

## **BAB V**

### **PENUTUP**

#### **5.1 Kesimpulan**

Berdasarkan uraian mengenai pengujian dan analisis pada bab sebelumnya, hasil dari proses penelitian dapat disimpulkan sebagai berikut.

1. Sistem pengenalan suara pada penelitian ini dirancang untuk melakukan klasifikasi terhadap sinyal audio dengan durasi maksimal satu detik. Sistem akan mengambil representasi dua dimensi dari sinyal audio (*spectrogram*) setiap satu detik, dan kemudian langsung melakukan klasifikasi terhadap data tersebut. Sistem pengenalan suara pada penelitian ini dapat bekerja secara *stand alone* pada sebuah komputer komersial tanpa harus menggunakan komputer dengan spesifikasi khusus. Ini dapat dicapai karena sistem pengenalan suara pada penelitian ini dibuat menggunakan jaringan *feed forward* dengan variasi *multi layer perceptrons* yang dirancang untuk melakukan *preprocessing* dalam jumlah minimal.
2. Jenis algoritma *deep learning* yang digunakan pada penelitian ini adalah *Convolutional Neural Network (CNN)* yang merupakan pengembangan *multi layer perceptron* yang didesain untuk mengolah data dua dimensi. Data dua dimensi yang diproses oleh CNN pada penelitian ini adalah bentuk *spectrogram* dari *input* sinyal audio. Arsitektur CNN pada sistem ini terdiri dari 24 *layer* yang saling terintegrasi satu sama lain untuk melakukan proses *feature learning* dan *classification* terhadap *input spectrogram* yang diberikan.

3. Sistem pengenalan suara menggunakan DL-CNN pada penelitian ini dirancang untuk dapat diaplikasikan pada *home automation*. Perintah-perintah yang dapat dikenali oleh sistem pengenalan suara pada sistem ini adalah sebagai berikut.

- Nyalakan Lampu
- Matikan Lampu
- Kunci Pintu
- Buka Pintu
- Kipas Mati
- Kipas Menyala

Sistem pengenalan suara pada penelitian ini memiliki persentase keberhasilan sebesar 100% pada kondisi ruangan dengan intensitas *background noise* 24dB (senyap), sebesar 67,67% pada kondisi ruangan dengan intensitas *background noise* 42dB, dan sebesar 51,67% pada kondisi ruangan dengan intensitas *background noise* 52dB (bising). Dari hasil tersebut dapat disimpulkan bahwa semakin senyap kondisi suatu ruangan, maka akan semakin baik pula akurasi dari sistem pengenalan suara pada penelitian ini. Sebaliknya semakin bising kondisi suatu ruangan, maka akurasi dari sistem pengenalan suara akan semakin menurun. Oleh karena itu, untuk memperoleh hasil yang optimal, sistem ini lebih cocok digunakan pada ruangan dengan intensitas *background noise* yang rendah.

## 5.2 Saran

Dengan tujuan untuk perbaikan dan pengembangan penelitian ini lebih lanjut, penulis mengusulkan beberapa saran diantaranya:

1. Untuk penelitian selanjutnya, sistem pengenalan suara pada penelitian ini perlu dikembangkan lagi pada segi akurasi. Sehingga dengan demikian diharapkan bahwa, sistem tidak hanya memiliki akurasi yang baik pada kondisi ruangan yang senyap, namun juga memiliki akurasi yang baik pada kondisi ruangan yang bising.
2. Data *training* pada sistem ini perlu diperbanyak dengan data suara dari lebih banyak orang, sehingga nantinya sistem diharapkan dapat mendeteksi perintah suara dari berbagai jenis aksent atau pengucapan yang berbeda-beda.
3. Untuk saat ini, sistem ini hanya mampu memproses perintah suara dengan durasi pengucapan maksimal 1 detik. Pengembangan selanjutnya bisa dilakukan dengan menambah kemampuan sistem dalam mendeteksi perintah suara dengan durasi pengucapan yang lebih lama. Sehingga dengan demikian, perintah suara yang bisa dikenali oleh sistem ini bisa lebih beragam.
4. Arsitektur CNN yang digunakan pada penelitian ini hanya terdiri dari 5 layer konvolusi. Pengembangan selanjutnya dapat dilakukan dengan menambah layer konvolusinya sehingga sistem menjadi semakin akurat dan detail dalam memproses suatu *input* data.