

Article

Learning More with Less Data in Manufacturing: The Case of Turning Tool Wear Assessment through Active and Transfer Learning

Alexios Papacharalampopoulos * , Kosmas Alexopoulos , Paolo Catti, Panagiotis Stavropoulos 
and George Chryssolouris

Laboratory for Manufacturing Systems & Automation (LMS), Mechanical Engineering & Aeronautics Department, University of Patras, 26504 Patras, Greece; alexokos@lms.mech.upatras.gr (K.A.); catti@lms.mech.upatras.gr (P.C.); pstavr@lms.mech.upatras.gr (P.S.); xrisol@lms.mech.upatras.gr (G.C.)

* Correspondence: apapacharal@lms.mech.upatras.gr; Tel.: +30-2610-91-01-60

Abstract: Monitoring tool wear is key for the optimization of manufacturing processes. To achieve this, machine learning (ML) has provided mechanisms that work adequately on setups that measure the cutting force of a tool through the use of force sensors. However, given the increased focus on sustainability, i.e., in the context of reducing complexity, time and energy consumption required to train ML algorithms on large datasets dictate the use of smaller samples for training. Herein, the concepts of active learning (AL) and transfer learning (TL) are simultaneously studied concerning their ability to meet the aforementioned objective. A method is presented which utilizes AL for training ML models with less data and then it utilizes TL to further reduce the need for training data when ML models are transferred from one industrial case to another. The method is tested and verified upon an industrially relevant scenario to estimate the tool wear during the turning process of two manufacturing companies. The results indicated that through the application of the AL and TL methodologies, in both companies, it was possible to achieve high accuracy during the training of the final model (1 and 0.93 for manufacturing companies B and A, respectively). Additionally, reproducibility of the results has been tested to strengthen the outcomes of this study, resulting in a small standard deviation of 0.031 in the performance metrics used to evaluate the models. Thus, the novelty presented in this paper is the presentation of a straightforward approach to apply AL and TL in the context of tool wear classification to reduce the dependency on large amounts of high-quality data. The results show that the synergetic combination of AL with TL can reduce the need for data required for training ML models for tool wear prediction.

Keywords: tool wear; active learning; transfer learning; data management



Citation: Papacharalampopoulos, A.; Alexopoulos, K.; Catti, P.; Stavropoulos, P.; Chryssolouris, G. Learning More with Less Data in Manufacturing: The Case of Turning Tool Wear Assessment through Active and Transfer Learning. *Processes* **2024**, *12*, 1262. <https://doi.org/10.3390/pr12061262>

Academic Editors: Xiaosong Du and Devina P. Sanjaya

Received: 24 May 2024

Revised: 15 June 2024

Accepted: 18 June 2024

Published: 19 June 2024



Copyright: © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Traditionally, tool wear monitoring has been a manual procedure. As detailed in [1], one of the main traditional approaches to measuring tool wear is visually inspecting the tool for signs of wear such as crater wear, flank wear and notch wear. Microscopes have also been used to extract precise measurements of wear patterns on the tool, which has allowed for a detailed examination of the tool's surface [2]. However, such traditional techniques often rely on human expertise, making them time-consuming and potentially less reliable and limited in terms of precision [2,3].

Monitoring tool wear levels indirectly has been an active research objective in recent years, implicating many different sensors and machine learning methods [1–4]. However, it is not straightforward, since information is concealed well in the signals, rendering the uncertainty in monitoring high [5]. Various methodologies for this exist, involving statistical measures [6], multimodal inputs [7], deep networks [8] and intense image processing [9]. These procedures can be enriched with various types of physics models [10,11], or even

take into account accumulative information [12]. Computational techniques seem to be of great added value to this prediction procedure [13], as they provide insights into where the information lies. Despite all the above, tool wear monitoring is still open.

The use of ML algorithms towards classifying or predicting tool wear in cutting operations has significantly increased. In [14], the performance of six ML and deep learning algorithms was tested to identify the best performer in predicting the cutting tool's condition in milling operations. In [15], time-series images were used along with CNNs to predict tool wear in machining processes. Lastly, in [16], ML algorithms were tested to classify tool wear in milling-based operations using the cutting force and current signals.

The type of machining process is crucial to the way the signals are produced. In general, the contact of the cutting tool is that which produces these signals. Depending on the process and the number of cutting edges, the force evolution may be different.

The tool wear itself is also a complicated phenomenon. In general, there are two main types of tool wear—crater, which occurs when material from the tool is displaced and forms a crater on the tool's edge, and flank wear, which occurs on the flank of the cutting tool (Figure 1) [17].

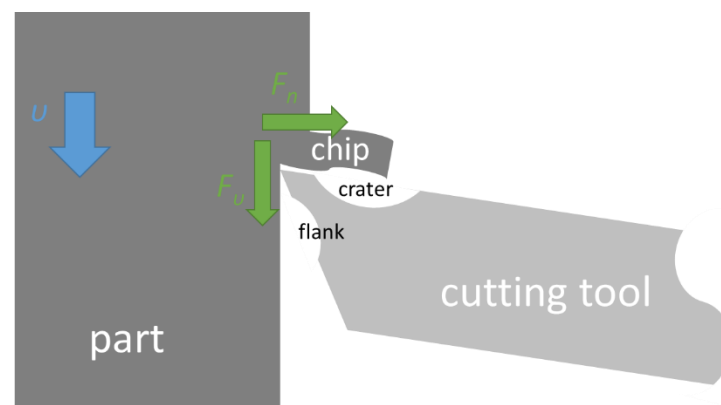


Figure 1. Methodology for investigating active learning in tool wear. The forces for this 2D configuration are also depicted (F_v and F_n , components for parallel and perpendicular to cutting speed v , respectively).

In addition to the machine configuration adding to the complexity of the signals retrieved, uncertainties obscure the mechanism that information related to wear is hiding within the signals. These are related to parameters' local (spatially or temporally) differentiations, which may occur due to randomness in previous manufacturing stages or intrinsic to the material. These uncertainties can be related to material property differentiations [18], process parameter variations [19], machine-related uncertainties, the transmission of signals and the sensors' capabilities. Additional variations are those of measurements [20,21] as well as microstructure [22]. All these add up, resulting in distortion of signals.

The main measurements involve cutting force estimation, and relevant quantities, such as electric current [23]. Fused data from different sensors have also been used to the same end [24]. This is because tool wear changes the interface between the tool and the part, changing the dynamics of the process. The process also affects the signals acquired.

The leading approaches in data-driven learning, such as deep learning, require huge amounts of quality data to generate meaningful results. However, many domains do not have access to such data because acquiring data is an expensive or time-consuming process or both [25]. Data-driven AI methods have been introduced to several manufacturing applications [4], enabled by the increasing degree of digitalization in manufacturing [26]. Several AI models for use in manufacturing applications are data hungry and acquiring data for their development might be expensive or error prone [27].

In the meantime, digital twins emerge, including multidisciplinary tools, ranging from machine learning [28] to systemic approaches [29]. This could help to respond to what-if

scenarios, but also to more elaborate modelling techniques. Modelling is expected to be the first step towards creating a nominal case and studying variations on top of it. This is highly relevant to the concept of milling in an adaptive environment [30,31].

All this has led to a need for more and more data, so that elaborated models, such as deep learning ones can achieve classification/regression without any overfitting. This is a typical case in manufacturing, i.e., in batteries' welding [32], for DED [33] and in digital manufacturing in general [34]. In addition, inference has also been studied under the context of tool wear [35]. This is expected to be achieved through statistical significance [36].

Simultaneously, the use of small data has been a global trend [37,38] and additionally, the research in [39] is targeting the creation of neural networks of N classes from $M < N$ samples. The topic is related to the so-called data-efficient algorithms [25].

Transfer learning can also be useful [40,41] as well as pretraining [42]. In addition, there are more focused techniques on limited data applications, like active learning [43] (a technique partly applied also in tool wear detection [44]), while physics-informed techniques [45,46] could be beneficial here, and more elaborated cases would require advanced techniques, such as neural operators [47]. In [48], online active learning has been employed for automated visual inspection, which has reduced the data labelling effort by 15% while keeping an acceptable classification performance.

Federated learning [49] allows the sharing of data and models between companies (Figure 2) to provide extra value to both of them, by allowing retraining of models locally and retrieving more elaborated machine learning models. It could be considered as an additional strategy attempting to reduce the data volume needed. However, generally in decentralized systems, there are privacy concerns that should be tackled before applying such a methodology [50]. Federated learning raises privacy concerns that include unintentional data leakage and model reconstruction attacks as presented in [51], where differential privacy and secure multi-party computation are suggested as a countermeasure. Additionally, in [52], challenges in privacy that include inference and model extraction attacks are documented and the countermeasures presented include several cryptographic methodologies, server cleaning and robust federated learning aggregation. Herein, a simplified direct form of simple transfer learning is applied.

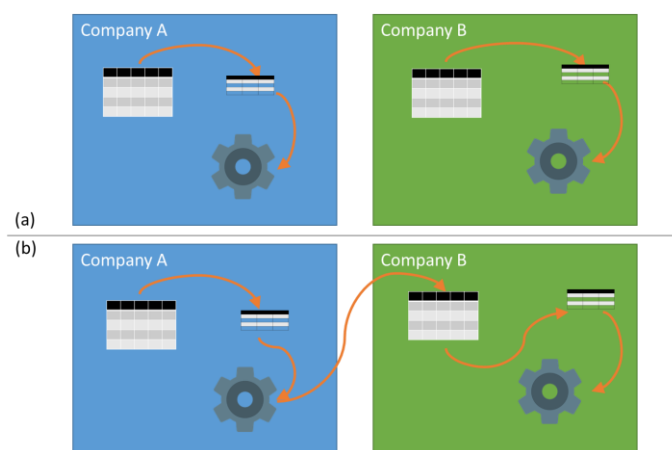


Figure 2. Relationship between two different companies. Independent models for active learning case (a) vs. shared data/models and transfer active learning case (b). The size of the data icon annotates the size of the dataset, while the gear is a symbol of ML training.

In the current work, the amount of data required for tool wear monitoring for the case of turning is studied in a two-fold way: firstly, active learning is investigated and then model transfer from one company to another is studied. The combination of applying AL and TL in the context of tool wear monitoring is unique.

2. Materials and Methods

The overall method is presented in Figure 3. The first step involves AL individually at each company, aiming at reducing the dataset size required for training the models. Figure 4 summarizes the procedure of AL. The idea behind this approach is straight forward; the data which have an impact to the utmost extent are used, with a free parameter used to control the percentage of uncertainty of data utilized. In principle, an ML model is trained using a low amount of data. To reduce the dependency on large datasets, 30% or less of the dataset is used for the initial training of the ML algorithm. The selection of 30% or less of the dataset enables the ML model to capture the underlying structure and complexity of the dataset, without using a large portion of it to reduce the computational time that would have been required if a larger portion of it was used. This selection is made randomly to ensure that no bias is introduced to the model during the initial training phase. To identify which data points in a dataset have the most impact on the model's performance, the model evaluates itself on the remaining 70% of the dataset. With each classification performed during the model evaluation, a percentage of uncertainty included in each classification is calculated. Data points that could provide the greatest insight to the algorithm are those of the ML algorithm that have a bigger than desired uncertainty in its performed classification during the validation step. Through this process the algorithm is fed with the most informative data points thus, reducing its dependency on low-value additional data. To increase the robustness of the final model while avoiding overfitting, the uncertainty threshold selected is 5%. When the individual classification uncertainty exceeds 5%, the data are fed to the training set and the model is retrained. The uncertainty level of 5% was selected to make sure that the model is provided with data whose uncertainty level, during the evaluation step, is of high statistical importance. Increasing the threshold would result in data points being omitted while they could provide meaningful information to the model while lowering the threshold would provide the model with more data points, which would increase the computational time during the retraining phase and potentially increase overfitting since data points whose uncertainty is low may show similarities to data points the model has been already been exposed to. Lastly, performance metrics are calculated, for the initial model (without AL) as well as for the final model (with AL).

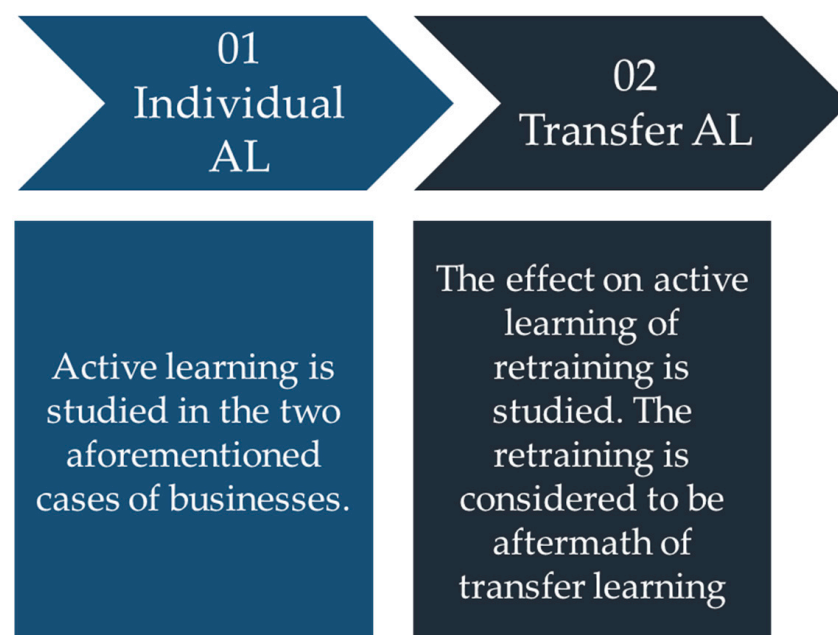


Figure 3. Methodology for investigating active learning in tool wear.

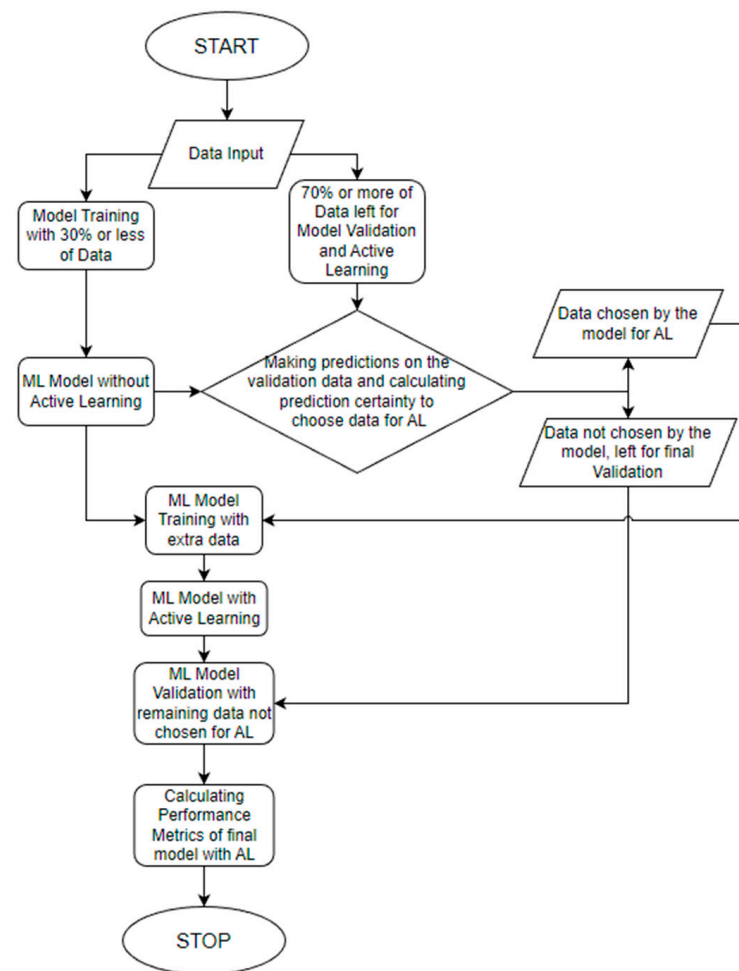


Figure 4. Active learning method.

In the second step of Figure 3, TL is applied from one company to another, using AL at each one of them; to this end, TL on top of AL is applied. More specifically, AL is applied in Company A and then the model is transferred to Company B, where AL is re-performed. This is performed to prove that active and transfer learning can work synergistically. It is noted that this has both business and technical implications, which are discussed hereafter; however, for the sake of simplicity, they are not taken into consideration.

3. Pilot Case

3.1. Overview

To test and validate the method proposed, an industrially relevant pilot case has been defined and executed. The pilot case considers two manufacturing companies, A and B. For each company, a dataset with tool wear-related information has been defined. The data included in both datasets were normalized and the min–max normalization method was used. Both datasets follow the same format as depicted in Table 1.

Table 1. The structure of the datasets.

Force 1 (Fx)	Force 2 (Fz)	Tool Wear
value 1	value 2	low or 0 tool wear
value 3	value 4	low or 0 tool wear
value 5	value 6	middle or 0.1 tool wear
value 7	value 8	high or 0.3 tool wear

The statistical pattern based on real data from the industry presented in the literature [5], is defined as Company A dataset. Since the axes of the diagram correspond to measured forces F_a (Force 1) and F_b (Force 2), the generated values can be on $F_a + r_1 = a(F_b + r_2) + b$, where r_1, r_2 are random numbers characterizing the deviation (randomness) and a, b are real numbers defining the average linear tendency of the data. Figure 5 displays a sample dataset, through statistical data generation. Data generation was done using MATLAB (R2023a, MathWorks, Natick, MA, USA). The two forces are components of the same force, as indicated in Figure 1. Regarding Company A, force components were measured in the feed direction, in the radial direction and in the cutting direction. The measurements were conducted using a dynamometer on a piezoelectric measuring platform located on the tool holder. Additionally, a similar case of turning, Company B dataset [50] has been considered, from real data, as well. The dataset has been tested experimentally in [53] and it consists of single measurements of forces for different process parameters and different tool wear levels. The forces correspond to different spatial force components. Two of them are visualized in Figures 6 and 7 and it seems that it is not clear what geometrical pattern the classes follow. For both companies, some indicative data regarding the experimental setups are given in Table 2.

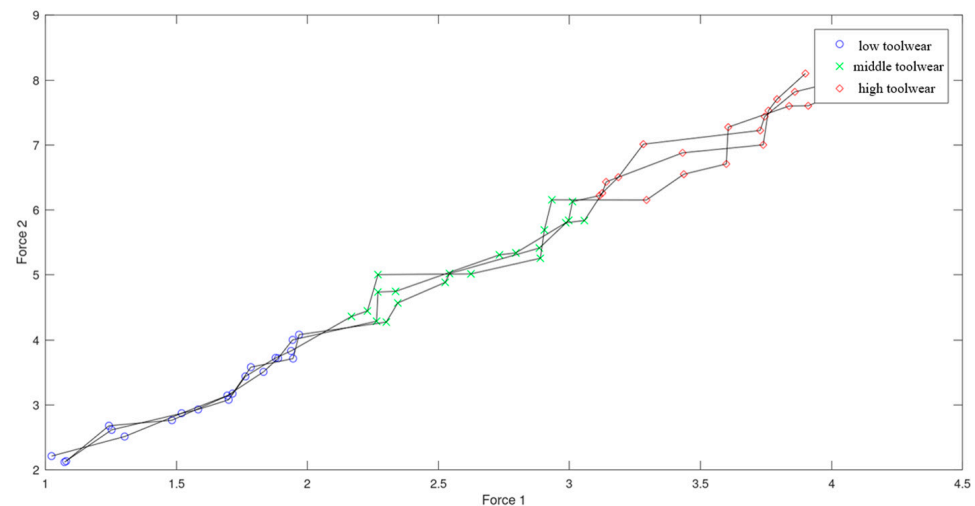


Figure 5. Generation of force measurements based on statistical patterns (presented as time series)—Company A data.

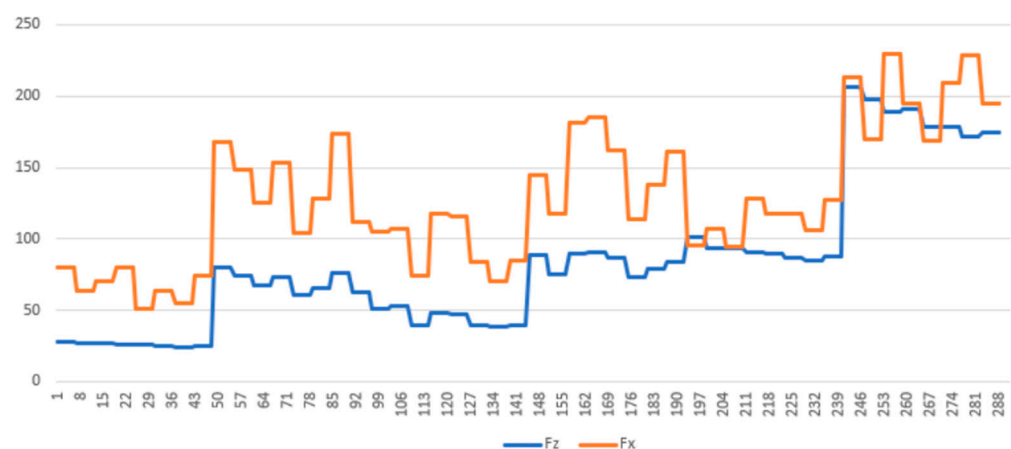


Figure 6. Company B dataset visualisation.

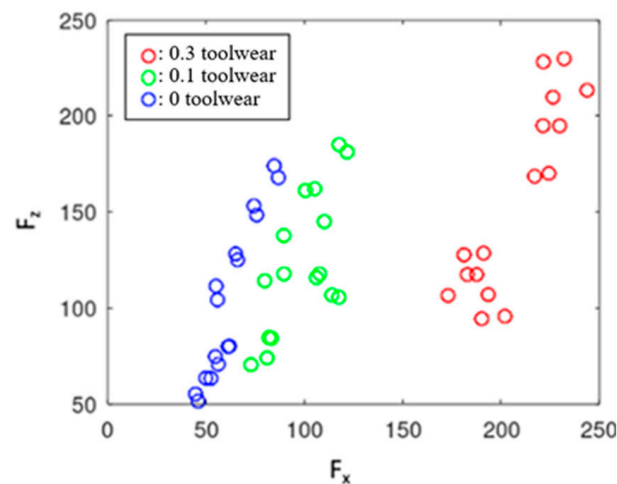


Figure 7. Company B classes visualization based on two components of force, including cutting force.

Table 2. Experimental Setup Information.

Experimental Setup	Company A [5]	Company B [50,53]
Machine tool/Cutting parameters	DMU/available in the documentation	ROMI/Available in the thesis
Material	Titanium alloy (TC4)	AISI H13
Cutting Tool/Coating/Fluid	Carbide/None/no	HC/CVD/Yes
Sensor	Spike sensory tool holder	Kistler (more details in the thesis)
Placement of the sensor	In the tool holder (schematic in the documentation)	Attached to cutting tool (schematic in the thesis)
Measurements	Mean value (configuration of path and sampling frequency can be found in the documentation)	Mean value of plateau in steady state (schematic in the thesis)

3.2. Machine Learning

Initially, typical machine learning (ML) was applied to both Company A and Company B datasets. Several ML methods were applied, more specifically the Decision Tree Classifier [54], the Support Vector Classifier [55], the Multi-Layer Perceptron Classifier [56] and a Custom Sequential Model (CSM). Based on the results, the CSM was selected as it outperformed the rest. The models were created in the Visual Studio Code source-editor (version 1.77, Microsoft, Redmond, WA, USA), using Python (version 3.9.12, Python Software Foundation, Beaverton, OR, USA) and its TensorFlow library (version 2.12, Google, Mountain View, CA, USA). Its architecture is presented in Figure 8. It consists of:

- A Dense layer with 64 units and its activation function set to 'elu';
- A Dense layer with 32 units and its activation function set to 'elu';
- A Dense layer with 16 units and its activation function set to 'elu';
- A last Dense layer with 3 units and its activation function set to 'softmax';
- The CSM was trained using the following parameters:
- An 80–20% split between training and validation data;
- Batch size was set to 4;
- The Adam [57] optimizer was utilized;
- The learning rate was set at 0.0001;
- The 'categorical_crossentropy' loss function was used;
- The maximum epoch size was set to 50, and an early stopping function was deployed to monitor the validation loss, with patience of 3 and the ability to restore the best weights.

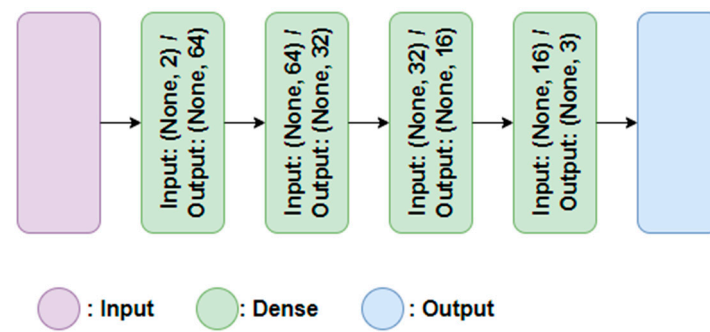


Figure 8. CSM architecture.

To evaluate the performance of the CSM, four performance metrics have been utilized. These include accuracy, precision, recall and F1 score. Accuracy represents the correctness of the model. Precision quantifies the truly positive classification among all positive instances. Recall identifies the truly negative identified classifications among all positive instances. Lastly, the F1 score is the harmonic mean of precision and recall. They are calculated based on the following equations (Equations (1), (2), (3), (4), respectively).

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \quad (1)$$

$$Precision = \frac{TP}{TP + FP} \quad (2)$$

$$Recall = \frac{TP}{TP + FN} \quad (3)$$

$$F1 \text{ score} = 2 \times \frac{Precision \times Recall}{Precision + Recall} \quad (4)$$

where

- TP : true positives;
- FP : false positives;
- FN : false negatives;
- TN : true negatives.

For Company A, whose dataset is balanced between the three classes as seen in Figure 5, the CSM was trained with 480 data points and validated using 120 unique data points. The model concluded training after the completion of 11 epochs. Similarly, the CSM was trained with 229 data points from Company B, whose dataset is also balanced with a similar amount of data points between classes as seen in Figure 6 and the rest 69 data points were used for validation. The training of the model was concluded after the completion of 7 epochs. In Table 3, the performance of the models can be seen.

Table 3. CSM performance metrics on Company A data (480 data points used) and Company B data (229 data points used).

Performance Metric	CSM on Company A Data	CSM on Company B Data
Accuracy	0.86	0.82
Precision	0.83	0.87
Recall	0.87	0.86
F1 score	0.85	0.83

3.3. Active Learning on Individual Datasets

The AL presented in Figure 4, was implemented on the CSMs and the performance of the models was evaluated. Using Company A data, AL was utilized to reduce the data

needed to train the model. The model was initially trained using 180 combinations of Force 1 and Force 2 (30% of the dataset). The rest of the dataset (i.e., 420 combinations) was used for the initial model's validation to give the model the ability to retrain with all possible data points whose classification uncertainty during validation exceeds the 5% threshold. After the validation, the model selected 30 additional data points from the validation set (removed from the validation set and transferred to the training set). Thus, it was ultimately trained using 210 combinations of Force 1 and Force 2. Similarly, the Company B dataset was considered. Forces F_x and F_z were used in conjunction with the tool wear. The tool wear could have the values 0, 0.1 or 0.3 (low, middle and high tool wear). The CSM was initially trained using 30% of the data. This means that the initial training was performed using 86 combinations of F_x and F_z . During the AL phase, the model utilized 40 additional combinations of F_x and F_z for training. The final model was trained with 126 combinations of forces (43.8% of the original dataset). The model was then validated using the remaining 56.2% of the original dataset. The models' performance metrics can be found in Table 4.

Table 4. Performance metrics of CSM trained on Company A data and AL (210 data points used) and Company B data and AL (126 data points used).

Performance Metric	Company A Data	Company B Data
Accuracy	0.92	1
Precision	0.85	1
Recall	0.92	1
F1 score	0.88	1

It is derived from the findings presented in Tables 3 and 4, that the introduction of AL significantly reduces the amount of data required for training of the CSM. Furthermore, an increase of 6.9% in the accuracy of the model when trained using Company A data is possible with AL while using 56.25% less data. Similarly, when comparing the results after the CSM has trained on Company B data using AL, a reduction of 45% on the amount of data used for training is possible, with an 18% increase in accuracy. This is evident due to the ability of the AL to select the most informative data points from each dataset for training. Data points whose inclusion could negatively affect the model's performance are excluded.

3.4. Transfer Active Learning

Herein, the AL method presented before has been extended by considering transfer learning (TL) between the CSMs of Company A and Company B. This was performed by using the weights of the CSM trained on the Company A dataset to train the CSM with the Company B dataset and vice versa.

In the first experiment, the CSM trained using AL was retrained using AL with 30% of Company B data to develop a new CSM for Company B. A total of 86 combinations of F_x and F_z forces were used for training, and the rest data points were used for validation. The model selected 60 more data points. Thus, it was ultimately trained using 146 data points. Similarly, TL was performed in the opposite direction, using the weights of the CSM trained with Company B data to train the new CSM with Company A data using AL. In the first experiment, the SCM trained using AL was retrained using AL with 30% of Company A data. After the initial training, the model chose only 20 additional informative combinations of Force 1 and Force 2. Thus, the final model was ultimately trained using 200 combinations of Force 1 and Force 2. The performance metrics of the new CSM for Company A and B that have been trained by combining AL and TL is shown in Table 5.

Table 5. Performance metrics of CSM trained with Company A data with AL and TL from CSM trained with Company B data (trained using 200 data points) and Company B data with AL and TL from CSM trained with Company A data (trained using 146 data points).

Performance Metric	Company A Data	Company B Data
Accuracy	0.93	1
Precision	0.86	1
Recall	0.93	1
F1 score	0.89	1

3.5. Results Summary

Overall, the application of AL in both Company A and Company B data has resulted in an average increase in the models' performance of approximately 1%. This increase, while not significant, was possible, with an overall reduction of 60.6% reduction in required data points. This signifies an immense potential towards reducing the computational load, time and cost required to train ML models.

The goal of simultaneously using AL and TL is to further reduce the dependency on large datasets for ML model training and improve the model's accuracy. Such a case is evident when performing AL with Company A data while using the weights of the CSM trained on Company B data with AL. The model manages to select 10 fewer data points during the AL process, a reduction of 4.7% of forces combinations used by the final model, while achieving a marginally higher accuracy. Even though the higher accuracy achieved is not high the reduction of 4.7% in data points used can be significant in the process of reducing computational time and cost, especially in applications where the size of the dataset is significantly higher. However, this contradicts the findings after performing AL with Company B data while using the weights of the CSM trained on Company A data with AL. In this case, although the model manages to achieve the same classification performance as previously (without TL), it has used 20 more data points. This indicates an overall 15.8% increase in required data to achieve a similar performance.

When applying TL from Company A to Company B data, the increase in the required data for achieving similar performance can be attributed to the fact that the source and target datasets show significant differences in the distribution of the data. Additionally, due to the higher complexity of Company A data (as shown in Figures 5 and 6, the extracted weights from a model trained on Company A data are of lower quality and can provide less useful information to a model trained with Company B data. This poses a challenge to the model to adapt to Company B data since the pre-acquired knowledge of the model is not directly applicable. Thus, to overcome this challenge, more data points are required by the method.

In addition, to study the coupling of TL with AL without common elements in between, the training procedures of Section 3.3 for AL and Section 3.4 for AL/TL were reproduced, reserving in all cases, several samples for new testing. These samples (30%) had not been used for training or validation, as shown by the data portions illustration, in Figure 9. The performance metrics of AL in Company B remained unchanged while using 122 data points. Regarding the performance metrics of AL in Company A they remained mostly unchanged while using 246 data points, in terms of accuracy the result on the testing set was 0.91, precision was 0.92, the recall was 0.91 and the F1 score was 0.90. Lastly, data delineated in Tables 3 and 4 represent the mean values of performance metrics obtained from both the training and testing phases of the models, conducted with and without AL, over ten iterations each. To also ensure the reproducibility of the results, some of the iterations were conducted using regenerated Company A data using the equation discussed in the literature [5]. In the case of using Company B data and AL, the performance metrics always result in 1, while in the case of using Company A data and AL, the standard deviation of the accuracy was 0.031. This was performed to evaluate the reproducibility of the methodology, as the number of remaining test samples was low.

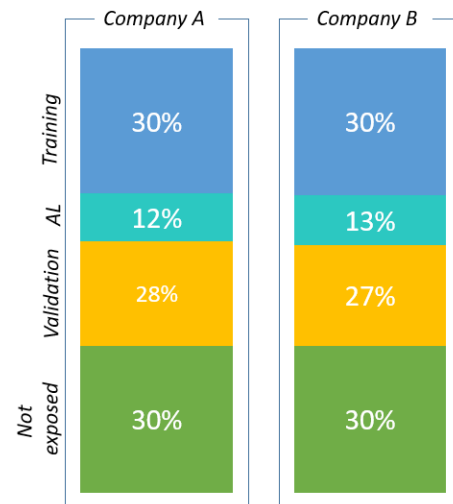


Figure 9. Data manipulation during AL/TL with 30% of the dataset not exposed for testing.

4. Discussion

Transfer learning on top of active learning is feasible, and, in some cases, it can lead to a reduction in the total amount of data required for training. In the case study presented in this work, the transfer from Company A to B did not affect the performance of the CSM model, while it enhanced it slightly in the case of transfer from Company B to A.

The area where the most misclassifications on Company A data were studied. As can be seen in Figure 10, where a portion of the data used to validate the CSM's performance is depicted, most of the misclassifications occur in the area of middle tool wear. This can be attributed to the underlying complexity of the data in the middle tool wear area when dealing with a high number of generated Force 1 and Force 2 combinations (600 force combinations in total). As depicted in Figure 10, point *c* is correctly classified as middle tool wear, while point *b* is misclassified as high tool wear as opposed to the correct middle tool wear label, while point *d* is misclassified as middle tool wear in contrast to the correct high tool wear label. This behaviour is not evident for data points near point *a* and *e* where the class differentiation (low and high tool wear respectively) is clearer. In addition, as observed from the two cases, the behavior of the datasets in terms of classes is quite different; in the case of company A, they are linear and non-separable, whilst in the case of company B, they are non-linear but fully separable. Also, the curves that could be used as threshold values for the two forces are non-linear. The differences can be seen in Figure 11. This depiction does not represent the actual data used, but it is conceptual and symbolically depicts the two differences that classes may appear to have. In any case, the two datasets are real and have been found as open data in the literature described above, in the overview of the pilot case.

As such, regarding the procedure of transfer learning, it is interesting to investigate in both scenarios, from company A to company B and vice versa. This checks the feasibility based on different criteria.

Furthermore, the results presented in this study affirm the belief that transfer learning can indeed improve an active learning ML model. As shown in [58], where a theory of transfer learning with applications to active learning is presented, and in [59,60], AL and TL are implemented, where informative data points are identified and manually selected through human intervention.

The presented approach is also characterized by its ease of reproducibility given its straightforward nature. This is due to the automated nature of selecting the most informative data points for further training. In [58], the presented AL approach requires human intervention to label new unlabeled data, which are then provided to the training set of the model. Furthermore, the architecture of the AL pipeline allows for a predefined number of unlabeled images to be labelled with human intervention as opposed to the approach

presented in Section 2, where AL is not constrained. Nevertheless, the approach presented in [58] highlights its potential for sample complexity reduction and efficiency gains. Similar findings are also reported in [60]. However, it is worth pointing out that the approaches discussed in the literature are focused on simplifying the labelling process of large datasets, in contrast to the presented approach which aims at reducing the dependency on large datasets, thus reducing the computational time and costs required for model training.

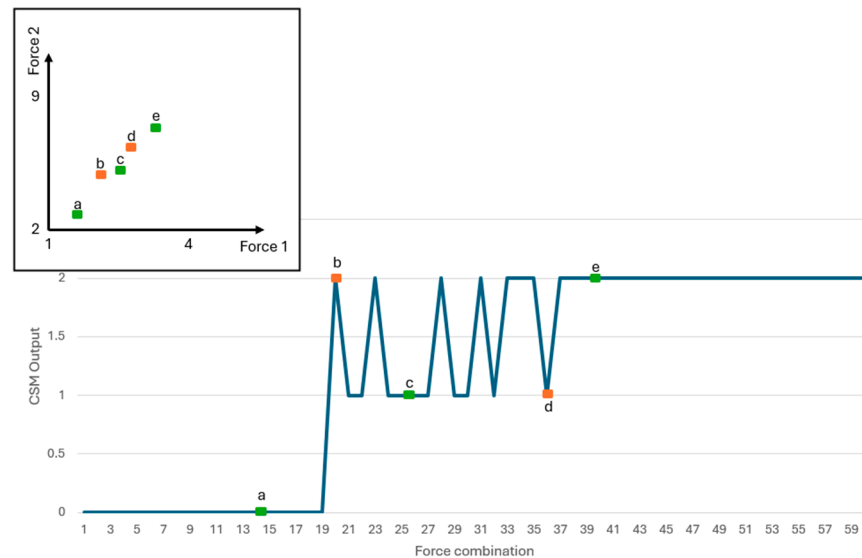


Figure 10. Sample of Company A data misclassifications.

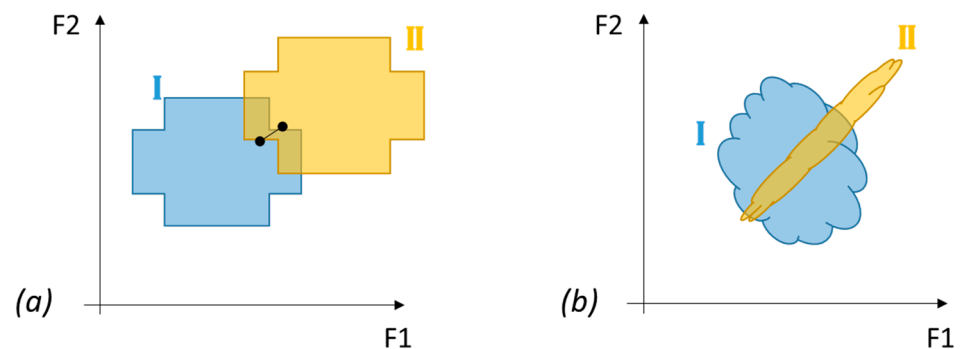


Figure 11. Factors of differentiation between two classes I and II: (a) separability and (b) geometry (linearity).

Moreover, alternative methodologies in the context of tool wear classification have been documented in literature which can potentially have similarly high accuracy to the accuracy achieved through AL/TL. In [9], a methodology utilizing images generated from sensor signals for tool degradation classification is presented. However, in contrast to the presented approach where through smaller datasets, high classification accuracy is achievable, the presented in [9] research utilized more than 20,000 images, resulting in a significant time needed for image labelling and image processing. In a similar manner, in [61], a CNN deep learning algorithm was used for tool wear classification using sensor signals. While the classifier was not capable of achieving high accuracy, the research signifies the computational burden that needed to be overcome during training due to the large number of data points required. This is again in contrast to the AL/TL approach presented, where computational time can be significantly reduced while retaining high classification accuracy.

However, additional characteristics should be considered in real applications, since sharing either models or data in the context of federated learning, is not straightforward.

Business issues, such as treating data as assets, need to be taken into consideration, especially if profit and confidentiality are in the way. More digital technologies should then be integrated, such as Blockchain and data encryption. This way, the network of engaged businesses taking part in this active and transfer learning will increase, leading to a more substantial final dataset. Also, the roles of the businesses in the network will determine the type of access to them. Categories can then be formed, i.e., companies with access to the model but not the data, partial access to the data and access to everything.

In addition, the value of a dataset [62] needs to be taken into consideration. This is highly related to the volume of the dataset. As such, the policy of “least data” seems promising towards creating high-value datasets.

Also, for the case of adding a slightly different process, such as milling, the procedure of creating transfer active learning is much more complicated. So, in the case of a dataset that can be quite complex in terms of predictability (Company C Dataset [63]), that has also been used previously in the literature [64], force signals (or time series, equivalently) constitute the force measurements. It is noted that taking simple metrics such as mean square value, mean value, or standard deviation, the classes defined by the tool wear levels are not separable in a straightforward way, at least in the context of utilizing the least data. Additionally, the evolution of tool wear in time is not the same for each one of the inserts.

Given the two last statements, it appears that transfer active learning can be the basis for knowledge building around similar applications, as it can be used to create a dataset that will provide maximum separability while keeping the data volume low. This knowledge could then be presented in alternative ways so that human operators can also participate in the procedure thus achieving collaborative human–machine decision making. Additional work is, nevertheless, expected towards this end.

5. Conclusions

In this study, an approach to utilizing a combination of AL and TL was presented. AL using Company B data shows that high classification accuracy is achievable while using 43.8% of the dataset. The implementation of active learning significantly reduces this amount while maintaining high accuracy, resulting in a significant reduction in computational time and power required for model training while also reducing the dependency on large amounts of data. Similarly, when using active learning with Company A data, the data required to achieve above-average classification performance are 35% of the dataset. However, due to the shallow nature of the NN model used and due to the linearity of the data, it is worth noting that with a dataset with higher complexity, the required data for achieving high classification accuracy could be higher, but this could be a topic of future work.

Additionally in this study, it is shown that transfer learning will not always achieve the desired results. This is indicative when performing transfer active learning from Company A data to Company B data due to the increase in required Company B data during training to achieve the same accuracy as that achieved when TL was solely deployed on Company B data. However, it is evident that through the proposed approach of utilizing AL in conjunction with TL, the dependency on large datasets can be significantly reduced.

The performance of machine learning alone, depending on the nature of the data, is less than 90%. However, the introduction of AL may increase the performance, even by far, especially if the dataset is of high value. Additionally, transfer learning combined with active learning seems to benefit the network of the companies involved, as the combined performance is even larger, resulting in a mean value of 95%.

Regarding future work, it would be highly interesting to compare classification from different processes, possibly with different features from multimodal monitoring. For instance, in the case of milling, additional signals could be utilized, beyond force. As the application herein is rather narrow (especially in the sense that it regards point-like measurements), it would be desired in the future to run this concept for a different case, i.e., milling, and more than two stakeholders, so that the complexity of the case study

is adequately large to act as a reference. Furthermore, it would be desirable to address encryption and hiding knowledge on top of such applications, to ensure business added value, while the integration of explainability would be extremely important, as extra benefits would be added to this decision-making procedure.

Author Contributions: Conceptualization, G.C.; methodology, K.A.; software, P.C.; validation, P.S.; formal analysis, A.P.; writing—original draft preparation, A.P. and P.C.; writing—review and editing, P.S., K.A. and G.C. All authors have read and agreed to the published version of the manuscript.

Funding: This research was partially supported by the EU project AIRISE (101092312).

Data Availability Statement: The original contributions presented in the study are included in the article, further inquiries can be directed to the corresponding author.

Conflicts of Interest: The authors declare no conflict of interest.

References

- Herrera-Granados, G.; Misaka, T.; Herwan, J.; Komoto, H.; Furukawa, Y. An Experimental Study of Multi-Sensor Tool Wear Monitoring and Its Application to Predictive Maintenance. *Int. J. Adv. Manuf. Technol.* **2024**, 1–19. [\[CrossRef\]](#)
- Bagga, P.J.; Makhesana, M.A.; Patel, K.M. A Novel Approach of Combined Edge Detection and Segmentation for Tool Wear Measurement in Machining. *Prod. Eng. Res. Devel.* **2021**, *15*, 519–533. [\[CrossRef\]](#)
- Pimenov, D.Y.; Gupta, M.K.; da Silva, L.R.; Kiran, M.; Khanna, N.; Krolczyk, G.M. Application of measurement systems in tool condition monitoring of Milling: A review of measurement science approach. *Measurement* **2022**, *199*, 11503. [\[CrossRef\]](#)
- Chryssolouris, G.; Alexopoulos, K.; Arkouli, Z. *A Perspective on Artificial Intelligence in Manufacturing*; Springer: Cham, Switzerland, 2023.
- Twardowski, P.; Wiciak-Pikuła, M. Prediction of tool wear using artificial neural networks during turning of hardened steel. *Materials* **2019**, *12*, 3091. [\[CrossRef\]](#) [\[PubMed\]](#)
- Niu, B.; Sun, J.; Yang, B. Multisensory based tool wear monitoring for practical applications in milling of titanium alloy. *Mater. Today Proc.* **2020**, *22*, 1209–1217. [\[CrossRef\]](#)
- Bagga, P.J.; Makhesana, M.A.; Patel, H.D.; Patel, K.M. Indirect method of tool wear measurement and prediction using ANN network in machining process. *Mater. Today Proc.* **2021**, *44*, 1549–1554. [\[CrossRef\]](#)
- Ma, J.; Luo, D.; Liao, X.; Zhang, Z.; Huang, Y.; Lu, J. Tool wear mechanism and prediction in milling TC18 titanium alloy using deep learning. *Measurement* **2021**, *173*, 108554. [\[CrossRef\]](#)
- Martínez-Arellano, G.; Terrazas, G.; Ratchev, S. Tool wear classification using time series imaging and deep learning. *Int. J. Adv. Manuf. Technol.* **2019**, *104*, 3647–3662. [\[CrossRef\]](#)
- Awasthi, U.; Wang, Z.; Mannan, N.; Pattipati, K.R.; Bollas, G.M. Physics-based modeling and information-theoretic sensor and settings selection for tool wear detection in precision machining. *J. Manuf. Process.* **2022**, *81*, 127–140. [\[CrossRef\]](#)
- Liu, T.; Wang, Q.; Wang, W. Micro-Milling Tool Wear Monitoring via Nonlinear Cutting Force Model. *Micromachines* **2022**, *13*, 943. [\[CrossRef\]](#)
- Xi, T.; Benincá, I.M.; Kehne, S.; Fey, M.; Brecher, C. Tool wear monitoring in roughing and finishing processes based on machine internal data. *Int. J. Adv. Manuf. Technol.* **2021**, *113*, 3543–3554. [\[CrossRef\]](#)
- Matos, F.; Silva, T.E.F.; Marques, F.; Figueiredo, D.; Rosa, P.A.R.; de Jesus, A.M.P. Machinability assessment of Inconel 718 turning using PCBN cutting tools. *Procedia CIRP* **2023**, *117*, 468–473. [\[CrossRef\]](#)
- Shurrab, S.; Almshnanah, A.; Duwairi, R. Tool Wear Prediction in Computer Numerical Control Milling Operations via Machine Learning. In Proceedings of the 2021 12th International Conference on Information and Communication Systems (ICICS), Valencia, Spain, 24 May 2021; pp. 220–227.
- Zhou, X.; Yu, T.; Wang, G.; Guo, R.; Fu, Y.; Sun, Y.; Chen, M. Tool Wear Classification Based on Convolutional Neural Network and Time Series Images during High Precision Turning of Copper. *Wear* **2023**, *522*, 204692. [\[CrossRef\]](#)
- Schwenzer, M.; Miura, K.; Bergs, T. Machine Learning for Tool Wear Classification in Milling Based on Force and Current Sensors. *IOP Conf. Ser. Mater. Sci. Eng.* **2019**, *520*, 012009. [\[CrossRef\]](#)
- Stavropoulos, P.; Papacharalampopoulos, A.; Vasiladis, E.; Chryssolouris, G. Tool wear predictability estimation in milling based on multi-sensorial data. *Int. J. Adv. Manuf. Technol.* **2016**, *82*, 509–521. [\[CrossRef\]](#)
- da Silva, L.S.; Rebelo, C.; Nethercot, D.; Marques, L.; Simões, R.; Real, P.V. Statistical evaluation of the lateral–torsional buckling resistance of steel I-beams, Part 2: Variability of steel properties. *J. Constr. Steel Res.* **2009**, *65*, 832–849. [\[CrossRef\]](#)
- Singh, K.K.; Singh, R. Process mechanics based uncertainty modeling for cutting force prediction in high speed micromilling of Ti6Al4V. *Procedia Manuf.* **2020**, *48*, 273–282. [\[CrossRef\]](#)
- Xu, T.; Wang, K.; Song, S. Measurement uncertainty and representation of tensile mechanical properties in metals. *Metals* **2021**, *11*, 1733. [\[CrossRef\]](#)
- Choi, M.K.; Huh, H.; Jeong, S.; Kim, C.G.; Chae, K.S. Measurement uncertainty evaluation with correlation for dynamic tensile properties of auto-body steel sheets. *Int. J. Mech. Sci.* **2017**, *130*, 174–187. [\[CrossRef\]](#)

22. Tao, W.; Zhu, P.; Xu, C.; Liu, Z. Uncertainty quantification of mechanical properties for three-dimensional orthogonal woven composites. Part II: Multiscale simulation. *Compos. Struct.* **2020**, *235*, 111764.
23. Liu, Y.; Xiong, Z.; Liu, Z. Stochastic Cutting Force Modeling and Prediction in Machining. *J. Manuf. Sci. Eng.* **2020**, *142*, 121004. [\[CrossRef\]](#)
24. Stavropoulos, P.; Papacharalampopoulos, A.; Souflas, T. Indirect online tool wear monitoring and model-based identification of process-related signal. *Adv. Mech. Eng.* **2020**, *12*, 1687814020919209. [\[CrossRef\]](#)
25. Adadi, A. A survey on data-efficient algorithms in big data era. *J. Big Data* **2021**, *8*, 24. [\[CrossRef\]](#)
26. Alexopoulos, K.; Sipsas, K.; Xanthakis, E.; Makris, S.; Mourtzis, D. An industrial Internet of things based platform for context-aware information services in manufacturing. *Int. J. Comput. Integr. Manuf.* **2018**, *31*, 1111–1123. [\[CrossRef\]](#)
27. Chanda, S.S.; Banerjee, D.N. Omission and commission errors underlying AI failures. *AI Soc.* **2022**, 1–24. [\[CrossRef\]](#) [\[PubMed\]](#)
28. Papacharalampopoulos, A.; Michail, C.K.; Stavropoulos, P. Manufacturing resilience and agility through processes digital twin: Design and testing applied in the LPBF case. *Procedia CIRP* **2021**, *103*, 164–169. [\[CrossRef\]](#)
29. Zhang, L.; Chen, X.; Zhou, W.; Cheng, T.; Chen, L.; Guo, Z.; Han, B.; Lu, L. Digital Twins for Additive Manufacturing: A State-of-the-Art Review. *Appl. Sci.* **2020**, *10*, 8350. [\[CrossRef\]](#)
30. Gohari, H.; Mohamed, A.; Hassan, M.; M'Saoubi, R.; Attia, H. Hybrid Offline-Online Optimization, Monitoring and Control of Milling Processes. *CIRP Ann.-Manuf. Technol.* **2023**, *72*, 4. [\[CrossRef\]](#)
31. Hassan, M.; Sadek, A.; Attia, M.H. Novel sensor-based tool wear monitoring approach for seamless implementation in high speed milling applications. *CIRP Ann.* **2021**, *70*, 87–90. [\[CrossRef\]](#)
32. Wanner, J.; Weeber, M.; Birke, K.P.; Sauer, A. Quality modelling in battery cell manufacturing using soft sensing and sensor fusion-A review. In Proceedings of the 2019 9th International Electric Drives Production Conference (EDPC), Esslingen, Germany, 3–4 December 2019; pp. 1–9.
33. Dass, A.; Moridi, A. State of the art in directed energy deposition: From additive manufacturing to materials design. *Coatings* **2019**, *9*, 418. [\[CrossRef\]](#)
34. Harris, G.; Yarbrough, A.; Abernathy, D.; Peters, C. Manufacturing readiness for digital manufacturing. *Manuf. Lett.* **2019**, *22*, 16–18. [\[CrossRef\]](#)
35. Rizal, M.; Ghani, J.A.; Nuawi, M.Z.; Haron, C.H.C. Online tool wear prediction system in the turning process using an adaptive neuro-fuzzy inference system. *Appl. Soft Comput.* **2013**, *13*, 1960–1968. [\[CrossRef\]](#)
36. Coulson, M.; Healey, M.; Fidler, F.; Cumming, G. Confidence intervals permit, but don't guarantee, better inference than statistical significance testing. *Front. Psychol.* **2010**, *1*, 26. [\[CrossRef\]](#)
37. Strickland, E. Andrew Ng, AI minimalist: The machine-learning pioneer says small is the new big. *IEEE Spectr.* **2022**, *59*, 22–50. [\[CrossRef\]](#)
38. Xu, P.; Ji, X.; Li, M.; Lu, W. Small data machine learning in materials science. *NPJ Comput. Mater.* **2023**, *9*, 42. [\[CrossRef\]](#)
39. Sucholutsky, I.; Schonlau, M. Less than one-Shot Learning: Learning N classes from $M < N$ samples. *Proc. AAAI Conf. Artif. Intell.* **2021**, *35*, 9739–9746.
40. Kim, Y.; Kim, T.; Youn, B.D.; Ahn, S.H. Machining quality monitoring (MQM) in laser-assisted micro-milling of glass using cutting force signals: An image-based deep transfer learning. *J. Intell. Manuf.* **2022**, *33*, 1813–1828. [\[CrossRef\]](#)
41. Wang, J.; Yang, S.; Liu, Y.; Wen, G. Deep Subdomain Transfer Learning with Spatial Attention ConvLSTM Network for Fault Diagnosis of Wheelset Bearing in High-Speed Trains. *Machines* **2023**, *11*, 304. [\[CrossRef\]](#)
42. Ali, A.A.; Chramcov, B.; Jasek, R.; Katta, R.; Krayem, S.; Kadi, M. Detection of Steel Surface Defects Using U-Net with Pre-trained Encoder. In *Software Engineering Application in Informatics, Proceedings of the Computational Methods in Systems and Software, Online, Czech Republic, 1 October 2021*; Silhavy, R., Silhavy, P., Prokopova, Z., Eds.; Springer: Cham, Switzerland, 2021; pp. 185–196.
43. Pickering, E.; Guth, S.; Karniadakis, G.E.; Sapsis, T.P. Discovering and forecasting extreme events via active learning in neural operators. *Nat. Comput. Sci.* **2022**, *2*, 823–833. [\[CrossRef\]](#)
44. Martinez Arellano, G.; Ratchev, S. Towards an active learning approach to tool condition monitoring with bayesian deep learning. In Proceedings of the 33rd International ECMS Conference on Modelling and Simulation, Caserta, Italy, 11–14 June 2019.
45. Li, Y.; Wang, J.; Huang, Z.; Gao, R.X. Physics-informed meta learning for machining tool wear prediction. *J. Manuf. Syst.* **2022**, *62*, 17–27. [\[CrossRef\]](#)
46. Zou, Z.; Karniadakis, G.E. L-HYDRA: Multi-Head Physics-Informed Neural Networks. *arXiv* **2023**, arXiv:2301.02152.
47. Lu, L.; Jin, P.; Karniadakis, G.E. Deeponet: Learning nonlinear operators for identifying differential equations based on the universal approximation theorem of operators. *arXiv* **2019**, arXiv:1910.03193.
48. Rožanec, J.M.; Trajkova, E.; Dam, P.; Fortuna, B.; Mladenicić, D. Streaming Machine Learning and Online Active Learning for Automated Visual Inspection. *IFAC-Pap.* **2022**, *55*, 277–282. [\[CrossRef\]](#)
49. Kanagavelu, R.; Li, Z.; Samsudin, J.; Hussain, S.; Yang, F.; Yang, Y.; Goh, R.; Cheah, M. Federated learning for advanced manufacturing based on industrial IoT data analytics. In *Implementing Industry 4.0: The Model Factory as the Key Enabler for the Future of Manufacturing*; Springer: New York, NY, USA, 2021; pp. 143–176.
50. Hewage, C.; Rahulamathavan, Y.; Ratnayake, D. (Eds.) *Data Protection in a Post-Pandemic Society: Laws, Regulations, Best Practices and Recent Solutions*; Springer International Publishing: Cham, Switzerland, 2023. [\[CrossRef\]](#)
51. Gosselin, R.; Vieu, L.; Loukil, F.; Benoit, A. Privacy and Security in Federated Learning: A Survey. *Appl. Sci.* **2022**, *12*, 9901. [\[CrossRef\]](#)

52. Abad, G.; Picek, S.; Ramírez-Durán, V.J.; Urbieto, A. SoK: On the Security & Privacy in Federated Learning. *arXiv* **2021**, arXiv:2112.05423.
53. Canal, A.D. Surface Roughness Analysis in Turning Processes Using ANN. Realização de Instituto Tecnológico de Aeronáutica. São José dos Campos: ITA. 2022. Available online: http://www.bdlita.bibl.ita.br/tesesdigitais/lista_resumo.php?num_tese=78535 (accessed on 11 May 2023).
54. Sklearn.tree.DecisionTreeClassifier. Available online: <https://scikit-learn.org/stable/modules/generated/sklearn.tree.DecisionTreeClassifier.html> (accessed on 6 July 2023).
55. Sklearn.svm.SVC. Available online: <https://scikit-learn.org/stable/modules/generated/sklearn.svm.SVC.html> (accessed on 6 July 2023).
56. Sklearn.neural_network.MLPClassifier. Available online: https://scikit-learn.org/stable/modules/generated/sklearn.neural_network.MLPClassifier.html (accessed on 6 July 2023).
57. Kingma, D.P.; Ba, J. Adam: A Method for Stochastic Optimization. *arXiv* **2017**, arXiv:1412.6980.
58. Yang, L.; Hanneke, S.; Carbonell, J. A theory of transfer learning with applications to active learning. *Mach Learn.* **2013**, *90*, 161–189. [[CrossRef](#)]
59. Nakano, F.K.; Cerri, R.; Vens, C. Active learning for hierarchical multi-label classification. *Data Min. Knowl. Disc.* **2020**, *34*, 1496–1530. [[CrossRef](#)]
60. Kale, D.; Liu, Y. Accelerating Active Learning with Transfer Learning. In Proceedings of the 2013 IEEE 13th International Conference on Data Mining. Presented at the 2013 IEEE International Conference on Data Mining (ICDM), IEEE, Dallas, TX, USA, 7–10 December 2013; pp. 1085–1090. [[CrossRef](#)]
61. Terrazas, G.; Martínez-Arellano, G.; Benardos, P.; Ratchev, S. Online Tool Wear Classification during Dry Machining Using Real Time Cutting Force Measurements and a CNN Approach. *J. Manuf. Mater. Process.* **2018**, *2*, 72. [[CrossRef](#)]
62. Wuest, T.; Irgens, C.; Thoben, K.-D. An Approach to Monitoring Quality in Manufacturing Using Supervised Machine Learning on Product State Data. *J. Intell. Manuf.* **2014**, *25*, 1167–1180. [[CrossRef](#)]
63. IEEE Data Port. Available online: <https://ieee-dataport.org/open-access/toolwear-dataset-nuaaideahouse> (accessed on 10 May 2023).
64. Sayyad, S.; Kumar, S.; Bongale, A.; Kotecha, K.; Selvachandran, G.; Suganthan, P.N. Tool wear prediction using long short-term memory variants and hybrid feature selection techniques. *Int. J. Adv. Manuf. Technol.* **2022**, *121*, 6611–6633. [[CrossRef](#)]

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.