

BAB I

PENDAHULUAN

1.1. Latar Belakang

Saat ini, teknologi dan informasi telah berkembang sangat pesat karena masyarakat terus menginginkan segala aktivitasnya menjadi lebih mudah dan efisien. Salah satunya adalah teknologi *human-computer interaction* menjadi salah satu tuntutan dalam perkembangan teknologi. Dalam *human-computer interaction*, penggunaan *hand gesture* telah menarik perhatian masyarakat karena mudahnya manusia berinteraksi dengan mesin (Sharma dan Verma, 2015). *Human-computer interaction* memungkinkan pengguna berinteraksi dengan mesin melalui antarmuka pengguna grafis, suara, gambar, dan berbagai perangkat masukan lainnya. Salah satu contoh interaksinya adalah mengontrol perangkat melalui antarmuka perangkat keras masukan, yaitu tombol yang ditekan oleh pengguna untuk melakukan perintah kepada suatu perangkat. Seperti pada *lift*, dimana untuk mengontrolnya pengguna harus menekan tombol yang tersedia seperti buka/tutup pintu, berpindah ke lantai yang dimaksud (Perdana dan Sirait, 2019).

Lift merupakan salah satu fasilitas umum yang paling sering digunakan pada bangunan bertingkat, namun kontak langsung manusia dengan tombol *lift* akan meningkatkan kemungkinan terjadinya infeksi virus dan bakteri yang tersebar melalui udara dan *droplet* yang menempel pada permukaan benda. Pada tahun 2020, terdapat kasus pengguna *lift* menularkan *virus corona* (Covid-19) ke 71 orang, hal tersebut terjadi karena terdapat pengguna yang terinfeksi Covid-19 menyentuh

tombol *lift*, kemudian pengguna lainnya menyentuh tombol tersebut yang sudah terkontaminasi (Ferdian dan Panji, 2020). Secara umum untuk mengatasi masalah ini adalah dengan menempatkan *hand sanitizer* dan tisu di dalam *lift* sehingga pengguna dapat segera membersihkan tangannya setelah menyentuh *lift*. Solusi lainnya adalah menyediakan tongkat sekali pakai yang bersih atau benda lain sebagai alat untuk menekan tombol *lift*. Pengguna yang telah menggunakan tisu dan tongkat sekali pakai harus membuangnya ke tempat sampah. Meskipun terdapat tempat sampah, hal tersebut tidak dapat menjamin bahwa setiap orang menggunakan tempat sampah yang akan mempengaruhi kebersihan dan beberapa orang mungkin tidak suka menggunakan *hand sanitizer* dan tisu setiap naik *lift* (Lekki, 2021). Hal tersebut dapat diatasi dengan pengenalan *hand gesture* untuk kontrol perangkat sebagai pengganti tombol, seperti pada penelitiannya (A. Hande dan R. Chopde, 2020) tentang “*Implementation of Gesture Control Robotic Arm for Automation of Industrial Application*”. Penelitian tersebut, untuk mengendalikan atau kontrol perangkat dapat diimplementasikan dengan pengenalan *hand gesture*. Sistem pengenalan *hand gesture* tersebut memanfaatkan *flex sensor* yang ditempel pada sarung tangan kemudian dipakai oleh pengguna untuk mengontrol *Robotic Arm*. Namun, pendekatan ini terdapat masalah yaitu menyebabkan kerusakan kulit pada pengguna yang memiliki kulit sensitif (Oudah, Al-Naji dan Chahl, 2020).

Permasalahan diatas dapat diatasi dengan pendekatan *deep learning* dan *computer vision* yang diterapkan menggunakan kamera dan *hand gesture* sebagai *input* untuk kontrol perangkat. *Hand gesture* untuk kontrol perangkat dapat diimplementasikan dengan objek *hand gesture* yang ditangkap melalui kamera, lalu

setiap *frame* gambar yang berhasil ditangkap kemudian diproses oleh algoritma *deep learning* untuk mendeteksi *hand gesture* (Sharma dan Singh, 2021). Setelah *hand gesture* terdeteksi maka sistem mengirim sinyal kepada perangkat *output* untuk melakukan aksinya.

Penelitian ini bertujuan untuk mengimplementasikan deteksi *hand gesture* sebagai kontrol lift untuk berpindah lantai. Dalam penelitian ini, *hand gesture* digunakan sebagai *input* pengganti tombol untuk mengoperasikan *lift* tanpa kontak langsung dan *lift* yang digunakan dalam bentuk *prototype* yang dibangun sebanyak 3 lantai. *Lift* dapat mencapai lantai yang ditentukan berdasarkan jenis *gesture* yang terdeteksi. *Hand gesture* dideteksi menggunakan algoritma deteksi objek berbasis *deep learning*.

Beragam algoritma *deep learning* yang digunakan untuk deteksi *hand gesture*. Pertama, Faster R-CNN digunakan untuk mendeteksi *hand gesture*. Pada penelitian ini menggunakan *custom* dataset sebanyak 7500 gambar dengan 10 *class* yang berbeda, di mana setiap *class* memiliki rata-rata 750 gambar. Berdasarkan pengujian yang dilakukan, diperoleh hasil *precision* rata-rata 0,794, *recall* rata-rata 0,833, dan *F1-score* 0,813 (Bose dan Kumar, 2019). Kedua, *hand gesture* digunakan untuk *traffic control* menggunakan *Recurrent Neural Network* (RNN). Dataset yang digunakan sebanyak 100.000 gambar dengan 8 *class* yaitu *downward*, *frontward*, *backward*, *left*, *right*, *upward*, *diagonal*, *left and right*. Berdasarkan hasil pengujian pengenalan *hand gesture*, diperoleh nilai akurasi 91% dengan latar belakang yang berbeda (Baek dan Lee, 2022). Ketiga, YOLOv2 digunakan untuk mendeteksi dan mengklasifikasikan postur tangan statis. Dataset yang digunakan

terdiri dari 2000 gambar tangan dan 750 gambar tangan dengan noise. Berdasarkan hasil evaluasi model, diperoleh nilai akurasi rata-rata 97,46% dari 750 gambar tangan dengan *noise* dan 100% dari 200 gambar tangan tanpa *noise* (Tanmaie dan Rao, 2020). Keempat, YOLOv3 digunakan untuk mendeteksi bahasa isyarat Malaysia. YOLOv3 dikenal dengan kinerjanya yang lebih baik dari versi yang sebelumnya yaitu YOLOv2. Penelitian ini dimulai dengan pengumpulan gambar yang diperoleh dari *internet*, selanjutnya dilakukan anotasi untuk memberikan label pada setiap gambar. Berdasarkan pengujian yang telah dilakukan, sistem mencapai akurasi 63% dengan *learning saturation (overfitting)* pada iterasi ke-7000 dari *training* dan 72% dari pengujian sistem yang sesungguhnya (Asri dkk, 2019).

Penelitian yang diusulkan menggunakan salah satu algoritma deteksi objek berbasis *deep learning* yaitu YOLOv4. Berdasarkan penelitian terkait deteksi *hand gesture* menggunakan YOLOv4 seperti pada penelitiannya (Zheng, 2022) tentang *hand gesture* sebagai kontrol *music box*, berhasil mendeteksi 7 jenis *gesture* menghasilkan tingkat akurasi sebesar 97,8%. Pada penelitiannya (Chen dan Huang, 2021) tentang *hand gesture* sebagai kontrol lengan robot, berhasil mendeteksi 8 jenis *gesture* menghasilkan tingkat mAP sebesar 99,93%. Tingginya hasil yang diperoleh sehingga dapat disimpulkan bahwa *hand gesture* dapat dideteksi dengan baik menggunakan YOLOv4. Selain itu, algoritma ini mampu mendeteksi objek secara efisien, cepat, dan akurat. Hal ini terbukti pada penelitian sebelumnya dalam perbandingan YOLOv4 dengan algoritma deteksi objek lainnya bahwa YOLOv4 telah meningkatkan *average precision (AP)* dan *frame per second (FPS)* masing-masing sebesar 10% dan 12% (Bochkovskiy dkk, 2020). Penelitian ini

menggunakan *static hand gesture* sebagai objek deteksinya. Dalam *static gesture*, posisi tangan tidak berubah selama melakukan gerakan. Dengan kata lain, *static gesture* adalah konfigurasi dan postur tangan tertentu yang diwakili oleh satu citra (Chandra, Venu dan Srikanth, 2015).

1.2. Rumusan Masalah

Berdasarkan latar belakang di atas, maka dapat dirumuskan masalah penelitian sebagai berikut :

1. Bagaimana implementasi YOLOv4 berbasis *deep learning* sebagai model deteksi *hand gesture* untuk kontrol *prototype lift* ?
2. Bagaimana kinerja parameter model deteksi *hand gesture* yang telah dibangun ?

1.3. Tujuan Penelitian

Tujuan penelitian yang ingin dicapai adalah :

1. Mengimplementasikan YOLOv4 berbasis *deep learning* sebagai model deteksi *hand gesture* untuk kontrol *prototype lift*.
2. Mengetahui kinerja parameter model deteksi *hand gesture* yang telah dibangun.

1.4. Batasan Penelitian

Terdapat beberapa batasan masalah dalam penelitian ini, yaitu :

1. Dataset yang digunakan adalah gambar tangan ukuran dewasa yang diambil menggunakan kamera dengan jarak 0,3 sampai 0,5 meter, gambar berukuran 416x416 piksel yang telah di augmentasi.
2. Dataset terdiri dari 3 *class* yaitu *gesture* jari tangan angka satu, dua, dan tiga.
3. *Gesture* yang diperbolehkan pada masing-masing kelas adalah kelas Lantai 1 menggunakan jari telunjuk, Lantai 2 menggunakan jari telunjuk dan tengah, Lantai 3 menggunakan jari telunjuk, tengah, dan manis.
4. Jenis *gesture* yang digunakan adalah *static hand gesture*
5. Maksimal jarak deteksi adalah 0,5 meter.
6. Model yang digunakan algoritma deteksi objek berbasis *deep learning* yaitu YOLOv4.
7. Perangkat keluaran yang digunakan pada penelitian ini adalah *lift* dalam bentuk *prototype lift* sebanyak 3 lantai.
8. Aksi yang diberikan pada *prototype lift* hanya berpindah lantai, tidak dengan buka/tutup pintu.
9. Parameter yang diuji adalah *mean Average Precision* (mAP) dan jarak deteksi.

1.5. Manfaat Penelitian

Adapun beberapa manfaat yang diperoleh dalam penelitian ini adalah sebagai berikut :

1. Dapat menjadi referensi dalam penelitian di bidang *AI computer vision*, khususnya penelitian terkait implementasi *deep learning* untuk mendeteksi *hand gesture*.
2. Dapat digunakan sebagai pengendali alternatif pada *lift* yang sesungguhnya.