Journal of Information Analytics, 1(2), 21-28

Journal of Information Analytics

Submit	25.12.2025
Revision	06.01.2026
Accept	07.01.2026
Publish	08.01.2026

Open Access

Research Article

Comparative Evaluation of Video-Based Repetitive Movement Analysis Using Multiple Methods

Yasin Badur[1,*], *Faruk Serin*[2]

[1] *Web Design and Coding Program, Vocational School of Technical Sciences, Harran University, Şanlıurfa, Türkiye*

[2] *Department of Computer Engineering, Faculty of Engineering, Mersin University, Mersin, Türkiye*

Abstract

This study aims to analyze and compare repetitive human movements using skeleton-based keypoints extracted from videos. Using keypoints detected with the YOLOv11 algorithm, we apply the Discrete Fourier Transform (DFT), Principal Component Analysis (PCA), and Matrix Profile. The DFT provides frequency-domain characterizations of motion, PCA attenuates noise to yield more salient patterns, and the Matrix Profile enables precise detection of recurring motifs in the time series. Empirical findings indicate that Fourier-based analyses better capture global structure and provide more discriminative similarity measures, while the Matrix Profile complements them by detecting repeated motifs and onset/offset boundaries. This study aims to contribute to the development of more accurate and reliable approaches in the field of human motion analysis.

Keywords: Motion Analysis, Fourier Transform, Matrix Profile, Dynamic Time Warping,

Repetitive Movements, Skeleton-Based Features

Citation: Badur, Y., & Serin, F. (2025). Comparative Evaluation of Video-Based Repetitive Movement Analysis Using Multiple Methods, Journal of Information Analytics,1(2), 21-28.

*Corresponding Author: Yasin Badur yasin.badur.nk@gmail.com

A version of this work was presented in abstract form at *The Fibonacci International Conference on Engineering, Technology, and Mathematics*.

1. Introduction

Video-based human pose estimation is an active research topic in computer vision and AI, yet robust generalization remains challenging due to complex spatiotemporal structure, occlusions, viewpoint changes, heterogeneous capture conditions, and limited or imbalanced datasets. Although state-of-the-art models achieve high per-frame keypoint accuracy (e.g., Chen et al., 2023), the detection and comparison of recurring motion patterns and the measurement of cross-video similarity remain underexplored, reducing discriminative power and making reliable evaluation more difficult (e.g., Usman & Zhong, 2022).

We address this gap by modeling pose trajectories as multivariate time series and conducting pattern-based similarity analysis using dynamic time warping (DTW), Fourier analysis, principal component analysis (PCA), and the Matrix Profile to capture temporal regularities beyond visual matching. Our aims are to characterize the strengths and limitations of existing techniques, identify core gaps, and provide practical guidance—covering datasets, metrics, and protocols—for fair and reproducible assessment. The anticipated impact spans sports and rehabilitation (fine-grained performance/impairment analysis), security (recurrent/anomalous behavior detection), and human–machine interaction/AR.

The contributions: (1) we reframe the task of detecting and comparing recurring motions in skeleton-based video by specifying explicit assumptions and evaluation criteria; (2) we present a gap analysis that surfaces limitations of current methods and distills practitioner-oriented design principles; and (3) we propose actionable solution directions and experimental blueprints—covering dataset choice, metrics, and protocols—to enable fair and reproducible comparisons.

2. Proposed Method

2.1. Analyzed Pose-Estimation Algorithms

OpenPose provides open-source, multi-person pose estimation with optional hand/face tracking, though its computational load can limit real-time deployment (Cao et al., 2017). MediaPipe is lightweight and fast, runs in real time on mobile/low-power devices, and is widely used in healthcare, fitness, and AR (Lugaresi et al., 2019). PoseNet targets web/mobile via TensorFlow.js/TFLite and remains robust at low resolution, but is comparatively constrained in multi-person scenes (Papandreou et al., 2018). The YOLO family offers a strong speed–accuracy trade-off for detection and pose, with pose-optimized variants adopted in sports, rehabilitation, and entertainment.

We applied YOLOv11 for per-frame keypoint extraction owing to its real-time throughput and stability, high keypoint localization accuracy, and straightforward cross-platform deployment; it is well suited to repetitive-movement analysis. Our evaluation targets repetitive actions—rope skipping, hands-clapping, and mixed exercise sequences.

2.2. Video Data

Seven videos were selected to span movement types and conditions. Criteria included: (i) repetitive actions (jumping, running, warm-ups, therapy), (ii) sufficient resolution/quality for reliable keypoint detection, and (iii) varied lighting/backgrounds. Preprocessing trimmed relevant segments, standardized frame rates, and normalized formats/resolutions. Each video was processed with YOLOv11 to obtain per-frame keypoints; repetition frequency and accuracy were derived downstream from time-series analyses (Fourier, Matrix Profile, DTW).

2.3. Repetitive-Movement Analysis Methods

Per-frame poses are modeled as multivariate time series over 17 keypoints × (x, y) (34 dimensions). We apply complementary analyses to capture periodicity, alignment, and structural similarity:

- PCA projects the 34-D frame representation onto principal components (typically PC1), denoising and exposing dominant motion patterns.
- Fourier analysis quantifies periodic structure using (i) separate x/y channels and (ii) complex signals $x + iy$ to capture planar dynamics; dominant spectral peaks indicate repetition rates
- DTW aligns sequences under variable execution speeds and yields sequence-level similarity (Sakoe & Chiba, 2003).
- Matrix Profile discovers motifs and local anomalies via nearest-neighbor subsequence distances; we apply it to the PCA-reduced 1D series to recover repetitions and onset/offset boundaries.
- Similarity metrics include Euclidean distance (spatial proximity), cosine similarity (angular consistency of limb/trajectory vectors), and Pearson correlation (linear association for intra-class consistency and inter-class separability).

DTW addresses temporal alignment; Fourier captures periodicity; Matrix Profile localizes motifs and segment boundaries; geometric/angle/correlation metrics assess spatial and directional coherence.

3. Experimental Results

3.1. Time Series Analysis of Keypoint Coordinates

Across Rope Skipping 1–2 (Figure 1), the skeleton keypoint trajectories exhibit a clear and stable periodicity. Cranial keypoints show low-amplitude, near-parallel fluctuations, whereas distal joints (wrists, ankles, knees) display higher-amplitude oscillations that carry most of the discriminative signal. In Rope Skipping 1, brief transients—consistent with crossed-hands execution and occasional occlusion—appear predominantly in the wrist channels; these remain local and do not disrupt the global cycle structure. Rope Skipping 2 presents more regular, phase-locked cycles.

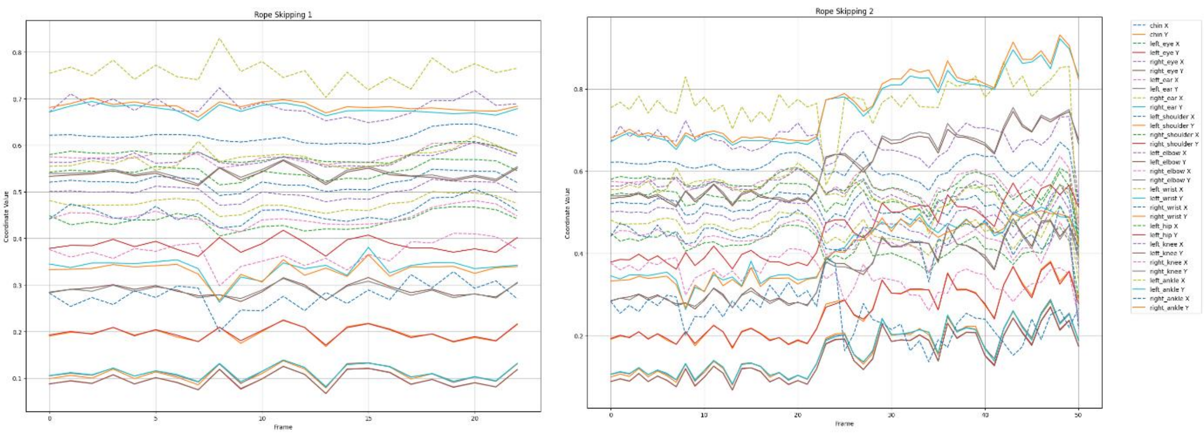

Figure 1. *Keypoint analysis of rope skipping movements*

For "Jumping with hands clapping" (Figure 2), onset/offset phases are demarcated by pronounced plateaus and transition regions. The first clip shows a short initial quiescence followed by highamplitude rhythmic oscillations; the second clip runs at a lower cadence, yielding longer periods and more salient phase transitions. Transient drops in wrist/shoulder signals are consistent with occlusion or temporary out-of-frame motion and do not alter the repetition structure.

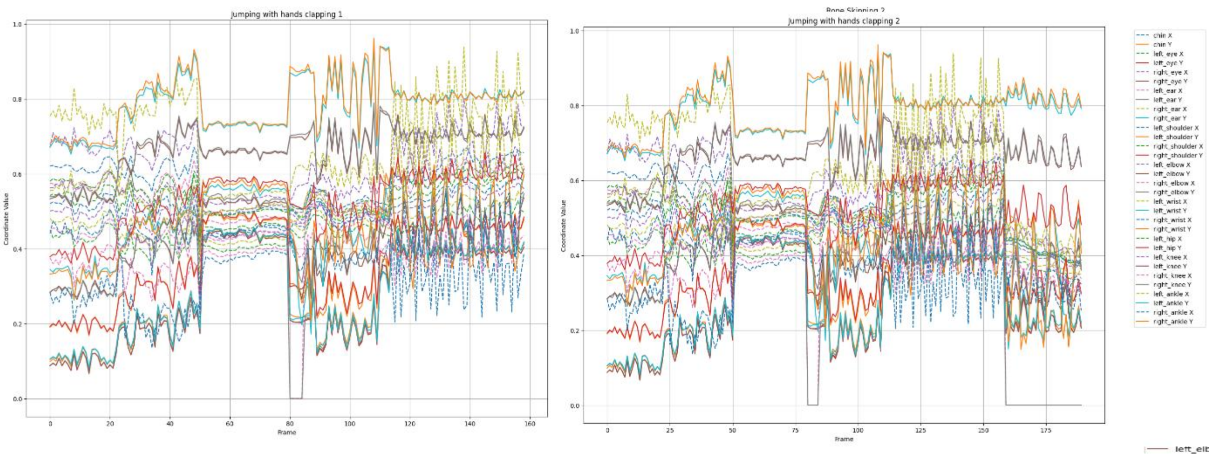

Figure 2*. Keypoint analysis of jumping with hands clapping*

The figures show that (i) head-region channels serve as a stable reference, (ii) distal joints constitute high-variance, information-rich channels, and (iii) despite local variability from execution style and cadence, the global periodic structure is preserved. These findings are consistent with—and reinforce—subsequent PCA, Fourier, and Matrix Profile analyses, indicating that the recordings furnish reliable, comparable time series for cross-video evaluation.

3.2. Analysis of Time Series of All Movement Coordinates Using PCA

We applied PCA to the x–y coordinates of 17 keypoints per motion, extracting a single principal component to obtain a one-dimensional, interpretable time series. Compared with the raw multi-channel traces (Figure 1–2), the PCA projection (Figure 3) makes periodicity, repetition counts, and segment boundaries more salient. Even in the clip containing successive, heterogeneous actions, recurring patterns were clearly exposed. Overall, PCA reduced noise and dimensionality, improved cross-record comparability, and provided a compact, information-dense representation for downstream analyses (Fourier analysis, DTW, Matrix Profile).

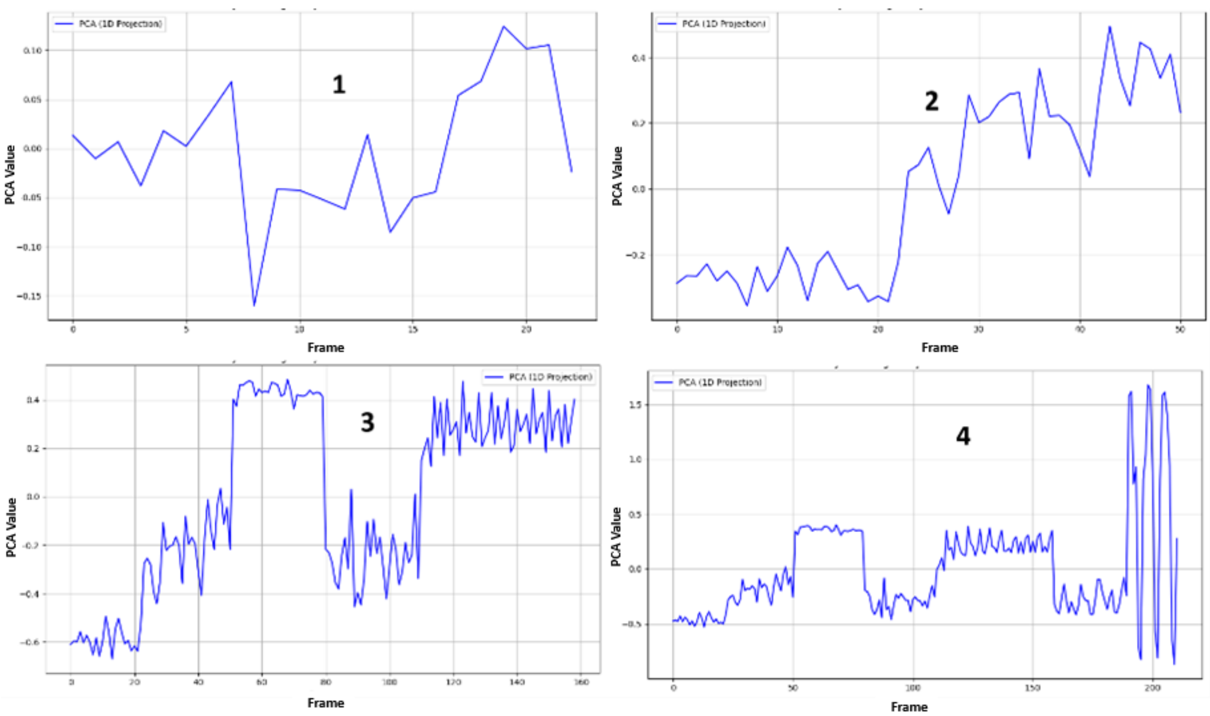

Figure 3. *PCA-based analysis of all movement coordinates (1: Rope Skipping-1, 2: Rope Skipping-2, 3: Jumping with Hands Clapping-1, 4: Mixed Exercises).*

3.3. Pattern Analysis of Keypoints Using Fourier Transform (With and Without PCA)

Across all datasets, we computed Fourier spectra separately for the PCA-reduced 1D series and for the original multivariate series. While Fourier analysis—especially after PCA—provides a stable basis for between-sequence similarity, it is insufficient on its own for accurate repetition counting; accordingly, we use Matrix Profile to recover repetition structure and boundaries, and Fourier (often after PCA) for discriminative inter-sequence similarity. As shown in Table 1, Fourier-based period estimates provide an approximate repetition count..

Table 1. Pattern Frame Count and Approximate Repetition Number of Videos (Fourier)

Video Name	Total Frames	Actual Repetitions	Approx. Period (frames)	Approx. Repetitions (Without PCA / With PCA)
rope skipping-1	380	14	190	2/2
rope skipping-2	620	19	207	3/24
rope skipping-3	283	9	28	10/10
rope skipping-4	305	12	17	18/18
jumping with hands clapping-1	276	9	276	1/9
jumping with hands clapping-2	208	4	26	2/4
different exercises	327	3	47	7/3

For pattern similarity, Table 2 (without PCA) indicates generally high cosine similarity—reflecting directional proximity—alongside weak Pearson correlations and middling Euclidean/DTW distances. In contrast, Table 3 (with PCA) shows marked reductions in Euclidean and DTW distances and more coherent association patterns: similar pairs exhibit high positive correlation, whereas out-of-phase pairs show strong negative correlation. Notably, "Jumping with Hands Clapping 1–2" and "Rope Skipping 1–

4” emerge as strongly similar (low distances, high correlation), while “Rope Skipping 2–3” displays a pronounced negative correlation indicative of phase/misalignment effects.

Table 2. Pattern Similarity Ratios of Keypoint Averages of Repetitive Movements in Videos Obtained by Fourier Transform (Without PCA)

Motion Pair	Euclidean Distance	Pearson Correlation	Cosine Similarity	DTW Distance
rope skipping 1 - 2	20.17	0	0.88	9.65
rope skipping 1 - 3	9.35	-0.01	0.82	4.71
rope skipping 1 - 4	5.80	-0.04	0.88	3.03
rope skipping 2 - 3	8.76	-0.01	0.84	4.68
rope skipping 2 - 4	5.18	-0.02	0.90	2.75
rope skipping 3 - 4	6.40	0.02	0.85	3.38
hands clapping 1 - 2	7.47	-0.06	0.89	4.32
hands clapping 1 - different exercises	8.46	-0.01	0.93	4.81
rope skipping 2 - different exercises	9.28	-0.01	0.90	4.64

Table 2 shows that, in the analyses conducted without PCA, cosine similarity scores were generally high, whereas Pearson correlation remained low; nevertheless, the motions were found to be similar in terms of directionality.

Table 3. Pattern Similarity Ratios of Keypoint Averages of Repetitive Movements in Videos Obtained by Fourier Transform with PCA

Motion Pair	Euclidean Distance	Pearson Correlation	Cosine Similarity	DTW Distance
rope skipping 1 - 2	2.12	-0.63	-0.63	1.64
rope skipping 1 - 3	2.07	0.54	0.54	0.96
rope skipping 1 - 4	1.00	0.83	0.82	0.86
rope skipping 2 - 3	2.65	-0.98	-0.98	2.14
rope skipping 2 - 4	0,8	-0.33	-0.33	0.47
rope skipping 3 - 4	1.70	0.31	0.31	1.21
hands clapping 1 - 2	0.58	0.84	0.84	0.29
hands clapping 1 - different exercises	3.35	0.35	0.35	2.98
rope skipping 2 - different exercises	3.36	-0.40	-0.39	3.04

In sum, Fourier analysis effectively reveals the periodic nature of the motions and—especially when combined with PCA—provides a reliable and consistent basis for similarity assessment. However, it is not sufficient on its own for accurate repetition counting; complementary approaches such as Matrix Profile are recommended for that purpose.

3.4. Keypoint Pattern Analysis with Stumpy (Matrix Profile) and PCA

As summarized in Table 4, Matrix Profile was applied to the PCA-reduced (1D) keypoint time series to detect motifs, onset/offset boundaries, and repetition structure. Matrix Profile was applied to the PCA-reduced (1D) keypoint time series to detect motifs, segment boundaries (onset/offset), and repetition structure. Across all datasets, the method reliably exposed periodicity and localized repeating segments, while also differentiating heterogeneous motion sequences. Similar pairs exhibited low DTW distances and high positive correlation, whereas out-of-phase pairs showed strong negative correlation indicative of phase effects. Euclidean distance tracked structural proximity, and cosine similarity remained consistent with correlation under z-scoring.

Table 4. Keypoint Pattern Analysis with Stumpy (Matrix Profile) and PCA

Video Pairs	Best Pattern Window Size 1	Estimated Repetition Count 1	Best Pattern Window Size 2	Estimated Repetition Count 2	Actual Repetition Count	DTW Distance	Correlation	Euclidean Distance	Cosine Similarity
rope skipping 1-2	20	18.33	20	12.40	4	0.36	-0.63	0.99	0.35
rope skipping 1-3	20	18.33	20	2.42	9	0.38	-0.90	0.99	0.38
rope skipping 1-4	20	18.33	20	5.87	9	0.36	0.01	0.99	0.29
rope skipping 2-3	20	12.40	20	2.42	2	0.09	0.56	0.99	0.07
rope skipping 2-4	20	12.40	20	5.87	9	0.09	0.69	0.99	0.07
rope skipping 3-4	20	2.42	20	5.87	5	0.17	-0.01	0.99	0.10
rope skipping 1-different exercises	20	18.33	20	3.37	4	0.69	0.73	0.99	0.68

Overall, Matrix Profile is well suited for repetition counting and segmentation and for revealing temporal motif topology, but it is less discriminative than Fourier-based analysis for fine-grained structural similarity. Accordingly, we recommend using Matrix Profile to recover repetition structure and boundaries, in conjunction with Fourier (and related metrics) for precise inter-motion similarity assessment.

4. Conclusion

This study evaluated skeleton-based analyses of repetitive human motion in video. A YOLOv11-based inference pipeline achieved high detection accuracy and throughput, underscoring its suitability for applied settings such as sports performance analysis, rehabilitation monitoring, and human–robot interaction. Among the downstream analyses, Fourier-based representations most effectively quantified inter-motion similarity, yielding lower dynamic time warping (DTW) distances and higher cosine-similarity scores, while PCA improved signal quality by attenuating noise. The Matrix Profile method excelled at recovering repetition structure—localizing motifs and delineating onset/offset boundaries—but was less precise than Fourier analysis for fine-grained structural similarity. Taken together, Fourier is preferred for discriminative similarity measurement, whereas the Matrix Profile is a complementary tool for robust motif and boundary detection; using both in tandem provides a comprehensive characterization of repetitive motion.

Generalizability is limited by the scope and diversity of the available datasets. Future research should employ larger, better-balanced corpora spanning varied motion types and environmental conditions (e.g., viewpoints, occlusions, illumination) and pursue real-time–oriented enhancements, including computational optimization, multi-view fusion, angle/topology-invariant features, and domain adaptation.

References

Agarwal, S., Gupta, M., Khandelwal, S., Jain, P., Aggarwal, A., Singh, D., & Mishra, V. K. (2021). FitMe: a fitness application for accurate pose estimation using deep learning. 2021 2nd International Conference on Secure Cyber Computing and Communications (ICSCCC), 232–237.

Andriluka, M., Pishchulin, L., Gehler, P., & Schiele, B. (2014). 2d human pose estimation: New benchmark and state of the art analysis. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 3686–3693.

Chen, Y., Xie, X., Yin, W., Li, B., & Li, F. (2023). Structure guided network for human pose estimation. *Applied Intelligence*, 53(18), 21012–21026. https://doi.org/10.1007/S10489-023-04521-8/FIGURES/7

Cao, Z., Simon, T., Wei, S. E., & Sheikh, Y. (2017). 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).

Dantone, M., Gall, J., Leistner, C., & Van Gool, L. (2013). Human pose estimation using body parts dependent joint regressors. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 3041–3048.

Lugaresi, C., Tang, J., Nash, H., McClanahan, C., Uboweja, E., Hays, M., ... & Grundmann, M. (2019). Mediapipe: A framework for building perception pipelines. *arXiv preprint arXiv:1906.08172*.

Papandreou, G., Zhu, T., Chen, L. C., Gidaris, S., Tompson, J., & Murphy, K. (2018). Personlab: Person pose estimation and instance segmentation with a bottom-up, part-based, geometric embedding model. In *Proceedings of the European conference on computer vision (ECCV)* (pp. 269-286).

Pham, H. H., Khoudour, L., Crouzil, A., Zegers, P., & Velastin, S. A. (2022). Video-based human action recognition using deep learning: A review. https://arxiv.org/abs/2208.03775v1

Sakoe, H., & Chiba, S. (2003). Dynamic programming algorithm optimization for spoken word recognition. *IEEE transactions on acoustics, speech, and signal processing*, *26*(1), 43-49.

Samkari, E., Arif, M., Alghamdi, M., & Al Ghamdi, M. A. (2023). Human pose estimation using deep learning: a systematic literature review. *Machine Learning and Knowledge Extraction*, 5(4), 1612–1659.

Usman, M., & Zhong, J. (2022). Skeleton-based motion prediction: A survey. *Frontiers in Computational Neuroscience*, 16, 1051222. https://doi.org/10.3389/FNCOM.2022.1051222/BIBTEX

Wang, C., Zhang, F., & Ge, S. S. (2021). A comprehensive survey on 2D multi-person pose estimation methods. *Engineering Applications of Artificial Intelligence*, 102, 104260.

Yan, S., Xiong, Y., & Lin, D. (2018). Spatial temporal graph convolutional networks for skeleton-based action recognition. *32nd AAAI Conference on Artificial Intelligence*, AAAI 2018, 7444–7452. https://arxiv.org/abs/1801.07455v2