



SKRIPSI

PENINGKATAN *ROBUSTNESS* CNN PADA DETEKSI LAGU AI MENGGUNAKAN *DATA AUGMENTATION*

FARIS NUR TSANI
NPM 22083010076

DOSEN PEMBIMBING

Wahyu Syaifullah Jauharis Saputra, S. Kom., M. Kom.
Alfan Rizaldy Pratama, S. Tr. T., M. Tr. Kom.

**KEMENTERIAN PENDIDIKAN TINGGI, SAINS, DAN TEKNOLOGI
UNIVERSITAS PEMBANGUNAN NASIONAL VETERAN JAWA TIMUR
FAKULTAS ILMU KOMPUTER
PROGRAM STUDI SAINS DATA
SURABAYA
2026**

LEMBAR PENGESAHAN

PENINGKATAN *ROBUSTNESS* CNN PADA DETEKSI LAGU AI MENGUNAKAN *DATA AUGMENTATION*

Oleh:

FARIS NUR TSANI
NPM. 22083010076

Telah dipertahankan di hadapan dan diterima oleh Tim Penguji Sidang Skripsi Program Studi Sains Data Fakultas Ilmu Komputer Universitas Pembangunan Nasional Veteran Jawa Timur pada tanggal 27 Februari 2026:

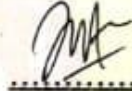
Menyetujui,

Wahyu Syaifullah Jauharis Saputra, S. Kom., M. Kom.
NIP. 19860825 202121 1 003



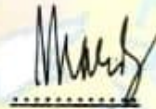
(Pembimbing I)

Alfan Rizaldy Pratama, S.Tr.T., M.Tr.Kom.
NIP. 19990606 202406 1 001



(Pembimbing II)

Dr. I Gede Susrama Mas Diyasa, ST., MT.
NIP. 19700619 202121 1 009



(Ketua Penguji)

Muhammad Nasrudin, M.Stat.
NIP. 19960909 202406 1 002



(Penguji I)

Mengetahui,
Dekan Fakultas Ilmu Komputer



Prof. Dr. Ir. Novirina Hendrasarie, MT.
NIP. 19681126 199403 2 001

LEMBAR PERSETUJUAN

**PENINGKATAN *ROBUSTNESS* CNN PADA DETEKSI LAGU AI
MENGUNAKAN *DATA AUGMENTATION***

Oleh:
FARIS NUR TSANI
NPM. 22083010076

Telah disetujui untuk mengikuti Ujian Skripsi

Menyetujui,

**Koordinator Program Studi Sains Data
Fakultas Ilmu Komputer**



Dr. Eng. Ir. Dwi Arman Prasetya, S.T., M.T., IPU., ASEAN Eng.
NIP. 19801205 200501 1 002

SURAT PERNYATAAN BEBAS PLAGIASI

Saya yang bertanda tangan di bawah ini:

Nama Mahasiswa : Faris Nur Tsani
NPM : 22083010076
Program : Sarjana (S1)
Program Studi : Sains Data
Fakultas : Fakultas Ilmu Komputer

Menyatakan bahwa dalam dokumen ilmiah Skripsi ini tidak terdapat bagian dari karya ilmiah lain yang telah diajukan untuk memperoleh gelar akademik di suatu lembaga Pendidikan Tinggi, dan juga tidak terdapat karya atau pendapat yang pernah ditulis atau diterbitkan oleh orang/lembaga lain, kecuali yang secara tertulis disitasi dalam dokumen ini dan disebutkan secara lengkap dalam daftar pustaka.

Dan saya menyatakan bahwa dokumen ilmiah ini bebas dari unsur-unsur plagiasi. Apabila di kemudian hari ditemukan indikasi plagiat pada Skripsi ini, saya bersedia menerima sanksi sesuai dengan peraturan perundang-undangan yang berlaku.

Demikian surat pernyataan ini saya buat dengan sesungguhnya tanpa ada paksaan dari siapapun juga dan untuk dipergunakan sebagaimana mestinya.



Surabaya, 4 Maret 2026
Yang Membuat Pernyataan,



FARIS NUR TSANI
NPM. 22083010076

ABSTRAK

Nama Mahasiswa / NPM : Faris Nur Tsani / 22083010076
Judul Skripsi : Peningkatan *Robustness* CNN Pada Deteksi Lagu AI Menggunakan *Data Augmentation*
Dosen Pembimbing : 1. Wahyu Syaifullah Jauharis Saputra, S. Kom., M. Kom.
2. Alfian Rizaldy Pratama, S. Tr. T., M. Tr. Kom

Perkembangan teknologi *Artificial Intelligence* (AI) telah memungkinkan produksi lagu secara otomatis dengan kualitas yang menyerupai karya manusia, sehingga menimbulkan tantangan dalam membedakan lagu AI-generated dan human-generated. Penelitian ini bertujuan untuk meningkatkan *robustness* model *Convolutional Neural Network* (CNN) dalam mendeteksi lagu AI melalui penerapan *data augmentation* serta menganalisis pengaruh kompleksitas arsitektur *Residual Network* (ResNet-18, ResNet-34, dan ResNet-50) terhadap performa dan generalisasi model. Representasi audio menggunakan spektrogram berbasis *Short-Time Fourier Transform* (STFT) sebagai input model. Evaluasi dilakukan pada data utama serta dua skenario *robustness*, yaitu manipulasi sinyal (*noise*, *pitch shifting*, dan *time stretching*) dan pengujian lintas genre sebagai bentuk *out-of-distribution testing*. Hasil penelitian menunjukkan bahwa peningkatan kedalaman arsitektur tidak selalu berbanding lurus dengan peningkatan *robustness*. Penerapan data augmentation secara konsisten meningkatkan ketahanan model terhadap manipulasi sinyal pada seluruh arsitektur, dengan peningkatan paling signifikan terjadi pada ResNet-18 (64,79% menjadi 73,71%). Pada pengujian lintas genre, performa model relatif stabil dengan atau tanpa augmentasi. Temuan ini menunjukkan adanya *trade-off* antara kompleksitas model dan generalisasi, serta menegaskan bahwa *data augmentation* berperan penting dalam memperluas distribusi pembelajaran sehingga meningkatkan *robustness* model terhadap variasi kondisi audio. Penelitian ini juga mengintegrasikan metode *Explainable Artificial Intelligence* (XAI) berbasis Grad-CAM untuk meningkatkan interpretabilitas hasil klasifikasi.

Kata kunci : *Convolutional Neural Network, Residual Network, Robustness Model, Data Augmentation, Lagu AI*

ABSTRACT

Student Name / NPM : Faris Nur Tsani / 22083010076
Undergraduate Thesis Title : *Improving The Robustness Of CNN In AI Song Detection Using Data Augmentation*
Advisors : 1. Wahyu Syaifullah Jauharis Saputra, S. Kom., M. Kom.
2. Alfian Rizaldy Pratama, S. Tr. T., M. Tr. Kom

The rapid advancement of Artificial Intelligence (AI) has enabled the automatic generation of music with quality comparable to human compositions, creating new challenges in distinguishing AI-generated songs from human-produced music. This study aims to improve the robustness of Convolutional Neural Network (CNN) models in detecting AI-generated songs through the application of data augmentation and to analyze the impact of architectural complexity within the Residual Network family (ResNet-18, ResNet-34, and ResNet-50) on model performance and generalization. Audio signals are represented using Short-Time Fourier Transform (STFT)-based spectrograms as model inputs. Evaluation is conducted on the primary dataset and under two robustness scenarios: signal manipulation (noise addition, pitch shifting, and time stretching) and cross-genre testing as an out-of-distribution assessment. The results indicate that increasing architectural depth does not necessarily lead to improved robustness. The implementation of data augmentation consistently enhances model resilience against signal manipulation across all architectures, with the most significant improvement observed in ResNet-18 (from 64.79% to 73.71%). In cross-genre testing, model performance remains relatively stable with or without augmentation. These findings highlight a trade-off between model complexity and generalization and demonstrate that data augmentation plays a crucial role in expanding the learning distribution and improving robustness against variations in audio conditions. Additionally, Explainable Artificial Intelligence (XAI) using Grad-CAM is employed to enhance the interpretability of classification outcomes.

Keywords : Convolutional Neural Network, Residual Network, Robustness Model, Data Augmentation, AI-Generated Song

KATA PENGANTAR

Puji syukur kehadiran Allah SWT atas segala rahmat, hidayah dan karunia-Nya kepada penulis sehingga skripsi dengan judul “**Peningkatan *Robustness* CNN Pada Deteksi Lagu AI Menggunakan *Data Augmentation***” dapat terselesaikan dengan baik.

Penulis menyampaikan rasa terima kasih yang sebesar – besarnya kepada berbagai pihak yang telah memberikan dukungan dan bantuan selama proses pengerjaan skripsi baik dalam bentuk moril, spiritual, maupun materil. Setiap arahan, nasihat, dan informasi yang diberikan telah menjadi landasan yang kuat dalam penyusunan skripsi ini. Oleh karena itu, penulis ingin menyampaikan rasa terima kasih kepada:

1. Prof. Dr. Ir. Akhmad Fauzi, M.MT., IPU, selaku Rektor Universitas Pembangunan Nasional “Veteran” Jawa Timur.
2. Ibu Prof. Dr. Ir. Novirina Hendrasari, MT., selaku Dekan Fakultas Ilmu Komputer Universitas Pembangunan Nasional “Veteran” Jawa Timur.
3. Bapak Dr. Eng. Ir. Dwi Arman Prasetya, ST., MT., IPU., ASEAN Eng., selaku Ketua Program Studi Sains Data Fakultas Ilmu Komputer Universitas Pembangunan Nasional “Veteran” Jawa Timur.
4. Wahyu Syaifullah Jauharis Saputra, S. Kom., M. Kom. selaku Dosen Pembimbing 1 dan Alfian Rizaldy Pratama, S. Tr. T., M. Tr. Kom., selaku Dosen Pembimbing 2 yang telah meluangkan waktunya untuk memberikan bimbingan, nasehat serta motivasi kepada penulis.
5. Keluarga LA, Ayah, Mama, Mas Haikal, dan Mbak Dinda atas segala doa, kasih sayang, dukungan moral maupun material, serta ketulusan dalam kebersamaan setiap langkah perjuangan penulis hingga dapat menyelesaikan pendidikan ini.
6. Para sobat penulis dan rekan CP (Cangkruk Produktif) Khafid, Wahyu, Budi, Daffin, Ayak, Mei, Chelin yang telah banyak memberikan dukungan, motivasi, serta menjadi tempat berbagi cerita dan semangat selama proses penyusunan skripsi ini.

7. Teman-teman Bismillah Sukses Lulus (Kadafi, Nashif, Hauzan, Ajeng, Fira, Sela, dan Melinda) yang telah kebersamai dalam proses penyusunan skripsi, pengurusan administrasi, berbagi cerita, serta saling memberikan dukungan hingga tahap akhir.
8. Terakhir, kepada diri sendiri yang sudah berjuang dan bertahan sampai pada titik ini.

Penulis menyadari bahwa di dalam penyusunan skripsi ini banyak terdapat kekurangan. Untuk itu kritik dan saran yang membangun dari semua pihak sangat diharapkan demi kesempurnaan penulisan skripsi ini. Akhirnya, dengan segala keterbatasan yang penulis miliki semoga laporan ini dapat bermanfaat bagi semua pihak umumnya dan penulis pada khususnya.

Surabaya, Februari 2026

Penulis

DAFTAR ISI

HALAMAN SAMPUL.....	i
LEMBAR PENGESAHAN	ii
LEMBAR PERSETUJUAN	iii
SURAT PERNYATAAN BEBAS PLAGIASI.....	iv
ABSTRAK	v
<i>ABSTRACT</i>	vi
KATA PENGANTAR.....	vii
DAFTAR ISI.....	ix
DAFTAR GAMBAR.....	xi
DAFTAR TABEL	xiii
DAFTAR LAMPIRAN	xiv
DAFTAR NOTASI.....	xv
BAB I PENDAHULUAN.....	1
1.1. Latar Belakang	1
1.2. Rumusan Masalah	6
1.3. Batasan Masalah.....	7
1.4. Tujuan Penelitian	7
1.5. Manfaat Penelitian	8
BAB II TINJAUAN PUSTAKA.....	9
2.1. Penelitian Terdahulu	9
2.2. Kerangka Teori.....	15
2.2.1. Kecerdasan Buatan (<i>Artificial Intelligence</i>) dalam Musik.....	16
2.2.2. <i>Deep Learning</i>	17
2.2.3. <i>Convolutional Neural Network</i>	21
2.2.5. <i>Residual Network</i>	25
2.2.6. Pengolahan Data Suara dan Representasi Audio Spektogram.....	28
2.2.7. <i>Explainable Artificial Intelligence (XAI)</i>	30
2.2.8. <i>Gradient-weighted Class Activation Mapping (Grad-CAM)</i>	31
2.2.9. <i>Robustness</i> dalam Model <i>Deep Learning</i>	33

2.2.10. <i>Data Augmentation</i> dalam Audio	35
2.2.11. Evaluasi Model	37
BAB III METODOLOGI PENELITIAN	41
3.1. Variabel Penelitian dan Sumber Data	41
3.2. Langkah Analisis	42
3.2.1. <i>Data Collecting</i>	43
3.2.2. <i>Data Preprocessing</i>	45
3.2.3. <i>Data Augmentation</i>	49
3.2.4. <i>Feature Extraction</i>	55
3.2.5. <i>Modelling</i>	57
3.2.6. <i>Robustness Model Testing</i>	60
3.2.7. Implementasi <i>Explainable Artificial Intelligence</i>	63
3.2.8. Pembuatan Sistem Deteksi (<i>Streamlit</i>)	67
3.3. Desain Sistem	68
BAB IV HASIL DAN PEMBAHASAN	71
4.1. Pelatihan dan Evaluasi Model	71
4.1.1. Hasil <i>Training</i> dan Validasi Model	71
4.1.2. Evaluasi Kinerja Model	77
4.2. <i>Robustness Model Testing</i>	84
4.2.1. Hasil Robust Test (<i>Signal Manipulating</i>)	85
4.2.2. Hasil Robust Test (<i>Cross-Genre</i>)	86
4.3. Perbandingan Performa dan <i>Robustness Model</i>	87
4.4. Tampilan Antarmuka (GUI)	89
BAB V PENUTUP	93
5.1. Kesimpulan	93
5.2. Saran Pengembangan	94
DAFTAR PUSTAKA	97
LAMPIRAN	101

DAFTAR GAMBAR

Gambar 2.1. Ilustrasi proses <i>forward pass</i> dalam Deep Learning [24].....	18
Gambar 2.2. Ilustrasi proses <i>back propagation</i> dalam Deep Learning [24]	20
Gambar 2.3. Contoh alur arsitektur CNN untuk klasifikasi gambar	22
Gambar 2.4. Lapisan Konvolusi	23
Gambar 2.5. Teknik <i>pooling layer</i> dengan <i>max pooling</i> [27]	24
Gambar 2.6. <i>Residual Learning: a Building Block</i> [28]	25
Gambar 2.7. Fungsi residual yang lebih dalam. Kiri: <i>a building block</i> (56×56 <i>feature maps</i>). Kanan: “ <i>bottleneck</i> ” <i>building block</i> untuk ResNet-50/101/152 [28]	27
Gambar 2.8. Representasi audio berupa gambar spektogram	29
Gambar 2.9. Transformasi audio ke spektogram.....	30
Gambar 2.10. Cara Kerja Grad-CAM [32].....	32
Gambar 2.11. Ilustrasi pengaruh <i>robustness</i> model pada nilai akurasi	34
Gambar 3.1. Alur penelitian	42
Gambar 3.2. Pengunduhan audio dari platform dengan extention CRX	43
Gambar 3.3. Lagu hasil generate menggunakan Suno AI	44
Gambar 3.4. Prompt untuk membuat lagu menggunakan Udio AI	45
Gambar 3.5. Hasil <i>Grad-CAM</i> salah satu segmen lagu <i>Human</i>	66
Gambar 3.6. Hasil <i>Grad-CAM</i> salah satu segmen lagu <i>AI</i>	66
Gambar 3.7. Desain sistem	68
Gambar 4.1. Grafik <i>training vs validation (accuracy & loss)</i> ResNet-18 dataset tanpa augmentasi	72
Gambar 4.2. Grafik <i>training vs validation (accuracy & loss)</i> ResNet-18 dataset dengan augmentasi	73
Gambar 4.3. Grafik <i>training vs validation (accuracy & loss)</i> ResNet-34 dataset tanpa augmentasi	74
Gambar 4.4. Grafik <i>training vs validation (accuracy & loss)</i> ResNet-34 dataset dengan augmentasi	75

Gambar 4.5. Grafik <i>training vs validation (accuracy & loss)</i> ResNet-50 dataset tanpa augmentasi	76
Gambar 4.6. Grafik <i>training vs validation (accuracy & loss)</i> ResNet-50 dataset dengan augmentasi	77
Gambar 4.7. <i>Confusion Matrix</i> ResNet-18 pada data tanpa augmentasi	78
Gambar 4.8. <i>Confusion Matrix</i> ResNet-18 pada data tanpa augmentasi	79
Gambar 4.9. <i>Confusion Matrix</i> ResNet-34 pada data tanpa augmentasi	80
Gambar 4.10. <i>Confusion Matrix</i> ResNet-34 pada data dengan augmentasi	81
Gambar 4.11. <i>Confusion Matrix</i> ResNet-50 pada data tanpa augmentasi	82
Gambar 4.12. <i>Confusion Matrix</i> ResNet-50 pada data dengan augmentasi	83
Gambar 4.13 Tampilan Awal Sistem.....	90
Gambar 4.14 Tampilan sistem saat file telah diunggah	90
Gambar 4.15 Tampilan sistem berupa hasil deteksi	91
Gambar 4.16 Tampilan hasil sistem deteksi (Penerapan Grad-CAM).....	91

DAFTAR TABEL

Tabel 2.1. Penelitian terdahulu	9
Tabel 2.2. Tabel Arsitektur ResNet-18, ResNet-34, ResNet-50 [29].....	27
Tabel 2.3. Tabel <i>Confusion matrix</i>	37
Tabel 3.1. Ringkasan perolehan lagu AI	45
Tabel 3.2. Ringkasan perubahan data dalam tahap preprocessing	49
Tabel 3.3. Kombinasi Teknik Augmentasi pada Robust Test.....	61
Tabel 4.1. Matriks evaluasi ResNet-18 pada data tanpa augmentasi	78
Tabel 4.2. Matriks evaluasi ResNet-18 pada data dengan augmentasi	79
Tabel 4.3. Matriks evaluasi ResNet-34 pada data tanpa augmentasi	80
Tabel 4.4. Matriks evaluasi ResNet-34 pada data dengan augmentasi	81
Tabel 4.5. Matriks evaluasi ResNet-50 pada data tanpa augmentasi	83
Tabel 4.6. Matriks evaluasi ResNet-50 pada data tanpa augmentasi	84
Tabel 4.7. Metrik evaluasi <i>robust test (signal manipulating)</i>	85
Tabel 4.8. Metrik evaluasi <i>robust test (cross-genre)</i>	86
Tabel 4.9. Perbandingan nilai akurasi kedua skema model.....	88

DAFTAR LAMPIRAN

Lampiran 1. Metadata dan Data Penelitian	101
Lampiran 2. Kode Script	102
Lampiran 3. GUI Code	103
Lampiran 4. <i>Letter of Acceptance</i> Jurnal.....	104

DAFTAR NOTASI

$A(i, j)$	:	Nilai pada <i>activation map</i>
I	:	Citra <i>input</i>
K	:	<i>Kernel</i> berukuran $M \times N$
b	:	Paramater bias
y	:	Nilai neuron
σ	:	<i>Input</i> hasil <i>flattening</i>
x_i	:	Parameter bias
$\{W_i\}$	:	Kumpulan bobot dari lapisan di dalam blok residual
$f(x, \{W_i\})$	:	Transformasi nonlinier
$\frac{\partial y_c}{\partial A_{ij}^k}$	:	Gradien skor kelas c terhadap elemen pada posisi (i, j)
α_k^c	:	Bobot kepentingan feature map ke- k terhadap kelas c
Z	:	Jumlah total elemen (<i>pixel</i>) dalam feature map A^k
$L_c^{Grad-CAM}$	:	Peta aktivasi Grad-CAM untuk kelas <i>heatmap</i> akhir
$\sum_k \alpha_k^c A^k$	:	Kombinasi linear dari semua <i>feature map</i> dengan bobotnya
$x(t)$	:	Sinyal waktu
$g(t - \tau)$	:	Fungsi <i>window</i>
τ	:	Waktu (time position)
ω	:	Frekuensi sudut
$e^{-j\omega t}$	:	Basis <i>Fourier</i> (Gelombang Kompleks)