


A Novel Image Recognition-Based Assessment System for Elderly Independent Living Ability and Its Applications



Yi Zhou 

Business College, Hunan International Economics University, Changsha 410000, China

Corresponding Author Email: sxy@hieu.edu.cn

<https://doi.org/10.18280/ts.400328>

ABSTRACT

Received: 28 January 2023

Accepted: 6 May 2023

Keywords:

image processing, the elderly, independent living ability assessment, behavior recognition

In an aging population, the gradual decline in cognitive, comprehensive, and mobility abilities poses a significant challenge for the independent living of the elderly. Accurate assessment of such abilities becomes crucial, not only for facilitating care-giving but also for policy-making related to elder care. To address this need, this study presents a novel image recognition-based assessment system designed to accurately evaluate the independent living ability of the elderly. The system construction and subsequent real-world applications constitute the initial focus of this study. Significant efforts have been made to detect key points of the skeletal structure in elderly individuals. The skeleton extraction process has been methodically divided into five distinct steps: detection of key points, matching and connecting joint points, generating coordinates of joint points and their connection maps, measuring correlations between key point pairs, and calculating optimal matching results using a bipartite graph approach. In the latter part of the study, an advanced model integrating an attention mechanism with a Graph Convolutional Neural Network (GCNN) has been developed and implemented for elder behavior recognition. The effectiveness and validity of this approach have been assessed through rigorous experimental validation. This study's findings can potentially revolutionize the quality of life assessment for the elderly and provide valuable insights for relevant policy-making. Further research in this direction is deemed necessary for enhancing the assessment system and expanding its applications.

1. INTRODUCTION

The escalating global trend of population ageing has ushered in the need for a comprehensive examination of senior citizens' capabilities for independent living [1-3]. Changing family structures and the increasing scarcity of familial caregivers for the elderly underpin the urgency of this exploration [4-9]. Aging invariably accompanies a decline in cognitive, comprehension, and mobility skills, subsequently impacting the independent living ability of the elderly [10, 11].

A crucial element in gauging the life quality of elderly citizens is the assessment of their ability to live independently. This evaluation holds immense importance not only in the sphere of elderly care but also in the drafting of related policies [12-14]. In the light of remarkable advancements in computer vision and artificial intelligence over the past decades, there has been a growing interest in image recognition-based independent living ability assessment systems for the elderly in the associated research field [15-20].

The successful operation of such systems heavily relies on the precise recognition of elderly behavior. Analyzing the daily life activities of the elderly and subsequently evaluating their ability to live independently can yield vital evidence for medical treatment, healthcare, and policy formulation [21, 22]. Yet, the application of extant image-based behavior recognition methods for assessing the independent living ability of the elderly is impeded by several limitations [23]. Conventional manual feature extraction methods, for instance, display suboptimal accuracy and robustness in dealing with

complex elderly behaviors. Additionally, notwithstanding the significant strides made by deep learning in image recognition, its application in elderly behavior recognition is beset with problems such as data imbalance and insufficient model generalization.

In an endeavor to mitigate these challenges, the current research delineates a novel method for an independent living ability assessment system based on existing image-based behavior recognition techniques. The system capitalizes on computer vision and artificial intelligence, providing a potent solution for the task of elderly behavior recognition. The research further endeavors to enhance the feature extraction of elderly behavior by refining existing image recognition techniques, aiming to bolster the assessment system's accuracy. Concurrently, it attempts to augment the model's generalization capability, thereby improving the system's practicality.

The following sections of this research elaborate on the application of the independent living ability assessment system for the elderly in real-world scenarios. Chapter two introduces the system and its applications in a variety of situations. Chapter three discerns critical points in elderly bone structure and dissects the skeleton extraction process into five distinct steps: key point detection, joint point matching and connection, generation of joint point coordinates and joint pair connection maps, correlation measurement of key point pairs, and the computation of optimal matching result based on bipartite graph. Chapter four delineates the construction and application of an attention mechanism-integrated GCNN

model for elderly behavior recognition. The research ultimately validates the proposed method experimentally, establishing its efficacy.

2. THE PROPOSED SYSTEM

In recent years, the family structures have become smaller, leading to a reduced number of family members available to care for the elderly. Consequently, the need for elderly care has become increasingly significant. Furthermore, as the elderly population ages, their cognitive, comprehension, and mobility abilities decline, reducing their capacity to live independently. In response to these challenges, the concept of smart elderly care has emerged, which integrates intelligent control in the field of elderly care using technologies such as the Internet, computers, 3D vision, and data transmission systems. This approach aims to provide efficient, timely care services for the elderly in communities and nursing homes by connecting communities, hospitals, and relatives through a home care service system.

A critical component of the smart elderly care implementation is the independent living ability assessment system. This system evaluates the physiological needs of the elderly by assessing their daily living abilities, assisting them in selecting suitable care modes and products. Image recognition-based independent living ability assessment systems offer a more objective and accurate assessment method, providing robust support for smart elderly care.

Mainstream elderly care modes include elderly care institutions, community care, and home care, each with distinct functional requirements. In the community care mode, for example, the technique for assessing independent living ability can be applied for service management and elderly monitoring. In service management, personalized services such as living aid and rehabilitation training can be provided for the elderly through independent living ability assessment. In elderly monitoring, the assessment system can track elderly activities in real-time, detecting abnormalities such as falls or disorientation, and providing timely assistance. In the home care mode, this technique also has significant application value, including daily behavior recording and analysis of the elderly in dynamic document management, and living status monitoring in intelligent call and video care.

Establishing such a system is highly meaningful, as it can improve assessment accuracy, reduce information transmission costs, and provide personalized care services for the elderly to enjoy a higher quality of life in their later years.

The construction and application of the assessment system in real-world scenarios rely on accurate recognition of the elderly's behavior. Independent living ability refers to the capacity to perform daily tasks, such as dressing, eating, bathing, and toileting without external assistance. Precise behavior recognition can assist in determining the appropriate level of care services and products for the elderly. Accurate behavior recognition enables the assessment system to provide nurses with information about elderly living ability and needs, allowing them to offer personalized services such as life care, rehabilitation training, and psychological support, ultimately improving the elderly's quality of life. In addition, accurate behavior recognition can assist in detecting abnormalities in the elderly, such as falls, disorientation, or sudden illness. Upon detecting an abnormality, the system can immediately raise an alarm to provide timely assistance and reduce potential safety risks. Furthermore, accurate behavior

recognition can aid in assessing the mental health status of the elderly, employing facial expression and behavior pattern analysis to identify possible mental health issues such as depression or anxiety. This information can be used to develop appropriate psychological intervention measures, enhancing the mental well-being of the elderly.

To achieve more accurate recognition of elderly behavior, image recognition technology must be further optimized. This optimization includes refining image processing algorithms, enhancing the model's recognition accuracy, and developing recognition schemes for specific scenarios. Only with accurate behavior recognition can the assessment system realize its full potential and provide high-quality care services for the elderly.

3. SKELETON EXTRACTION METHOD

In the current study, skeleton extraction was leveraged to discern behaviors performed by elderly individuals, such as walking, sitting, or standing, through the detection and analysis of critical skeletal key points. This information potentially aids in assessing the senior individuals' living abilities and requirements, thereby minimizing possible safety risks. Moreover, abnormalities such as falls or walking difficulties can be promptly identified, allowing for immediate alerts and the provision of timely assistance.

Further application of this real-time skeleton extraction and analysis was found within the rehabilitation process. Medical professionals, such as doctors or therapists, could utilize this method to comprehend the recovery progress of the elderly, thereby enabling potential adjustments to the training regimen based on the skeleton analysis during rehabilitation movements.



(1)



(2)

Figure 1. Examples of key point detection of the skeleton of the elderly

In the analysis presented herein, key skeletal points on the elderly were detected prior to the execution of elder behavior recognition via image processing. Figure 1 illustrates some examples of this detection. The process of skeleton extraction was conducted in five stages: 1) key point detection, 2) joint point matching and connection, 3) joint point coordinates and joint pair connection maps generation, 4) correlation measurement of key point pairs, and 5) the computation of the optimal matching result grounded on a bipartite graph.

To commence, human body key points in the input image were identified, including critical areas such as the head, shoulders, wrists, and knees, which served as foundational data for subsequent steps. Supposing A to represent the confidence map of each joint point in the elderly behavior image, O to represent the position of the key point in the image, u to signify the joint point, k to denote the k -th person in the image, $Z_{u,k}$ to represent the u -th joint point of the k -th person, and δ_2 to represent the variance, then the 2D confidence map of body part positions of the elderly can be formally expressed as follows:

$$A_{u,k}^*(O) = \exp\left(-\frac{\|O - Z_{u,k}\|_2^2}{\delta^2}\right) \quad (1)$$

In the process of skeleton extraction, the adherence to the normal distribution is observed, specifically in the case of image processing involving the presence of multiple elderlies. Here, numerous peaks of the normal distribution are discerned, necessitating a comparison and subsequent selection of the maximum value [Eq. (2)].

$$A_u^*(O) = \text{MAX}_k A_{u,k}^*(O) \quad (2)$$

Subsequent to the identification of key points, the next step involves their pairing and connection to construct a skeletal representation of a human figure. This is facilitated through predetermined joint pairs such as head-shoulder, and shoulder-wrist. During this process, the optimal connection mode can be ascertained, based on spatial distance and orientation information pertaining to the key points. If the b -th joint pair is represented by b , a two-dimensional vector field describing the position and orientation of joint points is constructed [Eq. (3)].

$$M_{b,k}^*(O) = \begin{cases} c, & \text{If } O \text{ is on joint pair } b \text{ of the } k\text{-th person} \\ 0, & \text{Otherwise} \end{cases} \quad (3)$$

The key point O on the b -th joint pair of the k -th person is represented by c . If c is valid, the vector is denoted by c ; if not, it is denoted by 0. The position coordinates of u_1 and u_2 are represented by $Z_{u_1,k}$ and $Z_{u_2,k}$ respectively. If $m_{b,k}$ represents the length of a joint, and δ_m signifies the thickness of the joint skeleton, the presence of a key point O on the b -th joint pair of the k -th person can be assessed using [Eqns. (4) and (5)].

$$0 \leq c \bullet (O - Z_{u_1,k}) \leq m_{b,k} \quad \text{and} \quad |c \times (O - Z_{u_1,k})| \leq \delta_m \quad (4)$$

$$m_{b,k} = \|Z_{u_2,k} - Z_{u_1,k}\|, c = \frac{(Z_{u_2,k} - Z_{u_1,k})}{m_{b,k}} \quad (5)$$

The acquired coordinate information of key points and their connection information are integrated to generate a coordinate graph and a connection graph. The coordinate graph pertaining to the joint logs spatial location information of each key point, whereas the connection graph depicts the relational aspects of key points. The recorded data in these graphs aid subsequent analysis.

To further refine the matching of key points and their connections, a measure of correlation between any two key points is determined. This is accomplished through the calculation of geometric distance, orientation information, or similarities in the feature space between key points. Higher correlation indices suggest the key points more likely belong to the same human skeleton. If f_{u_1} and f_{u_2} represent positions of two potential skeleton points, and $O(i)$ denotes positions on the line linking these points, the confidence of correlation between the two candidate skeleton points is assessed by $M_b(\cdot)$ [Eqns. (6) and (7)].

$$O(i) = (1-i)f_{u_1} + if_{u_2} \quad (6)$$

$$R = \int_{i=0}^{i=1} M_b(O(i)) \bullet \frac{f_{u_2} - f_{u_1}}{\|f_{u_2} - f_{u_1}\|_2} f_i \quad (7)$$

The correlation degree between key points is graphically represented in the form of a bipartite graph, with nodes symbolizing key points and edges denoting the correlation degree between key points. To identify the optimal matching result within the bipartite graph, the Hungarian algorithm or other similar graph matching algorithms can be utilized. This allows for the selection of connections that exhibit the highest correlation degree between key points, hence enhancing the accuracy of skeleton extraction.

This method represents a significant contribution to the field of skeleton extraction, providing a rigorous and comprehensive technique for determining skeletal structures from images. It should prove to be an invaluable tool for further studies in this area.

4. MODEL STRUCTURE

Traditional GCNNs often fall short in accurately capturing dependencies between central and neighbouring nodes during graph convolution operations. Consequently, such a limitation may lead to the overlooking of crucial local structural information during the processing of graph data, thus detrimentally affecting the model's recognition performance. Moreover, uniform weight assignment to all neighbouring nodes in GCNNs implies the model's inability to discern the importance amongst neighbouring nodes, thereby compromising its effectiveness in prioritizing critical nodes for prediction and classification tasks.

To address these challenges, an advanced GCNN model integrating an attention mechanism is proposed. The incorporation of the attention mechanism enables adaptive learning of dependencies between the central and neighbouring nodes. As a result, the model can effectively capture local structural information in the graph data, thereby significantly enhancing its recognition performance, especially in the context of assessing the independent living ability of the elderly. Furthermore, by enabling the model to weigh neighbouring nodes based on their importance, the

attention mechanism can prioritize critical nodes, leading to a boost in the model's generalization capability and accuracy.

However, in the context of image recognition-based independent living ability assessment for the elderly, the mere introduction of the graph attention mechanism may be insufficient due to differences between spatio-temporal graphs and standard graphs. To address this, an innovative approach was adopted which introduces the *Squeeze-and-Excitation Network (SENet)* module, integrating it with the graph attention mechanism. This module facilitates the modeling of relationships between feature channels, enabling adaptive weighting of different channels. Consequently, this approach enhances the model's capability to emphasize important features and suppress irrelevant or redundant information, thereby improving its recognition performance. By introducing the *SENet* module, the model can simultaneously consider node dependencies and relationships between feature channels, which bolsters its generalization and accuracy in complex scenarios. Consequently, the model's performance in human behavior recognition is significantly enhanced. A comprehensive diagram detailing the overall structure of the network model is presented in Figure 2.

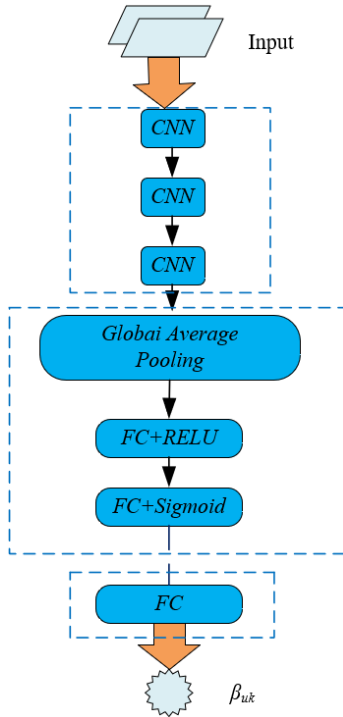


Figure 2. Structure of the network model

Given the set of eigenvectors $g = \{g_1, g_2, \dots, g_B\}$ for skeleton key points in the images of elderly behavior introduced into the graph attention layer, where g_u is in the domain of E^D and represents the eigenvector of key point u , and B represents the total number of key points in the graph. The output set of eigenvectors from the graph attention layer is denoted as $g' = \{g'_1, g'_2, \dots, g'_B\}$, where g'_u is in the domain of $R^{D'}$ and represents the transformed eigenvector of key point u . The transformation process can be represented as $g'_u = Qg_u$, where, Q is a matrix in the domain of $R^{D \times D}$ symbolizing the transformation matrix.

The attention mechanism can be expressed by function $s: R^{D'} \times R^{D'} \rightarrow E$. Given that r_{uk} indicates the importance of key point k for key point u , the linear transformation based on the eigenvectors of two skeleton key points can be written as:

$$r_{uk} = s(Qg_u, Qg_k) \quad (8)$$

where, it is worth noting that r_{uk} is calculated solely for node $k \in B_u$, where B_u represents the set of neighboring key points of key point u in the image of elderly behavior.

In the graph attention layer of the developed elderly behavior recognition model, the attention mechanism is implemented through a fully connected module. The input for this module is the concatenated result of the eigenvectors of the two aforementioned skeleton key points. Hence, the parameters of the fully connected module are considered to be the weight vector $s \rightarrow$ in the domain of $R^{2D'}$. The function s is then calculated as per the following equation:

$$s(Qg_u, Qg_k) = LR(\vec{s}^T [Qg_u \parallel Qg_k]) \quad (9)$$

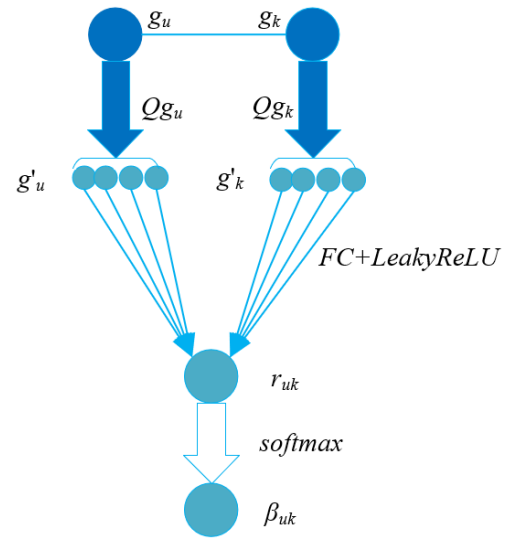


Figure 3. Calculation process of weight coefficient

The calculation process for the weight coefficient is illustrated in Figure 3. Assuming T symbolizes the transpose operation, and \parallel represents the concatenation operation between eigenvectors, the calculation formula for r_{uk} is:

$$r_{uk} = LR(\vec{s}^T [Qg_u \parallel Qg_k]) \quad (10)$$

Subsequent analysis of the correlation between two key points indicates that the number of neighboring key points of key point u in the images of elderly behavior is typically greater than one. Consequently, normalization processing on the attention coefficients between multiple neighboring key points is performed, based on the *SoftMax* function. The attention weight coefficients are then calculated as per the following formula:

$$\beta_{uk} = \frac{\exp(LR(\vec{s}^T [Qg_u \parallel Qg_k]))}{\sum_{j \in B_u} \exp(LR(\vec{s}^T [Qg_u \parallel Qg_k]))} \quad (11)$$

The eigenvector associated with neighboring key points is overlaid with the computation result from the preceding formula. Subsequently, this superimposed outcome undergoes non-linear processing, culminating in the generation of a new eigenvector corresponding to the central key point.

In the developed model for elderly behavior recognition, the *SENet* module is primarily used to map input $Z \in R^{G \times Q \times V}$ to output $I \in R^{G \times Q \times V}$. The mapping conversion function, represented by D_{ye} , is considered a convolution operator in this context. Let's denote c_u as the u -th filter kernel, with $C=[c_1, c_2, \dots, c_V]$ embodying the set of filter kernels in the convolution operator. The data information contained in the u -th channel of input Z is represented as $z_u \in R^{G \times Q}$, with $Z=[z_1, z_2, \dots, z_V]$ symbolizing the input $Z \in R^{G \times Q \times V}$. In a similar vein to the input, $i_u \in R^{G \times Q}$ denotes the data information contained in the u -th channel of output I , with $I=[i_1, i_2, \dots, i_V]$ representing the output $i_u \in R^{G \times Q}$. When $*$ symbolizes the convolution operation, and c_v^a signifies the two-dimensional filter kernel parameter, the formula for calculating the convolution operator is written as:

$$i_v = c_v * Z = \sum_{a=1}^V c_v^a * z_a \quad (12)$$

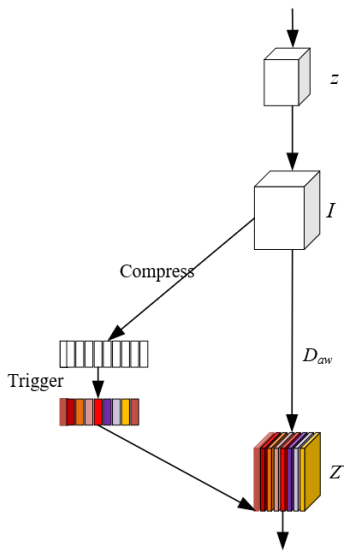


Figure 4. Calculation flow of the *SE* network module

The *SE* network module's calculation flow is depicted in Figure 4. Conventional convolutional operations exhibit certain limitations in sensing multi-channel data, potentially failing to adequately capture correlation information between channels. The *SE* network module, however, is capable of adaptively adjusting the convolutional operation results through modeling data relationships between channels. This ability significantly enhances the information display and improves the model's recognition performance in the context of elderly independent living ability assessment.

The configured *SE* network module consists of two procedures: compression and triggering. The compression operation is designed to produce a channel descriptor, denoted by a V -dimensional $x \in R^V$. The calculation formula for the v -th element of x is as follows:

$$x_v = D_{aw}(i_v) = \frac{1}{G \times Q} \sum_{u=1}^G \sum_{k=1}^Q i_v(u, k) \quad (13)$$

The triggering operation involves two steps: linear transformation and activation function processing. Here, $Q_1 \in R^{V/e \times V}$ and $Q_2 \in R^{V \times V/e}$ denote two linear transformation matrices, σ stands for the *ReLU* activation function, δ

represents the sigmoid activation function, and e signifies the hyper parameter. Accordingly, the descriptor $a \in R^V$ is constructed as:

$$a = D_{rc}(x, Q) = \delta(Q_2 \sigma(Q_1 x)) \quad (14)$$

Elements in a are linearly combined with the data in each channel of I , leading to the output $\tilde{z}=[\tilde{z}_1, \tilde{z}_2, \dots, \tilde{z}_V]$ of the *SE* network module. Each element's calculation formula is given as:

$$\tilde{z}_v = D_{sc}(i_v, a_v) = a_v i_v \quad (15)$$

5. EXPERIMENTAL RESULTS AND ANALYSIS

Analysis of the results outlined in Table 1 reveals that irrespective of whether the *Temporal Segment Network (TSN)* or the proposed model is utilized, adopting the proposed strategy for skeleton extraction has exhibited superior performance over the *VNect* strategy in terms of Top-1 and Top-5 accuracy. This finding underscores the enhanced capability of the proposed skeleton extraction strategy in capturing human body postures and behavioral features.

Moreover, under the *VNect* skeleton extraction strategy, the proposed model registered a 9.8% and 14.1% increase in Top-1 and Top-5 accuracy, respectively, compared to the *TSN*. This indicates a more effective use of skeleton information extracted by *VNect* for behavior recognition by the proposed model.

Utilizing the proposed skeleton extraction strategy, the proposed model delivered a 2.1% and 1.9% increase in Top-1 and Top-5 accuracy, respectively, compared to the *TSN*. While these increments are relatively modest, they nonetheless reaffirm the superior performance of the proposed model in extracting skeleton information under the proposed strategy. Therefore, it is confirmed that the proposed skeleton extraction strategy surpasses the *VNect* strategy in the task of behavior recognition, effectively augmenting the recognition accuracy of various models.

Table 1. Comparative analysis of behavioral recognition models under different skeleton extraction strategies

Skeleton extraction strategy	TSN		The proposed model	
	Top-1	Top-5	Top-1	Top-5
<i>VNect</i>	17.8%	36.2%	27.6%	50.3%
The proposed strategy	28.3%	61.3%	30.4%	63.2%

Table 2. Experimental outcomes under varying hyperparameter values

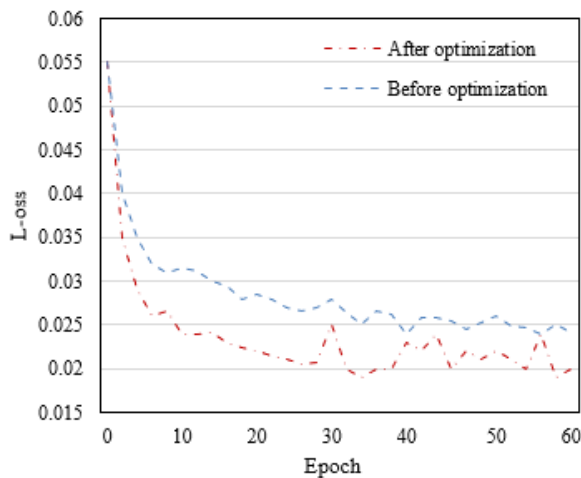
Parameter value e	Top-1	Top-5
2	32.4%	64.3%
4	32.6%	63.5%
8	32.5%	65.3%

Table 2 presents the experimental outcomes under varying hyperparameter values. When the hyperparameter e was set at 2, 4, and 8, Top-1 accuracy stood at 32.4%, 32.6%, and 32.5% respectively, indicating marginal variation under these conditions. The highest Top-1 accuracy was obtained when e was set at 4. For Top-5 accuracy, when e was set at 2, 4, and

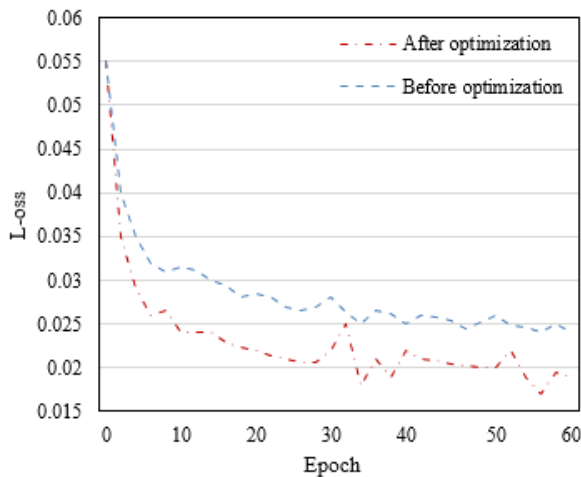
8, the results were 64.3%, 63.5%, and 65.3% respectively, with the maximum accuracy observed when e was set at 8.

Based on these experimental outcomes, it can be inferred that the model demonstrated the highest Top-1 accuracy when e was set at 4, albeit with minimal variance under the three parameter values, demonstrating model robustness against changes in the hyperparameter e . Top-5 accuracy was maximized when e was set at 8, indicating that considering the first two most likely behavior categories, a larger value of the hyperparameter 'e' can yield superior recognition results. Thus, in practical applications, a suitable value for hyperparameter e can be chosen based on specific needs and performance metrics. If Top-1 accuracy is prioritized, e can be set at 4; if Top-5 accuracy is emphasized, e can be set at 8.

In the analysis of the experimental results, the alteration of the Mean Squared Error (MSE) loss in the model, pre- and post-optimization, is highlighted in Figure 5. It is observed that the MSE loss fluctuates in line with the training epoch for the entire duration of the training process. A marked improvement in the MSE loss is noted following optimization, indicative of a faster convergence speed and superior generalization performance. Particularly in higher training epochs, the MSE loss of the optimized model is significantly less than the pre-optimization model, corroborating the benefits of model optimization in reducing prediction errors.



(a) Training



(b) Verification

Figure 5. Comparison of MSE loss of the model before and after optimization

Table 3 contains the outcomes of various behavior recognition models. Noteworthy is the proposed method's Top-1 accuracy of 35.7%, surpassing all other methods. Comparatively, it improved the Top-1 accuracy by 3.1 percentage points over TSN, at 32.6%. The proposed method also showcased superior performance in the behavior recognition task. In terms of Top-5 accuracy, the proposed method achieved 65.1%, again surpassing all other models. The improvement over the Two-Stream CNN, at 52.9%, was 12.2 percentage points, further establishing the superiority of the proposed method in complex behavior recognition tasks. Consequently, the proposed method outperformed other behavior recognition models in both Top-1 and Top-5 accuracy, indicating an aptitude for capturing behavior features and performing accurate recognition.

Table 3. Experimental results of different behavior recognition models

Method	Top-1	Top-5
CNN	21.5%	41.0%
3D-CNN	28.4%	51.2%
Two-Stream CNN	32.5%	52.9%
TSN	32.6%	43.5%
The proposed method	35.7%	65.1%

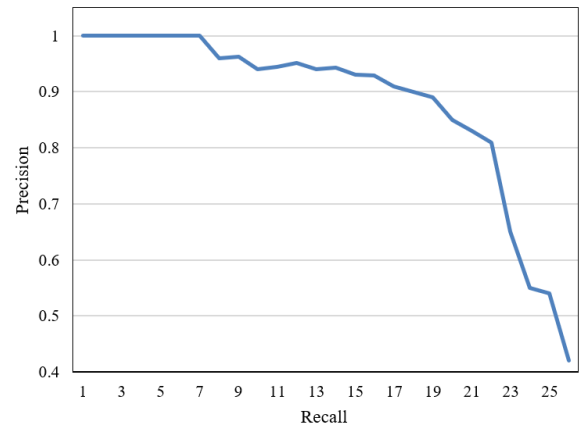


Figure 6. P - R plot of abnormal behaviors of the elderly

The Precision-Recall (P - R) plot of abnormal behaviors of the elderly is displayed in Figure 6. As the recall rate increases, precision generally decreases, an observation consistent with an increase in false positives associated with higher recall rates. Notably, the precision rate remained at 1 for a recall rate of less than or equal to 0.24, indicating accurate identification of all abnormal behaviors within this limited range. As the recall rate increased, the precision rate began to drop, although it remained at a respectable 0.943 for a recall rate of 0.52. Beyond this point, precision rates fell at a faster rate, particularly beyond a recall rate of 0.86, revealing a decrease in model recognition performance at high recall rates. Consequently, it can be concluded that the model performed optimally at moderate recall rates, such as 0.52, with a notable decrease in performance at high recall rates due to an increase in false positives.

In evaluating the experimental results shown in Figure 7, it was discerned that the model incorporating all optimization measures yielded superior recognition accuracy across all behavior categories for elderly subjects. Such results denote the efficacy of each optimization measure implemented in the model.

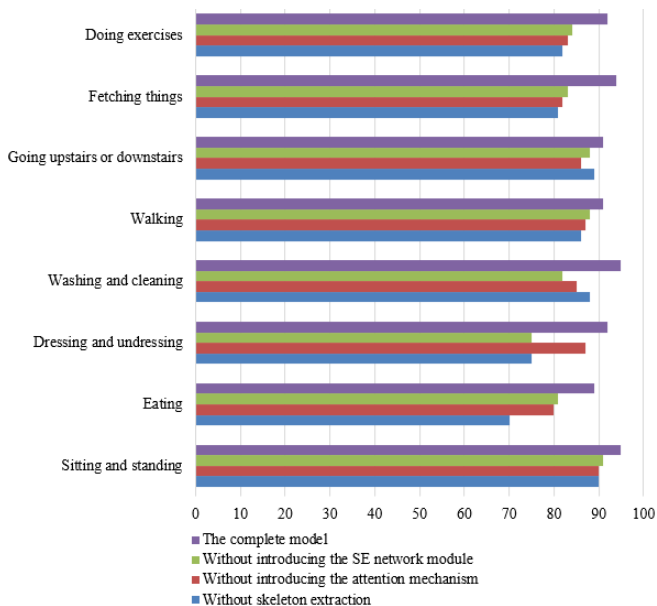


Figure 7. Statistics of behavior categories of the elderly

Further assessment revealed a marked underperformance of the model devoid of skeleton extraction, especially in the recognition of actions including eating, dressing, and undressing. This observation underscores the instrumental role of skeleton extraction in the successful capture of the relevant features of these specific behaviors.

The model bereft of the attention mechanism displayed diminished recognition capacity in the same set of behaviors, indicating the crucial contribution of the attention mechanism to bolstering focus on salient features, thereby augmenting recognition accuracy.

Analogously, the model without the *SENet* module exhibited suboptimal performance in the recognition of actions such as dressing, undressing, washing, and cleaning, suggestive of the influential role of the SE network module in these particular behaviors' recognition.

In summary, three primary inferences can be drawn. Firstly, the model complete with all optimization measures achieved superior recognition accuracy in all behavior categories, underlining the significance of these measures. Secondly, the role of skeleton extraction, the attention mechanism, and the SE network module is crucial for enhancing recognition accuracy of the elderly's behaviors. While the skeleton extraction aids in identifying behavior features, the attention mechanism amplifies focus on key features, and the SE network model assists in heightening the recognition of specific behaviors.



Figure 8. Examples of abnormal behavior recognition results

Further insights can be gleaned from Figure 8, presenting several instances of abnormal behavior recognition outcomes. Evidently, the skeleton extraction technique was successful in isolating key information pertaining to human body joint points, furnishing potent feature representations for the model.

Concurrently, the GCNN model, which integrates the attention mechanism, demonstrated high recognition accuracy, aptly distinguishing between normal and abnormal behaviors. These findings highlight the proposed skeleton extraction technique's proficiency, coupled with the GCNN model integrated with the attention mechanism, in recognizing elderly behaviors.

6. CONCLUSION

In the recent investigation into an image recognition-based system for assessing the independent living ability of the elderly, certain key findings were unearthed. This research initially applied the system in real-life contexts, followed by the detection of essential points in the skeletal structure of older individuals. Subsequently, a GCNN model was constructed with the inclusion of the attention mechanism, applied towards the recognition of elderly behavior.

Key conclusions drawn from the experimental findings can be summarized as follows:

(1) For the detection of unusual behavioral patterns in the elderly, a combination of the proposed skeletal extraction technique and the GCNN model with an attention mechanism demonstrated high accuracy. It was found that an increase in the recall rate was associated with a downward trend in model precision, emphasizing the need to balance these two parameters to optimize the recognition of abnormal behaviors in practical applications.

(2) A comparative analysis of different model optimization strategies suggested that the skeletal extraction method, attention mechanism, and the *SE* network module all contributed significantly towards enhancing the accuracy of elderly behavior recognition. The fully integrated model outperformed all others across behavior categories in terms of recognition accuracy.

(3) The proposed methodology offers significant insights for assessing the independent living ability of the elderly. By enabling precise identification and analysis of the daily behaviors of the elderly, it provides a mechanism for anticipating potential issues they may encounter during their independent living.

This research not only contributes to the field of automated health monitoring systems, but it also has potential implications for the design and implementation of home care services for the elderly. It opens up opportunities for the development of robust, effective, and personalized health monitoring systems, ensuring enhanced quality of life for the elderly. Furthermore, it prompts future exploration of other aspects of independent living for the elderly, including cognitive capabilities and social interactions, to design a comprehensive assessment system.

This study acts as a stepping stone for future research in this direction, pushing the boundaries of current knowledge and offering a novel way to address the challenge of caring for our aging population.

Please note that further investigation is warranted to validate these findings in different cultural and socio-economic contexts, as well as to address any potential ethical and privacy concerns associated with the use of image recognition technology in elder care.

The investigation underscores the critical need for the adoption of advanced technological solutions in addressing the challenges associated with the growing elderly population and their desire for independent living. By integrating advanced

image recognition techniques and deep learning models, it provides a promising avenue towards enhancing the care and wellbeing of our elders.

REFERENCES

- [1] Kang, S. (2020). A study on smart homecare for daily living ability and safety management of the elderly. In *Information Science and Applications: ICISA 2019*, pp. 707-710. https://doi.org/10.1007/978-981-15-1465-4_72
- [2] Muneeb, M., Rustam, H., Jalal, A. (2023). Automate appliances via gestures recognition for elderly living assistance. In *2023 4th International Conference on Advancements in Computational Sciences (ICACS)*, Lahore, Pakistan, pp. 1-6. <https://doi.org/10.1109/ICACS55311.2023.10089778>
- [3] Ruengtam, P. (2022). The factors of environmental living design for elderly well-being in Thai spiritual environments. In *Proceedings of the 2nd International Civil Engineering and Architecture Conference: CEAC 2022*, Singapore, pp. 409-418. https://doi.org/10.1007/978-981-19-4293-8_43
- [4] Chen, Y., Guo, Y., Liu, Q., Liu, Y., Lei, Y. (2023). Therapeutic lighting in the elderly living spaces via a daylight and artificial lighting integrated scheme. *Energy and Buildings*, 285: 112886. <https://doi.org/10.1016/j.enbuild.2023.112886>
- [5] Zhang, H., Duong, T.T., Rao, A.K., Mazzoni, P., Agrawal, S.K., Guo, Y., Zanotto, D. (2022). Transductive learning models for accurate ambulatory gait analysis in elderly residents of assisted living facilities. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 30: 124-134. <https://doi.org/10.1109/TNSRE.2022.3143094>
- [6] Xu, R., Mai, Y., Xu, T., Yang, Y., Zhang, W., Zhang, Z. (2022). Intelligent warning system for the elderly living alone based on OpenPose. In *2022 5th International Conference on Pattern Recognition and Artificial Intelligence (PRAI)*, Chengdu, China, pp. 799-803. <https://doi.org/10.1109/PRAI55851.2022.9904281>
- [7] Lim, C.K. (2022). A seamlessly integrated smart living system at a solitary elderly home. In *Proceedings of the 27th International Conference of the Association for Computer-Aided Architectural Design Research in Asia (CAADRIA)*, Hong Kong, pp. 445-454.
- [8] Zhang, H., Zhang, F., Zhang, Y., Cheng, H., Gao, R., Li, Z., Zhao, J.K., Zhang, M. (2022). An elderly living-alone guardianship model based on wavelet transform. In *2022 4th International Conference on Power and Energy Technology (ICPET)*, Beijing, China, pp. 1249-1253. <https://doi.org/10.1109/ICPET55165.2022.9918289>
- [9] Zhu, M. (2022). Living with offspring surely brings happiness to the elderly?—the heterogeneity in the effect of living arrangements on the life satisfaction of the Chinese elderly. *Procedia Computer Science*, 214: 359-366. <https://doi.org/10.1016/j.procs.2022.11.186>
- [10] Tanaka, K., Kudo, M., Kimura, K. (2022). Sensor data simulation with wandering behavior for the elderly living alone. In *2022 26th International Conference on Pattern Recognition (ICPR)*, Montreal, QC, Canada, pp. 885-891. <https://doi.org/10.1109/ICPR56361.2022.9956332>
- [11] Li, Y., Lin, Z., Huang, Z., Cai, Z., Huang, L., Wei, Z. (2022). A channel hopping lora technology based emergency communication system for elderly people living alone. In *2022 21st International Symposium on Communications and Information Technologies (ISCIT)*, Xi'an, China, pp. 19-26. <https://doi.org/10.1109/ISCIT55906.2022.9931253>
- [12] Alsaeedi, A., Jabeen, S., Kolivand, H. (2022). Ambient assisted living framework for elderly care using Internet of medical things, smart sensors, and GRU deep learning techniques. *Journal of Ambient Intelligence and Smart Environments*, 14(1): 5-23. <https://doi.org/10.3233/AIS-210162>
- [13] Miyazaki, Y., Hirano, K., Kitamura, K., Nishida, Y. (2022). Analysis of relationship between natural standing behavior of elderly people and a class of standing aids in a living space. *Sensors*, 22(3): 1178. <https://doi.org/10.3390/s22031178>
- [14] Irisawa, H., Mizushima, T. (2022). Assessment of changes in muscle mass, strength, and quality and activities of daily living in elderly stroke patients. *International Journal of Rehabilitation Research. Internationale Zeitschrift Fur Rehabilitationsforschung. Revue Internationale De Recherches De Readaptation*, 45(2): 161-167. <https://doi.org/10.1097/MRR.0000000000000523>
- [15] He, J., Xiang, M., Zhao, X. (2022). An elderly indoor behavior recognition method based on improved slowfast network. *Journal of Physics: Conference Series*, 2216(1): 012102. <https://doi.org/10.1088/1742-6596/2216/1/012102>
- [16] He, Y., Huang, H., Wu, Y., Zhu, G. (2022). Research on abnormal behavior recognition of the elderly based on spatial-temporal feature fusion. In *Proceedings of the 3rd International Symposium on Artificial Intelligence for Medicine Sciences*, Amsterdam, Netherlands, pp. 85-92. <https://doi.org/10.1145/3570773.3570823>
- [17] Shang, C., Chang, C.Y. (2021). Behavior recognition algorithm using unsupervised learning for home elderly. In *2021 IEEE International Conference on Consumer Electronics-Taiwan (ICCE-TW)*, Penghu, Taiwan, pp. 1-2. <https://doi.org/10.1109/ICCE-TW52618.2021.9603268>
- [18] Li, Y., Cao, Y., Lin, Q., Wang, W. (2021). Research on the recognition of abnormal behaviors in the elderly based on Wi-Fi signals. *Journal of Physics: Conference Series*, 1848(1): 012077. <https://doi.org/10.1088/1742-6596/1848/1/012077>
- [19] Lentzas, A., Vrakas, D. (2020). Non-intrusive human activity recognition and abnormal behavior detection on elderly people: A review. *Artificial Intelligence Review*, 53(3): 1975-2021. <https://doi.org/10.1007/s10462-019-09724-5>
- [20] Anitha, G., Baghavathi Priya, S. (2019). Posture based health monitoring and unusual behavior recognition system for elderly using dynamic Bayesian network. *Cluster Computing*, 22, 13583-13590. <https://doi.org/10.1007/s10586-018-2010-9>
- [21] Zhang, X., Liu, H., Zhang, M. (2018). A new behavior recognition method of nursing-care robots for elderly people. In *Recent Developments in Mechatronics and Intelligent Robotics: Proceedings of the International Conference on Mechatronics and Intelligent Robotics (ICMIR2017)*, Kunming, China, pp. 547-553. https://doi.org/10.1007/978-3-319-65978-7_82
- [22] Endelin, R., Renouard, S., Tiberghien, T., Aloulou, H.,

Mokhtari, M. (2013). Behavior recognition for elderly people in large-scale deployment. In *Inclusive Society: Health and Wellbeing in the Community, and Care at Home: 11th International Conference on Smart Homes and Health Telematics, ICOST 2013, Singapore*, pp. 61-68. https://doi.org/10.1007/978-3-642-39470-6_8

[23] Ito, Y., Ueno, M., Takahashi, H., Chiba, S., Abe, T.,

Suganuma, T. (2020). Implementation of a behavioral recognition sensor for understanding interactive communication situation of the elderly. In *2020 IEEE 9th Global Conference on Consumer Electronics (GCCE)*, Kobe, Japan, pp. 806-807. <https://doi.org/10.1109/GCCE50665.2020.9292059>