

# Deteksi Gerakan *Fitness* Menggunakan *Pose Estimation* Dan *YOLOv11*

<sup>\*1</sup>Mealdi Arwintoro, <sup>2</sup>Ardi Sanjaya, <sup>3</sup>Rizky Aswi Ramadhan

<sup>1,2,3</sup> Teknik Informatika, Universitas Nusantara PGRI Kediri

E-mail: <sup>1</sup>\*aldiarwin25@gmail.com, <sup>2</sup>dersky@gmail.com, <sup>3</sup>riskyaswiramadhani@gmail.com

**Penulis Korespondens : Mealdi Arwintoro**

**Abstrak**—Penelitian ini bertujuan untuk mengembangkan sistem deteksi gerakan *Fitness* secara otomatis menggunakan metode *pose estimation* dan implementasi *YOLOv11*. Tujuan utama dari penelitian ini adalah mengidentifikasi dua kelompok gerakan utama dalam aktivitas *Fitness* yaitu latihan otot kaki dan latihan otot punggung. *Pose estimation* digunakan untuk mendeteksi titik-titik tubuh dari video gerakan *Fitness*, sedangkan *YOLOv11* digunakan untuk klasifikasi gerakan berdasarkan pola pergerakan tubuh. Dataset terdiri dari 13 jenis latihan yang dibagi menjadi dua kelompok utama: otot kaki (*cable front raise*, *cable row*, *deltoid press*, *dumbbell row*, *lat pulldown*, *t-bar row*) dan otot punggung (*hack squat*, *lunges*, *leg extension*, *leg press*, *squat*, *standing calf raise*, *sumo squat*). Hasil penelitian menunjukkan bahwa kombinasi metode ini mampu mencapai akurasi klasifikasi rata-rata di 73,3%. Sistem ini dapat dimanfaatkan sebagai alat bantu dalam memonitor dan mengevaluasi gerakan *Fitness* secara mandiri.

**Kata Kunci**— Deteksi Gerakan, *Fitness*, *Pose estimation*, *YOLOv11*

**Abstract**— This study aims to develop an automatic *Fitness* motion detection system using *pose estimation* methods and the implementation of *YOLOv11*. The main goal is to identify two major movement groups in *Fitness* activity, namely leg muscle and back muscle training. *Pose estimation* is used to detect keypoints from the video, while *YOLOv11* classifies movements based on body motion patterns. The dataset consists of 13 types of exercises categorized into two major groups: leg muscles (*cable front raise*, *cable row*, *deltoid press*, *dumbbell row*, *lat pulldown*, *t-bar row*) and back muscles (*hack squat*, *lunges*, *leg extension*, *leg press*, *squat*, *standing calf raise*, *sumo squat*). The results show that the combination of these methods can achieve classification accuracy above 73,3%. This system can be used as a supporting tool to monitor and evaluate *Fitness* exercises independently..

**Keywords**— *Fitness*, Motion Detection, *Pose estimation*, *YOLOv11*

This is an open access article under the CC BY-SA License.



## I. PENDAHULUAN

Kegiatan *Fitness* semakin diminati oleh masyarakat luas karena memberikan manfaat besar terhadap kebugaran tubuh [1]. Namun, pelaksanaan gerakan *Fitness* yang salah dapat menimbulkan cedera atau tidak memberikan hasil maksimal [2]. Oleh karena itu, diperlukan sistem yang mampu mendeteksi dan mengklasifikasikan gerakan secara otomatis.

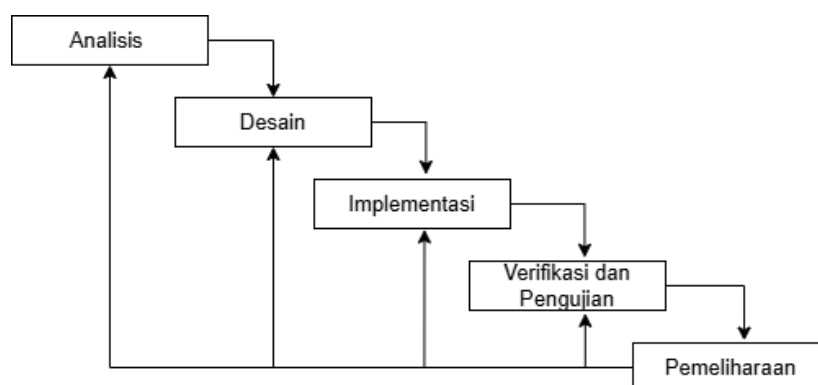
Beberapa penelitian sebelumnya telah menggunakan metode vision-based untuk mengenali gerakan tubuh [3], Zhang dkk (2023) mengembangkan sistem deteksi gerakan *fitness* menggunakan RGB-D camera dengan akurasi 89.2% untuk klasifikasi gerakan dasar seperti

push-up, sit-up, dan jumping jacks [4]. Kumar dan Sharma (2022) menerapkan MediaPipe Holistic untuk ekstraksi keypoints tubuh dan menggunakan Random Forest classifier untuk mengklasifikasikan 8 jenis gerakan fitness dengan akurasi rata-rata 85.6% [5]. Sementara itu, Rodriguez et al. (2023) mengusulkan pendekatan berbasis deep learning menggunakan 3D-CNN untuk menganalisis gerakan squat dengan fokus pada deteksi kesalahan postur, mencapai akurasi 92.1% [6]. Namun kebanyakan masih terbatas pada klasifikasi umum tanpa spesifikasi terhadap kelompok otot tertentu. Teknologi terbaru seperti *pose estimation* dan object detection memberikan peluang untuk mengembangkan sistem yang lebih akurat dan spesifik [7].

Pada penelitian ini, fokus diarahkan pada klasifikasi dua kelompok gerakan utama yaitu otot kaki dan otot punggung. Kedua kelompok ini memiliki karakteristik pergerakan yang berbeda sehingga membutuhkan pendekatan klasifikasi yang tepat. Dengan menggabungkan *MediaPipe Pose estimation* untuk ekstraksi titik-titik tubuh dan *YOLOv11* sebagai model klasifikasi [8], sistem ini diharapkan mampu mendeteksi gerakan *squat* dengan akurasi tinggi. Penelitian ini juga bertujuan memberikan kontribusi dalam pengembangan teknologi kebugaran berbasis AI [9].

## II. METODE

Metode penelitian ini dimulai dengan analisis sistem. Selanjutnya, desain sistem dilakukan dengan menjabarkan alur proses. Kemudian diimplementasikan sesuai dengan desain sistem. Setelah itu, verifikasi dan pengujian. Terakhir, pemeliharaan sistem dilakukan untuk memastikan sistem berfungsi dengan baik, sebagaimana dijelaskan pada gambar 1



Gambar 1 Alur Waterfall

Metode ini dirancang dengan terstruktur dan cermat, bertujuan untuk menghasilkan temuan yang jelas dan signifikan.

### A. Analisis Sistem

#### 1. Pengambilan data

Data berupa video yang dilakukan seorang individu yang melakukan gerakan *Fitness* dalam pengambilan video dengan jarak 2 meter antara posisi laptop dan pengguna.

#### 2. Deteksi Objek dengan *YOLOv11*

Sistem menggunakan *YOLOv11* untuk mendeteksi keberadaan objek manusia dalam video. Model ini menghasilkan *bounding box* untuk mengisolasi area tubuh

pengguna dari latar belakang, sehingga proses berikutnya dapat difokuskan pada individu yang melakukan gerakan [10].seperti pada gambar 2.



Gambar 2 Objek YOLOv11

### 3. Ekstraksi *pose* dengan *Mediapipe pose*

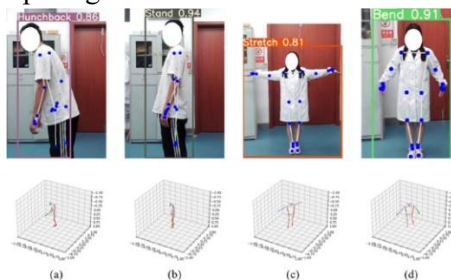
Setelah objek manusia terdeteksi, proses dilanjutkan dengan ekstraksi *pose* menggunakan *MediaPipe Pose*. Dengan memanfaatkan model *BlazePose* [11], sistem mengidentifikasi 33 titik kunci (keypoints) pada tubuh manusia, seperti pergelangan, lutut, pinggul, dan bahu, yang akan digunakan untuk analisis postur. Seperti pada gambar 3 [12].



Gambar 3 Index Keypoint pada *Mediapipe pose*

### 4. YOLOv11 dan *Mediapipe*

Menggabungkan dua teknologi utama, yaitu *YOLOv11* untuk deteksi objek (tubuh manusia) dan *MediaPipe* untuk *pose estimation* (titik-titik tubuh). Dengan melibatkan deteksi tubuh menggunakan *YOLOv11*, diikuti dengan pelacakan *pose* tubuh menggunakan *MediaPipe* [13]. Kombinasi ini memungkinkan sistem mengenali gerakan. Dapat dilihat pada gambar 4.



Gambar 4 Menggabungkan *YOLOv11* dan *Mediapipe pose*

## 5. Validasi Gerakan

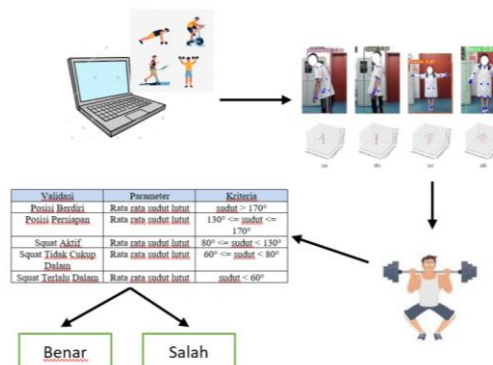
Sistem akan melakukan validasi dengan memeriksa apakah sudut tersebut memenuhi kriteria atau batasan yang telah ditentukan berdasarkan data yang telah didapatkan. Validasi ini akan menghasilkan tanda validasi yang menampilkan pesan "salah" jika tidak memenuhi kriteria dan pesan "benar" jika memenuhi kriteria. seperti yang diperlihatkan pada tabel 1.

Tabel 1 Kriteria Validasi gerakan *Squad* Berdasarkan Lutut

Validasi	Parameter	Kriteria
Posisi Berdiri	Rata rata sudut lutut	sudut > 170°
Posisi Persiapan	Rata rata sudut lutut	130° <= sudut <= 170°
<i>Squat</i> Aktif	Rata rata sudut lutut	80° <= sudut < 130°
<i>Squat</i> Tidak Cukup Dalam	Rata rata sudut lutut	60° <= sudut < 80°
<i>Squat</i> Terlalu Dalam	Rata rata sudut lutut	sudut < 60°

## 6. Tampilkan Evaluasi *Feedback* Gerakan

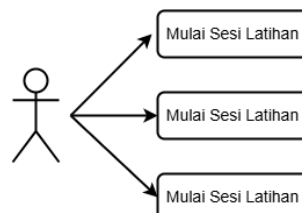
Sistem Evaluasi secara otomatis dengan menampilkan *feedback* Gerakan *Squad* yang benar. Dengan desain sistem ini, diharapkan dapat memberikan solusi yang efektif untuk memvalidasi. Proses dapat dilihat dari gambar 5.



Gambar 5 Proses sistem Validasi *Squad*

## B. Desain sistem (Arsitektur)

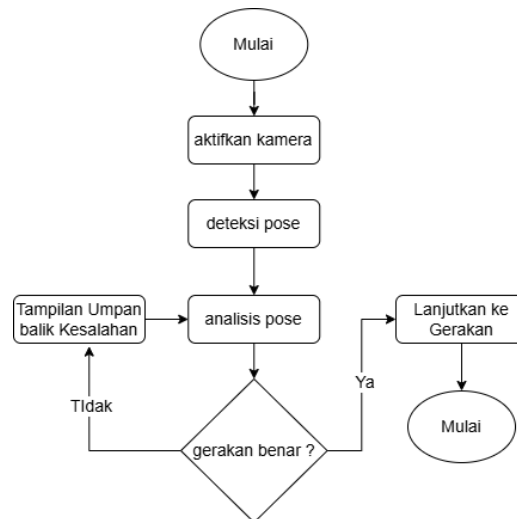
### 1. Use Case Diagram



Gambar 6 Use Case

Diagram pada gambar 6 menggambarkan interaksi pengguna dengan sistem aplikasi latihan fisik yang memiliki tiga fungsi utama. Pengguna dapat memulai sesi latihan untuk mengawali program olahraga sesuai kebutuhan mereka. Selama proses latihan berlangsung, sistem akan menampilkan berbagai gerakan *Fitness* beserta instruksi teknik yang benar untuk membantu pengguna melakukan latihan dengan optimal.

## 2. Alur Sistem



Gambar 7 Alur sistem

Alur sistem pada gambar 7 menggambarkan alur kerja sistem deteksi dan koreksi *pose* estimation gerakan olahraga yang dimulai dengan mengaktifkan kamera untuk menangkap gerakan pengguna. Sistem kemudian melakukan deteksi *pose* estimation awal untuk mengidentifikasi posisi tubuh pengguna, dilanjutkan dengan analisis mendalam terhadap *pose* estimation gerakan yang dilakukan.

## III. HASIL DAN PEMBAHASAN

Dalam Penelitian ini Memilih Gerakan *Squat* untuk Pengujian Mendalam karena *squad* gerakan dasar yang melibatkan otot punggung, paha, bokong dan kaki kesalahan dalam melakukan *Squad* bisa menimbulkan cedera serius pada lutut dan punggung [14]. sistem mendeteksi dua kategori utama yaitu *squat* dengan teknik benar dan *squat* dengan teknik salah yang mencakup berbagai jenis kesalahan seperti kedalaman kurang dari 60°, lutut terlalu jauh ke depan, dan kombinasi *error*. Data training yang digunakan berupa video dari enam partisipan dengan tingkat keahlian berbeda. Semakin banyak data training yang digunakan semakin baik hasil dari sebuah pengujian *machine learning* [15]. Maka pada penelitian ini menggunakan dataset dengan 30 video yang terdiri dari 15 video *squat* benar dan 15 video *squat* salah.

Tabel 2. Hasil Evaluasi Sistem Deteksi *Squat* pada 30 Video Dataset

Metrik Evaluasi	Nilai	Interprestasi
<i>akurasi</i>	73,3%	Sistem benar memprediksi 22 dari 30 video
<i>presisi</i>	73,3%	Dari 15 prediksi "benar", 11 memang betulan benar
<i>recall</i>	73,3%	Dari 15 <i>squat</i> benar, sistem berhasil deteksi 11
<i>spesifisitas</i>	73,3%	Dari 15 <i>squat</i> salah, sistem berhasil deteksi 11
<i>F1-Score</i>	73,3%	Balance optimal antara precision dan recall

Hasil evaluasi sistem Pada Tabel 2 deteksi *squat* menunjukkan performa yang konsisten dengan nilai *akurasi*, *precision*, *recall*, *specificity*, dan *F1-score* yang seragam sebesar 73,3%. Konsistensi ini mengindikasikan bahwa sistem memiliki balance yang baik antara kemampuan mendeteksi *squat* yang benar dan *squat* yang salah. Nilai akurasi 73,3% menunjukkan bahwa sistem berhasil mengklasifikasikan dengan benar 22 dari 30 video dalam dataset, yang merupakan performa yang cukup memuaskan untuk aplikasi *Fitness*.

Tabel 3 .*Confusion Matrik* Sistem Deteksi *Squat*

	Prediksi Benar	Prediksi Salah	Total
<i>Ground Truth Benar</i>	11 (TP)	4 (FN)	15
<i>Ground Truth Salah</i>	4 (FP)	11 (TN)	15
Total	15	15	30

Analisis Pada Tabel 3 confusion matrik mengungkapkan distribusi *error* yang seimbang dengan 4 kasus *False Positive* (FP) dan 4 kasus *False Negative* (FN). *False Positive* yang terjadi ketika sistem memprediksi *squat* salah sebagai benar berpotensi berbahaya karena dapat menyebabkan pengguna mengikuti form yang salah dan berisiko mengalami cedera. Sebaliknya, *False Negative* yang terjadi ketika sistem menolak *squat* yang sebenarnya benar dapat menyebabkan frustrasi pengguna dan menurunkan motivasi dalam belajar teknik *squat* yang tepat.

Tabel 4 .Perbandingan Performa Berdasarkan Level Keahlian

Keahlian	TP	TN	FP	FN	Total Video	Akurasi
Mahir	5	6	2	1	15	78,6%
Pemula	6	5	2	3	15	68,8%
Keseluruhan	11	11	4	4	30	73,3%

Perbandingan performa berdasarkan level keahlian menunjukkan perbedaan yang signifikan antara partisipan mahir dan pemula. Partisipan mahir mencapai akurasi 78,6% yang lebih tinggi dibandingkan partisipan pemula dengan akurasi 68,8%. Perbedaan ini dapat dijelaskan oleh

konsistensi gerakan yang lebih baik pada partisipan mahir, sedangkan partisipan pemula cenderung memiliki variasi teknik yang lebih besar dan *movement instability* yang menyulitkan sistem dalam melakukan deteksi yang akurat. Pola *error* juga berbeda dimana partisipan mahir lebih sering mengalami *False Positive*, sementara partisipan pemula lebih rentan terhadap *False Negative*

Tabel 5 Analisis Deteksi *Error* Berdasarkan Jenis Kesalahan

Jenis Kesalahan	Total Kasus	Terdeteksi	Terlewatkan	Persentase
Kedalaman <60°	6	6	0	100%
Lutut Kedepan	4	2	2	50%
Kombinasi salah	5	3	2	60%

Analisis Pada Tabel 5 berdasarkan jenis kesalahan menunjukkan variasi performa yang signifikan dalam mendeteksi berbagai tipe *error squat*. Sistem menunjukkan performa excellent dalam mendeteksi *error* kedalaman kurang dari 60° dengan Deteksi Rata rata 100%, mengindikasikan bahwa *algoritme angle detection* yang digunakan sangat efektif. Namun, sistem mengalami kesulitan dalam mendeteksi *error* lutut terlalu jauh ke depan dengan Deteksi Rata rata hanya 50%, dan kombinasi salah dengan Deteksi Rata rata 60%. Rendahnya Deteksi Rata rata untuk *error* lutut dapat disebabkan oleh *subtlety* dalam perubahan posisi lutut yang membutuhkan *tracking* yang lebih sensitif dan akurat.

#### IV. KESIMPULAN

Penelitian ini berhasil mengembangkan sistem deteksi gerakan Fitness secara otomatis menggunakan kombinasi metode pose estimation dan implementasi YOLOv11 sesuai dengan tujuan yang ditetapkan. Sistem mampu mengidentifikasi dua kelompok gerakan utama dalam aktivitas Fitness yaitu latihan otot kaki dan latihan otot punggung dengan akurasi klasifikasi rata-rata 73,3%.

Implementasi pose estimation menggunakan MediaPipe berhasil mendeteksi titik-titik tubuh dari video gerakan Fitness, sementara YOLOv11 efektif melakukan klasifikasi gerakan berdasarkan pola pergerakan tubuh. Sistem yang dibangun telah memenuhi fungsinya sebagai alat bantu untuk memonitor dan mengevaluasi gerakan Fitness secara mandiri.

Dengan demikian, tujuan penelitian untuk mengembangkan sistem deteksi gerakan Fitness otomatis yang mampu mengidentifikasi kelompok gerakan otot kaki dan punggung serta berfungsi sebagai alat monitoring mandiri telah tercapai.

#### DAFTAR PUSTAKA

- [1] dr. Rizal Fadli, “Kebugaran Jasmani: Pengertian, Unsur, Bentuk Latihan, dan Manfaatnya,” 2023. [Online]. Available: <https://www.halodoc.com/artikel/kebugaran-jasmani-pengertian-unsur-bentuk-latihan-dan-manfaatnya>

- [2] STRONGBEE, “Apa Itu Personal Trainer dan Manfaatnya?,” 2023. [Online]. Available: <https://strongbee.co.id/blog/apa-itu-personal-trainer>
- [3] wiley, “Pengenalan Gerakan Manusia Berbasis Computer Vision,” 2021, [Online]. Available: <https://onlinelibrary.wiley.com/doi/10.1155/2021/6679746>
- [4] dkk. Zhang, Y., “RGB-D Based Fitness Exercise Recognition Using Deep Learning Approach,” *IEEE Trans. Multimed.*, 2023.
- [5] R. Kumar, A., & Sharma, “MediaPipe-Based Human Exercise Recognition System,” *Int. J. Comput. Vis.*, 2022.
- [6] dkk. Rodriguez, M., “3D-CNN for Automated Squat Form Analysis in Fitness Applications,” *Comput. Vis. Image Underst.*, 2023.
- [7] K. Ragil, K. Dyansyah, S. D. Purwantoro, and M. Ilmi, “Penggunaan Computer Vision untuk Estimasi Pose Squat sebagai Solusi Alternatif Latihan Kebugaran di Gym,” vol. 4, pp. 199–207, 2025.
- [8] learnopencv, “Pose YOLOv7 vs MediaPipe dalam Estimasi Pose Manusia,” 2025, [Online]. Available: <https://learnopencv.com/yolov7-pose-vs-mediapipe-in-human-pose-estimation/>
- [9] J. Terven, D. M. Córdova-Esparza, and J. A. Romero-González, “A Comprehensive Review of YOLO Architectures in Computer Vision: From YOLOv1 to YOLOv8 and YOLO-NAS,” 2023. doi: 10.3390/make5040083.
- [10] Ultralytics., “Ultralytics YOLO11 Documentation,” 2024, [Online]. Available: <https://docs.ultralytics.com/models/yolo11/>
- [11] A. Fauziah and Y. Saragih, “Sistem Identifikasi Pengukuran Baju Menggunakan Human Body Estimation Dataset Mediapipe Dengan Metode Euclidean Distance,” *Aisyah J. Informatics Electr. Eng.*, vol. 5, no. 2, pp. 127–134, 2023, doi: 10.30604/jti.v5i2.151.
- [12] J. S. Mochamad Yuda Trinurais, Ardi Sanjaya, “Validasi Gerakan Sit Up Menggunakan Trigonometri Berbasis OpenCV,” vol. Agustus 20, 2024, [Online]. Available: url: <https://proceeding.unpkediri.ac.id/index.php/inotek/>
- [13] Medium, “Comparative study of SOTA Human pose estimation deep learning models,” 2023, [Online]. Available: <https://medium.com/@arumugamg/deep-learning-based-human-pose-estimation-6357c13acfe7>
- [14] Musclefirst, “Hati-Hati Cedera Punggung Karena Gerakan Squat Yang Salah!,” 2023, [Online]. Available: <https://musclefirst.co.id/blog/cedera-punggung-karena-gerakan-squat>
- [15] I. Husain, P. Purwantoro, and C. Carudin, “Analisis Performa State Management Provider Dan Getx Pada Aplikasi Flutter,” *JATI (Jurnal Mhs. Tek. Inform.*, vol. 7, no. 2, pp. 1417–1422, 2023.