

KLASIFIKASI SPESIES BURUNG BERDASARKAN SUARA MENGGUNAKAN DEEP LEARNING DENGAN FITUR MEL-SPEKTOGRAM DAN MEL FREQUENCY CEPTRAL COEFFICIENTS (MFCC)

Daniswara Aditya Putra¹ dan Kartika Fithriasari²

^{1,2}Departemen Statistika, Institut Teknologi Sepuluh Nopember
e-mail: ¹daniswara.dap@gmail.com

Latar Belakang



Burung sebagai indikator alami yang mencerminkan kesehatan ekosistem



Identifikasi spesies lewat suara memungkinkan pemantauan tanpa mengganggu habitat



Tantangan utama adalah variasi suara, noise lingkungan, dan durasi yang berbeda-beda



Deep learning dapat mengatasi tantangan ini dengan mengenali pola suara secara akurat

Tujuan Penelitian

- Membangun model klasifikasi suara burung dengan CNN, MobileNet, dan VGGish menggunakan fitur Mel-Spektrogram dan MFCC
- Menentukan model terbaik berdasarkan performa evaluasi seperti accuracy, precision, recall, dan f1-score

Spesies Burung yang Diteliti



Burung Kutilang



Burung Perkutut



Burung trucidan



Burung Gereja



Burung Tekukur



Burung Cendet

Alur Penelitian

Audio

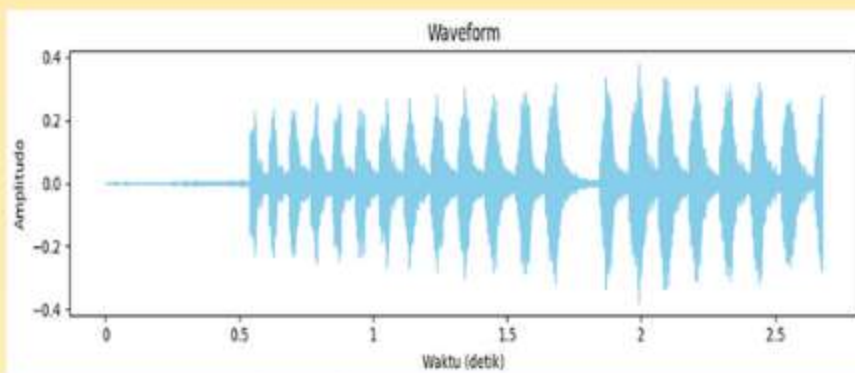
Trimming

Ekstraksi Fitur Mel-Spektrogram dan MFCC

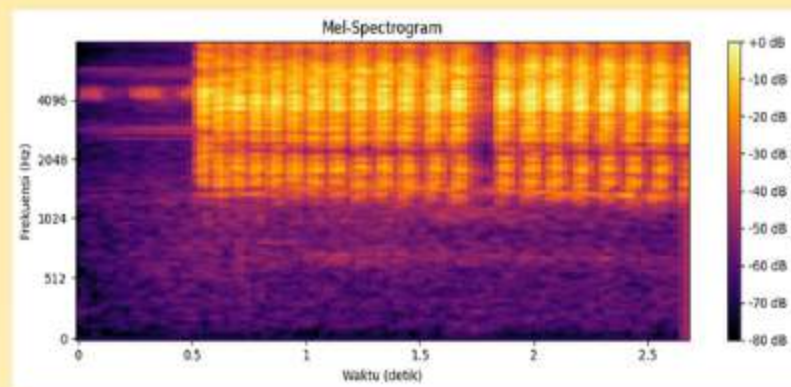
Split Data (Training, Validation, Test)

Pelatihan Model

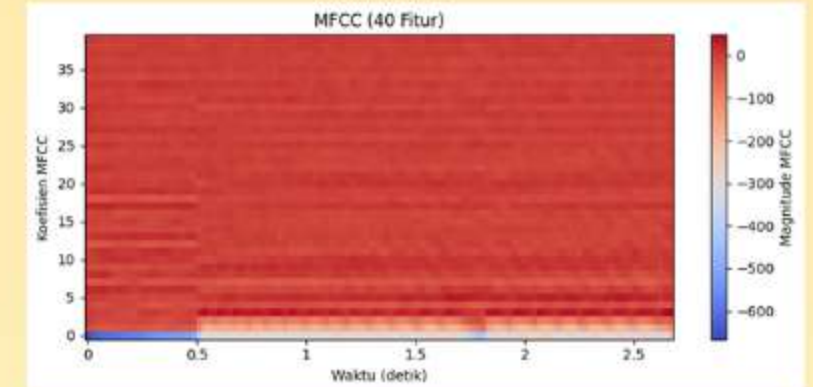
Evaluasi Model



Contoh Visualisasi Gelombang Suara



Contoh Visualisasi Mel-Spektrogram



Contoh Visualisasi MFCC

Penelitian diawali dengan **pengumpulan rekaman suara burung** dari situs **xeno-canto.org**, kemudian dilakukan **trimming** menggunakan **Voice Activity Detection (VAD)** untuk memisahkan segmen suara aktif. Setiap segmen diubah menjadi fitur **Mel-Spektrogram (128 fitur)** dan **MFCC (40 fitur)**, yang kemudian di-resize dengan **zero padding** agar sesuai dengan input model. **Data tidak berupa gambar, melainkan nilai numerik fitur audio yang digunakan sebagai input untuk pelatihan model.** Dataset dibagi secara stratifikasi menjadi 64% training, 16% validation, dan 20% test. Tiga model deep learning **CNN, MobileNet, dan VGGish** dilatih menggunakan kombinasi hyperparameter terbaik. Kemudian setiap model diuji pada data test dan hasilnya dianalisis menggunakan **accuracy, precision, recall, dan F1-score.**

Hasil Penelitian

Model	Test Accuracy	Test Recall	Test Precision	Test F1-Score
CNN (Mel-Spektrogram)	0.9161	0.9179	0.9080	0.9102
CNN (MFCC)	0.9790	0.9827	0.9778	0.9796
MobileNet (Mel-Spektrogram)	0.9693	0.9712	0.9659	0.9673
MobileNet (MFCC)	0.9661	0.9684	0.9637	0.9649
VGGish (Waveform)	0.9503	0.9522	0.9474	0.9485

Berdasarkan hasil evaluasi, model **CNN dengan input MFCC** menunjukkan performa terbaik dalam mengklasifikasikan enam jenis suara burung dengan **akurasi test sebesar 97,97%**, serta **precision, recall, dan F1-score di atas 97%**. Model ini mampu **mengenali pola suara secara akurat dan seimbang**, serta memiliki kemampuan **generalisasi yang kuat terhadap data baru**, mengungguli model lain seperti MobileNet dan VGGish.

Untuk pengembangan ke depan, disarankan agar **jumlah dan variasi data diperbanyak**, serta mempertimbangkan **arsitektur lain seperti EfficientNet atau CNN-LSTM**. Selain itu, hasil penelitian ini **berpotensi diimplementasikan dalam sistem monitoring otomatis berbasis audio** untuk mendukung upaya konservasi burung secara **efisien dan non-invasif**.

Scan for More Information

