# Causality-Informed Fusion Network for Automated Assessment of Parkinsonian Body Bradykinesia

Yuyang Quan[1], Chencheng Zhang[2], Rui Guo[1(✉)], and Xiaohua Qian[1(✉)]

[1]Medical Image and Health Informatics Lab, School of Biomedical Engineering, Shanghai Jiao Tong University, Shanghai, 200230, China
graymm@sjtu.edu.cn
xiaohua.qian@sjtu.edu.cn
[2]Department of Functional Neurosurgery, Ruijin Hospital, Shanghai Jiao Tong University School of Medicine, Shanghai, 200025, China

**Abstract.** Body bradykinesia, a prominent clinical manifestation of Parkinson's disease (PD), characterizes a generalized slowness and diminished movement across the entire body. The assessment of body bradykinesia in the widely employed PD rating scale (MDS-UPDRS) is inherently subjective, relying on the examiner's overall judgment rather than specific motor tasks. Therefore, we propose a graph convolutional network (GCN) scheme for automated video-based assessment of parkinsonian body bradykinesia. This scheme incorporates a causality-informed fusion network to enhance the fusion of causal components within gait and leg-agility motion features, achieving stable multi-class assessment of body bradykinesia. Specifically, an adaptive causal feature selection module is developed to extract pertinent features for body bradykinesia assessment, effectively mitigating the influence of non-causal features. Simultaneously, a causality-informed optimization strategy is designed to refine the causality feature selection module, improving its capacity to capture causal features. Our method achieves 61.07% accuracy for three-class assessment on a dataset of 876 clinical case. Notably, our proposed scheme, utilizing only consumer-level cameras, holds significant promise for remote PD bradykinesia assessment.

**Keywords:** Parkinson's disease, body bradykinesia, graph convolution network, causality guidance, feature fusion.

## 1 Introduction

Parkinson's disease (PD) stands as a prevalent neurodegenerative condition among the elderly [1]. Currently, the Movement Disorder Society-sponsored revision of the Unified Parkinson's Disease Rating Scale (MDS-UPDRS) serves as the predominant clinical tool for PD assessment [2]. In particular, body bradykinesia (i.e. global spontaneity of movement, item 3.14 in MDS-UPDRS) plays a pivotal role in indicating the global slowness in patients. Severity is graded on a scale from 0 to 4, correlating with the extent of global slowness and poverty of spontaneous movements. The subjective nature of this assessment, attributed to the absence of specific examination tasks for assessing body bradykinesia, results in significant variability among different examiners. Therefore, an automated assessment system for body bradykinesia is needed to ensure consistent ratings. To the best of our knowledge, there is currently no existing research on automated assessment for body bradykinesia.

Various initiatives have been undertaken to develop video-based automated assessments for different motor examination tasks within MDS-UPDRS. These approaches offer notable advantages, including convenience, minimal equipment requirements, and non-contact interaction. The fundamental process involves utilizing a pose estimator to derive human skeletons from the video, extracting motion features, and employing a classifier for automated assessment [3-10]. For example, Lu *et al.* [3] used DD-Net, a structure based on convolutional neural network (CNN), for extracting motion features from gait task. Notably, graph convolutional networks (GCNs) can explicitly model spatial relationships in human body skeletons [11], making it well-suited for video-based assessments of PD examination tasks [4, 5, 9]. Zhang *et al.* [5] employed GCN with a pyramidal attention module to enhance the capture of parkinsonian tremor signals. Guo *et al.* [4] designed a sparse adaptive GCN to improve the modeling of spatiotemporal logical connections for more accurate leg-agility assessment.

To achieve the automated assessment of PD body bradykinesia, we propose a GCN-based scheme, marking the first attempt to combine motion features extracted from leg-agility and gait videos for indirect rating. The leg-agility task, involving rapid foot raising and stomping, effectively reveals slowness severity [12]. Meanwhile, the gait task, where patients walk towards and away from the examiner, allows for the observation of spontaneous movement poverty, such as the lack of arm swing [13, 14]. However, challenges arise from the presence of non-causal features, such as those irrelevant to body bradykinesia assessment. Therefore, the key challenge in this study is how to enhance the fusion of causal features while suppressing non-causal ones.

To tackle this challenge, our approach involves guiding the model to unearth causal features before the fusion stage. Therefore, we design a causality-informed fusion network based on GCNs for the automated assessment of PD body bradykinesia. Specifically, human skeletons for leg-agility and gait are firstly extracted using an advanced pose estimator, followed by motion feature extraction through GCN encoders. Subsequently, a novel adaptive causal feature selection module is developed to learn the contribution of each channel, driving the model to capture causal features effectively. Furthermore, a causality-informed optimization strategy is designed to guide the iterative optimization of GCNs and causal feature selection modules through opposing optimization objectives, further enhancing the ability to capture causal features.

In summary, the contributions of our work are as follows:

- Clinically, we pioneer a framework for video-based automated assessment of PD body bradykinesia, requiring only consumer-level cameras.
- Technically, we propose a causality-informed fusion network, enhancing the fusion of causal features through a novel adaptive causal feature selection module and a causality-informed optimization strategy.

## 2 Method

To achieve automated assessment of PD body bradykinesia, we propose a causality-informed fusion network which effectively enhance the fusion of causal features across different tasks. The technical details will be described in this section.

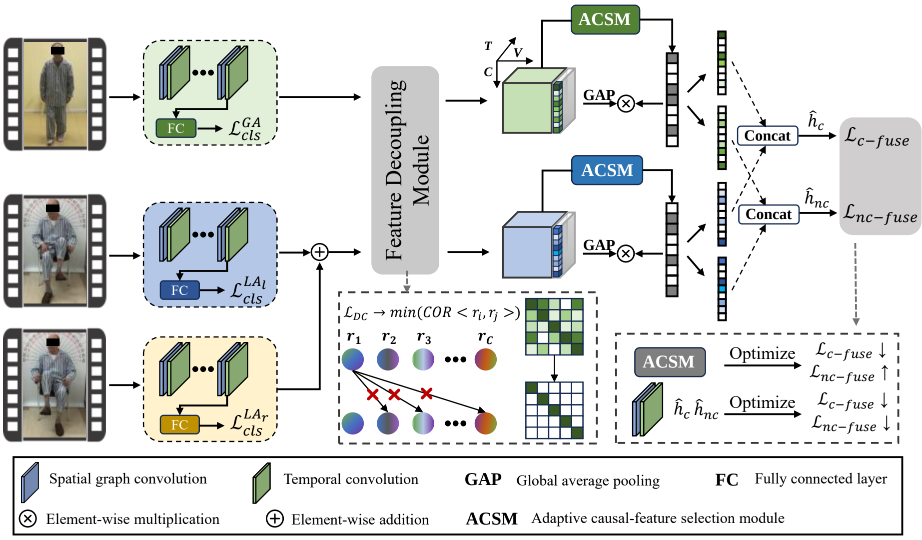


**Fig. 1**. The overview of our proposed causality-informed fusion network.

### 2.1 Pipeline

The overview of the proposed scheme is illustrated in Fig. 1. In the preprocessing stage, three skeleton sequences will be obtained by OpenPose [15] from RGB videos of gait and leg-agility of both legs. Every skeleton sequence is defined by a feature matrix $\mathcal{X} \in \mathbb{R}^{C_{in} \times T_{in} \times V}$ and an adjacency matrix $\mathcal{A} \in \mathbb{R}^{V \times V}$ representing the physical connections of human body with self-connection, where $C_{in}$ is the number of channels, $T_{in}$ is the length of sequence and $V$ is the number of human joints.

In the encoding stage, given the skeleton sequences, the feature maps $F_{GA}, F_{LA_l}, F_{LA_r} \in \mathbb{R}^{C \times T \times V}$ are then extracted by three independent GCN encoders $E = \{E_{GA}, E_{LA_l}, E_{LA_r}\}$ respectively, where $C$, $T$ and $V$ represent channel, temporal and joint dimension. The update rule of the $l$-th layer in GCN can be described as:

$$Z^{(l+1)} = \sigma\left(D^{-1}\mathcal{A}X^{(l)}W^{(l)}\right), \tag{1}$$

where $Z^{(l+1)} \in \mathbb{R}^{C^{(l+1)} \times T \times V}$ represents the output feature matrix of the current layer, $D \in \mathbb{R}^{V \times V}$ is the degree matrix of $\mathcal{A}$ for normalization [16], and $W^{(l)}$ denotes the learnable weights of this layer. In the meanwhile, to learn the spatiotemporal information of each task and speed up the model convergence, we perform multi-class classification loss $\mathcal{L}_{cls}$ for each task to supervise the learning of GCN encoders:

$$\mathcal{L}_{cls} = \sum\nolimits_{t \in T} \ell(\hat{h}_t(GAP(F_t), y_t)\,, \tag{2}$$

where $T$ denotes the set of motor examination tasks, i.e. $T = \{GA, LA_l, LA_r\}$, $\hat{h}_t, y_t$ is the classifier and label for the corresponding task, GAP means global average pooling, $\ell$ represents cross-entropy loss function.

In the fusion stage, the feature maps of left and right leg-agility are firstly element-wise summed to obtain $F_{LA} \in \mathbb{R}^{C \times T \times V}$ representing features of both legs. Afterwards, $F_{GA}$ and $F_{LA}$ are fed into the adaptive causal feature selection module, where causal features are separated from non-causal features. Finally, causal and non-causal features are concatenated separately. Classification loss of fused causal features $\mathcal{L}_{c-fuse}$ and fused non-causal features $\mathcal{L}_{nc-fuse}$ are obtained by two classifiers $\hat{h}_c$ and $\hat{h}_{nc}$.

### 2.2 Adaptive Causal Feature Selection Module

An adaptive causal feature selection module is designed to extract discriminative causal features for fusion. The entire process can be divided into three steps: feature decoupling, contribution learning, and causality-driven feature sampling.

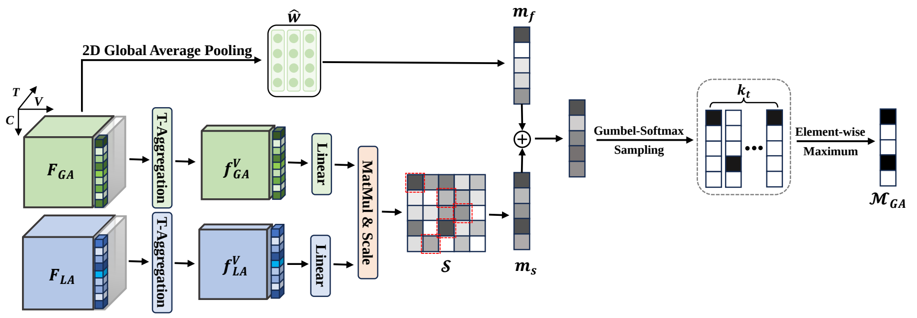


**Fig. 2**. Illustration of contribution learning and causal feature sampling.

**Feature Decoupling:** To better distinguish between causal and non-causal components at the channel dimension, each feature channel should be jointly independent, i.e. there is minimal information overlap between different channels [17]. Therefore, we employ a feature decoupling loss to obtain independent channels. We illustrate the details using the gait branch as an example, and the same loss is also applied to the leg-agility branch.

For motion features $F_{GA} \in \mathbb{R}^{C \times T \times V}$, information across frames and joints is aggregated by global average pooling, followed by z-score normalization in the batch dimension to get feature matrix $\boldsymbol{\mathcal{R}}_{GA} = [r_1, r_2, \cdots, r_N]^T \in \mathbb{R}^{N \times C}$, where $N$ is the batch size, $r_i \in \mathbb{R}^{C \times 1}$ is the feature of the $i$-th sample. We construct channel correlation matrix $\boldsymbol{\mathcal{C}}$:

$$\boldsymbol{\mathcal{C}}_{ij} = \frac{<\tilde{r}_i,\ \tilde{r}_j>}{||\tilde{r}_i|| \cdot ||\tilde{r}_j||}, i, j \in 1, 2, \ldots, C, \tag{3}$$

where $\tilde{r}_i$ represents the $i$th column of $\boldsymbol{\mathcal{R}}_{GA}$, $<\cdot>$ denotes the inner product operation. The non-diagonal elements of $\boldsymbol{C}$ measure the correlation between different channels, thus they need to be minimized. The feature decoupling loss $\mathcal{L}_{DC}$ can be defined as:

$$\mathcal{L}_{DC}^{GA} = \frac{1}{2C(C-1)}||\boldsymbol{C} - diag(\boldsymbol{C})||_F^2, \quad \mathcal{L}_{DC} = \mathcal{L}_{DC}^{GA} + \mathcal{L}_{DC}^{LA}, \tag{4}$$

where $diag(\cdot)$ constructs a diagonal matrix containing only the principal diagonal elements of the input matrix, $||\cdot||$ denotes Frobenius norm. The overall decoupling loss is the sum of gait branch decoupling loss and leg-agility branch decoupling loss.

**Contribution Learning Module:** To identify which feature channels are causal, we propose a contribution learning module $\mathcal{G}$ to obtain the contribution of each channel. Two contribution learning modules $\mathcal{G} = \{\mathcal{G}_{GA}, \mathcal{G}_{LA}\}$ with the same structure but independent parameters are utilized for gait and leg-agility branch, respectively. Here we introduce this module using gait branch as an example. The process of contribution learning and causal feature sampling is illustrated in Fig. 2.

Firstly, a multi-layer perceptron (MLP) network denoted as $\widehat{w}$ is utilized to learn the contribution of each channel based on its motion features:

$$m_f = softmax\left(\widehat{w}\big(GAP(F_{GA})\big)\right) \in \mathbb{R}^C. \tag{5}$$

Additionally, body bradykinesia can result in some similar movement abnormalities when patients perform gait and leg-agility tasks [18]. These features are also beneficial for body bradykinesia assessment. The motion feature for gait $F_{GA}$ and leg agility $F_{LA}$ are aggregated over the time dimension using different linear layers, resulting in features representing each body joint $f_{GA}^V, f_{LA}^V \in \mathbb{R}^{V\times C}$. Since the arrangement of joints for both tasks is the same, similarity matrix $\boldsymbol{S}$ is defined as:

$$Q = f_{GA}^V W_1, \; K = f_{LA}^V W_2, \; \boldsymbol{S} = Softmax\left(\frac{Q^T K}{\sqrt{C}}\right) \in \mathbb{R}^{C\times C}, \tag{6}$$

where $\sqrt{C}$ is a scaling factor, $W_1$ and $W_2$ denote linear layers. $\boldsymbol{S}_{ij}$ represents the similarity between the $i$-th channel of gait and the $j$-th channel of leg-agility. Therefore, the maximum value of each row also serves as a measure of contribution of different channels. The contribution of all channels of gait features is the summation of $m_f$ and $m_s$.

$$m_s = max\big(\boldsymbol{S}_{i,:}\big) \in \mathbb{R}^C, \quad \mathcal{G}_{GA}(F_{GA}) = m_f + m_s \in \mathbb{R}^C. \tag{7}$$

**Causality-Driven Feature Sampling:** During training, the derivable Gumbel-Softmax algorithm [19] is employed to sample $k_t$ one-hot vectors based on the contribution values. Channels selected are considered as causal ones. The maskers for sampling are defined as:

$$\begin{aligned} \mathcal{M}_{GA} &= max_e\{\mathrm{Gumbel-Softmax}(\mathcal{G}_{GA}(F_{GA}), k_t)\} \in \mathbb{R}^C \\ \mathcal{M}_{LA} &= max_e\{\mathrm{Gumbel-Softmax}(\mathcal{G}_{LA}(F_{FA}), k_t)\} \in \mathbb{R}^C, \end{aligned} \tag{8}$$

where $max_e$ denotes element-wise maximum operation. Finally, the fusion classification losses for causal and non-causal features are represented as:

$$\mathcal{L}_{c-fuse} = \ell\big(\hat{h}_c\big((f_{GA}^C \odot \mathcal{M}_{GA}) \oplus (f_{LA}^C \odot \mathcal{M}_{LA})\big), y\big)$$
$$\mathcal{L}_{nc-fuse} = \ell\big(\hat{h}_{nc}\big((f_{GA}^C \odot (1-\mathcal{M}_{GA})) \oplus (f_{LA}^C \odot (1-\mathcal{M}_{LA}))\big), y\big), \quad (9)$$

where $f_{GA}^C = GAP_{2D}(F_{GA}) \in \mathbb{R}^C$, $f_{LA}^C = GAP_{2D}(F_{LA}) \in \mathbb{R}^C$, $\odot$ represents element-wise multiplication, $\oplus$ denotes concatenation operation, $\hat{h}_c$ and $\hat{h}_{nc}$ are MLP-based classifiers with independent parameters.

### 2.3 Causality-informed Optimization Strategy

To better extract causal features, a causality-informed optimization strategy is designed by dividing each epoch into two stages. In the first stage, we fix the parameters of contribution learning modules $\mathcal{G}$ and optimize the encoders $E$ and classifiers $\hat{h}_c$, $\hat{h}_{nc}$ by minimizing $\mathcal{L}_{cls}$, $\mathcal{L}_{DC}$ and two fusion losses $\mathcal{L}_{c-fuse}$, $\mathcal{L}_{nc-fuse}$. In the second stage, we fix the parameters of encoders and classifiers and optimize the contribution learning modules $\mathcal{G}$ by minimizing causal fusion loss $\mathcal{L}_{c-fuse}$ while maximizing non-causal fusion loss $\mathcal{L}_{nc-fuse}$. This iterative strategy can enhance the capacity of the causal feature selection module because 1) with an optimized $\hat{h}_{nc}$ to minimize $\mathcal{L}_{nc-fuse}$ based on current maskers, optimizing $\mathcal{G}$ to select channels for maximizing $\mathcal{L}_{nc-fuse}$ can find channels with less contribution; 2) since causal channels and non-causal channels are complementary, better selection of non-causal channels will be beneficial for causal channel extraction, and 3) the objective of $\mathcal{G}$ is opposite to that of $E$ and $\hat{h}$, therefore the ability to extract causal channels will increase synchronously with the optimization of $E$ and $\hat{h}$. The overall optimization objective can be summarized as:

$$\min_{E,\hat{h}_c,\hat{h}_{nc},\hat{h}_t} \mathcal{L}_{c-fuse} + \mathcal{L}_{nc-fuse} + \lambda_1 \mathcal{L}_{cls} + \lambda_2 \mathcal{L}_{DC}, \quad \min_{\mathcal{G}} \mathcal{L}_{c-fuse} - \mathcal{L}_{nc-fuse}, \quad (10)$$

where $\lambda_1$ and $\lambda_2$ are hyperparameters for balancing various loss items. During inference, we directly select the top $k_e$ channels with the highest contribution values from $\mathcal{G}_{GA}(F_{GA})$ and $\mathcal{G}_{LA}(F_{LA})$ as causal features for fusion, and $\hat{h}_c$ is used for classification.

## 3 Experiments

### 3.1 Datasets

The video dataset utilized in our study was compiled by Department of Functional Neurosurgery at Ruijin Hospital, Shanghai Jiao Tong University School of Medicine in China between 2017 and 2020. During the data acquisition process, a consumer camera was positioned in front of the patients, who were instructed to perform gait and leg-agility tasks for both left and right legs. The gait, leg-agility and bradykinesia were scored by experienced neurosurgeons. According to the MDS-UPDRS scores for body bradykinesia, patients with a score 0 indicating no bradykinesia were found to be rare, while those with a score 4 faced challenges in completing gait and leg-agility tasks. Therefore, aligning with clinical requisites, we classified body bradykinesia into three categories: mild bradykinesia (⩽1), moderate bradykinesia (2) and severe bradykinesia (⩾3). Totally 876 sets of videos from 293 patients were collected, comprising 297 cases of mild bradykinesia, 346 cases of moderate bradykinesia and 233 cases of severe bradykinesia.

## 3.2 Implementation Details

In our experiments, the training epochs was set to 70. The initial learning rate was set to 0.005 and decays with a cosine scheduler. For the first 5 epochs, we did not utilize $\mathcal{M}_{GA}$ and $\mathcal{M}_{LA}$ to extract causal features, i.e. all features were directly concatenated and fed into $\hat{h}_c$ and $\hat{h}_{nc}$. $\lambda_1$ and $\lambda_2$ in Eq. (10) were experimentally set to 2 and 0.01, respectively. The proportion of causal channels was set as $k_t = C/2$, $k_e = C/3$. The batch size was set to 32, and the stochastic gradient descent strategy was employed to tune the parameters. EfficientGCN [20] was utilized as the encoders $E$. We employ five-fold cross-validation to evaluate the performance, and accuracy, macro average precision, recall and F1-score are used as the evaluation metrics for classification.

## 3.3 Assessment Performance

The performance and confusion matrix of our proposed method is illustrated in Table 1 and Fig. 3 (a). Compared to recent studies [4, 21], this work tackles a more challenging task due to the absence of specific examination actions. As a pioneering attempt to utilize multimodal video information for indirect assessment of body bradykinesia, the proposed method achieved satisfactory performance in PD body bradykinesia assessment and demonstrated balanced performance across all categories.

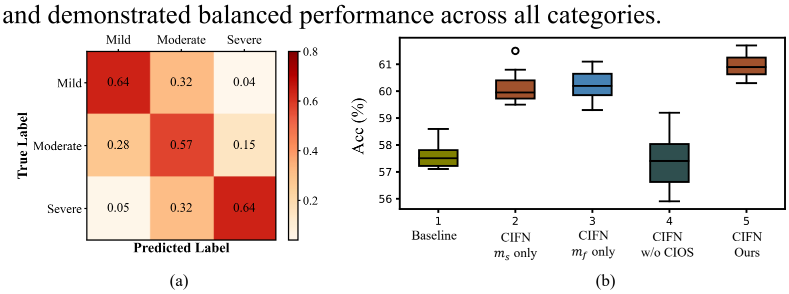


**Fig. 3.** (a) Confusion matrix under cross-validation; (b) repeated experimental results of the baseline, ablation methods and the proposed method.

## 3.4 Ablation Studies

The comparison results of the proposed method with baseline and ablation methods are summarized in Table 2 and Fig. 3 (b). We took the direct concatenation of gait and leg-agility features as the baseline. To assess the impact of the proposed optimization strategy, we optimized all components using a unified objective aimed at minimize $\mathcal{L}_{c-fuse}$, which is denoted as "w/o CIOS". The results indicate that the absence of the optimization strategy leads to a significant decrease in the ability to extract causal channels. Additionally, we conducted ablation experiments on the two sub-components $m_f$ and $m_s$ of the contribution learning module $\mathcal{G}$, confirming that including both parts yields the optimal performance.

**Table 1.** Performance of the proposed method.

| Body-bradykinesia grading | Acc (%) | Prec (%) | Rec (%) | F1 (%) |
|---|---|---|---|---|
| Mild (1-) | 63.64 | 63.64 | 63.64 | 63.64 |
| Moderate (2) | 57.23 | 53.80 | 57.23 | 55.46 |
| Severe (3+) | 63.52 | 70.14 | 63.52 | 66.67 |
| Total | 61.07 | 62.53 | 61.46 | 61.99 |

**Table 2**. Comparison results with baseline and ablation methods (CIFN: causality-informed fusion network; CIOS: causality-informed optimization strategy).

| Methods | Acc (%) | Prec (%) | Rec (%) | F1 (%) |
|---|---|---|---|---|
| Baseline | 57.62±0.48 | 59.42±1.32 | 58.24±0.59 | 58.81±0.69 |
| CIFN ($m_s$ only) | 60.14±0.59 | 61.16±0.25 | 61.00±0.71 | 61.08±0.29 |
| CIFN ($m_f$ only) | 60.21±0.55 | 61.23±0.34 | 61.40±0.83 | 61.32±0.42 |
| CIFN (w / o CIOS) | 57.38±0.96 | 58.54±0.45 | 58.62±0.47 | 58.58±0.45 |
| CIFN (Ours) | **60.95±0.43** | **62.32±0.68** | **61.76±0.29** | **62.04±0.47** |

**Table 3**. Stability analysis with various encoders (Vanilla fusion: all features are directly concatenated for fusion; CI: Causality-Informed).

| Encoder structure | Fusion mode | Acc (%) | Prec (%) | Rec (%) | F1 (%) |
|---|---|---|---|---|---|
| ST-GCN [11] | Vanilla fusion | 57.08±0.51 | 58.28±0.48 | 57.72±0.86 | 58.00±0.56 |
| | CI fusion | 60.84±0.62 | 62.14±0.74 | 61.36±0.74 | 61.74±0.66 |
| 2s-AGCN [22] | Vanilla fusion | 59.80±0.49 | 60.83±0.98 | 60.64±0.70 | 60.73±0.52 |
| | CI fusion | 60.90±0.76 | 62.44±0.59 | 61.40±0.94 | 61.91±0.72 |
| MS-G3D [23] | Vanilla fusion | 60.24±0.65 | 61.66±0.83 | 60.86±0.84 | 61.25±0.69 |
| | CI fusion | 61.62±0.57 | 63.14±0.80 | 62.08±0.75 | 62.60±0.58 |
| EfficientGCN [20] | Vanilla fusion | 57.62±0.48 | 59.42±1.32 | 58.24±0.59 | 58.81±0.69 |
| | CI fusion | 60.95±0.43 | 62.32±0.68 | 61.76±0.29 | 62.04±0.47 |

### 3.5 Stability analysis

To further validate the stability of our proposed causality-informed fusion scheme, we conducted additional experiments with several state-of-the-art GCN encoders including ST-GCN [11], 2s-AGCN [22] and MS-G3D [23]. The results are summarized in Table 3, with all experiments randomly repeated for five times. The results indicate that 1) our proposed causality-informed (CI) fusion scheme outperforms vanilla fusion scheme without causality guidance under the same encoder architecture (p-values < 0.05 in t-test), demonstrating an enhancement in performance, and 2) under the CI fusion scheme, there is no statistically significant difference (p-values > 0.05 in t-test) in performance among these four encoder architectures. This observation underscores the robustness of our proposed fusion scheme to encoder architecture variations, effectively enhancing the stability of feature fusion. Even though MS-G3D with CI fusion achieved better performance, EfficientGCN is selected as the backbone in this study due to its advantages in deployability (1/9 FLOPs compared to 2s-AGCN and 1/12 FLOPs compared to MS-G3D).

## 4 Conclusion

The automated assessment of body bradykinesia is crucial for PD diagnosis and treatment. In this study, we develop a video-based assessment scheme for body bradykinesia, which only requires consumer-level cameras. Technically, we design a novel causality-informed fusion network to enhance the fusion of causal features while mitigating non-causal components. Our proposed scheme presents a powerful tool for PD bradykinesia assessment, demonstrating immense potential for widespread applications.

**Acknowledgments.** This study was supported by the Guangxi Key Research and Development Project (grant number 2023AB22106), Special Fund for Transformation of Scientific and Technological Achievements of Inner Mongolia Autonomous Region of China (grant number 2021CG0052), National Natural Science Foundation of China (grant number 62371286 and 62171273), and Natural Science Foundation of Shanghai (grant number 22ZR1432100).

**Disclosure of Interests.** The authors have no competing interests to declare that are relevant to the content of this article.

## References

1. Jankovic, J.: Parkinson's disease: clinical features and diagnosis. Journal of neurology, neurosurgery & psychiatry 79, 368-376 (2008)
2. Goetz, C.G., Tilley, B.C., Shaftman, S.R., Stebbins, G.T., Fahn, S., Martinez-Martin, P., Poewe, W., Sampaio, C., Stern, M.B., Dodel, R.: Movement Disorder Society-sponsored revision of the Unified Parkinson's Disease Rating Scale (MDS-UPDRS): scale presentation and clinimetric testing results. Movement disorders: official journal of the Movement Disorder Society 23, 2129-2170 (2008)
3. Lu, M., Poston, K., Pfefferbaum, A., Sullivan, E.V., Fei-Fei, L., Pohl, K.M., Niebles, J.C., Adeli, E.: Vision-based estimation of MDS-UPDRS gait scores for assessing Parkinson's disease motor severity. In: Medical Image Computing and Computer Assisted Intervention–MICCAI 2020: 23rd International Conference, Lima, Peru, October 4–8, 2020, Proceedings, Part III 23, pp. 637-647. Springer, (2020)

4. Guo, R., Shao, X., Zhang, C., Qian, X.: Sparse adaptive graph convolutional network for leg agility assessment in Parkinson's disease. IEEE Transactions on Neural Systems and Rehabilitation Engineering 28, 2837-2848 (2020)
5. Zhang, H., Ho, E.S., Zhang, X., Shum, H.P.: Pose-based tremor classification for parkinson's disease diagnosis from video. In: International Conference on Medical Image Computing and Computer-Assisted Intervention, pp. 489-499. Springer, (2020)
6. Guo, Z., Zeng, W., Yu, T., Xu, Y., Xiao, Y., Cao, X., Cao, Z.: Vision-based finger tapping test in patients with Parkinson's disease via spatial-temporal 3D hand pose estimation. IEEE Journal of Biomedical and Health Informatics 26, 3848-3859 (2022)
7. Morinan, G., Peng, Y., Rupprechter, S., Weil, R.S., Leyland, L.-A., Foltynie, T., Sibley, K., Baig, F., Morgante, F., Wilt, R.: Computer-vision based method for quantifying rising from chair in Parkinson's disease patients. Intelligence-Based Medicine 6, 100046 (2022)
8. Novotny, M., Tykalova, T., Ruzickova, H., Ruzicka, E., Dusek, P., Rusz, J.: Automated video-based assessment of facial bradykinesia in de-novo Parkinson's disease. NPJ digital medicine 5, 98 (2022)
9. Guo, R., Sun, J., Zhang, C., Qian, X.: A self-supervised metric learning framework for the arising-from-chair assessment of parkinsonians with graph convolutional networks. IEEE Transactions on Circuits and Systems for Video Technology 32, 6461-6471 (2022)
10. Liu, W., Lin, X., Chen, X., Wang, Q., Wang, X., Yang, B., Cai, N., Chen, R., Chen, G., Lin, Y.: Vision-based estimation of MDS-UPDRS scores for quantifying Parkinson's disease tremor severity. Medical Image Analysis 85, 102754 (2023)
11. Yan, S., Xiong, Y., Lin, D.: Spatial temporal graph convolutional networks for skeleton-based action recognition. In: Proceedings of the AAAI conference on artificial intelligence. (2018)
12. Dunnewold, R., Jacobi, C., Van Hilten, J.: Quantitative assessment of bradykinesia in patients with Parkinson's disease. Journal of neuroscience methods 74, 107-112 (1997)
13. Samà, A., Pérez-López, C., Rodríguez-Martín, D., Català, A., Moreno-Aróstegui, J.M., Cabestany, J., de Mingo, E., Rodríguez-Molinero, A.: Estimating bradykinesia severity in Parkinson's disease by analysing gait through a waist-worn sensor. Computers in biology and medicine 84, 114-123 (2017)
14. Creaby, M.W., Cole, M.H.: Gait characteristics and falls in Parkinson's disease: A systematic review and meta-analysis. Parkinsonism & related disorders 57, 1-8 (2018)
15. Cao, Z., Simon, T., Wei, S.-E., Sheikh, Y.: Realtime multi-person 2d pose estimation using part affinity fields. In: Proceedings of the IEEE conference on computer vision and pattern recognition, pp. 7291-7299 (2017)
16. Kipf, T.N., Welling, M.: Semi-supervised classification with graph convolutional networks. arXiv preprint arXiv:1609.02907 (2016)
17. Peters, J., Janzing, D., Schölkopf, B.: Elements of causal inference: foundations and learning algorithms. The MIT Press (2017)
18. Mentzel, T.Q., Mentzel, C.L., Mentzel, S.V., Lieverse, R., Daanen, H.A., van Harten, P.N.: Instrumental assessment of bradykinesia: a comparison between motor tasks. IEEE journal of biomedical and health informatics 20, 521-526 (2015)
19. Jang, E., Gu, S., Poole, B.: Categorical Reparameterization with Gumbel-Softmax. In: International Conference on Learning Representations (2016)

20. Song, Y.-F., Zhang, Z., Shan, C., Wang, L.: Constructing stronger and faster baselines for skeleton-based action recognition. IEEE transactions on pattern analysis and machine intelligence 45, 1474-1488 (2022)
21. Guo, R., Shao, X., Zhang, C., Qian, X.: Multi-scale sparse graph convolutional network for the assessment of Parkinsonian gait. IEEE Transactions on Multimedia 24, 1583-1594 (2021)
22. Shi, L., Zhang, Y., Cheng, J., Lu, H.: Two-stream adaptive graph convolutional networks for skeleton-based action recognition. In: Proceedings of the IEEE/CVF conference on computer vision and pattern recognition, pp. 12026-12035 (2019)
23. Liu, Z., Zhang, H., Chen, Z., Wang, Z., Ouyang, W.: Disentangling and unifying graph convolutions for skeleton-based action recognition. In: Proceedings of the IEEE/CVF conference on computer vision and pattern recognition, pp. 143-152 (2020)