

BAB V

KESIMPULAN DAN SARAN

5.1 Kesimpulan

Setelah analisis dan evaluasi dilakukan terhadap hasil eksperimen dan sistem yang dikembangkan serta dengan merujuk pada rumusan masalah dan tujuan yang ingin dicapai dalam penelitian ini, dapat disimpulkan bahwa:

1. *Cost DTW* yang berasal dari pencocokan data uji terhadap *template* 1, 5, dan 10 dari model yang sama memiliki nilai yang hampir stabil, ini dapat membuktikan bahwa orang yang sama memiliki kecenderungan untuk melafalkan ayat dengan variasi yang lebih konsisten.
2. *Dynamic Time Warping* sensitif terhadap suara yang berasal dari orang yang berbeda, hal ini dibuktikan dari hasil eksperimen kedua yaitu eksperimen *cross speaker*, namun pada beberapa kasus pencocokkan, akurasi yang dihasilkan masih tinggi karena belum adanya *threshold*.
3. MFCC dapat melakukan ekstraksi fitur pada data suara dengan durasi yang tidak terbatas, namun semakin panjang durasi data suara maka *frame* yang dihasilkan akan semakin banyak. Hal ini dapat berpengaruh pada proses komputasi DTW di mana durasi pengucapan berbanding lurus dengan waktu komputasi. Jika data suara semakin panjang durasinya, maka *frame* yang dihasilkan akan semakin banyak dan waktu komputasi DTW akan semakin lama. Sehingga untuk pengembangan sistem pengenalan ucapan dengan durasi ucapan yang lebih panjang, sebaiknya tidak menggunakan algoritma DTW.

5.2 Saran

Setelah analisis dan evaluasi terhadap hasil eksperimen dan sistem yang dikembangkan, terdapat beberapa saran yang dapat dipertimbangkan untuk penelitian selanjutnya, yaitu:

1. Dalam menghafal Al-Quran menggunakan kata awal lebih baik menggunakan merekam suara sendiri untuk menghasilkan pencocokan yang lebih akurat.

2. Pada penelitian ini, *threshold* atau nilai ambang untuk suatu kata dapat dikenali sebagai ayat yang benar belum dapat dilakukan. Untuk penelitian lanjutan dapat menambahkan *threshold* sebagai rumusan masalah dan untuk mengembangkan program pengenalan ucapan ini menjadi lebih baik.
3. Menambahkan variasi kata awal yang lebih banyak untuk mendapatkan nilai *threshold*.
4. Pada setiap perekaman suara yang dilakukan, selalu terdapat sedikit jeda diam selama beberapa milisekon pada bagian awal pengucapan dan akhir pengucapan, hal ini akan mempengaruhi keakuratan pada proses pengenalan ucapan sehingga diperlukan proses penghilangan bagian yang diam atau *silent* dari data suara. Dalam penelitian ini proses penghilangan tersebut masih dilakukan secara manual menggunakan aplikasi penyunting suara sehingga cukup memakan banyak waktu dan sangat tidak efisien. Untuk penelitian lanjutan diharapkan dapat menambah fitur untuk proses penghilangan bagian *silent* dari data suara dalam sistem pengenalan ucapan agar dapat dilakukan secara otomatis.

