BAB V

PENUTUP

5.1 Kesimpulan

Berdasarkan hasil perancangan, implementasi, pengujian, dan pembahasan sistem deteksi kantuk berbasis multimodal, diperoleh beberapa kesimpulan sebagai berikut.

1. Sistem deteksi kantuk berbasis multimodal berhasil dikembangkan dengan menggabungkan informasi visual dan audio. Fitur visual diperoleh dari nilai *Eye Aspect Ratio* (EAR) dan *Mouth Aspect Ratio* (MAR) menggunakan *MediaPipe BlazeFace dan FaceMesh*, sedangkan fitur audio diperoleh melalui ekstraksi *Mel-Frequency Cepstral Coefficients* (MFCC). Keluaran kedua modalitas digabungkan menggunakan logika fuzzy untuk menentukan kondisi pengemudi menjadi Terjaga, Lelah, atau Mengantuk.
2. Model ANN visual pada skenario awal pembagian data 80% latih dan 20% validasi memperoleh nilai *accuracy* sebesar 91%. Setelah dilakukan pembagian data menjadi 80% latih, 10% validasi, dan 10% uji berdasarkan kelompok video, hasil evaluasi pada data uji internal menunjukkan *accuracy* sekitar 99% untuk klasifikasi kondisi mata dan mulut. Pengujian menggunakan data eksternal dari 10 responden juga menghasilkan *accuracy* sekitar 99%, sehingga menunjukkan bahwa performa model visual dapat dipertahankan pada data yang berasal dari subjek di luar dataset pengembangan.
3. Model ANN audio menggunakan *threshold* 0,33 yang ditentukan dari data validasi. Pada pengujian internal memperoleh *accuracy* sebesar 88% pada 221 data uji. Model berhasil mengklasifikasikan 195 sampel dengan benar, terdiri atas 104 sampel bukan menguap dan 91 sampel menguap. Pada pengujian eksternal terhadap 200 sampel dari 10 responden model menghasilkan *accuracy* sebesar 90%. Hasil tersebut menunjukkan bahwa model audio tetap mampu membedakan suara menguap dan bukan menguap pada data eksternal yang tidak digunakan dalam proses pelatihan maupun penentuan *threshold*.
4. Evaluasi sistem logika fuzzy terhadap 27 kombinasi aturan menghasilkan tingkat kesesuaian fungsional sebesar 100%. Hasil ini menunjukkan bahwa proses fuzzifikasi, inferensi, dan defuzzifikasi telah berjalan sesuai dengan basis aturan yang dirancang.
5. Validasi sistem multimodal pada 20 skenario menghasilkan 18 keputusan yang sesuai dengan kondisi aktual, sehingga tingkat kesesuaian tegas mencapai 90%. Dua kesalahan yang terjadi hanya berbeda satu tingkat kelas, sehingga skor

kesesuaian bertingkat mencapai 95%. Tidak ditemukan kesalahan langsung antara kondisi Terjaga dan Mengantuk.

6. Pemrosesan visual dan fuzzy membutuhkan waktu rata-rata 51,77 ms per frame, sedangkan pemrosesan audio membutuhkan 86,86 ms. Total waktu pemrosesan sistem multimodal adalah 135,09 ms atau sekitar 0,135 detik per siklus. Hasil tersebut menunjukkan bahwa sistem dapat dijalankan pada lingkungan simulasi, meskipun optimasi masih diperlukan terutama pada ekstraksi fitur audio serta inferensi fuzzy dan defuzzifikasi.

5.2 Saran

Berdasarkan hasil penelitian yang telah dilakukan, terdapat beberapa saran yang dapat digunakan untuk pengembangan penelitian selanjutnya, yaitu sebagai berikut.

1. Pengujian sistem dapat dilakukan pada kondisi berkendara nyata atau semi-nyata dengan jumlah responden dan variasi kondisi yang lebih beragam. Hal ini diperlukan untuk mengevaluasi kinerja sistem terhadap perubahan pencahayaan, posisi wajah, karakteristik pengguna, serta kebisingan lingkungan yang belum sepenuhnya terwakili pada pengujian simulasi.
2. Dataset multimodal dapat dikembangkan dengan merekam data visual dan audio secara bersamaan pada subjek dan waktu yang sama. Data yang tersinkronisasi diharapkan dapat memberikan representasi hubungan kondisi mata, mulut, dan suara menguap yang lebih baik dalam proses penggabungan informasi dan pengambilan keputusan akhir.
3. Pemrosesan sistem, terutama pada jalur audio dan logika fuzzy, dapat dioptimalkan untuk meningkatkan kecepatan dan kestabilan keputusan. Pengembangan selanjutnya dapat difokuskan pada efisiensi ekstraksi fitur dan inferensi ANN audio serta penyempurnaan parameter dan aturan fuzzy berdasarkan data pengujian yang lebih luas.