

ABSTRAK

ANALISIS PERBANDINGAN SISTEM PENGENALAN SUARA PADA ROBOT MENGGUNAKAN ALGORITMA TRADISIONAL DAN MODERN

(2026: 102 Halaman + 33 Gambar + 19 Tabel + Daftar Pustaka + Lampiran)

NABILA IMESA

062240342186

SARJANA TERAPAN TEKNIK ELEKTRO

TEKNIK ELEKTRO

POLITEKNIK NEGERI SRIWIJAYA

Sistem pengenalan suara merupakan salah satu teknologi yang dapat mendukung komunikasi suara antara manusia dan robot. Penelitian ini bertujuan membandingkan performa empat model pengenalan suara, yaitu *Hidden Markov Model* (HMM), *HMM-Gaussian Mixture Model* (HMM-GMM), *HMM-GMM-Deep Neural Network* (HMM-GMM-DNN), dan *Transformer Vanilla* pada dataset suara bahasa Indonesia. Dataset terdiri atas 1.250 rekaman audio dari 50 kalimat, yang meliputi 25 kalimat tanya dan 25 kalimat perintah. Evaluasi dilakukan menggunakan *Word Error Rate* (WER) dan *Character Error Rate* (CER). Hasil pelatihan menunjukkan bahwa *Transformer Vanilla* memperoleh WER sebesar 0,0222 dan CER sebesar 0,0164. Sementara itu, pada pengujian menggunakan audio di luar dataset, HMM-GMM-DNN memperoleh WER dan CER sebesar 0,00 pada kedua jenis kalimat. Hasil tersebut menunjukkan adanya perbedaan performa antara tahap pelatihan dan pengujian, sehingga pengujian menggunakan data baru diperlukan untuk mengetahui kemampuan generalisasi model.

Kata Kunci: *Sistem Pengenalan Suara, Bahasa Indonesia, Hidden Markov Model, Gaussian Mixture Model, Deep Neural Network, Transformer Vanilla, WER, CER.*

ABSTRACT

COMPARATIVE ANALYSIS OF SPEECH RECOGNITION SYSTEMS ON ROBOTS USING TRADITIONAL AND MODERN ALGORITHMS

(2026: 102 Pages + 33 Figures + 19 Tables + References + Appendices)

NABILA IMESA

062240342186

BACHELOR OF APPLIED ELECTRICAL ENGINEERING

ELECTRICAL ENGINEERING DEPARTMENT

STATE POLYTECHNIC OF SRIWIJAYA

Speech recognition systems are one of the technologies that can support voice communication between humans and robots. This study aims to compare the performance of four speech recognition models, namely Hidden Markov Model (HMM), HMM-Gaussian Mixture Model (HMM-GMM), HMM-GMM-Deep Neural Network (HMM-GMM-DNN), and Vanilla Transformer using an Indonesian speech dataset. The dataset consists of 1,250 audio recordings from 50 sentences, including 25 interrogative sentences and 25 command sentences. The evaluation was conducted using Word Error Rate (WER) and Character Error Rate (CER). The training results showed that the Vanilla Transformer achieved a WER of 0.0222 and a CER of 0.0164. Meanwhile, testing using audio outside the dataset showed that HMM-GMM-DNN achieved a WER and CER of 0.00 for both types of sentences. These results indicate a difference in model performance between the training and testing stages, indicating that testing with new data is necessary to evaluate the model's generalization ability.

Keywords: *Automatic Speech Recognition, Indonesian Language, Hidden Markov Model, Gaussian Mixture Model, Deep Neural Network, Vanilla Transformer, Word Error Rate, Character Error Rate.*