

Klasifikasi Pose pada Olahraga Panahan Berbasis YoloV8 Menggunakan Metode SVM dan Random Forest

Tugas Akhir

**diajukan untuk memenuhi salah satu syarat
memperoleh gelar sarjana**

dari Program Studi Informatika

Fakultas Informatika

Universitas Telkom

1301213054

Yuridikta A. M.



Program Studi Sarjana Informatika

Fakultas Informatika

Universitas Telkom

Bandung

2024

LEMBAR PENGESAHAN

Klasifikasi Pose pada Olahraga Panahan Berbasis YoloV8 Menggunakan Metode SVM dan Random Forest **Pose Classification in Archery Sports Based on YoloV8 Using SVM and Random Forest Methods**

NIM : 1301213054

Yuridikta Adha Muslim

Tugas akhir ini telah diterima dan disahkan untuk memenuhi sebagian syarat memperoleh gelar pada Program Studi Sarjana Informatika

Fakultas Informatika

Universitas Telkom

Bandung, 13 Januari 2025

Menyetujui

Pembimbing I,

Bedy Purnama. S.Si., M.T., Ph.D

NIP : 10790050

Pembimbing II,

Dr. Bayu Erfianto, S.Si., M.Sc.

NIP: 05730058

Ketua Program Studi
Sarjana Informatika

Dr. Erwin Budi Setiawan, S.Si., M.T.

NIP: 00760045

LEMBAR PERNYATAAN

Dengan ini saya, Yuridikta Adha Muslim, menyatakan sesungguhnya bahwa Tugas Akhir saya dengan judul **Klasifikasi Pose pada Olahraga Panahan Berbasis YoloV8 Menggunakan Metode SVM dan Random Forest** beserta dengan seluruh isinya adalah merupakan hasil karya sendiri, dan saya tidak melakukan penjiplakan yang tidak sesuai dengan etika keilmuan yang berlaku dalam masyarakat keilmuan. Saya siap menanggung resiko/sanksi yang diberikan jika di kemudian hari ditemukan pelanggaran terhadap etika keilmuan dalam buku TA atau jika ada klaim dari pihak lain terhadap keaslian karya,

Bandung, 12 Januari 2025

Yang Menyatakan



Yuridikta Adha Muslim

Klasifikasi Pose pada Olahraga Panahan Berbasis YoloV8 Menggunakan Metode SVM dan Random Forest

Yuridikta Adha Muslim¹, Bedy Purnama², Bayu Erfianto³

^{1,2,3}Fakultas Informatika, Universitas Telkom, Bandung

¹adhamuslim@students.telkomuniversity.ac.id,

²bedy.purnama@telkomuniversity.ac.id, ³erfianto@telkomuniversity.ac.id.

Abstrak

Mengklasifikasikan gerakan atlet panahan dari rekaman video masih menjadi tantangan dalam pengembangan sistem analisis olahraga. Identifikasi yang tepat pada setiap fase gerakan memanah seperti Setup, Draw and Aim, dan Expand and Shoot membutuhkan sistem yang dapat mengenali pola gerakan yang kompleks dengan tingkat presisi yang tinggi. Penelitian ini mengembangkan sistem klasifikasi pose berbasis YOLOv8 yang dikombinasikan dengan metode Support Vector Machine (SVM) dan Random Forest (RF) untuk menganalisa dan mengklasifikasikan gerakan-gerakan atlet panahan. Penelitian ini memberikan pendekatan baru dalam analisis gerakan memanah dengan mengintegrasikan deteksi pose menggunakan YOLOv8 dan teknik machine learning untuk klasifikasi yang lebih akurat. Sistem dikembangkan melalui beberapa tahap yaitu pengumpulan data video, perancangan dan implementasi sistem, serta analisis hasil implementasi. Prosesnya meliputi ekstraksi fitur gabungan menggunakan YOLOv8 dan klasifikasi menggunakan SVM dan RF untuk kategori Recurve dan Barebow. Hasil pengujian menunjukkan performa yang berbeda antara kedua metode klasifikasi. Untuk kategori Recurve, SVM mencapai akurasi 94% pada data training dan 90% pada data testing, sementara RF mencapai akurasi 100% pada training namun menurun menjadi 87% pada testing. Pada kategori Barebow, SVM menunjukkan akurasi 88% untuk training dan 76% untuk testing, sedangkan RF mencapai 100% untuk training tetapi menurun menjadi 75% pada testing. Kedua metode menunjukkan kecenderungan yang berbeda dalam hal generalisasi, dengan SVM menampilkan stabilitas yang lebih baik antara performa training dan testing. Kedua metode tersebut menunjukkan kecenderungan yang berbeda dalam hal generalisasi, dengan SVM menunjukkan stabilitas yang lebih baik antara kinerja pelatihan dan pengujian... Hasil penelitian menunjukkan bahwa kombinasi YOLOv8 dengan SVM memberikan performa yang lebih stabil dan dapat diandalkan untuk klasifikasi pose memanah, meskipun RF menunjukkan akurasi pelatihan yang lebih tinggi namun cenderung mengalami overfitting pada data testing.

Kata kunci : Panahan, Klasifikasi Pose, RF, SVM, YoloV8

Abstract

Classifying archery athlete movements from video recordings is still a challenge in the development of sports analysis systems. Proper identification of each phase of archery movements such as Setup, Draw and Aim, and Expand and Shoot requires a system that can recognize complex movement patterns with a high degree of precision. This research develops a YOLOv8-based pose classification system combined with Support Vector Machine (SVM) and Random Forest (RF) methods to analyze and classify the movements of archery athletes. This research provides a new approach in archery motion analysis by integrating pose detection using YOLOv8 and machine learning techniques for more accurate classification. The system was developed through several stages including video data collection, system design and implementation, and analysis of implementation results. The process includes joint feature extraction using YOLOv8 and classification using SVM and RF for Recurve and Barebow categories. The test results show different performance between the two classification methods. For the Recurve category, SVM achieved 94% accuracy on training data and 90% on testing data, while RF achieved 100% accuracy on training but decreased to 87% on testing. In the Barebow category, SVM showed 88% accuracy for training and 76% for testing, while RF achieved 100% for training but decreased to 75% in testing. The two methods showed different tendencies in terms of generalization, with SVM displaying better stability between training and testing performance.. The results show that the combination of YOLOv8 with SVM provides a more stable and reliable performance for archery pose classification, although RF shows higher training accuracy but tends to experience overfitting on testing data.

Keywords: Archery, Pose Classification, RF, SVM, YoloV8

1. Pendahuluan

1.1 Latar Belakang

Perkembangan teknologi di bidang visi komputer pada era digital ini telah membuka jalan baru untuk meningkatkan analisis dan pemahaman dalam setiap aspek kehidupan, baik itu analisa kepada makhluk hidup maupun benda mati. Analisa pada makhluk hidup tidak terbatas pada kondisi aktivitas pada umumnya seperti berjalan, memasak, tidur, dan lainnya, melainkan juga pada aktivitas lainnya yang bisa dilakukan oleh makhluk hidup tersebut. Aktivitas - aktivitas tersebut memiliki pergerakan yang bisa dideteksi melalui pendeteksian pose. Pose yang bisa dideteksi diantaranya adalah tangan, kepala, wajah, kaki, dan setiap persendian[1]. Setiap pose tersebut ketika diproses ke dalam komputer akan ditentukan setiap *keypoints* pose agar pengukuran yang dilakukan lebih tepat dan akurat[2]. Pendeteksian pose tersebut juga bisa dilakukan pada aktivitas olahraga. Sama seperti pendeteksian pose yang sebelumnya, pose yang dideteksi berupa kepala, tangan, badan, dan kaki. Akan tetapi ada beberapa kondisi dimana pendeteksian untuk *keypoints* yang telah dilakukan tidak terbaca dengan baik, maka dari itu diperlukan deteksi pose yang berkelanjutan agar hasil bisa maksimal [3]. Untuk pengembangan deteksi pose pada olahraga bisa dilakukan pada berbagai jenis olahraga yang ada, tidak terkecuali pada olahraga panahan.

Panahan sendiri merupakan sebuah cabang olahraga yang mulai menarik banyak orang, baik itu untuk hiburan semata ataupun untuk berkompetisi secara *intens*. Olahraga panahan juga merupakan olahraga yang mana setiap atletnya merasakan perubahan gerak akan tetapi perubahan tersebut tidak terlihat secara kasat mata [2]. Perubahan tersebut terdiri dari beberapa gerakan seperti *Set-up*, *Draw and Aim*, dan *Expand and Shoot*[4]. Metode tradisional yang melibatkan analisis manual sudah digantikan dengan pendekatan visi komputer yang lebih efisien dan otomatis. Pada penelitian sebelumnya yang telah diangkat oleh Kawaguchi dan Saptiaawan [2][4], pendeteksian pose tidak dilakukan menggunakan metode yang ada pada pengujian visi computer. Maka dari itu penelitian ini berusaha mencari dan mengungkap dengan menggabungkan teknologi YOLOv8, yang mana memungkinkan dalam pendeteksian pose tubuh manusia yang lebih akurat dari vide, dengan mempertimbangkan cara pemrosesan data oleh Support Vector Machine (SVM) dan Random Forest (RF) untuk menghasilkan nilai klasifikasi pose pada atlet panahan yang belum pernah dilakukan penelitian sebelumnya.

Pendekatan visi computer ini sendiri banyak jenis yang bisa digunakan [5][6]. Pendekatan yang bisa digunakan untuk pengujian model *deep learning* tersebut diantaranya SVM dan RF. Pengujian *deep learning* tidak hanya terfokus kepada kedua metode tersebut, akan tetapi ada algoritma lain yang bisa digunakan seperti *AlexNet*, LSTM, *Convolutional Neural Network* (CNN), dan *Decision Tree* (DT)[6]. Dari setiap algoritma yang telah disebutkan memiliki kelebihan dan kekurangannya masing-masing, seperti halnya dengan LSTM. Pada dasarnya memang LSTM dirancang untuk mendeteksi gambar, akan tetapi pada implementasinya bisa digunakan juga pada video seperti yang dijelaskan oleh Siddiqui et al [6]. Jika melihat juga pada perbandingan penggunaan SVM dan RF dengan LSTM, maka bisa dilihat bahwa perbedaan nilai yang dihasilkan cukup signifikan. Terlihat pada *Precision*, *Recall*, dan *F1 Score* bahwa nilai yang didapatkan oleh SVM itu sendiri berturut-turut di angka 0.867, 0.846, dan 0.846. Kemudian untuk RF sendiri berada diangka 0.998, 0.995, dan 0.998. Nilai tersebut sangat signifikan jika dibandingkan dengan LSTM yang hanya berada diangka 0.452, 0.502, dan 0.502. Hal tersebut bisa menjadi indikasi awal untuk menguji kedua model tersebut untuk mengklasifikasikan suatu pose.

Disaat yang bersamaan, SVM dan RF tersebut berjalan berdampingan dengan algoritma estimasi pose yang lainnya, yaitu YOLOv8. YOLOv8 adalah algoritma pose yang dirilis pada tahun 2023 untuk pendeteksi objek real-time YOLO dengan akurasi dan kecepatan yang luar biasa. Dibangun di atas kemajuan dari versi sebelumnya dari YOLO, YOLOv8 memiliki fitur dan pengoptimalan baru yang membuatnya pilihan yang sempurna untuk berbagai tugas pendeteksian objek dalam berbagai aplikasi [7]. Sejumlah penelitian sebelumnya telah mengungkapkan dan menerapkan model *deep learning* dan estimasi pose pada olahraga, ada beberapa metode yang lain seperti *MoveNet*, *OpenPose*, dan *MediaPipe*. Jika ditelusuri lebih lanjut memang setiap metode ada kelebihan dan kekurangan, seperti yang ada pada OpenPose contohnya. Perbedaan tersebut bisa dilihat pada penelitian yang dilakukan oleh Dong et al [8] yang mana nilai OpenPose untuk AP berada di angka 48.0, sedangkan untuk YoloV8 sendiri diangka 70.9. Kemudian untuk *latency* yang dihasilkan juga cukup signifikan perbedaan keduanya, untuk OpenPose sendiri memiliki *latency* diangka 368 ms dan untuk YoloV8 sendiri berada diangka 78. Perbedaan tersebut bisa terjadi dikarenakan OpenPose cenderung lebih berat jika dibandingkan dengan YoloV8 dalam hal komputasi pendeteksian pose.

Penelitian ini menggunakan YoloV8 untuk mengekstraksi fitur dari setiap pose dengan tingkat detail yang tinggi. Kemudian penerapan SVM dan Random Forest ini digunakan untuk analisis dan klasifikasi gerakan. SVM dipilih karena kemampuannya dalam menangani data dengan dimensi tinggi serta menghasilkan margin keputusan yang optimal sehingga cocok untuk mengklasifikasikan pola gerakan yang kompleks. Di sisi lain, Random Forest digunakan karena keandalannya dalam menangani data dengan karakteristik non-linier, kepekaan terhadap fitur penting dan kemampuannya untuk mengurangi risiko overfitting melalui pendekatan ensemble. Hal tersebut merupakan kombinasi yang belum banyak dilakukan maupun dijelajahi secara menyeluruh dalam hal analisis gerakan olahraga panahan. Harapann *output* dari penelitian ini adalah sistem yang digunakan bisa mendeteksi gerakan pada panahan secara akurat dan bisa mengklasifikasikan setiap gerakan yang diujikan.

1.2. Topik dan Batasannya

Topik dan batasan masalah pada penelitian ini adalah pengujian performa dari model klasifikasi berdasarkan data video atlet panahan bertangan kanan. Penelitian ini berfokus pada pengenalan gerakan atlet panahan menggunakan fitur sendiri yang telah diekstrak dari video. Pengklasifikasian dilakukan berdasarkan 3 kelas, yaitu *Set-up*, *Draw and Aim*, dan *Expand and Release* dengan membandingkan hasil dari *Random Forest* dan *Support Vector Machine*. Data yang digunakan berasal dari video rekaman atlet panahan kategori *recurve* dan *barebow*.

1.3. Tujuan

Tujuan pada penelitian ini adalah mengembangkan sistem yang dapat mengenali gerakan atlet panahan dengan akurasi yang tinggi berdasarkan data video rekaman. Penelitian ini juga bertujuan untuk melakukan klasifikasi video panahan dari kategori *recurve* dan *barebow* dengan mengacu pada tiga kelas, yaitu *Set-up*, *Draw and Aim*, dan *Expand and Release*. Klasifikasi tersebut dilakukan dengan memanfaatkan data fitur sendiri yang telah dianalisis menggunakan metode *Random Forest* dan *Support Vector Machine*.

1.4 Organisasi Tulisan

Pada Bab 2 membahas studi yang relevan. Pada Bab 3 membahas teori dan rancangan sistem penelitian. Pada Bab 4 membahas hasil dan analisis, dan Bab 5 membahas kesimpulan penelitian

2. Studi Terkait

Penelitian yang berjudul “Klasifikasi Pose pada Olahraga Panahan Berbasis YoloV8 Menggunakan Metode SVM dan Random Forest” ini diperlukan pendalaman literatur terkait evolusi dan aplikasi dari teknologi YoloV8, SVM, dan Random Forest perihal pengenalan gerakan manusia dalam olahraga, khususnya pada olahraga panahan.

Pengeksplorasian pada sisi *Human Activity Recognition* (HAR) merupakan suatu tantangan tersendiri dari pengolahan data. Menurut Jaouedi et al [9] tantangan itu sendiri terjadi ketika melakukan suatu *scene* yang kompleks dan menuntut kecepatan yang tinggi. Kemudian menurut Munea et al [10] tantangan itu sendiri bisa terjadi pada bagaimana penentuan *keypoints* yang digunakan dimana penentuan tersebut harus sesuai pada pada titik pengujian agar mendapatkan hasil yang diinginkan. Penentuan *keypoints* tersebut dilakukan pada YoloV8.

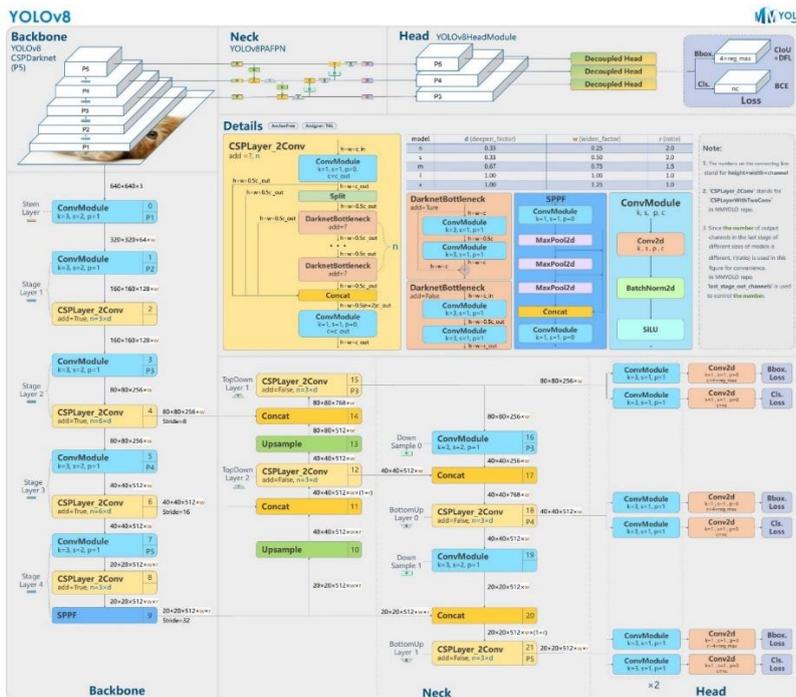
Jika dibandingkan dengan YOLO versi lainnya, YoloV8 menghasilkan nilai yang mumpuni. Penelitian yang dilakukan oleh Muhammad Husain [11] dapat berjalan secara efektif jika dibandingkan dengan YOLO generasi sebelumnya. Penggunaan YoloV8 dalam pendeteksian pose sendiri telah dilakukan oleh Tan et al [12] yang mana telah menjabarkan bagaimana penggunaan setiap fitur yang ada pada YoloV8 tersebut. Pada penelitian tersebut menggunakan robot sebagai representasi manusia yang dideteksi posenya.

Kemudian SVM dan *Random Forest* sendiri merupakan algoritma pengujian yang biasa digunakan pada penelitian lain, seperti yang dilakukan Avci et al [13]. Pada penelitian tersebut menunjukkan pada RF menghasilkan nilai yang lebih baik dibandingkan SVM itu sendiri. Hal ini juga selaras dengan penelitian yang dilakukan oleh Siddiqui et al [6] dimana hasil RF lebih baik jika dibandingkan dengan SVM itu sendiri.

Di sisi lain, pendeteksian pada olahraga panahan memiliki peran yang sangat krusial dalam penelitian kali ini. Penentuan krusial tersebut berdasarkan penelitian yang telah dilakukan oleh Kawaguchi et al [2] yang mana menampilkan bagaimana titik yang optimal dalam pengambilan gambar pengujian dan teknik apa saja yang akan menjadi acuan dalam pengujian kali ini. Kemudian dalam studi yang dilakukan oleh Saptiawan et al [4] menunjukkan titik mana saja yang memungkinkan untuk dijadikan pengujian pada penelitian kali ini.

2.1. YoloV8

YoloV8 merupakan pengembangan lanjutan yang dilakukan oleh ultralytic pada tahun 2023 dengan lima versi skala, yaitu *nano*, *small*, *medium*, *large*, dan *extra large*[14]. YoloV8 sendiri merupakan model pendeteksian objek yang sangat dihormati dikarenakan keseimbangan yang sangat baik antara akurasi dan *real time performance* nya [12]. Berbagai studi telah dilakukan untuk menguji seberapa efektif dan presisi dari penggunaan YoloV8 ini. Secara keseluruhan dari struktur YoloV8 bisa dilihat pada gambar 1.

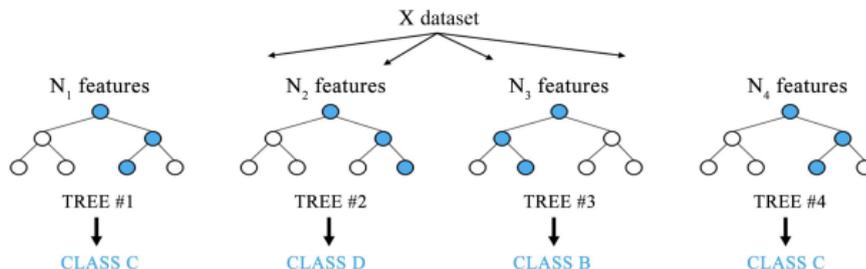


Gambar 1. Arsitektur YoloV8 secara keseluruhan [14]

Dari gambar tersebut bisa dilihat bagaimana YoloV8 mendeteksi setiap fitur yang ada pada gambar maupun video yang dimasukkan pada sistem.

2.2. Random Forest.

Random Forest merupakan suatu metode supervised learning yang menggabungkan beberapa decision tree untuk mengurangi error yang ada pada data[15]. Pengklasifikasian Random Forest terdiri dari sejumlah pohon dengan setiap pohon menggunakan semacam bentuk pengacakan[16]. Random Forest bisa dibidang salah satu metode yang sangat baik dalam menangani data yang kompleks dan besar serta dapat mengatasi masalah overfitting data. Secara garis besar proses kerja yang dilakukan oleh random forest bisa dilihat pada gambar 2.

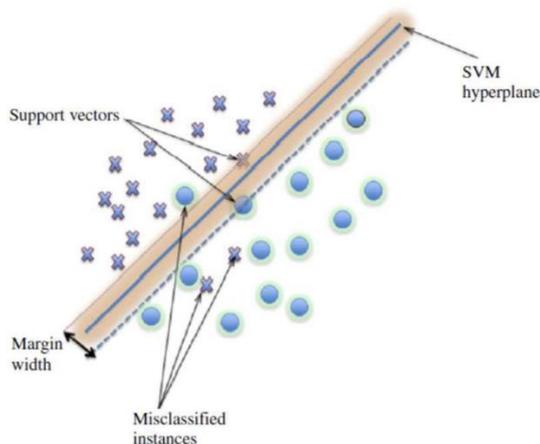


Gambar 2. Arsitektur Random Forest [16].

Bisa dilihat pada gambar 2, Random Forest membangun beberapa decision tree yang dilatih secara independen menggunakan subset data yang dipilih secara acak. Random Forest mengubah cara klasifikasi atau regresi dibangun. Pada pohon standar, setiap node terpisah memiliki pemisahan terbaik dari semua variabel, sedangkan pada RF, setiap node terpisah memiliki pemisahan terbaik dari semua variabel[16].

2.3. Support Vector Machine

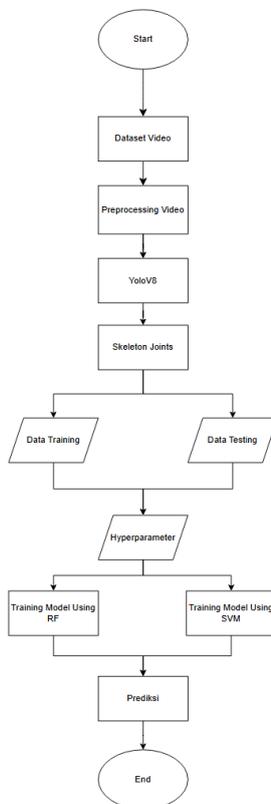
Support Vector Machine atau SVM merupakan suatu metode yang digunakan untuk menyelesaikan masalah klasifikasi data. Dalam penggunaannya, model SVM ini dapat diartikan sebagai algoritma yang bertujuan untuk mendapatkan nilai margin hyperplane paling optimal[17]. Nilai margin tersebut berguna untuk mengukur seberapa jauh hyperplane tersebut terhadap titik-titik terdekat dari setiap kelas yang ada. Secara garis besar proses kerja yang dilakukan oleh SVM bisa dilihat pada gambar 3.



Gambar 3. Ilustrasi SVM[14]

3. Sistem yang Dibangun

3.1. Desain Sistem



Gambar 4. Rancangan Sistem

Gambar 4 merupakan tahapan yang dilakukan pada penelitian ini. Metode tersebut dibagi menjadi 8 tahapan, yaitu pengumpulan dataset, *preprocessing video*, ekstraksi pose menggunakan YoloV8, data split, menentukan *hyperparameter*, training model, dan prediksi serta evaluasi. Dari tahapan-tahapan tersebut dilakukan secara berurutan untuk mengetahui seberapa jauh permasalahan yang terjadi di setiap tahapan yang ada.

3.2. Dataset Video

Langkah yang dilakukan adalah mengumpulkan video gerakan 20 atlet panahan *recurve* dan 20 atlet panahan *barebow*. Setiap pemanah *recurve* dan *barebow* melakukan gerakan secara natural yang mengandung gerakan *Setup*, *Draw and Aim*, dan *Expand and Shoot*. *Dataset* dibagi kepada 3 *output class* yang mana secara keseluruhan berupa *Setup*, *Draw and Aim*, dan *Expand and Shoot*. Setiap video yang ada dilakukan proses *trimming* sehingga didapatkan video singkat dengan panjang lebih kurang 2(dua) detik per *output class*. Pada gerakan *Setup* pada kategori *barebow* dapat terjadi perbedaan waktu dikarenakan pada gerakan tersebut ada tahapan *crowling* atau mencari titik bidik pada *string*, hal ini tidak ada pada kategori *recurve* dikarenakan pada kategori tersebut sudah ada tempat untuk membidik target. Pada gerakan *Expand and Shoot* dari kedua kategori juga terjadi perbedaan waktu dimana proses *shoot* bergantung pada lama atlet melakukan *holding* untuk membidik

target.

3.3. Pre-Processing

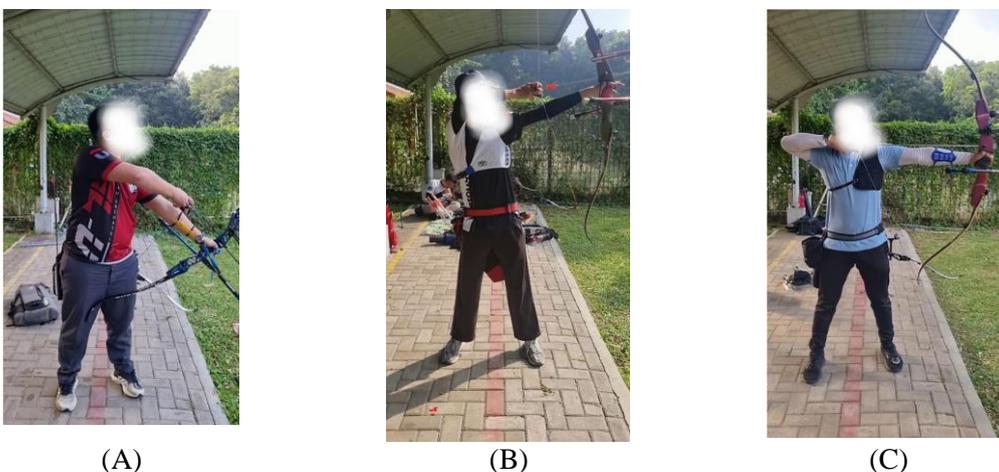
Pada tahap ini, dataset yang digunakan akan disaring dan dilakukan tahap-tahap pra-proses, tahapan-tahapan yang dilakukan seperti berikut :

- Penyesuaian FPS

Video yang telah dilakukan proses *trimming* memasuki tahapan penurunan FPS atau *frame per second*. Video-video tersebut memiliki FPS yang berbeda-beda pada saat pengumpulan data dikarenakan video-video tersebut berasal dari 3 sumber yang berbeda, yaitu dari Instagram, Youtube, dan rekam secara langsung di lapangan. Penurunan FPS ini dilakukan untuk mendapatkan ukuran yang sama pada setiap video yang ada.

- Ekstraksi Frame

Video yang telah disesuaikan ukuran FPS-nya diekstrak menjadi *frame* dengan jumlah maksimal 50 *frame per video*. Pembatasan dari jumlah *frame per video* dilakukan untuk menyamakan jumlah *frame* yang digunakan dikarenakan setiap video yang ada memiliki durasi yang berbeda-beda. Untuk memastikan bahwa setiap video bisa mendapatkan setiap gerakan yang ada tanpa perlu ada data yang hilang, maka proses tersebut sudah diantisipasi pada tahapan penyesuaian FPS. Hasil ekstrak tersebut kemudian disimpan dalam 1 folder dengan nama sesuai dengan kelas yang telah ditentukan sebelumnya. Hasil dari ekstraksi *frame* kategori *recurve* dan kaegori *barebow* bisa dilihat pada gambar 5 dan gambar 6.



Gambar 5. Hasil ekstraksi *frame* kategori *recurve*
 (A) Setup (B) Draw and Aim (C) Expand and Shoot



Gambar 6. Hasil ekstraksi *frame* kategori *barebow*
 (A) Setup (B) Draw and Aim (C) Expand and Shoot

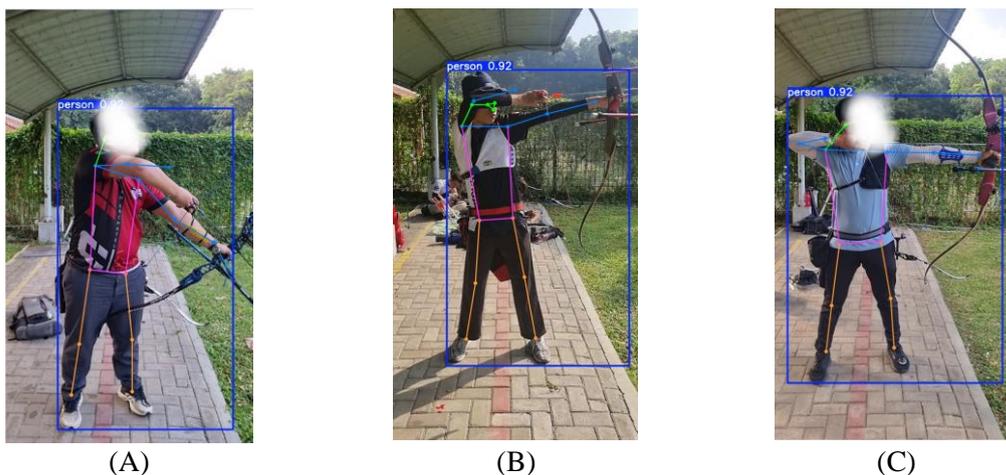
3.4. Deteksi Pose

Setelah *frame* terekstrak pada tahapan sebelumnya, selanjutnya setiap *frame* dideteksi setiap *keypoints* berdasarkan YoloV8 yang menjadi basis dari ekstraksi pose pada pengujian kali ini. *Keypoints* sendiri merupakan suatu fitur untuk mendeteksi setiap titik-titik yang unik atau menonjol dalam gambar atau video yang dapat digunakan untuk mengidentifikasi, mendeskripsikan, atau mencocokkan objek atau fitur yang ada. Pada YoloV8, terdapat 17 titik-titik *keypoints* yang bisa digunakan, titik-titik tersebut bisa dilihat pada tabel 1.

Tabel 1. Titik-titik keypoints yang bisa dideteksi pada YoloV8 [7]

No	Nama Keypoints
1.	Nose
2.	Left Eye
3.	Right Eye
4.	Left ear
5.	Right ear
6.	Left Shoulder
7.	Right Shoulder
9.	Left Elbow
9.	Right Elbow
10.	Left Wrist
11.	Right wrist
12.	Left Hip
13.	Right Hip
14.	Left Knee
15.	Right Knee
16.	Left Ankle
17.	Right Ankle

Pada penelitian ini terdapat *keypoints* yang tidak digunakan, yaitu *left ear*, *left knee*, *right knee*, *left ankle*, dan *right ankle*. *Keypoints-keypoints* tersebut tidak digunakan pada pengujian ini dikarenakan tidak adanya pergerakan yang terjadi selama aksi dilakukan sehingga tidak perlu dilakukan pendeteksian lebih lanjut. Selain itu *keypoints left ear* juga tidak dimasukkan karena pada video yang ada setiap atlet telah mengarah kepada target yang mana *keypoints* tersebut tidak terlihat kasat mata. Hasil *frame* yang telah dideteksi untuk kategori *recurve* dan kategori *barebow* bisa dilihat pada gambar 7 dan gambar 8.



Gambar 7. Hasil deteksi *keypoints* kategori *recurve*
 (A) Setup (B) Draw and Aim (C) Expand and Shoot



Gambar 8. Hasil deteksi *keypoints* kategori *barebow*
 (A) Setup (B) Draw and Aim (C) Expand and Shoot

3.5. Labeling dan Data Split.

Setelah deteksi pose dilakukan, maka masuk pada tahap *labelling*. Setiap titik data yang sudah ditentukan, kemudian dilabeli dengan angka 1 untuk kelas *Setup*, angka 2 untuk kelas *Draw and Aim*, dan angka 3 untuk kelas *Expand and Shoot*. Data yang telah dilabeli kemudian dibagi menjadi 2(dua), yaitu 75% termasuk pada data *training* dan 25 % termasuk pada data *testing*. Hal ini berlaku tidak hanya kepada kategori *recurve*, tetapi juga berlaku pada kategori *barebow*. Data tersebut berisikan setiap *keypoints* dalam bentuk titik koordinat x dan y dimana nilai-nilai tersebut disimpan dalam bentuk .csv untuk mempermudah dalam melakukan pengujian.

3.6. Hyperparameter Tuning

Hyperparameter Tuning atau HT merupakan suatu proses optimasi paramater yang tidak langsung dipelajari oleh model selama proses *training*. Pada penelitian ini, HT yang digunakan diantaranya adalah *gridsearch* untuk SVM dan *random state* untuk RF. Proses ini melakukan pengujian setiap kemungkinan sampai ditemukannya parameter yang terbaik dalam memberikan nilai skor tertinggi. Parameter SVM yang digunakan dengan *gridsearch* menggunakan range bisa dilihat pada tabel 2 dan parameter RF bisa dilihat pada tabel 3.

Tabel 2. *Hyperparamet Tuning SVM*

No	Parameter	Value
1	Kernel	'rbf'
2	C	0.01, 0.1, 1, 10, 100, 1000
3	Gamma	'scale', 'auto', 0.001, 0.01, 0.1, 1

Tabel 3. *Hyperparamet Tuning RF*

No	Parameter	Value
1	n_estimator	'100'
2	random_state	42

3.7. Training dan Pembuatan Model

Dalam penelitian ini dilakukan dua model training, yaitu menggunakan Support Vector Machine(SVM) dan Random Forest. Untuk Random Forest sendiri menggunakan 100 estimator dan disetiap estimator memiliki 42 random state. Kemudian untuk SVM sendiri menggunakan 2 paramater, yaitu parameter C dan parameter gama. Penetapan parameter dari setiap metode ini dimaksudkan untuk memberikan gambaran mendalam tentang bagaimana setiap konfigurasi mempengaruhi kemampuan model untuk mengklasifikasikan gerakan olahraga panahan.

3.8. Prediksi dan Evaluasi

Random Forest dan SVM diuji dengan data pelatihan, seperti yang dilakukan sebelumnya dengan data pelatihan dan validasi yang telah diuji. Untuk memastikan bahwa model yang dilatih dapat membuat prediksi yang tepat dan dapat dipercaya untuk membedakan antara video asli dan prediksi yang dapat dipercaya dan akurat untuk membedakan antar kategori, pengujian ini sangat penting. *Confusion Matrix*, akurasi, presisi, recall, dan skor f1 adalah ukuran kinerja sistem.

- *Confusion Matrix*

Confusion Matrix adalah tabel evaluatif yang digunakan untuk menilai keefektifan model klasifikasi dengan menyandingkan prediksi algoritme dengan data aktual. Tabel ini terdiri dari empat komponen utama: *True Positive* (TP) dan *True Negative* (TN), yang masing-masing mewakili prediksi akurat untuk kelas positif dan negatif, serta *False Positive* (FP) dan *False Negative* (FN), yang menunjukkan prediksi yang salah untuk setiap kelas. *Confusion Matrix* memungkinkan penghitungan beberapa metrik evaluasi, termasuk akurasi, presisi, recall, dan F1-Score untuk memberikan penilaian yang komprehensif terhadap kinerja model, terutama dalam set data yang ditandai dengan ketidakseimbangan kelas.

- *Accuracy*

Accuracy validation adalah metode statistik yang digunakan untuk mengevaluasi seberapa baik model klasifikasi berfungsi dengan membandingkan prediksi yang benar dengan jumlah total prediksi jurnal yang dibuat [18]. Penghitungan dari *accuracy* bisa dilihat pada rumus berikut :

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \tag{1}$$

Keterangan :

- TP = *True Positive*
- TN = *True Negative*
- FP = *False Positive*
- FN = *False Negative*

- *Precision*

Precision merupakan rasio prediksi positif metrik yang akurat terhadap semua prediksi positif model[19]. Penghitungan dari *precision* bisa dilihat pada rumus berikut :

$$Precision = \frac{TP}{TP+FP} \tag{2}$$

Keterangan :

- TP = True Positive
- FP = False Positive

- Recall

Recall merupakan presentase data yang menguntungkan dari metrik yang terdeteksi oleh model dengan benar [20]. Penghitungan dari recall bisa dilihat pada rumus berikut :

$$Recall = \frac{TP}{TP+FN} \tag{3}$$

Keterangan :

- TP = True Positive
- FN = False Negative

- F1-Score.

F1-Score merupakan rata-rata dari metric precision dan recall [20]. Penghitungan F1-Score bisa dilihat pada rumus berikut :

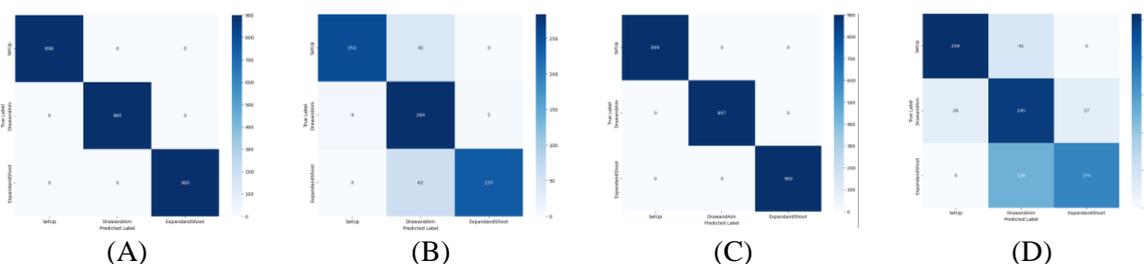
$$F1-score = 2 * \frac{Precision * Recall}{Precision+Recall} \tag{4}$$

4. Evaluasi

4.1 Hasil Pengujian

Data yang telah dilakukan *training validation* diuji untuk mengukur seberapa baik hasil dari pengujian yang telah dilakukan sebelumnya. Data *testing* adalah data yang berbeda dengan data *training and validasion* seperti yang telah disampaikan pada point data split. Hal ini dilakukan untuk melihat seberapa jauh keberhasilan pengujian model tersebut. Tujuan dari pengujian ini juga memastikan bahwa model yang dihasilkan berfungsi dengan baik pada data baru maupun data *trainig*

- Random Forest



Gambar 7. Confusion Matrix RF

(A) Training Recurve (B) Testing Recurve (C) Training Barebow (D) Testing Barebow

Berdasarkan tabel 4, tabel 5, gambar 7(A), dan gambar 7(B) hasil pelatihan model RF menunjukkan performa yang sempurna dengan nilai akurasi, presisi, recall, dan F1-score sebesar 100% untuk ketiga kelas (SetUp, DrawAndAim, dan ExpandAndShoot), seperti yang ditunjukkan pada tabel performa dan confusion matrix yang berbentuk diagonal. Hal ini mengindikasikan bahwa model dapat mengklasifikasikan semua data pelatihan dengan benar tanpa kesalahan. Hasil pengujian RF menunjukkan akurasi keseluruhan sebesar 87% untuk ketiga fitur, dengan kinerja metrik presisi, recall, dan F1-score yang bervariasi. Fitur 1 memiliki presisi yang tinggi (97%) tetapi recall yang lebih rendah (86%), yang mengindikasikan bahwa model sering kali benar dalam memprediksi kelas positif tetapi melewatkan beberapa contoh yang relevan. Sebaliknya, fitur 2 memiliki recall yang tinggi (95%) tetapi presisi yang lebih rendah (73%), mengindikasikan model cenderung lebih sering memprediksi kelas positif, meskipun tidak semuanya benar. jika dilihat pada confusion matrix, terdapat perbedaan pada expand and shoot, tetapi masih aman untuk hasil pengujian.

Tabel 4. Random Forest Training Recurve.

Class	Performance			
	Accuracy	Precision	Recall	F1-Score
1	100 %	100 %	100 %	100 %
2	100 %	100 %	100 %	100 %
3	100 %	100 %	100 %	100 %

Tabel 5. Random Forest Testing Recurve.

Class	Performance			
	Accuracy	Precision	Recall	F1-Score
1	87 %	97 %	86 %	91 %
2	87 %	73 %	95 %	83 %
3	87 %	98 %	79 %	88 %

Untuk kategori barebow dimana bisa dilihat pada tabel 6, tabel 7, gambar 7(C), dan gambar 7 (D) model RF menunjukkan performa yang sangat baik pada data training dengan akurasi 100%, F1-score yang sempurna untuk setiap kelas, dan confusion matrix tanpa kesalahan prediksi. Namun, performa pada data uji menurun secara signifikan, dengan akurasi 75% dan performa yang tidak seimbang antar kelas (precision, recall, dan F1-score terendah pada kelas “DrawandAim”). Matriks kebingungan pada data uji menunjukkan bahwa model memiliki kesalahan prediksi yang tinggi, terutama untuk kelas “ExpandandShoot,” yang sering salah diprediksi sebagai kelas “DrawandAim.”

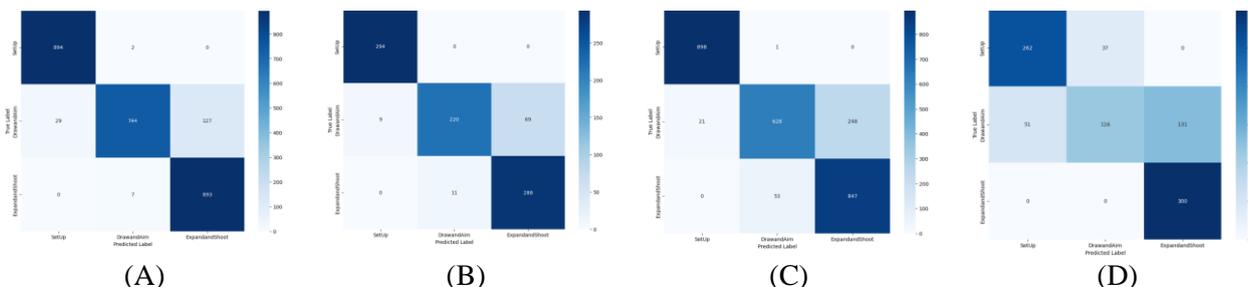
Tabel 6. Random Forest Training Barebow

Class	Performance			
	Accuracy	Precision	Recall	F1-Score
1	100 %	100 %	100 %	100 %
2	100 %	100 %	100 %	100 %
3	100 %	100 %	100 %	100 %

Tabel 7. Random Forest Testing Barebow

Class	Performance			
	Accuracy	Precision	Recall	F1-Score
1	75 %	91 %	86 %	89 %
2	75 %	59 %	82 %	69 %
3	75 %	87 %	58 %	69

- SVM



Gambar 8 Confusion Matrix SVM

(A) Training Recurve (B) Testing Recurve (C) Training Barebow (D) Testing Barebow

Untuk kategori *recurve* dimana bisa dilihat pada tabel 8, tabel 9, gambar 8(A), dan gambar 8(B) hasil evaluasi model SVM menunjukkan performa yang cukup baik, dengan akurasi sebesar 94% pada data training dan 90% pada data testing. Pada data training, kelas “SetUp” memiliki performa paling tinggi, sedangkan kelas “DrawandAim” menunjukkan performa yang lebih rendah dengan recall sebesar 83%. Pada data pengujian, kelas “DrawandAim” tetap menjadi kelemahan model dengan nilai F1-score hanya 0.83 karena presisi dan recall yang rendah. Matriks kebingungan menunjukkan bahwa sebagian besar kesalahan terjadi pada kelas “DrawandAim,” yang sering salah diprediksi sebagai “ExpandandShoot.” Masih ada ruang untuk perbaikan dalam menangani ketidakseimbangan antar kelas.

Tabel 8. SVM Training Recurve

Class	Performance			
	Accuracy	Precision	Recall	F1-Score
1	94 %	97 %	100 %	98 %
2	94 %	99 %	83 %	90 %
3	94 %	88 %	99 %	93 %

Tabel 9. SVM Testing Recurve

Class	Performance			
	Accuracy	Precision	Recall	F1-Score
1	90 %	97 %	100 %	98 %
2	90 %	95 %	74 %	83 %
3	90 %	81 %	96 %	88 %

Untuk kategori *barebow* Dimana bisa dilihat pada tabel 9, tabel 10, gambar 8(C), dan gambar 8(D) model SVM menunjukkan performa yang sangat baik pada data training dengan akurasi 88%, terutama pada kelas Setup (precision 0.98, recall 1.00). Namun, performa pada data testing mengalami penurunan yang cukup signifikan dengan akurasi 76%. Hal ini terlihat pada confusion matrix, di mana model kesulitan untuk membedakan kelas DrawandAim (recall 0.39), sering kali salah diklasifikasikan sebagai Setup atau

ExpandandShoot.

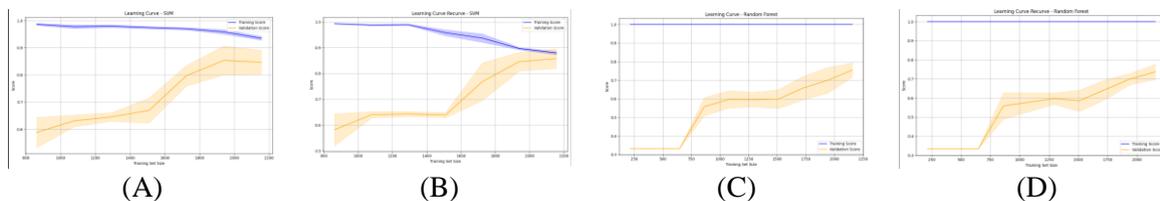
Tabel 9. SVM Training Barebow

Class	Performance			
	Accuracy	Precision	Recall	F1-Score
1	88%	98%	100%	99%
2	88%	92%	70%	80%
3	88%	77%	94%	85%

Tabel 10. SVM Testing Barebow

Class	Performance			
	Accuracy	Precision	Recall	F1-Score
1	76%	84%	88%	86%
2	76%	76%	39%	51%
3	76%	70%	100%	82%

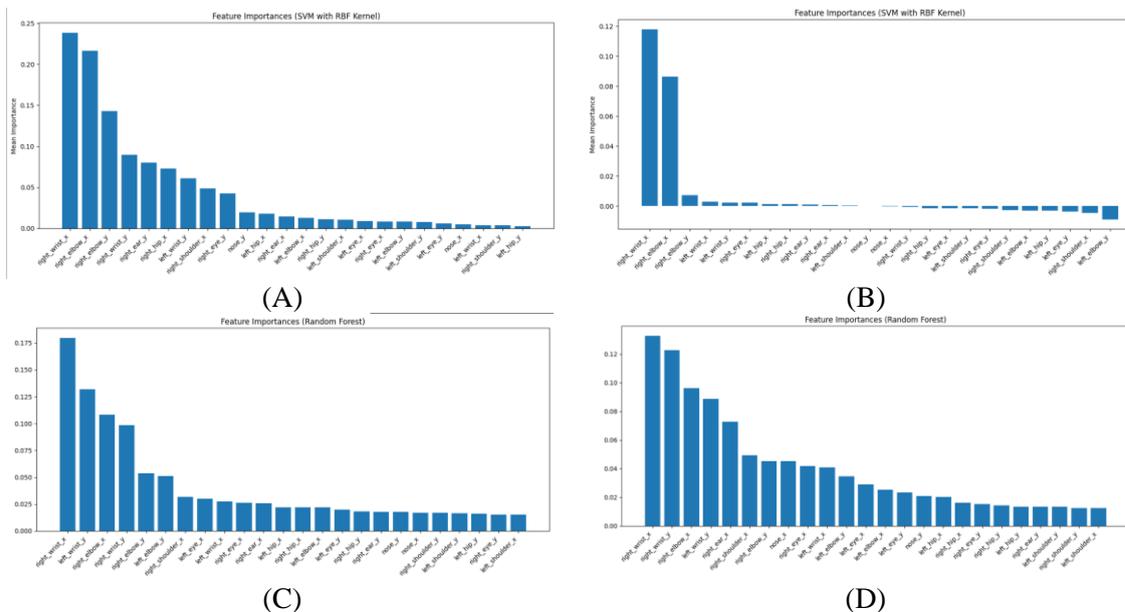
- Learning Curve



Gambar 9. Learning curve (A) SVM Recurve (B) SVM Barebow (C) RF Recurve (D) RF Barebow

Pada gambar 9 bisa dilihat pembelajaran yang dilakukan oleh setiap metode. Pada gambar 9(A) dan 9(B) menunjukkan konvergensi yang patut dipuji, dengan kurva *training* dan *testing* yang tetap berdekatan. Hal ini menunjukkan bahwa model-model ini telah mencapai keseimbangan optimal antara bias dan varians, sehingga menghindari *overfitting* yang substansial. Perbedaan minimal antara kinerja *training* dan *testing* menandakan bahwa model-model tersebut menggeneralisasi secara efektif untuk data baru. Akan tetapi pada gambar 9(C) dan 9(D) terjadi perbedaan yang signifikan antara kurva pelatihan dan pengujian untuk RF Recurve dan RF Barebow menandakan *overfitting*. Performa *testing* (garis jingga) secara nyata jauh dibawah performa *training* (garis biru), yang menunjukkan bahwa model-model ini terlalu terspesialisasi pada data *training* dan tidak menggeneralisasi secara efektif ke contoh-contoh baru.

- Feature Importance



Gambar 10. Feature Importance (A) SVM Recurve (B) SVM Barebow (C) RF Recurve (D) RF Barebow

Pada gambar 10 bisa dilihat koordinat *keypoints* mana saja yang paling berpengaruh pada pengujian ini. Semakin tinggi nilai dari koordinat *keypoints*, maka semakin berpengaruh pula koordinat *keypoints* tersebut. Pada gambar 10(A) dan 10(B) bisa dilihat bahwa *right_wrist_x* dan *right_elbow_x* memiliki pengaruh nilai tertinggi pada metode SVM. Untuk gambar yang sama, nilai *left_hip_y* dan *right_shoulders_y* pada gambar 10(A) dan *left_elbow_y* dan *right_shoulders_x* pada gambar 10(B)

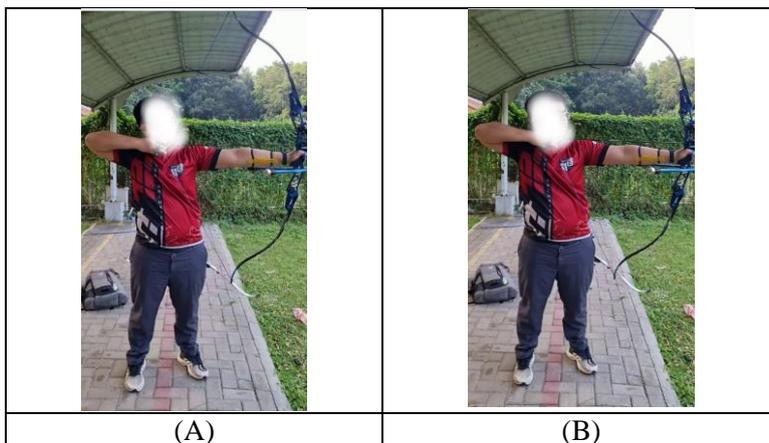
memiliki nilai terendah.

Hal ini juga bisa dilihat pada gambar 10(C) dan 10(D) dimana *right_wrist_x* dan *right_wrist_y* memiliki nilai tertinggi. Untuk gambar yang sama pula, nilai *left_shoulder_x* dan *right_eyes_y* pada gambar 10(C) dan *left_shoulder_x* dan *right_shoulders_y* pada gambar 10(B) memiliki nilai terendah.

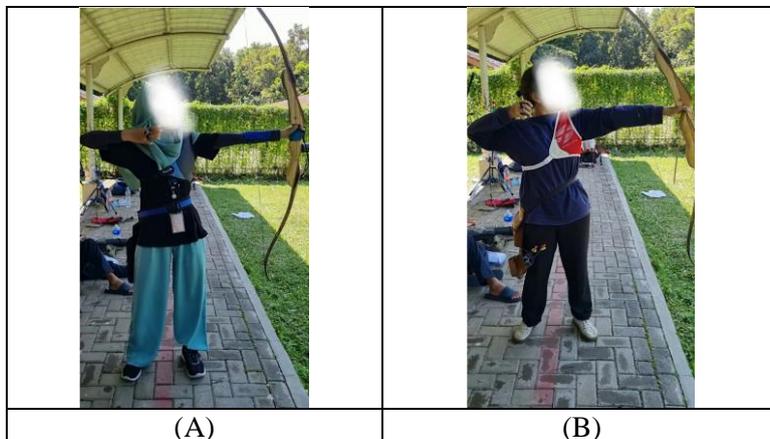
4.2 Analisis Hasil Pengujian

Berdasarkan hasil pada point sebelumnya, maka dapat dilihat bahwa kedua kategori pada metode yang digunakan menghasilkan nilai yang berbeda-beda pada *training* dan *testingnya*. Metode RF menunjukkan nilai terbaik pada *training* dimana menghasilkan nilai 100% pada kedua kategori. Hal tersebut tidak berlaku pada *testing* dimana pada kategori kedua kategori menunjukkan nilai SVM lebih baik dengan nilai akurasi sebesar 90% pada kategori *recurve* dan nilai akurasi sebesar 76 % pada kategori *barebow*. Jika merujuk pada *confution matrix* yang ada pada gambar 7 dan gambar 8, terjadi misdeteksi pada *testing* dimana misdeteksi tersebut banyak terjadi pada kedua metode di kategori *barebow*. Misdeteksi terjadi pada *class 3* untuk metode RF dan *class 2* untuk metode SVM. Tidak dapat dipungkiri pada kategori *recurve* juga terdapat misdeteksi pada kedua metode, hanya saja misdeteksi yang terjadi tidak semasif pada kategori *barebow*.

Misdeteksi pada kategori *recurve* bisa terjadi karena terjadinya tumpang tindih dalam proses penembakan. Bisa dilihat pada gambar 11, jika dilihat secara kasat mata memang tidak ada perbedaan sama sekali pada gambar 11(A) dan 11(B). Hanya saja, kedua gambar tersebut tidak dalam satu *class* yang sama sehingga misdeteksi tersebut bisa terjadi. Kemudian pada misdeteksi kategori *barebow* bisa terjadi karena dua hal, yaitu sama tumpang tindih seperti yang telah dijelaskan pada kategori *recurve* sebelumnya dan gaya menembak atlet yang berbeda-beda. Bisa dilihat pada gambar 12 dimana digunakan contoh pada *class Expand and Shoot*, kedua atlet tersebut sama-sama telah melepaskan anak panahnya. Terlihat pada gambar 12(A) atlet 1 melakukan gerakan *shooting* yang mirip seperti gerakan *Draw and Aim* dimana tangan masih berada di sekitar area wajah, sedangkan gambar 12(B) atlet 2 melakukan gerakan *shooting* dengan tangan yang berada di belakang kepala. Pada dasarnya tidak ada yang salah pada kedua teknik tersebut pada panahan kategori *barebow*, hanya saja ketika dilakukan pendeteksian pada komputer, kemungkinan terjadi misdeteksi sangat besar sehingga mempengaruhi hasil dari *training* maupun *testing* yang ada.



Gambar 11. Contoh gambar yang tumpang tindih (A) Draw and Aim (B) Expand and Shoot



Gambar 12. Contoh gambar gaya menembak atlet (A) Atlet 1 (B) Atlet 2

Disaat yang bersamaan, pada gambar 9 dan 10 terlihat hasil yang cukup jelas secara kasat mata. Dalam hal performa pembelajaran dan tingkat kepentingan fitur, ada perbedaan yang signifikan antara model SVM dan

Random Forest. Sementara model Random Forest menunjukkan tanda-tanda overfitting yang ditandai dengan jarak besar antara kinerja pengujian dan pelatihan, model SVM (baik Recurve maupun Barebow) menunjukkan kinerja yang lebih baik, menunjukkan kemampuan generalisasi yang baik pada data baru. Kedua model menemukan pola serupa dalam analisis tingkat kepentingan fitur, di mana beberapa fitur awal memiliki pengaruh dominan, tetapi dengan distribusi yang berbeda. SVM menunjukkan penurunan bertahap dalam nilai kepentingan fitur, sementara Random Forest menunjukkan penurunan yang lebih tajam. Meskipun masalah overfitting pada Random Forest dapat diatasi dengan teknik pruning atau optimasi hyperparameter yang tepat, temuan ini menunjukkan bahwa dalam situasi ini, model SVM mungkin merupakan pilihan yang lebih baik. Selain itu, penting untuk dicatat bahwa kedua model secara konsisten menunjukkan bahwa metrik performa dan fitur teknik menembak adalah faktor yang paling memengaruhi pembuatan prediksi.

5. Kesimpulan

Berdasarkan hasil analisis, Penelitian lengkap menunjukkan bahwa model SVM menunjukkan kinerja yang lebih unggul dan konsisten dibandingkan dengan model RF dalam mengklasifikasikan gerakan memanah untuk kategori recurve dan barebow. Meskipun Random Forest mencapai akurasi *training* yang sempurna (100%) untuk kedua kategori, RF menunjukkan *overfitting* yang cukup besar, yang dibuktikan dengan kinerja *testing* yang menurun. Sebaliknya, SVM menunjukkan keseimbangan yang lebih baik, dengan akurasi testing sebesar 90% untuk kategori recurve dan 76% untuk kategori barebow.

Faktor-faktor penentu yang memengaruhi akurasi deteksi terutama berkaitan dengan kerumitan gerakan memanah itu sendiri. Kesalahan deteksi dapat dikaitkan dengan dua faktor utama: pertama, tumpang tindihnya gerakan di antara kelas-kelas yang menunjukkan kemiripan visual yang signifikan, dan kedua, perbedaan dalam metode pemotretan di antara para atlet, khususnya dalam kategori barebow. Hal ini paling jelas terlihat di kelas Expand and Shoot, di mana variasi dalam teknik individu atlet dapat menyebabkan postur tubuh yang sangat bervariasi, bahkan dalam fase teknik yang sama.

Learning Curve dan *Feature Importance* mendukung kesimpulan bahwa SVM menawarkan generalisasi yang unggul, yang dibuktikan dengan perbedaan yang lebih sempit antara kinerja pelatihan dan pengujian. Data ini menunjukkan bahwa untuk sistem klasifikasi keterampilan memanah, SVM adalah pilihan yang lebih unggul, terutama dalam menangani banyak gerakan dan kompleksitas yang melekat pada olahraga panahan.

Daftar Pustaka

- [1] BeomJun, J., & SeongKi, K. Comparative analysis of OpenPose, PoseNet, and MoveNet model for pose estimation in mobile devices. *Treatment du signal* 39.1 (2022): 119.
- [2] Kawaguchi, M., Mitake, H., & Hasegawa, S. Archery shots visualization by clustering and comparing from angular velocities of bows. *Proceedings of the Augmented Humans International Conference*. 2020.
- [3] Kitamura, T., Teshima, H., Thomas, D., & Kawasaki, H. Refining OpenPose with a new sports dataset for robust 2D pose estimation. In: *Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision* (2022). p. 672-681.
- [4] Saptiawan, L. H., Purboyo, T.W., & Ansor, A. S. Pose Effect Analysis with Accuracy of Archery using Manhattan Distance Algorithm. *Journal of Engineering and Applied Sciences* 15.3 (2020): 819-825.
- [5] Li, M., Zhou, Z., & Liu, X. Multi-person pose estimation using bounding box constraint and LSTM. *IEEE Transactions on Multimedia* 21.10 (2019): 2653-2663.
- [6] Siddiqui, H. U., Younas, F., Rustam, F., Flores, E. S., Ballester, J. B., Dies, I. d., ... Ashraf, I. Enhancing Cricket Performance Analysis with Human Pose Estimation and Machine Learning. *Sensors* 23.15 (2023): 6839.
- [7] Glenn Jocher et al. YOLOv8. <https://github.com/ultralytics/ultralytics>, 2023 [Diakses November 2024].
- [8] Dong, C., & Du, G. (2024). An enhanced real-time human pose estimation method based on modified YOLOv8 framework. *Scientific Reports*, 14(1), 8012.
- [9] Jaouedi, N., Perales, F. J., Buades, J. M., Boujnah, N., & Bouhleb, M. S. Prediction of human activities based on a new structure of skeleton features and deep learning model. *Sensors* 20.17 (2020): 4944
- [10] Munea, T. L., Jembre, Y. Z., Weldegebriel, H. T., Chen, L., Huang, C., & Yang, C. The progress of human pose estimation: A survey and taxonomy of models applied in 2D human pose estimation. *IEEE Access* 8 (2020):

133330-133348.

- [11] Hussain, M. (2024). Yolov5, yolov8 and yolov10: The go-to detectors for real-time vision. *arXiv preprint arXiv:2407.02988*.
- [12] Tang, X., & Zhao, S. (2024). The application prospects of robot pose estimation technology: exploring new directions based on YOLOv8-ApexNet. *Frontiers in Neurorobotics*, 18, 1374385.
- [13] Avcı, C., Budak, M., Yağmur, N., & Balçık, F. (2023). Comparison between random forest and support vector machine algorithms for LULC classification. *International Journal of Engineering and Geosciences*, 8(1), 1-10.
- [14] Terven, J., Córdova-Esparza, D. M., & Romero-González, J. A. (2023). A comprehensive review of yolo architectures in computer vision: From yolov1 to yolov8 and yolo-nas. *Machine Learning and Knowledge Extraction*, 5(4), 1680-1716.
- [15] E. I. Verdiguier dan R. Z. Milla, "An evaluation of Guided Regularized Random Forest for classification and regression tasks in remote sensing," *Int J Appl Earth Obs Geoinformation*, vol. 88, no. 28, pp. 1-13, 2020.
- [16] Boateng, E. Y., Otoo, J., & Abaye, D. A. (2020). Basic tenets of classification algorithms K-nearest-neighbor, support vector machine, random forest and neural network: A review. *Journal of Data Analysis and Information Processing*, 8(4), 341-357.
- [17] Lestari, D., & Subekti, L. Implementasi Chatbot pada Telegram sebagai Monitoring Assistant dengan Analisis Teks Klasifikasi Menggunakan Metode Support Vector Machine. *Journal of Internet and Software Engineering*, 5(2), 68-74.
- [18] Tran, G. S., Nghiem, T. P., Nguyen, V. T., Luong, C. M., & Buric, J. C. Improving Accuracy of Lung Nodule Classification Using Deep. *Journal of healthcare engineering*. (2019)
- [19] Chicco, D., Tötsch, N., & Jurman, G. (2021). The Matthews correlation coefficient (MCC) is more reliable than balanced accuracy, bookmaker informedness, and markedness in two-class confusion matrix evaluation. *BioData mining*, 14, 1-22.
- [20] Yilmaz, A., Demircali, A. A., Kocaman, S., & Uvet, H. (2020). Comparison of deep learning and traditional machine learning techniques for classification of pap smear images. *arXiv preprint arXiv:2009.06366*