

BAB V PENUTUP

5.1 Kesimpulan

Setelah keseluruhan implementasi dijalankan, dapat disimpulkan bahwa dengan menggunakan algoritma CNN yang mengombinasikan perbedaan dimensi dari hasil ekstraksi fitur dan juga model CNN itu sendiri yakni 1 dan 2 dimensi pada 3 *dataset* yang digunakan pada penelitian ini yaitu RAVDESS, ASVP-ESD, TESS dapat dibuktikan bahwa implementasi dapat dijalankan. Selain itu, berdasarkan rumusan masalah dan tujuan dari penelitian ini maka dapat diuraikan sebagai berikut.

1. Implementasi dari Algoritma CNN pada kasus *Speech Emotion Recognition* dengan kombinasi 1 dan 2 dimensi dapat berjalan dengan baik mulai dari persiapan data, *pre-processing*, *augmentation*, *segmentation* hingga ke pelatihan model seperti yang telah dijelaskan pada BAB IV.
2. Model kombinasi yang sudah dilatih oleh peneliti dapat melakukan prediksi dengan rata-rata akurasi hingga 95% pada *train set*, 80% pada *validation set*, dan pada saat evaluasi mendapat hasil akurasi sebesar 75%.
3. Pada Tabel 4. 2 terlihat bahwa dengan menggunakan kombinasi CNN 1 dan 2 dimensi meningkatkan kecepatan proses pelatihan data dibandingkan dengan pelatihan pada masing-masing model CNN secara terpisah.

5.2 Saran

Pada penelitian ini, peneliti sadar bahwa masih ada banyak sekali kekurangan yang dilakukan. Maka dari itu saran untuk penelitian-penelitian selanjutnya adalah sebagai berikut.

1. Menggunakan *dataset* yang lebih banyak lagi agar model dapat mengenali emosi yang ada pada suara dengan lebih baik dan juga memperkecil kemungkinan terjadinya *overfitting* karena ditemukannya indikasi *overfitting* pada model yang telah dilatih dengan akurasi dari *validation* dan *test* yang tidak meningkat pada saat akurasi dari *training* meningkat.
2. Menyeimbangkan data yang akan diproses berdasarkan label.
3. Menggunakan metode normalisasi agar data dapat lebih terukur pada rentang data.
4. Menggunakan metode fitur ekstraksi yang lebih lagi pada ekstraksi fitur terutama pada fitur 2 dimensi.