



A novel platform to enable the future human-centered factory

Andrea Generosi¹ · Thomas Agostinelli¹  · Silvia Ceccacci² · Maura Mengoni¹

Received: 13 July 2022 / Accepted: 28 July 2022

© The Author(s) 2022

Abstract

This paper introduces a web-platform system that performs semi-automatic compute of several risk indexes, based on the considered evaluation method (e.g., RULA—Rapid Upper Limb Assessment, REBA—Rapid Entire Body Assessment, OCRA—Occupational Repetitive Action) to support ergonomics risk estimation, and provides augmented analytics to proactively improve ergonomic risk monitoring based on the characteristics of workers (e.g., age, gender), working tasks, and environment. It implements a body detection system, marker-less and low cost, based on the use of RGB cameras, which exploits the open-source deep learning model CMU (Carnegie Mellon University), from the tf-pose-estimation project, assuring worker privacy and data protection, which has been already successfully assessed in standard laboratory conditions. The paper provides a full description of the proposed platform and reports the results of validation in a real industrial case study regarding a washing machine assembly line composed by 5 workstations. A total of 15 workers have been involved. Results suggest how the proposed system is able to significantly speed up the ergonomic assessment and to predict angles and perform a RULA and OCRA analysis, with an accuracy comparable to that obtainable from a manual analysis, even under the unpredictable conditions that can be found in a real working environment.

Keywords Motion capture · Ergonomics risk assessment · Machine learning · Extended reality · Human-centered manufacturing

1 Introduction

Industry 4.0 renews the focus on human-centered manufacturing: human still represents an extraordinary driver for flexibility in high automation manufacturing environments, ensuring creativity and problem-solving. To improve factories' efficiency, productivity, and social sustainability, one

of the challenges is to put workers in human-suitable working conditions, adapting the workspace to the individual capabilities going behind the concept of “the worker” as a homogeneous group. To implement this worker-centric manufacturing model, more effective health and safety management systems are needed, based on objective, automatic, and real-time MSDs (musculoskeletal disorders) risks assessment. In fact, in a rapidly changing scenario like the manufacturing environment is today, the workload a worker is subjected to is also constantly changing. This situation, in addition to aggravating the mental workload (i.e., because they must remember the steps of many different assembly stations), increases the likelihood of MSDs. This happens because workers are frequently moved to a new workstation and do not have time to memorize correct postures and any corrections made to improve posture. Obviously, given the high frequency with which workers change workstations or tasks, it is impractical for an ergonomist to perform a new ergonomic risk analysis each time. Current practices make ergonomic assessment costly and time-consuming, highly affect the intra- and inter-observer results variability, and may lead to low accuracy of such evaluations [1].

✉ Thomas Agostinelli
t.agostinelli@pm.univpm.it

Andrea Generosi
a.generosi@univpm.it

Silvia Ceccacci
silvia.ceccacci@unimc.it

Maura Mengoni
m.mengoni@univpm.it

¹ Department of Industrial Engineering and Mathematical Sciences, Università Politecnica delle Marche, Via Brecce Bianche, 12, Ancona 60131, Italy

² Department of Education, Cultural Heritage and Tourism, University of Macerata, Piazzale Luigi Bertelli (Contrada Vallebona), Macerata 62100, Italy

It can be seen that there are now many simulation tools that support ergonomic analyses during the workstation design process, including digital human mannequins, which can potentially eliminate the need for field analyses if well used (e.g., software such as CATIA, CREO, DELMIA, JACK). In any case, simulations are not without drawbacks: in order to ensure a realistic process simulation, they require expert users and a lot of effort again to assign the mannequins proper posture settings [2]. One way to partly speed up the process is to integrate the virtual mannequins with real user posture information, through the use of motion capture technologies. However, the use of such systems to acquire data in real working environments is still scarce [3] since it requires very complex procedures that are not immediately applicable [4], so that their application remains limited to virtual simulation contexts.

In this context, the introduction of new methods and tools for semi-automated ergonomic postural assessment, which by exploiting the input of Motion Capture (MoCap) systems based on one or more RGB images enables the identification of key points of the human body, is the solution of choice. Significant advances in the field of computer vision have been made in recent years, leading to the definition of new body tracking systems based on supervised machine learning techniques, which could become a breakthrough for markerless postural assessment, as the assessment would eventually be available from photos and videos taken from any common RGB camera, improving the accessibility and applicability of the system [1, 5].

Research has produced so far either 2D and 3D RGB MoCap systems [1]. Systems based on the acquisition of 3D data are more accurate of their 2D counterparts, but to reach such an high accuracy they have drawbacks: before they can be used, they need to be calibrated to obtain intrinsic camera parameters [6]. On the other hand, while being less accurate, 2D MoCap systems present the advantage of being more portable and more flexible, given they can adapt to various positioning configurations without requiring recalibration each time. Although 2D systems have proven to have interesting potential for the future, the validation of these tools is still mostly done under controlled conditions in the laboratory [1, 5], while experimentation in real working conditions is quite scarce.

To support the ergonomist performing an ergonomic analysis in a more efficient and reliable way, this paper introduces a novel web-platform system able to perform semi-automatic ergonomic risk assessment based on several evaluation methods (i.e., RULA, REBA, OWAS, OCRA). It implements a 2D motion capture tool that needs at least one common RGB camera (even that of a smartphone), exploiting an open-source deep learning model with low computational requirements, already validated in laboratory conditions [5]. In order to test the effectiveness of such a system

in supporting ergonomic risk assessment in a real industrial production environment, we discuss the results of an experiment conducted in an assembly line of an Italian washing machine and dryer manufacturing plant.

The paper is organized as follows: in Sect. 2, we present the research background, while in Sect. 3 we describe the proposed system that is used in the experimentation, which is detailed in Sect. 4. Section 5 shows the numeric results and the equations that support data processing. Finally, in Sect. 6, discussion and conclusions are presented.

2 Research background

2.1 Ergonomic risk assessment methods

Musculoskeletal disorders (MSDs) are a critical work-related health problem [7, 8] and are caused by the accumulation, due to inappropriate postures assumed by workers during manual labor, of musculoskeletal loads. To reduce MSDs, ergonomists have been utilizing ergonomic risk assessment protocols that help them evaluate the risk, which are based on direct on-site observation or posthumous video analysis of workers when performing their jobs. They are easy to use and applicable for many work scenarios, although they demand a time-consuming, and thus costing, activity of on-site observation or manual video analysis. Among those, RULA—Rapid Upper Limb Assessment [9], REBA—Rapid Entire Body Assessment [10] (whose popularity is increasing over the years according to [3]), OWAS—Ovako Working Posture Analysis System [11], and OCRA—Occupational Repetitive Action [12] are four of the most popular choices. In summary, the characteristics of each are as follows:

- RULA, it is a so-called “rapid” assessment tool that assigns a score ranging from 1 (lowest risk) to 7 (highest risk) to the task analyzed. It does so by comparing the angles assumed by several body’s joints (e.g., the neck in relation to the back, the elbow, the shoulders, the knees) with predefined angular ranges. It also takes into account muscle load and frequency of repetition of the posture within the task to calculate the score but does not consider the total duration of the cycle.
- REBA, it indicates the need to change the workstation layout by assigning each posture a score. This score is calculated by analyzing, through a checklist of predefined limit positions, the position of several body parts, such as the neck, upper limbs, trunk, lower limbs, and wrist. This method takes also in consideration the presence of loads applied to the worker’s body.
- OWAS, with reference to some predefined levels of danger, the worker’s postures are designated by a 4-digit code representing its danger level. This code is computed

considering the arms, back, and legs as well as the weight lifted by the worker. The single postures are then analyzed taking into account the sequence that binds them within the task.

- OCRA, was developed to assess the MSDs in a more detailed manner, not only considering awkward upper limb postures and hand gestures, but also other risk factors such as lack of appropriate recovery periods or the task duration.

From the previous description of the four methods, it is clear that, whichever of them the choice falls on, the ergonomist must spend a large amount of time analyzing the task of interest either in person or through photo and video recordings, with the latter becoming more common due to the ability to record video or photos via a smartphone, now present in the pockets of virtually any adult in a developed country [13].

2.2 Ergonomic risk assessment tools

To reduce the time spent analyzing video captures, researchers have been proposing automatic or semi-automatic ergonomic risk assessment tools. Initially, those systems were based on the acquisition of the human body postures through high-end motion capture hardware, such as optical or inertial wearable sensors as in [14–16]. The limitations of these instruments are, on the one hand, the high economic investment required to acquire them and, on the other hand, the impractical use in real manufacturing environments, for example because they require the worker to wear invasive wearable sensors during the acquisition. To reduce costs and get rid of wearable sensors, new tools were proposed. Many studies proposed systems based on a low-cost motion capture tool, the Microsoft Kinect v2, e.g. [17, 18], to automatically collect joint's angles necessary for ergonomic risk assessment based on RULA. However, although Microsoft Kinect v2 is a low-cost equipment when compared to the high-end gear, it suffers of occlusion and self-occlusion problems [1, 19]. Occlusion of the human body is the normality in a manufacturing environment, and since it leads to partial human body acquisition, the Kinect cannot be the universal solution of choice. Combining the data acquired from a depth camera with Machine Learning algorithms is an effective solution to predict those joints' angles that are occluded by the obstacles; however, this can only be obtained at a cost of a high computational demand from the tool [18]. So, to overcome the occlusion and the high computational costs, more advanced techniques were proposed. They are based on the acquisition of a video recording through a simple RGB camera and on the processing of these videos with Machine Learning algorithm, to predict joint's location and thus to compute joints' angles [20–22].

2.3 RGB motion capture technologies

As previously mentioned in the last section, new Motion Capture (MOCAP) technologies, based on the acquisition of common photos or video recordings and on the subsequent processing by means of Machine Learning (ML) techniques, are gaining momentum. Several new tools are being proposed. The first example of such a technology was OpenPose [20], a powerful tool developed by the Carnegie Mellon University. It can detect and track the body, hands, and face with a remarkable accuracy and without suffering of occlusion problems [23] and has proved itself to be useful in performing an ergonomic risk assessment [24, 25]. However, this high performance is paid in high computational resources needed [26].

A novel tool, namely tf-pose-estimation [21], was later proposed. It was developed as a lightweight tool to enable real-time computation. Because of its light weight, it can also work on mobile devices [27]; its accuracy as a tool to help ergonomists assess ergonomic risk has been validated in [5].

Later, Google proposed its own framework, MediaPipe [22]. This framework contains several tools to detect body features, such as pose detection, hand detection, iris detection, and so on. However, its best strength is the hand detection module, which accuracy has been assessed in several works [28–30].

All the works presented here have a common trait: they were tested only under controlled conditions in a laboratory environment. Validation of these tools in a real-world environment (e.g., manufacturing plant) is the next step that may or may not decree their wide adoption.

Thus, this paper proposes a semi-automatic tool that leverages the work presented in [31] to perform several ergonomic risk assessments, extending the previous work by adding some features and a user-friendly interface, and validating it in a real manufacturing environment.

3 The proposed system

This research work proposes a platform that, by means of Deep Learning models and algorithms, provides the analysis of video taken during manufacturing work operations, with the main purpose of assisting the ergonomist one hand in the collection of data related to postures, hand grip types, and body segments angles assumed by the worker, and on the other hand in providing insights through a portable and user-friendly interface.

The proposed architectural constraints/objectives concern:

- Modular architecture
- Simple data acquisition method with no technological constraints
- High configurability with the possibility of working with different ergonomic protocols
- Intuitive and portable user interface for results analysis

To perform the data acquisition, the proposed platform requires in input one or more video shots, retrieved by pointing a camera respectively parallel to the three anatomical planes (i.e., Sagittal Plane, Frontal Plane, and Transverse Plane). Each video must capture the operator(s) during the entire work cycle. The necessary shots can be recorded in succession, using even only one camera, as the system does not require that they are simultaneously collected.

In this way, users have at their disposal a data (i.e., video of the worker's work cycle) acquisition method that obviously increases the accuracy of the system as the number of the used cameras increases, but that does not constrain them in any case. This choice was made necessary by the fact that working environments hardly allow ergonomists to work in ideal conditions, and most of the time it is unthinkable that they will be able to recreate an environment in which the worker moves in a confined area, always staying

within the frame without changing his or her angle to the camera(s). An architectural scheme is shown in Fig. 1.

The main component, the Motion Analysis System, provides an analysis of the people framed by the camera. To this end, it performs Data Collection by means of two different modules. The first one implements the tf-pose RGB MAS system described in [5], which exploits the open-source deep learning model CMU. It enables body tracking and associates a proper ID to each person framed by the camera. The second one uses the Google Mediapipe hand tracking model [21] to enable hand landmarks recognition. Mediapipe is an APIs package for multiple programming languages that Google makes available to developers to implement image analysis functionalities by means of pre-trained Deep Learning models, including hand palm detection. This module has been added to the previous configuration for a twofold purpose: to allow the calculation of the angle assumed by the wrist and to allow the recognition of the hand grip type. Both these parameters are in fact fundamental for the evaluation of some ergonomic protocols, especially for OCRA, and were not considered in [5] which instead made use of RULA for the experimental tests.

Mediapipe can recognize from a video frame the coordinates of the landmarks shown in Fig. 2, but despite the

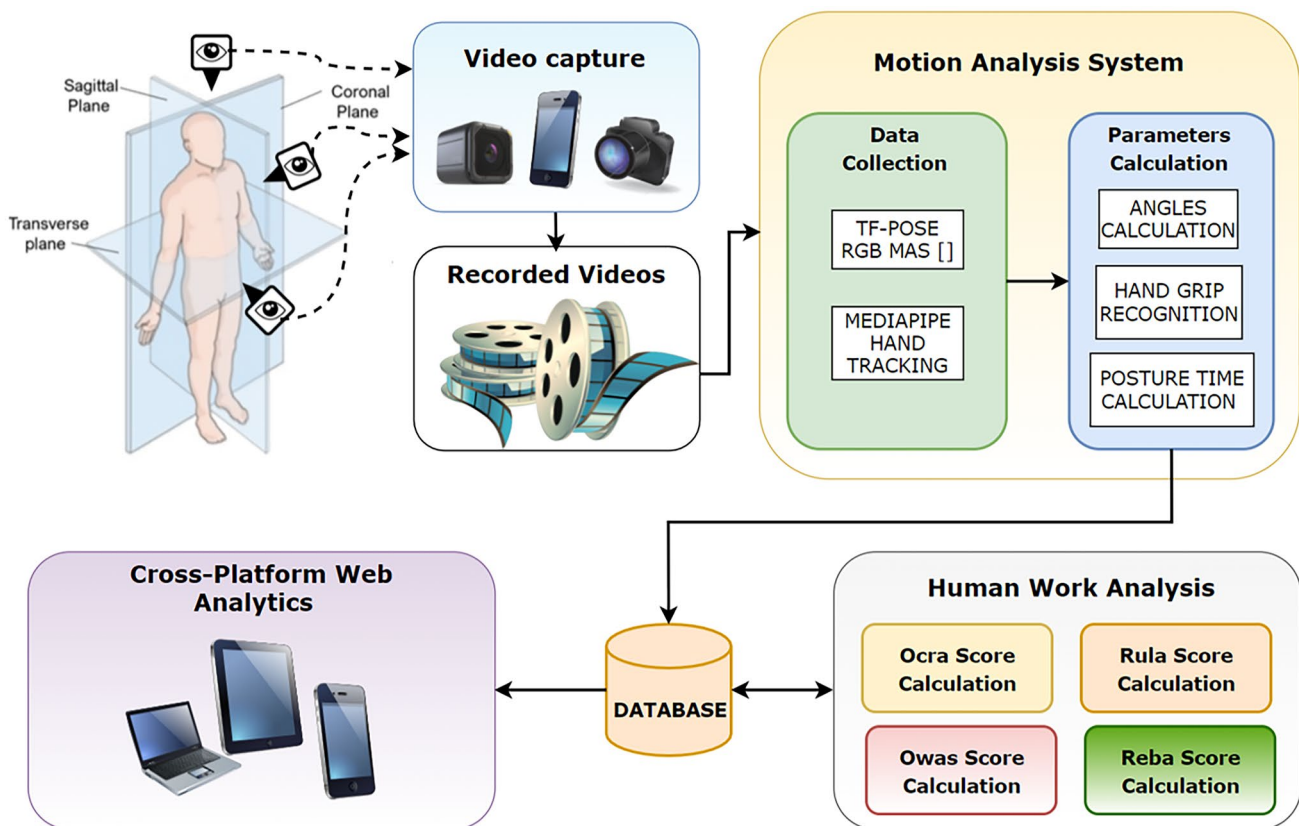
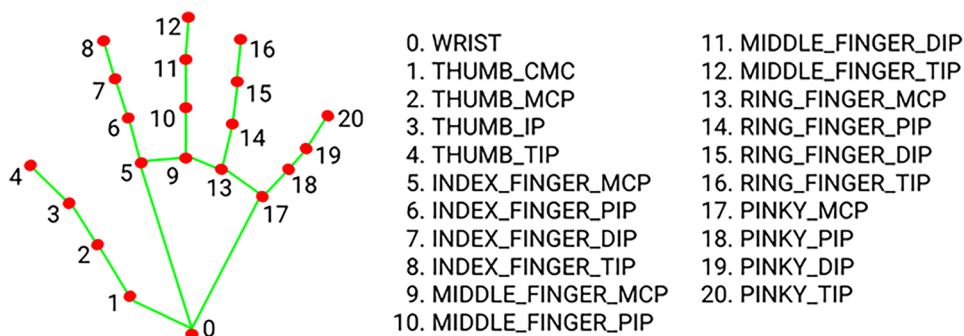


Fig. 1 The proposed platform's architecture

Fig. 2 Mediapipe recognised hand landmarks



notable performances achievable at a computational level compared to systems such as OpenPose, it suffers from other problems, first of all the difficulty of detecting a hand if it is covered by a glove. This is a problem of utmost importance in a manufacturing context where workers are forced to wear gloves. With the intention to use an off-the-shelf tool, this library is however the most performing and reliable for this kind of tasks among the existing ones.

Collected data are then processed through the Parameters Calculation module in order to compute posture angles, recognize hand grips, and determine the time the traced worker spends in each detected posture. Specifically, the calculation of angles is carried out by the algorithms described in [5], while hand grip recognition and posture time calculation are performed using solutions similar to that described in [31].

In particular, with respect to the systems described in [5] and [31], among the calculated parameters there is also the wrist angle, obtained with the calculation of the angle between the segments formed by the coordinates of elbow and wrist obtained from tf-pose and by those between wrist and middle finger MCP obtained from Mediapipe. In this case, being able to understand if the wrist is bent following a flexion or extension movement, or even if we have a radial or ulnar deviation, is quite complicated, since the hand can rotate with more degrees of freedom than the other body segments. By evaluating the position of the x-coordinate of the hand landmarks, it has been possible to discriminate the case in which the palm of the hand is directly facing the camera or it is the back of the hand that is facing the camera: in the first case, for example with the right hand, if the x-coordinate of the middle finger MCP has a greater value than that of the wrist (i.e., the middle finger MCP is more to the left of the wrist from the viewpoint of the camera) it is possible to assume that the user is making a radial deviation of the wrist; otherwise, he/she is making an ulnar deviation. The references are reversed if the back of the hand is facing the camera or if the left hand is being analyzed. Similar reasoning can be used to discriminate a flexion movement from an extension movement. Figure 3 is an excerpt of the followed algorithm, in pseudocode, for the right wrist.

Obviously, that algorithm can only work if the hand palm (or its back) is directly facing the camera. In case the system detects that the hand palm or the back is not facing the camera, (i.e., when the x,y coordinates of the middle finger MCP landmark fall in the same neighborhood, empirically derived, of the pinky finger MCP coordinates) it stops calculating wrist radial/ulnar deviation angles. By using two cameras, the problem of missing data is greatly mitigated, until it disappears altogether if three cameras are used. After all the parameters have been calculated by the Motion Analysis System, they are stored in a Database.

The system provides a specific module, the Human Work Analysis, to support the compute of ergonomic indexes respectively based on the RULA, REBA, OCRA, and OWAS checklists. Each sub-module provides the specific checklist partially filled with all the posture-related data, computed in an automatic way through the Motion Analysis System, and requires the ergonomist to insert the other data required to

```

# the palm is facing the camera
if (rightMiddleMcp_x > rightPinkyMcp_x)
{
    if (rightMiddleMcp_x < rightWrist_x)
        rightUlnarDeviation = RightWristAngle
    else
        rightRadialDeviation = RightWristAngle
}

# the back is facing the camera
if (rightMiddleMcp_x < rightPinkyMcp_x)
{
    if (rightMiddleMcp_x > rightWrist_x)
        rightUlnarDeviation = RightWristAngle
    else
        rightRadialDeviation = RightWristAngle
}

if (rightMiddleMcp_y <= rightWrist_y)
    rightFlexion = RightWristAngle
else
    rightExtension = RightWristAngle
    
```

Fig. 3 The pseudocode for the followed algorithm

complete the indexes estimation. After analyzing the entire video, this module retrieves the parameters stored so far and uses them to calculate the scores of the implemented ergonomic evaluation methods. These scores are calculated frame by frame, and only after analyzing the entire sequence a weighted average is calculated to obtain the overall score for the entire task. As an example, in the case of the RULA A and C scores, the computation is as follows:

$$A = \frac{\sum_{i=1}^N w_i A_i}{\sum_{i=1}^N w_i} \quad C = \frac{\sum_{i=1}^N w_i C_i}{\sum_{i=1}^N w_i} \quad (1)$$

where the i_{th} weight w_i is computed considering the amount of time t_i that each single i_{th} score A_i and C_i (i.e., 1,2, ..., 7)

lasted over the total task's duration t , as shown in the following formula:

$$w_i = \frac{t_i}{t} \quad (2)$$

Then, following the RULA calculation rules, the single RULA's Left and Right scores are computed for each worker, and the RULA Grandscore is the mean values of the former. A similar reasoning applies to the other implemented evaluation methods.

The system provides a web interface (Fig. 4), through which the ergonomist can perform a semi-automatic analysis of the postures found in the various tasks that characterize the work cycle and assess the risk related to MSDs. In

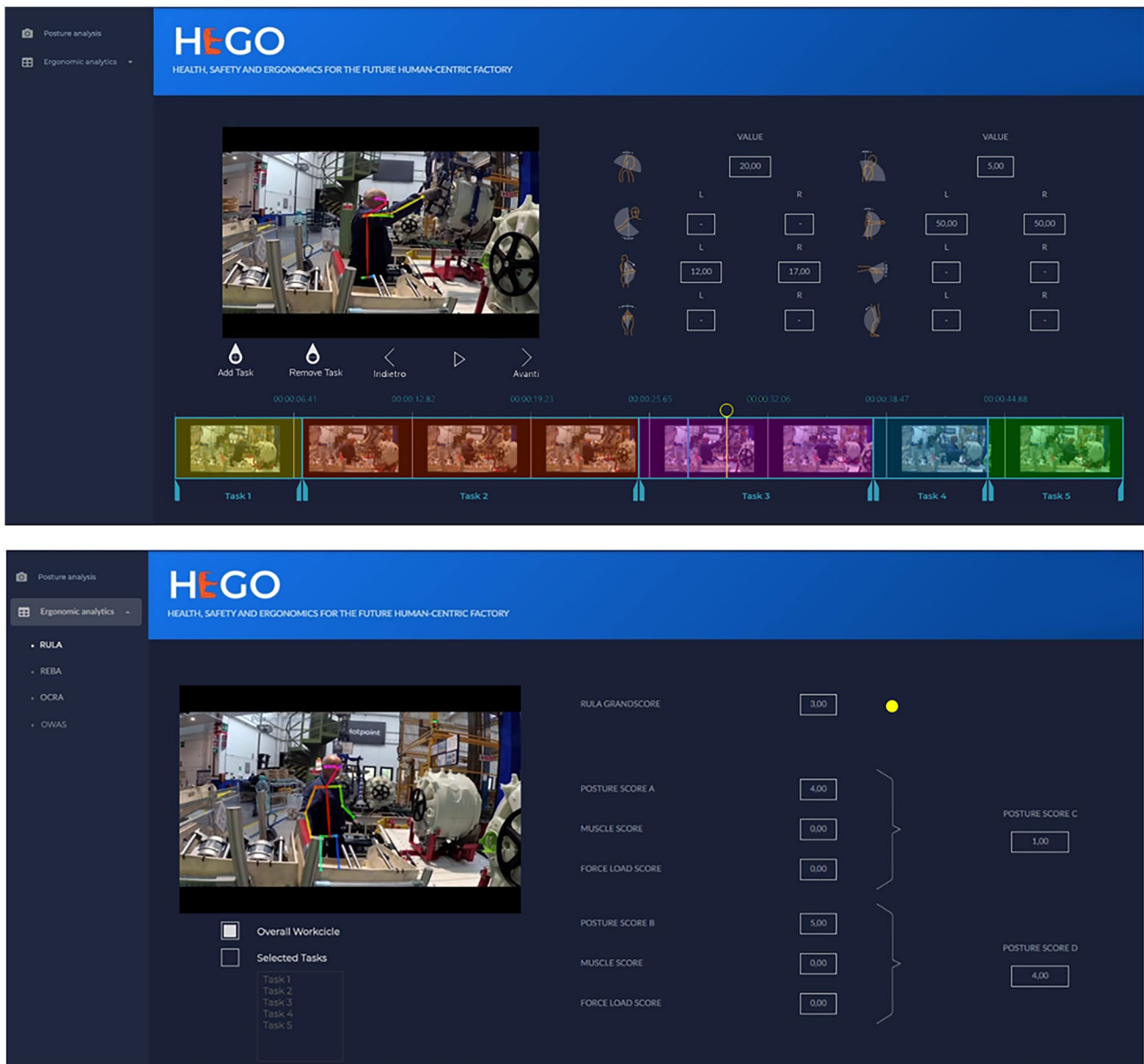


Fig. 4 The proposed platform's web interface

particular, the system provides basic video editing functionalities to enable the ergonomist to define and select the working cycle or working task to be analyzed, in a very quick and easy way.

4 Experimental case study

In order to test its effectiveness and accuracy, the proposed system has been adopted to support the assessment of the ergonomics risk, based on the RULA and OCRA indexes, of workers performing assembly operations in an assembly line of an Italian washing machine and dryer production plant, which produces 570 pieces per work shift, consisting of five workstations (Fig. 5).

The workstations, which follow one after the other during the entire manufacturing cycle, are the following:

- **Workstation 1**, the subjects (standing for the whole working cycle), with the aid of a mechanized hoist, picked up the basket of a washing machine from a conveyor belt on their left and positioned it, still suspended from the hoist, in front of them. At this point, they installed a pipe and the motor on the basket, and they fixed them with screws, using an electric screwdriver. They then placed the assembly on a conveyor belt in front of them and repositioned the hoist on the next basket on their left. They finished installing two springs and a component on the first assembly, and the working cycle began again. The total duration was 51 s
- **Workstation 2**, the subjects, sitting on a rotating chair, had a conveyor belt in front of them, running from left to right. Starting on their far left, they simultaneously tightened four screws on a washing machine basket using a special electric screwdriver, then in order they placed a belt on a pulley, checked its free rotation, and verified the tension of the belt. At this point, by rotating the basket 90° thanks to a rotary table, they installed a first pipe and then installed a second one after another 90° rotation of the basket. Finally, they tightened a screw that secured the second pipe with a suspended electric screwdriver and began the cycle again. All of this took place within 42 s
- **Workstation 3**, the subjects, standing, installed on baskets that slide in front of them, from left to right on a conveyor belt, the door seal. To do this, they took the gasket to their left and greased it, as they did on the seat where it will be installed, with the aid of a brush. Finally, they secured the gasket by placing a metal screw clamp around it, which they tightened with a suspended electric screwdriver. The cycle took 44 s
- **Workstation 4**, the subjects were standing in front of a conveyor belt that carried washing machine baskets, left to right. They took a plastic component from a box to their left, placed it on the basket, and locked it in place by tightening a screw, helping themselves with an overhead electric screwdriver. Then they repeated the same action but with a component taken from their right and secured it with three screws and a plastic band, all in 42 s



Fig. 5 The considered workstations

Table 1 RULA A and C scores resulted from automatic and manual analyses^a, $\Delta\% = -25.83\%$, $\sigma = 37.17\%$

Workstation	Automatic		Manual	
	A score	C score	A score	C score
WS1	2 (-50.00%)	3 (+50.00%)	4	2
WS2	2 (-33.33%)	3 (-25.00%)	3	4
WS3	1 (-75.00%)	3 (+0.00%)	4	3
WS4	2 (-50.00%)	4 (+0.00%)	4	4
WS5	1 (-75.00%)	3 (+0.00%)	4	3

^aThe figures in brackets next to the automatic values represent the percentage variation from the corresponding manual value

- **workstation 5**, the subjects hooked the basket of the washing machine with a mechanized hoist, but before lifting it they equipped it with two springs and a plastic component. After that, they lifted the basket and placed it inside of the frame of the washing machine, where they secured it using the two previously positioned springs. The procedure took 45 s

4.1 Experimental procedure

The experiment was carried out considering three work shifts. A total of 15 workers (11 males and 4 females), who gave their informed consent to participate, were involved.

The repetitive cycle for each operator was about 45 s.

Workers have been filmed with an iPhone XS (12 MP f/1.8 aperture), mounted on a tripod, from a single point of view: while the camera should be ideally placed on one of the subject's sides, at their pelvis height, the actual positioning was dictated from the workstation layout (even if, at some degree, the recording point was still located at the side of the subjects). This situation is the norm when performing an ergonomic analysis in a real working environment and, combined with the occlusions that naturally occur in such an environment, leads to the ideal harsh context in which to assess the system's true accuracy.

Table 2 RULA's Left, Right, and Grandscore values resulted from automatic and manual analyses^a, $\Delta\% = -8.33\%$, $\sigma = 14.57\%$

Workstation	Automatic			Manual		
	L	R	Grandscore	L	R	Grandscore
WS1	3 (+0.00%)	3 (+0.00%)	3 (+0.00%)	3	3	3
WS2	3 (+0.00%)	3 (-25.00%)	3 (+0.00%)	3	4	3
WS3	3 (+0.00%)	3 (+0.00%)	3 (+0.00%)	3	3	3
WS4	3 (-40.00%)	4 (-20.00%)	3 (-40.00%)	5	5	5
WS5	3 (+0.00%)	3 (+0.00%)	3 (+0.00%)	3	3	3

^aThe figures in brackets next to the automatic values represent the percentage variation from the corresponding manual value

RULA score variability comparison

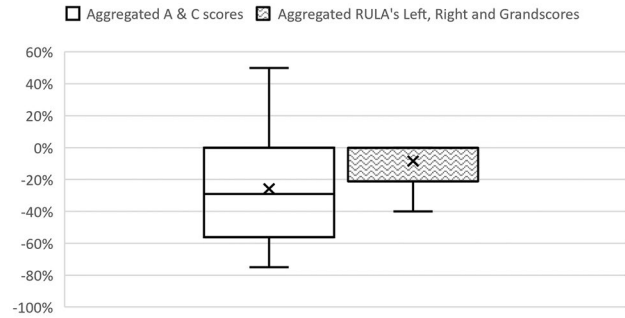


Fig. 6 Comparison of the variability between the RULA A and C partial scores and the RULA's Left, Right, and Grand score

4.2 Data collection

Angles were extracted from the video recordings either automatically by the proposed tool and manually from still frames by a panel of expert ergonomists. The tool was running on a PC workstation with an Intel(R) Core (TM) i7-7700 K CPU at 4.20 GHz and 32 GB RAM, and a GTX 1080 Ti GPU, running Windows 10 Pro. Once collected, angles were processed by the proposed tool, running on the same PC workstation just mentioned. To avoid possible differences in results given by approximation errors, the system has been used to compute RULA and OCRA scores either considering the angle automatically extracted, and those manually measured through video analysis.

5 Results

For each of the workstations (WS), the score is calculated as the median between the scores obtained by the three workers involved in that WS. The partial scores of the experimentation for the RULA (A and C scores) either for automatic

and manual angles extraction are reported in Table 1, while Table 2 shows the final RULA's Left, Right, and Grand score.

In addition, for each of the following categories of automatic scores obtained:

- aggregated A & C scores
- aggregated RULA Left, Right, and Grand score

for every *i*th score, the percentage variation ($\Delta\%$) from the corresponding manual score was calculated as in Eq. (3), followed by the mean values $\Delta\%$ and the standard deviations σ (Eq. 4) that were calculated for each category:

$$\Delta\%_i = \frac{(V_{ai} - V_{mi})}{V_{mi}} \cdot \% \tag{3}$$

$$\sigma = \sqrt{\frac{\sum_{i=1}^N (\Delta\%_i - \overline{\Delta\%})^2}{N}} \tag{4}$$

with:

- V_{ai} is the *i*th automatic value.
- V_{mi} is the *i*th manual value.
- N is the total number of items in that category.

Table 3 OCRA upper limbs incongruous postures scores^a, $\Delta\% = -24.56\%$, $\sigma = 17.65\%$

			Score I ^b	Score II ^c	Score III ^d	Score IV ^e
WS1	A	L	5.40 (-40.00%)	5.80 (-3.33%)	17.40 (-13.00%)	1.40 (-30.00%)
		R	13.40 (-33.00%)	7.90 (-12.22%)	19.40 (-19.17%)	2.10 (-47.50%)
	M	L	9.00	6.00	20.00	2.00
		R	20.00	9.00	24.00	4.00
WS2	A	L	20.59 (-29.00%)	2.04 (+2.00%)	8.60 (-14.00%)	1.30 (-35.00%)
		R	23.39 (-31.00%)	2.10 (+5.00%)	21.24 (-10.00%)	3.05 (-39.00%)
	M	L	29.00	2.00	10.00	2.00
		R	33.90	2.00	23.60	5.00
WS3	A	L	14.58 (-39.00%)	1.64 (-18.00%)	16.48 (+3.00%)	2.52 (-37.00%)
		R	15.84 (-45.00%)	1.52 (-24.00%)	17.75 (-9.00%)	1.71 (-43.00%)
	M	L	23.90	2.00	16.00	4.00
		R	28.80	2.00	19.50	3.00
WS4	A	L	13.66 (-46.00%)	18.31 (+9.00%)	14.75 (+1.00%)	2.24 (-44.00%)
		R	8.36 (-45.00%)	15.66 (-11.00%)	22.68 (-16.00%)	3.85 (-45.00%)
	M	L	25.30	16.80	14.60	4.00
		R	15.20	17.60	27.00	7.00
WS5	A	L	13.66 (-46.00%)	2.00 (+0.00%)	10.24 (-11.00%)	1.40 (-44.00%)
		R	8.03 (-27.00%)	1.64 (-18.00%)	18.72 (-22.00%)	2.64 (-56.00%)
	M	L	25.30	2.00	11.50	2.50
		R	11.00	2.00	24.00	6.00

A stands for Automatic, M stands for Manual

^aThe figures in brackets next to the automatic values represent the percentage variation from the corresponding manual value

^bHand in pinch grip or handheld or hook posture

^cArm almost at shoulder height or other incongruous postures

^dExtreme deviation of wrist in flexion and/or in radio/ulnar deviation

^eComplete rotation of objects and/or large elbow flexions-extensions

Table 4 OCRA partial and final indexes^a. Partial index: $\Delta\% = -9.88\%$, $\sigma = 6\%$. Final index: $\Delta\% = -10.74\%$, $\sigma = 8\%$

Workstation	Index	Automatic		Manual	
		L	R	L	R
WS1	Partial index	4.00 (-16.67%)	8.00 (-11.11%)	4.80	9.60
	Final index	4.80 (-17.24%)	9.6 (-11.11%)	5.80	10.80
WS2	Partial index	11.18 (-14.00%)	13.00 (+0.00%)	13.00	13.00
	Final index	11.99 (-19.00%)	15.54 (+5.00%)	14.80	14.80
WS3	Partial index	8.10 (-19.00%)	11.88 (-1.00%)	10.00	12.00
	Final index	9.96 (-17.00%)	14.69 (+2.00%)	12.00	14.40
WS4	Partial index	10.44 (-13.00%)	12.96 (-4.00%)	12.00	13.50
	Final index	11.81 (-18.00%)	15.07 (-7.00%)	14.40	16.20
WS5	Partial index	12.09 (-7.00%)	6.09 (-13.00%)	13.00	7.00
	Final index	13.73 (-12.00%)	7.31 (-13.00%)	15.60	8.40

^aThe figures in brackets next to the automatic values represent the percentage variation from the corresponding manual value

As it can be seen from Table 1, there is quite a remarkable percentage variation in the automatic values compared to the manual reference ones, with a mean value $\Delta\% = -25.83\%$ and a standard deviation $\sigma = 37.17\%$. However, looking at Table 2, the RULA’s Left, Right, and Grandscores suffer of a lower variability, with the mean value $\Delta\% = -8.33\%$ and a standard deviation $\sigma = 14.57\%$. A visual comparison of the RULA scores variability can be found in Fig. 6.

OCRA scores are computed for each WS, considering the median score collected for the analyzed workers. The estimated time cycle considered for the OCRA scores computation is the ones reported above in the workstations’ description.

Finally, for each of the following categories of automatic scores obtained:

- aggregated OCRA upper limbs incongruous posture scores
- OCRA Partial (OP) indexes
- OCRA Final (OF) indexes

The same percentage variation and standard deviation dealt with in the RULA paragraph are calculated.

All the results are reported in Tables 3 and 4. Table 3 shows the scores obtained through manual and automatic assessment with regard to OCRA upper limbs incongruous postures score. The outcomes clearly demonstrate strong variability in the results, as evidenced by the percentage changes in the automatic scores from the corresponding manual values (i.e., the values shown in parentheses next to the scores). The mean value of these variations is $\Delta\% = -24.56\%$, with a standard deviation amounting to $\sigma = 17.65\%$.

However, looking at Table 4, where the Partial and Final OCRA indexes are shown, we get a better picture. In fact, the variations are limited: the mean value of the variation for the OCRA Partial index is $\Delta\% = -9.88\%$, and the standard deviation is $\sigma = 6\%$, while for the OCRA Final index we have $\Delta\% = -10.74\%$ and $\sigma = 8\%$. A quick visual comparison of the variability of the OCRA incongruous posture scores and the OCRA Partial and Final indexes can be found in Fig. 7.

Again, the system provides an automatic score that is more than sufficient to conduct a reliable ergonomic risk analysis.

OCRA score variability comparison

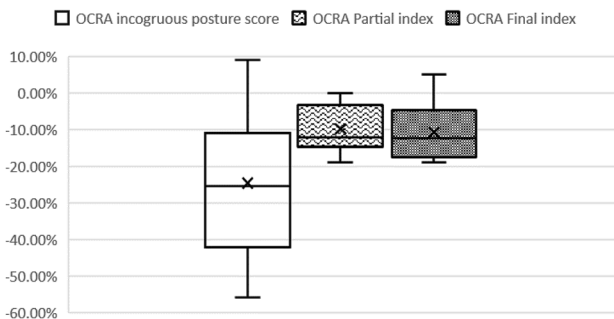


Fig. 7 Comparison of the variability between the OCRA upper limbs incongruous posture score and the OCRA partial

6 Discussion and conclusion

In this paper, a new semi-automatic ergonomic risk assessment system, with its web interface, is presented. Experimentation in a real manufacturing context showed that the system is robust enough to withstand the severe condition that may occur in such an environment (e.g., occlusion, varying lighting conditions, sub-optimal location of the acquisition cameras). This confirmed the results reported in previous studies, which compared the performances of other RGB-based MoCap systems, similar to that here considered

(e.g., OpenPose), with other body tracking techniques, such as Microsoft Kinect [1] or Vicon Nexus [5].

The ergonomists using it may benefit from a significant decrease in the time it takes to perform an ergonomic risk assessment: while manually determining the posture angles from video analysis required on average an hour per workstation, the system demonstrated to be able to provide, within a minute, automatic scores that mostly match those supplied by expert ergonomists, so that it seems reliable enough to support an ergonomic risk analysis, at least as far as RULA and OCRA are concerned. However, the system evidenced some limitations mainly due to the hand tracking and gesture recognition pipeline. In particular, the hand tracking system (based on Google Mediapipe hand recognition model) highlighted severe problems in recognizing gloved hands. The shape and size of the gloves seem not to influence the results: problems happen even if very tight and thin gloves are used. Given that in a manufacturing environment almost every worker wears gloves, this can be a severe limitation.

This hand recognition problem may have affected RULA scores (in particular, the hand and wrist analysis sections) although the final results are almost matching those provided by expert ergonomists, but it strongly affected OCRA results related to the hand. In fact, the large discrepancy between the times estimated by the system and by the experts is mainly due to the large number of frames in which the system failed to recognize the hand. Based on the best of our knowledge, there is no tool yet able to recognize and track landmarks hands, in a robust way, even when they are covered with gloves. Further studies that aim to overcome the limitations of current tools are needed. One possible solution could be the use of only the hand landmarks recognition model from Mediapipe, while the palm detection (the task of recognizing a hand within a frame which takes place before landmarks recognition) is performed using tf-pose-estimation: the idea is to retrieve the landmark coordinates relative to the wrist and crop the image on the space around it. The image will then be passed to Mediapipe. In this way, accuracy in recognizing hand coordinates even when gloves are worn could be greatly increased.

Another problem that causes variability in score estimation is that the proposed system is not always capable to correctly predict angles from body joints with 2 or more degrees of freedom with non-calibrated cameras, e.g., when the arms are simultaneously showing flexion in the lateral plane and abduction in the frontal plane. In fact, at the present time the proposed system is not able to manage spatial correlation between two or more cameras, so as it does not implement solution to compensate perspective error due to camera positioning. The situation gets worse when only a camera is used, as in the case studied in this paper.

However, since the percentage variation between the automatic RULA and OCRA final scores with respect to the

corresponding manual ones is, respectively, $\Delta\% = -8.33\%$ and $\Delta\% = -10.74\%$ and therefore very contained, one can argue that the inability to provide accurate hands assessment of a motion capture acquisition tool does not compromise the validity of the ergonomic risk assessment. Despite the proposed system demonstrating a mean error equal to 10% in the compute of final indexes compared to the manual analysis, no significant differences have been found in terms of risk class estimation between ergonomic assessments performed by an expert ergonomist, with or without the support of the system. This should make us reflect on the limits of the considered ergonomic risk indexes (i.e., RULA and OCRA), and on the usefulness and opportunity of an ergonomic analysis carried out only by means of observational methods. Since they provide an estimation of risk based on classes, they have the tendency to equalize different situations placed at the extremes of the same class, while at the same time diversify two situations that are very similar but placed very close to the classes' boundaries. Probably, the integration with other tools that include assessing the psychological aspects of the job (e.g., Quick Exposure Check [32]) or with questionnaires (e.g., Standardized Nordic Musculoskeletal Questionnaire [33]) can provide different perspectives from which deeply analyze the results of an ergonomic assessment [34].

Future works should consider extending the features of such system to support the workers themselves, somehow replacing the ergonomist in following them during their daily activities. The proposal should be a tool that not only performs ergonomic risk analyses, but also merges with the work environment and becomes an integral part of it, much like a screwdriver.

The solution identified would therefore constantly monitor workers, alerting and correcting them in real time if they assume incorrect postures. Moreover, it could be implemented in novel or already existing systems that aid the worker, by showing them the assembly steps that they must perform, like the systems developed in [35, 36].

Finally, this solution allows ergonomic risks related to all workstations to be assessed in a tight time frame, opening the possibility to constantly monitoring the operators working along a production line, and to collect a large amount of historic data. Such data would have a great impact in the workstations' design phase. The availability of historic data about results of many ergonomic analyses can be useful to better guide the improvement of the workstations or to support the design of innovative solutions, in order to attempt the reduction of work-related musculoskeletal disorders. In particular, the design can be carried out at the level of the worker-station coupling, going to customize the workstation on the needs of the individual. However, this possibility is currently hampered by the difficulty of providing a frequent ergonomic risk monitoring without incurring in huge

costs. Moreover, in a smart manufacturing context, Machine Learning and AI techniques could be employed to analyze this data history, with the aim of optimizing shifts and task assignments to the workers, based on their characteristics and skills.

Author contribution The authors contributed equally to this work.

Funding Open access funding provided by Università Politecnica delle Marche within the CRUI-CARE Agreement. This project has been funded by Marche Region in implementation of the financial program POR MARCHE FESR 2014–2020, project “Miracle” (Marche Innovation and Research Facilities for Connected and sustainable Living Environments), CUP B28119000330007, and has been supported by the EMOJ srl startup within the program “HEGO: a novel enabling framework to link health, safety and ergonomics for the future human-centric factory toward an enhanced social sustainability,” POR MARCHE FESR 2014–2020 – ASSE 1 – OS 1 – AZIONE 1.1. INT 1.1.1.

Declarations

Competing interests The authors declare no competing interests.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Kim W, Sung J, Saakes D, Huang C, Xiong S (2021) Ergonomic postural assessment using a new open-source human pose estimation technology (OpenPose). *Int J Ind Ergon* 84:103164
- Chaffin DB, Peacock B (2001) Digital human modeling for vehicle and workplace design. Society of Automotive Engineers, Warrendale, PA
- Lowe BD, Dempsey PG, Jones EM (2019) Ergonomics assessment methods used by ergonomics professionals. *Appl Ergon* 81:102882
- Peruzzini M, Grandi F, Cavallaro S, Pellicciari M (2021) Using virtual manufacturing to design human-centric factories: an industrial case. *Int J Adv Manuf Technol* 115(3):873–887
- Agostinelli T, Generosi A, Ceccacci S, Khamaisi RK, Peruzzini M, Mengoni M (2021) Preliminary validation of a low-cost motion analysis system based on RGB cameras to support the evaluation of postural risk assessment. *Appl Sci* 11(22):10645
- Mehta D, Sotnychenko O, Mueller F, Xu W, Elgharib M, Fua P, Seidel HP, Rhodin H, Pons-Moll G, Theobalt C (2019) Xnect: real-time multi-person 3d human pose estimation with a single rgb camera. arXiv preprint [arXiv:1907.00837](https://arxiv.org/abs/1907.00837)
- Kim IJ (2017) The role of ergonomics for construction industry safety and health improvements. *J Ergon* 7(2):2–5
- Punnett L, Wegman DH (2004) Work-related musculoskeletal disorders: the epidemiologic evidence and the debate. *J Electromyogr Kinesiol* 14(1):13–23
- McAtamney L, Corlett EN (1993) RULA: a survey method for the investigation of work-related upper limb disorders. *Appl Ergon* 24(2):91–99
- Hignett S, McAtamney L (2000) Rapid entire body assessment (REBA). *Appl Ergon* 31(2):201–205
- Karhu O, Kansi P, Kuorinka I (1977) Correcting working postures in industry: a practical method for analysis. *Appl Ergon* 8(4):199–201
- Ochipinti E (1998) OCRA: a concise index for the assessment of exposure to repetitive movements of the upper limbs. *Ergonomics* 41(9):1290–1311
- Statista. Available online. <https://www.statista.com/statistics/539395/smartphone-penetration-worldwide-by-country/>. Accessed 20 Jan 2022
- Battini D, Persona A, Sgarbossa F (2014) Innovative real-time system to integrate ergonomic evaluations into warehouse design and management. *Comput Ind Eng* 77:1–10
- Huang C, Kim W, Zhang Y, Xiong S (2020) Development and validation of a wearable inertial sensors-based automated system for assessing work-related musculo-skeletal disorders in the workspace. *Int J Environ Res Public Health* 17(17):6050
- Vignais N, Miezal M, Bleser G, Mura K, Gorecky D, Marin F (2013) Innovative system for real-time ergonomic feedback in industrial manufacturing. *Appl Ergon* 44(4):566–574
- Manghisi VM, Uva AE, Fiorentino M, Gattullo M, Boccaccio A, Evangelista A (2020) Automatic ergonomic postural risk monitoring on the factory shopfloor—the ergosentinel tool. *Procedia Manuf* 42:97–103
- Abobakr A, Nahavandi D, Hossny M, Iskander J, Attia M, Nahavandi S, Smets M (2019) RGB-D ergonomic assessment system of adopted working postures. *Appl Ergon* 80:75–88
- Plantard P, Auvinet E, Pierres ASL, Multon F (2015) Pose estimation with a kinect for ergonomic studies: evaluation of the accuracy using a virtual mannequin. *Sensors* 15(1):1785–1803
- Cao Z, Simon T, Wei SE, 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
- TF-Pose. Available online. <https://github.com/tryagainconcepts/tf-pose-estimation>. Accessed 20 Jan 2022
- MediaPipe. Available online. <https://google.github.io/mediapipe/>. Accessed 20 Jan 2022
- Ota M, Tateuchi H, Hashiguchi T, Kato T, Ogino Y, Yamagata M, Ichihashi N (2020) Verification of reliability and validity of motion analysis systems during bilateral squat using human pose tracking algorithm. *Gait Posture* 80:62–67
- He Ling WANG, Yun-Ju LEE (2019) Occupational evaluation with Rapid Entire Body Assessment (REBA) via imaging processing in field. In: Proceedings of the Human Factors Society Conference. Elsinore, Denmark, pp 25–28
- Li L, Martin T, Xu X (2020) A novel vision-based real-time method for evaluating postural risk factors associated with musculoskeletal disorders. *Appl Ergon* 87:103138
- Da Silva Neto JG, Teixeira JMXN, Teichrieb V (2020) Analyzing embedded pose estimation solutions for human behaviour understanding. In: Anais Estendidos do XXII Simpósio de Realidade Virtual e Aumentada. SBC: Porto Alegre, Brasil, pp 30–34
- Lindera. Available online. <https://www.lindera.de/technologie/>. Accessed 20 Jan 2022

28. Zhang F, Bazarevsky V, Vakunov A, Tkachenka A, Sung G, Chang CL, Grundmann M (2020) Mediapipe hands: on-device real-time hand tracking. arXiv preprint [arXiv:2006.10214](https://arxiv.org/abs/2006.10214)
29. Chundururu V, Roy M, Chittawadigi RG (2021) Hand tracking in 3d space using mediapipe and pnp method for intuitive control of virtual globe. In: 2021 IEEE 9th Region 10 Humanitarian Technology Conference (R10-HTC). IEEE, pp 1–6
30. Halder A, Tayade A (2021) Real-time vernacular sign language recognition using mediapipe and machine learning. *Int J Res Anal Rev* 2(5):9–17
31. Altieri A, Ceccacci S, Talipu A, Mengoni M (2020) A low cost motion analysis system based on RGB cameras to support ergonomic risk assessment in real workplaces. In: International Design Engineering Technical Conferences and Computers and Information in Engineering Conference, vol 83983. American Society of Mechanical Engineers, p V009T09A067
32. David G, Woods V, Li G, Buckle P (2008) The development of the Quick Exposure Check (QEC) for assessing exposure to risk factors for work-related musculoskeletal disorders. *Appl Ergon* 39(1):57–69
33. Kuorinka I, Jonsson B, Kilbom A, Vinterberg H, Biering-Sørensen F, Andersson G, Jørgensen K (1987) Standardised Nordic questionnaires for the analysis of musculoskeletal symptoms. *Appl Ergon* 18(3):233–237
34. Ayub Y, Shah ZA (2018) Assessment of work related musculoskeletal disorders in manufacturing industry. *J Ergon* 8(3):1–5
35. Mengoni M, Ceccacci S, Generosi A, Leopardi A (2018) Spatial Augmented Reality: an application for human work in smart manufacturing environment. *Procedia Manuf* 17:476–483
36. Werrlich S, Nitsche K, Notni G (2017) Demand analysis for an augmented reality based assembly training. In: Proceedings of the 10th International Conference on PErvasive Technologies Related to Assistive Environments, pp 416–422

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.