

## INTISARI

### ***CONVOLUTIONAL NEURAL NETWORK SEBAGAI FEATURE EXTRACTOR PADA KLASIFIKASI SUARA UNTUK PEMANTAUAN HUTAN DENGAN REPRESENTASI MEL-SPEKTOGRAM***

Oleh

Rizqi Fathin Fadhillah  
18/427506/PA/18466

Hutan memiliki peran yang penting bagi kehidupan di muka bumi. Kebutuhan untuk memonitori hutan terhadap adanya aktivitas illegal dan satwa-satwa yang ada disana perlu dilakukan untuk menjaga kondisi hutan agar tetap baik. Akan tetapi, kondisi hutan yang luas dan keterbatasan sumber daya membuat pemantauan kondisi hutan yang dilakukan secara langsung oleh petugas (manusia) menjadi terbatas. Dalam hal ini, suara dengan pemrosesan sinyal digital dapat dijadikan sebagai salah satu sarana untuk membantu mengetahui adanya tindakan illegal yang ada di hutan, seperti *illegal logging* dan perburuan liar. Selain itu, suara juga dapat dijadikan sebagai indikator untuk mengetahui suatu jenis satwa yang ada di hutan.

Proses pengolahan sinyal suara terdiri dari 5 tahapan. Tahap akuisisi data dilakukan 4 jenis augmentasi, yaitu *add noise*, *time stretch*, *time shift*, dan *pitch shift*. Pada tahap pra-pengolahan dilakukan konversi mel-spektogram dan perhitungan *snr*. Tahap ekstraksi fitur dilakukan dengan menggunakan *transfer learning* VGG-16 dan MobileNetV3. Tahapan terakhir adalah klasifikasi yang dilakukan dengan metode Random Forest, SVM, KNN, dan MLP. Model yang dihasilkan kemudian akan diimplementasikan pada Raspberry Pi 3B+.

Hasil penelitian ini menunjukkan bahwa metode klasifikasi *machine learning* dengan *transfer learning* arsitektur CNN sebagai *feature extractor* dapat diimplementasikan untuk mengklasifikasikan suara pada Raspberry Pi 3B+. Berdasarkan performa yang dihasilkan, didapatkan bahwa model MobileNetV3-Small + MLP dengan data latih *time stretch* + *time shift* memberikan performa yang paling bagus untuk diimplementasikan pada sistem ini, dengan durasi *inference* 0.8 detik, akurasi sebesar 93.96%, presisi sebesar 94.1%, dan penggunaan daya sebesar 2892 mW.

**Kata kunci:** Klasifikasi suara, Mel-spektogram, CNN, MLP

## ABSTRACT

### CONVOLUTIONAL NEURAL NETWORK AS FEATURE EXTRACTOR IN SOUND CLASSIFICATION FOR FOREST MONITORING WITH MEL- SPECTROGRAM REPRESENTATION

by

Rizqi Fathin Fadhillah  
18/427506/PA/18466

*Forest has an important role on earth. The need to monitor forest from illegal activities and the animals in there is needed to keep the forest in good condition. However, the condition of the vast forest and limited resource make direct forest monitoring by officer (human) is limited. In this case, Sound with digital signal processing can be used as a mean to detect illegal actions in the forest, such as illegal logging and poaching. In addition, sound can also be used as an indicator to find out a type of animal in the forest.*

*The audio signal processing consists of 5 stages. In the data acquisition stage, 4 types of augmentation were carried out, namely add noise, time stretch, time shift, and pitch shift. In the pre-processing stage, mel-spectrogram conversion and snr calculations are carried out. The feature extraction stage was carried out using transfer learning VGG-16 and MobileNetV3. The final stage is classification which is carried out using the Random Forest, SVM, KNN, and MLP methods. The resulting model will then be implemented on the Raspberry Pi 3B+.*

*The results of this study indicate that the machine learning method with transfer learning CNN can be implemented to classify sounds on the Raspberry Pi 3B+. Based on the model performance, it was found that the MobileNetV3-Small + MLP model with combined training data from time stretch and time shift augmentation provide the best performance to be implemented in this system, with an inference duration of 0.8 seconds; 93.96% accuracy; and 94.1% precision.*

**Keywords:** *Sound classification, Mel-spectrogram, CNN, MLP*