2024 International Electrical Engineering Congress (iEECON 2024)
March 6-8, 2024, Pattaya Chonburi, THAILAND

Classification of Snatch Weightlifting Phases

Ming Qi
College of Engineering
Rangsit University
Pathum Thani, Thailand

Rong Phoophuangpairoj
Department of Computer Engineering
College of Engineering
Rangsit University
Pathum Thani, Thailand

2024 12th International Electrical Engineering Congress (iEECON) | 979-8-3503-8359-1/24/$31.00 ©2024 IEEE | DOI: 10.1109/IEECON60677.2024.10537843

***Abstract*—Weightlifting is a sport in which competitors compete by lifting a barbell loaded with weight plates from the floor with the goal of successfully lifting the heaviest weights. The snatch is the first event completed in weightlifting competitions. This work studied the classification of snatch weightlifting phases using images, posture landmarks and object detection. We used the Open Source Computer Vision Library (OpenCV) to extract the images from weightlifting videos. CNN, ResNet50 and VGG16 were applied to classify the phases of snatch weightlifting from the images. In addition, MediaPipe and You Only Look Once (YOLO) were studied to extract posture landmarks of athletes' bodies and barbell features from weightlifting images. The accuracy of 91.96% was achieved using the Support Vector Machine (SVM) with the posture landmarks and barbell features. The VGG16 classifier achieved the highest accuracy of 92.26%. The accuracies from the first to the sixth phase were 98.50%, 98.05%, 99.25%, 99.32%, 94.59% and 94.82%, respectively.**

Keywords—machine learning, MediaPipe, object detection, posture landmarks, snatch weightlifting

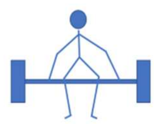
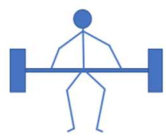
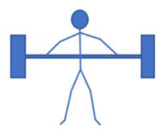
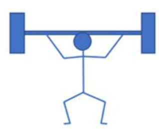
I. INTRODUCTION

Weightlifting is considered as one of the greatest tests of strength and power [1]. Weightlifting is a global sport that is part of the Olympics. People from all over the world follow the weightlifting tournament and cheer on their athletes in the fight for gold. Weightlifting athletes attempt to successfully lift the heaviest weights by lifting a barbell with weight plates up from the floor to overhead. The snatch is one of the two main lifts used in weightlifting. The athlete lifts the barbell up from the floor in a single movement [2].

Skeleton behavioral recognition is based on the principles of computer vision and deep learning which aims to detect and analyze human behavior by analyzing and interpreting the positions and movements of human skeletal joints. Object recognition is a computer vision technology that automatically detects, identifies, and locates specific objects, items, or patterns in visual data from images or videos. MediaPipe technology does not require special sensors and gloves, but can provide information about the points, characteristics and position of the legs, body, face and hands using a simple camera. MediaPipe is used to detect sports movements. Mediapipe and YOLO have been studied to recognize hand signals from sign language and martial arts fisting techniques. The successful applications of Mediapipe and YOLO in various computer vision tasks enhance our understanding and assessment of complex movements, such as those in weightlifting [3].

A CNN has tens or hundreds of layers, each of which learn to detect different features of an image. Filters are applied to each training image at different resolution. CNNs are composed of three types of layers. These are convolutional layers, pooling layers and fully-connected layers. When these layers are stacked, a CNN architecture is formed. The fully-connected layers then perform the same tasks found in Artificial Neural Networks (ANNs) and attempt to produce class scores from the activations that are used for determining a classification result. MobileNet efficiently trades off between latency and accuracy. To reduce the computing time, MobileNet replaced standard convolution filters with two layers: depthwise convolution and pointwise convolution to build a depthwise separable filter [4]. VGG16 is a convolution neural network model supporting 16 layers. VGG16 is a CNN architecture that is considered one of the best vision model architectures [5]. The SVM is a very powerful and versatile machine learning model [6]. ANN and SVM are efficient classifiers applied to various pattern recognition tasks in the field of computer vision, demonstrating their effectiveness in areas such as image classification, object detection, and face recognition.

The phases of the snatch weightlifting consisted of (1) the first pull, (2) the transition from the first to the second pull, (3) the second pull, (4) turnover under the barbell, (5) the catch phase, (6) rising from the squat position (and fully stand), as shown in Fig 1 [7,8]. The classification of the six snatch phases can be used as a part of determining success in weightlifting. Therefore, this work investigated methods for feature extraction and classification of images according to snatch lifting phases.

(1) the first pull	
(2) the transition from the first to the second pull	
(3) the second pull	
(4) turnover under the barbell	

(5) the catch phase	
(6) rising from the squat position (and fully stand)	

Fig. 1. Snatch weightlifting phases

II. Materials and Methods

A. Data:

The images were extracted from front-view videos of various Olympic weightlifting competitions, including the 2008 Beijing Summer Olympics, the 2016 Rio Olympic Weightlifting, the 2018 Buenos Aires Youth Olympic Games, and the Tokyo Weightlifting. These videos provided valuable visual data for analyzing and studying weightlifting performances and techniques. Data collection and preparation was divided into three parts: 1) extracting the images from snatch weightlifting videos, 2) annotating the position of barbells, and 3) categorizing the image files into six phases. In the first part, the images were extracted from 66 videos and resized to 224x224 pixels. In the second part, the barbell position was annotated to train Yolo for barbell detection. In the third part, the images and feature files were categorized according to the six different phases of the weightlifting motion.

For training machine learning models, the dataset was derived from 50 successful snatch weightlifting videos. A total of 3644 images were extracted from these videos distributed across the six phases. The number of the images of phases 1 to 6 were 500, 314, 257, 208, 806, and 1559, respectively. For testing purposes, the dataset included 16 successful snatch weightlifting videos. A total of 1331 images were extracted from these videos. The number of images from phases 1 to 6 were 157, 122, 93, 83, 279, and 597, respectively.

B. Instuments

Python version 3.8, OpenCV version: 4.5, Tensorflow version: 2.6) and Keras version: 2.7 were applied to create CNN, MobileNet, ResNet50 and VGG16 image classifiers. MediaPipe version: 0.9.1.0 was used to extract posture landmarks while YOLO version: 7.0 was used to detect the positions of a barbell.

C. Methods

1) Features

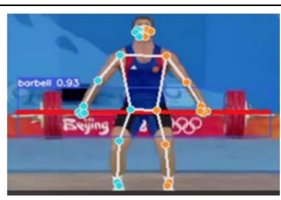
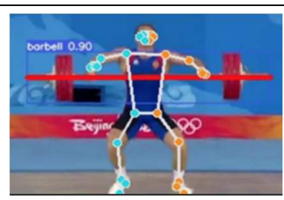
Features related to skeleton joints [3], barbell joint positions are essential for analyzing and understanding weightlifting movements. Table 1 shows the features used in this work. They consisted of 33 features of MediaPipe posture landmarks and 2 Yolo (barbell) features. Each feature consisted of x and y values representing its position.

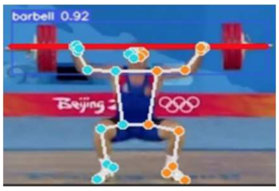
TABLE I. Features

No.	Features	No.	Features
1	Nose Positions (NOP)	19	Right Pinky Positions (RPP)
2	Left Eye Inner Positions (LEIP)	20	Left Index Positions (LIP)
3	Left Eye Positions (LEP)	21	Right Index Positions (RIP)
4	Left Eye Outer Positions (LEOP)	22	Left Thumb Positions (LTP)
5	Right Eye Inner Positions (REIP)	23	Right Thumb Positions (RTP)
6	Right Eye Positions (REP)	24	Left Hip Positions (LHP)
7	Right Eye Outer Positions (REOP)	25	Right Hip Positions (RHP)
8	Left Ear Positions (LEARP)	26	Left Knee Positions (LKP)
9	Right Ear Positions (REARP)	27	Right Knee Positions (RKP)
10	Mouth Left Positions (MLP)	28	Left Ankle Positions (LAP)
11	Mouth Right Positions (MRP)	29	Right Ankle Positions (RAP)
12	Left Shoulder Positions (LSP)	30	Left Heel Positions (LHP)
13	Right Shoulder Positions (RSP)	31	Right Heel Positions (RHP)
14	Left Elbow Positions (LEP)	32	Left Foot Index Positions (LFIP)
15	Right Elbow Positions (REP)	33	Right Foot Index Positions (RFIP)
16	Left Wrist Positions (LWP)	34	Left Barbell (LB)
17	Right Wrist Positions (RWP)	35	Right Barbell (RB)
18	Left Pinky Positions (LPP)		

2) Find barbell features [7,8]

YOLO was used to detect a barbell and the features representing the left side of a barbell (LB) and the right side of a barbell (RB) were extracted. To detect a barbell using YOLO, barbells were annotated [9-11]. Then the specialized YOLO model was trained to recognize a barbell. Finally, LB and RB could be derived from the position of the left and right sides of a barbell [10-13]. Examples of the features extracted using MediaPipe and YOLO were illustrated in Fig 2.

	
(a) the first pull	
	
(b) the transition from the first to the second pull	(c) the second pull



(d) turnover under the barbell

(e) the catch phase

(f) rising from the squat position (and fully stand)

Fig. 2. Points used as features for classifying the snatch weightlifting phases

3) Classification

CNN, MobileNetV2, ResNet50 and VGG16 were used for classifying images while ANN and SVM were used for classifying posture landmarks of athletes' bodies and barbell features. In the classification, ANN consisted of two hidden layers with 128 nodes in each layers and a output layer consisting of 6 nodes was used. The polynomial kernel function was used for classification using SVM because, after testing, it was able to provide better results than linear and radial basis functions. The number of nodes and the number of layers used in CNN, MobileNetV2, ResNet50 and VGG16 were tuned to achieve the best results. To calculate the accuracy of weightlifting phases, a multiclass confusion matrix was transformed into separeated binary confusion matrices for each class. The accuracies were then computed using Equation 1.

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \qquad (1)$$

TP, TN, FP and FN are True Positive, True Negative, False Positive and False Negative, respectively.

III. RESULTS

Table 2 shows the classification accuracy of the snatch weightlifting phases. The results were divided into two parts: 1) classifcation of images using CNN, MobileNet, ResNet50 and VGG16, 2) classifcation of posture landmarks and barbell features.

TABLE II. ACCURACY OF WEIGHTLIFTING PHASES CLASSIFICATION USING DIFFERENT METHODS

Classifier	Accuracy using images	Accuracy using posture landmarks and barbell features
ANN	-	89.86%
SVM	-	91.96%
CNN	88.13%	-
MobileNet	88.58%	-
ResNet50	78.14%	-
VGG16	92.26%	-

When using image classification, the VGG16 achieved the high accuracy of 92.26%. Meanwhile, the ANN and SVM with posture landmarks and barbell features, the accuracy of 89.86% and 91.96% was obtained, respectively. As indicated in the result, using VGG16 to classify images could achieve slightly better results than using SVM with posture landmarks and barbell features. Nonetheless, the landmarks and barbell features were independent or less dependent on the background and athletes.

1) Classifcation of images using CNN, MobileNet, ResNet50 and VGG16

The confusion matrices obtained by classifying images using CNN, MobileNet, ResNet50 and VGG16 were shown in Tables 3-6. The results showed that when using CNN, MobileNet and ResNet50, significant misclassification errors were found between weightlifting phases.

TABLE III. CONFUSION MATRIX WHEN CLASSIFYING IMAGES USING CNN

Actual / Classified	phase 1	phase 2	phase 3	phase 4	phase 5	phase 6
phase 1	99.36% (156)	0.64% (1)	0%	0%	0%	0%
phase 2	14.75% (18)	77.87% (95)	7.38% (9)	0%	0%	0%
phase 3	0%	0%	95.70% (89)	2.15% (2)	2.15% (2)	0%
phase 4	0%	0%	6.02% (5)	86.75% (72)	4.82% (4)	2.41% (2)
phase 5	2.86% (8)	1.08% (3)	1.08% (3)	0%	93.55% (261)	1.43% (4)
phase 6	1.84% (11)	0.50% (3)	1.84% (11)	0.34% (2)	11.73% (70)	83,75% (500)

TABLE IV. CONFUSION MATRIX WHEN CLASSIFYING IMAGES USING MOBILENET

Actual / Classified	phase 1	phase 2	phase 3	phase 4	phase 5	phase 6
phase 1	95.54% (150)	4.46% (7)	0%	0%	0%	0%
phase 2	4.10% (5)	90.98% (111)	4.92% (6)	0%	0%	0%
phase 3	4.30% (4)	4.30% (4)	90.32% (84)	1.08% (1)	0%	0%
phase 4	3.61% (3)	0%	1.21% (1)	89.16% (74)	2.41% (2)	3.61% (3)
phase 5	5.02% (14)	0.36% (1)	0%	0%	83.15% (232)	11.47% (32)
phase 6	4.52% (27)	0%	0%	0%	7.04% (42)	88.44% (528)

Table 5 shows that ResNet50 performed poorly in classification. Significant classification errors were found in phases 1, 4 and 5. The percentages of correct classification of the weightlifting images of phases 1, 4 and 5 were only 65.60%, 55.42% and 55.56%, respectively.

TABLE V. CONFUSION MATRIX WHEN CLASSIFYING IMAGES USING RESNET50

Actual / Classified	phase 1	phase 2	phase 3	phase 4	phase 5	phase 6
phase 1	65.60% (103)	21.02% (33)	0%	0%	0%	13.38% 21
phase 2	0%	73.77% (90)	12.30% (15)	0%	0%	13.93% (17)
phase 3	0%	5.38% (5)	79.57% (74)	2.15% (2)	0%	12.90% (12)
phase 4	0%	0%	4.82% (4)	55.42% (46)	1.21% (1)	38.55% (32)
phase 5	0%	0%	0%	1.43% (4)	55.56% (155)	43.01% (120)
phase 6	0%	0%	0%	0.67% (4)	3.52% (21)	95.81% (572)

Table 6 shows that when using VGG16, the percentages of correct classification of the first to sixth phase images were 87.26%, 98.36%, 93.55%, 92.77%, 87.10% and 94.74%, respectively. Compared to the CNN, MobileNet and ResNet50, the VGG16 could provide better classification results.

TABLE VI. CONFUSION MATRIX WHEN CLASSIFYING FROM IMAGES USING VGG16

Actual / Classified	phase 1	phase 2	phase 3	phase 4	phase 5	phase 6
phase 1	87.26% (137)	12.74% (20)	0%	0%	0%	0%
phase 2	0%	98.36% (120)	1.64% (2)	0%	0%	0%
phase 3	0%	4.30% (4)	93.55% (87)	2.15% (2)	0%	0%
phase 4	0%	0%	2.41% (2)	92.77% (77)	3.61% (3)	1.21% (1)
phase 5	0%	0%	0%	0.36% (1)	87.10% (243)	12.54% (35)
phase 6	0%	0%	0%	0%	5.53% (33)	94.47% (564)

2) Classifcation of posture landmarks and barbell features.

Tables 7-8 show the confusion matrices obtained by classifying posture landmarks and barbell features extracted using MediaPipe and YOLO. Table 7 shows that when using ANN, the percentages of correct classification of the first to sixth phase images were 96.18%, 87.70%, 94.62%, 85.54%, 81.36% and 92.46%, respectively.

TABLE VII. CONFUSION MATRIX WHEN CLASSIFYING WEIGHTLIFTING PHASES USING POSTURE LANDMARKS, BARBELL FEATURES AND ANN

Actual / Classified	phase 1	phase 2	phase 3	phase 4	phase 5	phase 6
phase 1	96.18% (151)	3.82% (6)	0%	0%	0%	0%
phase 2	5.74% (7)	87.70% (107)	6.56% (8)	0%	0%	0%
phase 3	0%	4.30% (4)	94.62% (88)	1.08% (1)	0%	0%
phase 4	0%	1.21% (1)	3.01% (3)	85.54% (71)	0%	9.64% (8)
phase 5	0%	0%	0%	0.36% (1)	81.36% (227)	18.28% (51)
phase 6	0%	0%	0%	0%	7.54% (45)	92.46% (552)

Table 8 shows when using SVM with posture landmarks and barbell features, the percentages of correct classification of the first to sixth phase images were 94.27%, 95.90%, 97.84%, 95.18%, 90.68% and 89.78%, respectively. The results showed that the SVM classifier outperformed the ANN classifier.

TABLE VIII. CONFUSION MATRIX WHEN CLASSIFYING WEIGHTLIFTING PHASES USING POSTURE LANDMARKS, BARBELL FEATURES AND SVM

Actual / Classified	phase 1	phase 2	phase 3	phase 4	phase 5	phase 6
phase 1	94.27% (148)	5.73% (9)	0%	0%	0%	0%
phase 2	2.46% (3)	95.90% (117)	1.64% (2)	0%	0%	0%
phase 3	0%	1.08% (1)	97.84% (91)	1.08% (1)	0%	0%
phase 4	0%	1.21% (1)	2.40% (2)	95.18% (79)	1.21% (1)	0%
phase 5	0%	0%	0%	1.43% (4)	90.68% (253)	7.89% (22)
phase 6	0%	0%	0%	0%	10.22% (61)	89.78% (536)

Tables 7-8 shows that misclassification still occurred in various phases. This demonstrated the necessity to find the efficient features in order to accomplish better classification of the phases of the snatch weightlifting.

IV. Conclusion

The work presented methods for classifying the phases of snap weightlifting. The machine learning methods consisting of SVM and ANN were examined with posture and barbell features extracted using MediaPipe and YOLO, while CNN, ResNet50, VGG16 and MobileNet were applied to directly classify images of snatch weight lifting phases. This research paves the way to analyzing whether athletes lift weights successfully at each phase. In the future, methods will be examined using additional weightlifting images and video data. Images should also be added to provide a better understanding of the accuracy that can be achieved using each method.

References

[1] G. Liu et al., “A biomechanical study on failed snatch based on the human and bar combination barycenter,” Applied Bionics and Biomechanics, vol. 2022, 9279638, pp. 1–14, May 2022, doi: 10.1155/2022/9279638.

[2] S. M. A. Rahmati and M. Mallakzadeh, “Prediction of weightlifter’s motor behavior to evaluate snatch weightlifting techniques based on a new method of investigation of consumed energy,” Human Movement Science, vol. 38, pp. 58–73, Dec. 2014, doi: 10.1016/j.humov.2014.08.008.

[3] J. Ma, L. Ma, W. Ruan, H. Chen, and J. Feng, “A wushu posture recognition system based on MediaPipe,” in 2022 2nd International Conference on Information Technology and Contemporary Sports (TCS), Guangzhou, China, pp. 10–13, Jun. 2022, doi: 10.1109/TCS56119.2022.9918744.

[4] A. G. Howard, M. Zhu, B. Chen, D. Kalenichenko, W. Wang, T. Weyand., M. Andreetto, and H. Adam. “MobileNets: Efficient convolutional neural networks for mobile vision applications,” 2017, Retrieved from http://arxiv.org/abs/1704.04861

[5] K. Simonyan and A. Zisserman, “Very deep convolutional networks for large-scale image recognition,” arXiv, Apr. 10, 2015. Accessed: Oct. 14, 2023. [Online]. Available: http://arxiv.org/abs/1409.1556

[6] A. Géron, “Hands-on machine learning with Scikit-Learn, Keras, and TensorFlow,” Concepts, tools, and techniques to build intelligent systems (2nd ed.) (p. 155), O'Reilly Media., pp. 155–174, 2019.

[7] S. Korkmaz and E. Harbili, “Biomechanical analysis of the snatch technique in junior elite female weightlifters,” Journal of Sports Sciences, vol. 34, no. 11, pp. 1088–1093, Jun. 2016, doi: 10.1080/02640414.2015.1088661.

[8] G. Liu et al., “Comparative 3-dimensional kinematic analysis of snatch technique between top-elite and sub-elite male weightlifters in 69-kg category,” Heliyon, vol. 4, no. 7, pp. 1–17, Jul. 2018, doi: 10.1016/j.heliyon.2018.e00658.

[9] A. Karunaratne, C. Jayasooriya, S. Deegalla, and R. Navarathna, “Objectively measure player performance on Olympic weightlifting,” 10th International Conference on Information and Automation for Sustainability (ICIAfS), Negambo, Sri Lanka, pp. 410–415, 2021, doi: 10.1109/ICIAfS52090.2021.9605963.

[10] A. I. Al-Khleifat, M. Al-Kilani, and H. A. Kilani, “Biomechanics of the clean and jerk in weightlifting national Jordanian team,” Journal of Human Sport and Exercise, vol. 14, no. 5(Proc), pp. S2429–S2434, 2019, doi: 10.14198/jhse.2019.14.Proc5.58.

[11] C. L. Ang and P. W. Kong, “Field-based biomechanical assessment of the snatch in Olympic weightlifting using wearable in-shoe sensors and videos—A preliminary report,” Sensors, vol. 23, no. 3:1171, pp. 1–15, Jan. 2023, doi: 10.3390/s23031171.

[12] A. J. Cunanan et al., “Survey of barbell trajectory and kinematics of the snatch lift from the 2015 world and 2017 Pan-American weightlifting championships,” Sports, vol. 8, no. 9:118, Aug. 2020, doi: 10.3390/sports8090118.

[13] B. Pueo, J. J. Lopez, J. M. Mossi, A. Colomer, and J. M. Jimenez-Olmedo, "Video-based system for automatic measurement of barbell velocity in back squat," Sensors, vol. 21, no. 3:925, 2021, doi:10.3390/s21030925.