

**PENGEMBANGAN MODEL HYBRID CNN-LSTM UNTUK
PENGENALAN AKOR OTOMATIS PADA SEGMENTASI
AUDIO MUSIK DIGITAL BERBASIS
LOG-MEL SPECTROGRAM**

SKRIPSI

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



disusun oleh

TIGOR NEILSON SINAGA

22.11.4725

Kepada

**FAKULTAS ILMU KOMPUTER
UNIVERSITAS AMIKOM YOGYAKARTA
YOGYAKARTA**

2026

**PENGEMBANGAN MODEL HYBRID CNN-LSTM UNTUK
PENGENALAN AKOR OTOMATIS PADA SEGMENTASI
AUDIO MUSIK DIGITAL BERBASIS
LOG-MEL SPECTROGRAM**

SKRIPSI

untuk memenuhi salah satu syarat mencapai derajat Sarjana
Program Studi Informatika



disusun oleh

TIGOR NEILSON SINAGA

22.11.4725

Kepada

**FAKULTAS ILMU KOMPUTER
UNIVERSITAS AMIKOM YOGYAKARTA
YOGYAKARTA**

2026

HALAMAN PERSETUJUAN

SKRIPSI

**PENGEMBANGAN MODEL HYBRID CNN-LSTM UNTUK
PENGENALAN AKOR OTOMATIS PADA SEGMENTASI
AUDIO MUSIK DIGITAL BERBASIS
LOG-MEL SPECTROGRAM**

yang disusun dan diajukan oleh

Tigor Neilson Sinaga

22.11.4725

telah disetujui oleh Dosen Pembimbing Skripsi
pada tanggal 12 Mei 2026

Dosen Pembimbing,


Ike Verawati, M.Kom.
NIK. 190302237

HALAMAN PENGESAHAN
SKRIPSI
PENGEMBANGAN MODEL HYBRID CNN-LSTM UNTUK
PENGENALAN AKOR OTOMATIS PADA SEGMENT
AUDIO MUSIK DIGITAL BERBASIS
LOG-MEL SPECTROGRAM

yang disusun dan diajukan oleh

Tigor Neilson Sinaga

22.11.4725

Telah dipertahankan di depan Dewan Penguji
pada tanggal 12 Mei 2026

Susunan Dewan Penguji

Nama Penguji

Tanda Tangan

Anna Balta, S.Kom., M.Kom.
NIK. 190302290

Fauzia Anis Sekarningrum, S.T., M.T.
NIK. 190302683

Ike Verawati, S.Kom., M.Kom.
NIK. 190302237



Skripsi ini telah diterima sebagai salah satu persyaratan
untuk memperoleh gelar Sarjana Komputer
Tanggal 12 Mei 2026

DEKAN FAKULTAS ILMU KOMPUTER



Prof. Dr. Kusriani, M.Kom.
NIK. 190302106

HALAMAN PERNYATAAN KEASLIAN SKRIPSI

Yang bertandatangan di bawah ini,

Nama mahasiswa : TIGOR NEILSON SINAGA
NIM : 22.11.4725

Menyatakan bahwa Skripsi dengan judul berikut:

PENGEMBANGAN MODEL HYBRID CNN-LSTM UNTUK PENGENALAN AKOR OTOMATIS PADA SEGMENT AUDIO MUSIK DIGITAL BERBASIS LOG-MEL SPECTROGRAM

Dosen Pembimbing : Ike Verawati, 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, 12 Mei 2026

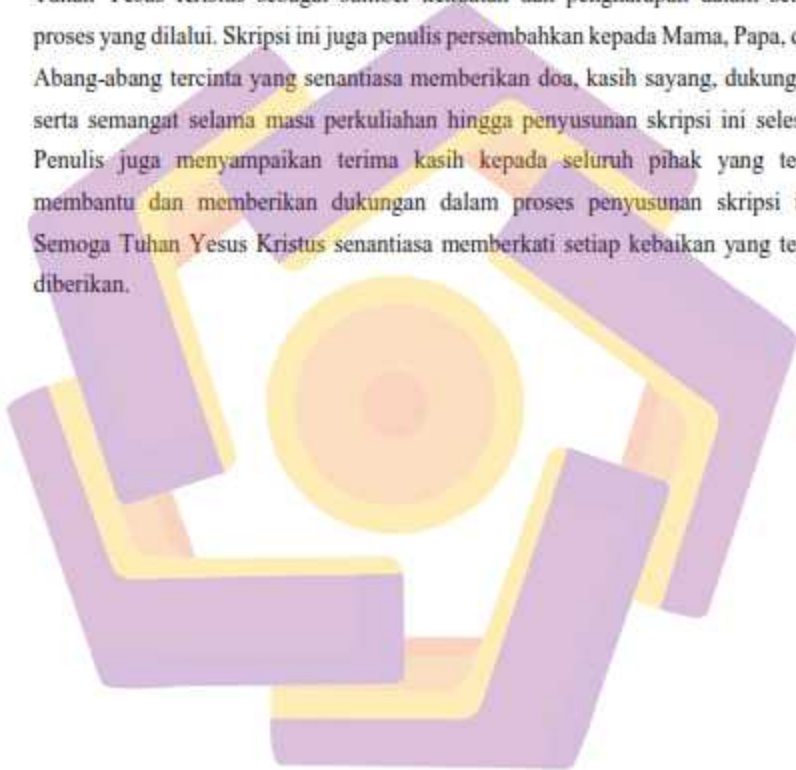
Yang Menyatakan,



Tigor Neilson Sinaga

HALAMAN PERSEMBAHAN

Segala puji dan syukur penulis panjatkan ke hadirat Tuhan Yesus Kristus atas kasih, berkat, dan penyertaan-Nya sehingga skripsi ini dapat diselesaikan dengan baik. Dengan penuh rasa syukur, skripsi ini penulis persembahkan kepada Tuhan Yesus Kristus sebagai sumber kekuatan dan pengharapan dalam setiap proses yang dilalui. Skripsi ini juga penulis persembahkan kepada Mama, Papa, dan Abang-abang tercinta yang senantiasa memberikan doa, kasih sayang, dukungan, serta semangat selama masa perkuliahan hingga penyusunan skripsi ini selesai. Penulis juga menyampaikan terima kasih kepada seluruh pihak yang telah membantu dan memberikan dukungan dalam proses penyusunan skripsi ini. Semoga Tuhan Yesus Kristus senantiasa memberkati setiap kebaikan yang telah diberikan.



KATA PENGANTAR

Puji dan syukur penulis panjatkan ke hadirat Tuhan Yesus Kristus atas kasih, berkat, dan penyertaan-Nya, sehingga penulis dapat menyelesaikan skripsi yang berjudul **“Pengembangan Model Hybrid CNN-LSTM untuk Pengenalan Akor Otomatis pada Segmen Audio Musik Digital Berbasis Log-Mel Spectrogram”** dengan baik. Skripsi ini disusun sebagai salah satu syarat untuk memperoleh gelar Sarjana pada Program Studi Informatika, Fakultas Ilmu Komputer, Universitas AMIKOM Yogyakarta.

Penyusunan skripsi ini tidak terlepas dari dukungan, bimbingan, dan bantuan dari berbagai pihak. Oleh karena itu, penulis menyampaikan terima kasih yang sebesar-besarnya kepada:

1. Prof. Dr. M. Suyanto, M.M., selaku Rektor Universitas AMIKOM Yogyakarta.
2. Prof. Dr. Kusriani, M.Kom., selaku Dekan Fakultas Ilmu Komputer Universitas AMIKOM Yogyakarta.
3. Ibu Eli Pujastuti, M.Kom., selaku Ketua Program Studi Informatika Universitas AMIKOM Yogyakarta.
4. Ibu Ike Verawati, M.Kom., selaku dosen pembimbing yang telah memberikan arahan, bimbingan, masukan, serta waktu selama proses penyusunan skripsi ini.
5. Ibu Anna Baita, M.Kom., dan Ibu Fauzia Anis Sekar Ningrum, S.T., M.T., selaku tim dosen penguji yang telah memberikan saran dan masukan untuk penyempurnaan skripsi ini.
6. Seluruh dosen Program Studi Informatika Universitas AMIKOM Yogyakarta yang telah memberikan ilmu, pengalaman, dan wawasan selama masa perkuliahan.
7. Orang tua dan keluarga penulis yang senantiasa memberikan doa, dukungan, semangat, dan motivasi selama proses perkuliahan hingga penyusunan skripsi ini selesai.

8. Efionsie Indrisari May Kihi dan Darwin Mezakh Loduwik Rumpumbo yang senantiasa memberi semangat, dorongan, serta menemani disaat proses penyusunan skripsi ini selesai.
9. Kepada tim Slepet FC Jogja-Sleman, Bantul atas hiburan dan dukungan selama proses penulisan naskah penelitian ini.
10. Semua pihak yang telah membantu dalam proses penyusunan skripsi ini, baik secara langsung maupun tidak langsung.

Penulis menyadari bahwa skripsi ini masih memiliki keterbatasan dan belum sepenuhnya sempurna. Oleh karena itu, penulis terbuka terhadap kritik dan saran yang membangun agar karya ini dapat menjadi lebih baik. Semoga skripsi ini dapat memberikan manfaat bagi pembaca serta menjadi referensi bagi penelitian selanjutnya, khususnya dalam bidang pengenalan akor otomatis dan analisis musik digital berbasis kecerdasan buatan.

Yogyakarta, 12 Mei 2026

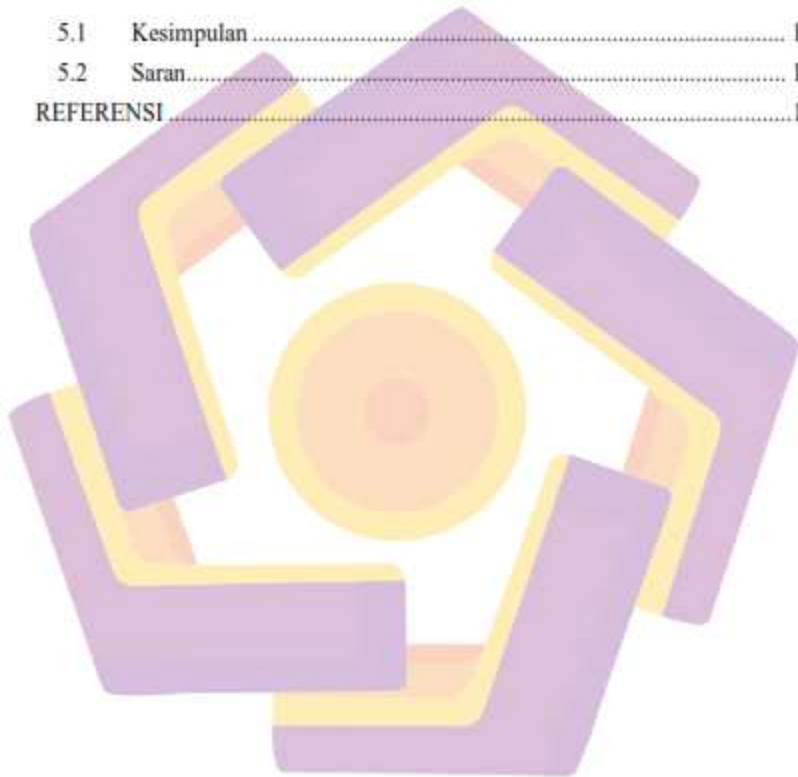
Penulis

DAFTAR ISI

HALAMAN JUDUL.....	i
HALAMAN PERSETUJUAN.....	ii
HALAMAN PENGESAHAN.....	iii
HALAMAN PERNYATAAN KEASLIAN SKRIPSI.....	iv
HALAMAN PERSEMBAHAN.....	v
KATA PENGANTAR.....	vi
DAFTAR ISI.....	viii
DAFTAR TABEL.....	xi
DAFTAR GAMBAR.....	xii
DAFTAR LAMBANG DAN SINGKATAN.....	xiii
DAFTAR ISTILAH.....	xiv
INTISARI.....	xvi
ABSTRACT.....	xvii
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.....	5
1.6 Sistematika Penulisan.....	6
BAB II TINJAUAN PUSTAKA.....	8
2.1 Studi Literatur.....	8
2.2 Dasar Teori.....	16
2.2.1 Artificial Intelligence.....	16
2.2.2 Machine Learning.....	17
2.2.3 Music Information Retrieval (MIR).....	19

2.2.4	Akor (Chord).....	20
2.2.5	Progresi Akor	21
2.2.6	Deep Learning.....	22
2.2.7	Audio Processing	24
2.2.8	Short-Time Fourier Transform (STFT).....	25
2.2.9	Mel-Scale	27
2.2.10	Mel-Spectrogram	28
2.2.11	Log-Mel Spectrogram.....	30
2.2.12	Convolutional Neural Network (CNN).....	31
2.2.13	Recurrent Neural Network (RNN).....	35
2.2.14	Long Short-Term Memory (LSTM)	36
2.2.15	Arsitektur Hybrid CNN-LSTM.....	39
2.2.16	Optimasi dan Parameter Pelatihan Model.....	42
2.2.17	Metrik Evaluasi Model.....	43
BAB III METODE PENELITIAN		46
3.1	Objek Penelitian.....	46
3.2	Alur Penelitian	47
3.3	Alat dan Bahan.....	50
3.3.1	Data Penelitian	51
3.3.2	Alat/Instrumen	52
BAB IV HASIL DAN PEMBAHASAN		56
4.1	Akuisisi Dataset IDMT-SMT-CHORDS	56
4.2	Ekstraksi File Audio dan File Anotasi	59
4.3	Parsing Label Akor dari File Anotasi	61
4.4	Pemetaan Label Menjadi 24 Kelas Akor.....	64
4.5	Segmentasi Audio Berdasarkan Anotasi Akor.....	65
4.6	Resampling dan Konversi Audio ke Mono.....	67
4.7	Ekstraksi Fitur Mel-Spectrogram dan Log-Mel Spectrogram	68
4.8	Normalisasi, Padding, Cropping, dan Pembentukan Tensor Input	71
4.9	Encoding Label dan Pembagian Data	74
4.10	Perancangan Model Hybrid CNN-LSTM.....	76
4.11	Pelatihan Model dan Pemilihan Model Terbaik	79

4.12	Evaluasi Model pada Data Test	83
4.13	Perbandingan CNN-LSTM Berbasis Log-Mel Spectrogram dan CNN-LSTM Berbasis Mel-Spectrogram	93
4.14	Pengujian dengan File Audio WAV	96
4.15	Prediksi Akor per Segmen dan Pembentukan Progresi Akor	100
BAB V PENUTUP.....		104
5.1	Kesimpulan	104
5.2	Saran.....	105
REFERENSI		107



DAFTAR TABEL

Tabel 2.1 Keaslian penelitian	12
Tabel 3.1 Seleksi kelas akor pada dataset	52
Tabel 3.2 Tipe perangkat keras	55
Tabel 3.3 Perangkat tambahan untuk pembuatan audio uji	55
Tabel 4.1 Ringkasan dataset dan metadata penelitian.....	59
Tabel 4.2 Hasil ekstraksi file audio dan file anotasi	61
Tabel 4.3 Metadata hasil parsing label akor dari file anotasi.....	63
Tabel 4.4 Daftar 24 kelas akor yang digunakan.....	64
Tabel 4.5 Ringkasan hasil segmentasi audio berdasarkan anotasi akor.....	66
Tabel 4.6 Parameter resampling dan konversi audio ke mono	67
Tabel 4.7 Parameter ekstraksi fitur pada dua eksperimen	71
Tabel 4.8 Bentuk akhir tensor input CNN-LSTM	73
Tabel 4.9 Ringkasan encoding label dan pembagian data	75
Tabel 4.10 Bentuk data pembagian hasil dataset	76
Tabel 4.11 Ringkasan hasil pelatihan model CNN-LSTM berbasis log-mel spectrogram.....	82
Tabel 4.12 Ringkasan hasil pelatihan model CNN-LSTM berbasis mel-spectrogram.....	83
Tabel 4.13 Hasil evaluasi model CNN-LSTM berbasis log-mel spectrogram	84
Tabel 4.14 Hasil evaluasi model CNN-LSTM berbasis mel-spectrogram	88
Tabel 4.15 Perbandingan hasil CNN-LSTM berbasis log-mel spectrogram dan mel-spectrogram.....	94

DAFTAR GAMBAR

Gambar 2.1 Mel-spectrogram	29
Gambar 2.2 Arsitektur convolutional neural network.	32
Gambar 2.3 Anatomi LSTM.	37
Gambar 2.4 Arsitektur hybrid CNN-LSTM.	41
Gambar 3.1 Alur penelitian.	48
Gambar 4.1 Struktur folder proyek dan dataset hasil ekstraksi.	57
Gambar 4.2 Hasil identifikasi file audio dan file anotasi.	60
Gambar 4.3 Contoh isi file anotasi mentah.	62
Gambar 4.4 Waveform segmen audio hasil segmentasi.	65
Gambar 4.5 Representasi mel-spectrogram.	69
Gambar 4.6 Representasi log-mel spectrogram.	70
Gambar 4.7 Bentuk tensor input CNN-LSTM.	72
Gambar 4.8 Ringkasan arsitektur model CNN-LSTM.	78
Gambar 4.9 Grafik loss dan accuracy model log-mel spectrogram.	80
Gambar 4.10 Grafik loss dan accuracy model mel-spectrogram.	81
Gambar 4.11 Classification report model log-mel spectrogram.	85
Gambar 4.12 Confusion matrix model log-mel spectrogram.	87
Gambar 4.13 Classification report model mel-spectrogram.	90
Gambar 4.14 Confusion matrix model mel-spectrogram.	92
Gambar 4.15 Tampilan pengujian file audio WAV pada sistem.	97
Gambar 4.16 Hasil prediksi akor per segmen pada file audio WAV.	99
Gambar 4.17 Hasil penyusunan prediksi akor menjadi progresi akor.	102

DAFTAR LAMBANG DAN SINGKATAN



AI	Artificial Intelligence
ML	Machine Learning
DL	Deep Learning
MIR	Music Information Retrieval
STFT	Short-Time Fourier Transform
CNN	Convolutional Neural Network
RNN	Recurrent Neural Network
LSTM	Long Short-Term Memory
WAV	Waveform Audio File Format
IDMT	Institut für Digitale Medientechnologie
SMT	Symbolic Music Transcription
Hz	Hertz (satuan frekuensi)
dB	Decibel
log	Logaritma

DAFTAR ISTILAH



Artificial Intelligence	Bidang ilmu komputer yang berfokus pada pengembangan sistem yang mampu meniru kecerdasan manusia
Machine Learning	Metode pembelajaran mesin yang memungkinkan sistem belajar dari data tanpa pemrograman eksplisit
Deep Learning	Cabang machine learning berbasis jaringan saraf tiruan dengan banyak lapisan
Music Information Retrieval (MIR)	Bidang yang mempelajari ekstraksi dan analisis informasi dari data musik
Akor (Chord)	Kombinasi beberapa nada yang dimainkan secara bersamaan
Progresi Akor	Urutan akor dalam suatu bagian musik
Audio Processing	Proses pengolahan sinyal audio untuk ekstraksi informasi
STFT	Metode transformasi sinyal untuk melihat perubahan frekuensi terhadap waktu
Mel-Scale	Skala frekuensi yang disesuaikan dengan persepsi pendengaran manusia
Mel-Spectrogram	Representasi visual spektrum frekuensi berbasis skala Mel
Log-Mel Spectrogram	Mel-spectrogram yang telah ditransformasi logaritmik untuk menyeimbangkan energi frekuensi
CNN	Model deep learning yang digunakan untuk ekstraksi fitur spasial
RNN	Model neural network untuk data berurutan (sequential)



LSTM	Varian RNN yang mampu menangkap dependensi jangka panjang
CNN-LSTM	Arsitektur hybrid yang menggabungkan CNN dan LSTM
Dataset	Kumpulan data yang digunakan dalam penelitian
Segmentasi Audio	Proses membagi sinyal audio menjadi bagian-bagian kecil
Ekstraksi Fitur	Proses mengambil informasi penting dari data mentah
Tensor	Struktur data multidimensi yang digunakan dalam deep learning
Training	Proses pelatihan model menggunakan data
Validation	Proses evaluasi model selama pelatihan
Testing	Pengujian model menggunakan data baru
Akurasi	Tingkat ketepatan prediksi model
Precision	Rasio prediksi benar terhadap total prediksi positif
Recall	Rasio prediksi benar terhadap total data aktual
F1-Score	Rata-rata harmonik antara precision dan recall
Confusion Matrix	Matriks evaluasi untuk melihat performa klasifikasi
Loss	Nilai kesalahan model selama pelatihan
Epoch	Satu siklus penuh pelatihan model
Hyperparameter	Parameter yang ditentukan sebelum pelatihan model

INTISARI

Identifikasi akor dan progresi akor pada musik digital masih sering dilakukan secara manual melalui pendengaran, sehingga membutuhkan waktu, ketelitian, dan kemampuan musikal yang cukup tinggi. Kondisi ini dapat menjadi kendala dalam proses pembelajaran musik, latihan ansambel, maupun pengembangan aplikasi analisis musik otomatis. Penelitian ini bertujuan mengembangkan model hybrid CNN-LSTM berbasis log-mel spectrogram untuk mengenali akor secara otomatis pada segmen audio musik digital, lalu menyusun hasil prediksi tiap segmen menjadi representasi progresi akor. Penelitian menggunakan dataset IDMT-SMT-CHORDS dengan fokus pada 24 kelas akor mayor dan minor. Tahapan penelitian meliputi akuisisi dataset, pembacaan anotasi, pemetaan label, segmentasi audio, praproses sinyal, ekstraksi fitur, pembentukan tensor input, pelatihan model, evaluasi, serta pengujian menggunakan file audio WAV buatan sendiri. Hasil penelitian menunjukkan bahwa model CNN-LSTM berbasis log-mel spectrogram memberikan performa yang lebih baik dibandingkan model berbasis mel-spectrogram, dengan nilai test accuracy 0,9398, macro precision 0,9426, macro recall 0,9393, dan macro F1-score 0,9388. Selain itu, sistem berhasil mengenali akor pada tiap segmen audio dan menyusunnya menjadi progresi akor yang sesuai dengan urutan permainan. Hasil penelitian ini dapat dimanfaatkan oleh musisi, pelajar musik, pengajar musik, dan pengembang aplikasi musik digital sebagai dasar pengembangan sistem analisis harmoni musik yang lebih cepat, objektif, dan efisien.

Kata kunci: pengenalan akor, progresi akor, CNN-LSTM, log-mel spectrogram, musik digital.

ABSTRACT

Chord and chord progression identification in digital music is still often carried out manually by listening, requiring considerable time, accuracy, and musical skill. This condition can become an obstacle in music learning, ensemble practice, and the development of automatic music analysis applications. This study aims to develop a hybrid CNN-LSTM model based on log-mel spectrogram to automatically recognize chords in digital music audio segments and then arrange the segment-level predictions into a chord progression representation. The study uses the IDMT-SMT-CHORDS dataset with a focus on 24 major and minor chord classes. The research stages include dataset acquisition, annotation reading, label mapping, audio segmentation, signal preprocessing, feature extraction, input tensor construction, model training, evaluation, and testing using a self-created WAV audio file. The results show that the CNN-LSTM model based on log-mel spectrogram performs better than the mel-spectrogram-based model, achieving a test accuracy of 0.9398, macro precision of 0.9426, macro recall of 0.9393, and macro F1-score of 0.9388. In addition, the system successfully recognizes chords in each audio segment and arranges them into a chord progression that matches the played sequence. The results of this study can be utilized by musicians, music students, music teachers, and digital music application developers as a basis for developing music harmony analysis systems that are faster, more objective, and more efficient.

Keyword: chord recognition, chord progression, CNN-LSTM, log-mel spectrogram, digital music.