

**SEPARASI SUARA GAMELAN MENGGUNAKAN DEEP
NEURAL NETWORK**

SKRIPSI

Diajukan untuk memenuhi salah satu syarat mencapai derajat Sarjana
Program Studi S1 Informatika



disusun oleh

BIMA WIRATAMA

19.11.3022

Kepada

**FAKULTAS ILMU KOMPUTER
UNIVERSITAS AMIKOM YOGYAKARTA
YOGYAKARTA**

2023

**SEPARASI SUARA GAMELAN MENGGUNAKAN DEEP
NEURAL NETWORK**

SKRIPSI

untuk memenuhi salah satu syarat mencapai derajat Sarjana
Program Studi S1 Informatika



disusun oleh

BIMA WIRATAMA

19.11.3022

Kepada

**FAKULTAS ILMU KOMPUTER
UNIVERSITAS AMIKOM YOGYAKARTA
YOGYAKARTA**

2023

HALAMAN PERSETUJUAN

SKRIPSI

**SEPARASI SUARA GAMELAN MENGGUNAKAN DEEP NEURAL
NETWORK**

yang disusun dan diajukan oleh

Bima Wiratama

19.11.3022

telah disetujui oleh Dosen Pembimbing Skripsi
pada tanggal 23 Juni 2023

Dosen Pembimbing,



Agit Amrullah, S.Kom, M.Kom

NIK. 190302356

HALAMAN PENGESAHAN
SKRIPSI
SEPARASI SUARA GAMELAN MENGGUNAKAN DEEP NEURAL
NETWORK

yang disusun dan diajukan oleh

Bima Wiratama

19.11.3022

Telah dipertahankan di depan Dewan Penguji
pada tanggal 20 Juli 2023

Susunan Dewan Penguji

Nama Penguji

Tanda Tangan

Arif Akbarul Huda, S.Si, M.Eng
NIK. 190302287

Arifiyanto Hadinegoro, S.Kom, MT
NIK. 190302289

Agit Amrullah, S.Kom, M.Kom
NIK. 190302356



Skripsi ini telah diterima sebagai salah satu persyaratan
untuk memperoleh gelar Sarjana Komputer
Tanggal 20 Juli 2023

DEKAN FAKULTAS ILMU KOMPUTER



Hanif Al Fatta, S.Kom., M.Kom.
NIK. 190302096

HALAMAN PERNYATAAN KEASLIAN SKRIPSI

Yang bertandatangan di bawah ini,

Nama mahasiswa : Bima Wiratama
NIM : 19.11.3022

Menyatakan bahwa Skripsi dengan judul berikut:

SEPARASI SUARA GAMELAN MENGGUNAKAN DEEP NEURAL NETWORK

Dosen Pembimbing : Agit Amrullah, S.Kom, M.Kom

1. Karya tulis ini adalah benar-benar ASLI dan BELUM PERNAH diajukan untuk mendapatkan gelar akademik, baik di Universitas AMIKOM Yogyakarta maupun di Perguruan Tinggi lainnya.
2. Karya tulis ini merupakan gagasan, rumusan dan penelitian SAYA sendiri, tanpa bantuan pihak lain kecuali arahan dari Dosen Pembimbing.
3. Dalam karya tulis ini tidak terdapat karya atau pendapat orang lain, kecuali secara tertulis dengan jelas dicantumkan sebagai acuan dalam naskah dengan disebutkan nama pengarang dan disebutkan dalam Daftar Pustaka pada karya tulis ini.
4. Perangkat lunak yang digunakan dalam penelitian ini sepenuhnya menjadi tanggung jawab SAYA, bukan tanggung jawab Universitas AMIKOM Yogyakarta.
5. Pernyataan ini SAYA buat dengan sesungguhnya, apabila di kemudian hari terdapat penyimpangan dan ketidakbenaran dalam pernyataan ini, maka SAYA bersedia menerima SANKSI AKADEMIK dengan pencabutan gelar yang sudah diperoleh, serta sanksi lainnya sesuai dengan norma yang berlaku di Perguruan Tinggi.

Yogyakarta, 20 Juli 2023

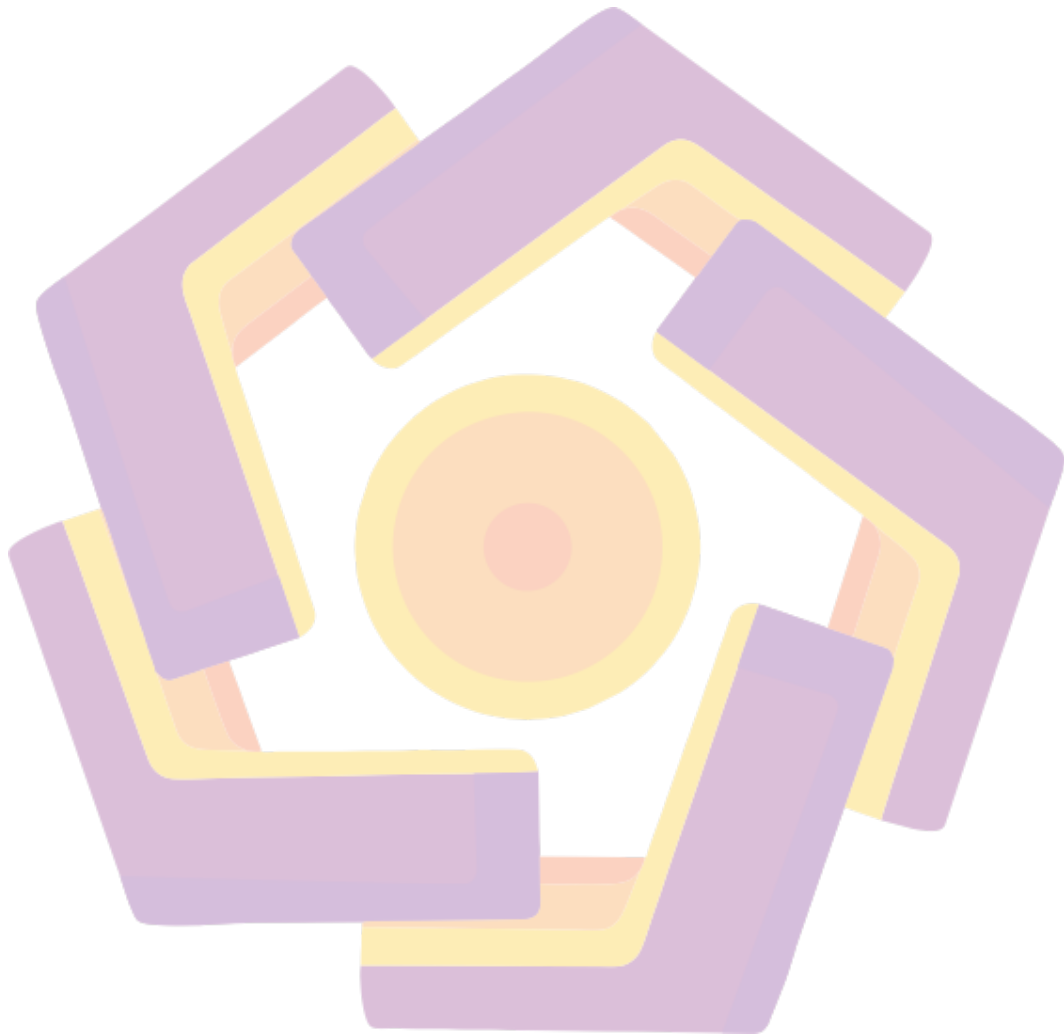
Yang Menyatakan,



Bima Wiratama

HALAMAN PERSEMBAHAN

Skripsi ini saya persembahkan sebagai hasil akhir perkuliahan kepada Bapak. Drs. Suratmono yang telah membiayai proses akademik saya di Universitas AMIKOM Yogyakarta.



KATA PENGANTAR

Segala puji bagi Tuhan yang Maha Esa yang telah memberikan rahmat, kesempatan, sehingga penulis dapat menulis skripsi dengan judul “**SEPARASI SUARA GAMELAN MENGGUNAKAN DEEP NEURAL NETWORK**”. Penulisan skripsi ini menjadi salah satu bagian untuk memenuhi persyaratan kelulusan program Sarjana di Universitas AMIKOM Yogyakarta pada jurusan Informatika.

Pada penyusunan skripsi ini tidak lepas dari dukungan berbagai pihak, oleh karena itu pada kesempatan ini penulis mengucapkan terima kasih kepada:

1. Bapak Wawan Purnawan dari sanggar Omah Seni Tanjung Anom di Semanu, Gunung Kidul yang telah memberikan data rekaman gamelan yang dijadikan sebagai dataset dalam penelitian ini.
2. Bapak Agit Amrullah, M.Kom selaku dosen pembimbing yang telah meluangkan waktunya untuk memberikan arahan dan bimbingan untuk menyusun skripsi.
3. Kedua orang tua penulis yang telah memberi dukungan secara terus menerus.
4. Bapak ibu dosen dan pegawai Universitas AMIKOM Yogyakarta.
5. Seluruh pihak yang telah membantu proses penyusunan skripsi yang tidak dapat penulis sebutkan satu persatu.

Yogyakarta, 31 Mei 2023

Penulis

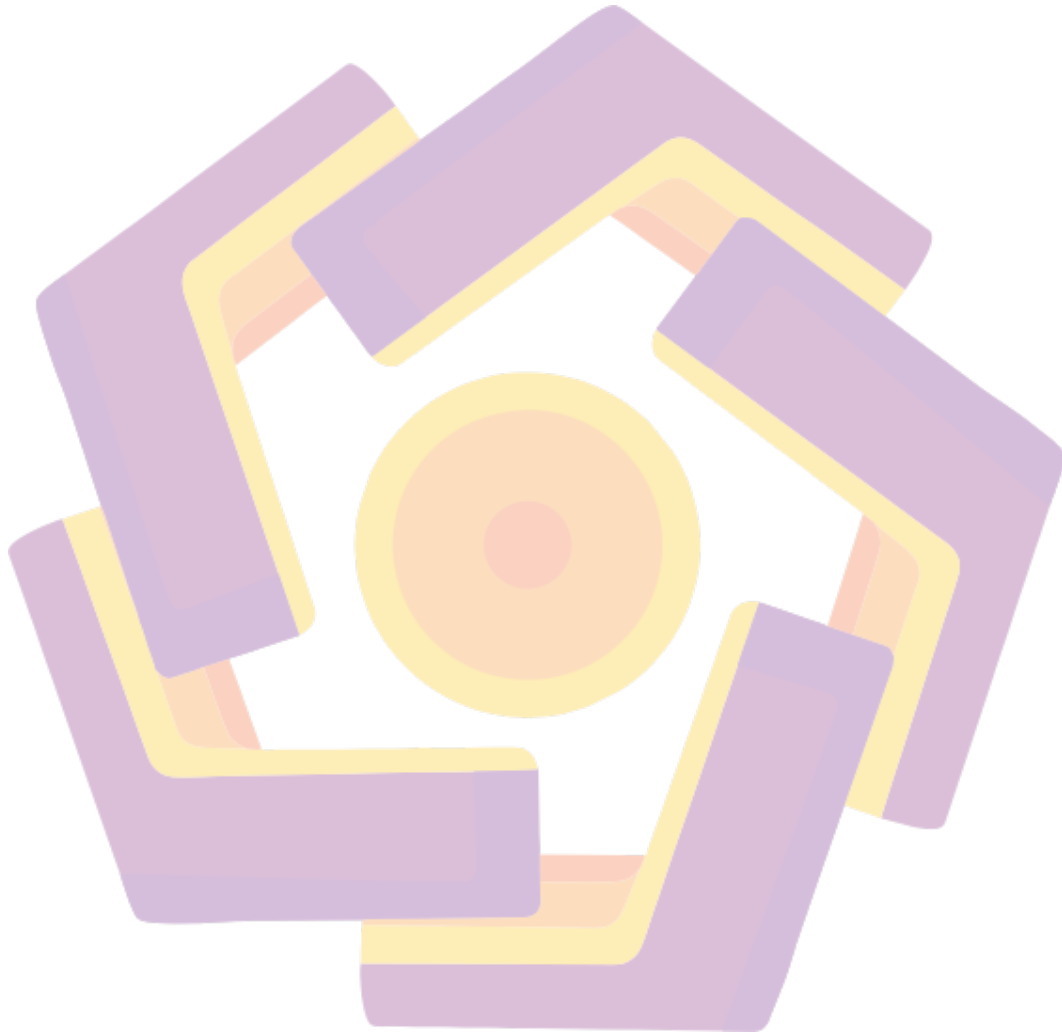
DAFTAR ISI

HALAMAN JUDUL	i
HALAMAN PERSETUJUAN	ii
HALAMAN PENGESAHAN	ii
HALAMAN PERNYATAAN KEASLIAN SKRIPSI	iv
HALAMAN PERSEMBAHAN	v
KATA PENGANTAR	vi
DAFTAR ISI	vii
DAFTAR TABEL	ix
DAFTAR GAMBAR	x
INTISARI	xi
ABSTRACT	xii
BAB I PENDAHULUAN	1
1.1 Latar Belakang	1
1.2 Rumusan Masalah	3
1.3 Batasan Masalah	3
1.4 Tujuan Penelitian	4
1.5 Manfaat Penelitian	4
1.6 Sistematika Penulisan	4
BAB II TINJAUAN PUSTAKA	6
2.1 Studi Literatur	6
2.2 Landasan Teori	14
2.2.1 Music Source Separation (MSS)	14
2.2.2 Gamelan Jawa	15
2.2.3 Deep Neural Network	18
2.2.4 Arsitektur U-Net	21
2.3 Bottleneck	24
2.3.1 LSTM	24
2.3.2 DPRNN	27
2.3.3 Hybrid Demucs	28
2.4 Representasi Input	30
2.4.1 Audio Digital	30

2.4.2	STFT	31
2.4.3	Wiener Filtering	33
2.4.4	Complex As Channels	35
2.4.5	Inverse STFT	35
2.5	Bentuk & Alur Penelitian	36
BAB III METODOLOGI PENELITIAN		37
3.1	Jenis, Sifat dan Pendekatan Penelitian	37
3.1.1	Jenis Penelitian	37
3.1.2	Sifat Penelitian	37
3.1.3	Pendekatan Penelitian	37
3.2	Alur Penelitian	38
3.2.1	Studi Literatur	38
3.2.2	Perumusan Masalah	39
3.2.3	Pengumpulan Data	39
3.2.4	Analisis Data	40
3.2.5	Membuat Testbed	42
3.2.6	Training Model	44
3.2.7	Evaluasi Model	45
3.2.8	Analisis Komparatif Terhadap Hasil Evaluasi	46
3.2.9	Membuat Kesimpulan	47
BAB IV HASIL DAN PEMBAHASAN		48
4.1	Hasil Training	48
4.2	Hasil SDR, SIR dan SAR	51
BAB V PENUTUP		57
5.1	Kesimpulan	57
5.2	Saran	57
DAFTAR PUSTAKA		59

DAFTAR TABEL

Tabel 2.1. Keaslian Penelitian	9
Tabel 3.1. Spesifikasi dataset yang digunakan pada penelitian.	39
Tabel 4.1. Hasil evaluasi training pada berbagai model yang diuji	49
Tabel 4.2. Hasil SDR pada berbagai konfigurasi model pada data test	52
Tabel 4.3. Hasil SAR pada berbagai konfigurasi model pada data test	53
Tabel 4.4. Hasil SIR pada berbagai konfigurasi model pada data test	54



DAFTAR GAMBAR

Gambar 1.1. Spektogram berbagai instrumen gamelan dalam rekaman.	2
Gambar 2.1. Instrumen dalam Gamelan Jawa. (JT Titon, 2009)	16
Gambar 2.2. Gambaran dari Artificial Neuron.	19
Gambar 2.3. Pipeline arsitektur U-Net. (Rachel Bittner, 2017)	22
Gambar 2.4. Input dan kernel pada konvolusi. (Vincent Dumoulin, 2018)	23
Gambar 2.5. Unit memory sel LSTM. (Aston Zhang 2021)	26
Gambar 2.6. Dua layer RNN pada topologi DPRNN. (Yi Luo 2020)	27
Gambar 2.7. Arsitektur network pada Hybrid Demucs. (Défossez 2021)	29
Gambar 2.8. Gelombang suara yang direpresentasikan secara digital. (Meinard Muller, 2021)	31
Gambar 2.9. Spektogram. Dimensi pertama menunjukkan indeks dari frame, dimensi kedua menunjukkan indeks dari frekuensi bin. Matrix berisi nilai energy pada frame ke-x dan frekuensi y. (Ethan Manilow, 2020)	33
Gambar 2.10. Contoh soft mask.	34
Gambar 3.1. Alur Penelitian	38
Gambar 3.2. Contoh dari salah satu sampel <i>magnitude spectrogram</i> .	43
Gambar 3.3. Alur Arsitektur U-Net.	44
Gambar 4.1. Nilai loss pada data training	50
Gambar 4.2. Nilai loss pada data validasi	51
Gambar 4.3. Spektogram slenthem pada LSTM_WF dan LSTM_CAC	55
Gambar 4.4. Noise pada berbagai model yang diuji.	56

INTISARI

Penggunaan instrumen gamelan atau idiom permainan dalam gamelan banyak digunakan dalam berbagai komposisi film, video game dan musik kontemporer. Mayoritas instrumen gamelan yang terbuat dari logam menjadi tantangan untuk memisahkan sinyal antar instrumen dalam melakukan evaluasi komposisi musik gamelan. Music Source Separation merupakan suatu proses untuk memisahkan sinyal antar sumber suara menggunakan model neural network yang bersifat supervised. Arsitektur U-Net digunakan untuk menguji berbagai metode dan fitur. Penelitian ini bertujuan untuk melakukan studi komparatif mencari metode dan fitur neural network yang paling baik untuk Music Source Separation (MSS) pada musik gamelan jawa metode neural network yang diuji dipilih dari metode yang sering digunakan untuk data sequential, seperti Long Short-Term Memory (LSTM), Dual Path Recurrent Neural Network (DPRNN) dan Hybrid Demucs. Ada dua representasi input spektogram yang diuji dalam network yaitu spektogram magnitudo dengan Wiener Filtering (WF) dan Complex-as-Channels (CAC). Pada pengujian yang dilakukan menggunakan metrik evaluasi berupa Source-To-Distortion Ratio (SDR), Source to Artifact Ratio (SAR), dan Source to Interference Ratio (SIR). Secara rata-rata SDR, Hybrid Demucs memiliki nilai yang paling baik yaitu 8,2753 dB. Secara rata-rata SAR, Hybrid Demucs juga mempunyai nilai yang paling tinggi yaitu 3,2944 dB. Untuk rata-rata SIR, DPRNN CAC mempunyai nilai yang paling tinggi yaitu 5,1595 dB. Hybrid Demucs mempunyai performa sekitar 27,5% lebih baik daripada DPRNN CAC dan 28,2% lebih baik dari LSTM CAC.

Kata Kunci: Music Source Separation, U-Net, Gamelan Jawa

ABSTRACT

The use of gamelan instruments or its playing idioms is widely used in various film compositions, video games and contemporary music. The vast majority of gamelan instruments are made of metal, making it a hard challenge to separate signals between instruments in evaluating gamelan music compositions. Music Source Separation is a process for separating signals between sound sources using a supervised neural network model. The U-Net architecture is used to test various methods and features. The aim of this study is to conduct a comparative study to find the most suitable neural network method and features for Music Source Separation (MSS) in gamelan music. The neural network methods tested were selected from methods that are often used for sequential data, such as Long Short-Term Memory (LSTM), Dual Path Recurrent Neural Network (DPRNN) and Hybrid Demucs. There are two representations of input spectrograms tested in the network: magnitude spectrogram with Wiener Filtering (WF) and Complex-as-Channels (CAC). There are three evaluation metrics that used in testing, such as Source-To-Distortion Ratio (SDR), Source to Artifact Ratio (SAR), and Source to Interference Ratio (SIR). On average SDR, Hybrid Demucs has the highest value of 8.2753 dB. Hybrid Demucs also has the highest value of 3.2944 dB for average SAR. For the average SIR, DPRNN CAC has the highest value of 5.1595 dB. Hybrid Demucs performs about 27.5% better than DPRNN CAC and 28.2% better than LSTM CAC.

Keywords: Music Source Separation, U-Net, Javanese Gamelan