https://doi.org/10.5607/en24008
Exp Neurobiol. 2024 Jun;33(3):119-128.
pISSN 1226-2560 • eISSN 2093-8144

Original Article

Experimental Neurobiology

Egocentric 3D Skeleton Learning in a Deep Neural Network Encodes Obese-like Motion Representations

Jea Kwon[1†], Moonsun Sa[1†], Hyewon Kim[1,2†], Yejin Seong[1] and C. Justin Lee[1*]
[1]Center for Cognition and Sociality, Institute for Basic Science (IBS), Daejeon 34126,
[2]Department of Pre-Medicine, Eulji University School of Medicine, Daejeon 34824, Korea

Obesity is a growing health concern, mainly caused by poor dietary habits. Yet, accurately tracking the diet and food intake of individuals with obesity is challenging. Although 3D motion capture technology is becoming increasingly important in healthcare, its potential for detecting early signs of obesity has not been fully explored. In this research, we used a deep LSTM network trained with individual identity (identity-trained deep LSTM network) to analyze 3D time-series skeleton data from mouse models with diet-induced obesity. First, we analyzed the data from two different viewpoints: allocentric and egocentric. Second, we trained various deep recurrent networks (e.g., RNN, GRU, LSTM) to predict the identity. Lastly, we tested whether these models effectively encode obese-like motion representations by training a support vector classifier with the latent features from the last layer. Our experimental results indicate that the optimal performance is achieved when utilizing an identity-trained deep LSTM network in conjunction with an egocentric viewpoint. This approach suggests a new way to use deep learning to spot health risks in mouse models of obesity and should be useful for detecting early signs of obesity in humans.

Key words: Obesity, Egocentric, 3D skeleton, Identity, LSTM, Behavior

INTRODUCTION

Recent advancements in deep learning for motion capture systems show a promise in medical healthcare by accurately capturing 3D skeleton data over time [1-4]. These data are valuable for early diagnosis and remote monitoring, particularly for conditions related to motor movements, such as Parkinson's disease [5-7]. However, their potential for understanding health conditions not directly related to motor movements such as obesity is yet to be evaluated.

Obesity, a prevalent lifestyle disease, is primarily linked to poor dietary habits [8, 9]. While the transition from an unhealthy diet to an obesity diagnosis can take considerable time, the traditional Body Mass Index (BMI) system, which bases obesity diagnosis on weight and height, falls short in accuracy by failing to differentiate between muscle mass and body fat [10]. Predicting the risk of diet-induced obesity (DIO) through early behavioral pattern observation presents a promising avenue for healthcare systems. While 3D skeleton data provide detailed spatiotemporal information on dynamic movements, the complexity of data structure makes it challenging to extract meaningful insights, such as disease traits [11, 12].

This study aims to detect early signs of obesity-like motion representations using time-series 3D skeleton data from DIO mouse models. Since diet tracking on a daily basis in clinical settings is infeasible, we explore the potential of deep recurrent networks (DRNs) to capture obese-like motion representations without dietary information. Inspired by previous works on deep convolutional networks (DCNs) [13, 14], we leverage DRNs to predict the identity of DIO models based on their motion data. Moreover, we applied the concept of object-centered view (i.e., allocentric)

Submitted April 16, 2024, Revised May 28, 2024,
Accepted June 4, 2024

*To whom correspondence should be addressed.
TEL: 82-42-878-9150, FAX: 82-42-878-9151
e-mail: cjl@ibs.re.kr
†These authors contributed equally to this article.

or self-centered view (i.e., egocentric) [15-17] on this dataset. We show that combining the egocentric viewpoint with deep LSTM networks trained to identify individual identities, identity-trained networks (note identity-training differs from the concept of identity-aware learning [18, 19].) effectively differentiates movement patterns between mice on a standard chow diet and those on a high-fat diet (HFD). This approach offers a promising way to identify obesity-related motion characteristics without invasive or continuous dietary monitoring.

MATERIALS AND METHODS

Animals housing and induction of diet-induced obesity (DIO) mouse model

All experiments involving animals in this study were conducted following the guidelines approved by the Institutional Animal Care and Use Committee of both the Institute for Basic Science (IBS) in Daejeon, South Korea. The mice used in this research were housed in a facility ensuring a specific pathogen-free environment. The housing conditions for these mice included a consistent 12-hour cycle of light and dark, with lights turning on at 8:00 AM. The environment was maintained at a temperature of 21°C and a humidity level between 40% and 60%. Mice had unrestricted access to both water and food throughout the study.

To establish the DIO mouse model, male C57BL/6J mice, starting at six weeks of age, were fed either a high-fat diet (HFD; consisting of 60% of calories from fat, sourced from D12492, Research Diets Inc.) or a standard chow diet (Teklad, 2018S, Envigo) for a period of 6 to 18 weeks, following protocols previously established (Fig. 1, 2a) [20-23]. All experimental procedures were carried out using age-matched control groups to ensure accuracy and reliability. These mice were subjected to weekly measurements at the same time every week, including body weight measurements, and AVATAR experiments were conducted.

Time-series data

For the acquisition of 3D obesity skeleton dataset, we used AVATAR system [24], a YOLO-based 3D pose estimation with multi-view images that extracts $D^1 \times V^2$ time series data from multiple joint movements from freely moving mice. Both the chow group (12 mice) and the HFD group (12 mice) are subjected to weekly

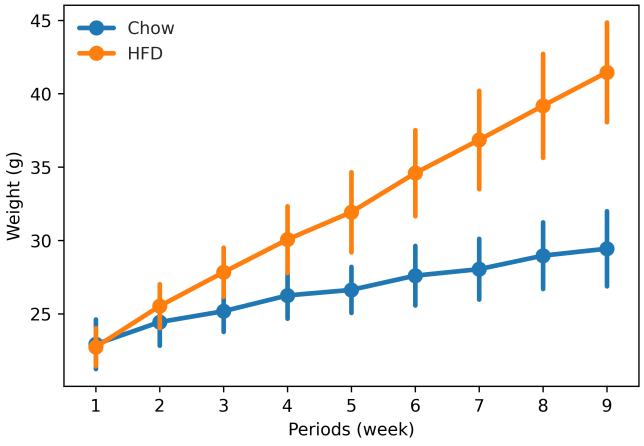

Fig. 1. Weekly weight changes in different animal groups. Comparison of weight between 12 mice in the chow diet (blue) and 12 mice in the HFD diet (orange) for duration of 9 weeks.

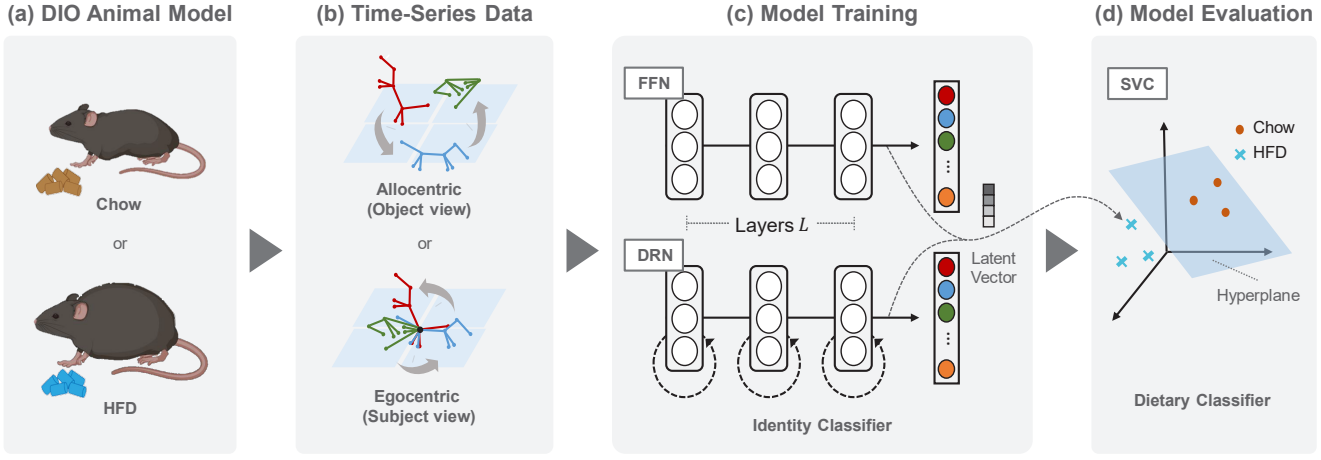

Fig. 2. Schematic illustration of our experimental design. (a) Mouse model: chow diet (top) and HFD (bottom). (b) 3D skeleton data, collected from the AVATAR system, are processed into different viewpoints: allocentric (top) or egocentric (bottom) data. (c) FFN (top) and DRNs (bottom) were trained to perform identity classification tasks. (d) The latent vectors from the last hidden layer were used to test the separability of diet types (chow vs. HFD) with a linear SVC.

https://doi.org/10.5607/en24008

measurements at the same time for 9 weeks, including body weight measurements, and recorded for 10 minutes (20 frames per second; 12,000 frames per session).

For generation of allocentric dataset (object view), the skeletons were adjusted on spatiotemporal centroid offset. For generation of egocentric dataset (subject view), the skeletons were adjusted on anus node offset (Fig. 2b). Each trace was randomly split with a chunk size of T. Then the dataset was randomly grouped into *train*, *valid*, and *test* datasets with 8:1:1 ratio, with different sequence lengths[3].

Model training

To address time series 3D skeleton data, we have explored the potential of DRNs (Fig. 2c): we compared feed-forward network (FFN), recurrent neural network (RNN) [25], gated recurrent unit (GRU) [26] and long-short term memory (LSTM) [27]. With batch size of N the inputs for FFN are given by 2D tensor (N,T×D×V), and for DRNs are given by 3D tensor (N,T,D×V). Both FFN and DRNs consist of L=[1,2,3] number of hidden layers with 256 units per layer and one fully connected linear layer for identity classification. ReLU was employed as the activation function, and AdamW served as the optimizer, with cross-entropy loss utilized as the loss function. The objective of these models is to predict the accurate identity (finding the source animal) from provided 3D skeleton sequences. During this phase, *train* and *valid* datasets were used.

Model evaluation

After training models, the last layer hidden activations (or latent vectors) were extracted to detect obese-like motion representations, similar to previous approach using support vector classifier (SVC) (Fig. 2d) [14]. The objective of SVC is to find a linear hyperplane that discriminates the diet group (chow vs HFD). During this phase, *valid* dataset was used to train and *test* dataset was used to evaluate.

RESULTS

Obese-like motion representations emerge with identity classification

One interesting property of deep neural networks is their ability to capture features that are seemingly unrelated to the original task at hand during the process of learning from large datasets. For example, when trained to recognize the identity of individuals from their photos, the DCN could also develop representations related to facial expressions [14]. Similarly, when trained with natural sounds, the DCN could also detect rudimentary music [13]. In clinical settings, tracking the precise composition of subjects' diets is considerably more challenging than tracking their individual identity. Consequently, utilizing individual identity offers practical advantage to deep neural networks. Therefore, we examined whether identity-trained DRNs trained with 3D skeleton datasets can effectively capture seemingly unrelated dietary information using SVC.

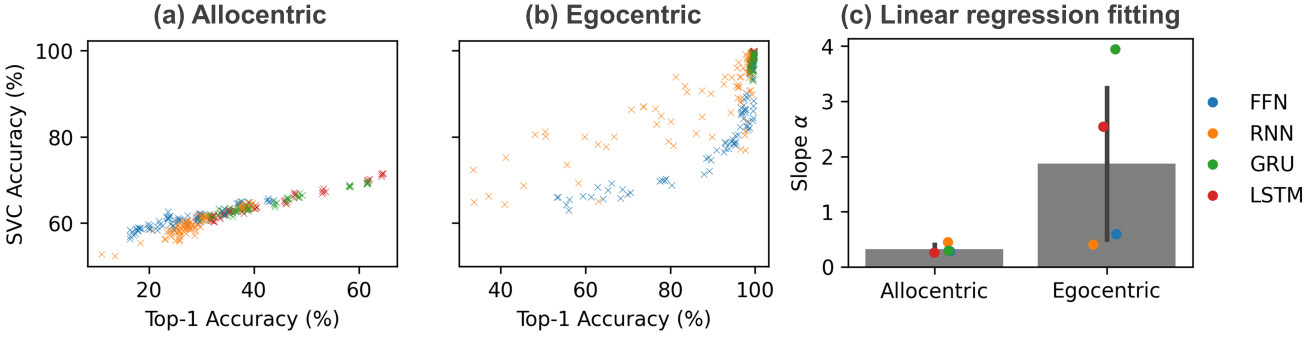

Fig. 3. Effect of allocentric and egocentric viewpoint differences. (a, b) Scatter plots of top-1 identity classification accuracy trained with (a) allocentric or (b) egocentric 3D skeletons and their diet group prediction accuracy from linear SVC; Each color represent FFN (blue), RNN (orange), GRU (green) or LSTM (red); Each dots represent a different number of hidden layers, sequences, and trials (Total 75=3×5×5). (c) Summary bar graphs for the slopes from linear regression (y=αx+β) on scatter plots for each network.

[1] *D*: Number of dimensions (x, y and z). Here, D=3
[2] *V*: Number of joints (head, limbs, tail and etc). Here, V=9
[3] T: Number of frames (consecutive number of skeletons). Here, *T*=[10,20,30,40,50]

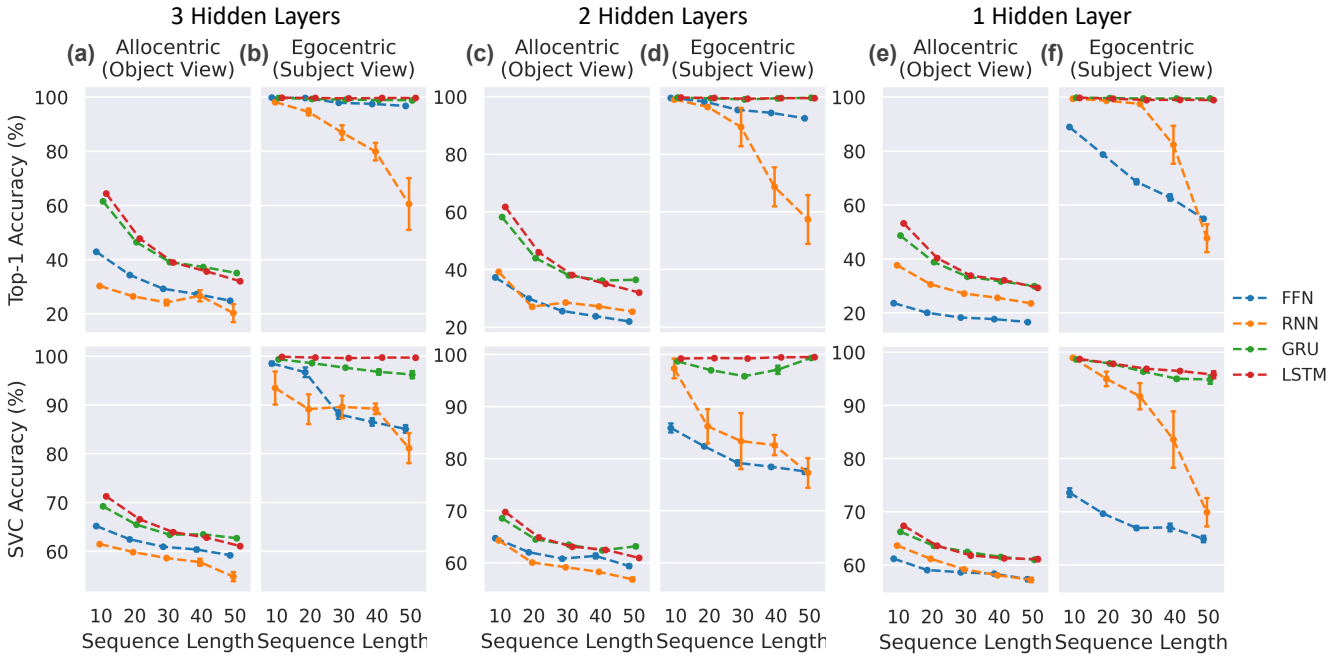

Fig. 4. Evaluation of various architectures with hidden layers. Top-1 accuracy of identity classification task (top) and SVC accuracy of dietary classification task (bottom) in allocentric (a) or egocentric (b) viewpoints with three hidden layers, two hidden layers (c, d), one hidden layer (e, f). Error bar, standard deviation of 5 trials.

We found that deep neural networks, when models were trained to the task of identity classification, significantly enhance the ability to differentiate between two dietary groups, chow vs HFD (Fig. 3). The positive correlations (Pearson's correlation analysis, Top-1 identity classification accuracy vs SVC diet classification accuracy; allocentric, r=0.94, p<0.001; egocentric, r=0.78, p<0.001) imply that the shared latent feature vectors (Fig. 2c, d) encode both identity and dietary information necessary for differentiating data. Moreover, egocentric viewpoint of 3D skeletons consistently outperformed allocentric viewpoint (Fig. 3a, b), suggesting that egocentric viewpoint is more beneficial in capturing underlying data structure of 3D skeletons. Our finding is consistent with previous studies, supporting the idea that egocentric motion representations can contain effective information for understanding and classifying complex behaviors and traits [28].

Next, we sought to explore the impact of viewpoint difference on the effectiveness of DRNs in time series data analysis. Through linear regression analysis, we demonstrated that allocentric data consistently exhibited strong linear relationships regardless of the model architecture used, with high goodness of fit values across different architectures (FFN, r^2=0.85; RNN, r^2=0.84; GRU, r^2=0.96; LSTM, r^2=0.97). In contrast, egocentric data generally showed a decreased effectiveness (FFN, r^2=0.70; RNN, r^2=0.65; GRU, r^2=0.22; LSTM, r^2=0.74), with a significantly larger variance in the slope values across models (Fig. 3c). These findings suggest that the choice of model architecture plays a more pronounced role in dietary feature emergence when learned from egocentric 3D motion data. In summary, combining egocentric 3D skeleton data with modern DRNs can be effective in both identity and dietary classification, highlighting the delicate interplay between the viewpoint and model architecture in analyzing time series data.

Identity-trained deep LSTM encodes obese-like motion representations

While we found that an egocentric viewpoint plays a significant role in capturing shared features of both identity and dietary information with a combination of DRNs, it remains unclear how architectural difference affects the dietary classification. To investigate this, we have compared the dietary classification accuracy across models with different numbers of hidden layers and sequence lengths (Fig. 4).

The most significant impact on identity accuracy was observed to be due to differences in the viewpoint (Fig. 4). Regarding sequence length, we noted a general trend where increasing lengths tended to decrease accuracy for both identity and dietary classifications. However, deep LSTM networks demonstrated a remarkable resilience to this performance degradation compared to other networks (Fig. 4). Furthermore, this resilience was most pronounced with three hidden layers (Fig. 4a). Surprisingly, this identity-trained LSTM was more effective compared to end-to-

https://doi.org/10.5607/en24008

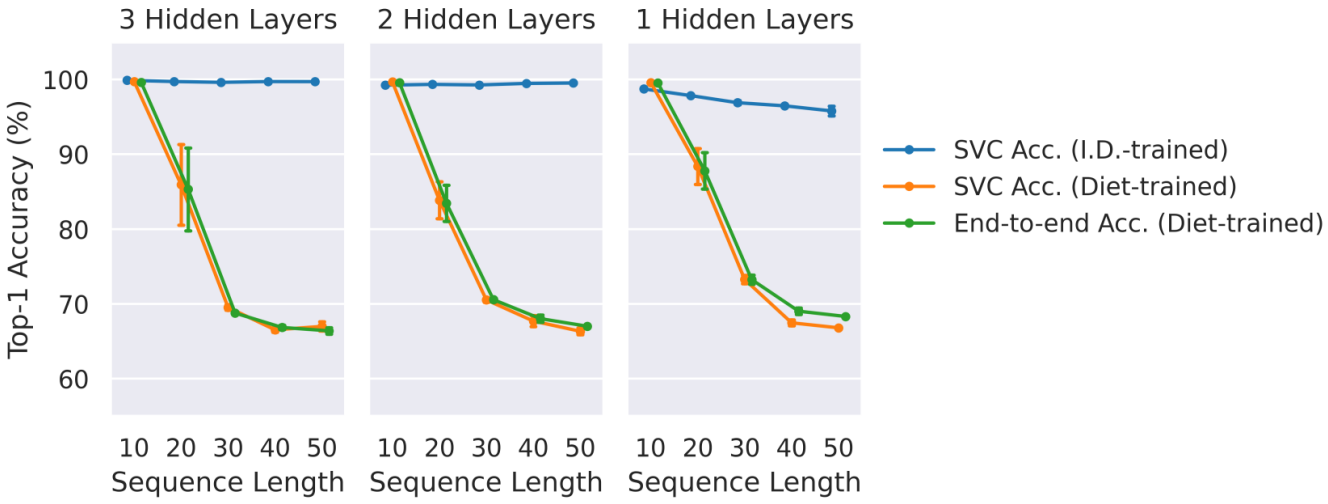

Fig. 5. Comparison with diet-aware LSTM. Prediction of diet using LSTM with varying numbers of hidden layers (3, 2, 1) trained on either identity (I.D.) or diet. As shown in the figure, SVC accuracy of I.D.-trained LSTM outperforms the SVC accuracy of diet-trained LSTM or end-to-end test accuracy.

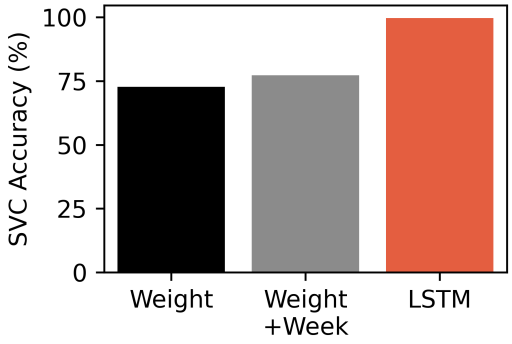

Fig. 6. LSTM effectively distinguish different dietary groups with linear SVC. Identity trained LSTM with 3 hidden layers (red) outperforms dietary group prediction when using only weight information (black) or weight and period information (grey).

end diet-trained LSTM (Fig. 5). These results collectively suggest that the memory cells in deep LSTM networks trained with identity may play a key role in capturing the underlying data structure which is beneficial for accurately predicting both identity and dietary habits.

Another intriguing observation is that the performance of the identity-trained LSTM, utilizing animals' 3D behavioral data, was more accurate than when using the animals' weight and period information (Fig. 1, 6). This implies that in the progression of diet-induced obesity, behavioral changes may precede weight changes.

Increased linear separability in the latent feature embedding space

In our proposed two-step method, an identity-trained LSTM followed by a diet-trained linear SVC demonstrates superior performance when the dataset is represented in an egocentric skeleton format. To elucidate how linear SVC effectively captures the diet feature representation extracted from the latent vectors of deep models, we have visualized the linear separability in the latent feature embedding space (Fig. 7).

For this visualization, we first applied linear discriminant analysis (LDA) to identify an axis (LD1) that maximizes the separation between two classes. LDA achieves this by maximizing the ratio of the between-class variance to the within-class variance. After determining the LD1 axis, we then utilized a modified principal component analysis (PCA) technique to identify principal components within the subspace orthogonal to LD1. This modification involved projecting the data onto a plane orthogonal to LD1 before applying PCA, thereby ensuring that the variance maximized by the first principal component is orthogonal to LD1.

While there are notable differences in how LDA and SVC approach the task of finding a linear hyperplane between classes (LDA assumes a normal distribution of data within each class, whereas SVC does not), combining these linear transformations significantly enhances our understanding of the feature distribution and relationships within the latent space. This method not only facilitates clearer visualizations but also provides deeper insights into how the model distinguishes between different classes, thus improving our understanding of the model's effectiveness and behavior.

As demonstrated in Fig. 7, the egocentric representation notably improves linear separability between classes across various model architectures. This enhancement is primarily due to the shift from allocentric to egocentric skeletons, characterized by the absence of global geometric information. We believe this key factor signifi-

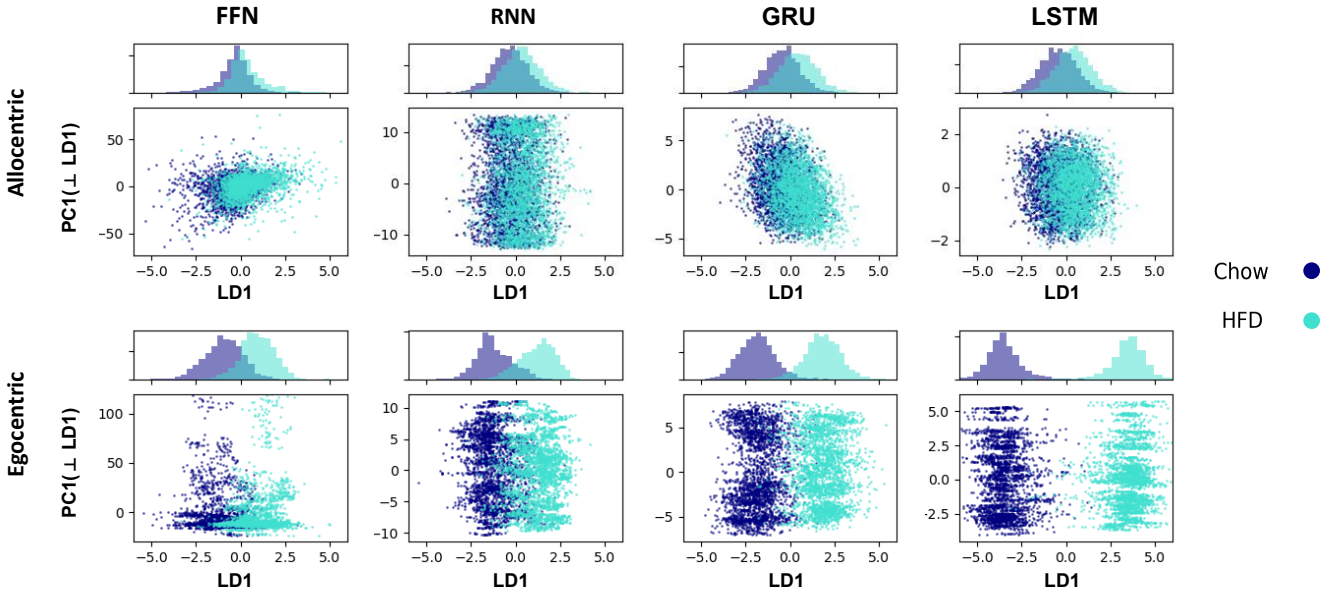

Fig. 7. Data Separability by models and data types. Each dot represents Chow (navy) and HFD (turquoise) data points 2D visualized by LD1 as x-axis, and PC1 as y-axis, and histogram above shows data point density along LD1 component. Eight pairs of two plots are shown according to four types of models and two types of data, Allocentric (top) and Egocentric (bottom).

cantly boosts linear separability in the latent feature embedding space by allowing the model to focus more on motion differences.

Furthermore, we evaluated the identity-trained LSTM to determine its effectiveness in classifying motion changes induced by diet. Although the results are intriguing, the mechanism by which identity-specific motion training enhances the predictive capabilities of a linear SVC remains unclear. Linear SVC aims to find a linear hyperplane that maximizes the separation between two different diet groups. *If the identity-specific motion features are well-represented in the latent feature embedding space, it simplifies the task for the SVC to identify and apply the linear hyperplane that distinctly separates these classes*. We substantiated this hypothesis by analyzing the data distribution along the LD1 axis across various model architectures, as shown in Fig. 7. These findings suggest that the unique architecture of the LSTM, characterized by recurrent connectivity and memory cells (Fig. 8, vs FFN), presumably enables effective aggregation of identity-specific motion features within the latent embedding space.

Assessing the impact of body weight and size on motion recognition performance in diet-induced obesity models

During the progression of diet-induced obesity, both weight and size can vary over time between the chow and high-fat diet (HFD) groups. To ensure that our identity-trained deep LSTM accurately learns motion differences rather than size or weight differences, we divided the original datasets into two sub-datasets: Similar Weight (SW) and Different Weight (DW), as shown in Fig. 9a. In the SW dataset, both forebody and hindbody skeleton lengths were comparable, while in the DW dataset, these measurements were significantly different, as depicted in Fig. 9b. We then cross-validated our method under four different conditions of 'LSTM-trained' & 'SVC-predicted' pairs: SW & SW, DW & DW, DW & SW, and SW & DW (Fig. 9c). In allocentric conditions, the SVC prediction performance significantly declined in cross-condition pairs (i.e., SW & DW, DW & SW), whereas in egocentric conditions, the performance was much more stable, indicating robustness to variations in body size and weight. This result suggests that the latent features encoded by the identity-trained deep LSTM are derived from motion differences, not from differences in size or weight.

DISCUSSION

In this study, our focus centered on the detection and prediction of obesity, a growing concern in modern society. Employing novel methods, we aimed to uncover motion features indicative of dietary information through the process of identity classification. By developing a 3D skeleton identity classification network for both chow and HFD models, we extracted latent vectors and utilized an SVC to evaluate the representation of dietary features within these vectors. Our findings underscore the potential of identity-trained deep LSTM networks in identifying obese-like features from time-series 3D skeleton data, shedding light on the previously elusive

https://doi.org/10.5607/en24008

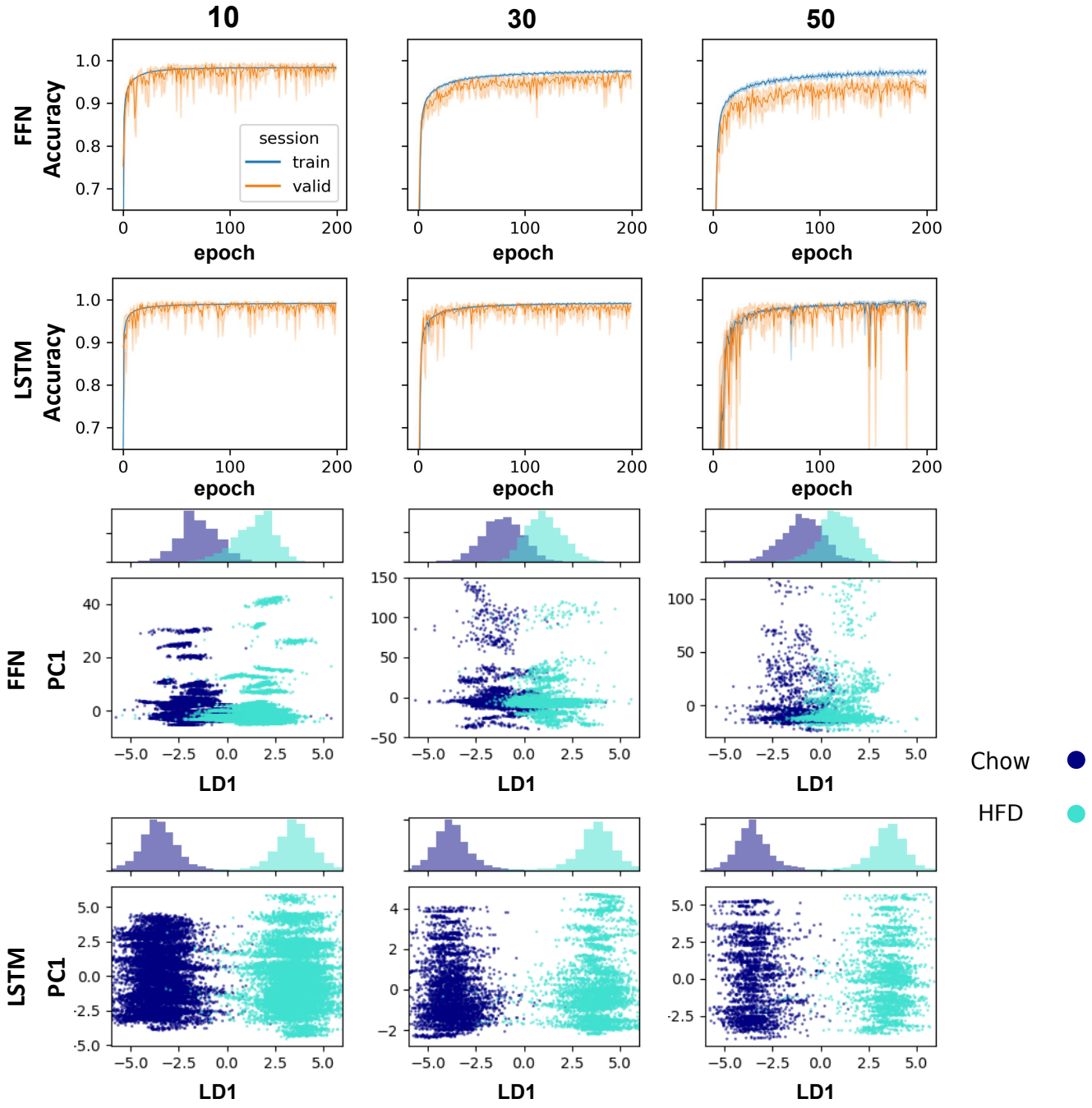

Fig. 8. Data separability and accuracy by epoch for FFN and LSTM across various sequence lengths. The training accuracy (blue) and validation accuracy (orange) of the identity-trained model are depicted for each epoch (top). The data points are visualized in a 2-dimensional image (bottom) using LD1 from LDA and PC1 from PCA. In the LDA visualization, data points are categorized into two groups: Chow (navy) and HFD (turquoise).

association between dietary habits and skeletal movement. Our research also marks the first demonstration that DRNs can capture seemingly unrelated features from original tasks, extending previous DCN studies. This indicates a broader capability of deep learning models. We foresee these models transforming clinical practice, fostering deep-learning solutions for obesity prevention and enhancing healthcare outcomes, thus promoting societal well-being.

Despite the remarkable performances demonstrated by deep LSTM networks in both identity and dietary classification tasks, our study acknowledges some limitations. Firstly, the absence of cross-subject validation. While our results showcase the LSTM's

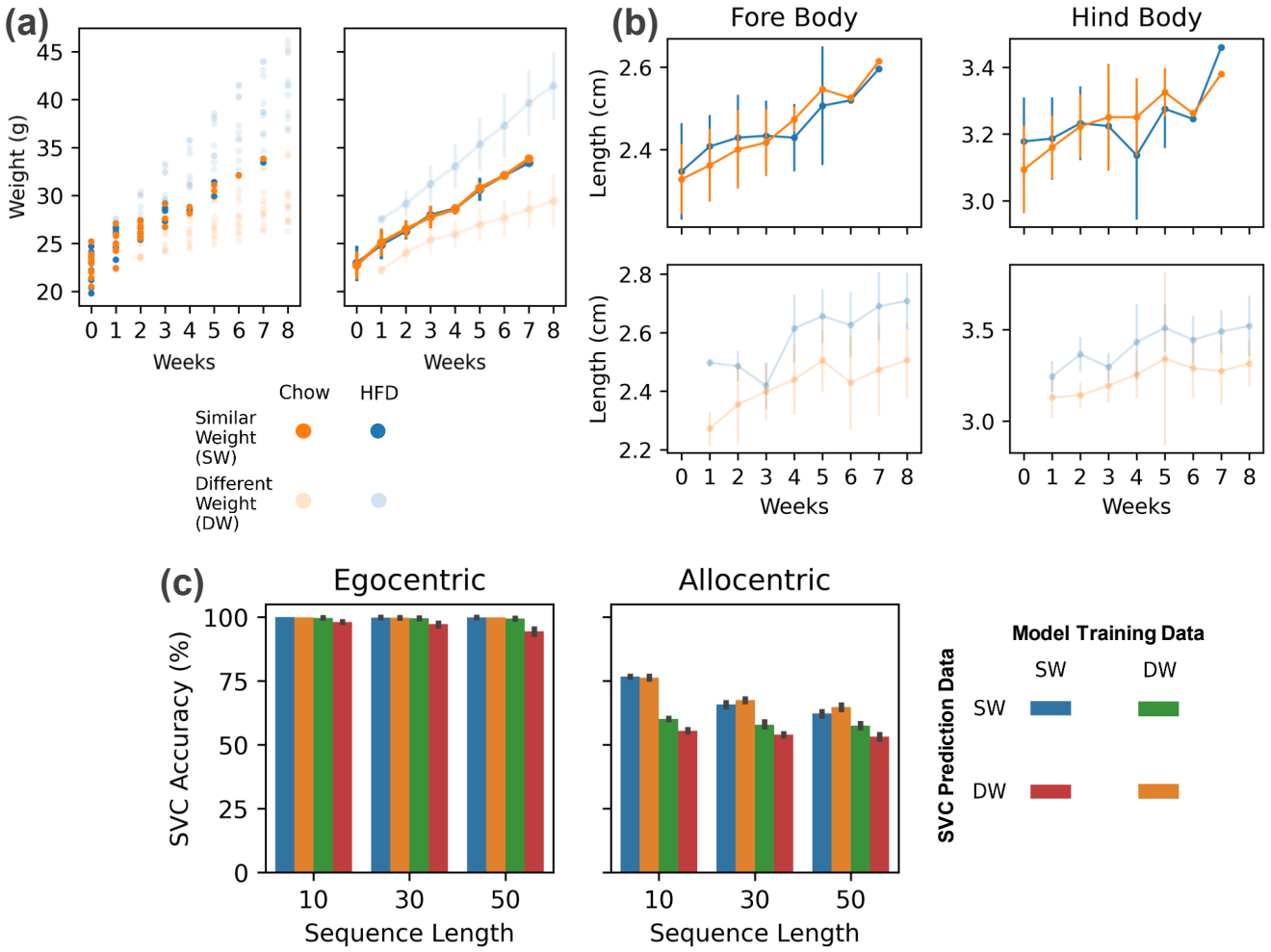

Fig. 9. Data groups classified by weight show accuracy differences depending on the data type combinations. These groups are categorized by significant weight differences and are indicated with varying transparency (a). Their skeleton lengths are also compared between HFD vs Chow group depending on the presence of weight differences (b). Consequently, different accuracies were observed depending on which data were used for training and which data were put into the SVC (c).

proficiency, questions persist regarding its reliance on memorization versus genuine generalization capabilities. Additionally, the identity-trained diet prediction model unexpectedly outperforms the end-to-end model, a surprising deviation from typical expectations that highlights the need for further research. Lastly, our dataset and task formulation have not yet been compared across various structures such as graph-based, convolutional, or transformer-based networks research endeavors should prioritize investigating the LSTM's ability to generalize across subjects, or the potential of different model architectures, thus providing further validation of its effectiveness in discerning dietary-related motion patterns beyond the confines of our study's experimental scope.

Linear separability in the latent feature embedding space

In this study, we explore the combination of an identity-trained LSTM with a Linear SVC to effectively distinguish differences in diet-induced egocentric motion. One key observation is that the linear separability of the latent feature embeddings, extracted from the last layer of the DRNs, is significantly influenced by the model architecture and the type of skeleton representation used (Fig. 7). The LSTM's distinctive recurrent connectivity and memory cells contribute to a consistent performance improvement across varying sequence lengths (Fig. 8). These results underscore the critical role of data representation and model architecture selection in achieving robust performance.

Translational insights to human clinical applications

Controlling genetic diversity, living environments, and dietary habits in clinical human studies is challenging, and within this context, simultaneously tracking 3D motion and dietary data to

https://doi.org/10.5607/en24008

study diet-induced obesity's effects adds a significant layer of complexity. In contrast, animal models enable the acquisition of tightly controlled datasets for studying long-term obesity progression. In the context of human clinical settings, obtaining personal identity information is far simpler than accurately documenting dietary composition. An identity-trained deep LSTM model can be a potential method for tracing back to the dietary patterns of patients who suffer from nutritional issues but are not necessarily obese, such as those with diabetes or fatty liver disease, etc. This enables the redesign of future dietary and behavioral strategies. However, this approach requires a significant amount of human data to set criteria in humans based on results obtained from mice. Furthermore, if advancements were made in related wearable technology capable of monitoring human joint movements, it could apply to everyone to monitor daily dietary patterns and health. These aspects underscore the capacity of the identity-trained deep LSTM model to accurately identify dietary patterns without the need to explicitly categorize diet group classifications, laying a crucial groundwork for translating these findings into human clinical practices.

ACKNOWLEDGEMENTS

This work was supported by the Institute for Basic Science (IBS), Center for Cognition and Sociality (IBS-R001-D2). Graphical abstract and Fig. 2a were created via Biorender.com.

REFERENCES

1. Redmon J, Divvala S, Girshick R, Farhadi A (2016) You only look once: unified, real-time object detection. In: 2016 IEEE conference on computer vision and pattern recognition (CVPR), pp 779-788. IEEE, Los Alamitos, CA.
2. Lee LH, Braud T, Zhou P, Wang L, Xu D, Lin Z, Kumar A, Bermejo C, Hui P (2021) All one needs to know about metaverse: a complete survey on technological singularity, virtual ecosystem, and research agenda. ArXiv. doi: 10.48550/arXiv.2110.05352.
3. Wang J, Tan S, Zhen X, Xu S, Zheng F, He Z, Shao L (2021) Deep 3D human pose estimation: a review. Comput Vis Image Underst 210:103225.
4. Lam WWT, Tang YM, Fong KNK (2023) A systematic review of the applications of markerless motion capture (MMC) technology for clinical measurement in rehabilitation. J Neuroeng Rehabil 20:57.
5. Monje MHG, Domínguez S, Vera-Olmos J, Antonini A, Mestre TA, Malpica N, Sánchez-Ferro Á (2021) Remote evaluation of Parkinson's disease using a conventional webcam and artificial intelligence. Front Neurol 12:742654.
6. Delrobaei M, Memar S, Pieterman M, Stratton TW, McIsaac K, Jog M (2018) Towards remote monitoring of Parkinson's disease tremor using wearable motion capture systems. J Neurol Sci 384:38-45.
7. Tian H, Li H, Jiang W, Ma X, Li X, Wu H, Li Y (2024) Cross-spatiotemporal graph convolution networks for skeleton-based parkinsonian gait MDS-UPDRS score estimation. IEEE Trans Neural Syst Rehabil Eng 32:412-421.
8. Liberali R, Kupek E, Assis MAA (2020) Dietary patterns and childhood obesity risk: a systematic review. Child Obes 16:70-85.
9. San-Cristobal R, Navas-Carretero S, Martínez-González MÁ, Ordovas JM, Martínez JA (2020) Contribution of macronutrients to obesity: implications for precision nutrition. Nat Rev Endocrinol 16:305-320.
10. Romero-Corral A, Somers VK, Sierra-Johnson J, Thomas RJ, Collazo-Clavell ML, Korinek J, Allison TG, Batsis JA, Sert-Kuniyoshi FH, Lopez-Jimenez F (2008) Accuracy of body mass index in diagnosing obesity in the adult general population. Int J Obes (Lond) 32:959-966.
11. Kwon J, Kim S, Kim DK, Joo J, Kim S, Cha M, Lee CJ (2023) Subtle: an unsupervised platform with temporal link embedding that maps animal behavior. bioRxiv. doi: 10.1101/2023.04.12.536531.
12. Su Y, Lin G, Sun R, Hao Y, Wu Q (2021) Modeling the uncertainty for self-supervised 3D skeleton action representation learning. In: MM '21: proceedings of the 29th ACM international conference on multimedia, pp 769-778. Association for Computing Machinery, New York, NY.
13. Kim G, Kim DK, Jeong H (2024) Spontaneous emergence of rudimentary music detectors in deep neural networks. Nat Commun 15:148.
14. Zhou L, Yang A, Meng M, Zhou K (2022) Emerged human-like facial expression representation in a deep convolutional neural network. Sci Adv 8:eabj4383.
15. Islam MS, Algosaibi A, Rafaqat W, Bakhat K (2023) Employing FGP-3D, a fully gated and anchored methodology, to identify skeleton-based action recognition. Appl Sci 13:5437.
16. Caetano C, Brémond F, Schwartz WR (2019) Skeleton image representation for 3D action recognition based on tree structure and reference joints. In: 2019 32nd SIBGRAPI conference on graphics, patterns and images (SIBGRAPI), pp 16-23. IEEE, Los Alamitos, CA.
17. Ke Q, Bennamoun M, An S, Sohel F, Boussaid F (2017) A new representation of skeleton sequences for 3D action recogni-

tion. In: 2017 IEEE conference on computer vision and pattern recognition (CVPR), pp 4570-4579. IEEE, Los Alamitos, CA.

18. Du W, Ding S, Guo L, Zhang J, Ding L (2024) Expressive multi-agent communication via identity-aware learning. Proceedings of the AAAI Conference on Artificial Intelligence 38:17354-17361.
19. Liu S, Zhou Y, Song J, Zheng T, Chen K, Zhu T, Feng Z, Song M (2023) Contrastive identity-aware learning for multi-agent value decomposition. Proceedings of the AAAI Conference on Artificial Intelligence 37:11595-11603.
20. Sa M, Yoo ES, Koh W, Park MG, Jang HJ, Yang YR, Bhalla M, Lee JH, Lim J, Won W, Kwon J, Kwon JH, Seong Y, Kim B, An H, Lee SE, Park KD, Suh PG, Sohn JW, Lee CJ (2023) Hypothalamic GABRA5-positive neurons control obesity via astrocytic GABA. Nat Metab 5:1506-1525.
21. Sa M, Lee JM, Park MG, Lim J, Kim JMJ, Koh W, Yoon BE, Lee CJ (2022) Unaltered tonic inhibition in the arcuate nucleus of diet-induced obese mice. Exp Neurobiol 31:147-157.
22. Casimiro I, Stull ND, Tersey SA, Mirmira RG (2021) Phenotypic sexual dimorphism in response to dietary fat manipulation in C57BL/6J mice. J Diabetes Complications 35:107795.
23. Wang CY, Liao JK (2012) A mouse model of diet-induced obesity and insulin resistance. Methods Mol Biol 821:421-433.
24. Kim DG, Shin A, Jeong YC, Park S, Kim D (2022) Avatar: ai vision analysis for three-dimensional action in real-time. bioRxiv. doi: 10.1101/2021.12.31.474634.
25. Bengio Y, Simard P, Frasconi P (1994) Learning long-term dependencies with gradient descent is difficult. IEEE Trans Neural Netw 5:157-166.
26. Cho K, van Merriënboer B, Bahdanau D, Bengio Y (2014) On the properties of neural machine translation: encoder-decoder approaches. arXiv. doi: 10.48550/arXiv.1409.1259.
27. Hochreiter S, Schmidhuber J (1997) Long short-term memory. Neural Comput 9:1735-1780.
28. Dhamanaskar A, Dimiccoli M, Corona E, Pumarola A, Moreno-Noguer F (2023) Enhancing egocentric 3D pose estimation with third person views. Pattern Recognit 138:109358.