WearRAcon Europe 2023
European Exoskeleton Conference
October 25-26, 2023

WEARABLE SYSTEMS FOR ACTION RECOGNITION OF INDUSTRIAL ASSEMBLY TASKS

Jan Kuschan[1], Niklas Grambow[1], Marcel Radke[1], Sophie Irmer[1], Jörg Krüger [1,2]

[1] Fraunhofer Institute for Production System and Design Technologies, Pascalstraße 8-9 10357 Berlin
[2]Technical University of Berlin, Pascalstraße 8-9 19357 Berlin

ABSTRACT

The evolution from Industry 4.0 to Industry 5.0 emphasizes the need for a human-centric approach, referred to as "Resilient Operator 5.0." Despite remarkable advancements in digitized production and logistics systems, the integration of human workers has been insufficiently addressed. This paper responds to the call for Industry 5.0 by exploring the applicability of Inertial Measurement Units (IMUs) for recognizing assembly tasks in an automotive factory. A dataset was collected using the XSens MVN system, capturing activities at three distinct workstations. The dataset, involving ten sensors placed on various body parts, was meticulously labeled and analyzed using a deep convolutional bi-directional long-short term memory (DC-LSTM) architecture. Various parameters impacting accuracy were evaluated, including batch size, label count, window size, training epochs, and stability over iterations. The results exhibited considerable promise, revealing an approximate categorical accuracy and F_1-Score of 0.91. The study unveiled nuanced effects of different parameters on performance. Optimal performance was achieved with a reduced label set from 18 to four, showcasing an accuracy of 0.955.

Keywords: IMU-based action recognition, Industry 5.0, Industrial Assembly Recognition

1. INTRODUCTION

Human beings, due to their outstanding cognitive and sensorimotor abilities and flexibility, are an integral part of almost all production and logistics processes (including production, storage, and distribution logistics). However, humans have so far been only minimally represented ("humans as a source of uncertainty") in digitized Industry 4.0 production and logistics systems, and almost exclusively in a workpiece- or order-related manner (e.g., qualification data, work steps, and process specifications, machine assignments). These deficiencies regarding worker integration and a digital representation of their condition within the framework of Industry 4.0 have been clearly identified by Frost & Sullivan [1] and, above all, the EU Directorate for Research/Innovation [2], highlighting the research and development needs in terms of workforce health prevention and comprehensive resilience of production and logistics processes. Both have introduced the term "Industry 5.0" for this purpose: "A human-centric approach is the core element of transitioning from Industry 4.0 to 5.0: Resilient Operator 5.0" [1].

2. MATERIALS AND METHODS

To make a next step towards Industry 5.0, we made a study regarding the suitability of Inertial Measurement Units (IMUs) to recognize assembly tasks in an automotive factory. Therefore, we recorded a dataset with XSens MVN system [3] and recorded three different work stations; (1) load and unload an AGV, (2) assembling and screwing objects on a carrier plate and (3) plugging and unplugging cables. A detailed overview of the actions is shown in Figure 1. For the recording 10 XSens sensors were used, placed on the head, upper body, arms, hands and legs but not the feet. The recording consists of a dataset of four participants (~30 min each), one a professional from the automotive industry and three from a research institute.After the recording, the data was manually labeled with a synchronized video for the training deep convolutional bi-directional long-short term memory (DC-LSTM) [4]. Regarding the accuracy the following impacts were evaluated: batch-size, number of labels, window size, training epochs and stability over iterations.

Nr	Reduced Label Set	Full Label Set
0	Load AGV	Unload top layer Unload bottom layer Load top layer Load bottom layer
1	Tool Change	Place power drill Pick up power drill Take screwdriver Place screwdriver
2	Position Parts	Set aside part Pack part in Insert part Insert screws Set aside screws
3	Screwing	Loosen screw Insertion screw
4	Plugging	Insert plug
5	Unplugging	Disconnect plug Pull out plug

Figure 1: Detailed overview of the actions

3. RESULTS AND DISCUSSION

The results on the reduced label set are very promising – one exemplary confusion matrix is shown in Figure 2 – and show for both, the categorial accuracy and F_1-Score, approximately 0.91. A deeper investigation shows the impact of the aforementioned parameters.

Interesting to see was that a larger number of datasets lead to slightly worse results in the categorical accuracy, possibly due to higher variances in the data set. As expected, a high number of labels show worse results in the categorical accuracy, for 18 labels the accuracy dropped to 0.75, a reduction to 6 labels shows very good results, reason for this could be the complexity due to many labels. Reducing it even further to four improves the accuracy even to 0.955. By increasing the number of epochs improves results a bit further, but for the current investigation 20 epochs seem reasonable to reduce the computation time of the DC-LSTM. The batch-size has a big influence, a very high number of batches (240) gives bad results, but good results for 4-12 batches. The window-size also has a big influence, for high window-sizes (128) we achieved better results but with the cost of much higher computer load, so for 256 training stops. A compromise seems reasonable, therefore window-size of 64 favored for further training. So far, the variation within the iterations is quite small, but it is still important to prove results cleanly, even though preliminary tests are also possible with few iterations but are yet not the final ones

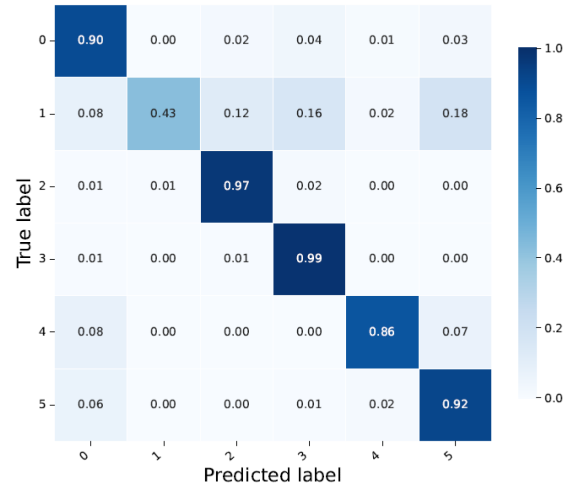

Figure 2: Confusion matrix of the reduced class configuration: (0) Load AGV, (1) Tool Change, (2) Position Parts, (3) Screwing, (4) Plugging, (5) Unplugging

4. CONCLUSION

For six classes we achieved an accuracy of 0.91 for a two-hour dataset consisting of four people. During the next steps we will investigate the influence of the sensor amount and the placement to design a system for an online classifier. The information from the classifier can then be used to intergrade the human actions into a Manufacturing Execution Systems (MES).

ACKNOWLEDGEMENTS

The research and development project BioFusion 4.0 is funded by the German Federal Ministry of Education and Research (BMBF) within the framework concept "Research for the production of tomorrow”.

REFERENCES

[1]https://msu.euramet.org/current_calls/documents/EC_Industry5.0.pdf –2021-10-22
[2] H. Kagermann, W. Wahlster, J. Helbig, “Deutschlands Zukunft als Produktionsstandort sichern, Umsetzungsempfehlungen für das Zukunftsprojekt Industrie 4.0: Abschlussbericht des Arbeitskreises Industrie 4.0”
[3] D. Roetenberg, H. Luinge and P. Slycke, “Xsens MVN: Full 6DOF human motion tracking using miniature inertial sensors”, Xsens Motion Technol. BV Tech. Rep., 2009
[4] F. J. Ordóñez and D. Roggen, “Deep Convolutional and LSTM Recurrent Neural Networks for Multimodal Wearable Activity Recognition”, Sensors 2016, https://www.mdpi.com/1424-8220/16/1/115, ISSN 1424-8220, DOI 10.3390/s16010115