

BAB 1

PENDAHULUAN

1.1 Latar Belakang Masalah

Musik menurut Kamus Besar Bahasa Indonesia memiliki arti (1) ilmu atau seni menyusun nada atau suara dalam urutan, kombinasi, dan hubungan temporal untuk menghasilkan komposisi (suara) yang mempunyai kesatuan dan kesinambungan; (2) nada atau suara yang disusun demikian rupa sehingga mengandung irama, lagu, dan keharmonisan (terutama yang menggunakan alat-alat yang dapat menghasilkan bunyi-bunyi itu), sedangkan Jamalus mengartikannya sebagai bentuk karya seni seseorang atau sekelompok orang yang mengekspresikan pikiran dan perasaannya ke dalam bentuk lagu atau komposisi musik [1]. Masing-masing musik yang dihasilkan oleh seseorang tentunya akan berbeda satu dengan yang lain karena musik merupakan suatu karya seni. Perbedaan ini kemudian dikelompokkan berdasarkan kemiripannya sehingga kelompok-kelompok itu disebut dengan genre musik.

Label yang digunakan oleh manusia untuk mengelompokkan dan mendeskripsikan dunia musik yang luas dikenal sebagai genre musik [2]. Terdapat banyak sekali genre musik, di antaranya yaitu *rock*, *metal*, klasik, *blues*, *jazz*, *pop*, *hiphop*, dan masih banyak lagi. Setiap orang tentunya memiliki preferensi tersendiri mengenai genre musik yang disukai. Genre lagu tidak dapat ditentukan hanya dengan mendengarkan sepenggal bagiannya saja, tetapi lagu perlu didengarkan secara utuh agar genre dapat ditentukan. Oleh karena itu, dibutuhkan sebuah alat bantu yang dapat mengelompokkan genre musik secara otomatis agar pengelompokan dapat dilakukan dengan lebih efektif dan efisien.

Music Information Retrieval (MIR) adalah sebuah bidang penelitian yang berfokus pada analisis dan ekstraksi informasi musik, seperti deteksi ketukan (*beat-detection*), deteksi onset, transkripsi musik, dan klasifikasi genre [3]. Penelitian mengenai klasifikasi genre musik mulai berkembang, terutama penelitian yang menggunakan *deep learning*. Penelitian [4] membandingkan algoritma CNN (*Convolutional Neural Network*) dan LSTM (*Long Short Term Memory*) dengan menggunakan dua *dataset* yaitu GTZAN *dataset* dan FMA (*Free Music Archive*) *dataset*. Hasil yang didapatkan saat menggunakan GTZAN *dataset* yaitu akurasi mencapai 56% menggunakan algoritma CNN dan 42% menggunakan algoritma

LSTM, sedangkan saat menggunakan FMA *dataset*, akurasi yang didapatkan hanya mencapai 50,5% menggunakan algoritma CNN dan 33,5% menggunakan algoritma LSTM. Penelitian [5] membuat sistem klasifikasi genre menggunakan GTZAN *dataset* dan algoritma CNN dengan akurasi mencapai 83%. Penelitian [6] membandingkan beberapa algoritma *deep learning* dengan menggunakan FMA *dataset*. Akurasi algoritma CRNN (*Convolutional-Recurrent Neural Network*) mencapai 90%, melebihi akurasi yang dimiliki algoritma lain.

Berdasarkan beberapa literatur yang telah dijabarkan sebelumnya, akurasi yang didapatkan oleh algoritma *deep learning* yang sering digunakan (seperti CNN dan RNN atau *Recurrent Neural Network*) belum maksimal. Sementara itu, algoritma CRNN mendapatkan akurasi yang paling tinggi. Algoritma CRNN adalah algoritma yang menggabungkan CNN dengan RNN dalam arsitekturnya. Telaah literatur mengenai penggunaan algoritma CRNN dalam mengklasifikasi genre musik dilakukan lebih lanjut. Penelitian [6], yang sudah disebutkan sebelumnya, melakukan komparasi antara beberapa algoritma *machine learning* (SVC atau *Support Vector Classifier*, regresi logistik, dan *ensemble learning* menggunakan AdaBoost) dengan algoritma *deep learning* (ANN atau *Artificial Neural Network*, CNN, CRNN dengan gabungan CNN-LSTM, dan PCRNN atau Paralel CRNN). *Input* yang digunakan pada penelitian ini adalah *feature matrix* untuk model *machine learning* dan Mel-spectrogram untuk model *deep learning* dengan *dataset* yang digunakan adalah FMA *dataset*. Algoritma yang menghasilkan akurasi tertinggi pada penelitian ini yaitu algoritma CRNN sebesar 90%. Akurasi yang dihasilkan pada penelitian ini mencapai angka 90% karena jumlah data yang digunakan hanya 480 audio dari total 8000 audio sehingga *dataset* belum digunakan secara maksimal. Penelitian [7] melakukan klasifikasi genre musik dengan beberapa kombinasi algoritma CRNN, yaitu CNN-LSTM, CNN-GRU, CNN-BiLSTM, dan CNN-BiGRU. *Input* yang digunakan adalah Mel-spectrogram dan MFCC (*Mel-Frequency Cepstral Coefficient*) dengan menggunakan GTZAN *dataset*. Akurasi tertinggi yang didapatkan ketika menggunakan *input* Mel-spectrogram adalah CNN-BiGRU sebesar 89,3%, sedangkan akurasi tertinggi ketika menggunakan *input* MFCC adalah CNN-LSTM sebesar 76,4%. Penelitian [8] melakukan komparasi hasil algoritma CNN dan CRNN (CNN-LSTM dan CNN-BiLSTM) dalam mengklasifikasi genre musik dan hasil tersebut digunakan untuk mendapatkan rekomendasi genre musik. *Input* yang digunakan adalah Mel-spectrogram yang didapatkan menggunakan FMA *dataset*. Akurasi tertinggi yang didapatkan pada penelitian ini yaitu 72% menggunakan algoritma CNN-LSTM.

Berdasarkan literatur mengenai algoritma CRNN yang sudah dijelaskan sebelumnya, penelitian ini akan menggunakan algoritma CRNN dengan sumber datanya yaitu GTZAN *dataset*. Algoritma ini dapat mengekstrak fitur-fitur lokal dan melakukan agregasi pola temporal dari hasil ekstrak tersebut. *Dataset* akan diolah menjadi MFCC yang kemudian akan dijadikan *input* model. Jenis arsitektur CRNN yang digunakan pada penelitian adalah CNN-BiLSTM (*Bidirectional Long Short Term Memory*). Penelitian ini menggunakan salah satu algoritma yang digunakan dari penelitian [7], yaitu CNN-BiLSTM, dengan arsitektur yang lebih sederhana. Penelitian [7] menggunakan lima lapisan CNN dan tiga lapisan RNN (LSTM-BiLSTM-LSTM) serta menggunakan *feature extraction* MFCC yang dibagi menjadi 10 segmen yang mendapatkan hasil akurasi 73,69%. Penelitian ini akan menggunakan tiga lapisan CNN dan dua lapisan RNN yang terdiri dari BiLSTM-BiLSTM dan *feature extraction* yang digunakan adalah MFCC yang dibagi menjadi 10 segmen. Penelitian ini bertujuan untuk mengembangkan penelitian mengenai klasifikasi genre musik menggunakan algoritma CRNN. Selain itu, penelitian ini dilakukan untuk melihat apakah dengan arsitektur yang lebih sederhana daripada penelitian sebelumnya dapat menghasilkan akurasi yang lebih baik.

1.2 Rumusan Masalah

Berdasarkan latar belakang masalah tersebut, maka ditentukan rumusan masalah sebagai berikut.

1. Bagaimana cara mengimplementasikan algoritma CRNN dalam mengklasifikasi genre musik?
2. Bagaimana hasil akurasi, *precision*, *recall*, dan F1 *score* saat mengklasifikasi genre musik?

1.3 Batasan Permasalahan

Penelitian ini memiliki batasan mengenai luasan dan cakupan yang hendak diselesaikan. Berikut merupakan batasan-batasan yang telah ditentukan sebelumnya.

1. Model *deep learning* hanya terbatas dalam mengklasifikasi 10 genre musik, yaitu *blues*, *classical*, *country*, *disco*, *hiphop*, *jazz*, *metal*, *pop*, *reggae*, dan *rock*.

2. Data yang digunakan untuk melatih model diambil dari GTZAN Dataset.
3. Algoritma yang digunakan yaitu CRNN dengan menggunakan *library* Keras.

1.4 Tujuan Penelitian

Berdasarkan rumusan masalah yang telah ditetapkan, terdapat beberapa tujuan yang hendak dicapai oleh penelitian ini, antara lain sebagai berikut.

1. Mengetahui cara mengimplementasikan algoritma CRNN agar dapat mengklasifikasi genre musik.
2. Mengetahui hasil akurasi, *precision*, *recall*, dan *F1 score* dari model *deep learning* dalam mengklasifikasi genre musik.

1.5 Manfaat Penelitian

Penelitian ini diharapkan dapat bermanfaat bagi beberapa pihak, yaitu penulis, peneliti lain, masyarakat umum, dan universitas.

1. Bagi penulis, penelitian ini diharapkan dapat menambah kontribusi pada bidang *Music Information Retrieval* (MIR), spesifiknya yaitu mengenai klasifikasi genre musik.
2. Bagi masyarakat umum, penelitian ini diharapkan dapat menambah wawasan dan pengetahuan mengenai klasifikasi genre musik dan bidang *deep learning*.
3. Bagi universitas, penelitian ini diharapkan dapat menambah kepustakaan literatur milik Universitas Multimedia Nusantara.

1.6 Sistematika Penulisan

Berikut merupakan sistematika penulisan yang digunakan untuk menyusun laporan penelitian.

- Bab 1 PENDAHULUAN

Bab ini menjelaskan mengenai latar belakang masalah, rumusan masalah, batasan permasalahan, tujuan penelitian, manfaat penelitian, dan sistematika penulisan.

- Bab 2 LANDASAN TEORI

Bab ini menjelaskan mengenai teori-teori yang digunakan pada penelitian ini, seperti GTZAN dataset, Mel-Frequency Cepstral Coefficient (MFCC), dan algoritma *Convolutional-Recurrent Neural Network* (CRNN).

- Bab 3 METODOLOGI PENELITIAN

Bab ini menjelaskan mengenai metodologi penelitian yang digunakan dalam penelitian ini.

- Bab 4 HASIL DAN DISKUSI

Bab ini menjelaskan mengenai proses dan hasil yang didapatkan pada penelitian ini.

- Bab 5 SIMPULAN DAN SARAN

Bab ini menjelaskan mengenai kesimpulan dari penelitian ini dan saran yang dapat diterapkan untuk penelitian serupa ke depannya.

