The International Journal of Advanced Manufacturing Technology
https://doi.org/10.1007/s00170-023-12506-7

ORIGINAL ARTICLE

Visual analysis of fatigue in Industry 4.0

David Alfavo-Viquez[1] · Mauricio-Andres Zamora-Hernandez[1] · Jorge Azorín-López[2] · Jose Garcia-Rodriguez[2]

Received: 3 November 2022 / Accepted: 10 October 2023

Abstract

The performance of manufacturing operations relies heavily on the operators' performance. When operators begin to exhibit signs of fatigue, both their individual performance and the overall performance of the manufacturing plant tend to decline. This research presents a methodology for analyzing fatigue in assembly operations, considering indicators such as the EAR (Eye Aspect Ratio) indicator, operator pose, and elapsed operating time. To facilitate the analysis, a dataset of assembly operations was generated and recorded from three different perspectives: frontal, lateral, and top views. The top view enables the analysis of the operator's face and posture to identify hand positions. By labeling the actions in our dataset, we train a deep learning system to recognize the sequence of operator actions required to complete the operation. Additionally, we propose a model for determining the level of fatigue by processing multimodal information acquired from various sources, including eye blink rate, operator pose, and task duration during assembly operations.

Keywords Automatic control · Computer Vision · Deep Learning · Fatigue recognition

1 Introduction

Fatigue in manufacturing processes is a crucial factor that warrants analysis as it leads to a decline in operator performance during their daily work. This decline in performance not only affects the operators themselves but also negatively impacts the overall productivity of the manufacturing process.

The objective of this work is to contribute to the research on fatigue by creating a dataset comprising assembly operations of a simple manufacturing process. Multiple operators were recorded performing the same process under different levels of fatigue.

The recorded videos capture the assembly operations from three different perspectives: top, front, and side views. The front view enables the detection of fatigue through facial cues such as facial expressions, eye blinking, and ear movements. The top view allows for the analysis of hand movements and action recognition during the assembly process. The side view facilitates the evaluation of operator posture, including bending and the assessment of non-ergonomic postures. This dataset can be utilized in investigating work design, as it allows for the measurement of time taken to perform specific manufacturing actions through the recorded videos. Each video in the dataset is accompanied by corresponding labels for actions and objects.

The main contribution of this work is the proposal of a model that determines the fatigue level of an operator by processing multimodal data. This model evaluates the flicker rate, operator posture, and time spent performing various assembly operations.

The subsequent sections of this research paper are organized as follows: Section 2 provides a review of related works, including literature on fatigue in manufacturing environments, automatic process monitoring, and action recognition. Section 3 describes the proposed system architecture for fatigue analysis. Section 4 presents the details of the dataset

David Alfavo-Viquez, Mauricio-Andres Zamora-Hernandez, Jorge Azorín-López and Jose Garcia-Rodriguez are authors contributed equally to this work.

✉ Jose Garcia-Rodriguez
jgarcia@dtic.ua.es

David Alfavo-Viquez
david.alfaro@ucr.ac.cr

Mauricio-Andres Zamora-Hernandez
mauricio.zamorahernandez@ucr.ac.cr

Jorge Azorín-López
jazorin@ua.es

1 University of Costa Rica, San José, Costa Rica

2 University of Alicante, Alicante, Spain

Published online: 02 December 2023

created for the training process of assembly activities, allowing for the identification of fatigue and other manufacturing elements. Section 5 analyzes the experimental design and showcases the final product of the assembly. Finally, the last section presents the conclusions drawn from the research and outlines potential future lines of investigation.

2 Related Work

In this section, we review scientific articles that examine the impact of fatigue in production environments, including research on human factors affecting manufacturing processes and the development of automatic monitoring systems.

Several studies have investigated the consequences of fatigue on worker performance, productivity, and safety. Fatigue has been found to diminish cognitive abilities, leading to an increase in errors and a reduction in the efficiency of manufacturing tasks.

Human factors, which focus on how ergonomics influences operators, productivity, and safety, play a significant role in the efficiency of manufacturing processes. For instance, the design of ergonomically optimized workstations can reduce physical strain and fatigue, thereby enhancing physical well-being and worker productivity.

In recent years, automatic monitoring systems have emerged as valuable tools for fatigue detection in manufacturing environments. These systems employ computer vision and machine learning techniques to analyze worker behavior and identify early signs of fatigue. By continuously monitoring vital parameters and behavioral patterns, they can provide timely alerts and interventions.

Digital twin technology has shown promise in the field of human factors and fatigue studies. Sharotry [1] employed digital twin technology to assess fatigue in manufacturing operations. The study utilized the dynamic time warping (DTW) algorithm to analyze changes in joint angles. By using a digital twin, it becomes possible to evaluate the presence of fatigue in operators. The research findings suggest that fatigue can occur in different groups of joints in individuals, necessitating the adjustment of digital twin configurations to reflect these characteristics.

Virtual reality (VR) technology has also been employed in the study of workspace ergonomics. VR allows for the creation of virtual workspaces that simulate real-life working conditions, enabling the evaluation of a person's movements. Grajewski [2] discussed the use of virtual reality in developing two workstation models: a soldering station and a drill bench.

Within the field of human factors, fatigue is a sub-area of particular importance, as it informs the design of workstations and processes. In the present research, we will further explore the analysis of fatigue in automatic control systems, as discussed in the following section.

2.1 Fatigue in Production Environments

Berti et al. [3] emphasized the role of ergonomics in detecting fatigue in operator performance during manufacturing tasks. They proposed a model that considers fatigue as a cumulative factor that increases after completing each task. As operators accumulate fatigue, their performance gradually declines, leading to a reduction in the overall throughput of the production process.

Lambay [4] conducted research confirming that increased operator fatigue directly correlates with decreased efficiency in the production process. The study of human factors revealed that fatigue levels can vary among different work centers and operators. Factors such as workloads, individual fatigue status, and demographic characteristics were taken into account. A machine learning algorithm was developed to predict fatigue levels in operators, introducing a new method for fatigue level prediction.

Li [5] designed a helmet equipped with inertial measurement sensors and EEG (electroencephalography) to determine operator fatigue. The helmet includes a buzzer to alert the operator and sends a signal for fatigue detection, which can prompt the stopping of processes or machines.

Facial markers have been employed to estimate fatigue in individuals, particularly in drivers. Studies by Savas et al. [6] and Zhu et al. [7] focused on facial markers such as the percentage of eye-opening frequency (PERCLOS) and the mouth aspect ratio (MAR) to assess fatigue levels.

In related research on fatigue monitoring in industrial environments, various sensors and techniques have been employed. For instance, Lambay (2021) mentions the use of invasive techniques like EMG sensors, as well as non-invasive techniques such as computer vision, for fatigue monitoring [8].

Regarding the utilization of deep learning techniques for fatigue monitoring, it is common to find the application of recurrent neural networks (RNNs) for fatigue prediction, as demonstrated in the work by Lambay (2021) [8]. It is worth noting that studies such as Escobar-Linero (2022) and Lambay (2021) trained their systems using datasets obtained from IMU sensors [8, 9].

Our proposed architecture is based on three-dimensional convolutional neural networks (3D CNNs) for training a CNN model for action recognition. This type of CNN effectively captures temporal information in video sequences.

The 3D CNN neural network comprises multiple 3D convolutional layers, followed by clustering layers and fully connected layers. Each 3D convolutional layer performs convolutions in the spatial and temporal dimensions of the input volume, capturing relevant features at each frame and throughout the video sequence. Clustering layers reduce dimensionality and extract essential features, while fully connected layers classify the learned features into different action classes.

2.2 Automated Inspection Systems

A crucial aspect of our research involves the use of automated inspection systems employing computer vision to analyze the influence of fatigue on inspection processes in Industry 4.0.

In manufacturing, inspections are commonly performed either manually or through automated processes. The use of computer vision has significantly accelerated these tasks in various industries. Lukinac et al. [10] implemented a computer vision system for quality control in the beer industry, focusing on characteristics such as physical appearance and fermentation. This non-destructive inspection process has proven valuable for the food industry.

Villalba-Diez and Kazemian [11] discuss vision systems for monitoring product quality in manufacturing environments, specifically for quality monitoring in printing systems. They developed a computer vision algorithm to implement a closed-loop extrusion system, using recorded videos to segment and analyze each element of the extrusion system.

Reich et al. [12] configured a vision system to monitor a production process and determine the demands on staff time and resources. They proposed a programming block ontology that reduced the implementation time of the vision system by allowing the representation of a computer vision algorithm.

Deep learning techniques have been employed to enhance automated inspection processes. Riedel et al. [13] developed a modular assembly assistant for real-world products using RGB camera detection to reduce errors. Other researchers, such as Chang et al. [14], Cheng et al. [15], Denkena et al. [16], Kousi et al. [17], Tao and Zamora-Hernandez [18, 19], proposed intelligent assistants that utilize computer vision or data fusion to enable real-time interaction and control in assembly processes.

Furthermore, Böllhoff et al. [20], Denkena et al. [16], Qeshmy et al. [21], and Riedel et al. [13] specialized their assistants for manufacturing, focusing on operator interaction with visual systems that can detect or prevent errors, improve the assembly process, or enhance operator protection.

Regarding the inspection of the operators and the interaction with objects, deep learning techniques have facilitated the development of applications that can identify both. The recognition of actions is an ongoing topic in computer vision research. Perera et al. [22] are working on developing datasets that enable the development of action recognition systems.

Varol et al. [23] utilize synthetic videos for activity tracking and recognition. They reconstruct a 3D model of the human body to generate synthetic videos and label the performed activities. They also introduce a novel methodology for data generation, enabling the training of spatiotemporal CNNs (Convolutional Neural Networks) for action classification.

Jones et al. [24] propose the use of a Latent Convolutional Skip Chain Conditional Random Field (LC-SC-CRF) time series model, which can learn a set of primitive actions based on interpreted sensor data, such as accelerometers.

Action recognition is vital for applications such as robot interaction, action automation, and surveillance systems. Wang et al. [25] emphasize the importance of timely knowledge of the operating context to improve safety and efficiency in human-robot collaboration. They utilize machine learning, specifically a deep convolutional network adapted from AlexNet, for continuous motion analysis in human-robot collaboration and future predictions of movements. Their experiment, focused on a motor assembly task, achieved a 96% accuracy in action recognition.

3 Manufacturing Fatigue Recognition Architecture

This paper introduces an arquitecture designed to monitor and analyze fatigue levels in individuals during work activities. The primary objective of this arquitecture is to enhance the well-being of operators, optimize their performance, and effectively address fatigue-related concerns. The system is based on the architecture of the Fig.1. It comprises multiple interconnected modules, including the Action Recognition Module, Pose Recognition Module, Face Recognition Module, and Fatigue Evaluation Module. These modules collaborate to detect indicators of fatigue, facilitate smooth data transfer, offer decision support, present analyzed data in a user-friendly format, and enable seamless integration with other systems. Figure 1 illustrates the proposed architecture.

- **Action Recognition Module**.This module is designed to identify the actions carried out by the operator during their work. It relies on a top view camera positioned on the workbench to capture the necessary data. By utilizing this perspective, the module is capable of recognizing tools, objects, hands, and their corresponding actions on the workbench. To train the system, a dataset was created involving both expert and inexperienced users who performed assembly actions. This dataset was utilized to train the module effectively. Once deployed, this module

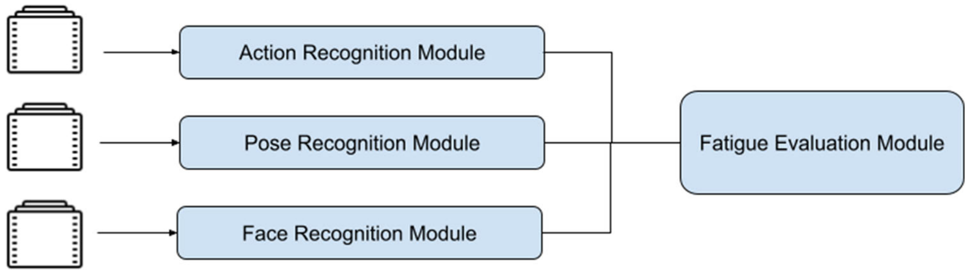

Fig. 1 This figure shows the Components of the System Proposal

is able to identify the actions performed by the user and determine if any deviations from the prescribed procedure have occurred. Moreover, it actively monitors the production quality and provides warnings in the event of potential errors.

- **Pose Estimation** Module. This module serves the purpose of posture recognition for users during their work. It utilizes a camera positioned at the side view to facilitate this task. The module has been trained on videos captured from the side view, which are annotated with the individual's skeletal structure while they are engaged in work activities. As a result, it is capable of assessing and monitoring posture changes throughout the workday. Evaluating posture is instrumental in identifying indicators of fatigue.
- **Blink Rate Estimation Ratio** Module.This module is specifically designed to recognize the face of the individual engaged in the activity. It leverages the video feed captured by a camera positioned in front of the operator on the worktable. By doing so, the system is able to accurately identify the frequency of the person's blinking while they are working. One of the fundamental premises of this research is that the blink frequency increases over time, serving as an indicator of fatigue resulting from the workload. It is important to note that this study solely focuses on evaluating physical fatigue, while aspects such as mental fatigue are not considered. The primary objective of this module is to assess the blinking patterns as a means to gauge the level of fatigue experienced by the individual during the task, thus providing insights into their workload-related tiredness. It was found that the Eye Aspect Ratio (EAR) is a widely used indicator to analyze the presence of fatigue in drivers. In this work, we propose to utilize this EAR indicator to analyze fatigue in manufacturing operations, specifically in assembly tasks. The system measures the eye-opening ratio of the operator as an indicator. To measure the EAR, we utilize landmarks of the eye, as shown in Figure 2. By analyzing the information extracted from these landmarks, our system can detect signs of fatigue in the operator.
- **Fatigue Evaluation Module**. The information obtained from the aforementioned modules is utilized within the Fatigue Evaluation Module to determine the factors employed in assessing operator fatigue. Consequently, the evaluation encompasses an analysis of the actions performed by the operator, the frequency of blinking, and the changes in posture. By considering these key factors, the Fatigue Evaluation Module can derive insights regarding the level of fatigue experienced by the operator during their work activities. This comprehensive evaluation enables a more accurate assessment of the operator's fatigue state, contributing to the overall goal of promoting operator well-being and ensuring optimal performance. This module generates the final result of our proposal by integrating the outputs of the remaining components.

The global fatigue level of our model is defined using the following equation:

$$fs = off \cdot \alpha + pff \cdot \beta + \frac{\sum_{1}^{n\_task}(tts)}{n\_task} \cdot \gamma + b \qquad (1)$$

where:

$$\begin{aligned}
fs &= \text{fatigue scoring}\\
\alpha &= \text{fatigue scoring adjustment factor}\\
off &= \text{ocular fatigue factor}\\
\beta &= \text{ocular fatigue factor adjustment factor}\\
pff &= \text{positional fatigue factor}\\
n_task &= \text{number of tasks}\\
tts &= \text{time task scoring}\\
\gamma &= \text{positional fatigue factor adjustment factor}\\
b &= \text{bias for personal adjustment of the operator}
\end{aligned}$$

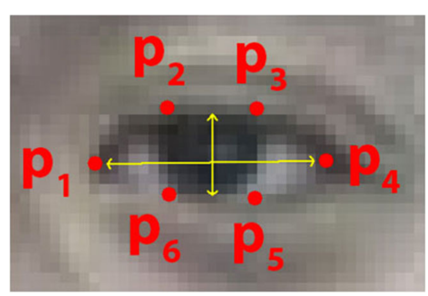

Fig. 2 This figure illustrates the eye landmarks

The equation consists of the following parts, each of which will be detailed below:

- Ocular Fatigue Component
- Component of positional fatigue
- Component of the contribution of task time score weights
- Operator Adjustment Bias

It is important to consider the limitations and sensitivity of our manufacturing action recognition and fatigue monitoring system. The accuracy of the system can be affected by the quality of the images or data streams. Blurred images or inadequate illumination can pose challenges in accurate feature extraction or action recognition.

The system's sensitivity may vary depending on the diversity of actions performed in our dataset. This limitation arises from the limited set of actions, which restricts the generalization and recognition of less frequent actions.

Another aspect to consider is the adaptability of the system to different manufacturing environment configurations. Since production environments may vary in facility layouts, equipment used, and specific tasks, additional adjustments and adaptations to the model architecture and hyperparameters may be required to ensure optimal performance in different contexts.

4 FATIGACTION: Dataset for Action Recognition and Fatigue Evaluation in Manufacturing Environments

This section provides an overview of our dataset, which we have developed specifically for the purpose of human action recognition in manufacturing environments.

There are currently limited datasets available that are specifically tailored for action recognition in industrial settings. For instance, Dallel et al. [26] utilized digital twins (DT) to generate synthetic self-labeled data, resulting in the creation of the inHard-DT dataset. The DT simulates assembly actions to generate synthetic self-labeled data for the purpose of creating an action recognition dataset.

Another notable industrial dataset is the Human Action Multimodal Monitoring in Manufacturing (HA4M) dataset developed by Cicirelli et al. [27]. This dataset encompasses a variety of data types, including RGB images, depth maps, IR images, depth-aligned RGB images, point clouds, and skeleton data.

4.1 Dataset Overview

The foundation of our research lies in the assumption that the operator will be working at a workstation and following specific instructions. To effectively analyze fatigue levels, it is important to capture relevant elements such as facial expressions and general body postures in the recorded videos. We have recorded these videos from three synchronized points of view: front, top, and side.

Our dataset comprises three distinct views: the top view, the front view, and the side view. Each view offers unique perspectives of the worker, allowing for the capture and analysis of actions in manufacturing environments. The top view provides a comprehensive recording of the worker's hands, tools, and spatial relationships between objects. This detailed view enables a thorough understanding of the worker's interactions with the tools and their manipulation techniques. The front view focuses on capturing the worker's face, which allows us to analyze factors such as blinking and the Eye Aspect Ratio (EAR) as indicators of fatigue. Lastly, the side view records the skeleton of the person, enabling analysis of their posture.

Figure 3 illustrates two different moments of an assembly captured from the three views (front, top, and side).

The recorded videos feature individuals performing product assemblies, serving as training data for our system. It is crucial for the videos to closely resemble a manufacturing environment. The workstations in the videos have dimensions ranging from 100 cm to 120 cm by 60 cm to 70 cm. The operator is standing and utilizes their hands or tools during the assembly process.

To ensure proper lighting conditions during recording, we employ artificial light. The videos are recorded in high definition (720 lines), and all three video views are captured from the same distances to the workstation.

During recording, a green desktop is utilized on the workstations, accompanied by white backgrounds. This setup facilitates the identification of hands, parts, components, and tools, among other elements, during the training process.

When selecting volunteers for the recordings, we aimed to generate variability in terms of the operators' physical characteristics, including physical build, height, gender, and skin color.

4.2 Experimental Environment

The recording and production of the videos took place in a designated area within the Human Factors Laboratory of the School of Industrial Engineering at the University of Costa Rica. Workstations were specifically assigned and set up with a green background. The positions, types of cameras, and means of holding the cameras were carefully defined.

The laboratory is equipped with artificial white light illumination and provides ample space and ventilation for the work teams. To mitigate the impact of temperature-related fatigue, we maintain control over the ambient temperature. Additionally, the laboratory is designed to minimize external

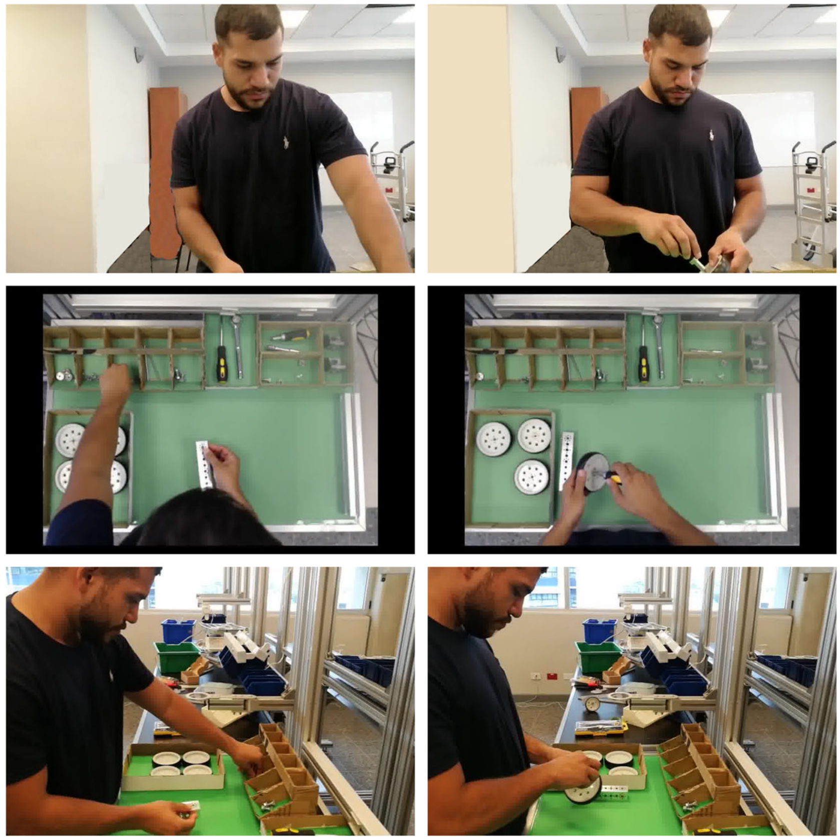

Fig. 3 This figure showcases two different moments of an assembly captured from the three views (front, top, side)

noise, preventing noise fatigue and aiding operators in maintaining concentration. Access to the laboratories is controlled to minimize distractions during video recordings.

We utilized Logitech C310 and Microsoft high-definition webcams to capture the videos. All videos were recorded in Full HD (1920 x 1080 pixels) and UHD (3840 x 2160 pixels) resolutions. Each camera was positioned using a holder to ensure stability.

The videos were recorded in HD (1280 x 720 pixels) at a frame rate of 30 frames per second. All audio tracks were removed from the videos. To achieve synchronization across the three views, we used the OBS Project application and the Record source plugin running on Ubuntu.

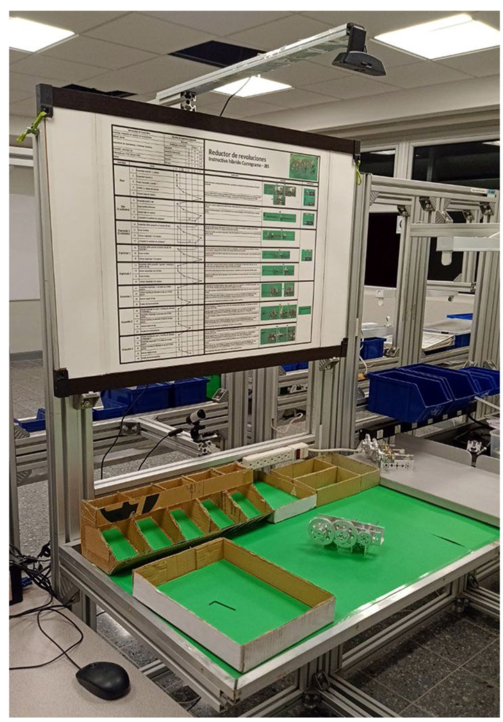

Fig. 4 This figure illustrates an example of the workstations used for video recording

5 Experimentation and Results

This section outlines the experimentation phase, which was conducted in two stages. The first stage involved validating the inputs for the system. This stage focused on the detection of objects, actions, skeleton and hands, time measurement, and fatigue signs. The second stage integrated all these elements and generated the results. We will provide detailed explanations of these stages below.

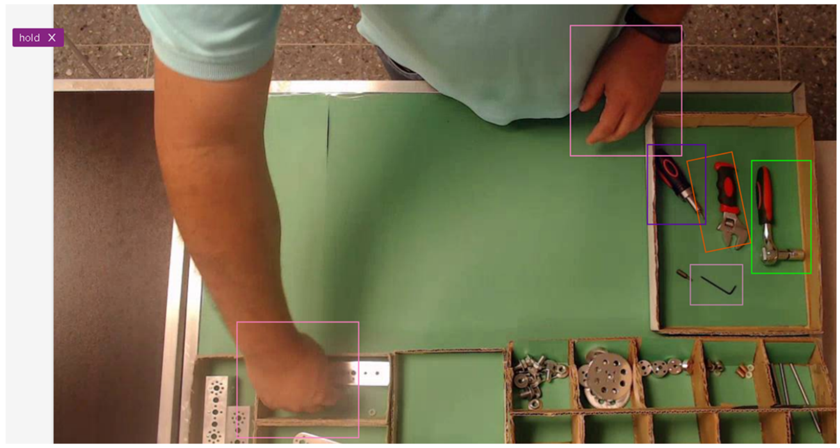

Fig. 5 This figure demonstrates object and action recognition

5.1 Action Recognition Module

The primary purpose of this module is to verify the operator's actions and determine their level of fatigue. This module incorporates the action processing engine based on the Deep Activity Description Vector, developed by the University of Alicante [28]. This application utilizes a sequence of images to accurately recognize actions. We validated the dataset's ability to correctly identify actions and used our dataset for the validation process (Fig. 5).

In this module, we adjusted the training data to recognize the following actions:

1. hold
2. tight
3. screw
4. locknut
5. hit
6. drilling
7. put
8. release

We utilized the entire dataset of videos for conducting the experiments. We employed Scikit-learn to analyze the results and generate the confusion matrix analysis. The mean values for recall and F1-score metrics were found to be 96.13% and 96.25%, respectively. Recall indicates a high level of positive classification by the model, while the F1-score indicates a high level of correct responses. The confusion matrix is presented in Figure 6, and the results can be visualized in Figure 7.

Figure 5 illustrates an example of the frames used to train the system. The frames are labeled using the CVAT software [29]. For action recognition, a label is placed at the beginning and end of each frame, indicating the action performed by the worker during that time.

5.2 Pose Estimation Module

To evaluate fatigue, we utilize complementary measures based on human factors characteristics, which will be further detailed in the following section.

For skeleton and hand detection, we employ OpenCV, which greatly facilitates these tasks in our experiment. Our team conducted visual tests to verify the correct operation, where we visually inspected the videos and determined whether the detection was performed accurately or not. The detection of hands follows a similar procedure. This information serves as a complement to the action recognition. An example of the skeleton identification process from different views can be seen in Figure 8.

By analyzing the operator's posture, we can determine if they slouch while performing their job. When operators engage in repetitive actions and remain in one position for extended periods, they may adopt a non-upright posture, indi-

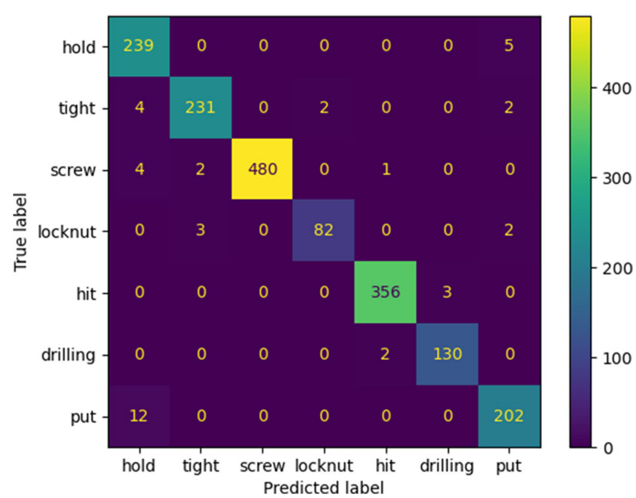

Fig. 6 Action confusion matrix

Fig. 7 Results of the action confusion matrix

	precision	recall	f1-score	support
drilling	0.98	0.98	0.98	132
hit	0.99	0.99	0.99	359
hold	0.92	0.97	0.95	246
locknut	0.98	0.91	0.94	90
put	0.94	0.94	0.94	214
release	0.91	0.97	0.94	116
screw	1.00	0.98	0.99	490
tight	0.98	0.95	0.97	242
accuracy			0.97	1889
macro avg	0.96	0.96	0.96	1889
weighted avg	0.97	0.97	0.97	1889

Fig. 8 Pose Estimation

Fig. 9 Results of fatigue using skeleton positions confusion matrix

	precision	recall	f1-score	support
fatigue	0.99	0.98	0.99	339
non_fatigue	0.96	0.99	0.98	215
accuracy			0.98	554
macro avg	0.98	0.98	0.98	554
weighted avg	0.98	0.98	0.98	554

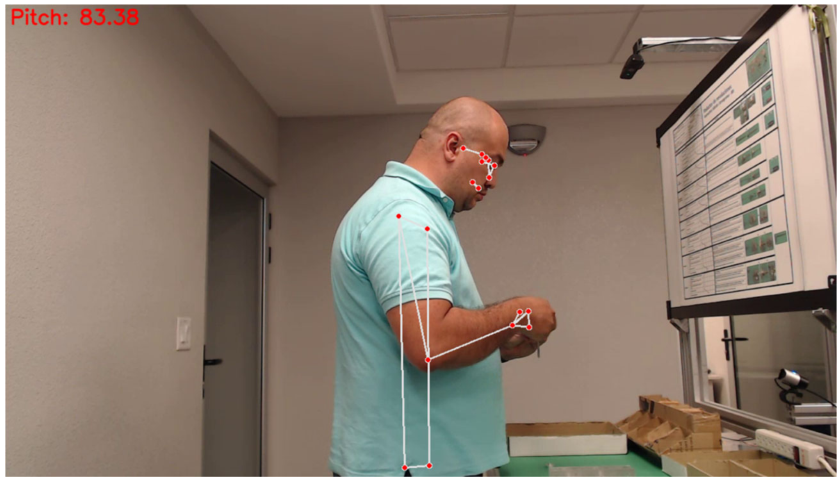

Fig. 10 Positional Estimation

cating signs of fatigue due to work accumulation. The system can detect these changes in position by analyzing the computerized skeletons. To estimate the presence of fatigue, we also consider the EAR indicator. Additionally, we measure the operation time to determine if the person takes longer to perform the same assembly operation, which can be an indication of fatigue. We use the duration taken by an expert individual to perform the assembly as a reference.

The results of fatigue detection using skeleton positions are presented in Figure 9, showing the confusion matrix. The precision, recall, and f1-score metrics are as follows:

This component represents one of our innovative proposals. While the existing literature predominantly focuses on fatigue detection related to blinking or drowsiness, our research introduces the concept of dynamic fatigue analysis in manufacturing. To determine the adjustment factor β, extensive testing has been conducted by the team (Figs. 10 and 11).

To accurately identify this crucial factor, computer vision techniques have been leveraged for skeleton monitoring and the estimation of operators' body inclination variations. It is hypothesized that a fatigued operator may exhibit a compromised upright posture during work, thus necessitating the consideration of this factor to effectively detect the onset of fatigue. This information holds significant importance for decision-makers as it not only impacts the overall product quality but also plays a vital role in ensuring occupational safety. By incorporating these findings, proactive measures can be implemented to effectively address concerns related to fatigue.

The accuracy of positional fatigue detection is 98%, indicating that the system is able to accurately detect fatigue based on changes in posture and operation time.

5.3 Blink Rate Estimation Ratio

The system generates an analysis primarily based on the operators' blinking frequency, which contributes to the rating used in the equation.

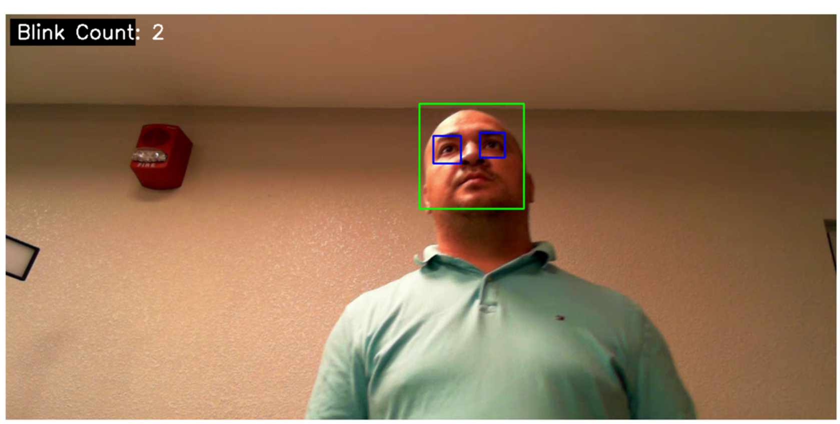

Fig. 11 Blink Rate Estimation Ratio

To discern the blink frequency of individuals engaged in assembly tasks and gather insights into ocular fatigue, the application of OpenCV and computer vision techniques has been employed. These advanced methodologies facilitate the identification of eyes and their corresponding landmarks, thereby enabling the detection of blink incidents. This research significantly contributes to the field of fatigue monitoring by leveraging the power of computer vision approaches.

In the time measurement experimentation, the research team compared the results generated by the system. The system was used to detect actions, and the time was determined by counting consecutive frames. The task's time was calculated by multiplying the frames per second (FPS) rate by the number of frames.After the videos were generated, they were divided into frames for labeling, and then training activities were performed, these videos were processed at 30 FPS.

The evaluation of the time measurements was done based on the following rules: - If the difference between the elapsed time measured by the system and the researcher was less than one second, it was labeled as "timed_ok". - If the difference was greater than one second but less than three seconds, it was labeled as "timed_acceptable". - If the time difference was longer than three seconds, it was labeled as "timed_bad".

The results of the time measurement analysis are presented in Figure 12 as the confusion matrix. The precision, recall, and f1-score metrics are as follows:

The accuracy of the timed task detection is 98%, indicating that the system can accurately measure the time taken to perform assembly tasks. The precision and recall values are high, indicating a reliable identification of the task times, which are important inputs for various calculations in the system.

5.4 Fatigue Evaluation Module

The table 5.4 illustrates the outcomes obtained through the application of various monitoring modules to assembly tasks within the scope of this experiment. The results indicate that an experienced operator requires approximately 8 minutes and 45 seconds to complete the assembly actions. In contrast, based on a study involving 30 recorded videos, it was found that an inexperienced user takes an average of 10 minutes and 45 seconds to accomplish the same tasks.

User Type	Blinking Freq	Positional Dev	Time Execution (min)
Expert	74	3	8.45
Normal	115	5	10.45

In terms of blink frequency and positional deviation, it was observed that experienced users blink approximately 74 times and exhibit a positional deviation of 3 degrees while performing these activities. Conversely, inexperienced users display an average positional deviation of 5 degrees and a blink frequency of 115 times. It is worth noting that these observations were made during daytime work sessions, and both expert and non-expert users reported experiencing fatigue.

To determine the level of fatigue, we calculate a fatigue score. As there were no predefined limits, we established the thresholds for determining fatigue based on the following approach:

After testing the system with 15 individuals, we sought evaluation from industrial engineering experts to assess our metrics. Out of the 15 tests, they confirmed that 13 correctly identified the operator fatigue levels. Consequently, we defined a score of 850 or higher as indicative of a fatigued operator.

6 Conclusions and Future Work

In the manufacturing industry, measuring productivity is essential for assessing performance based on metrics such as total production time, product damage, or rework. Among the various factors that can affect these metrics, operator fatigue has been identified as a significant contributor. However, there is currently no real-time instrument for measuring operator fatigue.

This paper proposes an innovative solution for assessing operator fatigue, providing a means to determine when productivity is degraded due to fatigue. Additionally, our sys-

Fig. 12 Results of timed task confusion matrix

	precision	recall	f1-score	support
timed_acceptable	1.00	1.00	1.00	34
timed_bad	0.00	0.00	0.00	0
timed_ok	1.00	0.98	0.99	393
accuracy			0.98	427
macro avg	0.67	0.66	0.66	427
weighted avg	1.00	0.98	0.99	427

tem estimates the total time required to complete production tasks.

The main contribution of this work is the development of subsystems for visually analyzing operator fatigue during assembly operations in the manufacturing industry. By combining multimodal information, including eye blink rate, operator pose, and task duration, we achieve accurate fatigue detection.

The findings of this research have direct implications for manufacturing companies and their managers, enabling them to make better-informed decisions. By controlling production rework and minimizing losses due to manufacturing errors, companies can realize benefits such as cost reduction.

Moving forward, we plan to refine our model to assess the impact of different submodules on the final classification, thereby improving the system's overall accuracy. As well as the creation of a digital twin to be able to work with synthetic data and obtain better results in the proposed model.

Acknowledgements The authors thank the Industrial Engineering career of the University of Costa Rica for their collaboration in passing the tests and generating the data set.

Author Contributions The authors contributed equally to this work.

Funding Open Access funding provided thanks to the CRUE-CSIC agreement with Springer Nature. "A way of making Europe" European Regional Development Fund (ERDF) and MCIN/AEI/10.13039/501100011033 for supporting this work under the MoDeaAS project (grant PID2019-104818RB-I00).

Code availability Not applicable.

Declarations

Ethics approval Authors agree with paper publication and consent to it. There are not ethical issues with the work.

Consent to participate Authors agree to the authorship order.

Consent to publish The publisher has the permission of the authors to publish the given article.

Conflict of interest The authors declare no competing interests.

References

1. Sharotry A, Jimenez J, Mendez F, Wierschem D, Koldenhoven R, Valles D (2022) Manufacturing operator ergonomics: A conceptual digital twin approach to detect biomechanical fatigue. IEEE Access 10:12774–12791
2. Grajewski D, Górski F, Zawadzki P, Hamrol A (2013) Application of virtual reality techniques in design of ergonomic manufacturing workplaces. Procedia Computer Science 25:289–301. https://doi.org/10.1016/j.procs.2013.11.035
3. Berti N, Artigues C, Battaia O, Guillaume R, Battini D (2019) Heuristic approaches for scheduling manufacturing tasks while taking into account accumulated human fatigue. IFAC-PapersOnLine 52(13):963–968. https://doi.org/10.1016/j.ifacol.2019.11.319
4. Lambay A, Liu Y, Ji Z, Morgan P (2022) Effects of demographic factors for fatigue detection in manufacturing. IFAC-PapersOnLine 55(2):528–533. https://doi.org/10.1016/j.ifacol.2022.04.248
5. Li, P., Meziane, R., Otis, M.J.-D., Ezzaidi, H., Cardou, P.: A smart safety helmet using imu and eeg sensors for worker fatigue detection. In: 2014 IEEE International Symposium on Robotic and Sensors Environments (ROSE) Proceedings, pp. 55–60 (2014). https://doi.org/10.1109/ROSE.2014.6952983
6. Savas BK, Becerikli Y (2020) Real time driver fatigue detection system based on multi-task connn. IEEE Access 8:12491–12498. https://doi.org/10.1109/ACCESS.2020.2963960
7. Zhu T, Zhang C, Wu T, Ouyang Z, Li H, Na X, Liang J, Li W (2022) Research on a real-time driver fatigue detection algorithm based on facial video sequences. Applied Sciences 12:2224. https://doi.org/10.3390/app12042224
8. Lambay, A., Liu, Y., Morgan, P., Ji, Z.: A data-driven fatigue prediction using recurrent neural networks. In: 2021 3rd International Congress on Human-Computer Interaction, Optimization and Robotic Applications (HORA), pp. 1–6 (2021). https://doi.org/10.1109/HORA52670.2021.9461377
9. Escobar-Linero E, Domínguez-Morales M, Sevillano JL (2022) Worker's physical fatigue classification using neural networks. Expert Systems with Applications 198:116784. https://doi.org/10.1016/j.eswa.2022.116784
10. Lukinac J, Mastanjevic K, Mastanjevic K, Nakov G, Jukic M (2019) Computer vision method in beer quality evaluation a review. Beverages 5:1–21. https://doi.org/10.3390/beverages5020038
11. Villalba-Diez, J., Schmidt, D., Gevers, R., Ordieres-Mere, J., Buchwitz, M., Wellbrock, W.: Deep learning for industrial computer vision quality control in the printing industry 4.0. Sensors (Switzerland) 19, 1–23 (2019). https://doi.org/10.3390/s19183987
12. Reich, S., Teich, F., Tamosiunaite, M., Worgotter, F., Ivanovska, T.: A data-driven approach for general visual quality control in a robotic workcell. Journal of Physics: Conference Series 1335 (2019). https://doi.org/10.1088/1742-6596/1335/1/012013
13. Riedel, A., Gerlach, J., Dietsch, M., Herbst, S., Engelmann, F., Brehm, N., Pfeifroth, T.: A deep learning-based worker assistance system for error prevention: Case study in a real-world manual assembly. Advances in Production Engineering And Management 16, 393–404 (2021). https://doi.org/10.14743/apem2021.4.408
14. Chang MML, Ong SK, Nee AYC (2017) AR-guided product disassembly for maintenance and remanufacturing. Procedia CIRP 61:299–304. https://doi.org/10.1016/j.procir.2016.11.194
15. Cheng H, Xu W, Ai Q, Liu Q, Zhou Z, Pham DT (2017) Manufacturing capability assessment for human-robot collaborative disassembly based on multi-data fusion. Procedia Manufacturing 10:26–36. https://doi.org/10.1016/j.promfg.2017.07.008

16. Denkena B, Dittrich M-A, Wilmsmeier S (2019) Automated production data feedback for adaptive work planning and production control. Procedia Manufacturing 28:18–23. https://doi.org/10.1016/j.promfg.2018.12.004
17. Kousi N, Stoubos C, Gkournelos C, Michalos G, Makris S (2019) Enabling human robot interaction in flexible robotic assembly lines: an augmented reality based software suite. Procedia CIRP 81:1429–1434. https://doi.org/10.1016/j.procir.2019.04.328
18. Tao W, Lai ZH, Leu MC, Yin Z, Qin R (2019) A self-aware and active-guiding training & assistant system for worker-centered intelligent manufacturing. Manufacturing Letters 21:45–49. https://doi.org/10.1016/j.mfglet.2019.08.003
19. Zamora-Hernandez, M.-A., Castro-Vargas, J.A., Azorin-Lopez, J., Garcia-Rodriguez, J.: Deep learning-based visual control assistant for assembly in industry 4.0. Computers in Industry 131, 103485 (2021). https://doi.org/10.1016/j.compind.2021.103485
20. Böllhoff J, Metternich J, Frick N, Kruczek M (2016) Evaluation of the human error probability in cellular manufacturing. Procedia CIRP 55:218–223. https://doi.org/10.1016/j.procir.2016.07.080
21. Qeshmy DE, Makdisi J, da Silva EHDR, Angelis J (2019) Managing human errors: Augmented reality systems as a tool in the quality journey. Procedia Manufacturing 28:24–30. https://doi.org/10.1016/j.promfg.2018.12.005
22. Perera AG, Law YW, Ogunwa TT, Chahl J (2020) A multiviewpoint outdoor dataset for human action recognition. IEEE Transactions on Human-Machine Systems 50:405–413. https://doi.org/10.1109/THMS.2020.2971958
23. Varol G, Laptev I, Schmid C, Zisserman A (2021) Synthetic humans for action recognition from unseen viewpoints. International Journal of Computer Vision 129:2264–2287. https://doi.org/10.1007/s11263-021-01467-7
24. Jones JD, Cortesa C, Shelton A, Landau B, Khudanpur S, Hager GD (2021) Fine-grained activity recognition for assembly videos. IEEE Robotics and Automation Letters 6:3728–3735. https://doi.org/10.1109/LRA.2021.3064149
25. Wang P, Liu H, Wang L, Gao RX (2018) Deep learning-based human motion recognition for predictive context-aware human-robot collaboration. CIRP Annals 67:17–20. https://doi.org/10.1016/j.cirp.2018.04.066
26. Dallel M, Havard V, Dupuis Y, Baudry D (2023) Digital twin of an industrial workstation: A novel method of an auto-labeled data generator using virtual reality for human action recognition in the context of human–robot collaboration. Engineering Applications of Artificial Intelligence 118:105655. https://doi.org/10.1016/j.engappai.2022.105655
27. The ha4m dataset: Multi-modal monitoring of an assembly task for human action recognition in manufacturing. Scientific Data 9 (2022). https://doi.org/10.1038/s41597-022-01843-z
28. Borja-Borja, L.F.: Arquitectura de vision y aprendizaje para el reconocimiento de actividades de grupos usando descriptores de movimiento. Doctoral Thesis, University of Alicante (2020)
29. Sekachev, B., Manovich, N., Zhiltsov, M., Zhavoronkov, A., Kalinin, D., Hoff, B., TOsmanov, Kruchinin, D., Zankevich, A., DmitriySidnev, Markelov, M., Johannes222, Chenuet, M., A-Andre, Telenachos, Melnikov, A., Jijoong Kim, Ilouz, L., Glazov, N., Priya4607, Tehrani, R., Seungwon Jeong, Skubriev, V., Yonekura, S., Vugia Truong, Zliang7, Lizhming, Tritin Truong: opencv/cvat: v1.1.0. Zenodo (2020). https://doi.org/10.5281/ZENODO.4009388. https://zenodo.org/record/4009388