

Pemanfaatan *Mediapipe Body Pose Estimation* dan *Dynamic Time Warping* untuk Pembelajaran Tari Remo

Yusuf Effendi¹, Yosi Kristian², Lukman Zaman P. C. S. W³, Hariadi Yutanto⁴

^{1,2,3}Informatic Department, Institut Sains dan Teknologi Terpadu Surabaya

⁴ Informatic Department, Universitas Hayam Wuruk Perbanas Surabaya

Article Info

Article History

Received: 23-06-2023

Revises : 04-12-2023

Accepted: 19-12-2023

Keywords

Mediapipe;

Dtw;

Tari Remo;

Computer Vision

✉ Corresponding Author

Yusuf Effendi,

Institut Sains dan Teknologi

Terpadu Surabaya,

Tel. +62 85736262787

yusuf.effendi39@gmail.com

ABSTRACT

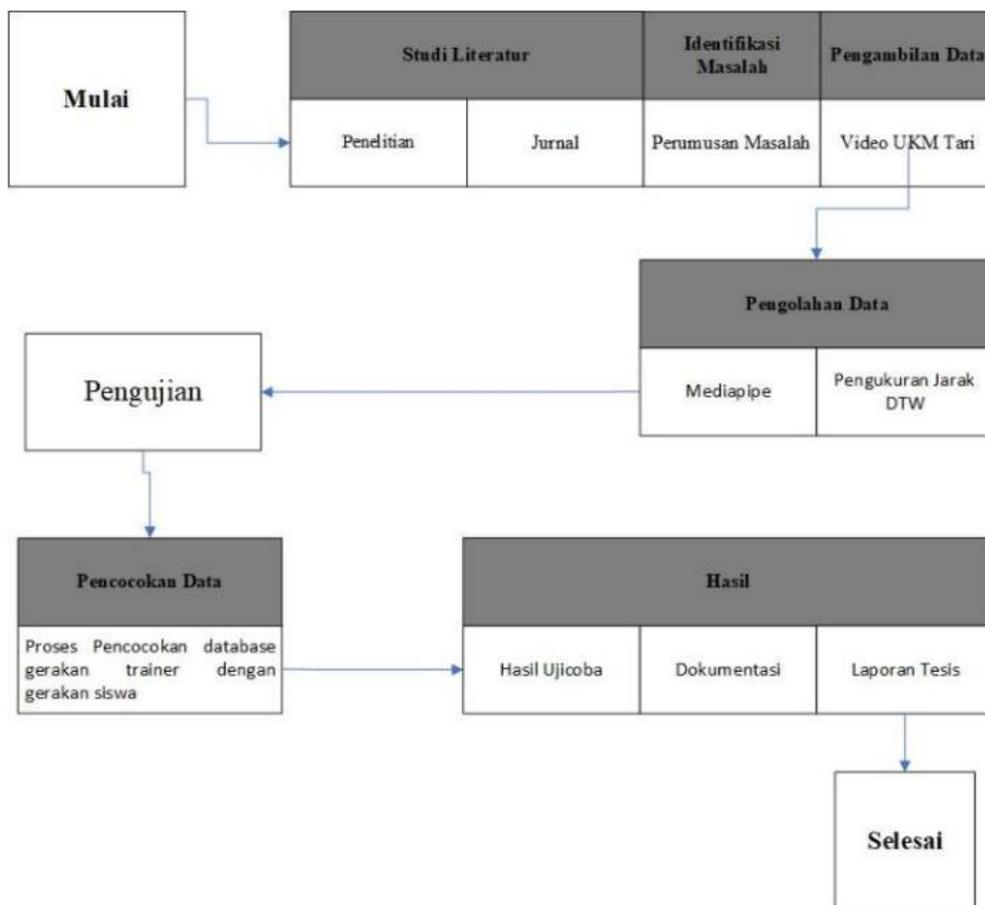
Video can be used as a learning medium for various purposes. In this research, the object of study is the movements of the traditional dance "remo." Thus, as a substitute for an absent coach or instructor, videos can take on the role of a dance instructor. Therefore, the author aims to create a system to assist coaches in correcting the dance movements of their learners. Using the MediaPipe module and the Dynamic Time Warping algorithm, the author developed a system to correct the learners' dance movements. This system can detect deviations from the coach's instructional video and provide notifications about which body angles do not match the coach's video instructions. The system operates by having users upload a video of their dance movements, and then it identifies which movements deviate from the correct remo dance. The accuracy is measured by comparing the angle distances between the master's movements and the test data. If the angle exceeds a predetermined threshold, the movement is considered incorrect. The system's output is validated by the coach, and it achieves 90% accuracy in identifying movement errors in videos. With this accuracy, the system can assist coaches in evaluating their learners' remo dance movements.

PENDAHULUAN

Kesenian tari remo pada dasarnya mempunyai keistimewaan yaitu para penarinya menggunakan gelang lonceng pada kakinya[1][2]. Gerakan tari yang rumit membutuhkan latihan yang disiplin untuk menguasai gerakannya. Terkadang kehadiran pelatih dibutuhkan untuk mengarahkan Gerakan yang benar. Namun ketika anak didiknya ingin belajar secara mandiri, mereka menemukan kendala yaitu tidak adanya instruktur yang bisa mengevaluasi gerakannya. Karena instruktur hanya bisa datang pada saat jadwalnya saja. Fleksibilitas waktu pelatih yang menjadikan kendala bagi anak didik yang masih belajar diawal.

Hal ini sama halnya dengan virtual *gym Instructor*, penelitian menyebutkan peran dan tugas utama instruktur adalah memotivasi, membimbing dan memberikan feedback untuk peserta didiknya [3]. Padatnya jadwal yang menyusahkan instruktur untuk memberikan pengarahan kepada anak didiknya. Dan dari sudut pandang anak didik jg merasa tidak tahu gerakan yang di tiru dari video instruktur itu sudah benar atau tidak. Hal ini yang mendasari peneliti untuk membuat aplikasi yang bisa mendeteksi gerakan dari seseorang sudah benar atau tidak. Dengan video yang direkam oleh peserta didik, nantinya akan diproses oleh computer dan nantinya akan direkomendasikan oleh *soft computing* bahwa Gerakan tersebut sudah benar atau tidak. Dengan aplikasi ini dapat menggantikan peran kehadiran instruktur tari remo dalam mengoreksi gerakan peserta didiknya.

METODE



Gambar 1. Metodologi Penelitian

Dalam alur tersebut, dimulai dari peneliti mencari literatur atau referensi untuk menunjang penelitian saat ini. Setelah mencari literatur, merumuskan masalah yang telah dijabarkan pada latar belakang penelitian ini. Setelah itu baru mencari data untuk dijadikan data training. Data tersebut diambil dari unit organisasi mahasiswa (ormawa) tari di universitas Hayam Wuruk Perbanas Surabaya. Data tersebut mendapatkan data *training*, data tersebut diolah untuk menentukan data master gerakan tari yang benar. Setelah itu melakukan uji coba dengan video gerakan dari orang yang akan belajar gerakan tari remo. Dari video orang latihan tersebut, nantinya akan ada notifikasi atau label yang memberikan informasi bahwa gerakan tersebut sudah sesuai atau tidak. Adapun langkah-langkah yang digunakan peneliti dalam mengerjakan penelitian ini adalah sebagai berikut:

Teknik Pengambilan Data

Dalam pengambilan data ini, perlu berkoordinasi dengan Unit Kegiatan Mahasiswa (UKM) Tari Universitas Hayam Wuruk Perbanas. Penulis mengambil video mahasiswa dan mahasiswa yang melakukan tari remo. Dengan data mahasiswa berjumlah 4 orang sedangkan mahasiswi berjumlah 6 orang. Masing-masing orang melakukan gerakan tari remo, dengan 6 gerakan meliputi gerakan tari sembahan, agem, tanjak, iket, ngelawung, dan sabetan. Masing-masing gerakan diambil sebanyak 3 kali *take* video. Dan masing-masing orang juga melakukan gerakan yang benar dan gerakan yang dibuat salah.

Tabel 1. Gerakan Tari Remo

No.	Nama Gerakan	Gambar
1	Sembahan	
2	Agem	
3	Tanjak	
4	Iket	
5	Ngelawung	
6	Sabetan	

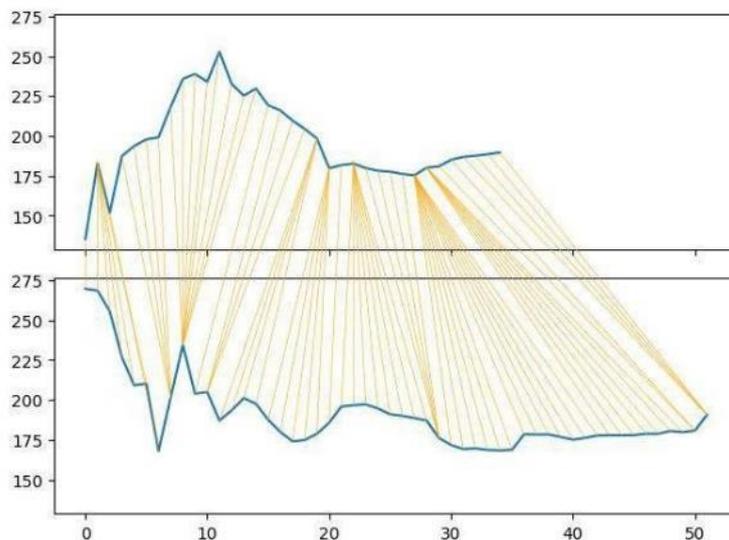
Teknik Pengolahan Data

Dalam pengolahan data ini menggunakan modul *Mediapipe* untuk menentukan titik *skeleton*. *Mediapipe* merupakan *framework* untuk membangun *machine learning* dalam lintas platform[4]. Pada saat ini penelitian tentang *Human pose estimation* sedang ramai dibicarakan diseluruh dunia[5][6]. Modul *mediapipe* ini dapat membaca titik-titik sudut yang ada di tubuh manusia sebanyak 32 titik sudut seperti gambar 2.



Gambar 2. *Skeleton*

Gambar 2 merupakan hasil dari modul *mediapipe*. Setiap frame video akan dideteksi titik skeletonnya[7]. Sehingga titik *skeleton* tersebut dapat di ukur sudut dari tiap-tiap gerakan. Sudut tersebut di tampung dalam *variable array* dari masing masing titik sudut *skeleton*. Setelah itu baru menentukan 3 video sebagai master pada masing-masing gerakan. 3 video master itu diujikan dengan semua video gerakan benar pada masing-masing gerakan. Menurut Halder A dalam penelitiannya yang tentang pengenalan bahasa isyarat menggunakan *framework mediapipe* dan deteksi akurat *real time* menggunakan *Support Vector Machine (SVM)* dapat digunakan dengan mudah[8]. Sehingga *framework mediapipe* sangat cocok untuk mendeteksi *skeleton* tubuh manusia. Jika sudah dideteksi *skeletonnya*, pada masing-masing gerakan dapat dihitung jarak DTW nya.



Gambar 3. Jarak DTW Antara Gerakan Master (8.mov) dengan Gerakan Benar (1.mov) pada Sudut Siku Kiri

Pada gambar 3 merupakan hasil dari pengukuran jarak DTW (*Dynamic Time Warping*) untuk sudut siku kiri, Dari pengukuran tersebut gerakan master dengan file 9.mov diukur dengan gerakan benar dengan nama file 1.mov. Pada terapannya DTW dapat dipake ketika obyek data menggunakan data file video mentah. Namun pada kenyataanya DTW juga dapat dipraktekkan untuk data yang secara *real time*[9][10]. Jarak DTW nya dirata-rata pada masing-masing sudut, sehinga dapat dijadikan acuan untuk menentukan jarak DTW pada gerakan yang benar. Jarak tersebut dijadikan acuan untuk menentukan gerakan yang diujikan salah atau benar dapat dilihat pada Tabel 2.

Tabel 2. Video Master Sembahan yang Diujikan dengan Gerakan Benar Gerakan Sembahan

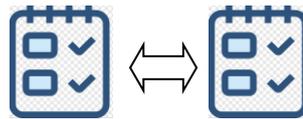
File	Lutut Kanan	Lutut Kiri	Siku Tangan Kanan	Siku Tangan Kiri	Ketiak Kanan	Ketiak Kiri	Pinggang Kanan	Pinggang Kiri
9.mov	269,48	50,89	327,84	272,79	1145,36	1145,36	302,88	302,88
10.mov	217,86	31,14	345,77	310,69	329,83	329,83	50,29	50,29
11.mov	191,79	41,66	388,15	234,44	97,68	97,68	51,13	51,13

Selanjutnya untuk pengujian sistem, dicoba menginputkan gerakan yang dianggap yang dengan sengaja disalahkan. Video tersebut terlebih dahulu dihitung DTW nya dengan data video master. Setelah itu, dari 3 video master tersbeut, jika lebih dari 1 video yang menghasikan jarak DTW diatas jarak DTW yang benar. Maka sudut dari video yang diujikan dianggap salah oleh sistem. Pada penelitian terdahulu yang dilakukan oleh candra dinata, ketika melakukan

penelitian menggunakan Dynamic Time Warping pada pengujian kata memiliki akurasi 95.85% dan untuk pengujian pada kalimat memiliki akurasi 94% [11]. Lain halnya dengan Tiar Anggita yang melakukan pengujian pada bahasa isyarat secara *real time* menggunakan DTW memiliki akurasi sebesar 93% [10]. Sedangkan pada penelitian terdahulu yang ditulis oleh Ahmed W menunjukkan akurasi 90% dalam mengenali 24 *gesture* berdasarkan bahasa isyarat India [12]. Dari ketiga penelitian tersebut, penelitian yang menghitung jarak *sequence* lebih cocok menggunakan algoritma *Dynamic Time Warping*. Akan tetapi untuk penelitian menggunakan *computer vision* akan dilakukn pengujian pada penelitian saat ini

Pengujian

Pada pengujian ini membandingkan hasil yang ditulis oleh sistem dengan penilaian dari pakar atau pelatih tari. Jika sistem dan pakar memberikan penilaian yang sama, maka sistem dapat dikatakan sudah benar.



Gambar 4. Perbandingan Hasil Sistem Dengan Hasil Dari Pakar

Pada gambar 4 merupakan serangkaian proses untuk pengujian sistem. Dimana pengujian ini diperlukan untuk memvalidasi proses pembacaan sistem. Apakah sistem sudah dapat memberikan notifikasi dengan benar atau tidak. Untuk hasil pengujiannya dapat dilihat pada Tabel 3.

Tabel 3. Pengujian Video Ujicoba yang Diukur Jarak DTWnya dengan 3 Video Master

	Lutut Kanan ▼	Lutut Kiri	Siku Tangan Kanan	Siku Tangan Kiri	Ketiak Kanan	Ketiak Kiri	Pinggang Kanan ▼	Pinggang Kiri ▼
19.mov	309,9406796	365,713259 3	509,1925372	306,8128959	888,1960 083	888,1960 083	436,5925226	436,59252 26
19.mov	451,1174644	391,064893 1	1030,033245	316,896105	957,9705 489	957,9705 489	498,6608403	498,66084 03
19.mov	444,4608014	400,203216 8	608,2235924	244,264521	788,9746 154	788,9746 154	481,4406489	481,44064 89

Tabel 3 adalah hasil pengukuran jarak DTW antara 3 video master dengan data uji coba. Hal ini dikarenakan pada dasarnya penggunaan DTW merupakan sebuah algoritma atau metode yang menghitung jalur pembengkokan optimal antara dua deretan waktu [11]. Sehingga DTW dalam riset ini digunakan dalam menghitung jarak antara gerakan master yang benar dibandingkan dengan gerakan yang disalahkan. Setelah didapat hasil jarak tersebut, selanjutnya hasil pada tiap sudut tersebut dibandingkan dengan jarak DTW master dengan gerakan benar. Jika lebih besar dari jarak yang sudah ditentukan, maka sistem akan menyimpan sebagai sudut yang salah. Dan jika dari 3 video master tersebut mengatakan lebih dari 1 video yang mengatakan salah, maka sistem akan memberikan label sudut tersebut dianggap gerakan yang tidak benar.

Selanjutnya melakukan wawancara dengan pakar atau pelatih tari untuk mengoreksi data uji coba tersebut. Pada hasil wawancara didapatkan hasil seperti pada Tabel 4.

Tabel 4. Video Master Sembahan yang Diujikan dengan Gerakan Salah Gerakan Sembahan

File	Lutut Kanan	Lutut Kiri	Siku Tangan Kanan	Siku Tangan Kiri	Ketiak Kanan	Ketiak Kiri	Pinggang Kanan	Pinggang Kiri
22.mov	Benar	Benar	Salah	Salah	Salah	Salah	Benar	Benar
23.mov	Benar	Benar	Salah	Salah	Benar	Benar	Benar	Benra
24.mov	Salah	Benar	Salah	Salah	Salah	Salah	Benar	Benar
25.mov	Salah	Benar	Salah	Salah	Salah	Salah	Salah	Salah
26.mov	Salah	Salah	Salah	Salah	Salah	Salah	Salah	Salah
27.mov	Salah	Benar	Salah	Benar	Salah	Salah	Salah	Salah
28.mov	Salah	Benar	Salah	Benar	Salah	Salah	Salah	Salah
29.mov	Salah	Benar	Salah	Benar	Benar	Benar	Benar	Benar
30.mov	Salah	Benar	Salah	Salah	Salah	Salah	Salah	Salah
31.mov	Salah	Benar	Salah	Salah	Salah	Salah	Salah	Salah
32.mov	Salah	Benar	Salah	Salah	Benar	Benar	Salah	Salah
33.mov	Salah	Benar	Salah	Salah	Salah	Salah	Salah	Salah
34.mov	Salah	Benar	Salah	Salah	Salah	Salah	Salah	Salah
35.mov	Salah	Benar	Salah	Salah	Salah	Salah	Salah	Salah
36.mov	Salah	Benar	Salah	Salah	Salah	Salah	Salah	Salah

Tabel 4 merupakan wawancara dengan pakar tari untuk mengkoreksi video uji coba. Dalam wawancara tersebut dihasilkan sudut sudut mana yang dinilai kurang benar gerakannya. Dan sudut sudut mana yang masih dianggap batas toleransi. Setelah itu hasil wawancara dicocokkan dengan hasil yang sudah dihasilkan oleh sistem. Pencocokan tersebut dilakukan pada masing-masing sudut. Masing-masing sudut dilakukan pengecekan dan dibandingkan dengan hasil wawancara. Untuk contoh pencocokan dapat dilihat pada Tabel 5.

Tabel 5. Proses Pencocokan Hasil Sistem dengan Hasil Wawancara

File	Lutut Kanan dari sistem	Hasil Lutut Kanan Dari Pakar	Hasil Pengujian Lutut Kanan
22.mov	Benar	Benar	Benar
23.mov	Benar	Benar	Benar
24.mov	Salah	Salah	Benar
25.mov	Salah	Salah	Benar
26.mov	Salah	Salah	Benar
27.mov	Salah	Salah	Benar
28.mov	Salah	Salah	Benar
29.mov	Salah	Salah	Benar
30.mov	Salah	Salah	Benar
31.mov	Salah	Salah	Benar
32.mov	Salah	Salah	Benar
33.mov	Salah	Salah	Benar
34.mov	Salah	Benar	Salah
35.mov	Salah	Salah	Benar
36.mov	Salah	Salah	Salah

Pada tabel 5 menunjukkan bahwa file video dengan nama file 27.mov diproses oleh sistem menunjukkan bahwa gerakan sembah dengan melihat lutut kanan, sistem menilai sudut tersebut tidak benar. Sedangkan menurut pakar tari juga sama bahwa di sudut lutut kanan dianggap kurang sesuai gerakannya. Pada tabel 6 dibawah ini menunjukkan rule yang dibuat peneliti dalam melakukan validasi sistem.

Tabel 6. Rule Validasi

Hasil Pengujian Sistem	Hasil Pengujian Wawancara	Hasil Pengujian
Salah	Salah	Benar
Benar	Benar	Benar
Benar	Salah	Salah
Salah	Benar	Salah

Pada tabel 6 menjelaskan jika sistem dan hasil wawancara menunjukkan hasil yang sama, maka hasil pengujian itu benar. Dan jika sistem dan hasil wawancara menunjukkan hasil yang berbeda, maka pengujian tersebut dianggap tidak benar.

HASIL DAN PEMBAHASAN

Hasil

Hasil dari penelitian yang dilakukan pada UKM Seni Tari yang menggunakan tari remo sebagai obyek penelitiannya. Sehingga mendapatkan data dengan rincian sebagai berikut: 6 mahasiswi dan 4 mahasiswa. Dari masing masing orang tersebut diminta untuk melakukan 6 gerakan tari remo. Dari masing-masing gerakan, setiap orang diminta untuk melakukan sebanyak 3 kali gerakan. Dengan gerakan yang benar dan gerakan yang dianggap salah. Setelah itu penulis memilih dari masing-masing gerakan tari remo dipilih gerakan yang dianggap benar dan dijadikan sebagai gerakan master. Pada gerakan master tersebut, penulis mengukur jarak DTW nya dengan gerakan tari remo yang benar.

Selanjutnya dari pengukuran jarak tersebut, membuat batas *threshold* untuk mengukur batas toleransi dari jarak DTWnya. Setelah didapatkan batas *threshold* tersebut, menghitung juga jarak DTW antara gerakan master dengan gerakan yang akan diujikan. Jika lebih kecil dari batas *threshold*, maka sistem akan memberikan label bahwa sudut dari gerakan tersebut benar. Dan begitu juga sebaliknya, jika hasil jarak DTW antara gerakan master dengan gerakan yang diujikan lebih besar dari batas *threshold*, maka sistem akan memberikan label salah.

Setelah sistem memberikan label gerakan benar atau salah, selanjutnya s memverifikasi hasil sistem dengan hasil yang diberikan oleh instruktur. Jika hasilnya sama dengan hasil di instruktur, maka penilaian dari sistem tersebut sudah benar. Dan jika hasilnya berbeda dengan hasil yang diberikan instruktur, maka penilaian sistem tersebut salah. Pada hasil penelitian ini telah menghitung total pengujian yang benar, dibandingkan dengan total pengujian keseluruhan. Hasil menunjukkan tingkat akurasi dapat mencapai 90% untuk tingkat kecocokan dengan pendapat dari pakar atau instuktur.

Pembahasan

Pada penelitian ini merupakan penelitian terbaru sampai saat penulis melakukan penelitian. Walaupun pada saat penulis mengerjakan penelitian ini masih tergolong penelitian terbaru. Namun penulis juga masih membutuhkan beberapa referensi untuk menunjang penyelesaian penelitian ini. Beberapa referensi penunjang tersebut meliputi, pengenalan tari remo sampai dengan penggunaan *mediapipe* sebagai pengenalan skeleton [4] [1].

SIMPULAN DAN SARAN

Aplikasi ini dapat digunakan untuk membantu proses latihan tari remo ketika instruktur tidak dapat hadir di tempat. Sehingga ketergantungan kehadiran instruktur untuk tari remo dapat digantikan dengan aplikasi ini. Pada saat penelitian ini penulis melakukan pengambilan video tari remo ini dilakukan oleh 4 mahasiswa dan 6 mahasiswi. Dan menggunakan algoritma DTW untuk menentukan tingkat kecocokan video gerakan dapat menggunakan algoritma DTW.

DTW ini digunakan untuk mengukur jarak antar 2 obyek berdasarkan satuan waktu. Algoritma ini dapat mengukur tingkat akurasi bisa mencapai 90% untuk mengukur gerakan tari remo. Namun nilai tersebut dapat dicapai dengan catatan, gerakan yang dilakukan di kedua

video itu adalah gerakan yang sama. Meskipun kedua video gerakan dengan durasi waktu yang berbeda, algoritma DTW ini dapat mengukur tingkat kecocokan gerakan di kedua video tersebut. Dalam pengukuran jarak sudut tersebut, posisi kamera saat mengambil video juga berpengaruh terhadap sudut titik skeleton. Sehingga posisi kamera pada saat pengambilan video bagi instruktur dan peserta didik diharapkan sama sehingga aplikasi ini bisa berjalan dengan lancar. Jadi sebaiknya posisi kamera sama ketika pengambilan video antara video instruktur dengan video peserta didik.

DAFTAR PUSTAKA

- [1] Viral Food Travel, “Mengenal Gerakan Tari Remo dan Sikap Dasarnya,” <https://kumparan.com/>, 2021. <https://kumparan.com/viral-food-travel/mengenal-gerakan-tari-remo-dan-sikap-dasarnya-1vaBKJCqPyB/full> (accessed Jun. 10, 2023).
- [2] M. Hardi, “Tari Remo dari Jawa Timur: Asal-Usul, Makna, dan Komposisinya,” <https://www.gramedia.com/>. <https://www.gramedia.com/literasi/tari-remo/> (accessed Nov. 10, 2022).
- [3] D. Brown and M. Ndleve, “Virtual Gym Instructor,” *22nd South. Africa Telecommun. Networks Appl. Conf.*, no. September, pp. 246–251, 2019.
- [4] F. Zhang *et al.*, “MediaPipe Hands: On-device Real-time Hand Tracking,” 2020, [Online]. Available: <http://arxiv.org/abs/2006.10214>
- [5] J. W. Kim, J. Y. Choi, E. J. Ha, and J. H. Choi, “Human Pose Estimation Using MediaPipe Pose and Optimization Method Based on a Humanoid Model,” *Appl. Sci.*, vol. 13, no. 4, 2023, doi: 10.3390/app13042700.
- [6] G. Dsouza, D. Maurya, and A. Patel, “Smart gym trainer using Human pose estimation,” *2020 IEEE Int. Conf. Innov. Technol. INOCON 2020*, pp. 4–7, 2020, doi: 10.1109/INOCON50539.2020.9298212.
- [7] “MediaPipe Pose,” <https://google.github.io/>. <https://google.github.io/mediapipe/solutions/pose.html> (accessed Nov. 28, 2022).
- [8] A. Halder and A. Tayade, “Real-time Vernacular Sign Language Recognition using MediaPipe and Machine Learning,” *Int. J. Res. Publ. Rev.*, no. 2, pp. 9–17, 2021, [Online]. Available: www.ijrpr.com
- [9] S. N. Budiman, S. Lestanti, S. M. Evvandri, and R. K. Putri, “Pengenalan Gestur Gerakan Jari Untuk Mengontrol Volume Di Komputer Menggunakan Library Opencv Dan Mediapipe,” *Antivirus J. Ilm. Tek. Inform.*, vol. 16, no. 2, pp. 223–232, 2022.
- [10] T. K. Anggita Wijayanti Nurul; Suciati, Nanik, “Pengenalan Bahasa Isyarat Indonesia dengan Metode Dynamic Time Warping (DTW) menggunakan Kinect 2.0,” *J. Tek. ITS*, vol. 7, no. Vol 7, No 1 (2018), pp. 199–202, 2018, [Online]. Available: <http://ejurnal.its.ac.id/index.php/teknik/article/view/28943>
- [11] C. Dinata, D. Puspitaningrum, and E. Erna, “Implementasi Teknik Dynamic Time Warping (Dtw) Pada Aplikasi Speech To Text,” *J. Tek. Inform.*, vol. 10, no. 1, pp. 49–58, 2018, doi: 10.15408/jti.v10i1.6816.
- [12] W. Ahmed, K. Chanda, and S. Mitra, “Vision based Hand Gesture Recognition using Dynamic Time Warping for Indian Sign Language,” *Proc. - 2016 Int. Conf. Inf. Sci. ICIS 2016*, pp. 120–125, 2017, doi: 10.1109/INFOSCI.2016.7845312.