimbing yang telah memberikan bimbingan, arahan, dan motivasi atas terselesainya tugas akhir ini.
5. Keluarga saya yang telah memberikan bantuan dukungan material dan moral, sehingga saya dapat menyelesaikan tugas akhir ini.
6. Semua pihak yang tidak dapat disebutkan penulis satu per satu yang telah memberikan bantuan, doa, dan dukungan hingga pada akhirnya penyusunan laporan skripsi ini dapat terselesaikan dengan baik.

Semoga skripsi ini dapat bermanfaat, baik untuk sumber informasi, sumber inspirasi, ataupun sebagai referensi untuk para pembaca.

Tangerang, 04 Juli 2025



Jason Sebastian Tjoang

DETEKSI VIDEO DEEFAKE MENGGUNAKAN INCEPTIONV3 DAN LSTM BERDASARKAN METODE PEMBUATAN VIDEO DEEFAKE

Jason Sebastian Tjoang

ABSTRAK

Pesatnya perkembangan teknologi telah mendorong peningkatan penyebaran video *deepfake*, yaitu video manipulasi wajah berbasis kecerdasan buatan. Meskipun memiliki manfaat potensial, teknologi ini menimbulkan risiko seperti disinformasi, pencemaran nama baik, dan eksploitasi non-konsensual. Tantangan utama dalam deteksi *deepfake* adalah rendahnya kemampuan model untuk menggeneralisasi terhadap berbagai teknik manipulasi. Penelitian ini mengusulkan arsitektur deteksi berbasis InceptionV3 dan LSTM. InceptionV3 (dibekukan tanpa *top layer*) dengan GlobalAveragePooling2D mengekstraksi fitur spasial dari setiap *frame*, yang kemudian diproses secara temporal oleh LSTM 128 unit, lalu diklasifikasikan melalui lapisan Dense 64-neuron dan Dense 1-neuron beserta lapisan *dropout* dan *kernel regularizer* untuk mengurangi *overfitting*. Enam model dilatih: lima pada masing-masing metode manipulasi dari *dataset* FaceForensics++ (Deepfakes, FaceSwap, Face2Face, NeuralTextures, FaceShifter), dan satu model tambahan pada gabungan seluruh metode. Evaluasi dilakukan menggunakan lima metrik, dengan fokus pada AUC-ROC untuk skenario *cross-subset testing*. Hasil menunjukkan bahwa model spesifik Deepfakes mencapai AUC-ROC tertinggi (0,9007), namun menurun drastis pada metode lain. Pada *cross-subset testing*, model gabungan lebih seimbang namun tidak lebih unggul dari model yang dilatih dengan *deepfakes*. Semua model mengalami *overfitting* karena kurangnya variansi data, arsitektur terlalu kompleks, serta jumlah *frame* yang tidak cukup untuk menangkap pola temporal yang kompleks. Secara umum, performa InceptionV3-LSTM berada di bawah arsitektur lain seperti Xception.

Kata kunci: Deteksi *Deepfake*, FaceForensics++, Generalisasi Model, InceptionV3, LSTM

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

DEEPPAKE VIDEO DETECTION USING INCEPTIONV3 AND LSTM BASED ON DEEPPAKE GENERATION METHODS

Jason Sebastian Tjoang

ABSTRACT

The rapid development of technology has driven the proliferation of deepfake videos, which are artificial intelligence-based facial manipulation videos. Despite its potential benefits, this technology poses risks such as disinformation, defamation, and non-consensual exploitation. A major challenge in deepfake detection is the model's poor generalizability across various manipulation techniques. This study proposes a detection architecture based on InceptionV3 and LSTM. InceptionV3 (frozen without the top layer) with GlobalAveragePooling2D extracts spatial features from each frame, which are then temporally processed by a 128-unit LSTM, and then classified using a 64-neuron Dense layer and a 1-neuron Dense layer along with dropout and kernel regularizer layers to reduce overfitting. Six models were trained: five on each manipulation method from the FaceForensics++ dataset (Deepfakes, FaceSwap, Face2Face, NeuralTextures, FaceShifter), and one additional model combining all methods. The evaluation was conducted using five metrics, focusing on AUC-ROC for the cross-subset testing scenario. The results showed that the Deepfakes-specific model achieved the highest AUC-ROC (0.9007), but decreased significantly compared to the other methods. In cross-subset testing, the combined model was more balanced but not superior to the model trained on deepfakes. All models experienced overfitting due to a lack of data variance, overly complex architectures, and an insufficient number of frames to capture complex temporal patterns. Overall, the performance of InceptionV3-LSTM was lower than that of other architectures such as Xception.

Keywords: *Deepfake Detection, FaceForensics++, InceptionV3, LSTM, Model Generalization,*

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 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	xii
DAFTAR GAMBAR	xiii
DAFTAR KODE	xv
DAFTAR RUMUS	xvi
DAFTAR LAMPIRAN	xvii
BAB 1 PENDAHULUAN	1
1.1 Latar Belakang Masalah	1
1.2 Rumusan Masalah	4
1.3 Batasan Permasalahan	4
1.4 Tujuan Penelitian	4
1.5 Manfaat Penelitian	5
1.6 Sistematika Penulisan	5
BAB 2 LANDASAN TEORI	6
2.1 Penelitian Terdahulu	6
2.2 <i>Deepfake</i>	9
2.2.1 <i>Generative Adversarial Network (GAN)</i>	9
2.2.2 <i>Autoencoder</i>	11
2.2.3 <i>FaceForensics++</i>	13
2.3 <i>Convolutional Neural Network (CNN)</i>	15
2.3.1 <i>Convolutional Layer</i>	16
2.3.2 <i>Pooling Layer</i>	17
2.3.3 <i>Fully Connected Layer</i>	18
2.3.4 <i>Fungsi Aktivasi (Non-Linier)</i>	18
2.3.5 <i>Loss Function</i>	19
2.3.6 <i>Regularisasi</i>	20
2.4 <i>Adaptive Moment Estimation (ADAM)</i>	20
2.5 YuNet	21
2.5.1 <i>Backbone</i>	22
2.5.2 <i>Neck</i>	22
2.5.3 <i>Head</i>	23
2.6 GoogleNet (InceptionV1)	24
2.7 InceptionV2	26
2.8 InceptionV3	27
2.9 <i>Recurrent Neural Network (RNN)</i>	28
2.9.1 <i>Input Layer</i>	28
2.9.2 <i>Recurrent Layer</i>	29
2.9.3 <i>Output Layer</i>	29
2.10 <i>Long Short-Term Memory (LSTM)</i>	32
2.10.1 <i>Forget Gate</i>	33

2.10.2	<i>Input Gate</i>	33
2.10.3	<i>State Computation</i>	34
2.10.4	<i>Output Gate</i>	34
2.11	Metrik Evaluasi Model	35
2.11.1	<i>Confusion Matrix</i>	35
2.11.2	<i>AUC (Area Under the Curve) & ROC (Receiver Operating Characteristic)</i>	36
BAB 3	METODOLOGI PENELITIAN	38
3.1	Gambaran Umum Penelitian	38
3.2	Studi Literatur	38
3.3	Pengumpulan Data	39
3.3.1	Sumber Dataset	39
3.3.2	Karakteristik Dataset	39
3.3.3	Struktur <i>Dataset</i> FaceForensics++	39
3.4	<i>Pre-processing Data</i>	41
3.4.1	Ekstraksi <i>Frame</i> dan Wajah dari Video	42
3.4.2	<i>Resize Frame</i> Wajah	45
3.4.3	Pembagian Subset Dataset untuk Pelatihan dan Evaluasi	46
3.5	Pelatihan Model	48
3.5.1	<i>Data Splitting</i>	48
3.5.2	Perancangan Arsitektur Model	49
3.5.3	Konfigurasi <i>Hyperparameter</i>	50
3.5.4	Strategi <i>Loading</i> Input Data	51
3.5.5	Proses Pelatihan Model	52
3.6	Evaluasi Model	54
3.6.1	Evaluasi <i>Intra-Subset</i>	54
3.6.2	Evaluasi <i>Cross-Subset</i>	55
BAB 4	HASIL DAN DISKUSI	56
4.1	Spesifikasi Sistem	56
4.2	<i>Dataset</i>	57
4.3	<i>Pre-processing Data</i>	57
4.3.1	<i>Import Library</i> yang Diperlukan	57
4.3.2	Pembentukan <i>Dataset</i>	59
4.3.3	Bentuk <i>Folder</i> Setelah <i>Preprocessing</i>	71
4.4	<i>Model Training</i>	71
4.4.1	<i>Video Frame Generator</i>	71
4.4.2	<i>Hyperparameter</i>	74
4.4.3	Penyusunan Model	75
4.4.4	Pengambilan Data <i>PreProcessed</i>	77
4.4.5	Inisialisasi Pelatihan	78
4.5	Evaluasi Model	82
4.5.1	Deepfakes	82
4.5.2	Face2Face	86
4.5.3	FaceSwap	90
4.5.4	NeuralTextures	94
4.5.5	FaceShifter	98
4.5.6	<i>Combined Methods</i>	102
4.5.7	Ringkasan Performa Model pada <i>Test Set</i>	106
4.5.8	<i>Cross-Dataset Testing</i> AUC-ROC	106
4.6	Diskusi	107
4.6.1	Performa <i>Training</i> Model	107

4.6.2	Performa Model pada <i>Test Set</i>	108
4.6.3	Performa <i>Cross-Dataset</i>	109
4.6.4	Perbandingan dengan Penelitian Terdahulu	110
BAB 5	SIMPULAN DAN SARAN	112
5.1	Simpulan	112
5.2	Saran	113
DAFTAR PUSTAKA		115



DAFTAR TABEL

Tabel 2.1	Penelitian Terdahulu Terkait Deteksi Video <i>Deepfake</i> . . .	6
Tabel 2.2	Deskripsi <i>Dataset</i> FaceForensics++	13
Tabel 2.3	<i>Confusion Matrix</i> Berukuran 2×2	35
Tabel 4.1	Ringkasan Metriks Performa Model pada <i>Test Set</i> dari Metode Manipulasi yang Sama dengan Data Pelatihan . . .	106
Tabel 4.2	Matriks Evaluasi Silang Model (AUC-ROC)	106



DAFTAR GAMBAR

Gambar 2.1	Contoh Wajah Asli (Kiri) dan Wajah Hasil <i>Deepfake</i> (Kanan)	9
Gambar 2.2	Struktur <i>Generative Adversarial Network</i> (GAN)	10
Gambar 2.3	Menghasilkan <i>Deepfake</i> Menggunakan <i>Autoencoder</i>	12
Gambar 2.4	Contoh Teknik Pembuatan <i>Deepfake</i> pada <i>Dataset</i> FaceForensics++	13
Gambar 2.5	Contoh Arsitektur CNN Untuk Klasifikasi Gambar	15
Gambar 2.6	Operasi Konvolusi	16
Gambar 2.7	Arsitektur YuNet	21
Gambar 2.8	InceptionV1 atau GoogleNet	24
Gambar 2.9	Struktur <i>Inception Module</i>	25
Gambar 2.10	<i>Old Inception Module</i>	26
Gambar 2.11	<i>New Inception Module</i> dengan Pemecahan Konvolusi Besar	26
Gambar 2.12	<i>Inception Module</i> dengan <i>Factorized Convolution</i>	27
Gambar 2.13	<i>Recurrent Neural Network Layers</i>	28
Gambar 2.14	Pemrosesan Data pada Model RNN Berdasarkan Urutan Waktu	30
Gambar 2.15	Arsitektur LSTM	32
Gambar 3.1	Alur Penelitian	38
Gambar 3.2	<i>Subset Dataset</i> FaceForensics++	39
Gambar 3.3	Tahapan <i>Preprocessing</i>	41
Gambar 3.4	Ekstraksi <i>Frame</i> dan Wajah dari Video	42
Gambar 3.5	<i>Flow</i> Ekstraksi 15 <i>Frame</i> dari Video Secara Merata	43
Gambar 3.6	<i>Flow</i> Deteksi Wajah pada Setiap <i>Frame</i>	44
Gambar 3.7	<i>Flow Resize Wajah</i>	46
Gambar 3.8	<i>Flow Data Splitting</i>	48
Gambar 3.9	Arsitektur Model	49
Gambar 3.10	<i>Flow</i> Proses Pelatihan Model	53
Gambar 4.1	Struktur <i>Folder Dataset</i> Awal	57
Gambar 4.2	Ilustrasi Ekstraksi <i>Frame</i> dari Video	63
Gambar 4.3	Ilustrasi Ekstraksi Wajah dari <i>Frame</i>	65
Gambar 4.4	Ilustrasi <i>Resize Image</i>	67
Gambar 4.5	Struktur <i>Dataset</i> Setelah <i>Preprocessing</i>	71
Gambar 4.6	10 Lapisan Akhir Inceptionv3	76
Gambar 4.7	Struktur Akhir Model yang Digunakan	77
Gambar 4.8	Akurasi <i>Subset Deepfakes</i>	82
Gambar 4.9	<i>Loss Subset Deepfakes</i>	83
Gambar 4.10	<i>Confusion Matrix Subset Deepfakes</i>	84
Gambar 4.11	<i>AUC-ROC Subset Deepfakes</i>	85
Gambar 4.12	Akurasi <i>Subset Face2Face</i>	86
Gambar 4.13	<i>Loss Subset Face2Face</i>	87
Gambar 4.14	<i>Confusion Matrix Subset Face2Face</i>	88
Gambar 4.15	<i>AUC-ROC Subset Face2Face</i>	89
Gambar 4.16	Akurasi <i>Subset FaceSwap</i>	90
Gambar 4.17	<i>Loss Subset FaceSwap</i>	91
Gambar 4.18	<i>Confusion Matrix Subset FaceSwap</i>	92
Gambar 4.19	<i>AUC-ROC Subset FaceSwap</i>	93
Gambar 4.20	Akurasi <i>Subset NeuralTextures</i>	94
Gambar 4.21	<i>Loss Subset NeuralTextures</i>	95

Gambar 4.22	<i>Confusion Matrix Subset NeuralTextures</i>	96
Gambar 4.23	<i>AUC-ROC Subset NeuralTextures</i>	97
Gambar 4.24	<i>Akurasi Subset FaceShifter</i>	98
Gambar 4.25	<i>Loss Subset FaceShifter</i>	99
Gambar 4.26	<i>Confusion Matrix Subset FaceShifter</i>	100
Gambar 4.27	<i>AUC-ROC Subset FaceShifter</i>	101
Gambar 4.28	<i>Akurasi Gabungan Subset</i>	102
Gambar 4.29	<i>Loss Gabungan Subset</i>	103
Gambar 4.30	<i>Confusion Matrix Gabungan Subset</i>	104
Gambar 4.31	<i>AUC-ROC Gabungan Subset</i>	105



UMN
UNIVERSITAS
MULTIMEDIA
NUSANTARA

DAFTAR KODE

Kode 4.1	<i>Import Library</i>	57
Kode 4.2	Mengekstraksi <i>Frame</i> Wajah dari Video	59
Kode 4.3	Mengekstraksi 15 <i>Frame</i> dari Video 10 Detik	63
Kode 4.4	Mendeteksi dan Mengekstraksi Wajah dalam Suatu <i>Frame</i>	64
Kode 4.5	Memanggil Fungsi <i>process_paired_multiple_fakes</i>	66
Kode 4.6	<i>Resize Image</i>	67
Kode 4.7	<i>Resize</i> dengan <i>Padding Black Bars</i> pada Gambar	69
Kode 4.8	Melakukan <i>Resize</i> dengan <i>Padding Black Bars</i> pada Gambar	69
Kode 4.9	<i>Video Frame Generator</i>	71
Kode 4.10	<i>Hyperparameter</i>	74
Kode 4.11	Inisialisasi Model	75
Kode 4.12	Pengambilan Data <i>PreProcessed</i>	77
Kode 4.13	<i>Splitting Data</i>	78
Kode 4.14	Contoh Key	80
Kode 4.15	Inisialisasi Generator dan Callback	80
Kode 4.16	Proses Pelatihan Model	81



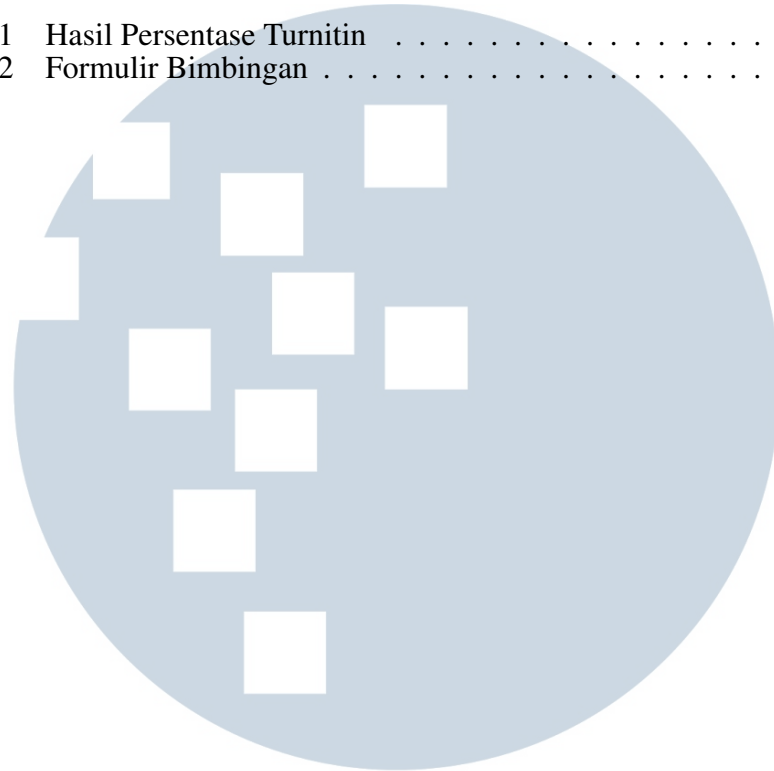
DAFTAR RUMUS

Rumus 2.1	<i>Minimax GAN</i>	10
Rumus 2.2	<i>Autoencoder</i>	11
Rumus 2.3	<i>Sigmoid</i>	19
Rumus 2.4	<i>TanH</i>	19
Rumus 2.5	<i>ReLU</i>	19
Rumus 2.6	<i>Softmax</i>	20
Rumus 2.7	<i>ADAM</i>	21
Rumus 2.8	<i>FPN</i>	23
Rumus 2.9	<i>Loss Function Head YuNet</i>	23
Rumus 2.10	<i>Forget Gate LSTM</i>	33
Rumus 2.11	<i>Input Gate LSTM</i>	33
Rumus 2.12	<i>State Computation LSTM</i>	34
Rumus 2.13	<i>State Computation LSTM</i>	34
Rumus 2.14	<i>Output Gate LSTM</i>	34
Rumus 2.15	<i>Output Gate LSTM</i>	34
Rumus 2.16	<i>Akurasi</i>	35
Rumus 2.17	<i>Presisi</i>	36
Rumus 2.18	<i>Recall</i>	36
Rumus 2.19	<i>F1-Score</i>	36
Rumus 2.20	<i>True Positive Rate</i>	37
Rumus 2.21	<i>False Positive Rate</i>	37



DAFTAR LAMPIRAN

Lampiran 1	Hasil Persentase Turnitin	118
Lampiran 2	Formulir Bimbingan	130



UMN
UNIVERSITAS
MULTIMEDIA
NUSANTARA