Lift It Up Right: A Recommender System for Safer Lifting Postures

Gaetano Dibenedetto
University of Bari Aldo Moro
Bari, Italy
gaetano.dibenedetto@uniba.it

Pasquale Lops
University of Bari Aldo Moro
Bari, Italy
pasquale.lops@uniba.it

Marco Polignano
University of Bari Aldo Moro
Bari, Italy
marco.polignano@uniba.it

Helma Torkamaan
Delft University of Technology
Delft, Netherlands
H.Torkamaan@tudelft.nl

Abstract

Work-related musculoskeletal disorders, often caused by poor lifting posture and unsafe manual handling, continue to pose a significant threat to worker health and safety. This paper presents a health recommender system designed to prevent injury by assessing and correcting posture for lifting techniques. Leveraging monocular video input, our method estimates key ergonomic parameters to compute the Lifting Index based on the Revised NIOSH Lifting Equation. When the computed Lifting Index exceeds a predefined safety threshold, the system automatically generates graphical and textual recommendations to guide the worker towards safer postural strategies. This safety-aware recommender system provides interpretable and actionable feedback without requiring wearable sensors or multi-camera setups, making it suitable for deployment in real-world workplace environments. By integrating ergonomics with recommender system design, we contribute to a new class of context-aware, safety-oriented recommendation technologies tailored for occupational health.

CCS Concepts

• Computing methodologies → Computer vision; **• General and reference → Measurement**; **• Applied computing → Health care information systems**; **• Information systems → Recommender systems**.

Keywords

Health Recommender Systems, Postural Correction, Ergonomics, Lifting, Niosh Lifting Equation, Monocular Camera, Video-based Recommendation,

ACM Reference Format:

Gaetano Dibenedetto, Pasquale Lops, Marco Polignano, and Helma Torkamaan. 2025. Lift It Up Right: A Recommender System for Safer Lifting Postures. In *Proceedings of the Nineteenth ACM Conference on Recommender Systems (RecSys '25), September 22–26, 2025, Prague, Czech Republic*. ACM, New York, NY, USA, 6 pages. https://doi.org/10.1145/3705328.3759314

ACM ISBN 979-8-4007-1364-4/25/09
https://doi.org/10.1145/3705328.3759314

1 Introduction

Musculoskeletal disorders (MSDs) are among the most common work-related health problems in Europe, affecting millions of workers and costing employers billions annually [12]. These disorders often arise from poor posture, repetitive movements, and unsafe lifting techniques, leading to long-term injuries and reduced worker efficiency [29] with repeated incorrect behaviors being hard to unlearn even when risks are known [23]. Early intervention and early-stage training are key to prevention.

The U.S. National Institute for Occupational Safety and Health (NIOSH) developed the Revised NIOSH Lifting Equation (RNLE) [32, 33], widely used—including in ISO 11228 [15]—to assess manual lifting risk via the Lifting Index (LI). Traditionally, ergonomists manually assess these variables, which is time-consuming. Based on these assessments and the resulting LI, ergonomists recommend strategies for workers and companies to reduce injury risks, relying primarily on the LI outcome.

We propose a novel, safety-aware recommender system aimed at preventing injury risk through posture correction, based on a computer vision approach. The system extracts lifting parameters directly from video footage, eliminating the need for manual measurements. We provide automated, quantitative feedback on lifting posture by detecting key moments in a lifting sequence and estimating essential ergonomic variables. If the computed LI exceeds the accepted ergonomic limit, the system generates posture recommendations, presented in both textual and visual formats, to help workers identify unsafe movements and adopt safer lifting techniques. The code and the dataset are available on GitHub[1]. Our paper contributes to the field of health recommender systems [14, 26, 28] which often differ from traditional RecSys approaches due to the nature of health domain [26] and constraints like data scarcity, privacy, and safety requirements [7, 20]. In particular, RecSys papers such as [4, 16] support the relevance of our methodology and position of our contribution within this growing area.

2 Related Works

Health-focused recommender systems. These systems span food choices [22], drugs [17], general nutrition [25], mental health support [27], physical activity guidance [8, 16], and runner coaching [4, 11]. To provide proper recommendations, most of these systems rely on user-provided data, such as food diaries, symptom logs, self-assessments of mood, and manual training reports. Sports-oriented

[1] https://github.com/GaetanoDibenedetto/recsys25/

recommender systems incorporate wearable sensor signals, such as GPS and heart rate [4, 11]. Sensor-based approaches cannot deliver real-time feedback on complex movements captured in video. To address this, Jaiswal et al. [16] introduced a pipeline for diagnosing fitness technique error, using MediaPipe [3] to extract body keypoints and an Interaction Network [2] to learn motion dynamics. In their approach, the system provides corrective recommendations based on the correct learned exercise, relying on repetition counting and movement analysis, without modeling external loads or ergonomic risk scores, and assuming known exercise types. As such, they do not apply to the working environment demanding spatial relationships between the user and object, and their corrective feedback can feel like a *black box* to safety engineers who must justify interventions. Safety-aware recommender systems for ergonomics and posture work environments remain an important but novel application area.

Ergonomics, pose detection, and lifting-risk assessment. In parallel, the ergonomics community has pursued vision-based lifting-risk assessment under formal frameworks like the Revised NIOSH Lifting Equation, adherent to international standards for working environments. Early work by Wang et al. [30] employed wearable markers and biomechanical calculations to estimate lifting risks. Later studies relied on multiple camera placements, focusing either on classifying safe and unsafe lifts [35], or on estimating the force and moment at the lumbosacral joint [19]. In [35], data are processed using an existing human pose estimator [5], an optical flow estimator [18], and facial expression recognition [1]; while in [19] one of the first studies utilizing an ad hoc solution, training a human pose estimator from scratch based on the hourglass network architecture [21] is presented. More recently, Sabetta et al. [24] leveraged a single sagittal-plane camera with MediaPipe [3] and a fine-tuned object detector to semi-automate RNLE reporting as a conceptual work—yet without validating distances, lift-phase detection, real-time LI computation, end-to-end automation, or a recommendation component.

Our contribution builds on the single-camera ambition of [24, 30], but removes wearables, uses freely placed monocular cameras (e.g., a smartphone), and integrates open-vocabulary object detection with state-of-the-art pose estimation to automatically detect lift start/end via package velocity. We compute the full Lifting Index (LI), expose each RNLE component in real time, and generate transparent, interpretable, standards-compliant textual and graphical recommendations—delivering a modular, adaptable safety recommender suited for real-world workplace deployment.

3 Methodology

3.1 Problem Formulation

Our system is a posture-based lifting risk analyzer and recommender. Its primary objective is to estimate the LI, as defined by the RNLE[31], directly from the subject's posture P_{subj} and to generate posture corrections if the LI exceeds a safety threshold, as suggested by ergonomics [13]. In our scenario, the LI can be expressed as:

$$LI = f(H(P_{\text{subj}}), V(P_{\text{subj}}), D(P_{\text{subj}}))$$

where H, V, and D represent horizontal reach, vertical height, and vertical travel distance, respectively (Fig. 1). RNLE defines

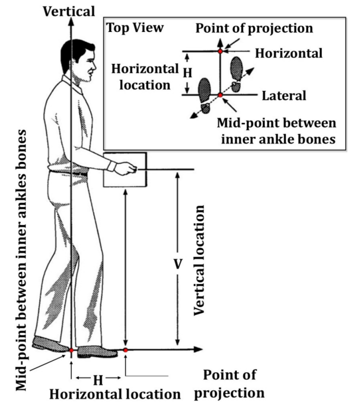


Figure 1: Graphical representation of V and H values (adapted from [31, p.536]).

LI as inversely proportional to the Recommended Weight Limit (RWL). While the RNLE additionally accounts for asymmetry angle (A), frequency over duration (F), and hand-to-object coupling (C) [31], the present study considers only the posture-dependent parameters—H, V, and D—together with the load constant (LC), a posture-independent term that varies with age and sex [15].

To detect lifting events, we identify the start (I_S) and end (I_E) frames within a video frame sequence $\mathcal{I} = \{I_1, \ldots, I_N\}$ by analyzing the object velocity $V_{\text{obj}}(t) = dP_{\text{obj}}(t)/dt$. ($I_S$) and ($I_E$) are determined by detecting when $V_{\text{obj}}(t)$ transitions from zero to nonzero (lifting begins) and nonzero to zero (lifting ends). The term *nonzero* is defined using a threshold value of γ.

From the posture data in these frames, we extract H, V, and D to compute the LI. If the LI exceeds a safety threshold τ, the system generates corrective feedback in visual G_{rec} and textual T_{rec} forms:

$$G_{\text{rec}}, T_{\text{rec}} = \begin{cases} g(H(P_{\text{subj}}), V(P_{\text{subj}}), D(P_{\text{subj}}), P_{\text{subj}}), & \text{if } LI > \tau \\ P_{\text{subj}}, & \text{otherwise} \end{cases}$$

The function g iteratively adjusts posture until $LI \leq \tau$.

3.2 Video Analysis

Our method offers a geometric and trigonometric solution for lifting analysis, avoiding the complexity of ad hoc approaches seen in related works. We use YOLOv8x-world [6] as object detector and ViTPose [34] as human pose estimator. YOLOv8x-world supports open-vocabulary detection, allowing to recognize boxes or packages not included in standard object categories, enhancing generalizability without fine-tuning. ViTPose, a Vision Transformer-based model, achieves state-of-the-art accuracy [10] and is highly scalable.

Frame by Frame Estimation. Accurate lifting parameter extraction requires precise localization of both the subject and the object across frames, even in occlusions, perspective distortions, and imperfect detections. Our frame-by-frame estimation pipeline addresses these challenges by combining object detection, human pose estimation, and side selection strategies to compute the required ergonomic parameters. Each frame is processed using YOLO to detect the subject and the package. If the subject is successfully detected, ViTPose is applied to estimate body keypoints. To compute values H, V and D, we use body landmarks extracted by the human pose estimator.

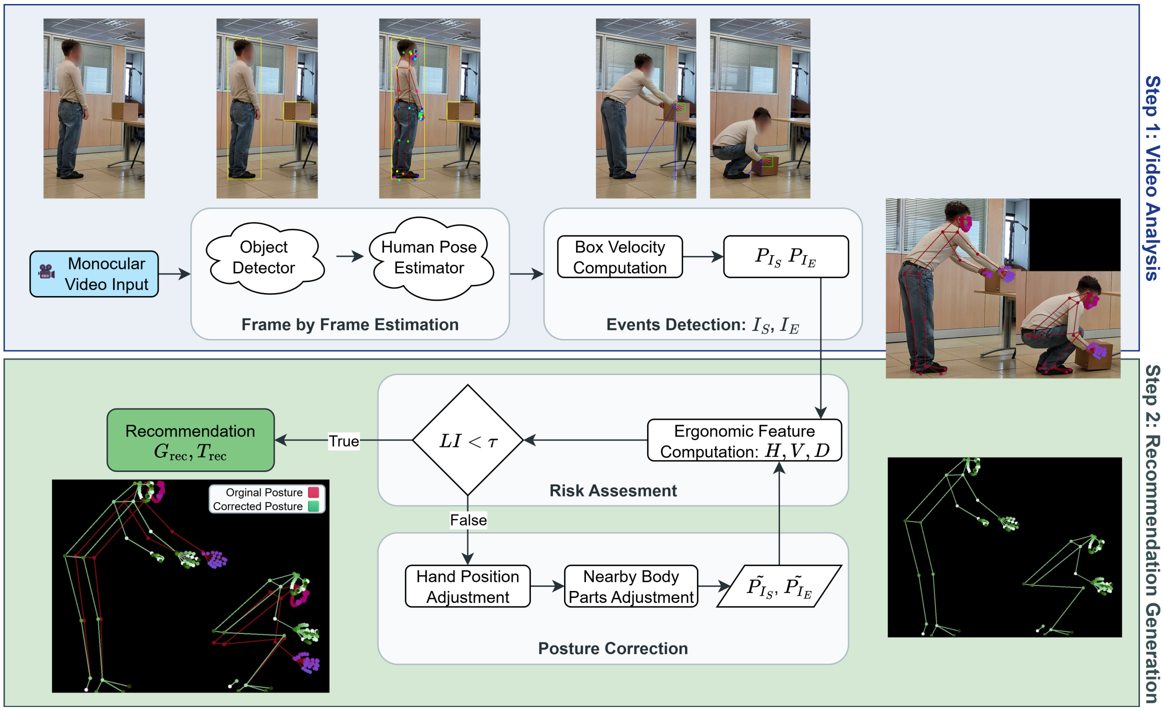


Figure 2: Pipeline for ergonomic recommendation generation from video analysis. The system detects the original lifting postures at the start (P_{I_S}) and end (P_{I_E}) of the movement, then generates corrected postures ($\tilde{P}_{I_S}$, $\tilde{P}_{I_E}$) if the Lifting Index exceeds a threshold. The final visualization overlays the original and recommended postures to illustrate the suggested adjustments.

V (i.e., hand-to-floor distance) is calculated from the landmark corresponding to the small toe up to the hand root. H is computed as the distance between the hand root and the average position of the ankles. D is defined following [31], as the change in V between key moments (I_S, I_E). All measurements, obtained in pixels, are then converted into real-world units using a scaling factor derived from the estimated height of the subject.

Detecting the Start and End of the Lift. To detect lift start and end, we estimate the box velocity by computing the difference in position between the current frame (t) and the frame five steps earlier ($t-5$), averaging over five frames to smooth noise. If the velocity exceeds a threshold of $\gamma = 0.3cm/sec$, the first and last frames exceeding γ are marked as the start and end of the lift (Fig. 2).

Averaging and Handling Missing Data. To handle missing or unreliable detections, distances from key frames are averaged with those from neighboring frames. The hand distance is chosen based on the leading side, and missing values are interpolated using the average of the previous and next frames. If multiple consecutive frames are missing, a wider temporal window is used for interpolation.

3.3 Recommendation Generation

Our system provides *visual* and *textual* lifting posture recommendations by adjusting key parameters to keep the LI below a threshold τ. To enhance interpretability and inspired by the visual feedback method in [9], we complement the graphical feedback with textual descriptions, making the recommendations easier to understand for both experts and non-experts.

To generate recommendations, we gradually adjust the subject's hand position—the main driver of changes in the lifting equation—until the LI drops below τ. Following Fox et al. [13]'s observation, who highlight the absence of a universal LI threshold, we adopt $\tau = 0.7$ for evaluation purposes, assuming a lifted load of 10 kg. Our analysis focuses on the two key frames: the start and end of the lifting movement. For each frame, the parameters H, V, and D are examined to identify their impact on the overall LI, and proper adjustments are recommended by applying horizontal, vertical, or combined corrections to the hand position accordingly. To maintain anatomical consistency, surrounding keypoints are displaced proportionally to the hand movement. Fig. 2 shows a full example of this recommendation process, corresponding to the following textual recommendation associated with that visual output:

```
Lifting Index Original Pose: 1.10;
Recommended Pose Lifting Index: 0.62;
At the start of the lifting: Move your hands at least 19cm closer
      to your body. This adjustment will likely also move your
      shoulders (8cm) and elbows (15cm) closer to your body. You
      should actively shift your hips approximately 6-10cm in the
      same direction.
At the end of the lifting: Move your hands at least 20cm up. This
      adjustment will likely also move your shoulders (8cm) and
      elbows (16cm) up. You should actively shift your hips
      approximately 6-10cm in the same direction. This posture
      change leads to a decrease in knee flexion of about 10.0°.
```

Table 1: Errors in the estimation of distances across environments and camera angles. Reported values (in cm) include vertical (V), horizontal (H), and displacement (D) errors, computed from the positions of the box and the hands. Errors in RWL estimation are reported using MAE and RMSE. The table also includes the number of subjects and videos analyzed in each environment subset. Results are grouped by environment and camera angle (side, rear side), with combined rows aggregating both views.

Env	Camera Angle	#Subjects	#Videos	Box Errors (cm)			Hand Errors (cm)			RWL Errors	
				V	H	D	V	H	D	MAE	RMSE
1	Side	3	23	1.69	29.47	1.17	39.08	-	-	7.27	7.60
	Rear side	3	18	2.72	23.05	2.66	38.55	-	-	5.76	6.63
	Combined	3	41	2.14	26.36	1.82	38.85	-	-	5.28	6.18
2	Side	2	8	1.00	15.62	6.12	17.50	-	18.50	0.00	0.00
	Rear side	2	8	4.12	1.12	0.75	23.12	-	26.12	8.13	8.22
	Combined	2	16	2.56	8.37	3.43	20.31	-	22.31	4.06	5.81
3	Side	8	37	1.40	9.29	1.59	48.43	-	58.76	1.22	1.43
	Rear side	8	39	2.74	1.92	1.61	42.07	-	51.15	3.84	4.35
	Combined	8	76	2.09	5.51	1.60	45.17	-	54.88	2.57	3.28
AVG	-	13	133	2.16	12.37	1.89	40.23	-	47.10	3.73	4.87

4 Experimental Evaluation

To evaluate our system, we collected a dataset[1] of lifting tasks performed in three different environments: *Env* #1 a lab, *Env* #2 an office, and *Env* #3 an aula (a classroom setting), with no constraints on camera placement. The dataset includes 13 distinct subjects, each performing two lifting actions: lifting a box from the ground to a desk and vice-versa. For some subjects, multiple recordings of the same action from the same camera angle are available, with subtle variations between them. The camera angles used are: (1) rear-side and (2) side views, which align with common setups in related works [19, 24, 35]. Front and direct rear views were excluded due to occlusion of the object or critical body parts.

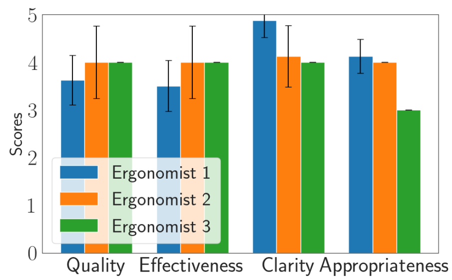
Distance Estimation Error. According to ergonomics literature, the ideal grip position during lifting is with the hands placed at the lower center of the box. To reflect this, we annotated distances from both the hand root and the box itself. Box-based annotations were easier to measure reliably—especially when referencing fixed surfaces like the floor or desk—and more consistent across repetitions.

However, estimating distances from body parts like the midpoint of the hand can introduce inaccuracies due to variation in grip positioning across repetitions. In environments where hand root distance data is missing, we only used the box-based measurements. Note that some annotation noise was present in *Env* #1 and #2, where measurements weren't collected by the authors. For instance, in *Env* #1, horizontal distance was measured from the front of the feet instead of mid-ankles, introducing a systematic error of about 15–20 cm. We also evaluated RWL error impact on LI using MAE and RMSE. Dataset details and results are in Table 1.

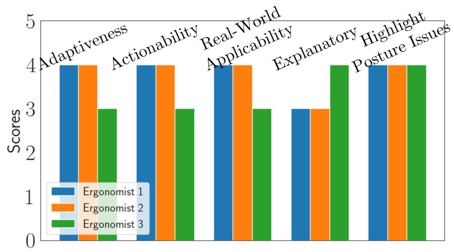
While MAE and RMSE are standard and interpretable metrics for quantifying estimation errors, their application to RWL prediction warrants careful interpretation [31]. Nonetheless, they remain valuable tools for highlighting the overall sensitivity of RWL to posture-related errors and identifying which parameters contribute most to variability. Among all posture-derived parameters, H emerges as the primary source of error. As shown in Table 1, estimation errors in H are the highest across environments, even in *Env* #3. This matters because the Horizontal Multiplier (HM)—a multiplier in RWL computed from the horizontal distance H—introduces a hard threshold at 63 cm: when H exceeds this value, HM drops from 0.40 to 0.00, causing the RWL to collapse to zero [31]. This binary behavior inflates the error captured by MAE and RMSE, even for small variations in H. *Env* #2 illustrates this phenomenon well. In the side camera setup, all annotated RWL values are zero due to consistently high H, leading to zero error metrics—not necessarily indicating accurate estimates, but rather consistent threshold-induced outcomes. Conversely, in the rear-side setup (still in *Env* #2), small average errors in H lead to substantial RWL fluctuations when estimates fall near the HM threshold, resulting in higher MAE and RMSE.

Expert Review of Recommendation. To assess the recommendation quality and its practical relevance, we conducted a structured expert evaluation with three professional ergonomists. Each ergonomist independently examined eight video recordings—six involving lifting postures with identifiable ergonomic risks. Four videos were recorded in the lab (two with correct, two with deliberately flawed posture), while the remaining were field recordings provided by the ergonomists. For each video, ergonomists first viewed the output of our tool with distance estimations, enabling them to manually compute the LI and derive corrective actions as they typically do for their clients. Only afterward, they reviewed and evaluated the system-generated recommendations, by answering four questions related to the recommendation *quality*, *effectiveness*, *clarity* and *appropriateness* (Fig. 3a). After reviewing all videos at the end, they answered five additional questions to assess the system's overall performance, focusing on *adaptiveness*, *actionability*, *explanatory power*, *real-world applicability*, and *ability to highlight posture issues* that are otherwise difficult to detect (Fig. 3b). The full list of questions is available in the GitHub repository[1].

Overall, recommendations were rated positively, with no score below 3 out of 5 (Fig. 3a). Clarity was the highest-rated dimension (AVG=4.3), indicating that the system's corrective actions were clear and easily understood by evaluators. Quality (AVG=3.9), effectiveness (AVG=3.8), and appropriateness (AVG=3.7) received slightly



(a) Per Video Evaluation



(b) Overall Evaluation

Figure 3: Recommendation Evaluation

lower but still favorable scores, suggesting that, while the recommendations were generally perceived as helpful and suitable, there is room for refinement to meet expectations more consistently.

In the system-level evaluation (Fig. 3b), adaptability, actionability, and real-world applicability each received an average score of 3.7, suggesting that ergonomists generally found the system flexible for different postures, functional, and practical for real-world use. The explanatory dimension received the lowest average score (3.3), indicating that although the system's outputs were clear, more detailed explanations of why the subject posture was risky and how the recommendation helps mitigate injury risk. The highest score (4.0) was for the system's ability to highlight posture issues, affirming the system's strength in identifying ergonomic risks that may be otherwise overlooked.

Limitations. Expert review confirmed that the recommendations are actionable, but larger-scale field studies are still required to evaluate and validate the proposed system further. The system uses an object detector not trained for typical packages—side-viewed boxes may appear flat—leading to missed detections and then inaccurate *hand* distance estimates. Fine-tuning could improve accuracy but may reduce generalizability. To comply with current ergonomic practices, the recommendation engine currently uses a fixed, rule-based approach that lacks environmental awareness (e.g., using a stand). Future work will explore integrating environmental context, leveraging visual cues for more adaptive, context-aware feedback.

5 Impact Statement

This work enables automated, interpretable lifting posture recommendations from video, helping reduce musculoskeletal injury risk in occupational settings. By providing both visual and textual feedback grounded in ergonomic principles and standards, the system supports safer lifting behavior. Our method introduces a sensor-free approach to ergonomic risk assessment without requiring any specialized hardware that is non-intrusive and applicable across diverse real-world settings. Its passive, smartphone-based, and scalable nature makes it well-suited for workplace training, health assessments, and long-term monitoring in logistics, manufacturing, and warehouse environments. However, overreliance on automated feedback poses risks. The system lacks full environmental and contextual awareness e.g., space constraints, urgency, or worker fatigue. As such, it should be used to complement, not replace, professional ergonomist evaluations. Its interpretability and ease of use help make feedback more accessible to both experts and non-experts, supporting more informed and safer workplace practices.

6 Conclusions

We presented the first vision-only, safety-aware recommendation system for lifting posture assessment and correction. Using pose estimation and open-vocabulary object detection, our system computes RNLE parameters with centimeter-level accuracy and provides interpretable feedback when the Lifting Index exceeds a configurable threshold. We evaluated our approach on 133 lifting sequences across three environments, demonstrating low distance estimation error and positive expert feedback on the clarity and usefulness of the recommendations. These results establish a strong proof of concept and open the door to future integration in real-world occupational safety workflows.

Acknowledgments

The research is partially funded by PNRR - Mission 4 ("Education and research") – Component 2 ("From research to business"), Investment 3.3 ("Introduction of innovative doctorates that respond to the innovation needs of companies and promote the hiring of researchers by companies") D.M.n. 117/2023 - CUP: H91I23000170007 and is supported by the co-funding of the European Union - Next Generation EU: NRRP Initiative, Mission 4, Component 2, Investment 1.3 – Partnerships extended to universities, research centers, companies, and research D.D. MUR n. 341 del 15.03.2022 – Next Generation EU (PE0000013 – "Future Artificial Intelligence Research – FAIR" - CUP: H97G22000210007).
We extend our sincere gratitude to Naps Lab S.r.l.s. for their support and collaboration in the realisation of this research.

References

[1] Tadas Baltrusaitis, Amir Zadeh, Yao Chong Lim, and Louis-Philippe Morency. 2018. OpenFace 2.0: Facial Behavior Analysis Toolkit. In *2018 13th IEEE International Conference on Automatic Face & Gesture Recognition (FG 2018)*. 59–66. doi:10.1109/FG.2018.00019

[2] Peter W. Battaglia, Razvan Pascanu, Matthew Lai, Danilo Jimenez Rezende, and Koray Kavukcuoglu. 2016. Interaction Networks for Learning about Objects, Relations and Physics. In *Advances in Neural Information Processing Systems 29: Annual Conference on Neural Information Processing Systems 2016, December 5-10, 2016, Barcelona, Spain*, Daniel D. Lee, Masashi Sugiyama, Ulrike von Luxburg, Isabelle Guyon, and Roman Garnett (Eds.). 4502–4510.

[3] Valentin Bazarevsky, Ivan Grishchenko, Karthik Raveendran, Tyler Zhu, Fan Zhang, and Matthias Grundmann. 2020. BlazePose: On-device Real-time Body Pose tracking. *CoRR* abs/2006.10204 (2020). arXiv:2006.10204

[4] Jakim Berndsen, Barry Smyth, and Aonghus Lawlor. 2019. Pace my race: recommendations for marathon running. In *Proceedings of the 13th ACM Conference on Recommender Systems* (Copenhagen, Denmark) *(RecSys '19)*. Association for Computing Machinery, New York, NY, USA, 246–250. doi:10.1145/3298689.3346991

[5] Zhe Cao, Gines Hidalgo, Tomas Simon, Shih-En Wei, and Yaser Sheikh. 2021. OpenPose: Realtime Multi-Person 2D Pose Estimation Using Part Affinity Fields. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 43, 1 (2021), 172–186. doi:10.1109/TPAMI.2019.2929257

[6] Tianheng Cheng, Lin Song, Yixiao Ge, Wenyu Liu, Xinggang Wang, and Ying Shan. 2024. YOLO-World: Real-Time Open-Vocabulary Object Detection. In *IEEE/CVF Conference on Computer Vision and Pattern Recognition, CVPR 2024, Seattle, WA, USA, June 16-22, 2024*. IEEE, 16901–16911. doi:10.1109/CVPR52733.2024.01599

[7] Ine Coppens, Luc Martens, and Toon De Pessemier. 2023. Analyzing Accuracy versus Diversity in a Health Recommender System for Physical Activities: a Longitudinal User Study. In *Proceedings of the 17th ACM Conference on Recommender Systems, RecSys 2023, Singapore, Singapore, September 18-22, 2023*, Jie Zhang, Li Chen, Shlomo Berkovsky, Min Zhang, Tommaso Di Noia, Justin Basilico, Luiz Pizzato, and Yang Song (Eds.). ACM, 1146–1151. doi:10.1145/3604915.3610650

[8] Gaetano Dibenedetto, Elio Musacchio, Marco Polignano, and Pasquale Lops. 2025. Fine-Tuning Large Multimodal Models for Fitness Action Quality Assessment. In *Adjunct Proceedings of the 33rd ACM Conference on User Modeling, Adaptation and Personalization, UMAP Adjunct 2025, New York City, NY, USA, June 16-19, 2025*. ACM, 39–44. doi:10.1145/3708319.3733684

[9] Gaetano Dibenedetto, Marco Polignano, Pasquale Lops, and Giovanni Semeraro. 2024. Human Pose Estimation for Explainable Corrective Feedbacks in Office Spaces. In *Adjunct Proceedings of the 32nd ACM Conference on User Modeling, Adaptation and Personalization, UMAP Adjunct 2024, Cagliari, Italy, July 1-4, 2024*. ACM. doi:10.1145/3631700.3665184

[10] Gaetano Dibenedetto, Stefanos Sotiropoulos, Marco Polignano, Giuseppe Cavallo, and Pasquale Lops. 2025. Comparing human pose estimation through deep learning approaches: An overview. *Computer Vision and Image Understanding* (2025), 104297. doi:10.1016/j.cviu.2025.104297

[11] Ciara Feely, Brian Caulfield, Aonghus Lawlor, and Barry Smyth. 2024. Recommending Personalised Targeted Training Adjustments for Marathon Runners. In *Proceedings of the 18th ACM Conference on Recommender Systems* (Bari, Italy) *(RecSys '24)*. Association for Computing Machinery, New York, NY, USA, 1051–1056. doi:10.1145/3640457.3688192

[12] European Agency for Safety and Health at Work. 2019. Musculoskeletal disorders. https://osha.europa.eu/en/themes/musculoskeletal-disorders. Accessed: 2025-04-16.

[13] Robert R. Fox, Ming-Lun Lu, Enrico Occhipinti, and Matthias Jaeger. 2019. Understanding outcome metrics of the revised NIOSH lifting equation. *Applied Ergonomics* 81 (2019), 102897. doi:10.1016/j.apergo.2019.102897

[14] Hanna Hauptmann, Christoph Trattner, and Helma Torkamaan. 2024. The 6th International Workshop on Health Recommender Systems. In *Proceedings of the 18th ACM Conference on Recommender Systems* (Bari, Italy) *(RecSys '24)*. Association for Computing Machinery, New York, NY, USA, 1232–1236. doi:10.1145/3640457.3687113

[15] ISO. 2021. ISO 11228-1:2021 Ergonomics — Manual handling Part 1: Lifting, lowering and carrying. https://www.iso.org/standard/76820.html. Accessed: 2025-04-01.

[16] Abhishek Jaiswal, Gautam Chauhan, and Nisheeth Srivastava. 2023. Using Learnable Physics for Real-Time Exercise Form Recommendations. In *Proceedings of the 17th ACM Conference on Recommender Systems* (Singapore, Singapore) *(RecSys '23)*. Association for Computing Machinery, New York, NY, USA, 688–695. doi:10.1145/3604915.3608816

[17] Haoxuan Li, Taojun Hu, Zetong Xiong, Chunyuan Zheng, Fuli Feng, Xiangnan He, and Xiao-Hua Zhou. 2023. ADRNet: A Generalized Collaborative Filtering Framework Combining Clinical and Non-Clinical Data for Adverse Drug Reaction Prediction. In *Proceedings of the 17th ACM Conference on Recommender Systems* (Singapore, Singapore) *(RecSys '23)*. Association for Computing Machinery, New York, NY, USA, 682–687. doi:10.1145/3604915.3608813

[18] Pengpeng Liu, Michael Lyu, Irwin King, and Jia Xu. 2019. SelFlow: Self-Supervised Learning of Optical Flow. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*.

[19] Rahil Mehrizi, Xi Peng, Dimitris N. Metaxas, Xu Xu, Shaoting Zhang, and Kang Li. 2019. Predicting 3-D Lower Back Joint Load in Lifting: A Deep Pose Estimation Approach. *IEEE Transactions on Human-Machine Systems* 49, 1 (2019), 85–94. doi:10.1109/THMS.2018.2884811

[20] Fedelucio Narducci, Cataldo Musto, Marco Polignano, Marco de Gemmis, Pasquale Lops, and Giovanni Semeraro. 2015. A Recommender System for Connecting Patients to the Right Doctors in the HealthNet Social Network. In *Proceedings of the 24th International Conference on World Wide Web* (Florence, Italy) *(WWW '15 Companion)*. Association for Computing Machinery, New York, NY, USA, 81–82. doi:10.1145/2740908.2742748

[21] Alejandro Newell, Kaiyu Yang, and Jia Deng. 2016. Stacked Hourglass Networks for Human Pose Estimation. In *Computer Vision – ECCV 2016*, Bastian Leibe, Jiri Matas, Nicu Sebe, and Max Welling (Eds.). Springer International Publishing, Cham, 483–499.

[22] Alessandro Petruzzelli, Cataldo Musto, Michele Ciro Di Carlo, Giovanni Tempesta, and Giovanni Semeraro. 2024. Recommending Healthy and Sustainable Meals exploiting Food Retrieval and Large Language Models. In *Proceedings of the 18th ACM Conference on Recommender Systems* (Bari, Italy) *(RecSys '24)*. Association for Computing Machinery, New York, NY, USA, 1057–1061. doi:10.1145/3640457.3688193

[23] Ana Lucy Rodrigues Ferreira Faisting and Tatiana de Oliveira Sato. 2019. Effectiveness of ergonomic training to reduce physical demands and musculoskeletal symptoms - An overview of systematic reviews. *International Journal of Industrial Ergonomics* 74 (2019), 102845. doi:10.1016/j.ergon.2019.102845

[24] Nicolò Sabetta, Margherita Bernabei, Silvia Colabianchi, Davide Colangelo, Francesco Costantino, et al. 2024. Ergonomic training tool: a pose detection-based digitalization of ISO/TR 12295 and ISO 11228-1. In *Proceedings of the Summer School Francesco Turco*.

[25] Hanna Schäfer. 2016. Personalized Support for Healthy Nutrition Decisions. In *Proceedings of the 10th ACM Conference on Recommender Systems* (Boston, Massachusetts, USA) *(RecSys '16)*. Association for Computing Machinery, New York, NY, USA, 455–458. doi:10.1145/2959100.2959105

[26] Hanna Schäfer, Santiago Hors-Fraile, Raghav Pavan Karumur, André Calero Valdez, Alan Said, Helma Torkamaan, Tom Ulmer, and Christoph Trattner. 2017. Towards Health (Aware) Recommender Systems. In *Proceedings of the 2017 International Conference on Digital Health, London, United Kingdom, July 2-5, 2017*, Patty Kostkova, Floriana Grasso, Carlos Castillo, Yelena Mejova, and Arnold Bosman (Eds.). ACM, 157–161. doi:10.1145/3079452.3079499

[27] Helma Torkamaan and Jürgen Ziegler. 2022. Recommendations as Challenges: Estimating Required Effort and User Ability for Health Behavior Change Recommendations. In *IUI 2022: 27th International Conference on Intelligent User Interfaces, Helsinki, Finland, March 22 - 25, 2022*, Giulio Jacucci, Samuel Kaski, Cristina Conati, Simone Stumpf, Tuukka Ruotsalo, and Krzysztof Gajos (Eds.). ACM, 106–119. doi:10.1145/3490099.3511118

[28] Thi Ngoc Trang Tran, Alexander Felfernig, Christoph Trattner, and Andreas Holzinger. [n. d.]. Recommender systems in the healthcare domain: state-of-the-art and research issues. 57, 1 ([n. d.]), 171–201. doi:10.1007/s10844-020-00633-6

[29] U.S. Bureau of Labor Statistics. 2018. *Injuries, illnesses, and fatalities (IIF)*. Washington, D.C., USA.

[30] M.J. Wang, G.J. Huang, W.Y. Yeh, and C.L. Lee. 1996. Manual lifting task risk evaluation using computer vision system. *Computers & Industrial Engineering* 31, 3 (1996), 657–660. doi:10.1016/S0360-8352(96)00254-9 18th International Conference on Computers and Industrial Engineering.

[31] Thomas R Waters. 2007. Revised NIOSH lifting equation. In *Biomechanics in Ergonomics*. CRC Press, 549–580.

[32] Thomas R Waters, Vern Putz-Anderson, and Arun Garg. 1994. Applications manual for the revised NIOSH lifting equation. (1994).

[33] Thomas R. Waters, Vern Putz-Anderson, Arun Garg, and Lawrence J. Fine. 1993. Revised NIOSH equation for the design and evaluation of manual lifting tasks. *Ergonomics* 36, 7 (1993), 749–776. doi:10.1080/00140139308967940 PMID: 8339717.

[34] Yufei Xu, Jing Zhang, Qiming Zhang, and Dacheng Tao. 2022. ViTPose: Simple Vision Transformer Baselines for Human Pose Estimation. In *Advances in Neural Information Processing Systems 35: Annual Conference on Neural Information Processing Systems 2022, NeurIPS 2022, New Orleans, LA, USA, November 28 - December 9, 2022*, Sanmi Koyejo, S. Mohamed, A. Agarwal, Danielle Belgrave, K. Cho, and A. Oh (Eds.). http://papers.nips.cc/paper_files/paper/2022/hash/fbb10d319d44f8c3b4720873e4177c65-Abstract-Conference.html

[35] Guoyang Zhou, Vaneet Aggarwal, Ming Yin, and Denny Yu. 2022. A Computer Vision Approach for Estimating Lifting Load Contributors to Injury Risk. *IEEE Transactions on Human-Machine Systems* 52, 2 (2022), 207–219. doi:10.1109/THMS.2022.3148339