Tool Wear Monitoring Using Multi-Sensor Time Series and Machine Learning

Jonathan Dreyer[1,2][0009−0007−8665−3094], Stefano Carrino[1][0000−0001−5171−6541], Hatem Ghorbel[1][0000−0001−5501−9807], and Paul Cotofrei[2][0000−0002−4103−5467]

[1] Haute Ecole Arc Ingénierie, University of Applied Sciences and Arts Western Switzerland (HES-SO), St. Imier, Switzerland
{jonathan.dreyer,stefano.carrino,hatem.ghorbel}@he-arc.ch
[2] Information Management Institute, University of Neuchatel, Neuchatel, Switzerland
{jonathan.dreyer,paul.cotofrei}@unine.ch

Abstract. In the milling process of micro-machining, the optimization process is one of the keys to reduce production cost. By monitoring the tool wear and detecting when it is no longer acceptable, the machining process can be adjusted more accurately. This research explores four approaches using different machine learning models to predict machining tool wear during the milling process. The study is based on a dataset created with a face milling operation on stainless steel (AISI 303) round material. The machining is divided into a number of stairs and is performed with a 3mm tungsten carbide. Three different types of sensors are used to measure the wearing process, with acoustic emission, accelerometers and axis currents. The better approach achieved a f1-score of 73% on five classes with a Extra Trees Classifier.

Keywords: Tool wear monitoring · Milling machining · Multi-sensors time series · Machine learning.

1 Introduction

In this article, we explore the possibility of improving machining performance by monitoring and evaluating the tool wear without interrupting the machining process. To be able to perform the tool wear detection, a mix method is used by implementing a state of the art methodology and compare it with other algorithms not used in papers. This can be achieved by measuring in-process with sensors, which help to improve the productivity and the control of the manufacturing [6]. The custom production implies a reduction in the demand for large volumes of production, and the needs of small-sized machine tools satisfy the requirements for micro manufacturing which enables the miniaturization of products and high accuracy [9].

The computer numerical control machining field includes several types of machining such as turning, milling and grinding. The first one is a machining technique where a cutting tool moves longitudinally while the workpiece rotates on itself. The second one is a machining process which uses a rotating cutter tool to remove material by moving the cutter tool into a workpiece. The last one is a machining process which uses a circular abrasive wheel to remove material from the surface while creating a smooth surface texture. In this study, we will focus on the milling process with a micro-milling machining.

The low consumption of these micro-machining machine has been possible by scaling down the dimension [22]. This size reduction enables the reduction of the moving masses and a higher dynamic of motion [9]. The high-speed machining (HSM) is of benefit for micro machining and reduces the requirement of cooling during the milling, because a major part of heating is dissipated by chips. The HSM also helps to increase the productivity by speeding up the milling speed and the cutting speed [13].

Actually, the prediction of tool wear in micro milling context with sensors is challenging. To address this problem, we propose various approaches based on several machine learning classification models, including Convolutional Neural Networks (CNN). The machine learning algorithms can enhance the tool wearing detection by analyzing large volumes of datasets and detecting hidden patterns. The objective is to study different machine learning techniques to improve the tool wear detection without interrupting the machining process, by measuring in-process with external sensors.

The rest of the paper is structured as follows. The section 2 is focused on the state of the art of the machining, tool wear sensing & prediction and artificial intelligence models.

The milling process, the necessary materials (e.g., the sensors), and the global methodology in the context of a machine learning approach (from dataset to model evaluation) are detailed in the following section. The section 4 is dedicated to the results for each approach, and, in the final section, the discussion of the results is presented along with the conclusion of the achieved results.

2 State of the art

The goal of this paper is to study the link between tool wear and the measures recorded by different sensors in the field of micro milling and high-speed milling. We aim to push forward the detection of tool wear and be able to identify the most interesting sensors for tool wear detection.

In the machining field, the wear of tools is due to multiple parameters, such as physical constraints, milling materials, etc. [5]. Tool wear management is one of the keys to optimize the milling process. Physical constraints (e.g., cutting forces, accelerations, vibrations, etc.) can be measured with sensors and used to determine the condition of the tool [6].

2.1 Tool wear sensing and prediction

Adaptive control systems aim at estimating the remaining useful life (RUL) of a tool in order to enable a fine control and optimization of the machining process. The lifetime of a tool can be decomposed into three consecutive stages designated as: break-in, steady-state, and failure [7]. After a relatively short break-in period characterized by a fast wear, the tool will pass into the steady-state in which the wear slowly increases. The last phase of the tool life cycle, failure, is characterized by a rapid deterioration of the tool at the end of which, the tool is no more usable. This degradation can be indirectly observed without stopping the production by using external sensors. This non-destructive method allows estimating the current wear of the tool. It can also be used to detect when the tool is too marked and needs to be replaced [21].

Different sensors can capture various aspects of the milling process and the related physical parameters. These sensors are used to model the behavior of one or multiple physical elements. For instance, the acoustic emission (AE) sensors are indicated as relevant sources of information [14,15]. A drawback of AE sensors is the noisy signal, and the influence of unrelated sounds and vibrations present in a machining environment. Less sensitive to the environment noise than AE, the accelerometers placed on the machine axes can also provide significant results to correlate the tool wear process and estimate the wear [11,23], that allow them to be less sensitive to the environmental noise than AE. Accelerometers provide information about vibration on the axes caused by the interaction between the tool and the material. The physical effort on an axis can be related to the power used by this axis. This power consumption can be measured by the electric current drawn by each motor. The relation between consumption and tool wear has been demonstrated by multiple authors, such as [11, 17]. Those various data sources can be used individually or can be merged to create a multivariate dataset as demonstrated by [11, 15].

To extract relevant information from data sources and reduce data size, few authors proposed to use features instead of raw data. Authors have demonstrated that spectrograms can be used instead of raw acoustic emission timeseries [2, 4]. The idea behind is to compute features that represents the frequency and his amplitude on a specific period. Statistical features can also be used in place of those computed ones which are less computational cost [11].

2.2 Machine Learning models

According to the literature, the detection of tool wear can be achieved by several approaches. Machine learning and artificial intelligence methods are widely investigated. Krishnakumar et al. [15, 16] used statistical features extracted from vibration signals to train classification models (decision tree, artificial neural network and support vector machine) to predict the stage of the tool wear. Cao et al. has established that features extracted from derived wavelet frames with Convolutional Neural Network (CNN) provide strong results [3].

Modeling tool wear states using deep learning approach was considered by several authors. Dou et al. [7] used a sparse auto-encoder model trained on vibration and force signals, Von Hahn et al. [10] chosed a disentangled-variational-autoencoder, with a temporal convolutional neural network, and Liu et al. [18] proposed a transformer-based neural network and a long-short memory network using temporal features extracted from raw signals.

3 Methodology

3.1 Context

The project focuses on the analysis of a *micro* milling machine, called micro5. The machine has 5-axis with less than 10 Kg of moving masses with high-speed machining (up to 60'000 revolutions/minutes). The kinematics of the machine combine three linear axes (X-Y-Z) and two rotary axes (B-C). It is corresponding to the type 57 from ISG-kernel [3]. All the data used in this project are produced by sensors integrated into this machine. This dataset is similar to the one introduced in [4] but has some major differences regarding the raw material used for machining and the sensors. The three principal differences are: the shape of the raw material (a cylindrical shape is used here instead of a cubic one), the number of AE sensors (reduced to one) and the use of multiple accelerometers.

Sensors The dataset is composed by signals acquired via three different types of sensors: electric current sensors (one for each axis), accelerometers (placed on different positions inside the machine) and an acoustic emission sensor.

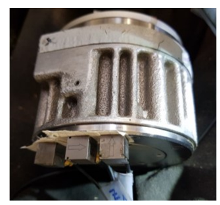

Fig. 1: C axis with accelerometers sensors

The AE sensor is located inside the milling machine, and positioned close to the raw material where the machining process is performed but not directly glued on the raw material as in [4]. The acoustic sensor is a Vallen VS45-H and the acquisition is realized with an Advantech PCIe 1840/L. The acquisition sampling is limited to 200 kHz. The sensor is positioned as close as possible that the raw material on the B axis, which allow replacing the raw material without moving the AE sensor.

Multiple accelerometer sensors are located on the spindle axis and the stator of C axis. Two different types of sensors have been used both from *Brüel & Kjær* (triaxial & uniaxial sensor / type 4520 & 4507).

Fig. 2: Spindle axis with accelerometers sensors

One 3-axis accelerometer sensor is located on the bottom of the spindle (for XYZ axis). Two one-axis accelerometers are located on the top of the spindle (XY axis). On the stator of C axis, three one-axis accelerometers are located which also measure XYZ axis. All sensors are acquired with National Instruments NI-9234 modules and sampled at 2 kHz. Figures 1 & 2 illustrates the sensor placement.

Both acoustic and accelerometers signals are synchronized with an external trigger which is recorded by the acquisition material. For those both data sources, the acquisition frequencies have been chosen by using equipment with a high acquisition frequency and wide bandwidth, in

[3] https://www.isg-stuttgart.de/en/

order to retain as much as possible of the information in the signal. The recorded signal are not digitialy filtered to keep as much as possible of embedded data. When some of the values are missing, the ongoing milling operation is ignored, as explained in Section 3.1.

All axis motors are driven by separated drive and the currents of each are stored during the milling. The acquisition sampling is limited to 1 kHz. For this data source, a digital filter is used to reduce noise from power supply and motor drive. The motor itself also performs as a filter, because it is a large inductive element.

Milling process The machining process consists in milling the raw material in multiple stairs; to create each stair, the milling path is again divided in several linear passes of identical width to remove the same quantity of material (see Figure 3). In the following, we will use the term of **experience** for one block of raw material machined where at the beginning the tool is new and at the end the tool breaks or is not more usable. The dataset is composed with a total of six experiences of the same machining path. The cutting technique used for manufacturing experiences parts is conventional milling.

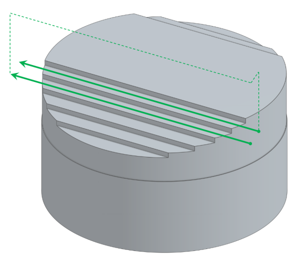

Fig. 3: Multiple linear passes

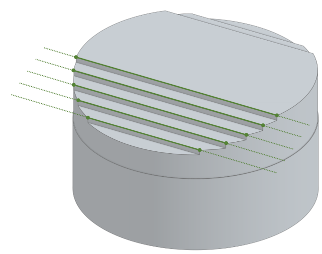

Fig. 4: Different lengths of linear passes

The linear milling path is achieved in three directions of the machine, in X (experiences one and four), in Y (experiences two and three) and in X+Y (experiences five and six). The experience parts are milled from a round bar and are sliced into cylinders. As the shape is a cylinder, the length of the linear passes tangential to the shape is shorter than the length of passes near the center of the raw material (as illustrated in Figure 4). This means that with this cylindric shape, the distance of the tool out of the material is not uniform and evolve between linear passes.

The tool has a diameter of 3mm and is made of tungsten carbide. It was replaced at the beginning of each experience. The rotation speed of the spindle and tool was set to 35'000 revolutions per minute (RPM) and was the same for all experiences. The milling of each part was achieved with pressurized air lubrication.

To automatically detect and segment when the tool is milling the raw material, and not just moving without contact, a variable threshold on the intensity of the acoustic signal is used. As first step to dynamically fix the threshold the signal is filtered using a rolling mean. Then, the difference between the maximum value of the signal in the central part of the pass and the maximal value at the beginning and ending of the pass is measured (the machining process design assures that the tool is not touching the material at the beginning and the end of each pass). The mid-point of the difference is used as the threshold to segment each linear pass. In detail, beginning and ending of a pass are defined as the first 5000 values and the 5000 latest values, which is approximately corresponding to 1ms (see Figure 5). This procedure is repeated for each linear pass.

In the Figure 5, three different signals are illustrated: the acoustic emission (grey), one of nine accelerometers (blue) and the spindle electric current (red). The transient phases of a milling pass is indicated in green.

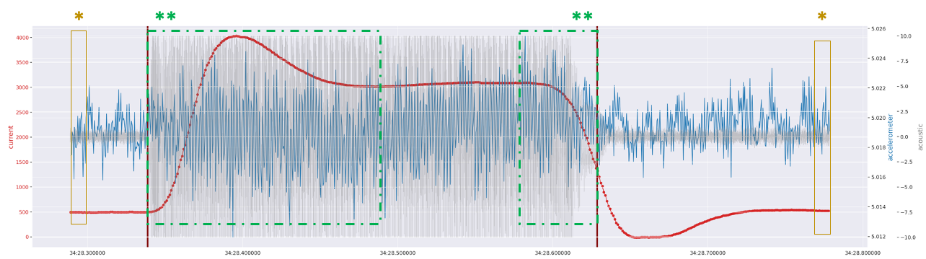

Fig. 5: Signals detected for a typical machining pass. * time window for maximal value detection (in yellow); ** transient phases during milling (in green). Three different signals are illustrated: the acoustic emission (grey), one of nine accelerometers (blue) and the spindle electric current (red).

Objectives of the study In this article, we focus on the milling process and specifically on the step when the tool removes material. An experience is considered completed, if, at the end of the milling process, a tool is considered as no more usable (broken, burned out, etc.). The main goal is to construct a model able to predict the remaining useful life (RUL) of the tool. In order to apply classification models, the continuous variable RUL must be firstly discretized. We propose to use five states (or labels), denoted from "0" (corresponding to a new tool) to "4" (corresponding to a non-usable tool), to perform a finer tool wear evaluation. We have chosen that at the end of the milling operations, the tool is no longer usable and receives the label "4" (or "worst"). The label "0" (or "the best") corresponds to the first passes when the tool is new. During the milling process, the label of the tool passes from "0" (beginning of the process) to "4" (end of the process). Two different time related methods are used to distribute the remaining labels. The methods are detailed for each approach in each section below.

Table 1 presents several approaches (or strategies) that we considered and evaluated in this paper. The split between train set and test set follows the k-fold cross-validation methodology [12].

Table 1: Approaches evaluated in this article

Section	Title
3.2.	Initial approach: dataset validation and baseline
3.3.	Focus on machine learning algorithms using statistical features
3.4.	Focus on stationary components of the signals
3.5.	Mixing spectrogram approach and stationary components of the signals

As the dataset contains outliners and peculiar values due to the implicit process variability, the dataset is shuffled multiple times and split into multiple folds in order to mitigate such effects. To evaluate the generalized performance of the classification model we apply a cross-validation methodology using three folds. Therefore, the ratio between training and test set is fixed to 66%-33% (train and test respectively). The performance measure used to compare different classification models is the mean of F1 score, returned by the cross-validation process. The F1 score is a measure that balances the precision and recall of a classification model, providing a measure of its overall precision. The processing time of classification is not considered because it is insignificant (few milliseconds).

It is worth to notice that initially in this multi-source/experience dataset, there are some missing data in different experiences. As missing values into one data source can impact the comparison among other data sources, it has been chosen to mitigate this problem by removing those linear passes in all experiences. For instance, if at some point data values are missing for an accelerometer the whole pass is discarded and this is propagated to the other data sources. For the acoustic emission and electric current signals, the same methodology is used. Those removed passes represent around 5.5% of the whole dataset. This technique has been applied to all approaches in Table 1.

3.2 Initial approach: dataset validation and baseline

This initial approach has the goal to validate the dataset and create a baseline to be used as initial comparison point among each data sources. The dataset is processed to extract spectrograms from the acoustic emission signal with a length of around 200ms (39936 values). Only acoustic emission during the milling is used. Depending on the length of the pass one or two spectrograms are extracted from each milling pass. Spectrograms are converted into images and are used to train a CNN. As an image used by the CNN has a fixed size (144px by 144px), only the linear passes allowing them to be a multiple of the spectrogram size are used. Too short linear passes are not considered for the creation of the spectrograms. The training of the CNN in this initial approach is repeated 10 times which allows us to stabilize the learning process cross-validation.

In this initial approach, the labeling of the five classes is based on a discrete uniform distribution. In other terms, for each experience, the sequence of all passes is divided in five equal size classes. This methodology will be considered as the baseline performances achieved by analyzing an acoustic emission dataset in the following sections. In addition, we have tried to improve the model performance by forcing the learning process to diagonalize the previous multiclass classification problem, using the implicit class order ("0" > "1" > .. > "4"). This method is also known as *ordinal classification* or *ranking learning* . The idea is more in line with the final objective which is to determine the remaining life of the tool. In this case, the impact of the prediction error is greater if the prediction is further from the truth. For instance, mispredicting the class 0 (the tool is new) for a class 4 (the tool is completely burnt out) should be evaluated considerably worse than mistaking the class 0 with a class 1 (new tool vs tool slightly worn out). As explained by Gaudette and Japkowicz [8], the metrics RMSE (Root-Mean-Square Error) or MSE (Mean Square Error) perform better for this kind of classification.

3.3 Focus on machine learning algorithms using statistical features

This approach, decomposed in two steps, is based on the definition and usage of statistical features to investigate performance of machine learning algorithms using the different data sources. The initial step is, indeed, dedicated to exploring a wide range of machine learning algorithms and statistical features; the second step is devoted to optimizing the most promising algorithms by tweaking hyperparameters of models.

During the exploration phase, multiple classification algorithms with default parameters are compared. In this phase, the *lazypredict* library [19] is used to evaluate a wide variety of algorithms which are based on scikit-learn [20] models.

The features used by algorithms have an important impact on the models performance. A previous project [4] using the same database provided a reliable performance to determine the level of wear of a tool. Therefore, the choice of the features is made such that to obtain at least a similar performance. In this project, we evaluated the correlation between acoustic emission and the tool wear for a set of features including: mean, std, min, max, first quartile, second quartile, third quartile. All features are extracted from the window provided from the raw signals. The data is divided into three sub-datasets (acoustic emission, accelerometers and currents) and the extraction process generates three separated data sources. Each feature sub-dataset is composed of features of each signal in the data source.

The labeling process used in this approach is also based on five classes, but the label affectation method is different. Instead of equally distributing the sequence of linear passes in the five classes, in this approach we consider that only the actual machining time affect the tool wear (when the tool is not touching the material there is no wear). Practically, the time interval between the beginning and the end of the experience is equally divided in five sub-intervals, and all passes occurred in a given sub-interval receive the same label. This labeling method is more consistent with the tool wear profile (as presented in Section 3.2). This approach slightly impacts the label balancing, but it allows conserving an accurate distribution between each class.

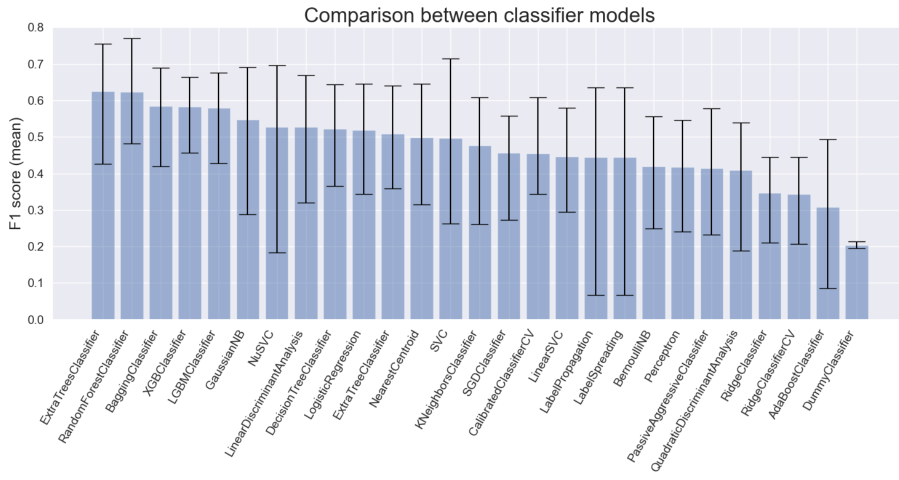

Fig. 6: Result of algorithms comparison using currents as data source

Figure 6 presents the performance (90% confidence interval for F1 score) of different classification models using the electric current signals. On other data sources (accelerometers and acoustic emission), algorithms' performances are weaker.

Based on F1 score performance metric, we selected the two best performing algorithms for fine tuning: Extra trees classifier and Random Forest classifier. Fine tuning mostly consisted in hyperparameters optimization. For this second step, the data (in our case the features) used for the training is the same as for the previous step, only the hyperparameters are modified (results are presented below in the section 4).

3.4 Focus on stationary components of the signals

Each linear pass in the material is composed of transient phases (when the tool impacts, enters the material and when the tool gradually leaves the material) and a stationary phase (the tool is completely in the material) – as illustrated in Figure 5. After machining signals analysis, the transient phase affects the shape of the signals by generating spikes and noise. It was chosen to exclude the transient phases and focus on the stationary components of the signals to improve the prediction performance. Accordingly, the approach presented in this section focuses on the stationary components of the signals for feature extraction. The insight behind this approach is to leave out the transient regimes of each pass since they generate the noisy signals in the machining process. This choice has been driven by the observation that the spindle electric current and the shape of the transient regimes is totally different between the transient states and the stationary state (Figure 5).

To suppress the transient signals, the timeframe of both transient states (entering and exiting the material) has been measured during all the milling process. We assessed that the duration of the "entering" transient is less than 150ms and "exiting" transient is around 50ms. By removing transient states, only the stationary part of each linear pass of the signal is kept for feature extraction and classification. Concerning the electric current signals, the values of the spindle and nonmoving axes are steady during the stable milling interval. For the acoustic emission, we expect to get a cleaner signal with stable frequencies. Regarding the accelerometer values, the vibration should be steadier as the effort is nearly constant. As a final step, to conserve only meaning full signals, it was chosen to retain only stationary signal longer than 100ms.

3.5 Mixing spectrogram approach and stationary components of the signals

The latest enhancement explored in this article is to use spectrograms and CNN as in [4] by improving the signal pre-processing phase (detailed in the previous sections). In this approach we intend to associate the stationary components of the signals [presented in Section 3.4] with the spectrogram. The idea behind this mixed approach is to quantify the impact of transient phase which impacts the acoustic emission.

Using only the stationary components of the signal means that the spectrogram duration will correspond to the new size of the signal which is around half of the length of the original signal. With this approach, the spectrogram representation does not contain the frequencies generated by the transient states of milling, which can pollute spectrograms.

The labeling method used is the same as the one applied in the stationary signals section (labeling on milling duration).

4 Results

The result section is divided into four parts, one for each approach. For all the approaches, the same dataset is used but the data sources and pre-processing change. In each section, the result is detailed with short discussion.

Initial approach: dataset validation and baseline - The f1-score achieved for this approach is 35% on five classes, which we cannot be used in the field. This result could come from the faulty acoustic data. By modifying the loss function to minimize the diagonal spread (*ordinal classification*), the result is worse with a f1 score decreasing to 25%.

Focus on machine learning algorithms using statistical features - The best result provides by the algorithms comparison (Figure 6) is the *ExtraTreesClassifier* with a f1-score of 62% (ex-aequo with the *RandomForestClassifier*). For the other data sources, only the score of the best algorithm is presented. This score is achieved with the electric current signals. The second data source is the AE with a f1-score of 43%. The f1-score for the latest data source (accelerometers) is 42%. The merged data sources obtain a f1-score of 58%. By optimizing all algorithm parameters described in scikit-learn over 1000 runs with

Optuna [1], the f1-score is increased up to 66% on both two top algorithms described below. Optuna is a hyperparameter optimization framework designed to automate the hyperparameter search of machine learning models. It uses efficient algorithms to explore the search space and find optimal configurations, making it an interesting tool for improving model performance.

Focus on stationary components of the signals - In this approach, the same two steps methodology is followed. For the classification of all algorithms, once again the *ExtraTreesClassifier* and the *RandomForestClassifier* reach first with a f1-score of 69% with the electric current signals. For the other data sources, the merged data sources obtain a f1-score of 68%, followed by accelerometers with a f1-score of 49%. Finally, the f1-score of AE data source is 47%. The Optuna optimized f1-score for the best algorithm is increased up to 73% with the confusion matrix shown in Figure 7.

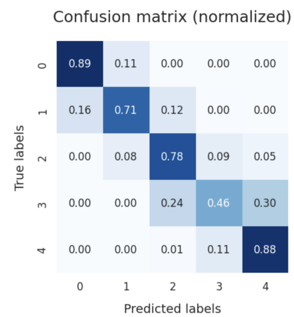

Fig. 7: Classification performance of ExtraTreesClassifier with Optuna optimization

Mixing spectrogram approach and stationary components of the signals - The f1-score of this approach is 38%, and it can be compared to the baseline approach. It can be established that there is no significant difference.

Based on the results presented in the section, it is evident that the spectrogram approaches did not provide satisfying results, while the approaches based on features showed promising results. By optimizing the hyperparameters of the algorithms used in these approaches, the f1-score was increased up to 66% and 73%.

5 Conclusion – Discussion

In this article, we present different approaches to predict machining tool wear in the milling process of micro-machining. This optimization process can significantly improve the milling process and reduce manufacturing costs. The objective of this article is to explore different machine learning techniques to enhance tool wear detection by measuring in-process with three different data sources (acoustic emission, accelerometers and axis currents), without interrupting the machining process.

Between three different data sources, the electrical currents of the machine perform the best. The AE dataset did not provide a reliable result in opposite to other papers. This could be due to an issue during the recording or to a noisy environment. The extra trees algorithm and random forest algorithm provide the best results and are largely on the top during algorithm comparisons. By optimizing the hyperparameters of the extra trees algorithm, the model achieved a f1-score of 73% on five classes. To clarify the issue with AE dataset, new additional machinings needs to be conducted to establish if the adoption of this type of data is relevant.

No filtering of acoustic emission or accelerometer signals has been performed to keep as much as possible of embedded information. A specific study could be interesting to determine the impact on classification performance.

Funding: This research was funded by the CHIST-ERA program (SOON project) and the University of Applied Sciences and Arts Western Switzerland (HES-SO).

Acknowlegements: These results could not have been achieved without the help of the technical team that collected the datasets (instrumented of the machine, design and implemented the machining plan), and the data analysis team that pre-processed and explored datasets. In particular, the authors want to thank Massimo De Santis, Edouard Goffinet, Jonathan Guerne and Célien Donzé.

References

1. Akiba, T., Sano, S., Yanase, T., Ohta, T., Koyama, M.: Optuna: A Next-generation Hyperparameter Optimization Framework. In: Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (2019)
2. Arslan, M., Kamal, K., Sheikh, M.F., Khan, M.A., Ratlamwala, T.A.H., Hussain, G., Alkahtani, M.: Tool Health Monitoring Using Airborne Acoustic Emission and Convolutional Neural Networks: A Deep Learning Approach. Applied Sciences **11**(6), 2734 (2021). https://doi.org/10.3390/app11062734
3. Cao, X., Chen, B., Yao, B., Zhuang, S.: An Intelligent Milling Tool Wear Monitoring Methodology Based on Convolutional Neural Network with Derived Wavelet Frames Coefficient. Applied Sciences **9**(18), 3912 (2019). https://doi.org/10.3390/app9183912
4. Carrino, S., Guerne, J., Dreyer, J., Ghorbel, H., Schorderet, A., Montavon, R.: Machining Quality Prediction Using Acoustic Sensors and Machine Learning. Proceedings **63**(1), 31 (2020). https://doi.org/10.3390/proceedings2020063031
5. Dehen, S., Segebade, E., Gerstenmeyer, M., Zanger, F., Schulze, V.: Milling parameter and tool wear dependent surface quality in micro-milling of brass. Procedia CIRP **87**, 95–100 (2020). https://doi.org/10.1016/j.procir.2020.02.024
6. Dornfeld, D., Lee, Y., Chang, A.: Monitoring of Ultraprecision Machining Processes. The International Journal of Advanced Manufacturing Technology **21**(8), 571–578 (2003). https://doi.org/10.1007/s00170-002-1294-2
7. Dou, J., Xu, C., Jiao, S., Li, B., Zhang, J., Xu, X.: An unsupervised online monitoring method for tool wear using a sparse auto-encoder. The International Journal of Advanced Manufacturing Technology **106**(5-6), 2493–2507 (2020). https://doi.org/10.1007/s00170-019-04788-7
8. Gaudette, L., Japkowicz, N.: Evaluation Methods for Ordinal Classification. In: Gao, Y., Japkowicz, N. (eds.) Advances in Artificial Intelligence, vol. 5549, pp. 207–210. Springer Berlin Heidelberg, Berlin, Heidelberg (2009). https://doi.org/10.1007/978-3-642-01818-3_25, series Title: Lecture Notes in Computer Science

9. Grimske, S., Kong, N., Röhlig, B., Wulfsberg, J.P.: Square Foot Manufacturing—A Modular and Mutable Desktop Machine Tool System. Mechanics Based Design of Structures and Machines **42**(3), 386–397 (2014). https://doi.org/10.1080/15397734.2014.908728
10. Hahn, T.V., Mechefske, C.K.: Self-supervised learning for tool wear monitoring with a disentangled-variational-autoencoder. International Journal of Hydromechatronics **4**(1), 69 (2021). https://doi.org/10.1504/IJHM.2021.114174
11. Harris, K., Triantafyllopoulos, K., Stillman, E., McLeay, T.: A Multivariate Control Chart for Autocorrelated Tool Wear Processes. Quality and Reliability Engineering International **32**(6), 2093–2106 (2016). https://doi.org/10.1002/qre.2032
12. Hastie, T., Tibshirani, R., Friedman, J.: The Elements of Statistical Learning. Springer Series in Statistics, Springer New York, New York, NY (2009). https://doi.org/10.1007/978-0-387-84858-7
13. Jain, A., Bajpai, V.: Introduction to high-speed machining (HSM). In: High Speed Machining, pp. 1–25. Elsevier (2020). https://doi.org/10.1016/B978-0-12-815020-7.00001-1
14. Kishawy, H.A., Hegab, H., Umer, U., Mohany, A.: Application of acoustic emissions in machining processes: analysis and critical review. The International Journal of Advanced Manufacturing Technology **98**(5-8), 1391–1407 (2018). https://doi.org/10.1007/s00170-018-2341-y
15. Krishnakumar, P., Rameshkumar, K., Ramachandran, K.I.: Acoustic Emission-Based Tool Condition Classification in a Precision High-Speed Machining of Titanium Alloy: A Machine Learning Approach. International Journal of Computational Intelligence and Applications **17**(03), 1850017 (2018). https://doi.org/10.1142/S1469026818500177
16. Krishnakumar, P., Rameshkumar, K., Ramachandran, K.: Tool Wear Condition Prediction Using Vibration Signals in High Speed Machining (HSM) of Titanium (Ti-6Al-4V) Alloy. Procedia Computer Science **50**, 270–275 (2015). https://doi.org/10.1016/j.procs.2015.04.049
17. Lauro, C., Brandão, L., Baldo, D., Reis, R., Davim, J.: Monitoring and processing signal applied in machining processes – A review. Measurement **58**, 73–86 (2014). https://doi.org/10.1016/j.measurement.2014.08.035
18. Liu, H., Liu, Z., Jia, W., Lin, X., Zhang, S.: A novel transformer-based neural network model for tool wear estimation. Measurement Science and Technology **31**(6), 065106 (2020). https://doi.org/10.1088/1361-6501/ab7282
19. Pandala, S.R.: Lazy Predict (2022), https://github.com/shankarpandala/lazypredict
20. Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M., Duchesnay, E.: Scikit-learn: Machine Learning in Python. Journal of Machine Learning Research **12**, 2825–2830 (2011)
21. Teti, R., Jemielniak, K., O'Donnell, G., Dornfeld, D.: Advanced monitoring of machining operations. CIRP Annals **59**(2), 717–739 (2010). https://doi.org/10.1016/j.cirp.2010.05.010
22. Wulfsberg, J.P., Redlich, T., Kohrs, P.: Square Foot Manufacturing: a new production concept for micro manufacturing. Production Engineering **4**(1), 75–83 (2010). https://doi.org/10.1007/s11740-009-0193-x
23. Yu, J., Liang, S., Tang, D., Liu, H.: A weighted hidden Markov model approach for continuous-state tool wear monitoring and tool life prediction. The International Journal of Advanced Manufacturing Technology **91**(1-4), 201–211 (2017). https://doi.org/10.1007/s00170-016-9711-0