

BAB 1

PENDAHULUAN

1.1 Latar Belakang

Perkembangan teknologi robotika pada era Revolusi Industri 4.0 telah mendorong terciptanya sistem otomatis yang semakin cerdas, adaptif, dan interaktif. Robot kini tidak hanya digunakan dalam industri manufaktur, tetapi juga telah merambah ke sektor pendidikan, layanan publik, serta kehidupan rumah tangga. Salah satu jenis robot yang paling banyak dikembangkan adalah *robot arm*, yaitu lengan robotik dengan beberapa derajat kebebasan (Degree of Freedom) yang dirancang untuk melakukan manipulasi objek secara otomatis (Siciliano et al., 2010). Namun, sebagian besar sistem kendali robot arm masih bergantung pada *tombol fisik*, *joystick*, atau pemrograman langsung, yang kurang intuitif dan tidak efisien bagi pengguna (Saad & Salah, 2020).

Seiring meningkatnya kebutuhan interaksi manusia dan mesin yang lebih alami, teknologi *speech recognition* atau pengenalan suara mulai banyak digunakan dalam sistem robotika. Teknologi ini memungkinkan pengguna memberikan instruksi verbal tanpa menggunakan alat kendali tambahan. Dengan pengenalan suara, interaksi antara manusia dan robot menjadi lebih efisien, alami, dan inklusif, terutama bagi pengguna dengan keterbatasan fisik (Martínez et al., 2013).

Bersamaan dengan itu, kemajuan pada bidang *computer vision* juga memunculkan sistem pengenalan visual yang semakin akurat. Salah satu bentuk implementasinya adalah *colour recognition* atau pengenalan warna, yang memungkinkan robot mengenali dan mengelompokkan objek berdasarkan warna dominan. Teknologi ini banyak diterapkan dalam sistem penyortiran otomatis, navigasi visual, dan aplikasi robotik pendidikan (Gonzalez & Woods, 2008).

Meskipun demikian, penelitian-penelitian terdahulu menunjukkan bahwa metode pengenalan warna berbasis ruang warna RGB memiliki kelemahan mendasar, yaitu sensitif terhadap perubahan intensitas cahaya. Penelitian yang dilakukan oleh (Samsudin, 2016) misalnya, menggunakan metode RGB dan ANFIS untuk deteksi warna pada robot arm, namun hasilnya tidak stabil ketika pencahayaan berubah. Hal ini menunjukkan perlunya pendekatan baru yang lebih tahan terhadap kondisi lingkungan yang bervariasi.

Sementara itu, dari sisi pengenalan suara, metode klasik seperti *Hidden Markov Model* (HMM) dan *Recurrent Neural Network* (RNN) sering digunakan, tetapi memiliki keterbatasan dalam efisiensi komputasi. RNN dan turunannya seperti LSTM cenderung membutuhkan waktu pelatihan yang lama dan tidak efisien ketika dijalankan pada perangkat dengan sumber daya terbatas. Seiring kemunculan arsitektur *Transformer*, metode ini menawarkan kemampuan untuk memproses urutan data secara paralel dengan mekanisme *self-attention* (Vaswani et al., 2017). Namun, arsitektur *Transformer* penuh memiliki kompleksitas tinggi dan sulit diimplementasikan pada sistem berskala kecil seperti mikrokontroler.

Untuk mengatasi keterbatasan tersebut, digunakanlah *Lightweight Transformer Encoder Block* (LTEB), yaitu versi ringan dari arsitektur *Transformer* yang tetap mempertahankan keunggulan dalam mempelajari pola temporal tetapi dengan jumlah parameter yang jauh lebih sedikit. LTEB memungkinkan proses pengenalan suara dilakukan secara efisien pada perangkat berdaya rendah seperti Arduino atau Raspberry Pi (Tay et al., 2022).

Dari sisi penglihatan komputer, pendekatan berbasis ruang warna *Hue, Saturation, Value* (HSV) memberikan hasil yang lebih stabil terhadap perubahan pencahayaan dibandingkan ruang warna RGB. Metode ini mampu memisahkan komponen warna dan intensitas

cahaya, sehingga lebih akurat untuk aplikasi *real-time* dan lebih mudah diimplementasikan tanpa memerlukan pelatihan model yang kompleks (Bradski, 2000).

Berdasarkan kajian literatur, sebagian besar penelitian sebelumnya hanya berfokus pada satu jenis modalitas, baik audio atau visual. Misalnya, (A. Hammadi et al., 2021) mengembangkan sistem pengenalan warna berbasis CNN, sedangkan (Xu & Zhang, 2023b) menggunakan *Transformer* untuk pengenalan suara. Keduanya menghasilkan akurasi tinggi, tetapi belum mengintegrasikan kedua pendekatan tersebut secara bersamaan. Oleh karena itu, diperlukan pendekatan multimodal yang ringan secara komputasi, mampu bekerja secara *real-time*, dan dapat dijalankan pada perangkat keras berbiaya rendah.

Penelitian ini dilakukan untuk menjawab kebutuhan tersebut, dengan menggabungkan sistem pengenalan suara berbasis LTEB dan sistem deteksi warna berbasis HSV. Integrasi kedua metode ini diharapkan menghasilkan sistem kendali robot arm yang intuitif, efisien, dan mudah diimplementasikan. Pendekatan ini tidak hanya menawarkan peningkatan efisiensi dalam pengendalian robot, tetapi juga membuka peluang aplikasi lebih luas seperti *assistive robotics* untuk penyandang disabilitas dan *home automation* berbasis suara.

Teknologi bantu berbasis suara untuk penyandang disabilitas sebenarnya telah dikembangkan dalam berbagai bentuk, seperti kursi roda yang dikendalikan suara dan robot asisten sederhana untuk aktivitas sehari-hari (Raj et al., 2015; Ahmed et al., 2019). Namun, penerapan sistem yang menggabungkan kendali suara, pengenalan visual, dan kemampuan manipulatif seperti robot arm masih jarang ditemukan dalam aplikasi nyata. Oleh karena itu, penelitian ini memiliki posisi yang strategis dalam pengembangan sistem robotika cerdas yang inklusif dan efisien.

Dengan demikian, penelitian ini menempati posisi penting di antara studi-studi sebelumnya karena menggabungkan dua pendekatan ringan dan adaptif LTEB untuk domain suara dan HSV untuk domain visual ke dalam satu sistem kendali robot arm. Sistem ini diharapkan menjadi langkah awal menuju pengembangan robotika multimodal generasi baru yang tidak hanya efisien, tetapi juga mampu memberikan dampak sosial nyata dalam bidang pendidikan, industri, dan teknologi asistif.

1.2 Rumusan Masalah

Berdasarkan latar belakang yang telah dijelaskan, maka permasalahan dalam penelitian ini dapat dirumuskan sebagai berikut:

1. Bagaimana merancang dan mengimplementasikan metode *Lightweight Transformer Encoder Block* (LTEB) untuk melakukan pengenalan suara (*speech recognition*) pada sistem kendali Robot Arm?
2. Bagaimana membangun sistem pendeteksian warna berbasis pemrosesan citra menggunakan ruang warna HSV dan deteksi centroid sebagai penentu posisi objek?
3. Bagaimana cara mengintegrasikan hasil prediksi suara menggunakan LTEB dan hasil pendeteksian warna berbasis HSV ke dalam satu sistem kendali Robot Arm berbasis Arduino UNO?
4. Bagaimana kinerja sistem Robot Arm dalam proses pendeteksian warna, akurasi perintah suara, serta ketepatan posisi objek berdasarkan hasil pengujian?

1.3 Batasan Masalah

Agar penelitian tetap terarah dan tidak melebar, maka batasan masalah ditetapkan sebagai berikut:

1. Sistem hanya mengenali tiga warna, yaitu merah, biru, dan kuning.
2. Model pengenalan suara menggunakan arsitektur *Lightweight Transformer Encoder Block* (LTEB) dengan dataset suara yang dibuat secara mandiri.
3. Deteksi objek menggunakan segmentasi warna dalam ruang HSV tanpa model *deep learning* untuk citra.
4. Pengujian dilakukan pada lingkungan terkontrol dengan pencahayaan stabil.
5. Robot arm yang digunakan berbasis Arduino UNO dan servo SG90, sehingga jangkauan dan kekuatan gerak terbatas.

1.4 Tujuan Penelitian

Adapun tujuan dari penelitian ini adalah:

1. Membangun dan mengimplementasikan model *Lightweight Transformer Encoder Block* (LTEB) untuk mendeteksi dan mengklasifikasikan warna berdasarkan perintah suara.
2. Mengembangkan sistem pendeteksi warna berbasis HSV dan analisis kontur untuk menentukan posisi warna pada grid 3x3.
3. Mengintegrasikan pengenalan suara dan pendeteksian warna ke dalam satu sistem kendali Robot Arm berbasis Arduino UNO.
4. Melakukan evaluasi terhadap performa sistem berdasarkan tingkat akurasi prediksi warna, ketepatan penentuan posisi, serta keberhasilan gerakan Robot Arm.

1.5 Manfaat Penelitian

Penelitian ini diharapkan memberikan beberapa manfaat berikut:

1. Memberikan kontribusi dalam pengembangan sistem Robot Arm berbasis *speech recognition* menggunakan arsitektur *Lightweight Transformer Encoder Block* yang lebih efisien dan ringan.
2. Menyediakan referensi dalam penerapan metode pendeteksian warna berbasis HSV dan centroid untuk aplikasi robotika.
3. Menjadi dasar pengembangan lebih lanjut untuk sistem robotika cerdas yang mampu merespons instruksi suara sekaligus mengenali objek secara visual.
4. Meningkatkan efisiensi dan efektivitas interaksi antara manusia dan robot dengan menggunakan metode pemrosesan suara dan visual yang cepat, ringan, dan dapat diimplementasikan pada perangkat berbiaya rendah.

1.6 Sistematika Penulisan

Penulisan Skripsi ini disusun ke dalam enam buah bab dengan sistematika sebagai berikut.

BAB I : PENDAHULUAN

Memberikan pemahaman mengapa penelitian ini dilakukan, dengan menguraikan latar belakang yang memicu penelitian, rumusan masalah yang menjadi fokus, batasan-batasan yang mengapit penelitian, tujuan yang ingin dicapai, manfaat yang diharapkan dari hasil penelitian, dan rangkuman singkat tentang struktur keseluruhan penulisan.

BAB II : TINJAUAN PUSTAKA

Merinci pengetahuan terkait dengan penelitian ini. Ini mencakup ulasan literatur yang relevan dan konsep-konsep yang menjadi dasar bagi penelitian ini, serta menjelaskan prinsip dasar dari Robot Arm berbasis *Speech Recognition* dan *colour Recognition*

BAB III : METODE PENELITIAN

Menggambarkan secara detail tentang bagaimana penelitian ini dilakukan. Ini mencakup lokasi dan waktu penelitian, daftar peralatan dan bahan yang digunakan, serta langkah-langkah prosedur penelitian.

BAB IV : HASIL DAN PEMBAHASAN

Menjelaskan hasil pengujian dan pembahasan metode yang di gunakan secara *real time* untuk mengontrol gerak Robot Arm.

BAB V : PENUTUP

Merangkum hasil penelitian dan mengidentifikasi potensi kekurangan dalam penelitian ini. Rekomendasi akan diberikan untuk pengembangan penelitian lebih lanjut di masa depan.