Sistem *Tracking* Posisi Kamera Menggunakan Pengolahan Citra Untuk Pemusatan Posisi Pengambilan Video di *Automation Academy*

**Joko Endrasmono1, Adam Kusumah2, Agus Khumaidi3, M. Basuki Rahmat4** *Program Studi Teknik Otomasi, Jurusan Teknik Kelistrikan Kapal  
Politeknik Perkapalan Negeri Surabaya*Jalan Teknik Kimia, Kampus ITS Sukolilo, Surabaya 60111adamkusumah@student.ppns.ac.id

*Abstract*— Kamera merupakan komponen yang sangat penting pada *Automation Academy*, dimana pada aktivitas yang dilakukan dalam perusahaan ini adalah berhubungan dengan video dan isi dari video adalah materi yang di unggah pada website *automationacademy.com*, sehingga kamera sangat mengambil peranan yang tidak kalah besarnya dengan peranan lain. Dalam pengambilan video biasanya harus ada paling sedikit 2 orang untuk menjadi *cameraman*, ini menjadi masalah karena melihat keterbatasan sumber daya manusia yang ada di *Automation Academy*. Penelitian ini bertujuan untuk meminimalisir masalah tersebut dengan membuat Sistem Tracking Posisi Kamera Menggunakan Pengolahan Citra Untuk Pemusatan Posisi Pengambilan Video di *Automation Academy*, pada penelitian ini, *framework* *MediaPipe* digunakan sebagai pengolahan citra untuk pengenalan posisi pemateri yang akan direkam menggunakan kamera DSLR. Mekanik dari Sistem *Tracking* Posisi Kamera Menggunakan Pengolahan Citra Untuk Pemusatan Posisi Pengambilan Video di *Aautomation Academy* bekerja sesuai dengan sistem yang telah direncanakan. Akurasi dari pendeteksian menggunakan *framework MediaPipe* sangat bagus, dapat terdeteksi dengan jarak antara 1,5 meter hingga 8 meter. Kemudian intensitas cahaya yang ideal adalah antara 125 lux hingga 190 lux. Lalu agar posisi kamera dapat mengikuti manusia dengan cara mengubah nilai *pixels* menjadi *pulse* dari motor *stepper.*

Kata kunci: *Automation Academy, video, cameraman, auto-tracking*

# Pendahuluan (*Heading 1*)

Pada era Revolusi Industri 4.0 ini, perkembangan teknologi sangatlah pesat baik di sektor industri maupun non industrinya. Di sekeliling kita banyak sekali hasil dari perkembangan teknologi, salah satunya adalah kamera. Kamera adalah alat paling populer dalam aktivitas [fotografi](https://id.wikipedia.org/wiki/Fotografi" \o "Fotografi). Nama ini didapat dari *camera obscura*, [bahasa Latin](https://id.wikipedia.org/wiki/Bahasa_Latin" \o "Bahasa Latin) untuk "ruang gelap", yang ditemukan oleh Al Haitam atau Alhazen sekitar tahun 1000 Masehi, mekanisme awal untuk memproyeksikan tampilan di mana suatu ruangan berfungsi seperti cara kerja kamera fotografis yang modern, kecuali tidak ada cara pada waktu itu untuk mencatat tampilan gambarnya selain secara manual mengikuti jejaknya (Wikipedia). Fungsi kamera juga bisa untuk mengambil objek berupa video, yang mana hasil dari objek tersebut di tampilkan tidak hanya berupa gambar yang diam, namun gambar yang bergerak sesuai durasi yang diambil.

Saat ini kamera tidak hanya digunakan untuk keperluan mengabadikan momen langka, pada sektor pembelajaran juga kamera adalah hal yang sangat penting, terlebih saat ini dunia sedang menghadapi pandemi *COVID-19*, pembelajaran dalam jaringan (daring) yang mana peserta didik dan juga tenaga didik diharuskan untuk meneruskan kegiatan belajar mengajar di rumah masing masing. Namun untuk peserta didik vokasi yang notabene pembelajarannya di dominasi dengan pembelajaran praktek akan kesusahan jika tidak dapat praktek langsung ke sekolah atau kampus mereka, maka dari itu peran kamera sebagai visualisasi praktikum sangatlah besar sehingga peserta didik tidak sampai ketinggalan studi mereka meskipun tidak bisa belajar di sekolah maupun kampus.

Pada beberapa waktu lalu, penulis dan tim mengikuti Program Kewirausahaan Mahasiswa Vokasi (PWMV) tahun 2021, pada program tersebut tim kami mengusulkan produk yang bernama *Automation Academy*.

*Automation Academy* adalah *platform* media ajar yang tersedia pada *website automationacademy.co.id*, di situs tersebut para peserta yang telah mendaftar dapat mengakses dan membeli produk bahan ajar yang telah disediakan oleh tenaga didik *professional*. Banyak dari materi yang di unggah di situs ini menggunakan video, yang mana penggunaan video sebagai konsep dari pembelajaran di *Automation Academy* adalah salah satu cara yang efektif digunakan untuk memaksimalkan proses belajar mengajar.

Pada proses pengembangannya, dikarenakan bisnis ini masih pemula, maka untuk mengurangi biaya yang dibutuhkan untuk menyewa juru kamera, pada penelitian mengenai Sistem Tracking Posisi Kamera Menggunakan Pengolahan Citra Untuk Pemusatan Posisi Pengambilan Video di *Automation Academy* dilakukan untuk mengatasi masalah tersebut. Dengan mengusung konsep pengambilan video dengan cara *auto tracking* menggunakan metode *MediaPipe*, penelitian ini diharapkan dapat memaksimalkan kinerja dari alat dan juga dapat membantu untuk mengurangi biaya yang dibutuhkan untuk menyewa juru kamera.

.

# Penelitian

## Metode yang Digunakan

Dalam Tugas Akhir ini, peneliti menggunakan *Framework MediaPipe* sebagai pengolahan citra untuk pergerakan kamera otomatis.

## MediaPipe

*MediaPipe* adalah *framework* yang digunakan untuk menerapkan dalam alur *machine learning*, dan ini merupakan framework sumber terbuka *Google. Framework MediaPipe* berguna untuk pengembangan lintas platform karena *framework* dibangun menggunakan *time series data. Framework* dari *MediaPipe* adalah multimodal, dimana *framework* ini dapat diterapkan ke berbagai audio dan video. *Framework MediaPipe* yang digunakan oleh pengembang untuk membangun dan menganalisis sistem melalui grafik, dan juga telah digunakan untuk mengembangkan sistem untuk tujuan aplikasi. Langkah-langkah yang terlibat dalam sistem yang menggunakan *MediaPipe* dilakukan dalam *pipeline configuration. Pipeline* yang dibuat bisa dapat berjalan di berbagai platform yang memungkinkan skalabilitas di seluler dan desktop. *Framework MediaPipe* didasarkan pada tiga bagian dasar, mereka adalah evaluasi kinerja, *framework* untuk mengambil data sensor, dan kumpulan komponen yang disebut kalkulator dan dapat digunakan kembali. Pipeline adalah grafik yang terdiri dari komponen yang disebut kalkulator, setiap kalkulator dihubungkan oleh aliran dimana paket data mengalir. Pengembang dapat mengganti atau menentukan kalkulator khusus dimana saja dalam grafik yang membuat aplikasi mereka sendiri. Kalkulator dan aliran digabungkan membuat diagram aliran data, grafik dibuat dengan MediaPipe dimana setiap node adalah kalkulator dan node dihubungkan oleh aliran (Shriram et al., 2021).

Gambar Landmarks/Keypoints MediaPipe Holistic

Teori dari penelitian sebelumnya mengenai *MediaPipe*, menurut (Anilkumar et al., 2021) dengan judul “*Pose Estimated Yoga Monitoring System*”, gambar input diumpankan ke perpustakaan *MediaPipe* untuk deteksi *keypoint* dari tubuh pengguna. Outputnya adalah daftar koordinat pada sumbu X, Y dan Z untuk 33 titik kunci utama tubuh manusia. Daftar koordinat ini menentukan lokasi setiap bagian tubuh utama dalam gambar masukan. Dengan menggunakan koordinat ini, dapat membangun orientasi kerangka pengguna yang akurat. Pada Gambar 2.1, penanda menunjukkan sendi utama dan lokasi pada tubuh manusia. Mereka diindeks dari 0 hingga 32 untuk menunjukkan total 33 *landmarks* yang dihasilkan dari *library* *MediaPipe*. 11 penanda pertama dari 0 hingga 10 digunakan untuk prosedur penandaan wajah. Dengan menggunakan *landmarks* atau *key points* ini, kita dapat mendeteksi wajah dalam gambar serta orientasinya. 11 *landmarks* berikutnya dari 11 hingga 22 digunakan dari deteksi tubuh bagian atas. Tubuh bagian atas meliputi bahu, siku, pergelangan tangan, tangan dan perkiraan 3 jari yaitu jari kelingking, jari telunjuk dan ibu jari pada kedua tangan. 11 poin *landmarks* terakhir dari 23 hingga 32 digunakan untuk menentukan tubuh bagian bawah yang terdiri dari pinggul, lutut, tungkai, dan kaki. Mereka bersama-sama memberikan perkiraan tidak hanya struktur tubuh manusia dalam gambar tetapi juga orientasi tubuh dalam ruang 3D.



# Hasil dan Pembahasan

## Pengujian Keakuratan Deteksi Manusia

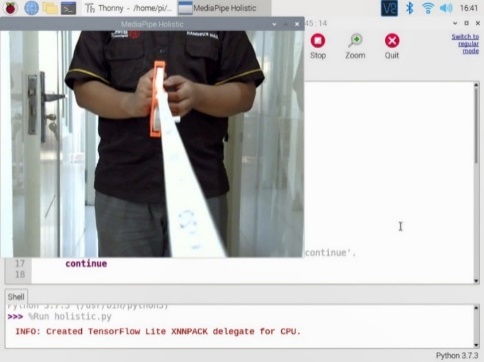
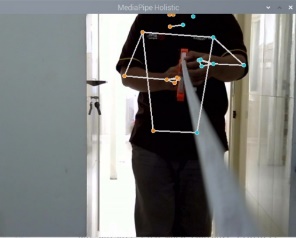
Pengujian *framework* yang akan dibahas adalah *framework* *MediaPipe Holistic*, dimana pada pengujian ini untuk mengetahui keakuratan deteksi dari *framework* tersebut. Indikator dari apakah *framework* berhasil mendeteksi adalah munculnya *landmarks* dan garis penghubung antar *landmarks* pada tubuh manusia berdasarkan posisi tubuh.

.

| No. | Pengujian | Posisi | Keterangan |
| --- | --- | --- | --- |
| 1 |  | 90° | Terdeteksi |
| 2 |  | 135° | Terdeteksi |
| 3 |  | 180° | Terdeteksi |
| 4 |  | 225° | Terdeteksi |
| 5 |  | 270° | Terdeteksi |
| 6 |  | 315° | Terdeteksi |
| 7 |  | 360° | Terdeteksi |
| 8 |  | 45° | Terdeteksi |

Tabel Pengujian Keakuratan Deteksi Manusia

## Pengujian Jangkauan Deteksi

Pada tahap ini dilakukan pengujian jangkauan deteksi *framework MediaPipe* dalam mendeteksi manusia. Hasil pengujian yang didapatkan menggunakan *roll meter* adalah *framework* dapat mendeteksi dari jarak 1.5 meter hingga 8 meter

Gambar pemberian anotasi

Gambar Pengujian Jarak 1 Meter

Gambar Pengujian Jarak 1.5 Meter

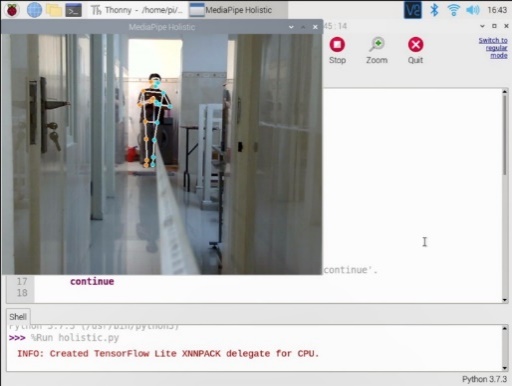
## Proses Training

1. Pengumpulan *dataset*

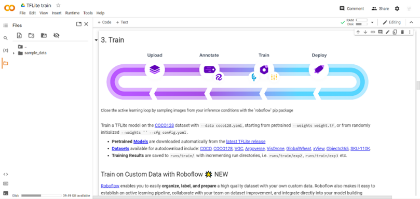
Pengumpulan *dataset* dilakukan dengan mengumpulkan gambar yang berisi objek manusia dengan berbagai gestur. Dataset yang digunakan untuk *MediaPipe Holistic* menggunakan *cocodataset keypoint*. Dataset tersebut dapat di unduh di <https://cocodataset.org/#keypoints-2020>.

Gambar website www.cocodataset.com

1. Anotasi

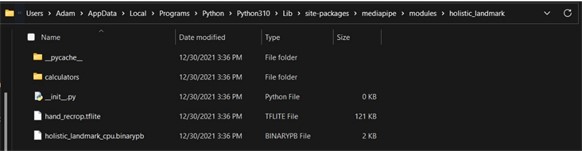
 Anotasi adalah proses memberikan *ROI (Region of Interest)* dan label dari objek yang akan dideteksi. Dalam hal ini anotasi dilakukan dengan menentukan titik-titik *landmark* pada objek.

1. *Training*

 *Training* adalah proses pembuatan model yang cocok untuk pendeteksian objek yang telah diberikan pada proses sebelumnya. Proses *training* dilakukan menggunakan *Google Colab.*

Gambar training dataset

Gambar Pengujian Jarak 8 Meter

Proses *training* yang dilakukan memakan waktu sebanyak 500 epoch. Waktu tersebut tergantung pada banyaknya dataset yang digunakan. Hasil dari *training* berupa file yang berekstensi (.tflite) yang digunakan untuk mendeteksi. File tersebut di *rename* ke *hand\_recrop.tflite* dan simpan kedalam *folder* pada Gambar dibawah ini.

Gambar folder penyimpanan dataset

## Proses Deteksi dan Gerakan

Pada tahap ini dilakukan penentuan perintah arah gerak kamera dan kondisi berjalan atau berhentinya motor *stepper*. Penentuan tersebut berdasarkan area aman.

Gambar Index 0 kembali masuk frame

Jika titik acuan berada diluar area aman maka harus bergerak. Jika berada di kanan area aman, maka arah gerak ke kanan, begitu juga dengan yang lain. Jika status bergerak, maka dilakukan perhitungan jarak terhadap titik aman dengan rumus berikut:

(1)

(2)

Dengan *d* adalah jarak titik acuan dengan titik aman dengan satuan piksel. Setelah didapatkan nilai *d*, maka dikonversi menjadi nilai *step* yang digunakan untuk menjalankan motor *stepper*, menggunakan rumus berikut:

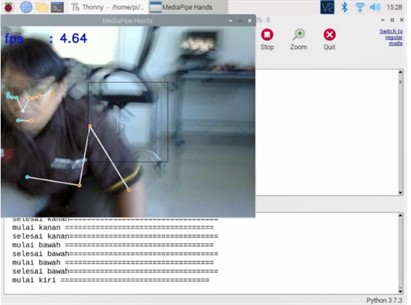
(3)

B (4)

Dengan *step* adalah nilai step yang digunakan untuk gerakan motor *stepper*. *Step A* putaran motor *stepper* horizontal, dan *step B* adalah putaran motor *stepper* vertical, dan *k* adalah nilai konstanta yang merupakan jumlah step yang mewakili satu piksel. Nilai *k* didapatkan dari membandingkan jumlah *step* untuk berpindah dari sisi paling kiri ke sisi yang paling kanan dengan resolusi kamera.

(5)

(6)

Pengaplikasian dari rumus terhadap alat didokumentasikan pada Gambar dibawah ini.

# KESIMPULAN DAN SARAN

Dari hasil pengujian yang telah dilakukan dapat Berdasarkan penelitian diatas dapat diambil kesimpulan sebagai berikut:

1. Dari hasil pengujian dapat diketahui bahwa posisi terbaik untuk meningkatkan akurasi sistem deteksi manusia dengan menggunakan framework MediaPipe yaitu dengan jarak antara 1,5 meter hingga 8 meter. Kemudian intensitas cahaya yang ideal adalah antara 125 lux hingga 190 lux. Dan framework MediaPipe dapat mendeteksi seluruh posisi tubuh mulai dari 0° hingga 360°.
2. Dari hasil pengujian peforma *framework MediaPipe Holistic*, dapat disimpulkan bahwa beberapa bagian tubuh memiliki *keypoints* yang saling berkesinambungan, artinya saat hanya beberapa *keypoints* saja yang dideteksi pada bagian tubuh tertentu, maka hasilnya tidak terdeteksi manusia.
3. Dari hasil pengujian *framework MediaPipe* dapat diketahui bahwa agar posisi kamera dapat mengikuti manusia dengan cara mengubah nilai *pixels* menjadi *pulse* dari motor *stepper*.

##### Daftar Pustaka

1. Shriram, S., Nagaraj, B., Jaya, J., Shankar, S., & Ajay, P. (2021). Deep Learning-Based Real-Time AI Virtual Mouse System Using Computer Vision to Avoid COVID-19
2. Anilkumar, A., K.T., A., Sajan, S., & K.A., S. (2021). Pose Estimated Yoga Monitoring System. *SSRN Electronic Journal*, *Icicnis*, 1–8. https://doi.org/10.2139/ssrn.3882498
3. Bazarevsky, V., & Grishchenko, I. (2020). *On-device, Real-time Body Pose Tracking with MediaPipe BlazePose*. Google AI. https://ai.googleblog.com/2020/08/on-device-real-time-body-pose-tracking.html
4. MediaPipe. (2020). *MediaPipe*. MediaPipe. www.mediapipe.dev
5. Wahyudi, Z. N. (2020). Rancang bangun monitoring inventaris barang menggunakan motor stepper nema 17.
6. Comaro, J. (2020). PERANCANGAN DAN PENGEMBANGAN ALAT UJI TARIK MINI BERBASIS ARDUINO UNTUK SPESIMEN NON-FERRO. 7–29.

Gambar Index 0 keluar dari frame