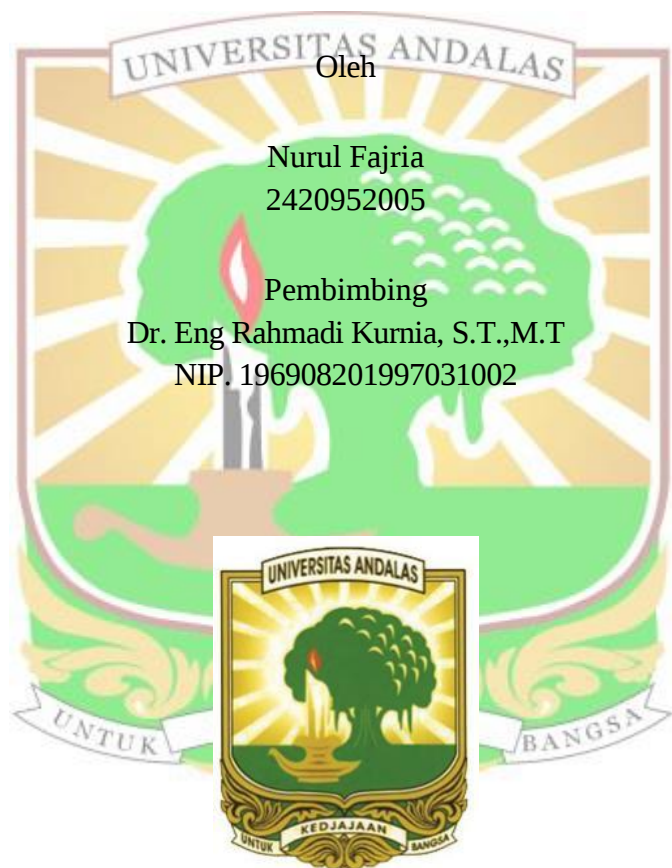


**PENGEMBANGAN DETEKSI KANTUK MULTIMODAL BERBASIS
LANDMARK WAJAH DAN SUARA MENGUAP DENGAN KLASIFIKASI
MENGUNAKAN ARTIFICIAL NEURAL NETWORK DAN
PENGAMBILAN KEPUTUSAN BERBASIS LOGIKA FUZZY**

TESIS

Karya Ilmiah sebagai salah satu syarat untuk menyelesaikan jenjang strata dua
(S-2) di Departemen Teknik Elektro, Fakultas Teknik, Universitas Andalas



**Program Studi Magister
Teknik Elektro Fakultas Teknik
Universitas Andalas
2026**

Judul	Pengembangan Deteksi Kantuk Multimodal Berbasis Landmark Wajah dan Suara Menguap dengan Klasifikasi Menggunakan Artificial Neural Network dan Pengambilan Keputusan Berbasis Logika Fuzzy	Nurul Fajria
Program Studi	Teknik Elektro	2420952005
Fakultas Teknik Universitas Andalas		
ABSTRAK Kantuk pada pengemudi dapat menurunkan kewaspadaan, memperlambat respons, dan meningkatkan risiko kecelakaan lalu lintas. Sistem deteksi satu modalitas memiliki keterbatasan karena informasi visual dapat dipengaruhi pencahayaan, posisi wajah, dan aktivitas mulut, sedangkan informasi audio dapat dipengaruhi kebisingan lingkungan. Penelitian ini mengembangkan sistem deteksi kantuk berbasis multimodal dengan menggabungkan informasi visual dari landmark wajah dan informasi audio dari suara menguap untuk menentukan kondisi pengemudi menjadi Terjaga, Lelah, atau Mengantuk. Fitur visual diperoleh menggunakan Eye Aspect Ratio (EAR) dan Mouth Aspect Ratio (MAR), sedangkan fitur audio diekstraksi menggunakan Mel-Frequency Cepstral Coefficients (MFCC). Fitur visual dan audio diklasifikasikan terpisah menggunakan model Artificial Neural Network (ANN). Keluaran ANN visual dan ANN audio digabungkan menggunakan logika fuzzy dengan 27 aturan. ANN visual dengan pembagian 80% latih, 10% validasi, dan 10% uji menghasilkan akurasi 99,92% pada kondisi mata dan 99,97% pada kondisi mulut. Pengujian eksternal pada 10 responden menghasilkan akurasi 99,90% pada mata dan 99,95% pada mulut. Model ANN audio dengan threshold 0,33 memperoleh akurasi 88,24% pada data uji internal, sedangkan pengujian eksternal memperoleh akurasi 92%. Evaluasi fungsional logika fuzzy terhadap 27 kombinasi aturan menghasilkan tingkat kesesuaian 100%. Validasi sistem multimodal pada 20 skenario menghasilkan 18 keputusan yang sesuai dengan kondisi aktual, sehingga tingkat kesesuaian tegas mencapai 90% dan skor kesesuaian bertingkat mencapai 95%. Sistem membutuhkan waktu pemrosesan 51,77 ms untuk proses visual dan fuzzy, 86,86 ms untuk proses audio, serta 135,09 ms untuk keseluruhan proses multimodal. Hasil menunjukkan bahwa sistem mampu menggabungkan kondisi mata, mulut, dan suara menguap secara bertahap dalam menentukan tingkat kewaspadaan pengemudi. Pendekatan ini berpotensi dikembangkan sebagai sistem peringatan dini noninvasif berbasis kamera dan mikrofon standar. Kata Kunci: Deteksi kantuk multimodal, Landmark wajah, Suara menguap, ANN, Logika fuzzy.		

Title	<i>Development of a Multimodal Drowsiness Detection System Based on Facial Landmarks and Yawning Sounds Using Artificial Neural Network Classification and Fuzzy Logic-Based Decision Making</i>	Nurul Fajria
Major	Electrical Engineering Department	2420952005
Engineering Faculty Andalas University		
ABSTRACT <i>Driver drowsiness can reduce alertness, slow reaction time, and increase the risk of traffic accidents. Single-modality detection systems have limitations because visual information can be affected by lighting conditions, facial position, and mouth movements, while audio information can be influenced by environmental noise. This study develops a multimodal drowsiness detection system by combining visual information from facial landmarks and audio information from yawning sounds to classify the driver's condition into three states: Alert, Fatigued, and Drowsy. Visual features are obtained using the Eye Aspect Ratio (EAR) and Mouth Aspect Ratio (MAR), while audio features are extracted using Mel-Frequency Cepstral Coefficients (MFCC). The visual and audio features are classified separately using Artificial Neural Network (ANN) models. The outputs of the visual ANN and audio ANN are then combined using fuzzy logic with 27 rules. The visual ANN, using an 80% training, 10% validation, and 10% testing split, achieved accuracies of 99.92% for eye-condition classification and 99.97% for mouth-condition classification. External testing involving 10 respondents achieved accuracies of 99.90% for eye-condition classification and 99.95% for mouth-condition classification. The audio ANN model, using a threshold of 0.33, achieved an accuracy of 88.24% on the internal test data, while external testing achieved an accuracy of 92%. Functional evaluation of the fuzzy logic system across all 27 rule combinations achieved a 100% agreement rate. Validation of the multimodal system using 20 scenarios produced 18 decisions that matched the actual conditions, resulting in an exact agreement rate of 90% and a graded agreement score of 95%. The system required an average processing time of 51.77 ms for visual and fuzzy processing, 86.86 ms for audio processing, and 135.09 ms for the overall multimodal process. The results indicate that the system is capable of progressively integrating information from eye condition, mouth condition, and yawning sounds to determine the driver's level of alertness. This approach has the potential to be further developed as a non-invasive early warning system using standard cameras and microphones.</i> Keywords: multimodal drowsiness detection, facial landmarks, yawning sound, ANN, fuzzy logic.		