SINGLE-CAMERA SYSTEM FOR HAND INJURY ASSESSMENT

DAINING MA[1], YANZI ZHOU[2], XUJIN JIANG[1], PATRICK CHAN[1]*

[1]Shien-Ming Wu School of Intelligent Engineering, South China University of Technology, China
[2]School of Future Technology, South China University of Technology, China
E-MAIL: madaining325@outlook.com, 15706720989@163.com, 202264643041@scut.edu.cn, patrickchan@scut.edu.cn
*Corresponding author: Patrick Chan

2024 International Conference on Machine Learning and Cybernetics (ICMLC) | 979-8-3315-2804-1/24/$31.00 ©2024 IEEE | DOI: 10.1109/ICMLC63072.2024.10935119

Abstract:

The hand plays a vital role in human functionality but is frequently vulnerable to injuries, particularly in occupational settings, with around 1 million workers visiting emergency rooms annually for hand-related incidents. Recovery assessment post-injury is complex and costly, often involving significant time and financial burdens on patients. While advanced surgical techniques are available, optimal recovery necessitates effective rehabilitation and functional monitoring, as neglecting follow-up can lead to complications that impair hand function. Traditional evaluation methods for hand injuries often rely on subjective assessments and lack precision, hindering rehabilitation progress. There is an increasing demand for automated hand function evaluation, which requires accurate measurement of hand movements. This paper introduces a cost-effective computer vision-based system using a single camera to assess hand recovery levels. Employing MediaPipe for joint motion analysis, the system calculates finger bending angles and evaluates motion smoothness. Additionally, the Jebsen-Taylor Hand Function Test (JTHFT) is incorporated, analyzed through the YOLOv8s object detection model to monitor task execution. This innovative approach aims to enhance the accuracy and accessibility of hand function assessments, ultimately supporting better recovery outcomes for patients with hand injuries.

Keywords:

Hand Injury Assessment; Single Camera; Skeleton Extraction; MediaPipe; Object Detection

1. Introduction

The hand, a crucial component of human functionality, is often susceptible to injury during everyday tasks. Additionally, medical conditions like stroke, arthritis, and severe burns can exacerbate this risk. Each year, approximately 1 million workers visit emergency rooms for hand injuries, according to the U.S. Bureau of Labor Statistics [1]. Post-injury recovery assessment is complex and costly [2], imposing significant burdens on patients both in terms of time and finances. Despite advanced surgical techniques, optimal recovery requires rehabilitation and functional monitoring. Timely follow-up is essential, as neglecting recovery can lead to complications like muscle atrophy and tendon adhesions, ultimately impairing hand function [3]. Active diagnostic and therapeutic interventions are vital to enhance recovery of an injured hand, reduce complications, and improve daily functioning.

According to the trial standards for upper limb functional assessment established by the Chinese Medical Association of Hand Surgery [4], evaluations of hand injuries focus primarily on joint motion angles, muscle strength, and functional performance. Examples of traditional evaluation methods are Muscle strength measurement, joint range of motion assessment and mobility and participation level quantitative scale assessment. Many medical assessment scales [5] used in the methods lack precision. Moreover, they often depend on self-reporting from patients and subjective judgments from healthcare providers. Furthermore, these assessments can be time-consuming, and their inaccuracies and inconsistencies hinder rehabilitation progress and patient well-being.[6]

There is a growing demand for automatic hand function evaluation, which necessitates accurate measurement of hand movements. Various detection methods have been developed, including wearable devices like data gloves using resistive bend sensors [7] and optical fiber sensors [8], which offer high precision but can be expensive. Color recognition gloves [9] provide a sensor-free alternative, though they may restrict natural movement and pose challenges for users with hand injuries. Depth cameras, for example, Ultraleap [10] and Kinect [11], enhance gesture recognition by capturing three-dimensional shapes. However, depth image processing requires substantial computational power, potentially limiting real-time recognition and application due to high costs. Multi-camera systems [12] demonstrate stable detection of finger bending through skeletal images, but the complicated setup and high hardware cost limit feasibility.

To reduce costs and improve flexibility, this paper presents a highly accessible method for evaluating hand recovery levels that maintains high accuracy. We introduce a computer vision-based system utilizing a single camera to assess the range of motion and functional integrity of hand joints. The system employs MediaPipe [13] to evaluate joint motion, calculate finger bending angles, and assess motion smoothness. For hand function evaluation, we utilize the Jebsen-Taylor Hand Function Test (JTHFT) [14], analyzed by the YOLOv8s object detection model [15] for task execution. The three functionalities of the system are evaluated experimentally on the dataset we collected.

The remainder of the paper is organized as follows: The second section introduces the principle and components of a single-camera based hand injury assessment system. The third part introduces experimental test and result evaluation, including system function test and data analysis. The fourth section summarizes the full essay, and discusses the future development of the project.

2. A single-camera based hand injury assessment system

This section covers the system implementation, focusing on the use of a single camera to evaluate the range of motion and functional integrity of hand joints.

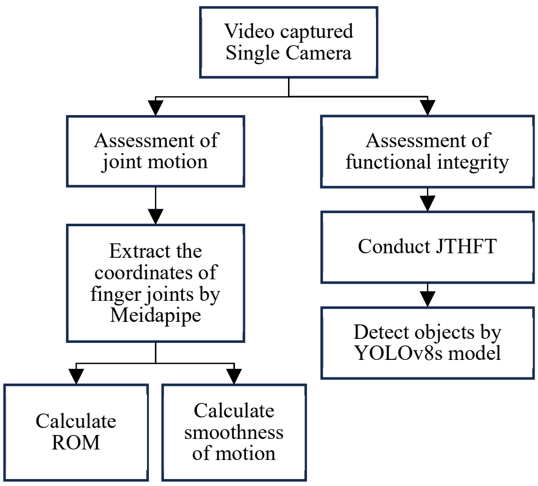

FIGURE 1. Work flow of our assessment system

2.1. Range of motion (ROM) Assessment

Range of motion (ROM) describes the arc of movement a joint can achieve, typically measured in degrees. Mediapipe [13] is used to implement a hand skeleton detection model to calculate the active ROM of each finger joint, scoring points based on a reference table.

Active ROM indicates the degrees patients can reach independently, without external assistance. Accurate ROM measurement is vital for evaluating impaired hand function and rehabilitation effectiveness. By quantitatively analyzing ROM, we can assess joint movement, muscle strength, skeletal stability, and neurological coordination, revealing the presence and severity of hand injuries. This data is crucial for diagnosing hand trauma and directly influences the development and adjustment of rehabilitation programs to prevent further dysfunction. The Total ROM in Table 1 is the sum of the bending angle values of the MCP and PIP joints of the thumb, and the Total TAM (total active movement) is the sum of the bending angle values of the MCP, PIP and DIP joints of each of the other four fingers.

TABLE 1. Active range of motion of the thumb joints [16]

Total ROM	Score
>90°	20 points
<90	10 points
90°	0 points

TABLE 2. Active range of motion of the thumb joints [16]

Total TAM	Score
260°~200°	20~16 points
190°~130°	15~11 points
130°~100°	10~6 points
<100°	5~0 points

MediaPipe offers a real-time hand tracking module capable of detecting 21 key points of the hand with a single camera, categorizing finger joints into metacarpophalangeal (MCP), proximal interphalangeal (PIP), and distal interphalangeal (DIP) joints, with the primary assessment focusing on the absolute value of the flexion angle. Deep convolutional neural networks (CNNs) are used to process input images. During the image processing, the CNN outputs a set of coordinates representing key points that indicate the positions of various joints in the human body. By connecting the detected key points, a complete skeletal diagram is formed. We extract the two-dimensional coordinates of the joints using Mediapipe and calculate the bending angles of each finger joint.

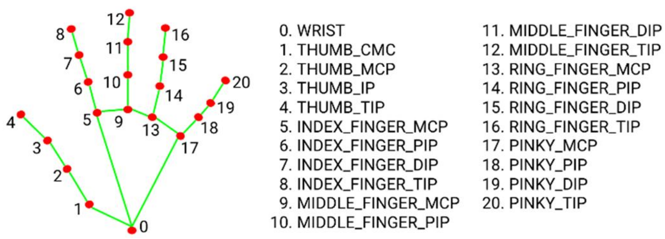

FIGURE 2. Mediapipe Hand detection [17]

Let A, B, and C be the 3 adjacent finger joints, and B is the middle joint located between A and C. The angle between ABC ($\angle B$) is defined as：

$$\overrightarrow{BA} = (x_A - x_B, y_A - y_B) \tag{1}$$

$$\overrightarrow{BC} = (x_C - x_B, y_C - y_B) \tag{2}$$

$$\angle B = cos^{-1}(cosB) = cos^{-1}\left(\frac{\overrightarrow{BA} \cdot \overrightarrow{BC}}{|\overrightarrow{BA}| \cdot |\overrightarrow{BC}|}\right) \tag{3}$$

A patient is asked to open and clench their fists to the best of his/her ability in order to capture the maximum and minimum values of the angle of a finger joint, denoted by $\boldsymbol{\theta_{max}}$ and $\boldsymbol{\theta_{min}}$ respectively.

$$\mathrm{ROM} = \theta_{\mathrm{max}} - \theta_{\mathrm{min}} \tag{4}$$

$$Total\ ROM = ROM_{MCP} + ROM_{PIP} \tag{5}$$

$$Total\ TAM = ROM_{MCP} + ROM_{PIP} + ROM_{DIP} \tag{6}$$

where ROM_{MCP}, ROM_{PIP}, and ROM_{DIP} denote the ROM of the metacarpophalangeal, proximal interphalangeal, and distal interphalangeal joint respectively. Finally, the score obtained by a patient is calculated according to Table I and II.

2.2. Smoothness of motion [18] Assessment

Smoothness of motion is a critical indicator for evaluating hand function and the degree of impairment. It is primarily used to assess finger motor control performance and reflects the continuity of motor control. A high level of motion smoothness signifies good motor control and coordination, while irregular changes in speed may suggest impaired hand function. For individuals with hand injuries, smoothness provides objective quantitative measures of movement regularity, aiding in injury severity assessment and rehabilitation program development.

To calculate the smoothness of joint movement, users are instructed to place their hands flat on a table, with the camera positioned perpendicularly at a suitable distance. The coordinate of finger joints extracted by Mediapipe is used to calculate the bending angles. The movement of the fingertips is monitored over time using the Euclidean distance formula to determine finger movement speed. Subsequently, the variation in the flexion angle of the PIP joint is calculated to derive the angular velocity of joint flexion. We developed two metrics for assessing smoothness: the average velocity divided by the maximum velocity, which applies to both linear and bending motions. The Linear Smoothness S_L is defined as:

$$v_{i,j} = \sqrt{\left(x_{i,j} - x_{i-1,j}\right)^2 + \left(y_{i,j} - y_{i-1,j}\right)^2} \tag{7}$$

where i denotes the ith frames, j denotes the jth fingers, $v_{max,j} = \max\left(v_{i,j}\right)$, $v_{mean,j} = mean\left(v_{i,j}\right)$, and $S_L = mean\left(\frac{v_{mean,j}}{v_{max,j}}\right)$.

On the other hand, the Bending Smoothness S_B is defined as:

$$\omega_{i,j} = \sqrt{\left(\theta_{i,j} - \theta_{i-1,j}\right)^2} \tag{8}$$

where θ denotes the bending angle of the PIP joint of each finger, $\omega_{max,j} = \max\left(\omega_{i,j}\right)$, $\omega_{mean,j} = mean\left(\omega_{i,j}\right)$, and $S_B = mean\left(\frac{\omega_{mean,j}}{\omega_{max,j}}\right)$.

By employing these metrics, we can analyze hand motion in a 3D space using only 2D images, leading to more accurate and reliable assessments. The formula involves subtracting the fingertip coordinates from the previous frame to estimate finger movement velocity. Static frames are excluded from the calculation, as they are considered outliers when computing average velocity. Once we determine the maximum and average velocities throughout the entire motion, a ratio closer to 1 indicates that the action is performed more coherently and smoothly.

2.3. Jebsen-Taylor Hand Function Test (JTHFT)

JTHFT [14] is a functional test providing a comprehensive assessment of a person's overall activity abilities, addressing some limitations of single physiological index evaluations. It is widely regarded as a robust tool for evaluating hand function, as its assessment items align closely with the demands of daily activities and work tasks. JTHFT consists of seven tests that evaluate various hand activities: 1) simulated eating, 2) simulated page turning, 3) picking up small objects, 4) moving a large empty can, 5) moving a large heavy can, 6) stacking checkers, and 7) writing a sentence. The handwriting test is not included in our system as it is optional in JTHFT.

YOLOv8s [15] is integrated into our system for object identification due to its high accuracy and complexity. YOLOv8s is fine-tuned according to the requirements of the quantification table. A total of 1,361 photos were collected using crawlers, encompassing all six classes of items specified in JTHFT. Manual annotations were completed using the LabelMe software [19]. The details of the dataset are shown in Table III. The image size is resized to 640x640, and the batch size is set to 8. Training is executed for 100 epochs.

TABLE 3. Detail of the dataset

Category	Number of pictures
Bean	202
Chess	245
Paper clip	222
Cap	216
Coin	240
Can	236

By recording the time taken by the patient to complete each test and comparing it with the reference values for their dominant hand and gender in Table 4, deficits in specific hand functions can be determined.

TABLE 4. Comparison of mean and standard deviation on JTHFT subtest and total score between healthy individuals and patient with hand injury. The first and second lines for each subtest indicate the average (standard deviation) of nondominant and dominant hand scores respectively [20]

TJHFT subtests	Healthy individuals	Individuals with hand injury
Stacking checkers	2.50 (0.75) 2.37 (0.77)	4.37 (1.89) 3.71 (1.47)
Simulated page turning	4.70 (1.19) 4.58 (1.30)	8.88 (5.37) 7.04 (4.0)
Lifting small objects	5.31 (1.36) 4.97 (1.12)	8.80 (4.23) 7.18 (2.49)
Simulated feeding	11.64 (4.83) 8.44 (2.07)	12.54 (5.93) 11.51 (5.02)
Lifting large light objects	3.59 (0.78) 3.59 (0.78)	5.56 (1.15) 4.62 (0.92)
Lifting large light objects	3.74 (0.8) 3.75 (0.76)	6.06 (1.28) 5.00 (0.89)

To enhance imaging clarity, we established specific requirements for the card flipping task. The cards are colored black and white on the front and back to facilitate identification of the flipped side. The camera should be positioned directly overhead, unobstructed, at approximately 0.5 meters from the tabletop, which must have a solid color background. Five cards, each measuring 13 cm x 18 cm, should be arranged in a straight line on the left side of the tabletop, with 5 cm spacing between them. The setting is shown in Figure 3. Participants will begin with their non-dominant hand, flipping the first card on the left as quickly as possible and returning it to its original position, allowing for slight displacement. Afterward, the task is repeated using the dominant hand.

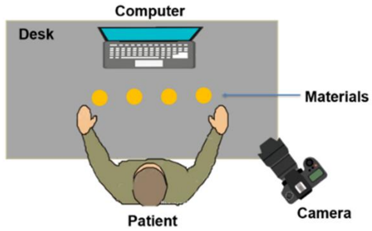

FIGURE 3. Diagram of the Testing Process of JTHFT

3. Experiments

3.1. Results of ROM Assessment

A professional finger joint protractor was initially utilized to measure the range of motion (ROM) of each joint in the hand model to establish ground truth, which was subsequently compared with the angle values calculated by our system. Figure 4 illustrates the measurement process. Four different hand models, including left and right hands of both men and women, were assessed, with each finger joint measured twice to minimize measurement error. The ground truth is defined as the average of the two measurements.

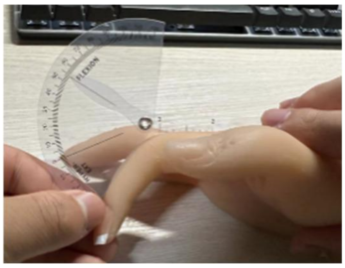

FIGURE 4. The joint angles of the hand model are measured using a professional finger joint protractor

The performance of our system is satisfactory, achieving RMSE values of 3.42 and variance of 8.45, respectively. The maximum error observed is 8.1, occurring at the DIP joint of the little finger. This may be attributed to frequent occlusions that occur in the little finger. An error of less than 10 degrees is deemed negligible in ROM assessment, as the broad range of medical scales suggests that such errors are unlikely to affect the final score.

Effectiveness testing for the scoring component was also conducted. Simulated video recordings of both healthy individuals and patients with restricted joint mobility were analyzed, revealing no differences of more than 10 points in the final calculated scores. In conclusion, although significant errors in angle measurement may occur due to occlusion or suboptimal shooting angles, it is indicated that this does not substantially impact the final score.

3.2. Results of Smoothness Assessment

20 volunteers with varying hand sizes and skin tones participated in the video test on smoothness, simulating both healthy hands and hands with restricted joint mobility. A healthy hand is expected to perform movements quickly and steadily, while an injured hand tends to move more slowly and irregularly due to reduced strength and tendon damage. Examples of samples and the outputs of our system are showed in Figure 5.

a) Fast and stable ($S_L = 0.87$, $S_B = 0.81$)

b) Slow but stable ($S_L = 0.82$, $S_B = 0.77$)

c) Unstable ($S_L = 0.44$, $S_B = 0.34$)

FIGURE 5. Examples of the video samples used for smoothness assessments

Although there is no definitive smoothness value for reference, statistical analysis and variance testing are employed to assess whether significant differences exist among various types of data. The results are shown in Table 5. Due to the small sample size, the Shapiro-Wilk test was conducted to verify whether the data conforms to a normal distribution. The p-values for each data type were greater than 0.05, indicating that the null hypothesis cannot be rejected; thus, the data satisfies the normal distribution assumption. An independent samples t-test was then performed. The test statistic for S_L was found to be 15.694, and for S_B it was 16.625, both with 19 degrees of freedom. Since both t-statistics exceed the α = 0.01 confidence level, it indicated that there is a significant difference in smoothness between the injured hand and the healthy hand.

TABLE 5. Static Analysis of Injured and Normal Samples [16]

Class	number of samples	Median	Mean	SD	P-value
Injured S_L	20	0.46	0.465	0.067	0.731
Healthy S_L	20	0.84	0.844	0.037	0.796
Injured S_B	20	0.5	0.495	0.044	0.435
Healthy S_B	20	0.795	0.804	0.039	0.665

3.3. Results of JTHFT

The ROC curve of our re-trained YOLOv8s model is illustrated in Figure 6. Generally, the detection performance on all objects is good, except the beans. This may be due to the diversity of beans and their small size, which is difficult to be correctly captured by YOLO. This could be improved by further enhancing the sample number of beans.

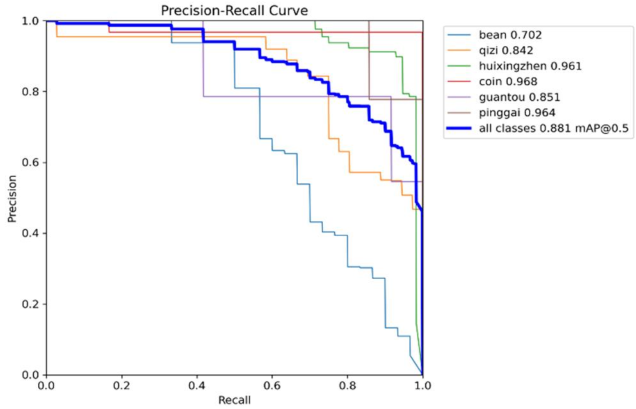

FIGURE 6. ROC curve of the trained YOLOv8s model

Figure 7 shows a screenshot of our system, which detects beans and cans. It assesses in real time whether all the beans in each frame have been placed into the cans. Once the task is identified, it displays the task complete time taken by patients.

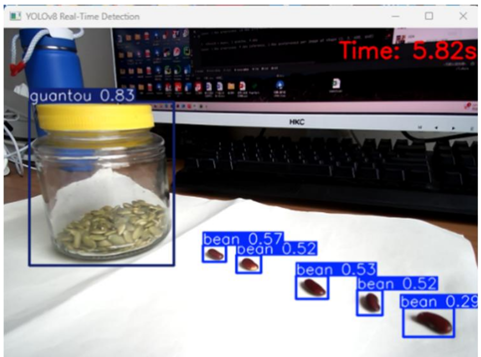

FIGURE 7. Object detection in simulated feeding test

Different types of coins, paper clips, bottle caps, etc. were found to be detectable by YOLO, demonstrating the excellent robustness of the re-trained model. 20 volunteers participated in the JTHFT and had a professional physician and a computer record their time spent completing each test respectively. Comparing the recorded time between the two, it was found that it achieves RMSE values of 0.44 and variance of 0.72, respectively.

4. Conclusions

In this paper, we developed a cost-effective computer vision system using a single camera to assess hand recovery levels. By utilizing MediaPipe for joint motion analysis, we calculated range of motion (ROM) and motion smoothness, integrating the JTHFT and YOLOv8s models for hand function assessment. Our experiments demonstrated that this intelligent detection system offers high accuracy and efficiency in hand function detection. This system significantly impacts hand rehabilitation by reducing reliance on extensive medical resources and minimizing the need for frequent follow-ups. It also lowers time and financial costs for patients with a comprehensive, low-cost assessment approach. Importantly, it simplifies equipment requirements, utilizing only a standard RGB camera, which enhances its potential for widespread adoption and development. In JTHFT, the YOLOv8s model struggles with bean identification due to limited training data and diverse bean appearances, especially in low light or complex backgrounds. This leads to false detections and affects measurement accuracy. Similarly, measuring joint motion angles suffers when fingers are blocked or shooting angles are not optimal, resulting in significant errors. These challenges can hinder accurate assessments of hand function and impact users' rehabilitation plans, highlighting the need for improvements in future work.

Future research will enhance model generalization by diversifying the dataset for varying conditions and improving architecture and training. Additionally, efforts will focus on developing a mobile application for the intelligent detection system, increasing user convenience and expanding accessibility for hand function testing on smartphones.

References

[1] https://www.facilities.udel.edu/safety/5382/

[2] Zhenxia Huang, Le Xu, Lin Li, et al., "Research Progress on Clinical Rehabilitation Evaluation Methods for Motor Function in Patients with Hand Burns," Chinese Nursing Journal, Vol. 49, No. 5, pp. 583-587, 2014.

[3] Yunhai Yao, Yaoqi Hu, Min Gu, et al., "Effects of Early Rehabilitation Intervention on Motor Function after Hand Tendon Injury," Chinese Journal of Physical Medicine and Rehabilitation, Vol. 24, No. 2, pp. 102-103, 2002.

[4] Dade Pan, Yudong Gu, De Shi, et al., "Trial Standards for Functional Assessment of the Upper Limb by Chinese Medical Association of Hand Surgery," Chinese Journal of Hand Surgery, Vol. 16, No. 3, pp. 130-135, 2000.

[5] Feng Zhou, Lifang Wang, Guimin Xie, et al., "Application of High-Frequency Ultrasound Dynamic Assessment in Rehabilitation Treatment after Hand Flexor Tendon Rupture Repair," Chongqing Medical Journal, Vol. 49, No. 18, pp. 3053-3057, 2020.

[6] Dongrui Zhang, Yanjuan Geng, Lisheng Xu, et al., "Design and Implementation of a Virtual Reality Hand Rehabilitation Training System," Integrated Technology, Vol. 4, pp. 32-38, 2013.

[7] Bandara, D.S.V., Arata, J., "Active Range of Motion Measurement System Using an Optical Sensor to Evaluate Hand Functions," in Proc. 45th Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. (EMBC), vol. 979-8, pp. 3503-2447-1, 2023.

[8] E. Fujiwara, M. F. M. d. Santos, and C. K. Suzuki, "Flexible Optical Fiber Bending Transducer for Application in Glove-Based Sensors," IEEE Sensors J., vol. 14, no. 10, pp. 3631-3636, Oct. 2014.

[9] R. Y. Wang and J. Popović, "Real-time hand-tracking with a color glove," ACM Trans. Graph., vol. 28, no. 3, Art. 63, pp. 1-8, August 2009

[10] Q. Fang, S. S. Mahmoud, X. Gu, and J. Fu, "A novel multistandard compliant hand function assessment method using an infrared imaging device," IEEE Journal of Biomedical and Health Informatics, vol. 23, no. 2, pp. 758–765, 2018.

[11] C. Xi, J. Chen, C. Zhao, Q. Pei, and L. Liu, "Real-time Hand Tracking Using Kinect," in Proc. 2nd Int. Conf. Digital Signal Process., Tokyo, Japan, pp. 37-42, Feb. 25-27, 2018.

[12] A. Utsumi, T. Miyasato, F. Kishino and R. Nakatsu, "Hand gesture recognition system using multiple cameras," Proceedings of 13th International Conference on Pattern Recognition, Vienna, Austria, pp. 667-671 vol.1, 1996.

[13] C. Lugaresi, J. Tang, H. Nash, C. McClanahan, E. Uboweja, M. Hays, F. Zhang, C. Chang, M. G. Yong, J. Lee, W. Chang, W. Hua, M. Georg, and M. Grundmann, "MediaPipe: A Framework for Building Perception Pipelines," arXiv:1906.08172 [cs.DC], 2019.

[14] M. Tipton-Burton, "Jebsen–Taylor Hand Function Test," in Encyclopedia of Clinical Neuropsychology, J. S. Kreutzer, J. DeLuca, and B. Caplan, eds., Springer, New York, NY, 2011.

[15] R. Varghese and S. M., "YOLOv8: A Novel Object Detection Algorithm with Enhanced Performance and Robustness," 2024 International Conference on Advances in Data Engineering and Intelligent Computing Systems (ADICS), Chennai, India, pp. 1-6, 2024.

[16] D. Panda, Y. Gu, D. Shi, and K. Shou, "Preliminary trial standard for the functional assessment of the upper limb

by the Chinese Medical Association Hand Surgery Society," Chinese J. Hand Surg., vol. 8, no. 6, pp. 345-349, 1992.

[17] https://chuoling.github.io/mediapipe/solutions/hands

[18] A. de los Reyes-Guzmán, I. Dimbwadyo-Terrer, F. Trincado-Alonso, F. Monasterio-Huelin, D. Torricelli, and A. Gil-Agudo, "Quantitative assessment based on kinematic measures of functional impairments during upper extremity movements: A review," Clin. Biomech., vol. 29, no. 7, pp. 719-727, Jul. 2014

[19] B. C. Russell, A. Torralba, K. P. Murphy, et al., "LabelMe: A Database and Web-Based Tool for Image Annotation," Int. J. Comput. Vis., vol. 77, no. 2, pp. 157-173, Apr. 2008.

[20] İ. Ç. Sığırtmaç, Ç. Öksüz, "Investigation of reliability, validity, and cutoff value of the Jebsen-Taylor Hand Function Test," J. Hand Ther., vol. 34, no. 3, pp. 396-403, Jul.-Sep. 2021.