Available online at www.sciencedirect.com

ScienceDirect

Procedia CIRP 120 (2023) 649–654

www.elsevier.com/locate/procedia

56th CIRP Conference on Manufacturing Systems, CIRP CMS ‘23, South Africa

Generalizability of an Identification Approach for Machine Control Signals in Brownfield Production Environments

Philipp Gönnheimer[a,]*, Robin Ströbel[a], Roman Dörflinger[a], Marcel Mattes[a], Philipp Alexander[a], Thorsten Wuest[b], Jürgen Fleischer[a]

[a]*wbk Institute of Production Science, Karlsruhe Institute of Technology, Kaiserstraße 12, 76131 Karlsruhe, Germany*

[b]*Industrial and Management Systems Engineering, West Virginia University, Morgantown, WV 26506-6070, USA*

* Corresponding author. Tel.: +49 1523 950 2578. E-mail address: philipp.goennheimer@kit.edu

Abstract

Digital transformation has been a central aspect of optimizing processes in manufacturing companies for several years now. A basic prerequisite of successful transformation is the vertical integration of all machines and machine tools to capture data at all levels. This can create further applications that enable more sustainable and resource-saving processes. At the same time cost- and quality-optimizing analyses such as failure detection, predictive maintenance or general process optimization represent major incentives for companies. While the necessary interfaces are now integrated in state-of-the-art machine tools, companies with older legacy machines face the problem that no such interfaces are readily available. Brownfield machine tools feature outdated technology that does not allow direct networking connectivity without further effort. To participate in the technological progress, a system was developed that allows to extract machine control signals from machine tools and identify them automatically as time series. This is compatible with several communication protocols (e.g., OPC UA) to be as universally applicable as possible. Since machine control signals are often not interpretable for the user due to different naming conventions, the extracted time series are analyzed by machine learning and analytical rule bases, these are based on expert knowledge, and assign a specific signal type in each case. With regard to a cross-machine generalization capability, several aspects have to be considered. Due to different data sources, the identification system must still function reliably with varying sampling frequency. Another challenge is the diversity of different types of machines and production equipment. Therefore, this publication investigates the different influences of data sources and machine types on the machine control signal identification system.

Keywords: Artifical intelligence; Automation; Digital manufacturing system; Machine Tool; Industry 4.0

1. Introduction

With the objective of a high Overall Equipment Effectiveness (OEE), production system operators are pursuing a variety of approaches to increase the availability, performance, and generated quality of production equipment in a data-driven manner. There is already a large number of exemplary Industry 4.0 / Industrial Internet of Things (IIoT) solutions such as pattern recognition, anomaly detection or comprehensive condition monitoring systems [1, 2]. This stands true in research in particular, but also in industrial applications. In many cases successful examples are based on machine component data, for example motor currents or position data of feed axes [3, 4].

Despite the growing prevalence of Industry 4.0, IIoT and the respective digitalization solutions, the actual implementation of these solutions, primarily through heterogeneous control and communication systems based on different manufacturers and machine and control generations in brownfield production environments, represents a large and often manual control

10.1016/j.procir.2023.09.053

signal identification and mapping effort for operators. Signal identification also requires extensive knowledge and is therefore usually the responsibility of experts only.

To reduce the time and effort of today's time-consuming signal identification process, which additionally requires a lot of expertise, and to automate it, a prototype system has been developed. This prototype can extract signals from different data sources and identify them based on a hybrid machine learning (ML) / analytical identification approach [5].

The developed system was designed for use in milling machines and its functionality was tested on different milling machine models. However, it has not been determined how well the system works under more difficult conditions, for example with lower-frequency data. In addition, transferability to different production equipment such as other machine tools or entirely dissimilar systems such as industrial robots is also an open question. These generalization aspects were investigated in the work underlying this paper.

Nomenclature

CUR	Current
CMX	DMG MORI CMX 600 V
DMC	DECKEL MAHO DMC 60H – HDM
FCN	Fully Convolutional Network
HF	High frequency
IIoT	Industrial Internet of Things
LSTM	Long-Short-Term-Memory
ML	Machine Learning
NC	Numerical Control
OEE	Overall Equipment Effectiveness
ResNet	Residual Network
RF	Random Forest
TD	Test dataset
TOR	Torque
umati	Universal Machine Technology Interface
VEL	Velocity

2. State of the art and preliminary work

The underlying problem is the connection of machines from the brownfield environment and generally machines with heterogeneous internal data sources and protocols.

There are already various approaches for integrating these machines into an information system [6, 7]. However, no solution exists for the general and automatic identification of high frequency (HF) signals such as motor currents (CUR).

Additionally, there are approaches such as the universal machine technology interface (umati), which combine signals into a standardized information model from different manufacturers, but this also does not include HF signals from machine components [8]. In addition, due to their novelty, applications are rarely implemented on brownfield machines, due to their age. Regarding [4], 81.00 % of the machines in southern Germany are older than six years, and, according to [9], the average age in France is 19 years.

Existing approaches, which already deal with the automated identification of signals, such as [10, 11], are aimed at other applications such as vehicles and are not suitable for machine tools or production equipment in general. Other publications often only propose frameworks and are not real-time capable. Therefore, to develop a suitable solution for applications in production equipment, an own approach was necessary.

The own approach developed in preliminary work and presented in [5] consists of a hybrid model for the identification of machine control signals and is shown in Figure 1. The goal of this approach is to analyze classification data in order to identify which specific signal it is based on its values. The approach consists of analytical rules and a machine learning model. In a first stage, read-in datasets are examined for signals that are easy to recognize. For the most part, signals without information content are extracted. Subsequently, the remaining signals are divided into position and non-position data by an ML model. In the third stage, these are further divided using analytical rules based on basic physical and kinematic relationships until specific signals can finally be identified.

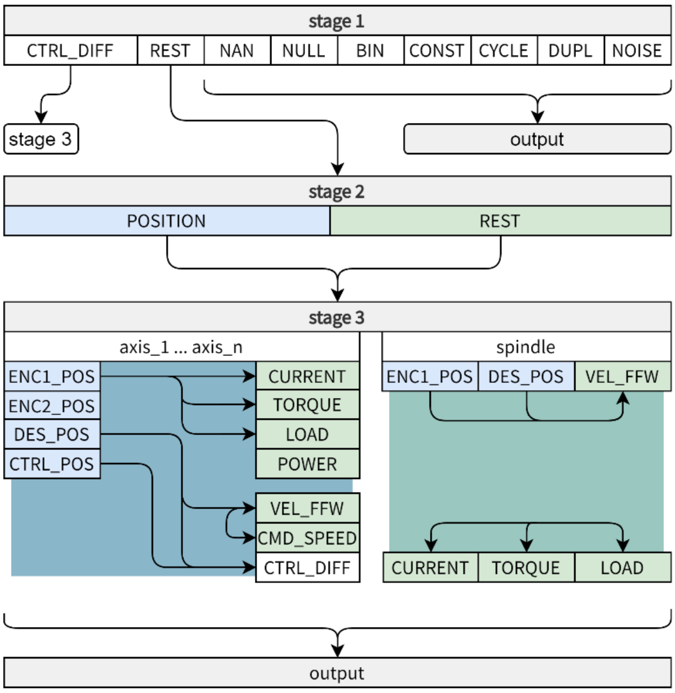

Figure 1: Structure of the 3-stage hybrid model [5]

The generalization capability of the hybrid model in Figure 1 has already been verified in [5] at the level of similar milling machines. This paper now investigates the transferability of the system.

3. Approach and application schematic

The underlying problem statement of this work is to take up the recently published hybrid model and to carry out extensive investigations regarding the generalizability of the model.

The presented hybrid model should also be applicable as independently as possible of different manufacturers and machine types, e.g. milling machines, industrial robots and so forth. This extended level of generalizability is now to be comprehensively investigated in this work. Therefore, it is necessary to collect new data from other heterogeneous machine types and devices. The problem with the status quo is therefore that not much can be said about the generalizability of the hybrid model of [5] so far. Even though the results for the milling machines in the previous work turned out well, this only gives an initial impression for a first degree of generalization. For this reason, more comprehensive investigations are necessary.

This is why this publication is mainly focused on more datasets from different machine types and test it with different settings with the hybrid model. For this, a dataset from a honing machine is used which represents a different type of cutting machine. Furthermore, industrial robots are now more widely used than ever, so, as a benchmark for the hybrid model, a dataset from an industrial robot is used as well. With these datasets, the ability of the existing hybrid model is analyzed in terms of generalizability.

Another important property which has to be taken into account, not just between machine types, but within one type, is the sampling rate. Regarding the fact that the sampling rate of a signal can be different since manufactures use different systems in their machines the effect of the different sampling rates needs to be investigated. In previous work, HF signals with 500 Hz were used. For this reason, sampling rates will be reduced to smaller and more common frequencies.

Since the datasets are partly very small, only partial modules of the hybrid model are used under these circumstances. Furthermore, all the ML models described in [5] are used instead of following only the best. The basic rules from stage 3, which are also to be used here, are shown in the table below.

Table 1. Analytical rules to identify signals and signal groups [5]

	Correlation
Cor. 1	$CTRL\_DIFF \sim CTRL\_POS - ENC2\_POS$
Cor. 2	$VEL\_FFW \sim \frac{d}{dt}(DES\_POS)$
Cor. 3	$CMD\_SPEED \sim VEL\_FFW$
Cor. 4	$\left\lvert\frac{d^2}{dt^2}(ENC1\_POS)\right\rvert \sim \lvert CURRENT\rvert \sim \lvert TORQUE\rvert \sim \lvert LOAD\rvert$

Before the signals get processed with the correlations from Table 1 they have to be modified at first. Depending on the rule the signals will be smoothed with moving average and moving standard deviation. Furthermore, conspicuous values in the time series will be extracted from the original data. To evaluate the correlation between signals, Bravais-Pearson and the Minkowski distance are used. In [12], a Residual Network (ResNet), Fully Convolutional Network (FCN), Long-Short-Term-Memory (LSTM), and a Random Forest (RF) have already been trained with data from milling machines. In subsequent work, the models were continuously improved, which is why they are also used in this work.

To perform the described tests, all datasets are split in different experiments. The first experiment analyzes all existing datasets for the first stage of the hybrid model where signals are extracted with simple rules. The second experiment describes the correlation between different sampling rates and the accuracies of the ML models and the overall result of the hybrid model. The third and fourth experiments are the analyses of two datasets regarding the ML models and the third stage of the hybrid model. For the third stage where signals are identified by physical and kinematical rules a selection of these rules must be made at first. This is since the required signals are not present in the datasets for each rule. Furthermore, the effect of the reduction of the sampling rate will be analyzed for each rule as well.

4. Datasets and machines

The objective of this chapter is to sufficiently characterize the research data and materials used. Relevant machines are listed within the machine setup. Machine-specific metadata as well as associated datasets are presented in the following chapter.

4.1. Machine setup

A four-axis (DECKEL MAHO DMC 60H – HDM (DMC)) and a three-axis milling machining center (DMG MORI CMX 600 V (CMX)), which are both equipped with a SINUMERIK 840D control represent the baseline of the validation schematic. The three translatory axes (X, Y, and Z) of the CMX enable working movements in the range of 600 mm, 560 mm, and 510 mm. The same applies for the DMC with 620 mm, 560 mm and 560 mm with a supplementary rotatory B-axis. In addition, a honing machine (Nagel VARIOHONE center) is considered, which represents a machine concept for the machining of cylinder tracks. The equipment includes a rotary table for non-productive time reduction and vertically arranged electromechanical spindles. Servo-motorized stroke and rotary drives enable the honing of cylinder bores for internal combustion engines. Detached from machining processes, a vertically articulated robot (KUKA KR6 R900-2) is included into the scope. The 6-axis serial kinematic robot has a maximum reach of 901 mm and a maximum payload of 6.70 kg. The rotary axes enable handling processes in the field of pick-and-place applications.

4.2. Datasets

The relevant test datasets (TD) were collected from the data sources introduced in the machine setup. Data specifications are shown in Table 2 and allow comparability.

Table 2. Data specification

Test dataset	Process	Duration	Signals	Sampling rate
TD1	Empty run	13.00 s	100	500 Hz
TD2	Milling	276.00 s	92	500 Hz
TD3	Honing	231.40 - 402.50 s	24	500 Hz
TD4	Handling	504.00 s.	12	83.30 Hz

TD1 was generated on the DMC on the basis of unique and identifiable reference runs [13] with 100 signals at a rate of 500 Hz. Further 92 signals at 500 Hz were sampled under arising process forces (TD2) on the CMX. In both cases a Siemens industrial edge including a numerical control (NC) interface has been utilized. The process data of the honing machine were recorded at 500 Hz by using an industrial edge device. A total of 24 different signals could be collected during the machining process of the cylinder tracks (TD3). The total data pool was finally extended by the dataset of the vertical articulated arm robot (TD4) using a KUKA smartPAD with an active trace function. During a handling process, 12 signals were sampled with 83.30 Hz. The entire recorded signals can roughly be subdivided into the classes position, control differences, velocity (VEL), current, torque (TOR), power, load and trivial signals. Trivial signals are e.g. binary and cycle signals. The absolute number of groups varies depending on the dataset. Additionally, the records are labeled and are further available in this format. The effects of the domain- as well as machine-bound conditions on the generalization ability of the model - in particular on the model levels - are analyzed in the following chapter.

5. Results

In order to select a reasonable sampling rate reduction, the datasets were analyzed beforehand. For TD1, TD2 and TD3 a sampling rate reduction from 500 Hz to 250 Hz and 125 Hz leads to reasonable steps depending on the recording lengths.

Due to the short recording length of the present dataset TD4, the synthetic sampling frequency cannot be set arbitrarily. Since the time series for the ML models must meet certain requirements, a reduction of the sampling rate to a maximum of 20.83 Hz can be made with these datasets.

5.1. First Experiment – prefilter with all TD

Table 3 lists the accuracies of the first stage of the hybrid model of all datasets with different sampling frequencies. For all data sets, the easy-to-detect signals could be identified with at least 98.84 %.

Table 3. Accuracies of stage 1 with different sampling rates

TD	500 Hz	250 Hz	125 Hz	83.30 Hz	41.65 Hz	20.83 Hz
TD1	99.59 %	99.16 %	98.84 %	-	-	-
TD2	100.00 %	100.00 %	100.00 %	-	-	-
TD3	100.00 %	100.00 %	100.00 %	-	-	-
TD4	-	-	-	100.00 %	100.00 %	100.00 %

5.2. Second Experiment – classification and overall results with TD1 and TD2

In the following Table 4 the accuracies of the ML models with the dataset TD1 are shown. It can be seen that the accuracies at 500 Hz are at least 99.72 % and the reduction in sampling rate has little effect on this. At 125 Hz, the RF loses the most percentage points, but is still at 98.26 %.

Table 4. ML results with TD1

ML model	500 Hz	250 Hz	125 Hz
ResNet	99.97 %	99.95 %	99.98 %
FCN	99.72 %	99.51 %	99.94 %
LSTM	99.87 %	99.59 %	99.05 %
RF	99.99 %	99.82 %	98.26 %

Table 5 lists the accuracies of the ML models with the TD2 dataset. The LSTM model has the highest accuracy at 500 Hz with 95.71 %. The FCN and RF achieve just above and the ResNet below 90.00 %. For all models, a reduction in accuracy is accompanied by a reduction in sampling rate. The exception is the ResNet, which also has lower values, but is 0.95 % higher at 125 Hz than at 250 Hz. For the LSTM, the maximum loss in accuracy is -1.06 %, for the ResNet -7.23 %, for the RF -7.49 %, and for the FCN -15.90 %.

Table 5. ML results with TD2

ML model	500 Hz	250 Hz	125 Hz
ResNet	89.52 %	82.29 %	83.34 %
FCN	90.33 %	89.73 %	74.43 %
LSTM	95.71 %	95.66 %	94.65 %
RF	90.37 %	89.91 %	82.88 %

The effect of sampling rate reduction of TD1 and TD2 on the overall hybrid model is shown in Table 6. TD1 has the highest accuracy at 500 Hz with 99.01 % but loses percentage points with sampling rate reduction. TD2 has almost constant good results over the different sampling rates. While these are 87.84 % at 500 Hz, they increase up to 88.75 % when the sampling rate is reduced.

Table 6. Accuracies of the hybrid overall model with TD1 and TD2 at different sampling rates

Test dataset	500 Hz	250 Hz	125 Hz
TD1	99.01 %	98.76 %	93.85 %
TD2	87.84 %	88.75 %	88.43 %

5.3. Third Experiment – classification and correlation results with TD3

Table 7 shows the accuracies of the ML models with the TD3 dataset. The ResNet and FCN behave very similarly with my maximum value at 500 Hz between 87.00 % and 88.00 % and with sampling rate reduction decreasing accuracies to 82.84 % and 82.69 %. The LSTM shows a rather increasing tendency with decreasing sampling rate. The maximum value is 83.38 %. The RF is between 84.25 % and 86.69 %.

Table 7. ML results of TD3

ML model	500 Hz	250 Hz	125 Hz
ResNet	87.20 %	86.76 %	82.69 %
FCN	87.94 %	86.99 %	82.84 %
LSTM	80.96 %	83.38 %	83.17 %
RF	86.69 %	84.25 %	85.77 %

Table 8 shows the results of stage 3 of the hybrid model with TD3. Based on the available data, only speed and torque signals are examined with the rules Cor. 2 and Cor. 4. Both rules cannot identify any signals. Cor. 2 (mod) shows a slightly adapted rule of Cor. 2, where all signals can be identified.

Table 8. Analysis of analytical rules and different sampling rate with TD3

Analytical rule	Signal	500 Hz	250 Hz	125 Hz
Cor. 2	VEL	0.00 %	0.00 %	0.00 %
Cor. 2 (mod)	VEL	100.00 %	100.00 %	100.00 %

5.4. Fourth Experiment – classification and correlation results with TD4

Table 9 lists the accuracies of the ML models with TD4. At 83.30 Hz, the LSTM has the highest accuracy at 97.88 %, while the other models range from 91.00 % to 92.00 %. As the sampling rate is reduced, the values of all models increase. ResNet, FCN and RF already have 100.00 % at 41.65 Hz.

Table 9. ML results of TD4

ML model	83.30 Hz	41.65 Hz	20.83 Hz
ResNet	91.56 %	100.00 %	100.00 %
FCN	91.51 %	100.00 %	100.00 %
LSTM	97.88 %	99.01 %	99.42 %
RF	91.67 %	100.00 %	100.00 %

For the evaluation of individual rules from stage 3, only Cor. 2 and Cor. 4 are analyzed since sufficient signals are available for these. The results are shown in Table 10. From the current signals all signals are identified based on the rule. For the torque signals there are differences between the X and Y signals. In the X direction, the accuracy values are between 75.76 % and 78.79 %, which are thus higher than those in the Y direction. It is noticeable that with Cor. 2, depending on the axis, either almost all or no signals are identified. Cor. 2 (mod) shows the results of a slightly adapted rule of Cor.2 whereby the pattern just described can be recognized again with the difference that now signals can be identified in X instead of Y direction.

Table 10. Analysis of analytical rules and different sampling rate with TD4

Analytical rule	Signal	83.30 Hz	41.65 Hz	20.83 Hz
Cor. 2	VEL_X	0.00 %	0.00 %	0.00 %
Cor. 2 (mod)	VEL_X	100.00 %	96.97 %	75.76 %
Cor. 4	CUR_X	100.00 %	100.00 %	100.00 %
Cor. 4	TOR_X	78.79 %	78.79 %	75.76 %
Cor. 2	VEL_Y	100.00 %	100.00 %	71.11 %
Cor. 2 (mod)	VEL_Y	0.00 %	0.00 %	0.00 %
Cor. 4	CUR_Y	100.00 %	100.00 %	100.00 %
Cor. 4	TOR_Y	57.78 %	57.78 %	57.78 %

6. Evaluation of results

In [5] it was shown that the hybrid model performs very well at a sampling rate of 500 Hz with the datasets of two milling machines. Now the model was evaluated in the second experiment with the same dataset but based on 250 Hz and 125 Hz. The dataset is synthetic, which was generated using a computational transformation from the original 500 Hz dataset.

For 250 Hz, only marginally smaller accuracy values are found in the mode evaluation for both machines compared to 500 Hz. The change in accuracy for the different scenarios ranges from -0.25 % to 0.91 %.

For 125 Hz, however, stronger decreases in accuracy can be observed. Here, the interval increases to a range of -0.59 % to -5.16 %. A closer look reveals significant differences between the two milling machines. The decrease in accuracy primarily affects TD1. For TD2 the accuracies remain almost the same over the reduction of the sampling rate. This is quite unexpected, since there is a reduction of the sampling rate by a factor of 4 from 500 Hz to 125 Hz.

6.1. Prefilter

As can be seen in Table 3, the easy-to-detect signals were identified with up to 100 % accuracy in all datasets. Also, the reduction of the sampling rate has only a very small influence on the accuracy of the results. This can be justified with the simple rules from [5] and with the small number of signals in TD3 and TD4. Since the rules look for binary numbers, constants and non-numerical values, it was to be expected that the reduction of the sampling rate has no or if so, only very small effects on the results.

6.2. ML models

When evaluating different ML models, it can basically be recognized that all of them can achieve good results. With TD1, the best results are achieved overall. All ML models have accuracies of at least 98.26 %. However, since the training dataset TD1 was recorded on the same milling machine with which the ML models were trained, the results are to be expected. Table 5 is more interesting, since this data was recorded on a different milling machine. Basically, all models almost exclusively achieve lower accuracies with the reduction of the sampling rate. However, the LSTM is particularly robust. At 500 Hz, 95.71 % is achieved and up to 125 Hz, this value drops only 1.06 %.

The investigation with TD3 shows the generalization ability in respect of very different machine tools. Here, the tendency of decreasing accuracy with the reduction of the sampling rate can be seen for ResNet and FCN. Both models have high accuracies of 87.20 % and 87.94 % at 500 Hz and drop by about 5.00 % at 125 Hz. The LSTM achieves slightly higher values than at 500 Hz with the reduction of the sampling frequency but is at a maximum of 83.38 %. The RF is not the best model with 86.69 % but varies only by 2.44 %.

With TD4, a machine type is also analyzed that differs greatly from the original lathes. It is noticeable that initially at 83.30 Hz all ML models have an accuracy of at least 91.51 %. The LSTM is the best model with 97.88 %, showing that in this case all ML models have a very high generalization capability. When the sampling rate is reduced, all accuracies increase again up to 100.00 %. Thus, the ML models are stable, but must be evaluated with further datasets, since the recorded data TD4 is not large enough. Due to the predefined sample length for ML models, the sampling frequency cannot be reduced further. However, it would be expected that with further reduction, a tipping point would occur at which the accuracy would drop sharply. This will be further analyzed in future work.

6.3. Rule bases

TD1 and TD2 are not analyzed again separately for the respective rules from stage 3, since their dataset could already be completely validated with the hybrid model anyway. In addition, the results of different sampling rates of the entire hybrid model with TD1 and TD2 were already presented at the beginning of the chapter.

In TD3 there are several position signals, but only one of them was recognized particularly well. The others were only partially identified as position signals depending on the ML model. This may be since different axis controls were used. While NC axes are mainly used for the milling machines, it is quite possible that this is only occasionally the case for the honing machine. The NC controlled signal was then examined with Cor. 2 and Cor. 4. Initially, no signal was detected with Cor. 2, which is since the sign of the velocity signal is atypical. With the modification that the time-derived position signal is then given a magnitude, all velocity signals were identified. At Cor. 4, modified ones are analyzed based on their Bravais-Pearson correlation. For this, a threshold of 0.80 must be exceeded, which was not reached in this dataset. However, all results are between 0.51 and 0.75, which means that there are certainly meaningful correlations that can be strengthened by further adjustments if necessary.

For TD4, the two rules Cor. 2 and Cor. 4 could also be used. Here, too, anomalies were detected with respect to the sign of signals. While for VEL_X the rule Cor. 2 without modification no velocity signals could be identified, with a sign reversal at 83.30 Hz all signals were found. When reduced to 20.83 Hz, only 75.76 % of the velocity signals are identified with the rule. For VEL_Y an inverted behavior can be seen. Here the unmodified rule Cor. 2 works. When detecting the current signals with Cor. 4, all signals could be detected at all sampling rates. For torque, between 75.76 % to 78.79 % of the signals were detected with TOR_X and 57.78 % with TOR_Y. From this it can be concluded that both minor modifications of the rules must be made with respect to the independence of the signs. Furthermore, the rules are mostly not dependent on sampling rate, which is a promising finding since the rules thus do not have to be adjusted at different frequencies.

7. Conclusion and outlook

It could be shown that the generalization capability of the hybrid model within the milling machine type leads to very good results. In general, all datasets were successful processed in stage 1 with accuracies no less than 98.84 %. Also, the ML models lead to very good results. Especially the LSTM model comes to accuracies of 95.71 % and is very stable with respect to the sampling rate reduction.

Furthermore, the models are highly generalizable also in consideration of other machine types. With accuracies of 87.94 %, position signals and non-position signals were correctly identified in TD3. In TD4, an accuracy of 97.88 % was achieved with the LSTM at 500 Hz, which was further increased when the sampling rate was reduced. However, this needs to be examined in more detail since, in general, the information content is reduced at lower sampling rates. In order to analyze the assignment of the signals in the ML models, embedding must be performed. Thereby it can be determined by means of dimension reduction, how far the respective samples are distributed in a low-dimensional space, on which conclusions can be led to the results of this publication.

Using the established rules from Table 1 it can be said that they build a solid basis to identify a variety of different signals. Due to inconsistent sign of convention only minor adjustments are required to make the sign of signal values independent. At the same time, it was shown that rule Cor. 2 and Cor. 4 work reliably and are only weakly dependent on the sampling rate.

Acknowledgements

We extend our sincere thanks to the German Federal Ministry of Economic Affairs and Climate Action (BMWK) for supporting this research project 13IK001ZF "Software-Defined Manufacturing for the automotive and supplying industry" (SDM4FZI).

References

[1] M. A. Farahani et al., "Time-Series Pattern Recognition in Smart Manufacturing Systems: A Literature Review and Ontology," Jan. 2023. [Online]. Available: https://arxiv.org/pdf/2301.12495.

[2] M. Netzer, J. Michelberger, and J. Fleischer, "Intelligent Anomaly Detection of Machine Tools based on Mean Shift Clustering," 2020, doi: 10.5445/IR/1000130459.

[3] T. Bauernhansl, J. Krüger, G. Reinhart, and G. Schuh, "WGP-Standpunkt Industrie 4.0," Wissenschaftliche Gesellschaft für Produktionstechnik - WGP-, 2016.

[4] M. Netzer, E. Begemann, P. Gönnheimer, and J. Fleischer, "Digitalisierung im deutschen Maschinen- und Anlagenbau," vol. 2021, 07-08, pp. 526–530, doi: 10.37544/1436-4980-2021-07-08-58.

[5] P. Gönnheimer, R. Ströbel, R. Dörflinger, M. Mattes, and J. Fleischer, "Interoperable System for Automated Extraction and Identification of Machine Control Data in Brownfield Production: [Paper accepted]," Manufacturing Letters, 2023.

[6] M.-A. Tnani, M. Feil, and K. Diepold, "Smart Data Collection System for Brownfield CNC Milling Machines: A New Benchmark Dataset for Data-Driven Machine Monitoring," Procedia CIRP, vol. 107, pp. 131–136, 2022.

[7] D. Etz, H. Brantner, and W. Kastner, "Smart Manufacturing Retrofit for Brownfield Systems," Procedia Manufacturing, vol. 42, pp. 327–332, 2020.

[8] A. Wohlfeld, T. Heinemann, and G. Görisch, "The Global Production Language - OPC UA for Machine Tools," 2020.

[9] Jan Larosse, "Analysis of national initiatives on digitising european industry," 2017. Accessed: Oct. 24 2022. [Online]. Available: https://ec.europa.eu/futurium/en/system/files/ged/fr_country_analysis.pdf.

[10] S. Lestyan, G. Acs, G. Biczok, and Z. Szalay, "Extracting vehicle sensor signals from CAN logs for driver re-identification," Feb. 2019. [Online]. Available: https://arxiv.org/pdf/1902.08956.

[11] H. Hozhabr Pour, L. Wegmeth, A. Kordes, M. Grzegorzek, and R. Wismüller, "Feature Extraction and Classification of Sensor Signals in Cars Based on a Modified Codebook Approach," Progress in Computer Recognition Systems, pp. 184–194, 2019, doi: 10.1007/978-3-030-19738-4_19.

[12] P. Gönnheimer, A. Karle, L. Mohr, and J. Fleischer, "Comprehensive machine data acquisition through intelligent parameter identification and assignment," CIRP Conference on Manufacturing Systems, 2021.

[13] P. Gönnheimer, R. Ströbel, M. Netzer, and J. Fleischer, "Generation of identifiable CNC reference runs with high information content for machine learning and analytic approaches to parameter identification," Procedia CIRP, vol. 107, pp. 734–739, 2022, doi: 10.1016/j.procir.2022.05.054.