

# **PENGENALAN SUARA PADA SISTEM NOTULEN RAPAT MENGUNAKAN *CONVOLUTIONAL NEURAL NETWORK* (CNN)**

## **ABSTRAK**

Notulen rapat merupakan salah satu hal yang dapat dikatakan cukup penting untuk pendokumentasian pada sebuah kelompok atau instansi. Biasanya pendokumentasian sebuah notulen dapat ditangani oleh notulis. Namun seringkali terdapat kesalahpahaman saat notulis menuliskan hasil rapat. Salah satu format penulisan notulen rapat adalah menuliskan siapa yang berbicara saat rapat berlangsung. Salah satu solusi untuk menangani permasalahan tersebut adalah dengan melakukan pengenalan suara terhadap peserta rapat. Pengenalan suara ini dapat diimplementasikan menggunakan metode-metode tentang pengolahan suara. Penelitian ini mengimplementasikan salah satu model dari *Deep Learning* yaitu *Convolutional Neural Network* (CNN). Penelitian ini ditujukan untuk menganalisis dan mengenali kecocokan pada setiap pola dari hasil *spectrogram image* yang berupa sinyal suara berdasarkan *frekuensi* dan waktu. Pola sinyal suara seseorang tentunya berbeda-beda bergantung pada setiap karakteristik dari suara masing-masing. Karakteristik suara manusia misalnya suara keras, suara lembut, suara serak, dll. Untuk mencapai proses CNN tersebut, sebelumnya akan dilakukan pra-proses pada data suara. Data suara dengan format .wav akan disamakan durasinya terlebih dahulu kemudian dikonversi kedalam bentuk *spectrogram image*. Setelah diperoleh data *spectrogram image*, kemudian akan dilakukan *scaling* serta *thresholding*. Proses perhitungan akurasi pada setiap data dilakukan menggunakan *confusion matrix*. Hasil eksperimen terhadap 120 data rekaman suara dengan format .wav yang dikumpulkan menunjukkan bahwa akurasi tertinggi yang dihasilkan berdasarkan *voice recognition* menggunakan model CNN adalah 92,5% dengan rata-rata akurasi adalah 80,33%.

**Kata Kunci:** *Voice Recognition, Spectrogram Image, Deep Learning, Convolutional Neural Network.*

## ***SOUND RECOGNITION SYSTEM USING THE PROCEEDINGS CONVOLUTIONAL NEURAL NETWORK (CNN) IN MEETING MINUTES***

### ***ABSTRACT***

*Meeting minutes are the most important thing for documentation in a group or agency. Usually documenting minutes is being handled by a court reporter. however, there will frequently a miscomprehension when the court reporter writes the result of the meeting. One of the writing format of a meeting minutes is by noting who is speaking during the ongoing meeting. One of the solution to solve this problem is by having a voice recognition system for each participant of the meeting. This voice recognition system can be implemented by using voice recognition methods, this research implemented one of the models of beep learning called Convolutional Neural Network. This research is dedicated to analyze and recognize the matches for each patterns from the spectrogram image's results in the form of voice signal based on frequency and time. Voice pattern of every person is definitely different depending on each characteristics of the voice. for example the characteristics such as loud, soft, hoarse and many more. to achieve the CNN process a preprocessing of voice data must be done before. Voice data with a format of .wap will have an equated duration first and then converted into a form of spectrogram image. After a spectrogram image data can be achieved, a scaling and thresholding will be done. The process of calculation for the accuracy of each data will be done by using confission matrix. The result of experiment to 120 voice records data with .wap format shows that the highest accuracy value produced using CNN speech recognition model is 92,5%. while the average value of the accuracy is 80,33%.*

***Keyword:*** *Voice Recognition, Spectrogram Image, Deep Learning, Convolutional Neural Network.*