



PROPOSAL TUGAS AKHIR - TD123456

**KENDALI *MOBILE ROBOT* BERBASIS POSE TANGAN
MENGUNAKAN *CONVOLUTIONAL NEURAL
NETWORK (CNN)***

Alfan Miftah Arzaqi

NRP 0721 19 4000 0003

Dosen Pembimbing I

Ahmad Zaini, S.T., M.Sc.

NIP 19750419200212 1 003

Dosen Pembimbing II

Dr. Eko Mulyanto Yuniarno, S.T., M.T.

NIP 19680601199512 1 009

Program Studi Strata 1 (S1) Teknik Komputer

Departemen Teknik Komputer

Fakultas Teknologi Elektro dan Informatika Cerdas

Institut Teknologi Sepuluh Nopember

Surabaya

2022

LEMBAR PENGESAHAN

KENDALI *MOBILE ROBOT* BERBASIS POSE TANGAN MENGGUNAKAN *CONVOLUTIONAL NEURAL NETWORK (CNN)*

PROPOSAL TUGAS AKHIR

Diajukan untuk memenuhi salah satu syarat memperoleh gelar Sarjana Teknik pada Program Studi S-1
Teknik Komputer

Departemen Teknik Komputer
Fakultas Teknologi Elektro dan Informatika Cerdas
Institut Teknologi Sepuluh Nopember

Oleh: **Alfan Miftah Arzaqi**
NRP 0721 19 4000 0003

Disetujui oleh Tim Penguji Proposal Tugas Akhir:

Ahmad Zaini, S.T., M.Sc.
NIP 19750419200212 1 003

(Pembimbing 1)

Dr. Eko Mulyanto Yuniarno, S.T., M.T.
NIP 19680601199512 1 009

(Pembimbing 2)

-
NIP: -

(Penguji I)

-
NIP: -

(Penguji II)

-
NIP: -

(Penguji III)

SURABAYA
Desember, 2022

KENDALI *MOBILE ROBOT* BERBASIS POSE TANGAN MENGUNAKAN *CONVOLUTIONAL NEURAL NETWORK* (CNN)

Nama Mahasiswa / NRP: **Alfan Miftah Arzaqi / 0721 19 4000 0003**

Departemen : **Teknik Komputer FTEIC - ITS**

Dosen Pembimbing : **1. Ahmad Zaini, S.T., M.Sc.
Dr. Eko Mulyanto Yuniarno, S.T., M.T.**

Abstrak

Pose tangan biasa digunakan oleh teman-teman tuli untuk berkomunikasi, namun dengan perkembangan teknologi *Human Computer Interaction* atau HCI yang memungkinkan untuk berkomunikasi dengan komputer menggunakan pose tangan. Pose tangan yang akan digunakan disini ada 5 pose yang akan menyimbolkan maju, mundur, berhenti, belok kanan, dan belok kiri. Proses kalsifikasi akan menggunakan CNN dalam melakukan training dan akan menggunakan citra yang sudah disimpan dari tahapan dataset. Dari data yang telah di klasifikasikan dan sudah didapatkan posenya maka pose tersebut akan diterjemahkan ke dalam suatu perintah untuk dapat menggerakkan *mobile robot*. Robot akan menerima data dari hasil klasifikasi dan selanjutnya akan bergerak sesuai data yang diberikan.

Kata Kunci: **CNN, *Mobile Robot*, Pose,**

DAFTAR ISI

LEMBAR PENGESAHAN	ii
ABSTRAK	iii
DAFTAR ISI	iv
DAFTAR GAMBAR	v
DAFTAR TABEL	vi
1 PENDAHULUAN	1
1.1 Latar Belakang	1
1.2 Rumusan Masalah	1
1.3 Batasan Masalah atau Ruang Lingkup	1
1.4 Tujuan	2
1.5 Manfaat	2
2 TINJAUAN PUSTAKA	2
2.1 Hasil Penelitian Terdahulu	2
2.2 Teori Dasar	2
2.2.1 Mobile Robot	2
2.2.2 Estimasi Pose Tangan	3
2.2.3 <i>Convolutional Neural Network (CNN)</i>	3
3 METODOLOGI	4
3.1 Metode yang Digunakan	4
3.1.1 Perangkat Lunak	4
3.1.2 Perangkat Keras	6
3.2 Urutan Pelaksanaan Penelitian	7
4 HASIL YANG DIHARAPKAN	7
4.1 Hasil yang Diharapkan dari Penelitian	7
4.2 Hasil Pendahuluan	7
5 DAFTAR PUSTAKA	7

DAFTAR GAMBAR

2.1	21 titik <i>keypoint</i> pada tangan (“MediaPipe Hands”, 2021)	3
2.2	Arsitektur CNN (Desphande, 2016)	4
3.1	Blok diagram perangkat lunak	4
3.2	Blok diagram pembuatan dataset	5
3.3	Pose tangan yang digunakan	5
3.4	Blok diagram perangkat keras	6
4.1	Hasil estimasi pose menggunakan <i>Mediapipe</i>	7

DAFTAR TABEL

3.1	Urutan pelaksanaan penelitian	7
-----	---	---

1 PENDAHULUAN

1.1 Latar Belakang

Pose adalah potongan-potongan dari gerakan yang bertujuan untuk komunikasi non verbal yang digunakan untuk menyampaikan pesan penting, baik secara langsung maupun tidak langsung tanpa menggunakan kata-kata. Setiap pose dapat memiliki arti tersendiri sesuai dengan kesepakatan umum ataupun personal yang melakukan komunikasi. Pose tangan berarti suatu sikap yang diberikan oleh tangan untuk berkomunikasi (Wibowo, 2022).

Pose tangan ini biasa digunakan oleh teman-teman tuli untuk berkomunikasi, namun dengan perkembangan teknologi dimana adanya teknologi *human computer interaction* yang memungkinkan untuk berkomunikasi dengan komputer menggunakan pose tangan. *Human Computer Interaction* atau biasa disingkat HCI adalah suatu bidang studi yang merancang teknologi komputer untuk lebih interaktif saat dipakai. Adanya HCI ini dapat memudahkan manusia untuk menyelesaikan tugas-tugasnya dengan bantuan komputer. HCI memungkinkan manusia untuk berinteraksi lebih dengan benda-benda disekitarnya, salah satu contohnya adalah penggunaan gestur tangan pada saat menggunakan *handphone*, mematikan atau menyalakan lampu dengan perintah suara, serta menggerakkan robot (Risetya, 2022).

Pada saat ini perkembangan HCI cukup pesat yang diakibatkan dari perkembangan pada teknologi *machine learning* dan robotika. Teknologi *machine learning* memungkinkan komputer untuk belajar dari data-data yang diberikan oleh manusia dan juga mengambil keputusan dari data-data tersebut. Algoritma machine learning dirancang agar dapat belajar dan kemampuannya meningkat seiring waktu ketika terdapat data baru tanpa diprogram secara eksplisit. Pada salah satu cabang *machine learning* yaitu *computer vision* memungkinkan untuk komputer mengerti gerakan atau pose yang diberikan manusia melalui gambar atau video (Hidayatullah, 2021). Tidak hanya *machine learning* yang berkembang namun pada robotika juga berkembang terutama pada sistem kendalinya. Awalnya robot hanya dapat dikendalikan secara dekat, namun beberapa tahun berikutnya robot sudah bisa dikendalikan dengan jarak yang jauh dengan tanpa kabel atau wireless dan dikendalikan dengan remote control. Teknologi kendali robot telah dikembangkan yang dapat bergerak sesuai dengan inputan dari gestur manusia melalui bantuan sensor-sensor yang diletakkan pada tangan atau bagian tubuh lainnya dan juga dapat menggunakan *computer vision* untuk mengolah gestur tubuh manusia.

1.2 Rumusan Masalah

Berdasarkan latar belakang di atas, pada Tugas Akhir ini diajukan rancangan mendeteksi pose tangan dengan menggunakan Mediapipe dan diimplementasikan untuk mengendalikan mobil robot Arduino. Langkah pertama yaitu mengumpulkan dataset melalui ekstraksi citra dari webcam menggunakan Mediapipe. Kemudian dilakukan proses learning menggunakan machine learning serta pengujian akurasi deteksi pose tangan. Setelah itu data pembacaan pose tangan tersebut akan dikirimkan kepada *bluetooth* yang ada pada mobil dan akan diterjemahkan menjadi input perintah logika pemrograman untuk menentukan aksi gerakan robot.

1.3 Batasan Masalah atau Ruang Lingkup

Dalam pembuatan Tugas Akhir ini, pembahasan masalah dibatasi beberapa hal berikut :

1. Pengambilan citra akan dilakukan dengan menggunakan webcam yang terdapat pada Laptop.

2. Pendeteksian tangan akan menggunakan *framework mediapipe*.
3. Pembatasan pose tangan yang menjadi dataset hanya akan ada 5 yakni maju, mundur, berhenti, belok kiri, dan belok kanan.
4. Pengujian yang dilakukan adalah pengujian akurasi klasifikasi pada pose tangan.

1.4 Tujuan

Berdasarkan latar belakang dan rumusan masalah diatas, mendapatkan tujuan pada Tugas Akhir ini sebaga berikut :

1. Mendeteksi tangan menggunakan Mediapipe.
2. Mengetahui pose tangan menggunakan Convolutional Neural Network.
3. Membuat sistem kendali untuk mobil robot arduino menggunakan pose tangan.

1.5 Manfaat

Manfaat dari penelitian ini yaitu membuat kendali untuk *mobile robot* berbasis pose tangan yang akurat dan dapat digunakan oleh masyarakat secara umum. Dari penelitian ini diharapkan dapat sebagai opsi lain dalam mengontrol robot yang lebih mudah dan tanpa alat lain.

2 TINJAUAN PUSTAKA

2.1 Hasil Penelitian Terdahulu

Pada November 2020, Jati Widyo Leksono, Agung Samudra, Nanndo Yannuansa, dan Ahmad Fauzi membuat jurnal mengenai sistem kendali mobil robot menggunakan isyarat tangan. Dalam penelitian ini akan dilakukan pengembangan mobil robot yang mampu berjalan sesuai dengan isyarat tangan atau gestur tangan yang diberikan oleh user. Pada telapak tangan user nantinya akan ditempelkan suatu sensor yang mampu membaca pergerakan telapak tangan menurun, naik, kiri, kanan, dan mendarat. Dari pembacaan sensor tersebut data akan dikirimkan menuju receiver yang terdapat pada robot secara wireless (Leksono et al., 2020).

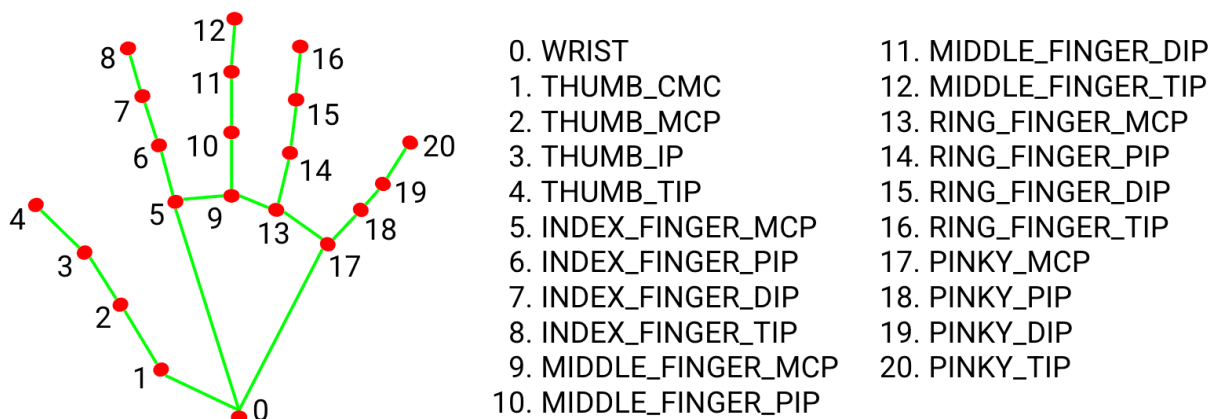
2.2 Teori Dasar

2.2.1 Mobile Robot

Robot merupakan suatu perangkat otomatis yang dapat menjalankan suatu tugas yang berasal dari manusia (Jaya, 2016). Mobile Robot merupakan salah satu jenis robot yang dapat berpindah posisi dari satu titik ke titik yang lain. Perpindahan robot ini dapat dimungkinkan dengan adanya suatu aktuator yang memungkinkan untuk menggerakkan keseluruhan badan robot. Salah satu contoh dari mobile robot adalah robot beroda yang terdiri dari 2 buah roda yang berpasangan pada kanan dan kiri robot serta digerakkan memutar oleh aktuator. Pergerakan robot beroda akan diatur oleh arah serta kecepatan putaran robot. Pada saat arah serta kecepatan putaran roda sama maka robot akan bergerak lurus, namun jika arah atau kecepatan putaran roda ada yang berbeda pada satu sisi roda maka akan membuat robot berbelok (Sulistyo, 2017).

2.2.2 Estimasi Pose Tangan

Pengenalan gestur atau pose tangan merupakan topik yang menarik dalam ilmu komputer dan teknologi yang bertujuan untuk komunikasi antara manusia dengan komputer menggunakan gestur tanpa menyentuhnya secara langsung. Saat ini banyak *framework* atau *library* pembelajaran mesin untuk mengenali gerakan tangan, salah satu yaitu Mediapipe. Mediapipe merupakan suatu *framework* yang dirancang oleh Google untuk menghasilkan audio atau video dari membangun *pipelines* untuk mengolah data persepsi. Dengan bantuan dari Mediapipe tangan yang terdapat pada citra akan dapat dideteksi. Proses pertama yaitu akan mendeteksi telapak tangan atau disebut *palm detection* agar program dapat mendeteksi keberadaan tangan. Selanjutnya yaitu mendeteksi 21 titik *keypoint* dalam tangan yang sudah terdeteksi (Nautica, 2022).



Gambar 2.1: 21 titik *keypoint* pada tangan (“MediaPipe Hands”, 2021)

2.2.3 Convolutional Neural Network (CNN)

Convolutional Neural Network atau biasa dikenal CNN merupakan salah satu dari *neural network* yang digunakan untuk memproses data berupa grid. Salah satu contoh dari data grid adalah citra. Citra dianggap sebagai grid karena berbentuk piksel 2 dimensi. CNN dianggap sebagai salah satu solusi untuk permasalahan pada bidang *Computer Vision* untuk mendeteksi dan mengenali sebuah image (Perdananto, 2019). Secara nama *Convolutional Neural Network* menunjukkan bahwa operasi matematika yang akan digunakan yaitu konvolusi dengan cara mengkalikan data 2 dimensi dengan kernel. Pada CNN akan memiliki 3 lapisan utama yaitu :

1. Convolutional Layer

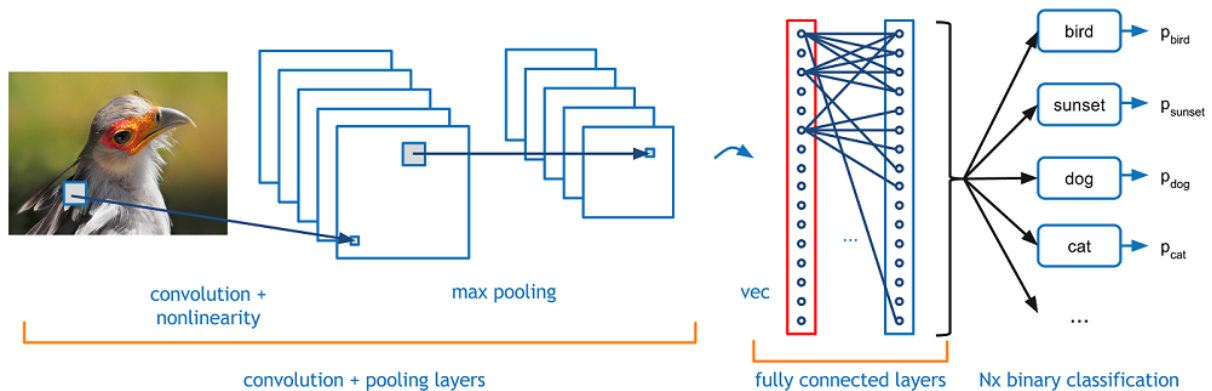
Layer ini merupakan inti dari CNN karena pada layer ini akan dilakukannya operasi konvolusi. Operasi konvolusi pada layer ini akan dilakukan antara input data dengan *filter* atau *kernel*. *Filter* akan melintasi input data dan akan melakukan operasi ”dot” antara input data dengan nilai dari *filter* sehingga akan menghasilkan *feature map* (Qolbiyatul, 2019).

2. Pooling Layer

Layer ini digunakan untuk melakukan *downsampling* yang berguna untuk mengurangi dimensi dan jumlah parameter. Terdapat 2 jenis *pooling* yang sering digunakan yaitu *max pooling* dan *average pooling*. *Max pooling* adalah mengambil citra maksimum yang dicakup oleh kernel, sedangkan pada *average pooling* adalah mengambil nilai rata-rata dari citra yang dicakup kernel. Namun tidak hanya itu *pooling layer* juga digunakan untuk mengekstraksi fitur dominan (Hidayatullah, 2021).

3. Fully Connected Layer

Layer ini digunakan untuk melakukan transformasi pada dimensi data agar dapat diklasifikasikan secara liner. Hasil dari *pooling layer* masih berbentuk *multidimensional array*, maka dari itu pada *fully connected layer* akan dilakukan *reshape* input dari *multidimensional array* menjadi *vector*. Setiap *neuron* pada *convolutional layer* akan ditransformasikan menjadi data satu dimensi sebelum dapat dimasukkan dalam sebuah *fully connected layer* (Wijaya et al., 2016).



Gambar 2.2: Arsitektur CNN (Desphande, 2016)

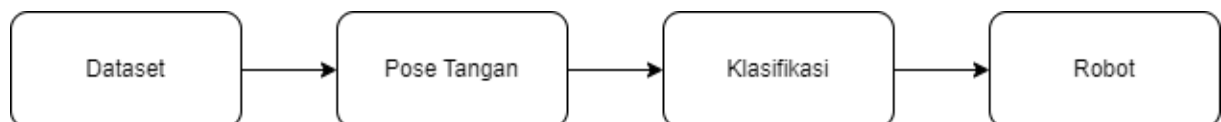
3 METODOLOGI

3.1 Metode yang Digunakan

Pada penelitian ini nantinya akan terdiri dari 2 langkah utama yaitu perancangan pada perangkat lunak (Software) dan perancangan pada perangkat keras (Hardware) :

3.1.1 Perangkat Lunak

Pada tahap ini akan dilakukan metode untuk membuat kendali pada mobil robot pada sisi perangkat lunak. Perangkat lunak ini akan dimulai dari pembuatan dataset dilanjutkan ke pose. Setelah pose akan dilanjutkan kepada tahapan klasifikasi dimana pada tahapan ini akan dilakukan pembelajaran mesin untuk dapat mengenali pose yang diberikan. Setelah tahap klasifikasi akan dilakukan tahap pembuatan program pada robot untuk dapat memahami data yang diterima dari tahap klasifikasi.



Gambar 3.1: Blok diagram perangkat lunak

1. Dataset

Pada tahapan dataset disini akan dilakukan mulai dari pengambilan data-data yang disini nantinya akan berupa gambar/citra sampai dengan menyeleksi citra tersebut dan siap melanjutkan ke tahapan selanjutnya.



Gambar 3.2: Blok diagram pembuatan dataset

(a) Input Citra

Pada dasarnya untuk mendeteksi suatu pose atau gerakan dari tangan membutuhkan untuk menemukan bagian tangan pada setiap Frame dan kemudian menganalisa pose atau gerakan tangan tersebut. Kamera digunakan untuk mendapatkan pose dari tangan pada setiap frame yang akan digunakan oleh user sebagai input untuk menentukan perintah. Dari input kamera ini yang nantinya akan digunakan untuk mendeteksi adanya tangan pada frame menggunakan Mediapipe.

(b) Deteksi Tangan

Untuk mendeteksi tangan pada frame-frame yang telah ditangkap oleh kamera akan digunakan *framework* Mediapipe. Mendeteksi tangan akan dimulai dari mendeteksi telapak tangan yang disebut sebagai *palm detection model*. Setelah melakukan pendeteksian telapak tangan pada keseluruhan gambar maka berikutnya yaitu mendeteksi 21 titik *keypoint* dalam wilayah tangan menggunakan regresi. Setelah 21 titik *keypoint* terdeteksi maka langkah selanjutnya yaitu menghubungkan titik-titik yang telah terdeteksi sehingga dapat membentuk seperti rangka tangan yang ditunjukkan oleh Gambar 2.1.

(c) Ekstraksi Citra

Tangan yang telah terdeteksi oleh mediapipe dan telah terdapat rangkanya pada setiap frame akan disimpan untuk menjadi dataset. Citra yang disimpan akan ada 2 macam yaitu citra yang ditangkap oleh kamera dan terdapat ditangannya dan juga terdapat citra dengan latar berwarna hitam dengan rangka tangan dari mediapipe didalamnya. Citra berwarna hitam sebelum disimpan akan dipotong sesuai luas dari kotak yang diambil dari koordinat terkecil dari 21 titik mediapipe. Citra yang telah dipotong akan di ubah ukurannya.

2. Pose Tangan



Gambar 3.3: Pose tangan yang digunakan

Pose tangan yang akan digunakan nantinya akan ada 5 pose yang akan sebagai simbol dari diam, maju, mundur, belok kanan, dan belok kiri. Diam akan disimbolkan dengan membuka telapak tangan serta meluruskan semua jari tangan. Maju akan disimbolkan dengan tangan mengepal. Mundur akan disimbolkan dengan meluruskan jari telunjuk, tengah, dan manis namun mengepalkan ibu jari dan jari kelingking. Belok kanan akan disimbolkan dengan meluruskan ibu jari dan mengepalkan semua jari kecuali ibu jari.

Belok kiri akan disimbolkan dengan meluruskan jari kelingking dan mengepalkan semua jari kecuali jari kelingking.

3. Klasifikasi

Data yang sudah didapatkan nantinya akan di inputkan ke dalam perhitungan menggunakan machine learning. Data yang telah diinputkan nantinya akan dilabeli untuk diklasifikasikan dan digunakan untuk memberikan action selanjutnya. Dari hasil kalsifikasi nantinya akan disimpan sebagai model yang nantinya akan digunakan untuk mendeteksi citra yang akan datang. Proses kalsifikasi akan menggunakan CNN dalam melakukan training dan akan menggunakan citra yang sudah disimpan dari tahapan dataset.

4. Robot

Dari data yang telah di klasifikasikan dan sudah didapatkan posenya maka pose tersebut akan diterjemahkan ke dalam suatu perintah untuk dapat menggerakkan mobil robot. Robot akan menerima data dari hasil klasifikasi dan selanjutnya akan bergerak sesuai data yang diberikan. Gerakan robot akan diatur melalui arah dan kecepatan putaran pada roda. Robot akan berjalan maju jikalau arah dan kecepatan putaran roda bernilai sama dan akan berbelok bilah salah satu dari arah dan kecepatan putaran berbeda.

3.1.2 Perangkat Keras

Pada tahap ini akan dilakukan metode untuk membuat kendali pada robot dari sisi perangkat keras. Perangkat yang akan digunakan yaitu Laptop, modul *bluetooth*, dan robot.



Gambar 3.4: Blok diagram perangkat keras

1. Laptop

Laptop disini nantinya akan digunakan untuk menjalankan program perangkat lunak yang telah dikembangkan untuk dapat mengklasifikasikan tangan yang terdeteksi dan juga kamera yang terdapat pada laptop ini juga yang nantinya akan digunakan sebagai input gambar/citra.

2. *Bluetooth*

Hasil klasifikasi yang ada pada laptop akan dikirim ke mobil robot untuk dapat memberikannya perintah sesuai dari hasil klasifikasi. Menghubungkan laptop dan robot ini akan menggunakan modul *bluetooth* yang akan dihubungkan dengan mobil robot dikarenakan mobil robot yang akan digunakan masih belum mempunyai modul *bluetooth*.

3. Robot

Pada bagian mobil robot ini terdapat beberapa komponen elektronik yang terpasang, diantaranya sebuah mikrokontroler Arduino uno dan mikrokontroler Arduino nano , 2 buah .FC-03 L298N motor driver dual, modul charger Tp 4056, dan modul bluetooth HC-05. Komponen mekanik terdapat 4 buah roda, 4 buah gearbox motor, dan 1 buah chassis.

3.2 Urutan Pelaksanaan Penelitian

Pada penelitian ini akan ada beberapa tahapan, dimulai dari pengambilan dan pemrosesan dataset hingga pada pengujian dan analisa. Berikut rencana waktu yang akan dibutuhkan dalam penelitian ini :

Tabel 3.1: Urutan pelaksanaan penelitian

Kegiatan	Minggu															
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
Pengambilan dan pemrosesan dataset																
Klasifikasi																
Perakitan mobil robot																
Pembuatan program gerakan mobil																
Pengujian dan analisa																

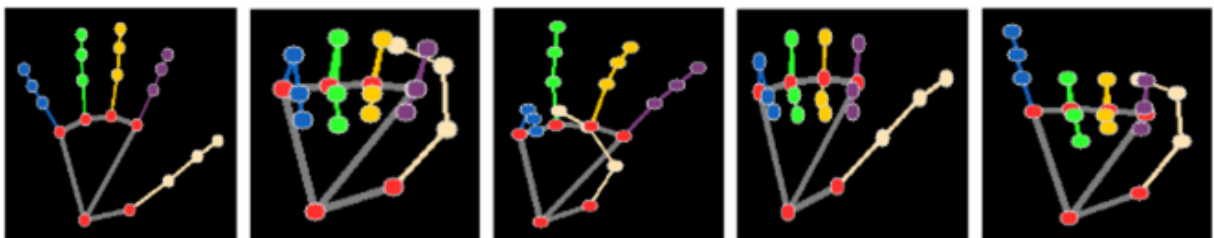
4 HASIL YANG DIHARAPKAN

4.1 Hasil yang Diharapkan dari Penelitian

Dari penelitian yang akan dilakukan, diharapkan dapat membuat kendali pada *mobile robot* yang berbasis pose tangan dengan menggunakan CNN untuk klasifikasi.

4.2 Hasil Pendahuluan

Sampai saat ini, penelitian telah berjalan sampai pembuatan dataset. Dataset yang dibuat yaitu citra berlatar gelap yang terdapat rangka tangan dari hasil mediapipe didalamnya. Sebelum disimpan citra berwarna gelap ini akan dipotong sesuai lebar dari koordinat titik terkecil dan koordinat titik terbesar dari 21 titik *keypoint* Mediapipe. Citra yang telah dipotong akan diresize menjadi 128×128 pixels.



Gambar 4.1: Hasil estimasi pose menggunakan *Mediapipe*

5 DAFTAR PUSTAKA

Desphande, A. (2016). A beginner's guide to understanding convolutional neural networks. Retrieved December 18, 2022, from <https://adeshpande3.github.io/A-Beginner's-Guide-To-Understanding-Convolutional-Neural-Networks/>

- Hidayatullah, P. (2021). *Buku sakti deep learning computer vision menggunakan yolo untuk pemula*. Stunning Vision AI Academy.
- Jaya, H. (2016). *Desain dan implementasi sistem robotika berbasis mikrokontroler*. Edukasi Mitra Grafika.
- Leksono, J. W., Samudra, A., Yannuansa, N., & Fauzi, A. (2020). Kendali mobil robot menggunakan isyarat tangan berbasis arduino. *Jurnal Electro Luceat*.
- Mediapipe hands. (2021). Retrieved December 18, 2022, from <https://google.github.io/mediapipe/solutions/hands.html#mediapipe-hands>
- Nautica, M. R. P. (2022). Hand gesture detection sebagai alat bantu ajar berhitung menggunakan mediapipe dan convolutional neural network secara realtime.
- Perdananto, A. (2019). Penerapan deep learning pada aplikasi prediksi penyakit pneumonia berbasis convolutional neural network. *ICT Vol 1 No PP001-010*.
- Qolbiyatul, L. (2019). Apa itu convolutional neural network? Retrieved December 18, 2022, from <https://medium.com/@16611110/apa-itu-convolutional-neural-network-836f70b193a4>
- Risetya, D. (2022). Human computer interaction (hci) : Pengertian, pentingnya, komponen dan ilmu yang berhubungan. <https://www.ekrut.com/media/hci-adalah>
- Sulistyo, E. (2017). Mobile robot dengan pengontrolan perintah suara berbasis android. *Jurnal Manutech Vol. 9 No 2*.
- Wibowo, A. A. (2022). Gestur tangan manusia dalam karya fotografi. *Jurnal Fotografi, Televisi, Animasi, Vol 17 No 2 Oktober 2021*.
- Wijaya, A. Y., P, I. W. S. E., & Soelaiman, R. (2016). Klasifikasi citra menggunakan convolutional neural network (cnn) pada caltech 101. *Jurnal Teknik ITS Vol 5*.