

BAB II

TINJAUAN PUSTAKA

Penelitian tentang pengenalan bahasa isyarat telah berkembang pesat sebagai akibat dari kemajuan teknologi di bidang *computer vision* dan *deep learning*. Abdullah *et al.* [2] membahas masalah utama dalam pembuatan sistem pengenalan bahasa isyarat. Ini termasuk keakuratan model, pemrosesan *gesture* dinamis, dan aplikasi *real-time*. Studi ini menekankan betapa pentingnya teknologi yang mendukung komunikasi bagi penyandang disabilitas. Seperti yang dibahas oleh Al-Qurishi *et al.* [3], teknik *deep learning* adalah salah satu metode yang paling banyak digunakan. Mereka menyoroti kemanjuran model LSTM dalam menangani data urutan waktu dan pentingnya keberagaman dataset untuk melatih model secara keseluruhan.

Pendekatan berbasis sensor juga digunakan untuk pengenalan *gesture*. Gu *et al.* [4] menggunakan *deep learning* untuk mengenali alfabet ASL secara akurat dengan menggunakan sistem *inertial motion capture*. Sementara itu, Haldera dan Tayade [5] menunjukkan metode yang dapat digunakan untuk mengenali bahasa isyarat lokal pada perangkat dengan spesifikasi rendah melalui sistem *real-time* berbasis MediaPipe. Dengan menggabungkan MediaPipe dan LSTM untuk mengenali alfabet ASL, Ba dan Bagyammal [6] memperkuat pendekatan berbasis *landmark* yang hasilnya menjanjikan.

Dalam hal teknis model di berbagai penelitian, seperti yang dilakukan oleh Sherstinsky [7] dan Bhadouria *et al.* [8], mempelajari struktur RNN, LSTM, dan BiLSTM, yang masing-masing memiliki kemampuan untuk memproses informasi temporal dengan baik. BiLSTM bahkan memiliki kemampuan untuk mempelajari konteks dua arah (maju dan mundur) dalam urutan data, yang membuatnya cocok untuk penggunaan *gesture* dinamis. MediaPipe yang dikembangkan oleh Lugaresi *et al.* [9], sangat penting untuk penelitian *gesture* karena dapat mendeteksi titik-titik *landmark* dengan cepat dan mudah.

MediaPipe digunakan bersama dengan model pembelajaran tambahan dalam sejumlah penelitian. Bora *et al.* [10] menggunakan MediaPipe dan *deep learning* untuk mengenali isyarat Assamese secara *real-time*. Sementara itu, Samaan *et al.* [11] mengintegrasikan *output landmark* MediaPipe dengan RNN untuk mengenali isyarat dinamis. Metode ini menunjukkan efektivitas *pipeline* yang bergantung pada fitur visual. Selain itu, Orovwode *et al.* [12] menekankan bahwa *preprocessing* dan rekayasa fitur sangat penting untuk sistem berbasis *machine learning* yang masih berkaitan untuk aplikasi di dunia nyata.

Metode *non-visual* yang ditawarkan dalam UltrasonicGS oleh Wang *et al.* [13] juga menarik sebagai alternatif, terutama dalam kondisi pencahayaan kurang. Abdulhamied *et al.* [14] membuktikan kemampuan LSTM untuk mengenali *gesture* secara langsung dengan mengembangkan sistem ASL *real-time* berbasis LSTM. Dari perspektif visual, Zhang *et al.* [15] menyarankan penggunaan CNN *dual-path* dengan penghapusan *background* untuk mengurangi gangguan visual dari latar belakang, dan Pathan *et al.* [16] menawarkan pendekatan multimodal yang menggabungkan gambar dan *landmark* menggunakan CNN *multi-head* untuk meningkatkan akurasi pengenalan.

Zhang *et al.* [17] mengembangkan pendekatan FFNN yang efektif bernama FNNS untuk pengolahan data skala besar dalam klasifikasi *gesture*. Kazbekova *et al.* [18] juga mengusulkan sistem *hybrid Deep CNN-BiLSTM* dengan Attention, yang menggabungkan kekuatan ekstraksi fitur spasial dan temporal dalam pengenalan *gesture* pada perangkat *edge*. Terakhir, Dabwan [19] menggunakan CNN dalam sistem penerjemah bahasa isyarat untuk menunjukkan bahwa metode konvolusi efektif dalam mengidentifikasi bentuk tangan pada gambar.

Tabel 2.1 di bawah ini merupakan tabel untuk membandingkan antara penelitian yang dikerjakan dengan penelitian serupa.

Tabel 2. 1 Perbandingan dengan Penelitian Serupa

Peneliti	Model	Dataset	Input	Akurasi	Kelas
S. Ba & B. Tb [6]	LSTM	ASL Alfabet	MediaPipe Hands	99%	26
G. H. Samaan <i>et al.</i> [11]	GRU, LSTM, BiLSTM	Custom Dataset	MediaPipe Hands, MediaPipe Holistic	100%, 99%, 99%	10
Abdulhamied <i>et al.</i> [14]	LSTM	ASL	MediaPipe Hands	99%	>20
G. Kazbekova <i>et al.</i> [18]	CNN + BiLSTM + Attention	Alfabet ASL	Frame gambar	99%	24
Penelitian ini	FNN, BiLSTM + Attention	Alfabet ASL + WLASL	MediaPipe Hands, MediaPipe Holistic	99%, 92%(cross- validation)	28 (ASL), 10 (WLASL)