

Implementasi Mediapipe Untuk Evaluasi Kualitas Gerakan *Squat* Dengan *Deep Learning*

Muhammad Reza Harsono, Mera Kartika Delimayanti

Teknik Informatika dan Komputer, Politeknik Negeri Jakarta

Depok, Jawa Barat

muhammad.reza.harsono.tik22@mhswn.pnj.ac.id

The squat is a fundamental fitness exercise that is prone to triggering joint injuries if performed with incorrect body posture. Direct supervision by professional trainers is often inaccessible due to cost and time constraints. This study aims to develop an automated squat posture evaluation system based on computer vision and deep learning that is fast and computationally efficient. The feature extraction method was built using the MediaPipe framework to track major joint coordinate points (shoulders, hips, knees, ankles), which were then transformed into spatial-temporal joint angle data. The movement quality classification process was executed using a 1-Dimensional Convolutional Neural Network (1D-CNN) architecture. Comparative hyperparameter testing was conducted to balance analytical performance and architectural load. The test results showed that the best model using a configuration of 2 convolutional layers, LeakyReLU activation function, and RMSprop optimization algorithm successfully achieved a Test Accuracy of 97.79%, Precision 97.71%, Recall 97.95%, and F1-Score 97.78%. In addition to high detection reliability, this final model proved to be extremely lightweight with a total of 15,586 computational parameters. These findings conclude that the combination of MediaPipe angle extraction and 1D-CNN networks is highly reliable in recognizing squat mechanical errors. Keywords: 3-5 keywords in the paper.

Keywords: *deep learning, squat evaluation, MediaPipe, 1D-CNN*

Gerakan *squat* merupakan latihan kebugaran fundamental yang rentan memicu cedera persendian apabila dilakukan dengan postur tubuh yang keliru. Pengawasan langsung oleh pelatih profesional seringkali tidak dapat diakses secara luas karena keterbatasan biaya dan waktu. Penelitian ini bertujuan untuk mengembangkan sistem evaluasi postur squat otomatis berbasis *computer vision* dan *deep learning* yang cepat serta efisien secara komputasi. Metode ekstraksi fitur dibangun menggunakan *framework* MediaPipe untuk melacak titik koordinat sendi utama (bahu, pinggul, lutut, pergelangan kaki), yang selanjutnya ditransformasikan menjadi data sudut persendian spasial-temporal. Proses klasifikasi kualitas gerakan dieksekusi menggunakan arsitektur 1-Dimensional Convolutional Neural Network (1D-CNN). Pengujian komparatif *hyperparameter* dilakukan guna menyeimbangkan performa analitis dan beban arsitektur. Hasil pengujian menunjukkan bahwa model terbaik yang menggunakan konfigurasi 2 lapisan konvolusi, fungsi aktivasi LeakyReLU, dan algoritma optimasi RMSprop berhasil mencapai Akurasi Uji sebesar 97%, Precision 97%, Recall 97%, dan F1-Score 97%. Selain keandalan deteksi yang tinggi, model final ini terbukti ringan dengan total 15.586 parameter komputasi. Temuan ini menyimpulkan bahwa perpaduan antara ekstraksi sudut MediaPipe dan jaringan 1D-CNN sangat andal dalam mengenali kesalahan mekanika squat.

Kata kunci: *deep learning, evaluasi squat, MediaPipe, 1D-CNN*

I. PENDAHULUAN

Seiring meningkatnya kesadaran akan pentingnya kebugaran fisik, gerakan *squat* menjadi salah satu latihan fungsional yang sangat krusial untuk dilakukan dengan postur yang benar. Gerakan *squat* adalah latihan utama yang sering digunakan dalam program performa olahraga maupun rehabilitasi lutut [1]. Aktivitas ini membutuhkan koordinasi otot yang rumit, sehingga cara pelaksanaannya sangat berpengaruh pada efektivitas latihan [2]. Meskipun sering dianggap berisiko bagi sendi lutut, tinjauan sistematis terbaru menunjukkan

bahwa *deep squat* (jongkok penuh) sebenarnya aman dan memberikan perlindungan terhadap cedera degeneratif asalkan dilakukan dengan teknik yang benar dan beban yang progresif [3]. Namun, dalam praktiknya, proses evaluasi kualitas gerakan *squat* masih sangat bergantung pada pengamatan visual subjektif oleh pelatih profesional. Proses ini memerlukan biaya yang relatif mahal dan seringkali menjadi hambatan bagi masyarakat umum yang ingin melakukan latihan mandiri. Padahal, latihan mandiri tanpa pengamatan ahli dapat berakibat pada penerapan teknik yang salah, seperti yaitu *lumbar*

hyperlordosis (punggung terlalu menekuk ke dalam) dan *knee valgus* (lutut kolaps ke dalam) [4], yang secara signifikan meningkatkan risiko cedera jangka panjang. Oleh karena itu, dibutuhkan pendekatan evaluasi gerakan yang lebih terjangkau, efisien, dan adaptif. Seiring perkembangan teknologi *computer vision*, *deep learning* mulai dimanfaatkan sebagai solusi untuk memantau postur tubuh secara otomatis, sehingga proses evaluasi latihan dapat dilakukan secara lebih objektif [5], menyatakan bahwa aplikasi *mobile* yang mengintegrasikan *computer vision* dapat memungkinkan pengguna memantau postur latihan secara mandiri.

Beberapa penelitian sebelumnya telah mengimplementasikan kecerdasan buatan untuk menganalisis dan mengklasifikasi gerakan *squat* [6], menerapkan metode *Support Vector Machine* (SVM) berbasis MediaPipe untuk mengklasifikasi gerakan *squat* dan menunjukkan tingkat akurasi mencapai 99,1%. Penelitian lainnya menggunakan pendekatan *deep learning* dengan arsitektur Bi-CGRU untuk analisis performa dan koreksi postur olahraga secara *real-time*, menghasilkan akurasi sebesar 96,1% [7]. Namun, penelitian pertama masih bergantung pada metode *machine learning* konvensional (SVM), sedangkan penelitian kedua menggunakan model *deep learning* dengan beban komputasi tinggi. Kesenjangan inilah yang menjadi dasar dari penelitian ini.

Penelitian ini bertujuan untuk membangun model *deep learning* yang mampu mengklasifikasikan kualitas postur gerakan *squat* serta menghitung jumlah repetisi yang benar secara otomatis. Dengan memanfaatkan sinyal *time-series* dari ekstraksi fitur sudut persendian yang telah diproses secara sistematis, model 1D-CNN yang dikembangkan diharapkan mampu memberikan evaluasi pergerakan secara lebih cepat, ringan secara komputasi, dan berbasis data. Pendekatan ini menjadi langkah awal menuju sistem asisten kebugaran mandiri yang lebih cerdas, otomatis, dan mudah diakses oleh pengguna awam.

Adapun batasan dalam penelitian ini adalah penggunaan input berupa video dengan format .MP4 yang diproses menjadi sinyal *time-series* dari sudut lutut, pinggul, dan punggung menggunakan MediaPipe Pose. Jumlah dataset masih terbatas sehingga pengujian pada variasi subjek yang lebih luas diperlukan pada penelitian lanjutan. Penelitian ini juga difokuskan pada pengembangan model *deep learning* berbasis 1D-CNN (*One-Dimensional Convolutional Neural Network*). Tujuan utama dari

penelitian ini adalah merancang model *deep learning* yang dapat memberikan evaluasi kualitas gerakan *squat* secara ringan dan cepat untuk membantu latihan mandiri. Dengan demikian, hasil penelitian ini diharapkan dapat memberikan kontribusi nyata dalam pengembangan sistem pencegahan cedera olahraga yang efisien, akurat, dan adaptif terhadap tren pemantauan kesehatan modern.

II. METODOLOGI PENELITIAN

Penelitian ini merupakan penelitian terapan dengan metode kuantitatif, yang berfokus pada pengembangan model *deep learning* untuk mendukung evaluasi latihan fisik mandiri. Penelitian ini dilakukan untuk membangun sistem berbasis 1D-CNN (*One-Dimensional Convolutional Neural Network*) yang mampu mengklasifikasikan postur gerakan *squat* serta menghitung jumlah repetisi yang benar secara otomatis. Untuk mencapai tujuan tersebut, tahapan dari penelitian ini mencakup pengumpulan data, *pre-processing*, perancangan arsitektur model, pelatihan model, dan evaluasi. Adapun langkah-langkah secara umum dijelaskan sebagai berikut.

A. PENGUMPULAN DATASET

Langkah pertama adalah pengumpulan dataset. Data yang digunakan dalam penelitian ini merupakan dataset sekunder berbasis rekaman video (.MP4) yang bersumber dari kumpulan sesi latihan fisik "MM-Fit Dataset - physical exercise workout sessions" dan video latihan dari platform Youtube dan Tiktok. Dataset ini terdiri dari total 28 video dengan durasi masing-masing 1 menit. Dataset yang sudah terkumpul juga dilakukan augmentasi berupa peningkatan kecepatan, penurunan kecepatan, rotasi kamera, dan zoom kamera. Sehingga total keseluruhan dataset adalah 1390 gerakan *squat*. Video gerakan *squat* ini kemudian diproses menggunakan teknologi MediaPipe Pose untuk mendeteksi dan mengekstraksi koordinat 33 titik *landmark* tubuh 3D. Data koordinat inilah yang menjadi bahan baku utama pembentukan sinyal *time-series* untuk melatih model pendeteksi kesalahan postur.

B. PRE-PROCESSING DATA

Langkah berikutnya adalah *pre-processing data*. Data mentah berupa *landmark* tubuh tidak dapat langsung dimasukkan ke dalam model *deep learning*. Data tersebut perlu dipersiapkan dan dibersihkan untuk menghindari bias dan

meningkatkan ketahanan (*robustness*) model. Tahapan dalam *pre-processing* data meliputi:

1. Ekstraksi Fitur (*Feature Extraction*)

Dari total 33 titik *landmark* tubuh yang dihasilkan MediaPipe, dilakukan proses pemilihan fitur yang secara spesifik berfokus pada sudut persendian utama yang relevan dengan mekanika *squat*, yaitu sudut sendi lutut, pinggul, dan punggung. Ekstraksi ini penting untuk mengurangi komputasi yang tidak perlu dan memfokuskan model pada area rentan cedera seperti *lumbar hyperlordosis* atau *knee valgus*.

2. *Filtering* Sinyal

Sinyal *time-series* dari koordinat tubuh sering kali terkontaminasi oleh *noise* dan *jitter* yang diakibatkan oleh getaran kamera atau ketidakstabilan deteksi. Oleh karena itu, diterapkan teknik *Low-Pass Filtering* (seperti filter Gaussian atau Butterworth) untuk menghaluskan sinyal tanpa merusak informasi temporal dari gerakan yang sebenarnya

3. Segmentasi Data (*Sliding Window*)

Model 1D-CNN memerlukan input data dengan panjang vektor yang tetap. Karena video memiliki durasi kontinu yang bervariasi, diterapkan teknik *Sliding Window* dengan ukuran jendela sebesar 30 *frame*. Teknik ini memotong aliran gerakan menjadi segmen-segmen waktu yang saling tumpang tindih (*overlap*), misalnya, *overlap* 50% berarti separuh data dari jendela pertama akan muncul kembali di jendela kedua [8], sehingga model dapat belajar mengenali transisi gerakan dari awal jongkok hingga kembali berdiri utuh

4. Normalisasi Data

Normalisasi merupakan proses transformasi skala numerik fitur ke dalam rentang yang seragam, guna mencegah dominasi fitur tertentu dan mempercepat konvergensi algoritma saat pelatihan [9]. Dalam penelitian ini, fitur sudut tubuh dinormalisasi

menggunakan teknik *Min-Max Normalization* ke dalam skala [0, 1].

5. *Splitting* Data Berbasis Video

Untuk mencegah terjadinya kebocoran data (*data leakage*) yang kerap memanipulasi nilai akurasi menjadi terlalu sempurna, pemisahan data (*train, validation, test*) dilakukan dengan teknik *K-Fold Cross Validation*. Metode ini bekerja dengan membagi dataset secara merata ke dalam *k* kelompok atau lipatan (*folds*) [10]. Pemisahan dilakukan berdasarkan ID video sebelum tahapan *Sliding Window* diaplikasikan, guna memastikan model benar-benar diuji menggunakan subjek yang belum pernah dikenali (*unseen data*).

C. PERANCANGAN ARSITEKTUR MODEL DEEP LEARNING

Penelitian ini secara spesifik menggunakan 1D-CNN (*One-Dimensional Convolutional Neural Network*) yang dibangun melalui *framework* TensorFlow/Keras. 1D-CNN adalah jenis algoritma *deep learning* yang dirancang khusus untuk memproses data yang berurutan atau berbasis waktu (*time-series*), seperti sinyal sensor gerak atau urutan frame video. Model 1D-CNN hanya memindai fitur secara satu arah (sepanjang garis waktu), menjadikannya sangat ringan dan cepat secara komputasi [11]. Karakteristik utama dari data *time-series* adalah data ini sering kali terkontaminasi oleh *noise*, *jitter*, dan *outlier* yang disebabkan oleh kamera saat pengambilan video [12]. Arsitektur dirancang secara berlapis, mulai dari *Convolutional Layers* (dengan jumlah filter menurun dari 128, 64, hingga 32), *MaxPooling* dan *Global Average Pooling* untuk kompresi data, lapisan *Dropout* untuk meredam *overfitting*, dan *Fully Connected Layer* untuk menghasilkan output klasifikasi akhir.

D. PELATIHAN DAN PENGUJIAN MODEL

Model 1D-CNN dilatih menggunakan dataset yang telah melalui tahap *pre-processing*. Model ini bertugas membaca sekumpulan *frame* berukuran 30 dan memetakan pola kinematikanya untuk mengklasifikasikan kelas "Benar" atau "Salah". Selama masa pelatihan, parameter seperti *learning rate* dan *batch size* dikonfigurasi secara hati-hati, dan fitur *callbacks* seperti *Early Stopping* dan *Model Checkpoint* digunakan untuk

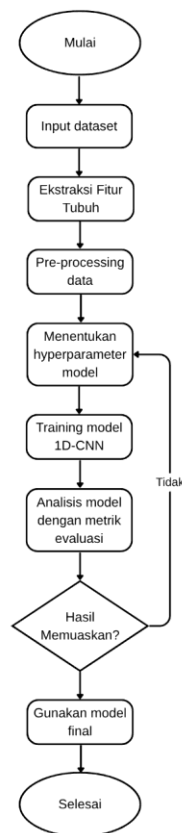
menghentikan proses pembelajaran tepat di titik optimal performa validasinya, sehingga jaringan terhindar dari menghafal data.

E. EVALUASI MODEL DAN PENGUJIAN APLIKASI

Evaluasi dilakukan untuk mengukur tingkat keandalan model dalam lingkungan simulasi nyata. Kinerja model 1D-CNN diukur menggunakan metrik *Accuracy*, *Precision*, *Recall*, *F1-Score*, dan *Confusion Matrix*. *Accuracy* digunakan untuk melihat proporsi ketepatan tebakan secara keseluruhan, sedangkan *Precision* dan *Recall* (yang direpresentasikan bersama dalam *F1-Score*) sangat vital untuk memastikan model tidak sembarangan meloloskan postur yang salah (mencegah *False Positive* tinggi yang berisiko cedera).

III. HASIL DAN PEMBAHASAN

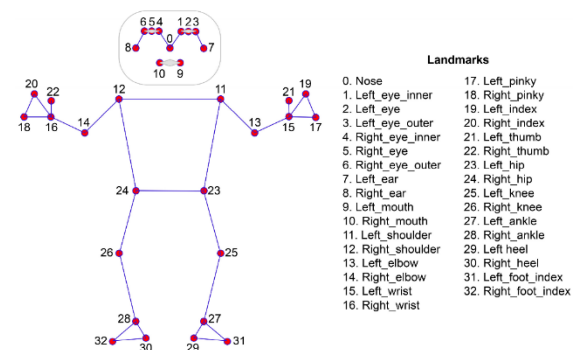
Pada penelitian ini, model 1D-CNN akan dilatih dengan pengaturan *hyperparameter* yang berbeda-beda, dengan tujuan untuk mencari model yang terbaik. Gambar 1 merupakan flowchart pencarian model 1D-CNN terbaik.



Gambar 1. Flowchart Pencarian Model 1D-CNN

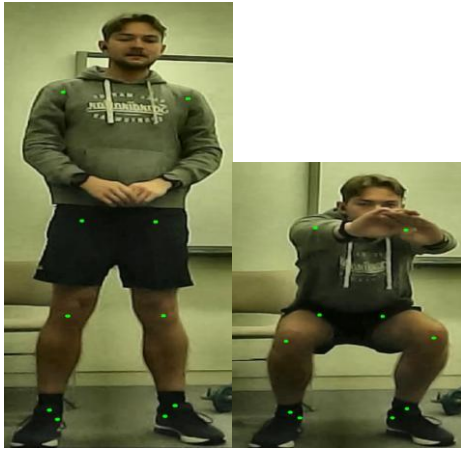
Dataset yang digunakan bersumber dari publik, yaitu MM-Fit Dataset - physical exercise workout sessions (video) dan pengumpulan video dari platform Youtube dan Tiktok. Data video mentah tidak dapat langsung diproses oleh model 1D-CNN. Oleh karena itu, diperlukan tahapan *pre-processing* data.

Proses ekstraksi fitur diawali dengan pemindaian video menggunakan *framework* MediaPipe Pose. MediaPipe merupakan *framework* berbasis *machine learning* yang dikembangkan oleh Google, dan menyediakan solusi deteksi fitur tubuh secara real-time dengan akurasi tinggi dan efisiensi komputasi yang baik [13] MediaPipe menganalisis setiap *frame* pada video untuk mengestimasi pose tubuh secara spasial, mengidentifikasi titik-titik kunci yang membentuk kerangka virtual tubuh. Secara keseluruhan, *framework* ini mengekstraksi koordinat dari 33 titik *landmark* tubuh secara spesifik.



Gambar 2. Koordinat Titik Tubuh Mediapipe [14]

Meskipun demikian, tidak seluruh data *landmark* diteruskan ke dalam model *deep learning*. Dilakukan proses pemilihan fitur (*feature selection*) yang secara spesifik hanya berfokus pada titik koordinat persendian utama yang berkorelasi langsung dengan mekanika gerakan *squat*. Fitur persendian yang dibutuhkan untuk mengevaluasi gerakan *squat* terdiri dari sendi pinggul (*hip*), sendi lutut (*knee*), dan pergelangan kaki (*ankle*) [7]. Sehingga titik-titik koordinat yang diperlukan adalah 10 titik koordinat, yaitu: 11, 12, 24, 23, 26, 25, 27, 28, 29, 30. Contoh dataset beserta *landmark* yang digunakan dapat dilihat pada Gambar 3.



Gambar 3. Contoh Gambar Dataset dan Landmark yang digunakan

Video diproses menggunakan MediaPipe Pose untuk mendeteksi 33 titik landmark tubuh 3D. Dari titik-titik ini, fitur yang diambil difokuskan pada sudut persendian utama yang relevan dengan squat, yaitu sudut lutut, pinggul, dan punggung. Teknik Low-Pass Filtering digunakan untuk memperhalus sinyal data time-series tanpa menghilangkan informasi gerakan penting. Teknik Sliding Window digunakan untuk menentukan ukuran jendela sebesar 30 frame. Teknik ini memotong aliran data kontinu menjadi segmen-segmen waktu yang tumpang tindih satu sama lain untuk menangkap transisi gerakan. Data sudut titik-titik landmark tubuh dinormalisasi ke dalam rentang 0 hingga 1 untuk mempercepat pelatihan model dan menghindari dominasi sudut titik-titik tertentu. Dua metode yang paling relevan dalam literatur pengenalan aktivitas manusia adalah *Min-Max Normalization* dan *Z-Score Standardization* [9].

Proses pencarian model 1D-CNN terbaik ini akan dibagi kedalam tiga skenario pengujian utama guna memvalidasi pengaruh masing-masing *hyperparameter* model terhadap akurasi prediksi:

1. Variasi Kedalaman Lapisan Konvolusi: Menguji performa model antara arsitektur kompleks (3 lapisan konvolusi dengan kapasitas filter hingga 256) melawan arsitektur ringkas (Micro 1D-CNN dengan 2 lapisan konvolusi berkapasitas 32 dan 64 filter).
2. Fungsi Aktivasi: Membandingkan fungsi aktivasi standar Rectified Linear Unit (ReLU) dengan Leaky ReLU ($\alpha=0.01$). Variasi ini diuji khusus untuk mengatasi fenomena Dying ReLU pada pemrosesan koordinat negatif sinyal *time-series*.

3. Algoritma Optimasi: Membandingkan efisiensi konvergensi nilai kerugian (loss) antara Adaptive Moment Estimation (Adam) dengan Root Mean Squared Propagation (RMSprop) dalam memproses fluktuasi data sekuensial penanda tubuh (landmark).

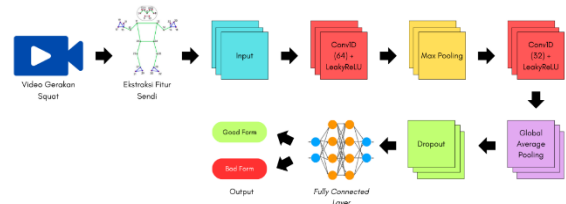
Model yang dibangun menggunakan arsitektur 1D-Convolutional Neural Network (1D-CNN). Arsitektur ini dipilih karena efektivitasnya dalam memproses data sekuensial (*time-series*) dan komputasinya yang ringan. Pembuatan model dilakukan menggunakan *library* TensorFlow/Keras. Proses ini mencakup perancangan layer konvolusi, *pooling*, dan *fully connected layer*, serta pengaturan *hyperparameter*. Berdasarkan hasil evaluasi ketiga skenario, model yang telah dilatih akan dievaluasi menggunakan data uji (testing data). Metrik evaluasi yang digunakan meliputi Accuracy, Precision, Recall, F1-Score, dan Confusion Matrix. Berdasarkan hasil evaluasi berbagai kombinasi *hyperparameter*, model dengan performa terbaik kemudian dipilih untuk digunakan dalam tahap pelatihan model. Hasil evaluasi berbagai *hyperparameter* ditunjukkan pada Tabel 1.

Tabel 1. Hasil Evaluasi Model dengan Berbagai Hyperparameter

Nilai K	Akurasi	Precision	Recall	F1-Score
Skenario A (3 Lapisan + ReLU + Adam)				
4	0.965	0.965	0.965	0.965
5	0.970	0.973	0.970	0.971
6	0.954	0.956	0.954	0.954
8	0.962	0.966	0.962	0.962
10	0.955	0.963	0.955	0.954
Skenario B (2 Lapisan + ReLU + Adam)				
4	0.959	0.960	0.959	0.959
5	0.963	0.966	0.963	0.963
6	0.962	0.961	0.962	0.962

8	0.964	0.968	0.964	0.964
10	0.965	0.970	0.965	0.965
Skenario C (3 Lapisan + LeakyReLU + RMSprop)				
4	0.951	0.952	0.951	0.951
5	0.962	0.964	0.962	0.962
6	0.952	0.955	0.952	0.952
8	0.946	0.954	0.946	0.946
10	0.965	0.968	0.965	0.965
Skenario D (2 Lapisan + LeakyReLU + RMSprop)				
4	0.962	0.963	0.962	0.962
5	0.976	0.977	0.976	0.976
6	0.951	0.955	0.951	0.951
8	0.955	0.952	0.955	0.955
10	0.955	0.959	0.955	0.954

Berdasarkan hasil evaluasi yang diperoleh melalui percobaan kombinasi *hyperparameter* yang ditunjukkan pada Tabel 1, model 1D-CNN diuji kinerjanya berdasarkan beberapa metrik, di antaranya adalah *Accuracy*, *Precision*, *Recall*, dan *F1-Score*. Dari hasil tersebut, diperoleh bahwa skenario D (2 lapisan + LeakyReLU + RMSprop) merupakan kombinasi *hyperparameter* terbaik. Kombinasi ini berhasil mencapai akurasi uji yang paling tinggi. Hal ini menunjukkan bahwa kombinasi fungsi aktivasi LeakyReLU dan algoritma optimisasi RMSprop efektif dalam mencegah terjadi *overfitting*. Sehingga model memahami pola dan pergerakan sendi daripada hanya sekadar menghafal data *training*. Model pada skenario D memiliki arsitektur model seperti di gambar 4 dan daftar parameter yang digunakan di tabel 2.



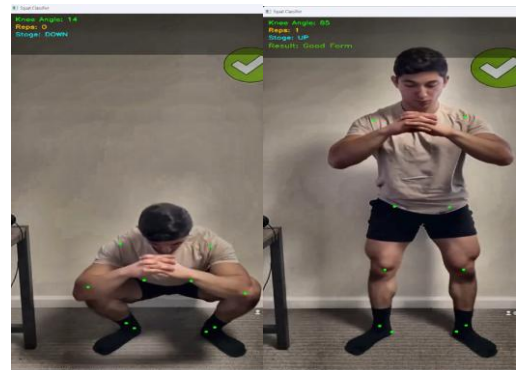
Gambar 4. Arsitektur Model 1D-CNN untuk Evaluasi Gerakan Squat

TABEL II. PARAMETER MODEL

1D-CNN	
Parameter	Value
Jumlah Epoch	41 (<i>Early Stop</i>)
Ukuran Batch	32
Fungsi Aktivasi	LeakyReLU
Algoritma Optimasi	RMSprop
Total Parameter	15.586

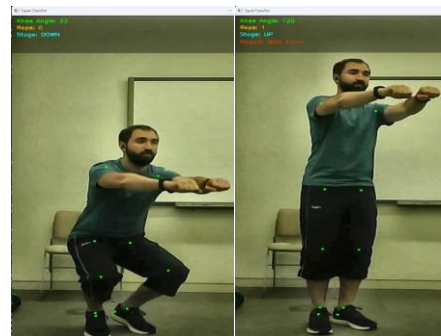
A. Hasil Klasifikasi

Hasil klasifikasi dari model yang telah dibuat dapat dilihat pada gambar 5 dan 6.



Gambar 5. Contoh Klasifikasi Gerakan Benar

Pada Gambar 5 menunjukkan contoh klasifikasi gerakan benar dimana klasifikasi dimulai pada saat pengguna berdiri, kemudian melakukan *squat* lalu kembali berdiri lagi, setelah subjek berdiri kembali maka hasil klasifikasi gerakan benar ditunjukkan di bagian kiri atas.



Gambar 6. Contoh Klasifikasi Gerakan Salah

Pada Gambar 6 menunjukkan contoh klasifikasi gerakan salah dimana subjek mengangkat tumit saat melakukan *squat*, klasifikasi dimulai pada saat subjek berdiri, kemudian melakukan *squat* lalu kembali berdiri lagi, setelah subjek berdiri kembali maka hasil klasifikasi ditunjukkan di bagian kiri atas.

B. Confusion Matrix

Confusion matrix merupakan salah satu metode evaluasi yang umum digunakan dalam sistem klasifikasi untuk mengukur kinerja model dalam membedakan kelas-kelas target. Matriks ini menampilkan perbandingan antara label yang diprediksi oleh model dengan label aktual dari data uji. Setiap elemen pada matriks menggambarkan jumlah prediksi yang benar maupun salah dalam tiap kategori, sehingga dapat memberikan pemahaman lebih rinci mengenai jenis kesalahan klasifikasi yang terjadi [15]. Gambar 7 merupakan confusion matrix untuk model 1D-CNN dalam mengevaluasi kualitas gerakan *squat*.



Gambar 7. Confusion Matrix

Hasil visualisasi *confusion matrix* pada Gambar 7 menunjukkan bahwa model 1D-CNN memiliki performa yang baik dalam mengenali kedua kategori gerakan *squat*. Pada bagian kiri atas merupakan metrik *true positive* untuk kelas *Good Form*. Angka ini menunjukkan bahwa terdapat 171 *windows* gerakan *squat* yang aslinya memang benar dan model ini berhasil mendeteksi 95.9% (164 *windows*) gerakan benar dengan tepat sebagai *Good Form*. Pada bagian kanan bawah merupakan metrik *true negative* untuk kelas *Bad Form*. Angka ini menunjukkan bahwa terdapat 146 *windows* gerakan *squat* yang aslinya salah dan model ini berhasil mendeteksi semuanya (100%) dengan tepat sebagai *Bad Form*.

Secara keseluruhan, model evaluasi ini mampu membedakan kedua kelas gerakan *squat* dengan akurasi yang tinggi. Hal ini menunjukkan bahwa model memiliki kemampuan yang baik dalam

membedakan fitur tubuh antara gerakan *squat* yang tepat dan keliru.

C. Matriks Evaluasi

Hasil evaluasi performa model 1D-CNN dalam mengevaluasi gerakan *squat* berdasarkan titik koordinat tubuh dalam suatu rentang waktu ditampilkan pada Tabel 3. Evaluasi dilakukan dengan menggunakan empat metrik, di antaranya adalah *Accuracy*, *Precision*, *Recall*, dan *F1-Score*. Perhitungan metrik ini dilakukan secara otomatis melalui Python dengan menggunakan *library* scikit-learn.

Tabel III. Matriks Evaluasi Model

Metrik Evaluasi	Skor
Accuracy	0.976
Precision	0.977
Recall	0.976
F1-Score	0.976

Nilai akurasi sebesar 0.97 menunjukkan bahwa model berhasil memprediksi 97% dari keseluruhan data uji dengan benar. Nilai *precision* sebesar 0.97 mencerminkan kemampuan model dalam memberikan label tebakan pada suatu gerakan *squat*, gerakan tersebut terbebas dari kesalahan *false positive*. Nilai *recall* sebesar 0.97 menunjukkan bahwa model mampu mendeteksi gerakan sebanyak 97%. Nilai *F1-Score* sebesar 0.97 menunjukkan bahwa model memiliki keseimbangan yang sangat kuat antara ketepatan prediksi dan kepekaan deteksi gerakan. Konsistensi nilai tinggi pada keempat metrik ini menunjukkan bahwa model memiliki performa yang baik dan stabil, baik dalam hal akurasi keseluruhan maupun dalam mengenali pola-pola spesifik pada koordinat gerakan tubuh. Hal ini menjadikan model sangat potensial untuk diterapkan dalam sistem evaluasi gerakan *squat*.

IV. KESIMPULAN DAN SARAN

A. Kesimpulan

Penelitian ini telah berhasil merancang model *deep learning* berbasis 1D-CNN yang andal dalam menilai kualitas gerakan *squat*. Kinerja model dalam mengklasifikasikan gerakan melalui titik koordinat tubuh terbukti optimal. Kesuksesan ini ditandai oleh perolehan skor yang tinggi (0.97) pada seluruh parameter evaluasi utama yakni *Accuracy*, *Precision*, *Recall*, dan *F1-Score* serta dikonfirmasi melalui *confusion matrix*. Temuan tersebut membuktikan bahwa pendekatan *deep learning* merupakan solusi yang menjanjikan untuk mewujudkan sistem evaluasi kebugaran yang otomatis, cerdas, dan efisien.

B. Saran

Berdasarkan hasil penelitian dan pengujian model yang telah dikembangkan, disarankan untuk melakukan pengumpulan dan pembaruan *dataset* secara berkala dengan mempertimbangkan berbagai kondisi fisik manusia, seperti jenis pakaian yang digunakan dan ukuran tubuh pengguna agar dapat meningkatkan kemampuan generalisasi model. Evaluasi model dengan berbagai rasio pembagian data yang beragam untuk mengetahui hasil model yang terbaik. Selain itu, perlu dilakukan peningkatan terhadap kemampuan model dalam mengevaluasi kualitas gerakan *squat* pada *angle* kamera dari samping dan jenis *squat* yang lain.

V. REFERENSI

- [1] L. V Slater and J. M. Hart, "MUSCLE ACTIVATION PATTERNS DURING DIFFERENT SQUAT TECHNIQUES," 2017. doi: 10.1519/JSC.0000000000001323.
- [2] B. A. Putra, B. Purnama, and B. Erfianto, "Klasifikasi Citra Gerakan Olahraga Dalam Gym Menggunakan Graph Convolutional Networ," 2025.
- [3] A. Rojas-Jaramillo *et al.*, "Impact of the deep squat on articular knee joint structures, friend or enemy? A scoping review," 2024, *Frontiers Media SA*. doi: 10.3389/fspor.2024.1477796.
- [4] Z. Papadakis *et al.*, "Addressing Biomechanical Errors in the Back Squat for Older Adults: A Clinical Perspective for Maintaining Neutral Spine and Knee Alignment," *J. Funct. Morphol. Kinesiol.*, vol. 9, no. 4, Dec. 2024, doi: 10.3390/jfmk9040224.
- [5] K. Contreras-Salazar, P. Costa-Mondragon, and W. Ugarte, "Mobile Application for Optimizing Exercise Posture Through Machine Learning and Computer Vision in Gyms," in *International Conference on Information and Communication Technologies for Ageing Well and e-Health, ICT4AWE - Proceedings*, Science and Technology Publications, Lda, 2025, pp. 360–367. doi: 10.5220/0013439300003938.
- [6] B. A. Putra, B. Purnama, and B. Erfianto, "Klasifikasi Analisis Gerakan Squat untuk Pemula dan Profesional Menggunakan Metode SVM Berbasis Mediapipe," 2025, doi: <https://doi.org/10.58344/locus.v4i12.4871>.
- [7] P. Rao, C. S. Asha, and P. R. Rao, "Real-Time Posture Correction of Squat Exercise: A Deep Learning Approach for Performance Analysis and Error Correction," *IEEE Access*, vol. 13, pp. 39557–39571, 2025, doi: 10.1109/ACCESS.2025.3545207.
- [8] T. H. Tan, J. H. Tian, A. K. Sharma, S. H. Liu, and Y. F. Huang, "Human Activity Recognition Based on Deep Learning and Micro-Doppler Radar Data," *Sensors*, vol. 24, no. 8, Apr. 2024, doi: 10.3390/s24082530.
- [9] J. L. He, J. H. Wang, C. M. Lo, and Z. Jiang, "Human Activity Recognition via Attention-Augmented TCN-BiGRU Fusion," *Sensors*, vol. 25, no. 18, Sep. 2025, doi: 10.3390/s25185765.
- [10] S. Prusty, S. Patnaik, and S. K. Dash, "SKCV: Stratified K-fold cross-validation on ML classifiers for predicting cervical cancer," *Frontiers in Nanotechnology*, vol. 4, Aug. 2022, doi: 10.3389/fnano.2022.972421.
- [11] S. Kiranyaz, O. Avci, O. Abdeljaber, T. Ince, M. Gabbouj, and D. J. Inman, "1D convolutional neural networks and applications: A survey," *Mech. Syst. Signal Process.*, vol. 151, Apr. 2021, doi: 10.1016/j.ymssp.2020.107398.
- [12] X. Kong *et al.*, "Deep Learning for Time Series Forecasting: A Survey," Mar. 2025, doi: 10.1007/s13042-025-02560-w.
- [13] A. Asvin Mahersatillah Suradi, T. Tri Funny Manguma, S. Alam, and A. Najiah Nurul Afifah, "Deteksi dan Peringatan Jarak Wajah Otomatis Menggunakan MediaPipe dan Computer Vision untuk Kesehatan Pengguna Komputer," 2025, doi: <https://mail.iainformatic.org/index.php/JUKI/article/view/1267>.
- [14] K. Y. Chen, J. Shin, M. A. M. Hasan, J. J. Liaw, O. Yuichi, and Y. Tomioka, "Fitness Movement Types and Completeness Detection Using a Transfer-Learning-Based Deep Neural Network," *Sensors*, vol. 22, no. 15, Aug. 2022, doi: 10.3390/s22155700.
- [15] S. Sathyanarayanan, "Confusion Matrix-Based Performance Evaluation Metrics," *African Journal of Biomedical Research*, pp. 4023–4031, Nov. 2024, doi: 10.53555/ajbr.v27i4s.4345.