Journal of Machine Engineering, 2026, Vol. 26
ISSN 1895-7595 (Print) ISSN 2391-8071 (Online)

Received*: 02 January 2026* / Accepted: *08 February 2026* / Published online: *24 February 2026*

neural network,
work system,
assembly,
verbal descriptions

Izabela KUTSCHENREITER-PRASZKIEWICZ[1*]

AUTOMATIC GENERATION OF WORK PROCESS DESCRIPTIONS

One of the critical issues in the analysis and design of work systems is the updating of work method descriptions and work process controls. Changes in work methods may originate from product or process engineers or from employees involved in production. Work methods should be documented and incorporated into company documentation in accordance with quality assurance procedures. Therefore, it is necessary to develop a method that supports the creation of production documentation and employs an efficient data analysis technique, such as neural networks. This article presents the application of a deep neural network (DNN) to develop verbal descriptions of production processes based on video recordings. The conversion of visual data into verbal descriptions utilizes a predefined vocabulary of the assembly process. The structure of the DNN and the results of experiments are presented. The proposed approach is useful in developing best practices for identifying working methods tailored to specific needs.

1. INTRODUCTION

Manual business process modelling can be a time-consuming and error-prone process, resulting in low-quality models [1]. Therefore, various approaches have been proposed in research and practice to partially or fully automate the creation of business process models, which differ in terms of the input data, generation methods, and modelling languages utilized. Despite active research into automatic model generation, a considerable scope for improvement remains. New technologies are continually being developed to support automatic model generation [1].

Organisations need to reduce costs, improve efficiency and achieve higher levels of quality. To meet these expectations, natural language processing (NLP) techniques and tools have emerged to generate process models and discover processes [2].

Understanding working methods helps to develop best practices in specific contexts and to understand which working methods best suit specific needs [3].

[1] Faculty of Mechanical Engineering and Computer Science, University of Bielsko-Biala, Poland
[*] E-mail: ipraszkiewicz@ubb.edu.pl
https://doi.org/10.36897/jme/217900

Process, procedure, and work instructions are created to standardize and control work through various teams [4].

Researchers [5] demonstrated that work time control and job crafting were positively associated with workers' work method control. Job crafting moderated the relationship between work time control and work method control [6].

The purpose of a work description document is to communicate to workers the scope of work they are authorized to perform, the main hazards, and the processes and methods used to ensure work is performed safely [7].

Standard work represents the most efficient, safe, and effective method to perform a task or process, based on current best practices. Standard work is documented and visual, making it easily accessible and understandable to everyone involved. It establishes a baseline from which continuous improvement efforts can be launched, serving as a reference point for measuring improvements [8].

Standard work consists of three key elements [8]:

1. The sequence of work: this outlines the specific order in which tasks should be performed to maximize efficiency and minimize waste. This sequence is determined by analyzing the process flow and identifying the most logical and efficient steps.
2. The standard work time: this defines the time required to complete a task or process, ensuring that there is a clear expectation of task duration, thereby promoting consistency and reliability in process outputs.
3. The standard work in process (WIP): this specifies the amount of work that should be in process at any given time, minimizing overproduction and excess inventory, which are forms of waste.

The implementation of standard work and procedures (Fig. 1) plays a pivotal role in achieving operational excellence and sustaining quality improvements. These components are essential in manufacturing control, aiming to minimize variability, reduce waste, and ensure that processes are efficient, predictable, and capable of consistently meeting customer expectations [8]. Researchers apply various methods for process monitoring [9, 10].

Fig. 1. Standard work and procedures

This article proposes solutions aimed at improving categorizing and recognizing assembly tasks, including feature extraction and neural network architecture design based on visual data analysis which facilitates data capture.

2. NATURAL LANGUAGE PROCESSING

NLP is related to information retrieval, knowledge representation and computational linguistics [11].

NLP is one of the most important interdisciplinary fields in computer and artificial intelligence research. Its goal is to enable computers not only to understand human language, but also to perform specific tasks [12,13].

Typical NLP tasks include text summarization, grammatical error correction, logic translation, machine translation, natural language understanding (NLU), natural language generation (NLG), dialogue management, question answering, text-to-image generation, text-to-scene generation, and text-to-video conversion.

Through NLP technology, electronic equipment has the ability to understand and process human language, which significantly improves human work efficiency [13,14]. NLP is involved in processes analysis and design by many researchers. Xu et al. [15] designed a dataset to support research on proactive robots that infer human needs from natural language conversations. Wang [13] noted that text classification technology can help people effectively manage big data, but it can also help people extract information hidden in the database. Yadav et al. [16] presented a method which employs a structured questionnaire to conduct cognitive evaluations based on users' natural language responses. Wang et al. [17] developed a rule-based system to represent the factors that are associated with the type of measurement as well as their interrelationships and then make decisions on the category of the measurement. The retrieval of instrument information is aided by a structural-based approach using the natural language processing (NLP) technique. Zhang et al. [18] proposed an intelligent decision-making system for complex products assembly process planning based on the machine learning (ML) method, providing a targeted decision-making scheme with an assembly process structure tree. Xinghai et al. [19] proposed a specific knowledge extraction method for assembly process document in tabular form. Wang et al. [20] focused on the significance of remote collaboration using virtual replicas, avatars, and gestures for procedural tasks in industry. An overview of the relationships between the methods discussed in the state of the art is presented in Fig. 2.

3. THE PROPOSED APPROACH

The sequence of work outlines the specific order in which tasks should be performed to maximize efficiency and minimize waste. This sequence is determined by analysing the process flow and identifying the most logical and efficient steps [8]; therefore, the work system evaluation procedure can lead to quality improvement (Fig. 2).

Mapping of product and process structure requires identification and representation of process-relevant product properties as well as detection and description of product-characteristic process patterns [21].

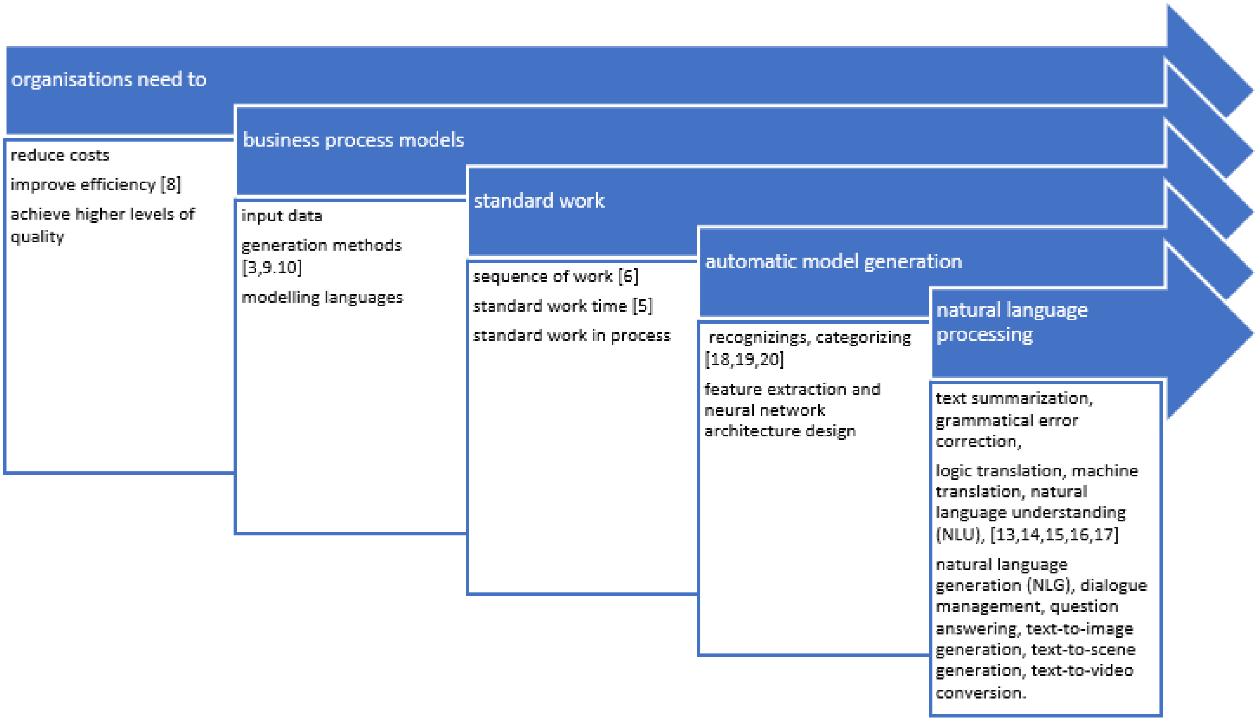

Fig. 2. Overview of the relationships between methods

The procedure for evaluating the work system is shown in Fig. 3.

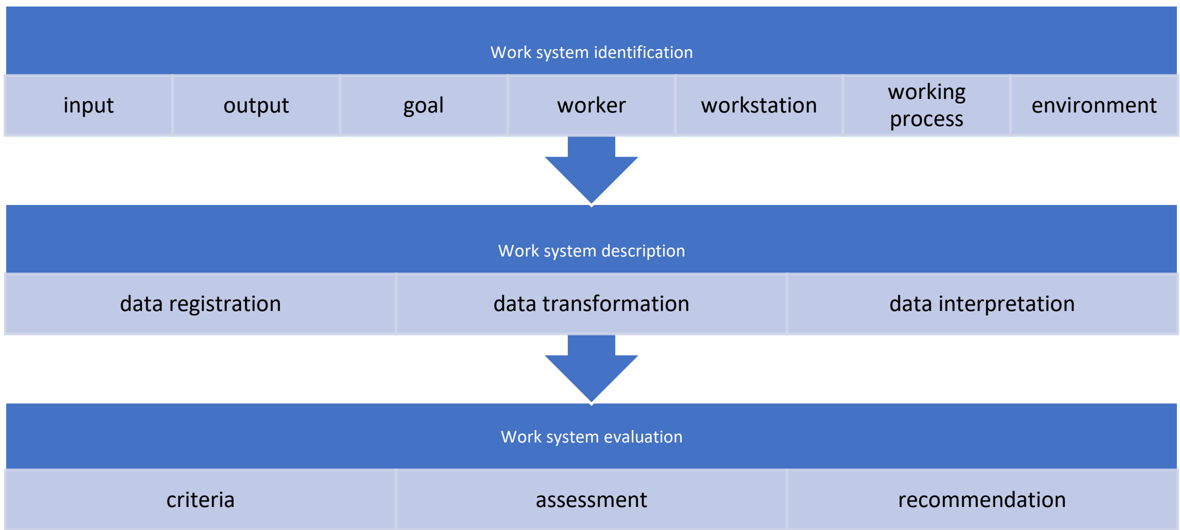

Fig. 3. Work system evaluation procedure

Data registration can use different forms and data types (Fig. 4). Visual Registration Systems (VRS) empower companies with smarter work management solutions [22], therefore visual data registration and analysis have become widely researched.

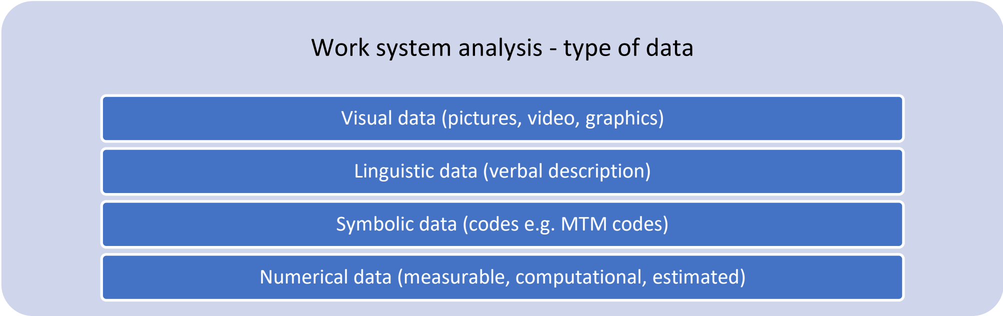

Fig. 4. Work system analysis - type of data

The transformation of the visual registration into a verbal description is shown in Fig. 5. The work description can be generated based on a video recording showing the tasks performed as part of the job.

The procedure for preparing a verbal description of the assembly process is shown in Fig. 5.

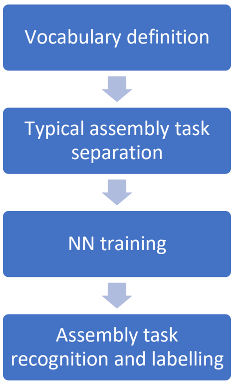

Fig. 5. Assembly process verbal description

Vocabulary definition in an assembly process can be defined based on the following guidelines:

- the description should contain at least a verb and a noun,
- the description should focus on an easily recognisable task,
- the task should be repeated in a similar manner in different assembly processes.

Therefore, assembly description should be related to typical workstation layouts.

Verbal description of assembly process can concern nouns such as:

- component assembled (component name, component characteristic – component weight, size, symmetry, etc.),
- tool or equipment (powered, manual).

Verbal description of assembly process can concern vers such as:

- activity type performed,
- activity characteristic performed (distances, force needed, etc.).

Typical assembly task separation means that the task begins and ends with a specific action that can be easily recognised by a clearly visible movement that repeats across work cycles.

The task should be repeated in a similar manner across assembly processes. Tasks performed by employees should consist of basic movements that are executed in the same sequence.

The next stage of the proposed approach focuses on identifying a neural network structure that recognises the assembly task.

Neural networks (NN) are challenging to apply in domains requiring high reliability due to their black-box nature [23].

Researchers have proposed several neural network (NN) algorithms, such as deep neural network (DNN), convolutional neural network (CNN), and recurrent neural network (RNN). Among them, DNN and CNN have gained popularity [24].

Researchers [25] suggested enhancing digital image categorization and identification capabilities through feature extraction using deep convolutional neural network and transfer learning.

The next stage of the proposed approach is focused on assembly task recognition and labelling.

Each typical movement should be named (labelling); therefore, unusual movements should be identified and flagged potential sources of product failure. The proposed approach can be integrated with product quality control procedures.

Assembly process know-how includes knowledge such as operating methods, sequences, requirements, and techniques [26]. Each assembly task that constitutes operations can be separated and named. The classification of complex assembly tasks has been discussed by Li et al. [26].

4. EXAMPLE

The analysed example is based on assembly work performed on a production line under mass production conditions. The work system used in the analysis included: one assembly station, one experienced worker, a subassembly consisting of four components, two hand tools, lubricant, and an assembly table with a mounting plate. The components are taken from containers. The employee is required to position parts and components for various assembly operations in order to complete the assembly task. The assembly process involves many assembly stages and diverse, complex assembly relationships. Therefore, it is possible to identify typical assembly tasks that are repeated across various assembly operations.

Methods of production process analysis is presented in Fig. 6. Analysis of the entire technological operation is used in time study and use short description e.g. “pick up and assemble the components”. The detailed description can be used to create work instructions and can be used to improve assembly.

Fig. 6. Production process analysis

In the example shown, the complex assembly operation has been divided into tasks and named. An overview of the analysed assembly workstation is presented in Fig. 7.

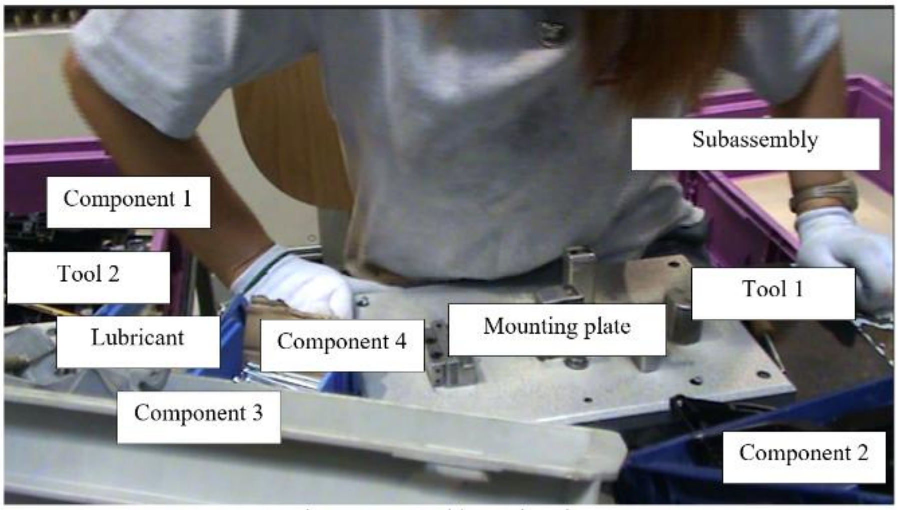

Fig. 7. Assembly workstation

The classification of complex assembly tasks uses vocabulary that includes the following items:

- put component (b),
- put tool (c),
- take component (d),
- take tool (e),
- lubricate component (f),
- connect elements (g),
- move the lever (h),
- immerse the element in grease (i),
- immerse the tool in lubricant (j),
- other (a).

Training and validation sets were created based on the vocabulary (Fig. 8), 3 work cycles and 834 observations.

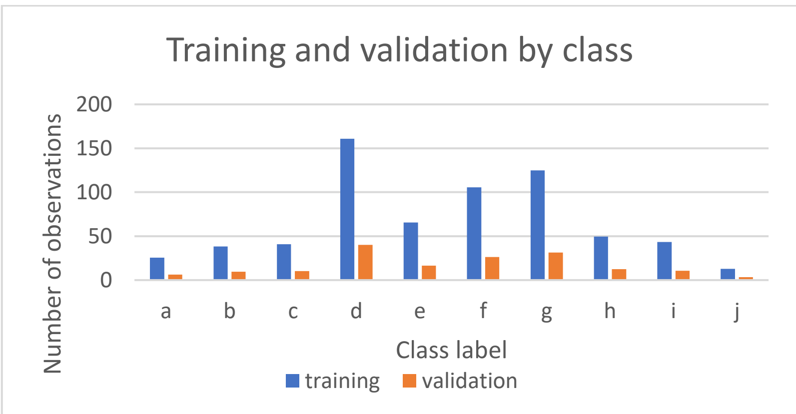

Fig. 8. Training and validation sets

Neural network structure used in the study is presented in Fig. 9. NN structure use layers consists of convolution which is used to extract features from input images. A convolutional layer consists of a set of learnable filters (also known as kernels), which are small matrices that slide over the input image and perform element-wise multiplication and summation operations at each position [30].

Input

```
super(SimpleCNN, self).__init__()
 self.features = nn.Sequential(
nn.Conv2d(3, 16, 3, padding=1), nn.ReLU(), nn.MaxPool2d(2), # 128 -> 64
nn.Conv2d(16, 32, 3, padding=1), nn.ReLU(), nn.MaxPool2d(2), # 64 -> 32
nn.Conv2d(32, 64, 3, padding=1), nn.ReLU(), nn.MaxPool2d(2)  # 32 -> 16)

self.classifier = nn.Sequential(
nn.Flatten(),
nn.Linear(64 * 16 * 16, 64),
nn.ReLU(),
nn.Linear(64, num_classes))
```

Output

put component (b)
put tool (c)
take component (d)
take tool (e)
lubricate component (f)
connect elements (g)
move the lever (h)
immerse the element in grease (i)
immerse the tool in lubricant (j)
other (a)

Fig. 9. NN structure

Convolution layer uses the following parameters [30]:

conv_layer = nn.Conv2d(in_channels, out_channels, kernel_size)

in_channels: The number of input channels in the input tensor. For example, if the input is a color image, in_channels would be 3 (for RGB channels).

out_channels: The number of output channels, which is equal to the number of filters in the convolutional layer.

kernel_size: The size of the convolutional kernel. It can be an integer or a tuple of two integers.

ReLu was used in the NN structure. The ReLU (Rectified Linear Unit) activation function is a popular choice in neural networks that outputs the input directly if it is positive, and zero otherwise. It helps introduce nonlinearity into the model and addresses issues like vanishing gradients, making it effective for deep learning tasks [31].

Pooling is a technique used in the CNN model to reduce the resolution of features from the previous layer and create new, summarising feature maps. In computer vision, it reduces the spatial dimensions of an image while preserving important features. The goal of pooling is to reduce the computational complexity of the model and reduce its sensitivity to small shifts in the input image. In deep learning, two main types of pooling are used: Max Pooling and Average Pooling. Max Pooling selects the maximum value from each set of overlapping filters and passes this maximum value to the next layer. This helps to preserve the most important feature information while reducing the size of the representation. Average Pooling calculates the average value of each set of overlapping filters and passes this average value to the next layer. This helps to preserve a more general form of feature information, but with reduced spatial resolution. [32]

Pooling is typically used after the convolution operation and helps reduce overfitting and improve overall model performance. [32]

Flattening is a crucial operation in neural networks, especially when transitioning from convolutional layers to linear layers. nn.Sequential combined with nn.Flatten in PyTorch provides a simple and effective way to implement this operation. [33]

Therefore using the built-in nn.Flatten or a custom module, flattening ensures that the data is in the right shape for subsequent layers, facilitating smooth model creation. [33]

In the world of deep learning, linear transformations are fundamental components of neural networks. PyTorch, one of the most popular deep learning platforms, provides the nn.Linear module for efficiently transform input features into a new set of features. [34]

The training set was divided into 10 classes. The results of the proposed approach in the final accuracy test reached 97.24%, which demonstrates that the proposed approach is effective. Accuracy was calculated using the formula (1):

$$\text{accuracy} = 100 * \text{correct} / \text{total} \tag{1}$$

A well-known method of error analysis is a confusion matrix, also known as the error table [35, 36], which is table used to measure how well a classification model is performing. It compares the predictions made by the model with the actual results and shows where the model was right or wrong. This helps understand where the model is making mistakes (Fig. 10).

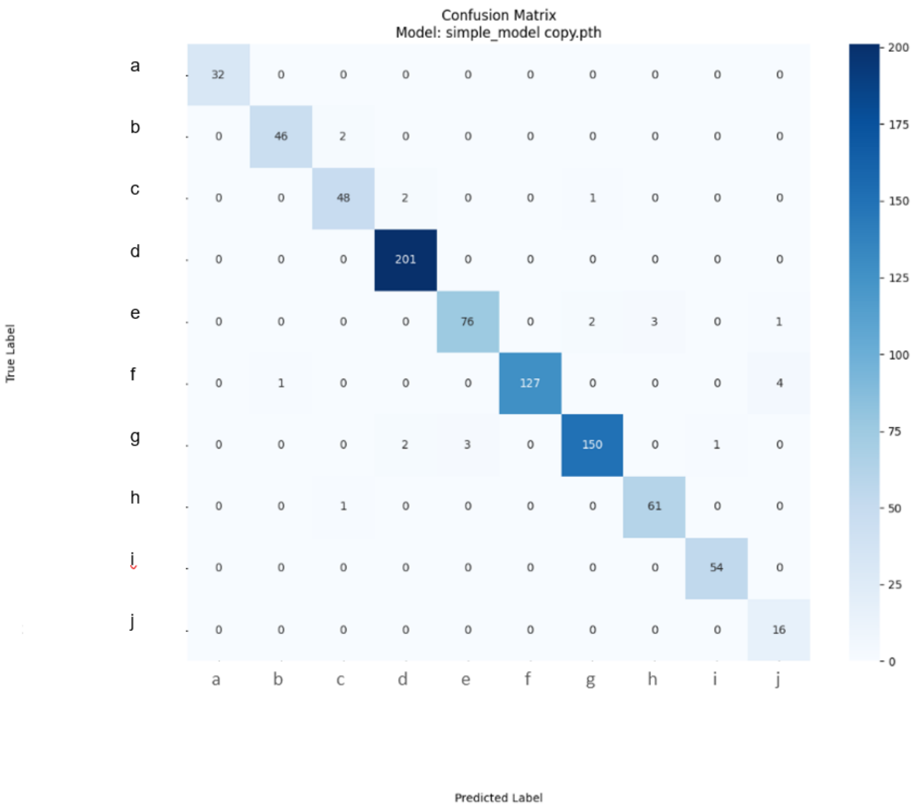

Fig. 10. Confusion matrix

5. DISCUSSION OF THE RESULTS

Automatic generation of work process description can be useful in various aspects of industrial engineering. Hence, in order to investigate the influence of different types of complex assembly tasks on the users' demands for augmented reality (AR) instructions, complex assembly tasks need to be classified [26].

The proposed approach uses DNNs for automatic classification of assembly tasks. Planning and control of assembly tasks require identification of assembly tasks categories. Verbal description of assembly tasks forms a basis for their classification. In the example presented, 10 assembly task categories have been identified. Based on video recordings, the DNN was trained with an accuracy of 97.24%.

Other methods of assembly work description include, for example, MTM (Methods Time Measurement) [27] and the AR-assisted assembly process proposed by Li et al. [26].

A comparison between the proposed approach and other methods used for assembly description is presented in Table.1.

Table 1. Comparison between the proposed approach and another method used for assembly description

No	Comparative criteria	Proposed approach	MTM	AR guided assembly process
1	Aim and scope	Work instruction and analysis	Work analysis	Work instructions
2	Labour intensity	Small (it works automatically)	Large (manual method)	Small (needs additional equipment)
3	Ease of use	Yes	Requires training	Yes
4	Type of work system	Mass production	Mass production	Non repeatable production

The proposed approach can be useful in designing human–robot collaboration in assembly tasks and in facilitating human–robot collaboration in customized assembly tasks. Human–Robot Collaboration (HRC) in industrial environments is in line with the principles of Industry 5.0, proposes an approach to support the integration of human operators' capabilities with advanced robotics, enhancing collaborative productivity and fostering a paradigm shift towards a more interactive and beneficial human–robot symbiosis [28]. Human motion prediction is crucial for facilitating human–robot collaboration in customized assembly tasks [29].

6. CONCLUSIONS

Verbal description of assembly tasks can be useful in various aspects of industrial engineering such as assembly instruction generation, assembly process control and monitoring, assembly instruction updating and facilitating human-robot collaboration.

The proposed approach uses a verbal description of the assembly process to classify assembly tasks and a DNN for its visual identification. Proposed approach is easy to use and does not require expensive equipment. The proposed approach expands the use of DNNs in industrial engineering and offers new opportunities in the field of process analysis. The predefined verbal characteristic of the production process can be flexibly adapted to various production processes. The presented approach helps in creating work standards. Future research will include more employees and samples, and the vocabulary will be expanded. The presented approach uses a given work system, and the example shown demonstrates that the proposed approach can use neural networks as a tool for analysing the work system. Future research will focus on finding parameters that affect the effectiveness of neural networks as a method of data analysis. Human activity can be modelled using different methods such as MTM or AG; each approach can be useful in certain circumstances. MTM, as a manual method of assembly process analysis, is universal but time-consuming. AR provides the opportunity to predict an assembly solution but requires additional equipment. The proposed approach is promising and combines both visual and verbal features of the assembly process.

REFERENCES

[1] SCHÜLER, S., ALPERS, S., 2024, *State of the Art: Automatic Generation of Business Process Models*, De Weerdt, J., Pufahl, L. (eds) Business Process Management Workshops, BPM 2023. Lecture Notes in Business Information Processing, vol 492, 161–173, Springer.
[2] BORDIGNON A.C., THOM L.H., SILVA T.S., DANI V.S., FANTINATO M., FERREIRA, R.C.B., 2018, *Natural Language Processing in Business Process Identification and Modeling: A Systematic Literature Review*. Proceedings of the XIV Brazilian Symposium on Information Systems, Caxias do Sul Brazil, ACM 1–8.
[3] MCCROSKY J., 2025, *The 2025 Guide to Work Management Methods and How They're Best Used*, https://www.mkrcenter.com/work-management/work-management-methods.
[4] https://clickhelp.com/clickhelp-technical-writing-blog/processes-procedures-and-work-instructions-differences-and-basic-principles/ (access 12.2025).
[5] UJOATUONU I.V.N., KANU G.C., OKAFOR C.O., OYIGA C.G., 2023, *Work Time Control and Job Crafting as Predictors of Work Method Control Among Health Workers*, SAGE Open Nurs. 18;9:23779608231189960, PMID: 37492558; PMCID: PMC10363868.
[6] https://pmc.ncbi.nlm.nih.gov/articles/PMC10363868/ (access 12.2025).
[7] https://training.lbl.gov › tutorials › templates › DOW_Template_Research_Work.pdf (access 12.2025).
[8] https://www.6sigmatool.com/blackbeltarticles/standard-work-and-procedures (access 12.2025).
[9] DEJA M., 2022, *Method of Monitoring of the Grinding Process with Lapping Kinematics Using Audible Sound Analysis*, Journal of Machine Engineering, 22/4,71–81.
[10] GOETZ S., SCHRAKNEPPER D., FAUSTINI G., BERGS T., 2020, *Process Monitoring in End Milling Using Polar Figures*, Journal of Machine Engineering, 20/3, 95–105.
[11] EISENSTEIN J., 2019, *Introduction to Natural Language Processing*, The MIT Press. p. 1, ISBN 978-0-262-04284-0.
[12] CROWSTON K., ALLEN EE., HECKMAN R., 2012, *Using Natural Language Processing Technology for Qualitative Data Analysis*, Int. J. Soc. Res. Methodol. 15/6, 523–543.
[13] WANG J., 2023, *Research of Natural Language Processing Based on Dynamic Search Corpus in Cultural Translation and Emotional Analysis*, Soft Computing 27, 7647–7655.
[14] YUE X., DI G., YU Y. et al., 2012, *Analysis of the Combination of Natural Language Processing and Search Engine Technology*, Procedia Eng. 29, 1636–1639.
[15] XU H., LI C., YUAN X., ZHI T., LIU H., 2025, *A Simulated Dataset for Proactive Robot Task Inference from Streaming Natural Language Dialogues*, Scientific Data, 12, 1405.
[16] YADAV A., DHANDA N., SINGH D., 2025, *Intelligent Cognitive Health Monitoring for Seniors Using Natural Language Processing-Integrated GRU Attention and Q-Reinforcement Learning*, International Journal of Speech Technology, 28, 281–298.
[17] WANG D., BAI Y., 2024, *Data Processing, Information Retrieval and Classification of Atmospheric Measurements*, Journal of Data, Information and Management, 6, 41–49.

[18] ZHANG S.,·WANG Z., CHENG D., FANG X., 2022, *An Intelligent Decision-Making System for Assembly Process Planning Based on Machine Learning Considering the Variety of Assembly Unit and Assembly Proces*, The International Journal of Advanced Manufacturing Technology, 121, 805–825.
[19] XINGHAI G.U., BAO H.U.A., YAHUI L.I.U., XUEMIN S.U.N., JINSONG B.A.O., 2024, *Semantic Entity Recognition and Relation Construction Method for Assembly Process Document*, J. Shanghai Jiao Tong Univ. (Sci.), 29/3, 537–556.
[20] WANG P., WANG Y., BILLINGHURST M., YANG H, XU P., LI Y., 2023, *BeHere: a VR/SAR Remote Collaboration System Based on Virtual Replicas Sharing Gesture and Avatar in a Procedural Task*, Virtual Reality, 27, 1409–1430.
[21] EROHIN O., KRETSCHMER R., 2017, *Application of Data Mining for Prospective Assembly Time Determination*, IDS 2017 Industrial Data Science Conference, Dortmund / Research project “Pro Mondi”, http://www.pro-mondi.de/.
[22] https://www.visualregistration.com/company.
[23] SHI, Z., MENG, Z., TUO, H. et al., 2025, *Attribution-Based Interpretable Classification Neural Network with Global and Local Perspectives*, Sci. Rep. 15, 24716.
[24] MAHBOOB, T., CHUNG, M.Y., 2025, *Neural Network-Based Encrypted Traffic Classification and Application Categorization Framework for the Tor Network*, Ann, Telecommun.
[25] ZHANG, H., 2025, *Improved Deep Convolutional Neural Network for Digital Art Image Classification and Identification*, Int. J. Comput. Intell. Syst. 18, 249.
[26] LI J., WANG S., WANG G., ZHANG J., FENG S., XIAO Y., WU S., 2024, *The Effects of Complex Assembly Task Type and Assembly Experience on Users' Demands for Augmented Reality Instructions*, The International Journal of Advanced Manufacturing Technology 131, 1479–1496.
[27] https://mtm.org/en/about-mtm/mtm (access 12.2025).
[28] BARRAVECCHIA, F., BARTOLOMEI, M., MASTROGIACOMO, L., et al., 2025, *Designing Symbiotic Human–Robot Collaboration in Assembly Tasks*, Prod. Eng. Res. Devel. 19, 629–661.
[29] DING, P., ZHANG, J., ZHENG, P., et al., 2025, *Dynamic Scenario-Enhanced Diverse Human Motion Prediction Network for Proactive Human–Robot Collaboration in Customized Assembly Tasks*, J. Intell. Manuf. 36, 4593–4612.
[30] https://www.codegenes.net/ (access 2.2026).
[31] https://www.geeksforgeeks.org/ (access 2.2026).
[32] https://www.geeksforgeeks.org/computer-vision/apply-a-2d-max-pooling-in-pytorch/ (access 2.2026).
[33] https://www.geeksforgeeks.org/deep-learning/how-to-flatten-input-in-nn-sequential-in-pytorch.
[34] https://www.codegenes.net/blog/how-nnlinear-works-in-pytorch/ (access 2.2026).
[35] https://www.geeksforgeeks.org/machine-learning/confusion-matrix-machine-learning/ (access 2.2026).
[36] STEHMAN S., 1997, *Selecting and interpreting measures of thematic classification accuracy*, Remote Sensing of Environment, 62/1, 77–89, https://doi.org/10.1016/S0034-4257(97)00083-7.