

# BAB 1 PENDAHULUAN

## 1.1. Latar Belakang

Otak manusia memiliki kemampuan untuk mempersepsi sumber suara dengan akurat. Hal ini mungkin dilakukan oleh mesin dengan memanfaatkan bidang keilmuan *machine learning* untuk membuat mesin mempelajari suatu pola yang ada pada sinyal suara, dan menjadikan hal yang telah dipelajari tersebut sebagai dasar untuk mengenali sumber suara yang diamati. Hal ini dilakukan dengan membandingkan dan mencari kesamaan pola pada suara yang ada pada kamus yang dimiliki oleh mesin dengan suara yang diamati. Sejumlah analisis dan metode yang melibatkan disiplin ilmu *machine learning* telah dikembangkan untuk mencapai tujuan ini.

*Machine Learning* adalah suatu cabang ilmu komputer dan merupakan pengembangan dari bidang kecerdasan artifisial. Cabang ilmu ini dapat dimanfaatkan untuk memungkinkan mesin atau komputer mempelajari suatu set data menggunakan algoritma tertentu sesuai kebutuhannya, dan kemudian melakukan aksi sesuai apa yang telah dipelajari dari set data tersebut. Hal-hal yang telah terlahir dari cabang ilmu ini diantaranya adalah mobil swa-kemudi, pengenalan pola pencarian pada suatu *e-commerce*, pencarian pada *web* yang lebih efektif, dan sebagainya. Ilmu dan implementasi *machine learning* terus berkembang dan memungkinkan mesin untuk mempelajari set data yang lebih kompleks untuk tugas yang kompleks pula. Salah satu tugas kompleks yang mungkin dicapai mesin dengan implementasi *machine learning* adalah separasi sumber suara pada *polyphonic audio*, di mana suara dari sejumlah sumber yang berbeda dimainkan secara bersamaan sehingga sinyal suara yang tercipta menjadi saling tumpang tindih [1].

Suatu metode yang dapat digunakan untuk analisis sinyal suara yang tumpang tindih adalah *non-negative matrix factorization* atau NMF. NMF mendekomposisi data untuk mencari pola yang terdapat pada suatu data. *Probabilistic latent component analysis* atau PLCA kemudian dikembangkan dari NMF [1], dengan mengadaptasi sifat probabilistik. Pada tugas akhir ini, masalah separasi sumber suara pada *polyphonic audio* diutarakan menggunakan disiplin ilmu *machine learning*, secara spesifik menggunakan PLCA sebagai metode untuk mencapai hal tersebut. Untuk dapat memanfaatkan metode ini, sinyal suara harus ditransformasikan terlebih dahulu ke bentuk distribusi energi berupa *spectral value* pada *spectrogram*.

*Spectrogram* ini didapatkan melalui proses transformasi sinyal dari domain waktu ke domain *time-frequency*. *Spectrogram* adalah representasi sinyal yang mengandung informasi distribusi energi akustik pada domain frekuensi/waktu. Dengan demikian, penggunaan teknik yang memiliki sifat probabilitas memungkinkan untuk kasus ini. Secara spesifik, teknik tersebut digunakan untuk mengekstraksi fitur berupa distribusi marginal terhadap waktu dan terhadap frekuensi. Kedua marginal ini digunakan untuk konstruksi objek suara. Ekstraksi distribusi marginal tersebut

dilakukan dengan iterasi algoritma *Expectation-Maximalization*. Untuk dapat menjalankan algoritma EM dan mengekstraksi distribusi marginal dari *spectrogram*, PLCA membutuhkan sejumlah parameter yang variatif, yang sangat berpengaruh terhadap hasil separasi. Di samping itu, parameter yang digunakan saat transformasi sinyal juga memiliki pengaruh langsung terhadap kemampuan separasi suara oleh PLCA. Tugas akhir ini menganalisis pengaruh parameter-parameter yang terkait dalam proses separasi sumber suara PLCA.

## 1.2. Rumusan Masalah

Berdasarkan uraian pada latar belakang, maka rumusan masalah pada tugas akhir ini dapat disimpulkan sebagai berikut.

1. Bagaimana pengaruh variabel-variabel yang terlibat pada proses separasi menggunakan PLCA?
2. Bagaimana kualitas *output* suara yang dihasilkan oleh pemisahan sumber dari campuran suara menggunakan PLCA?

Pada penelitian ini terdapat beberapa batasan masalah untuk menghindari meluasnya materi pembahasan. Adapun batasan masalah pada penelitian ini yaitu sebagai berikut.

- a. Dataset berupa objek audio atau *file* audio yang masing-masing mengandung suara satu alat music dan vocal.
- b. Format objek atau file audio yang digunakan adalah *Waveform Audio format* (.wav).
- c. Separasi pada penelitian ini berasumsi bahwa sumber suara atau suara tersebut telah dikenali sebelumnya oleh sistem.
- d. Penelitian pada tugas akhir ini berfokus pada separasi atau pemisahan sumber suara pada suatu *mixture* atau campuran audio menggunakan PLCA.
- e. Suatu campuran suara terdiri atas dua sumber suara.

## 1.3. Tujuan

Untuk menjawab rumusan masalah, tujuan dari tugas akhir ini antara lain:

1. Mengetahui pengaruh variabel-variabel yang terlibat dalam implementasi metode.
2. Mengetahui tingkat efektivitas PLCA dalam separasi atau pemisahan sumber suara terkait kualitas output.

## 1.4. Metodologi Penelitian

Tahapan yang akan dilakukan dalam pembuatan Tugas Akhir ini adalah:

a. Identifikasi Masalah

Pada tahap awal penelitian dilakukan identifikasi masalah mengenai separasi audio berdasarkan sumbernya, melalui literatur-literatur yang berhubungan dengan *audio processing* secara umum.

b. Studi Literatur

Pencarian literatur-literatur yang berhubungan dengan topik yang dibahas, yaitu tentang PLCA dan separasi sumber suara, dan langkah-langkah yang diperlukan sebelum dapat memanfaatkan PLCA untuk identifikasi dan separasi suara berdasarkan sumbernya.

c. Pengumpulan Data

Pada tahap ini dilakukan pengumpulan data yang akan digunakan untuk menguji sistem yang akan dibuat. Data yang digunakan berupa objek atau file audio dengan format .wav, dimana masing-masing file mengandung suara instrumen music atau vocal manusia.

d. Analisis Kebutuhan dan Perancangan Sistem

Pada tahap ini dilakukan perancangan sistem sesuai dengan penjelasan dari literatur acuan mengenai separasi audio dan metode PLCA. Selain itu pada tahap ini juga terjadi rancangan *preprocessing* untuk data yang akan digunakan dalam sistem.

e. Implementasi Sistem

Pembangunan sistem yang telah dirancang. Implementasi melibatkan *preprocessing* terhadap data, ekstraksi fitur objek atau file audio, transformasi sinyal suara yang didapatkan dari ekstraksi objek audio menjadi *spectrogram*, proses *learning dataset* yang telah ditransformasikan ke *spectrogram*, dan melakukan tahap-tahap pemisahan suara dari campuran suara.

f. Pengujian Sistem dan Analisis Hasil

Melakukan pengujian terhadap sistem yang dibuat menggunakan metode-metode untuk evaluasi terhadap kualitas *output* yaitu *mean opinion score* (MOS), dan *correlation*.

g. Penyusunan Laporan

Membuat laporan untuk dokumentasi seluruh kegiatan yang telah dilakukan beserta hasilnya dalam pengerjaan Tugas Akhir ini.

## 1.5. Sistematika Penulisan

Sistematika penulisan pada penelitian ini secara umum terdiri dari 5 (lima) dengan rincian sebagai berikut.

a. BAB 1 PENDAHULUAN

Pada bab ini membahas latar belakang, rumusan masalah, tujuan, batasan masalah, metodologi penelitian dan sistematika penulisan.

b. BAB 2 TINJAUAN PUSTAKA

Pada bab ini berisi pemaparan landasan-landasan teori yang berkaitan dengan penelitian ini.

c. BAB 3 PERANCANGAN SISTEM

Pada bab ini berisi tentang perancangan dari sistem yang akan dibuat yang terdiri dari *preprocessing* data, transformasi sinyal, implementasi PLCA untuk ekstraksi fitur *spectrogram*, dan implementasi PLCA untuk separasi suara dari campurannya.

d. BAB 4 PENGUJIAN DAN ANALISIS HASIL

Pada bab ini berisi tentang penjelasan hasil pengujian sistem dan analisis hasil yang didapatkan pada penelitian ini dengan membandingkan beberapa kondisi pengujian terhadap hasil performansi sistem.

e. BAB 5 KESIMPULAN DAN SARAN

Pada bab ini berisi kesimpulan-kesimpulan yang didapat dari hasil penelitian dan berisi saran-saran yang dapat dilakukan untuk mengembangkan penelitian ini lebih lanjut.