

BAB I

PENDAHULUAN

1.1 Latar Belakang Penelitian

Penggunaan gestur tangan sebagai antarmuka manusia–komputer (*human–computer interaction/HCI*) telah berkembang pesat dalam beberapa tahun terakhir. Gestur tangan menawarkan bentuk interaksi yang alami, intuitif, dan tidak memerlukan kontak fisik, sehingga cocok digunakan pada berbagai bidang, mulai dari *virtual reality* (VR), *augmented reality* (AR), robotika, industri kreatif, hingga alat bantu komunikasi bagi penyandang disabilitas. Salah satu penerapan yang memiliki dampak sosial signifikan adalah penerjemahan bahasa isyarat ke bentuk teks atau suara untuk membantu komunikasi penyandang tunarungu. WHO memperkirakan sekitar 2,5 miliar orang di dunia akan mengalami gangguan pendengaran pada tahun 2050, yang sebagian di antaranya membutuhkan bahasa isyarat sebagai sarana komunikasi utama (Lu et al., 2023). Dalam konteks ini, teknologi pengenalan gestur tangan berperan sebagai jembatan antara bahasa isyarat dan komunikasi verbal. Akurasi pergerakan jari merupakan faktor penentu keberhasilan sistem pengenalan gestur, karena perbedaan kecil pada posisi atau kelengkungan jari dapat mengubah makna suatu isyarat. Sistem pengenalan gestur yang tidak mampu menangkap perbedaan halus ini akan rentan menghasilkan kesalahan interpretasi (Mummadi et al., 2018; Lu et al., 2023).

Metode visi komputer memanfaatkan citra atau video dari kamera untuk mendeteksi dan melacak posisi tangan serta jari. Salah satu algoritma yang populer adalah YOLO (*You Only Look Once*), yang mampu melakukan deteksi objek secara *real-time*. Penelitian Mujahid et al. (2021) menunjukkan bahwa YOLOv3 mampu mengenali gestur tangan dengan akurasi hingga 97,68%, mengungguli metode lain seperti SSD dan VGG16. Namun, pendekatan ini memiliki keterbatasan, seperti sensitif terhadap perubahan pencahayaan, sulit mendeteksi tangan jika latar belakang kompleks, dan rentan terhadap *occlusion* (misalnya jari menutupi jari lain). Kondisi tersebut dapat menyebabkan hilangnya informasi penting, terutama pada gerakan yang melibatkan interaksi antar jari.

Sebaliknya, pendekatan berbasis sarung tangan pintar (*smart glove*) memanfaatkan *flex sensor* (sensor kelenturan jari) untuk mengukur sudut tekuk jari dan sensor IMU untuk melacak orientasi tangan, yang dapat bekerja tanpa bergantung pada kualitas visual atau pencahayaan. Kelebihan ini membuat *smart glove* banyak digunakan pada penelitian *gesture recognition*. Misalnya, Mummadi et al. (2018) melaporkan bahwa sarung tangan dengan IMU di setiap ujung jari berhasil mengenali abjad bahasa isyarat Perancis dengan akurasi 92%. Meskipun demikian, sarung tangan sensor juga memiliki kekurangan, seperti *flex sensor* yang hanya mengukur sudut tekuk tanpa konteks spasial, IMU yang mengalami *drift*, serta kesulitan membedakan gestur dengan pola tekukan mirip tanpa informasi tambahan.

Kelemahan dari kedua pendekatan tersebut mendorong pengembangan metode *hybrid*, yang menggabungkan sensor fisik dan visi komputer untuk saling melengkapi. Penelitian Lu et al. (2023) menunjukkan bahwa penggabungan data *curvature* jari dari sarung tangan dengan informasi visual dari kamera meningkatkan akurasi pengenalan bahasa isyarat dari 68,34% menjadi 84,13%. Pada metode ini, *flex sensor* memberikan data presisi sudut tekuk jari meskipun jari tidak terlihat kamera, kamera memberikan konteks visual posisi dan bentuk tangan, dan IMU menambah informasi orientasi tangan dalam ruang tiga dimensi.

Penelitian ini dimulai tanpa rancangan teknis yang pasti, melainkan melalui proses mengkaji berbagai penelitian terdahulu untuk memahami kelebihan dan kekurangan pendekatan yang ada. Dari analisis ini, muncul ide untuk merancang sistem yang memadukan data sarung tangan sensor dan kamera dalam satu alur pemrosesan, sehingga diharapkan dapat menghasilkan sistem yang lebih andal pada berbagai kondisi lingkungan. Hasil akhir dari sistem yang dirancang adalah animasi tangan *virtual* pada platform Unity, di mana model 3D tangan akan bergerak sesuai data gabungan dari sensor dan kamera. Mekanismenya adalah gerakan jari *real-time* dari *flex sensor* saat tidak ada gestur terdeteksi kamera, gerakan jari sesuai gestur kamera jika deteksi kamera tervalidasi oleh data sensor, serta rotasi pergelangan tangan dari IMU untuk memberikan orientasi yang akurat. Unity berperan sebagai

media visualisasi yang memudahkan pengujian dan demonstrasi hasil sistem secara interaktif. Sistem ini diharapkan menjadi referensi bagi pengembangan teknologi *hand tracking* yang akurat, responsif, dan dapat diintegrasikan pada berbagai aplikasi seperti VR, AR, robotika, atau komunikasi bahasa isyarat.

1.2 Rumusan Masalah

Berdasarkan latar belakang yang telah diuraikan, rumusan masalah dalam penelitian ini adalah sebagai berikut:

1. Bagaimana merancang sistem *occlusion* yang mampu menggabungkan data dari sarung tangan sensor (*flex sensor* dan IMU) dan deteksi gestur berbasis YOLOv8 dari kamera eksternal, untuk memperoleh pengenalan gerakan jari yang akurat?
2. Bagaimana merancang dan mengimplementasikan integrasi antar komponen sarung tangan sensor, webcam, pemrosesan data, dan Unity dengan komunikasi *real-time* menggunakan protokol UDP?
3. Bagaimana hasil visualisasi gerakan tangan *virtual* di Unity berdasarkan data gerakan tangan yang diperoleh dari sistem *hybrid sensor* dan visi tersebut?

1.3 Tujuan Penelitian

Berdasarkan pada rumusan masalah di atas, penelitian ini mempunyai beberapa tujuan, antara lain: Tujuan dari penelitian ini adalah:

1. Merancang sistem pengenalan gerakan tangan berbasis kombinasi sarung tangan sensor (*flex* dan IMU) serta kamera eksternal berbasis YOLOv8, yang dapat mendeteksi dan mengklasifikasikan gestur statis secara akurat.
2. Mengimplementasikan prototipe sistem dengan integrasi komunikasi UDP antara sarung tangan, modul deteksi kamera, program pemrosesan data, dan Unity sebagai visualisasi tangan *virtual*.

3. Memvisualisasikan hasil gerakan jari dan pergelangan tangan pengguna secara *real-time* pada model tangan 3D di Unity, baik berdasarkan *input sensor* maupun gestur yang terdeteksi kamera.

1.4 Manfaat Penelitian

1. Memberikan kontribusi terhadap pengembangan metode *hybrid gesture recognition* berbasis sensor dan *computer vision*, khususnya untuk pengenalan gerakan jari secara akurat pada pengguna yang memakai sarung tangan.
2. Menjadi referensi bagi penelitian selanjutnya dalam bidang *human-computer interaction* (HCI), VR, atau AR yang membutuhkan akurasi tinggi dalam pelacakan jari.
3. Menyediakan alternatif sistem *hand tracking* yang kompatibel dengan penggunaan sarung tangan, yang tidak didukung oleh sebagian besar sistem *vision-based* standar seperti MediaPipe.
4. Menjadi dasar pengembangan sistem kontrol gestur untuk aplikasi *virtual* seperti pelatihan, simulasi, atau interaksi tanpa kontak, khususnya di lingkungan yang memerlukan perlindungan tangan (misalnya industri, medis, atau robotik).
5. Memberikan solusi implementasi *connectionless* dan *real-time* yang dapat digunakan dalam aplikasi edukatif dan pengembangan teknologi *wearable* sederhana.

1.5 Batasan Masalah

Agar penelitian terarah, batasan masalah ditetapkan sebagai berikut:

1. Jenis Gestur: Sistem hanya menangani gestur tangan statis. Gestur dinamis atau sekuensial (yang melibatkan gerakan kontinu seperti melambai atau menunjuk) tidak menjadi fokus penelitian.
2. Jenis Input: Data input berasal dari dua sumber utama, yaitu:
 - a. Lima *flex sensor* dan satu IMU pada sarung tangan di masing-masing tangan.

- b. Citra video dari kamera eksternal untuk mendeteksi label gestur menggunakan model YOLOv8 hasil pelatihan pribadi.
3. Model yang Digunakan: Sistem menggunakan YOLOv8 (*varian yolov8n.pt*) yang telah dilatih melalui platform Roboflow menggunakan *dataset* tangan bersarung.
4. Keluaran Sistem: Hasil akhir sistem berupa pergerakan tangan *virtual* (model *Mixamo X Bot*) pada aplikasi Unity yang dikendalikan berdasarkan data gerakan sensor dan hasil deteksi kamera.
5. Protokol Komunikasi: Pengiriman data antar perangkat (ESP32, kamera, pemrosesan data, dan Unity) menggunakan protokol *User Datagram Protocol* (UDP) melalui jaringan lokal (Wi-Fi).
6. Evaluasi Sistem: Pengujian sistem hanya dilakukan pada aspek fungsionalitas, seperti sinkronisasi data, keakuratan visualisasi gerakan tangan, dan respons waktu sistem secara umum. Pengujian pengguna secara kuantitatif (*user study*) atau analisis statistik akurasi tidak dilakukan secara mendalam.
7. Spesifikasi Komputer Pengujian: Pengujian dilakukan pada komputer dengan spesifikasi berikut:
 - Prosesor: Intel(R) Core(TM) i5-7300HQ CPU @ 2.50GHz
 - RAM: 16 GB
 - GPU: NVIDIA GeForce GTX 1050 4 GB VRAM
 - Sistem Operasi: Windows 10 64-bit
 - Versi Python: 3.11
 - Versi Unity: 2022 ke atas
 - Arduino IDE untuk pemrograman ESP32
 - Visual Studio Code untuk pemrograman python dan c#