

BAB I PENDAHULUAN

1.1 Latar Belakang

Dalam lima tahun terakhir, perkembangan kecerdasan buatan telah membawa perubahan signifikan dalam analisis musik digital, khususnya pada bidang *Music Information Retrieval* (MIR). Teknologi ini memungkinkan komputer melakukan analisis terhadap berbagai karakteristik musik seperti tempo, melodi, dan harmoni secara otomatis. Salah satu pendekatan yang banyak digunakan dalam analisis sinyal audio adalah representasi *mel-spectrogram*, karena mampu memetakan distribusi energi frekuensi audio sesuai dengan karakteristik persepsi pendengaran manusia. Penelitian terdahulu menunjukkan bahwa penggunaan *mel-spectrogram* yang diproses menggunakan arsitektur *Convolutional Neural Network* (CNN) mampu menghasilkan performa klasifikasi musik yang stabil dan kompetitif pada berbagai tugas klasifikasi audio [1]. Hal ini menunjukkan bahwa representasi spektral berbasis *time-frequency* merupakan dasar yang kuat dalam pengembangan sistem analisis musik otomatis berbasis *deep learning*.

Musik tidak hanya memiliki dimensi spektral, tetapi juga dimensi temporal yang berperan penting dalam membentuk struktur musikal. Salah satu bentuk struktur temporal tersebut adalah progresi akor, yaitu rangkaian perubahan harmoni yang terjadi secara berurutan sepanjang waktu. Analisis progresi akor tidak cukup dilakukan hanya dengan mengamati pola frekuensi pada satu waktu tertentu, tetapi memerlukan model yang mampu memahami hubungan antar segmen audio secara berurutan. Penelitian oleh Budaya dkk. menunjukkan bahwa pendekatan *Long Short-Term Memory* (LSTM) efektif dalam menangkap pola temporal pada data audio berbasis fitur spektral [2]. Oleh karena itu, integrasi CNN sebagai pengekstrak fitur spektral dan LSTM sebagai pemodel hubungan temporal menjadi pendekatan yang relevan untuk menganalisis karakter progresi akor yang bersifat dinamis.

Sebagian besar penelitian yang memanfaatkan CNN dalam analisis musik masih berfokus pada klasifikasi kategori yang bersifat statis, seperti genre musik atau jenis instrumen. Penelitian oleh Pradnya dan Raharja menunjukkan bahwa CNN berbasis spektrogram mampu menghasilkan performa yang baik dalam klasifikasi genre musik, tetapi belum secara khusus mengkaji struktur harmonik yang berkembang secara sekuensial dalam waktu [3]. Di sisi lain, penelitian internasional mengenai pengenalan akor berbasis *deep learning* menunjukkan bahwa pendekatan jaringan saraf mampu meningkatkan akurasi dalam mengenali struktur harmoni musik [4]. Meskipun demikian, kajian yang secara khusus mengintegrasikan CNN dan LSTM berbasis *log-mel spectrogram* untuk pengenalan akor otomatis dan penyusunan progresi akor pada *dataset* yang memang dirancang untuk akor masih relatif terbatas. Penelitian dalam bidang *automatic chord recognition* menunjukkan bahwa sistem pengenalan akor umumnya melibatkan tahap ekstraksi fitur, klasifikasi, dan penyusunan urutan akor. Representasi waktu-frekuensi seperti *log-mel spectrogram* dan CQT juga banyak digunakan sebagai fitur dalam proses pengenalan akor karena mampu merepresentasikan karakteristik spektral audio yang berkaitan dengan struktur harmoni musik [5].

Berdasarkan kesenjangan tersebut, penelitian ini menggunakan pendekatan *hybrid CNN-LSTM* berbasis *log-mel spectrogram* untuk mengenali akor secara otomatis pada lagu digital. Dalam implementasinya, representasi fitur yang digunakan sebagai masukan model adalah *log-mel spectrogram*. Pendekatan ini mengintegrasikan kemampuan CNN dalam mengekstraksi pola spektral dari representasi waktu-frekuensi dengan kemampuan LSTM dalam memodelkan hubungan temporal antar-frame dalam setiap segmen audio. Dengan demikian, model difokuskan untuk mengklasifikasikan jenis akor pada tiap segmen audio secara otomatis. Hasil klasifikasi tersebut kemudian disusun berdasarkan urutan waktu sehingga membentuk representasi progresi akor sebagai keluaran akhir yang merepresentasikan susunan akor hasil prediksi.

Selain memberikan kontribusi metodologis, penelitian ini juga memiliki urgensi praktis dalam pengembangan teknologi analisis musik otomatis. Dalam

konteks pembelajaran musik, identifikasi akor dan progresi akor suatu lagu masih sering dilakukan secara manual melalui pendengaran, sehingga memerlukan waktu, ketelitian, dan keterampilan musikal yang cukup tinggi. Oleh karena itu, penelitian ini diarahkan untuk merancang dan membangun sistem yang mampu mengenali akor secara otomatis pada setiap segmen audio, kemudian menyusun hasil prediksi tersebut berdasarkan urutan waktu sehingga membentuk representasi progresi akor sebagai keluaran akhir. Dengan pendekatan *deep learning* berbasis CNN-LSTM dan representasi *log-mel spectrogram*, penelitian ini tidak hanya berfokus pada peningkatan performa klasifikasi akor, tetapi juga menghasilkan sistem yang dapat digunakan untuk membantu musisi, pelajar musik, maupun pengembang aplikasi musik digital dalam memahami struktur harmonik lagu secara lebih cepat, objektif, dan efisien. Dengan demikian, tujuan akhir penelitian ini adalah menghasilkan model pengenalan akor otomatis yang mampu bekerja pada segmen audio dan menyajikan susunan progresi akor sebagai bentuk keluaran yang lebih aplikatif.

1.2 Rumusan Masalah

Berdasarkan latar belakang dan permasalahan yang telah diuraikan sebelumnya, maka rumusan masalah dalam penelitian ini dapat dirumuskan sebagai berikut:

1. Bagaimana merancang dan mengimplementasikan model *hybrid* CNN-LSTM berbasis *log-mel spectrogram* untuk pengenalan akor otomatis pada segmen audio musik digital?
2. Bagaimana menyusun hasil prediksi akor per segmen menjadi representasi progresi akor pada file audio?
3. Bagaimana performa model *hybrid* CNN-LSTM dalam mengenali 24 kelas akor mayor dan minor berdasarkan metrik *accuracy*, *precision*, *recall*, dan *F1-score*?

1.3 Batasan Masalah

Dalam penelitian ini terdapat beberapa batasan untuk diteliti, berikut batasan masalah penelitian yang dilakukan yaitu:

1. Penelitian ini difokuskan pada perancangan, implementasi, dan evaluasi model *hybrid CNN-LSTM* untuk pengenalan akor otomatis pada segmen audio. Hasil prediksi antarsegmen kemudian disusun berdasarkan urutan waktu untuk membentuk representasi progresi akor.
2. *Dataset* yang digunakan adalah IDMT-SMT-CHORDS, sehingga model tidak diuji pada lagu komersial berdurasi penuh atau *dataset* lain di luar data tersebut.
3. Penelitian hanya memfokuskan pada pengenalan akor dasar seperti mayor dan minor dalam bentuk 24 kelas (12 *root major* dan 12 *root minor*), sesuai dengan jenis akor yang tersedia dalam *dataset*. Akor kompleks seperti seventh, sus, dim, atau augmented tidak disertakan.
4. Seluruh tahapan penelitian, mulai dari praproses data, ekstraksi fitur *log-mel spectrogram*, perancangan arsitektur, hingga pelatihan dan evaluasi model, dilaksanakan menggunakan bahasa pemrograman Python dengan dukungan pustaka seperti TensorFlow, Keras, dan Librosa.
5. Penelitian ini hanya menggunakan dua jenis representasi fitur audio, yaitu *log-mel spectrogram* sebagai fitur utama dan *mel-spectrogram* sebagai fitur pembanding. Penelitian ini tidak membandingkan representasi fitur audio lain seperti chromagram, MFCC, constant-Q transform (CQT), maupun spectrogram STFT secara langsung.

1.4 Tujuan Penelitian

Tujuan yang akan dicapai oleh peneliti dalam penelitiannya adalah untuk:

1. Mengembangkan tahapan segmentasi audio dan ekstraksi fitur *log-mel spectrogram* untuk merepresentasikan segmen audio secara sesuai pada tugas pengenalan akor.
2. Menganalisis dan mengukur performa model CNN-LSTM dalam mengenali akor pada segmen audio menggunakan indikator *accuracy*, *precision*, *recall*, *F1-score*, *confusion matrix*, serta evaluasi *training*

dan *validation loss*.

3. Menghasilkan representasi progresi akor dari urutan hasil prediksi akor pada segmen-segmen audio.
4. Menganalisis perbandingan performa model CNN-LSTM berbasis *log-mel spectrogram* dan *mel-spectrogram* dalam tugas pengenalan akor otomatis.

1.5 Manfaat Penelitian

Penelitian ini diharapkan dapat memberikan kontribusi yang signifikan baik dari segi pengembangan ilmu pengetahuan (teoretis) maupun dari sisi implementasi dan pemanfaatan di dunia nyata (praktis). Adapun manfaat yang dapat diperoleh adalah sebagai berikut:

1. Manfaat Teoretis

Secara teoretis, penelitian ini diharapkan dapat memperkaya khazanah ilmu pengetahuan dalam bidang *Music Information Retrieval* (MIR), khususnya pada penerapan model *hybrid CNN-LSTM* berbasis *log-mel spectrogram* untuk pengenalan akor otomatis, serta penyusunan hasil prediksi menjadi representasi progresi akor. Hasil penelitian ini memberikan kontribusi berupa bukti empiris mengenai efektivitas integrasi ekstraksi fitur spektral dan pemodelan temporal dalam analisis sinyal audio. Selain itu, penelitian ini dapat menjadi referensi akademik dalam pengembangan model pembelajaran mendalam untuk klasifikasi dan pengenalan pola audio, serta membuka peluang pengembangan konsep serupa pada analisis struktur musik yang lebih kompleks di masa mendatang.

2. Manfaat Praktis

Penelitian ini diharapkan dapat dimanfaatkan oleh musisi, pengajar musik, dan pengembang aplikasi digital dalam mendukung proses identifikasi akor secara otomatis, yang selanjutnya dapat disusun

menjadi progresi akor suatu lagu. Model yang dikembangkan berpotensi membantu meningkatkan efektivitas sesi latihan musik dengan menyediakan informasi harmonik secara cepat dan objektif. Selain itu, hasil penelitian ini juga dapat menjadi acuan teknis bagi industri kreatif dalam merancang sistem analisis musik berbasis kecerdasan buatan, sehingga mendorong inovasi teknologi dan optimalisasi proses produksi maupun pembelajaran musik digital.

1.6 Sistematika Penulisan

Penulisan skripsi ini disusun ke dalam lima bab utama yang masing-masing memuat uraian sebagai berikut:

- a. BAB I PENDAHULUAN, berisi latar belakang, rumusan masalah, batasan masalah, tujuan penelitian, manfaat penelitian, dan sistematika penulisan.
- b. BAB II TINJAUAN PUSTAKA, berisi studi literatur dan dasar teori, yang mencakup *Artificial Intelligence*, *Machine Learning*, *Music Information Retrieval* (MIR), akor, progresi akor, *Deep Learning*, *Audio Processing*, *Short-Time Fourier Transform* (STFT), *Mel-Scale, Mel-Spectrogram*, *Log-Mel Spectrogram*, *Convolutional Neural Network* (CNN), *Recurrent Neural Network* (RNN), *Long Short-Term Memory* (LSTM), arsitektur *hybrid CNN-LSTM*, optimasi dan parameter pelatihan model, serta metrik evaluasi model.
- c. BAB III METODE PENELITIAN, berisi objek penelitian, alur penelitian, serta alat dan bahan, yang mencakup data penelitian dan alat/instrumen.
- d. BAB IV HASIL DAN PEMBAHASAN, berisi akuisisi dataset IDMT-SMT-CHORDS, ekstraksi file audio dan file anotasi, *parsing* label akor dari file anotasi, pemetaan label menjadi 24 kelas akor, segmentasi audio berdasarkan anotasi akor, *resampling* dan konversi audio ke mono, ekstraksi fitur *mel-spectrogram* dan *log-mel spectrogram*, normalisasi, *padding*, *cropping*, pembentukan *tensor input*, *encoding*

label dan pembagian data, perancangan model *hybrid CNN-LSTM*, pelatihan model dan pemilihan model terbaik, evaluasi model pada data *test*, perbandingan model berbasis *log-mel spectrogram* dan *mel-spectrogram*, pengujian dengan file audio WAV, serta prediksi akor per segmen dan pembentukan progresi akor.

- e. BAB V PENUTUP, menyimpulkan hasil penelitian, menjawab rumusan masalah, dan memberikan saran untuk pengembangan selanjutnya.

