Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume 7, Number 3, July 2023
DOI : https://doi.org/10.33395/sinkron.v8i3.12571
e-ISSN : 2541-2019
p-ISSN : 2541-044X

Development Of Independent Taekwondo Training Machine Learning With 3D Pose Model Mediapipe

Billy Cahyo Santoso[1)*], **Handri Santoso**[2)], **Julio Sandjaya**[3)]
[1,2,3)]Universitas Pradita, Tangerang, Indonesia
[1)]billy.cahyo@student.pradita.ac.id [2)]handri.santoso@pradita.ac.id, [3)]julio.sanjaya@du-gi.com

Submitted : Jun 20, 2023 | **Accepted** : Jun 22, 2023 | **Published** : Jul 1, 2023

Abstract: Taekwondo is a martial art that focuses on punching and kicking movements while upholding the values of discipline, ethics, and good behavior. Discipline is built with routine training to make someone proficient in taekwondo martial arts. Training cannot be carried out flexibly because it must be accompanied by a sabouem to know the correct taekwondo movements. Machine learning can be used as a solution for taekwondo movement recognition by building a learning machine model that recognizes the correct taekwondo movement. The MediaPipe framework has the advantage of being able to recognize human posture with 33 points or landmarks. The research was carried out by conducting a literature study, where similar research was found but only based on the values of the x and y axes. So a problem arises where the majority of taekwondo movements require the z axis to know the correct taekwondo movements. The research was conducted to add z-axis values and change calculations, which were adjusted to reconstruct a training data object in the form of an image into a three-dimensional shape. From this study, it was found that machine learning using the x, y, and z axes is much better for its use, especially when detecting taekwondo movements from different viewpoints from the training data. This research can be developed by enlarging the image dataset and packaging the model into a mobile application so that it can be used for taekwondo training up to taekwondo movement assessment.

Keywords: Human Posture, Machine Learning, MediaPipe, Taekwondo, Three Dimensional

INTRODUCTION

Taekwondo is a martial art that comes from three words: *tae*, which means to destroy with the feet; *kwon*, which means attack with the hands; and *do*, which means discipline or art. Of the three words that make up the word taekwondo, taekwondo can be interpreted as a martial art from Korea that focuses on punching and kicking movements that uphold discipline, ethics, and good behavior (Puspodari & Muharram, 2020; Muharram & Puspodari, 2020). Taekwondo movement itself consists of five basic components consisting of body parts that are targeted (*keup so*), body parts that are used to attack or defend, stances (*seogi*), parrying or defending techniques (*makki*), and attack techniques (*kongkyok kisul*) (Muharram & Puspodari, 2020). Someone who wants to learn taekwondo or raise a taekwondo belt class needs to master the basic movements first with regular training (one of the principles of taekwondo, discipline) so that the position of the taekwondo movements is correct (Kim, et al., 2021). Taekwondo training often experiences problems because the training time cannot be carried out flexibly because it needs to be accompanied by a taekwondo or *saboeum* trainer. However,

*name of corresponding author

This is an Creative Commons License This work is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License.

Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume 7, Number 3, July 2023
DOI : https://doi.org/10.33395/sinkron.v8i3.12571
e-ISSN : 2541-2019
p-ISSN : 2541-044X

if it is done independently, one cannot know whether the taekwondo movement is correct or not. The objectives of the training in face-to-face meetings themselves, among others, are centered on the principles of readiness, progressiveness, and adaptation (Muharram & Puspodari, 2020). These three face-to-face training principles can be overcome by machine learning, which can assess taekwondo movements performed by a person.

Machine learning is a branch of computer science that studies algorithms and techniques to automate complex and difficult problem solutions that are difficult to program using conventional programming methods (Rebala, Ravi, & Churiwala, 2019). Machine learning can detect the human body in the form of landmark points (Tanjaya, Naufal, & Arwoko, 2023; Bala, et al., 2020). The framework used in building machine learning in this study is MediaPipe. Choosing MediaPipe as a framework for building a taekwondo training model independently is important because MediaPipe can reconstruct body postures into 33 landmarks (Suyudi, Sudadio, & Suherman, 2022; Chung, Ong, & Leow, 2022). Machine learning using MediaPipe is built by taking image data from the nine taekwondo moves as a reference for the nine correct taekwondo moves. MediaPipe is a machine learning framework that has a convolutional neural network architecture and has 33 key points or landmarks for reconstructing human poses (Kwon & Kim, 2022). If it is observed from previous research that is used to predict and return key points on the pose of the human body based on a single point of view, it is found that MediaPipe is able to predict poses on all low-latency platforms compared to other human pose algorithms such as YOLO (Dinh, et al., 2023).

Research conducted to create taekwondo training independently with 3D poses from MediaPipe led to the development of a real-time yoga training application using MediaPipe (urgonguyen, 2022). However, after trials using images of taekwondo movements, it was found that the machine learning that was built in previous studies had problems detecting taekwondo movements that require the z axis or rotation and from different angles. The research was developed by changing the image data, changing the values taken, and changing the calculations made to get the angle of a body posture so that machine learning can detect taekwondo movements with 3-dimensional reconstruction (Swain, et al., 2022).

METHOD

Research methods are used to obtain information for specific purposes and uses in a scientific way (Ramdhan, 2021). The research method used in this study was carried out in four stages. The four research stages are shown in the form of a flowchart in Fig. 1.

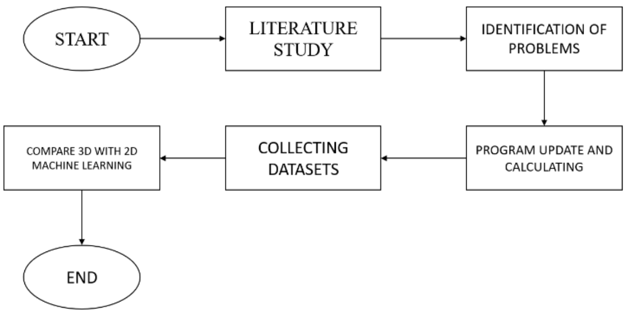

Fig. 1. Research Stages Flowchart

From the flowchart in Fig. 1, the four stages can be described as follows:

A literature study is carried out by collecting various references that are similar to the research that will be carried out. The reference started by collecting several research journals that discussed pose detection with MediaPipe and looking for similar research on social media based on YouTube video sharing. From the literature study conducted, it was found that MediaPipe is the right machine learning for detecting body poses by inferring 33 key points or landmarks with the convolutional neural network (CNN) architecture (Kwon & Kim, 2022).

*name of corresponding author

This is an Creative Commons License This work is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License.

Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume 7, Number 3, July 2023
DOI : https://doi.org/10.33395/sinkron.v8i3.12571
e-ISSN : 2541-2019
p-ISSN : 2541-044X

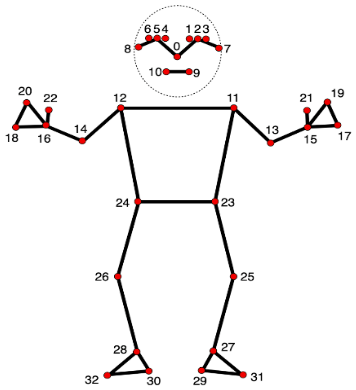

Fig. 2. 33 Key Points or Landmarks of the MediaPipe Model Pose

Then, a trial was conducted to identify the problem by changing the image data, which was originally in the form of yoga image data, to taekwondo. From these trials, it was found that machine learning has problems detecting movements from different angles and movements that require a rotation or z axis.

Programs that have been made in previous research are studied and understood to find out which parts of the program must be developed and modified to solve problems. In previous studies, the calculation of the 2-dimensional vector extension was used as in equation (1) below.

$$radians = \arctan 2(y_3 - y_2, x_3 - x_2) - \arctan 2\,(y_1 - y_2, x_1 - x_2) \quad (1)$$

The program is developed by taking the z axis value first. Then changes are made to the calculation using the 3-dimensional vector extension formula to determine each body angle. 3D calculations can be seen in equation (2) (Swain, et al., 2022; Bahukhandi & Gupta, 2021; Sidana, 2022).

$$\cos\alpha = \frac{\vec{a}\,\vec{b}}{|a|\,.|b|} \quad (2)$$

where:

$$a = x_2 - x_1, y_2 - y_1, z_2 - z_1 \quad (3)$$

$$b = x_2 - x_3, y_2 - y_3, z_2 - z_3 \quad (4)$$

After the program was modified by taking the z value and changing the angle calculation, taekwondo movement image data began to be collected as data used as movement references and data to test machine learning. The taekwondo movement data collected consists of nine basic taekwondo movements, namely *joochoom seogi*, *apkoobi*, *kyorugi joonbi*, *arae jireugi*, *momtong jireugi*, *eolgoul jireugi*, *arae makki*, *arae hecho makki*, and *eolgoul makki* (Sarinastiti, Setyowati, & Basuki, 2022). Then proceed with collecting the nine basic movements from a different point of view. Programs that use only x and y values and programs that use x, y, and z values will be run to get an output in the form of a score regarding the accuracy of the machine learning model built using taekwondo movement reference data and tester data. The test data is divided into two different viewpoints: the same viewpoint as the training data image, and a different viewpoint from the training data image. An example of test data can be seen in Fig. 3.

*name of corresponding author

This is an Creative Commons License This work is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License.

Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume 7, Number 3, July 2023
DOI : https://doi.org/10.33395/sinkron.v8i3.12571
e-ISSN : 2541-2019
p-ISSN : 2541-044X

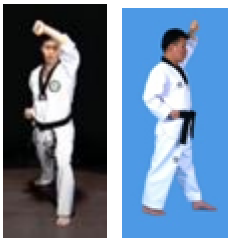

Fig. 3. Testing Data from the Same Viewpoint as the Training Data (Left) and a Different Viewpoint from the Training Data (Right)

The two programs will be compared to draw conclusions about which machine learning model is better for taekwondo movement training.

RESULT

Image Data

Comparisons using image data are broken down into nine taekwondo movements with the same perspective as the training data and with a different viewpoint from the training data. The results of the accuracy score of the program that has been run show the following results:

Seogi joochoom movement, with the same viewpoint of shooting, gets a score of 100 for the machine learning model that uses x and y values. For the machine learning model that uses x, y, and z values, it gets a score of 100. With a different image point of view, it gets a score of 51 for the machine learning model that uses x and y values. For learning models that use x, y, and z values, a score of 53 is obtained.

Apkoobi movement, with the same viewpoint of shooting, got a score of 88 for the machine learning model that uses x and y values. For the machine learning model that uses x, y, and z values, it gets a score of 84. With a different image point of view, it gets a score of 67 for the machine learning model that uses x and y values. For machine learning models that use x, y, and z values, a score of 83 is obtained.

Kyorugi joonbi movement, with the same viewpoint of shooting, got a score of 75 for the machine learning model that uses x and y values. For the machine learning model that uses x, y, and z values, it gets a score of 78. With a different image point of view, it gets a score of 63 for the machine learning model that uses x and y values. For machine learning models that use x, y, and z values, a score of 47 is obtained.

Arae jireugi movement, with the same viewpoint of shooting, gets a score of 82 for the machine learning model that uses x and y values. For the machine learning model that uses x, y, and z values, it gets a score of 71. With a different image point of view, it gets a score of 42 for the machine learning model that uses x and y values. For machine learning models that use x, y, and z values, a score of 55 is obtained.

Momtong jireugi movement, with the same viewing angle, gets a score of 85 for the machine learning model that uses x and y values. For the machine learning model that uses x, y, and z values, it gets a score of 87. With a different image point of view, it gets a score of 75 for the machine learning model that uses x and y values. For machine learning models that use x, y, and z values, a score of 89 is obtained.

Eolgoul jireugi movement, with the same viewpoint of shooting, gets a score of 75 for the machine learning model that uses x and y values. For the machine learning model that uses x, y, and z values, it gets a score of 87. With a different image point of view, it gets a score of 41 for the machine learning model that uses x and y values. For machine learning models that use x, y, and z values, a score of 79 is obtained.

Arae makki movement, with the same viewpoint of shooting, gets a score of 75 for the machine learning model that uses x and y values. For the machine learning model that uses x, y, and z values, it gets a score of 75. With a different image point of view, it gets a score of 81 for the machine learning model that uses x and y values. For machine learning models that use x, y, and z values, a score of 75 is obtained.

*name of corresponding author

This is an Creative Commons License This work is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License.

Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume 7, Number 3, July 2023 e-ISSN : 2541-2019
DOI : https://doi.org/10.33395/sinkron.v8i3.12571 p-ISSN : 2541-044X

Arae hecho makki movement, with the same viewpoint of shooting, gets a score of 95 for the machine learning model that uses x and y values. For the machine learning model that uses x, y, and z values, it gets a score of 80.

Arae hecho makki movement does not require an axis of rotation, such as the distance between the arms that are extended forward, one leg that is a different distance, and so on.

Eolgoul makki movement, with the same viewpoint of shooting, got a score of 81 for the machine learning model that uses x and y values. For the machine learning model that uses x, y, and z values, it gets a score of 91. With a different image point of view, it gets a score of 82 for the machine learning model that uses x and y values. For machine learning models that use x, y, and z values, a score of 76 is obtained.

A comparison of the results of machine learning scores using x and y values with those using x, y, and z values on the test data with the same perspective as the training data is visualized in Fig. 4.

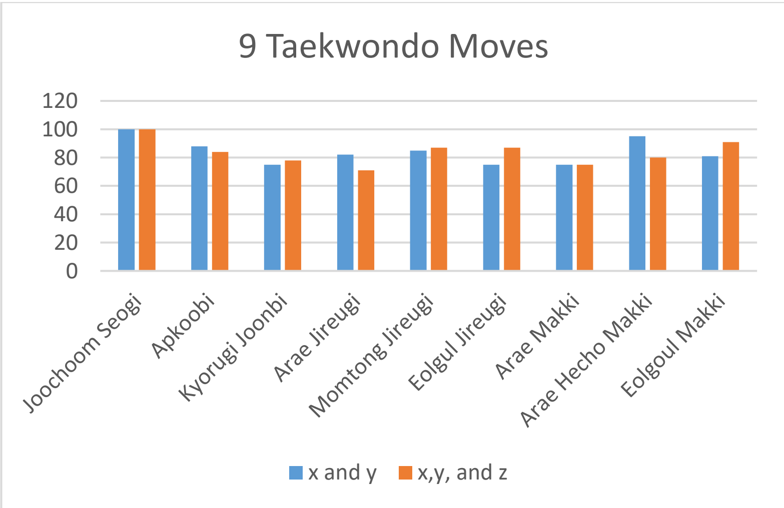

Fig. 4. Bar Graph of Comparison of Scores in the Same Viewpoint

In Figure 4, it is found that the results of machine learning using x and y values excel in three movements, namely apkoobi, arae jireugi, and arae hecho makki. Machine learning using x, y, and z values excels in four movements, namely kyorugi joonbi, momtong jireugi, eolgoul jireugi, and eolgoul makki. Whereas the two machine learning also have the same score results on two taekwondo movements, namely *joochoom seogi* and *arae makki*.

Then, for comparison with image data from a different perspective it is visualized in Fig. 5.

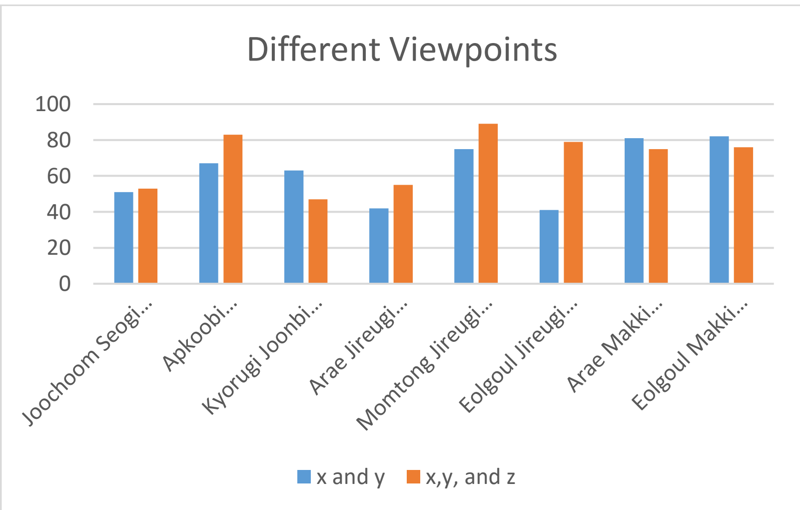

Fig. 5. Bar Graph of Comparison of Scores in Different Perspectives

*name of corresponding author

This is an Creative Commons License This work is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License.

In Fig. 5, it is found that the results of machine learning using x and y values excel in three movements, namely *kyorugi joonbi, arae makki*, and *eolgoul makki*. Machine learning using x, y, and z values excels in five movements, namely *joochoom seogi, apkoobi, arae jireugi, momtong jireugi,* and *eolgoul makki*.

From the results of these scores, it was found that machine learning using x, y, and z values produced a model that could better detect taekwondo movements compared to those using x and y values. From the same point of view, machine learning with x, y, and z values is not very different from machine learning with x and y values. Meanwhile, from a different point of view, machine learning with x, y, and z values has a striking difference compared to machine learning with x and y values.

Camera Data

The two programs were again compared by conducting a live camera test with someone demonstrating the nine taekwondo moves from five different angles. The five angles consist of the front view, facing left, facing right, half facing left, and half facing right. The form of testing using a camera can be seen in Fig. 6.

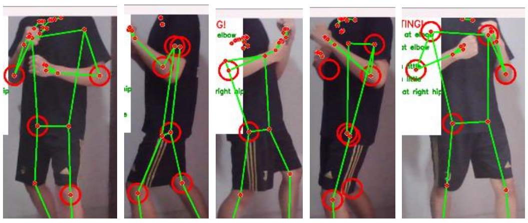

Fig. 6. The five angles consist of front view, left facing, right facing, half facing left, and half facing right

The score results for each movement are as follows:

Seogi joochoom movement scores 88, 59, 68, 75, and 66 in the machine learning model using x and y values. Machine learning uses the values x, y, and z to get scores of 80, 55, 63, 57, and 59. If averaged, machine learning with x and y gets a score of 71.2 and machine learning with x, y, and z scores 62.8.

The *apkoobi* movement scores 75, 61, 55, 62, and 58 on the machine learning model using x and y values. Machine learning uses the values x, y, and z to get scores of 67, 64, 66, 68, and 65. If averaged, machine learning with x and y gets a score of 62.2 and machine learning with x, y, and z scores 66.

The *kyorugi joonbi* movement scores 68, 36, 36, 62, and 43 on the machine learning model using x and y values. Machine learning uses the values x, y, and z to get scores of 81, 79, 81, 79, and 78. If averaged, machine learning with x and y gets a score of 49 and machine learning with x, y, and z scores 79.6.

The *arae jireugi* movement scores 75, 58, 51, 65, and 72 in the machine learning model using x and y values. Machine learning uses the values x, y, and z to get scores of 75, 55, 67, 68, and 85. If averaged, machine learning with x and y gets a score of 64.2 and machine learning with x, y, and z scores 70.

The *momtong jireugi* movement scores 71, 53, 50, 59, and 52 in the machine learning model using x and y values. Machine learning uses the values x, y, and z to get scores of 81, 75, 57, 67, and 72. A comparison of scores can be seen more clearly through the bar graph in Fig. 11. If averaged, machine learning with x and y gets a score of 57 and machine learning with x, y, and z scores 70.4.

The *eolgul jireugi* movement scores 68, 41, 51, 38, and 60 in the machine learning model using x and y values. Machine learning uses the values x, y, and z to get scores of 87, 87, 86, 87, and 80. A comparison of scores can be seen more clearly through the bar graph in Fig. 12. If averaged, machine learning with x and y gets a score of 51.6 and machine learning with x, y, and z scores 85.4.

*name of corresponding author

This is an Creative Commons License This work is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License.

SinkrOn
JURNAL & PENELITIAN TEKNIK INFORMATIKA
Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume 7, Number 3, July 2023
DOI : https://doi.org/10.33395/sinkron.v8i3.12571
e-ISSN : 2541-2019
p-ISSN : 2541-044X

The *arae makki* movement scores 68, 80, 81, 69, and 84 in the machine learning model using x and y values. Machine learning using x, y, and z scores of 65, 71, 71, 67, and 77. If averaged, machine learning with x and y gets a score of 76.4 and machine learning with x, y, and z scores 70.2.

The *arae hecho makki* movement scores 90, 81, 72, 77, and 77 in the machine learning model using x and y values. Machine learning uses x, y, and z scores to get scores of 85, 80, 68, 78, and 78. If averaged, machine learning with x and y gets scores of 79.4 and machine learning with x, y, and z scores of 77.8.

The *eolgoul makki* movement scores 62, 59, 50, 63, and 47 in the machine learning model using x and y values. Machine learning using the values x, y, and z gets scores of 85, 83, 67, 70, and 70. If averaged, machine learning with x and y gets a score of 56.2 and machine learning with x, y, and z scores 75.

DISCUSSIONS

In this study, problems were encountered when entering image data that needed the rotation or z axis. When the axis of rotation is not invoked in taekwondo movements that involve a straight-forward hand posture, the model detects that there is no distance between the hands and the shoulders. This problem is greatly assisted by the MediaPipe framework, which can reconstruct z values based on an image as can be seen from the visualization results in Fig. 7.

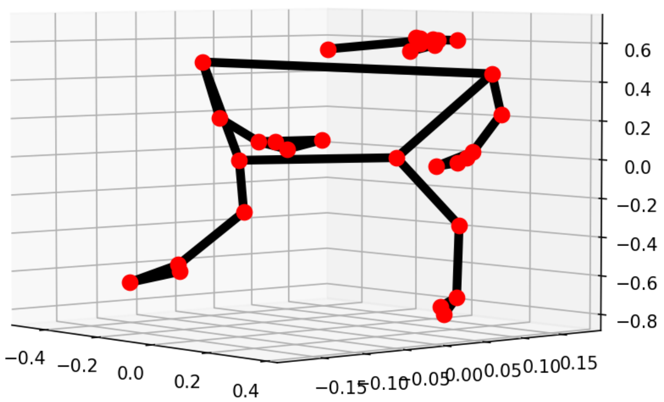

Fig. 7. Reconstructed Taekwondo Movement Visualization by MediaPipe

Research can be developed by collecting datasets consisting of more images from various viewpoints, like in other studies (Dewantara & Rhamadhaningrum, 2020). However, it is necessary to ensure computer performance to run programs with large enough datasets.

CONCLUSION

In this study, the calculation process was changed from previous research with reference to the angle formula between two space vectors. It is known that the striking difference lies in the time of testing using images with a different perspective from the training data. In the trial using the camera, it was found that the average overall score for machine learning using x and y values was 63 and for machine learning using x, y, and z values was 73. It can be concluded that machine learning using x, y, and z makes it easier to detect taekwondo moves. Research can be developed by enlarging the dataset and making the model enter into the application so that it can be developed for basic taekwondo training up to the assessment of taekwondo movements.

*name of corresponding author

This is an Creative Commons License This work is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License.

REFERENCES

Bahukhandi, U., & Gupta, S. (2021). YOGA POSE DETECTION AND CLASSIFICATION USING MACHINE LEARNING TECHNIQUES. *International Research Journal of Modernization in Engineering Technology and Science, 03*(12), 186-191.

Bala, P. C., Eisenreich, B. R., Yoo, S. B., Hayden, B. Y., Park, H. S., & Zimmermann, J. (2020). Automated markerless pose estimation in freely moving macaques with OpenMonkeyStudio. *Nature Communications*, *11*, 1-12. https://doi.org/10.1038/s41467-020-18441-5

Chung, J.-L., Ong, L.-Y., & Leow, M.-C. (2022). Comparative Analysis of Skeleton-Based Human Pose Estimation. *Futute Internet*, *14*(12), 1-19. https://doi.org/10.3390/fi14120380

Dewantara, B. S., & Rhamadhaningrum, D. T. (2020). Detecting Multi-Pose Masked Face Using Adaptive Boosting and Cascade Classifier. *International Electronics Symposium (IES)*, 436-441. https://doi.org/10.1109/IES50839.2020.9231934

Dinh, T. S., Luong, T. T., Do, H. S., Phan, V. N., Te, T. H., & Le, V. H. (2023). 3D HUMAN POSE ESTIMATION FROM SPORT VIDEOS USING MEDIAPIPE FRAMEWORK. *TẠP CHÍ KHOA HỌC ĐẠI HỌC TÂN TRÀO, 9*, 183-193. https://doi.org/10.51453/2354-1431/2023/975

Kim, Y.-J., Baek, S.-H., Park, J.-B., Choi, S.-H., Lee, J.-D., & Nam, S.-S. (2021). The Psychosocial Effects of Taekwondo Training: A Meta-Analysis. *International Journal of Environmental Research and Public Health*, *18*, 1-17. https://doi.org/10.3390/ijerph182111427

Kwon, Y., & Kim, D. (2022). Real-Time Workout Posture Correction using OpenCV and MediaPipe. *Journal of the Korean Institute of Information Technology, 20*, 199-208. http://dx.doi.org/10.14801/jkiit.2022.20.1.199

Muharram, N. A., & Puspodari. (2020). PENGEMBANGAN BUKU TEKNIK DASAR TAEKWONDO BERBASIS MOBILE LEARNING DAN MODEL TES KETERAMPILAN TENDANGAN AP HURIGI PADA ATLET TAEKWONDO KOTA KEDIRI. *Jurnal Kejaora: Jurnal Kesehatan Jasmani dan Olah Raga*, *5*(2), 41-46. https://doi.org/10.36526/kejaora.v5i2.1006

Puspodari, & Muharram, N. A. (2020). *BUKU S.O.P TEKNIK DASAR TAEKWONDO BERBASIS MOBILE LEARNING.* Jombang: CV.NAKOMU.

Ramdhan, M. (2021). *Metode Penelitian.* Surabaya: Cipta Media Nusantara.

Rebala, G., Ravi, A., & Churiwala, S. (2019). *An Introduction to Machine Learning.* Gewerbestrasse: Springer Nature Switzerland.

Sarinastiti, W., Setyowati, A. D., & Basuki, A. (2022). ANALISIS DATASET VIDEO GERAKAN DASAR TAEKWONDO DENGAN MOTION CAPTURE. *Jurnal Sistem Informasi Dan Bisnis Cerdas*, *15*, 45-53.

Sidana, K. (2022). REAL TIME YOGA POSE DETECTION USING DEEP LEARNING: A REVIEW. *International Journal of Engineering Applied Sciences and Technology*, *7*(7), 61-65.

Suyudi, I., Sudadio, & Suherman. (2022). Pengenalan Bahasa Isyarat Indonesia menggunakan Mediapipe dengan Model Random Forest dan Multinomial Logistic Regression (Introduction to Indonesian Sign Language Using Mediapipe With Random Forest Models and Multinomial Logistic Regression). *Jurnal Ilmu Siber dan Teknologi Digital (JISTED)*, *1*, 65-80. https://doi.org/10.35912/jisted.v1i1.1899

Swain, D., Satapathy, S., Acharya, B., Shukla, M., Gerogiannis, V. C., Kanavos, A., & Giakovis, D. (2022). Deep Learning Models for Yoga Pose Monitoring. *Algorithms*, *15*(11), 1-17. https://doi.org/10.3390/a15110403

Tanjaya, K. A., Naufal, M. F., & Arwoko, H. (2023). Pilates Pose Classification Using MediaPipe and Convolutional Neural Networks with Transfer Learning. *Jurnal Ilmiah Teknik Elektro Komputer dan Informatika (JITEKI)*, *9*, 212-222. http://journal.uad.ac.id/index.php/JITEKI/article/view/25975

urgonguyen. (2022, Januari 3). *GitHub - urgonguyen/yogatrainer*. Retrieved Februari 16, 2023, from https://github.com/urgonguyen/yogatrainer

*name of corresponding author

This is an Creative Commons License This work is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License.