

BAB 1

PENDAHULUAN

1.1 Latar Belakang

Ekologi merupakan ilmu yang mempelajari hubungan timbal balik antara makhluk hidup dengan lingkungannya (Odum, 1971). Dalam konteks ekologi dan konservasi, pemantauan keanekaragaman hayati menjadi aspek penting untuk menjaga keseimbangan ekosistem. Salah satu pendekatan dalam pemantauan ini adalah dengan mengidentifikasi jenis hewan berdasarkan suara yang dihasilkan. Pendekatan ini tidak hanya memungkinkan studi terhadap perilaku dan komunikasi hewan, tetapi juga dapat membantu dalam pelestarian spesies yang sulit diamati secara visual di alam liar. Identifikasi spesies berdasarkan suara telah menjadi metode yang umum digunakan dalam bidang ekologi akustik karena memungkinkan pemantauan fauna secara non-invasif dan berkelanjutan (Blumstein dkk., 2011).

Pengenalan suara hewan dapat memanfaatkan dua aspek utama, yaitu pengolahan sinyal suara dan kecerdasan buatan. Dalam beberapa tahun terakhir, teknologi pengenalan suara mengalami perkembangan pesat dan telah diterapkan dalam berbagai bidang seperti asisten virtual, sistem keamanan, serta interaksi manusia dan mesin (Wijaya, 2024). Dalam konteks ekologi, pengenalan suara hewan dapat digunakan untuk memantau aktivitas fauna di habitat aslinya secara non-invasif, serta memberikan informasi penting terkait dinamika populasi dan perubahan lingkungan.

Menurut penelitian yang dilakukan oleh LeBien dkk. (2020) menyatakan bahwa identifikasi multi-spesies di lingkungan tropis merupakan tugas yang kompleks karena tingginya kebisingan latar, tumpang tindih suara antar spesies, dan minimnya data vokalisasi sebagai referensi. Kesulitan ini disebabkan oleh berbagai faktor, seperti tumpang tindih frekuensi antar spesies, gangguan suara latar seperti angin dan hujan, serta keterbatasan dokumentasi vokalisasi dari banyak spesies hewan. Oleh karena itu, dibutuhkan sistem otomatis yang mampu melakukan pengenalan suara dengan tingkat akurasi yang tinggi untuk membantu peneliti dan pemerhati lingkungan.

Salah satu pendekatan untuk menyelesaikan permasalahan pengenalan suara adalah dengan menggunakan deep learning. Deep learning merupakan salah satu cabang dari machine learning yang menggunakan jaringan saraf tiruan (artificial neural network) dengan banyak lapisan untuk mengambil fitur kompleks dari data. Keunggulan utama deep learning adalah kemampuannya melakukan feature extraction secara otomatis tanpa memerlukan proses ekstraksi manual, sehingga mampu menghasilkan performa yang lebih baik dibanding metode konvensional, khususnya untuk data yang memiliki pola rumit seperti sinyal suara dan citra (LeCun dkk., 2015).

Dalam bidang pengenalan suara, berbagai penelitian telah membuktikan efektivitas deep learning. Misalnya, (Stowell, 2021) menggunakan CNN untuk mengklasifikasikan suara burung liar dari rekaman lingkungan, dan berhasil mencapai akurasi tinggi meskipun data mengandung noise. Penelitian lain oleh (Nanni dkk., 2020) menggabungkan CNN dengan teknik augmentasi data untuk mendeteksi suara amfibi di habitat alami, yang terbukti meningkatkan akurasi dibandingkan metode berbasis fitur tradisional. Temuan-temuan ini menunjukkan bahwa deep learning, khususnya arsitektur CNN, sangat sesuai untuk pengenalan suara dengan variabilitas tinggi.

Dalam penelitian ini, pengenalan suara hewan tidak dilakukan hingga tingkat spesies, tetapi dikelompokkan ke dalam kategori taksonomi besar. Teknologi ini, khususnya Convolutional Neural Network (CNN), telah terbukti efektif dalam menyelesaikan berbagai permasalahan klasifikasi gambar dan suara. Pada penelitian ini, akan digunakan arsitektur MobileNetV2, yaitu varian ringan dari CNN yang dirancang untuk efisiensi tinggi dan penggunaan pada perangkat dengan sumber daya terbatas seperti smartphone Android (Wang dkk., 2019). MobileNetV2 menggunakan teknik depthwise separable convolution untuk mengurangi jumlah parameter dan komputasi, sehingga cocok untuk implementasi di perangkat bergerak (Howard dkk., 2019).

Data yang digunakan dalam sistem ini adalah suara hewan yang diubah ke dalam bentuk visual berupa spectrogram. Spectrogram adalah representasi visual dari sinyal suara dalam domain waktu dan frekuensi, di mana intensitas atau amplitudo suara digambarkan dalam bentuk warna atau tingkat kecerahan. Untuk

menghasilkan spectrogram, terdapat beberapa metode transformasi sinyal yang umum digunakan, seperti *Short-Time Fourier Transform* (STFT), *Mel-Frequency Cepstral Coefficient* (MFCC), dan *Mel-Spectrogram*. Metode ini memungkinkan sistem untuk mengidentifikasi pola-pola khas dari suara hewan dalam bentuk visual, mirip dengan cara CNN mengenali pola dalam gambar (Li dkk., 2022).

Adapun dalam penelitian ini, proses transformasi suara menjadi spectrogram akan dilakukan dengan metode STFT dan *Mel-Spectrogram*. STFT bekerja dengan membagi sinyal menjadi segmen-segmen kecil (frame), kemudian menerapkan transformasi Fourier pada setiap segmen untuk mendapatkan distribusi frekuensi sepanjang waktu (Allen, 1977; Lawrence R. Rabiner, 2011). Sementara itu, *Mel-Spectrogram* menyesuaikan skala frekuensi berdasarkan persepsi pendengaran manusia, sehingga lebih sesuai untuk analisis sinyal suara alami (McFee, 2015; Müller, 2015; Stevens dkk., 1937)

Sistem pengenalan suara ini dirancang untuk berjalan pada perangkat Android, sehingga memungkinkan pengguna untuk merekam suara hewan langsung dari mikrofon perangkat. Suara tersebut kemudian diubah menjadi spectrogram dan diklasifikasikan menggunakan model MobileNetV2 untuk mengidentifikasi kategori kelas taksonomi hewan secara otomatis. Sistem ini tidak bertujuan mengidentifikasi spesies tertentu, melainkan mengelompokkan suara ke dalam kelas taksonomi. Dengan pendekatan ini, sistem diharapkan dapat menjadi solusi praktis dalam mendukung kegiatan pemantauan ekologis dan konservasi satwa liar, terutama di lapangan tanpa memerlukan peralatan khusus tambahan.

1.2 Perumusan Masalah

Berdasarkan latar belakang yang telah diuraikan, penelitian ini akan menjawab beberapa permasalahan berikut:

- 1) Bagaimana mengembangkan sistem berbasis deep learning yang dapat mengklasifikasikan suara hewan ke dalam kategori taksonomi (Aves, Insecta, Amphibia, Mammalia, Reptilia) secara otomatis?
- 2) Bagaimana perbandingan performa sistem pengenalan suara hewan berbasis deep learning antara metode ekstraksi fitur menggunakan STFT dan *Mel-Spectrogram*?

1.3 Batasan Masalah

Agar ruang lingkup penelitian tidak terlalu luas, batasan masalah yang diterapkan adalah sebagai berikut:

- 1) Sistem ini hanya akan mengklasifikasikan suara dari sejumlah spesies hewan tertentu yang tersedia dalam dataset iNaturalist Sounds Dataset (iNatSounds) yang dapat diakses melalui <https://cyl-umass.github.io/iNatSounds>.
- 2) Sistem hanya akan mengklasifikasikan pada tingkat kelas taksonomi (Aves, Mammalia, Amphibia, Insecta, dan Reptilia), bukan hingga spesies individu. Jumlah data pada setiap kategori disamakan mengikuti jumlah minimum data (kategori Mammalia) untuk menjaga keseimbangan distribusi dan menghindari bias pada model. Data suara yang digunakan dalam pelatihan dan pengujian berasal dari dataset publik.
- 3) Model CNN yang digunakan adalah MobileNetV2 yang telah diadaptasi untuk pengolahan citra *spectrogram*.
- 4) Evaluasi sistem dilakukan berdasarkan metrik akurasi, presisi, *recall*, dan F1-score.
- 5) Implementasi sistem akan dilakukan dalam lingkungan simulasi dan integrasi model ke aplikasi Android menggunakan TensorFlow Lite.

1.4 Tujuan dan Manfaat

1.4.1 Tujuan

Adapun tujuan dari penelitian ini adalah:

- 1) Mengembangkan sistem pengenalan suara hewan berbasis deep learning menggunakan CNN dengan arsitektur MobileNetV2.
- 2) Mengubah data suara hewan menjadi representasi fitur menggunakan dua metode, yaitu STFT dan *Mel-Spectrogram*, untuk keperluan pelatihan model CNN.
- 3) Membandingkan performa sistem pengenalan suara hewan yang menggunakan metode STFT dan *Mel-Spectrogram* berdasarkan metrik evaluasi seperti akurasi, presisi, *recall*, dan F1-score.

1.4.2 Manfaat

Hasil dari penelitian ini diharapkan dapat memberikan manfaat sebagai

berikut:

- 1) Mempermudah proses identifikasi kategori hewan berdasarkan suara hewan secara otomatis.
- 2) Mendukung penelitian dalam bidang ekologi dan konservasi satwa liar dengan menyediakan alat bantu berbasis kecerdasan buatan.
- 3) Memberikan informasi komparatif tentang efektivitas dua metode ekstraksi fitur suara (STFT dan *Mel-Spectrogram*) dalam sistem klasifikasi suara hewan.