

Identifikasi Suara sebagai Biometric dalam Berbagai Sampling Frekuensi = Voice Identification as Biometric in Various Sampling Frequency

Arifan Rahman, author

Deskripsi Lengkap: <https://lib.ui.ac.id/detail?id=20517900&lokasi=lokal>

Abstrak

Suara merupakan suatu hal yang penting pada diri manusia, dengan suara manusia bisa berkomunikasi dengan mudah. Selain itu, melalui suara kita bisa mengenali orang tersebut tanpa perlu melihat wajahnya. Hal itu disebabkan karena setiap suara manusia memiliki karakteristiknya masing-masing. Akan tetapi jika melalui perantara, misalnya melalui rekaman atau melalui pesawat telepon, suara yang dihasilkan kadang berbeda dengan suara aslinya tanpa melalui perantara. Hal ini dikarenakan tiap-tiap alat tersebut memiliki sample rate frekuensi yang berbeda. Oleh karena itu, perlu dikembangkan suatu sistem untuk identifikasi suara manusia dalam berbagai variasi sampling frekuensi. Dataset yang digunakan pada penelitian ini merupakan data rekaman dari empat orang dengan sample rate 8000Hz, 16000Hz, 32000Hz dan 48000Hz. Untuk melakukan identifikasi data tersebut, dibutuhkan metode ekstraksi fitur dan klasifikasi yang tepat untuk berbagai jenis data tersebut. Pada penelitian ini dilakukan penggabungan tiga fitur untuk identifikasi suara manusia dalam berbagai jenis sampling frekuensi yaitu fitur akustik MFCC, frekuensi formant, dan fitur prosodik dinamis. Proses klasifikasi dari fitur-fitur tersebut dan penggabungan fitur menggunakan Deep Neural Network (DNN). Hasil eksperimen menunjukkan bahwa penggabungan ketiga fitur dengan Deep Neural Network memperoleh f1-score yang lebih baik dibandingkan penggabungan fitur dengan menggunakan klasifikasi Support Vector Machine (SVM). Nilai f1-score yang didapatkan pada penggabungan fitur dengan menggunakan DNN pada data suara berdurasi 3 detik, 5 detik, dan 10 detik berturut-turut adalah 97.89%, 98.51%, 99.83%.

.....Voice is an essential thing in humans; with voice, humans can communicate easily. In addition, through voice, we can recognize the person without seeing his/her face. This is because every human voice has its characteristics. However, if through an intermediary, for example, through a recording device or a telephone, sometimes the sound or voice that is produced is different from the original sound without going through an intermediary. This is because each of these devices has a different sample rate of frequency. Therefore, it is necessary to develop a system for identifying the human voice in a variety of sampling frequencies. The dataset that is used in this study is recorded data from four people with sample rates of 8000Hz, 16000Hz, 32000Hz, and 48000Hz. To identify the data, the appropriate feature extraction and classification methods are needed for these various types of data. In this study, three features will be combined to identify the human voice in various types of frequency sampling, namely MFCC acoustic features, formant frequencies, and dynamic prosodic features. The process of classifying these features and combining features using Deep Neural Network (DNN). The experimental results show that combining the three features with Deep Neural Network obtain a better f1-score compared with Support Vector Machine (SVM). The f1-score value obtained by combining those features with DNN on sound data with 3 seconds, 5 seconds, and 10 seconds duration are 97.89%, 98.51%, and 99.83%, respectively.