The 12th IEEE International Conference on Intelligent Data Acquisition and Advanced Computing Systems: Technology and Applications
7-9 September, 2023, Dortmund, Germany

Evaluating an In-Home Exercise Program Using Vision-Based Edge AI for Elderly Healthcare

Sinan Chen[1] , Masahide Nakamura[2] , Kiyoshi Yasuda[3]
[1, 2, 3] Kobe University, 1-1 Rokkodai-cho, Nada, Kobe, 657-8501, Japan, chensinan@gold.kobe-u.ac.jp
[2] RIKEN Center for Advanced Intelligence Project, Tokyo, 103-0027, Japan, masa-n@cs.kobe-u.ac.jp

***Abstract* — In Japan, many in-home exercise programs have appeared to improve the motor function of the elderly through self-care. However, a challenging point always exists in evaluating the process of exercise programs quantitively and non-invasively for further reflecting the health status. As a preliminary progress, we present and discuss a vision-based novel method on edge that aims to evaluate an in-home exercise program in this paper. Our key idea is to integrate multiple different vision-based pre-trained models. In the proposed method, we mainly integrate and call multiple pre-trained models to recognize human facial, skeletal, and hand movements with edge computing. Also, we propose to evaluate an in-home exercise program process from the Euclidean distance, velocity, and angle of the two-dimensional coordinates of feature points in the time series. Through experiments, we estimate the physical health status from the analysis of body movements in the elderly.**

Keywords* — *exercise program; elderly healthcare; vision-based edge AI; edge computing

I. Introduction

Evaluating the exercise of elderly people at home for their healthcare is always a challenging research topic. Such as the home-based rehabilitation (PR) evaluation lacks direct supervision, and the existing complex nature of the rehabilitation model, most in-home exercises for elderly people often require being encouraged to achieve, and so on. Hawley-Hague et al. [1] developed motivational smartphone apps to support patients in performing exercises proven to aid fall reduction. However, some problems require improvement, including the complexity of operation and difficulty for individual elderly people of general households.

The invasiveness to the daily life of the monitoring device for evaluating the exercise of elderly people is also a task. Especially for elderly people requiring a long-term exercise program at home, devices with invasiveness and unique environments increase the evaluating complexity. Chan et al. [2] presented a wearable activity monitor system for patients with intermittent claudication. However, we consider that these wearable devices only apply to the short experiment to evaluate for a special purpose. For this, we must consider and study how to make a *non-invasive* evaluating process for elderly people at home, whether to retrieve or save data.

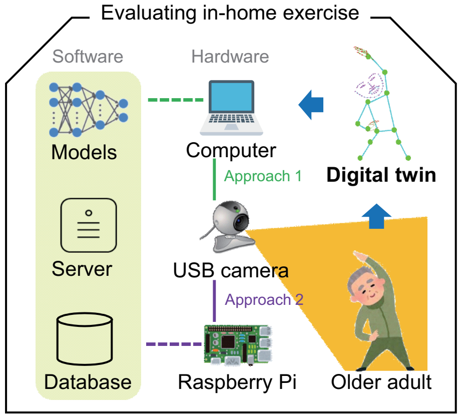

Figure 1. Overall architecture of the proposed method.

The goal of this study is to propose a non-invasive and fine-grained evaluation method for in-home exercise programs. Figure 1 shows the overall architecture of the proposed method. As an approach, we first collect live images at regular intervals using a fixed-point camera. Next, we install an edge-executable image recognition model on a general-purpose computer or Raspberry Pi. We subdivide and define the contents of in-home exercise programs and extract feature data from selected representative time-series images. In addition, we calculate from the Euclidean distance, velocity, and angle of the coordinate data (i.e., face, fingers, skeleton) as feature data and perform visualization analysis. In this way, it is possible to grasp the exercise status of the elderly at home, and estimating the physical health condition is promising.

II. Preliminaries

A. In-Home Exercise Program

In recent years, many exercise programs preventing falls in the home for the elderly were introduced into people's life in Japan. Decreased muscle strength is thought to be one of the causes of tumbles and falls. As an exercise that can finish at home, it puts less strain on the joints of the legs and lower back and trains the muscles around the knees and waist that can finish quickly. The existing exercise programs have become popular with the elderly, such as exercising while watching TV or listening to music as they can finish[1]. In this paper, we mainly focus on an exercise program while brushing teeth for elderly healthcare [2].

B. Technical Challenge

To evaluate an exercise program for elderly healthcare, we consider that the challenging point is the difficulty of retrieving *continuous* and *fine-grained* feature data with a non-invasive approach. Although some studies often rest on the subjective evaluation with questionnaires (e.g., Calculate the metabolic equivalents (METs) by asking the action contents and time), it is hard to avoid the individual differs from one household to another. Some related works also monitor elderly activity using wearable devices described in Section I. The invasiveness of daily life to elderly people from wearable devices exists, such as forgetting to wear the special device sometimes, being difficult to apply it to the physical body due to no good health status or suitable environments.

C. Vision-Based Edge AI

Vision-Based Edge Artificial Intelligence (AI) is one of the newest research topics. It uses computer vision-based pre-trained models with *transfer learning* to recognize images or video streams on edge (see Figure 2). For example, if we input an image into the target model that downloads from the cloud, the 2D coordinates or labels of the feature point will be returned locally. The typical recognition models include face recognition, pose estimation, and hand tracking, driven by the TensorFlow platform [3]. We consider that these models can improve users' security and privacy issues and communication overhead in real-time. In the previous study, although a non-invasive fine-grained home care monitoring system is proposed [4], it focuses on the pose estimation but lacks further descriptions of the process. As a preliminary study, this paper mainly integrates multiple models to evaluate an exercise program for elderly health at home.

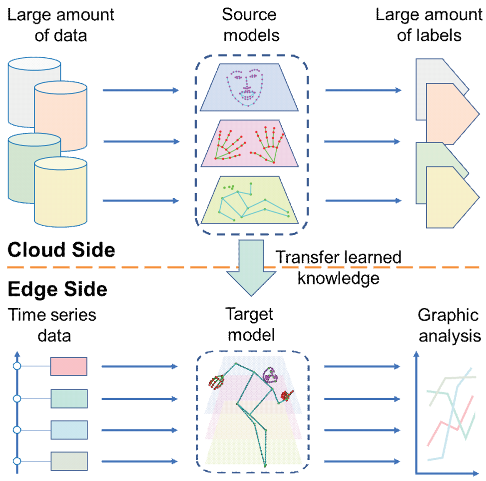

Figure 2. Overview of vision-based pre-trained models with transfer learning.

III. Proposed Method

In this section, we present a method to evaluate in-home exercise programs quantitatively for older adults. To acquire and analyze feature data (i.e., human facial, skeletal, and hand movements) changes with time series on edge, our key idea is to integrate multiple vision-based pre-trained models for a given set of exercise programs. More specifically, the proposed method consists of the following five steps.

A. Step 1: Acquiring Images on Time Series

In this step, the user first selects a specific space in the home for performing exercise programs. Next, the user connects one of the following user-selectable devices to the USB fixed-point camera: (1) a general-purpose computer browser and (2) a single-board computer (i.e., Raspberry Pi series). Then, we develop and execute a program that acquires live images in a fixed period (e.g., five seconds).

B. Step 2: Defining Exercise Programs to Recognize

In this step, the user defines the exercise programs $P = \{p_1, p_2, .., p_m\}$ to recognize for evaluating the health status. Generally, the exercise program often consists of slightly different body movements and postures. Also, the focus of strength training changes depending on the body part's orientation and the direction of movement.

C. Step 3: Selecting Continuous Representative Images

In this step, for each program $p_i \in P$, the user manually selects continuous representative n images $IMG(p_i) = \{img_{i1}, img_{i2}, ..., img_{in}\}$ that well expose p_i, from all images obtained in Step 1. Note that the user must choose a representative, consecutively acquired image for the program p_i to analyze feature changes before and after time-series images (see Step 5).

D. Step 4: Calling Vision-Based Pre-Trained Models

In this step, the user integrates a set $MODEL = \{model_1, model_2, \cdots, model_q\}$ of vision-based pre-trained models to perform the image recognition on edge, using the Web browser in a general computer, or the unique program in a Raspberry Pi device. For every $p_i \in P$, $img_{ij} \in IMG(c_i)$, and $model_k \in MODEL$, $model_k(img_{ij})$ is invoked, and a set $Feature(img_{ij}, model_k) = \{w_1, w_2, w_3, ...\}$ of output feature values is obtained. $Feature(img_{ij}, model_k)$ represents a recognition result for vision-based pre-trained models $model_k$ for an image img_{ij} belonging to a program p_i. The size of $Feature(img_{ij}, model_k)$ varies for img_{ij} and $model_k$. Since there are m programs, n images for each program, and q MODELs, this step creates totally $m \times n \times q$ sets of output feature values.

[1]http://www.city.chichibu.lg.jp/secure/8465/44gou.pdf (in Japanese)
[2]https://www.youtube.com/watch?v=u_MW0ai2oPI (in Japanese)

(a) (b) (c) (d)

Figure 3. Experiment Setup: (a∼c) Representative images of every exercise programs. (d) A Raspberry Pi and a USB camera.

E. Step 5: Analyzing Output Feature Values

In this step, the user first expands each point w_l of the feature data on the time series of n images obtained from each $model_k$ to two-dimensional coordinates $w_l = \{`x' : ..., `y' : ...\}$ for each program p_i. The user then quantitatively evaluates the exercise program from the following three computational approaches:

Euclidean distance: The user calculates the Euclidean distance d_t at time t using the coordinate data (x_t, y_t) at time t and the coordinate data (x_{t-1}, y_{t-1}) at time $t-1$ as follows:

$$d_t = \sqrt{(x_t - x_{t-1})^2 + (y_t - y_{t-1})^2} \tag{1}$$

Velocity: Since the images selected in Step 3 are representative and continuous with a constant time interval, the user calculates the velocity v_t by dividing with the time interval T_t from $t-1$ to t as follows:

$$v_t = \frac{d_t}{T_t} \tag{2}$$

Angle: To convert to angular velocity ω, the user selects coordinates of three feature points at time t, two vectors are obtained, and $cos\theta(t)$ is calculated. For example, suppose the user wants to check the degree of bending of the left leg. In that case, the user can calculate the angle with the following formula by selecting the left hip coordinate vector $\overrightarrow{d_1} = (x_1, y_1)$, the left knee coordinate vector $\overrightarrow{d_2} = (x_2, y_2)$, and the left leg coordinate vector $\overrightarrow{d_3} = (x_3, y_3)$.

$$cos\theta = \frac{\overrightarrow{d_{21}} \cdot \overrightarrow{d_{31}}}{d_{21} \cdot d_{31}} \tag{3}$$

IV. Experimental Evaluation

In this experiment, we set a single room as the target space. We aim to record and analyze the process of the specific exercise programs of an aged man (60's). Figure 3 shows representative images and experimental tools.

A. Experimental Setup

In step 1, we install a Raspberry Pi and a USB camera with a Node.js standalone program. It can collect live images automatically every five seconds. Moreover, we set the total image accumulation period to ten minutes and the image resolution to 320 × 240. In step 2, we define three kinds of in-home exercise programs while brushing the teeth (see Figure 3 (a∼c)): (p_1) Raising legs forward. (p_2) Raising legs left and right. (p_3) Raising legs behind. In step 3, we select 50 continuous representative images for each program p. Considering that raising the left or right legs differs, the number of images for raising the same leg also split into half of 50. In step 4, we integrate the Node.js standalone program of step 1 with the pre-trained model *PoseNet* [5] on edge. Hence, it also can output the 17 feature points of the human body from every image. In step 5, we first extract the output data (i.e., x and y coordinates) of the six specific feature points of the lower part of the body. Then, we calculate and visualize the Euclidean distance of the left knee, right knee, left ankle, and right ankle in the time series. Finally, we also calculate and visualize changes in the angle of raising different legs for each exercise program p.

B. Results

Figure 4 shows the representative results regarding Euclidean distance and angle for the three exercise programs. In the graph of results for Euclidean distance, on the one hand, the horizontal axis represents 250 seconds combined with 50 results at 5-second intervals. The vertical axis means the distance difference in the time-series images' pixel units. On the other hand, in the graph of the angular results, the horizontal axis is the same as above, and the vertical axis represents the angle in radians. These results show that in Program 1, the exercise movement of the left leg is more irregular than that of the right leg. In program 2, the lift distance of the left leg is lower, but the angle is similar to that of the right leg. Furthermore, in Program 3, the right foot had better dexterity than the left foot, and the range and angle of movement were more excellent.

V. Conclusion

To scientifically understand the motor function of older adults at home, this paper mainly focused on quantitative and non-invasive evaluation methods for in-home exercise programs. The main outcomes are as follows: (1) Proposing a method that integrates multiple vision-based pre-trained models and extracting features of human faces,

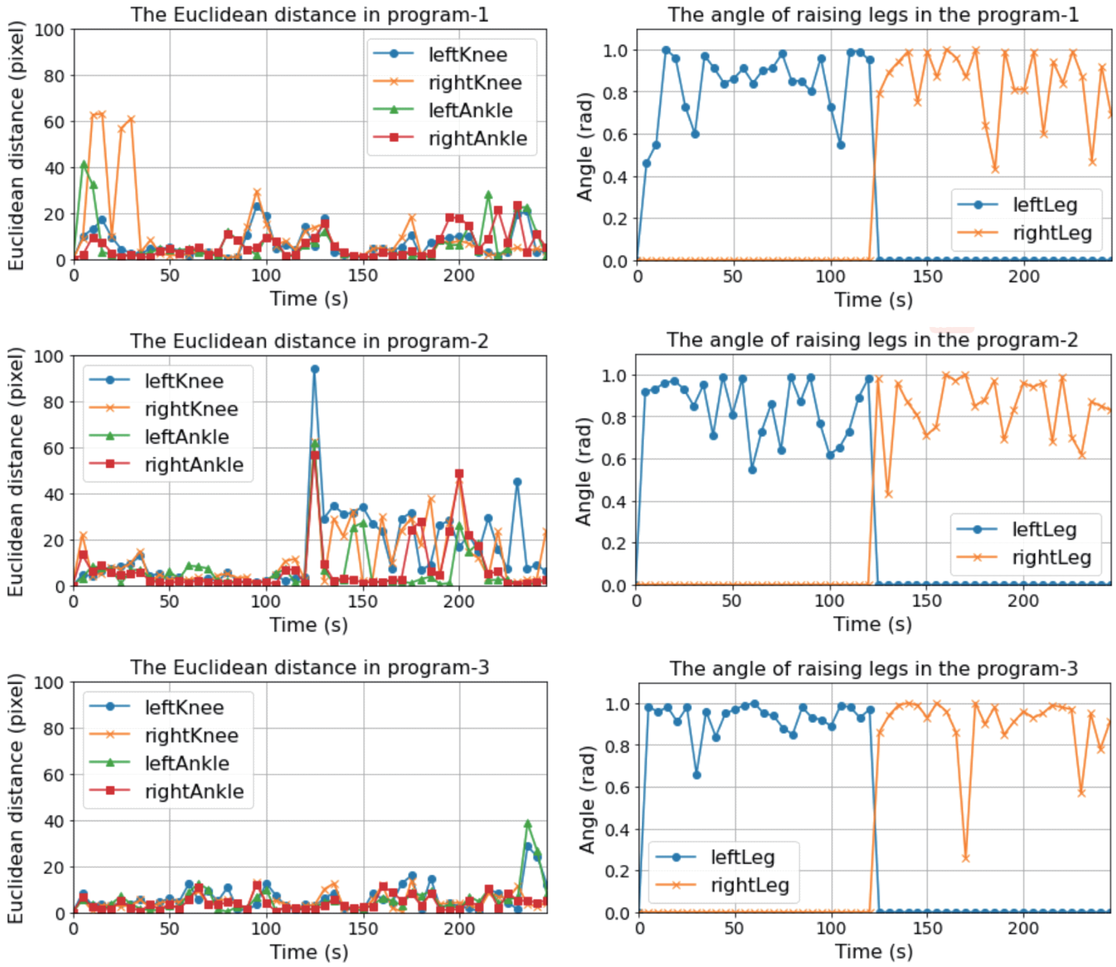

Figure 4. Representative results regarding Euclidean distance and angle for the three exercise programs.

skeletons, and fingers simultaneously on edge. (2) Experimenting with a method to evaluate the process of in-home exercise by calculating the Euclidean distance, velocity, and angle for the extracted two-dimensional coordinates on the time series. Though the experiment, we found that the movement frequency and flexibility of the subjects' right leg were better than that of the left leg. Meanwhile, by visualizing the irregular movement process of the left leg, it was presumed that there was a lack of motor function. In future work, we will develop a modeled home exercise evaluation system and further clarify its application's limitations and technical challenges (e.g., refer to studies [6] [7]).

Acknowledgment

This research was partially supported by JSPS KAKENHI Grant Numbers JP19H01138, JP20H05706, JP20H04014, JP20K11059, JP22H03699, JP19K02973, and Young Scientists (No.23K17006).

References

[1] H. Hawley-Hague, C. Tacconi, S. Mellone, E. Martinez, C. Ford, L. Chiari, J. Helbostad, C. Todd *et al.*, "Smartphone apps to support falls rehabilitation exercise: App development and usability and acceptability study," *JMIR mHealth and uHealth*, vol. 8, no. 9, p. e15460, 2020.

[2] C. Chan, V. Sounderajah, P. Normahani, A. Acharya, S. R. Markar, A. Darzi, C. Bicknell, and C. Riga, "Wearable activity monitors in home based exercise therapy for patients with intermittent claudication: a systematic review," *European Journal of Vascular and Endovascular Surgery*, vol. 61, no. 4, pp. 676–687, 2021.

[3] C. Carl, F. Schönfeld, I. Profft, A. Klamm, and D. Landgraf, "Automated detection of european wild mammal species in camera trap images with an existing and pre-trained computer vision model," *European journal of wildlife research*, vol. 66, no. 4, pp. 1–7, 2020.

[4] S. Chen, S. Saiki, and M. Nakamura, "Nonintrusive fine-grained home care monitoring: Characterizing quality of in-home postural changes using bone-based human sensing," *Sensors*, vol. 20, no. 20, p. 5894, 2020.

[5] B. Jo and S. Kim, "Comparative analysis of openpose, posenet, and movenet models for pose estimation in mobile devices," *Traitement du Signal*, vol. 39, no. 1, pp. 119–124, 2022.

[6] A. Saranya and R. Subhashini, "A systematic review of explainable artificial intelligence models and applications: Recent developments and future trends," *Decision analytics journal*, p. 100230, 2023.

[7] R. I. Mukhamediev, Y. Popova, Y. Kuchin, E. Zaitseva, A. Kalimoldayev, A. Symagulov, V. Levashenko, F. Abdoldina, V. Gopejenko, K. Yakunin *et al.*, "Review of artificial intelligence and machine learning technologies: Classification, restrictions, opportunities and challenges," *Mathematics*, vol. 10, no. 15, p. 2552, 2022.