

BAB 1. PENDAHULUAN

1.1 Latar Belakang

Komunikasi merupakan aspek penting dalam kehidupan manusia, digunakan untuk menyampaikan informasi, membangun hubungan sosial, serta memenuhi kebutuhan interaksi sehari-hari. Komunikasi bisa dilakukan secara verbal maupun nonverbal (Pohan & Fitria, 2021), bergantung pada kemampuan masing-masing individu.

Namun, bagi penyandang disabilitas tuna rungu dan tuna wicara, komunikasi verbal menjadi hambatan yang cukup signifikan. Mereka tidak mampu menangkap suara maupun mengeluarkan suara secara efektif, sehingga memerlukan bentuk komunikasi alternatif seperti bahasa isyarat (Nanda, 2022). Hambatan komunikasi ini dapat berdampak pada keterbatasan akses terhadap pendidikan, pekerjaan, dan layanan publik, serta menimbulkan ketimpangan sosial dalam masyarakat. Namun, bagi penyandang disabilitas tuna rungu dan tuna wicara, komunikasi verbal menjadi hambatan yang cukup signifikan. Mereka tidak mampu menangkap suara maupun mengeluarkan suara secara efektif, sehingga memerlukan bentuk komunikasi alternatif seperti bahasa isyarat (Nanda, 2022). Hambatan komunikasi ini dapat berdampak pada keterbatasan akses terhadap pendidikan, pekerjaan, dan layanan publik, serta menimbulkan ketimpangan sosial dalam masyarakat.

Salah satu sistem bahasa isyarat resmi yang digunakan di Indonesia adalah Sistem Isyarat Bahasa Indonesia (SIBI). SIBI dirancang mengikuti tata bahasa Indonesia dan digunakan dalam lingkungan pendidikan luar biasa untuk mempermudah proses belajar mengajar (Putra, 2025). Sistem ini terdiri atas huruf abjad manual serta berbagai bentuk isyarat untuk kata atau kalimat. Namun, dalam praktiknya, masih banyak masyarakat umum yang tidak memahami SIBI, sehingga penyandang disabilitas sering mengalami kesulitan berkomunikasi di luar lingkungan inklusif (Permana & Sutopo, 2023). Salah satu sistem bahasa isyarat resmi yang digunakan di Indonesia adalah Sistem Isyarat Bahasa Indonesia (SIBI). SIBI dirancang mengikuti tata bahasa

Indonesia dan digunakan dalam lingkungan pendidikan luar biasa untuk mempermudah proses belajar mengajar (Putra, 2025). Sistem ini terdiri atas huruf abjad manual serta berbagai bentuk isyarat untuk kata atau kalimat. Namun, dalam praktiknya, masih banyak masyarakat umum yang tidak memahami SIBI, sehingga penyandang disabilitas sering mengalami kesulitan berkomunikasi di luar lingkungan inklusif (Permana & Sutopo, 2023).

Seiring kemajuan teknologi, bidang *Computer Vision* dan *Deep Learning* mulai banyak digunakan dalam pengembangan sistem pengenalan bahasa isyarat secara otomatis. Salah satu pendekatan awal adalah menggunakan *Convolutional Neural Network* (CNN) yang bekerja dengan mengekstraksi fitur dari citra untuk mengenali bentuk tangan (Nurhayati dkk., 2022). Meskipun CNN efektif dalam mengenali citra statis, metode ini memiliki keterbatasan dalam mengenali gerakan atau pola dinamis, terutama pada huruf seperti “J” dan “Z” dalam SIBI yang mengharuskan sistem memahami urutan pergerakan tangan (*frame sequence*). Sayangnya, banyak penelitian sebelumnya hanya menangani huruf-huruf statis, sehingga belum mampu mengenali dua huruf tersebut secara akurat.

Hal tersebut menjadi dasar dalam penelitian ini untuk menambahkan mekanisme klasifikasi khusus terhadap huruf-huruf dinamis yang belum ditangani secara optimal oleh penelitian terdahulu.

Untuk mengatasi hal tersebut, digunakan metode deteksi objek *real-time* seperti YOLO (*You Only Look Once*). Algoritma ini mampu mengenali objek dalam satu tahap deteksi cepat dan efisien. YOLOv11, versi yang digunakan dalam penelitian ini, menghadirkan peningkatan signifikan dalam akurasi dan efisiensi dibanding versi sebelumnya seperti YOLOv5 dan YOLOv8 (Ultralytics, 2024). YOLO sangat baik dalam mengenali isyarat statis karena kecepatannya dalam memproses satu *frame* tanpa mengorbankan akurasi.

Namun, berdasarkan penelitian sebelumnya oleh Sholawati dkk. (2022), metode CNN yang digunakan untuk mengenali 26 huruf SIBI menghasilkan akurasi sebesar 80,76% dan gagal mengenali huruf dinamis seperti “J” dan “Z”. Begitu pula penelitian oleh Permana dan Sutopo (2023) yang menggunakan

YOLOv5 hanya mencapai akurasi 77% pada deteksi huruf secara *real-time*, namun tidak akurat pada huruf dinamis. Hal ini menunjukkan perlunya pendekatan tambahan yang bisa menangkap pola urutan gerak.

Berdasarkan hal tersebut, dapat disimpulkan bahwa masih terdapat celah dalam penelitian terdahulu, yaitu belum adanya sistem yang secara komprehensif menggabungkan deteksi huruf statis dan klasifikasi huruf dinamis seperti “J” dan “Z” dalam satu sistem terpadu dan *real-time*. Penelitian ini hadir untuk menutup celah tersebut dengan menambahkan mekanisme klasifikasi khusus terhadap huruf dinamis melalui model LSTM, sehingga mampu mengenali seluruh huruf dalam SIBI secara lengkap, baik statis maupun dinamis.

Oleh karena itu, penelitian ini menggabungkan dua metode yaitu YOLOv11 untuk mengenali huruf-huruf statis dan *Long Short-Term Memory* (LSTM) untuk menangani pengenalan gerakan huruf dinamis seperti “J” dan “Z”. LSTM adalah jenis *Recurrent Neural Network* (RNN) yang efektif dalam memproses data urutan waktu (*temporal sequence*) dan mengenali pola dalam pergerakan *frame* video (Altirika & Purnamasari, 2023). LSTM adalah jenis *Recurrent Neural Network* (RNN) yang efektif dalam memproses data urutan waktu (*temporal sequence*) dan mengenali pola dalam pergerakan *frame* video (Altirika & Purnamasari, 2023).

Penggabungan metode YOLOv11 dan LSTM memberikan keunggulan signifikan dalam membangun sistem deteksi huruf SIBI yang komprehensif dan efisien. YOLOv11 berperan dalam mendeteksi objek berupa tangan dan mengklasifikasikan huruf-huruf statis dengan cepat dan akurat secara *real-time*. Di sisi lain, LSTM digunakan untuk menangani huruf-huruf dinamis seperti “J” dan “Z”, dengan cara memproses urutan *frame* atau titik gerakan yang terekam dari video guna mengenali pola pergerakan yang khas dari huruf tersebut.

Dengan demikian, sistem yang dihasilkan mampu mengenali seluruh huruf abjad A hingga Z, baik yang bersifat statis maupun dinamis secara *real-time*. Sinergi antara kemampuan deteksi spasial dari YOLOv11 dan kemampuan

temporal dari LSTM menjadikan sistem ini lebih lengkap dan adaptif terhadap variasi bentuk serta gerakan tangan pengguna.

Dengan kombinasi ini, sistem mampu melengkapi kekurangan metode sebelumnya seperti CNN yang tidak mempertimbangkan dimensi waktu, maupun YOLOv5 murni yang kurang sensitif terhadap gerakan berurutan. Hasil dari penggabungan YOLOv11 dan LSTM diharapkan mampu memberikan solusi teknologi inklusif yang tidak hanya cepat, tetapi juga akurat dan andal dalam mengenali semua huruf dalam SIBI, sehingga dapat membantu penyandang disabilitas berkomunikasi secara lebih efektif di kehidupan sehari-hari.

1.2 Rumusan Masalah

- a. Bagaimana mengimplementasikan algoritma YOLO dan *Long Short Term Memory* (LSTM) untuk mengenali huruf-huruf dalam Sistem Isyarat Bahasa Indonesia (SIBI)?
- b. Bagaimana tingkat akurasi algoritma YOLO dan LSTM dalam mengenali huruf-huruf SIBI statis dan dinamis?

1.3 Tujuan Penelitian

- a. Mengembangkan sistem pendeteksi huruf-huruf SIBI berbasis algoritma YOLO dan LSTM.
- b. Menganalisis performa algoritma YOLO dan LSTM dalam mengenali huruf SIBI statis dan dinamis berdasarkan hasil evaluasi model serta pengujian sistem secara *real-time*.

1.4 Manfaat Penelitian

- a. Menyediakan solusi komunikasi berbasis teknologi yang dapat membantu penyandang disabilitas berkomunikasi lebih efektif, serta meningkatkan kesadaran masyarakat terhadap inklusi komunikasi.

- b. Memberikan kontribusi pada pengembangan teknologi pengenalan bahasa isyarat berbasis *Deep Learning*, khususnya menggunakan pendekatan YOLOv11.
- c. Menjadi dasar pengembangan produk berbasis *assistive technology*, seperti aplikasi penerjemah bahasa isyarat, sistem pembelajaran SIBI digital, atau fitur aksesibilitas berbasis *Computer Vision* yang dapat diterapkan pada bidang pendidikan, layanan publik, dan industri teknologi inklusif.

1.5 Batasan Penelitian

- a. Penelitian ini hanya mencakup pengenalan abjad SIBI (huruf A–Z).
- b. Huruf dinamis seperti “J” dan “Z” dikenali melalui pemrosesan urutan *frame* menggunakan algoritma LSTM.
- c. Model yang digunakan untuk deteksi huruf statis adalah YOLOv11n.
- d. *Dataset* yang digunakan berupa citra dan video yang dikumpulkan dari sumber publik dan pengambilan data mandiri.
- e. Sistem dibangun dan diuji dalam lingkungan dengan latar dan pencahayaan yang dikontrol.