UNIVERSITAS
GADJAH MADA
**Deteksi Crackle dan Wheeze pada Audio Respiratori menggunakan Augmentasi Roll Time Mixing-Hermit
Frequency Mixing dan Arsitektur Hybrid CNN-Multihead Transformer.**
Irfan Fadli Nugraha, Dr. Dyah Aruming Tyas, S.Si.
Universitas Gadjah Mada, 2025 | Diunduh dari http://etd.repository.ugm.ac.id/

INTISARI

Deteksi Crackle dan Wheeze pada Audio Respiratori menggunakan Augmentasi *Roll Time Mixing-Hermit Frequency Mixing* dan Arsitektur *Hybrid CNN-Multihead Transformer.*

Oleh
Irfan Fadli Nugraha 21/482012/PA/21010

Permasalahan dalam deteksi penyakit respiratori berbasis audio terletak pada keterbatasan data, keragaman pola suara pernapasan, serta gangguan noise dan ketidakseimbangan kelas dalam dataset medis. Deteksi suara abnormal seperti *crackle* dan *wheeze* memerlukan representasi spektral yang kuat serta model klasifikasi yang mampu menangkap karakteristik spasial dan temporal dari sinyal pernapasan. Tantangan ini semakin kompleks dalam konteks implementasi sistem diagnosis otomatis yang andal untuk lingkungan klinis nyata.

Untuk menjawab permasalahan tersebut, penelitian ini mengusulkan pendekatan deteksi fitur audio respiratori dengan menggabungkan teknik augmentasi *Roll Time Mixing* (RTM) dan *Hermit Frequency Mixing* (HFM) yang berfungsi memperkaya representasi log-mel spectrogram melalui manipulasi domain waktu dan frekuensi. Selanjutnya, digunakan arsitektur *Hybrid CNN-Multihead Transformer* untuk mengekstraksi fitur spasial dan temporal secara bersamaan. CNN digunakan untuk menangkap pola lokal, sedangkan transformer berfungsi memahami relasi jangka panjang antar fitur pada dimensi waktu.

Model yang diusulkan dievaluasi menggunakan dataset ICBHI 2017 dan menunjukkan performa yang lebih baik dibandingkan model baseline. Hasil eksperimen mencatat *balanced accuracy* sebesar 76,08%, *sensitivitas* sebesar 78,45%, dan *spesifisitas* sebesar 73,70%. Pencapaian ini mencerminkan peningkatan *balanced accuracy* sebesar 3,36% dan *sensitivitas* sebesar 12,14% dibandingkan model *state-of-the-art* sebelumnya.

UNIVERSITAS
GADJAH MADA
Deteksi Crackle dan Wheeze pada Audio Respiratori menggunakan Augmentasi Roll Time Mixing-Hermit Frequency Mixing dan Arsitektur Hybrid CNN-Multihead Transformer.
Irfan Fadli Nugraha, Dr. Dyah Aruming Tyas, S.Si.
Universitas Gadjah Mada, 2025 | Diunduh dari http://etd.repository.ugm.ac.id/

ABSTRACT

Detection of Crackles and Wheezes in Respiratory Audio Using Roll Time Mixing–Hermit Frequency Mixing Augmentation and a Hybrid CNN-Multihead Transformer Architecture

By

Irfan Fadli Nugraha 21/482012/PA/21010

The challenge in audio-based respiratory disease detection lies in the limitations of data, the diversity of breathing sound patterns, and the presence of noise and class imbalance in medical datasets. Detecting abnormal sounds such as crackles and wheezes requires strong spectral representations and classification models capable of capturing both spatial and temporal characteristics of respiratory signals. These challenges become more complex in the context of implementing reliable automatic diagnostic systems for real clinical environments.

To address these issues, this study proposes a respiratory audio feature detection approach by combining Roll Time Mixing (RTM) and Hermit Frequency Mixing (HFM) augmentation techniques, aimed at enriching log-mel spectrogram representations through time and frequency domain manipulations. A Hybrid CNN-Multihead Transformer architecture is employed to extract spatial and temporal features simultaneously, where CNN captures local patterns and the transformer models long-term relationships across the time dimension.

The proposed model was evaluated using the ICBHI 2017 dataset and demonstrated improved performance compared to baseline models. Experimental results showed a balanced accuracy of 76.08%, sensitivity of 78.45%, and specificity of 73.70%. These results reflect an improvement of 3.36% in balanced accuracy and 12.14% in sensitivity compared to previous state-of-the-art models.