

BAB V KESIMPULAN DAN SARAN

Bab ini dibagi menjadi dua bagian yaitu kesimpulan hasil penelitian dan saran yang diberikan untuk penelitian lebih lanjut.

5.1 Kesimpulan

Berikut merupakan kesimpulan yang dapat diambil untuk menjawab rumusan masalah yang ada pada penelitian ini.

1. Sistem pengenalan suara menggunakan metode *deep learning* model *Convolutional Neural Network* berhasil dilakukan. Adapun caranya adalah melakukan praproses audio yaitu penghapusan durasi suara gangguan (*noise*) dan durasi suara diam (*noiseless*). Konversi dan pelabelan yaitu dari data .mp3 menjadi .wav agar menjadikan data lebih detail, dan dari .wav menjadi *spectrogram image* untuk sebagai teknik yang dipergunakan dalam mengenali suara. Data tersebut selanjutnya diberi label untuk merepresentasikan perbedaan setiap data suara dan sebagai indikator dalam pengujian yang dilakukan. Dan praproses citra yaitu *scalling* 28x28 piksel, *grayscale* dan *thresholding*. Selanjutnya diimplementasikan metode *deep learning* model *Convolutional Neural Network* (CNN) menggunakan *Deep Learning Toolbox* untuk mempelajari data *input*, mengklasifikasikannya dan menghitung tingkat akurasi hasil klasifikasinya. Pada *Toolbox* tersebut menerapkan empat *hidden layer* CNN, yaitu dua *layer* konvolusi dan dua *layer subsampling*. Tahap pertama menentukan *output map* pada *convolution layer*. *Output map* adalah jumlah *neuron* di tiap *convolution layer*. Jumlah *output map* pada *layer* konvolusi yang pertama adalah 6 dan kedua 12. Langkah selanjutnya menentukan *kernel* yang digunakan pada 2 *layer* konvolusi dengan ukuran 5x5 piksel. Dan pada *subsampling layer* digunakan proses *max pooling* dengan skala 2x2.

2. Hasil metode *deep learning* model *Convolutional Neural Network* (CNN) yang implementasikan dapat mengenali data suara dengan tingkat rata-rata akurasi keberhasilan adalah 27.32% pada skenario 1, 40,59% pada skenario 2, 16,66% pada skenario 3, dan 43,70% pada skenario 4. Pada teknik pengucapan yang diterapkan, pengenalan teknik pengucapan secara normal menghasilkan tingkat pengenalan lebih tinggi disebabkan pengucapan secara normal menghasilkan data suara lebih banyak mendekati kesamaan. Sedangkan pengujian per suaranya, pada skenario 1 dan 2 menunjukkan suara pengucapan kata ‘tiga’ menghasilkan pengenalan tertinggi karena data suaranya lebih banyak mendekati kesamaan, dan kata ‘tujuh’ menghasilkan pengenalan terendah karena data suaranya lebih sedikit mendekati kesamaan. Adapun dari seluruh skenario yang telah dilakukan, pengenalan 10 kelas data suara menunjukkan adanya trend peningkatan rata-rata kemampuan pengenalan seiring perubahan skenario dengan peningkatan *epoch* yang digunakan. Pengujian 1 pada skenario 1 dan 2 membuktikan bahwa jika pengenalan menggunakan data yang sama, maka akurasi yang didapatkan lebih tinggi daripada menggunakan data suara yang berbeda seperti pengujian 2, yaitu akurasi keberhasilan pengenalan cenderung turun karena data yang berbeda dan kurang banyak jika dibandingkan pada bagian pengujian selanjutnya. Hasil tersebut juga menunjukkan penambahan nilai *epoch* dapat meningkatkan pengenalan. Hal ini dikarenakan semakin tinggi *epoch* membuat pengetahuan tentang data akan lebih baik secara detail karena lebih sensitif terhadap perubahan data tertentu, tetapi juga berarti akan meningkatkan kebutuhan memori dan komputasi. Dengan demikian, berdasarkan penelitian ini dapat dipastikan bahwa tingkat akurasi pengenalan adalah dipengaruhi oleh kualitas data, jumlah data dan jumlah *epoch* dan *batch* yang digunakan.

5.2 Saran

Berdasarkan kesimpulan tersebut diatas, disarankan melakukan hal-hal berikut pada penelitian selanjutnya:

- (1) Menambahkan jumlah data untuk proses *training* agar pengenalan suara dapat lebih bagus. Hal ini berdasarkan hasil penelitian skenario 1 dan 2 pada pengujian 2 yaitu akurasi pengenalan menjadi turun karena data yang berbeda, dan pengujian 3 pengenalan turun karena jumlah data yang kurang banyak jika dibandingkan pengujian selanjutnya.
- (2) Menambahkan atau menggunakan cara lain melakukan praproses. Pada penelitian ini praproses yang dilakukan hanya dapat menangani sebagian *noise*, oleh sebab itu praproses yang lebih detail menjadi bagian sangat penting untuk meminimalkan gangguan dan memperjelas data sehingga mendapatkan hasil pengenalan yang bagus.
- (3) Penelitian pengenalan suara ini terbatas pada data dan hanya dapat menerima *input* suara yang telah disiapkan sebelumnya, untuk penelitian selanjutnya dapat dikembangkan hingga ke pengenalan suku kata dan kalimat serta juga mengenali suara secara langsung sehingga lebih mempermudah kegunaanya dalam mengenali suara.