

Implementasi dan Analisis Sistem Pendeteksi Keaslian Suara Manusia Berbasis Convolutional Neural Networks (CNN) = Implementation and Analysis of Human Voice Authenticity Detection Based on Convolutional Neural Networks (CNN)

Syamsul Erisandy Arief, author

Deskripsi Lengkap: <https://lib.ui.ac.id/detail?id=9999920544067&lokasi=lokal>

Abstrak

Hadirnya beragam layanan penyintesis suara manusia di Internet memungkinkan siapa pun untuk melakukan sintesis suara manusia dengan memanfaatkan layanan ini. Di tangan yang salah, teknologi ini dapat merugikan masyarakat awam dan meningkatkan peluang keberhasilan penipuan. Maraknya layanan penyintesis suara manusia yang sudah hampir tidak dapat dibedakan oleh telinga manusia memberikan keluangan untuk menghadirkan sebuah sistem yang dapat membedakan suara manusia dengan suara manusia sintetis. Penelitian ini memanfaatkan teknologi pembelajaran mesin yang berupa *Convolutional Neural Networks* pada spektrogram suara manusia dari himpunan data pelatihan dengan 16 suara manusia yang berisikan 4 suara pria asli, 4 suara pria sintetis, 4 suara wanita asli, dan 4 suara wanita sintetis dengan jumlah 1.008 berkas rekaman suara manusia berformat WAV yang telah dirancang dan dibuat khusus untuk penelitian ini dengan pembagian pelatihan dan validasi sebesar 80% dan 20% secara berurut. Hasil akhir dari penelitian ini memberikan sebuah model CNN dengan bobotnya yang memberikan nilai *data loss* sekecil 0,00022 dan sebuah sistem yang dapat melakukan deteksi keaslian suara manusia berdasarkan berkas rekaman suara manusia dan model CNN serta bobot yang diberikan.

.....The presence of human voice synthesis services on the Internet allows everyone to create synthetic human voices by leveraging these services. In the wrong hands, this technology could harm unsuspecting citizens and promote chances of scams. The abundance of human voice synthesis service that is almost indistinguishable by human ears gave presence to a system that could distinguish between real and synthetic human voices. This study leverages machine learning technology in the form of Convolutional Neural Networks on a spectrogram from a training dataset with 16 different human voices consisting 4 authentic men voices, 4 synthetic men voices, 4 authentic women voices, and 4 synthetic women voices with the total of 1,008 WAV formatted human voice recording files that was designed and made specifically for this study with the splitting ratio for training and validation set to 80% and 20% respectively. The end result of this study produces a CNN model and its weights with a data loss score of 0.00022, as well as a system that can perform authenticity detection on a human voice based on the given human voice recording file and the CNN model with its weights.