DOI: 10.22146/jnteti.v13i1.8145

Identification of Fencing Athletes Based on Anthropometric Measurements Using MediaPipe Pose

Bagas Alif Fimaskoro[1], Suci Aulia[1], Dery Rimasa[2]

[1] School of Applied Science, Telkom University, Bandung, Indonesia
[2] Sports Physical Training Study Program, Universitas Pendidikan Indonesia, Bandung, Indonesia

[Received: 13 June 2023, Revised: 22 August 2023, Accepted: 8 December 2023]
Corresponding Author: Suci Aulia (email: suciaulia@telkomuniversity.ac.id)

ABSTRACT — Over time, numerous developments in digital technology have benefited people, including anthropometric measurements that provide information on an athlete's ability in sports. The use of digital technology in sports must continue, particularly in the National Sports Committee of Indonesia (Komite Olahraga Nasional Indonesia, KONI) of Bandung City. This study proposed a technique for classifying and identifying fencing athletes' talents. This work developed a methodology for evaluating sports talent based on anthropometric measurements of athletes' bodies using the posture detection approach. Fencing and nonfencing athletes in KONI Bandung City were categorized using this talent identification. This study used 36 datasets of body posture images from various skills of the sport. These images were in JPEG or JPG format with a resolution of 3,024 × 4,032 and were acquired using a Canon EOS 1300D camera. This study utilized four points landmarks, which are usually used as measurement components in KONI, to categorize fencing athletes. The four points are shoulder (S), elbow (E), index (I), and hip (H) landmarks. The testing was done using three different dataset settings. According to the test results of all scenarios, scenario 2 had the highest accuracy. This scenario was able to categorize fencing and nonfencing athletes with an accuracy rate of 89% and an average processing time of less than 3 s per image.

KEYWORDS — Anthropometry, Image Processing, Pose Detection, Fencing, Talent Identification.

I. INTRODUCTION

Sport is a movement carried out precisely according to the type of movement with various goals and directions. It is essential for everyone's life as it provides many benefits such as affecting the development of physical growth [1]. The manual system to measure the athletes' body anthropometry is still utilized today, especially in the National Sports Committee of Indonesia (Komite Olahraga Nasional Indonesia, KONI) of Bandung City. Therefore, innovation within KONI, particularly Bandung City KONI, should be made so that KONI is able to fulfill its responsibility on the athlete development. Taking advantage of advancements in technology throughout the digital age, several tools utilizing image processing have been developed to facilitate body anthropometric measurements in athletes and sport talent assistance. Anthropometry is a science related to measuring the dimensions of the human body [2]. Anthropometric measurements are useful for designing equipment and facilities for daily activities [3]. Previous studies have conducted anthropometric measurements of the body. Aligning to [4], measurements in this study were only applied to multiple specific areas. This study measured height and weight using the morphological image processing method. Another study has utilized MATLAB software to create an anthropometric measurement for all athletes [2].

As a continuation, this research proposed a design of an automatic anthropometric measurement system based on image processing to be implemented at the Bandung City KONI, mainly for fencing athletes. This identification can be utilized to differentiate body measurements between fencers and nonfencing athletes with focus on the upper body, namely arm length, shoulder width, and waist width. Human pose estimation technologies are currently being extensively researched in several fields such as sports, home training, surveillance, job monitoring, cultural activities, gesture control, home elder care, and virtual world avatars. Human pose estimation is usually separated into 2D and 3D model estimation methods, single-object and multiple-object methods based on the number of target subjects, monocular and multiview image methods based on the quantity of shooting cameras, and single image and video methods based on the input type [5]–[9]. Notably, human pose estimation is divided into single-stage and two-stage methods based on the structure of the deep learning process. Two groups of single-stage algorithms immediately convert input images into 3D body joint coordinates: regression-based approaches [10], [11] and detection-based methods [12], [13]. The regression-based methods directly estimate the location of connections to the location of the root joint [10], [11] or the joint angles utilizing a kinematic model with numerous joints and bones [12], [13]. On the contrary, the detection-based methods predict a likelihood heatmap for each joint, whose location is determined by using the heatmap's maximum likelihood. Lifting 2D to 3D, also known as leveraging 2D pose estimation discoveries for 3D human pose estimation, is an active area of research. It is because 2D pose estimation has a larger number of datasets in real-world scenarios that provide accurate joint coordinates compared to 3D pose estimate [14]. Long short-term memory (LSTM) has been used to derive advantage of the interactions between joints [15]. Meanwhile, generative adversarial networks (GANs) are frequently employed to create a more realistic 3D human pose [16]. The deep learning methods of 3D pose estimation from 2D images should overcome challenging issues such as a shortage of in-the-wild datasets, high demand for different posture data, depth ambiguities, and vast search state space for each joint despite ongoing technological advancements. A powerful personal computer (PC) with numerous graphic processing units (GPUs) is also required to run deep learning software. Due to the significant coupling between processes, changing an application of perception to include additional processing steps or inference models can be challenging. It also takes time to create the same application for several platforms because these processes typically need to be optimized for a particular device to function correctly and effectively. By interpreting and integrating several perception models into consistent pipelines, MediaPipe solves these problems. The pipeline architecture contains the instructions required to draw conclusions from the sensory data and produce the perceived outcomes. Since MediaPipe components have a similar interface focused on time-series data, reusing them in other pipelines throughout subsequent projects is simple [17]. Subsequently, every pipeline can function uniformly across multiple platforms, allowing the practitioner to create the application on computers and later deploy it on mobile devices.

This study employed a two-stage pose estimation method using MediaPipe Pose (MPP) to execute a human pose estimate package on a single-board computer. This study used the MPP, Google-provided open source and cross-platform framework, to estimate the 2D human joint coordinates in each image database. The MPP uses machine learning to construct pipelines and interpret cognitive data presented as images [17]–[19]. This fencing athlete's identification is expected to help the Bandung City KONI to take anthropometric body measurements quickly and precisely. The rest of this paper is organized as follows. Sections 2 and 3 present the proposed framework for pose estimation and talent identification. Section 4 presents the experiment's results. Finally, Section 5 presents our conclusions.

II. ANTHROPOMETRIC MEASUREMENT METHODS

A. ANTHROPOMETRY

Anthropometric measurements are used to determine a suitable body for several sports [20]. Anthropometry is the measurement of the human body, including the determination of length, width, diameter, and circumference; it also involves the calculation of ratios and proportions based on two or more measurements to identify body shape, size, and topography [21].

Anthropometry is a scientific field that focuses on the measurement and analysis of human body dimensions [22]. Anthropometry is widely applied in sports. One of the essential aspects in identifying sports talent and achieving good sports performance is strengthening the ideal body size at peak performance. Anthropometry utilizes information obtained from body measurements to distinguish people, groups, and other entities. An anthropomorphic system of proportions based on human body dimensions can be used to differentiate between adults and children. One method is static anthropometry, which involves measuring the body in conventional positions without doing any movements. Meanwhile, the other method is dynamic anthropometry, which is carried out when a person performs specific movements related to certain tasks.

B. RED, GREEN, BLUE (RGB) COLOR

Each color in a color image is represented by 4, 8, 16, or 24 bits of information for each pixel, with the number of colors ranging from 16; 256; 65,536; or 16 million. Red, green, and blue (RGB) make up the main hues of the actual color. RGB color refers to the combination of the three primary color-forming elements, which will eventually create a color image [23]. Red is a matrix that states the degree of brightness for the red color (for example, for a grayscale of 0–255, a value of 0 represents black and 255 represents red). Green is a matrix that states the degree of brightness for green. Blue is a matrix that states the degree of brightness for blue [24]. A 3 × 3 matrix transformation can convert the RGB data found in a pixel into the CIE XYZ color space. Tristimulus values, a configuration of three light-linear components that perform the CIE color-matching function, were used in this process. Specific colors are represented as always positive values in the XYZ color space [25].

C. GRAYSCALE

Greyscale refers to an image that possesses a grayscale level [26]. The grayscale image consists of the RGB colors, all of which possess an identical intensity [27]. In contrast to color images, where each pixel requires three intensities, grayscale images only require one intensity value. The degree of gray is represented by the 8-bit integer intensity of the grayscale image, which ranges from 0 to 255 with 0 indicating black and 255 indicating white; the value between them is the degree of gray [28].

D. ARM LENGTH CALCULATION

The measurement of the arm's length can be determined by calculating the number of pixels measured from the fingertips of the right hand to those of the left hand and multiplying them by the ratio of the hand's length, which is commonly denoted as $hand_l$. This study calculated the pixel length value of black objects represented by a plus sign (+) on a background and compared the actual length size of the object in centimeters to determine the $hand_l$ ratio. Meanwhile, $Object_{pl}$ denotes the number of object pixel lengths. The following is the equation for calculating arm length [29]:

$$Arm\ length = Object_{pl} \times hand_l. \tag{1}$$

E. WAIST CIRCUMFERENCE CALCULATION

The waist circumference can be determined by counting the number of pixels of the object's waist in the front view and the number of pixels of the object's waist in the side view. Determining the diameter 1 (d_1) and diameter 2 (d_2) waist on the object will yield the pixel values. Subsequently, the d_1 and d_2 values in pixels are converted into centimeters. Equation (2) shows the formula for calculating waist circumference [30]:

$$Waist\ circumference = \frac{1}{2} \times \pi \times (d_1 + d_2). \tag{2}$$

F. HUMAN POSE ESTIMATION

1) THE 2D POSE ESTIMATION

Human pose estimation has been thoroughly researched over the past few years, from pictorial structures [31], [32] to more current convolutional neural network (CNN) techniques [33]. Based on prior study, there are two groups of pose estimation algorithms: regression-based and detection-based methods. In detection-based approaches for posture estimation, each pixel in a heatmap represents a joint's detection score, which consider pose estimation as a heatmap prediction problem [34]. Prior study investigated the idea of stacked architectures, residual connections, and multiscale processing [35]. They subsequently developed the stacked hourglass network, which significantly improved the performance of 2D posture estimation issues. Since then, state-of-the-art techniques have proposed intricate modifications to the stacked hourglass structure. A pyramid residual module (PRM) was used in place of the residual unit in proposed attention model [36], which was based on the conditional random fields (CRF) [37]. GANs have been employed to enhance the capability of learning structural information [38] and to enhance heatmaps by learning more reasonable predictions [39]. However, joint coordinates are rarely provided by detection methods. the argmax function is frequently employed as a postprocessing step to recover the pose in the (x, y) coordinates.

On the other hand, the output of a regression-based strategy can be the joint coordinates since it uses a nonlinear function to translate the input to the desired output. Following this paradigm, prior research has offered a comprehensive method for body part detection based on cascade regression [40] and proposed iterative error feedback [41]. The softargmax function has been proposed to directly convert heatmaps to joint coordinates to address this shortcoming and afterward enable the transformation of detection methods into regression methods [42]. Regression methods have the primary benefit of being more frequently completely differentiable than detection approaches. It implies that the result of the posture estimation can be employed in further processing, thereby enabling the system-wide optimization.

2) THE 3D POSE ESTIMATION

Deep architectures are now more advanced than depth sensors [43] because they can learn detailed 3D models from RGB images [44] and have access to high-quality data [45]. The 3D posture estimation problem is distinguished into two components [46]. The initial component involves processing 2D pose estimation while considering the camera coordinates. Then, the second component involves the utilization of a nonparametric shape model to match the estimated poses for 3D representations. Reference [47] suggests to reduce the data variance and provide a bone representation of the human position. However, since the inaccuracy escalates with distance from the root joint, tasks that rely on the human body's extremities may be adversely affected by this structural change. The volumetric stacked hourglass architecture was then proposed [12].

G. MEDIAPIPE POSE (MPP)

This study employed MPP to estimate each database image's 2D human joint coordinates. Google offers MPP, an open-source and cross-platform framework. MediaPipe is a framework designed for artificial intelligence in an application. MPP utilizes machine learning to build pipelines (depicted by the blue lines in Figure 1) and decipher cognitive data shown as images. Additionally, MediaPipe makes integrating perception technology into applications and demos on various hardware platforms easy. MediaPipe's powerful configuration syntax and assessment tools enable gradual enhancements to perception pipelines.

The three main parts of MediaPipe are a framework for inference from sensory input, a set of tools for performance evaluation, and a library of reusable inference and processing pieces known as calculators. MediaPipe also contains TensorFlow, which supports a device's GPU and central processing unit (CPU) flow acceleration. The initial step that can be achieved from the MediaPipe is to use a camera or webcam. Next, the MediaPipe can display a 33-point skeleton, representing the human body [18]. Figure 1 displays the BlazePose, a lightweight machine learning architecture, that

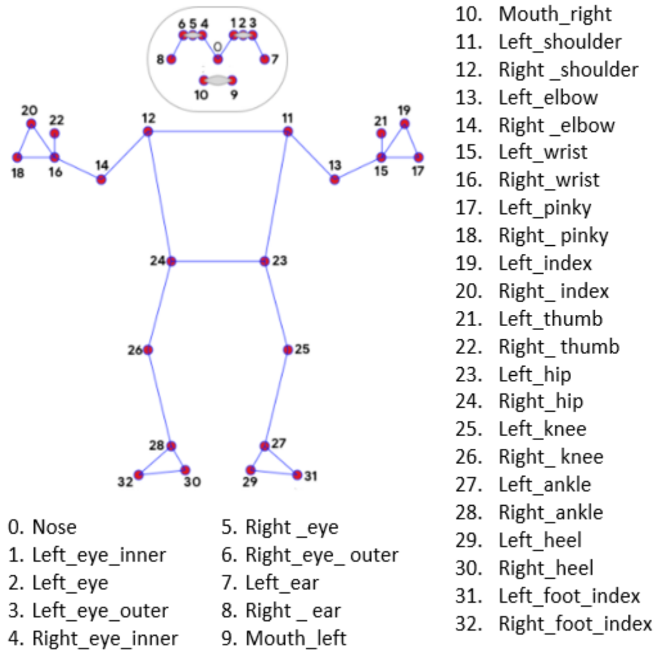


Figure 1. Landmarks in MPP [18], [19].

MPP uses to retrieve 33 2D landmarks from the human body [19]. BlazePose uses CPU inference to deliver real-time performance on mobile devices and desktop computers.

The inverse ratio must be multiplied by the pixel values on the y-axis when utilizing normalized coordinates for pose estimation. This study employed four of the calculated MPP landmarks to estimate arbitrary poses and motions. These landmarks were shoulder (X_1), index (X_2), hip (X_3), and elbow (X_4), which are depicted in Figure 2.

III. METHODOLOGY

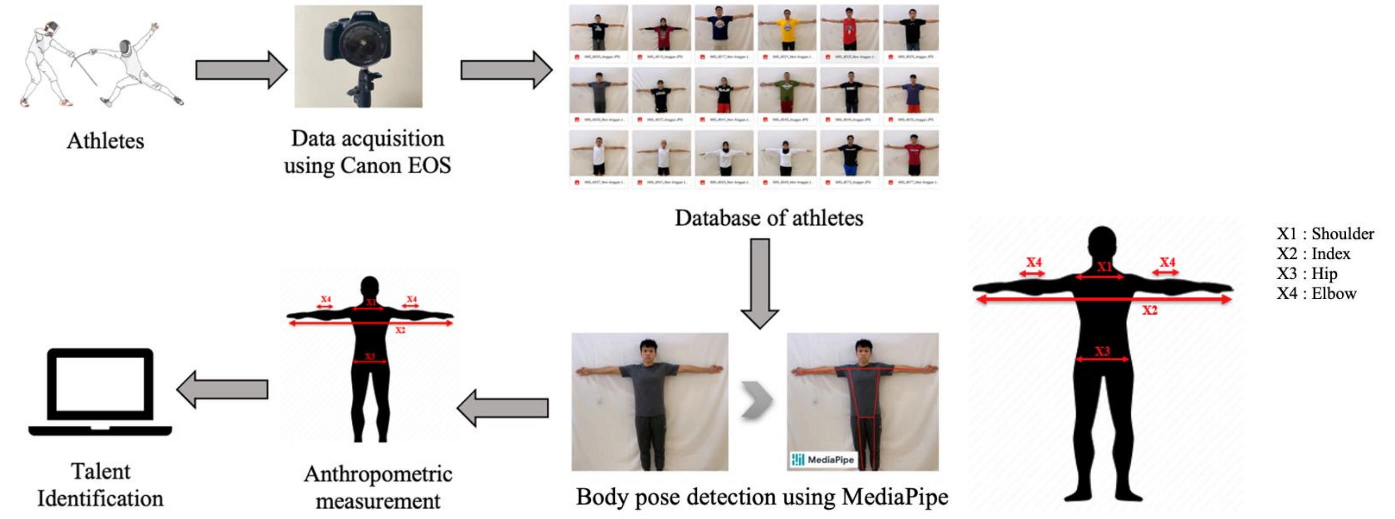
A. SYSTEM MODEL

The ultimate goal of this research is to analyze and compare the results of the original measurements with measurements made using software to determine the talent of budget athletes. In addition, this output of this research can offer recommendations to Ikatan Anggar Seluruh Indonesia (IKASI) and Bandung City KONI in carrying out automatic body anthropometric measurements to increase the convenience of trainers and manufacturing organizations. Figure 2 illustrates the system design started from image acquisition until talent identification (fencing and nonfencing).

The captured image underwent an image processing process to detect the upper body. This system used a white background with the size of 3 × 2 m. After the results of anthropometric measurements were obtained, data on the athlete's talent assistance were obtained as a benchmark for the ratio values that were used to identify sports talent in athletes. Figure 3 exhibits how images undergo digital image processing to obtain the required information. The following is a detailed explanation of Figure 3.

1) DATA COLLECTION

Data collection was the first step that must be done to determine the original data from an athlete. The dataset was collected using manual measurement tools that were done at Bandung City KONI. This dataset consisted of 36 images in JPG format and was distinguished in two classes: fencing and



(a) (b)

Figure 2. Images of (a) system identification block diagrams, (b) landmark in this research.

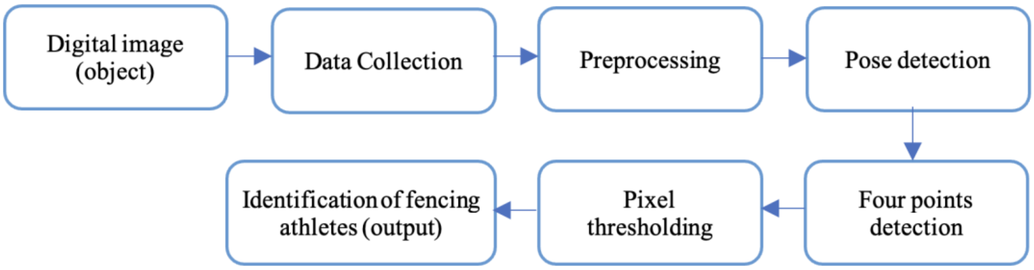


Figure 3. Block diagram of image processing.

nonfencing athletes. The participants, both fencing and nonfencing athletes, were students of the Faculty of Sport and Health Education, Universitas Pendidikan Indonesia, and consisted of men and women aged 20–22 years. Like most sports, fencing has a diverse range of participants. Physical traits, such as age, weight, height, power, speed, and general athleticism, are varied. In the dataset, fencing participants have a minimum height of 175 cm for men and 160 cm for women.

2) PREPROCESSING

Preprocessing was one of the first methods employed in image processing. Preprocessing seeks to minimize or increase data size, remove any existing noise, clarify image features, and change the original data into expected data that will later form a required object [16]. This process was carried out in this study for scenarios 2 and 3. In scenario 2, the contrast of the dataset was improved. In scenario 3, on the other hand, the original dataset was turned into black-and-white images. Preprocessing can also be done to improve an image from existing noise. It was done to remove parts of the input image that were not needed for the subsequent process.

3) POSE DETECTION

Pose detection using MediaPipe resulted in 33 points [17]–[19], as seen in Figure 1. Nevertheless, as previously mentioned, this study only utilized four points landmark as variables: shoulder (X_1), index (X_2), hip (X_3), and elbow (X_4) as seen in Figure 2.

4) PIXEL THRESHOLDING

Thresholding aims to determine the minimum and maximum pixel values in an image [20]. Pixel thresholding is one of the main techniques in clustering or classifying athletes into fencing and nonfencing categories. In this study, pixel thresholding segmented the measurement results (pixel) for each landmark (X_1), (X_2), (X_3), and (X_4) based on the actual size (pixel) of fencing athletes that had been simulated in the training process.

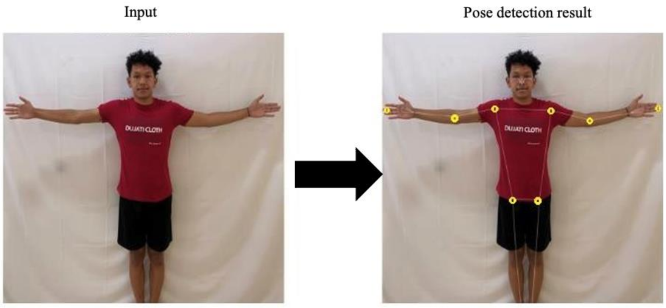
B. ACQUISITION DESCRIPTION

In this study, testing was carried out with 36 athletes of all types of sports in the Bandung City KONI, all of whom were of different body size. This testing was carried out with other measuring point parameters: shoulder (S), elbow (E), index (I), and hip (H). These images were taken using a camera lens with sufficient lighting so that there are no disturbing shadows when taking the picture. The camera used was Canon EOS 1300D with a height of ± 1 m from the floor surface, 2 m from the camera to the object, and angle of 90° between the camera and the object.

IV. RESULTS

In this research, testing was performed based on three dataset scenarios, each using 36 datasets taken directly at the Bandung City KONI. In scenario 1, the testing was conducted using original datasets. In scenario 2, the testing employed different datasets from the original ones because the contrast was changed. In scenario 3, the testing used datasets that were converted from their original RGB color into black and white.

The accuracy was calculated by comparing the measurement by the system proposed in this this research with the actual size by the Bandung City KONI team. The measure of a fencing athlete consisted of four points components: elbow length, shoulder width, index length, and hip width as seen in Figure 4.



No	Landmark Point	Distance (pixel)
1.	Shoulder	577
2.	Elbow	1377
3.	Hip	285
4.	Index	2441

Figure 4. Pose detection measurement process.

Figure 4 displays the process to obtain pose points on the identified image dataset. Using the pose detection method, pixels can be obtained from each measured point. To detect 2D landmarks from the human body on dataset, first step was to locate the pose region of interest (ROI) on image. Thirty-three landmark points were detected accurately in this measurement based on BlazePose [18]. However, as previously mentioned, only four points landmarks (S, E, I, and H) were used to classify the body pose image into two categories: fencing and nonfencing athletes. In Figure 4, these landmarks are demonstrated as yellow keypoints.

A. TESTING RESULTS OF SCENARIO 1 (ORIGINAL)

As shown in Table I, the nine-detection point scheme achieved the highest accuracy value in this test by simulating the effect of the measurement point parameters on the body pose detection method. The nine-detection points scheme is the variation of measurement among four points landmarks: S, E, I, and H. In scenario 1, the pose detection method yielded the highest overall accuracy of 64% when identifying the talent of a fencing athlete. Using the S, E, I, H and E, I, H detection point schemes, those values were acquired. Additionally, based on scenario 1, the system could differentiate fencing athletes from nonfencing athletes in an average of 76 s across all datasets, or ±2.89 seconds per image.

B. TESTING RESULTS OF SCENARIO 2 (CONTRAST ENHANCEMENT)

Similar to scenario 1, testing for scenario 2 used a nine-detection point scheme. As shown in Table I, this test aims to determine how contrast enhancement affects the accuracy rate. The overall test results in scenario 2 were better than those of scenario 1. Scenario 2 and scenario 1 yielded the highest accuracy of 89% and 64%, respectively. There was a difference in the fastest detection time of E, I, and H detection points, with the fastest being 69 s. Therefore, the best accuracy value test was obtained at the E, I, and H measurement points with a detection process time of 108 s. In other words, it could detect fencing athletes ± 3 s per image.

C. TESTING RESULTS OF SCENARIO 3 (BLACK AND WHITE IMAGES)

In scenario 3, the original dataset of RGB color was changed to a black-and-white color to imitate the effects of measuring point parameters using the body posture detection method as shown in Table I. The overall test results for scenario 3 indicated that the pose detection approach achieved the highest accuracy of 58% in identifying the talent of the fencing athletes. This resulut was lower than the outcome of scenario 2, which had the highest accuracy of 89%. Therefore, employing RGB images with preprocessing, such as contrast enhancement in scenario 2, is better suitable for classifying posture image into two categories, namely the talent of fencing and nonfencing athletes, using an anthropometric body measurement system. Figure 5 provides a comparison of all tests conducted.

TABLE I
ACCURACY RATE IN SCENARIO 1

Landmark	Detection Points	Accuracy Rate (%)			Process Length (s)		
		S1	S2	S3	S1	S2	S3
4 points	S, E, I, H	64	89	58	74	104	100
3 points	S, E, I	64	86	58	72	106	100
	E, I, H	64	89	58	69	108	91
	I, H, S	64	86	58	71	109	87
	H, S, E	64	81	56	79	106	97
2 points	S, E	64	78	56	84	97	99
	E, I	64	83	56	74	106	98
	I, H	64	83	58	79	106	101
	H, S	64%	78	58	86	107	100
Average		64	83	57	76	105	85

Note: S1: Scenario 1, S2: Scenario 2, S3: Scenario 3

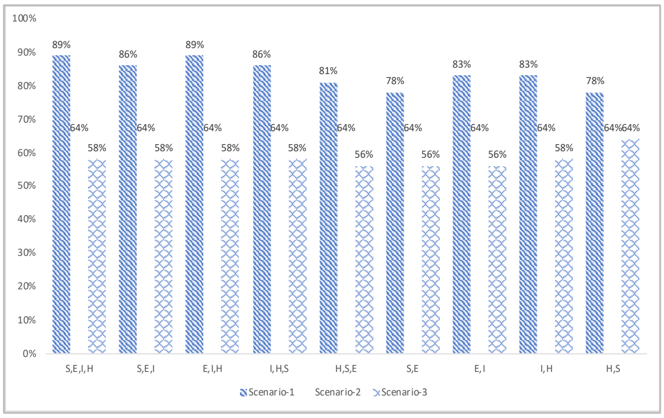


Figure 5. Comparison of test result.

V. CONCLUSION

This study highlights the potential of anthropometric measurements and posture detection for identifying sports talent. This talent identification classifies fencing and nonfencing athletes in KONI Bandung. This study used 36 datasets of body posture images with different talents from the KONI. Three dataset scenarios were used to conduct the tests. The types of images used in each scenario were varied. They were original colouring, contrast-enhancing, and converting RGB to BW images. With an average processing time of less than 3 s per image, scenario 2 had the highest accuracy in classifying fencing and nonfencing athletes for all scenarios, with an accuracy rate being 89%. This accuracy value must be increased in future work by considering the parameters of acquisition distance, number of datasets, gender-based clustering, and deep learning methods that can be classified in real time.

CONFLICTS OF INTEREST

The authors declare no conflict of interest.

AUTHORS' CONTRIBUTIONS

Conceptualization, Bagas Alif Fimaskoro, Suci Aulia, and Dery Rimasa; methodology, Bagas Alif Fimaskoro and Suci Aulia; formal analysis, Bagas Alif Fimaskoro and Suci Aulia; investigation, Bagas Alif Fimaskoro and Dery Rimasa; resources, Bagas Alif Fimaskoro and Dery Rimasa; data curation, Bagas Alif Fimaskoro and Suci Aulia; writing—original draft preparation, Bagas Alif Fimaskoro and Suci Aulia; writing—reviewing and editing, Bagas Alif Fimaskoro and Suci Aulia; visualization, Bagas Alif Fimaskoro and Suci Aulia; supervision, Suci Aulia and Dery Rimasa; project administration, Bagas Alif Fimaskoro and Suci Aulia; funding acquisition, Bagas Alif Fimaskoro.

REFERENCES

[1] I. Mahfud, R. Yuliandra, and A. Gumantan, "Model latihan dribling sepakbola untuk pemula usia SMA," *Sport Sci. Educ. J.*, vol. 1, no. 2, pp. 1–9, Jul. 2020, doi: 10.33365/ssej.v1i2.823.

[2] R.M. Kurniawan, T.N. Damayanti, and D. Rimasa, "Aplikasi pengukuran antropometri tubuh pada atlet di KONI Kota Bandung berbasis image processing," *eProc. Appl. Sci.*, vol. 7, no. 4, pp. 843–860, Aug. 2021.

[3] H. Purnomo, "Pengukuran antropometri tangan usia 18 sampai 22 tahun Kabupaten Sleman Yogyakarta," *Pros. Ind. Eng. Nat. Conf. (IENACO)*, 2014, pp. 106–112.

[4] S. Aulia, F.E. Satria, and R.D. Atmaja, "Sistem pengukur tinggi dan berat badan berbasis morphological image processing," *ELKOMIKA, J. Tek. Energi Elekt. Tek. Telekomun. Tek. Elektron.*, vol. 6, no. 2, pp. 219–231, May 2018, doi: 10.26760/elkomika.v6i2.219.

[5] N. Sarafianos, B. Boteanu, B. Ionescu, and I.A. Kakadiaris, "3D human pose estimation: A review of the literature and analysis of covariates," *Comput. Vis. Image Underst.*, vol. 152, pp. 1–20, Nov. 2016, doi: 10.1016/j.cviu.2016.09.002.

[6] Y. Chen, Y. Tian, and M. He, "Monocular human pose estimation: A survey of deep learning-based methods," *Comput. Vis. Image Underst.*, vol. 192, pp. 1–20, Mar. 2020, doi: 10.1016/j.cviu.2019.102897.

[7] J. Wang *et al.*, "Deep 3D human pose estimation: A review," *Comput. Vis. Image Underst.* vol. 210, pp. 1–21, Sep. 2021, doi: 10.1016/j.cviu.2021.103225.

[8] E. Alam, A. Sufian, P. Dutta, and M. Leo, "Vision-based human fall detection systems using deep learning: A review," *Comput. Biol., Med.*, vol. 146, pp. 1–22, Jul. 2022, doi: 10.1016/j.compbiomed.2022.105626.

[9] M.M.E. Yurtsever and S. Eken, "BabyPose: Real-time decoding of baby's non-verbal communication using 2D video-based pose estimation," *IEEE Sens. J.*, vol. 22, no. 14, pp. 13776–13784, Jul. 2022, doi: 10.1109/JSEN.2022.3183502.

[10] S. Li and A.B. Chan, "3D human pose estimation from monocular images with deep convolutional neural network," in *Computer Vision - ACCV 2014*, D. Cremers, I. Reid, H. Saito, and M.-H. Yang, Eds., Cham, Switzerland: Springer, 2015, pp. 332–347, doi: 10.1007/978-3-319-16808-1_23.

[11] X. Zhou et al., "Deep kinematic pose regression," in *Computer Vision - WCCV 2016 Workshops*, G. Hua and H. Jégou, Eds., Cham, Switzerland: Springer, 2016, pp. 186–201, doi: 10.1007/978-3-319-49409-8_17.

[12] G. Pavlakos, X. Zhou, K.G. Derpanis, and K. Daniilidis, "Coarse-to-fine volumetric prediction for single-image 3D human pose," *2017 IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, 2017, pp. 1263–1272, doi: 10.1109/CVPR.2017.139.

[13] D.C. Luvizon, D. Picard, and H. Tabia, "2D/3D pose estimation and action recognition using multitask deep learning," *2018 IEEE/CVF Conf. Comput. Vis. Pattern Recognit.*, 2018, pp. 5137–5146, doi: 10.1109/CVPR.2018.00539.

[14] D. Tome, C. Russell, and L. Agapito, "Lifting from the deep: Convolutional 3D pose estimation from a single image," *2107 IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, 2017, pp. 5689–5698, doi: 10.1109/CVPR.2017.603.

[15] J. Wang, S. Huang, X. Wang, and D. Tao, "Not all parts are created equal: 3D Pose estimation by modeling bi-directional dependencies of body parts," *2019 IEEE/CVF Int. Conf. Comput. Vis. (ICCV)*, 2019, pp. 7770–7779, doi: 10.1109/ICCV.2019.00786.

[16] B. Wandt and B. Rosenhahn, "RepNet: Weakly supervised training of an adversarial reprojection network for 3D human pose estimation," *2019 IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, 2019, pp. 7774–7783, doi: 10.1109/CVPR.2019.00797.

[17] V. Bazarevsky *et al.*, "BlazePose: On-device real-time body pose tracking," 2020, *arXiv: 2006.1024*.

[18] R. Soneja, S. Prashanth, and R. Aarthi, "Body weight estimation using 2D body image," *Int. J. Adv. Comput. Sci. Appl.*, vol. 12, no. 4, pp. 321–326, Apr. 2021, doi: 10.14569/IJACSA.2021.0120440.

[19] J.W. Kim, J.Y. Choi, E.J. Ha, and J.H. Choi, "Human pose estimation using mediapipe pose and optimization method based on a humanoid model," *Appl. Sci.*, vol. 13, no. 4, 2023, doi: 10.3390/app13042700.

[20] D.D. Prasetya and Sulistyorini, "Analisis antropometri dan daya tahan VO2Max peserta ekstrakulikuler bolavoli putri SMAN 1 Talun Blitar," *Indones. Perform. J.*, vol. 4, no. 1, pp. 46–55, Jun. 2019, doi: 10.17977/um077v4i12020p46-55.

[21] M. Maulina, "Profil antropometri dan somatotipe pada atlet bulutangkis," *AVERROUS J. Kedokt., Kesehat. Malikussaleh*, vol. 1, no. 2, pp. 69–74, Nov. 2015, doi:10.29103/averrous.v1i2.413.

[22] E.B. Setiawan and R. Herdianto, "Penggunaan smartphone android sebagai alat analisis kebutuhan kandungan nitrogen pada tanaman padi," *J. Nas. Tek. Elekt. Teknol. Inf.*, vol. 7, no. 3, pp. 273–280, Aug. 2018, doi: 10.22146/jnteti.v7i3.435.

[23] J. Ma *et al.*, "Contrast limited adaptive histogram equalization-based fusion in YIQ and HSI color spaces for underwater image enhancement," *Int. J. Pattern Recognit., Artif. Intell.*, vol. 32, no. 7, pp. 1–27, Jul. 2018, doi: 10.1142/S0218001418540186.

[24] Nofirza dan Z. Infi, "Perancangan alat belajar dan bermain yang ergonomis di Taman Kanak-Kanak Islam Permata Selat Panjang," *J. Ilm. Tek. Ind.*, vol. 10, no. 1, hal. 48–58, Jun. 2011.

[25] Sandhopi, L. Zaman P.C.S.W, and Y. Kristian, "Identifikasi motif jepara pada ukiran dengan memanfaatkan convolutional neural network," *J. Nas. Tek. Elekt., Teknol. Inf.*, vol. 9, no. 4, pp. 403–413, Nov. 2020, doi: 10.22146/jnteti.v9i4.541.

[26] R. Islamadina *et al.*, "Pengukuran badan ikan berupa estimasi panjang, lebar, dan tinggi berdasarkan visual capture," *J. Nas. Tek. Elekt., Teknol. Inf.*, vol. 7, no. 1, pp. 57–63, Feb. 2018, doi: 10.22146/jnteti.v7i1.401.

[27] R. Rulaningtyas, A.B. Suksmono, T.L.R. Mengko, and G.A.P. Saptawati, "Segmentasi citra berwarna dengan menggunakan metode clustering berbasis patch untuk identifikasi Mycobacterium Tuberculosis," *J. Biosains Pascasarj.*, vol. 17, no. 1, pp. 19–25, Jan. 2015, doi: 10.20473/jbp.v17i1.2015.19-25.

[28] S.E. Indraani, I.D. Jumaddina, S. Ridha, and S. Sinaga, "Implementasi edge detection pada citra grayscale dengan metode operator Prewitt dan operator Sobel," 2014.

[29] R. Sehgal, R. Gupta, and N. Anand, "Automatic extraction of 3D body measurements from 2D images of a female form," *IOSR J. Polym. Text. Eng.*, vol. 5, no. 3, pp. 7–17, May./Jun. 2018, doi: 10.9790/019X-05030717.

[30] M. Aslam, F. Rajbdad, S. Khattak, and S. Azmat, "Automatic measurement of anthropometric dimensions using frontal and lateral silhouettes," *IET Comput. Vis.*, vol. 11, no. 6, pp. 434–447, Jul. 2017, doi: 10.1049/iet-cvi.2016.0406.

[31] M. Dantone, J. Gall, C. Leistner, and L.V. Gool, "Human pose estimation using body parts dependent joint regressors," *2013 IEEE Conf. Comput. Vis. Pattern Recognit.*, 2013, pp. 3041–3048, doi: 10.1109/CVPR.2013.391.

[32] L. Pishchulin, M. Andriluka, P. Gehler, and B. Schiele, "Poselet conditioned pictorial structures," *2013 IEEE Conf. Comput. Vis., Pattern Recognit.*, 2013, pp. 588–595, doi: 10.1109/CVPR.2013.82.

[33] G. Ning, Z. Zhang, and Z. He, "Knowledge-guided deep fractal neural networks for human pose estimation," *IEEE Trans. Multimed.*, vol. 20, no. 5, pp. 1246–1259, May. 2018, doi: 10.1109/TMM.2017.2762010.

[34] A. Bulat and G. Tzimiropoulos, "Human pose estimation via convolutional part heatmap regression," in *Computer Vision – ECCV 2016*, B. Leibe, J. Matas, N. Sebe, and M. Welling, Eds., Cham, Switzerland: Springer, 2016, pp. 717–732, doi: 10.1007/978-3-319-46478-7_44.

[35] A. Newell, K. Yang, and J. Deng, "Stacked hourglass networks for human pose estimation," in *Computer Vision – ECCV 2016*, B. Leibe, J. Matas, N. Sebe, and M. Welling, Eds., Cham, Switzerland: Springer, 2016, pp. 483–499, dot: 10.1007/978-3-319-46484-8_29.

[36] W. Yang *et al*., “Learning feature pyramids for human pose estimation,” *2017 IEEE Int. Conf. Comput. Vis. (ICCV)*, 2017, pp. 1290–1299, doi: 10.1109/ICCV.2017.144.

[37] X. Chu *et al*., “Multi-context attention for human pose estimation,” *2017 IEEE Conf. Comput. Vis., Pattern Recognition (CVPR)*, 2017, pp. 5669–5678, doi: 10.1109/CVPR.2017.601.

[38] Y. Chen *et al*., “Adversarial PoseNet: A structure-aware convolutional network for human pose estimation,” *2017 IEEE Int. Conf. Comput. Vis. (ICCV),* 2017, pp. 1221–1230, doi: 10.1109/ICCV.2017.137.

[39] C.J. Chou, J.T. Chien, and H.T. Chen, “Self adversarial training for human pose estimation,” *2018 Asia-Pac. Signal Inf. Process. Assoc. Annu. Summit Conf. (APSIPA ASC)*, 2018, pp. 17–30, doi: 10.23919/APSIPA.2018.8659538.

[40] A. Toshev and C. Szegedy, “DeepPose: Human pose estimation via deep neural networks,” *2014 IEEE Conf. Comput. Vis., Pattern Recognit. (CVPR)*, 2014, pp. 1653–1660, doi: 10.1109/CVPR.2014.214.

[41] J. Carreira, P. Agrawal, K. Fragkiadaki, and J. Malik, “Human pose estimation with iterative error feedback,” *2016 IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, 2016, pp. 4733–4742, doi: 10.1109/CVPR.2016.512.

[42] D.C. Luvizon, H. Tabia, and D. Picard, “Human pose regression by combining indirect part detection and contextual information,” *Comput. Graph*., vol. 85, pp. 15–22, Dec. 2019, doi: 10.1016/j.cag.2019.09.002.

[43] D. Mehta *et al.*, “VNect: Real-time 3D human pose estimation with a single RGB camera,” *ACM Trans. Graph*., vol. 36, no. 4, pp. 1–14, Aug. 2017, doi: 10.1145/3072959.3073596.

[44] J. Martinez, R. Hossain, J. Romero, and J.J. Little, “A simple yet effective baseline for 3D human pose estimation,” *2017 IEEE Int. Conf. Comput. Vis. (ICCV)*, 2017, pp. 2659–2668, doi: 10.1109/ICCV.2017.288.

[45] C. Ionescu, D. Papava, V. Olaru, and C. Sminchisescu, “Human3.6M: Large scale datasets and predictive methods for 3D human sensing in natural environments,” *IEEE Trans. Pattern Anal. Mach. Intell*., vol. 36, no. 7, pp. 1325–1339, Jul. 2014, doi: 10.1109/TPAMI.2013.248.

[46] C.H. Chen and D. Ramanan, “3D human pose estimation = 2D pose estimation + matching,” *2017 IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, 2017, pp. 5759–5767, doi: 10.1109/CVPR.2017.610.

[47] S. Liang, X. Sun, and Y. Wei, “Compositional human pose regression,” *Comput. Vis., Image Underst.*, vol. 176–177, pp. 1–8, Nov./Dec. 2018, doi: 10.1016/j.cviu.2018.10.006.